

Methods in
Molecular Biology 2226

Springer Protocols

Florencia Cidre-Aranaz
Thomas G. P. Grünewald *Editors*

Ewing Sarcoma

Methods and Protocols

MOREMEDIA 

 Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor

John M. Walker

School of Life and Medical Sciences

University of Hertfordshire

Hatfield, Hertfordshire, UK

For further volumes:

<http://www.springer.com/series/7651>

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

Ewing Sarcoma

Methods and Protocols

Edited by

Florencia Cidre-Aranaz and Thomas G. P. Grünewald

*Division of Translational Pediatric Sarcoma Research, German Cancer Research Center (DKFZ)
Hopp-Children's Cancer Center (KITZ), Heidelberg, Germany*

Editors

Florencia Cidre-Aranaz
Division of Translational Pediatric
Sarcoma Research
German Cancer Research Center (DKFZ)
Hopp-Children's Cancer Center (KITZ)
Heidelberg, Germany

Thomas G. P. Grünewald
Division of Translational Pediatric Sarcoma Research
German Cancer Research Center (DKFZ)
Hopp-Children's Cancer Center (KITZ)
Heidelberg, Germany

ISSN 1064-3745

Methods in Molecular Biology

ISBN 978-1-0716-1019-0

<https://doi.org/10.1007/978-1-0716-1020-6>

ISSN 1940-6029 (electronic)

ISBN 978-1-0716-1020-6 (eBook)

© Springer Science+Business Media, LLC, part of Springer Nature 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Humana imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.

The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

Preface

Ewing sarcoma is a highly aggressive bone and soft tissue tumor, which mainly affects children, adolescents, and young adults. It is a clinically complex disease, as despite treatment by local resection and systemic chemotherapy, patients presenting with metastasis at diagnosis or relapse have still a dismal prognosis. Paradoxically, Ewing sarcoma is a genetically relatively simple cancer, characterized by a pathognomonic fusion gene (EWSR1-ETS) that, once established, completely rewires the epigenetic and transcriptional signatures of the tumors' cell(s)-of-origin. Although great advances have been achieved in the past decades, the exact mechanisms underlying Ewing sarcoma tumorigenesis and progression are still largely unknown.

One of the most challenging aspects of Ewing sarcoma research and clinical assessment is the scarcity of well-curated, highly standardized, and optimally working methods. Thus, this volume of *Methods in Molecular Biology*TM aims at comprising a selection of the most effective and widely employed protocols in the Ewing sarcoma field.

In order to accommodate a wide variety of *in vitro*, *in vivo*, and *in silico* techniques that may be applied to the analysis of Ewing sarcoma, this volume of the *Methods in Molecular Biology*TM series is organized into five sections: *Section I* deals with a selection of molecular biology techniques that have proven to be very helpful when analyzing Ewing sarcoma. This section starts with an introductory chapter explaining the key question of somatic alteration and genetic predisposition, which may aid to stratify patients according to risk in Ewing sarcoma. Additionally, this section describes Ewing sarcoma-specific protocols for molecular techniques including tissue preservation, nucleic acid isolation from formalin-fixed paraffin-embedded samples, which are usually the result of patient's solid biopsies and Western blotting. This chapter is complemented with a final comprehensive chapter on liquid biopsies.

In order to help Ewing sarcoma clinicians and physician-scientists navigate the intricacies of diagnosis, *Section II* comprises all techniques applied to the morpho-molecular diagnostics in Ewing sarcoma, including a chapter on routine immune-histological analysis which additionally describes all pre- and post-diagnostic procedures and three subsequent chapters on every molecular aspect that can be part of Ewing sarcoma diagnosis or confirmation (fluorescence in situ hybridization (FISH), RT-PCR, and targeted RNA sequencing).

Section III of this volume describes a collection of *in vitro* tools and cell culture functional assays to guide researchers on essential wet-lab techniques from Ewing sarcoma-specific re-expression models, to luciferase assays adapted to the analysis of regulatory sequences present in Ewing sarcoma, to a set of functional assays implemented to evaluate the potential effect of re-expression or downregulation of particular genes or pathways and/or determine the effectiveness of selected drugs.

Despite constant efforts since its first description in 1921, Ewing sarcoma remains a tumor entity of unclear cell of origin, which has underlined the importance of developing diverse methods that could recapitulate the disease *in vivo*. *Section IV* presents six chapters addressing these important animal models, starting with a method for a genetically engineered mouse model and covering other models such as classical subcutaneous xenografting and orthotopic implants. Due to the increasing need for systems that could better

recapitulate the complexity of patient's tumors, we have included in this *Section IV* a dedicated chapter on patient-derived xenografts (PDX) and orthotopic xenografts for metastasis assessment in Ewing sarcoma. In addition to these mouse models, a protocol for developing a zebrafish model is also detailed.

As a result of constant development of new projects, Ewing sarcoma researchers create a staggering increase of highly valuable data. However, these data usually require being properly bioinformatically analyzed in order to obtain its maximum profitability. Thus, *Section V* of this volume contains four bioinformatic-focused chapters that range from a thorough compilation and description of repositories and databases essential for Ewing sarcoma research to a detailed explanation on how to address ChIP-seq, epigenetic, and systems biology analyses by using both own or already published data.

Collectively, this volume of *Methods in Molecular Biology™* will be of interest to molecular biologists and physician-scientists working in the field of Ewing sarcoma or in pediatric fusion-driven tumors, physicians interested in diagnostic aspects, as well as to bioinformaticians approaching the analysis of Ewing sarcoma.

It has been our great pleasure to have collaborated in this volume with so many talented authors that are leaders in their respective fields. We are immensely thankful for their dedication to producing comprehensive chapters filled with extensively detailed tricks perfected at their laboratories, which will definitely help guide both the advanced and novice researchers.

It is our hope that this volume comprising a collective technical knowledge around Ewing sarcoma will bring us one step forward in the quest to fight this devastating disease.

Heidelberg, Germany

*Florencia Cidre-Aranaz
Thomas G. P. Grünewald*

Contents

<i>Preface</i>	<i>v</i>
<i>Contributors</i>	<i>ix</i>
PART I MOLECULAR BIOLOGY TECHNIQUES APPLIED TO EWING SARCOMA	
1 Germline Variation and Somatic Alterations in Ewing Sarcoma	3
<i>Mitchell J. Machiela and Thomas G. P. Grünewald</i>	
2 Western Blot Analysis in Ewing Sarcoma	15
<i>Aruna Marchetto and Laura Romero-Pérez</i>	
3 Tissue Preservation and FFPE Samples: Optimized Nucleic Acids Isolation in Ewing Sarcoma	27
<i>Laura Romero-Pérez and Thomas G. P. Grünewald</i>	
4 Liquid Biopsies in Ewing Sarcoma	39
<i>Manuela Krumbholz and Markus Metzler</i>	
PART II MORPHO-MOLECULAR DIAGNOSTICS IN EWING SARCOMA	
5 (Immuno)histological Analysis of Ewing Sarcoma	49
<i>David Marcilla, Isidro Machado, Thomas G. P. Grünewald, Antonio Llombart-Bosch, and Enrique de Alava</i>	
6 Molecular Approaches to Diagnosis in Ewing Sarcoma: Fluorescence In Situ Hybridization (FISH)	65
<i>Marcel Trautmann and Wolfgang Hartmann</i>	
7 Molecular Approaches to Diagnosis in Ewing Sarcoma: RT-PCR	85
<i>Carlos Rodríguez-Martín and Javier Alonso</i>	
8 Molecular Approaches to Diagnosis in Ewing Sarcoma: Targeted RNA Sequencing	105
<i>Carmen Salguero-Aranda and Juan Diaz-Martin</i>	
PART III CELLULAR BIOLOGY AND CELL CULTURE TECHNIQUES	
9 Ewing Sarcoma-Specific (Re)expression Models	119
<i>Maximilian M. L. Knott and Florencia Cidre-Aranaz</i>	
10 Analysis of Regulatory DNA Sequences by Dual-Luciferase Reporter Assays in Ewing Sarcoma	139
<i>Tilman L. B. Hölting and Maximilian M. L. Knott</i>	
11 Proliferation Assessment by Trypan Blue Exclusion in Ewing Sarcoma	151
<i>Cornelius Maximilian Funk and Julian Musa</i>	
12 Drug Screening by Resazurin Colorimetry in Ewing Sarcoma	159
<i>Julian Musa and Florencia Cidre-Aranaz</i>	

13	Analysis of Migration and Invasion in Ewing Sarcoma	167
	<i>Florencia Cidre-Aranaz</i>	
PART IV ANIMAL MODELS IN EWING SARCOMA		
14	Genetically Engineered Mouse Model in Ewing Sarcoma	183
	<i>Miwa Tanaka and Takuro Nakamura</i>	
15	Tumor Growth Analysis of Ewing Sarcoma Cell Lines Using Subcutaneous Xenografts in Mice	191
	<i>Florencia Cidre-Aranaz and Shunya Ohmura</i>	
16	Metastasis Assessment in Ewing Sarcoma Using Orthotopic Xenografts	201
	<i>Roser López-Aleman and Oscar M. Tirado</i>	
17	Orthotopic Implants in Mice	215
	<i>Elizabeth Stewart</i>	
18	Ewing Sarcoma PDX Models	223
	<i>Didier Surdez, Lorena Landuzzi, Katia Scotlandi, and Maria Cristina Manara</i>	
19	Using Zebrafish Larvae as a Xenotransplantation Model to Study Ewing Sarcoma	243
	<i>Susana Pascoal, Sarah Grissenberger, Eva Scheuringer, Rita Fior, Miguel Godinho Ferreira, and Martin Distel</i>	
PART V BIOINFORMATIC APPROACHES TO EWING SARCOMA		
20	Main Repositories and Databases Used in Ewing Sarcoma Research	259
	<i>Maria Debiec-Rychter</i>	
21	Chromatin Immunoprecipitation Followed by Next-Generation Sequencing (ChIP-Seq) Analysis in Ewing Sarcoma	265
	<i>Gwenneg Kerdivel and Valentina Boeva</i>	
22	Epigenetic Analysis in Ewing Sarcoma	285
	<i>Jeremy M. Simon and Nicholas C. Gomez</i>	
23	Systems Biology Analysis for Ewing Sarcoma	303
	<i>Marianyela Petrizzelli, Jane Merlevede, and Andrei Zinovyev</i>	
	<i>Index</i>	335

Contributors

- JAVIER ALONSO • *Unidad de Tumores Sólidos Infantiles, Instituto de Investigación de Enfermedades Raras, Instituto de Salud Carlos III, Madrid, Spain; Centro de Investigación Biomédica en Red de Enfermedades Raras, Instituto de Salud Carlos III (CB06/07/1009; CIBERER-ISCI3), Madrid, Spain*
- VALENTINA BOEVA • *INSERM, U1016, Cochin Institute, CNRS UMR8104, Paris Descartes University, Paris, France; Department of Computer Science, ETH Zurich, Institute for Machine Learning, Zurich, Switzerland; Swiss Institute of Bioinformatics (SIB), Zürich, Switzerland*
- FLORENCIA CIDRE-ARANAZ • *Division of Translational Pediatric Sarcoma Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany; German Cancer Consortium (DKTK), Partner Site Heidelberg, Heidelberg, Germany*
- ENRIQUE DE ALAVA • *Institute of Biomedicine of Sevilla (IBiS), Virgen del Rocío University Hospital/CSIC/University of Sevilla/CIBERONC, Seville, Spain; Department of Normal and Pathological Cytology and Histology, School of Medicine, University of Seville, Seville, Spain*
- MARIA DEBIEC-RYCHTER • *Department of Human Genetics, KU Leuven an, University Hospital Leuven, Leuven, Belgium*
- JUAN DIAZ-MARTIN • *Department of Pathology, Instituto de Biomedicina de Sevilla, CSIC-Universidad de Sevilla, Hospital Universitario Virgen del Rocío, Seville, Spain*
- MARTIN DISTEL • *Innovative Cancer Models, St. Anna Children's Cancer Research Institute, Vienna, Austria*
- MIGUEL GODINHO FERREIRA • *Institute for Research on Cancer and Aging of Nice (IRCAN), Université Côte d'Azur, Nice Cedex 2, France*
- RITA FIOR • *Champalimaud Centre for the Unknown, Champalimaud Foundation, Lisbon, Portugal*
- CORNELIUS MAXIMILIAN FUNK • *Division of Translational Pediatric Sarcoma Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany*
- NICHOLAS C. GOMEZ • *Laboratory of Mammalian Cell Biology and Development, The Rockefeller University, New York, NY, USA*
- SARAH GRISSENBERGER • *Innovative Cancer Models, St. Anna Children's Cancer Research Institute, Vienna, Austria*
- THOMAS G. P. GRÜNEWALD • *Division of Translational Pediatric Sarcoma Research, German Cancer Research Center (DKFZ), Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany*
- WOLFGANG HARTMANN • *Division of Translational Pathology, Gerhard-Domagk-Institute of Pathology, Münster, Germany*
- TILMAN L. B. HÖLTING • *Max-Eder Research Group for Pediatric Sarcoma Biology, Faculty of Medicine, Institute of Pathology, LMU Munich, Munich, Germany*
- GWENNEG KERDVEL • *Cochin Institute, INSERM U1016, CNRS UMR8104, University of Paris, Paris, France*

- MAXIMILIAN M. L. KNOTT • *Max-Eder Research Group for Pediatric Sarcoma Biology, Faculty of Medicine, Institute of Pathology, LMU Munich, Munich, Germany*
- MANUELA KRUMBHOLZ • *Department of Pediatrics, University Hospital Erlangen, Erlangen, Germany*
- LORENA LANDUZZI • *IRCCS—Istituto Ortopedico Rizzoli, Experimental Oncology Laboratory, Bologna, Italy*
- ANTONIO LLOMBART-BOSCH • *Department of Pathology, University of Valencia, Valencia, Spain*
- ROSER LÓPEZ-ALEMANY • *Sarcoma Research Group, Oncobell Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain*
- ISIDRO MACHADO • *Pathology Department, Instituto Valenciano de Oncología, Valencia, Spain; Pathology Department, Patologika Hospital Quirón Valencia, Valencia, Spain*
- MITCHELL J. MACHIELA • *Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA*
- MARIA CRISTINA MANARA • *IRCCS—Istituto Ortopedico Rizzoli, Experimental Oncology Laboratory, Bologna, Italy*
- ARUNA MARCHETTO • *Max-Eder Research Group for Pediatric Sarcoma Biology, Faculty of Medicine, Institute of Pathology, LMU Munich, Munich, Germany*
- DAVID MARCILLA • *Institute of Biomedicine of Sevilla (IBiS), Virgen del Rocío University Hospital/CSIC/University of Sevilla/CIBERONC, Seville, Spain*
- JANE MERLEVEDE • *Institut Curie, PSL Research University, Paris, France; INSERM U900, Paris, France; CBIO-Centre for Computational Biology, Mines ParisTech, PSL Research University, Paris, France*
- MARKUS METZLER • *Department of Pediatrics, University Hospital Erlangen, Erlangen, Germany*
- JULIAN MUSA • *Division of Translational Pediatric Sarcoma Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; Department of General, Visceral and Transplantation Surgery, University of Heidelberg, Heidelberg, Germany; Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany; German Cancer Consortium (DKTK), Partner Site Heidelberg, Heidelberg, Germany*
- TAKURO NAKAMURA • *Division of Carcinogenesis, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan*
- SHUNYA OHMURA • *Max-Eder Research Group for Pediatric Sarcoma Biology, Institute of Pathology, Faculty of Medicine, LMU Munich, Munich, Germany*
- SUSANA PASCOAL • *Innovative Cancer Models, St. Anna Children's Cancer Research Institute, Vienna, Austria*
- MARIANYELA PETRIZZELLI • *Institut Curie, PSL Research University, Paris, France; INSERM U900, Paris, France; CBIO-Centre for Computational Biology, Mines ParisTech, PSL Research University, Paris, France*
- CARLOS RODRÍGUEZ-MARTÍN • *Unidad de Tumores Sólidos Infantiles, Instituto de Investigación de Enfermedades Raras, Instituto de Salud Carlos III, Madrid, Spain*
- LAURA ROMERO-PÉREZ • *Division of Translational Pediatric Sarcoma Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; Hopp Children's Cancer Center (KiTZ), Heidelberg, Germany; German Cancer Consortium (DKTK), Heidelberg, Germany*
- CARMEN SALGUERO-ARANDA • *Department of Pathology, Instituto de Biomedicina de Sevilla, CSIC-Universidad de Sevilla, Hospital Universitario Virgen del Rocío, Seville, Spain*

- EVA SCHEURINGER • *Innovative Cancer Models, St. Anna Children's Cancer Research Institute, Vienna, Austria*
- KATIA SCOTLANDI • *IRCCS—Istituto Ortopedico Rizzoli, Experimental Oncology Laboratory, Bologna, Italy*
- JEREMY M. SIMON • *Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; UNC Neuroscience Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*
- ELIZABETH STEWART • *St. Jude Children's Research Hospital, Memphis, TN, USA*
- DIDIER SURDEZ • *INSERM U830, Équipe Labellisée LNCC, Genetics and Biology of Pediatric Cancers, PSL Research University, SIREDO Oncology Centre, Institut Curie Research Centre, Paris, France*
- MIWA TANAKA • *Division of Carcinogenesis, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan*
- OSCAR M. TIRADO • *Sarcoma Research Group, Oncobell Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain; CIBERONC, Carlos III Institute of Health (ISCIII), Madrid, Spain*
- MARCEL TRAUTMANN • *Division of Translational Pathology, Gerhard-Domagk-Institute of Pathology, Münster, Germany*
- ANDREI ZINOVYEV • *Institut Curie, PSL Research University, Paris, France; INSERM U900, Paris, France; CBIO-Centre for Computational Biology, Mines ParisTech, PSL Research University, Paris, France*

Part I

Molecular Biology Techniques Applied To Ewing Sarcoma



Chapter 1

Germline Variation and Somatic Alterations in Ewing Sarcoma

Mitchell J. Machiela and Thomas G. P. Grünewald

Abstract

Ewing sarcoma (EwS) is a rare bone or soft tissue tumor that occurs early in life and as such genetic variation is a major contributor to EwS risk. To date, genetic investigations have identified key somatic mutations and germline variants of importance for EwS risk. While substantial progress is being made in uncovering the genetic etiology of EwS, considerable gaps in knowledge remain. Herein, we provide a summary of methodological approaches for future genomic investigations of EwS. We anticipate this recommended analytical framework for germline and somatic investigations, along with genomic data from growing EwS case series, will aid in accelerating new genomic discoveries in EwS and expand knowledge of the genetic architecture of EwS.

Key words Ewing sarcoma, EWSR1-FL11, Genetics, Germline variants, SNPs, Somatic mutations, Translocation

1 Introduction

Ewing sarcoma (EwS) is a tumor of childhood, adolescence, and adulthood. As such, relatively few environmental exposures have accumulated during the life course to substantially affect risk of EwS and, to date, no exogenous exposures have been robustly associated with EwS risk [1, 2]. Rather, risk of EwS appears to be largely driven by genetic determinants [3]. Studies of germline genetic variation and acquired somatic alterations in EwS cases offer tremendous potential to better understand the etiologic mechanisms that lead to EwS development. The aim of this chapter is to provide a framework of common methods and approaches for performing genomic investigations of EwS.

Perhaps the largest motivator of genetic investigations of EwS is to find novel insights into EwS etiology. Whether these investigations identify inherited variants associated with EwS susceptibility or highlight common sets of acquired mutations shared across many EwS cases, such studies provide new clues into the underlying

genetic architecture of EwS. These genetic discoveries can provide important leads into promising research directions as well as accelerate advances into understanding the core regulatory networks key to EwS growth and survival. In addition, expanded knowledge of EwS genetic susceptibility can lead to improved genetic risk stratification and aid in identifying individuals or families at elevated risk of EwS for targeted screening or preventative measures.

Ample evidence suggests a genetic component to EwS. Well before modern genomic analyses of EwS, observational studies suggested a disparity in EwS incidence by ancestry in which higher rates of EwS were observed in populations of European ancestry [4, 5]. These differences in EwS incidence by ancestry are observed presently [6] with rates of EwS higher in Europeans (1.5 cases per million individuals) as compared to Asians (0.8 cases per million individuals) and Africans (0.2 cases per million individuals). Remarkably, African Americans retain the relatively low incidence of EwS despite being exposed to many of the same environmental factors as Americans of European ancestry. Additionally, there have been incidental reports of EwS clustering in families, suggesting rare instances of familial aggregation of EwS, although evidence is limited [7]. Interestingly, EwS does not appear to be an outcome of well-described inherited cancer predisposition syndromes [8], although further studies are needed to better investigate potential relationships.

The most well-characterized genomic component of EwS has been the characterization of pathognomonic somatic gene fusions seen in EwS cases. Evidence for recurrent EwS fusions between chromosome 11 and 22 were first reported in the 1988 [9] followed by evidence of gene fusions in 1992 by Delattre et al. [10]. Since these initial discoveries, chromosomal translocations involving a member of the FET gene family fused with an ETS family transcription factor have been demonstrated to be potent transcriptional dysregulators essential for the transformation of normal cells into EwS [11]. The most common of these fusions are *EWSR1-FLI1* translocations seen in approximately 85% of EwS cases. These fusions typically result in an oncoprotein with a DNA-binding domain of the FLI1 component combined with a strong transcriptional enhancer. Apart from these fusions, EwS tumors are relatively mutationally quiet and characterized by very few (mostly nonrecurrent) other mutations. Examples of identified recurrent somatic mutations include mutations in *STAG2*, *TP53*, and *CDKN2A* at frequencies of around 20%, 5%, and 3%, respectively [3].

Inherited genetic variation has also been identified to be an important contributor to EwS susceptibility. Recent genome-wide association studies have identified common germline variation associated with EwS susceptibility [12, 13]. Interestingly, these EwS susceptibility loci are enriched near polymorphic GGAA

microsatellite repeat sequences known to play an important role in *EWSR1-FLII* binding. As an example, a variant in the chromosome 10q21.3 region, rs79965208, associated with EwS susceptibility was found to disrupt a GGAA microsatellite sequence, thereby disrupting *EWSR1-FLII* binding to the region and impacting expression of *EGR2*, a gene with impacts on proliferation and clonogenicity in EwS cells [14]. Additionally, germline sequencing studies suggest a role of rare inherited inactivating mutations in DNA damage repair genes in approximately 13% of tumors, for example *BRCA1* mutations, although their pathogenic roles remain unclear [15].

Substantial progress in identifying genetic determinants of EwS has been made in the past decade. Despite recent advances, the genetic etiology of EwS remains poorly characterized. As series of EwS cases continue to expand and new genomic approaches are developed, our understanding of the genetic etiology of EwS is anticipated to expand. This chapter outlines key methods and approaches for future genomic investigations.

2 Identifying a DNA Source for Investigation

A key requirement for genetic studies of EwS is genomic material for investigation. While studies on both DNA and RNA provide insights into EwS etiology, this chapter will focus on current analysis approaches for studying variation in DNA important for EwS risk. Studies of DNA can generally be divided into two main categories: (1) studies on inherited germline genetic variation and (2) studies on acquired somatic variation. This chapter will expand on methods essential for both germline and somatic DNA analyses as they relate to EwS susceptibility.

The genomic question of interest often dictates the sample source for DNA collection. For example, studies of somatic mutational profiles of EwS tumors will require DNA derived from excised primary tumor specimens or metastases. Current approaches enable DNA extraction from both fresh frozen tumor tissue as well as archival tumor tissue stored in formalin fixed paraffin embedded (FFPE) blocks. While fresh frozen tissue is preferred as DNA yield and quality are superior, many common genomic applications can be modified to include input DNA from FFPE tissue. Knowing which genotyping or sequencing technologies will be used in a genomic investigation will dictate whether DNA from fresh frozen or FFPE tissue is appropriate. For example, some sequencing approaches may require additional quality control filters or expanded purification or amplification steps before FFPE derived DNA can be used as an input.

Germline genetic investigations of EwS generally focus on nontumor tissue. Typical sources of DNA include whole blood, peripheral blood mononuclear cells, buccal epithelial cells, hair follicles, or bone marrow, as these tissues generally reflect germline DNA as inherited from parents. As EwS tumors are mutationally quiet, some studies have used tumor-derived DNA as a substitute for germline DNA [12]. While such studies have produced important new advances with respect to inherited variation important for EwS susceptibility, care should be taken to ensure genomic findings of interest reflect germline variation and not variation introduced by somatic mutations.

3 Considerations When Selecting EwS Cases and Controls

One of the largest hurdles to investigating the genetic risk factors for EwS is the limited availability of EwS cases. As EwS is a rare tumor, acquiring large and well-powered series of EwS cases can be challenging. Most current genomic investigations of EwS are on only a few hundred EwS cases at most. While these sample sizes are large for a rare tumor, limitations in sample size impact power to detect key genetic relationships important to EwS etiology. The development of international EwS collaborations and consortia as well as sharing genomic data produced by these large efforts is vital for amassing large series of rare EwS cases for genomic investigation.

Selecting the type of EwS cases to obtain genetic material from is an important consideration in the design of genetic studies. As EwS is a tumor predominantly observed in individuals of European ancestry [16], most genetic studies of EwS focus on individuals with a European genetic background. The enrichment of EwS among Europeans, however, does not preclude EwS studies in other ancestries. For example, admixture mapping studies of EwS cases with admixed ancestral backgrounds could provide unique opportunities for localizing European risk haplotypes in a background of non-European ancestry.

Identifying a comparable control population is also essential for getting unbiased estimates for genetic effects. Researchers should ensure control series appropriately match the ancestral population from which the EwS cases arose. Many centers enrolling EwS cases fail to collect appropriate controls, often requiring controls to be obtained from independent series of cancer-free individuals. Furthermore, large age-matched populations of genotyped or sequenced disease-free controls often do not exist for the age range of individuals commonly affected by EwS. To circumvent the lack of collected controls for EwS genomic investigations, many researchers use cancer-free controls as a genomic comparison for EwS cases. Principal component matching can be used to ensure

EwS cases are genetically well-matched to controls [13]; however, researchers exercising this approach for genomic analysis should be cautious of any potential survival biases that may be introduced from using adult controls for a cancer that predominantly occurs in childhood, adolescence and early adulthood.

A final design consideration is the type of EwS cases included in the analysis. While EwS is well-defined by hallmark fusions, heterogeneity still exists with respect to age of diagnosis, tumor site, disease stage, and histological response to neoadjuvant therapy [3]. Investigators should be attentive to these individual and tumor characteristics when selecting cases for inclusion in genomic investigations. Imbalances in inclusion of certain subsets of EwS cases could result in findings that are not generalizable to all EwS cases. Alternatively, focusing on one particular subset of EwS cases could reduce heterogeneity in the analysis and aid in identifying novel genetic contributors to risk with application to a select subset of EwS cases.

4 Selection of Optimal Genomic Technologies

The analytical question of interest also dictates which genomic technologies are well-suited to address the question. Often the choice of genomic technology is a balance between cost, sample size, and the detail or amount of genomic data produced. Two general approaches are used to produce genetic data: (1) array-based approaches and (2) sequencing-based approaches; although other approaches also exist (e.g., TaqMan genotyping), they will not be the focus of discussion.

Studies of germline variation and EwS can use both array-based as well as sequencing-based approaches. In general, studies interested in systematically scanning for associations with low-frequency and common variation use genotyping arrays. Currently available genotyping arrays include several hundred thousand to a few million markers on a commercially available platform and when intensity data is analyzed produce genotypes for the germline variants on the array. There are several companies that offer these arrays and different arrays can be chosen based on focus (e.g., oncology-related arrays, multiethnic arrays). These arrays are generally designed to have a backbone of variants that tag common and low-frequency variation across the genome. While the array only includes a small subset of all germline variation, the design enables for identification of regions or specific haplotypes that may be associated with EwS risk. In most cases, the tagging variants themselves are not the functional variants, but are highly correlated with germline variation that may be important for disease risk. At present, genotyping arrays are typically much cheaper than exome-wide

or genome-wide sequencing approaches, although improvements in next-generation sequencing approaches and methods may soon decrease these cost differences.

Genotyping arrays perform well for identifying associations with common or low-frequency germline variation, but for germline investigations focused on rare variants, sequencing approaches are generally best. While Sanger sequencing [17] with chain terminating nucleotides still has application for regional sequencing experiments, most contemporaneous sequencing studies use massively parallel, short-read next-generation sequencing (NGS) approaches. There are several commercial vendors offering NGS products ranging in scope from single genes or regions of interest to exome-wide or genome-wide applications. Most NGS approaches require the construction of a library of fragmented DNA with adaptors to first be created followed by sequence-by-synthesis on a specially designed slide or flow cell within the sequencer. The sequencers produce large data files with sequence reads that are later aligned to a reference genome and eventually called for germline variants. Recently, long-read sequencing methods, in comparison to the short-read sequencing mentioned above, have also become widely available. These long-read sequencing approaches enable researchers to get better resolution of regional haplotype structure as well as improved sequencing of regions that are more difficult to sequencing by short-read sequencing (e.g., GGAA microsatellite regions that are of great importance to EwS pathogenesis [14, 18]). Costs for NGS is a function of the number of samples to sequence, the sequencing region(s) of interest and the target read depth. NGS is typically more expensive to perform than array-based genotyping.

Similar to investigations of rare germline variants, investigations of somatic mutations in EwS tumor tissue almost exclusively use NGS approaches. Regions of interest for genomic investigation are targeted for sequencing with studies ranging in scope from single genes to exome-wide or genome-wide investigations. Genotyping arrays are not suited to studying somatic variants as variation included on these arrays are germline variants and unable to detect somatic mutations. Recent expansion of methods that use raw genotyping array intensity data, however, have been fruitful in identifying large somatic structural copy number and copy neutral variation in a variety of tissue types [19, 20].

5 Key Considerations in EwS Germline Association Analyses

Equally important to generating genomic data is correctly analyzing this data. EwS germline association studies need to first ensure that genotype data is of sufficient quality before proceeding. A good rule of thumb is to run duplicate samples for a small

percentage of study participants (e.g., a mixture of 5% cases and controls) to ensure genotyping concordance rates in excess of 99% between samples. Each sample and genotyped variant should also be checked for high genotype completion rates (generally >95%). Other recommended sample quality checks include checking for sex concordance, unexpected duplicates or related samples as well as samples that have heterozygosity rates that substantially deviate from expected. Most studies also perform Hardy–Weinberg equilibrium tests for each variant to ensure genotypes among controls do not substantially deviate from expected population proportions. These data filtering steps are easy to perform using genomics software such as in PLINK [21].

Genotype imputation is common practice to fill in missing genotypes as well as infer genotypes of variants that are not directly genotyped or sequenced. As detailed above, only high-quality genotypes and samples are used as input for imputation. When performing imputation, sample genotypes are first phased to determine underlying haplotypes and then a reference panel of similar ancestry to the study population is used to fill in missing genotypes. A popular phasing program is EAGLE2 [22] and a popular imputation program is minimac [23], both of which can be implemented using an online pipeline through the Michigan Imputation Server [23]. Common imputation reference panels include the TOPMed project [24], the Haplotype Reference Consortium (HRC) [25] and the 1000 Genomes Project [26]. As with genotyping data, imputed data is filtered to remove poorly imputed variants. Generally, imputed variants with imputation scores or R^2 values in the range of 0.3–0.5 or below are removed from downstream analyses.

Genotype imputation permits germline data generated from different platforms or labs to be merged for large-scale analysis. While merging is attractive for expanding EwS sample size, caution should be taken to ensure each merged set has EwS cases and a representative set of controls and additionally has performed uniform quality control and imputation procedures. In addition, confounding by ancestry (also referred to as population stratification) is a key factor of concern for many genetic association studies. A simple way to reduce the concern for confounding by ancestry is to restrict to individuals of one ancestry for study (e.g., European ancestry as EwS is overrepresented among Europeans); although, studies that include multiple ancestries or admixed individuals can have unique advantages but usually require advanced analytical approaches. EwS germline studies can use self-reported ancestry or compute ancestry from genomic information (e.g., using SNPWEIGHTS [27]). Confounding by ancestry can still occur even when restricting to one ancestral group due to subtle differences in allele frequencies among individuals within a population. One common adjustment approach is a dimension reduction method called principal component analysis (PCA) in which

principal components representing dimensions of ancestry are used to adjust out the effects of residual population stratification [28]. Commonly used software for PCA analysis includes EIGENSTRAT [28] and PLINK [21].

Many software suites have been developed for performing genetic association analyses. Popular tools include PLINK [21], SNPTEST [29], BOLT-LMM [30], and SAIGE [31]. These packages have varying options and approaches for performing tests of genetic association. After association tests are performed data is commonly plotted in Manhattan plots, plots of genomic position by $-\log_{10} p$ -value, to identify regions of association with EwS. The qqman R package [32] offers good functionality for generating Manhattan plots. Following association analyses, signals are further interrogated by disentangling LD patterns [33, 34] and looking at functional annotation [35, 36] to inform future functional studies.

6 Analysis Elements for NGS Projects

The advent of NGS has resulted in expansive datasets of short-read sequencing data. Most modern sequencing platforms generate output files of many thousands of raw reads based on the collection of data during the sequencing process (e.g., Illumina HiSeq 4000 or NovaSeq). Initial quality control steps are first performed on these raw reads to remove poor quality reads as well as unnecessary adaptor or barcoding data. A key next step is the alignment of this short-read data to locations in the genome that these reads most optimally map to; although, not all reads will align to a location in the human genome. Current algorithms and alignment software have made substantial progress in speeding up genomic alignment of short read sequencing data. A common sequencing aligner is the Burrows–Wheeler Aligner available through the BWA-MEM software package [37]. For some EwS sequencing projects, more advanced alignment methods such as local realignment and deduplication of reads may be critical for generating high-quality sequencing alignments.

After sequencing alignment is complete, variant calling can be performed. There are typically two different objectives of variant calling: (1) detecting germline variation or (2) detecting somatic variation. The type of variant calling will depend on the design goal and area of interest of your study. Germline variant calling is most commonly used in studies of EwS susceptibility. Common germline callers include the Genome Analysis Toolkit (GATK) Haplotype-Caller [38], Free Bayes [39], and Strelka2 [40]. Newer, more sensitive approaches to calling germline variants also incorporate joint calling of multiple aligners into consensus calls with higher accuracy. In addition to standard germline callers, callers specialized

for detecting larger forms of structural variation are also coming online. While many of these callers are still in their infancy, platforms such as Lumpy [41], SnowmanSV, Manta [42], and SvABA [43] are helpful for detecting structural genetic variation.

Somatic variant calling is commonly used to look for mutations that arise in EwS tumor tissue. Generally, somatic variant callers require a study design of matched tumor and normal samples from the same individuals or a panel of “normal” samples for differentiating somatic variants from inherited germline variation. Mutect2 [44] and Strelka2 [40] are two widely used somatic variant callers. Detection of somatic variants by joint calling of germline callers with a focus on detected variants with variant allele frequencies that substantially deviate from 50:50 Mendelian proportions is also possible in instances where no germline DNA is available for the EwS patient. Likewise, germline structural variant callers can also be used to call somatic structural variation; but as with detection of germline structural variants, calling somatic structural variants from short read sequencing data can be challenging.

Linking detected genetic variation to EwS risk or progression is the primary aim of most EwS sequencing projects. Many of the analytic approaches mentioned above for array-based association studies also have relevance for sequencing-based association studies. A unique aspect of sequencing studies is that in addition to common variation, sequencing studies have the ability to detect low-frequency and rare variants not tagged well by array-based studies. These rare variants present challenges for statistical analysis as it is unlikely many EwS cases harbor exactly the same rare variant. Rather, it may be the case that several EwS cases carry different rare variants that all happen to cluster in a single gene, suggesting an important function of the gene for EwS. To investigate such instances of rare variant clustering specialized association tests have been developed that focus on aggregations of rare variants. An early developed rare variant aggregation test is SKAT [45], that allows for defined aggregations of rare variants while efficiently adjusting for other covariates. Other flavors of SKAT (e.g., SKAT-O [46]) as well as other extensions to burden tests have also become available for testing aggregations of rare variants from sequencing data [47].

A common goal of both germline and somatic EwS investigations is classifying associated variants based on predicted functional consequence. Most currently available variant annotation tools focus on variants within the coding region of genes as changes in nucleotide sequence in these regions are expected to have the greatest impact on protein function and subsequently EwS risk. Common variant annotation tools include COSMIC [48], ClinVar [49], CADD [50], REVEL [51], and Meta-SVM [52]. While these tools are helpful in predicting impact of protein coding changes or identifying associations with other diseases, the level of support for

predictions vary considerably and as such these tools should only be used to provide insights into potential relationships between genetic variants and disease risk. In most cases, further functional investigations and more refined association studies are needed to conclusively link genetic variants to altered EwS risk or response.

7 Final Considerations for EwS Genetic Investigations

This chapter on EwS genetic methods and protocols is intended to serve as a starting point for genomic investigations of EwS. Depending on the nature of specific EwS genomic investigations, further methodologic needs or more advanced analytic approaches may be required. It is recommended to collaborate with a diverse team of collaborators on EwS genomic investigations with individual expertise in areas such as biology, epidemiology, genetics, bioinformatics, biostatistics, and oncology.

Additionally, while many parts in the preceding sections made a distinction between somatic and germline investigations, recent EwS genomic investigations indicate germline–somatic interactions are likely key contributors to the genetic etiology of EwS [13, 14, 53]. As such, integrative studies that combine data on germline variation and somatic mutations along with rich genomic annotation data from gene expression, chromatin organization, DNA methylation and ChIP-seq are likely to provide important new insights for expanding knowledge of EwS risk.

References

1. Moore LE, Gold L, Stewart PA et al (2005) Parental occupational exposures and Ewing's sarcoma. *Int J Cancer* 114:472–478
2. Valery PC, McWhirter W, Sleigh A et al (2002) Farm exposures, parental occupation, and risk of Ewing's sarcoma in Australia: a national case-control study. *Cancer Causes Control* 13:263–270
3. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primer* 4:5
4. Fraumeni JF, Glass AG (1970) Rarity of Ewing's sarcoma among U.S. Negro children. *Lancet Lond Engl* 1:366–367
5. Jensen RD, Drake RM (1970) Rarity of Ewing's tumour in Negroes. *Lancet Lond Engl* 1:777
6. Jawad MU, Cheung MC, Min ES et al (2009) Ewing sarcoma demonstrates racial disparities in incidence-related and sex-related differences in outcome: an analysis of 1631 cases from the SEER database, 1973–2005. *Cancer* 115:3526–3536
7. Joyce MJ, Harmon DC, Mankin HJ et al (1984) Ewing's sarcoma in female siblings. A clinical report and review of the literature. *Cancer* 53:1959–1962
8. Rahman N (2014) Realizing the promise of cancer predisposition genes. *Nature* 505:302–308
9. Turc-Carel C, Aurias A, Mugneret F et al (1988) Chromosomes in Ewing's sarcoma. I. An evaluation of 85 cases of remarkable consistency of t(11;22)(q24;q12). *Cancer Genet Cytogenet* 32:229–238
10. Delattre O, Zucman J, Plougastel B et al (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 359:162–165
11. Sankar S, Lessnick SL (2011) Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet* 204:351–365
12. Postel-Vinay S, Véron AS, Tirode F et al (2012) Common variants near TARDBP and EGR2

- are associated with susceptibility to Ewing sarcoma. *Nat Genet* 44:323–327
13. Machiela MJ, Grünewald TGP, Surdez D et al (2018) Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility. *Nat Commun* 9:3184
 14. Grünewald TGP, Bernard V, Gilardi-Hebenstreit P et al (2015) Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite. *Nat Genet* 47:1073–1078
 15. Brohl AS, Patidar R, Turner CE et al (2017) Frequent inactivating germline mutations in DNA repair genes in patients with Ewing sarcoma. *Genet Med Off J Am Coll Med Genet* 19:955–958
 16. Worch J, Cyrus J, Goldsby R et al (2011) Racial differences in the incidence of mesenchymal tumors associated with EWSR1 translocation. *Cancer Epidemiol Biomarkers Prev* 20:449–453
 17. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467
 18. Gangwal K, Sankar S, Hollenhorst PC et al (2008) Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proc Natl Acad Sci U S A* 105:10149–10154
 19. Loh P-R, Genovese G, Handsaker RE et al (2018) Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature* 559:350–355
 20. Machiela MJ, Zhou W, Sampson JN et al (2015) Characterization of large structural genetic mosaicism in human autosomes. *Am J Hum Genet* 96:487–497
 21. Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575
 22. Loh P-R, Danecek P, Palamara PF et al (2016) Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* 48:1443–1448
 23. Das S, Forer L, Schönherr S et al (2016) Next-generation genotype imputation service and methods. *Nat Genet* 48:1284–1287
 24. Brody JA, Morrison AC, Bis JC et al (2017) Analysis commons, a team approach to discovery in a big-data environment for genetic epidemiology. *Nat Genet* 49:1560–1563
 25. McCarthy S, Das S, Kretzschmar W et al (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 48:1279–1283
 26. Auton A, Abecasis GR, Altshuler DM et al (2015) A global reference for human genetic variation. *Nature* 526:68–74
 27. Chen C-Y, Pollack S, Hunter DJ et al (2013) Improved ancestry inference using weights from external reference panels. *Bioinforma Oxf Engl* 29:1399–1406
 28. Price AL, Patterson NJ, Plenge RM et al (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909
 29. Marchini J, Howie B, Myers S et al (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39:906–913
 30. Loh P-R, Tucker G, Bulik-Sullivan BK et al (2015) Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* 47:284–290
 31. Zhou W, Nielsen JB, Fritsche LG et al (2018) Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet* 50:1335–1341
 32. Turner S (2018) qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. In: *J. Open Source Softw.* <https://doi.org/10.1101/005165>. Accessed 27 Nov 2019
 33. Machiela MJ, Chanock SJ (2015) LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinforma Oxf Engl* 31:3555–3557
 34. Machiela MJ, Chanock SJ (2018) LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. *Bioinforma Oxf Engl* 34:887–889
 35. Ward LD, Kellis M (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 40:D930–D934
 36. Boyle AP, Hong EL, Hariharan M et al (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 22:1790–1797
 37. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma Oxf Engl* 25:1754–1760
 38. DePristo MA, Banks E, Poplin R et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498

39. Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. *ArXiv12073907 Q-Bio*
40. Kim S, Scheffler K, Halpern AL et al (2018) Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods* 15:591–594
41. Layer RM, Chiang C, Quinlan AR, Hall IM (2014) LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol* 15: R84
42. Chen X, Schulz-Trieglaff O, Shaw R et al (2016) Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 32:1220–1222
43. Wala JA, Bandopadhyay P, Greenwald N et al (2018) SvABA: genome-wide detection of structural variants and indels by local assembly. *Genome Res* 28(4):581–591. <https://doi.org/10.1101/gr.221028.117>
44. Cibulskis K, Lawrence MS, Carter SL et al (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 31:213–219
45. Wu MC, Lee S, Cai T et al (2011) Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* 89:82–93
46. Lee S, Emond MJ, Bamshad MJ et al (2012) Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* 91:224–237
47. Lee S, Abecasis GR, Boehnke M, Lin X (2014) Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 95:5–23
48. Tate JG, Bamford S, Jubb HC et al (2019) COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res* 47:D941–D947
49. Landrum MJ, Lee JM, Riley GR et al (2014) ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* 42:D980–D985
50. Rentzsch P, Witten D, Cooper GM et al (2019) CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* 47:D886–D894
51. Ioannidis NM, Rothstein JH, Pejaver V et al (2016) REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* 99:877–885
52. Kim S, Jhong J-H, Lee J, Koo J-Y (2017) Meta-analytic support vector machine for integrating multiple omics data. *BioData Min* 10
53. Musa J, Cidre-Aranaz F, Aynaud M-M et al (2019) Cooperation of cancer drivers with regulatory germline variants shapes clinical outcomes. *Nat Commun* 10:1–10



Western Blot Analysis in Ewing Sarcoma

Aruna Marchetto and Laura Romero-Pérez

Abstract

Western blot is an experimental method used to analyze protein expression. In Ewing sarcoma, as in many other diseases, Western blot provides information about the level of protein expression in different cell conditions, in comparison with other tissues or upon induced molecular changes. Based on the specific pattern of protein expression of the tissue, as well as on the characteristics of the protein of interest, the antibodies and protocol of Western blot may be modified according to different specifications. Here we describe some of these peculiarities in frame of Ewing sarcoma field.

Key words Western blot, Ewing sarcoma, Protein expression

1 Introduction

Western blot is a common molecular method used to detect the expression of one or more proteins of interest. This technique was described in 1981 by W. Neal Burnette [1], who named it based on the similarity with previously described methods for the transference of DNA (Southern blot) or RNA (Northern blot) to a nitrocellulose membrane.

1.1 *Western Blot, the Technique*

In Ewing sarcoma, as in many other fields of biological and clinical research, Western blot is used to quantify the expression of specific proteins in different cell lines or even tissues, in an exploratory way or upon different experimental events. Examples of those experimental assays are: knocking out/down/in of certain genes; the induction of changes in a regulatory pathway; the inhibition/induction of protein expression by a specific drug treatment; the induction of protein modifications such as phosphorylation/dephosphorylation; or to detect target proteins obtained from previous assays such as immunoprecipitation.

In summary, the protocol of a Western blot is based on an electrophoretic process by which the proteins contained in a buffered lysate are charged in a gel (sodium dodecyl sulfate-

polyacrylamide gel, SDS-PAGE, **Note 1**) and separate according to their mass and electric charge (*see Note 2*). The pattern of separated proteins is then transferred to a membrane (*see Note 3*) to be then blocked (*see Note 4*) and incubated with the specific antibody (polyclonal or monoclonal; **Note 5**) against the protein of interest. Both the blocking step and the use of specific antibodies aim at minimizing nonspecific binding to other proteins.

Washing and incubation with a secondary antibody (reactive to the primary antibody host species) is required to amplify the signal through a conjugated enzyme such as horseradish peroxidase (HRP) or alkaline phosphatase (AP).

The most commonly used techniques to visualize the results are developing films or transillumination. By using transillumination, fluorescent, colorimetric or more frequently, chemiluminescent reagents can be detected depending on the desired sensitivity (*see Note 6*).

To analyze the results, the normalization of the expression of the protein among the different conditions is performed with the use of a loading control. It is based on the use of an antibody against a housekeeping protein, meaning a protein constitutively expressed in all cells of interest, which size should be different from the proteins that will be analyzed. In order to estimate the size of the bands (which is the same as to identify the detected proteins) a protein ladder must be charged in the gel. Color markers for certain sizes facilitate the identification of the protein of interest.

In the recent years, single cell-resolution Western blotting has been described [2, 3], even at the subcellular level [4]. This application is especially useful when avoiding background signal of surrounding cells is required. However, to our knowledge, these interesting methods have not been applied to the field of Ewing sarcoma yet probably due to the particularity of this kind of tumor showing “small round cells” with a little proportion of cytoplasm. Although this may present some difficulties, Ewing sarcoma researchers are encouraged to set up the conditions of these new methods in this field.

1.2 Ewing Sarcoma Specifications

Ewing sarcoma is characterized by recurrent balanced chromosomal translocations between *EWSR1* gene (22q12.2) and *FLII* (11q24.3) in around 85% of cases, but also *ERG* (10–15%), *ETVI*, *ETV4*, or *FEV* can fuse to *EWSR1* with phenotypically and clinically similar pattern? [5].

When the most common translocation is produced, exons 7 or 10 of *EWSR1* are fused to exon 6 (Type I), exon 5 (Type II), or exon 7 (Type IV) of *FLII*, or exons 6, 7, or 9 of *ERG*. Here the epitope of *EWSR1* becomes unrecognizable, as it maps between amino acids 2–19 at the N-terminus in humans. Other fusion points have been described to produce the same loss of recognition of *EWSR1* [6]. This fact explains why most specific antibodies used

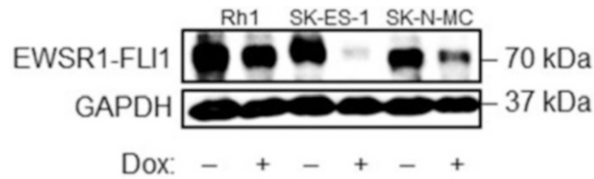


Fig. 1 Western blot showing endogenous EWSR1-FLI1 (–) in Ewing sarcoma cell lines (Rh1, SK-ES-1, and SK-N-MC), which harbor a dox-inducible shRNA against *EWSR1-FLI1* (+)

to identify Ewing sarcoma cells from a total cell extract are specific for either FLI1 or ERG (as the most common partners of *EWSR1*), and CD-99 (a cell surface glycoprotein highly expressed in Ewing sarcoma; [7]).

Importantly polyclonal antibodies against the transcription factor FLI1 may not only detect endogenous FLI1 but also FLI1 as part of the fusion protein. To ensure the detection of the fusion protein and not the endogenous FLI1, it is important to confirm the size of the obtained band and to use specific controls based on the presence or absence of EWS-FLI1 (Fig. 1) [8]. Figure 1 shows that upon the doxycycline induction of a shRNA against *EWSR1-FLI1* in Ewing sarcoma cell lines, the band of approx. 68 kDa corresponding to the fusion protein expected size [9] is reduced or completely lost.

Folpe et al. described that among normal tissues only endothelial cells and small lymphocytes express FLI1 [10]. This is nowadays debatable, since it is registered in The Human Protein Atlas [11, 12] that cells of the brain, lung, breast and female reproductive system or renal, urinary, and male reproductive system also express FLI1.

2 Materials

Prepare all the solutions for the Western blot analysis with distilled water and keep all the buffers at room temperature (RT) except for the RIPA buffer and the 1× Blotting buffer (4 °C). Never use sodium azide neither for the buffer nor for the antibody solutions as it is known to interfere with the HRP-coupled secondary antibody.

2.1 Reagents

1. *RIPA Buffer*: Mix 150 mM NaCl (3.5 g), 0.5% sodium deoxycholate (5 g), 50 mM Tris-HCl, pH 8, 1% Triton X-100 (4 mL), 0.1% SDS (1 g) to a total volume of 1 L H₂O and adjust the pH 8. Add every time fresh to the sample buffer: 1 mM Na₃VO₄ and 100 μL/mL Protease Cocktail inhibitor. Store at 4 °C.

2. *Bradford reagent*: Dilute in a ratio 1:5 the Bradford reagent with H₂O.
3. *BSA standard*: Dilute Bovine Serum Albumin (BSA) (0.125 µg, 0.25 µg, 0.5 µg, 0.75 µg, 1 µg, and 2 µg) with H₂O to generate the standard curve to determine protein concentration afterward.
4. *10× TBS buffer, pH 7.3*: Mix 24 g Tris base, 88 g NaCl and fill up to 1 L with H₂O. Adjust the pH 7.3. Store at RT.
5. *1× TBST (0.1%)*: 100 mL 10× TBS buffer, 900 mL H₂O and 1 mL Tween-20 Store at RT.
6. *10× Running/Blotting buffer (R/B)*: Mix 30 g Tris base and 144 g glycine to 1 L H₂O. Store at RT.
7. *1× Running buffer*: Mix 100 mL 10× R/B buffer with 10% SDS (100 g) and fill up with 900 mL H₂O. Store at RT.
8. *1× Blotting buffer*: Mix 100 mL 10× R/B buffer with 200 mL Methanol and fill up with 700 mL H₂O. Store at 4 °C.
9. *4× Loading dye*: Mix 312 mM Tris-HCl, pH 6.8, 10% SDS (100 mg), 50% (500 µL) glycerol and add a bit of Bromophenol Blue for the color. Prepare aliquots and store at -20 °C. Add fresh 250 mM DTT to 1 mL loading dye and keep up to 3 months in the fridge.
10. *Blocking buffer*: Mix 5% non-fat milk (2.5 g) or 5% BSA (2.5 g) with 20 mL of 1× TBST buffer, mix and fill up to 50 mL. Prepare fresh every time.
11. *Ammonium persulfate (APS)*: Dissolve 10% APS (100 mg) in 1 mL of H₂O. Prepare aliquots and store at -20 °C.
12. *Sodium dodecyl sulfate (SDS)*: Dissolve 10% SDS (100 mg) in 1 mL H₂O. Store at RT.
13. *N, N, N, N'-Tetramethyl-ethylenediamine (TEMED)*: Store at RT.
14. *30% bisacrylamide (37.5:1)*: Store at 4 °C.
15. *1.5 M Tris-HCl, pH 8.8*: Dissolve 45.4 g Tris-HCl in 100 mL H₂O and fill up to 250 mL. Adjust at pH 8.8 with HCl and store at RT.
16. *1 M Tris-HCl, pH 6.8*: Dissolve 30.3 g Tris-HCl in 100 mL H₂O and fill up to 250 mL. Adjust at pH 6.8 with HCl and store at RT.
17. Mouse monoclonal antibody against FLII (MRQ-1, Cell Marque). Observed band at ~70 kDa (primary antibody, **Note 7**).
18. Mouse monoclonal antibody against GAPDH (sc-32233, Santa Cruz). Observed band at ~37 kDa (primary antibody).
19. Polyclonal anti-mouse IgG (H + L) coupled to HRP (secondary antibody, **Note 8**).

2.2 Equipment

1. Spectrophotometer for measurement of protein absorption.
2. *SDS-polyacrylamide gel*: glass plates (15 mm), combs (10, 15 chambers), prestained molecular weight standards.
3. *Immunoblotting*: Nitrocellulose membranes, blotting paper.

3 Methods

All steps should be performed at room temperature, if not otherwise specified. This protocol is mainly focused on protein detection from Ewing sarcoma cell lines.

3.1 Sample Preparation

1. Prepare 100 μ L RIPA buffer supplemented with Protease cocktail inhibitor and Phosphatase inhibitor (Na_3VO_4) per 6-well sample and lyse the cells (*see Note 9*) on ice with a cell scraper (Fig. 2a).
2. Transfer the cell lysate into a reaction tube and incubate for 30 min on ice while shaking (Fig. 2b).
3. Spin down at $11,000 \times g$ for 15 min and transfer the supernatant into a new reaction tube. For long-term storage, keep lysates at -80°C .
4. Measure protein concentration (*see Note 10*), add $1 \times$ Loading dye supplemented with 250 mM DTT (*see Note 11*).
5. Incubate at 95°C for 10 min to disrupt S-S boundaries (*see Note 12*, Fig. 2c).
6. Shortly spin down the lysate to bring down the condensate. Cool down on ice and keep on ice or freeze.

3.2 SDS-PAGE Gel Electrophoresis

1. Resolving gel (7%): Add 3.7 mL of H_2O and 2 mL of 1.5 M Tris-HCl, pH 8.8 with 1.8 mL of 30% bisacrylamide in a 50 mL falcon. Mix thoroughly. Add 75 μ L of 10% SDS, 40 μ L of 10% APS and 10 μ L TEMED to the previous solution and mix well (*see Note 13*).
2. Cast the gel within the glass plates with 1.5 mm thickness (10 cm \times 7.5 cm). To avoid bubbles overlay the surface of the resolving gel with isopropanol. Wait until the remaining solution in the falcon is solidified and proceed with the stacking gel preparation.
3. Stacking gel (3%): Add 3.5 mL of H_2O and 860 μ L of 1 M Tris-HCl, pH 6.8 with 500 μ L of 30% bisacrylamide in a 5 mL falcon. Mix thoroughly. Add 47.5 μ L 10% SDS, 10% APS and 10 μ L TEMED to the previous solution and mix.
4. Remove the isopropanol by pouring the isopropanol from the glass plates into a paper towel, pour the stacking gel on top of

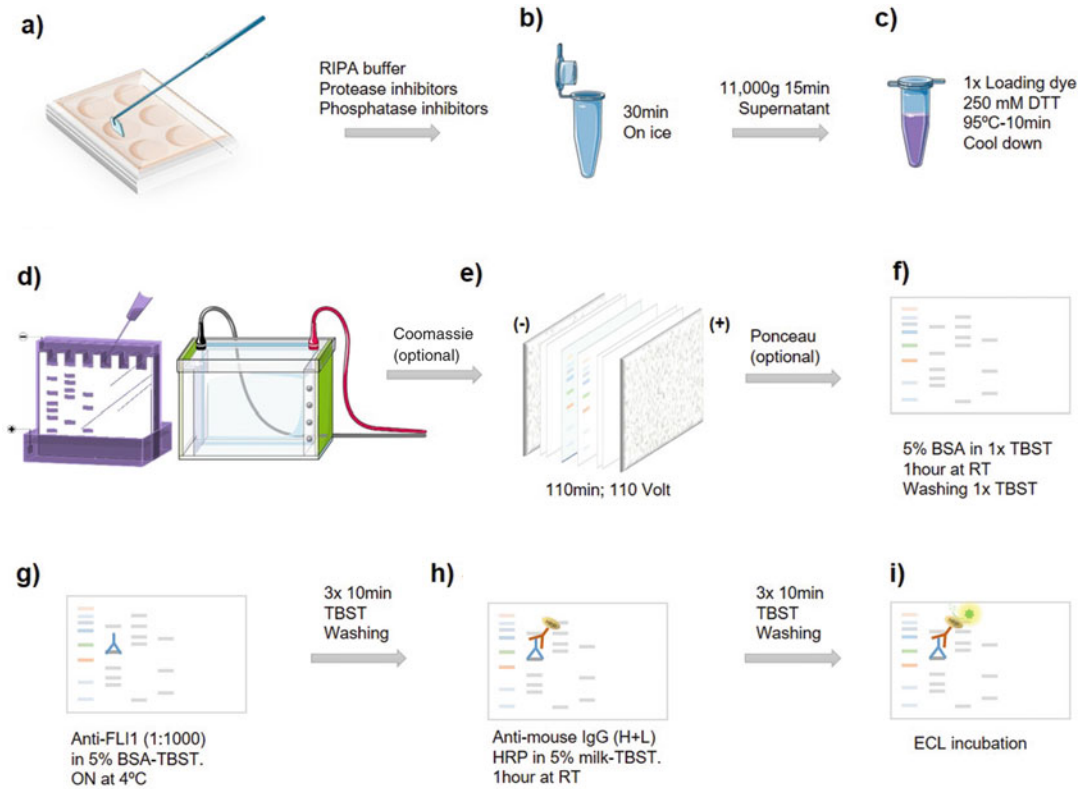


Fig. 2 Scheme of Western blot procedure. *RT*: room temperature, *ON*: overnight. (a) Cultured cells, (b) lysis, (c) denaturation, (d) gel electrophoresis, (e) protein transfer, (f) blocking, (g) primary antibody, (h) secondary antibody, (i) detection

the resolving gel and insert the corresponding 1.5 mm comb avoiding bubbles.

5. Insert the casted glass plates within the Western blot device and fill up with 1 × Running buffer until marked in the device.
6. Load the prepared samples (10–20 µg protein concentration) and the protein ladder on the outer position and in the middle of the lanes in the corresponding pockets of the gel (*see Note 14*).
7. Electrophorese at 90 V until the dye front reaches the bottom of the gel (Fig. 2d).

3.3 Protein Transference

1. Take the glass plates with the gel out of the device and pry the glass plates with the help of a spatula. Remove the stacking gel from one side of the glass plate with a spatula and cut out the upper part of the stacking gel. Be always aware of the order in which the samples have been loaded for future reference (*see Note 15*).

2. Cut the nitrocellulose membrane (*see Note 16*) and the blotting paper according to the size of the gel (10 cm × 7.5 cm) and bath them in cold 1× Blotting buffer.
3. Prepare the wet-blot device (*see Note 17*) on a magnet stirrer and load half of the chamber with a frozen thermal pack and 1× Blotting buffer.
4. Load the elements in the following exact order while immersed in 1× Blotting buffer to avoid bubble formation (Fig. 2e):
Cathode (−) sponge–two blotting papers–gel–nitrocellulose membrane–two blotting papers–sponge–anode (+).
5. Squeeze the air bubbles out of the *sandwich* with help of the spatula and insert it into the wet-blot device.
6. Let the proteins transfer for 110 min at 110 V while mixing on the magnetic stirrer (*see Note 18*).

3.4 Antibody Staining

1. Take out the *sandwich* from the wet-blot device and block the membrane by incubating for 1 h at RT with Blocking buffer (5% BSA dissolved in 1× TBST) while shaking (*see Note 19*, Fig. 2f).
2. Wash the membrane briefly with fresh 1× TBST buffer.
3. Cut the membrane according to the size of the proteins of interest and incubate each part with the corresponding primary antibody: anti-FLI1 (1:1000) in 5% BSA dissolved in 1× TBST and anti-GAPDH (1:2000) in 5% non-fat milk dissolved in 1× TBST overnight at 4 °C while rotating (Fig. 2g).
4. Wash the nitrocellulose membranes three times with fresh 1× TBST for 10 min while shaking.
5. Secondarily incubate with anti-mouse IgG (H + L) HRP (1:3000) with 5% non-fat milk TBST for 1 h (*see Note 20*, Fig. 2h).
6. Wash the membranes three times with fresh 1× TBST for 10 min each while shaking.
7. Drain the wash buffer from the membrane prior to incubating the membrane with ECL (time: 1 min). The protein amount is detected via chemiluminescence as HRP cleaves luminol (Fig. 2i).

3.5 Analysis

The analysis should be performed by comparing the intensity of the band of the protein of interest with respect to the intensity of the band of the housekeeping protein. The ratio should be calculated for the same specimen and also by comparison with the ratio corresponding to the rest of conditions or specimens analyzed in the same assay. The intensity estimation can be performed with different software such as Image J (NIH) [13].

4 Notes

1. The percentage of acrylamide will depend on the size of the protein of interest. Normally, 4–7% is recommended for proteins of 100–300 kDa and up to 10–12.5% for proteins <100 kDa.
2. Time, voltage and amperage for electrophoresis and transference can be modified. Our recommendation is to perform electrophoresis constant at 90 V and transference at constant 110 V for at least 110 min depending on the protein size.
3. The membrane can be made of nitrocellulose or PVDF. In the case of PVDF membranes, a methanol-activation step is required.
4. Blocking solution can be prepared with BSA or non-fat powder milk diluted in $1 \times$ TBST. Milk is commonly used because it is not expensive and it is ready to use. However it can affect the detection of phosphorylated proteins since milk contains casein, which is itself a phosphoprotein. Additionally, it should not be chosen if using avidin–biotin detection systems, since milk contains biotin. Conversely, BSA is only one protein, so there is less possibilities of cross-reaction and the results are normally clearer. However, it must be noted that it is more expensive than milk and it is not recommended in case of lectin probes as BSA contains carbohydrates that may increase background.
5. A specific antibody is the one recognizing only the epitope of interest. This is the definition of a monoclonal antibody. However, it is not always possible to find monoclonal antibodies for the proteins of interest. Therefore, a polyclonal antibody could be a good opportunity to perform the analysis (having carried out previous controls for similar proteins).
6. Sometimes the use of developing films may offer higher sensitivity for problematic antibodies.
7. Upon translocation the epitope of EWSR1 becomes unrecognizable. Thus, the primary specific antibody detects the fusion protein via binding to the C-terminal of FLI1 (~50 kDa) or ERG (~55 kDa) in EWSR1-FLI1 or EWSR1-ERG (~68 kDa) fusion positive tumors, respectively. Polyclonal antibodies may bind to more than one epitope of a target as compared to monoclonal antibodies which are highly specific and only bind to one epitope.
8. Primary antibodies detect a specific antigen (monoclonal) or multiple epitopes (polyclonal). Secondary antibodies bind primary antibodies based on their host species. This indirect labelling requires from coupling secondary antibodies to a

dye or an enzyme, most commonly HRP, which is measured via fluorescence or chemiluminescence, respectively.

9. We recommend to collect cells at ~80% confluence.
10. The total protein concentration can be measured by the Bradford protein assay [14]. First, dilute BSA and the Bradford reagent with H₂O as described above. In a 96-well plate, 99 μ L of diluted Bradford reagent was added to the corresponding amount of samples or BSA Standard (minimum two technical replicates). Subsequently, 1 μ L of each sample/BSA standard were added on top of the 99 μ L Bradford reagent (also including a “blank” sample with only Bradford reagent). This 96-well plate was incubated 5 min while shaking in the dark and subsequently absorption was measured at 595 nm with the spectrophotometer.
11. 1 \times Loading dye supplemented with 250 mM DTT is stable up to 3 months in the fridge without storing it at -20°C .
12. Cool down samples and subsequently centrifuge them after heating up at 95°C for 10 min. This helps to bring down condensate and to avoid that the sample clumps being not completely loaded on the gel.
13. Mix first the components H₂O, Tris-HCl, 30% bisacrylamide and then add 10% SDS, 10% APS, and TEMED. It will reduce clumps within the gel and lead to a more homogeneous protein detection.
14. By applying a bit of the 1 \times Loading dye, the gel chambers get visible and it is easier fill them with the samples.
15. Coomassie staining can be performed to confirm the presence, enough amount and integrity of proteins in the cell lysate. It is optional because these features are assumed in regular Western blot, however, when some previous conditions have been modified, it is usually performed to check if it makes sense to continue with the rest of the procedure. The staining solution is made with 0.25 g of Coomassie Blue, 10 mL of glacial acetic acid and 90 mL of MeOH-H₂O (1:1). The gel should be incubated in the staining solution with shaking for at least 30 min. The solution should be removed with water to check the protein bands and finally, a destaining solution should be added to eliminate the staining and incubated for 30–60 min (destaining solution is the same as staining solution but without the Coomassie Blue).
16. In order to keep orientation, cut or label with a pencil an edge of the membrane.
17. There is another possibility of rapidly transferring proteins from polyacrylamide gels to nitrocellulose or PVDF

membranes in 30–60 min by semi-dry-blot. This is only recommended for proteins of 10–100 kDa size.

18. Ponceau staining is used to detect protein bands after the transference. It may reveal also if there were bubbles in the transference process and whether they may affect the detection of the band of interest, or it can be even used as a first loading control. 5 mL of commercial Ponceau staining solution is enough to cover the membrane. Incubate in a shaker for 5 min at RT and wash with distilled water 2 or 3 times while shaking. After checking the bands of interest, Ponceau staining can be removed with distilled water and a final step of $1 \times$ TBST.
19. The blocking solution with 5% BSA-TBST is sterile filtered with 0.2 μm filter to reduce background during membrane blocking and primary antibody incubation. Additionally, this step increases the half-life of the primary antibody dilutions which are stored in the fridge. Primary antibodies dilutions can be kept without sodium azide addition at 4 °C for up to 1 month. All secondary antibodies must be always freshly prepared. In order to reduce clumps, dissolve 5% non-fat milk powder first in half of the $1 \times$ TBST volume and then filling up the rest afterward.

Acknowledgments

We would like to thank the service of the online tool SMART Servier Medical Art for providing some of the images with which we built Fig. 2. We would like to thank the Institute of Pathology of Ludwig Maximilians University of Munich and the Director Prof. Dr. Thomas Kirchner for the installations where to develop this experimental method carried out in the laboratory of the Max-Eder Research Group for Pediatric Sarcoma Biology headed by Dr. Thomas Grünewald. LR-P's position is funded by the Dr. Leopold und Carmen Ellinger Foundation and the Gert and Susanna Mayer Foundation. AM's position is funded by the Matthias-Lackas-Stiftung.

References

1. Burnette WN (1981) "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate—polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem* 112(2):195–203
2. Kang CC, Yamauchi KA, Vlassakis J et al (2016) Single cell-resolution Western blotting. *Nat Protoc* 11(8):1508–1530
3. Duncombe TA, Kang CC, Maity S et al (2015) Hydrogel pore-size modulation for enhanced single-cell Western blotting. *Adv Mater* 28(2):327–334
4. Yamauchi KA, Herr AE (2018) Subcellular Western blotting of single cells. *Microsyst Nanoeng* 3:16079. <https://doi.org/10.1038/micronano.2016.79>

5. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4(1):5. <https://doi.org/10.1038/s41572-018-0003-x>
6. Sankar S, Lessnick SL (2011) Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet* 204(7):351–365
7. Kovar H, Dworzak M, Strehl S et al (1990) Overexpression of the pseudoautosomal gene MIC2 in Ewing's sarcoma and peripheral primitive neuroectodermal tumor. *Oncogene* 5(7):1067–1070
8. Selvanathan SP, Graham GT, Grego AR et al (2019) EWS-FLI1 modulated alternative splicing of ARID1A reveals novel oncogenic function through the BAF complex. *Nucleic Acids Res* 47(18):9619–9636. <https://doi.org/10.1093/nar/gkz699>
9. Wang M, Nilsson G, Carlberg M et al (1998) Specific and sensitive detection of the EWS-FLI1 fusion protein in Ewing's sarcoma by Western blotting. *Virchows Arch* 432(2):131–134
10. Folpe AL, Hill CE, Parham DM et al (2000) Immunohistochemical detection of FLI-1 protein expression: a study of 132 round cell tumors with emphasis on CD99-positive mimics of Ewing's sarcoma/primitive neuroectodermal tumor. *Am J Surg Pathol* 24:1657–1662
11. Uhlén M, Fagerberg L, Hallström BM et al (2015) Tissue-based map of the human proteome. *Science* 347(6220):1260419
12. Uhlen M, Oksvold P, Fagerberg L et al (2010) Towards a knowledge-based human protein atlas. *Nat Biotechnol* 28(12):1248–1250
13. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675
14. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254



Tissue Preservation and FFPE Samples: Optimized Nucleic Acids Isolation in Ewing Sarcoma

Laura Romero-Pérez and Thomas G. P. Grünewald

Abstract

Different methods have been described for the preservation of biopsy or resection samples. In the routine pathology, the cheapest and most commonly used is fixation of samples in formalin and embedding in paraffin (FFPE samples). This method preserves tissue samples for a very long time and is suitable for several specialized techniques such as fluorescence in situ hybridization (FISH) and immunohistochemistry, the latter being the most frequent and often the only additional method used for establishment of final diagnosis. However, in light of the growing need of next-generation sequencing and microarray technologies that are often very helpful to establish and/or confirm diagnoses in the field of pediatric sarcoma (including Ewing sarcoma), preservation of high-quality and quantity of nucleic acids (DNA/RNA) is desirable. Herein, we describe how to ideally preserve samples, as well as how to proceed to isolate nucleic acids for successful subsequent molecular assays with a special focus on Ewing sarcoma samples.

Key words Preservation, FFPE samples, DNA, RNA, Molecular assays, Ewing sarcoma

1 Introduction

In pathology, the main aim when receiving a biopsy or resection sample is to be able to offer an accurate diagnosis to the patient to be treated accordingly. The macroscopic analysis can reveal many characteristics of the tissue, but it needs to be assessed microscopically and/or by using different molecular tests. Accordingly, tissue specimens should be managed and preserved in optimal conditions to avoid cell degeneration and degradation of DNA, RNA, or proteins. Thus, the first step to offer a precise diagnosis is to maximize the care of the sample since the moment it is collected from the patient. As Ewing sarcoma (EwS) is a solid tumor, we will focus on the different ways of preserving tissue samples, although liquid biopsy is an emerging and promising source for diagnosis.

2 Materials

2.1 *Fresh Tissue Sample Preservation*

1. Extracellular fluid type preservation solution: PBS or culture medium (i.e., DMEM).
2. Ice.

2.2 *Frozen Tissue Sample Preservation*

1. OCT (Optimal Cutting Temperature) solution.
2. Cryomold.
3. Isopentane.
4. Freezer ($-80\text{ }^{\circ}\text{C}$).
5. Liquid nitrogen.
6. 4% paraformaldehyde (PFA): 4 g of PFA in 100 ml of phosphate-buffered saline, PBS.
7. 30% sucrose: 30 g of sucrose in 100 ml of PBS
8. Cryovials.
9. Cryostat.

2.3 *FFPE Tissue Sample Preservation*

1. 10% Neutral-buffered formalin. It is composed by 37% formaldehyde and 10% methanol in distilled water with the addition of a buffer, typically sodium phosphate [1]. It is usually purchased as a prepared solution (*see Note 1*)
2. 70% ethanol (70 ml of absolute ethanol, 30 ml of distilled H_2O)
3. 80% ethanol (80 ml of absolute ethanol, 20 ml of distilled H_2O)
4. 95% ethanol (95 ml of absolute ethanol, 5 ml of distilled H_2O)
5. Ethanol (absolute).
6. Xylene.
7. Paraffin wax.
8. Metallic molds.
9. Plastic cassette ($3 \times 2.5\text{ cm}$).

2.4 *Optimized DNA/RNA Isolation from FFPE Samples*

1. Hematoxylin and eosin staining.
2. FFPE block.
3. 2 mm biopsy punches
4. Scalpel.
5. FormaPure total kit (Beckman Coulter).
6. Thermoblock.
7. Centrifuge.
8. Magnet for Eppendorf tubes 1.5 ml.

9. Nanodrop.
10. Qubit fluorimeter and reagents.
11. Bioanalyzer and reagents.

3 Methods

3.1 *Fresh Tissue Sample Preservation*

In routine surgical pathology, fresh tissue samples are relatively uncommon. The logistics for fresh tissue specimens requires some dedicated personal to immediately transport and process the tissue specimens after the surgical resection. However, sometimes it is possible to obtain fresh tissue specimens and they represent the most valuable source of information from the tissue.

The specific advantages of fresh tissue specimens are related to opportunities that maintenance of cell viability provides. These opportunities include among others the generation of in vitro primary cultures and in vivo patient-derived xenografts (PDX) to analyze the behavior and functional characteristics of the cells constituting the sample (tumor cells and stromal cells). The combination of these two techniques may yield different information such as the evaluation of the evolution of functional cell features, the analysis of a specific treatment response, and the capability of metastasize to other parts of the body. When using the whole sample only for diagnosis, the main advantage is that almost no degradation of nucleic acids or proteins takes place at that moment, allowing many molecular tests.

1. Immerse the tissue specimen in an extracellular fluid (ECF)-type preservation solution: usually Ca^{2+} and Mg^{2+} free phosphate-buffered saline (PBS) or high-glucose Dulbecco's modified Eagle's medium (DMEM) [2] (*see Note 2*).
2. Transport the immersed tissue specimen at 4 °C or below, in a refrigerated container or in ice for the shortest time possible.

Since the availability of fresh tissue specimens are still very infrequent, the most standardized protocols are usually designed for frozen, and mostly FFPE tissue samples.

3.2 *Frozen Tissue Specimen Preservation*

As it was described for fresh samples, also preservation of frozen tissue specimens requires the fast intervention of the personal after surgical resection and the presence of some specific preservative solutions, devices, and freezing equipment. This should be done as quickly as possible after the tissue is disconnected from its native vascular supply to avoid artifacts, autolysis, or even effects in differential gene expression due to the 'warm ischemia' time [3] (*see Note 3*).

Different methods are commonly used depending on the estimated time of cold storage and the subsequent analysis planned.

3.2.1 Long-Term Cold Storage

For long-term cold storage with the estimation of likely future histologic and molecular tests, the presence of cryoprotective solutions is necessary (*see* **Note 4**).

1. Place the tissue sample in a cryomold (*see* **Note 5**).
2. Cover the sample by adding OCT solution (*see* **Note 6**).
3. Introduce the cryomold in an isopentane container (cooled to -80°C ; **Note 7**).
4. Once it becomes a frozen block, store in a -80°C freezer or liquid nitrogen tank for long term.

When necessary for diagnostic or research purposes, OCT embedded frozen blocks can be sectioned in a cryostat (*see* **Note 8**) for histologic evaluation by IHC, immunofluorescence or in situ hybridization (*see* **Note 9**), as well as for isolation of nucleic acids for molecular tests.

3.2.2 Snap Freezing in Liquid Nitrogen

Water can be frozen so rapidly that it does not have time to form crystals and it stays in a vitreous form, this is what happens when inducing snap freezing in liquid nitrogen. This method is only recommended for biopsies or very small resection samples (*see* **Note 10**).

1. Place the sample in a container filled with liquid nitrogen, allowing for the most surface contact (*see* **Note 11**).
2. Transfer the frozen sample to a cryovial and store it in a liquid nitrogen tank for future molecular tests (*see* **Note 12**).

Tissue preservation in any of the different forms of frozen samples or formalin-fixed paraffin-embedded (FFPE) samples should be chosen depending on the resources and the estimated applications.

3.3 FFPE Sample Preservation

The most common way of storing biological specimens is in form of FFPE samples. It is the most cost-efficient method because they can be stored at room temperature in nonsterile conditions and to be histologically reanalyzed many years after the first diagnosis with a relative low rate of degradation (except for RNA). FFPE samples work well for IHC and they are the best option for morphology analyses because of the fixation process. However, FFPE samples show some important disadvantages: preparation of FFPE samples implies the longest protocol for tissue preservation; it also involves toxicity due to reagents such as formalin; last, FFPE samples are usually the less suitable for molecular analysis since the fixation process induce cross-linking and degradation of nucleic acids.

Even so, the vast majority of cohorts in molecular pathology research are composed by FFPE samples, including most of the works published in the field of EwS.



Fig. 1 Scheme of the standard procedure for FFPE samples preparation

3.3.1 FFPE Sample Preparation

In this case, also the time between the resection of the sample and the fixation is crucial to avoid artifacts that are secondary to the processing. There are many different modifications of the same protocol to prepare FFPE samples, but automation of the main steps helps to eliminate variations affecting experimental results. Main steps are represented in Fig. 1.

1. Place the tissue sample in a cassette and introduce the tissue sample in formalin (*see Note 13*) for 24 h (maximum 48 h **Note 14**) in order to fixate the sample to provide rigidity, protection against autolysis or decomposition, and to respect cell architecture and tissue cell composition.
2. Remove all water present in the fixed sample by performing an alcoholic dehydration consisting in introducing the sample in fresh ethanol solutions at crescent concentration from two times 70% (60 min), 80% (60 min), 95% (60 min), and three times 100% (60 min) (*see Note 15*).
3. Introduce the tissue in xylene (60 min) \times 2–3 times (*see Note 16*).
4. Introduce the tissue in paraffin at approximately 58 °C during 60 min, \times two times to allow the paraffin wax to infiltrate the tissue (*see Note 17*).
5. Impregnate the definitive mold with a bit of molten paraffin, open the cassette with the tissue and place it in the mold letting a free frame of around 2 mm to be surrounded by paraffin.
6. Transfer the mold to a cold surface to let the paraffin solidify making some pressure and adding enough paraffin to cover the height of the cassette.
7. Once it is solidified, the FFPE block can be stored long term at room temperature.

3.3.2 FFPE Samples Issues for Molecular Analysis

Whereas FFPE samples retain most of tissue proteins expression that can be evaluated by IHC many years after the diagnosis, nucleic acids are highly degraded because of the crosslinking and chemical modifications induced by the fixation process. Even after having been described some formalin free tissue preservatives solutions, with less toxicity and considered as ‘nucleic acids friendly’ [4, 5], the reality is that the most common method used is still the

one described in the nineteenth century [6]. This makes necessary to adapt every molecular protocol for working with FFPE samples and to assume that a certain percentage of samples will show non-acceptable quality values, thus the original series size will be reduced.

For EwS diagnosis, sometimes it is necessary to use next generation techniques for which the integrity of nucleic acids should be the greatest possible. Herein we described how to improve the isolation DNA and RNA from FFPE samples to optimize molecular analysis outputs.

3.4 Optimized DNA/RNA Isolation from FFPE Samples

3.4.1 Tumor Cells Enrichment

The minimum percentage of tumor cells present in a sample to be considered as informative is around 70–80%, however EwS samples are sometimes limited in this regard. On the one hand, tumor cells are usually surrounded by normal cells, necrosis, lymphocytes, etc. On the other hand, many times tumor cells are included in a very small biopsy sample from a pediatric patient.

By performing microtome slides of the whole tissue and paraffin frame, most of the times the real percentage of tumor cells is much lower. For this reason, we recommend to perform microdissection of the tissue to specifically isolate the cells of interest. Microdissection can be performed in different ways, the most precise one is laser microdissection. However, routine laser microdissection is very time-consuming and the amount of cells obtained from each trial is very small. Instead, the following protocol is recommended:

1. Select the tumor cells enriched area (>70–80%) in the hematoxylin–eosin staining corresponding to each FFPE block (Fig. 2).
2. Localize this area in the FFPE block.
3. Isolate cylinders of tissue by puncturing with a 2-mm biopsy punch (Fig. 2).
4. Cut the tissue core into very small pieces with the scalpel, or triturate/homogenize, to facilitate the subsequent digestion with Proteinase-K required for DNA/RNA isolation.

This is very useful when working with small biopsies from pediatric patients such as those with suspected EwS.

3.4.2 Deparaffination

Nucleic acids contained in FFPE samples are inevitably highly degraded due to the traditional method of fixation in formalin. The procedures of deparaffination and isolation of nucleic acids from FFPE samples should be the less aggressive possible to offer clean and of high-quality results.

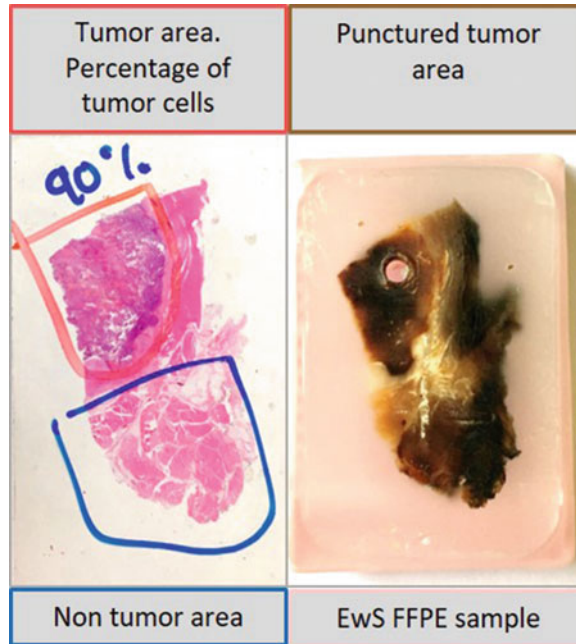


Fig. 2 Tumor cell enrichment by localizing and puncturing the tumor area of an Ewing sarcoma (EwS) FFPE sample. Tumor area is marked in red and the percentage of tumor cells is annotated. The corresponding FFPE block has been punctured with a 2 mm biopsy punch in the selected area

1. Use mineral oil (from the FormaPure total kit; **Note 18**) to dissolve the paraffin from the tissue with the help of the temperature for a short period of time (80 °C, 5 min; **Note 19**) according to the manufacturer's instructions.
2. Add the lysis buffer, centrifuge and add Proteinase-k according to the manufacturer instructions. An interface will be formed and the mineral oil will stay in the upper part without contacting the sample.

3.4.3 DNA and RNA Isolation

Many kits are commercially available for the isolation of DNA and/or RNA from FFPE samples. Most of these kits include steps of centrifugation to make the sample be filtered through a column, retaining the nucleic acids in the filter and letting the washing solution pass to the lower part to be discarded. In this way, some of the nucleic acids remain retained in the filter even after the last elution step, obtaining less amount of nucleic acids than the one potentially contained in the sample.

On the contrary, we propose to use isolation methods based on nucleic acids binding to magnetic beads (SPRI beads), like the FormaPure total kit. It has been optimized for using before downstream next generation sequencing (NGS) and genotyping assays, which are sometimes required for complex sarcoma cases. In

addition, this kit offers the possibility of isolating both DNA and RNA from the same FFPE sample, which is an advantage when the amount of tumor tissue is reduced.

1. For RNA isolation a binding solution is added to immobilize the nucleic acids to the surface of the magnetic beads. Contaminants are rinsed away using a simple washing procedure, DNA is removed from the sample, and RNA is again immobilized on the surface of the beads before eluting with water (*see Note 20* for enzymatic treatment).
2. For DNA isolation, RNA is removed from the sample, and a binding solution is added to immobilize the DNA to the surface of the magnetic beads. Contaminants are rinsed away using a simple washing procedure, and the DNA is eluted with water (*see Note 20*).

3.4.4 Measurement of the Concentration and Quality of Nucleic Acids

Nanodrop is usually chosen as the fastest and easiest method to check at the same time the concentration and estimated quality of isolated DNA or RNA. We recommend to pay attention to both, 260/280 and 260/230 absorbance ratios to calculate an index of quality (which should be next to 1, and separately each of them should be close to 2), in addition to the curve pattern. However, this method is not the most precise and it could just help to decide if a sample should be discarded or not from the cohort.

For more accurate nucleic acids concentration estimation, we recommend the Qubit fluorimeter rather than UV-absorbance quantification methods, being able to distinguish between DNA and RNA in the same sample. It can quantify the concentration from very low concentrated samples (10 pg/ μ l) to 1 μ g/ μ l, but it is not able to detect contaminants as it can be revealed or estimated by the NanoDrop curve. Moreover, Qubit measurements require to perform a protocol consisting in preparing a working solution where to dilute each sample, as well as the nucleic acids standards to build a concentration curve. Sometimes, very low or highly concentrated samples may offer an out-of-range result, for which it will be necessary to repeat the process.

Since Qubit provides very precise information about nucleic acids concentration, but it does not provide any information about the quality of the sample, we recommend to additionally use the Bioanalyzer for subsequent NGS analysis. This method, based in electrophoresis, may result expensive because of the use of especial DNA- or RNA-chips, reagents and device, but it is very informative. It provides information about sizing, quantitation, integrity and purity from DNA and RNA, expending minimal sample volumes. Information about the amount of DNA/RNA fragments and the average size is crucial when working with FFPE samples. For example, DV₂₀₀ values (the percentage of fragments >200 nucleotides) are of especial interest for sequencing assays.

4 Notes

1. Methanol is a stabilizer that slows the process by which the polymerized PFA would precipitate in water. Buffers (usually phosphate buffers, but also magnesium carbonate, citrate or Tris) are added to formalin to maintain the pH close to 7 and to avoid the oxidation of formaldehyde to formic acid, which would produce pigment granules precipitation [6].
2. The method of fresh tissue preservation is similar to the one usually used for organ transplantation. Preservation solutions have been classified as: intracellular fluid (ICF)-type solution with a high potassium–low sodium ratio; or extracellular fluid (ECF)-type preservation solution with a low-potassium–high-sodium ratio, being this last the one recently demonstrated to improve the functionality of the organs. The most commonly used ECF-type solutions are PBS or DMEM, although it has been recently described another particular ECF-type preservation solution with better results in terms of viability of tissue stem cells for longer period of time [2].
3. From the frequent type of samples received in the pathology routine, frozen samples are the most valuable sources for molecular assays. If these tissue samples cannot be processed immediately after the resection, they can be immersed in a tube with *RNAlater* and to store overnight at 2–8 °C in an upright position, to ensure the conservancy of the RNA [7]. To do so, the size should not exceed 0.5 × 0.5 cm in height and width. After overnight incubation, samples can be stored in *RNAlater* at room temperature (RT), 4 °C or –20 °C, for some time, but it is much better to store directly in liquid nitrogen or –80 °C for long term.
4. The use of cryopreservative solutions helps to avoid the damage of tissue due to ice crystal formation during freezing. Ice crystals would alter the morphology, damage cell membranes, and finally affect histologic evaluation. The appearance at microscope of a specimen with ice crystals will be like a perforated surface of tissue, losing the architecture of potentially informative areas.
5. When working with a frozen block, the side that will be first sectioned is the one touching the bottom of the cryomold, so the sample should be placed accordingly and properly orientated. While adding OCT, try to avoid the formation of air bubbles.
6. Instead of the only used of OCT, other methods include a previous step of fixation in 4% paraformaldehyde (PFA),

followed by cryoprotection in sucrose (usually 30% in phosphate buffer). This allows to use a slower freezing method but generates fixative cross-linking, making necessary to perform antigen retrieval before immunohistochemistry (IHC).

7. It is not recommended to use methanol, ethanol or acetone instead of isopentane, because when OCT contacts any of these reagents, it is not able to freeze down properly.
8. A cryostat is a special microtome maintaining constant temperature ($-10\text{ }^{\circ}\text{C}$ and below) during the whole process, so the freeze-thaw cycles should be taken into account, but the degradation is not as extreme as the one caused by freezing and thawing repeated times at RT.
9. OCT is a water-soluble mixture of glycols and resins, which does not leave any residues that may interfere during staining procedures. Usually, as no cross-linking is produced by any fixative, antigen retrieval is not required and epitopes are shown in an almost native state.
10. Even if liquid nitrogen seems to be the best option for snap freezing because it does not mix with the tissue, there are some specific problems related with the boiling effect in contact with the tissue at RT. This effect generates a vapor barrier next to the tissue surface inducing a slow penetration of cold, becoming the rate of freezing imprecise. This, together with the fact that biological tissues are poor thermal conductors, make possible that a thermal gradient can be established from the surface to the core of the tissue during the freezing process, especially if sample is not very thin. Thus, this method is only recommended to freeze biopsies or very small resection samples.
11. Although the method of surrounding the sample with powdered dry ice could be also an option, it is less recommended. Even to introduce the sample in a freezer is possible but any of these methods can freeze the tissue quickly enough.
12. For molecular tests liquid nitrogen snap-frozen samples can be stored in cryovials for long term. The only limitation is that in this case, the whole sample is thawed for DNA/RNA/protein isolation, so no different procedures or isolations will be performed in the future from the same sample.
13. When fixating a sample it is important to take into account the fixative solution chosen, the time of fixation and the thickness and properties of the tissue. Fixatives can be divided in four groups including aldehydes, oxidizing agents, alcohol-based and metallic group of fixatives [8]. Formalin, made with formaldehyde, can penetrate into the tissue at an average of 1 mm/h [1], depending on the tissue type. Once inside the tissue, the reaction of cross-linking between formaldehyde and proteins

starts and triggers the formation of stable methylene bridges, which provides harshness to the tissue.

14. Usually, 24 h is enough for the formalin fixation reaction, depending on the thickness of the tissue and the temperature [8], but sometimes 36–48 h may be necessary, increasing the fragmentation of the nucleic acids by over fixation.
15. Since formalin is composed by a proportion of water, which is immiscible with paraffin, a step of dehydration should be performed.
16. Alcohols are either not miscible with paraffin, so they should be cleared from the sample with xylene.
17. There are different protocols applying different paraffin mixtures and melting points temperatures (usually from 55–63 °C). Paraffin wax is solid at RT, but once melted, it can infiltrate the tissue.
18. Contrary to the standardized methods, we strongly recommend not to use organic solvents such as xylene for the deparaffination. In addition to the effect on the sample (if the washing steps with ethanol cannot completely remove it), the hazards of xylene for the individual are important if this is inhaled. It is very well documented to be toxic mainly at the level of central nervous system, but also in lungs, liver and kidney, blood, skin, and other systems [9].
19. The use of Ultra-Turrax is the best option to homogenize the tissue after standard xylene-mediated deparaffination when the sample is immersed in lysis buffer from standard kits. In case of using mineral oil for deparaffination, we do not recommend the use of Ultra-Turrax because of the bubbles problem generated when mixing the lysis buffer and the oil.
20. 80% ethanol for washing steps should be freshly prepared to ensure efficient cleaning of the samples. Before adding DNase-I or RNase-A, respectively, be sure that these treatments will not interfere with downstream analysis as happens for fusion analysis by RNA-seq in case of EwS.

References

1. Hewitt SM, Lewis FA, Cao Y et al (2008) Tissue handling and specimen preparation in surgical pathology: issues concerning the recovery of nucleic acids from formalin-fixed, paraffin-embedded tissue. *Arch Pathol Lab Med* 132:1929–1935. <https://doi.org/10.1043/1543-2165-132.12.1929>
2. Suzuki T, Ota C, Fujino N et al (2019) Improving the viability of tissue-resident stem cells using an organ-preservation solution. *FEBS Open Bio* 9(12):2093–2104. <https://doi.org/10.1002/2211-5463.12748>
3. Dash A, Maine IP, Varambally S et al (2002) Changes in differential gene expression because of warm ischemia time of radical prostatectomy specimens. *Am J Pathol* 161:1743–1748. [https://doi.org/10.1016/S0002-9440\(10\)64451-3](https://doi.org/10.1016/S0002-9440(10)64451-3)

4. Vincek V, Nassiri M, Nadji M, Morales AR (2003) A tissue fixative that protects macromolecules (DNA, RNA, and protein) and histomorphology in clinical samples. *Lab Invest* 83:1427–1435
5. Stanta G, Mucelli SP, Petrera F et al (2006) A novel fixative improves opportunities of nucleic acids and proteomic analysis in human archive's tissues. *Diagn Mol Pathol* 15:115–123
6. Fox CH, Johnson FB, Whiting J, Roller PP (1985) Formaldehyde fixation. *J Histochem Cytochem* 33:845–853. <https://doi.org/10.1177/33.8.3894502>
7. Florell SR, Coffin CM, Holden JA et al (2001) Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol. *Mod Pathol* 14:116–128. <https://doi.org/10.1038/modpathol.3880267>
8. Thavarajah R, Mudimbaimannar VK, Elizabeth J et al (2012) Chemical and physical basics of routine formaldehyde fixation. *J Oral Maxillofac Pathol* 16:400–405. <https://doi.org/10.4103/0973-029X.102496>
9. Kandyala R, Raghavendra SPC, Rajasekharan ST (2010) Xylene: an overview of its health hazards and preventive measures. *J Oral Maxillofac Pathol* 14:1–5. <https://doi.org/10.4103/0973-029X.64299>



Liquid Biopsies in Ewing Sarcoma

Manuela Krumbholz and Markus Metzler

Abstract

Liquid biopsies enable noninvasive therapy monitoring in patients with solid tumors. Specific serum markers such as proteins, hormones, or enzymes released from tumor cells or in response to tumor growth can be used for quantification of the tumor burden. However, only a fraction of pediatric tumors has none of these serum markers, but tumor-specific genetic alterations represent reliable alternatives. Here we describe a method for using genomic fusion sequences as liquid biopsy markers in Ewing sarcoma patients.

Key words Liquid biopsy, Serum marker, Genomic fusion sequences, Ewing sarcoma, Cell-free circulating DNA, Droplet digital PCR, Double-quenched probes

1 Introduction

Liquid biopsies in Ewing sarcoma enable noninvasive therapy monitoring in addition to imaging techniques, that is, X-ray, magnetic resonance tomography (MRT), or positron emission tomography-computed tomography (PET-CT) [1].

Tumor-specific genomic fusion sequences (e.g., *EWSR1-FLII*; present in ~85% of Ewing sarcoma patients or *EWSR1-ERG*; present ~10% of Ewing sarcoma patients) represent reliable serum markers for quantification of tumor growth [2, 3]. Theoretically, they can be detected in circulating tumor cells (CTCs), exosomal vesicles (tumor derived RNA) or as cell-free circulating tumor DNA (ctDNA) isolated from patient's blood samples [4]. However, only very low numbers of CTCs in the background of millions of hematopoietic cells are detectable in patients with advanced disease stages [5] and fusion transcript (mRNA) positive exosomal vesicles have only been quantified in cell culture and in Ewing sarcoma mouse models in detectable copy numbers so far [6, 7].

In contrast, genomic *EWSR1-FLII* or *EWSR1-ERG* ctDNA fusion sequences have recently been presented as reliable serum markers in Ewing sarcoma patients [1, 8]. Quantification of ctDNA copy numbers enables a more frequent therapy monitoring

and a better evaluation of imaging data which is an important step toward personalized therapy optimization and improved relapse diagnostics.

Here we describe a detailed protocol for plasma sample collection, isolation of cell-free circulating DNA from frozen plasma samples and quantification of patient-specific ctDNA copy numbers. Due to the high fragmentation of ctDNA and the large background of non-tumor-derived cell-free DNA copy numbers, primer-probe assays for therapy assessment must be well optimized and applied in a highly sensitive quantification technique like digital PCR. Furthermore, we recommend the use of double quenched probes for better differentiation between negative and positive samples using quantitative PCR.

2 Materials

2.1 Plasma Sample Collection and Storage

1. EDTA tubes.
2. Centrifuge.
3. -80°C freezer
4. 1.5 ml reaction tubes.

2.2 Isolation of Cell-Free Circulating DNA

1. Centrifuge.
2. Kits for isolation of cell-free nucleic acids:
 - (a) NucleoSpin Plasma XS Kit (Macherey-Nagel).
 - (b) QIAamp MinElute ccfDNA Kit (Qiagen).
 - (c) QIASymphony Circulating DNA Kit (Qiagen) for isolation with a QIASymphony SP Instrument.
3. PBS buffer (phosphate-buffered Saline).

2.3 Quantification of ctDNA in Ewing sarcoma Patients

1. Primers-probe assay for quantification of the tumor-specific fusion sequence.
2. Primers-probe assay for quantification of a single copy control gene.
3. QX200 droplet generator (Bio-Rad).
4. QX200 droplet reader.
5. Heat sealer.
6. Pierceable foil for heat seal.
7. PCR thermocycler.
8. Electronic multichannel pipette.
9. ddPCR Supermix for probes (no dUTP).
10. Droplet generator oil for probes.
11. DG8 cartridges and gaskets.

12. ddPCR droplet reader oil.
13. 96-well plate
14. 27G needles
15. Quantasoft Analysis Software.
16. Nuclease-free water.
17. Tumor DNA (genomic DNA) for positive control.
18. DNA from healthy individuals (genomic DNA and/or cell-free circulating DNA) for negative control.

3 Methods

3.1 Plasma Sample Collection and Storage

1. Collect blood sample in EDTA-tubes and start processing within the next 6 h (*see Note 1*).
2. Centrifuge EDTA-tubes at $1200 \times g$ for 10 min at room temperature.
3. Carefully transfer the plasma into 1.5 ml reaction tubes (*see Note 2*).
4. Optional: Centrifuge plasma samples at $11,000 \times g$ for 3 min at room temperature to remove residual cells (*see Note 3*).
5. Store plasma samples at -80°C .

3.2 Isolation of Cell-Free Circulating DNA

1. Thaw plasma samples at room temperature.
2. Isolate cell-free DNA from plasma using the appropriate commercially available kit (*see Note 4*).
3. Store ctDNA samples at -20°C .

3.3 Quantification of ctDNA in Ewing Sarcoma Patients

Patient-specific genomic *EWSR1-FLII* or *EWSR1-ERG* fusion sequences are used as highly tumor-specific markers for quantification of the ctDNA in plasma samples (*see Note 6*). For quantification of minimal amounts of ctDNA copy numbers high-sensitive digital PCR techniques have to be applied.

1. For each patient, design an individual primer-probe assay for the quantification of the ctDNA level by locating the probe directly on the fusion site, and primers at the 5' and 3' side of the genomic breakpoint. The ideal fragment length of the resulting amplification product should be 80–100 bp (*see Note 7*).
2. Use double quenched probes as they allow a better differentiation between positive and negative samples (*see Fig. 1*).
3. For droplet digital PCR (ddPCR) use ddPCR Supermix for Probes (no dUTP) to perform primer-probe base quantification of amplicons (*see Notes 8 and 9*).

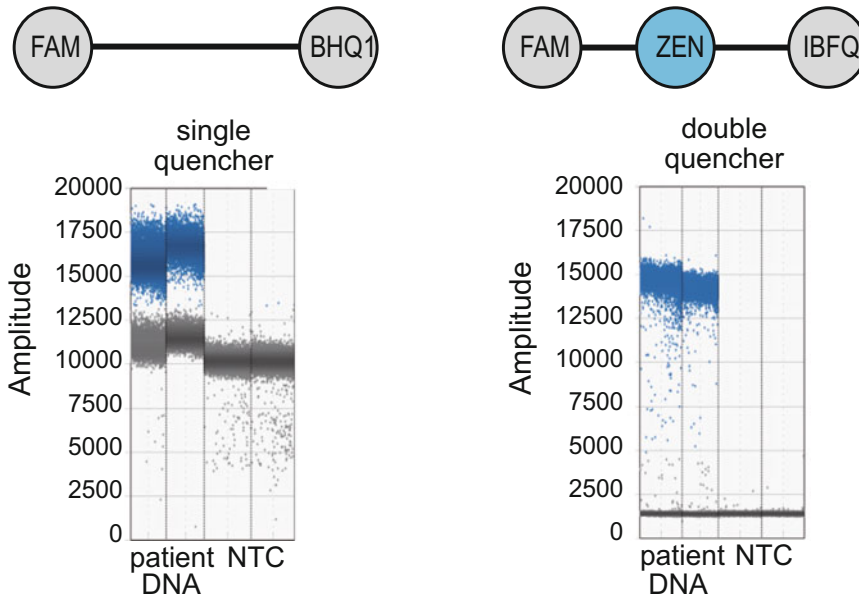


Fig. 1 Double quenched probes reduce background signal in quantitative PCR assays. Amplitudes of amplicon positive (patient DNA) and negative (nontemplate control, NTC) samples using identical droplet digital PCR assays but differently quenched probes

4. Pipet the reaction mix according to the manufacturer's instructions (*see Note 10*).
5. Prepare a positive control reaction mix using genomic tumor DNA isolated from tumor biopsy (*see Note 11*).
6. Incubate the DNA at 95 °C for 10 min.
7. Prepare negative control reaction mixes using human genomic DNA or cell-free plasma DNA of healthy individuals and nontemplate control (nuclease-free water).
8. Denature human genomic DNA at 95 °C for 10 min.
9. Before the sample cartridge is placed in the droplet generator remove all air bubbles using 27G needles.
10. After droplet generation, carefully transfer the droplets into a 96-well plate using an electronic multichannel pipette at minimum speed.
11. Perform PCR reaction according manufacturer's instructions.
12. Incubate the 96-well plate overnight at 4 °C to increase the numbers of final accepted droplets (*see Note 12*).
13. Measure samples at the droplet reader.
14. Copies of genomic fusion sequences can be calculated per ml of plasma or per copies of a single copy control gene (e.g., albumin) (*see Note 13*).

4 Notes

1. Prolonged storage time of blood samples at room temperature results in a lysis of white blood cells and a large increase of wild-type DNA fragments in the plasma. The higher background of wild-type DNA will consequently reduce the sensitivity of subsequent assays for quantification of tumor-specific ctDNA fusion fragments.
2. Plasma samples should be stored in small aliquots to avoid repeated thawing–freezing cycles during use. For reliable ctDNA quantification plasma samples can be thawed and refroze only once.
3. This step is recommended if isolation of cell-free circulating DNA will be performed with column based isolation kits. Residual cells or cell debris that could clog the column will be removed.
4. Various commercially kits are available. We recommend the use of NucleoSpin Plasma XS Kit (Macherey-Nagel) for small plasma volumes ($\leq 500 \mu\text{l}$) and the QIAamp MinElute ccfDNA Kit (Qiagen) for larger plasma volumes ($500 \mu\text{l}$ – 5 ml). For high-throughput sample preparation ($500 \mu\text{l}$ – 4 ml) use the QIASymphony Circulating DNA Kit with the QIASymphony SP (Qiagen) instrument.

Use the kits according manufacturer’s instructions with following modifications:

(a) NucleoSpin Plasma XS Kit (Macherey-Nagel)

Use the “High sensitivity protocol for the isolation of DNA from plasma”.

- For final elution of DNA pipet $30 \mu\text{l}$ instead of $20 \mu\text{l}$ elution buffer to the bottom of the column and incubate column with loaded elution buffer for 5 min at room temperature before centrifugation (see **Note 5**).

(b) QIAamp MinElute ccfDNA Kit (Qiagen).

- Plasma volumes $< 1 \text{ ml}$ (minimum $500 \mu\text{l}$) can be used. Fill up samples with PBS solution to a final volume of 1 ml .
- To remove the supernatant from the magnet beads place the tube with the magnetic bead solution into a magnetic rack and remove the supernatant with a thin glass pipette. After addition of $200 \mu\text{l}$ Bead Elution Buffer to the bead pellet, vortex the beads and briefly centrifuge the tube to remove drops from the rim.

- (c) QIAasympphony Circulating DNA Kit (Qiagen).
- Plasma volumes <2 ml (minimum 500 μ l) can be used. Fill up the samples with PBS solution to a final volume of 2 ml or 4 ml.
 - Adjust the protocol of the QIAasympphony SP that cell-free DNA will be finally extracted with 60 μ l Elution Buffer.
5. Incubation at room temperature for 5 min will increase DNA yield. 30 μ l elution buffer are required to finally obtain 20 μ l DNA solution.
 6. The tumor-specific genomic fusion sequence is individual for each patient. For identification of the genomic fusion site use nested multiplex long-range PCR assays [1, 9] or next generation sequencing technologies. DNA isolated from fresh or fresh frozen cryopreserved tumor biopsies are most suitable for detection of genomic breakpoints. Direct identification of genomic fusion sequences from cell-free circulating plasma DNA is only possible if proportion of ctDNA is high enough (>20%).
 7. Cell-free circulating DNA is highly fragmented, most fragments are 150–180 bp long [10]. Hence, using primer-probe assays resulting in small-length amplicons (80–100 bp) increase the probability to detect tumor-specific fragments.
 8. Although different digital PCR systems are available, we used the droplet digital PCR system QX200 (Bio-Rad) because 48 samples can be analyzed in one run, it is easy to handle, measurements achieve a high sensitivity, and costs per sample are moderate.
 9. Do not perform a quantification based on ddPCR EvaGreen Supermix as the maximum amount of template DNA that can be used per reaction is much lower, resulting in lower assay sensitivity.
 10. If low levels of ctDNA are expected include a maximal amount of template ctDNA (7 μ l) per reaction. Due to the high fragmentation of ctDNA an initial restriction digest of the DNA sample is not necessary.
 11. If using DNA isolated from fresh or fresh-frozen cryopreserved tumor biopsies DNA has to be fragmented by heat denaturation.
 12. Incubation of samples overnight at 4 °C stabilizes generated droplets resulting in an increased number of well-formed droplets which results in an increased number of droplets that will be accepted for final analysis in the droplet reader.

13. In addition to the tumor-specific fusion gene a single copy control gene (e.g., albumin) should be quantified for calculation of ctDNA copy numbers in relation to wild-type cell-free circulating DNA.

Furthermore, the supplementary quantification of a single copy control gene enables an evaluation of plasma DNA quality. Quantification of very few copy numbers of the control gene indicates that the isolation of cell-free DNA from plasma did not work sufficiently. Unusually high copy numbers of the control gene indicate lysis of white blood cells and contamination of cell-free circulating plasma DNA with genomic leucocytes DNA. The increase of background DNA from nontumor cells significantly reduce the sensitivity for the detection of low copy numbers of ctDNA fragments in the plasma and could generate false-negative results.

References

1. Krumbholz M, Hellberg J, Steif B, Bauerle T, Gillmann C, Fritscher T, Agaimy A, Frey B, Juengert J, Wardelmann E, Hartmann W, Juergens H, Dirksen U, Metzler M (2016) Genomic EWSR1 fusion sequence as highly sensitive and dynamic plasma tumor marker in Ewing sarcoma. *Clin Cancer Res* 22 (17):4356–4365. <https://doi.org/10.1158/1078-0432.CCR-15-3028>
2. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G et al (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 359 (6391):162–165. <https://doi.org/10.1038/359162a0>
3. Sankar S, Lessnick SL (2011) Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet* 204(7):351–365. <https://doi.org/10.1016/j.cancergen.2011.07.008>
4. Joosse SA, Pantel K (2015) Tumor-educated platelets as liquid biopsy in cancer patients. *Cancer Cell* 28(5):552–554. <https://doi.org/10.1016/j.ccell.2015.10.007>
5. Benini S, Gamberi G, Cocchi S, Garbetta J, Alberti L, Righi A, Gambarotti M, Picci P, Ferrari S (2018) Detection of circulating tumor cells in liquid biopsy from Ewing sarcoma patients. *Cancer Manag Res* 10:49–60. <https://doi.org/10.2147/CMAR.S141623>
6. Tsugita M, Yamada N, Noguchi S, Yamada K, Moritake H, Shimizu K, Akao Y, Ohno T (2013) Ewing sarcoma cells secrete EWS/Flt-1 fusion mRNA via microvesicles. *PLoS One* 8 (10):e77416. <https://doi.org/10.1371/journal.pone.0077416>
7. Miller IV, Raposo G, Welsch U, Prazeres da Costa O, Thiel U, Lebar M, Maurer M, Bender HU, von Luettichau I, Richter GH, Burdach S, Grunewald TG (2013) First identification of Ewing's sarcoma-derived extracellular vesicles and exploration of their biological and potential diagnostic implications. *Biol Cell* 105 (7):289–303. <https://doi.org/10.1111/boc.201200086>
8. Hayashi M, Chu D, Meyer CF, Llosa NJ, McCarty G, Morris CD, Levin AS, Wolinsky JP, Albert CM, Stepan DA, Park BH, Loeb DM (2016) Highly personalized detection of minimal Ewing sarcoma disease burden from plasma tumor DNA. *Cancer* 122 (19):3015–3023. <https://doi.org/10.1002/cncr.30144>
9. Berger M, Dirksen U, Braeuningner A, Koehler G, Juergens H, Krumbholz M, Metzler M (2013) Genomic EWS-FLI1 fusion sequences in Ewing sarcoma resemble breakpoint characteristics of immature lymphoid malignancies. *PLoS One* 8(2):e56408. <https://doi.org/10.1371/journal.pone.0056408>
10. Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, Wong GL, Chan SL, Mok TS, Chan HL, Lai PB, Chiu RW, Lo YM (2015) Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc Natl Acad Sci U S A* 112(11):E1317–E1325. <https://doi.org/10.1073/pnas.1500076112>

Part II

Morpho-Molecular Diagnostics in Ewing Sarcoma



(Immuno)histological Analysis of Ewing Sarcoma

David Marcilla, Isidro Machado, Thomas G. P. Grünewald,
Antonio Llombart-Bosch, and Enrique de Álava

Abstract

The diagnosis of Ewing sarcoma requires the integration of the information generated from numerous techniques, some of them being very sophisticated. However, the first steps of the diagnostic process are crucial to achieve the maximum possible diagnostic performance. In this chapter we will review how to handle the diagnostic specimen from its collection, how to prepare it for diagnosis, how to make a complete pathology report, and provide guidance for the reasonable use of immunohistochemical techniques in this malignancy.

Key words Sarcoma, Ewing sarcoma, Decalcification, Immunohistochemistry, Neoadjuvant therapy, Diagnostic report, Differential diagnosis

1 Introduction

Ewing sarcoma (EwS) is a small round cell sarcoma defined by gene fusions involving one member of the FET family of genes (usually *EWSR1*) and a member of the ETS family of transcription factors [1, 2].

In most cases, EwS tumors are composed of uniform small round cells with round nuclei containing finely stippled chromatin and inconspicuous nucleoli, scant clear or eosinophilic cytoplasm, and indistinct cytoplasmic membranes (classic EwS) [1–3]. In others, the tumor cells are larger, with prominent nucleoli and irregular contours (atypical EwS) [3]. Sometimes, a higher degree of neuroectodermal differentiation (ill-defined groups of as many as 10 cells oriented towards a central space and/or with a consistent immunophenotype) is present (historically termed primitive neuroectodermal tumor) [4]. After induction chemotherapy, EwS cells show a variable degree of necrosis and are replaced by loose connective tissue. Complete pathological response to neoadjuvant chemotherapy is a favorable prognostic factor [5].

EwS expresses several proteins, which can be detected immunohistochemically, and help in the differential diagnosis, such as BCL11B, CAV-1, CD99, ERG, FLI-1, GLG1, NKX2.2, and PAX7 [1, 2].

Multidisciplinary groups consisting of radiologists, orthopedic surgeons, pathologists, radiotherapists, molecular biologists, and oncologists are essential to achieve an appropriate diagnostic and therapeutic approach to EwS. In both bone and soft tissue tumor pathology, the diagnostic procedure of choice is a 14G or 16G thickness biopsy (trocar, TRU-CUT) which must be representative of the tumor, avoiding areas of necrosis. Several biopsy procedures are available: fine needle aspiration biopsy, FNA (almost in disuse as it has a limited role in the diagnosis of bone and soft tissue tumors and may be performed in cases of suspected recurrence, but it helps in obtaining material suitable for molecular studies [6–9], frozen section (it helps determine the status of the resection margins and if the sample is adequate and enough for deferred study [9, 10], core biopsy or TRU-CUT (it is the method of choice for diagnosis of bone and soft tissue tumors, where the biopsy path should be included in the definitive surgery [6, 9], curettage (can be used as a diagnostic-therapeutic procedure in some benign or low-grade neoplasms of the bone (giant cell tumor, chondroblastomas, etc.) but it is not a procedure of choice in the initial diagnosis of EwS [6, 9], incisional biopsy (indicated when after two closed biopsies, the diagnosis is not conclusive, or there is a disagreement between the histological and clinical-radiological findings [6, 9, 10] and marginal resection, “en bloc” resection or radical resection (amputation/disarticulation, performed in referral centers by experienced and specialized surgeons).

This chapter aims at summarizing the different materials and macroscopic, microscopic, and immunohistochemical methods in the diagnostic management of EwS. In the preparation of this document, a review of the relevant published literature has been performed. The objective is to provide a series of protocols and recommendations that can contribute to improving the management and pathological evaluation of EwS in terms of *prediagnostic activities* (freezing, fixation, decalcification, grossing, sampling), *diagnostic activities* (histopathology, immunohistochemistry, assessment of response to neoadjuvant therapy), and *postdiagnostic activities* (reporting).

2 Materials

1. Positively charged glass slides.
2. Recommended fixatives: Laboratory practice employs a 10% formalin or 4% formaldehyde water solution. 95% ethyl alcohol is used to fix tumor imprints.

3. Hematoxylin and eosin (H&E).
4. Liquid nitrogen.
5. Digital photo camera, preferably connected to the information system of the Pathology lab.
6. China Ink (preferably black).
7. Small painting brush (sizes 2 to 6).
8. Bone cutting saw (*see* **Note 1**).
9. Scalpel.
10. Decalcifying solutions: Generic decalcifiers include chelating agents (EDTA), 5–15% formic acid and 5–10% nitric acid.
11. Microscope.
12. Antibodies for immunohistochemistry in EwS differential diagnosis: BCL11B, BCOR, CAV-1, CCNB3, CD99, ERG, ETV4, FLI-1, GLG1, INI-1, NKX2.2, and PAX7 (*see* Table 1).

3 Methods

A bone biopsy will need several prediagnostic methods, namely, grossing (*see* Subheading 3.1), fixation (*see* Subheading 3.2), and decalcification (*see* Subheading 3.3). Pathological diagnosis and assessment of response to neoadjuvant therapy (*see* Subheading 3.4), immunohistochemistry and its interpretation (*see* Subheading 3.5), reporting of core biopsies (*see* Subheading 3.6) or of resections (*see* Subheading 3.7), are relevant diagnostic and post diagnostic procedures in EwS specimens.

3.1 Grossing

1. All biopsy specimens (*see* Subheading 1) must be submitted fresh, without any fixatives, allowing them to be frozen, and cryopreserved.
2. Take cytological imprints, whenever possible if it does not interfere with the diagnosis by gently pressing a freshly cut surface of the specimen on a positively charged glass slide, with a gliding movement.
3. Fix the slide with 95% ethyl alcohol for 5 s.
4. Stain the slide with H&E [6, 9, 10].
5. Snap-freeze fragments of both tumor and nontumor tissue using liquid nitrogen.
6. If the tissue material is to be included in the biobank for subsequent studies, complete the relevant informed consent.
7. For Core biopsies, TRU-CUTS, and curettages then *proceed directly to* fixation (*see* Subheading 3.2) and decalcification (*see* Subheading 3.3).

Table 1
Technical recommendations for antibodies useful for differential diagnosis of EwS

Antibodies	Source	Clone	Dilution	Pretreatment condition	Staining pattern
CD99	Dako	HBA-71	1/50	Citrate buffer (pH 6.0) pressure cooker boiling	M
NKX2.2	Developmental studies Hybridoma Bank	74.5A5	1/250	Autoclave, low pH	N
PAX-7	Developmental studies Hybridoma Bank, Iowa City, IA	Anti-PAX-7	1/200	Envision system with mouse LINKER	N
FLI-1	Santa Cruz	Polyclonal	1/50	Citrate buffer (pH 6.0) pressure cooker boiling	N
CAV-1	Santa Cruz	Polyclonal	1/200	Citrate buffer (pH 6.0) pressure cooker boiling	M, C
ERG	Epitomics	EPR3864(2), C-terminus	1/1000	Combine Autostainer with PT link	N
ERG	Biocare	9FY	1/200	Combine Autostainer with PT link	N
ETV4	Santa Cruz biotechnology	16	1/50	Autoclave, high pH	N
BCOR	Santa Cruz biotechnology	C-10	1/200	Autoclave, low pH	N
CCNB3	SIGMA-ALDRICH HPA000496	Polyclonal	1/500	Autoclave, low pH	N
GLG-1	HPA010815, atlas antibodies, Bromma, Sweden	Polyclonal	1/250	Microwave and antigen retrieval AR-10 solution (HK057-5 K, DCS innovative, Hamburg, Germany)	Perinuclear
BCL11B	ab70453, Abcam, Cambridge, UK	Polyclonal	1/1000	Microwave and Dako target retrieval solution (S1699, Agilent, Santa Clara, CA, USA)	N

Legend for staining pattern: *M* membrane, *N* nucleus, *C* cytoplasm

8. For open biopsies and tumor resections, visualize the presurgical imaging studies in order to determine the section axis (sagittal, coronal, or axial).

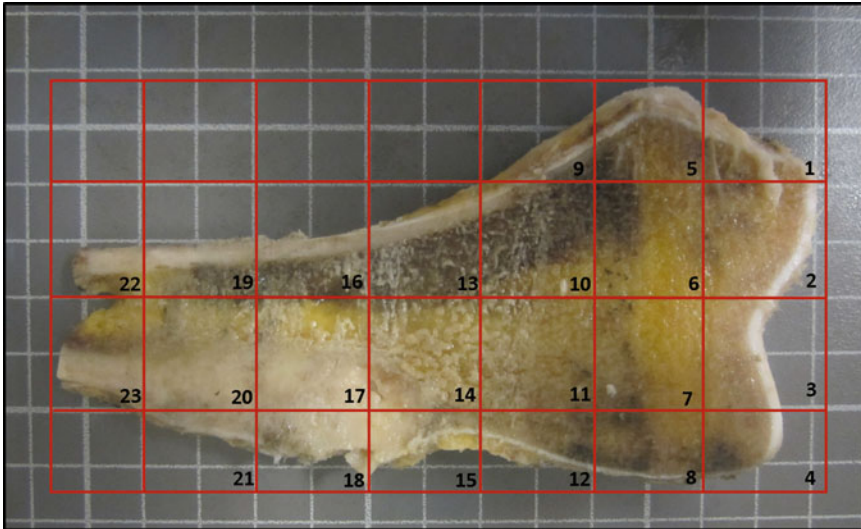


Fig. 1 EwS of the femur of a 17-year-old boy treated with neoadjuvant chemotherapy. A section of the tumor has been prepared; the picture was taken as an inclusion map; numbers indicate the paraffin blocks to indicate which area of the bone corresponds to each section taken for histological analysis

9. Before grossing, take photographs of the resection, and use China ink and a painting brush to paint all surgical margins.
10. In order to obtain the most representative tumor section, make an incision using a scalpel or prepare a section with a bone cutting saw (*see Note 1*) seeking the maximum tumor volume and the greatest relationship with the adjacent structures.
11. Take a photograph of the obtained section that will serve as an inclusion map (mapping) for the different sections (Fig. 1, *see Note 2*).
12. Prepare a diagram with the different paraffin blocks that will be obtained, indicating which area of the bone corresponds to each section taken for the subsequent histological analysis [6, 9–11].
13. Determine which sections should be submitted for study including the most representative tumor complete section [10–13].
14. In case of neoadjuvant treatment, include the complete section with mapping according to the outline of the photograph taken after the section (*see Note 2*).
15. In case of involvement of adjacent soft tissues, include additional sections covering those areas.
16. In cases that have not received neoadjuvant treatment, include one section for each cm of tumor, considering its maximum diameter, up to a maximum of 12 sections and taking the closest resection margins.

17. Take sections of any abnormal zone located in bone, soft tissue and skin.
18. Take sections including lymph nodes and vascular resection margin.
19. Take a section that includes the path of the diagnostic biopsy (for which it is important to tattoo the entrance of the biopsy to locate it on the resection).

3.2 Fixation

1. In cases in which fresh material is obtained for preservation in a biobank, fixate the sample within 30 min after resection in 10% neutral buffered formalin.
2. In cases where fresh material is available but there is no need to freeze any tissue, fixate the sample immediately after the biopsy is taken by submerging it in the adequate solution during the appropriate time period (*see Note 3*).

3.3 Decalcification

Bone biopsies must be adequately fixed and decalcified. Most conventional (H&E, histochemistry) and immunohistochemistry (IHC) techniques can be successfully performed on decalcified tissues even with acids. However, the use of strong acids affects the preservation of nucleic acids and therefore interferes in molecular pathology studies because they generate DNA and RNA fragmentation.

1. Determine and apply the decalcification agent according to the type of sample (biopsy or surgical specimen). Generic decalcifiers include chelating agents (EDTA), 5–15% formic acid, and 5–10% nitric acid [14].
2. Determine the time of decalcification according to the sample size and the type of decalcifier chosen [8, 9, 11] (*see Note 4*).

3.4 Pathological Diagnosis and Assessment of Histological Response to Neoadjuvant Therapy

H&E-stained sections are examined under the microscope. Although most EwS tumors can be recognized with a classical H&E stain, immunohistochemistry (*see* Subheading 3.5) and molecular confirmation by PCR, FISH, or RNAseq are mandatory to confirm diagnosis and differential diagnosis [1, 2, 13]. To assess histological response to neoadjuvant therapy in resection specimens follow these steps:

1. Select all viable areas measured by H&E observation.
2. Divide the total viable area by the total cross-sectional area of the tumor and determine the percentage [10].
3. Describe the response to therapy on the surgical pathology report (*see* Subheading 3.7) as the proportion between viable and necrotic tumor areas [10, 15–17].

4. Determine the response to neoadjuvant treatment according to Salzer-Kuntschik or by the Picci system (reviewed in 16) (*see Note 5*).
5. Alternatively, score the response according to the European Organization for research and Treatment of Cancer-Soft tissue and bone sarcoma Group (EORTC-STBSG) (*see Note 6*).

3.5 Immuno-histochemistry and Interpretation

Immunohistochemistry analysis provides an additional tool to confirm the morphological diagnosis of EwS [1, 4, 17, 18]. Nevertheless, pathologists must be aware that the specificity and sensitivity of many of the traditional (CD99, FII-1, CAV-1) and emerging (NKX2.2, PAX-7, ETV4, BCOR, BCL11B, GLG1) antibodies are not perfect [4, 19–38]. A table is provided with all relevant conditions for each antibody of interest in the differential diagnosis of EwS (Table 1).

CD99 expression constitutes a very sensitive, but poorly specific marker for the diagnosis of EwS, and many genetically confirmed cases reveal strong or moderate membranous staining [1, 2, 4, 17–20] (Fig. 2). CD99 cytoplasmic expression or lack of CD99 immunoreactivity is uncommon in EwS, and thus prompts the search for alternative diagnoses, even though decalcification procedures in bone tumor samples may result in inadequate CD99 staining [4, 17–20, 22, 25]. Regarding specificity, weak and focal CD99 immunoreactivity may also be expressed in a number of other round/ovoid cell tumors (rhabdomyosarcoma, synovial sarcoma, lymphoblastic lymphoma, small cell osteosarcoma, other EWSR1-rearranged sarcomas, CIC-rearranged sarcomas, or BCOR-rearranged sarcomas) [4, 17, 18, 22, 25, 26, 33–35, 37].

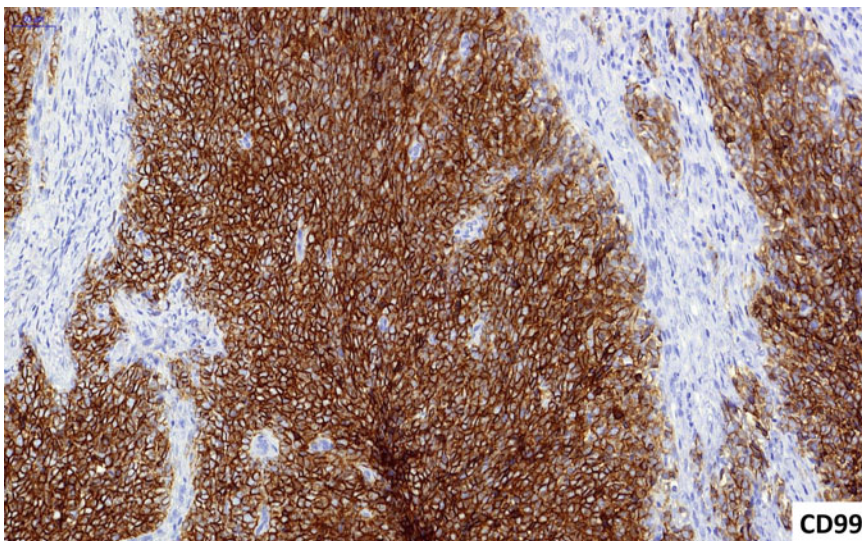


Fig. 2 CD99 expression in EwS has a strong and diffuse membrane pattern

FLI-1 and *ERG* genes are typical fusion partners for the *EWSR1* gene in EwS [1, 2]. Accordingly, it is not unusual for FLI1 and/or ERG immunoreactivity to be observed in EwS [4, 17–32]. Both antibodies display a nuclear staining pattern which facilitates the interpretation, in addition both antibodies stain the cell nuclei of endothelial cells and the lymphocytes of the stroma, thus an internal positive control indicates the tissue quality, especially in decalcified samples [4, 17–32]. Like CD99, both markers lack specificity and their sensitivity is quite variable [1, 4, 17–25, 30–32]. EwS with *EWSR1-ERG* gene fusion usually shows positive immunostaining for the ERG antibody, but many of the cases with *EWSR1-FLI-1* also reveal ERG immunoreactivity [31, 32]; therefore, nuclear ERG immunoreactivity is not completely specific for EwS with an *EWSR1-ERG* fusion. Although molecular studies to search for a specific gene fusion is the most appropriate method to determine ERG or FLI-1 rearrangement, knowing the specific gene fusion in EwS so far lacks prognostic significance [39, 40].

CAV-1, a direct target of the *EWSR1-FLI1* chimeric proteins, has been proposed as a potential diagnostic marker for EwS diagnosis [4]. CAV1 immunoreactivity has been observed in around 96% of EwS and is usually expressed in CD99-negative genetically confirmed ES [4]. Despite its high sensitivity, CAV-1 has limited specificity since CAV-1 immunoreactivity has been communicated in several tumors (sarcomas and carcinomas) [4].

Nuclear NKX2.2 (Fig. 3) and/or PAX-7 immunoreactivity is frequent in EwS and nuclear immunoreactivity offers some advantages for their assessment [21–24, 26–28]. However, the lack of

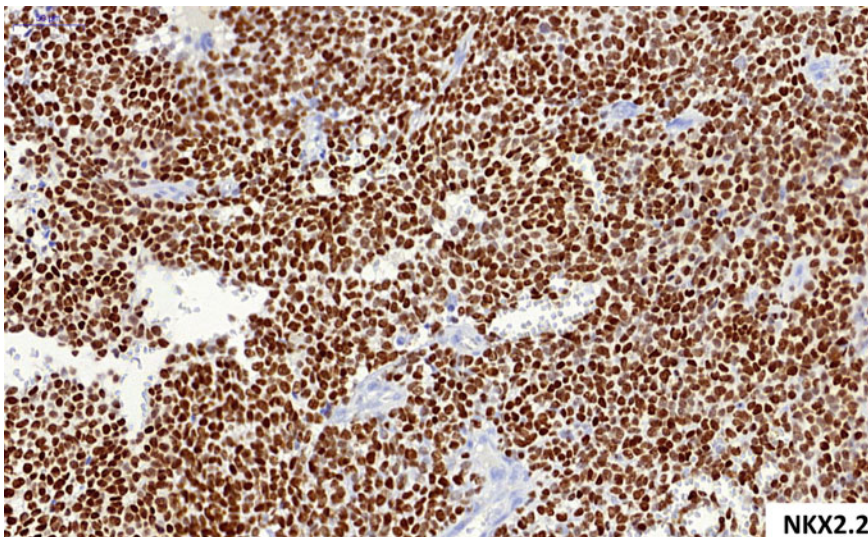


Fig. 3 Nuclear expression of NKX2.2 is observed in about 93% of EwS and constitutes a useful diagnostic biomarker

internal positive control makes it difficult to assure optimal tissue quality for immunohistochemical evaluation, particularly in decalcified tissues [21–24, 26–28].

The *NKX2.2* gene has also been identified as a target of *EWSR1-FLII* [23, 24, 26]. Yoshida et al. reported NKX2.2 nuclear staining in 93% of EwS, with these results being confirmed in additional series [26]. Yoshida et al. found diffuse staining in all positive cases and moderate or strong staining for most other cases [26]. With a sensitivity of 93%, NKX2.2 is a valuable marker for EwS diagnosis, although the specificity is only 89% since it can also be positive in other round cell tumors (olfactory neuroblastomas, extraskeletal myxoid chondrosarcoma, mesenchymal chondrosarcomas, small cell carcinomas, Merkel cell carcinoma, and synovial sarcomas) [22–25]. Shibuya R et al. reported that specificity of NKX2.2 when combined with CD99 increased to 98%, with expression of both markers in only one mesenchymal chondrosarcoma and one small cell carcinoma [25]. This study supports the combination of CD99 and NKX2.2 as a powerful diagnostic tool when dealing with EwS [25].

PAX-7 is a novel marker with documented expression in EwS (Fig. 4) [21, 27, 28]. Toki S et al. demonstrated that although sensitivity of NKX2.2 and PAX-7 is similar in EwS diagnosis, PAX-7 showed more extensive and strong reactivity [21]. However, PAX-7 was also expressed in other round cell sarcomas, including EWSR1-NFATC2 sarcoma, alveolar rhabdomyosarcomas, poorly

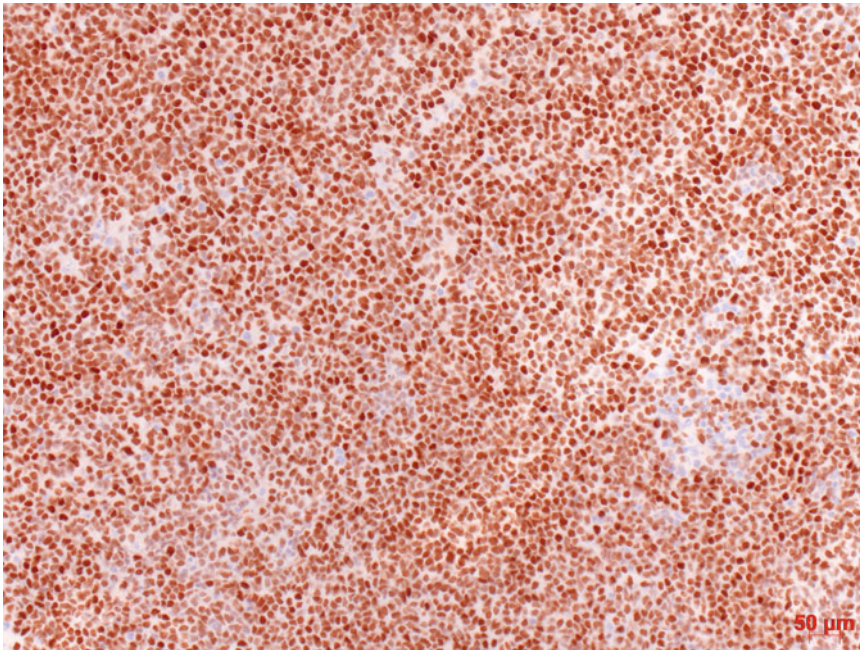


Fig. 4 Most EwS show nuclear expression of PAX7

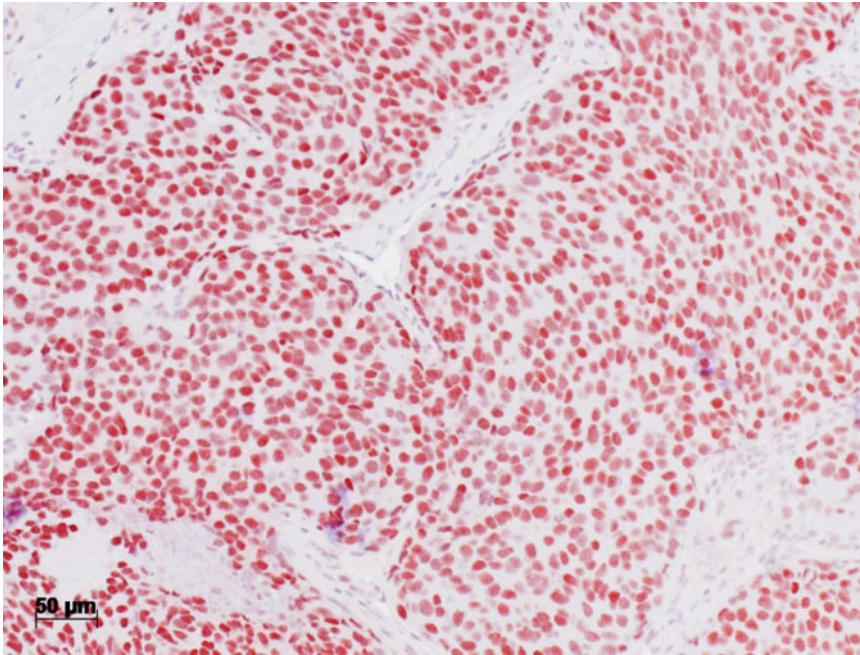


Fig. 5 BCL11B is a novel diagnostic biomarker which is expressed in the nuclei of EwS cells

differentiated synovial sarcomas, BCOR-CCNB3 sarcomas, small-cell osteosarcoma, and desmoplastic small round cell tumor [21, 27–29]. Although PAX-7 is a sensitive marker for EwS diagnosis, its specificity is like that of NKX2.2. Nevertheless, the PAX-7 antibody may stain several EwS mimics, whose spectrum is somewhat different from the NKX2.2-positive non-Ewing entities [20]. Finally, the combination of CD99, NKX2.2, and PAX-7 staining is helpful in narrowing down the differential diagnosis of EwS and can provide important information to streamline a further molecular workup.

Recently, also BCL11B (Fig. 5) and GLG1 (Fig. 6) have been identified and validated as useful and highly specific auxiliary diagnostic markers for EwS, which can be used in conjunction with CD99 [19, 20].

ETV4 nuclear immunoexpression facilitates the identification of some small round cell sarcomas with Ewing-like morphology (CIC-rearranged sarcomas and EWSR1-NFATc2 sarcomas), although ETV4 may be positive in EwS with *EWSR1-ETV4* gene fusion [22, 33–37]. In fact, EwS with the *EWSR1-ETV4* gene fusion represents a very low proportion of cases of this entity [22, 33–37].

BCOR immunoreactivity is rare in EwS and a diffuse and nuclear BCOR immunoexpression is very suggestive of BCOR-rearranged sarcomas as well as other round cell sarcomas, for instance most synovial sarcomas may reveal BCOR immunoreactivity [38, 41].

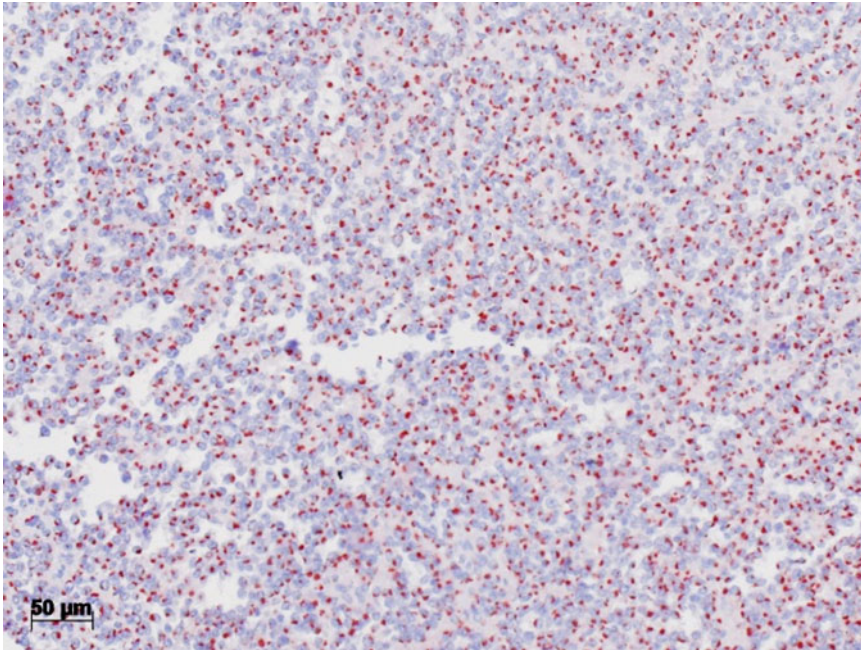


Fig. 6 GLG-1 expression in EwS cells has a characteristic paranuclear pattern

The combination of CD99, NKX2.2, PAX-7, ETV4, and BCOR facilitates the subclassification of small round/ovoid cell sarcomas as a presumptive diagnosis of EwS in which the molecular studies do not obtain a definitive diagnosis due to negative or noninformative results.

3.6 Reporting
Template of a Core
Biopsy with EwS
[10–13]

1. Clinical data including familial syndromes and preexisting skeletal diseases.
2. Imaging findings.
3. Anatomical location.
4. Specify the involved bone(s).
5. Location within the affected bone (s): diaphysis, epiphysis, metaphysis, medullary, cortical, and periosteal.
6. Specify whether the tumor extends to soft tissues.
7. Type of procedure: biopsy and needle diameter (14G–16G).
8. Use of decalcifier and type of agent.
9. Gross pathology: length and diameter of the core(s).
10. Pathology diagnosis-tumor subtype.
11. Necrosis (optional).
12. Lymphovascular permeation (optional).
13. Results of immunohistochemical studies.

14. Results of Molecular Pathology studies (FISH, PCR, massive parallel sequencing).
15. Microscopic description (optional; recommended only when providing relevant or clarifying information).

3.7 Reporting
Template of a Surgical
Resection Ewing
Sarcoma [10–13]

1. Clinical data.
2. Imaging techniques. The pathologist must have access to the imaging techniques or their results.
3. Neoadjuvant therapy (type of treatment administered).
4. Specify affected bone and tumor location: diaphysis, epiphysis, metaphysis. Cortical, medullary, or periosteal.
5. Extension to adjacent soft tissues and/or joint involvement.
6. Type of procedure: marginal resection, wide resection, radical resection.
7. Tumor size (mm), at least two main axes.
8. Pathology diagnosis-tumor subtype.
9. Tumor grade.
10. Lymphovascular permeation (optional).
11. Response to neoadjuvant treatment: percentage of changes attributable to treatment.
12. Surgical margins:
 - (a) Negative: distance (in mm) to closest margin.
 - (b) Positive: specify which margin is positive.
13. Lymph nodes:
 - (a) Number of lymph nodes studied.
 - (b) Number of positive lymph nodes.
14. Results of immunohistochemical studies.
15. Results of molecular pathology studies.
16. Tumor stage (TNM eighth edition).

4 Notes

1. Modern, electrical equipment should be available in bone pathology referral centers to prepare bone resections for histopathological analysis. Safety issues are important in the choice of such an instrument (low risk of serious injury during operation, no tissue aerosols during operation, etc.). Diamond bands help to minimize risks.
2. One should bear in mind that practically all surgical specimens will correspond to post-neoadjuvant resections that will require the histological study of a complete section that must be the

most representative of the tumor, to give a degree of pathological response. Therefore, photos should be taken before and after the dissection or cutting.

3. Formaldehyde fixation time will be determined by the sample size and type of procedure [11]. Recommended fixation times are:
 - (a) Core biopsy, curettage, and incisional biopsy: between 6 and 24 h.
 - (b) Surgical resection: between 24 and 48 h [7, 8, 10, 11].
 - (c) Radical resection: Complete compartment resections or amputations are included in this section. In the event of limb amputation, the limb can be frozen in freezers -80°C for no more than 24 h. Subsequently, it must be dissected and subsequently fixed. This cutting system after freezing allows us to better observe the relationship of the tumor with the adjacent anatomic structures. It is not advisable to keep the piece frozen for more than 24 h, in order to avoid freezing artifacts in the tissue, which can subsequently hinder histological evaluation (mainly in cases of neoadjuvant treatment) [11].

If the resection piece includes metal prostheses (mainly in cases of recurrence), the presence of the orthopedic surgeon is required, who will provide his/her experience and the appropriate material to facilitate the handling of the piece [11].

4. For diagnostic biopsies (trocar, incisional), the use of decalcifying agents with EDTA is recommended, which, although require a longer time compared to formic or nitric acid to achieve an optimal decalcifying status, preserves nucleic acids and allows for molecular studies to be carried out (FISH, PCR, massive parallel sequencing, etc.). For a better performance of EDTA salts, it is recommended to renew the agent every 24–48 h. The process can also be speed up by using microwave or stirring.

In resection specimens of already diagnosed cases that do not require molecular pathology studies, we can use faster-acting agents such as formic acid or nitric acid. The decalcification with nitric acid is very fast, so we must control it properly since it can produce an exaggerated decalcification of the tissue, with denaturation and destruction of nucleic acids and alteration of cell morphology [8, 9, 11].

Obviously, these are general recommendations, and each laboratory must choose the method with which it achieves the best results [10].

5. Both systems agree in that a higher than 90% extensive necrosis is usually associated with better survival and better prognosis [10, 15–17]. The Salzer-Kuntschik system defines six grades:

grade 1 = no vital tumor cells; grade 2 = very few single viable tumor cells; grade 3 = less than 10% viable tumor tissue; grade 4 = 10–50% viable tumor tissue; grade 5 = more than 50% viable tumor tissue; grade 6 = completely viable tumor [16]. Alternatively, Picci P et al. defined three grades: grade I (at least one residual macroscopic nodule of viable tumor), grade II (only isolated microscopic nodules of viable tumor cells) and grade III (no viable nodules of tumor cells identified within the specimen) [10, 16].

6. Although primary soft tissue EwS are less frequent in comparison with bone tumors, the European Organization for research and Treatment of Cancer-Soft tissue and bone sarcoma Group (EORTC-STBSG) have been proposed a new response score for soft tissue tumors defines as follows: (a) no stainable tumor cells (STC); (b) single STC or small clusters (overall below 1% of the whole specimen), (c) $\geq 1\%$ - $< 10\%$ STC, (d) $\geq 10\%$ - $< 50\%$ STC; E. $\geq 50\%$ STC. They proposed classifying the residual tumor cells as “STC” instead of viable cells, where STC means that well-discernible nuclei are visualized by H&E [16]. EwS has no osteogenic matrix, and consequently this entity may have more similarities with soft tissue sarcoma about pathological response compared to osteosarcomas.

References

1. de Alava E, Lessnick SL, Stamenkovic I (2020) Ewing sarcoma. In: WHO classification of tumours of soft tissue and bone, 5th edn. IARC Press, Lyon
2. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4:5
3. Machado I, Noguera R, Mateos EA et al (2011) The many faces of atypical Ewing’s sarcoma. A true entity mimicking sarcomas, carcinomas and lymphomas. *Virchows Arch* 458:281–290
4. Llombart-Bosch A, Machado I, Navarro S et al (2009) Histological heterogeneity of Ewing’s sarcoma/PNET: an immunohistochemical analysis of 415 genetically confirmed cases with clinical support. *Virchows Arch* 455:397–411
5. Albergo JI, Gaston CL, Laitinen M et al (2016) Ewing’s sarcoma: only patients with 100% of necrosis after chemotherapy should be classified as having a good response. *Bone Joint J* 98-B:1138–1144
6. Wu JS, Goldsmith JD, Horwich PJ, Shetty SK, Hochman MG (2008) Bone and soft-tissue lesions: what factors affect diagnostic yield of image-guided core-needle biopsy? *Radiology* 248:962–970
7. Mc Carthy EF (2007) CT-guided needle biopsies of bone and soft tissue tumors: a pathologist’s perspective. *Skelet Radiol* 36:181–182
8. Traina F, Errani C, Toscano A et al (2015) Current concepts in the biopsy of musculoskeletal tumors: AAOS exhibit selection. *J Bone Joint Surg Am* 97:1–6
9. Errani C, Traina F, Perna F et al (2013) Current concepts in the biopsy of musculoskeletal tumors. *Sci World J* 2013:538152
10. Laurini J; Antonescu C, Cooper K et al (2017) Protocol for the examination of specimens from patients with primary tumors of bone. CAP cancer and CAP pathology electronic reporting committees. Accessed 12 Dec 2019
11. Machado I, Pozo JJ, Marcilla D et al (2017) Protocol for the study of bone tumours and standardization of pathology reports. *Rev Esp Patol* 50(1):34–44
12. Mangham DC, Athanasou NA (2011) Guidelines for histopathological specimen examination and diagnostic reporting of primary bone tumors. *Clin Sarcoma Res* 1:1–6

13. ESMO Guidelines Committee, PaedCan and ERN EURACAN (2018) Bone sarcomas: ESMO PaedCan and ERN EURACAN clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 29(Suppl 4):iv79–iv95
14. Mangham DC, Williams A, McMullan DJ et al (2006) Ewing's sarcoma of bone: the detection of specific transcripts in a large, consecutive series of formalin-fixed, decalcified, paraffin-embedded tissue samples using the reverse transcriptase-polymerase chain reaction. *Histopathology* 48:363–376
15. Bacci G, Ferrari S, Bertoni F et al (2000) Prognostic factors in nonmetastatic Ewing's sarcoma of bone treated with adjuvant chemotherapy: analysis of 359 patients at the Istituto Ortopedico Rizzoli. *J Clin Oncol* 18:4–11
16. Wardelmann E, Haas RL, Bovée JV et al (2016) Evaluation of response after neoadjuvant treatment in soft tissue sarcomas; the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC-STBSG) recommendations for pathological examination and reporting. *Eur J Cancer* 53:84–95
17. Folpe AL, Goldblum JR, Rubin BP et al (2005) Morphologic and immunophenotypic diversity in Ewing family tumors: a study of 66 genetically confirmed cases. *Am J Surg Pathol* 29:1025–1033
18. Folpe AL, Hill CE, Parham DM et al (2000) Immunohistochemical detection of FLI-1 protein expression: a study of 132 round cell tumors with emphasis on CD99-positive mimics of Ewing's sarcoma/primitive neuroectodermal tumor. *Am J Surg Pathol* 24:1657–1662
19. Baldauf MC, Orth MF, Dallmayer M et al (2017) Robust diagnosis of Ewing sarcoma by immunohistochemical detection of super-enhancer-driven EWSR1-ETS targets. *Oncotarget* 9:1587–1601
20. Orth MF, Hölting TLB, Dallmayer M et al (2020) High specificity of BCL11B and GLG1 for EWSR1-FLI1 and EWSR1-ERG positive Ewing sarcoma. *Cancers (Basel)* 12(3):644. <https://doi.org/10.3390/cancers12030644>
21. Toki S, Wakai S, Sekimizu M et al (2018) PAX7 immunohistochemical evaluation of Ewing sarcoma and other small round cell tumours. *Histopathology* 73:645–652
22. Machado I, Yoshida A, López-Guerrero JA et al (2017) Immunohistochemical analysis of NKX2.2, ETV4, and BCOR in a large series of genetically confirmed Ewing sarcoma family of tumors. *Pathol Res Pract* 213:1048–1053
23. McCuiston A, Bishop JA (2018) Usefulness of NKX2.2 immunohistochemistry for distinguishing Ewing sarcoma from other sinonasal small round blue cell tumors. *Head Neck Pathol* 12:89–94
24. Hung YP, Fletcher CD, Hornick JL (2016) Evaluation of NKX2.2 expression in round cell sarcomas and other tumors with EWSR1 rearrangement: imperfect specificity for Ewing sarcoma. *Mod Pathol* 29:370–380
25. Shibuya R, Matsuyama A, Nakamoto M et al (2014) The combination of CD99 and NKX2.2, a transcriptional target of EWSR1-FLI1, is highly specific for the diagnosis of Ewing sarcoma. *Virchows Arch* 465:599–605
26. Yoshida A, Sekine S, Tsuta K et al (2012) NKX2.2 is a useful immunohistochemical marker for Ewing sarcoma. *Am J Surg Pathol* 36:993–999
27. Charville GW, Wang WL, Ingram DR et al (2019) PAX7 expression in sarcomas bearing the EWSR1-NFATC2 translocation. *Mod Pathol* 32:154–156
28. Charville GW, Varma S, Forgó E et al (2016) PAX7 expression in Rhabdomyosarcoma, related soft tissue tumors, and small round blue cell neoplasms. *Am J Surg Pathol* 40:1305–1315
29. Baldauf MC, Gerke JS, Orth MF et al (2018) Are EWSR1-NFATc2-positive sarcomas really Ewing sarcomas? *Mod Pathol* 31:997–999
30. Chen S, Deniz K, Sung YS et al (2016) Ewing sarcoma with ERG gene rearrangements: a molecular study focusing on the prevalence of FUS-ERG and common pitfalls in detecting EWSR1-ERG fusions by FISH. *Genes Chromosomes Cancer* 55:340–349
31. Machado I, Mayordomo-Aranda E, Scotlandi K et al (2014) Immunoreactivity using anti-ERG monoclonal antibodies in sarcomas is influenced by clone selection. *Pathol Res Pract* 210:508–513
32. Tomlins SA, Palanisamy N, Brenner JC et al (2013) Usefulness of a monoclonal ERG/-FLI1 antibody for immunohistochemical discrimination of Ewing family tumors. *Am J Clin Pathol* 139:771–779
33. Yoshida A, Arai Y, Kobayashi E et al (2017) CIC break-apart fluorescence in-situ hybridization misses a subset of CIC-DUX4 sarcomas: a clinicopathological and molecular study. *Histopathology* 71:461–469

34. Kao YC, Sung YS, Chen CL et al (2017) ETV transcriptional upregulation is more reliable than RNA sequencing algorithms and FISH in diagnosing round cell sarcomas with CIC gene rearrangements. *Genes Chromosomes Cancer* 56:501–510
35. Smith SC, Palanisamy N, Martin E et al (2017) The utility of ETV1, ETV4 and ETV5 RNA in-situ hybridization in the diagnosis of CIC-DUX sarcomas. *Histopathology* 70:657–663
36. Hung YP, Fletcher CD, Hornick JL (2016) Evaluation of ETV4 and WT1 expression in CIC-rearranged sarcomas and histologic mimics. *Mod Pathol* 29:1324–1334
37. Le Guellec S, Velasco V, Pérot G et al (2016) ETV4 is a useful marker for the diagnosis of CIC-rearranged undifferentiated round-cell sarcomas: a study of 127 cases including mimicking lesions. *Mod Pathol* 29:1523–1531
38. Sbaraglia M, Righi A, Gambarotti M, Dei Tos AP (2020) Ewing sarcoma and Ewing-like tumors. *Virchows Arch* 476:109–119
39. Barr FG, Meyer WH (2010) Role of fusion subtype in Ewing sarcoma. *J Clin Oncol* 28:1973–1974
40. Le Deley MC, Delattre O, Schaefer KL (2010) Impact of EWS-ETS fusion type on disease progression in Ewing's sarcoma/peripheral primitive neuroectodermal tumor: prospective results from the cooperative Euro-E.W.I.N.G. 99 trial. *J Clin Oncol* 28:1982–1988
41. Kao YC, Owosho AA, Sung YS et al (2018) BCOR-CCNB3 fusion positive sarcomas: a Clinicopathologic and molecular analysis of 36 cases with comparison to morphologic Spectrum and clinical behavior of other round cell sarcomas. *Am J Surg Pathol* 42:604–615



Molecular Approaches to Diagnosis in Ewing Sarcoma: Fluorescence In Situ Hybridization (FISH)

Marcel Trautmann and Wolfgang Hartmann

Abstract

The differential diagnosis of small round cell tumors (SRCT) crucially relies on the synoptic evaluation of morphology, immunohistochemical patterns, and molecular features. Though the implementation of broad RNA sequencing in diagnostic molecular pathology routines has substantially changed the standards of molecular affirmation of diagnoses, fluorescence in situ hybridization (FISH) on formalin-fixed, paraffin-embedded (FFPE) tissue sections is still an elementary tool to provide a rapid molecular corroboration of diagnoses, essentially required for therapeutic decisions. We discuss here the major FISH approaches currently employed in diagnostic molecular pathology, addressing classic Ewing sarcoma and differential diagnoses among SRCT which cannot sufficiently be ruled out by immunohistochemistry. This chapter will approach technical issues but particularly strategies and pitfalls in the interpretation of FISH patterns.

Key words Small round cell tumor, Ewing sarcoma, FISH, *EWSR1*, *FUS*, *CIC*, *BCOR*

1 Introduction

The group of malignant small round cell tumors (SRCT) comprises a multitude of different tumor entities. Within a given clinical context, conventional histology and immunohistochemical characteristics are of predominant importance in the differential diagnosis of SRCT; however, morphological overlap and the lack of specific markers to confirm a certain lineage may assign a major role to genetic diagnostic substantiation. Given its largely nonspecific immunohistochemical profile comprising positivity for CD99, FLI1, Caveolin, and NKX-2.2 [1–5], molecular confirmation of Ewing sarcoma is mandatory, even in a typical clinical setting. Consistent with the initial description of the t(11;22)(q24;q12) chromosomal translocation the vast majority of Ewing sarcomas (~85%) are characterized by an *EWSR1-FLI1* gene fusion [6] which displays a relevant variability [7]. In the 15–20% of Ewing sarcomas negative for *EWSR1-FLI1* fusions, variant fusions between *EWSR1* and other members of the *ETS* family occur, most commonly *ERG*

and very rarely *ETV1*, *ETV4*, or *FEV* [7–11] occur. As shown later, *FUS*, another member of the *FET* gene family, may replace *EWSR1* in cases of morphologically typical Ewing sarcoma [12, 13]. The common pattern of recurrent genomic translocations involving the *FET* gene *EWSR1* (or rarely *FUS*) provided the opportunity to establish diagnostic assays focusing on these “constantly” involved fusion partners in genomic translocations in Ewing sarcomas [14], which has then been implemented in laboratory routines as dual-color break-apart FISH. Though these diagnostic approaches do not identify the concrete gene fusion, they are suitable to indicate a chromosomal rearrangement characteristic of Ewing sarcoma which, in the context of a typical morphological and immunohistochemical profile, may suffice as molecular confirmation. However, care must be taken since chromosomal rearrangements in *EWSR1* and *FUS* are not specific for Ewing sarcoma and may occur in biologically very different tumors [15]. Beyond molecular confirmation of histologically characteristic Ewing sarcoma, FISH may be relevant and even essential in the diagnosis of SRCT morphologically mimicking Ewing sarcoma. These tumors typically lack *FET-ETS* gene fusions but harbor other characteristic rearrangements affecting the *CIC* gene [16–19] or the *BCOR* gene, each with variable fusion partners [20–23]. Some but not all of these alterations may be detectable by FISH, and some may be missed by RNA sequencing. Though RNA-based analytics are increasingly implemented in diagnostic molecular pathology routines and continuously change the standards of confirmatory molecular diagnostics, there is still a major role for FISH on formalin-fixed, paraffin-embedded (FFPE) tissue sections, given its broad availability, the rapidness of the diagnostic procedure and its fundamentally unbiased character which occasionally turns out to be advantageous. FISH will therefore continue to be employed as an elementary tool in the molecular corroboration of diagnoses.

Given the elaborate literature on technical issues of FISH in FFPE tissue [24] and the broad availability of commercially available fluorescence-labeled FISH probes, this chapter will give a short survey on technical procedures but particularly focus on the interpretation of FISH patterns in the differential diagnosis of Ewing sarcoma, suggesting diagnostic algorithms and discussing pitfalls.

2 Materials

Different laboratory protocols for all routine molecular applications are broadly available. We here describe a slightly modified protocol employing the Histology FISH Accessory Kit (Dako, Glostrup, Denmark) suitable for the analysis of FISH probes of different manufacturers. We do not claim this protocol to be superior to

other procedures and briefly comment on the relevance of the important steps which are comparable in other manufacturer's procedures.

1. 4% neutral-buffered formalin, commercially available as ready-to-use solution
2. Tissue cassettes.
3. Rotary microtome.
4. Poly-L-lysine-coated or adhesive slides (e.g., Superfrost Plus).
5. ThermoBrite hybridization system (Abbott) or a comparable device to warrant standardized incubation temperatures and time.
6. Standard xylene.
7. Ethanol series (to be prepared freshly every 200 slides):
 - 99% ethanol
 - 85% ethanol (add 15 mL deionized water to 85 mL ethanol)
 - 70% ethanol (add 30 mL deionized water to 70 mL ethanol).
8. Fluorochrome-labelled dual-color break-apart FISH probe for *EWSR1*.
9. Fluorochrome-labeled dual-color break-apart FISH probe for *FUS*.
10. Fluorochrome-labeled dual-color break-apart FISH probe for *CIC*.
11. Fluorochrome-labeled dual-color break-apart FISH probe for *BCOR*.
12. Wash Buffer (Tris/HCl buffer) supplied as concentrate to be diluted 1:20 in deionized water.
13. Pretreatment Solution (MES (2-[*N*-morpholino]ethanesulfonic acid) buffer) supplied as concentrate to be diluted 1:20 in deionized water.
14. Pepsin solution supplied as ready-to-use solution.
15. Coverslip sealant.
16. Stringent Wash Buffer (SSC (saline-sodium citrate) buffer with Tween 20) supplied as concentrate to be diluted 1:20 in deionized water.
17. Mounting Medium (4',6-diamidino-2-phenylindole with 500 µg/L DAPI).
18. Fluorescence microscope equipped with appropriate filters and 63 × oil immersion objectives.
19. Immersion oil.

20. High-quality camera.
21. Computer hardware and software for image acquirement (e.g., Leica DM5500 B).

3 Methods

3.1 Technical Procedure

1. Routinely preserved fresh tissue biopsy samples should be rapidly cut in 3–4 mm thick slices.
2. Transfer the cut biopsy to neutral-buffered formalin and fix for 18–24 h (*see Note 1*).
3. Dehydrate and process using standard procedures.
4. Embed the sample in paraffin in an appropriately sized tissue cassette.
5. If possible, cut FFPE tissue sections freshly at a thickness of 2–3 μm on a rotary microtome and mount them on positively charged slides.
6. A further section for standard H&E staining needs to be prepared in parallel (*see Note 2*).
7. Dry the slides by keeping them at 37 °C overnight.
8. For tissue maturation (reducing background signals), incubate the slides at 60 °C for 60 min.
9. Recool the slides to room temperature.
10. Proceed to deparaffinate and rehydrate at room temperature by immersing each slide in the following solutions:
 - 2 \times xylene for 15 min
 - 2 \times 99% ethanol for 2 min
 - 2 \times 85% ethanol for 2 min
 - 2 \times 70% ethanol for 2 min
 - 2 \times Wash Buffer for 2 min
11. Add 150 μL Pretreatment Solution to the deparaffinized sections and cover with a coverslip (*see Note 3*).
12. Incubate at 97 °C in a ThermoBrite slide processor.
13. Cool down to 30 °C.
14. Incubate 2 \times 3 min in Wash Buffer.
15. Add 250 μL of pepsin to cover the specimen (*see Note 4*).
16. Add a coverslip and incubate at room temperature for 8–14 min.
17. Drain the pepsin solution.
18. Incubate for 2 \times 3 min in Wash Buffer.

19. Perform subsequent dehydration at room temperature:
 - 70% ethanol for 2 min
 - 85% ethanol for 2 min
 - 99% ethanol for 2 min
20. Air-dry the tissue sections.
21. Apply the fluorochrome-labeled FISH probes (2–10 μL , according to the size of the specimen), to the center of the tissue section and cover with a coverslip.
22. Seal it applying Coverslip Sealant around the entire periphery of the coverslip.
23. For hybridization, place the slides in the ThermoBrite hybridizer and subject them to a procedure of denaturation and hybridization specific for each manufacturer under protection from light (*see Note 5*).
24. Carefully remove the Coverslip Sealant.
25. Incubate for 2×5 min in Wash Buffer at 37°C until the coverslips loosen.
26. Incubate the slides in Stringent Wash Buffer at 65°C for 10 min.
27. Incubate the slides for 2×3 min in Wash Buffer at room temperature.
28. Perform the following dehydration steps at room temperature:
 - 70% ethanol for 2 min
 - 85% ethanol for 2 min
 - 99% ethanol for 2 min.
29. Air dry the tissue sections.
30. For mounting, apply 15 μL of Fluorescence Mounting Medium containing DAPI to the specimen.
31. Apply a coverslip to each slide.
32. Read the slides 15 min after mounting.
33. Store at 4°C .

3.2 Microscopic Evaluation and Interpretation

Microscopic evaluation and interpretation of break-apart FISH patterns in SRCT is described using the example of *EWSR1*, given its frequent involvement in SRCT (*see Note 6*). The basic principles are, of course, valid for the other relevant gene loci (*FUS*, *CIC*, and *BCOR*) as well (*see Note 7*).

1. Evaluate the slides using a fluorescence microscope with appropriate filters to the fluorochromes employed; for example, excitation at 543 nm induces Cy3 red emission, excitation at 530 nm induces FITC green emission, and excitation at

364 nm induces DAPI fluorescence (blue emission). Employ $63 \times$ oil immersion lenses.

2. Capture digital microscopic images employing appropriate software (*see* **Notes 7** and **8**).
3. Include only intranuclear signals (DAPI control) (*see* **Note 9**).
4. Optimally, a minimum of 100 distinct tumor cells should be analyzed to warrant statistically relevant results. In principle, only cells with at least two green and two red signals should be evaluated. Unusual signal distributions may be observed, possibly resulting in a different signal pattern and indicating variant chromosomal rearrangements (*see* **Notes 10** and **11**).
5. In case of lacking signals in the tumor tissue, internal positive controls (stromal tissue, vasculature, inflammatory bystanders) should be evaluated to exclude technical reasons.
6. Following the schematic example of the *EWSR1* gene locus in Fig. 1, the hybridization signals of the DNA probes are red (5' [centromeric] to the *EWSR1* breakpoint region) and green

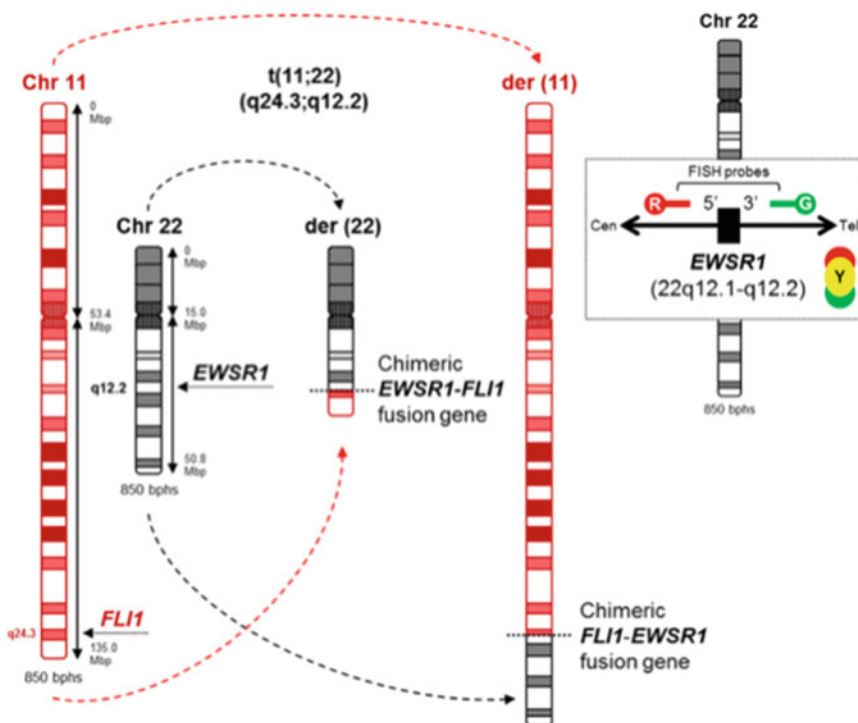


Fig. 1 (left) Schematic view of the most frequent recurrent balanced chromosomal translocation in Ewing sarcoma, $t(11;22)(q12;q24)$, leading to a fusion of the *FET* family gene *EWSR1* with the *ETS* family gene *FLI1*. Given the common involvement of *EWSR1* in Ewing sarcoma, a break-apart FISH assay addressing the *EWSR1* gene locus represents a rational diagnostic tool in diagnostic molecular pathology. **(right)** Basic principle of *EWSR1* FISH using a break-apart probe: Two fluorescence-labeled probes ((Green and (R)ed) hybridize to the telomeric and centromeric flanking regions of *EWSR1*. Cells negative for *EWSR1* rearrangement display a pattern of fused green and red signals with the spectral overlap creating a yellow signal

(3' [telomeric] to the *EWSR1* breakpoint region). In interphases of cells without a translocation involving the *EWSR1* gene region, two green/red fusion signals (optically appearing as yellow in case of overlapping signals) are detected. If the *EWSR1* gene region is affected by a chromosomal translocation, *in the typical situation* separated green and red signals (most frequently displaying a distance >2 signal diameters) are detected apart from a green/red fusion signal (intact *EWSR1* locus) (Fig. 2) (*see Note 12*).

7. Threshold values to define positivity in interpretation need to be defined per laboratory, taking into account the individual

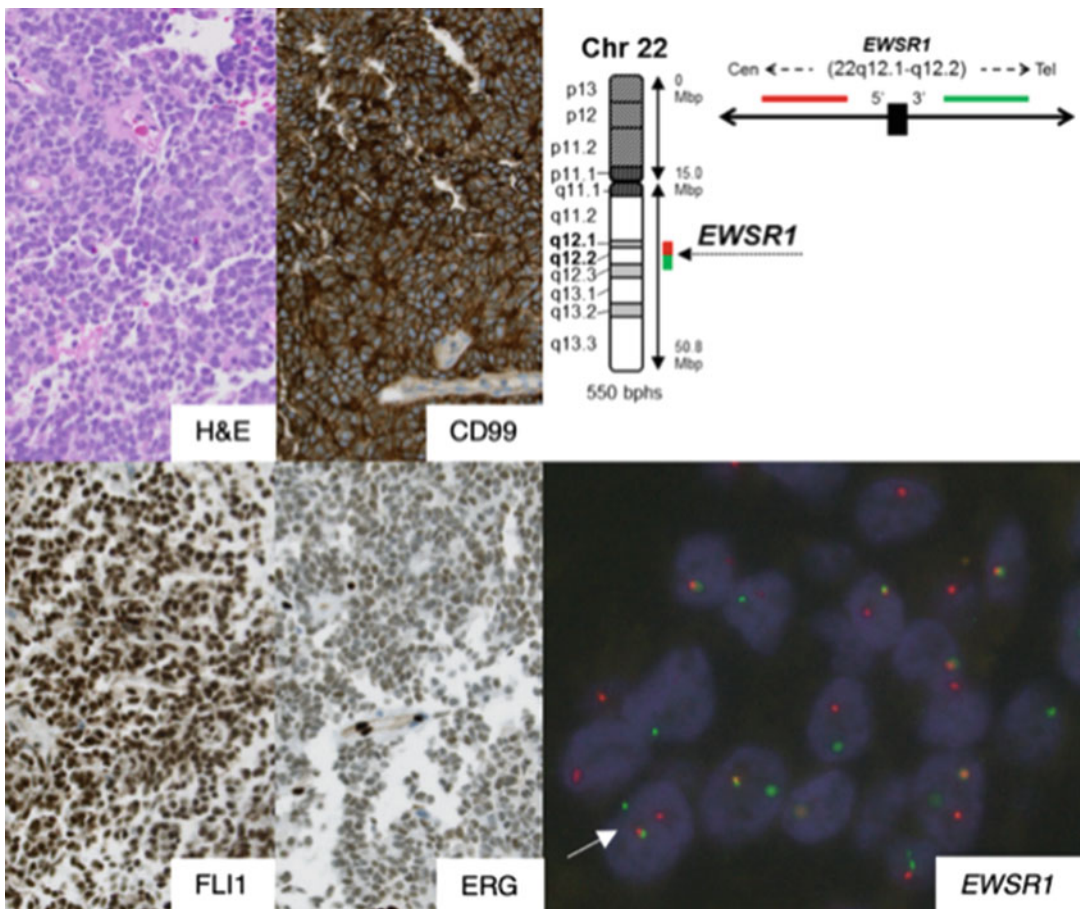


Fig. 2 (left) Case of Ewing sarcoma with characteristic uniform small round cell morphology, strong positivity for CD99 and FLI1 and only weak ERG expression (note strong expression of ERG in endothelial cells). (right, top) Schematic view of the *EWSR1* break-apart FISH approach employed in molecular routine diagnostics. (right, bottom) *EWSR1* break-apart FISH showing multiple cells with an “incomplete” set of signals, with, however, frequent occurrence of separate red or green signals in combination with one yellow (fusion) signal. Arrows indicate cells with a classic break-apart pattern with one yellow signal combined with separate red and green signals at a distance >2 signal diameters

technical settings (e.g., thickness of tissue sections, characteristics of employed probes).

8. The final *interpretation* of FISH results needs to be done in synopsis with the clinical setting, conventional histology and the immunohistochemical profile. Given the broad use of break-apart FISH assays in laboratory routines, it is important to be aware of the following:
 - (a) Classic break-apart assays do not prove a specific chromosomal translocation but only indicate a chromosomal rearrangement in a single genomic region. The detection of aberrant signals may also be due to undirected genomic instability randomly affecting a particular locus.
 - (b) Absence of aberrant signals above the threshold does not necessarily exclude a genomic translocation. Specific molecular backgrounds or (random) secondary molecular changes may dissimulate an aberrant FISH pattern (*see Notes 13 and 14*). Defined cutoff values may therefore require reconsideration in certain contexts to avoid false negativity.
 - (c) Attention needs to be paid to nontypical FISH patterns (*see Note 15*).
 - (d) Morphologically and immunohistochemically typical Ewing sarcoma lacking an *EWSRI* rearrangement need to be checked for *FUS* rearrangements (*see Note 16*).
 - (e) Its lacking restriction to one particular translocation (resulting, e.g., in the *EWSRI-FLII* gene fusion) is one of the major strategic advantages of break-apart FISH in terms of diagnostic pragmatism; however, awareness has to be raised to the fact that translocations involving genes commonly involved in SRCT (e.g., *EWSRI*, or *FUS*) frequently occur in very different tumors than Ewing sarcoma. It is therefore mandatory to regard a given FISH result as one of many different diagnostic parameters only and to interpret the finding in the context of clinics, radiology, conventional histology, and immunohistochemistry.
 - (f) SRCT with nontypical morphology and immunophenotype should be considered to belong to the groups of *CIC*- or *BCOR*-rearranged sarcomas. Detection of rearrangements in these gene loci by FISH may be challenging (*see Notes 17 and 18*).

4 Notes

1. For fresh tissue processing, if required, ethylenediaminetetraacetic acid (EDTA) should be employed as decalcifying agent. Valid FISH on tissue subjected to acid calcification is most often technically impossible due to tissue (including DNA) damage.
2. To provide the opportunity of an optimal orientation through correlation of the FISH image with conventional morphology, and to ensure that the FISH section contains representative tumor tissue, hematoxylin and eosin (H&E) sections should always be cut in parallel to FISH sections and be made available to the FISH reader.
3. Pretreatment: It is performed to maximize the accessibility of the target to the DNA probe and will break DNA cross-links formed during tissue fixation.
4. Pepsin treatment leads to a removal of the cytoplasm and tissue permeabilization, preserving cellular morphology and nuclear structure. Overdigestion may lead to low-quality morphology (dark intranuclear areas), increased background, and loss of signal. Pepsin incubation time may need to be optimized dependent on tissue fixation and/or decalcification.
5. Generally, denaturation is performed at temperatures 73–83 °C for 3–10 min and hybridization is performed at 37–45 °C overnight (14–20 h).
6. In molecular pathology routine applications concerning SRCT, there is no major role for dual-fusion assays, which may, however, be used in individual laboratories for the final confirmation of a given gene fusion.
7. Evaluation of FISH as a diagnostic marker is only reliable in the context of a given tumor tissue sample and a specific diagnostic question. Prior to analysis, the FISH reader has to be made familiar with the tumor's conventional histological appearance since accuracy of FISH evaluation essentially relies on the unequivocal identification of tumor cells. This may be challenging in monomorphic round cell tumors since tumor cells may show only slight nuclear irregularities and may not differ clearly from, for example, nuclei of (damaged) lymphocytes in the DAPI image. However, information including clinical setting and immunohistochemical profile may be beneficial but may also bias FISH evaluation.
8. Optimal filters are crucial to obtain maximum signal intensity. Parameters may vary according to particular fluorochromes used in different probes.

9. Overlapping nuclei should be omitted as should be overdigested cells (showing dark intranuclear areas) or nuclei with a strong background fluorescence (to be identified by signals in more than one channel).
10. In tiny biopsies with clearly aberrant signals evaluation of <100 cells may suffice if the observed pattern is consistent.
11. Principally, the evaluation of break-apart FISH assays requires the whole set of signals (i.e., at least two green and two red signals) to be present in a nucleus. However, it may occur, that the specific molecular background of a particular chromosomal rearrangement dissimulates the classic pattern and that only incomplete sets of signals are detected. It is of importance to become aware of such abnormal patterns, to interpret them appropriately and, as required, to confirm a supposed rearrangement by complementary FISH assays or RNA sequencing.
12. In the classic situation of a $t(11;22)(q24;q12)$ chromosomal translocation resulting in an *EWSR1-FLII* gene fusion, the evaluation of the *EWSR1* break-apart FISH is usually straightforward, particularly in the context of a typical histological appearance and immunohistochemical profile with strong CD99 positivity which is characteristic to almost all cases of classic Ewing sarcoma (Fig. 2). Consistent with the *EWSR1-FLI* fusion protein being the crucial pathogenic driver in Ewing sarcoma, virtually all cells show an aberrant signal pattern here, even though several cells are formally non-evaluable due to “incompleteness” of the set of signals which may also be due to the chosen thickness of the section. However, frequently the evaluation of break-apart FISH assays is easier in thinner sections, since less focusing is required to avoid overlooking signals that may be in different focal planes.
13. Detection of a rearrangement in the *EWSR1* gene locus may be difficult in case of a $t(21;22)(q22;q12)$ chromosomal translocation resulting in an *EWSR1-ERG* gene fusion. In such cases, the FISH pattern may be only modestly aberrant or may even appear entirely normal. This is due to a more complex mechanism of rearrangement which is not a “simple” balanced translocation as known from *EWSR1-FLII*. Due to *EWSR1* and *ERG* having opposite directions of transcription, the generation of an *EWSR1-ERG* fusion associates with two breaks in the *EWSR1* locus and an *EWSR1* inversion before *EWSR1* is inserted in the *ERG* locus 21q22.3 (Fig. 3). Depending on the distance between the two breaks in 22q12, *EWSR1* break-apart FISH assays may fail to detect the *EWSR1* gene rearrangement. As the resolution of FISH is not optimal for breaks <2 Mb apart, smaller *EWSR1* breaks in Ewing sarcoma with *EWSR1-*

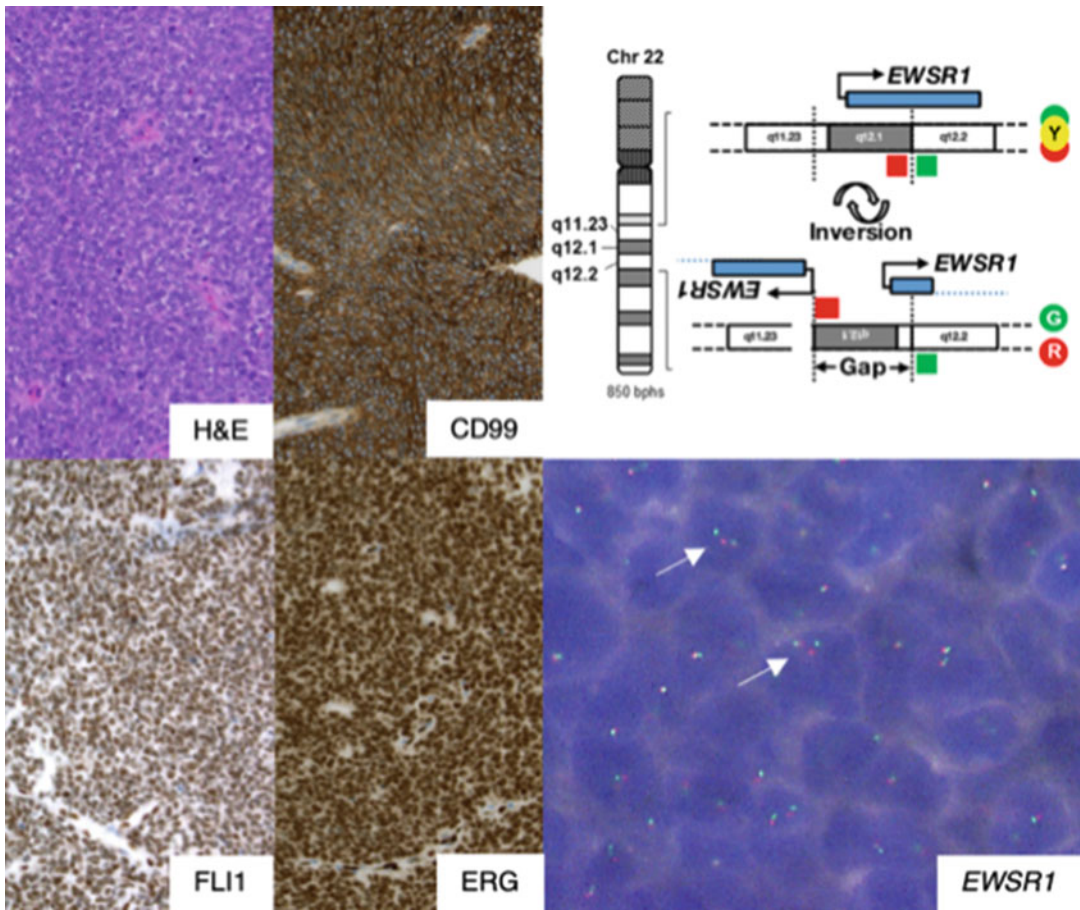


Fig. 3 (left) Ewing sarcoma with characteristic morphology, positivity for CD99 and FLI1, and strong ERG expression (comparable with ERG levels in endothelial cells). **(right, top)** Schematic view of the complex mechanism of the t(21;22) translocation, involving two 22q12 breaks (including the *EWSR1* break) followed by an inversion. Depending on the size of the inverted segment, *EWSR1* break-apart FISH may or may not be able to detect the rearrangement. **(right, bottom)** *EWSR1* break-apart FISH with the majority of cells showing yellow fusion signals without primary evidence of a rearrangement; particular attention must be paid to cells with discernible separated red and green signals at a distance <2 signal diameters or other aberrant patterns (e.g., fusion signal plus consistent isolated red or separate green signal). In such situations, ERG immunopositivity (restricted to Ewing sarcoma with *ERG* translocation) may provide a valuable clue helping to avoid a false negative evaluation

ERG fusion may not be detected by FISH [25]. The inclusion of ERG immunohistochemistry may help in cases with Ewing-typical morphology and bright CD99 positivity as strong ERG positivity in Ewing sarcoma has been reported to be restricted to *ERG*-rearranged tumors. *ERG* FISH or RNA-based analyses to document *EWSR1-ERG* rearrangement (RT-PCR or RNASeq) may be added to molecularly substantiate the diagnosis.

14. Another type of *EWSRI*- non-*ETS* rearranged SRCT biologically differing from Ewing sarcoma is *EWSRI-PATZI* rearranged small round and/or spindled sarcoma, often displaying a fibrous stroma and diverse immunophenotypes, including evidence of neural or myogenic differentiation [26]. The *EWSRI* rearrangement may be missed in *EWSRI* break-apart assays due to the short distance of ~2 Mb between *EWSRI* and *PATZI* (both located 22q12.2). However, polysomy 22q12, a finding very uncommon in Ewing sarcoma, has been reported in *EWSRI-PATZI* sarcoma. Secondary driver mutations in cell-cycle genes affecting *CDKN2A* in 71% of the cases have been reported in *EWSRI-PATZI* sarcomas [26].
15. In *EWSRI*-non-*ETS* rearranged SRCT an aberrant pattern to be observed employing *EWSRI* break-apart FISH assays is numerous copies of the proximal region flanking the *EWSRI* gene which was shown to be due to an amplification of the *EWSRI-NEATc2* gene fusion [27–29] (Fig. 4). Morphologically, apart from their (inconsistent) CD99 positivity, *EWSRI-NEATc2* rearranged tumors differ from classic Ewing sarcoma showing a pattern of cells predominantly arranged in cords or small nests in a myxohyaline stromal background so that misinterpretation as classic Ewing sarcoma can be avoided. Very rarely *FUS* may substitute for *EWSRI* in *NEATc2*-rearranged tumors. Interestingly, however, there is no evidence of an amplification of *FUS*-related signals in *FUS-NEATc2* rearranged sarcomas, and the mRNA expression profiles of *EWSRI-NEATc2* and *FUS-NEATc2* sarcomas were shown to differ [30].
16. Cases with typical morphology of Ewing sarcoma lacking an *EWSRI* rearrangement should be considered to harbor translocations of the *FUS* gene which occurs in <1% of Ewing sarcomas [13, 25, 31]. *FUS* rearrangements in Ewing sarcoma may lead to a fusion of *FUS* with the *ETS* genes *ERG* or *FEV*, but not *FLI1* (Fig. 5). As principally valid for all break-apart FISH assays, unbalanced rearrangements may occur, leading to aberrant signals which do not comply with the criterion of “signal completeness” [25]. Importantly, *FUS-ERG* fusion genes have also been described in acute myeloid leukemia [32], and in cases of suboptimal morphology and limited marker expression, a manifestation of myeloid sarcoma with CD99 expression may be misinterpreted as Ewing sarcoma.
17. The group of small round cell sarcomas of bone and soft tissue has recently been extended by the identification *CIC*-rearranged sarcomas which were formerly subsumed as Ewing-like sarcoma, a term which is not recommended anymore. A classic hint at a *CIC*-rearrangement is the aspect of a round cell (and also spindle cell or epithelioid) tumor showing increased

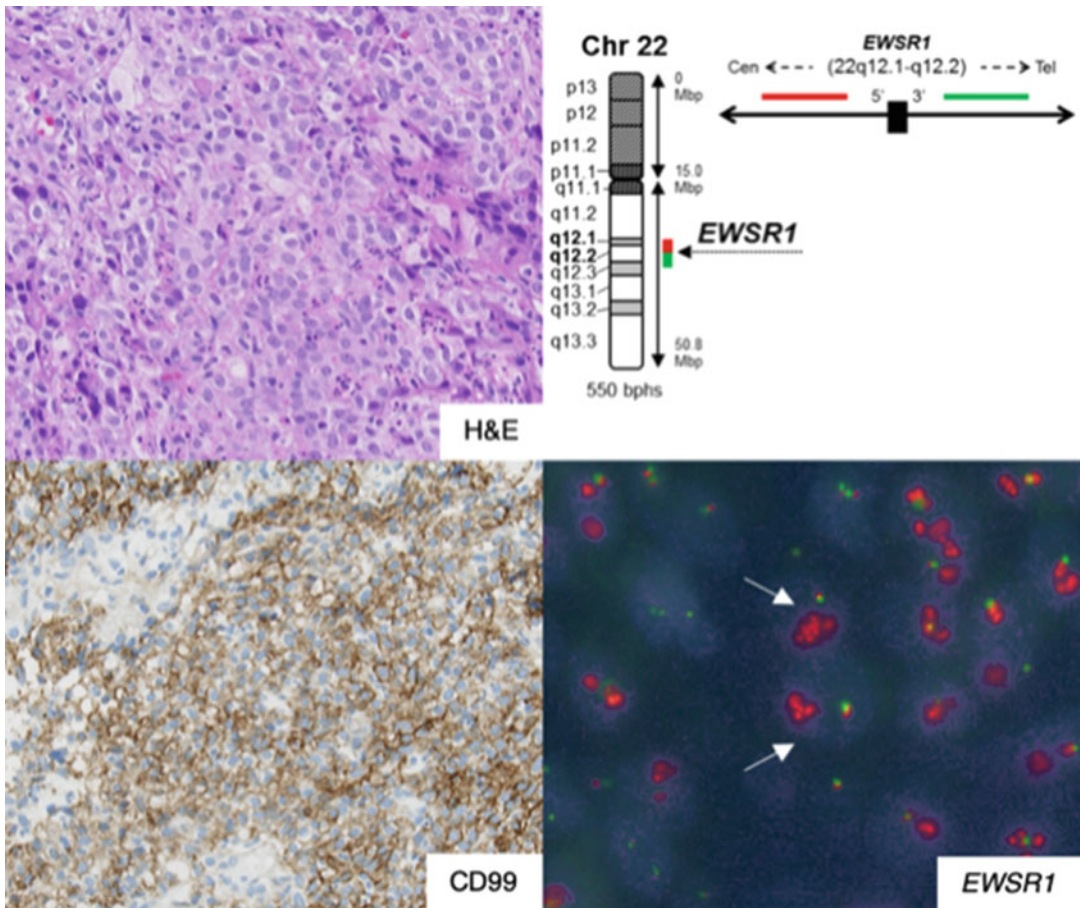


Fig. 4 (left) Case of an *EWSR1-NFATc2* translocated sarcoma showing a non-round cell morphology with isomorphic epithelioid cells, often in a myxohyaline background, immunohistochemically positive for CD99. **(right, top)** Schematic view of the *EWSR1* break-apart FISH approach. **(right, bottom)** *EWSR1* break-apart FISH showing multiple cells (arrows) with one (yellow) fused signal and low-grade amplified red signals, suggesting rearrangement of the *EWSR1* gene associated with additional chromosomal aberrations. Against the background of a typical morphology, this particular pattern is characteristic of *EWSR1-NFATc2* rearranged sarcomas and is not suggestive of classic Ewing sarcoma

nuclear pleomorphism with inconsistent patchy CD99 positivity and coexpression of WT1 and ETV4 [23]. Making up more than 95%, $t(4;19)(q35;q13)$ or $t(10;19)(q26;q13)$ represent the most frequent *CIC*-rearrangements and involve one of the *DUX4* genes present in the human genome [17]. Less than 5% of *CIC*-sarcomas cases harbor fusions of *CIC* to non-*DUX4* partners as *FOXO4*, *NUTM*, and *NUTM2A* [18, 19, 33]. Applying break-apart FISH approaches, a subgroup of 14% of *CIC*-rearranged sarcomas has been shown to be missed [34], which may partly be due to incomplete signal patterns subsequent to *CIC* telomeric deletion [16] (Fig. 6). It is therefore mandatory to synoptically evaluate microscopic and

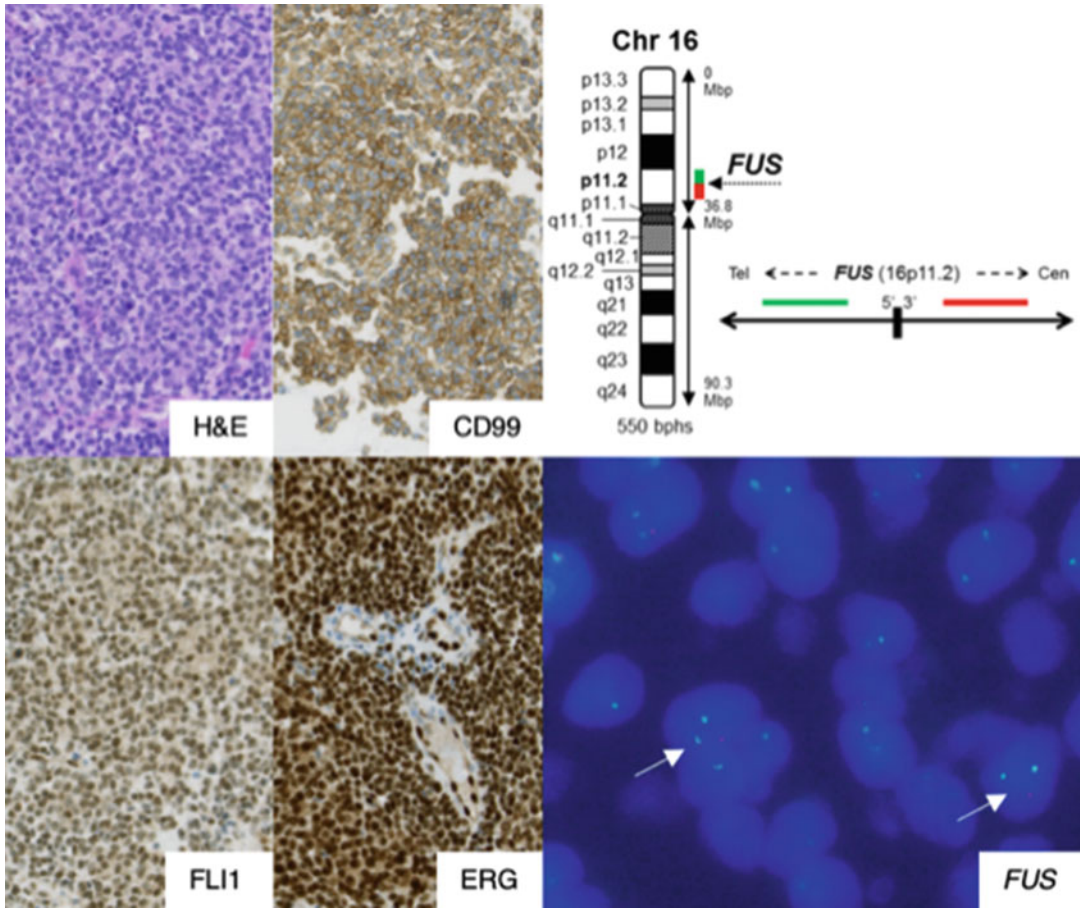


Fig. 5 (left) Ewing sarcoma with characteristic morphology, positivity for CD99, FLI1 and strong ERG expression (comparable with ERG levels in endothelial cells). **(right, top)** Schematic view of the *FUS* break-apart FISH approach employed in molecular routine diagnostics. **(right, bottom)** *FUS* break-apart FISH showing cells with a classic break-apart pattern with one yellow signal combined with separated red and green signals with a distance >2 signal diameters (arrows). Note that “incomplete” sets of signals may occur due to unbalanced rearrangements

immunohistochemical findings and to consider the possibility of false negativity of *CIC* break-apart FISH assays. Confirmatory RT-PCR approaches are difficult given the diversity of possible *CIC-DUX4* gene fusions on the one hand and the rare occurrence of non-*DUX4* fusions on the other hand [34]. NGS-based RNA sequencing approaches are therefore valuable as alternative diagnostic tool in *CIC*-rearrangements but may also be leaky in *CIC-DUX4* rearrangements due to aberrant bioinformatics algorithms (own observation).

18. Sarcomas with *BCOR* (Xp11.4) alterations represent another member of the group of small round cell sarcomas of bone and soft tissue. Characteristic *BCOR* overexpression may either be

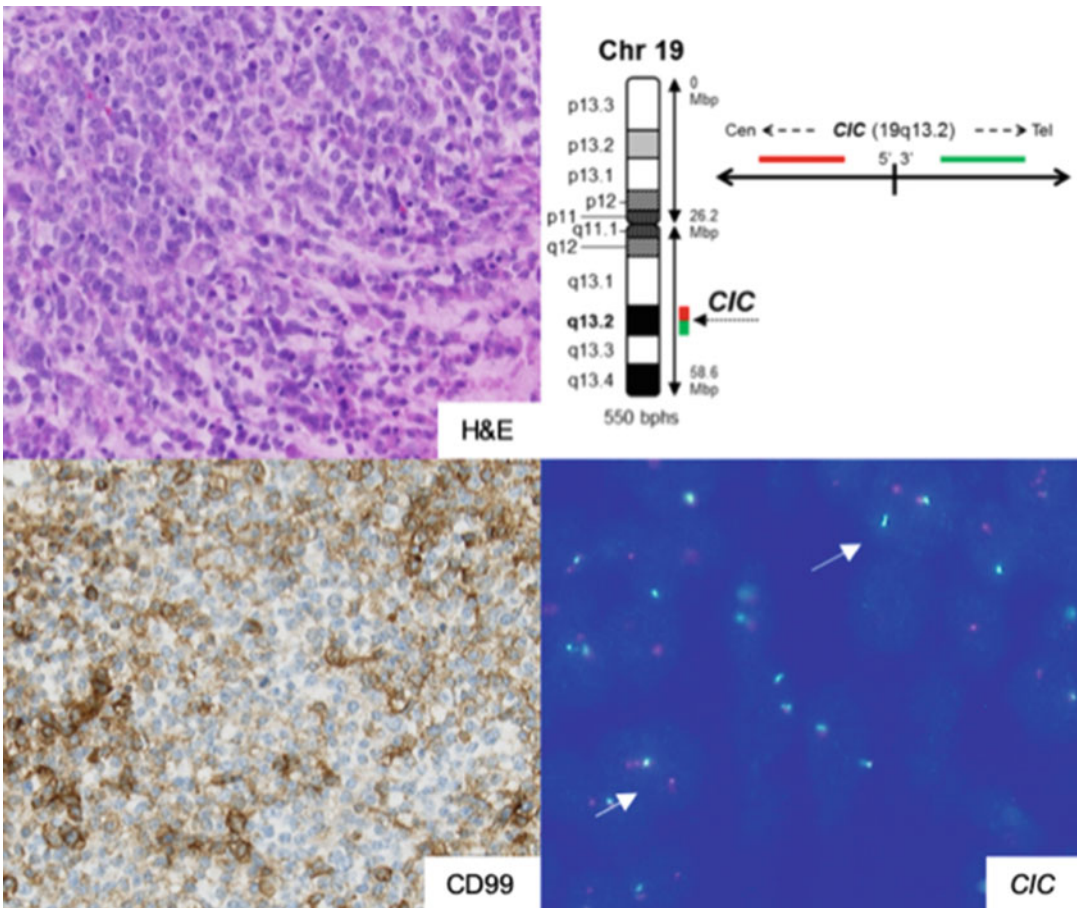


Fig. 6 (left) *CIC*-rearranged sarcomas showing a higher degree of histological and cytological variability compared to classic Ewing sarcoma and inconsistent patchy CD99 positivity. (right, top) Schematic view of the *CIC* break-apart FISH approach employed in molecular routine diagnostics. (right, bottom) *CIC* break-apart FISH showing cells with a classic break-apart pattern with one yellow signal combined with separated red and green signals with a distance >2 signal diameters (closed arrowhead). Note that “incomplete” sets of signals may occur due to unbalanced rearrangements (open arrowhead). It has been shown that up to 14% of *CIC*-rearrangements may be missed by FISH approaches

realized through classic *BCOR* rearrangements (most frequently intrachromosomal *BCOR-CCNB3* inversions, followed by rarer interchromosomal *BCOR-MAML2* or *ZC3H7B-BCOR* gene fusions) or *BCOR* internal tandem duplication [20–22, 35, 36]. Morphological patterns hinting at a possible *BCOR* alteration is the aspect of primitive round to spindle cells arranged in sheets or showing fascicular growth with a variably myxoid stroma with delicate vasculature, associated with immunohistochemical positivity for *BCOR*, and in case of a *BCOR-CCNB3* paracentric inversion, Cyclin B3. Given the proximity of the *BCOR* and *CCNB3* genes on chromosome Xp which approximates 10 Mb, it is not

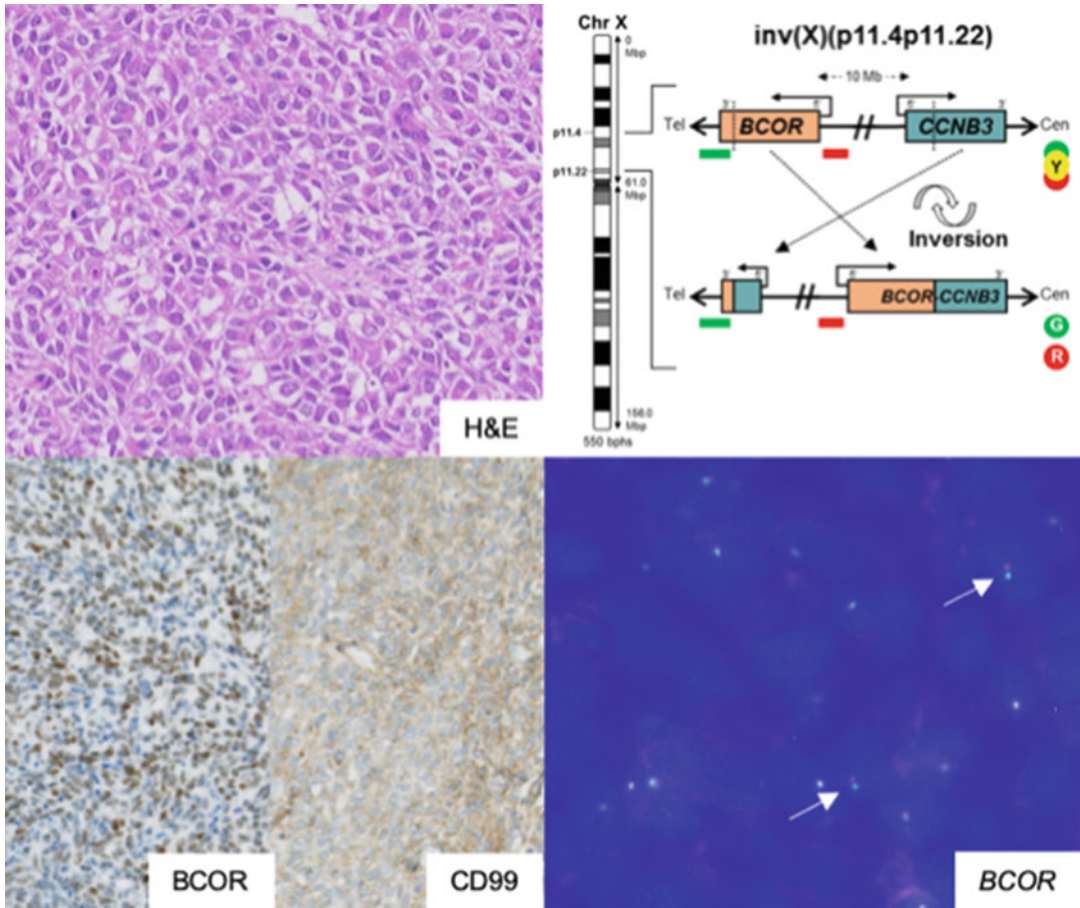


Fig. 7 (left) Case of a *BCOR-CCNB3* translocated sarcoma with primitive round to spindle cells arranged in sheets, associated with characteristic delicate vasculature. CD99 expression is heterogeneous, BCOR immunohistochemistry may be of help as screening tool. **(right, top)** Schematic view of the *BCOR-CCNB3* inversion, which, due to the proximity of the *BCOR* and *CCNB3* genes on chromosome Xp, may be missed by *BCOR* break-apart FISH. Arrows indicated cells with only slightly increased distances between red and green signals pointing to a rearrangement; *BCOR-CCNB3* RT-PCR should be performed in such a situation to confirm the rearrangement

surprising that a *BCOR-CCNB3* inversion may not readily be detected by *BCOR* break-apart FISH (Fig. 7); it has furthermore to be considered that *BCOR* is located on the X chromosome so that only one copy is present in the male genome. Due to the uniformity of the *BCOR-CCNB3* rearrangement it may easily be detected by RT-PCR assays [37].

Acknowledgments

We thank Gay Naber, Inka Buchroth, Christin Fehmer, Simone Paisdzior, and Lydia Falcker for excellent technical assistance.

References

- Ambros IM, Ambros PF, Strehl S, Kovar H, Gadner H, Salzer-Kuntschik M (1991) MIC2 is a specific marker for Ewing's sarcoma and peripheral primitive neuroectodermal tumors. Evidence for a common histogenesis of Ewing's sarcoma and peripheral primitive neuroectodermal tumors from MIC2 expression and specific chromosome aberration. *Cancer* 67(7):1886–1893. [https://doi.org/10.1002/1097-0142\(19910401\)67:7<1886::aid-cncr2820670712>3.0.co;2-u](https://doi.org/10.1002/1097-0142(19910401)67:7<1886::aid-cncr2820670712>3.0.co;2-u)
- Hornick JL (2014) Novel uses of immunohistochemistry in the diagnosis and classification of soft tissue tumors. *Mod Pathol* 27(Suppl 1):S47–S63. <https://doi.org/10.1038/modpathol.2013.177>
- Llombart-Bosch A, Machado I, Navarro S, Bertoni F, Bacchini P, Alberghini M, Karzeladze A, Savelov N, Petrov S, Alvarado-Cabrero I, Mihaila D, Terrier P, Lopez-Guerrero JA, Picci P (2009) Histological heterogeneity of Ewing's sarcoma/PNET: an immunohistochemical analysis of 415 genetically confirmed cases with clinical support. *Virchows Arch* 455(5):397–411. <https://doi.org/10.1007/s00428-009-0842-7>
- Russell-Goldman E, Hornick JL, Qian X, Jo VY (2018) NKX2.2 immunohistochemistry in the distinction of Ewing sarcoma from cytomorphologic mimics: diagnostic utility and pitfalls. *Cancer Cytopathol* 126(11):942–949. <https://doi.org/10.1002/cncy.22056>
- Shibuya R, Matsuyama A, Nakamoto M, Shiba E, Kasai T, Hisaoka M (2014) The combination of CD99 and NKX2.2, a transcriptional target of EWSR1-FLI1, is highly specific for the diagnosis of Ewing sarcoma. *Virchows Arch* 465(5):599–605. <https://doi.org/10.1007/s00428-014-1627-1>
- Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G et al (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 359(6391):162–165. <https://doi.org/10.1038/359162a0>
- Zucman J, Melot T, Desmaze C, Ghysdael J, Plougastel B, Peter M, Zucker JM, Triche TJ, Sheer D, Turc-Carel C et al (1993) Combinatorial generation of variable fusion proteins in the Ewing family of tumours. *EMBO J* 12(12):4481–4487
- Jeon IS, Davis JN, Braun BS, Sublett JE, Rousel MF, Denny CT, Shapiro DN (1995) A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* 10(6):1229–1234
- Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H, Delattre O (1997) A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene* 14(10):1159–1164. <https://doi.org/10.1038/sj.onc.1200933>
- Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, Denny CT (1994) A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat Genet* 6(2):146–151. <https://doi.org/10.1038/ng0294-146>
- Urano F, Umezawa A, Yabe H, Hong W, Yoshida K, Fujinaga K, Hata J (1998) Molecular analysis of Ewing's sarcoma: another fusion gene, EWS-ELAF, available for diagnosis. *Jpn J Cancer Res* 89(7):703–711. <https://doi.org/10.1111/j.1349-7006.1998.tb03274.x>
- Ng TL, O'Sullivan MJ, Pallen CJ, Hayes M, Clarkson PW, Winstanley M, Sorensen PH, Nielsen TO, Horsman DE (2007) Ewing sarcoma with novel translocation t(2;16) producing an in-frame fusion of FUS and FEV. *J Mol Diagn* 9(4):459–463. <https://doi.org/10.2353/jmoldx.2007.070009>
- Shing DC, McMullan DJ, Roberts P, Smith K, Chin SF, Nicholson J, Tillman RM, Ramani P, Cullinane C, Coleman N (2003) FUS/ERG gene fusions in Ewing's tumors. *Cancer Res* 63(15):4568–4576
- Desmaze C, Zucman J, Delattre O, Thomas G, Aurias A (1992) Unicolor and bicolor in situ hybridization in the diagnosis of peripheral neuroepithelioma and related tumors. *Genes Chromosomes Cancer* 5(1):30–34. <https://doi.org/10.1002/gcc.2870050105>
- Antonescu CR, Dal Cin P (2014) Promiscuous genes involved in recurrent chromosomal translocations in soft tissue tumours. *Pathology* 46(2):105–112. <https://doi.org/10.1097/PAT.0000000000000049>
- Antonescu CR, Owosho AA, Zhang L, Chen S, Deniz K, Huryn JM, Kao YC, Huang SC, Singer S, Tap W, Schaefer IM, Fletcher CD (2017) Sarcomas with CIC-rearrangements are a distinct pathologic entity with aggressive outcome: a clinicopathologic and molecular study of 115 cases. *Am J Surg Pathol* 41(7):941–949. <https://doi.org/10.1097/PAS.0000000000000846>
- Italiano A, Sung YS, Zhang L, Singer S, Maki RG, Coindre JM, Antonescu CR (2012) High prevalence of CIC fusion with double-

- homeobox (DUX4) transcription factors in EWSR1-negative undifferentiated small round cell sarcomas. *Genes Chromosomes Cancer* 51(3):207–218. <https://doi.org/10.1002/gcc.20945>
18. Sugita S, Arai Y, Tonooka A, Hama N, Totoki Y, Fujii T, Aoyama T, Asanuma H, Tsukahara T, Kaya M, Shibata T, Hasegawa T (2014) A novel CIC-FOXO4 gene fusion in undifferentiated small round cell sarcoma: a genetically distinct variant of Ewing-like sarcoma. *Am J Surg Pathol* 38(11):1571–1576. <https://doi.org/10.1097/PAS.000000000000286>
 19. Le Loarer F, Pissaloux D, Watson S, Godfraind C, Galmiche-Rolland L, Silva K, Mayeur L, Italiano A, Michot A, Pierron G, Vasiljevic A, Ranchere-Vince D, Coindre JM, Tirode F (2019) Clinicopathologic features of CIC-NUTM1 sarcomas, a new molecular variant of the family of CIC-fused sarcomas. *Am J Surg Pathol* 43(2):268–276. <https://doi.org/10.1097/PAS.0000000000001187>
 20. Kao YC, Owosho AA, Sung YS, Zhang L, Fujisawa Y, Lee JC, Wexler L, Argani P, Swanson D, Dickson BC, Fletcher CDM, Antonescu CR (2018) BCOR-CCNB3 fusion positive sarcomas: a clinicopathologic and molecular analysis of 36 cases with comparison to morphologic spectrum and clinical behavior of other round cell sarcomas. *Am J Surg Pathol* 42(5):604–615. <https://doi.org/10.1097/PAS.0000000000000965>
 21. Kao YC, Sung YS, Zhang L, Jungbluth AA, Huang SC, Argani P, Agaram NP, Zin A, Alaggio R, Antonescu CR (2016) BCOR overexpression is a highly sensitive marker in round cell sarcomas with BCOR genetic abnormalities. *Am J Surg Pathol* 40(12):1670–1678. <https://doi.org/10.1097/PAS.0000000000000697>
 22. Pierron G, Tirode F, Lucchesi C, Reynaud S, Ballet S, Cohen-Gogo S, Perrin V, Coindre JM, Delattre O (2012) A new subtype of bone sarcoma defined by BCOR-CCNB3 gene fusion. *Nat Genet* 44(4):461–466. <https://doi.org/10.1038/ng.1107>
 23. Specht K, Sung YS, Zhang L, Richter GH, Fletcher CD, Antonescu CR (2014) Distinct transcriptional signature and immunoprofile of CIC-DUX4 fusion-positive round cell tumors compared to EWSR1-rearranged Ewing sarcomas: further evidence toward distinct pathologic entities. *Genes Chromosomes Cancer* 53(7):622–633. <https://doi.org/10.1002/gcc.22172>
 24. Summersgill BM, Shipley JM (2010) Fluorescence in situ hybridization analysis of formalin fixed paraffin embedded tissues, including tissue microarrays. *Methods Mol Biol* 659:51–70. https://doi.org/10.1007/978-1-60761-789-1_4
 25. Chen S, Deniz K, Sung YS, Zhang L, Dry S, Antonescu CR (2016) Ewing sarcoma with ERG gene rearrangements: a molecular study focusing on the prevalence of FUS-ERG and common pitfalls in detecting EWSR1-ERG fusions by FISH. *Genes Chromosomes Cancer* 55(4):340–349. <https://doi.org/10.1002/gcc.22336>
 26. Bridge JA, Sumegi J, Druta M, Bui MM, Henderson-Jackson E, Linos K, Baker M, Walko CM, Millis S, Brohl AS (2019) Clinical, pathological, and genomic features of EWSR1-PATZ1 fusion sarcoma. *Mod Pathol* 32(11):1593–1604. <https://doi.org/10.1038/s41379-019-0301-1>
 27. Bode-Lesniewska B, Fritz C, Exner GU, Wagner U, Fuchs B (2019) EWSR1-NFATC2 and FUS-NFATC2 gene fusion-associated Mesenchymal tumors: clinicopathologic correlation and literature review. *Sarcoma* 2019:9386390. <https://doi.org/10.1155/2019/9386390>
 28. Sadri N, Barroeta J, Pack SD, Abdullaev Z, Chatterjee B, Puthiyaveetil R, Brooks JS, Barr FG, Zhang PJ (2014) Malignant round cell tumor of bone with EWSR1-NFATC2 gene fusion. *Virchows Arch* 465(2):233–239. <https://doi.org/10.1007/s00428-014-1613-7>
 29. Szuhai K, Ijszenga M, de Jong D, Karseladze A, Tanke HJ, Hogendoorn PC (2009) The NFATc2 gene is involved in a novel cloned translocation in a Ewing sarcoma variant that couples its function in immunology to oncology. *Clin Cancer Res* 15(7):2259–2268. <https://doi.org/10.1158/1078-0432.CCR-08-2184>
 30. Watson S, Perrin V, Guillemot D, Reynaud S, Coindre JM, Karanian M, Guinebretiere JM, Freneaux P, Le Loarer F, Bouvet M, Galmiche-Rolland L, Larousserie F, Longchamps E, Ranchere-Vince D, Pierron G, Delattre O, Tirode F (2018) Transcriptomic definition of molecular subgroups of small round cell sarcomas. *J Pathol* 245(1):29–40. <https://doi.org/10.1002/path.5053>
 31. Sankar S, Lessnick SL (2011) Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet* 204(7):351–365. <https://doi.org/10.1016/j.cancergen.2011.07.008>
 32. Panagopoulos I, Aman P, Fioretos T, Hoglund M, Johansson B, Mandahl N, Heim S, Behrendtz M, Mitelman F (1994) Fusion of the FUS gene with ERG in acute

- myeloid leukemia with t(16;21)(p11;q22). *Genes Chromosomes Cancer* 11(4):256–262. <https://doi.org/10.1002/gcc.2870110408>
33. Sugita S, Arai Y, Aoyama T, Asanuma H, Mukai W, Hama N, Emori M, Shibata T, Hasegawa T (2017) NUTM2A-CIC fusion small round cell sarcoma: a genetically distinct variant of CIC-rearranged sarcoma. *Hum Pathol* 65:225–230. <https://doi.org/10.1016/j.humpath.2017.01.012>
 34. Yoshida A, Arai Y, Kobayashi E, Yonemori K, Ogura K, Hama N, Mukai W, Motoi T, Kawai A, Shibata T, Hiraoka N (2017) CIC break-apart fluorescence in-situ hybridization misses a subset of CIC-DUX4 sarcomas: a clinicopathological and molecular study. *Histopathology* 71(3):461–469. <https://doi.org/10.1111/his.13252>
 35. Kao YC, Sung YS, Zhang L, Huang SC, Argani P, Chung CT, Graf NS, Wright DC, Kellie SJ, Agaram NP, Ludwig K, Zin A, Alaggio R, Antonescu CR (2016) Recurrent BCOR internal tandem duplication and YWHAE-NUTM2B fusions in soft tissue undifferentiated round cell sarcoma of infancy: overlapping genetic features with clear cell sarcoma of kidney. *Am J Surg Pathol* 40(8):1009–1020. <https://doi.org/10.1097/PAS.0000000000000629>
 36. Specht K, Zhang L, Sung YS, Nucci M, Dry S, Vaiyapuri S, Richter GH, Fletcher CD, Antonescu CR (2016) Novel BCOR-MAML3 and ZC3H7B-BCOR gene fusions in undifferentiated small blue round cell sarcomas. *Am J Surg Pathol* 40(4):433–442. <https://doi.org/10.1097/PAS.0000000000000591>
 37. Puls F, Niblett A, Marland G, Gaston CL, Douis H, Mangham DC, Sumathi VP, Kindblom LG (2014) BCOR-CCNB3 (Ewing-like) sarcoma: a clinicopathologic analysis of 10 cases, in comparison with conventional Ewing sarcoma. *Am J Surg Pathol* 38(10):1307–1318. <https://doi.org/10.1097/PAS.0000000000000223>



Molecular Approaches to Diagnosis in Ewing Sarcoma: RT-PCR

Carlos Rodríguez-Martín and Javier Alonso

Abstract

Ewing sarcoma is a rare and aggressive tumor that affects children and young adults. Ewing sarcomas are characterized by specific chromosomal translocations that give rise to fusion transcripts that codify for aberrant transcription factors. More than 95% of Ewing sarcoma harbor translocations that produce the fusion of the EWSR1 gene with the transcription factors FLI1 or ERG. This feature can be used to diagnose this entity unambiguously.

In this chapter we describe a RT-PCR method that allows for the detection of the most frequent alterations with elevated specificity and sensitivity which is able to distinguish among the different types of fusions. The method is fast and economical, and can be carried out with the conventional equipment available in any molecular biology laboratory.

Key words Ewing sarcoma, EWSR1-FLI1, EWSR1-ERG, Molecular diagnosis, RT-PCR

1 Introduction

Ewing sarcoma is a rare tumor arising in bone that affects children and young adults. Histologically, Ewing sarcomas are characterized by small round cells with no apparent signs of differentiation arranged in dense cell sheets. Thus, Ewing sarcoma belong to the so-called small round cell tumors that include several tumors with similar apparent histology (e.g., alveolar rhabdomyosarcoma, synovial sarcoma, desmoplastic small round cell tumors, and others) [1, 2]. Therefore, molecular markers beyond conventional histology are necessary to diagnose accurately this type of tumors. CD99 membrane staining is routinely used in the differential diagnosis. However, this membrane protein is also expressed in other round cell sarcomas and lymphoblastic lymphoma and leukemia [1, 2], so despite being useful as an initial approach, this marker is not definitive. Since proper diagnosis is essential to provide a tumor-specific treatment to patients, other markers must be used in the differential diagnosis.

Table 1
List of chromosomal translocations and associated fusion transcripts identified in Ewing sarcoma

Chromosomal translocations	Genes involved	Frequency (%)
t(11;22)(q24;q12)	EWSR1-FLI1	85
t(21;22)(q22;q12)	EWSR1-ERG	10
t(2;22)(q33;q12)	EWSR1-FEV	<1
t(7;22)(p22;q12)	EWSR1-ETV1	<1
t(17;22)(q21;q12)	EWSR1-ETV4	<1
t(16;21)(p11;q22)	FUS-ERG	<1
t(2;16)(q35;p11)	FUS-FEV	<1

The unambiguous molecular characteristic that defines Ewing sarcoma is the presence of chromosomal translocations that give rise to aberrant transcription factors that drive the oncogenic process. These aberrant transcription factors are formed by the fusion of the N-terminal region of a gene of the FET family (usually EWSR1, and more rarely FUS) and the C-terminal region of a gene of the ETS family of transcription factors (most common being FLI1 and ERG) (Table 1) [2–4]. Thus, it is widely accepted that identification of these specific alterations is the bona fide approach to diagnosing Ewing sarcoma. In addition, for each chromosomal translocation type, there are different subtypes depending on where breakpoints take place. This variable breakpoints cause different exon combinations between the genes involved in the fusion transcript. The most frequent subtype is formed by exon 7 of EWSR1 gene and exon 6 of FLI1 gene (Fig. 1) [3].

Two methods can be used to identify these alterations in a molecular diagnostic context: (a) FISH (fluorescent in situ hybridization) which can be used to detect the chromosomal alteration itself and (b) a specific RT-PCR assay (reverse transcription-polymerase chain reaction) used to detect the fusion mRNA expressed from the fused genes. Each of these methods has its advantages and disadvantages.

- Specificity: Most diagnostic kits used to detect this type of translocations by FISH use break apart probes located on both sides of the EWSR1 breakpoint. Thus, these probes detect EWSR1 break in Ewing sarcoma, but also other fusion genes characteristics of other unrelated tumors in which EWSR1 is also involved, such as EWSR1-WT1 in desmoplastic small round cell tumor, EWSR1-DDIT3 in myxoid liposarcoma or EWSR1-ATF1 in clear cell sarcoma [4, 5]. By contrast, RT-PCR is specific for each tumor-specific gene fusion since primers can

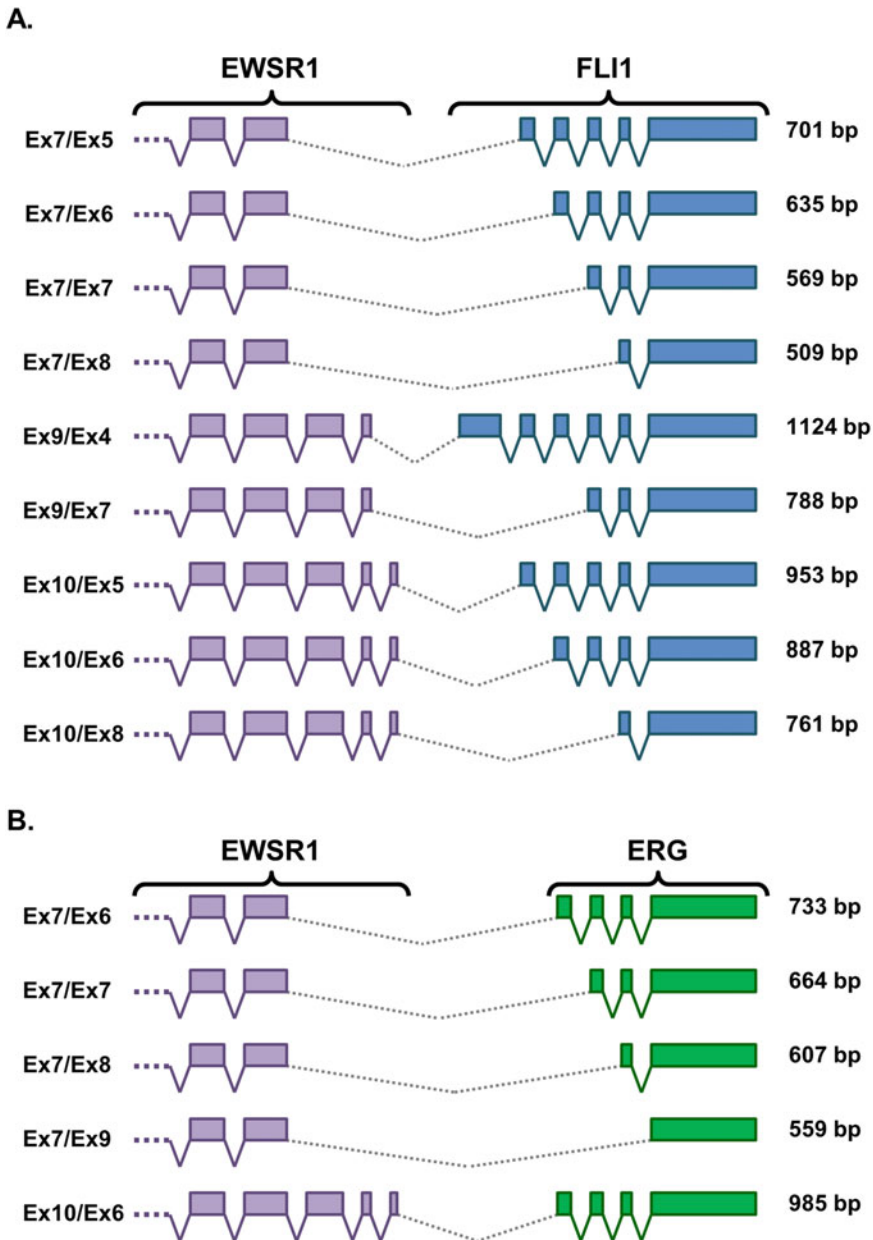


Fig. 1 Schematic representation of the different EWSR1-FLI1 (a) and EWSR1-ERG (b) fusion transcripts that can be detected with the assay described in this chapter. The exons involved in the fusion transcripts and the predicted amplicon size obtained in the RT-PCR assay are indicated for each variant with the set of specific primers

be specifically designed in each case. Furthermore, RT-PCR provides information about the exact exons involved in the fusion gene, which cannot be done using FISH probes.

- Sensitivity: The sensitivity of FISH methods depends on the number of cells that could be analyzed. Normally, it is recommended to analyze a minimum of 50 cells, considering a positive results if at least 10% of the cells show signs of chromosomal translocations [6]. The number of tumor cells present in Ewing sarcoma biopsies is normally not the limiting factor in the majority of the cases. However, some samples such as tumor-infiltrated bone marrow could have a scarce number of tumor cells in relation to normal cells. By contrast, RT-PCR is sensitive enough to detect tumor cells in low proportion tumor cell-normal cell ratios, 1:1.000–10.000).
- Equipment: Different scientific equipment is necessary to carried out each technique. Conventional equipment available in a standard molecular laboratory, such as PCR-cabinets for sample processing and setting PCR reactions, centrifuges, programmable thermal cycles, electrophoresis devices and gel imaging systems, are necessary for RT-PCR. For FISH, hybridization ovens, fluorescence microscopy, and preferably specialized imaging software are required.
- Expertise: For FISH, highly specialized personnel are necessary to adequately interpret images. Interpretation of RT-PCR results is significantly simpler.
- Required sample type: It is highly recommended to use frozen biopsies when isolating RNA for RT-PCR. Although RT-PCR can be performed from RNA extracted from paraffin blocks, in our experience, RNA extracted for these samples tends to render low-quality results. By contrast, excellent results are obtained with FISH performed on samples fixed and embedded in paraffin blocks.

In this chapter we describe a specific, sensitive, fast and economical method to identify the characteristic gene fusions found in Ewing sarcoma. Here we explain a method to isolate RNA from hard bone samples and a specific RT-PCR assay to identify most frequent fusion transcripts. This protocol has been used for the last 20 years to analyze more than 400 tumor samples referred to our laboratory for molecular diagnosis.

2 Materials

All reagents used should be molecular biology grade or analytical grade. Molecular biology reactions should be set up with RNase and DNase-free water. Buffers for electrophoresis should be prepared using ultrapure water. We make small aliquots of all molecular biology buffers and solutions and keep them frozen at -20°C to avoid contaminations.

A.



B.

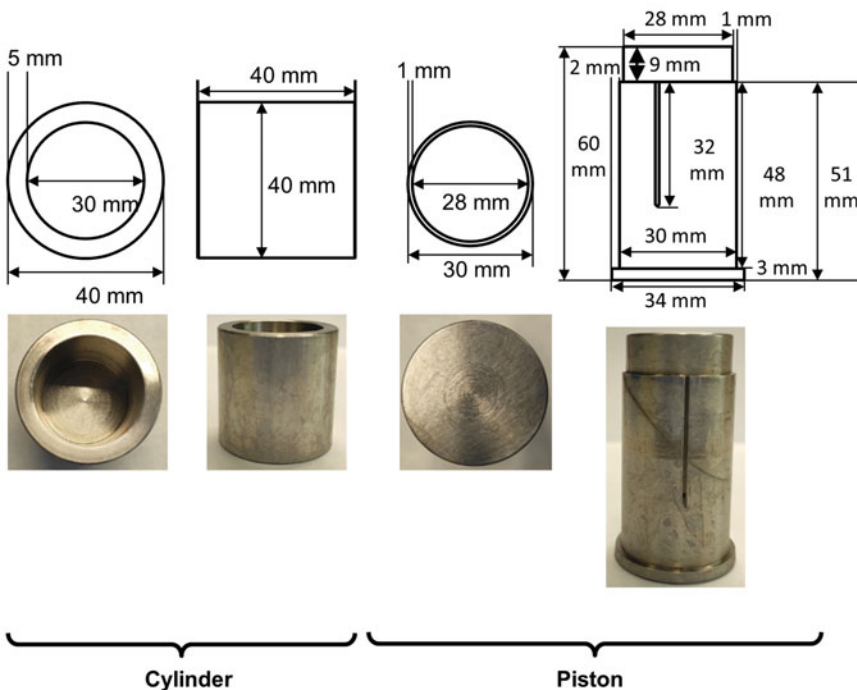


Fig. 2 (a) Materials used to powder tumor fragments: mortar, piston-cylinder set, spatula, tissue forceps. (b) Detailed description of the piston-cylinder set

2.1 RNA Extraction

1. Piston-cylinder set (Fig. 2).
2. Mortar.
3. Spatula.
4. Tissue forceps.
5. Guanidine thiocyanate and phenol mix, as for example TRI Reagent[®] (Sigma-Aldrich).
6. Chloroform.

7. Ethanol 70%.
8. Ethanol 80%.
9. RNA cleanup kit as for example RNeasy™ MinElute Cleanup Kit (Qiagen).
10. Tabletop centrifuge.
11. Refrigerated microcentrifuge.
12. Spectrophotometer as for example NanoDrop™ for RNA quantification.

2.2 Reverse Transcription (RT)

1. Reverse Transcriptase and associated buffer (e.g., M-MLV Reverse Transcriptase from Invitrogen, including reverse transcriptase buffer and 100 mM DTT solution).
2. 10 mM dNTPs (2.5 mM each)
3. 25 mM MgCl₂
4. 50 μM random hexamers
5. 20 U/μL RNase inhibitor
6. Molecular biology grade RNase and DNase-free water.
7. Thermocycler.

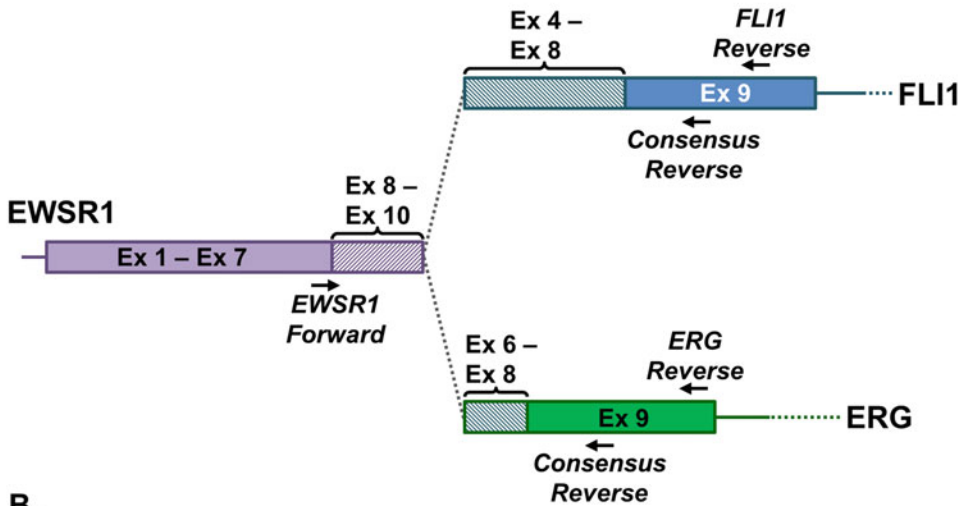
2.3 Polymerase Chain Reaction (PCR)

1. Taq DNA polymerase and associated buffers. In our case, we use AmpliTaq Gold™ DNA polymerase from Applied Biosystems which include polymerase buffer and MgCl₂ solution. Other Taq DNA polymerases can work equally well.
2. 10 mM dNTPs (2.5 mM each)
3. 25 mM MgCl₂
4. Primers diluted at 10 μM in 10 mM Tris-HCl, pH 8 (Table 2, Fig. 3).
5. Molecular biology grade RNase and DNase-free water.
6. Thermocycler.

Table 2
Primer sequences used in PCR reactions

Primer	5' > 3' sequence
EWSR1, forward	TCCTACAGCCAAGCTCCAAGTC
FLI1/ERG consensus, reverse	TGTTGGGCTTGCTTTTCCGCTC
FLI1 specific, reverse	GTGACAGGCATGGAGGATGGA
ERG specific, reverse	GAGAAGGCATATGGCTGGTGG
TBP, forward	CAGCCGTTCCAGCAGTCAAC
TBP, reverse	TGTTGGTGGGTGAGCACAAG

A.



B.

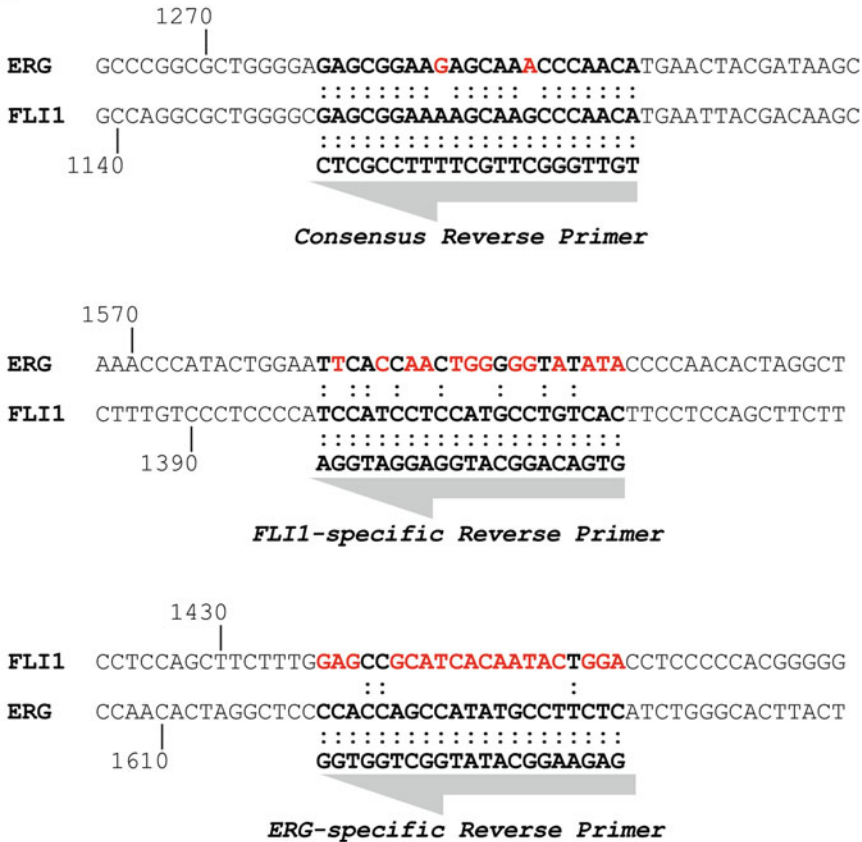


Fig. 3 (a) Schematic representation of EWSR1-FLI1 and EWSR1-ERG fusion transcripts indicating the relative location of PCR primers. **(b)** Fragments of the nucleotide sequence of FLI1 and ERG where the primers are located, indicating the base pairs with which they hybridize. Nucleotide position according to NM_002017.5 (FLI1) and NM_004449.4 (ERG) are indicated

2.4 Agarose Gel Electrophoresis

1. Electrophoresis buffer (TBE 5×): 0.45 M Tris, 0.45 M boric acid, 10 mM EDTA, pH 8.4. Dilute to 1× with ultrapure water.
2. Agarose for gel electrophoresis.
3. SYBR™ Safe or equivalent DNA gel stain.
4. Loading buffer.
5. TE Buffer (1×): 10 mM Tris–HCl, 1 mM EDTA, pH 8.
6. DNA molecular weight marker (e.g., a 100 bp ladder). Prepare working stock solution at 0.5 mg/mL in 1× TE buffer, 1× loading buffer.
7. Gel casting tray, combs, cell and power supply for horizontal electrophoresis.

3 Methods

3.1 RNA Extraction

1. Remove the cryovial containing the tumor from the ultrafreezer and transport it in a container with dry ice to the pre-PCR area (*see* **Notes 1** and **2**).
2. Fill a container with liquid nitrogen (*see* **Note 3**). Liquid nitrogen will be used to cool all the instruments used to homogenize the tumor in the next step.
3. Fill a mortar and the piston–cylinder set with liquid nitrogen to cool it (Fig. 2, **Note 4**). Introduce the spatula and forceps in the mortar that we will use to manipulate the piece of tumor. Fill with liquid nitrogen a 50 mL tube (do not close the tube).
4. Introduce the tumor fragment (30–100 mg) into the cylinder when the liquid nitrogen is completely evaporated. Immediately insert the piston into the cylinder and hit it hard several times with a hammer until the tumor is fully powdered.
5. Transfer the powdered tissue with the help of a spatula to the 50 mL tube filled with liquid nitrogen. Allow the liquid nitrogen to evaporate.
6. Add 1.5 mL TRI Reagent™ per 100 mg of tumor (minimum 3 mL).
7. Vortex at full speed until the powdered tissue is totally dissolved.
8. Incubate for 3 min at room temperature (*see* **Note 5**).
9. Centrifuge the lysate at $700 \times g$ for 1 min in a tabletop centrifuge at room temperature in order to precipitate bone debris.

10. Distribute the lysate in two 2 mL microcentrifuge tubes, 1.5 mL per tube, being careful not to transfer the insoluble material from the bottom of the 50 mL tube.
11. Add 300 μ L of chloroform to each 2 mL microcentrifuge tube.
12. Vortex at maximum speed 1 min.
13. Incubate for 5 min at room temperature (*see Note 6*).
14. Centrifuge tubes at $13,000 \times g$ for 30 min at 4 °C in a microcentrifuge.
15. Carefully remove the tubes from the microcentrifuge. At this stage two clearly differentiated phases can be observed. Transfer the upper phase (750 μ L) of each tube to two 1.5 mL microcentrifuge tube, 750 μ L per tube (*see Note 7*).
16. Add 750 μ L of 70% ethanol to each tube and mix by vortex. Keep at room temperature.
17. Place a column of the RNA cleanup kit in a 1.5 mL microcentrifuge tube (one column per tumor sample). Add 700 μ L of the mixture aqueous phase/ethanol of the previous step to the column.
18. Centrifuge at $8000 \times g$ for 15 s in a microcentrifuge at room temperature. Discard the eluate and repeat the same steps as many times as necessary so that all the mixture from the same sample has passed through the same column. Discard the eluate and change the column to a clean 1.5 mL microcentrifuge tube.
19. Add 500 μ L of Buffer RPE to the column. Centrifuge at $8000 \times g$ for 15 s. Discard the eluate.
20. Add 500 μ L of 80% ethanol to the column. Centrifuge at $8000 \times g$ for 2 min.
21. Discard the eluate and change the column to a clean 1.5 mL microcentrifuge tube. Centrifuge at full speed for 5 min to dry the column completely.
22. Place the column in a clean 1.5 mL microcentrifuge tube. Carefully add 14 μ L of RNase-free water to the center of the column being careful not to touch the membrane (*see Note 8*).
23. Incubate the column for 5 min at room temperature.
24. Centrifuge at full speed for 1 min to elute pure RNA.
25. Quantify the amount of RNA by spectrophotometric methods (e.g., NanoDrop™) and prepare a working RNA stock at 200 ng/ μ L using RNase-free water. Store RNA stocks at -80 °C until use.

3.2 Reverse Transcription

1. Thaw reagents for reverse transcription reaction on ice. Enzymes will be removed from the freezer at the time of use. Reverse transcription (RT) setup should be performed in the pre-PCR area (*see Note 2*), preferably in a PCR cabinet.

2. Prepare the following PCR tubes for each tumor RNA to be analyzed (keep the tubes on ice):
 - One tube for the EWSR1-FLI1 positive control (*see Note 9*).
 - One tube for the EWSR1-ERG positive control (*see Note 9*).
 - One tube for the RNA tumor sample.
 - One tube for the negative control (containing tumor RNA but not reverse transcriptase).
 - One tube for the no template control (containing only reagents).
3. Set up the reverse transcriptase reaction according to the following table. All volumes are indicated in microliters.

Reagent	Positive control EWSR1-FLI1	Positive control EWSR1-ERG	RNA tumor sample	No RT control	No template control
5× reverse transcriptase buffer	4	4	4	4	4
10 mM dNTPs (2.5 mM each)	2	2	2	2	2
25 mM MgCl ₂	2	2	2	2	2
0.1 M DTT	2	2	2	2	2
50 μM random hexamers (<i>see Note 10</i>)	0.5	0.5	0.5	0.5	0.5
20 U/μL RNase inhibitor	0.5	0.5	0.5	0.5	0.5
50 U/μL M-MLV reverse transcriptase	0.5	0.5	0.5	–	0.5
RNase/DNase free water	3.5	3.5	3.5	4.0	8.5
200 ng/μL RNA EWSR1-FLI1 control	5	–	–	–	–
200 ng/μL RNA EWSR1-ERG control	–	5	–	–	–
200 ng/μL RNA tumor sample	–	–	5	5	–
Total volume	20	20	20	20	20

4. Vortex the tubes gently and centrifuge. Place the tubes in the thermocycler and launch the following program:
 - 25 °C, 10 min.
 - 42 °C, 30 min.
 - 95 °C, 5 min.
 - 4 °C
5. Store PCR tubes in a freezer at –20 °C in the pre-PCR area or continue immediately with next step.

3.3 Polymerase Chain Reaction (PCR)

1. Thaw reverse transcription reactions and reagents for polymerase chain reaction (PCR) on ice. Enzymes will be removed from the freezer at the time of use. PCR setup should be performed in the pre-PCR area (*see Note 2*), preferably in a PCR cabinet.
2. For each RT reaction prepare four PCR tubes corresponding to the following PCR mixes (5 RT tubes × 4 PCR master mixes = 20 PCR tubes):
 - “Consensus”: to detect both EWSR1-FLI1 and EWSR1-ERG fusions (*see Note 11*).
 - “EWSR1-FLI1 specific”: to detect only EWSR1-FLI1 fusions.
 - “EWSR1-ERG specific”: to detect only EWSR1-ERG fusions.
 - “TBP”: to detect the housekeeping gene TBP (TATA binding protein).
3. Set up the PCR master mixes according to the following table. Mix briefly by vortex and keep on ice. Volumes are indicated in microliters and are enough for five tubes plus one additional tube for pipetting errors.

Reagent	Consensus	EWSR1-FLI1 specific	EWSR1-ERG specific	TBP
10× PCR buffer	13.5	13.5	13.5	13.5
10 mM dNTPs (2.5 mM each)	10.5	10.5	10.5	10.5
25 mM MgCl ₂	10.5	10.5	10.5	10.5
10 μM EWSR1 primer, forward	1.5	1.5	1.5	–
10 μM FLI1/ERG consensus primer, reverse	1.5	–	–	–
10 μM FLI1 specific primer, reverse	–	1.5	–	–

(continued)

Reagent	Consensus	EWSR1-FLI1 specific	EWSR1-ERG specific	TBP
10 μ M ERG specific primer, reverse	–	–	1.5	–
10 μ M TBP primer, forward	–	–	–	1.5
10 μ M TBP primer, reverse	–	–	–	1.5
5 U/ μ L Taq DNA polymerase (<i>see Note 12</i>)	1.5	1.5	1.5	1.5
RNase/DNase-free water	96	96	96	96
Total volume	135	135	135	135

4. Add 2.5 μ L of the corresponding reverse transcription reactions to the tubes prepared in **step 2**.
5. Add 22.5 μ L of the corresponding master mix to the appropriate tube. Mix well by pipetting.
6. Place the tubes in the thermocycler and launch the following program:

Step	Temperature ($^{\circ}$ C)	Time	Number of cycles
Initial denaturation	94	10 min	1
Denaturation	94	30 s	40
Annealing	62	30 s	
Extension	72	45 s	
Final extension	72	7 min	1
Hold	4	∞	1

7. Store tubes in a freezer at -20° C in the post-PCR area or continue immediately with electrophoresis.

3.4 Gel Electrophoresis

1. Electrophoresis should be performed in the post-PCR area (*see Note 2*).
2. Prepare a 1.5% agarose-TBE gel including the recommended concentration of a suitable DNA gel stain (*see Note 13*). Preferably the gel should contain enough wells to allow all samples to run simultaneously.
3. Thaw the PCR reactions and add 5 μ L of $6\times$ loading buffer to each tube. Mix by pipetting.
4. Load the entire volume of the PCR tubes (*see Note 14*). Load a well with 5–10 μ L of the molecular weight standard.

5. Run the gel for 45–60 min at 6 V/cm until the front reaches 6–8 cm.
6. Visualize and photograph the gel under ultraviolet light. Record the result in an appropriate form (*see Note 15*).

3.5 Interpretation of the Results

Figure 4a shows a Ewing sarcoma tumor positive for EWSR1-FLI1 fusion. Specific bands are observed in the “consensus” well and in the EWSR1-FLI1 specific well, but not in the EWSR1-ERG specific well. Figure 4b shows a Ewing sarcoma tumor positive for EWSR1-ERG fusion. In this case, specific bands are observed in the “consensus” well and in the EWSR1-ERG specific well, but not in the EWSR1-FLI1 specific well.

Furthermore, in order to properly interpret the results the following considerations must be taken into account:

- Positive controls (RNA isolated from Ewing sarcoma cell lines expressing EWSR1-FLI1 or EWSR1-ERG fusions): A band at the expected size should be clearly visible in the “consensus” reaction, EWSR1-FLI1 and EWSR1-ERG specific reactions as appropriate and TBP reaction, indicating that RT and PCR reactions have worked correctly. If no bands are observed in these wells, then RT and/or PCR must be repeated.
- Negative controls (no reverse transcriptase and no template control): No bands at all should be visible in these wells. If bands are visible, it could indicate amplicon contamination at any step. In this case the assay should be discarded and repeated. Physical and temporal separation of pre- and post-PCR activities and dedicated equipment helps to reduce the risk of amplicon contamination (*see Note 2*).
- Tumor sample: A band at the expected size should be observed in the TBP, RT positive reaction, indicating a sufficient quality of the tumor RNA. If no band is observed, then the RNA is of poor quality and, therefore, it is not possible to provide a definitive molecular diagnostic result. In that case it is recommended to repeat the RNA isolation from a second sample of the same tumor if available.

If all the above requirements are met, and no bands are observed in the tumor sample in the consensus and EWSR1-FLI1/ERG specific reactions, then it can be concluded that the sample analyzed does not express EWSR1-FLI1/ERG fusions (*see Note 16*).

With this assay, it is possible to deduce the fusion type and the exons involved in the fusion transcripts calculating the size of the PCR bands in relation to the molecular weight standard (Fig. 1).

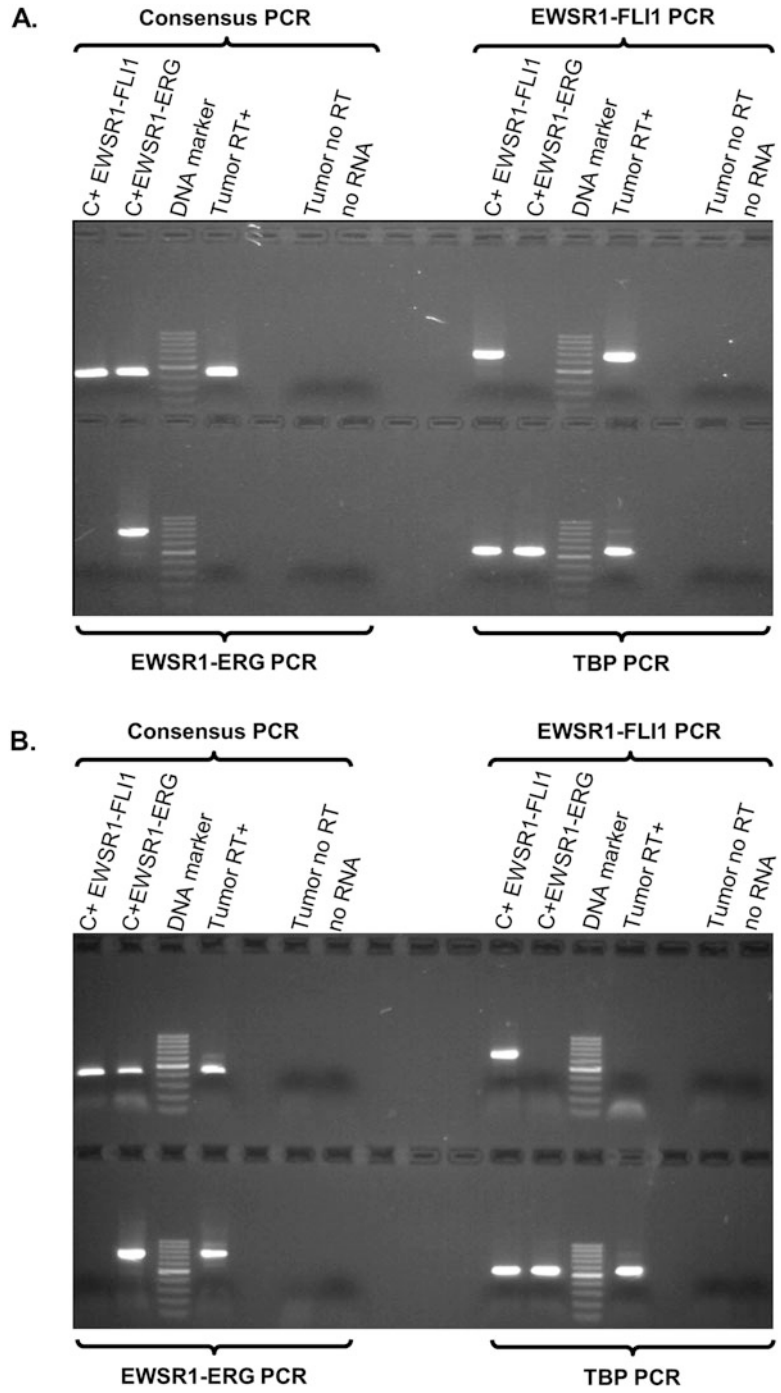


Fig. 4 Examples of the final results obtained in the test. (a) Tumor sample EWSR1-FLI1 positive (Exon 7/Exon 6). (b) Tumor sample EWSR1-ERG positive (Exon 7/Exon 6)

4 Notes

1. The protocol described in this chapter is intended to be used with frozen samples. To achieve the best results, biopsies should be frozen in liquid nitrogen or in isopentane at -80°C as soon as possible to avoid RNA degradation. In this sense, cooperation between surgeons and pathologists is critical to reduce the time between obtaining the biopsy and freezing. If tissue samples have to be transported in the hospital or outside it for analysis in other laboratory, they must be transported in a container suitable for transporting biological samples in dry-ice.
2. Due to the high sensitivity of the PCR reaction it is recommended to take a series of measures to reduce the risk of cross-contamination and therefore the appearance of false positive results. It is a good idea to perform the pre-PCR (RNA extraction and RT-PCR set up) and post-PCR (gel electrophoresis) tasks in separated rooms. Laboratory clothes dedicated for each area are also strongly recommended. In addition, the following recommendations could be also useful to avoid cross-contaminations.
 - Never perform RNA extraction or RT-PCR reaction setup after loading a gel or handling PCR products. It is advisable to perform different tasks on different days or by different technician, particularly gel electrophoresis.
 - Never open or store tubes with PCR products in the pre-PCR area. Store the tubes with PCR products always in the freezers of the post-PCR area.
3. Liquid nitrogen is used to keep the sample frozen during powdering. Liquid nitrogen must be handled with extreme care as it can cause severe burns in contact with the skin. It is essential to wear appropriate gloves and a cryoprotection face shield. If you have doubts about how to handle liquid nitrogen, contact your occupational risk prevention service.
4. The piston–cylinder set is made of stainless steel by an industrial turner. After use, clean thoroughly with soap water, rinse well with distilled water and dry thoroughly before sterilizing. If a piston–cylinder set is not available, a mechanical homogenizer can be used to disintegrate the tumor tissue. In our experience, the piston–cylinder set gives very good results with bone samples, which can be difficult to homogenize. Mortar, piston–cylinder set, spatulas, and forceps must be sterilized at 185°C for 4 h to eliminate nucleases.
5. If powdered tissue is hard to dissolve, more TRI Reagent[®] solution can be added in multiples of 1.5 mL. In many cases, some insoluble white precipitate can remain, corresponding to insoluble bone.

6. At this point the samples can be stored at -20°C in a freezer in the pre-PCR area. To continue with the extraction thaw the samples, vortex at maximum speed 1 min and incubate for 5 min at room temperature.
7. At this stage, two clearly differentiated phases will be formed. The upper phase contains the RNA, while proteins and DNA are distributed in the interphase and organic phase. Be careful to not disturb the phases when manipulating the tubes. To collect the upper phase (aqueous phase containing the RNA) we use a 200 μL micropipette setting at 150 μL and aspirate very slowly in order not to disturb the interphase. If you notice that you have aspirated part of the interface it is preferable to discard that aspiration. Approximately, 750 μL for each tube should be transferred to a 1.5 mL microcentrifuge tube. If the volume is smaller, adjust the volume of ethanol of the next step accordingly to maintain the proportions. Note that solutions and tubes containing phenol/chloroform should be manipulated in a suitable fume hood and residuals disposed according to local biosafety guidelines.
8. The elution volume can be increased if the amount of RNA that is expected to be obtained is significantly higher than 10 μg . In our experience, 14 μL elution is suitable for most of the samples we receive in the laboratory.
9. Positive control for the most frequent fusions EWSR1-FLI1 and EWSR1-ERG can be prepared from RNA isolated from several Ewing sarcoma cell lines that can be obtained from ATCC (<https://www.lgcstandards-atcc.org/>) or DSMZ (<https://www.dsmz.de/>) cell biobanks (Table 3). We prepare

Table 3
List of Ewing sarcoma cell lines that can be used as positive control in RT-PCR reactions

Gene fusion	Cell line	Cell biobank	Exons involved in the fusion	Reference
EWSR1-FLI1	A673	ATCC	Exon 7/Exon 6	[7] [8]
	RD-ES	ATCC	Exon 7/Exon 5	[9, 10]
	SK-ES-1	ATCC	Exon 7/Exon 5	[10]
		DSMZ		[11]
	SK-N-MC	ATCC	Exon 7/Exon 6	[7, 10]
		DSMZ		[12]
	TC-71	DSMZ	Exon 7/Exon 6	[10]
	SK-PN-DW	ATCC	Exon 7 /Exon 6	[10]
	MHH-ES-1	DSMZ	Exon 10/Exon 6	[13] [14]
EWSR1-ERG	CADO-ES1	DSMZ	Exon 7/Exon 6	[15] [16]

5 μL aliquots from a 200 ng/ μL RNA working stock. Aliquots are stored at -80°C and thawed on ice just before preparing the reverse transcriptase reaction.

10. In our experience, better results are obtained using random hexamers instead oligo(dT) primer in the cDNA synthesis reaction.
11. The reverse consensus primer hybridizes with a region highly conserved between FLI1 and ERG. In this way, both fusions can be amplified with the same primer pair (Fig. 3). Additionally specific primers for FLI1 and ERG are used in separate PCR reactions to confirm the existence of fusion transcripts. In this way two independent assays are performed in the same test. By using these primer sets it is possible to identify all EWSR1-FLI1 and EWSR1-ERG fusion transcripts, which occur in the 95% of Ewing sarcoma patients. To identify other fusions that are observed in a minority of Ewing sarcoma, specific primers should be designed [17–19].
12. In our experience, this concentration of AmpliTaq Gold™ polymerase is optimal to render good results. If other Taq DNA polymerase is used, it could be necessary to adjust the concentration to obtain an acceptable sensibility.
13. Nowadays, there are several safer alternatives to ethidium bromide (EtBr) for staining agarose DNA gels. For instance, SYBR™ Safe, which displays acceptable sensibility in these assays. There are other DNA stains that provide higher sensitivity to that of EtBr, so these dyes can also be used.
14. We recommend leaving an empty well between the positive RT wells and the negative controls. In this way, unintended sample contamination from well to well during gel loading is avoided (Fig. 4).
15. It is a good idea to design standardized forms to annotate the different steps performed, including dates, reagents lots, technical staff who performed the test, and result interpretation.
16. Negative EWSR1-FLI1/ERG results do not absolutely discard that the tumor was a Ewing sarcoma since other fusions involving EWSR1 and other ETS transcription factors are possible (Table 1). To confirm this, the use of other primers pairs specific for the most infrequent variants could be necessary (*see Note 11*). When the assay is negative for EWSR1-FLI1/ERG fusions we recommend reassessing the patient in the context of a multidisciplinary team (pathologists, clinical oncologists, and molecular biologists) to evaluate whether other diagnostic options are possible.

Acknowledgments

This work has been supported by grants from the Instituto de Salud Carlos III (PII2/00816; PII6CIII/00026; DTS18CIII/00005); Asociación Pablo Ugarte (TPY-M 1149/13; TRPV 205/18), ASION (TVP 141/17), Fundación Sonrisa de Alex & Todos somos Iván (TVP 1324/15), Asociación Candela Riera (TVP 333-19). CRM was supported by contracts from the Ministry of Economy and Competitiveness (PTA2012-7562-I) and Asociación Pablo Ugarte.

References

- de Alava E (2017) Ewing sarcoma, an update on molecular pathology with therapeutic implications. *Surg Pathol Clin* 10(3):575–585. <https://doi.org/10.1016/j.path.2017.04.001>
- Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, Sorensen PH, Delattre O, Dirksen U (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4(1):5. <https://doi.org/10.1038/s41572-018-0003-x>
- Sankar S, Lessnick SL (2011) Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet* 204(7):351–365. <https://doi.org/10.1016/j.cancer.2011.07.008>
- Cantile M, Marra L, Franco R, Ascierio P, Liguori G, De Chiara A, Botti G (2013) Molecular detection and targeting of EWSR1 fusion transcripts in soft tissue tumors. *Med Oncol* 30(1):412. <https://doi.org/10.1007/s12032-012-0412-8>
- Noujaim J, Jones RL, Swansbury J, Gonzalez D, Benson C, Judson I, Fisher C, Thway K (2017) The spectrum of EWSR1-rearranged neoplasms at a tertiary sarcoma centre; assessing 772 tumour specimens and the value of current ancillary molecular diagnostic modalities. *Br J Cancer* 116(5):669–678. <https://doi.org/10.1038/bjc.2017.4>
- Sugita S, Hasegawa T (2017) Practical use and utility of fluorescence in situ hybridization in the pathological diagnosis of soft tissue and bone tumors. *J Orthop Sci* 22(4):601–612. <https://doi.org/10.1016/j.jos.2017.02.004>
- Aryee DN, Sommergruber W, Muehlbacher K, Dockhorn-Dworniczak B, Zoubek A, Kovar H (2000) Variability in gene expression patterns of Ewing tumor cell lines differing in EWS-FLI1 fusion type. *Lab Invest* 80(12):1833–1844
- Giard DJ, Aaronson SA, Todaro GJ, Arnstein P, Kersey JH, Dosik H, Parks WP (1973) In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J Natl Cancer Inst* 51(5):1417–1423. <https://doi.org/10.1093/jnci/51.5.1417>
- Bertolotti A, Melot T, Acker J, Vigneron M, Delattre O, Tora L (1998) EWS, but not EWS-FLI-1, is associated with both TFIID and RNA polymerase II: interactions between two members of the TET family, EWS and hTAFII68, and subunits of TFIID and RNA polymerase II complexes. *Mol Cell Biol* 18(3):1489–1497. <https://doi.org/10.1128/mcb.18.3.1489>
- Giovannini M, Biegel JA, Serra M, Wang JY, Wei YH, Nycum L, Emanuel BS, Evans GA (1994) EWS-erg and EWS-Flil fusion transcripts in Ewing's sarcoma and primitive neuroectodermal tumors with variant translocations. *J Clin Invest* 94(2):489–496. <https://doi.org/10.1172/JCI117360>
- Bloom ET (1972) Further definition by cytotoxicity tests of cell surface antigens of human sarcomas in culture. *Cancer Res* 32(5):960–967
- Biedler JL, Helson L, Spengler BA (1973) Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture. *Cancer Res* 33(11):2643–2652
- Hayashi M, Chu D, Meyer CF, Llosa NJ, McCarty G, Morris CD, Levin AS, Wolinsky JP, Albert CM, Steppan DA, Park BH, Loeb DM (2016) Highly personalized detection of minimal Ewing sarcoma disease burden from plasma tumor DNA. *Cancer* 122(19):3015–3023. <https://doi.org/10.1002/ncr.30144>
- Pietsch T, Göttert E, Feickert HJ, Riehm H, Blin N, Kovacs G (1989) MHH-ES-1, a new Ewing sarcoma cell line. *Cancer Genet Cytogenet* 38(2):167. [https://doi.org/10.1016/0165-4608\(89\)90568-2](https://doi.org/10.1016/0165-4608(89)90568-2)

15. Kodama K, Doi O, Higashiyama M, Mori Y, Horai T, Tateishi R, Aoki Y, Misawa S (1991) Establishment and characterization of a new Ewing's sarcoma cell line. *Cancer Genet Cytogenet* 57(1):19–30. [https://doi.org/10.1016/0165-4608\(91\)90185-w](https://doi.org/10.1016/0165-4608(91)90185-w)
16. Ladanyi M (2002) EWS-FLI1 and Ewing's sarcoma: recent molecular data and new insights. *Cancer Biol Ther* 1(4):330–336
17. Lewis TB, Coffin CM, Bernard PS (2007) Differentiating Ewing's sarcoma from other round blue cell tumors using a RT-PCR translocation panel on formalin-fixed paraffin-embedded tissues. *Mod Pathol* 20(3):397–404. <https://doi.org/10.1038/modpathol.3800755>
18. Shing DC, McMullan DJ, Roberts P, Smith K, Chin SF, Nicholson J, Tillman RM, Ramani P, Cullinane C, Coleman N (2003) FUS/ERG gene fusions in Ewing's tumors. *Cancer Res* 63(15):4568–4576
19. Ng TL, O'Sullivan MJ, Pallen CJ, Hayes M, Clarkson PW, Winstanley M, Sorensen PH, Nielsen TO, Horsman DE (2007) Ewing sarcoma with novel translocation t(2;16) producing an in-frame fusion of FUS and FEV. *J Mol Diagn* 9(4):459–463. <https://doi.org/10.2353/jmoldx.2007.070009>



Molecular Approaches to Diagnosis in Ewing Sarcoma: Targeted RNA Sequencing

Carmen Salguero-Aranda and Juan Diaz-Martin

Abstract

Molecular testing of pathognomonic gene fusions is mandatory for small round cell tumor diagnosis, including Ewing sarcoma which is indeed defined by a variety of chimeric genes. Reference laboratories are increasingly implementing NGS-based techniques to overcome several limitations of conventional singleplex determinations. We have been early adopters of a targeted-RNA sequencing method based on Anchored multiplex PCR, which allows assessing several fusion transcripts simultaneously with previous knowledge of only one partner gene. Here we describe in detail our protocol and tips for nucleic acid extraction, library preparation, sequencing, and reporting of gene fusions.

Key words Gene fusion, Targeted RNA-seq, Precision diagnostics, Anchored Multiplex PCR (AMP), FFPE sample, Nucleic acid extraction, Ewing sarcoma

1 Introduction

Bone and soft tissue tumors are arguably among the most challenging neoplasms for precision diagnostic. Like many of these mesenchymal tumors, Ewing sarcoma (ES) is characterized by recurrent gene fusions which play a major role in neoplastic transformation. The most frequent gene rearrangement in ES involves *EWSRI* (*FET gene family*) and *FLII* (*ETS transcription factor*) as 5' and 3' partners respectively. However, *EWSRI* can be replaced by *FUS*, and a variety of different *ETS* genes (*ERG*, *ETV1/4*, *FEV*, and *ELA-F*) have been described as 3' partners [1]. Occasionally, *EWSRI* is fused to *non-ETS* gene partners (*PATZ1*, *SP3*, *NFATc2*, and *SMARCA5*) in atypical ES [2, 3]. Moreover, the so-called Ewing-like sarcomas (ELS) present non-*FET* fusions involving *CIC* or *BCOR* genes [3, 4]. ES, atypical ES and ELS often exhibit deceptive and overlapping histomorphologic features, but can show a different clinical behavior [4, 5].

Hence, accurate differential diagnosis of ES requires assessing a sizable variety of gene fusions (GF) with different exonic variants as

well. Molecular testing with traditional gold standards techniques, such as *fluorescence in situ hybridization* (FISH) and/or *reverse transcriptase-PCR* (RT-PCR) can be challenging mainly because these methods do not allow evaluation of multiple GF simultaneously and have poor performance for detection of particular GF [6, 7]. Furthermore, repeated FISH probing has to deal with sample exhaustion, a common issue since sampling techniques usually minimize tissue availability.

Therefore, a multiplexed approach would shorten turnaround time and save sample material. We then implemented in our routine clinical setting a targeted-RNA sequencing method based on Anchored multiplex PCR [8]. RNA is preferred for searching for GF because most of them arise due to breaks within large introns, and after splicing of non-coding sequences, the mRNA molecule is much smaller than the corresponding DNA sequence. Targeted libraries are prepared using the Archer FusionPlex Sarcoma kit, which generates short amplicons enabling the use of paraffin-embedded (FFPE) material. The panel covers different exons of 26 genes involved in GF defining soft tissue sarcoma subtypes. Amplification using both universal and gene specific primers elicits gene fusion identification without prior knowledge of fusion partners. We have integrated the use of this targeted panel (general workflow is depicted in Fig. 1) within our diagnostic algorithm for differential diagnosis of ES and ELS, in which *EWSR1* FISH positive cases with atypical presentation as well as *EWSR1* FISH negative cases are subject to targeted RNA-seq for precise identification of gene fusion.

2 Materials

2.1 Total Nucleic Acid Extraction

1. Microtome.
2. RNase inhibiting surfactant solution.
3. Agencourt FormaPure[®] (Beckman Coulter).
4. DynaMag[™]-2 Magnet (Thermo Fisher Scientific).
5. 1.7 mL microcentrifuge tubes.
6. Heat block.
7. 100% isopropanol.
8. Ultrapure distilled water, DNase/RNase free.
9. 70% ethanol: 3 mL of ultrapure water for each 7 mL of 100% ethanol. Prepare enough volume considering that 1 mL is needed for each sample to be processed.
10. 90% isopropanol: 1 mL of ultrapure water for each 9 mL of 100% ethanol. Prepare enough volume considering that 1.5 mL is needed for each sample to be processed.

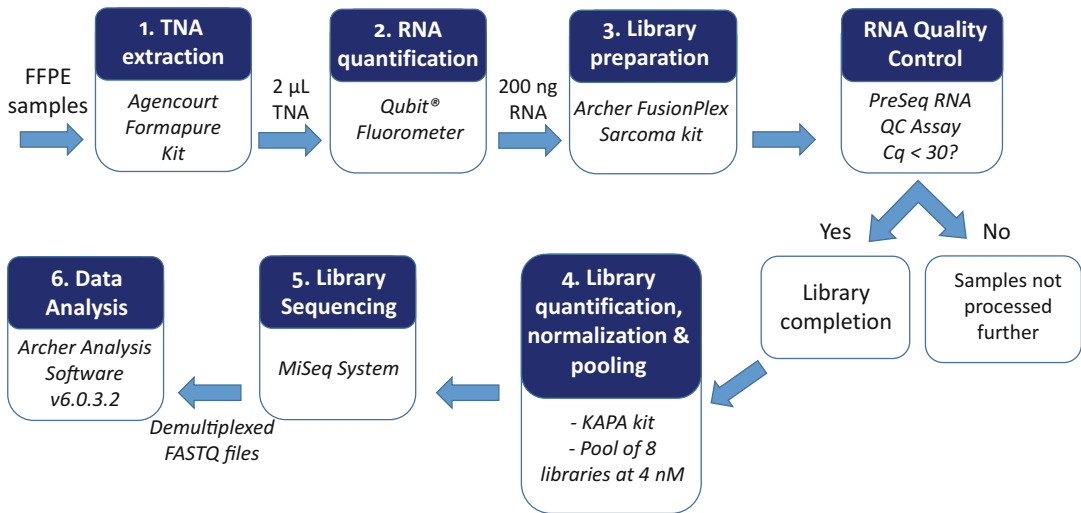


Fig. 1 General workflow for targeted RNA-seq with FFPE samples

2.2 RNA Quantification

1. Qubit[®] fluorometer (Invitrogen[™]).
2. Qubit[®] RNA HS Assay Kit (Invitrogen[™]).
3. 1.7 mL or 15 mL tubes for mixing the Qubit[®] working solution.
4. Qubit[®] assay tubes (Invitrogen[™]).

2.3 Library Preparation, Quantification, Normalization, and Sequencing

1. FusionPlex[®] Sarcoma Kit, for Illumina[®] (ArcherDX).
2. Molecular Barcode (MBC) Adapters, for Illumina[®] (ArcherDX).
3. Agencourt AMPure[®] XP beads (Beckman Coulter).
4. iTaq[™] Universal SYBR Green SuperMix (Bio-Rad Laboratories).
5. Ultrapure distilled water, DNase/RNase free.
6. Nucleic acid decontamination reagent.
7. Real-time PCR thermal cycler.
8. qPCR plates.
9. 0.2 mL PCR tube strips.
10. 1.7 mL microcentrifuge tubes.
11. DynaMag[™]-96 Side Magnet.
12. Microcentrifuge.
13. Plate centrifuge.
14. Vortex mixer.
15. KAPA Universal Library Quantification kit (KAPA Biosystems).

16. Dilution buffer: 10 mM Tris–HCl pH 8.0. Add 200 μ L 500 mM Tris–HCl pH 8.0 to 9.8 mL ultrapure water. Then, add 5 μ L of Tween 20 and mix by vortexing. Prepare the buffer just before use.
17. PhiX Control Kit v3 (Illumina).
18. MiSeq[®] Reagent kit v2 300-cycle (Illumina).
19. HT1 (Hybridization Buffer).
20. 2.5 L ice bucket.
21. MiSeq system (Illumina).

2.4 Data Analysis

1. Archer Analysis Software v6.0.3.2.

3 Methods

3.1 Extraction of Total Nucleic Acid from FFPE Samples

1. A trained pathologist will choose an appropriate FFPE block and estimate the proportion of tumor cells (>10%).
2. Wipe down bench surface, pipettes, and the equipment with RNase inhibiting surfactant solution.
3. FFPE sections contained in 1.5 mL tubes are processed with Agencourt FormaPure kit.
4. Add 200 μ L of Lysis buffer to each tube containing the FFPE sections (*see Note 1*) and incubate in a 70 °C heat block for 60 min to break down paraffin (*see Note 2*).
5. Add 20 μ L of proteinase K solution (provided with the kit, *see Note 3*) and incubate the tubes in a heat block at 55 °C for 60 min. Ensure the proteinase K goes directly into the sample. Pipette to mix three times or mix by vortexing.
6. Spin down and incubate in a heat block at 80 °C for 60 min to perform reverse cross-linking.
7. Cool the tubes on ice for 2 min to solidify any excess of paraffin.
8. Transfer 200 μ L of lysate to a new 1.7 mL tube avoiding any solidified paraffin.
9. Add 150 μ L of Bind 1 buffer (provided with the kit) to each sample. Pipette mix 5 times with a mix volume of 300 μ L.
10. Prepare fresh Bind2/isopropanol solution: Vortex Bind 2 buffer (provided with the kit) vigorously to resuspend magnetic particles. For each individual isolation, combine 20 μ L of Bind 2 Buffer with 300 μ L of 100% isopropanol. Mix well by vortexing. Discard unused Bind2/isopropanol solution.
11. Add 320 μ L Bind 2/Isopropanol solution to each sample, mix by pipetting 5 times with a mix volume of 500 μ L.
12. Incubate the tubes in 55 °C heat block for 5 min.

13. Move the tubes onto the DynaMag™-2 Magnet and separate for 5 min. Discard the cleared supernatant without disturbing the separated magnetic beads.
14. Off the magnet, add 300 µL of wash buffer (*see Note 4*) and pipette 5 times to resuspend the beads. Incubate for 1 min off the magnet.
15. Separate on the magnet for 3 min and carefully aspirate and discard the cleared supernatant. Avoid bead carryover.
16. Off the magnet, add 750 µL of 70% ethanol and pipet 5 times with a volume of 500 µL to resuspend the beads.
17. Separate on the magnet for 5 min and carefully aspirate and discard the cleared supernatant. Avoid bead carryover.
18. Off the magnet, add 500 µL of 90% isopropanol and pipette 5 times with a volume of 400 µL to resuspend the beads. Incubate in a 70 °C heat block for 3 min.
19. Separate on the magnet for 3 min and carefully aspirate and discard the cleared supernatant.
20. Repeat the isopropanol wash for a total of 2 heated isopropanol washes.
21. Off the magnet, add 750 µL of 70% ethanol and pipet 5 times with a volume of 500 µL to resuspend the beads.
22. Separate on the magnet for 5 min and carefully aspirate and discard the cleared supernatant.
23. Let the beads air-dry for 10 min. Be sure to dry the beads before proceeding to the elution step (*see Note 5*).
24. Off the magnet, add 40 µL of nuclease-free water, mix well by pipetting and heat the tubes at 65 °C for 30 s to elute the total nucleic acid from the beads.
25. Separate the samples on the magnet for 1 min and transfer the clear solution to a new 1.7 mL tube. The samples can be stored at -20 °C when not in use, or proceed directly to their quantification.

3.2 RNA Quantification

1. Extracted RNA is quantified using the Qubit® RNA HS Assay Kit and the Qubit® Fluorometer.
2. Prepare the Qubit Working solution by diluting the Qubit RNA HS Reagent 1:200 with Qubit RNA HS Buffer in a 1.7 or 15 mL tube depending on the number of samples to be quantified (*see Note 6*).
3. Add 190 µL of Qubit Working solution to each of the Qubit® assay tubes used for standards and 198 µL of Qubit Working solution to each of the Qubit® assay tubes used for samples.

4. Add 10 μL of each prediluted RNA Qubit standard and 2 μL of each sample to the appropriate tube. Then mix by vortexing 2–3 s. Be careful not to create bubbles.
5. Allow all tubes to incubate at room temperature for 2 min.
6. Proceed to read the standards and samples with Qubit[®] Fluorometer. Ensure to select RNA High Sensitivity in the Qubit device and also select 2 μL as the volume of sample used (*see Note 7*).

3.3 Library Preparation

1. Target-enriched libraries are prepared with Archer FusionPlex Sarcoma kit in conjunction with Molecular Barcode (MBC) Adapters following the manufacturer's protocol. We typically process 8 samples and use 200 ng of RNA as input quantity.
2. To evaluate the integrity of RNA, a real-time PCR is performed during library prep workflow (after first-strand cDNA synthesis). This assay determines the relative amount of amplifiable mRNA derived from the *VCP* gene, which is included as a control in the library panel. Cq value lower than 30 is predictive of sequencing success. Samples with Cq above 30 are not processed further since probably will render analysis failure due to poor RNA integrity (*see Note 8*).

3.4 Library Quantification, Normalization, and Pooling

1. Libraries are quantified using the qPCR-based assay KAPA Universal Library Quantification kit. qPCR is widely regarded as the gold standard for NGS library quantification.
2. Prepare serial dilutions of the targeted libraries in 0.2 mL PCR tube strips as follows (*see Notes 9 and 10*):
 - (a) 1:100 dilutions: Add 2 μL of each library to 198 μL of dilution buffer (*see step 16* in Subheading 2).
 - (b) 1:10,000 dilutions: Add 2 μL of each 1:100 dilution to 198 μL of dilution buffer.
 - (c) 1:100,000 dilutions: Add 10 μL of each 1:10,000 dilution to 90 μL of dilution buffer.
 - (d) 1:500,000 dilutions: Add 20 μL of each 1:100,000 dilution to 80 μL of dilution buffer.

Triplicates of 1:100,000 and 1:500,000 dilutions are used for qPCR.
3. Prepare the required volume of Master Mix for triplicates and dispense 16 μL into wells of PCR plate (*see Note 11*).

- Dispense in this order: 4 μL of PCR-grade water to NTC wells, 4 μL of each prediluted DNA Standard (from the most diluted to the most concentrated), and 4 μL of each dilution of libraries.

Transfer the plate to the real-time PCR thermal cycler and initiate the run as absolute quantification using the following cycling conditions:

Step	Temperature	Time	Cycles
Activation	95 °C	5 min	1
Denaturation	95 °C	15 s	35
Primer, annealing and extension	60 °C	1 min	

- Analyze the qPCR results to determine the library concentration transferring Cq values to the available template at (*see* Fig. 2 and **Note 12**). <https://www.kapabiosystems.com/document/kapa-library-quantification-data-analysis-template/>
- Normalize libraries to 4 nM by placing 5 μL of each into a new 1.7 mL tube and adding ultrapure water up to the volume required.
- Take 10 μL of each 4 nM diluted library and add it to a new 1.7 mL tube. The new tube will now have 80 μL (eight pooled libraries).

3.5 Library Sequencing

- In an ice bucket, prepare an ice–water bath by combining 3 parts ice and 1 part water.
- Remove a MiSeq reagent cartridge (*see* **Note 13**) from $-15\text{ }^{\circ}\text{C}$ to $-25\text{ }^{\circ}\text{C}$ storage and thaw it at ice bucket.
- In order to denature the DNA, combine the following volumes of pooled final DNA library and fresh 200 mM NaOH in a microcentrifuge tube:
 - 10 μL of the 4 nM pooled library.
 - 10 μL of 200 mM NaOH.
- Vortex briefly to mix the sample solution, and then centrifuge the sample solution at $280 \times g$ at $20\text{ }^{\circ}\text{C}$ for 1 min.
- Incubate for 5 min at room temperature to denature the DNA into single strands.
- Add the following volumes to the tube containing the 20 μL of denatured DNA to make a 40 pM denatured library:
 - 10 μL of 200 mM Tris HCl pH 8.0.
 - 970 μL of Prechilled HT1.
- Invert several times to mix and then pulse-centrifuge the DNA solution.

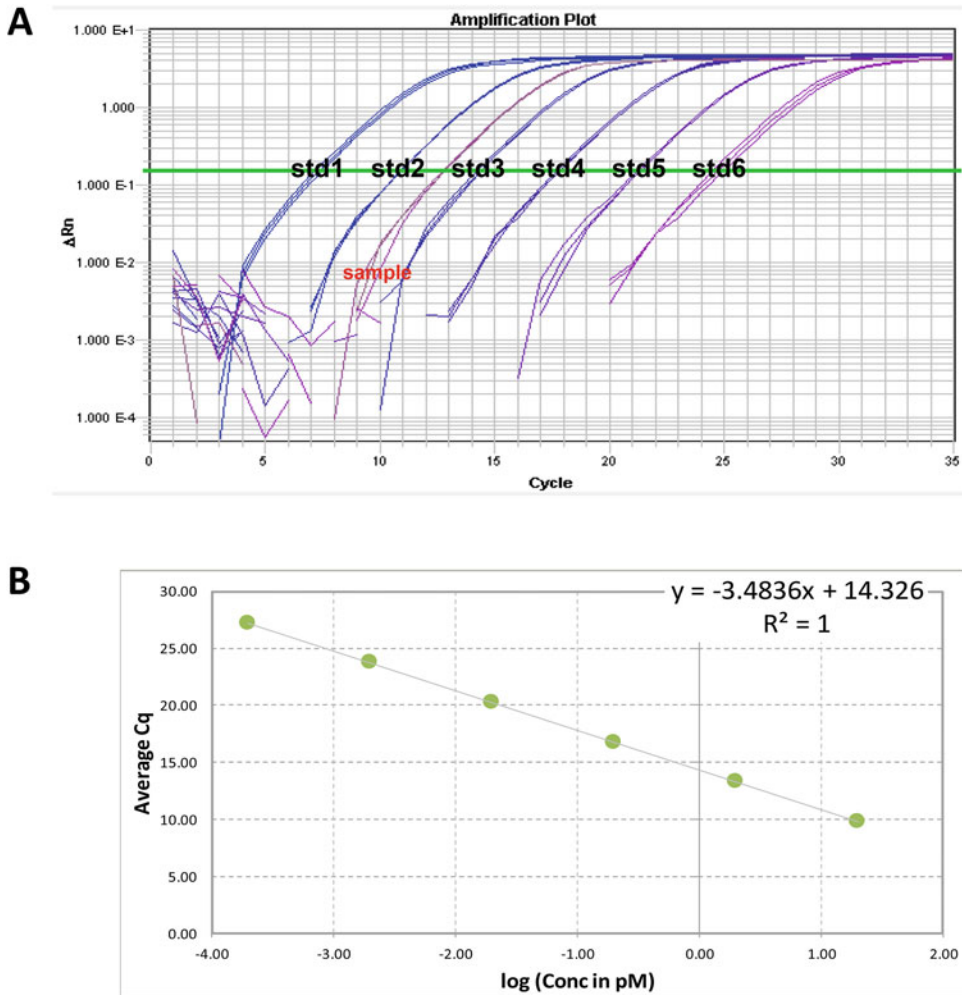


Fig. 2 Library quantification. (a) Absolute qPCR plot displaying amplification of standards and one library sample. (b) Standard curve produced with data from panel A

8. Place the denatured DNA (40 pM) on ice until proceeding to the final dilution.
9. In order to denature and dilute PhiX Control, combine the following volumes to dilute the PhiX library to 4 nM (*see Note 14*):
 - 2 μ L of 10 nM PhiX library.
 - 2 μ L of 10 mM Tris pH 8.0.
10. Add 5 μ L of 200 mM NaOH to the 4 nM PhiX library mixture.
11. Vortex briefly to mix the 2 nM PhiX library solution.
12. Incubate for 5 min at room temperature to denature the PhiX library into single strands.
13. Add 990 μ L of prechilled HT1 to the tube containing 10 μ L of denatured PhiX library to result in a 20 pM PhiX library.

14. Invert several times to mix and then pulse centrifuge the DNA solution.
15. Dilute the denatured 20 pM PhiX library to 10 pM by combining 50 μ L of 20 nM PhiX and 50 μ L of HT1.
16. Invert several times to mix and then pulse centrifuge the DNA solution.
17. Place the denatured and diluted PhiX library (10 pM) on ice.
18. In order to combine Amplicon Library and PhiX Control, combine the following volumes of denatured PhiX control library (10 pM), the denatured amplicon library (40 pM) and HT1 in a microcentrifuge tube (*see Note 15*):
 - 79 μ L of denatured and diluted PhiX control.
 - 375 μ L of denatured and diluted amplicon library.
 - 546 μ L of prechilled HT1.
19. Set the combined sample library and PhiX control aside on ice until loading the entire volume (1 mL) into the appropriate well of MiSeq v2 cartridge and run MiSeq system (*see Note 16*).
20. Prepare a sample sheet specifying sample names and indexes (MBC) as well as read length (151 bp) (*see Note 17*).

3.6 Data Analysis

1. Upload demultiplexed FASTQ files to Archer Analysis Software v6.0.3.2. Both R1 and R2 files are required for samples that are paired-end sequenced. All sample files must be in the same folder and selected at the same time (*see Note 18*).
2. Select RNA Fusion/Isoform analysis pipeline and Sarcoma target region.
3. The criteria to be met for confident GF calling are at least five breakpoint spanning reads, the proportion of breakpoint spanning reads that support the candidate GF relative to the total number of RNA reads spanning the breakpoint needs to be at least 10%, and there must be at least three unique start sites within the population of breakpoint spanning reads.
4. If no GF is found, in order to be consider as a true negative the sample run should meet the following criteria: the number of total reads per sample must be at least 1.5 M, the Fusion Quality Control metric, defined as the average number of unique start sites (from RNA reads) calculated per control GSP2, must be at least 10.
5. Reporting of GF transcripts should specify transcript accession numbers and exons involved. Chromosome coordinates of the fusion breakpoint must be also detailed referred to a specific version of human genome assembly.

4 Notes

1. In our experience 1–2 sections of 10 μm will render enough RNA for the study.
2. Push the sections with a pipette tip until they are completely submerged.
3. Proteinase K (40 mg/mL) solution: Add the required PK Buffer volume directly to the PK vial. Write the date of assembly on the vial, mix well by gently shaking the vial and store the PK solution at $-20\text{ }^{\circ}\text{C}$ when not in use.
4. Wash buffer: Add 100% isopropanol to the Wash Buffer according to the kit acquired:
 - 50 reactions: 15 mL of isopropanol 100%.
 - 96 reactions: 28 mL of isopropanol 100%.
 - 384 reactions: 112 mL of isopropanol 100%.
5. Liquid droplets in the tube walls can be spread across with a pipette tip in order to accelerate drying.
6. The two standards require 190 μL each of Qubit Working solution while samples require 198 μL each. Thus, prepare enough Qubit Working solution for the samples plus the two standards assuming that each tube need 200 μL and add one extra volume to compensate for pipetting error.
7. If RNA concentration is too high for the standard curve, sample dilutions must be prepared (1:5 or 1:10).
8. FusionPlex[®] Sarcoma Kit (Archer) includes enough reagents to test the quality of 16 samples (PreSeq RNA QC assay), to help ensure that 8 high-quality samples are identified for each batch of targeted library prep and sequencing.
9. Clean all the surfaces, pipettes and the equipment to be used (vortex, minicentrifuge, thermal cycler) with a nucleic acid decontamination reagent to prevent contamination with high concentrated libraries.
10. Be sure to mix well by pipetting up and down or vortexing and spin down the tubes before transferring the samples between tubes.
11. When preparing the Master Mix for library quantification, make sure that the 6 standards, samples and NTC are counted in triplicated and to include in the calculation some extra wells for pipetting error. For example, for eight samples and two dilutions of each (100,000 and 500,000) plus the six standards and NTC, that makes a total of 23 samples. Taking into account the triplicates, prepare a mix for 80 samples.

Component	Reaction mix (<i>n</i> = 1) (μL)	Master mix i.e. (<i>n</i> = 80) (μL)
SYBR mix	12.4	992
Ultrapure water	3.6	288
Standards, library dilutions (100,000 and 500,000) or NTC	4	–
Total	20	1280

12. For completing the KAPA Library Quantification Data Analysis Template for Illumina, go first to *Readme* sheet and follow the instructions. In this case, two library dilutions were done (100,000 and 500,000), so the *2 dilution analysis* sheet is the one that need be completed and the *2 dilution summary* will be directly generated.
13. The recommended sequencing depth for FusionPlex Sarcoma libraries is 1,5 million reads per sample. We usually run 8 pooled samples using MiSeq[®] Reagent kit v2 300-cycle which provides 12–15 million reads.
14. The PhiX library must be diluted to the same loading concentration as the Amplicon library (in this case, 4 nM).
15. The final library mixture containing the denatured PhiX control library, the denatured amplicon library and HT1 is done in order to reach 5% PhiX and 14 pM as the final loading concentration of the library.
16. Cluster density should ideally be 750–1100 k/mm².
17. Sample sheet forms for different MBC can be found at <https://archerdx.com/documents/>
18. FASTQ files can be either uncompressed (with extension “.fastq” or “.fq”) but it is recommended to use compression (using the GZIP algorithm, and have the extension “.gz”). ZIP compression is not supported.

References

1. Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, Sorensen PH, Delattre O, Dirksen U (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4(1):5. <https://doi.org/10.1038/s41572-018-0003-x>
2. Sankar S, Lessnick SL (2011) Promiscuous partnerships in Ewing’s sarcoma. *Cancer Genet* 204(7):351–365. <https://doi.org/10.1016/j.cancergen.2011.07.008>
3. Renzi S, Anderson ND, Light N, Gupta A (2019) Ewing-like sarcoma: an emerging family of round cell sarcomas. *J Cell Physiol* 234(6):7999–8007. <https://doi.org/10.1002/jcp.27558>
4. Carter CS, Patel RM (2019) Important recently characterized non-Ewing small round cell tumors. *Surg Pathol Clin* 12(1):191–215. <https://doi.org/10.1016/j.path.2018.10.008>
5. Le Loarer F, Pissaloux D, Coindre JM, Tirode F, Vince DR (2017) Update on families of round cell sarcomas other than classical Ewing

- sarcomas. *Surg Pathol Clin* 10(3):587–620. <https://doi.org/10.1016/j.path.2017.04.002>
6. Chen S, Deniz K, Sung YS, Zhang L, Dry S, Antonescu CR (2016) Ewing sarcoma with ERG gene rearrangements: a molecular study focusing on the prevalence of FUS-ERG and common pitfalls in detecting EWSR1-ERG fusions by FISH. *Genes Chromosomes Cancer* 55(4):340–349. <https://doi.org/10.1002/gcc.22336>
 7. Noujaim J, Jones RL, Swansbury J, Gonzalez D, Benson C, Judson I, Fisher C, Thway K (2017) The spectrum of EWSR1-rearranged neoplasms at a tertiary sarcoma centre; assessing 772 tumour specimens and the value of current ancillary molecular diagnostic modalities. *Br J Cancer* 116(5):669–678. <https://doi.org/10.1038/bjc.2017.4>
 8. Zheng Z, Liebers M, Zhelyazkova B, Cao Y, Panditi D, Lynch KD, Chen J, Robinson HE, Shim HS, Chmielecki J, Pao W, Engelman JA, Iafrate AJ, Le LP (2014) Anchored multiplex PCR for targeted next-generation sequencing. *Nat Med* 20(12):1479–1484. <https://doi.org/10.1038/nm.3729>

Part III

Cellular Biology and Cell Culture Techniques



Ewing Sarcoma-Specific (Re)expression Models

Maximilian M. L. Knott and Florencia Cidre-Aranaz

Abstract

Gene expression and knockdown systems are powerful tools to study the function of single genes and their pathway interaction. Plasmid transfection and viral transduction have revolutionized the field of molecular biology and paved the ground for various gene-editing strategies such as TALEN, zinc finger nucleases, and ultimately CRISPR. In Ewing sarcoma (EwS), almost as many genes are repressed by the expression of EWSR1-FLI1 as are upregulated by the fusion oncogene. Here we present a useful point-to-point protocol for the generation of transgene expression systems in EwS that allow (conditional) reexpression of a gene of interest. We provide an extensive instruction on molecular cloning, plasmid generation, viral transduction, and expression validation. Finally, we address common problems and highlight potential pitfalls, which can easily be avoided by thoughtful guidance.

Key words Ewing sarcoma, Gene expression, Overexpression, Inducible vector, Microsatellites, Enhancer, Promoter, Fusion-driven sarcoma

1 Introduction

pSC101 was the first plasmid ever used for cloning purposes by Herbert Boyer and Stanley Cohen in 1973 [1]. Since then, molecular cloning strategies have evolved significantly and by now there are more than 80,000 plasmids stored in the most common plasmid repository addgene.org [2]. Originally, plasmids were isolated from certain bacterial strains such as *Salmonella* spp. or *Escherichia* spp. and modified by restriction digest and subsequent ligation [3]. Today, a plethora of molecular cloning strategies exist, such as PCR-based cloning, isothermal assembly (i.e., Gibson Assembly), gateway cloning, ligation-independent cloning (LIC), and Golden Gate-Assembly [4]. Lastly, plasmids can even be fully synthesized de novo [5].

In contrast to the wide variety of strategies that allow suppression of gene expression (i.e., complete knockout by CRISPR, siRNA-mediated knockdown, stable or conditional shRNA-mediated knockdown), gene reexpression requires the introduction of a third, either constitutively or conditionally expressed copy of

the gene of interest (transgene), either by transfection (episomal) or viral transduction (episomal or integrating).

Viral transduction is the most common gene delivery strategy. Gene trafficking by viruses was first observed in bacteriophages by Norton Zinder and Joshua Lederberg in 1952 [6]. In 1976, Stephen Goff and Paul Berg were the first to use a modified SV40 virus for viral gene transfer in mammalian cells [7]. Since then, retroviruses, γ -retroviruses, lentiviruses, adenoviruses, adeno-associated viruses, and herpes-simplex viruses have been used for gene delivery both in vitro and in vivo [8]. Viral transduction systems offer great versatility, high efficiency, and stable gene expression, especially when integrating viruses are employed. Among them, lentiviruses are the most common vectors in molecular biology. Lentiviruses are ssRNA+ viruses that integrate into the host cell's genome and are capable of transducing even nondividing, terminally differentiated cells such as neurons. For in vitro transductions, modified, replication-defective lentiviruses are usually used.

Gene expression systems can be categorized into constitutive or conditional: in constitutive expression systems, transgene expression is always turned on (driven by a strong promoter (i.e., CMV or EF1 promoter)); in contrast, conditional expression systems harbor a minimal promoter, which needs additional transactivation factors to drive gene expression. In conditional expression systems, transgene expression is often induced by the tetracycline analogue doxycycline (tetracycline-controlled operator system), tamoxifen (CreER-system), or cumate (cumate-controlled operator system) and can thereby be turned on and off (for review *see* [9]). Inducible systems are advantageous over constitutive ones since transgene expression can be tightly controlled and initiated with some delay (which is especially helpful when inoculating tumors in vivo). Moreover, no single-clone effects have to be considered, as no additional (untransduced) control clones are required. However, doxycycline-dependent regulation of certain *off-target* genes has been described [10, 11] and most of the available vectors show weak transgene expression even in absence of the appropriate stimulus (leakiness).

Gene overexpression systems have been an important tool in oncologic research for years. They have been employed to allow fluorescence or chemiluminescence-based bioimaging of tumors in vivo [12], for the study of tumor-suppressor genes [13] and reexpression of suppressed genes [14]. Moreover, they are often used for rescue experiments, where the expression of single genes can rescue the loss of an upstream regulator [15].

EwS is regulated by a chimeric translocation (in 85% of cases EWSR1-FLI1) that completely rewires the cell's transcriptional network [16], thus either repressing potentially tumor suppressing genes or reexpressing oncogenes. Analyzing the in vitro and in vivo impact of reexpressing EWSR1-FLI1-repressed genes has proven to

be an important tool for discovering new potential therapeutic targets and/or druggable pathways in EwS [17, 18].

Here we describe a complete protocol for the development of transgene-expressing EwS cell lines with special regards to potential pitfalls and caveats. In short, the protocol consists of a detailed instruction on the vector design, cloning of the transfer plasmid, production and titering of lentiviral particles, transduction of EwS cells, and the validation of transgene expression. Figure 1 depicts the workflow of this protocol.

2 Materials

2.1 Vector Design

1. Cloning software (benchling.com, Snapgene, etc.)

2.2 Cloning of the Transfer Plasmid

1. Gene expression vector of interest (*see Note 1*).
2. Genomic DNA, cDNA, or commercial cDNA clone containing the gene of interest (*see Note 2*).
3. PCR primers flanking the region of interest (*see Note 3*).
4. PCR primers flanking the region of interest with 5'-overhangs (*see Note 4*).
5. Phusion High-Fidelity DNA-Polymerase (Thermo Scientific) (*see Note 5*).
6. Polymerase-specific buffer.
7. 10 mM dNTP mix.
8. PCR-grade DMSO.
9. ddH₂O.
10. 0.2 ml PCR tube strips with caps.
11. Agarose.
12. 10× TAE buffer (0.4 M TRIS (48.456 g), 0.2 M acetic acid (12.0104 g), 10 mM EDTA (2.9224 g) in 1 L ddH₂O).
13. 1× TAE buffer (dilute 100 ml of 10× TAE in 900 ml ddH₂O).
14. DNA gel stain (e.g., ethidium bromide or SYBR Safe).
15. 6× loading dye for electrophoresis gels.
16. Molecular weight ladder for electrophoresis (1 kb).
17. Horizontal gel electrophoresis chamber.
18. Scalpels.
19. UV Table.
20. 1.5 ml reaction tubes.
21. DNA gel extraction and PCR clean-up kit (e.g., Macherey-Nagel).
22. Spectrophotometer.



Vector design (1 day)



Cloning of transfer plasmids (3-5 days)



Lentivirus production (4 days)



Titration of viral particles (3 days)



Viral transduction and selection (3-7 days)



Validation of transgene expression (2-3 days)

Fig. 1 Schematic overview of the protocol for the generation of transgene expressing EwS cell lines

23. High-fidelity restriction enzymes and Cutsmart buffer (New England Biolabs).
24. T4 DNA Ligase (New England Biolabs).
25. Chemically competent *E. coli*.
26. Agar plates containing 100 µg/ml ampicillin (*see Note 6*).
27. LB broth containing 100 µg/ml ampicillin (add 1 ml of ampicillin stock solution (100 mg/ml) to 1 l of LB broth).
28. Bunsen burner.
29. 50% glycerol in sterile ddH₂O (mix 50 ml glycerol with 50 ml sterile ddH₂O).
30. Plasmid DNA midi prep kit.

2.3 Lentivirus Production

1. Class II laminar flow hood.
2. Appropriate safety gloves.
3. Mask.
4. Lab coat.
5. HEK293T cells of low passage number (*see Note 7*).
6. 6-well cell culture plates.
7. 1.5 ml reaction tubes.
8. DMEM with 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 50 ml FCS, 5 ml L-glutamine (200 mM) and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
9. Opti-M I Reduced Serum Medium (Gibco).
10. pCD/NL-BH*DDD plasmid (Addgene) for second generation lentiviral vectors.
11. pCEF-VSV-G plasmid (Addgene) for second generation lentiviral vectors.
12. Qubit fluorometer or spectrometer.
13. PEI Max Transfection Grade Linear Polyethylenimine Hydrochloride MW 40.000 (Polysciences).
14. Syringe filter (0.45 µm pore size, PES or PVDF membrane).
15. Cryo tubes.

2.4 Titering of Viral Particles Using Flow Cytometry

1. 12-well plate.
2. RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 50 ml FCS, 5 ml L-glutamine (200 mM), and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
3. Cell line of interest.
4. Viral supernatant.

5. DMEM with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 5 ml FCS, 5 ml L-glutamine (200 mM) and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
6. Flow cytometer.

2.5 Titering of Viral Particles Using qPCR

1. 12-well plate.
2. 1.5 ml reaction tubes.
3. RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 50 ml FCS, 5 ml L-glutamine (200 mM) and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
4. Cell line of interest.
5. Viral supernatant.
6. DMEM with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 5 ml FCS, 5 ml L-glutamine (200 mM) and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
7. DNaseI.
8. PBS.
9. 0.25% trypsin–EDTA
10. gDNA extraction kit.
11. Transfer plasmid used for the generation of transduced cells.
12. Qubit fluorometer.
13. ddH₂O.
14. qPCR primers for nonhuman sequences of the plasmid (inside the 5'- and 3'-LTRs, e.g., WPRE) (*see Note 8*).
15. qPCR primers for a housekeeping gene (e.g., RPLPO).
16. SYBR Select Mastermix (e.g., ThermoFisher Scientific).
17. qPCR plates.
18. Optical adhesive film.
19. Real-time qPCR cyclor.

2.6 Transduction and Selection

1. 12-well plate.
2. RPMI 1640 with 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 50 ml FCS, 5 ml L-glutamine (200 mM) and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
3. Cell line of interest.
4. Viral supernatant.
5. Fluorescence-activated cell sorter (FACS).

6. Antibiotic-selection medium (e.g., 0.5–2 µg/ml puromycin in RPMI 1640 with 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 2.5–10 µl puromycin (10 mg/ml), 5 ml FCS, 0.5 ml L-glutamine (200 mM), and 0.5 ml penicillin–streptomycin (10,000 IU/ml/ 10 mg/ml) to 50 ml of medium) (*see Note 1*).

2.7 Validation of Gene Expression

1. 6-well plate.
2. 1.5 ml reaction tube.
3. RPMI 1640 with 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (add 50 ml FCS, 5 ml L-glutamine (200 mM), and 5 ml penicillin–streptomycin (10,000 IU/ml/10 mg/ml) to a 500 ml bottle of medium).
4. Transduced cell line of interest.
5. PBS.
6. mRNA extraction kit (e.g., Macherey-Nagel).
7. cDNA transcription kit (e.g., ThermoFisher Scientific).
8. qPCR primers specific for the gene of interest.
9. qPCR primers for a housekeeping gene (e.g., RPLPO).
10. SYBR Select Mastermix (e.g., ThermoFisher Scientific).
11. qPCR plates.
12. Optical adhesive film.
13. Real-time qPCR cyclers.

3 Methods

3.1 Vector Design

This is the most crucial step that requires focused attention in planning your cloning design. Use appropriate software tools, such as benchling.com (free) or commercial cloning software.

1. Decide on the gene expression system you want to use (*see Note 1*).
2. Design and order primers for your transgene sequence by using the first 16–25 nt of the 5' strand of your sequence of interest (cDNA or genomic DNA) (*see Note 2*) as the forward primer and the last 16–25 nt of the 3' strand as the reverse primer (*see Note 3*).
3. Identify restriction sites in your transgene sequence and exclude them from your cloning strategy (*see Note 9*).
4. Choose appropriate restriction sites from the multiple cloning site (MCS) of your transfer plasmid and use them for the design of your 5'-overhang primers (*see Note 4*).

3.2 Cloning of Transfer Plasmids

Perform all steps at room temperature unless specified otherwise. When handling bacteria or medium to grow them in, work near a Bunsen burner and pay attention on using proper sterile technique. Here, we describe the most common, restriction-based cloning. For other cloning strategies, such as Gibson Assembly, please refer to relevant literature and the manufacturer's protocol.

1. To amplify the sequence of interest from your template using Phusion high-fidelity polymerase, prepare a PCR master mix by adding 10 μl of HF buffer, 2 μl of DNA template (0.5–500 ng/ μl), 1 μl 10 mM dNTP, 0.5 μl Phusion polymerase, and 34 μl of PCR-grade ddH₂O per sequence of interest/primer pair (*see Note 10*).
2. Vortex mildly, spin down and add 47.5 μl to a 0.2 ml PCR tube for each reaction.
3. To each reaction add 1.25 μl of 10 μM forward primer and 1.25 μl of 10 μM reverse primer (*see Note 11*).
4. Vortex mildly and spin down.
5. Place the reaction tubes into a PCR cyclor and start a PCR program following cycling instructions given in the Phusion high-fidelity user guide (*see Note 12*).
6. Prepare a 1.3% agarose gel by measuring 1.3 g of agarose and mixing it with 100 ml of 1 \times TAE buffer. Microwave until all agarose is dissolved (approximately 2 min at 900 W).
7. Let solution cool down to approximately 65 °C under the fume hood and add DNA gel stain of choice to a final concentration of 0.5 $\mu\text{g}/\text{ml}$.
8. Pour mixture into a gel tray with an already inserted comb that can hold up to 60 μl of PCR product per line.
9. Mix the obtained PCR products with 10 μl 6 \times loading dye.
10. Add a 12.5 μl of molecular weight ladder into first well of the gel and 50 μl of PCR product mixed with 6 \times loading dye to the remaining wells of your gel.
11. Separate PCR products by applying 100 mV over 30 min.
12. Inspect gel under UV light. If DNA bands corresponding to expected lengths are visible, neatly cut out the gel containing the DNA band using a scalpel and store it in 1.5 ml reaction tube (*see Note 13*).
13. Extract DNA from cut out gel blocks using a commercial kit and following the manufacturer's protocol.
14. Measure DNA concentration using a spectrophotometer.
15. Create DNA fragments with digestible overhangs by following PCR protocol as described in **steps 1–5** with primers containing overhangs. Use 1 μl of DNA (10 ng/ μl) extracted in **step**

- 13** as template and prepare at least two 50 μ l reactions for each insert to generate adequate amounts of your desired PCR product (*see Note 14*).
16. Repeat gel electrophoresis and gel extraction as described in **steps 6–13**. Use a big comb to allow loading of at least 100 μ l PCR product per lane.
 17. Measure DNA concentration using a spectrophotometer.
 18. In order to digest extracted DNA as well as your plasmid of choice with two suitable restriction enzymes (*see Note 15*), mix 1 μ g of DNA with 5 μ l Cutsmart buffer, and fill up with DNase-free ddH₂O to 48 μ l. Add 1 μ l of enzyme 1 and 1 μ l of enzyme 2 at last.
 19. Vortex mildly and spin down.
 20. Incubate at 37 °C for 60 min.
 21. Perform PCR cleanup using commercial kit following manufacturer's protocol to retrieve the digested plasmid and digested inserts ready for sticky-end ligation.
 22. Measure DNA concentration using a spectrophotometer.
 23. To perform sticky-end ligation of 25 ng of digested plasmid with digested inserts, calculate a mass of digested insert corresponding to a molecular plasmid–insert ratio of 1:5 (*see Note 16*).
 24. Mix plasmid and insert DNA with 2 μ l of T4 ligase buffer and 1 μ l of T4 ligase in a total reaction volume of 20 μ l (filled up with DNase-free ddH₂O).
 25. Mix by pipetting up and down. Spin down.
 26. Incubate at room temperature for 15 min (*see Note 17*).
 27. Transform competent cells with ligation mix following manufacturer's protocol.
 28. Streak transformed bacteria on an LB agar plate containing 100 μ g/ml ampicillin.
 29. Incubate overnight (12–16 h) at 37 °C.
 30. Pick eight single colonies by touching them with a sterile pipette tip and swirling the tip in 10 μ l of LB containing 100 μ g/ml ampicillin in a 0.2 ml PCR tube (*see Note 18*).
 31. In order to perform colony PCR to detect colonies containing the desired insert, calculate a PCR master mix as described in **step 1** but omitting the DNA template.
 32. For each picked colony, add 49 μ l of PCR master mix to a fresh 0.2 ml PCR tube.
 33. Add 1 μ l of LB containing each picked bacteria from **step 30** to each tube.

34. Use the same cycling program that lead to visible bands at expected length in the previous PCRs.
35. Add 91 μl of LB containing 100 $\mu\text{g}/\text{ml}$ ampicillin to each of the tubes containing the remaining 9 μl of inoculated LB.
36. Store in a shaking incubator at 37 °C and 200 rpm until the results of the colony PCR are obtained.
37. Perform gel electrophoresis as described in **steps 6–12**.
38. Write down the colonies in which DNA bands of expected lengths are visible.
39. Cultivate colonies noted in **step 38** by adding the entire 100 μl of the colony culture from **step 35** to 100 ml of LB containing 100 $\mu\text{g}/\text{ml}$ ampicillin.
40. Leave overnight in a bacterial incubator at 37 °C at 200 rpm.
41. The next morning the LB should be turbid due to bacterial growth. Create glycerol stock of bacterial culture by mixing 500 μl of turbid LB with 500 μl of 50% glycerol.
42. Store immediately at -80 °C.
43. Extract plasmid from the remaining culture using a commercial midi prep kit and following the manufacturer's instructions.
44. Assess prep purity by spectrophotometry.
45. Measure DNA quantity by use of Qubit.
46. Evaluate plasmid integrity by agarose gel electrophoresis.
47. Confirm correct plasmid sequence by Sanger sequencing using a commercial sequencing service (*see Note 19*).

3.3 Lentivirus Production

To achieve stable gene expression in EwS cell lines, cells are lenti-
virally transduced with the cloned plasmid. Please note that lenti-
viruses fall within NIH Biosafety Level 2 criteria and must be
handled with care. Always work in a class 2 laminar flow hood,
wear appropriate safety gear and observe local guidelines on how to
decontaminate cultures and waste. We strongly recommend to use
the Qubit fluorescence-based DNA quantification system to deter-
mine plasmid concentrations of your transfer plasmid, pCD/NL-
BH*DDD and pCEF-VSV-G, since their ratio and amount are
crucial for good virus titers.

1. 24 h before transfection seed 530.000 HEK293T cells per well into 6-well plates in 2 ml DMEM containing 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (if needed) (*see Note 20*).
2. On the next day, check the confluence of your HEK293T cells. Best yields of viral titers will be obtained with a 60–70% confluence on day of transfection (*see Note 21*).

3. On day of transfection, for each well to be transfected mix 510 ng of the cloned transfer plasmid (100 ng/ μ l) with 340 ng of pCD/NL-BH*DDD (100 ng/ μ l) and 170 ng pCEF-VSV-G (100 ng/ μ l) in Opti-MEM for a final volume of 100 μ l in a 1.5 ml reaction tube.
4. In a separate 1.5 ml reaction tube mix 13.2 μ l PEI (0.6 mg/ml) and 91.8 μ l Opti-MEM for each well to be transfected.
5. Incubate both solutions for 10 min individually at room temperature.
6. Add 100 μ l of the PEI mix on top of the DNA mix and pipet up and down 5–10 times without producing bubbles.
7. Incubate for 10 min at room temperature.
8. In the meantime, change medium of the HEK293T cells and replace it with 2 ml fresh DMEM containing 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (if needed).
9. Add 200 μ l of DNA–PEI mix dropwise to each well to transfect while gently swirling the dish.
10. Incubate at 37 °C for 16 h.
11. After 16 h, change medium and replace it with 2 ml fresh DMEM containing 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin to avoid PEI-mediated toxicity.
12. Incubate at 37 °C for additional 48 h.
13. Harvest the supernatant and filter it through a syringe filter to remove cell debris.
14. Aliquot the virus supernatant into 100 μ l fractions and store it in cryo tubes at –80 °C immediately. Alternatively, use the virus for direct transduction of the cell lines of interest (*see Note 22*).

3.4 Titering of Viral Particles Using Flow Cytometry

To prevent exceeding or insufficient transduction rates, the produced virus should be titered. If your plasmid harbors a fluorescent selection marker, viral titers can be analyzed using flow cytometry. Otherwise, genomic titering using qPCR is recommended (*see Subheading 3.5*). Optimal titers range from 1×10^6 to 1×10^7 viral particles per ml. When even higher titers are needed (e.g., for in vivo use), viral particles can be concentrated using PEG precipitation or ultracentrifugation (*see Note 23*).

1. Seed 6 wells of a 12-well plate with 50,000 cells of interest each in 1 ml RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin.
2. 24 h later, count two of the seeded wells and determine the exact number of cells at the time of transduction.

3. Change medium to DMEM with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin in the remaining four wells.
4. Add drop-wise 0 μ l, 1 μ l, 5 μ l, or 50 μ l of your virus supernatant to each of the four wells.
5. 24 h later, change medium for fully supplemented RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin.
6. Incubate for additional 48 h.
7. Harvest cells and analyze by flow cytometry.
8. Calculate titer using the following formula: $TU/ml = (P \times N \times 1000)/V$ where P = percentage of fluorescent cells, N = number of cells at time of transduction and V = volume (μ l) of virus added into each well for transduction. Accurate titers can be determined in a range of fewer than 40% transduced cells.

3.5 Titering of Viral Particles Using qPCR

When lentiviral transfer plasmids that do not contain any fluorescent marker are used, titering has to be performed by genomic qPCR.

1. Seed 6 wells of a 12-well plate with 50,000 cells of interest each in 1 ml RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin.
2. 24 h later, count two of the seeded wells to obtain the exact number of cells at the time of transduction.
3. Change medium to DMEM with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin in the remaining four wells.
4. Add dropwise 0 μ l, 1 μ l, 5 μ l, or 50 μ l of your virus supernatant to each of the four wells.
5. After 20 h, remove the medium and add 500 μ l medium containing DNaseI in a final concentration of 10 U/ml. Incubate for 15 min on 37 °C.
6. Remove the DNaseI and add 2 ml fully supplemented RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin (*see Note 24*).
7. Incubate for additional 48 h.
8. Wash the cells with 1 ml PBS and harvest the cells by trypsinization.
9. Isolate the genomic DNA using a gDNA extraction kit and follow the manufacturer's protocol.
10. Measure the gDNA concentration of your untransduced sample and your transfer plasmid using a fluorescence based (Qubit) system.
11. Prepare a serial dilution of your transfer plasmid by calculating its molecular weight, calibrating it to 2×10^{11} molecules per

100 μl and diluting it serially 1:10 in ddH₂O seven times in order to obtain a range from 2×10^3 to 2×10^{11} molecules. Use this as a plasmid standard.

12. Prepare a serial dilution of your untransduced gDNA starting with 1 $\mu\text{g}/100 \mu\text{l}$ and diluting it serially to 200 ng/100 μl , 100 ng/ μl , 40 ng/100 μl , 8 ng/100 μl , and 1.6 ng/100 μl in ddH₂O. 1 ng of gDNA corresponds to 333 genomic copy equivalents. Use this as a gDNA standard.
13. Prepare a qPCR mastermix for the plasmid sequence and the housekeeping gene using 0.375 μl of each of the respective primers, 7.5 μl SYBR Green Mastermix, and 1.75 μl ddH₂O per well.
14. Add 5 μl of each transduced sample (in quadruplicates) into the wells of the qPCR plate (*see Note 25*).
15. Add 5 μl of your diluted plasmid standard (in duplicate) into the wells of the qPCR plate.
16. Add 5 μl of your untransduced genomic DNA standard (in duplicate) into the wells of the qPCR plate.
17. Add 10 μl of the mastermix for the plasmid sequence to each transduced sample (in duplicate).
18. Add 10 μl of the mastermix for the housekeeping gene to each transduced sample (in duplicate).
19. Add 10 μl of the mastermix for the plasmid sequence to the plasmid standard (in duplicate).
20. Add 10 μl of the mastermix for the housekeeping gene to the untransduced gDNA standard (in duplicate).
21. Seal the plate with an optical adhesive film and transfer it to a q-RT-PCR cycler.
22. Choose the appropriate settings for SYBR Green fluorescence and set up a protocol as follows:
 - Step 1:** 2:00 min 95 °C.
 - Step 2:** 0:10 min 95 °C.
 - Step 3:** 0:30 min 60 °C.
 Repeat **steps 2** and **3** for 50 cycles.
23. Calculate the titer using the following formula: $\text{TU}/\text{ml} = (C \times N \times 1000)/V$, where C = plasmid copies per genome, N = number of cells at time of transduction and V = volume (μl) of virus added into each well for transduction. Vector copies per genome can be calculated using the plasmid and gDNA standard curves. Accurate titers are determined in a range of fewer than five vector copies per genome.

3.6 *Viral Transduction and Selection*

To obtain stable gene expression, EwS cell lines can now be transduced using the previously produced and titered virus.

1. Seed 2 wells of a 12-well plate with 50,000 cells of interest each in 1 ml RPMI 1640 with 1% FCS, 1% L-glutamine, and 1% penicillin–streptomycin.
2. 24 h later, change the medium in both wells and add drop-wise the calculated amount of virus supernatant (MOI 1 to 5 (*see Note 26*)) to one of the wells (*see Note 27*).
3. 24 h later, change the medium again.
4. Incubate for additional 48 h.
5. Select for successfully transduced cells using fluorescence activated cell sorting (FACS) or by adding the respective drug (e.g., puromycin) in its minimal effective dose (as previously determined (*see Note 28*)).

3.7 *Validation of Gene Expression*

After sorting or selecting successfully transduced cells the expression of the transgene can be easily assessed by q-RT-PCR. When working with inducible expression vectors, transgene expression has to be induced by adding the respective stimulus (doxycycline, estrogen, cumate, etc.). Here, we describe a SYBR Green–based q-RT-PCR protocol.

1. Seed 1×10^6 cells in a 6-well for each condition (transduced v. s. untransduced or induced vs. noninduced) in 2 ml RPMI 1640 with 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin.
2. Incubate for 48 h (*see Note 29*).
3. Wash cells twice with 1 ml PBS.
4. Add respective lysis buffer (*see Note 30*).
5. Freeze the plate at $-20\text{ }^{\circ}\text{C}$ for later RNA extraction or proceed directly according to the manufacturer's protocol.
6. Measure your RNA with a spectrophotometer or Qubit system (*see Note 31*).
7. Transcribe 1 μg RNA of each sample into cDNA according to the manufacturer's protocol.
8. Add 7.5 μl SYBR Green mastermix, 0.375 μl of each primer (10 μM), and 6.75 cDNA template (in duplicate) per well of the qPCR plate (*see Note 25*).
9. Seal the plate with an optical adhesive film and transfer it to a q-RT-PCR cycler.
10. Choose the appropriate settings for SYBR Green fluorescence and set up a protocol as follows:
 - Step 1:** 2:00 min $95\text{ }^{\circ}\text{C}$.
 - Step 2:** 0:10 min $95\text{ }^{\circ}\text{C}$.

Step 3: 0:30 min 60 °C.

Repeat **steps 2** and **3** for 50 cycles.

11. Normalize your target gene expression on the expression levels of the respective housekeeping gene used (*see* **Note 32**).

4 Notes

1. If you are unsure which vector to choose, addgene.org is an excellent start and a great repository for many gene expression vectors, both for constitutive and conditional gene expression (i.e., doxycycline-inducible). In general, lentiviral transfer plasmids can be categorized into second and third generation vectors. Whereas second generation plasmids can generate higher viral titers, third generation plasmids show an even better safety profile. Be careful to choose your packaging and envelope plasmids accordingly. In this protocol, second generation transfer plasmids are used.
2. Usually cDNA is used to express a transgene from viral vectors, but sometimes you may want to use the original gene (i.e., to study alternative splicing, etc.). In this case you have to make sure that the resulting plasmid size does not exceed 15 kbp, since the transduction rate will drop significantly above 10 kbp. The cDNA of interest can be easily cloned from a cDNA library, if the transcript is sufficiently expressed in the cell. Otherwise, there are many commercial suppliers for cDNA clones.
3. Sometimes cloning directly from the genome or from a cDNA library does not work well when primers are used that already contain a 5'-overhang. Therefore, we recommend to order both: primers with perfect alignment and primers with a 5'-overhang, and to perform sequential amplification using the product of the first PCR as a template for the second one. Be careful to choose the right polymerase when calculating the annealing temperature of your primers. In our hands, annealing temperatures between 62.0 °C and 65.0 °C work best when using the Phusion polymerase. Furthermore, more than 2 consecutive G/C residues should be avoided both at the 5' and 3' end of your primer. A GC-content of ~50% is optimal. Finally, it's recommended to check primer specificity using Primer-BLAST, especially when cloning from cDNA or genomic DNA libraries.
4. The design of the 5'-overhang depends on the restriction enzymes you want to use. Most enzymes show a very low cut rate when the restriction site resides at the very end of a DNA strand, so include at least four additional bases upstream or downstream, respectively, of your restriction site.

Furthermore, when protein coding genes are cloned, a Kozak sequence should be included directly upstream of the start codon. Omit overhang sequences when calculating annealing temperatures for primers containing 5'-overhangs.

5. For cloning purposes, polymerases with proofreading function should be employed. Q5 polymerase can serve as an alternative to the Phusion polymerase.
6. Ampicillin has a limited half-life. When in solution, ampicillin should be stored at $-20\text{ }^{\circ}\text{C}$ for a maximum of 12 months. Ampicillin-selection plates should be stored at $4\text{ }^{\circ}\text{C}$ and be used within 2 months after preparation. Moreover, the ampicillin should not be added to boiling agar, but when it has already cooled down to lukewarm.
7. The quality of the HEK293T cells is the most crucial factor for good viral titers. You should always aim at cells of low passage number and be careful to avoid full confluence and overgrowth when culturing them.
8. To calculate the vector copies per genome it is important to use primers that do not bind to human components inside the plasmid, that is, the transgene, but truly identify the plasmid. Otherwise, the titers might be overestimated due to the endogenous two copies of each gene inside the genome.
9. We recommend the use of single cutting enzymes that ideally generate sticky ends. To avoid sequential digests due to buffer incompatibility, check for the availability of high fidelity versions of your respective enzymes.
10. If the PCR reaction does not give the expected results, PCR additives such as DMSO (up to 5%) or betaine (up to 2.5 mM) can help to achieve a successful reaction. Alternatively, switch to the provided GC Buffer (Phusion) or add the Q5 GC Enhancer (Q5).
11. The primer concentration can severely affect the specificity of your PCR reaction. When you do not have success cloning your sequence of interest, reducing the primer concentration can help to increase the specificity of your reaction.
12. If conventional PCR protocols do not work, a touchdown approach may be worth considering. For setting up a touchdown reaction, set the annealing temperature $5\text{ }^{\circ}\text{C}$ above your actual annealing temperature and reduce it by $0.5\text{ }^{\circ}\text{C}$ degrees per cycle for the first 10 cycles. Then perform 25 cycles on the actual annealing temperature.
13. Most gel extraction systems are column-based. Try to cut pieces of 200 mg or less to avoid multiple rounds of column-loading and increase DNA yield.

14. Only minimal amounts of DNA are required for this PCR. In fact, exceeding template concentrations can severely interfere with your reactions.
15. Check the respective digestion protocol for each of your enzyme. Some enzymes require unique buffers or higher incubation temperatures (i.e., BssHII).
16. Use appropriate software to calculate the insert mass for a given vector mass and insert–vector ratio (i.e., www.nebiocalculator.neb.com).
17. When performing blunt-end ligation, we recommend an overnight ligation protocol (16 h, 16 °C). We have observed significant mutagenesis at restriction sites when incompletely ligated plasmids were incorporated into competent bacteria.
18. When working with different bacterial clones, open only one tube at a time and ensure proper sterile technique.
19. We recommend a test-digest (digest your plasmid with at least two different restriction enzymes and compare the obtained band sizes on a gel with the expected band sizes) before sending your plasmid to sequencing. A test digest can detect incorporation of bacterial transposons, transgene truncations and restriction site integrity. Furthermore, use transgene-spanning primers for sequencing purposes to not only analyze transgene integrity, but also the correct insertion of the respective transgene. In case of low viral titers or low plasmid concentration in Mini- or Midi-Preps consider full-length sequencing of your plasmid.
20. You might want to seed different densities of HEK293T cells (e.g., 480,000, 530,000, 580,000 cells) to have at least one optimal condition for subsequent transfection on the following day.
21. If the confluence of your cells is below this range, wait a couple more hours before proceeding with **step 3**. If the confluence is too high, repeat **step 1** seeding less cells per well.
22. We highly recommend viral titering; however, for some initial experiments it can be omitted. Please be aware that viral titers of frozen virus is lower than of freshly harvested viral particles. When titering, freeze all aliquots and thaw one for titering purposes. Using fresh virus for titering will lead to significant overestimation of the titer of your remaining, frozen virus.
23. Due to the small volume of ultracentrifugation tubes, PEG precipitation is much better suited for the production of big amounts of viral particles.
24. When qPCR is employed for titering of viral particles, a DNaseI treatment should be applied after removing the viral particles to digest any remaining plasmids that could interfere with your qPCR.

25. Each sample should have at least one technical replicate on the same plate. q-RT-PCR is very sensitive toward even low pipetting inaccuracy. Find a pipetting system that suits you best and work as exact as possible.
26. The MOI is calculated using the following formula: $MOI = V/C$ where V = viral particles and C = cells. In our hands, EwS cell lines are easily transduced in the proposed range and show stable expression of transgenes.
27. An untransduced negative control should always be included.
28. When a chemical selection marker is used (i.e., puromycin resistance), the minimal cytotoxic dose of the respective drug should be determined in each cell line individually before starting the selection process. Therefore, plate 1×10^6 cells of interest in fully supplemented RPMI 1640 with 10% FCS, 1% L-glutamine, and 1% penicillin–streptomycin in each well of a 6-well plate and incubate for 24 h before adding 0 $\mu\text{g}/\text{ml}$, 0.5 $\mu\text{g}/\text{ml}$, 1 $\mu\text{g}/\text{ml}$, 1.5 $\mu\text{g}/\text{ml}$, 2 $\mu\text{g}/\text{ml}$, and 2.5 $\mu\text{g}/\text{ml}$ puromycin to the cells. Incubate for 72 h. If the untransduced cells have not been completely killed in this time, repeat the experiment with a higher dosage. Always try to use the lowest possible dose to avoid selection for overtransduced cells. In our hands, most of the EwS cell lines can be readily killed by puromycin doses of 0.5–2.0 $\mu\text{g}/\text{ml}$.
29. When conditional gene expression systems are used, a time course experiment is recommended to investigate the dynamics of the transgene expression. Therefore, seed 5 different plates and harvest the cells after 24 h, 48 h, 72 h, 96 h, and 120 h. In our hands, transgene expression reaches its maximum after 48–72 h in EwS.
30. When using phenol–chloroform for RNA extraction, make sure to use appropriate protection gear and work under a chemical hood.
31. An accurate measurement of the RNA concentration is not crucial for the subsequent transcription. Some researches do not measure the RNA concentration at this step. However, measurement with a spectrophotometer can give you information about RNA purity and possible contamination with residues of phenol, and so on.
32. When you face supraphysiologic or insufficient expression levels of your gene of interest, single cell cloning can be a useful technique to obtain cells that express your transgene to the desired level.

Acknowledgments

M.M.L.K. acknowledges support from the Friedrich-Baur-Stiftung, as well as from the Dr. Rolf M. Schwiete Stiftung and the Verein zur Förderung von Wissenschaft und Forschung an der Medizinischen Fakultät der Ludwig-Maximilians-Universität München e.V. F.C.A. acknowledges support from the Barbara und Hubertus Trettner Stiftung as well as from the German Cancer Aid in frame of the Max-Eder program (DKH-70112257 granted to T. Grünewald).

References

1. Cohen SN, Chang ACY, Boyer HW, Helling RB (1973) Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci* 70(11):3240–3244. <https://doi.org/10.1073/pnas.70.11.3240>
2. Kamens J (2015) The Addgene repository: an international nonprofit plasmid and data resource. *Nucleic Acids Res* 43(Database issue):D1152–D1157. <https://doi.org/10.1093/nar/gku893>
3. Balbás P, Soberón X, Merino E, Zurita M, Lomeli H, Valle F et al (1986) Plasmid vector pBR322 and its special-purpose derivatives—a review. *Gene* 50(1–3):3–40. [https://doi.org/10.1016/0378-1119\(86\)90307-0](https://doi.org/10.1016/0378-1119(86)90307-0)
4. Celie PH, Parret AH, Perrakis A (2016) Recombinant cloning strategies for protein expression. *Curr Opin Struct Biol* 38:145–154. <https://doi.org/10.1016/j.sbi.2016.06.010>
5. Hughes RA, Ellington AD (2017) Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb Perspect Biol* 9(1). <https://doi.org/10.1101/cshperspect.a023812>
6. Zinder ND, Lederberg J (1952) Genetic exchange in Salmonella. *J Bacteriol* 64(5):679–699
7. Goff SP, Berg P (1976) Construction of hybrid viruses containing SV40 and λ phage DNA segments and their propagation in cultured monkey cells. *Cell* 9(4):695–705. [https://doi.org/10.1016/0092-8674\(76\)90133-1](https://doi.org/10.1016/0092-8674(76)90133-1)
8. Dunbar CE, High KA, Joung JK, Kohn DB, Ozawa K, Sadelain M (2018) Gene therapy comes of age. *Science* 359(6372):eaan4672. <https://doi.org/10.1126/science.aan4672>
9. Kallunki T, Barisic M, Jäättelä M, Liu B (2019) How to choose the right inducible gene expression system for mammalian studies? *Cell* 8(8). <https://doi.org/10.3390/cells8080796>
10. Ahler E, Sullivan WJ, Cass A, Braas D, York AG, Bensinger SJ et al (2013) Doxycycline alters metabolism and proliferation of human cell lines. *PLoS One* 8(5):e64561. <https://doi.org/10.1371/journal.pone.0064561>
11. Chatzispyrou IA, Held NM, Mouchiroud L, Auwerx J, Houtkooper RH (2015) Tetracycline antibiotics impair mitochondrial function and its experimental use confounds research. *Cancer Res* 75(21):4446–4449. <https://doi.org/10.1158/0008-5472.CAN-15-1626>
12. Aalipour A, Chuang H-Y, Murty S, D’Souza AL, Park S-M, Gulati GS et al (2019) Engineered immune cells as highly sensitive cancer diagnostics. *Nat Biotechnol* 37(5):531–539. <https://doi.org/10.1038/s41587-019-0064-8>
13. Maxwell PH, Wiesener MS, Chang GW, Clifford SC, Vaux EC, Cockman ME et al (1999) The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* 399(6733):271–275. <https://doi.org/10.1038/20459>
14. Anz D, Rapp M, Eiber S, Koelzer VH, Thaler R, Haubner S et al (2015) Suppression of intratumoral CCL22 by type I interferon inhibits migration of regulatory T cells and blocks cancer progression. *Cancer Res* 75(21):4483–4493. <https://doi.org/10.1158/0008-5472.CAN-14-3499>
15. Pradhan AK, Talukdar S, Bhoopathi P, Shen X-N, Emdad L, Das SK et al (2017) Mda-7/IL-24 mediates cancer cell-specific death via regulation of miR-221 and the beclin-1 axis. *Cancer Res* 77(4):949–959. <https://doi.org/10.1158/0008-5472.CAN-16-1731>
16. Grünewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4(1):1–22. <https://doi.org/10.1038/s41572-018-0003-x>

17. Agra N, Cidre F, García-García L, de la Parra J, Alonso J (2013) Lysyl oxidase is downregulated by the EWS/FLI1 oncoprotein and its propeptide domain displays tumor suppressor activities in Ewing sarcoma cells. *PLoS One* 8 (6):e66281. <https://doi.org/10.1371/journal.pone.0066281>
18. Cidre-Aranaz F, Grünewald TGP, Surdez D, García-García L, Carlos Lázaro J, Kirchner T et al (2017) EWS-FLI1-mediated suppression of the RAS-antagonist Sprouty 1 (SPRY1) confers aggressiveness to Ewing sarcoma. *Oncogene* 36(6):766–776. <https://doi.org/10.1038/onc.2016.244>



Analysis of Regulatory DNA Sequences by Dual-Luciferase Reporter Assays in Ewing Sarcoma

Tilman L. B. Hölting and Maximilian M. L. Knott

Abstract

Reporter gene assays allow for examining the influence of regulatory DNA sequences on the transcription of target genes. In Ewing sarcoma, the study of these DNA sequences is especially paramount for its main driver mutation is a fusion transcription factor that binds different motifs than its wild-type constituents. Here, we describe the process of analyzing the enhancer activity of regulatory DNA sequences using transfection-based dual-luciferase reporter assays in Ewing sarcoma cell lines. To this end, we provide a protocol for cloning sequences of interest from genomic DNA into a firefly luciferase-containing plasmid, transfecting Ewing sarcoma cells with plasmids and measuring luciferase expression by luminescence. The entire procedure can be completed in 14 days.

Key words Reporter assays, Luciferase assay, Gene expression, Regulatory DNA, Regulatory sequence, Enhancer activity, Promoter activity, Fusion-driven sarcoma, Ewing sarcoma

1 Introduction

Every cell's phenotype is primarily defined by the genes it expresses and thus the proteins it produces. This is also true for the malignant phenotypes of cancer cells, which are caused by the expression of mutated genes, directly or indirectly resulting in altered gene expression patterns. While most tumors harbor many genetic mutations, Ewing sarcoma is instead characterized by a notable scarcity of somatic mutations. This entity's key driver mutation is the disease-defining gene fusion between *EWSR1* and a member of the ETS-family of transcription factors (*FLI1* or *ERG*), resulting in the expression of the fusion transcription factor *EWSR1-FLI1* [1]. Apart from binding to canonical ETS-motifs, *EWSR1-FLI1* additionally binds aberrantly to GGAA-microsatellites that are interspersed throughout the human genome and uses them as de novo enhancers to induce the transcription of nearby genes [2]. As the interaction between the fusion transcription factor and GGAA microsatellites seems to be dependent on the number of motif

repeats and since variations in the length of certain microsatellites have been found to have an impact on the phenotype, functional analysis of regulatory DNA elements is of great interest in Ewing sarcoma research [3–5]. Therefore, techniques to examine the influence of noncoding DNA regions as promoters or enhancers on the transcription of neighboring genes are an important tool, especially for researchers working on understanding the pathophysiology of and generating new therapeutic approaches for Ewing sarcoma.

A simple method to quantitatively assess the transcription-activating function of potential regulatory DNA sequences are transfection-based reporter gene assays. In these assays, the promoter or enhancer activity of DNA regions of interest is evaluated based on their ability to induce the expression of a reporter gene. To this end, these regions of interest are cloned into plasmids upstream of a reporter gene whose expression is then evaluated in cells transfected with this plasmid [6]. Most commonly, firefly luciferase, an enzyme from the firefly *Photinus pyralis* catalyzing a light producing reaction, is used as the reporter gene for several reasons [7]. First, it allows for rapid quantitative readouts over a broad intensity range when using modern luminometers. Additionally, by cotransfecting with a plasmid containing *Renilla* luciferase from the sea pansy *Renilla reniformis* regulated by a constitutive promoter, an easy internal control for normalizing differences in sample mass and transfection efficiency is available [8].

This protocol first describes the process of cloning putative enhancers from genomic DNA into a firefly luciferase reporter plasmid (Figs. 1 and 2). Afterwards, Ewing sarcoma cells are transfected with the generated plasmid containing the sequence of interest, or a control plasmid that does not contain any additional sequence. Both are cotransfected with a plasmid containing *Renilla* luciferase for normalization. Finally, the firefly and *Renilla* luciferase expression is determined by use of a Dual-Luciferase[®] Reporter Assay System. Following this approach, different sequences' transcription-inducing activities in Ewing sarcoma cell lines can be compared. In addition to that, by performing luciferase assays in combination with knockdown of certain transcription factors, such as the fusion oncogene *EWSR1-FLII*, these transcription factors' impact on the transactivation activity of the cloned regulatory sequence can be evaluated.

2 Materials

1. pGL3 containing firefly luciferase and SV40 promoter (Promega).
2. pRL-SV40 containing *Renilla* luciferase and SV40 promoter (Promega).

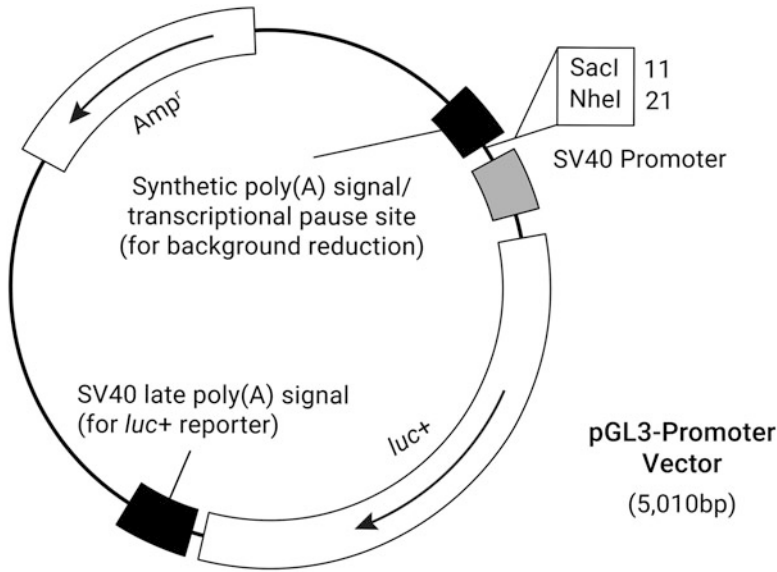


Fig. 1 Vectormap of pGL3-Promoter Vector plasmid. SacI and NheI restriction sites used in this protocol are located upstream of SV40 promoter that controls the expression of the firefly luciferase gene

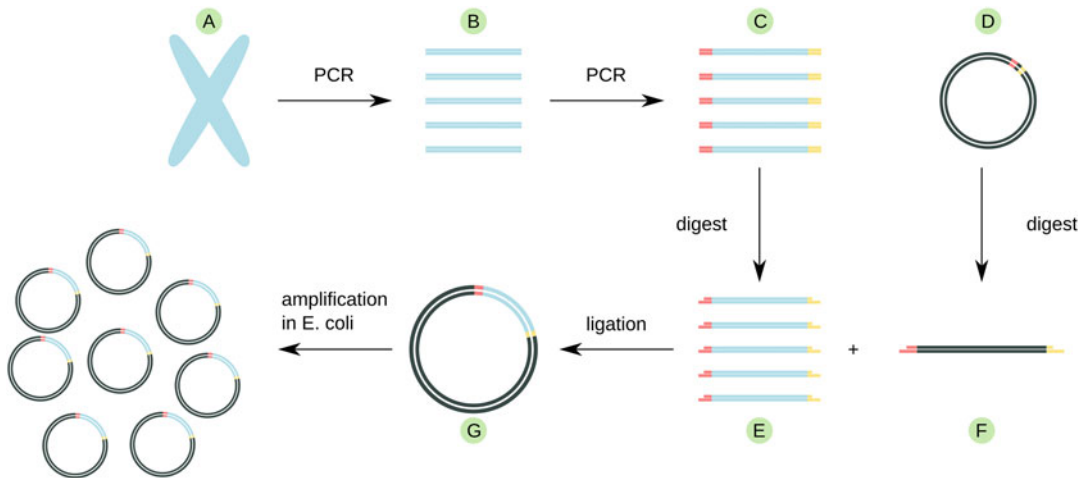


Fig. 2 Workflow of cloning process. Amplify sequence of interest from genomic DNA (a). Amplify resulting DNA (b) with primers with overhangs containing restriction sites for SacI (red) and NheI (yellow). Digest sequence flanked by cutting sites (c) and pGL3 plasmid (d). Ligate digested sequence of interest (e) and plasmid (f). Transform competent *E. coli* with resulting plasmid (g) and cultivate a successfully transformed clone for plasmid amplification

3. Genomic DNA containing the sequence of interest.
4. PCR primers flanking the region of interest (*see Note 1*).
5. PCR primers flanking the region of interest with 5'-overhangs (*see Note 2*).

6. Phusion High-Fidelity DNA Polymerase, Thermo Scientific.
7. 10 mM dNTP mix.
8. 8 well 0.2 ml PCR tube strips with caps.
9. Agarose for agarose gel electrophoresis.
10. TRIS, acetic acid, EDTA buffer (10× TAE buffer): 0.4 M TRIS, 0.2 M acetic acid, and 10 mM EDTA in deionized water. (48.4 g TRIS base, 11.4 mL acetic acid, and 3.7 g EDTA in 800 mL water. Stir at room temperature until all compounds are in solution. Add H₂O to 1 L.)
11. DNA gel stain (e.g., ethidium bromide or SYBR Safe).
12. One horizontal gel electrophoresis chamber.
13. DNA gel extraction and PCR cleanup kit.
14. Restriction enzymes: SacI-HF and NheI-HF, New England Biolabs.
15. T4 DNA Ligase, New England Biolabs.
16. Chemically competent *E. coli*.
17. Agar plates containing 100 µg/mL ampicillin.
18. LB broth containing 100 µg/mL ampicillin.
19. Plasmid DNA midi prep kit.
20. Spectrophotometer for DNA purity assessment and quantification.
21. Qubit fluorometer for DNA quantification.
22. Ewing sarcoma cell lines of interest (e.g., with inducible KD of the fusion oncogene).
23. RPMI-1640 cell culture medium containing 10% fetal bovine serum (FBS).
24. 24-well cell culture plates.
25. Opti-MEM™ I Reduced Serum Medium, Gibco.
26. Phosphate buffered saline (PBS).
27. Lipofectamine™ LTX Reagent with PLUS™ Reagent, Invitrogen.
28. Dual-Luciferase® Reporter Assay System, Promega. Prepare Luciferase Assay Reagent II and Stop & Glo® Reagent according to manufacturer's protocol and store in aliquots of 1 mL at -80 °C.
29. White 96-well plates, polystyrene, opaque bottom.
30. Microplate luminometer, optionally with two reagent injectors (e.g., Orion II Microplate Luminometer, Titertek-Berthold).

3 Methods

3.1 Cloning of Reporter Plasmids

Perform all steps at room temperature unless specified otherwise. When handling bacteria or medium to grow them in, work near a Bunsen burner and pay attention on using proper sterile technique.

1. To amplify the sequence of interest from genomic DNA using Phusion high-fidelity polymerase, prepare a PCR master mix by adding 10 μl of HF buffer, 2 μl of genomic DNA (50–125 ng/ μl), 1 μl dNTP (10 mM), 0.5 μl Phusion polymerase, and 35.25 μl of PCR-grade H_2O per sequence of interest–primer pair.
2. Vortex mildly, spin down and add 48.75 μl to a 0.2 ml PCR tube for each reaction.
3. To each reaction, add 0.65 μl of forward primer and 0.65 μl of reverse primer (10 μM).
4. Vortex mildly and spin down.
5. Put reaction tubes into PCR cycler and start PCR following cycling instructions given in the Phusion high-fidelity user guide (*see* **Notes 3** and **4**).
6. To prepare a 1% agarose gel, measure 1 g of agarose and mix with 100 ml of 1 \times TAE buffer.
7. Microwave until all agarose is dissolved (around 2 min at 900 W).
8. Let solution cool down to approximately 50 $^\circ\text{C}$.
9. Under the fume hood, add ethidium bromide to a final concentration of 0.5 $\mu\text{g}/\text{mL}$.
10. Pour mixture into a gel tray with an already inserted comb.
11. Perform agarose gel electrophoresis: Mix PCR products with 6 \times loading dye delivered with the restriction enzymes.
12. Add a 12.5 μl of a molecular weight ladder into first well of the gel. Add 50 μl of PCR product mixed with loading dye to wells of your gel.
13. Separate PCR products by applying 80 V over 30 min.
14. Inspect gel under UV light. If DNA bands corresponding to expected lengths are visible, neatly cut out the gel containing the DNA and store in 1.5 ml reaction tube.
15. Extract DNA from cut out gel blocks using commercial kit and following manufacturer's protocol.
16. Measure DNA concentration using spectrophotometer.
17. Create DNA fragments with digestible overhangs by following PCR protocol as described in **steps 1–5** with primers containing overhangs. As template, use 1 μl of DNA created in **step 3** diluted to 10 ng/ μl (*see* **Note 5**).

18. Repeat gel electrophoresis and gel extraction as described in **steps 6–14**.
19. Measure DNA concentration using spectrophotometer.
20. Digest extracted DNA as well as pGL3 plasmid with SacI-HF and NheI-HF: Mix 1 µg of DNA with 5 µl Cutsmart buffer, 1 µl of SacI-HF, and 1 µl of NheI-HF.
21. Fill up with DNase-free water to 50 µl.
22. Vortex mildly and spin down.
23. Incubate at 37 °C for 30 min.
24. Perform PCR cleanup using commercial kit following manufacturer's protocol to create digested pGL3 and digested inserts ready for sticky-end ligation. Measure DNA concentration using spectrophotometer.
25. Perform sticky-end ligation of 25 ng of digested pGL3 with digested inserts using a mass of digested insert corresponding to a molecular plasmid–insert ratio of 1:5: Mix plasmid and insert DNA with 2 µl of T4 ligase buffer and 1 µl of T4 ligase in a total reaction volume of 20 µl (filled up with DNase-free H₂O).
26. Mix by pipetting up and down and spin down.
27. Incubate at room temperature for 15 min (*see Note 6*).
28. Transform competent cells with ligation mix following manufacturer's protocol and streak transformed bacteria on an LB agar plate containing 100 µg/ml ampicillin.
29. Incubate inoculated agar plates overnight (12–16 h) at 37 °C.
30. Pick eight single colonies by touching them with a sterile pipette tip (or toothpick) and swirl the tip in 10 µl of LB containing 100 µg/ml ampicillin in a 0.2 ml PCR tube. Open only one tube at a time and ensure proper sterile technique! (*see Note 7*).
31. Perform colony PCR to detect colonies containing the desired insert: Create PCR master mix as described in **step 1** but omitting the genomic DNA.
32. For each picked colony, add 49 µl of PCR master mix to a fresh 0.2 ml PCR tube.
33. Add 1 µl of LB containing picked bacteria from **step 9** to each tube.
34. Perform PCR using cycling program that lead to visible bands at expected length in the previous PCRs.
35. Add 91 µl of LB containing 100 µg/ml ampicillin each of the tubes containing the remaining 9 µl of inoculated LB and Store in a shaking incubator at 37 °C and 200 rpm.

36. Separate the products of the colony by gel electrophoresis as described above and note the numbers of colonies in which DNA bands of expected lengths are visible.
37. Cultivate colonies noted in **step 36** by adding the entire 100 μl of the colony to 100 ml of LB containing 100 $\mu\text{g}/\text{ml}$ ampicillin.
38. Incubate overnight (12–16 h) in bacterial shaking incubator at 37 °C and 200 rpm.
39. Next morning the LB should be turbid due to bacterial growth. Create glycerol stock of bacterial culture by mixing 500 μl of turbid LB with 500 μl of 50% glycerol. Store immediately at –80 °C.
40. Extract plasmid from the remaining culture using commercial midi prep kit and following manufacturer's instructions.
41. Assess prep purity by spectrophotometry.
42. Measure DNA quantity by use of Qubit.
43. Confirm correct plasmid sequence by sanger sequencing using commercial sequencing service (*see Note 8*).

3.2 Transfection of Ewing Sarcoma Cells

To assess the influence of the cloned sequences on luciferase expression, Ewing sarcoma cells are transfected with each of the cloned plasmids in separate wells. To compare to the baseline firefly luciferase expression, each cell line used should also be transfected with unaltered pGL3-Promoter in another well. All pGL3 plasmids are cotransfected with pRL as an internal control for differences in transfection efficiencies and cell mass. Cells are cultivated in an incubator at 37 °C with 5% CO₂ in a humidified atmosphere. This protocol uses the transfection agent Lipofectamine[®] LTX (*see Note 9*).

1. Twenty-four hours before transfection seed 5×10^5 Ewing sarcoma cells into 24 well plates in 0.5 ml RPMI containing 10% FBS.
2. On day of transfection, for each well to be transfected mix 212 ng of pGL3 plasmid (100 ng/ μl) containing sequence of interest with 2.12 ng of pRL (10 ng/ μl) in Opti-MEM[®] for a final volume of 55 μl .
3. Add 0.45 μl of PLUS[™] Reagent and mix gently by pipetting up and down 10 times.
4. Incubate solution for 10 min at room temperature.
5. Add 0.58 μl of Lipofectamine[®] LTX. Mix thoroughly by pipetting up and down at least 10 times without producing bubbles.
6. Incubate for 25 min at room temperature.

7. Add 50 μl of transfection per well and move plate in 8-like motion to equally disperse it.
8. After 4–6 h, change medium to prevent reduction in cell viability.

3.3 Assessing Luciferase Expression

In this protocol, luciferase levels are determined using the Dual-Luciferase[®] Reporter Assay System which allows quick measurement of firefly and *Renilla* luciferase activity in the same well. Alternative reagent providers are available (*see Note 10*). This protocol describes the procedures for luminometers with and without reagent injectors.

1. 48 h after transfection, dilute 5 \times Passive Lysis Buffer (from Dual-Luciferase[®] Reporter Assay System) with distilled water to a sufficient volume of 1 \times working solution. Per well of a 24-well plate, 100 μl of 1 \times working solution are required.
2. Aspirate medium from each well.
3. Wash each well with 200 μl PBS without detaching adherent cells.
4. Add 100 μl of 1 \times Passive Lysis Buffer from the Dual-Luciferase[®] Reporter Assay System to each well. Make sure that the entirety of each well's cell layer is covered in lysis buffer.
5. Incubate at room temperature on an orbital shaker for 15 min.
6. Transfer resulting lysate from the well into a 0.2 μl microtube. Spin down with tabletop centrifuge for 30 s. Store samples on ice (*see Note 11*).
7. For each sample to be measured, thaw 200 μl of each LARII and Stop & Glo[®] Reagent (both from Dual-Luciferase[®] Reporter Assay System). If you intend to use a luminometer with automatic reagent injectors, thaw enough additional buffer volume for priming of the injectors, skip the next 5 steps and continue with the alternative assay protocol. Volume required for priming depends on your luminometer and can usually be found in the manufacturer's protocol.
8. As technical duplicates add 10 μl of each sample to two wells of a white 96-well plate. Additionally, as a background control, add 10 μl of 1 \times Passive Lysis Buffer to 5 wells.
9. Add 100 μl of LARII to each well and mix by pipetting up 2–3 times (*see Note 12*).
10. Insert plate into luminometer and perform 10 s measurements of luminescence for each well containing sample or Lysis Buffer.
11. Remove plate from luminometer. Add 100 μl of Stop & Glo[®] Reagent to each well and mix by pipetting up 2–3 times (*see Note 12*).

12. Insert plate into luminometer and perform 10s measurements of luminescence for each well containing sample or Lysis Buffer (*see* **Note 13**).

Alternative assay protocol: If your luminometer is equipped with at least two reagent injectors, follow **steps 1–7** and then continue as follows:

1. Create the following measuring/injection protocol for your luminometer:
 - (a) Injection of 100 μ l of LARII.
 - (b) Delay for 2.05 s.
 - (c) Measurement of luminescence for 10 s.
 - (d) Delay for 10 s.
 - (e) Injection of 100 μ l of Stop & Glo[®] Reagent.
 - (f) Delay for 2.05 s.
 - (g) Measurement of luminescence for 10 s.
2. Add 10 μ l of each sample to two wells of a white 96-well plate. These duplicates will serve as technical replicates. Additionally, add 10 μ l of 1 \times Passive Lysis Buffer to 5 wells for measurements of background luminescence.
3. Insert plate into luminometer and run the created protocol.
4. After completion of measurements, make sure to clean in the injectors according to the machine's user guide (*see* **Note 12**).

3.4 Data Analysis

1. Calculate background luminescence by taking the mean of firefly and *Renilla* luminescence of the five control wells in which lysis buffer was added instead of sample.
2. Subtract firefly and *Renilla* background luminescence from samples' values for firefly and *Renilla* luminescence respectively.
3. Calculate the mean of firefly and *Renilla* luminescence for the samples' technical replicates.
4. Divide mean firefly luminescence by the *Renilla* luminescence for each sample. The higher the resulting ratio, the higher the transcriptional induction that is caused by the sequence cloned into the pGL3.

4 Notes

1. When designing primers for amplification from genomic DNA, make sure to confirm the specificity of the PCR reaction by using tools such as UCSC Genome Browser and UCSC In-Silico PCR. Design primers with a T_m of 60 °C \pm 4 °C, a

length of at least 20 nucleotides and with a GC content between 45% and 55%. The difference in annealing temperature of forward and reverse primer should be lower than 5° C. The annealing temperatures can easily be calculated using a web-based T_m Calculator, for example T_m Calculator on thermofisher.com. For this protocol, annealing temperatures should be determined using calculator settings for Phusion DNA polymerase which requires higher annealing temperatures for the same primer sequences as compared calculations for use with Taq polymerase.

2. 5'-overhang for the forward primer should be TATGAGCTC (buffer and **cutting site** for SacI-HF). 5'-overhang for the reverse primer should be TATGCTAGC (buffer and **cutting site** for NheI-HF). These overhang containing primers are used to add appropriate restriction sites to the resulting DNA molecules. Most restriction enzymes don't work if their cutting site is directly at the end of a linear DNA molecule. Hence, for the enzymes used in this protocol three filler nucleotides have been included in the proposed overhangs. Other enzymes may require longer fillers.
3. To increase the specificity of your PCR reactions, starting with an annealing temperature of 5 °C higher than the calculated one and reducing the annealing temperature by 0.5 °C over the next 10 cycles may be beneficial (9). After this so-called touch-down, continue for another 25 cycles with an annealing temperature calculated based on the primer sequences.
4. In case of difficulties when amplifying from genomic DNA, addition of 3–5% DMSO may improve results. Additionally, try digesting the genomic DNA with an enzyme, that is known not to cut within the sequence of interest and separate the resulting DNA fragments by gel electrophoresis. Cut out the gel at the height you expect the fragment containing your sequence of interest to run at and extract DNA to use as alternative template.
5. Omit overhang sequences when calculating annealing temperatures for primers containing 5'-overhangs.
6. For quick calculation of the required mass of insert, the Ligation Calculator provided online by New England Biolabs as part of the NEBioCalculator is recommended.
7. When in solution, ampicillin should be stored at –20 °C for a maximum of 6 months.
8. A well working sequencing primer 5' of the multiple cloning site is pGL3-RV3 (5'-CTAGCAAAATAGGCTGTCC-3').
9. For even cheaper assays, especially when analyzing many different sequences, most Ewing sarcoma cell lines can be transfected

efficiently enough for luciferase assays using PEI MAX (Transfection Grade Linear Polyethylenimine Hydrochloride (MW 40,000), Cat-no. 24765-1, Polysciences). Per well to be transfected prepare DNA solution by mixing 231 ng of pGL3 plasmid (100 ng/ μ l) containing sequence of interest with 2.31 ng of pRL (10 ng/ μ l) in Opti-MEM[®] for a final volume of 27.5 μ l. Prepare PEI solution by mixing 1.9 μ g of PEI MAX (1 μ g/ μ l) in Opti-MEM[®] for a final volume of 31.25 μ l. Vortex both mixes and incubate for 10 min at room temperature. Add 27.5 μ l of PEI solution to DNA solution and mix by pipetting up and down 10 times without producing bubbles. Incubate for 5 min at room temperature. Add 50 μ l of DNA-PEI mix dropwise to the well.

10. As an alternative to the reagents provided by Promega, Beetle-Juice and Renilla-Juice by PJK GmbH have successfully been used in our lab.
11. For longer term storage, lysates should be kept at -80°C .
12. For manually adding reagents to the multi-well plate, using a multichannel pipette is recommended as it reduces the time difference between addition of reagent and measuring of luminescence between your wells.
13. After readout, plates should be washed with distilled water (filling all the used wells at least 3 times). When completely dried, the plate can be reused for future assays.

References

1. Delattre O, Zucman J, Plougastel B et al (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 359:162–165. <https://doi.org/10.1038/359162a0>
2. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4:5. <https://doi.org/10.1038/s41572-018-0003-x>
3. Johnson KM, Taslim C, Saund RS, Lessnick SL (2017) Identification of two types of GGAA-microsatellites and their roles in EWS/FLI binding and gene regulation in Ewing sarcoma. *PLoS One* 12:e0186275. <https://doi.org/10.1371/journal.pone.0186275>
4. Musa J, Cidre-Aranaz F, Aynaud M-M et al (2019) Cooperation of cancer drivers with regulatory germline variants shapes clinical outcomes. *Nat Commun* 10:4128. <https://doi.org/10.1038/s41467-019-12071-2>
5. Grünewald TGP, Bernard V, Gilardi-Hebenstreit P et al (2015) Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite. *Nat Genet* 47:1073–1078. <https://doi.org/10.1038/ng.3363>
6. Brasier AR, Ron D (1992) Luciferase reporter gene assay in mammalian cells. *Methods Enzymol* 216:386–397
7. Gould SJ, Subramani S (1988) Firefly luciferase as a tool in molecular and cell biology. *Anal Biochem* 175:5–13. [https://doi.org/10.1016/0003-2697\(88\)90353-3](https://doi.org/10.1016/0003-2697(88)90353-3)
8. Sherf BA, Navarro SL, Hannah RR, Wood KV (1996) Dual-Luciferasetm reporter assay: an advanced co-reporter technology integrating firefly and Renilla luciferase assays. *Promega Notes Mag* 57:02



Proliferation Assessment by Trypan Blue Exclusion in Ewing Sarcoma

Cornelius Maximilian Funk and Julian Musa

Abstract

Cell proliferation is broadly defined as a process leading to an increase of cell number, essentially depending on a balance between cell cycle progression/cell division, cell death, and cellular senescence. Deregulation of cell proliferation is a key feature of cancer cells, making assessment of proliferation a central methodological issue in cancer research. Especially in Ewing sarcoma (EwS) that exhibit a high proliferative capacity, experimental assessment of proliferation in preclinical research plays an important role. Among the variety of applicable methods, trypan blue exclusion is described here as a robust, easy-to-perform, and cost-effective method to assess cell proliferation in an experimental setting.

Key words Ewing sarcoma, Proliferation, Cell cycle progression, Cell division, Cell death, Cellular senescence, Trypan blue

1 Introduction

Cell proliferation is broadly defined as a process leading to an increase of cell number. This process depends on a balance between cell cycle progression/cell division, cell death, and cellular senescence and is physiologically important for maintenance and regeneration of tissues [1, 2]. Deregulation of this balance, for example by alterations in gene expression caused by inherited or acquired mutations, may lead to severe tissue dysfunction and disease development [3–5]. “Limitless replicative potential” and “evading programmed cell death” have been described as classical hallmarks of cancer by Hanahan and Weinberg in 2000 [4, 5]. Therefore, assessment of cell proliferation is methodologically essential in cancer research.

A variety of methods exist to analyze cell proliferation which include “direct” and “indirect” approaches and have different advantages and disadvantages [6]. “Direct” approaches measure the amount of cells actively dividing in a cell population, whereas “indirect” approaches extrapolate from cell number/viability or

metabolic activity to cell proliferation [6]. “Direct” approaches use either incorporation of nucleoside analogs during DNA synthesis (such as BrdU) [6, 7], detection of cell cycle-associated proteins (such as Ki-67) with antibodies [6, 8], or photometric methods using cytoplasmic proliferation dyes that dilute with each cell division equivalently to daughter cells (such as carboxyfluorescein diacetate succinimidyl ester (CFSE)) [6, 9]. “Indirect” methods include cell counting (with or without cell viability stains, such as trypan blue) [6, 10, 11] and metabolic activity assays (such as resazurin) [6, 12].

The terms “direct” and “indirect” refer to the strict definition of cell proliferation in terms of cell cycle progression. However, seeing cell proliferation not only as a result of cell cycle progression, but also of cell death and cellular senescence, suggests using a “screening” method for initial proliferation assessment that does not solely measure cell cycle progression. Although several methods are suitable for assessment of proliferation in EwS, and the choice of method will depend on the specific question and experimental background, cell counting using trypan blue cell viability staining shall be reviewed here, as the results do not only include information about total cell number, but also about cell viability. When staining cells with trypan blue, viable cells with intact membranes do not incorporate the dye, whereas dead cells without intact membranes do, which makes it possible to distinguish viable from dead cells [10, 11]. The assay measures proliferation in a broader sense and gives a hint on whether the observed phenotype is more related to cell cycle progression or cell death, which can then be specified in subsequent assays (such as propidium iodide (PI) staining and Annexin V/PI staining using flow cytometry, or further “direct” proliferation assays). Further advantages are that the trypan blue exclusion assay is easy-to-perform and cost-effective, and leads to robust results. Compared to “indirect” metabolic assays the readout is also not disturbed by potential metabolic changes due to different treatment conditions. Trypan blue cell viability staining has been frequently used in EwS research to assess cell proliferation [13–15].

2 Materials

1. Growth media: Appropriate medium for each cell line (500 ml) supplemented with stable glutamine, fetal bovine serum (FBS, 50 ml), and penicillin (100 U/ml)–streptomycin (100 µg/ml) (5 ml). For most commonly used Ewing sarcoma cell lines RPMI medium is appropriate.
2. T25 or T75 flasks.
3. Sterile 6- and 96-well plates.

4. Phosphate buffered saline (PBS).
5. Trypsin.
6. Accutase.
7. Reaction tube (15 ml).
8. Trypan blue (0.4%).
9. Hemocytometer (Neubauer improved chamber, preferably standardized single-use hemocytometer).
10. Hand counter (preferably two, one for viable cells and one for dead cells).

3 Methods

3.1 Preparation of Cells and Seeding

1. Grow cells under standard cell culture conditions (37 °C, 5% CO₂, humidified atmosphere) until about 70% confluency. Make sure the cell have sufficient growth medium and are mycoplasma negative) (*see Note 1*).
2. Remove growth medium, wash the cells with PBS, and trypsinize cells (add 3 ml Trypsin for T75, 1 ml for T25) incubating at 37 °C for 2–5 min depending on the cell line.
3. Mix trypsinized cells with growth medium in a 2:1 ratio of medium:trypsin and spin down the cells at $314 \times g$ for 4 min.
4. Remove supernatant and resuspend the cells in fresh growth medium (5–10 ml, depending on the pellet size).
5. Invert cell suspension 5 times (flicking of the tube is additionally necessary in case of pellet formation when leaving the cells in suspension for several minutes) and determine cell concentration for each cell line (*see Note 2*).
6. Choose the number of cells you want to seed per well for every cell line. The number of cells seeded needs to be adapted to the growth rate of the used cell lines, the duration of the experiment, and the experimental conditions (e.g., application of cell stress). Table 1 gives a rough estimation of recommended cell numbers to seed per well (6-well plate), specifically for Ewing sarcoma cell lines (*see Note 3*).

3.2 Experimental Procedure

1. Carefully check your cells a short time after seeding to evaluate consistent attachment of cells in all wells of an experimental unit.
2. Monitor your cells throughout the experiment, especially regarding obvious differences between experimental conditions.
3. Change medium at least one time consistently in every well of an experimental unit after seeding, in order to eliminate cells

Table 1

Estimation of recommended cell numbers to seed per well (6-well plate) depending on the cell line, growth condition, and duration of the experiment

Growth condition	Normal conditions			Stress conditions		
	72 h	96 h	120 h	72 h	96 h	120 h
Faster growing cell lines (e.g., A673, TC-32, RDES)	9×10^4	7×10^4	6×10^4	14×10^4	12×10^4	11×10^4
Slower growing cell lines (e.g., SK-N-MC, MIC, CHLA-10)	11×10^4	9×10^4	8×10^4	16×10^4	14×10^4	13×10^4

that died during the process of detachment/seeding. The time-point of changing medium depends on the individual experimental design (*see* **Notes 4** and **5**). When changing medium, make sure to do so very gently in order to not lose attached cells (*see* **Note 4**).

3.3 Readout

1. Collect the supernatant of each well in a collection tube (*see* **Note 6**).
2. Wash cells in each well with 1 ml of PBS (*see* **Note 7**).
3. Trypsinize cells for detachment (500 μ l). When cell stress conditions were applied during the experimental procedure (e.g., nutrient starvation or transfection reagent toxicity) use Accutase (500 μ l) instead of Trypsin as a gentler alternative (*see* **Note 7**).
4. In case of Trypsin usage incubate cells at 37 °C for 2–5 min depending on the cell line. In case of Accutase usage incubate cells for 2–5 min at room temperature.
5. Rinse the wells in a standardized way 5 times with applied Trypsin or Accutase (*see* **Note 8**).
6. Transfer the detached cells with the Trypsin/Accutase into the collection tube containing the supernatant (counting suspension).
7. Prepare 10–40 μ l of trypan blue, depending on the amount of cells in the sample, in a well of a 96-well plate per cell count. Count every sample of the experiment at least 2 times as technical replicates. Keep the amount of trypan blue constant for each experimental unit (*see* **Notes 9** and **10**).
8. Mix the counting suspension by inverting the reaction tube 5 times. Additional flicking of the tube is necessary in case of pellet formation (*see* **Note 11**).
9. Transfer 10 μ l of each sample (counting suspension) into the respective wells of the 96-well plates containing trypan blue.

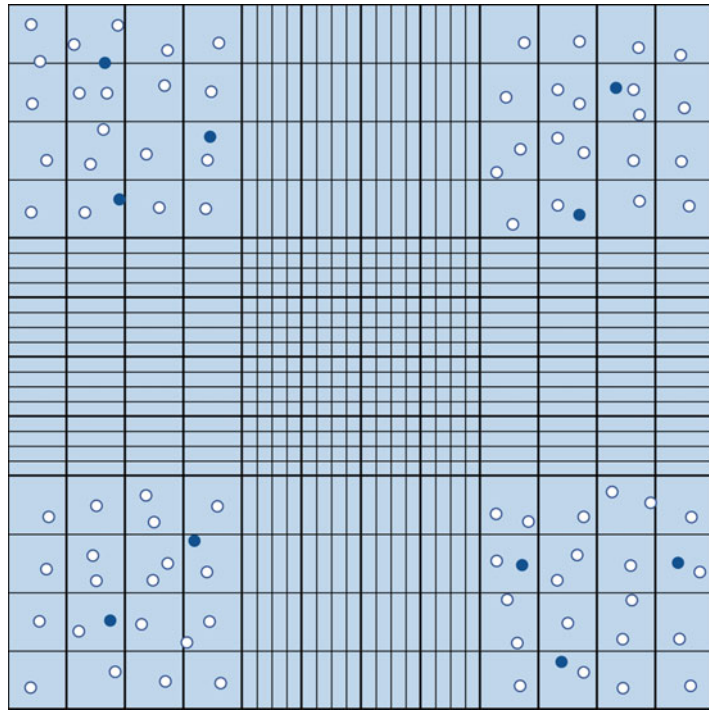


Fig. 1 Scheme of a Neubauer chamber with viable (light gray) and dead (blue) cells stained with trypan blue

10. Mix the trypan blue/counting suspension mix by pipetting the total volume up and down 5 times.
11. Transfer 10 μl of the trypan blue/counting suspension mix into a chamber of a standardized single-use hemocytometer (*see Note 12*).
12. Count viable cells (light gray) and dead cells (blue) in at least two technical replicates (*see Note 13*) (Fig. 1).

4 Notes

1. Make sure the cells are in a good condition (about 70% confluency, sufficient growth medium, mycoplasma negative) when seeding, because this can (unsystematically) effect cell survival during the process of detachment, counting, and seeding and can lead to potential variations of seeded cell numbers per well.
2. When initially seeding the cells, standardize inverting of reaction tubes for every tube and limit the number of inversions to a reasonable number to avoid application of (unsystematic) mechanical stress to the cells, which could potentially lead to variations of seeded cell numbers per well.

3. Make sure that the cell number seeded is adapted to the growth rate of the used cell lines, the duration of the experiment, and the experimental conditions (e.g., application of cell stress). It is important to choose a number of cells that allows sufficient (exponential) growth with which a potential effect between experimental conditions can emerge properly, but with which the confluency does not reach 100% of the well, since this could limit further proliferation and may influence cell numbers finally counted in different experimental conditions.
4. When changing medium during the experimental procedure, make sure to not touch the bottom of the wells in order not to scratch any cells. Also, aspiration must be done very gently, preferably only at the surface edge of the medium after putting the plate in a 45° angle in order to not lose any attached cells. Pipet fresh medium only very gently directly to the walls of the wells.
5. Make sure the amount of medium on top of the wells (supernatant) when ending the experiment and performing the read-out is as low as reasonable, meaning the lowest amount of medium with which the cells do not starve (medium should stay red and not turn yellow), in order to achieve a high cell concentration of the counting suspension (*see* Subheading 3.3, step 6).
6. Be careful to not unintentionally aspirate the supernatant of your wells containing the dead cells.
7. For detachment of cells, make sure you use Accutase instead of Trypsin when your experiment contains conditions leading to cell stress (e.g., nutrient starvation or transfection reagent toxicity). This reduces the probability of cell death caused by the process of detachment and counting, as Accutase detaches cells more gently. Washing the wells with PBS before Accutase or Trypsin application helps to achieve faster and complete detachment. In order to not significantly detach cells before application of Trypsin/Accutase, perform washing with PBS very gently as described for medium in **Note 4**.
8. Rinse the wells in a standardized number of times (5 times recommended, in order to avoid application of unsystematic mechanical stress to the cells) with the Trypsin or Accutase you used to detach the cells in order to transfer all cells of the well into your counting suspension.
9. Spin down the trypan blue before usage, as usually some debris can be found in a bottle/tube of trypan blue. This debris can disturb cell counting and, in some instances, can be confused with dead cells while counting.
10. Make sure to adjust the amount of trypan blue with which you mix 10 µl of your counting solution (*see* Subheading 3.3,

steps 7 and 9) for every functional unit of your experiment, in order to optimally count between 70 and 120 cells per chamber in the control condition. This allows a compromise between precision and time when counting.

11. When finally counting cells of your wells, standardize the number of inversions of the reaction tubes and limit it to a reasonable number to avoid application of (unsystematic) mechanical stress to the cells, which could lead to unsystematic cell death immediately before and/or while counting, and affect the number of trypan blue positive cells you count.
12. Especially when performing the readout, preferably use standardized single-use hemocytometers to increase precision and save time. If not available, normal multi-use glass Neubauer improved chambers can be used instead. Alternatively, automated cell counters may be used.
13. When counting, use two different hand counters (one in each hand) for simultaneous separated viable and dead cell counting in order to save time.

Acknowledgments

J.M. was supported by a scholarship of the Kind-Philipp Foundation and the Rudolf und Brigitte Zenner Foundation.

References

1. Soteriou D, Fuchs Y (2018) A matter of life and death: stem cell survival in tissue regeneration and tumour formation. *Nat Rev Cancer* 18:187–201. <https://doi.org/10.1038/nrc.2017.122>
2. Gurtner GC, Werner S, Barrandon Y, Longaker MT (2008) Wound repair and regeneration. *Nature* 453:314–321. <https://doi.org/10.1038/nature07039>
3. Marte B (2004) Cell division and cancer. In: *Nature*. <https://www.nature.com/articles/432293a>. Accessed 19 Oct 2019
4. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* 100:57–70. [https://doi.org/10.1016/S0092-8674\(00\)81683-9](https://doi.org/10.1016/S0092-8674(00)81683-9)
5. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144:646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
6. Romar GA, Kupper TS, Divito SJ (2016) Research techniques made simple: techniques to assess cell proliferation. *J Invest Dermatol* 136:e1–e7. <https://doi.org/10.1016/j.jid.2015.11.020>
7. Matatall KA, Kadmon CS, King KY (2018) Detecting hematopoietic stem cell proliferation using BrdU incorporation. *Methods Mol Biol* 1686:91–103. https://doi.org/10.1007/978-1-4939-7371-2_7
8. Scholzen T, Gerdes J (2000) The Ki-67 protein: from the known and the unknown. *J Cell Physiol* 182:311–322. [https://doi.org/10.1002/\(SICI\)1097-4652\(200003\)182:3<311::AID-JCP1>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4652(200003)182:3<311::AID-JCP1>3.0.CO;2-9)
9. Lyons AB, Blake SJ, Doherty KV (2013) Flow cytometric analysis of cell division by dilution of CFSE and related dyes. *Curr Protoc Cytom Chapter 9:Unit9.11*. <https://doi.org/10.1002/0471142956.cy0911s64>
10. Strober W (2001) Trypan blue exclusion test of cell viability. *Curr Protoc Immunol Appendix 3:Appendix 3B*. <https://doi.org/10.1002/0471142735.ima03bs21>
11. Strober W (2015) Trypan blue exclusion test of cell viability. *Curr Protoc Immunol* 111:A3.B.1–A3.B.3. <https://doi.org/10.1002/0471142735.ima03bs111>

12. Czekanska EM (2011) Assessment of cell proliferation with resazurin-based fluorescent dye. *Methods Mol Biol* 740:27–32. https://doi.org/10.1007/978-1-61779-108-6_5
13. Musa J, Cidre-Aranaz F, Aynaud M-M et al (2019) Cooperation of cancer drivers with regulatory germline variants shapes clinical outcomes. *Nat Commun* 10:4128. <https://doi.org/10.1038/s41467-019-12071-2>
14. Marchetto A, Ohmura S, Orth MF et al (2020) Oncogenic hijacking of a developmental transcription factor evokes vulnerability toward oxidative stress in Ewing sarcoma. *Nat Commun* 11(1):2423. <https://doi.org/10.1038/s41467-020-16244-2>
15. Dallmayer M, Li J, Ohmura S et al (2019) Targeting the CALCB/RAMP1 axis inhibits growth of Ewing sarcoma. *Cell Death Dis* 10:116. <https://doi.org/10.1038/s41419-019-1372-0>



Drug Screening by Resazurin Colorimetry in Ewing Sarcoma

Julian Musa and Florencia Cidre-Aranaz

Abstract

In Ewing sarcoma (EwS), development of new therapeutic strategies is crucial in order to refine treatment and improve patient survival, especially in metastatic or recurrent disease stages. Thus, preclinical drug screening is a key issue in EwS research. As especially in such drug screening assays, the cell viability aspect of cell proliferation is important, resazurin colorimetry shall be reviewed here as a fast, high-throughput method with automated readout to efficiently screen for potency of drugs via measurement of cell viability.

Key words Ewing sarcoma, Cell viability, Drug screening, Resazurin

1 Introduction

The process ranging from discovery to commercialization of drugs is often lengthy, work- and cost-intensive [1]. Efficient methods, especially regarding preclinical testing of drugs, are therefore essential to optimize this process. In EwS, development of new therapeutic strategies is crucial in order to improve survival rates of patients, particularly in metastasized and recurrent disease stages [2]. When screening for effectivity of new potential drugs, assessment of cell viability is exceptionally important, as the major aim of anticancer drug treatment is to kill a cancer cell rather than merely stopping them from dividing. Methodologically, this requires a fast, high-throughput method with an automated readout for assessment of cell viability. Thus, the resazurin cell viability assay shall be reviewed here as a suitable method for such drug screening purposes. This method belongs to the “indirect” methods for proliferation assessment as measurements depend on the metabolic activity of the cells [3, 4]. Viable cells reduce resazurin in an NADH- or NADPH-dependent manner to resorufin, which is accompanied by a color change (and fluorescence emission) that can be photometrically detected [5, 6]. As the production of the electron donors NADH and NADPH is dependent on the metabolic activity of the cells, only viable cells are able to induce such color change

[5, 6]. However, resazurin colorimetry can also be used for proliferation assessment in other contexts but harbors limitations for this purpose, as only assessment of viable cells is possible and measurements can be altered by changes of cell metabolism that do not necessarily reflect viable cell numbers [3]. Since in initial drug screening assays, cell viability is a desirable experimental readout, such limitations of this assay can be accepted in favor of time and cost efficiency. After identification of promising candidate drugs, further evaluation needs to follow, either by in vitro 3D culture drug screens or in vivo models [1]. Resazurin-based assays for evaluation of drug effectivity and potency have been frequently used in EwS research [7–11].

2 Materials

1. Growth media: Appropriate medium for each cell line (500 ml) supplemented with stable glutamine, fetal bovine serum (FBS, 50 ml), and penicillin (100 U/ml)–streptomycin (100 µg/ml) (5 ml). For most commonly used Ewing sarcoma cell lines RPMI medium is appropriate.
2. T25 or T75 flasks.
3. Sterile 96-well plates.
4. Phosphate buffered saline (PBS).
5. Trypsin or Accutase.
6. Reaction tube (15 ml, 50 ml).
7. Aluminum foil.
8. 0.2 µm filter.
9. Resazurin stock solution: 1 g/l Resazurin sodium salt in PBS (50 ml). Mix thoroughly and filter-sterilize it using a 0.2 µm filter. Store it in a sterile 50 ml reaction tube covered with aluminum foil at –20 °C for long-term storage, or 4 °C for short-term storage.
10. Resazurin working solution: 1:10 dilution of the Resazurin stock solution in PBS. Prepare right before use. The remaining solution can be stored at 4 °C for up to a week.
11. Trypan blue (0.4%, commercially available already in solution).
12. Hemocytometer (Neubauer chamber, preferably standardized single-use hemocytometer).
13. Hand counter.
14. Centrifuge.
15. 200 µl sterile pipette tips.
16. Multidispenser pipette (suitable for 200 µl pipette tips).

17. Sterile or autoclaved reservoirs.
18. Microplate reader equipped with a 560 nm excitation–590 nm emission filter set.
19. GraphPad Prism software (<https://www.graphpad.com/scientific-software/prism/>).

3 Methods

All methods must be carried out under sterile conditions, which by default implies working under a S1 cell culture hood. However, when attempting a drug screening, special attention must be paid to hazardous compounds potentially exerting severe toxicity that must be handled under S2 cell culture hoods with appropriate filters. Additionally, residues and materials that have been in contact with such substances must be appropriately discarded according to specific waste regulations applying to the respective institute.

3.1 Preparation and Seeding of Cells

1. At the beginning of the procedure, cells should grow at 70–80% confluency in a T75 or T25 flask (*see Note 1*).
2. Prepare the cells for harvesting by aspirating the medium, washing the cells with 5 ml of PBS and incubating them at 37 °C for approximately 5 min with 2–3 ml of trypsin (alternatively, Accutase may be used at this step instead of trypsin, whereby in this case, cells should be incubated at room temperature, *see Note 2*).
3. Harvest the cells by adding 5 ml of culture medium to the flask and repeatedly flushing the bottom of the flask until all cells are collected in a 15 ml reaction tube.
4. Centrifuge at $314 \times g$ for 4 min in order to pellet the cells.
5. Aspirate the supernatant, add 5–10 ml of fresh culture medium (depending on pellet size), and flick the tube to resus. Invert multiple times.
6. Count the cells using trypan blue, a hemocytometer, and a hand counter (*see Note 3*).
7. Seed cells in a 96-well plate following the plate design in Fig. 1a by using a sterile reservoir and a multidispenser pipette with 200 μ l tips (*see Note 4*): the outer wells should be filled with 100 μ l of culture medium without cells (*see Note 5*), the inner wells should contain 90 μ l (*see Note 6*) of culture medium with 3000 cells per well (*see Note 7*).
8. Allow the cells to attach to the plate by placing it in a 37 °C incubator overnight (*see Note 8*).

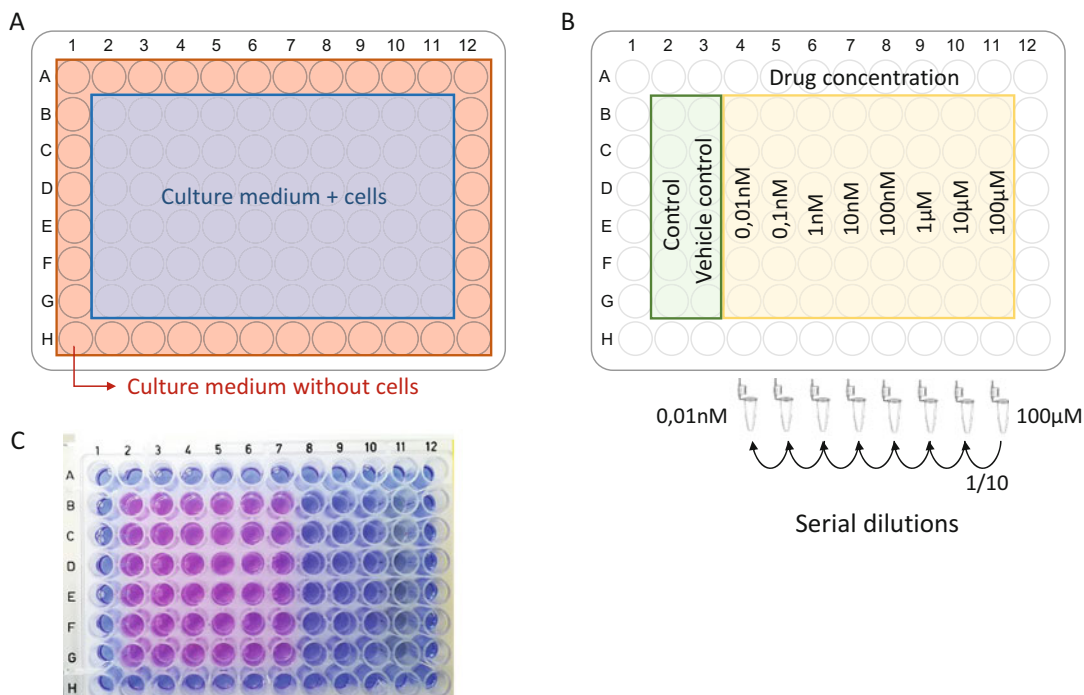


Fig. 1 Setup of resazurin colorimetry using Ewing sarcoma cells. (a) The scheme shows how to organize the wells containing cells (blue area) and the wells containing only medium (red area). This setup allows for six technical replicates of each condition per column. (b) The scheme illustrates a suggested plate setup and the preparation of serial dilutions of any compound. Green wells represent control conditions and yellow wells represent incrementing doses of the compound being tested. (c) The plate shows A673 Ewing sarcoma cells treated with serial dilutions of a toxic compound. Once the resazurin has been incubated for 3 h, the extent of pink colouring indicates the amount of viable cells in the respective wells

3.2 Stimulation with Compounds

1. Check the cells the following day under a microscope for proper attachment and to detect potential mistakes in plating (*see Note 9*).
2. Dissolve the drugs to be tested in the adequate vehicle (i.e., DMSO or water) and prepare serial dilutions (*see Note 10*, Fig. 1b).
3. For a drug screening approach, plate the desired compounds in 10 μl volume per well (the final volume per well will be 100 μl) using a 10 μl pipette in incremental concentrations as depicted in Fig. 1b (*see Note 11*). It is essential to leave untreated wells as controls (no-drug control) and also to keep some wells to test the potential effect of the vehicle in which each compound has been resuspended (vehicle control).
4. Incubate the cells for the desired amount of time to observe a potential drug effect. Normally 48–72 h is a proper amount of time for an initial analysis. During this incubation time, check

the plate daily under the microscope to evaluate a drug effect by eye and to rule out potential bacterial contaminations (**Note 12**).

3.3 Resazurin Assay

1. Check the plate under the microscope.
2. Add 20 μ l of the Resazurin working solution (*see Note 13*) to each well using a multidispenser pipette and a sterile reservoir container (*see Note 14*).
3. Incubate 1–5 h at 37 °C (*see Notes 15 and 16*). The metabolism of resazurin into resorufin by viable cells creates a color change from purple to pink in the wells (Fig. 1c). The intensity of pink fluorescence (until saturation) and the speed of color change is positively correlated to the number of viable cells in the wells.
4. Record fluorescence using a microplate reader with a 560 nm excitation–590 nm emission filter set (*see Note 17*).

3.4 Analysis

1. In order to yield acceptable results, the recorded fluorescence values should range from 600 to 800 rlu (*see Note 18*).
2. Wells without seeded cells are used for background subtraction.
3. Results can be depicted as the amount of viable cells in treatment conditions relative to the vehicle control condition.
4. If necessary, IC₅₀ calculations can be obtained by using Graph-Pad Prism (or any other suitable software) and adjusting a logarithmic inhibitor vs. response curve to the amount of viable cells in treatment conditions relative to the vehicle control condition.

4 Notes

1. Prior to the beginning of any experiment it is always recommended to test cells for potential mycoplasma contamination that may obscure the results of the assay. Additionally, cell lines should be routinely authenticated (i.e., using STR-analysis) to confirm their identity.
2. The majority of Ewing sarcoma cell lines can be easily detached by using trypsin. In case of difficulties or if a more gentle way of detaching is required, Accutase may be attempted.
3. It is recommended to use exclusion dyes such as trypan blue to determine the viability of the resuspended cells and thereby to avoid using suspensions that have a high proportion of dead cells.
4. Although the assay can be plated using normal pipettes (or multistep pipettes), multidispenser pipettes are recommended

since they allow for a faster, more even pipetting throughout the plate that, in turn, ensures a higher reproducibility of the results.

5. This outer circle of wells filled with medium builds an outer “wall” in which medium preferentially evaporates, which protects the wells containing the cells from evaporation of medium.
6. Final volume of each well will be 100 μl . We recommend using 90 μl to plate the cells and 10 μl to add any compound on top. These volumes are large enough to allow for low error and high reproducibility in pipetting.
7. The number of cells per experienced can be modified depending on the final incubation time. We experienced that seeding 3000 cells per well works for most Ewing sarcoma cell lines for an incubation time of 72 h. Shorter incubation times (i.e., 24 h) can work with 5000–7000 cells per well. For longer incubation times (i.e., 96 h), 1000 cells per well are recommended.
8. Some EwS cell lines grow in suspension (i.e., EW3 or EW24) and thus this step may be ignored in such case.
9. It is very important to early detect potential inaccuracies in pipetting (for example wells that contain more cells than others). If such inaccuracy only randomly affects a few wells, they can be properly marked and omitted from the analysis. If a more general inaccuracy is detected it will be more beneficial to re-start the protocol and plate a new batch of cells.
10. Serial dilutions are an appropriate approach when analyzing for instance IC50s, which is a very frequent first exploratory step when dealing with a drug screening. It is very important to consider that although the concentration of each tested drug will be diluted, the final vehicle concentration (e.g., DMSO) must be equal in every dilution. This can be achieved by adding additional DMSO to the mix in each Eppendorf tube while preparing each serial dilution.
11. When adding the 10 μl on top of each well, a new pipette tip must be used each time to avoid the risk of cross contamination between wells (and the stock drug solution).
12. If a bacterial contamination is pronounced enough it can be detected *de visu* by the change of medium color from pink to yellow. However, wells should be inspected under the microscope to completely rule out a contamination before adding resazurin. This is especially important since it can heavily impact on the colorimetric measurements due to bacterial metabolism affecting the resazurin turnover. If a general bacterial contamination is detected, the experiment has to be restarted.

13. The resazurin solution is deep blue colored and stains the supernatant of the wells it was applied to (*see* Fig. 1c). If the cells have been plated in a different final volume, an adjusted volume of resazurin working solution must be applied. Alternatively, commercial kits containing resazurin are available from several companies. They are very easy to use and results are highly reproducible. The main disadvantage is that they tend to be more expensive than buying resazurin sodium salt and preparing the reagents. In our hands, the protocol presented here works properly for Ewing sarcoma cell lines, which argues against the use of such kits.
14. A normal pipette can also be used at this step. However, the task is very time consuming and using a normal pipette normally leads to lower reproducibility. When using a multidispenser pipette, special attention must be paid to avoid forming bubbles that may disturb pipeting of correct volumes.
15. As a general approach, it is recommended to record the fluorescence for the first time after 1.5 h of incubation time. If recorded values are lower than 600–800 rlu, additional incubation time is required. Therefore, the plate must be placed in the incubator again and remeasured again in 1 h increments until the adequate range is achieved (in the case of most Ewing sarcoma cells, this should happen in maximum 5 h).
16. A general disadvantage of resazurin assays is that the resazurin substrate needs to be incubated with the cells at 37 °C for a period of time before the signal is generated. The longer the incubation time, the more likely it is that unwanted or unpredicted chemical/biological interactions between the compounds being tested, the substrate, and the cells may take place which could potentially influence the results of the assay.
17. The amount of resorufin generated (until saturation) and the speed of color change is positively correlated with the number of viable cells in the wells, which can be quantified using a microplate fluorometer equipped with a 560 nm excitation–590 nm emission filter. Additionally, resorufin can be quantified by measuring absorbance. However, this approach is normally not used because it is less sensitive than measuring fluorescence.
18. Since the supernatant of the wells changes its color as a consequence of resazurin being metabolized to resorufin by the cells (from dark purple to bright pink), it is recommended to visually inspect the plate for this color shift before placing it into the microplate reader.

Acknowledgments

J.M. acknowledges support from the Kind-Philipp Foundation and the Rudolf und Brigitte Zenner Foundation. F.C.A. acknowledges support from the Barbara und Hubertus Trettner Stiftung as well as from the German Cancer Aid in frame of the Max-Eder program (DKH-70112257 granted to T. Grünewald).

References

1. Suggitt M, Bibby MC (2005) 50 years of pre-clinical anticancer drug screening: empirical to target-driven approaches. *Clin Cancer Res* 11:971–981
2. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4:5. <https://doi.org/10.1038/s41572-018-0003-x>
3. Romar GA, Kupper TS, Divito SJ (2016) Research techniques made simple: techniques to assess cell proliferation. *J Invest Dermatol* 136:e1–e7. <https://doi.org/10.1016/j.jid.2015.11.020>
4. Czekanska EM (2011) Assessment of cell proliferation with resazurin-based fluorescent dye. *Methods Mol Biol* 740:27–32. https://doi.org/10.1007/978-1-61779-108-6_5
5. Präbst K, Engelhardt H, Ringgeler S, Hübner H (2017) Basic colorimetric proliferation assays: MTT, WST, and resazurin. *Methods Mol Biol* 1601:1–17. https://doi.org/10.1007/978-1-4939-6960-9_1
6. Chen JL, Steele TWJ, Stuckey DC (2018) Metabolic reduction of resazurin; location within the cell for cytotoxicity assays. *Biotechnol Bioeng* 115:351–358. <https://doi.org/10.1002/bit.26475>
7. Grünewald TGP, Bernard V, Gilardi-Hebenstreit P et al (2015) Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite. *Nat Genet* 47:1073–1078. <https://doi.org/10.1038/ng.3363>
8. Musa J, Cidre-Aranaz F, Aynaud M-M et al (2019) Cooperation of cancer drivers with regulatory germline variants shapes clinical outcomes. *Nat Commun* 10:4128. <https://doi.org/10.1038/s41467-019-12071-2>
9. Dallmayer M, Li J, Ohmura S et al (2019) Targeting the CALCB/RAMP1 axis inhibits growth of Ewing sarcoma. *Cell Death Dis* 10:116. <https://doi.org/10.1038/s41419-019-1372-0>
10. Cidre-Aranaz F, Grünewald TGP, Surdez D et al (2017) EWS-FLI1-mediated suppression of the RAS-antagonist Sprouty 1 (SPRY1) confers aggressiveness to Ewing sarcoma. *Oncogene* 36:766–776. <https://doi.org/10.1038/onc.2016.244>
11. Marchetto A, Ohmura S, Orth MF et al (2020) Oncogenic hijacking of a developmental transcription factor evokes vulnerability toward oxidative stress in Ewing sarcoma. *Nat Commun* 11(1):2423. <https://doi.org/10.1038/s41467-020-16244-2>



Analysis of Migration and Invasion in Ewing Sarcoma

Florencia Cidre-Aranaz

Abstract

The metastasis is a complex, well-orchestrated process, which includes migration from the primary tumor and invasion into secondary locations as main features. In Ewing sarcoma, metastasis is the main determinant of malignancy, with ~30% of patients presenting with metastatic disease at diagnosis. Therefore, analyzing migration and invasion in different experimental settings *in vitro* is key to understanding this disease. Among the variety of possible techniques to study migration, this chapter described the methods of wound healing (migration in 2D) and transwell (migration through a porous membrane in response to a given stimulus). Additionally, this chapter includes a variation of the transwell protocol that allows for the analysis of cell invasion through a gel matrix in response to stimulus.

Key words Ewing sarcoma, Migration, Invasion, Transwell, Wound healing

1 Introduction

Ewing sarcoma is the second most common sarcoma of children and young adults [1]. Approximately 30% of patients present metastatic disease at diagnosis, at which point therapy resistance often leads to high mortality [2]. Additionally, even patients without overt metastases usually present micrometastases that would only later become obvious [1].

Metastasis-formation is a complex, well-orchestrated, multistep process which involves invasion, intravasation, migration, extravasation, and tumor growth at the secondary location [3]. However, in a research setting, the study of metastasis is frequently compartmentalized due to the inherent limitations of *in vitro* and *in vivo* models. Thus, two of the main features of metastasis that are usually analyzed in Ewing sarcoma studies *in vitro* are migration and invasion. There are several protocols that can help us analyze the migratory and invasive capacities of cell lines [4]. Here two of the most commonly used assays in Ewing sarcoma research are described: wound healing assay for determining the migratory capacity in a 2D monolayer of cells, and *transwell* assay to study

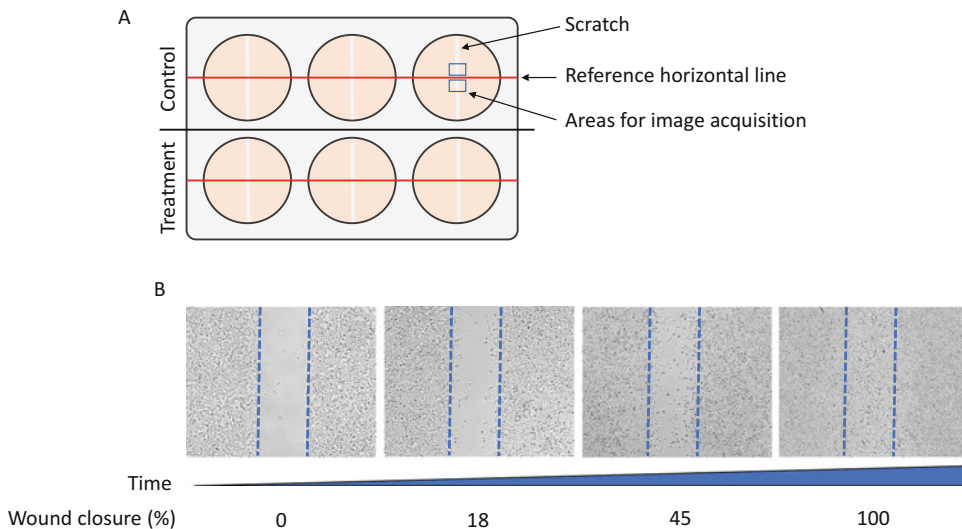


Fig. 1 Wound healing assay preparation and analysis. **(a)** Schematic plate design where a reference horizontal line is drawn perpendicular to the scratched areas (white lines) allowing for the selection of two potential areas for image acquisition (blue squares). **(b)** Example of a wound healing migration assay with the determined wound area (blue dashes) being covered by migrating TC32 Ewing sarcoma cells over time. Wound closure percentage is calculated as compared to control (picture at initial time point)

the migration in response to stimuli across a porous membrane with/without the invasion component making use of a gel layer.

The wound healing assay is based on plating cells in a monolayer at confluency followed by performing a “wound” or scratch on the area covered by the cells using a pipette tip. During the experiment, the cells at the edge of the scratch will start migrating to fill the wound, while producing new cell–cell contacts until the wound is fully closed (Fig. 1b). The evaluation of the process is performed by the periodic acquisition of images and subsequent analysis of the migration speed by using the appropriate software.

Wound healing assays have been widely used in Ewing sarcoma research as they are easy, label-free, and inexpensive, and the obtained data can be quickly analyzed [5–8]. These assays can be performed using supplies that are usually available in every cell culture laboratory such as multiwell plates and microscopes equipped with 4×–10× magnification glasses and a camera attachment. Main caveats of this technique include the precision required when manually performing the scratch in order to obtain consistency among replicates and different cell conditions, and the importance of avoiding damaging the underlying extracellular matrix, which would impede a correct migratory behavior [9]. Of note, migration assays normally need to be corrected for proliferation, especially when using cell lines with high proliferative rates.

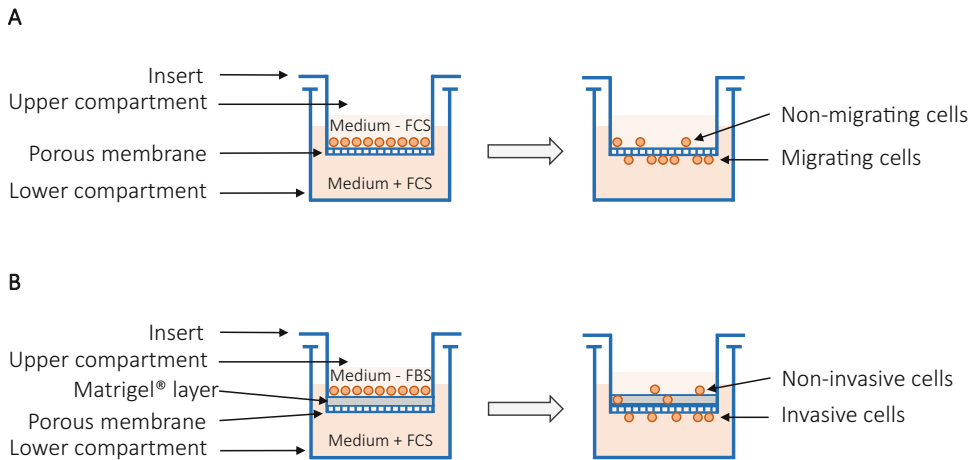


Fig. 2 Transwell assay Scheme. **(a)** Transwell migration assay where the lower compartment is filled with 10% FCS (+FCS), and the upper compartment is filled with 1% FCS (–FCS) and Ewing sarcoma cells. After incubation, cells migrate in response to serum (chemoattractant) through the porous membrane into the lower compartment. **(b)** Transwell invasion assay following the same scheme but adding a layer of gel matrix on top of the porous layer. After incubation, cells migrate in response to serum (chemoattractant) and invade the gel matrix to finally go through the porous membrane and into the lower compartment

Transwell assays are commonly used to study both migration and invasion *in vitro* in Ewing sarcoma cells [6, 10–12]. They consist of a well divided into two compartments by a porous membrane. Cells are plated in the upper compartment, normally containing a lower concentration of the selected chemoattractant, and they migrate through the porous membrane to the lower compartment, which contains a higher concentration of the chemoattractant (Fig. 2a). A variation of this technique allows for the analysis of invasion by placing a gel matrix coating on top of the porous membrane, which the cells would have to go through (“invade”) in order to migrate to the lower chamber (Fig. 2b). In any case, at the end of the experiment, migrating/invading cells are attached to the bottom part of the porous membrane, and can be subsequently fixated and stained for visualization and quantification.

Transwell assays are relatively more complex methods as compared to wound healing assays, but they allow for the analysis of the effect of chemoattractants, which is not possible otherwise.

2 Materials

2.1 Wound Healing Assay

1. Ewing sarcoma cell line.
2. Phosphate Buffer Saline (PBS): 10 mM NaH_2PO_4 , 10 mM Na_2HPO_4 , 120 mM NaCl, pH 7.4.

3. Growth media: Appropriate media for each cell line (500 ml) supplemented with stable glutamine, fetal calf serum (FCS) (50 ml), and 100 U/ml penicillin–100 µg/ml streptomycin (5 ml) (*see Note 1*).
4. Sterile pipette tips (20 µl and 200 µl).
5. 6-well cell culture plates.
6. Marker and ruler.
7. Microscope with 4× and 10× magnification, and a camera attachment.
8. Software for image analysis (i.e., ImageJ, <https://imagej.nih.gov/ij/>).

2.2 Transwell Assay

1. Ewing sarcoma cell lines.
2. Phosphate Buffer Saline (PBS): 10 mM NaH₂PO₄, 10 mM Na₂HPO₄, 120 mM NaCl, pH 7.4.
3. Growth medium (supplemented with 10% FCS): Appropriate medium for each cell line (500 ml) supplemented with stable glutamine, FCS (50 ml), and 100 U/ml PENICILLIN–100µG/ML STREPTOMYCIn (5 ml) (*see Note 1*).
4. Growth medium (supplemented with 1% FCS): Appropriate medium for each cell line (500 ml) supplemented with stable glutamine, FCS (5 ml) and 100 U/ml Penicillin/100 µg/ml Streptomycin (5 ml) (*see Note 1*).
5. 6-well plates with Transwell[®] inserts with 8 µm pores (*see Notes 2 and 3*).
6. Gel matrix (i.e., Matrigel).
7. 5 ml and 10 ml serological pipettes.
8. Crystal violet solution (0.1% crystal violet in 20% ethanol).
9. Neutrally buffered formalin.
10. Laboratory forceps.
11. Microscope and camera attachment.
12. Spectrophotometer.
13. Ice.
14. Pipettes and sterile pipette tips in all sizes.
15. Cotton swabs.
16. 96-well plates.
17. 10% acetic acid in PBS (v/v).
18. ddH₂O.
19. Sterilized container to place the transwell inserts.
20. Glass beakers.
21. 6-well plates.

3 Methods

3.1 Wound Healing Assay

3.1.1 Experimental Procedure

1. Prior to the experiment, estimate the number of cells required by calculating the number of desired replicates and determine the appropriate plate size (*see* **Notes 4** and **5**).
2. If a specific long pretreatment is required (i.e., doxycycline addition to the medium to express a specific construct), add the required compounds to the cultured cells.
3. The day of beginning of the experiment, plate the cell line of interest in a 6-well plate using the appropriate cell culture medium, accounting for at least triplicates for each condition (control and treatment; Fig. 1a).
4. Add to the culture medium any specific treatment required.
5. Using a marker and a ruler, draw a horizontal line across the bottom of each well to use as a reference during image acquisition (Fig. 1a).
6. When each well's culture is 80–90% confluent in a monolayer, using a p20 or p200 sterile pipette tip, manually perform a vertical scratch across the cell monolayer in each individual well by holding the pipette tip on a 45° angle (not perpendicular to the plate) (Fig. 1a, **Notes 6** and **7**). The scratch must be as straight as possible (*see* **Note 8**).
7. Wash each well twice with PBS in order to eliminate excessive cellular debris resulting from the scratch (*see* **Note 9**).
8. Add appropriate culture medium with/without specific treatment (*see* **Note 10**).
9. Following the intersection between the previously drawn horizontal line and the vertically performed scratch, select an appropriate area for image acquisition (either below or above the line, Fig. 1a) and take pictures corresponding to the initial time point using a microscope. Smaller magnifications are recommended (4×–10×) (*see* **Note 11**).
10. Use a marker to identify the section that will be subsequently photographed to estimate the progress of the experiment.
11. Incubate the cells with/without the appropriate treatment until full wound closure (Fig. 1b) while photographing the selected sections using a microscope with the same magnification at every predetermined time point until the end of the experiment (*see* **Note 12**).

3.1.2 Wound Healing Analysis

Once all the images have been acquired, two main parameters can be analyzed to determine the migratory capacity of the cells: the number of cells that migrated into the area where the scratch was performed (wound), and/or the area of the scratch that has been covered by the migrating cells.

1. In order to determine the number of cells that have migrated into the scratched area, assess each picture (i.e., 0 h, 6 h, 12 h, 18 h, and 24 h) individually.
2. Select 5–10 random fields in each picture that must only belong to the scratched section where cells are migrating.
3. Count the number of cells that have migrated in each field.
4. Calculate the percentage of migrating cells vs. control.
5. In order to manually calculate the area of the wound that has been covered by the migrating cells, assess each picture (i.e., 0 h, 6 h, 12 h, 18 h, and 24 h) individually (*see Note 12*).
6. Upload the initial picture (0 h) to ImageJ.
7. Determine the area of the scratch (Fig. 1b) by using the line Freehand Selection drawing tool and tracing the perimeter of the scratched area.
8. Measure the selected area by using the commands Analyze>Measure (*see Note 13*).
9. Save the calculated area.
10. Proceed with analyzing the remaining pictures in the same fashion.
11. Determine the percentage of wound closure by using the area calculated on the 0 h picture as control (Fig. 1b).

3.2 Transwell Assay

3.2.1 Migration Analysis

1. Prior to the beginning of the experiment, estimate the number of cells required by calculating the number of desired replicates (estimate 3×10^5 cells per well, at least in triplicates, per condition) and culture the Ewing sarcoma cell line in a flask in the appropriate cell medium (*see Note 1*).
2. If a specific pretreatment is required (i.e., doxycycline addition to the medium to express a specific construct), add the required compounds to the cultured cells while they are still in the flask.
3. 24 h prior to the beginning of the experiment, starve the cell from serum by replacing the medium by the appropriate culture medium at 1% FCS.
4. The day of beginning of the experiment, open the 6-well plate and remove the inserts from the wells using sterile tongs. Place the inserts in a sterilized container.
5. Add 3 ml of appropriate culture medium (supplemented with 10% FCS) to each well (*see Note 14*).
6. Place the transwell inserts again on the plate (on top of the medium, Fig. 2a).
7. Prepare a cell suspension calculating 3×10^5 cells per well, resuspended in 2 ml of culture medium (supplemented with 1% FCS) per well.

8. Add 2 ml of culture medium containing the suspension of cells to the top compartment of the transwell (*see Note 15*, Fig. 2a).
9. Incubate for 6–8 h at 37 °C (*see Note 16*).

3.2.2 Invasion Analysis

1. Prior to the beginning of the experiment, estimate the number of cells required by calculating the number of desired replicates (estimate 3×10^5 cells per well, at least in triplicates per condition) and culture the Ewing sarcoma cell line in a flask in the appropriate cell medium (*see Note 1*).
2. If a specific pretreatment is required (i.e., doxycycline addition to the medium to express a specific construct), add the required compounds to the cultured cells while they are still in the flask.
3. 24 h prior to the beginning of the experiment, starve the cell from serum by replacing the medium by the appropriate culture medium at 1% FCS.
4. The day of beginning of the experiment, thaw the gel matrix (i.e., Matrigel) on ice until a complete liquid consistency is obtained (*see Note 17*).
5. Open the transwell 6-well plate, and add 150–200 μ l of gel matrix on top of each insert (until the surface of the porous membrane is fully covered). Close the plate and place it in an incubator for 30 min until a thin gel layer is formed on top of the membrane.
6. Take the plate from the incubator and using sterile tongs, remove each of the inserts from the plate and place them in a sterile container.
7. Add 3 ml of fresh culture medium (supplemented with 10% FCS) to each well in the plate (Fig. 2b).
8. Place the transwell inserts again on the plate (on top of the medium).
9. Prepare a cell suspension calculating 3×10^5 cells per well, resuspended in 2 ml of culture medium (supplemented with 1% FCS) per well (Fig. 2b).
10. Add 2 ml of culture medium containing the suspension of cells to the top compartment of the transwell (*see Note 15*, Fig. 2b).
11. Incubate for 24–48 h at 37 °C (*see Note 16*).

3.2.3 Transwell Analysis

1. For nonadherent Ewing sarcoma cells (*see Table 1*), migrated cells will be contained in the medium in the lower chamber. Quantification of the number of cells can be performed by harvesting the medium and using a hemocytometer. Final display of results should be number of migrated cells vs. control.
2. Adherent Ewing sarcoma cells (*see Table 1*) will be attached to the other side of the porous membrane. Thus, remove each

Table 1
Ewing cell lines growth type. Selection of most widely used Ewing sarcoma cell lines and their expected growth type (adherent or suspension)

Cell line	Growth type
A673	Adherent
CHLA-10	Adherent
CHLA-25	Adherent
CHLA-258	Adherent
CHLA-32	Adherent
CHLA-99	Adherent
COG-E-352	Adherent
ES7	Adherent
EW1	Adherent
EW16	Adherent
EW17	Adherent
EW18	Suspension
EW22	Adherent
EW24	Suspension
EW3	Suspension
EW7	Adherent
LAP-35	Adherent
MHH-ES1	Adherent
MIC	Adherent
ORS	Adherent
POE	Adherent
RDES	Adherent
RH1	Adherent
SB-KMS-KS1	Adherent
SK-ES1	Adherent
SK-N-MC	Adherent
SK-N-PLI	Adherent
SK-PN-DW	Adherent
STA-ET1	Suspension
TC-106	Adherent
TC-205	Adherent
TC-32	Adherent
TC-71	Adherent

insert from the 6-well plate with a pair of tongs and submerge it once in a glass beaker containing PBS to remove medium and debris.

3. Place each transwell in a new 6-well plate and add 2 ml of neutrally buffered formalin (until the membranes are fully covered). Incubate for 5 min.
4. Using a cotton swab per well, carefully remove the cells still present on the top layer of the porous membrane (nonmigrating cells, Fig. 2a, b).
5. Place each transwell in a new 6-well plate and add 2 ml of Crystal violet solution. Stain for 10 min.
6. Remove each transwell with a pair of tongs and carefully submerge it several times in a glass beaker full of ddH₂O (*see Note 18*). Allow the membrane to dry.
7. Use an inverted microscope with a camera attachment to visualize the porous membrane and to select 5–10 random fields.
8. Take pictures and count the number of migrated stained cells (*see Notes 19 and 20*).
9. Place each transwell membrane in a new 6-well plate and add 0.5 ml of acetic acid to each well in order to dissolve the crystal violet staining. Shake the plate until the crystal violet is completely dissolved.
10. Transfer 200 μ l of the solution to a 96 well plate and read the absorbance at 595 nm on a plate reader (*see Note 21*).
11. Calculate the relative absorbance of treatment vs. control groups.

4 Notes

1. For most commonly used Ewing sarcoma cell lines RPMI media is appropriate. However, always check the information from the repository to confirm which medium to use.
2. Other pore sizes are available in the market if required due to differences in cell size. For Ewing sarcoma cell lines, 8 μ m size works well. Different experimental assays, that is, transport studies or coculture analyses can also be performed with transwells, but they require from smaller pore sizes.
3. It is possible to use 12-well plates or 24-well plates by appropriately decreasing the volumes of media and number of plated cells per well. However, bigger plates always provide more surface, which helps with selecting better images for publication.

4. For wound healing assays, it is recommended to use 6-well plates as compared to smaller plates, as they provide with sufficient surface length to select a section of the scratched monolayer of cells that looks very consistent. However, these assays can be equally performed in 12-well or 24-well plates when, for instance, the effect of highly priced compounds is being tested, as a way to reduce the price per experiment.
5. A very common analysis performed in Ewing sarcoma cell lines is the study of the effect of the knockdown of tumorigenic genes or reexpression of tumor suppressing genes (i.e., knock-downs of the translocations are frequently analyzed). These variations in expression of functionally relevant genes may lead to a decrease in proliferation rates upon treatment. Therefore, when calculating the number of cells to plate, it is important to take this into account to avoid obtaining wells with different degrees of confluency prior to starting the experiment.
6. The use of each different size of pipette tip depends on the expected speed of migrations and on whether the objective of the experiment is to analyze the migration of individual cells. Bigger or smaller pipette tips can be used if necessary. Additionally, and although it is not recommended, several scratches can be attempted in each well as long as there is sufficient separation between them. This approach allows for a wider range of sections in which to find the perfect spot for image acquisition.
7. The scratch must be performed firmly, but without applying excessive pressure on the bottom of the plate, as this may damage the plate's coating and obscure the results. Special care must be taken when performing a scratch on a collagen/fibronectin coated plate to avoid damaging the coating layer. In this case, only minimal pressure should be applied.
8. Manually performing the assay is the most cost-effective approach. However, there are other commercially available alternatives for approaching a wound healing assay based on preformed plastic plate inserts that come with a straight line made of plastic in the middle where cells cannot be seeded. Once the insert is removed, that line without cells becomes the scratch from our model (i.e., Ibidi system).
9. Careful medium aspiration together with dropwise addition of PBS may ensure better results when dealing with easily detachable cells such as TC71 Ewing cell line.
10. When working with cell lines with high proliferation rate, it is advisable to decrease the percentage of fetal bovine serum in order to avoid obscuring the migration effect by proliferation.

For the same reason, shorter incubation times (<24–48 h) must be considered accounting for the proliferation rate of each cell line.

11. When taking the images, it is advisable to save them in TIFF format, as many of the image analysis software in the market usually require from this format.
12. Incubation times may range from 20 to 48 h, depending on the proliferation rate of each cell line and their migration speed. The number of images acquired during that time would depend on what the aim of the experiment is. There are mainly two approaches possible:
 - (a) Two time point analysis: it is used to assess general migration by taking photographs at the initial and final time point (normally 0 h and 24–48 h). It is normally useful as a first approach to the analysis. However, the main disadvantage of this setting is the loss of potentially relevant information of the migratory status of the cells at intermediate time points.
 - (b) Several time-point analysis: it consists of the periodic acquisition of images at a predetermined increment of time (normally every 6–8 h). It allows for an improved monitoring of all the intermediate states on the process of wound closure, as well as the final selection of the best representative pictures for analysis from a wider range of options. The main disadvantage is that it is more time consuming and it requires from a picture to be taken every 6–8 h for 24–48 h (or until full closure of the wound).
13. Manual assessment of wound healing is a good alternative as a first approach to analyzing the results or when working with a few images. If several experiments need to be analyzed or a more thorough analysis is required, it is recommended to use a specific software. There are several alternatives available; however, the most widely used are TScratch [13] and ImageJ, using the specifically developed tool “MRI Wound Healing Tool available at http://dev.mri.cnrs.fr/projects/imagej-macros/wiki/Wound_Healing_Tool). Both applications provide free access and are extremely easy to use. Additionally, they provide final images where the wound is already highlighted, which are useful for publication and/or presentation purposes.
14. FCS here acts as a chemoattractant to serum-deprived cells. Although alternative chemoattractant can be used, FCS is normally chosen, at least in preliminary experiments, because it is very rich in diverse chemoattractants and it is relatively inexpensive.

15. Make sure the medium in the bottom compartment is in contact with the porous membrane so a chemotactic gradient is formed.
16. Incubation time may vary depending on the cell type and the chosen chemoattractant.
17. Always keep the gel matrix on ice while manipulating it, as it goes from liquid to gel consistency when left at room temperature. If you have problems when manipulating it, try freezing overnight the pipette tips that will come into contact with the gel.
18. It is usually necessary to replace the water several times. The final wash should be completely clear, with no excess staining or debris. Be extremely careful when washing the membranes, as losing part of the migrated fixed cells must be avoided at all cost.
19. Counting the number of cells on the bottom part of the membrane once the pictures are taken can be performed manually or using image analysis software such as ImageJ.
20. Direct quantification of cells can also be done by staining nuclei with DAPI instead of using crystal violet.
21. Performing both direct (cell counting) and indirect (dissolving crystal violet and subsequent absorbance measurement) quantification is not always required. If the differences observed between treatment and control conditions are obvious, it performing simply a direct quantification may be enough.

Acknowledgments

F.C.A. acknowledges support from the Barbara und Hubertus Trettner Stiftung as well as from the German Cancer Aid in frame of the Max-Eder program (DKH-70112257 granted to T. Grünewald).

References

1. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4:5
2. Bosma SE, Lancia C, Rueten-Budde AJ et al (2019) Easy-to-use clinical tool for survival estimation in Ewing sarcoma at diagnosis and after surgery. *Sci Rep* 9:11000
3. Steeg PS (2016) Targeting metastasis. *Nat Rev Cancer* 16(4):201–218
4. Ashby WJ, Zijlstra A (2012) Established and novel methods of interrogating two-dimensional cell migration. *Integr Biol (Camb)* 4(11):1338–1350
5. Chaturvedi A, Hoffman LM, Welm AL et al (2012) The EWS/FLI oncogene drives changes in cellular morphology, adhesion, and migration in Ewing sarcoma. *Genes Cancer* 3(2):102–116
6. Cidre-Aranaz F, Grünewald TG, Surdez D et al (2017) EWS-FLI1-mediated suppression of the RAS-antagonist Sprouty 1 (SPRY1) confers aggressiveness to Ewing sarcoma. *Oncogene* 36(6):766–776
7. Kamura S, Matsumoto Y, Fukushi JI et al (2010) Basic fibroblast growth factor in the bone microenvironment enhances cell motility

- and invasion of Ewing's sarcoma family of tumours by activating the FGFR1-PI3K-Rac1 pathway. *Br J Cancer* 103(3):370–381
8. Hatano M, Matsumoto Y, Fukushi J et al (2015) Cadherin-11 regulates the metastasis of Ewing sarcoma cells to bone. *Clin Exp Metastasis* 32(6):579–591
 9. Kam Y, Guess C, Estrada L et al (2008) A novel circular invasion assay mimics in vivo invasive behavior of cancer cell lines and distinguishes single-cell motility in vitro. *BMC Cancer* 8:198
 10. Bailey KM, Airik M, Krook MA et al (2016) Micro-environmental stress induces Src-dependent activation of Invadopodia and cell migration in Ewing sarcoma. *Neoplasia* 18(8):480–488
 11. Li F, Wu JT, Wang PF et al (2019) NKAP functions as an oncogene in Ewing sarcoma cells partly through the AKT signaling pathway. *Exp Ther Med* 18(4):3037–3045
 12. El Beaino M, Liu J, Wasylshen AR et al (2020) Loss of Stag2 cooperates with EWS-FLI1 to transform murine Mesenchymal stem cells. *BMC Cancer* 20(1):3
 13. Gebäck T, Schulz MM, Koumoutsakos P et al (2009) TScratch: a novel and simple software tool for automated analysis of monolayer wound healing assays. *BioTechniques* 46(4):265–274

Part IV

Animal Models in Ewing Sarcoma



Genetically Engineered Mouse Model in Ewing Sarcoma

Miwa Tanaka and Takuro Nakamura

Abstract

Modeling Ewing sarcoma is challenging, since overexpression of *EWS-FLII* induces apoptosis and is not sufficient for tumor induction. It is therefore important to obtain the cell-of-origin of Ewing sarcoma that is tolerant of *EWS-FLII* expression. Here we describe the generation of the *EWS-FLII*-expressing mouse model for Ewing sarcoma by selecting embryonic chondrogenic progenitor, eSZ cells that contain Ewing sarcoma precursors.

Key words Ewing sarcoma, EWS-FLII, Embryonic superficial zone, Mouse model, Retrovirus-mediated gene transfer

1 Introduction

Bone and soft tissue sarcomas are rare cancer consisting of 1% of all cancer. They are subdivided into more than 50 different diseases and 40% of sarcomas show gene fusions caused by chromosomal alterations [1, 2]. Sarcomas of this category are characterized by low mutational burden in genome, onset at younger age than other sarcomas, and the unclear cell-of-origin [3]. Ewing sarcoma is one of such sarcomas in which generation of animal models is quite difficult [4]. The gene fusion between *EWSR1* and one of the *ETS* family transcription factor genes such as *FLII* and *ERG* is a genetic hallmark of Ewing sarcoma; however, overexpression of *EWS-FLII* induces apoptosis and/or senescence in normal cells [4–6]. Therefore, in order to obtain effective transformation, it is mandatory to select a cell with a background that is permissive to *EWS-FLII* expression. Thus, identification of the cell-of-origin is necessary to generate a reliable mouse model for Ewing sarcoma.

ERG, one of the fusion partner genes for *EWSR1* in Ewing sarcoma, is transiently expressed in chondrogenic progenitors of the perinatal period [7, 8], suggesting that the *ERG*-expressing cellular lineage may accept *EWS-ETS* expression to block further differentiation. Based on this hypothesis here we describe how to

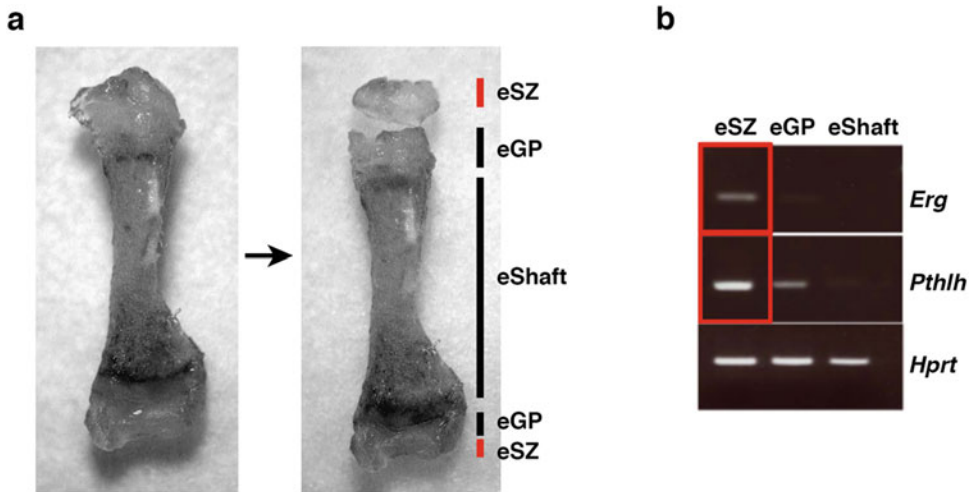


Fig. 1 Preparation of eSZ cells. **(a)** Femur is separated into two eSZ, two eGP and a eShaft fractions. **(b)** *Erg* and *Pthlh* expression in eSZ cells. *Erg* is expressed exclusively in eSZ. *Hprt* levels are shown as controls

purify chondrogenic progenitor cells and embryonic superficial zone (eSZ) cells from mouse embryos, introduce EWS-FLI1 or EWS-ERG, and transplant the transduced cells into nude or Balb/c mice (Fig. 1). This method efficiently induces Ewing sarcoma-like small round cell sarcoma [9]. The technique is also applicable to generate the models for other fusion gene-associated sarcomas such as alveolar soft part sarcoma and CIC-DUX4 sarcoma [10, 11]. In this chapter, we describe procedures to generate the ex vivo mouse model for Ewing sarcoma.

2 Materials

2.1 Animals

1. Pregnant BALB/c mice (*see Note 1*).
2. 6-week-old female nude mice or BALB/c mice (*see Note 2*).

2.2 Cell Culture

1. PLAT-E cells [12].
2. Cell culture medium: Dulbecco's modified Eagle's medium (DMEM) with high glucose supplemented with 10% fetal bovine serum (FBS) and penicillin–streptomycin.
3. Lipofectamine 2000 (Invitrogen).
4. Opti-MEM (Gibco).
5. 20 mg/ml polybrene in PBS.
6. Collagenase Type I: Prepare 100 mg/ml solution in sterile PBS right before use.
7. Cell Strainer (70 μm , Falcon).
8. Millex-HV (pore size 0.22 μm).

9. Trypsin–EDTA solution.
10. 10 mM Z-VAD-FMK in DMSO. Stock it at -20°C .
11. Matrigel Matrix HC Growth Factor Reduced (Corning): Mix 15 ml of the matrix and 5 ml of Opti-MEM at 4°C .

2.3 Plasmids

Entire coding sequences of human *EWS-FLII* and *EWS-ERG* cDNAs were generated from total RNA extracted from Ewing sarcoma clinical samples. cDNAs were FLAG-tagged and cloned into the retroviral vector pMY-IRES-GFP [7].

2.4 Antibodies

1. Mouse monoclonal anti-FLAG M2 (Sigma).
2. Rabbit polyclonal anti-PTHLH (Abcam).
3. Magnetic beads-coupled, biotinylated goat polyclonal anti-rabbit IgG.
4. CELLlection biotin binder kit (Invitrogen).

3 Methods

3.1 Preparation of Mouse Embryonic Superficial Zone (eSZ) Cells

1. At 18.5 dpc, euthanize the pregnant BALB/c mice using approved procedures (*see Note 3*).
2. Swab the entire body surface with 70% ethanol.
3. Remove embryos from the uterus using sterile forceps and scissors to a sterile petri dish.
4. Snip the placenta and the membrane, and cut limbs to obtain joints.
5. Remove skin and soft tissue around the long bones (femur, humerus, tibia and radius).
6. Collect eSZ with a pair of sterile knives under a stereomicroscope from 6 to 8 embryos (Fig. 1) and mince them into small pieces (*see Note 4*).
7. Incubate eSZ with collagenase type I (final concentration of 2 mg/ml) in PBS at 37°C for 30 min to 2 h. Dissociate the fragments by pipetting every 15 min (*see Note 5*).
8. Add 5 ml of DMEM with 10% FBS and pour the cells into Cell Strainer.
9. Centrifuge samples at $400 \times g$ at room temperature for 10 min.
10. Discard supernatant and resuspend the cells in 5 ml of DMEM with 10% FBS. Repeat centrifugation as above.
11. Discard supernatant and add 5 ml of PBS.
12. Count the cell number in a hemocytometer, centrifuge samples and discard supernatant (*see Note 6*).

3.2 Retroviral Infection

1. Day 1: Seed 2×10^6 PLAT-E cells onto a 10 cm dish in DMEM with 10% FBS without antibiotics (*see Note 7*).
2. Day 2: Transfect plasmid DNAs into PLAT-E cells using Lipofectamine 2000 according to the manufacturer's protocol.
3. Day 3: Replace the culture medium with fresh DMEM with 10% FBS with antibiotics.
4. Day 4: Collect the culture supernatant of PLAT-E cells and filtrate it using Millex-HV. Provide another 10 ml of culture medium to PLAT-E cells and keep culture for 24 h.
5. Resuspend eSZ cells at 2.5×10^5 ml in the viral supernatant and plate them into the type I collagen-coated 6-well plate.
6. Add 2.5 μ l of polybrene solution and 2.5 μ l of Z-VAD-FMK. Check the cell viability by a microscope.
7. Spin the 6-well plates at $1400 \times g$ at room temperature for 30 min.
8. Leave the cells in the CO₂ incubator at 37 °C overnight (*see Note 8*).
9. Day 5: Replace the culture medium with new retroviral supernatant collected from the PLAT-E culture with polybrene and Z-VAD-FMK, and spin them as above (*see Note 9*).

3.3 Transplantation

1. Digest infected eSZ cells with trypsin–EDTA solution at 37 °C for 5 min and count the cell number (*see Note 10*).
2. Resuspend the cells at 1×10^6 cells in 1 ml of ice-cold PBS.
3. Centrifuge the cell suspensions at $400 \times g$ for 10 min at 4 °C.
4. Discard supernatant and resuspend the cells in 200 μ l of Matrigel mixture on ice.
5. Inject 200 μ l of cell suspensions into the recipient subcutaneously (*see Note 11*).
6. Mice should be checked at least once a week for tumor formation (*see Note 12*).

4 Notes

1. We find that there is a significant strain difference in the tumor susceptibility by EWS-FLI1. Ewing sarcoma is induced in BALB/c mice at 100% frequency, whereas no mouse develops the sarcoma in C57Bl6/J.
2. Ewing sarcoma develops in nude mice at 100% penetrance. However, there are great merits to have Ewing sarcoma in the immune competent BALB/c mice to investigate drug response, tumor microenvironment, tumor invasion, and

metastasis. Alternatively, tumors developed in nude mice can be transplanted into BALB/c mice and be used for further analyses.

3. Embryos of 18.5 dpc provide the best result for eSZ purification, however, 17.5 dpc embryos and neonates before postnatal day 1 may be used.
4. It is better to collect other fractions such as growth plate or shaft (Fig. 1) to compare tumor frequencies and latencies with those of eSZ. Embryonic and adult mesenchymal cells can be also used as negative controls.
5. The activities of collagenase vary by lot. Determine the incubation period by checking dissociation of cells by microscope. Stop the reaction immediately after dissociation is completed.
6. Purification of PTHLH-positive eSZ cells (optional): Suspend 5×10^6 eSZ cells in 400 μ l of PBS. Add 2 μ l of the anti-PTHLH antibody in a 1.5 ml Eppendorf tube and incubate the mixture for 20 min on ice. Wash eSZ cells twice with 5% FBS in PBS. Add 2 μ l of the magnetic beads-coupled, biotinylated anti-rabbit IgG and incubate the tube for 20 min on ice. Wash the cells twice with 5% FBS in PBS. Collect the PTHLH-positive eSZ fraction using the CELlection Biotin Binder Kit and an appropriate magnetic apparatus. Average yield will be $2\text{--}4 \times 10^5$ cells from 8 embryos. Avoid using a cell sorter since it induces severe damages to eSZ cells.
7. Although PLAT-E is an excellent packaging cell to produce highly efficient ecotropic virus particles, it can be replaced with other 293T-based packaging cells.
8. Use fresh viral supernatant. Freezing and melting the ecotropic virus supernatant will decrease the infection efficiency.
9. It is recommended to keep transfected PLAT-E cells for confirming EWS-FLI1 or EWES-ERG expression by western blotting (Fig. 2).

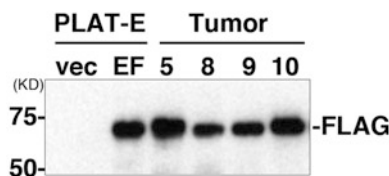


Fig. 2 Western blot analysis of EWS fusion protein expression. Expression of the FLAG-tagged EWS-FLI1 is positive in PLAT-E cells transfected with pMys-EWS-FLI1-IRES-GFP (EF) but negative in the cells with empty vector (vec). Expression of FLAG-EWS-ERG (5 and 8) and FLAG-EWS-FLI1 (9 and 10) is also shown in tumor samples

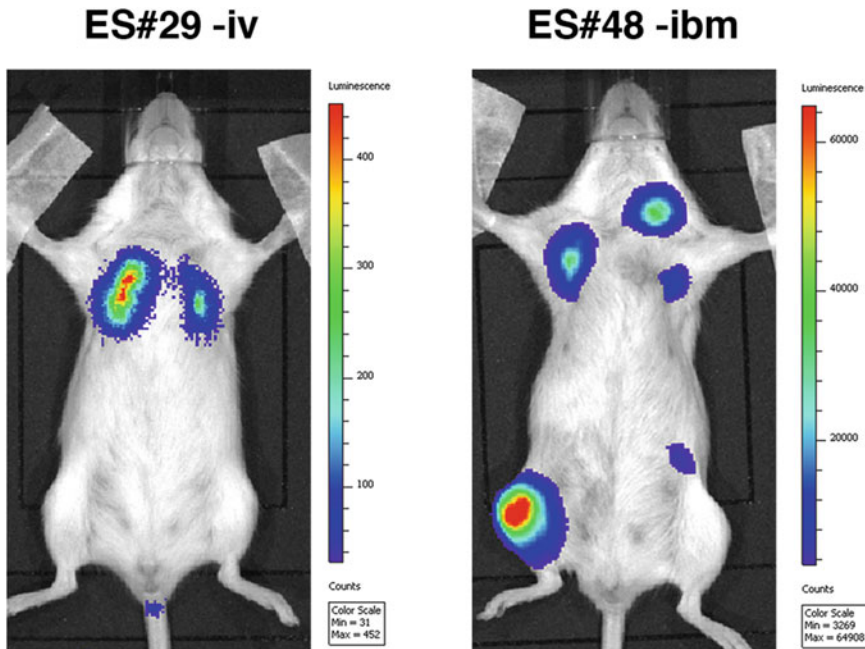


Fig. 3 Metastasis of mouse Ewing sarcoma. Metastases to bilateral lungs by intravenous injection (left), and multiple metastases to lungs and lymph nodes by intra-bone marrow transplantation (right). Sarcoma cells were introduced with an expression vector bearing luciferase. Firefly luciferin was given to tumor-bearing mice and bioluminescence images were captured using IVIS Lumina II (PerkinElmer)

10. Transduction efficiency of the retrovirus can be analyzed using small fractions of infected eSZ cells by FACS.
11. We recommend subcutaneous injection for the primary site to obtain constant tumor development. However, since subcutaneous tumors rarely metastasize, orthotopic transplantation to tibia as secondary transplantation is recommended to obtain systemic metastasis such as lungs, bone, or lymph nodes (Fig. 3). Intravenous injection of mouse Ewing sarcoma also shows efficient metastasis to lungs.
12. Subcutaneous tumors grow quickly. After tumor grows up larger than 1 cm in diameter, sacrifice the recipient and take tumor samples for pathological (Fig. 4), biochemical and cytological analyses.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research from Japan Society for the Promotion of Science (no. 26250029 to T. Nakamura), and by a Grant-in-Aid for Project for Cancer Research and Therapeutic Evolution from Japan Agency for Medical Research and Development (no. 17cmA002 to T. Nakamura).

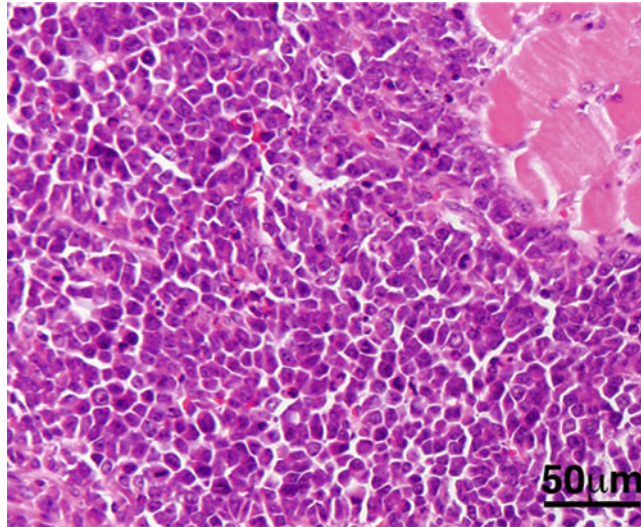


Fig. 4 Histology of mouse Ewing sarcoma. Typical small round cell morphology is shown. Scale bar = 50 μ m

References

1. Linch M, Miah AB, Thway K, Judson IR, Benson C (2014) Systemic treatment of soft-tissue sarcoma—gold standard and novel therapies. *Nat Rev Clin Oncol* 11:187–202
2. Mertens F, Tayebwa J (2014) Evolving techniques for gene fusion detection in soft tissue tumours. *Histopathology* 64:151–162
3. Genadry KC, Pietrobono S, Rota R, Linardic CM (2018) Soft tissue sarcoma cancer stem cells: an overview. *Front Oncol* 8:475
4. Minas TZ, Surdez D, Tahereh J, Tanaka M, Howarth M, Kang HJ et al (2016) Combined experience of six independent laboratories attempting to create an Ewing sarcoma mouse model. *Oncotarget* 8:34141–34163
5. Stoll G, Surdez D, Tirode F, Laud K, Barillot E, Zinovyev A et al (2013) Systems biology of Ewing sarcoma: a network model of EWS-FLI1 effect on proliferation and apoptosis. *Nucl Acids Res* 41:8853–8871
6. Tanaka M, Yamaguchi S, Yamazaki Y, Kinoshita H, Kuwahara K, Nakao K et al (2015) Somatic chromosomal translocation between *Ewsr1* and *Fli1* loci leads to dilated cardiomyopathy in a mouse model. *Sci Rep* 5:7826
7. Iwamoto M, Tamamura Y, Koyama E, Komori T, Takeshita N, Williams JA et al (2007) Transcription factor ERG and joint and articular cartilage formation during mouse limb and spine skeletogenesis. *Dev Biol* 305:40–51
8. Koyama E, Shibukawa Y, Nagayama M, Sugito H, Young B, Yuasa T et al (2008) A distinct cohort of progenitor cells participates in synovial joint and articular cartilage formation during mouse limb skeletogenesis. *Dev Biol* 316:62–73
9. Tanaka M, Yamazaki Y, Kanno Y, Igarashi K, Aisaki K, Kanno J et al (2014) Ewing's sarcoma precursors are highly enriched in embryonic osteochondrogenic progenitors. *J Clin Invest* 121:3061–3074
10. Tanaka M, Homme M, Yamazaki Y, Shimizu R, Takazawa Y, Nakamura T (2017) Modeling alveolar soft part sarcoma unveils novel mechanisms of metastasis. *Cancer Res* 77:897–904
11. Yoshimoto T, Tanaka M, Homme M, Yamazaki Y, Takazawa Y, Antonescu CR et al (2017) CIC-DUX4 induces small round cell sarcomas distinct from Ewing sarcoma. *Cancer Res* 77:2927–2937
12. Morita S, Kojima T, Kitamura T (2000) PLAT-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther* 7:1063–1066



Tumor Growth Analysis of Ewing Sarcoma Cell Lines Using Subcutaneous Xenografts in Mice

Florencia Cidre-Aranaz and Shunya Ohmura

Abstract

Subcutaneous murine xenograft models are one of the most commonly used *in vivo* experimental methods in the cancer research field. Due to the lack of appropriate animal models for Ewing sarcoma, subcutaneous murine xenograft models currently offer the simplest way to investigate antineoplastic effects of therapeutics or biological functions of target genes *in vivo*. In order to properly carry out tumor growth analysis via subcutaneous xenografts of Ewing sarcoma cells many factors should be taken into account beforehand at the planning phase of experiments. Therefore, in this chapter we describe in detail a widely used procedure for subcutaneous injection in mice, focusing on the specific characteristics of Ewing sarcoma cell lines.

Key words Subcutaneous injection, Immunocompromised mice, Xenografts, Ewing sarcoma

1 Introduction

Xenograft experiments, mostly with murine models, have been a major tool of cancer research as they allow the investigation of antineoplastic effects of therapeutics and of biological roles of oncogenes or tumor suppressor genes *in vivo*. The development of immunocompromised mice has obviously facilitated the task by enabling engraftments of human-derived cells or tissues [1, 2]. Xenograft models in general present several advantages over two-dimensional *in vitro* approaches. For instance, xenografts allow a more comprehensive analysis of experimental features such as three-dimensional biological effects, physiological environment by blood supply from the host, and hepatic or renal metabolic function in case of drug application [3].

In the case of Ewing sarcoma (EwS), xenograft models are particularly important because there seems to be no spontaneous occurrence of EwS in mice [4–6]. Additionally, the cell of origin for this tumor type is still unknown, which has so far hampered all attempts of experimentally inducing EwS by genetic engineering in

mice [6]. Thus, the use of xenograft models in the EwS research field remains inevitable.

Recently, xenograft models have advanced to include orthotopic application, which is thought to render a physiological environment relatively more similar to where the engrafted tumor cells are usually located [7–9]. However, orthotopic xenograft models are often more labor-intensive, expensive, and associated with increased stress levels for the animals due to medication, anesthesia and the performance of different invasive surgical procedures [7–9]. Thus, the pros and cons should be carefully considered in terms of potential gains in scientific knowledge while using orthotopic xenografts.

Alternatively, subcutaneous xenograft models are highly extended in the Ewing sarcoma research field because they recapitulate tumor features and allow for high reproducibility and low stress in terms of invasiveness to animals.

In this chapter, we are focusing on practical aspects of subcutaneous murine xenografts, specifically on procedures for subcutaneous injection in mice using commonly used EwS cell lines. Here we describe a detailed method for the preparation of a tumor cell suspension contained in a scaffold matrix, subcutaneous injection in mice by two operators, monitoring of tumor growth and extraction of xenografted tumor tissues.

2 Materials

1. Six to 12-week-old immunocompromised mice (NSG, NOD-*s-cid* IL2Rgamma^{null}, The Jackson Laboratory).
2. EwS cell lines frequently used in *in vivo* experiments are listed in Table 1.
3. Trypsin solution: 0.05% trypsin, 0.02% ethylenediamine tetraacetic acid (EDTA), diluted in phosphate buffered saline without Ca²⁺ and Mg²⁺.
4. Phosphate buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, double-distilled water.
5. Scaffold matrix: Geltrex™ LDEV-Free Reduced Growth Factor Basement Membrane Matrix aliquoted in 250 µl (*see* Notes 1 and 2).
6. Cell culture media: Appropriate media and supplements for each cell line as determined by the provider (*see* Note 3).
7. 10% (v/v) neutral buffered formalin.
8. Sterile 150 cm² or 175 cm² cell culture flasks.
9. Sterile 15 ml and 50 ml conical centrifuge tubes.
10. 0.4% Trypan Blue solution.

Table 1
Selection of Ewing sarcoma cell lines and approximate expected experimental timeframe in subcutaneous xenografts models

Cell line	Approximate time to treatment start after injection (days)	Approximate time to end of the experiment after injection (days)
A673	7–10	20–30
TC71	5–10	15–25
TC-32	5–10	15–20
RDES	6–10	25–50
SKNMC	7–12	35–40

11. Hematocytometer.
12. 26G and 18G sterile needles.
13. 1 ml sterile syringes.
14. Electronic shaver.
15. Caliper.
16. Disposable scalpel.
17. Ophthalmic scissors and forceps.
18. Liquid nitrogen.
19. Ice.

3 Methods

All procedures should be carried out sterilely. This protocol is designed for a mouse experiment analyzing two conditions, and therefore containing 8 versus 8 animals (total mouse number = 16) (*see Note 4*). Please make sure to adjust the quantity of materials required for different experimental settings containing more or less animals.

3.1 Preparation of Tumor Cell Suspension

1. One week before injection, thaw the selected cell line and expand it to two to five 150 cm² or 175 cm² flasks (*see Note 5*). The day of injection the cells should be at 70–80% confluence.
2. One day before injection precool PBS and 1 ml pipette tips by placing them at 4 °C overnight (*see Note 6*).
3. Prepare four 250 µl aliquots of scaffold matrix in microtubes. Keep them on ice.
4. Add 6 ml PBS onto the bottom of the cell culture flasks, wash gently the cell surface, aspirate PBS. Add 3 ml trypsin solution into each flask, incubate at 37 °C for 5 min (*see Note 7*).

5. Add 9 ml prewarmed media into each cell culture flask, mix well by pipetting with trypsin solution for inactivation and wash gently the cell surface to completely detach cells. Repeat gentle pipetting against the flask bottom, so that cells are completely dispersed each other. Replace the tumor cell suspension into a 15 ml centrifuge tube. Repeat this procedure for any additional flasks.
6. Centrifuge at $300 \times g$ for 4 min.
7. Aspirate the supernatant, tap vigorously the centrifuge tube tip to loosen the cell sediment. Add 6 ml PBS to each centrifuge tube, gently resuspend the cell sediment to remove completely components from media.
8. Centrifuge at $300 \times g$ for 4 min.
9. Aspirate the supernatant, resuspend cells in 2 ml PBS for each centrifuge tube. Gather all the cell suspension in one of the centrifuge tubes.
10. Count total cell number.
11. Calculate the required number of cells (*see Note 8*).
12. Take the calculated volume of tumor cell suspension required for injection in one centrifuge tube. Centrifuge at $300 \times g$ for 4 min.
13. Aspirate the supernatant, resuspend with approximately 900 μ l ice-cold PBS, so that the final cell suspension amounts to 1 ml in volume.
14. Mix well 250 μ l tumor cell suspension with 250 μ l of each aliquot of scaffold matrix using precooled 1 ml pipette tips in a microtube. Keep the microtubes on ice.

3.2 Subcutaneous Injection of Tumor Cells

All procedures should be carried out sterilely (appropriate disinfecting, in laminar flow hood).

1. Cool needles and syringes by keeping them on ice in order to avoid clotting of the scaffold matrix.
2. Prepare all the equipment and tools required for injection (26G and 18G sterile needles, 1 ml sterile syringes and electronic shaver) in a laminar flow hood and disinfect them with disinfectant solutions.
3. Carefully shave the mouse fur of the flank on the injection site with an electronic shaver (*see Note 9*).
4. Attach a cooled 18G needle to a 1 ml syringe, aspirate tumor cell suspension from a microtube, and mix the tumor cells and scaffold matrix very well in the tube by gently charging up and down the syringe.

5. Replace the 18G needle to a 26G and remove bubbles in the syringe by tapping. Carefully push out the suspension so that the needle is completely charged with the suspension.
6. The assistant takes out a mouse from a cage and fix it well by clasping the neck with two fingers from one hand and the tail with the other hand so that the operator can have perfect sight of the injection site.
7. The operator pulls up the preshaved flank skin with two fingers of one hand so that the skin is lifted up from the body.
8. The operator injects 100 μ l of cell suspension subcutaneously by holding the syringe with the other hand (*see Note 10*).
9. The operator removes the needle from the skin and the assistant puts the mouse in a different cage to prevent mixing the already injected mice with the rest.
10. Repeat the procedure for all the additional mice (*see Note 11*).

3.3 Monitoring of Tumor Growth

Assessment of tumor growth is carried out by measuring two dimensions of xenografted tumors (length and width) with a caliper at least every 2 days. One of the most common formulae for the three-dimensional volume approximation is given by $(\text{length}) \times (\text{width})^2/2$. The time point for starting treatment is normally not strictly defined, but empirically it should start when tumors are palpable (*see Notes 12 and 13*). At this point mice should be also randomized (*see Note 14*).

Once the defined experimental endpoint is achieved (e.g., average tumor diameter of 1.5 cm or tumor volume at 1500 cm³ as calculated with the formula), animals should be appropriately euthanized (*see Note 15*), according to experiment termination criteria defined by the animal ethics committee.

3.4 Extraction of Xenografted Tumor Tissue

Xenografted tumors should be extracted immediately after euthanization in order to avoid degradation of tumor tissues. After extraction, tumor tissues are further trimmed by removing mouse-derived tissues (Fig. 1). A small portion of the extracted tumor tissue can be cut in half, placed into two microtubes and snap-frozen in liquid nitrogen. One tissue piece serves for DNA or RNA extraction, the other can be used for protein extraction. The remaining tumor tissue can be fixated in neutral buffered formalin for histological analysis (*see Note 16*).

4 Notes

1. Subcutaneous injection of tumor cells suspended in PBS is the simplest variant, which requires no extra cost. However, this method may be associated with a risk of tumor cells diffusing

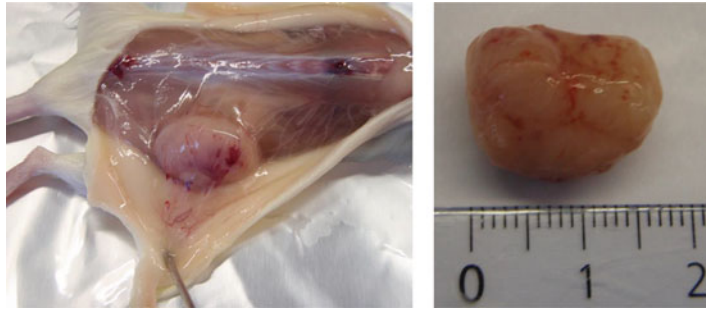


Fig. 1 Xenograft of subcutaneously injected TC32 in a NSG mouse. Three million TC32 EwS cells were subcutaneously injected on the right flank. Two weeks after injection the mouse was sacrificed by cervical dislocation. Left: Visualization of developed xenograft tumor tissue after removing the dorsal right part of the skin. Right: Resected and trimmed xenograft tumor tissue

through the subcutaneous space or even intraperitoneally if the peritoneum is injured during injection. With the aid of a scaffold matrix this risk can be reduced, since the matrix serves as an aggregator for tumor cells and prevents them from diffusing in the subcutis.

2. Scaffold matrix is mixed with the same volume of tumor cells suspended in PBS at the end, yielding doubled volume of the scaffold matrix aliquot. Normally, the subcutaneously injectable volume is 10 ml/kg body weight for mouse, corresponding 200–300 μ l for 20–30 g mouse body weight. However, since scaffold matrixes tend to be relatively expensive, we recommend a final volume of 100 μ l (50 μ l of scaffold matrix + 50 μ l of PBS/cells per mouse). In any case, the volume of the aliquots can be adjusted according to the desired final volume to be injected.
3. Most EwS cell lines can be cultured with Roswell Park Memorial Institute (RPMI) 1640 with stable glutamine supplemented with 100 U/ml penicillin/streptomycin and 10% (v/v) fetal bovine serum (FBS).
4. Since certain material is always lost during preparation or injection due to the viscosity of the scaffold matrix, we suggest to account for an increment of 20% in the final prepared injection volume. As an example, in this protocol we assume an experimental design with 16 mice. Therefore, we prepare a final volume of 2 ml, accounting for 20 times injection of 100 μ l per mouse.
5. The size and number of flasks needed for tumor cell expansion may range from 150 to 175 cm^2 and from two to five, respectively, depending on cell lines and cell numbers to be injected.

It is recommended to assess cell numbers at around 70–80% confluence in order to estimate how many flasks will be needed.

6. Since scaffold matrix is prone to solidify when maintained at room temperature, all tools which may come into contact with the scaffold matrix during the procedure should be cooled beforehand.
7. Incubation time for trypsinizing varies depending on cell lines and also how long cells are cultured in the flask. In general, most EwS cell lines can be readily detached by 5 min trypsinization.
8. In general, 2.5 to 5 million cells should be injected per mouse. The higher the injected cell number, the more quickly tumors will be formed. Note that higher cell numbers will increase the viscosity of tumor cell suspension.
9. Shaving of mouse fur makes injection of tumor cells and the subsequent measurement of tumors easier through direct observation of the skin. However, if an operator is sufficiently experienced, this procedure is not necessarily required. Alternatively, athymic nude mice can be used for the same purpose as their lack of hair facilitates the process.
10. It is extremely important to confirm that the needle edge is actually placed subcutaneously before injecting the tumor suspensions in order to avoid extra puncturing of the skin, leaking of tumor suspensions or misinjection into the peritoneum. To test the right positioning of the needle, leave the hold of the skin after inserting needle into the section of skin that was lifted by the two fingers. Then carefully flip the needle edge upward. If the needle is correctly positioned, a clear skin “ridge” becomes obvious covered by the skin, whereas any sharp edge should not be visible.
11. Once the injection of tumor cells finished, it is advisable to place the remaining cell suspension back in culture in order to make sure that the injected cells were alive at the moment of injection. This is especially important during the first trials of this protocol, and until the whole process is automatized by the experimenter.
12. In order to assess biological effects by a given treatment, tumor cells should be given sufficient time to be successfully established subcutaneously. Estimated durations for treatment start and experiment end when using EwS cell lines are summarized in Table 1.

13. Besides assessment of antineoplastic activity of drugs, one of major experimental methods used frequently in the EwS research field is induction of knockdown or overexpression of target genes to investigate their biological functions. Doxycycline-inducible target gene knockdown or overexpression plasmid constructs can be lentivirally transduced in EwS cell lines and induced *in vivo* by adding doxycycline in drinking water (ad libitum uptake). Due to the bitter taste of doxycycline, mice may hesitate to drink water. In this case, sucrose can be mixed to mitigate the bitter taste of doxycycline. It is recommended to assess the water consumption in the treatment and control group. In case that the water consumption between the two groups varies significantly, the sucrose concentration should be adjusted so that both the groups take similar water volume.

As an example 20 mg/ml doxycycline supplemented with 5% sucrose for the treatment group and 1.75% sucrose for the control group can be prepared as followed:

Preparation 10× doxycycline

- (a) Add 10.3 g Beladox (Beladox[®] 500 mg/g powder, belapharm) and 150 g sucrose in 500 ml sterile bottle.
- (b) Add sterilized water up to 300 ml.

Preparation 10× sucrose

- (a) Add 35 g sucrose in 500 ml sterile bottle.
 - (b) Add sterilized water up to 200 ml.
14. Since tumor cells often grow unevenly among individuals, mice should be evenly randomized between the treatment and control group at the moment of starting treatment according to the tumor size.
 15. At the experiment endpoint animals should be killed pain- and stress-free according to the regulation of animal experiment termination defined by the animal ethics committee. Cervical dislocation is particularly suitable for small experimental animals and may be most commonly used, since no additional equipment is required. Alternatively, overdose from inhalation anesthetics or injectable compounds may be used.
 16. Formalin-fixed paraffin-embedded xenograft tumor tissues can be conventionally sectioned and stained by hematoxylin and eosin staining, which enables microscopic evaluation of tumor morphology, necrosis, or apoptosis. For more detailed analyses, immunohistochemistry (IHC) can be employed. Proliferative cell fractions can be assessed by IHC with a Ki67 antibody, apoptotic cells with an activated caspase 3 antibody and EwS cells with a CD99 antibody, albeit CD99 is not entirely specific to EwS cells.

Acknowledgments

S.O is supported by the Deutsche Forschungsgemeinschaft (DFG 391665916). F.C.A. acknowledges support from the Barbara und Hubertus Trettner Stiftung as well as from the German Cancer Aid in frame of the Max-Eder program (DKH-70112257 granted to T. Grünewald).

References

1. Kerbel RS (2003) Human tumor xenografts as predictive preclinical models for anticancer drug activity in humans: better than commonly perceived-but they can be improved. *Cancer Biol Ther* 2:S134–S139
2. Kelland LR (2004) Of mice and men: values and liabilities of the athymic nude mouse model in anticancer drug development. *Eur J Cancer* 40:827–836
3. Ding Y, Cravero JD, Adrian K et al (2010) Modeling pancreatic cancer in vivo: from xenograft and carcinogen-induced systems to genetically engineered mice. *Pancreas* 39:283–292
4. Nielsen SW (1976) Comparative pathology of bone tumors in animals, with particular emphasis on the dog. In: Grundmann E (ed) *Malignant bone tumors. Recent results in cancer research/Fortschritte der Krebsforschung/Progrès dans les recherches Sur le cancer*, vol 54. Springer, Berlin, Heidelberg, pp 3–16
5. Sharkey FE, Fogh J (1979) Incidence and pathological features of spontaneous tumors in athymic nude mice. *Cancer Res* 39:833–839
6. Minas TZ, Surdez D, Javaheri T et al (2017) Combined experience of six independent laboratories attempting to create an Ewing sarcoma mouse model. *Oncotarget* 8:34141–34163
7. Mohammad RM, Al-Katib A, Pettit GR et al (1998) An orthotopic model of human pancreatic cancer in severe combined immunodeficient mice: potential application for preclinical studies. *Clin Cancer Res* 4:887–894
8. Stewart E, Goshorn R, Bradley C et al (2014) Targeting the DNA repair pathway in Ewing sarcoma. *Cell Rep* 9:829–841
9. Stewart E, Federico SM, Chen X et al (2017) Orthotopic patient-derived xenografts of paediatric solid tumours. *Nature* 549:96–100



Metastasis Assessment in Ewing Sarcoma Using Orthotopic Xenografts

Roser López-Alemaný and Oscar M. Tirado

Abstract

Orthotopic models are based on the implantation of tumor cells directly into the organ of origin, which allows interaction between the cells and the surrounding host tissues.

Here we describe a modified version of an orthotopic model that closely recapitulates the steps required for metastasis development in Ewing sarcoma: tumor cells are injected into the calf muscles of the mouse, and once the tumor reaches a certain volume, the muscles containing the tumor are surgically resected. This procedure involves a nonaggressive surgery of the muscle which allows for the survival of the mouse during a period of time that is long enough to enable the development of distant metastases. This spreading of tumor cells to metastatic sites in other organs takes place by a physiological mechanism similar to what occurs in human Ewing sarcoma.

Key words Ewing sarcoma, Metastasis detection, Orthotopic model, Xenograft model, Mouse surgery

1 Introduction

Ewing sarcoma (ES) is a bone and soft tissue sarcoma affecting mostly children and young adults. It is very aggressive and highly metastatic; approximately, one-third of ES patients present metastasis at diagnosis, being lung and bone marrow the most common sites [1]. The treatment and prognosis of patients are determined, among other factors, by the presence of metastasis at diagnosis [2]. Therefore, a full comprehensive understanding of ES metastatic process is required to help develop novel therapeutic strategies.

There is a lack of animal models that recapitulate the metastatic processes associated to Ewing sarcoma. Classic model involves intravenous tumor cell injection in the tail, which omits initial

Electronic supplementary material: The online version of this chapter (https://doi.org/10.1007/978-1-0716-1020-6_16) contains supplementary material, which is available to authorized users.

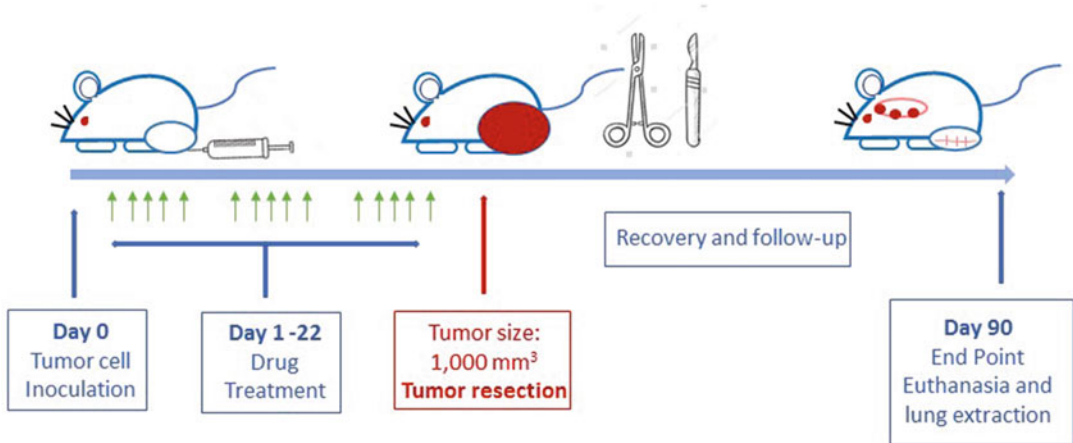


Fig. 1 Scheme of the main steps of the whole xenograft orthotopic metastatic assay. After cell inoculation on the gastrocnemius muscle, the tumor is left to grow until a critical point, when the tumor-containing muscle is surgically resected. The operated mice survive for a period long enough to develop distant metastasis

steps of metastasis, such as local invasion and intravasation [3]. Other models include severe surgery as hind limb amputation [4]. Here we propose a new orthotopic xenograft *in vivo* model that recapitulates the metastasis steps in Ewing sarcoma patients. Human Ewing sarcoma cells are injected intramuscularly in the gastrocnemius. Once the primary tumor has grown into the muscle, it is resected by a minor surgery that allows long time mice survival and the formation of distant metastasis (Fig. 1). Luciferin labeling of the tumor cells enables to follow the primary tumor development and the appearance of metastasis by a noninvasive method.

Thus, we present a model of spontaneous distant ES metastasis that mimics the clinical development of the human disease. Moreover, it allows to study the biology of the entire metastatic cascade. It can also be considered a useful tool to analyze the effect of drugs on metastasis formation, not only in ES but in other type of sarcomas.

We have already used the described model to analyze the role of p90 ribosomal kinase [5] and ephrin receptor EphA2 [6] in the generation of metastasis in xenografts of ES cells.

2 Materials

1. Nude athymic mice: Hsd:Athymic Nude *Foxn1^{nu}*.
2. Human ES cell line, A673, transfected with luciferase expressing construct, containing a puromycin resistance cassette (gift of Dr. Ibane Abasolo, VIHR, Barcelona).
3. Complete cell culture medium: RPMI-1640 GlutaMAX supplemented with 10% heat-inactivated fetal bovine serum and

1% penicillin–streptomycin and the selection antibiotic, in this case 0.6 µg/ml puromycin.

4. Antibiotics-free medium: RPMI-1640 GlutaMAX supplemented with 10% heat-inactivated fetal bovine serum.
5. 15 mg/ml D-Luciferin firefly dissolved in PBS.
6. Surgery material: tweezers, surgical scissors, scalpels, Kemmler forceps; clean and sterilized by autoclave.
7. Anesthesia apparatus: using an 5% isoflurane/2% O₂ mixture.
8. Surgical suture: Silk black suture thread (5/0) with swaged triangular needle (3/8).
9. Tissue adhesive as Histoacryl.
10. IVIS Luminar XR system apparatus.
11. Analgesia: 5 mg/ml meloxicam and 0.3 mg/ml buprenorphine solutions.
12. 10% povidone–iodine solution.
13. 70% alcohol.
14. Sterilized gauzes, cotton swabs.
15. 4% paraformaldehyde solution.
16. Sterile phosphate buffered saline (PBS).
17. 50 ml or 15 ml conical tubes.
18. 1.5 ml eppendorfs.
19. Needles (25G).

3 Methods

Animals must be handled according to the appropriate regulation for care and use of laboratory animals. Proper allowances must be obtained and approved prior to the beginning of the animal setup.

3.1 Cell Preparation

1. ES cell line A673 luciferase-labeled is cultured in complete medium.
2. Before inoculation, cells are tested for mycoplasma contamination and Luciferase activity.
3. Two days before inoculation, culture medium is changed to antibiotic-free medium.
4. Cells are harvested at exponential growth.
5. Cells are counted using an hemocytometer to prepare the appropriate number of cells in individual eppendorfs for each mouse (2×10^6 cells resuspended in 100 µl of PBS per mouse).
6. Cells must be kept on ice until inoculation and then moved to the Animal Facility.

3.2 Cells Inoculation in the Hind Limb

No anesthesia is required for this step.

1. Clean the right hind limb of the mouse with a 70% ethanol solution.
2. Charge the previously prepared cells (2×10^6 cells resuspended in 100 μ l of PBS) into a 1 ml syringe with a 25-diameter gauge.
3. Immobilize the mouse in supine position into the surgical surface.
4. Insert the gauge where the Achilles tendon separates into two, in parallel to the bone (the Achilles tendon has a characteristic white Y-shape and is easily visible under the skin (Fig. 2)). Make a 50 mm insertion between the two gastrocnemius muscles, inoculate the cells, and carefully remove the syringe.
5. No blood should be detected after injection, but if that was the case, clean the skin with a cotton swab sopped in 70% alcohol and gently press on the injection site.
6. Immediately move the mouse to the original cage (*see Note 1*).
7. Repeat the process for the remaining mice. Normally experiments are done with cohorts of 10 mice, and mice are randomly assigned to control or treatment group.
8. A single primary tumor should develop in >95% of mice over the ensuing 3 to 4 weeks.

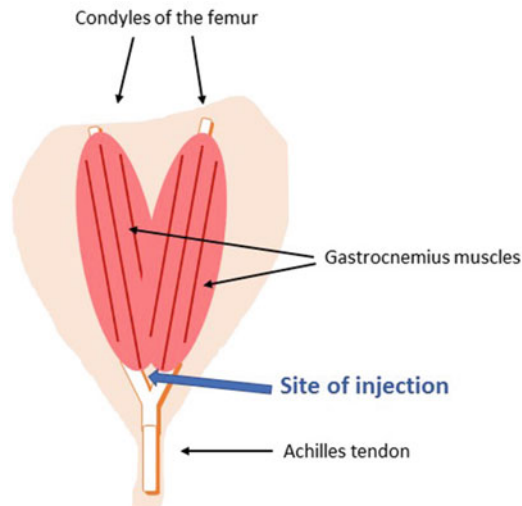


Fig. 2 Gastrocnemius muscle position in the hind limb. Tendons connecting the muscle to the bone are indicated. Blue arrow indicates the site of tumor cell inoculation

3.3 Tumor Development

After inoculation of ES cells, the general wellness and weight of each mouse is monitored twice or three times a week. Tumor growth is monitored by direct measure of the tumor-containing muscle using caliper or by IVIS imaging lecture (once a week).

3.3.1 Monitor of Tumor Growth by Caliper Measurement

1. Two diameters of the tumor are measured each 3–4 days using digital calipers.
2. Tumor volume is calculated using the formula: $(D \times d^2)/2$, where D , is the longer diameter of the hind limb and d , is the shorter diameter. Lower extremity without tumor are approximately 65 mm^3 . Figure 3a shows a typical curve of tumor growth.

3.3.2 IVIS Lecture Monitoring of Tumor Growth

1. 24 h after the inoculation of the cells, the mice are imaged using an in vivo 100 bioluminescence/optical imaging system. This monitoring should be repeated weekly.

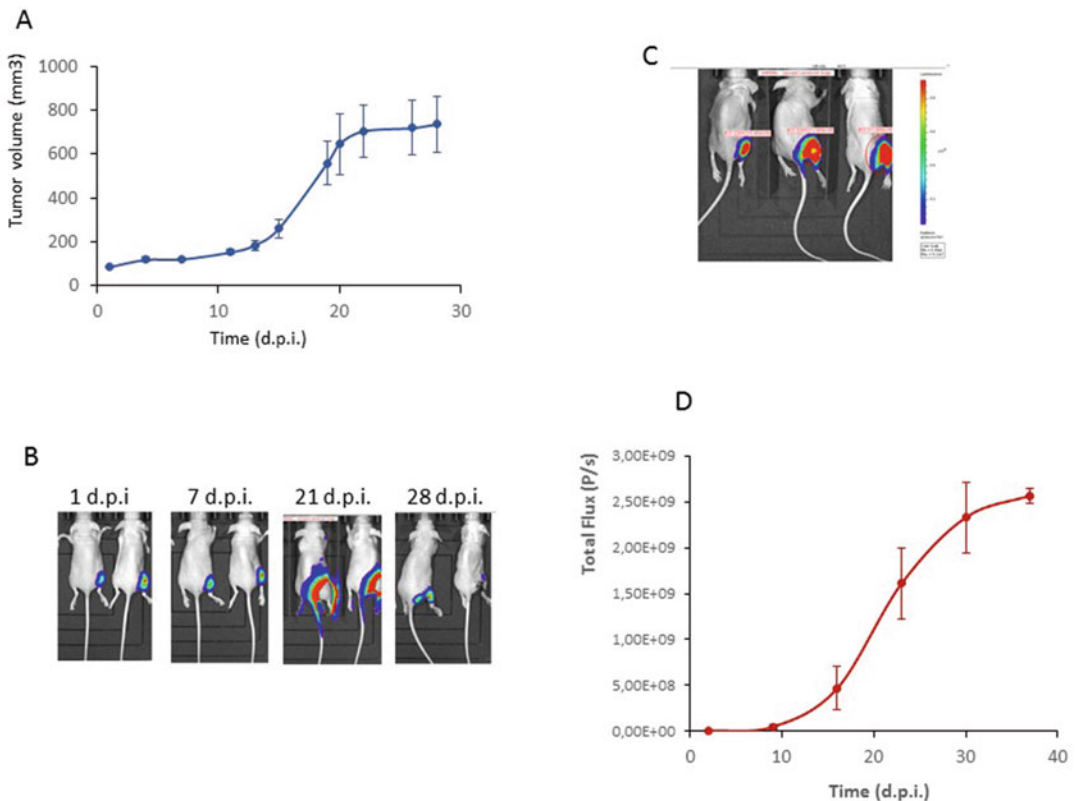


Fig. 3 (a) Tumor growth (in volume) of A673-luciferin labeled cells. A representative experiment is shown ($n = 8$) (b) Dorsal IVIS imaging pictures at different times after inoculation, where the growth of the tumor can be observed. 28 d.p.i is an image after surgery, when the luminescence has almost disappeared. (c) Example of ROIs detection and quantification. (d) Tumor growth quantification by luminescence detection of a representative experiment. d.p.i.: days postinoculation

2. D-Luciferin is intraperitoneally injected at a dose of 150 mg/kg.
3. 10 min before measuring the light emission, general anesthesia is induced with 5% isoflurane/2% O₂ mixture and continued during the procedure.
4. Once anesthetized, mice are placed on the IVIS apparatus, provided with a nose connector for continuous anesthesia.
5. Luminescent images are acquired at 60 s of exposure, and automatically superimposed by the IVIS living image software with photographic images. Dorsal and ventral images of mice are taken. Regions of interest (ROI) are automatically detected around the bodies of the mice to assess signal intensity emitted. Luminescence signal is expressed as photons per second emitted within the given ROI (Fig. 3b–d). Tumor bioluminescence in mice linearly correlates with tumor volume.

3.3.3 Drug Administration

The period in which the tumor develops after the inoculation of cancer cells is the appropriate moment to inject the drugs to be tested. Drugs can be administered by the most suitable method, depending on the compound: intraperitoneal, intravenous, intramuscular or oral gavage. Dose and pattern administration will depend on the tested drug.

3.4 Surgery

Mice are considered for surgical tumor resection when the tumor reaches the critical volume declared in the approved animal use application, normally 1000 mm³.

All the surgery procedure is carried out in a separate area in the Mouse Facility, under a laminar flow cabinet, if possible.

1. Before arranging the surgical area, clean the benchtop well with a disinfectant as Biospore 6.4%. Verify all the material is sterile or autoclaved (*see Note 2*).
2. Anesthesia: mouse is placed in an induction chamber using a 5% isoflurane/2% O₂ mixture.
3. Once anesthetized, the mouse is placed on a sterile surgical pad upside down, and a nose cone is used to give constant isoflurane/O₂ during the surgical procedure.
4. All the surgery is performed under a heating lamp, to avoid hypothermia. A lamp with a magnifier glass can be useful, but it is not essential.
5. The mouse is immobilized to the surgical pad using surgical tape: tape is placed on the back of the animal and on the right limb, ensuring that the tumor is clearly visible and accessible. Scotch the tail tip to the nose cone to allow easy access to keep the tumor-containing muscle.



Fig. 4 Immobilized mouse, showing the tumor containing hind limb, ready for surgery. Achilles tendon is visible under the skin

6. Clean the skin of the tumor area with 70% ethanol (Fig. 4).
7. Tumor extraction: using surgical scissors, carefully make an incision in the skin of the hind limb from the ankle to the knee. Open the skin around the tumor, leaving the entire gastrocnemius muscle accessible. Identify the Achilles tendon that is easily visible and introduce the forceps under the tendon and cut it carefully using surgical scissors. Keep the tendon grabbed during all the procedure, as a guide for the muscle extraction. Gently lift the gastrocnemius muscle by stretching the tendon to separate it from the limb. With the other hand, cut the membrane that wraps up the muscle. When the entire gastrocnemius containing the tumor is separated from the limb, cut the two superior tendons of the muscle.
8. Place the resected tumor into a petri dish.
9. Using scalpel, separate the remaining muscle tissue (red) from the tumor (white).
10. Weight the entire tumor and record the result.
11. Cut the tumor in two equal pieces. Submerge one of them into a 4% paraformaldehyde solution, for histologic analysis and reserve it at 4 °C. Cut the other half tumor in 2 equal pieces, for RNA and protein extraction. Conserve them into dry ice and later, stock them at a -80° freezer.
12. Clean the open limb with sterile PBS and sterile gauze. If some of the tumor cells are still remaining into the muscle, eliminate them using a surgical spatula. If some hemorrhage occurs, press carefully but steadily the wound with a sterile gauze for several minutes until the bleeding stops.

13. Suture the wound: keeping the mouse in the same position, start to sew the wound by the upper side of the limb. Use the black suture thread, equipped with a triangular needle and manipulate it using the curved Kemmler forceps. Using tweezers, keep together the skin of the two sides of the wound. Hold the needle of the suture using the Kemmler forceps. Make a double knot and cut the thread, leaving at least 0.5 cm by each side. Make the next stitch at 0.4 cm approximately of the first one. Continue all the wound long until the wound is sealed. Usually 6 or 7 stitches are enough to cover all the wound. If at some point the skin has been damaged and it is not possible to sew it, use a drop or two of a tissue adhesive. Clean all the wound and the limb with povidone–iodine solution (Fig. 5).
14. Right after the end of surgery use analgesia, by intraperitoneal injection of 200 μ l of meloxicam.
15. Recovery after surgery: place the animal in a clean cage, over a sterile gauze, under a warm red light for recovery. Control the temperature and regular breathing of the mouse. Usually the mouse awakes after 10 or 15 min. When the mouse completely recovers from the anesthesia and starts to move, place it in a regular new clean cage (*see Note 3*). Several operated animals can be caged together. Put some wet food on the bottom of the cage to ensure that the mouse can reach it. Some environmental enrichment with nesting material (cardboard rolls or tissues) can be added to the cage. There is no perioperative mortality associated with hind limb surgery.
16. Follow-up of mice at short term: during the following days, carefully control the general health state of the animal. Weight each mice the day after. Usually an important weight loss is associated with surgery (around 1.5 g), but it is recovered in few days (*see Note 4*). Use further analgesia for the following 2 days: intraperitoneal injection of 100 μ l of buprenorphine.
After 10–12 days mice should have completely recovered and lead a normal life, they should be active and be able to move, walk and reach the food (*see Video 1*). Operated mice should survive for a long period of time.

3.5 Follow Up of Mice at Long Term

After surgery, mice should survive a long time, enough to develop distant metastasis. In some cases, mice develop a recurrence tumor in the area next to the original primary tumor. If so, allow it to grow as long as the animal can use the leg and walk without difficulties, then euthanize them and proceed as described at Subheading 3.6.

1. During this period periodically control the animals (twice a week), weight the animals and identify any point of distress (*see Note 5*).

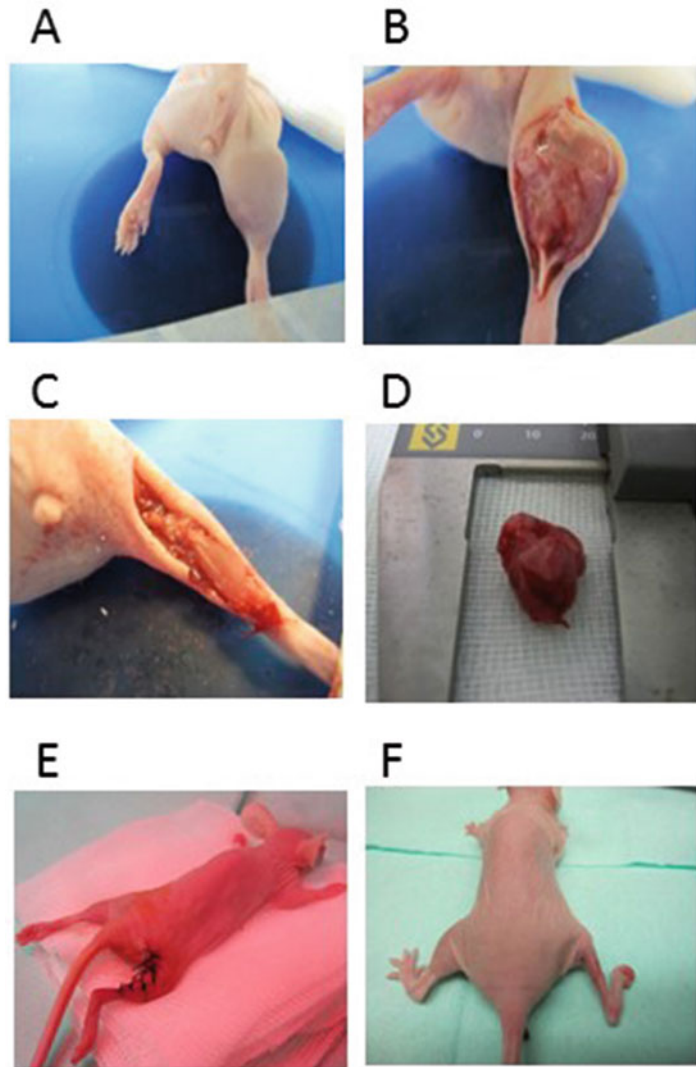


Fig. 5 Steps of surgery: (a) Immobilized mouse, where the tumor is in an accessible position; (b) opened hind limb showing the tumor, with the Achilles tendon clearly visible; (c) hind limb after tumor extraction; (d) resected muscle contain the tumor; (e) recovery after surgery and sewing; (f) 15 days after surgery, the leg completely recovered can be observed

2. Perform the IVIS lecture on weekly basis. This will inform of primary tumor recurrence or appearance of metastasis. Lung metastasis are detectable in a 38–50% of mice around day 40–50, depending on the cell line used.

3.6 End Point: Euthanization of Mice

The end point should be determined based on the cellular model used and the appearance of detectable metastasis. Usually 90 days after inoculation is a good point to let metastasis appear.

1. Before euthanasia, read IVIS of alive animals (in vivo lecture).
2. Euthanize them by cervical dislocation or using a gas chamber (carbon monoxide).
3. Lung extraction: place the animal in a surgical surface and proceed to lung extraction. Using scissors and tweezers cut the frontal side of the neck. Localize the trachea, pass a thread under it, and make a loose knot. Using surgical material, open the thoracic cage to expose the lungs. Cut the trachea over the knot and carefully pull it to completely extract the respiratory pack.
4. Place the lungs in a petri dish and proceed to a second IVIS evaluation of the lungs (ex vivo). Some metastasis are too small or are in an inaccessible position, and are not detectable in the whole animal reading, but become evident with the lungs IVIS evaluation (*see* **Notes 6** and **7**).
5. Lungs perfusion: load a 5 ml syringe with a 4% formaldehyde solution and connect it to a catheter (24 GA). Hold the trachea with tweezers and inject the catheter to it. Carefully inject the formaldehyde solution to the lungs, that will swell visibly. The volume injection will depend on the size of the lungs, but do not inject more than 2 ml of volume, otherwise the lungs will collapse. Take out the catheter of the trachea and secure firmly the knot. Conserve the lung in a 50 ml tube containing 4% formaldehyde.
6. Tissue extraction: other tissues can be extracted, to test the toxicity of the drugs used in the experiment. Usually kidney, liver, and spleen, of two animals of each of the drug treatments or controls are taken. Extract these tissues and conserve them in 4% paraformaldehyde.

3.7 Histological Analysis

Perform Eosin /Hematoxylin staining of sections of each extracted tissue: primary tumor, lungs, kidney, liver, and spleen.

3.8 Metastasis Detection

All along the experiment, presence of metastasis can be detected at three different levels:

1. In vivo detection: IVIS lecture of whole alive mice (Fig. 6a).
2. Ex vivo detection: IVIS lecture of lungs after euthanasia (Fig. 6b).
3. Histological analysis: Sections of the lung are stained with Eosin/Hematoxylin and metastasis are counted under an optical microscope (Fig. 6c).

Metastasis detected at level 1 and 2 are considered macro-metastasis. Metastasis detected at histological analysis (level 3) are considered micrometastasis.

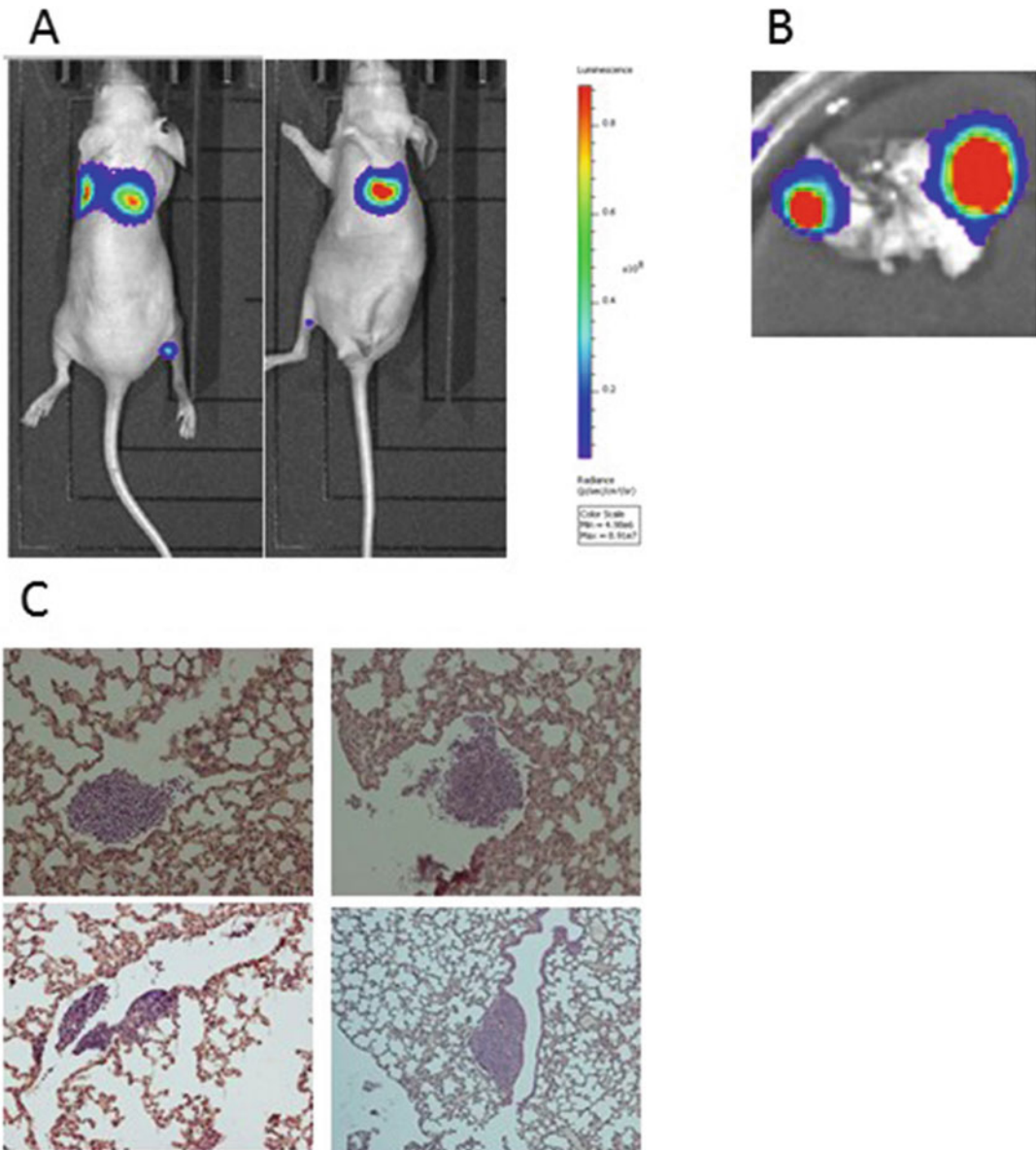


Fig. 6 Metastasis detection. **(a)** IVIS in vivo images, where two metastases are visible at the dorsal position, only one is detectable at ventral position. Some luminescence is also visible at the hind limb, remaining of the primary tumor. **(b)** Ex vivo IVIS images of lungs after euthanasia where two metastases can be detected. **(c)** Histological analysis of lungs section showing micrometastasis as groups of blue stained round cells in the normal parenchyma of the lungs

3.9 Final Analysis of the Data

This model can be used to identify differences in primary tumor growth, and specially to test differences in metastasis development. Once the experiment is finished and the histological analysis is performed, the following parameters must be considered:

1. Curves of primary tumor growth: in volume (mm^3) and light emission (P/s).
2. Time (days after inoculation) in which the surgery is performed.
3. Metastasis detection: macrometastasis and micrometastasis. Two parameters are taken into consideration: metastasis incidence (number of animals with metastasis) and number of metastasis/mouse.
4. Toxicity of the drugs tested: histological analysis of the tissues (kidney, liver, and spleen) will show indices of toxicity in the mouse.
5. Mouse weight during the treatments: Mouse weight should be stable during all the process, except for a few days after the surgery, when a loss of weight is normal. Loss of weight associated to some of the drugs will be considered as an indication of toxicity.
6. Statistical analysis: when the effect of drugs is tested on the model, compare the parameters (drugs vs. control) by current statistical analysis as Student's *t*-test or ANOVA covariance analysis.

4 Notes

1. Inoculation can be performed using patient derived xenografts (PDX), fragments of ES human tumoral tissue. PDX fragments (5×2 mm) thaw and embedded in Matrigel are used for implantation. Surgical implantation is performed under general anesthesia (as described in Subheading 3.4, **step 2**). Once the mouse is immobilized, a small incision in the hind limb skin is made at the Achilles tendon level with scissors. Carefully separate the tendon using forceps; with fine forceps create a “pocket” in the space between both gastrocnemius muscles. While holding the tendon with one hand, the PDX fragment is introduced in the “pocket” with forceps with the other hand. Once the PDX fragment is in the intermuscular space, sew the wound with two stitches (as described at Subheading 3.4, **step 13**). After implantation, clean the suture area with 70% alcohol and move the mouse to the original cage.
2. All the surgery can be conducted by a single investigator, but an assistant can be helpful, to assist with anesthesia, preparing animals and monitoring recovery, thus shortening the animal's time under anesthesia.
3. If the mouse does not recover properly after surgery, try a cardiac massage with your finger or an intraperitoneal injection of 500 μl of warm sterile PBS.

4. In some cases, some of the stitches of the wound may open. If so, try to sew them again. If it is not possible, use a drop of Histoacryl to close the wound. Most animals recover well after surgery. However, in the following days a small percentage of mice may present symptoms of gangrene on the toes. If so, these mice must be immediately euthanized to avoid possible complications.
5. During all the experiment, any sign of discomfort or disease must be taken into consideration, following the guidelines of the Animal Facility of your Institution.
6. We have never detected metastasis in other tissue than lungs, but to ensure the presence or not of metastasis in other tissues, an IVIS lecture of the whole animal after lung extraction can be performed.
7. If metastasis in the lungs are to be used for other uses than histological analysis, remove them before the lung perfusion, and conserve them at -80°C .

Acknowledgments

This work was supported by *Fundación Alba Pérez lucha contra el cáncer infantil. Instituto de Salud Carlos III (CES12/021; PI11/00038; PI15/00035; AC14/00026) and EU's Fondo Europeo de Desarrollo Regional (FEDER) "Una manera de hacer Europa/A way to achieve Europe."*

References

1. Williams RF, Fernandez-Pineda I (2016) Pediatric Sarcomas. *Surg Clin North Am* 96 (5):1107–1125
2. Windsor R, Strauss S, Seddon B, Whelan J (2019) Experimental therapies in Ewing's sarcoma. *Expert Opin Investi Drugs* 18 (2):143–159
3. Oskarsson T, Batlle E, Massagué J (2014) Metastatic stem cells: sources, niches, and vital pathways. *Cell Stem Cell* 14(3):306–321
4. Goldstein SD, Hayashi M, Albert CM, Jackson KW, Loeb DM (2015) An orthotopic xenograft model with survival hindlimb amputation allows investigation of the effect of tumor microenvironment on sarcoma metastasis. *Clin Exp Metastasis* 32:703–715
5. Lagares-Tena L, García-Monclús S, López-Alemán R, Almacellas-Rabaiget O, Huertas-Martínez J, Sáinz-Jaspeado M, Mateo-Lozano S, Rodríguez-Galindo C, Rello-Varona S, Herrero-Martín D, Tirado OM (2016) Caveolin-1 promotes Ewing sarcoma metastasis regulating MMP-9 expression through MAPK/ERK pathway. *Oncotarget* 7(35):56889–56903
6. García-Monclús S, López-Alemán R, Almacellas-Rabaiget O, Herrero-Martín D, Huertas-Martínez J, Lagares-Tena L, Alba-Pavón P, Hontecillas-Prieto L, Mora J, de Alava E, Rello-Varona S, Giangrande PH, Tirado OM (2018) EphA2 receptor is a key player in the metastatic onset of Ewing sarcoma. *Int J Cancer* 143(5):1188–1201



Orthotopic Implants in Mice

Elizabeth Stewart

Abstract

By directly implanting patient tumor cells into mice in a relevant location, we can mimic both the biology and tumor microenvironment of the original tumor. Here we describe the process of generating an orthotopic patient derived xenograft model by injecting a single cell suspension of Ewing sarcoma cells into the femur of a recipient mouse.

Key words Orthotopic xenograft, Patient-derived xenograft (PDX), Preclinical testing, Murine model, Ewing sarcoma

1 Introduction

Despite the genomic simplicity of Ewing sarcoma, efforts to develop a genetic mouse model incorporating the EWS-FLI1 fusion have been unsuccessful [1]. Researchers have relied historically on cell lines to study Ewing sarcoma; however, these approaches have limitations due to the genetic divergence that develops between a primary tumor and the corresponding cell line which has adapted to grow outside of a natural tumor environment [2, 3]. More efforts recently have been directed at generating well characterized animals models that can be used to advance our understanding of the basic biology and therapeutic vulnerabilities of these rare cancers [4–6].

One such approach is the utilization of orthotopic patient-derived xenograft models (PDX). One major advantage of PDX models of pediatric cancer is that the tumor cells tend to more consistently recapitulate the features of the original tumor, and the orthotopic implantation technique may better simulate the natural tumor environment [7]. These preclinical models are an essential tool for providing scientific justification and translational relevance for new therapeutic combinations in Ewing sarcoma and other pediatric cancers [7].

Here we describe the process of dissociating a tumor into a single cell suspension for orthotopic implantation into the femur of a mouse.

2 Materials

To utilize fresh patient tissue for establishing xenografts, follow institutional protocols to obtain necessary approvals. When using established cell lines or dissociated PDX cells (fresh or frozen) to generate xenografts, care should be taken to validate the authenticity by short tandem repeat analysis prior to implantation.

2.1 Tumor Dissociation

1. Surgical scalpels, #10 blade.
2. Dulbecco's modified Eagle's medium (DMEM): Store at 4 °C.
3. Trypsin: 10 mg/mL. Dissolve 100 mg vial of trypsin from bovine pancreas with 10 mL of phosphate-buffered saline without magnesium and calcium (PBS^{-/-}). Store at -20 °C (*see Note 1*).
4. Type II collagenase: Store at 4 °C (*see Note 2*).
5. Water bath, 37 °C.
6. Soybean trypsin inhibitor: 10 mg/mL. Dissolve 100 mg vial of soybean trypsin inhibitor with 10 mL of PBS^{-/-}. Store at -20 °C (*see Note 3*).
7. Deoxyribonuclease I: 2 mg/mL. Dissolve 10 mg vial of Deoxyribonuclease I from bovine pancreas with 5 mL of normal saline. Store at -20 °C (*see Note 4*).
8. Magnesium chloride: 1 M. Dissolve 20.3 g of magnesium chloride hexahydrate in 90 mL of water. Adjust the total volume to 100 mL using water and sterilize by autoclaving. Store at room temperature (*see Note 5*).
9. Cell strainer, 40 μm size.
10. Transfer pipettes.
11. Pipettes with tips, adjustable volume 60–1000 μL.
12. Centrifuge.
13. Red blood cell lysis solution: Store at room temperature (*see Note 6*).
14. Ninety percent PBS^{-/-}/10% fetal bovine serum (FBS): Add 5 mL of FBS to 45 mL of PBS^{-/-}. Store at room temperature (*see Note 7*).

2.2 Preparation of Cell Suspension

1. Matrigel basement membrane matrix: Store at -20 °C (*see Note 8*).
2. Wet ice.
3. Pipette tips, chilled on ice.

2.3 Orthotopic Injection

1. Povidone–iodine prep pads.
2. Alcohol wipes: 70% isopropyl alcohol.
3. Luer lock 50 μ L Hamilton glass syringe with 25 gauge needle.
4. Small animal isoflurane anesthesia machine with nose cone attachment.
5. Electric heating pad.
6. Immunocompromised mouse (*see Note 9*).

3 Methods

Carry out all procedures at room temperature, unless otherwise specified.

3.1 Tumor Dissociation

1. Weigh tumor and place in culture dish.
2. Rinse tumor with DMEM warmed to room temperature.
3. Use sterile scalpels to chop/mince tumor into small pieces.
4. Transfer tumor mixture to 50 mL conical tube.
5. Add additional warmed DMEM to fill the tube to ~45 mL (*see Note 10*).
6. Add trypsin thawed to room temperature to tube with tumor based on tumor weight. If tumor <5 g, add 600 μ L; if tumor is 5–10 g, add 900 μ L; if tumor is >10 g, add 1200 μ L.
7. Add 50 mg of type II collagenase to tube (*see Note 11*).
8. Invert tube several times to disperse enzymes within tumor mixture.
9. Place tube in warm 37 °C water bath for 60 min (*see Note 12*).
10. Remove tube from water bath check to see if dissociated properly. The solution should be viscous near the tumor but still able to fit through a transfer pipette (*see Note 13*).
11. Stop dissociation by adding thawed soybean trypsin inhibitor equal to the amount of trypsin used in **step 6**.
12. Invert the tube several times.
13. Add equal amounts of thawed deoxyribonuclease I and magnesium chloride in 60 μ L increments until tumor fragments easily settle at the bottom of the tube. Invert the tube several times between each addition (*see Note 14*).
14. Filter the tumor suspension into a new 50 mL conical tube with a 40 μ m cell strainer.
15. Centrifuge tube at 500 $\times g$ for 5 min.
16. Discard the supernatant above the cell pellet.

17. Add 10 mL of red blood cell lysis solution to the cell pellet and allow to incubate at room temperature for 10 min (*see Note 15*).
18. Add 40 mL of PBS^{-/-}/10% FBS to fill the 50 mL conical tube.
19. Centrifuge tube at $500 \times g$ for 5 min.
20. Aspirate and discard the supernatant.
21. Resuspend the cell pellet in DMEM for counting (*see Note 16*).

3.2 Preparation of Cell Suspension

1. Separate your desired volume of cells from your freshly dissociated tumor to be injected into mice (*see Note 17*).
2. Centrifuge tube at $500 \times g$ for 10 min.
3. Aspirate and discard the supernatant.
4. Gently resuspend cell pellet in thawed Matrigel on ice at a concentration of 1×10^6 per 10 μ L using cold pipette tip. Keep Matrigel cell suspension stored on ice (*see Note 18*).

3.3 Orthotopic Injection

1. To minimize distress and movement during the procedure, anesthetize the mouse with isoflurane gas.
2. Connect mouse to nose cone anesthesia on top of a heating pad.
3. Place the mouse in the supine position.
4. Prep the skin of the knee joint with three rounds of alternating povidone-iodine scrubs and 70% isopropyl alcohol wipes.
5. Load 10 μ L of tumor cell suspension into the Hamilton injection syringe with attached 25 gauge needle and place on wet ice (*see Note 19*).
6. The prepped leg is flexed at the knee joint by placing the index finger in the groin area and the securing the animal's foot to the work surface with the thumb. The femoral condyles become visible upon restraint of the leg (*see Note 20*).
7. Hold the injection syringe at a 45° angle to the mouse.
8. Insert the needle tip between the condyles and then swivel upward to align with the shaft of the femur (*see Note 21*).
9. Gently push the injection needle into the shaft of the femur. Piercing of the bone may require gentle rotation and forward pressure.
10. Advance the needle all the way down to the femoral head, which is approximately 5–10 mm depending on the size of the mouse. The needle is in the shaft of the femur when resistance is gone while moving the needle forward, and when the needle meets resistance when gently moving laterally (*see Note 22*).
11. Slowly inject the tumor cell suspension (*see Note 23*).

12. Gently remove the needle and apply a small amount of pressure to the insertion site (*see* **Note 24**).
13. Allow the mouse to wake from anesthesia and observe the mouse closely to make sure that movement of all limbs is intact. Mice should recover quickly with no visible limping or mobility limitation.

4 Notes

1. We find it best to aliquot trypsin (10 mg/mL) into eppendorf tubes with 600 μ L per tube ahead of time. This allows for quick thawing before tumor dissociation and minimizes the amount of freeze–thaw cycles of larger aliquots which may reduce trypsin activity.
2. We find it best to weigh and aliquot collagenase into 50 mg eppendorf tubes ahead of time.
3. We find it best to aliquot soybean trypsin inhibitor (10 mg/mL) into eppendorf tubes with 600 μ L per tube ahead of time. Concentrated solutions greater than 10 mg/mL may be hazy and have a yellow to amber color.
4. We find it best to aliquot deoxyribonuclease I (2 mg/mL) into eppendorf tubes with 500 μ L per tube ahead of time. Storage at temperatures greater than -20°C will cause decreased activity, and excessive heating above 60°C will cause solutions to lose activity in a number of hours. We recommend thawing aliquots at room temperature prior to use.
5. Magnesium chloride is extremely hygroscopic. We recommend making aliquots in eppendorf tubes with 500 μ L per tube ahead of time. Do not store opened bottles for long periods of time. Once the crystals become saturated with water, dispose of the chemical properly.
6. Several red blood cell lysis solutions are sold at $10\times$ concentrations. Follow manufacturer's instruction for dilution. We recommend storing in 10 mL aliquots.
7. The PBS^{-/-}/FBS solution is best made fresh that day. Keep in mind that you may need to thaw FBS ahead of time as this is typically stored at -20°C .
8. Matrigel typically comes in frozen 10 mL vials which can take several hours to thaw on ice. Special care should be taken to make sure that the Matrigel stays cool during the thawing process to prevent premature polymerization. We recommend thawing Matrigel overnight on ice in a cold room and aliquoting into 1 mL eppendorf tubes, stored at -20°C . These smaller aliquots will thaw quicker on ice on the day of use. Matrigel should not be rapidly thawed using heat.

9. Immunocompromised mice are needed when trying to engraft human tumor cells. Nod-scid or nod-scid gamma strains may be used. When using PDX models for preclinical testing with chemotherapy, nude mice tend to be preferred as they tolerate treatment better.
10. When mincing the tumor, try to create a tumor slurry with the DMEM. If you have a large tumor over 10 grams, it is best to split the tumor into multiple conical tubes or to use a larger glass or plastic container that will hold 100–200 mL with a lid. If a large tumor is packed into a tube, there will be less room for mixing/dissociation of cells.
11. Collagenase type II comes in the form of a brown powder. It is best to dissolve this powder first in a small amount of DMEM prior to adding to the tube with tumor. Once trypsin is added to the tumor tissue, the contents may become more viscous as the tumor starts to dissociate and the collagenase does not mix in as well if simply added to the tumor tube as a powder.
12. Ewing sarcoma tumors tend to dissociate well enzymatically in the water bath alone. If you have a tumor that has more of a firm or bony component, we recommend placing a stir bar in the tube and placing the tube on a magnetic stir plate in the water bath. This can help further mechanically dissociate the tumor.
13. If the solution is too thick, we recommend splitting the sample into 2 tubes, filling with additional DMEM, and incubating for an additional 30–60 min.
14. This is an important step to make sure that the tumor cells are well dissociated. If the tumor suspension is too thick, it will not move through the strainer, so it is best to make sure that enough equal amounts of deoxyribonuclease I and magnesium chloride are added prior to moving on to the next step.
15. You may need to gently tap the conical tube to break up the cell pellet prior to adding the red blood cell lysis solution. If you have a large pellet, we also suggest gently pipetting the cell pellet with a small amount of red blood cell lysis solution first to help break it up before adding the rest of the solution.
16. Automated cell counters can be used; however we recommend visually counting the cells to ensure that only viable appearing cells are counted. If you are unable to inject your tumor cells on the same day of tumor processing, you can resuspend cells in freezing media of FBS/10% DMSO in cryo vials to be thawed and injected at a later date.
17. We typically recommend injecting 1×10^6 cells per mouse. If you are wanting faster engraftment, some lines may grow more rapidly if you inject $1.5\text{--}2 \times 10^6$ cells per mouse. At this step,

alternatives to freshly dissociated cells may also be used such as thawed cryo-preserved cells or cells that have been grown in culture as cell lines.

18. We recommend placing a box of pipette tips on ice prior to resuspending cells in Matrigel. This step requires quick yet steady transfer of Matrigel to the cell pellet as you want to keep the Matrigel cool; however care should be taken not to pipet up and down too quickly. This causes bubbles in the Matrigel which are difficult to remove when trying to pull up the cells later in the injection syringe.
19. You will need a small amount of volume to charge your injection syringe. We suggest preparing an extra 100–200 μL of tumor cell suspension or utilizing blank Matrigel to charge the syringe before pulling up your tumor cells for injection. In addition, if you need to change the needle multiple times, you will likely need extra volume to charge your syringe.
20. As an alternative, a piece of tape can be used to help secure the foot of the leg instead of using your thumb.
21. Gentle swiveling up of the syringe helps to avoid ligament damage during the insertion step.
22. The insertion of the needle into the femur takes practice. We recommend first using euthanized mice to get the feeling for the needle inside the femur. Care should be taken not to move the needle too far laterally in either direction and to avoid using excessive force which will result in femur fracture. Utilization of dye for practice injections to ensure proper placement as well as post injection X-rays to verify the femur is intact can be helpful in perfecting the technique.
23. Care should be taken to inject the tumor cell suspension slowly. Since the bone marrow cavity is connected directly to circulation, pushing the tumor cells rapidly is more likely to cause rapid dissemination of the cells to metastatic sites rather than producing a localized tumor. An alternative approach includes utilizing a para-femoral injection. The injection technique is similar, however the needle tip is inserted adjacent to the femoral condyles and then pushed into the caudal thigh muscle tissue parallel to the femur shaft. When the tip of the needle scratches the surface of the periosteum, the tumor cell suspension is injected.
24. There is typically minimal to no bleeding after the injection. If you do have continued bleeding at the injection site that persists despite pressure, a small drop of tissue glue can be utilized to help close the skin. Mice found to have excessive bleeding, bruising, or limping following the procedure should be euthanized without delay.

Acknowledgments

The sharing of tumor processing and orthotopic injection protocols was supported by The Childhood Solid Tumor Network (www.stjude.org/CSTN).

References

1. Minas TZ et al (2017) Combined experience of six independent laboratories attempting to create an Ewing sarcoma mouse model. *Oncotarget* 8(21):34141–34163
2. Daniel VC et al (2009) A primary xenograft model of small-cell lung cancer reveals irreversible changes in gene expression imposed by culture in vitro. *Cancer Res* 69:3364–3373
3. Tentler JJ et al (2012) Patient-derived tumour xenografts as models for oncology drug development. *Nat Rev Clin Oncol* 9(6):338–350
4. Stewart E et al (2016) The childhood solid tumor network: a new resource for the developmental biology and oncology research communities. *Dev Biol* 411(2):287–293
5. Stewart E et al (2017) Orthotopic patient-derived xenografts of pediatric solid tumors. *Nature* 549(7670):96–100
6. Jones DTW et al (2019) Molecular characteristics and therapeutic vulnerabilities across paediatric solid tumours. *Nat Rev Cancer* 19:420–438
7. Langenau DM et al (2015) Preclinical models provide scientific justification and translational relevance for moving novel therapeutics into clinical trials for pediatric cancer. *Cancer Res* 75(24):5176–5186



Ewing Sarcoma PDX Models

Didier Surdez, Lorena Landuzzi, Katia Scotlandi,
and Maria Cristina Manara

Abstract

Ewing sarcoma (EWS) is a rare malignant pediatric tumor and patient derived xenografts (PDXs) could represent a possibility to increase the number of available models to study this disease. Compared to cell derived xenografts (CDX), PDXs are reported to better recapitulate tumor microenvironment, heterogeneity, genetic and epigenetic features and are considered reliable models for their better predictive value when comparing preclinical efficacy and treatment response in patients. In this chapter, we extensively describe a method for generating Ewing sarcoma PDX models, for their validation and molecular characterization.

Key words Ewing sarcoma, Patient-derived xenografts, Immunodeficient mice, Preclinical models, Pediatric tumors

1 Introduction

Historically, the discovery of athymic immunosuppressed mice (*nude* mice) in 1962 changed the paradigm of cancer research [1]. Indeed, since that time, it is possible to graft tumor cells or human tumor fragments on these T cells deficient mice and thus avoid transplant rejection. Using for the first time this approach, Rygaard and Povlsen [2] implanted under the skin of nude mice a fragment of colon cancer of a 71-year-old patient. This tumor developed as a differentiated adenocarcinoma like that of the donor. This model could be propagated over a period of 7 years, representing seventy-six successive transplants of the nude mouse tumor. Subsequently, other immunodeficient murine models (SCID, NOD-SCID, NSG, etc.) have been developed and also used for this purpose. During the 1980s, evidences for a good correlation between the response to certain chemotherapies in patients and in associated patient derived xenograft (PDX) models was reported but did not draw much attention [3]. In parallel, subcutaneous cell line derived xenograft models (CDXs) in

immunocompromised mice were also developed and emerged as an easier model to work with for preclinical studies. However, the predictive value of CDX models seemed, already at the time, less convincing than PDX models. Indeed, a study synthesizing the responses to numerous cytotoxic agents, in a panel of 39 CDX, had demonstrated the low correlation between the efficacy of these drugs in these models and in patients [4]. However, for reasons of accessibility to PDX models, ease of implementation of CDX models as well as the emergence of other mouse models (including transgenic models), PDX models have fallen into oblivion for almost three decades. During this period, most preclinical studies in mice, whether conducted by the academic laboratories or the pharmaceutical industry, have been based on these poorly predictive CDX models. However, in the last decade, a considerable effort has been made to rehabilitate PDX models for preclinical research in view of their faithfulness to recapitulate tumor heterogeneity, genetic and epigenetic features and for their better predictive value when comparing preclinical efficacy and treatment response in patient [5–11]. As any models, PDX present also some disadvantages that should be taken into consideration. For instance, PDX are generated into immunodeficient mice, which are, at the time of the expansion of immunotherapy-based approaches, not suitable for these preclinical studies. Humanized PDX models have been developed to circumvent to some extent this aspect [12]. Another illustration of potential caveats to consider in PDX models is for instance the cross-species signaling dysfunction. Indeed, murine hepatocyte growth factor (HGF) does not recognize and activate its human MET orthologous receptor [13] which is problematic when testing molecules targeting this pathway. However, this can be addressed using transgenic immunodeficient mice expressing human HGF [14] as illustrated above. Alternative approaches have been developed to circumvent some of the PDX drawbacks and scientists must carefully consider, if the use of PDX is the most suitable model to conduct their research.

In this regard, Ewing sarcoma PDX, could provide a better option to reproduce the pathogenesis of this rare pediatric and adolescence bone tumor. Indeed, a recent study showed that gene expression profile analysis of the PDX and of the cell culture obtained from the same patient demonstrated a higher concordance between the PDX and the human tumor than the cell culture [11].

With the idea of sharing our experience to generate Ewing sarcoma PDX models in the past years, in this chapter, we specifically focus on methodological aspects and also describe how to validate and characterize these models.

2 Materials

2.1 Tissue Collection and Preparation

1. Biological safety cabinet Biosafety Level 2 (BSL2).
2. Petri dishes 60 mm.
3. 50 ml sterile tubes.
4. Sterile fine scissors, scalpels, and tweezers.
5. Culture medium (i.e., RPMI or IMDM) plus 10% fetal bovine serum (FBS, optional) and penicillin–streptomycin.
6. Fetal bovine serum (certified or tested to comply with S(O)FP animal facility requirements).
7. Dimethyl sulfoxide (DMSO), sterile-filtered.
8. Freezing medium (90% FBS; 10% DMSO).
9. Sterile cryovials.
10. Mr. Frosty™ Freezing Container.
11. Vacutainer blood collection EDTA-treated tubes.
12. –80 °C ultra-freezer and liquid nitrogen storage tank.
13. Histopaque-1077.
14. Parafilm and specimen jar labels.
15. Ice bucket with prechilled cold packs.
16. Dry ice.
17. Isothermic box and absorbent material.

2.2 Tumor Engraftment

1. NSG (NOD.Cg-Prkdcscid IL2rgtm1Wjl/SzJ) mice (5–10 weeks old) are preferred for the establishment of PDX models but similar results can be achieved with SCID (CB17/Icr-Prkdcscid/IcrIcoCrl) or Swiss Nude (Crl:NU(Ico)-Foxn1nu) immunodeficient strains [15](*see Note 1*).
2. Anesthetic solution, for example a mixture of xylazine (20 mg/ml, 2% Rompun or 2%Xilor) and tiletamine+zolazepam (50/50 mg/ml Zoletil), should be prepared at the time of use by diluting the drugs in phosphate buffered saline (PBS). Xylazine should be diluted 12.6× and tiletamine+zolazepam 7.875×. For example, by adding to 250 µl of PBS, 25 µl of xylazine and 40 µl of tiletamine+zolazepam for a final volume of 315 µl.
3. 70% ethanol (v/v): 70 ml ethanol (absolute), 30 ml sterile deionized water.
4. 0.1% chlorhexidine gluconate solution.
5. Sterile surgical gloves.
6. Sterile gauze compress.
7. Mice trimmer.

8. Polystyrene disposable sterile forceps.
9. Dissecting scissors, forceps, and tweezers, sterile autoclaved.
10. Standard pattern forceps—straight 14.5 cm.
11. London forceps—angled 16 cm.
12. Adson forceps—serrated straight.
13. Michel suture clips and clip applicator.
14. Sterile 9 mm wound clips, wound clip applicator and wound clip remover autoclaved.
15. Sterile calipers for measurement of tumor size.

2.3 Passages and Tumor Collection from Mice

1. Ice.
2. 96% Ethanol.
3. Sterile tweezers and scissors.
4. Sterile Phosphate-buffered saline (PBS).
5. 10% neutral buffered formalin solution (*see Note 2*).
6. 1.5 ml sterile cryovial tubes.
7. Plastic petri dishes.
8. Liquid nitrogen.
9. Ultrafreezer -80°C .

2.4 Ewing Sarcoma PDX Validation

2.4.1 Histology (See Note 2)

1. Paraffin wax.
2. Microtome.
3. Hentellan jars.
4. Xylene, 100%; 96%, 70% ethanol.
5. Precoated slides.
6. Hematoxylin and Eosin.
7. Acid ethanol.
8. Deionized water (dH_2O).
9. Xylene-based mounting medium.
10. Microscope.

2.4.2 Immunohistochemistry (See Note 2)

1. Xylene, 100%; 96%, 70% ethanol.
2. dH_2O .
3. Hematoxylin.
4. Sterile Phosphate-buffered saline (PBS).
5. Antigen Unmasking: 10 mM Sodium Citrate Buffer (*see Note 3*).
6. Methanol (*see Note 2*).
7. 37% hydrogen peroxide (*see Note 4*).
8. ABC Reagent: (Vectastain ABC Kit, or equivalent).

9. Diaminobenzidine (DAB) stock solution 100× (*see Note 5*).
10. Xylene-based mounting medium.
11. Microscope.

2.4.3 PDX Chimerism

1. Mortar and pestle.
2. Liquid nitrogen.
3. RNase-free water.
4. DNase I (RNase free).
5. RNA extraction reagent or kit (TRIzol, RNeasy plus mini kit or equivalent).
6. Reverse transcription kit (High-Capacity cDNA Reverse Transcription Kit or equivalent).
7. Quantitative PCR reagents (Power SYBR™ Green PCR Master Mix or equivalent).
8. Primers:

TBP_Hs_forward	5'-AGAACAACAGCCTGCCACCTTAC-3'
TBP_Hs_reverse	5'-GGGAGTCATGGCACCCCTGAG-3'
Tbp_mm_forward	5'-CCCTTGTACCCTTCACCAATGAC-3'
Tbp_mm_reverse	5'-TCACGGTAGATAACAATATTTTGAAGCTG-3'
TBP_Hs+mm_forward	5'-TGCACAGGAGCCAAGAGTGAA-3'
TBP_Hs+mm_reverse	5'-CACATCACAGCTCCCCACCA-3'

3 Methods

3.1 Tissue Collection, Preparation and Delivery of Fresh and Frozen Samples

Before starting tissue collection, the research project involving PDX establishment must be approved by the Institutional Ethic Committee. Patients must receive exhaustive explanation regarding the project and authorize the use of their samples by signing an informed patient consent form. Furthermore, the use of animals for experimentation is strictly framed and all experiments made on animals must be tested in the respect of the guidelines established by Veterinary ethic Committee and National Law (*see Note 6*).

All tissue and blood preparation should be performed in a Biological Safety Cabinet (BSC) using sterile instruments and technique. Sterility must be kept during all procedures.

When receiving the sample for the establishment of PDX models, it is important to simultaneously preserve and collect patient samples (blood and tumor) for the subsequent molecular characterization procedure of these models. Therefore, if possible, it is recommended to keep at least one frozen piece and one formalin-

fixed, paraffin-embedded (FFPE) tumor piece as well as a blood sample. Since blood/bone marrow transplantation can occur in these patients, this information should be asked to the clinicians before considering using the blood sample for sequencing approaches (*see Note 7*).

Ideally, engraftment of the tumor sample should occur at the same location than the surgery/biopsy. However, if not possible, the following procedure can be used to transport these sample.

1. The tumor should be maintained at 4 °C in culture medium until implantation (*see Note 8*).
2. Transfer the tumor material into a sterile petri dish along with a small volume of the culture medium used for transportation to keep the tissue wet. Evaluate the material before cutting to make sure that viable tumor tissue is being implanted avoiding necrotic areas as well as normal tissues (bone, cartilage, connective tissue).
3. Cut the tissue into fragments using a sterile scalpel or fine scissors. If possible, the fragment dedicated to the PDX implantation should have a diameter of 3–4 mm (corresponding to 10–30 mm³ tissue fragment); engraftment of fragments of smaller dimensions may have lower probability to be successful.
4. For future characterization, patient's tumor tissue should be snap frozen as soon as possible (i.e., within a couple of minutes) in liquid nitrogen immediately following surgery and then stored at –80 °C.
5. For molecular characterization, a constitutional (germ line) sample (most frequently a patient's blood sample) prior to the initiation of chemotherapy, should be used (*see Note 9*).
6. If the tumor implantation needs to be delayed, samples should be suspended in 1 ml of freezing medium, place into the slow-rate freeze container and store at –80 °C overnight. Subsequently, vials should be transferred in a liquid nitrogen storage tank until delivery to the animal facility.
7. A fragment of tumor tissue must be available for histopathology and immunohistochemistry and should be fixed in a 10% formalin solution (*see Note 2*), routinely processed, and embedded in paraffin.
8. Any remaining tissue can be used for in vitro cultures.
9. In case of delivery of a patient tumor fresh sample to a distant animal facility, specimens must be placed in prechilled medium vials and shipped in sterile tubes, filled with culture medium +10% FBS.
10. Vials should be sealed with parafilm and wrapped with adsorbent material to avoid liquid leaks.

11. To maintain cold temperature, samples must be transported with prechilled cold packs in an isothermic box for implantation into mice.
12. In case of shipping of frozen patient tumor samples to a distant animal facility, cryotubes should be properly labelled and placed in a sealable plastic bag. The plastic bag should be placed in a screw cap plastic container, and the container should be placed in a polystyrene box containing enough dry ice for shipment (*see Note 10*).

3.2 Tumor Engraftment Over or Under the Inter-Scapular Brown Fat Pad

All experiments involving live animals must be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) and performed following National and International law.

1. Shave mice with a hair trimmer if necessary.
2. Prepare as many tumor fragments (4 mm × 4 mm) as mice to be engrafted in a petri dish and keep the tissue moist with some drops of PBS (Fig. 1).
3. Anesthetize immunodeficient mice with Xylazine, Tiletamine and Zolazepam solution mixture. 40 µl of the final solution should be administered by intramuscular injection in an anterior leg of the mouse (*see Note 11*).
4. Transfer the animal to the surgical platform in a sterile field and place it in a prone position.
5. When the mouse is fully anesthetized (*see Note 12*), disinfect the skin with sterile gauze compress embedded with chlorhexidine solution. Repeat this procedure twice with new embedded compresses.

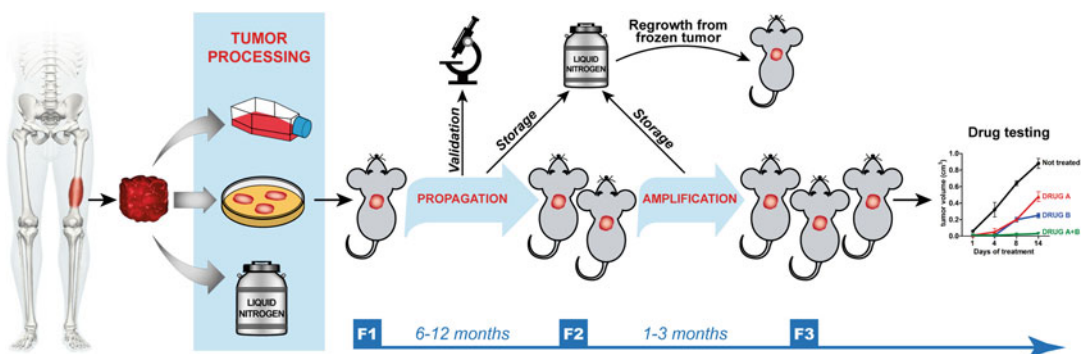


Fig. 1 Generation of EWS-PDX model. Tumor samples obtained from surgical or biopsy specimens are divided in small fragments and are used as follows: subcutaneously implanted in mice, snap-frozen and stored in biobank, and dissociated to establish primary cell cultures. When the tumor reaches the ethical size, mice will be euthanized. At every passage, tumor is divided into representative portions for different applications, such as model establishment or propagation (so-called passage), viable fragment freezing, histology, nucleic acid and protein isolation, and generation of viable cell lines. From the third passage, this model is defined as established and can be used for preclinical studies for instance

6. With sterile scissors, make a small (4–6 mm) incision (parallel to the spinal column) between the shoulder blades being careful not to cut any muscle. There should be little or no bleeding with this procedure.
7. Make a subcutaneous pocket by inserting the round tip of a tweezer into the center of the transplantation site and make a pocket about 4–6 mm long using blunt dissection technique and reaching the area of the inter-scapular brown fat pad.
8. Insert the tumor fragment into the pocket with the tweezer by placing it over the area of the brown fat pad. Carefully remove the tweezers once the tissue has been placed.
9. Close the incision with a wound clip by using the wound clip applier.

Alternatively, the tumor fragment can be implanted under the fat-pad (*see Note 13*) following the point 1–6 of the above procedure and continuing with the alternative procedure described below (point 10–17):

10. Using your left hand (if right-handed), hold and lift the mouse skin (approximately 1 cm from the incision site) with disposable sterile forceps.
11. With your right hand, introduce London forceps into the incision and gently pull out the brown fat pad, keep the fat pad grabbed outside the mouse and avoid touching the surface of the skin.
12. With your left hand, release the skin and put the disposable sterile forceps on the side.
13. With your left hand and Adson forceps, hold the fat pad a few millimeters away from the London forceps and hold it firmly.
14. With your right hand, release the contention of the London forceps on the fat pad (still hold with your left hand) and with the angled tips of the London forceps, drill a small hole in the membrane between the mouse spine and the fat pad and release slightly the London forceps to generate a cavity under the fat pad.
15. With your right hand and London forceps, take a tumor fragment and insert it in the cavity under the scapular fat pad.
16. With Adson forceps in your left hand, place the fat pad under the mouse skin.
17. Close the incision with a wound clip by using the wound clip applier.
18. After surgery, identify animals by the current method used in the laboratory (*see Note 14*).
19. Return the mouse to its sterile home cage and keep it in a warm place until it awakens.

20. Monitor the mice daily for 3–5 days post-surgery and remove wound clips within 7–10 days.
21. Check tumor growth weekly and measure its dimension with a sterile caliper.
22. When the tumor reaches the end-point size, mice are euthanized (*see Note 15*).

3.3 Passages and Tumor Collection from Mice

Necropsy must be performed very carefully recording tumor features such as color, vascularization, consistency, necrotic and viable areas. Any potential sites of metastasis can be identified by inspecting lungs, brain and all of the organs in the peritoneum. Ewing sarcoma tumors typically show white color and a very soft loose consistency.

1. Euthanize the mouse using the method approved by the appropriate Ethics Committee.
2. Lay the mouse in a supine position on the surgical platform and clean the incision area and the whole mouse with ethanol. The hair must be completely wet.
3. Make an incision to expose the tumor. Carefully detach the skin from the tumor using tweezers and scissors.
4. Carefully remove the tumor.
5. Put the tumor sample in a petri dish and place it on ice.
6. Divide the tumor into representative portions for different applications, such as passage in other mice, viable fragment freezing, histology, nucleic acid and protein isolation, and generation of viable cell lines.
7. Put sufficient fresh tumor samples in PBS for passage in other mice (*see Note 16*), for viable fragment freezing and generation of viable cell lines (Fig. 1).
8. Fix some tumor samples in 10% neutral buffered formalin solution for histology (*see Note 2*).
9. For nucleic acid and protein extraction, put some small (diameter 2–3 mm) tumor fragments in 1.5 ml conical polypropylene tubes and immediately snap-freeze the vials by submerging them in liquid nitrogen and then transfer the vials at -80°C in an ultra-freezer.

3.4 Histological Validation

The comparison between PDX models at different in vivo passages and patient's tumors should be done in all PDX obtained. As a first step, histopathologic features of the patient Hematoxylin-Eosin-stained slides and the PDX derived samples should be compared to gain information about histologic similarity (Fig. 2a).

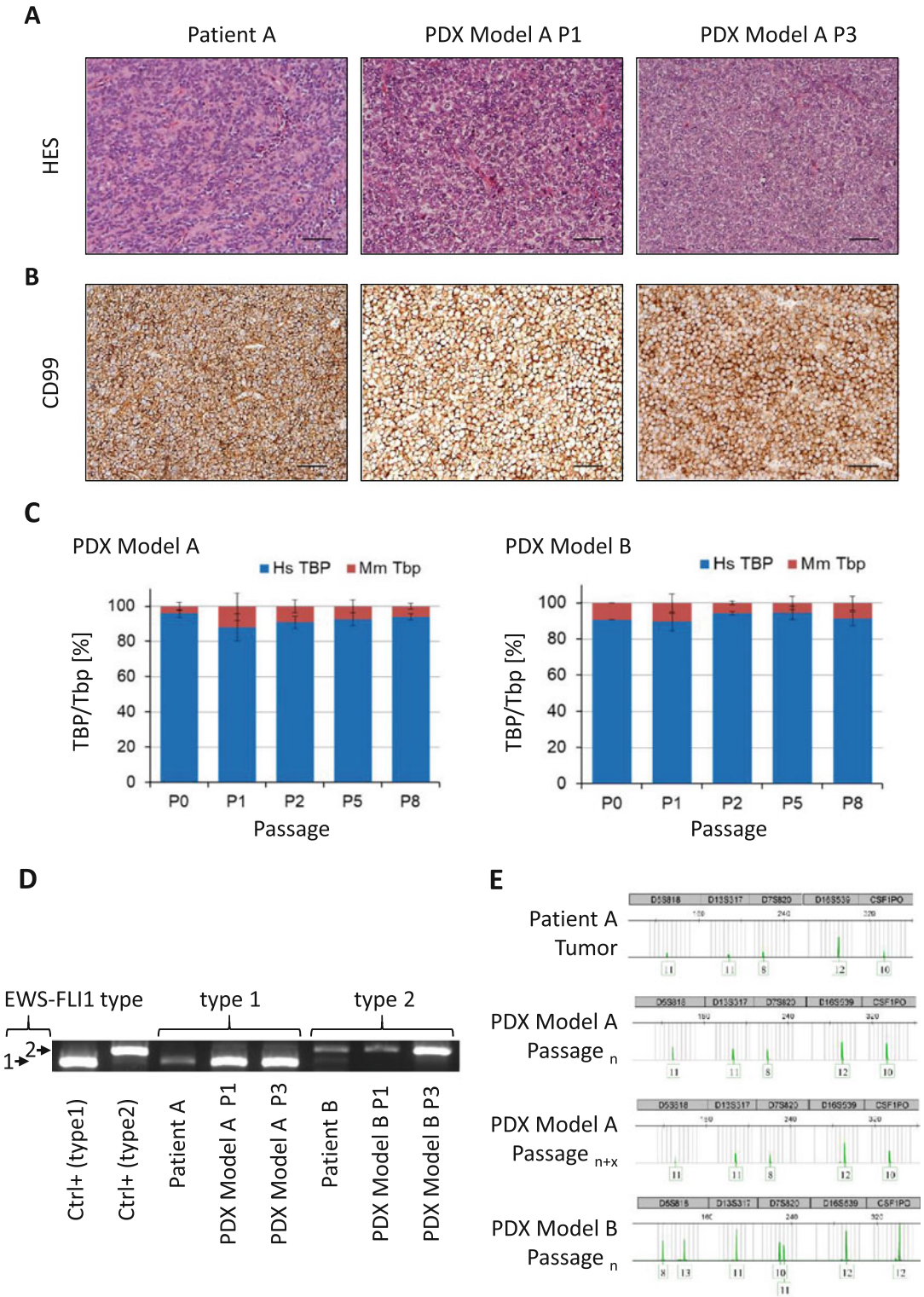


Fig. 2 (a and b) Histologic and immunohistochemical features of patient's tumors and corresponding PDX at different in vivo passages. EW PDXs consist of small round cell sheets, closely packed and without matrix, as patient's tumors. CD99 antigen expression of PDX reflected that of patient's tumor. (c) Absolute proportion of

1. The tissues from human tumor and PDX have to be fixed in 10% buffered formalin, routinely processed, and embedded in paraffin (*see Note 2*).
2. Serial, 4- μ m-thick, paraffin sections are mounted on precoated slides and processed as described in Subheading 3.5.1 according to standardized procedures [11].
3. After, incubate slides with Hematoxylin for 3 min and rinse with dH₂O and with tap water for 5 min.
4. Dip 8–12 \times (fast) in Acid ethanol to destain slides.
5. Rinse with tap water and for dH₂O for at least 5 min.
6. Stain the slides with Eosin (30 s).
7. Dehydrate slides as described in Subheading 3.5.6.
8. Coverslip slides using a xylene-based mounting medium (*see Note 17*).

The comparative morphological analysis between patient's tumor and corresponding PDX should include the evaluation of the following characteristics:

- (a) Tumor cellularity;
- (b) Pattern of growth;
- (c) Mitotic activity;
- (d) Cell morphology, including degree of pleomorphism and differentiation.

Histologically, conventional Ewing sarcoma is uniformly made of sheets of small round cells that are closely packed and without matrix [16]. The chromatin is finely stippled, and nucleoli are usually not evident. Usually, extensive deposits of glycogen are observed in the cytoplasm, that is generally scarce. A 'large cell', or 'atypical', variant of Ewing sarcoma has been reported [17]; the main difference of these cells from conventional Ewing sarcoma are larger-size nuclei with irregular contours. In the case of Peripheral Neuro Ectodermal Tumors (PNET), the presence of rosette pattern is common (*see Note 18*). The analysis and comparison of these morphological characteristics between the tumor of origin and the PDXs allows to classify the PDX models in three main categories:



Fig. 2 (continued) human derived tumor and murine stroma tissues in two EWS PDX models across several passages, **(d)** Evaluation of the EWSR1-ETS fusion transcript in EW PDXs and in patient's tumor samples. Positive controls are included. **(e)**. STR profiling between patient's tumor and PDX allows for the authentication of models along passages. PDX derived from tumor B displays a different STR profile than the one derived from tumor A

Complete identity: No difference between patient and xenograft-derived samples;

Similar pattern: Changes in growth pattern and overall cellularity allowed, but no differences in mitotic activity or/and pleomorphism;

Morphologic Shift: significant morphological differences concerning all the characteristics indicated above [7, 11].

The comparison should be done in all the generations of PDX obtained.

3.5 Immuno-histochemical Validation

Patient samples and the matched xenograft tissues should be assessed for markers of proliferation, resistance to apoptosis, or neural differentiation like Ki-67, Caspase 3 (total), Caspase 3 (cleaved), S-100, or Neuron-Specific-Enolase (NSE). In addition, a number of diagnostically delineating markers should be performed based on the histopathologic diagnosis like CD99 antigen (Fig. 2b), FLI1, CAV1, BCL11B or GLG1, NKX2-2 [7, 11, 18, 19]. Every analysis should be accompanied by appropriate positive and negative controls. Moreover, all stained sections should include nontumor mouse cells, such as endothelial cells, myopericytes, and fibroblasts which must be constantly negative.

The procedure here presented is an ABC avidin/biotin method, routinely used for IHC staining. However, in alternative, standard operative procedures of each laboratory could be followed.

Do not allow slides to get dry during all steps of this procedure.

3.5.1 Deparaffinization/Rehydration

1. Incubate slides in two washes of xylene for 30 min each.
2. Incubate slides in three washes of 100% ethanol for 5 min each.
3. Incubate slides in one wash of 95% ethanol for 5 min.
4. Incubate slides in one wash of 70% ethanol for 5 min.
5. Rinse slides with dH₂O for 5 min each.

3.5.2 Antigen Unmasking (if Necessary)

1. Put slides in Hellendal jars (10 slides/jar) with citrate buffer.
2. Bring slides to boil in sodium citrate buffer for 3 cycles of 5 min each at 750 W (replace buffer after every cycle).
3. Cool at room temperature for 20 min without changing the buffer of the last cycle.
4. Put slides in PBS for 3–5 min (until starting immunostaining).

3.5.3 Inhibition of Endogenous Peroxidase

1. Incubate slides with inhibition solution for 30 min at room temperature.
2. Wash twice (5 min each) with PBS and once with distilled water (5 min).

3.5.4 Immunostaining

1. Incubate slides for 15 min at room temperature with Horse normal serum (kit Mouse) or Goat normal serum (kit Rabbit) diluted in PBS.
2. Discard normal serum without allowing slides to dry.
3. Incubate slides overnight at +4 °C in a humidified chamber with the primary antibodies diluted in PBS (*see Note 19*).
4. Wash slides twice (5 min each) with PBS and once with dH₂O water (5 min).
5. Incubate slides for 30 min at room temperature with the biotinylated secondary Ab diluted in PBS (*see Note 20*).
6. Wash slides twice (5 min each) with PBS and once with dH₂O water (5 min).
7. Incubate slides for 30 min at room temperature with ABC.
8. Wash slides twice (5 min each) with PBS and once with dH₂O water (5 min).

3.5.5 Development of Immunoreaction

1. Defreeze diaminobenzidine (DAB) stock solution 100× (5 g/100 ml PBS) and prepare DAB working solution (*see Note 5*).
2. Incubate slides for 5 min at room temperature with DAB 1× and add 5–100 µl of H₂O₂ and incubate for additional 5–10 min (*see Note 21*).
3. Wash slides twice (5 min each) with PBS and once with dH₂O water (5 min).
4. Counterstain nuclei with prefiltered hematoxylin (30 s—3 min).
5. Wash slides with distilled water for 3–5 min.

3.5.6 Dehydration

1. Incubate slides in one wash of 70% ethanol for 2 min.
2. Incubate slides in one wash of 95% ethanol for 2 min.
3. Incubate slides in three washes of 100% ethanol for 5 min each.
4. Incubate slides in two washes of xylene for at least 2 min each.
5. Mount coverslips with xylene-based mounting medium.

3.6 Ewing Sarcoma PDX: Evaluating Human and Murine Chimerism

In PDX, the human component arises from the tumor fraction whereas the microenvironment derives from immunodeficient murine stromal cells (human stroma that is present at the time of engraftment is rapidly replaced by murine cells within the first passages) [10, 20]. Evaluating the proportion human and murine derived tissue in a PDX is a simple method that can be routinely used as a first step of model validation. This approach is also useful to promptly detect spontaneous murine tumors that occasionally develop in these immunodeficient strains [21, 22]. For this, RT-QPCR with a set of 3 primers recognizing respectively the

murine, the human or both orthologous transcripts of the *Tbp*, *TBP* genes can be used [23]. In Ewing sarcoma PDX, the murine stroma typically represents 5–15% of the tumor (Fig. 2c). In parallel, the presence of the *EWSR1-ETS* fusion transcript (Fig. 2d) can be used to confirm the proper derivation of a Ewing sarcoma model [24]. It is recommended to perform these RT-QPCR validation methods with all tumors growing on mice from P0 to P2 and later on, every third passage with one or two tumors per passage.

3.6.1 Isolation of RNA

1. Homogenize a snap frozen PDX tumor sample (10–30 mg) with a mortar and pestle, transfer sample powder into a 2 ml Eppendorf tube and add 500 μ l TRIzol reagent without thawing samples.
2. Incubate the homogenized samples for 5 min at room temperature.
3. Add 0.1 ml of pure chloroform, shake firmly for 15 s.
4. Incubate samples for 2–3 min at room temperature.
5. Centrifuge sample at $525 \times g$ for 15 min at 4 °C.
6. Transfer aqueous (clear upper) phase to a fresh tube.
7. Add 1 μ l glycogen as carrier to explant samples.
8. Precipitate RNA by adding 250 μ l isopropanol and incubate at –20 °C overnight.
9. Spin samples at $525 \times g$ for 30 min at 4 °C.
10. Remove isopropanol and wash RNA pellet once with 500 μ l 75% Ethanol.
11. Air dry pellet for 5 min.
12. Resuspend RNA in 20 μ l RNase free water H_2O and incubate 10 min at 55–60 °C.
13. Add 2 μ l DNase I (RNase free) to 20 μ l mRNA and incubate for 30 min at 37 °C. Samples can now be stored at –80 °C or used immediately for reverse transcription.

3.6.2 Reverse Transcription

1. Use 1 μ g of RNA for each reaction (extracted with a method described above).
2. Add reverse transcription buffer, dNTP, random primers, and reverse transcriptase and complete with water according to manufacturer's instructions.
3. Incubate reverse transcription sample according to manufacturer's instructions.

3.6.3 Quantitative PCR

1. Mix forward and reverse primers for each primer set (*TBP_Hs*, *Tbp_mm*, and *TBP_Hs+mm*) to prepare a 10 μ M stock solution (e.g., 20 μ l of 100 μ M *TBP_Hs_forward*, 20 μ l of 100 μ M *TBP_Hs_reverse*, and 160 μ l of water).

2. Add to the cDNA reaction 200 μ l of molecular grade water.
3. Perform duplicate QPCR reactions for all three primer sets as following:
 - (a) 9 μ l of diluted cDNA.
 - (b) 1 μ l of 10 μ M primer stock solution.
 - (c) 10 μ l of 2 \times Power SYBRTM Green PCR Master Mix.
4. Run RT-QPCR program and determine Ct and melting temperature for each reaction.
5. Determine the fraction of murine and human tissue with the following formula:

$$\text{Human fraction} = \frac{2^{-(\overline{\text{CT}}(\text{TBP_Hs})\overline{\text{CT}}(\text{TBP_Hs+mm}))}}{2^{-(\overline{\text{CT}}(\text{TBP_Hs})\overline{\text{CT}}(\text{TBP_Hs+mm}))} + 2^{-(\overline{\text{CT}}(\text{TBPmm})\overline{\text{CT}}(\text{TBP_Hs+mm}))}}$$

$$\text{Murine fraction} = 1 - \text{Human fraction}$$

3.7 Ewing Sarcoma PDX: STR Profiling

When generating or using several PDX simultaneously, the risk of inadvertent model exchange has to be considered and authentication through short-tandem repeat (STR) profiling is highly recommended to confirm their identity. It is recommended to perform STR validation with at least one tumor per passage and confirm its identity with the profile of the matched original patient tumor (Fig. 2c).

3.8 Ewing Sarcoma Comprehensive Molecular Characterization

Concurrently with morphological and immunohistochemical evaluation of markers of EWS, validation of EWS-PDX models should include assessment of type of fusion transcripts as well as the occurrence of the few recurrent secondary alteration present in Ewing Sarcoma (mutation in *STAG2*, *TP53* and deletion of *CDKN1A*) that is expected to mirror the profile of the original tumor samples [11, 25]. In addition, to gain a better insight into the similarity of the PDX with the original tumor, it would be a good practice to perform a global gene expression correlation analysis between gene expression profiles of EWS samples and the corresponding PDX, including the comparison of PDX of different in vivo passages. Moreover, due to the small number of EWS patients, an effort at national and European level should be done, such as that it is currently set up in the consortia dedicated to the generation of a large cohort of pediatric PDX solid tumor entities (ITCC-P4 project <https://www.itccp4.eu/>) and to generate an exhaustive molecular characterization of these models and their matched patient tumors. These comprehensive analyses will include low coverage whole genome sequencing (lcWGS), whole exome sequencing (WES), DNA methylation profiling, and RNA sequencing for all PDX and matched tumor as well as lcWGS and WES for patient germline DNA. These extensively characterized PDX

models shall provide to the scientific community and the pharmaceutical industry, state of the art pediatric PDX models (including for Ewing sarcoma) to transform preclinical investigation into successful clinical trial and innovative therapies against these aggressive pediatric cancers.

4 Notes

1. Human tumor fragments can be engrafted in mice with different degree of immunodeficiency. NOD-SCID-Il2rg^{-/-} (NSG) mice are strongly recommended, at least for the first engraftment, thanks to their high level of immunodeficiency and the long lifespan. Nude and SCID mice can be used as well, but lower immunodeficient mice or mice with a shorter survival should be employed only in later passages when PDX in vivo growth is stabilized and usually faster than initial passage. Nude mice can be useful when testing drug efficacy against PDX because they are less sensitive to cytotoxic and radiation therapies compared to NSG mice.
2. Chemicals must be handled under a fume hood accordingly to MSDS.
3. Citrate buffer: for 1 L add 2.94 g sodium citrate trisodium salt dihydrate and 2.10 g citric acid anhydrous to 1 L dH₂O. Adjust pH to 6.0.
4. Endogenous peroxidase inhibition solution (to be dissolved immediately before use): for 100 ml add 1 ml of H₂O₂ to 99 ml methanol.
5. To prepare 100× DAB working solution dilute 5 g of DAB in 100 ml of PBS and store 1 ml aliquots at -20 °C. Diaminobenzidine is highly dangerous and may cause cancer. Handle with care strictly following the MSDS.
6. In the planning of the experiment, the 3R guidelines should be applied: reducing the number of animals by protocol; refining, that is, improving animal welfare experimentation by limiting suffering and stress; replacing the use of animals by alternative methods as possible. The availability of alternative methods (in vitro systems, mathematics modeling) allowed to reduce number of animals used for biomedical research. However, there is no current model for the moment completely substitute PDX models.
7. All fresh human tissue, including whole blood and its components, must be handled under Biosafety Level 2 (BSL2) conditions. All work is conducted in a biological safety cabinet (BSC) using personal protective equipment and avoiding the use of sharp tools where possible. All tools potentially exposed to the

human material must be treated with a 10% sodium hypochlorite solution for a minimum of 10 min, double bagging for autoclaving and UV disinfection system. Follow all waste disposal regulation when disposing waste materials.

8. Tumor tissue should be implanted as soon as possible, an ideal time is within 2 h from surgery, but implantation in the 24 h post-surgery can yield successful PDX.
9. Blood could be collected and stored as whole blood sample in vials with EDTA (minimum 2 ml, preferentially 5–10 ml), at -80°C ; or in alternative, it is possible separate plasma from blood cells. In this case, whole blood needs to be centrifuged within 2 h of collection at 300 for 15 min at room temperature to separate blood cells from the plasma. After, blood cells should be diluted 1:2 in PBS, separated in the different sub-fractions with Histopaque-1077, according to manufacturer's instructions. Plasma (for cell free DNA extraction) and peripheral mononuclear cells PBMC need to be stored at -80°C .
10. In case of shipping by air transport, packaging must be done in compliance with IATA packaging instruction and properly labeled for biological substance category B (UN3373) and dry ice (class 9 hazard label, UN 1845, indicating weight of dry-ice content).
11. The volume of 40 μl of the final anesthetic solution is the dose recommended for 5–10 weeks old mice. Patient-derived tumor samples at the first implant in mice can have highly variable latency and growth rates. In the fastest growing PDX, tumor graft can become palpable around 2 weeks after implantation, but it is common for Ewing sarcoma tumor grafts to take several months or up to 1 year to become palpable and additional 2–3 months to be ready to be collected for the first time. For this reason, mice not older than 5–10 weeks should be used. If the tumor fragment is frozen prior to implantation, when retransplanting the tumor fragment, some additional months could be required. In our experience, in Ewing sarcoma often the retransplanted tissue will show reduced latency compared to the original patient derived sample. When cohorts of mice are transplanted with similar size fragments from the same stabilized PDX sample, tumors will usually grow with similar latency.
12. Make sure that the mouse is in plane of anesthesia before starting surgery and that the tongue is in the proper position. If some twitching is present, allow more time for the anesthesia to take effect.
13. The expected rate of successful engraftment with the implant over the interscapular fat pad for EWs should be around 24% [11], implant into the interscapular fat pad can give an uptake of 45–50%.

14. It is strongly advisable to keep in separate cages PDXs deriving from different patients. On each cage put a label indicating strain, gender and the unique identification number of the mouse, followed by the PDX unique identification code, number of passages in vivo, for example P1 for the first implant, P2 for the second implant and so on, and the date of the implant.
15. If the mice show any sign of discomfort or illness prior to tumor growth, or the tumor starts to ulcerate, the mice should be euthanized.
16. The procedure for a passage is identical to the one described under Subheading 3.2. In this case, PDX tumor fragments (typically 4 × 4 mm) are grafted in 2–4 immunodeficient mice (one tumor fragment per mouse).
17. Place a drop of mounting medium on the slide taking care to leave no bubbles and allow the mounting medium to cover all the tissue under the coverslip. Dry overnight in the hood.
18. Rosette is defined as a group of cells characterized by nuclei at the periphery and cytoplasm projections toward the center of this structure.
19. Refer to product data sheet for recommended antibody diluent. Optimal dilution should be standardized using proper controls.
20. During the incubation with the biotinylated secondary antibody, prepare the avidin–biotin peroxidase complex (ABC) according to manufacturer’s instructions.
21. Check microscopically the development of the reaction every 2–3 min. Stop the reaction once positive control included in the jar develops.

Acknowledgments

This work is supported by grants received from the Institut Curie; the INSERM; the Canceropôle Ile-de-France; the Ligue Nationale Contre le Cancer (Equipe labellisée) and projet de Recherche “Enfants, Adolescents et Cancer”; the Institut National du Cancer (PLBIO16-291), the Fondation ARC, the Agence Nationale de la Recherche (ANR-10-EQPX-03, Institut Curie Génomique d’Excellence (ICGex) and the société française de lutte contre les cancers de l’enfant et de l’adolescent. The European Union (ERANET TRANSCAN-2_TORPEDO ER-2015-2360405, to KS, *TRANSCAN-2_BRCAddict* TRANS-201801292 to DS and KS), H2020-IMI2-JTI-201 5-07 (116064—ITCC P4 to KS and DS). DS is supported by SiRIC (Grant « INCa-DGOS-4654).

We thank Cristina Ghinelli for the graphic support and Dr. Marianna Carrabotta (IRCCS-Istituto Ortopedico Rizzoli) for

her technical support with evaluation of the EWSR1-ETS fusion transcript. We also thank all parents, patients, and families that consented to provide samples to establish these models. The materials presented and views expressed here are the responsibility of the authors only. The sponsor takes no responsibility for any use made of the information set out.

References

1. Rygaard J (1969) Immunobiology of the mouse mutant “nude”. Preliminary investigations. *Acta Pathol Microbiol Scand* 77:761–762
2. Rygaard J, Povlsen CO (1969) Heterotransplantation of a human malignant tumour to “nude” mice. *Acta Pathol Microbiol Scand* 77:758–760
3. Fiebig HH, Schuchhardt C, Henss H et al (1984) Comparison of tumor response in nude mice and in the patients. *Behring Inst Mitt* 74:343–352
4. Johnson JI, Decker S, Zaharevitz D et al (2001) Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *Br J Cancer* 84:1424–1431. <https://doi.org/10.1054/bjoc.2001.1796>
5. Bruna A, Rueda OM, Greenwood W et al (2016) A biobank of breast cancer explants with preserved intra-tumor heterogeneity to screen anticancer compounds. *Cell* 167:260–274.e22. <https://doi.org/10.1016/j.cell.2016.08.041>
6. Pauli C, Hopkins BD, Prandi D et al (2017) Personalized in vitro and in vivo cancer models to guide precision medicine. *Cancer Discov* 7:462–477. <https://doi.org/10.1158/2159-8290.CD-16-1154>
7. Stewart E, Federico SM, Chen X et al (2017) Orthotopic patient-derived xenografts of paediatric solid tumours. *Nature* 549:96–100. <https://doi.org/10.1038/nature23647>
8. Izumchenko E, Paz K, Ciznadija D et al (2017) Patient-derived xenografts effectively capture responses to oncology therapy in a heterogeneous cohort of patients with solid tumors. *Ann Oncol* 28:2595–2605. <https://doi.org/10.1093/annonc/mdx416>
9. Hidalgo M, Bruckheimer E, Rajeshkumar NV et al (2011) A pilot clinical study of treatment guided by personalized tumorgrafts in patients with advanced cancer. *Mol Cancer Ther* 10:1311–1316. <https://doi.org/10.1158/1535-7163.MCT-11-0233>
10. Hidalgo M, Amant F, Biankin AV et al (2014) Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov* 4:998–1013. <https://doi.org/10.1158/2159-8290.CD-14-0001>
11. Nanni P, Landuzzi L, Manara MC et al (2019) Bone sarcoma patient-derived xenografts are faithful and stable preclinical models for molecular and therapeutic investigations. *Sci Rep* 9:12174. <https://doi.org/10.1038/s41598-019-48634-y>
12. Sanmamed MF, Chester C, Melero I, Kohrt H (2016) Defining the optimal murine models to investigate immune checkpoint blockers and their combination with other immunotherapies. *Ann Oncol* 27:1190–1198. <https://doi.org/10.1093/annonc/mdw041>
13. Rong S, Oskarsson M, Faletto D et al (1993) Tumorigenesis induced by coexpression of human hepatocyte growth factor and the human met protooncogene leads to high levels of expression of the ligand and receptor. *Cell Growth Differ* 4:563–569
14. Zhang Y-W, Su Y, Lanning N et al (2005) Enhanced growth of human met-expressing xenografts in a new strain of immunocompromised mice transgenic for human hepatocyte growth factor/scatter factor. *Oncogene* 24:101–106. <https://doi.org/10.1038/sj.onc.1208181>
15. Lu W, Chao T, Ruiqi C et al (2018) Patient-derived xenograft models in musculoskeletal malignancies. *J Transl Med* 16:107. <https://doi.org/10.1186/s12967-018-1487-6>
16. de Alava E, Lessnick SL, Sorensen PH (2013) In WHO classification of Tumours of soft tissue and bone chapter from Fletcher CDM. In: Bridge JA, PCW H, Mertens F (eds) WHO classification of tumours of soft tissue and bone, 4th edn. IARC Press, Lyon
17. Nascimento AG, Unii KK, Pritchard DJ et al (1980) A clinicopathologic study of 20 cases of large-cell (atypical) Ewing’s sarcoma of bone. *Am J Surg Pathol* 4:29–36
18. Baldauf MC, Orth MF, Dallmayer M et al (2018) Robust diagnosis of Ewing sarcoma by immunohistochemical detection of super-enhancer-driven EWSR1-ETS targets.

- Oncotarget 9:1587–1601. <https://doi.org/10.18632/oncotarget.20098>
19. Shibuya R, Matsuyama A, Nakamoto M, Shiba E, Kasai T, Hisaoka M (2014) The combination of CD99 and NKX2.2, a transcriptional target of EWSR1-FLI1, is highly specific for the diagnosis of Ewing sarcoma. *Virchows Arch* 465(5):599–605
 20. Schneeberger VE, Allaj V, Gardner EE et al (2016) Quantitation of murine stroma and selective purification of the human tumor component of patient-derived xenografts for genomic analysis. *PLoS One* 11:e0160587. <https://doi.org/10.1371/journal.pone.0160587>
 21. Moyer AM, Yu J, Sinnwell JP et al (2019) Spontaneous murine tumors in the development of patient-derived xenografts: a potential pitfall. *Oncotarget* 10:3924–3930. <https://doi.org/10.18632/oncotarget.27001>
 22. Huang P, Westmoreland SV, Jain RK, Fukumura D (2011) Spontaneous nonthymic tumors in SCID mice. *Comp Med* 61:227–234
 23. de Plater L, Vincent-Salomon A, Berger F et al (2014) Predictive gene signature of response to the anti-TweakR mAb PDL192 in patient-derived breast cancer xenografts. *PLoS One* 9:e104227. <https://doi.org/10.1371/journal.pone.0104227>
 24. Peter M, Gilbert E, Delattre O (2001) A multiplex real-time PCR assay for the detection of gene fusions observed in solid tumors. *Lab Invest* 81:905–912
 25. Tirode F, Surdez D, Ma X et al (2014) Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov* 4:1342–1353. <https://doi.org/10.1158/2159-8290.CD-14-0622>



Using Zebrafish Larvae as a Xenotransplantation Model to Study Ewing Sarcoma

Susana Pascoal, Sarah Grissenberger, Eva Scheuringer, Rita Fior, Miguel Godinho Ferreira, and Martin Distel

Abstract

Tumor models allowing for the *in vivo* investigation of molecular mechanisms driving tumor progression and metastasis are important to develop novel strategies for cancer treatment. Unfortunately, for Ewing sarcoma no adequate genetic animal models are currently available. Mouse xenograft models are the state of the art to model Ewing sarcoma *in vivo*. Here, we describe an alternative Ewing sarcoma xenograft model in embryonic and larval zebrafish. This xenograft model offers live imaging and easy compound testing opportunities hereby complementing mouse xenograft models. In this chapter, we provide a detailed protocol how to xenograft Ewing sarcoma cells (shSK-E17T) into 2-day-old zebrafish and how xenografted zebrafish can be imaged and analyzed over consecutive days to study tumor proliferation.

Key words Zebrafish, Xenotransplantation, Ewing sarcoma, Tumor model, Metastasis, Live imaging

1 Introduction

The management of Ewing sarcoma, a pediatric cancer of bone and soft tissues, is still challenging.

The development of effective therapies is hampered by the lack of adequate animal models for Ewing sarcoma. Due to the elusive cell of origin, attempts to genetically model Ewing sarcoma in mouse have failed so far [1]. Thus, xenotransplantation of human Ewing sarcoma cells into mice is currently the gold standard to characterize Ewing sarcoma cells *in vivo* and to develop and evaluate novel therapeutic strategies.

In recent years, xenotransplantation into zebrafish has emerged as an alternative approach to mouse xenografts in cancer research [2–5].

Xenotransplantation of human cancer cells into both adult and embryonic or larval zebrafish has been described [6–8]. Especially, transplantation of tumor cells into zebrafish embryos offers several advantages, most importantly live imaging and easy, cost-effective

compound screening opportunities as transparent embryos/larvae can be maintained in 96-well plate format and small compounds delivered into the fish water are readily absorbed by the embryos/larvae [9].

Several sites have been described for the injection of human tumor cells into zebrafish embryos, including the yolk, the hind-brain ventricle, the brain, the duct of Cuvier, the caudal vein, and the perivitelline space (PVS) [4, 10].

In this chapter, we will provide a general workflow of Ewing sarcoma cell xenotransplantation into zebrafish embryos at 48 h post fertilization (hpf) focusing on PVS injections. The PVS is an intermediate space between the yolk and its surrounding dermal cell layer, which offers enough space to easily introduce around 50 to 200 tumor cells at 48 hpf. The cells will be initially confined between the yolk and dermis, allowing for the investigation of tumor cell proliferation and/or migration and metastasis [3].

In this protocol, we describe the xenotransplantation of shSK-E17T Ewing sarcoma cells into transgenic zebrafish embryos with labeled vasculature [11]. We show that shSK-E17T cells survive and proliferate in the zebrafish host forming a tumor cell cluster at the injection site. Furthermore, blood vessels are attracted to the human Ewing sarcoma cells. We also describe how tumor growth can be followed and quantified over time using a high content imaging system.

Taken together, we provide a protocol for xenografting Ewing sarcoma cells into the perivitelline space of zebrafish embryos and we show possible options for image acquisition and analysis of zebrafish xenografts, which can be applied to quantify the effects of genetic or chemical perturbation of human Ewing sarcoma cells in zebrafish.

2 Materials

2.1 Cell Preparation

1. shSKE-17T Ewing sarcoma cells [11].
2. RPMI 1640 Medium.
3. Fetal bovine serum (FBS).
4. Penicillin–streptomycin (Pen/Strep) (10,000 U/ml).
5. RPMI supplemented with 10% FBS and 1% Pen/Strep.
6. RPMI supplemented with 2% FBS.
7. Dulbecco's phosphate buffered saline (PBS).
8. 1× PBS supplemented with 2% FBS.
9. Accutase Cell Detachment Solution (*see Note 1*).
10. Cell counting device, for example a Neubauer chamber or any automated cell counter.

11. 5 ml polystyrene round bottom tubes with cell strainer cap.
12. CellTracker™ CM-DiI C7001 (Molecular Probes Inc.).
13. Bovine pancreatic deoxyribonuclease I (DNase I) (*see Note 2*).
14. MgCl₂ or any other solution that contains bivalent cations (*see Note 3*).

2.2 Fish Breeding and Maintenance

1. Transparent zebrafish strains of choice (e.g., mitfa^{b692}/^{b692}; ednrba^{b140}/^{b140}).
2. Transparent zebrafish with labeled vasculature (e.g., tg(kdrl:mCherry-CAAX)^{y171} x mitfa^{b692}/^{b692}, ednrba^{b140}/^{b140}).
3. Fish tanks and a system for fish maintenance (e.g., Tecniplast).
4. Breeding tanks (e.g., Tecniplast).
5. Petri dishes (94 mm × 16 mm).
6. E3 embryo medium (*see Note 4*).
7. 1× PTU (To prepare a 25× PTU stock solution dissolve 0.152 g of 1-Phenyl-2-thiourea (PTU) in 200 ml double-distilled water (DDW). To prepare a 1× working solution dilute 1 ml of the 25× stock solution with 24 ml of E3 embryo medium.) (*see Note 5*).

2.3 Xenografting

1. Glass capillaries (Borosilicate glass with fire polished ends without filament; e.g., 0.78/1.00/80 mm GB100T-8P from Science Products).
2. Micropipette needle puller (e.g., Sutter Instruments, Model P-97).
3. Micromanipulator (e.g., World Precision Instruments, Model M3301R).
4. Microinjector (e.g., Eppendorf Femto Jet 4i).
5. Stereo microscope (e.g., Leica M80).
6. Tricaine stock solution 25×, used at 1× (Ethyl 3-aminobenzoate methanesulfonate).
7. E3 embryo medium.
8. Petri dish covered with 2% agarose (To prepare 100 ml of 2% agarose dissolve 2 g agarose in 100 ml E3 embryo medium).

2.4 Imaging

2.4.1 Manual Imaging

1. Fluorescence microscope (e.g., Zeiss Axio Zoom V16).
2. Camera (e.g., AxioCam 503 Color Zeiss).
3. Confocal microscope (e.g., Leica SP8 X WLL).
4. Adobe Photoshop CS6.

2.4.2 Automated Imaging/Quantification

1. Operetta CLS (PerkinElmer).
2. Harmony Software (PerkinElmer).
3. 96-well ZF plates (Hashimoto Electronic Industry Co).

2.5 Immunostaining

1. 4% PFA (To prepare 100 ml of 4% PFA mix 25 ml of 16% PFA with 75 ml of PBS).
2. 100% MetOH.
3. 1× phosphate buffered saline PBS (To prepare 1 L of 1× PBS dissolve 80 g NaCl, 2 g KCl, 14.4 g Na₂HPO₄·2H₂O, and 2.4 g KH₂PO₄ in 800 ml DDW. Adjust the pH to 7.4 with HCl and bring it up to a total volume of 1 L. Autoclave the stock solution and store the buffer at room temperature.
4. 75% MetOH–25% PBS (To prepare 100 ml mix 75 ml of 100% MetOH with 25 ml of 1× PBS).
5. 50% MetOH–50% PBS (To prepare 100 ml mix 50 ml of 100% MetOH with 50 ml of 1× PBS).
6. 25% MetOH–75% PBS (To prepare 100 ml mix 25 ml of 100% MetOH with 75 ml of 1× PBS).
7. Triton X-100.
8. PBS supplemented with Triton X-100 (PBSX) (To prepare 100 ml add 500 µl of 10% Triton X-100 to 50 ml of 1× PBS).
9. Bovine serum albumin (BSA).
10. Dimethyl sulfoxide (DMSO).
11. PBS supplemented with BSA, DMSO and Triton X-100 (PBDX) (To prepare 100 ml of PBDX dissolve 1 g BSA, 1 ml DMSO, and 0.5 ml 10% Triton X-100 in 50 ml 2× PBS (pH = 7.3) and adjust with DDW to a total volume of 100 ml.
12. Goat serum, normal donor herd (Sigma-Aldrich).
13. Ki-67 (8D5) mouse primary mAb #9449 (Cell signaling Technology, USA).
14. Cleaved Caspase-3 (D175) primary antibody (Cell Signaling Technology, USA).
15. Alexa 568 anti-mouse A-11019 (Invitrogen, USA).
16. Alexa 568 anti-rabbit A-21069 (Invitrogen, USA).
17. Dako Fluorescence mounting medium (Dako, Agilent, USA).
18. 15 µ-Slide 8 Well, uncoated (ibidi GmbH).
19. 1.2% Agarose, Type IX-A, ultralow gelling temperature (Sigma-Aldrich) (To prepare a 1.2% low melting agarose solution, 1.2 g agarose powder should be diluted in 100 ml E3 medium, heated up until dissolved and aliquoted. Aliquots can be stored at around 35 °C for some time.)

3 Methods

3.1 Overview

shSK-E17T Ewing sarcoma cells expressing GFP are injected into the PVS of zebrafish embryos at 2 days post fertilization (dpf) and observed over several consecutive days (*see* Fig. 1). To facilitate monitoring of xenografted cells, zebrafish pigmentation mutants (e.g., *mitfa*^{b692/b692}; *ednrba*^{b140/b140}) can be injected into. Alternatively, wild-type zebrafish can be treated with PTU starting around 22 hpf to prevent pigment formation (*see* Note 5). Several transgenic strains highlighting the zebrafish vasculature are

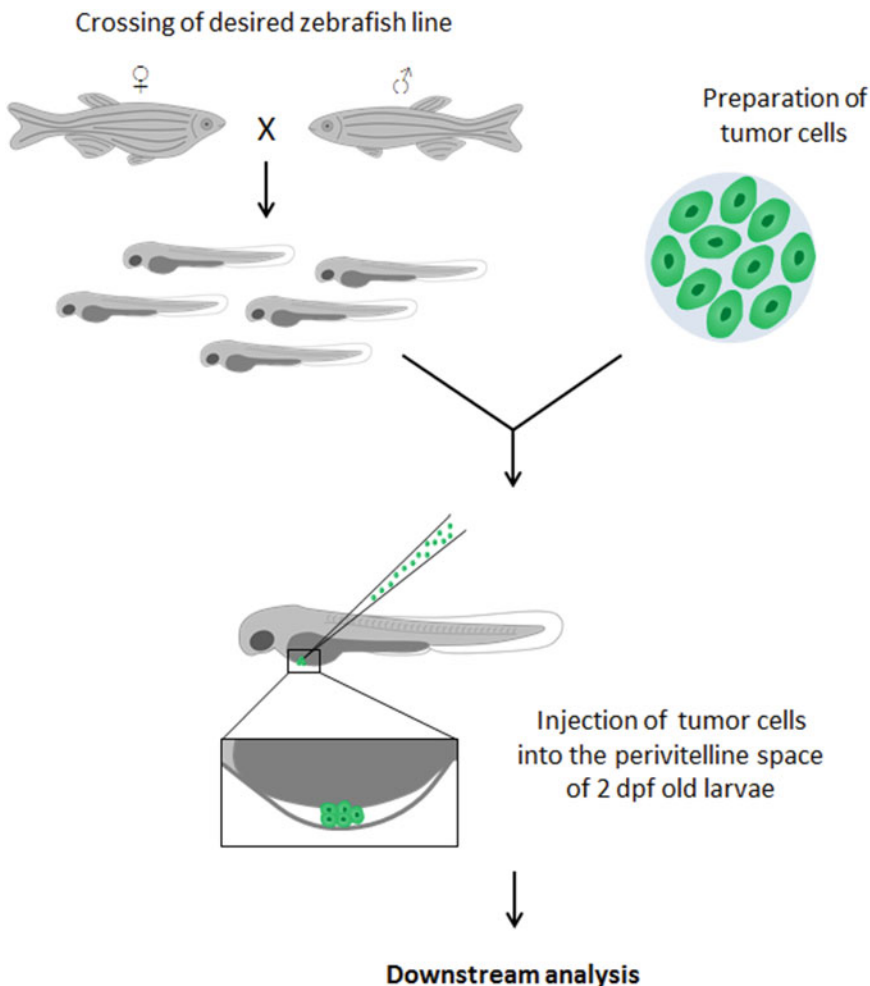


Fig. 1 Workflow of xenografting tumor cells into zebrafish embryos: Schematic representation of the xenotransplantation assay. Zebrafish of selected strains are mated and zebrafish embryos are collected for xenotransplantation. Embryos are kept at 28 °C until 2 dpf. Tumor cells, in this case carrying an intrinsic label (EGFP) are prepared following the protocol described under Subheading 3.2 and are injected into the PVS of zebrafish embryos. Xenografted embryos are kept at 35 °C and monitored. Proliferation and migration can be analyzed and the effect of small compounds can be investigated

available and allow for monitoring neoangiogenesis of transplanted cells or the circulation of human tumor cells within blood vessels and metastasis. After injection of tumor cells, zebrafish embryos are kept at 35 °C (*see Note 6*). Single xenografted embryos/larvae are followed over time, images are taken at desired time points (e.g., once every day) and tumor growth can be quantified. Several xenografted zebrafish larvae are fixed in 4% PFA, for example, around 48 hpi to perform immunostaining for a proliferation marker (e.g., Ki67) and an apoptosis marker (e.g., cleaved Caspase 3) to confirm that human tumor cells are proliferating in zebrafish. Using such zebrafish xenografts, small compounds can be tested to identify potential new therapies.

3.2 Preparing shSK-E17T Ewing Sarcoma Cells for Xenotransplantation

Here, we are applying shSK-E17T Ewing sarcoma cells for xenotransplantation into zebrafish embryos [11]. To be able to monitor the human tumor cells in the living zebrafish, it is necessary to incorporate a label, which can be followed, for example by fluorescence microscopy. The shSK-E17T Ewing sarcoma cells express EGFP. In case the tumor cells do not carry an intrinsic label, we also describe a labeling strategy using a fluorescent lipophilic dye to highlight the tumor cells (*see Subheading 3.2.2*).

3.2.1 Preparation of shSK-E17T Cell Solution

1. Grow shSK-E17T cells to ~80% confluence in RPMI + 10% FBS.
2. Remove medium from shSK-E17T cells and wash 1 × with 5 ml 1 × PBS. Remove PBS and add desired amount of Accutase into the culture flask (*see Note 7*). Incubate for 2 min at 37 °C.
3. Add 7 ml of RPMI + 10% FBS to the cells (total volume = 10 ml).
4. Count shSK-E17T cells.
5. Centrifuge shSK-E17T cells at 300 × *g* for 5 min at 4 °C to pellet the cells (*see Note 8*).
6. Resuspend shSK-E17T cells in 5 ml PBS + 2% FBS.
7. Put cells through cell strainer.
8. Count shSK-E17T cells.
9. Centrifuge shSK-E17T cells at 300 × *g* for 5 min at 4 °C to pellet the cells.
10. Resuspend shSK-E17T cells to a concentration of 50–100 cells/nl in PBS (without FBS) for subsequent xenotransplantation.

3.2.2 Alternative Labeling Strategy with DiI

In case the desired cell line does not express any fluorescent marker, cells can be labeled with DiI (Cell Tracker™ CM-DiI from Molecular Probes Inc.) (*see Note 9*).

1. Grow shSK-E17T cells to ~80% confluence in RPMI + 10% FBS.
2. Remove medium from shSK-E17T cells and wash 1 × with 5 ml 1 × PBS. Remove PBS and add desired amount of Accutase into the culture flask (*see Note 7*). Incubate for 2 min at 37 °C.
3. Add 7 ml of RPMI + 10% FBS to the cells (total volume = 10 ml).
4. Count shSK-E17T cells.
5. Centrifuge shSK-E17T cells at 300 × *g* for 5 min at 4 °C to pellet the cells (*see Note 8*).
6. Resuspend shSK-E17T cells in serum-free RPMI to a concentration of 1 × 10⁶ to 1 × 10⁷ cells/ml.
7. Add 2 µl/ml Cell Tracker™ CM-DiI to the cell suspension.
8. Incubate shSK-E17T cells with CM-DiI for 5 min at 37 °C in the dark.
9. Incubate shSK-E17T cells with CM-DiI for another 15 min on ice in the dark.
10. Wash labelled shSK-E17T cells twice by centrifugation at 300 × *g* for 5 min and resuspension in RPMI + 2% FBS to remove residual CM-DiI.
11. After the second wash step, resuspend shSK-E17T cells in PBS + 2% FBS to a concentration of 1 × 10⁶ to 1 × 10⁷ cells/ml.
12. Add MgCl₂ to a final concentration of 2 mM and DNase I to a final concentration of 50 µg/ml to the cell suspension (*see Note 3*).
13. Incubate shSK-E17T cells with DNaseI and MgCl₂ for 20 min at 37 °C.
14. Wash DNase I treated shSK-E17T cells twice by centrifugation and resuspension in PBS + 2% FBS, in order to remove all the residual DNaseI.
15. Count shSK-E17T cells.
16. Centrifuge shSK-E17T cells at 300 × *g* for 5 min at 4 °C to pellet the cells.
17. Resuspend shSK-E17T cells to a concentration of 50–100 cells/nl PBS (without FBS) for xenotransplantation.

3.3 Preparing Zebrafish Embryos for Xenografting

1. Set up multiple crosses of desired zebrafish strains in breeding tanks 4 days before the planned xenotransplantation experiment.
2. Fill the petri dish lids with 2% agarose and leave it to solidify at room temperature (RT).

3. Subsequently add 3 parallel lines of agarose on top of the solidified agarose (*see* **Note 10**).
4. Harvest eggs on the following morning and keep them in a petri dish with approximately 20 ml E3 medium at 28 °C.
5. Remove unfertilized eggs in the evening of the same day.
6. Remove remaining chorions of zebrafish embryos at 2 dpf using forceps before xenotransplantation.
7. Put zebrafish embryos into 1× Tricaine in E3 for sedation.
8. Align zebrafish embryos for injection by placing them in the agarose-covered petri dish lids along the agarose ridges with their dorsal side toward the ridges.

3.4 Xenografting shSK-E17T Ewing sarcoma Cells into 2 dpf Zebrafish Embryos

1. Place the petri dish with the aligned embryos under a stereomicroscope.
2. Backload an injection capillary with around 5 µl of tumor cell suspension.
3. Insert the injection needle into a needle holder mounted on a micromanipulator.
4. Break the needle open using forceps.
5. Adjust the injection pressure in a way that cells come out of the needle slowly, when pressing the pedal.
6. Introduce the needle at the perivitelline space (PVS) and release the cell suspension (50–100 cells/nl) into the PVS avoiding injections directly into the circulation.
7. Collect and place the injected zebrafish embryos in an incubator at 35 °C.
8. Around 1–2 h post injection (hpi) check injected embryos for injection into circulation using a fluorescent stereomicroscope and select the zebrafish embryos that present cells just in the PVS, but not in circulation. This step is only necessary for migration assays.
9. Maintain transplanted zebrafish in E3 embryo medium at 35 °C during the experiment.

3.5 Analysis of Ewing Sarcoma Xenografts

3.5.1 Manual Imaging of Xenotransplanted Zebrafish Larvae

This protocol is suitable for following each single larva over consecutive days and to acquire images every day (*see* Fig. 2).

1. Anesthetize the transplanted zebrafish larvae in 1× Tricaine and place them with the injection side up in the petri dish.
2. Acquire images of the larvae using, for example, a fluorescence stereomicroscope and put them back into a 12-well plate containing E3 embryo medium (1 larva per well).
3. Keep larvae at 35 °C after imaging.

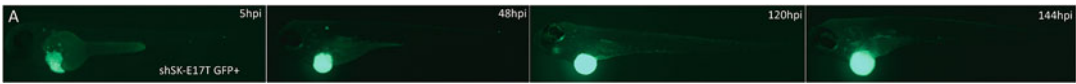


Fig. 2 Zebrafish embryos injected at 2 dpf with Ewing sarcoma cell line shSK-E17T. A single zebrafish embryo was followed over time after xenotransplantation of shSK-E17T Ewing sarcoma cells expressing GFP. Images were taken at 5, 48, 120, and 144 hpi. A rearrangement of tumor cells can be observed between 5 and 48 hpi. The tumor mass grows over time

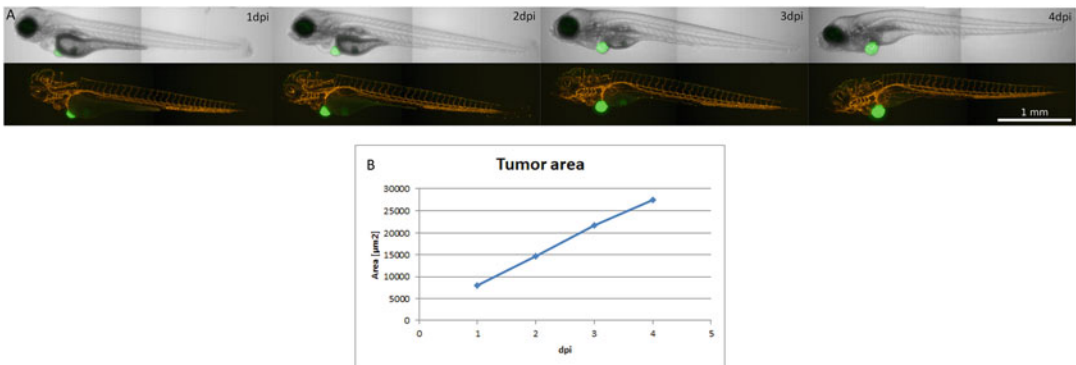


Fig. 3 Automated imaging and image analysis of zebrafish xenografts. (a) Transgenic zebrafish (*kdrl:mCherry-CAAX/mitfa^{b692/b692}; ednrba^{b140/b140}*) xenotransplanted with GFP+ shSK-E17T cells imaged over 4 consecutive days using a high content imaging system (Operetta CLS, PerkinElmer). (b) Tumor growth based on the area covered with fluorescence was quantified using the Harmony software (PerkinElmer)

3.5.2 Automated Image Acquisition of Xenotransplanted Zebrafish Larvae Using a High Content Imager

To record images of xenotransplanted zebrafish larvae at higher throughput, we make use of a high content imaging system (e.g., Operetta CLS, PerkinElmer) for automated image acquisition.

1. Anesthetize the zebrafish larvae in $1\times$ Tricaine and transfer them into ZF 96 well plates.
2. Images are acquired for 4 consecutive days once per day, which allows you to quantify tumor cell proliferation over time (Fig. 3). The quantification reveals an increase in tumor size over time based on the area covered with fluorescence. Quantification is done using the Harmony software (PerkinElmer).

3.5.3 Immunostaining and Confocal Analysis of Xenotransplanted Zebrafish Larvae

In order to confirm that shSK-E17T cells are proliferating after being transplanted into zebrafish embryos and being kept at $35\text{ }^{\circ}\text{C}$, immunostaining for the proliferation marker Ki67 is performed. In addition, an apoptosis marker (e.g., cleaved caspase 3) can be used to quantify apoptotic cells (*see* Fig. 4).

1. Xenografted zebrafish larvae are fixed in PFA (4%) at $4\text{ }^{\circ}\text{C}$ O/N.
2. 4% PFA is replaced for 100% MetOH and fixed larvae can be stored at $-20\text{ }^{\circ}\text{C}$ until immunostaining.

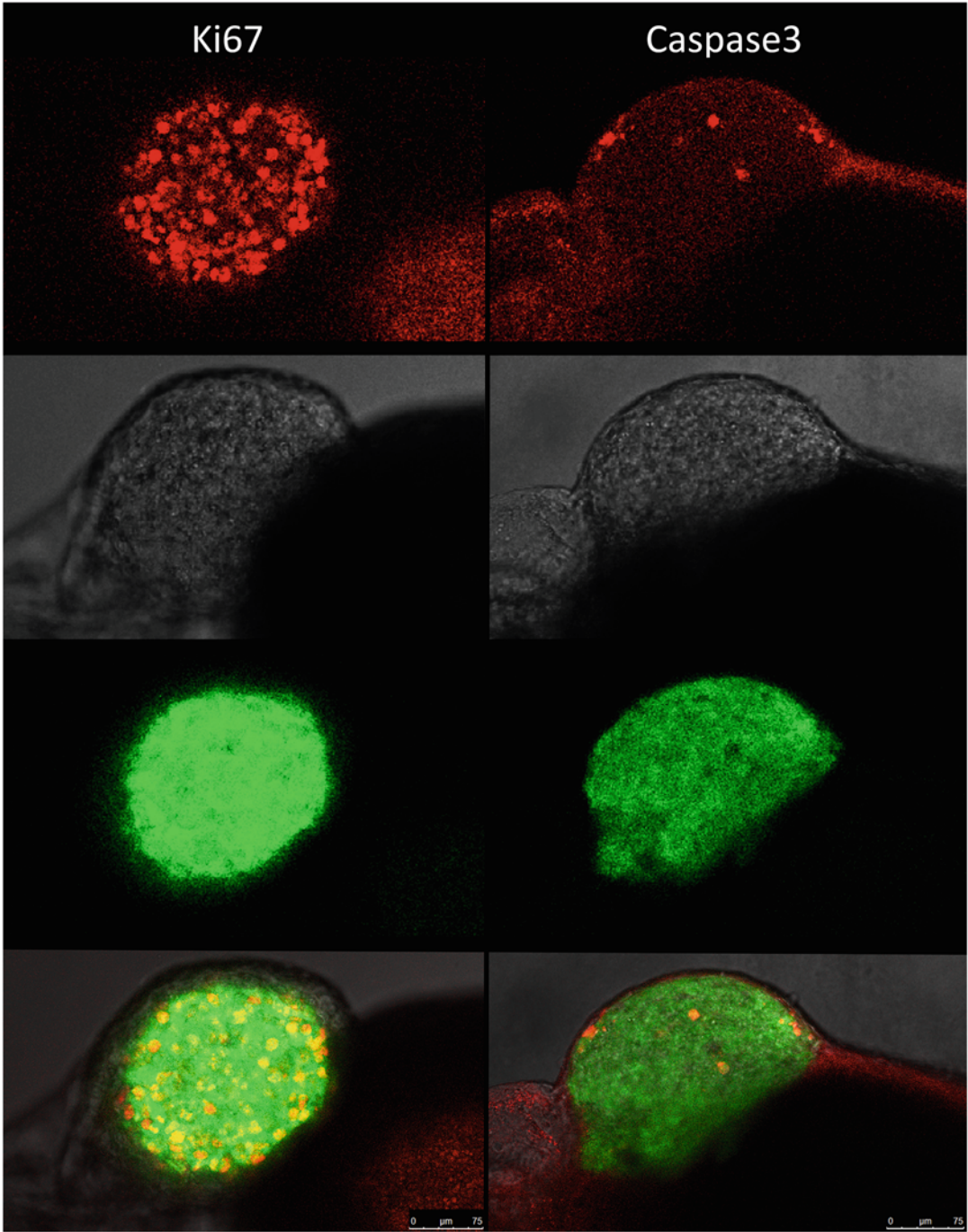


Fig. 4 Immunostaining of a 48 hpi zebrafish larva injected with shSK-E17T cells left column: Immunostaining with an antibody recognizing the proliferation marker Ki67 (red) or right column: cleaved Caspase 3 (red)

3. If embryos are stored in MetOH put them gradually in PBS by rinsing them 10 min each in 75% MetOH–25% PBS, 50% MetOH–50% PBS, 25% MetOH–75% PBS, and 100% PBS.
4. Wash 2× with PBSX for 5 min.
5. Wash once in distilled water.
6. Incubate in –20 °C acetone for 7 min.
7. Block for 1 h in PBDX supplemented with 15 µl of goat serum (GS) per ml of PBDX.
8. Wash 3× in PBDX for 15 min.
9. Dilute the primary antibody in PBDX plus 15 µl GS per ml PBDX and incubate over night at 4 °C (Ki67, 1:400; Caspase 3 1:100).
10. Wash 4× in PBDX for 30 min.
11. Dilute secondary antibody (1:500) in PBDX with 15 µl GS per ml PBDX and incubate for 1 h at RT (keep embryos protected from light from this point!).
12. Wash 2× in PBDX for 10 min.
13. Wash in PBS for 30 min.
14. Put the larvae in 4% PFA for 5 min.
15. Wash 3× in PBS for 5 min.
16. Store the larvae in Fluorescence Mounting Medium (e.g., Dako) at 4 °C.
17. Mount the larvae in 1.2% low melting agarose.
18. Image immunostained larvae using a confocal microscope.

4 Notes

1. We find using Accutase to be preferred over standard trypsin or mechanical detachment, since it is gentler and cells treated with Accutase show improved viability compared to other detachment methods.
2. We are using bovine pancreatic deoxyribonuclease I (DNase I) purchased from Merck. The DNase I digestion step is performed only in the context of DiI labelling, since the labelling procedure can lead to an increase in cell death. Dying cells release DNA, which leads to clumping of the cells. DNase I is used to digest DNA to avoid clumping of cells.
3. DNase I requires bivalent cations for maximum activity. We are using Mg²⁺ ions at a final concentration of 2 mM by adding MgCl₂ solution (50 mM) purchased from New England Biolabs to the PBS before adding DNase I.

4. To prepare 1 L of 50× E3 stock solution weigh 14.61 g NaCl, 0.63 g KCl, 1.83 g CaCl₂·2H₂O, and 1.99 g MgSO₄·7H₂O into a glass bottle and bring it up to 1 L with double distilled water (DDW). Adjust pH to 7.2 with NaOH and autoclave the stock solution.
5. The use of 1-phenyl-2-thiourea (PTU) to keep zebrafish larvae transparent should be considered carefully in the context of xenotransplantations, as it may have an effect on the tumor cells.
6. Since zebrafish are developing best at 28–29 °C, whereas human tumor cells prefer 37 °C, the xenotransplanted zebrafish larvae are kept at 35 °C, as compromise between the two temperatures. There is the possibility that some human cells behave differently at 35 °C compared to 37 °C. Therefore, some preliminary in vitro experiments, like comparing growth rates, and proliferation and apoptotic behavior, should ideally be performed at 35 °C.
7. The Accutase should cover the bottom of the culture flask completely. For a T-185 culture flask, we use 3 ml of Accutase, for smaller flasks volumes can be decreased.
8. We are performing all centrifugation steps in 15 ml conical centrifuge tubes.
9. Caution: When using DiI, please be aware that cell fragments are also labeled with the lipophilic dye. Thus, not all fluorescent signal might actually come from intact human cancer cells.
10. Plates can be kept at 4 °C for several days until use. Before starting the injection protocol, let the plates warm up.

Acknowledgments

We would like to thank Lisa Bierbaumer and Heinrich Kovar for providing shSK-E17T cells. This work received funding from the Austrian Research Promotion Agency (FFG) under project 7940628.

References

1. Minas TZ, Surdez D, Javaheri T, Tanaka M, Howarth M, Kang HJ, Han J, Han ZY, Sax B, Kream BE, Hong SH, Celik H, Tirode F, Tuckermann J, Toretsky JA, Kenner L, Kovar H, Lee S, Sweet-Cordero EA, Nakamura T, Moriggl R, Delattre O, Uren A (2017) Combined experience of six independent laboratories attempting to create an Ewing sarcoma mouse model. *Oncotarget* 8 (21):34141–34163. <https://doi.org/10.18632/oncotarget.9388>
2. Lee LM, Seftor EA, Bonde G, Cornell RA, Hendrix MJ (2005) The fate of human malignant melanoma cells transplanted into zebrafish embryos: assessment of migration and cell division in the absence of tumor formation. *Dev Dyn* 233(4):1560–1570. <https://doi.org/10.1002/dvdy.20471>

3. Nicoli S, Presta M (2007) The zebrafish/tumor xenograft angiogenesis assay. *Nat Protoc* 2(11):2918–2923. <https://doi.org/10.1038/nprot.2007.412>
4. Veinotte CJ, Delleire G, Berman JN (2014) Hooking the big one: the potential of zebrafish xenotransplantation to reform cancer drug screening in the genomic era. *Dis Model Mech* 7(7):745–754. <https://doi.org/10.1242/dmm.015784>
5. Barriuso J, Nagaraju R, Hurlstone A (2015) Zebrafish: a new companion for translational research in oncology. *Clin Cancer Res* 21(5):969–975. <https://doi.org/10.1158/1078-0432.CCR-14-2921>
6. Smith AC, Raimondi AR, Salthouse CD, Ignatius MS, Blackburn JS, Mizgirev IV, Storer NY, de Jong JL, Chen AT, Zhou Y, Revskoy S, Zon LI, Langenau DM (2010) High-throughput cell transplantation establishes that tumor-initiating cells are abundant in zebrafish T-cell acute lymphoblastic leukemia. *Blood* 115(16):3296–3303. <https://doi.org/10.1182/blood-2009-10-246488>
7. Ignatius MS, Chen E, Elpek NM, Fuller AZ, Tenente IM, Clagg R, Liu S, Blackburn JS, Linardic CM, Rosenberg AE, Nielsen PG, Mempel TR, Langenau DM (2012) In vivo imaging of tumor-propagating cells, regional tumor heterogeneity, and dynamic cell movements in embryonal rhabdomyosarcoma. *Cancer Cell* 21(5):680–693. <https://doi.org/10.1016/j.ccr.2012.03.043>
8. Konantz M, Balci TB, Hartwig UF, Delleire G, Andre MC, Berman JN, Lengerke C (2012) Zebrafish xenografts as a tool for in vivo studies on human cancer. *Ann N Y Acad Sci* 1266:124–137. <https://doi.org/10.1111/j.1749-6632.2012.06575.x>
9. Rennekamp AJ, Peterson RT (2015) 15 years of zebrafish chemical screening. *Curr Opin Chem Biol* 24:58–70. <https://doi.org/10.1016/j.cbpa.2014.10.025>
10. Fior R, Povoá V, Mendes RV, Carvalho T, Gomes A, Figueiredo N, Ferreira MG (2017) Single-cell functional and chemosensitive profiling of combinatorial colorectal therapy in zebrafish xenografts. *Proc Natl Acad Sci U S A* 114(39):E8234–E8243. <https://doi.org/10.1073/pnas.1618389114>
11. Franzetti GA, Laud-Duval K, van der Ent W, Brisac A, Irondele M, Aubert S, Dirksen U, Bouvier C, de Pinieux G, Snaar-Jagalska E, Chavrier P, Delattre O (2017) Cell-to-cell heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in Ewing sarcoma cells. *Oncogene* 36(25):3505–3514. <https://doi.org/10.1038/onc.2016.498>

Part V

Bioinformatic Approaches to Ewing Sarcoma



Main Repositories and Databases Used in Ewing Sarcoma Research

Maria Debiec-Rychter

Abstract

Within sarcomas 50 different histological subtypes exist, each with their own molecular and clinical characteristics. The combination of tumor subtype heterogeneity and often a limited number of clinical cases make detailed molecular sarcoma studies challenging, particularly when focusing on individual cohorts. However, the increasing number of publicly available genomics data opens inroads to overcome this obstacle. The international public repositories for high-throughput microarray and next-generation sequence functional genomic data sets submitted by the research community create resources that are freely available for download in a variety of formats. Here, we describe the selected web resources for sarcoma genomics research. These resources support archiving of raw data, processed data, and metadata which are indexed, cross-linked, and searchable.

Key words Genomic data, Databases, Repositories, Portals, Cancer genomics

Recent advances in technologies for high-throughput genome analysis, such as microarray-based methods and next-generation sequencing (NGS), have enhanced progress in the field of oncogenomics [1]. These tools were fundamental for the initiation and development of multicentered cancer genomic projects, such as (a) the Wellcome Trust Sanger Institute's Cancer Genome Project (CGP) [2], (b) The Cancer Genome Atlas (TCGA) [3], and (c) the International Cancer Genome Consortium (ICGC) projects [4]. These projects have been launched for genome-wide analyses of genetic, epigenetic, transcriptomic, and proteomic alterations in hundreds or even thousands of cancer samples. Their general aim was to provide publicly available genomic datasets for the better understanding of the cancer molecular mechanisms and for the assessment of the influence of specific alterations on clinical phenotypes. Consequently to their foundation and development, a number of cancer genomic web sites and portals were created to assist with accessing the abundant cancer datasets.

Hereby, we describe selected popular and effective web-based cancer genomics data repositories.

1 The Gene Expression Omnibus (GEO)

The Gene Expression Omnibus is NCBI gene expression and hybridization array data repository from the National Center for Biotechnology Information [5]. It delivers multiple individual data sets, including gene expression, noncoding RNA, chromatin immunoprecipitation (ChIP) and genome methylation profiling by microarray or NGS, high-throughput RT-PCR, genome variation profiling by arrayCGH, SNP arrays, serial analysis of gene expression (SAGE), and protein arrays from the curated DataSets in GEO repository. Of note, it enables search for specific profiles of interest based on gene annotation or precomputed profile characteristics. In addition, GEO also provides several web-based tools and strategies to assist users to query, analyze, and visualize data, including the release of GEO2R, an R-based web application (to identify and visualize differential gene expression). For submission GEO requires raw data, processed data, and metadata.

2 ArrayExpress

Over the past decade, high-throughput gene expression experiments have generated data from millions of assays. Data sets linked to publications are stored in functional genomics data archive ArrayExpress at the European Bioinformatics Institute Gene Expression Omnibus (<http://www.ebi.ac.uk/arrayexpress>), at the US National Center for Biotechnology Information and at the DNA Databank of Japan Omics Archive [6]. The reproducibility of published microarray-based studies is frequently limited, mostly owing to insufficient experiment annotation and sometimes to unavailability of the raw or processed data. In ArrayExpress, a stricter enforcement of Minimum Information About a Microarray Experiment (MIAME) requirements and also development of easy-to-use experiment annotation tools are implemented to achieve a better reproducibility.

3 cBioPortal

The cBioPortal was developed at the Memorial Sloan Kettering Cancer Center in New York City, NY USA [7]. It is an open source platform for exploring, visualizing, and analyzing multidimensional cancer genomics and clinical data. The public instance of the cBioPortal hosts more than 200 cancer genomics studies, including all

of the data from TCGA [8]. This portal contains genomic data, including copy number alterations, mRNA and microRNA expression, DNA methylation, and protein and phosphoprotein quantity, which were obtained for multiple types of cancer. Its biologist-friendly interface provides many rich analysis features, including a graphical summary of gene-level data across multiple platforms, correlation analysis between genes or other data types, survival analysis, and per-patient data visualization. Cancer-associated alterations deposited in the cBioPortal can be browsed as the overview of all of the genomic events that were detected in an individual cancer sample, alterations in a specific gene across all of the samples that were included in one study, and a comparison of the frequency of the alterations in a given gene across all studies (Fig. 1). For each study, it is also possible to inquire which genes are most frequently altered in the analyzed set of samples. The advantage of the tools that are available in the BioProfiling.de portal is that all of them provide results that are supported by appropriate statistical analysis (the R statistical package), which is not always available for the tools in the other portals. A false discovery rate control procedure is implemented to adjust the p -values when there is multiple testing.

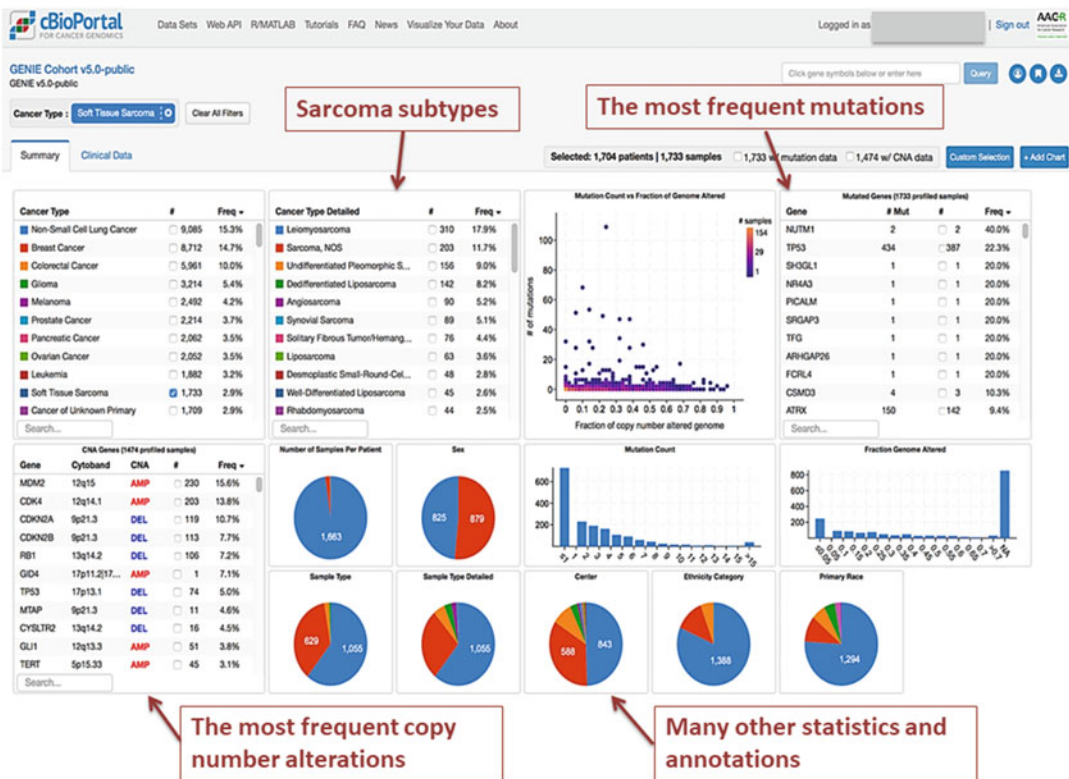


Fig. 1 Soft tissue sarcoma data overview at cBioPortal

A wide range of tools that are available makes the portal useful in various types of analyses, which has resulted in its popularity and applicability. For example, successful application of the cBioPortal in Ewing sarcoma (ES) associated research is described in the study of Liu and coworkers [9]. The objective of this study was to determine the utility of mutational burden in predicting outcomes in patients with localized ES. Clinical and genomic data from 99 patients with ES, of whom 63 had localized disease at diagnosis, were obtained from the cBioPortal for Cancer Genomics. Genomic data included the type and number of somatic mutations using cBioPortal mutation calling. Primary endpoints were overall survival and the time to progression. The findings suggested that the somatic mutation burden can be used to better risk stratify ES patients and to guide clinical decision-making.

4 OncoKB

OncoKB is a comprehensive precision oncology knowledge base that offers oncologists detailed, evidence-based information about individual somatic mutations and structural alterations present in patient tumors to support optimal treatment decisions. As for January 2018, over 3000 unique mutations, fusions, and copy number alterations in 418 cancer-associated genes have been annotated. OncoKB annotations are available through a public web resource [10] and are also incorporated into the cBioPortal for Cancer Genomics to facilitate the interpretation of genomic alterations by physicians and researchers [11].

5 The Cancer Genome Atlas (TCGA) Research Network

Initiated in 2005 by the National Cancer Institute and National Human Genome Research Institute to catalog genetic mutations causing cancer, using genome sequencing was primary focused on glioblastoma multiforme, lung and ovarian cancer [3]. In Phase II of development expanded to 20–25 different cancer types and complemented genome sequencing with broad genomic characterization, including gene expression profiling, copy number variations, mutations (NGS), DNA methylation profiling, mRNA and miRNA expression, pathology and medical images, and treatment and survival outcomes [12]. Raw and processed TCGA sarcoma data are available through NCI Genomic Data Commons [13].

6 AACR GENIE

Aiming to advance precision medicine in oncology and improve patient care, the American Association for Cancer Research has launched an international initiative known as AACR Project Genomics, Evidence, Neoplasia, Information, Exchange (GENIE) [14]. The overall goal of Project GENIE is to create a large-scale high-quality cancer genomics database, widely accessible to the global research community, in order to catalyze precision medicine research efforts and improve patient outcome. The project pools existing NGS data with longitudinal clinical outcomes and related pathology reports to find new mutations, assess potential biomarkers, and identify patient populations that might benefit from existing treatments, offering unparalleled insights into the applicability of cancer genomic profiling at a routine population level. It should be acknowledged that the GENIE project focuses on targeted cancer gene panel data, rather than whole-exome or whole-genome profiling. The first set of cancer genomic data available through GENIE was available in January 2017. The last update of GENIE 6.0-public took place in July 2019. The combined data set now contains over 80 major cancer types, including over 1750 sarcoma samples [15]. AACR GENIE Data is currently available via two mechanisms: Sage Synapse Platform [16] and cBioPortal for Cancer Genomics.

7 FusionHub

The chromosomal rearrangement event play a significant role in cancer due to the oncogenic potential of the chimeric protein generated through fusions. FusionHub is a unified web platform for large scale annotation and visualization of gene fusion events in human cancer [17]. FusionDatabase is updated regularly in every 3 months. The web server is freely available at <https://fusionhub.persistent.co.in>.

It is expected that the utilization of the websites for the analysis of expanding cancer genomic data will make a substantial contribution to our understanding of cancer molecular etiology and the translation of extended cancer genomic knowledge into clinical practice.

References

1. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458:719–724
2. Cancer Genome Project. <https://www.sanger.ac.uk/research/projects/cancergenome/>
3. The Cancer Genome Atlas. <http://cancergenome.nih.gov/>
4. Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, Bernabe RR et al (2010) International

- network of cancer genome projects. *Nature* 464:993–998
5. Clough E, Barrett T (2016) The gene expression omnibus database. *Methods Mol Biol* 1418:93–110
 6. Rung J, Brazma A (2013) Reuse of public genome-wide gene expression data. *Nat Rev Genet* 14:89–99
 7. cBioPortal. <http://www.cbioportal.org>
 8. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO et al (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6:11
 9. Liu KX, Lamba N, Hwang WL, Niemierko A, DuBois SG, Haas-Kogan DA (2019) Risk stratification by somatic mutation burden in Ewing sarcoma. *Cancer* 15:1357–1364
 10. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H (2017) Onco KB: a precision oncology Knowledge Base. *JCO Precis Oncol*. <https://doi.org/10.1200/PO.17.00011>. Epub 2017 May
 11. Cerami E, Gao J, Dogrusoz U, Dresdner G, Gross B, Sumer SO et al (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2:401–404
 12. Deng M, Brgelmann J, Schultze JL, Perner S (2016) Web-TCGA: an online platform for integrated analysis of molecular cancer data sets. *BMC Bioinformatics* 17:72
 13. NCI Genomic Data Commons. <https://portal.gdc.cancer.gov>
 14. AACR Project GENIE Consortium (2017) AACR project GENIE: powering precision medicine through an international consortium. *Cancer Discov* 7:818–831
 15. Micheel CM, Sweeney SM, LeNoue-Newton ML, André F, Bedard PL, Guinney J, AACR project GENIE consortium et al (2018) American Association for Cancer Research Project Genomics Evidence Neoplasia Information Exchange: From Inception to First Data Release and Beyond-Lessons Learned and Member Institutions' Perspectives. *JCO Clin Cancer Inform* 2:1–14
 16. Sage Synapse Platform. <http://synapse.org/genie>
 17. Panigrahi P, Jere A, Anamika K (2018) Fusion-Hub: a unified web platform for annotation and visualization of gene fusion events in human cancer. *PLoS One* 13:e0196588



Chromatin Immunoprecipitation Followed by Next-Generation Sequencing (ChIP-Seq) Analysis in Ewing Sarcoma

Gwenneg Kerdivel and Valentina Boeva

Abstract

ChIP-seq is the method of choice for profiling protein–DNA interactions, and notably for characterizing the landscape of transcription factor binding and histone modifications. This technique has been widely used to study numerous aspects of tumor biology and led to the development of several promising cancer therapies. In Ewing sarcoma research, ChIP-seq provided important insights into the mechanism of action of the major oncogenic fusion protein EWSR1-FLI1 and related epigenetic and transcriptional changes. In this chapter, we provide a detailed pipeline to analyze ChIP-seq experiments from the preprocessing of raw data to tertiary analysis of detected binding sites. We also advise on best practice to prepare tumor samples prior to sequencing.

Key words ChIP-seq, Binding sites, Transcription factors, Histone modifications, Motif analysis, Ewing sarcoma

1 Introduction

1.1 *ChIP-Seq: A Powerful Tool to Study Protein–DNA Interactions*

Protein–DNA interactions are implicated in numerous biological processes including chromatin organization, transcription, and DNA replication. In eukaryotes, genetic material is packaged into chromatin representing a complex of DNA, RNA, and proteins [1]. The primary units of chromatin are called nucleosomes and are composed of 147 bp of DNA wrapped around an octamer of proteins called histones. Histones possess a protein core that interacts with DNA and a ‘tail’ which contains dozens of amino acids that can be posttranslationally modified [2]. These modifications, among which are methylation and acetylation of lysine residues, are generally reversible; and the overall histone modification landscape is the result of action of many enzymes (chromatin writers and erasers). Histone modifications modulate the level of chromatin compaction and are linked to distinct transcriptional states of genes. They are recognized by specific DNA-binding proteins

(chromatin readers) [3]. For instance, acetylation of lysine K27 of histone H3 (H3K27ac) is the result of an interplay between histone acetyltransferases (HATs) and deacetylases (HDACs); it is mainly located in promoters and enhancers of active genes and is read by bromodomain-containing transcriptional co-activators Brd2 and Brd4 [4]. In contrast, promoters of silent genes are often marked by the repressive H3K27me3 modification, which is deposited by the Polycomb Repressive Complex 2 (PRC2) and read by chromodomain-containing proteins such as CBX2,4,6,7,8 forming mutually exclusive PRC1 complexes [5].

Transcription is tightly regulated by transcription factors (TFs) that bind DNA on specific locations (binding sites) within regulatory regions, such as promoters and enhancers. Upon binding, TFs recruit coactivator or corepressor protein complexes including those with histone acetyltransferase or histone deacetylase activity, thus regulating the recruitment and activity of the transcription-initiation complex and RNA polymerase [6]. Due to possessing DNA-binding protein domains, TFs are able to recognize certain DNA sequences, also called motifs. TF binding specificities lead to establishment of regulatory networks where each TF regulates expression of a distinct set of genes. Similarly to the regulation of DNA transcription, control of the DNA replication requires interaction of specific DNA-binding proteins with DNA, allowing for the recruitment of the prereplication complex for further DNA strand separation and DNA synthesis [7].

A method of choice to study protein–DNA interactions is chromatin immunoprecipitation (ChIP) introduced in the 1980s and since then being continuously improved [8–10]. This technique relies on the usage of antibodies to pull down specific protein–DNA complexes. The purified DNA can then be analyzed in a loci-specific manner by real-time PCR (ChIP-qPCR) or genome-wide by sequencing (ChIP-seq), allowing for identification of genomic loci bound by the protein of interest *in vitro* or *in vivo*. ChIP can be used in most of the organisms and applied to any DNA-binding protein for which an efficient and specific antibody exists. Examples of DNA-associated proteins, whose direct and, to a certain extent, indirect binding to DNA can be profiled with ChIP, include TFs, their cofactors, histones carrying or not carrying specific posttranslational modifications, proteins of the chromatin machinery, and replication factors. A ChIP extension, Re-ChIP, allows for detection of regions that are simultaneously cobound by two proteins of interest [11].

ChIP has a complex experimental setup with a number of critical steps. First, cells are fixed, most of the time with formaldehyde, to cross-link all proteins bound to DNA. Chromatin is then extracted and sheared into small DNA/protein fragments; this is achieved by sonication or enzymatic digestion with micrococcal nuclease. A portion of the fragmented DNA is kept as a control, named input, and the rest is used to perform immunoprecipitation

(IP). Briefly, an antibody specific to the protein of interest is added to the DNA/proteins complexes and captured by Protein G/A agarose or Protein G/A magnetic beads. After several steps of pull-down/washing, cross-links are reversed and the immunoprecipitated DNA is purified. These DNA samples are ready to be analyzed by qPCR or sequencing.

1.2 Applications of ChIP-seq in Cancer Research

ChIP-seq has been widely used to study mechanisms of tumorigenesis, determine new targets and link the genetic variation to changes in chromatin states. In fact, cancer cells exhibit very altered gene expression profiles as compared to their healthy counterparts. These transcriptomic differences are often induced by deregulation of TF activities associated with genomic variants or chromatin remodeling. In particular, mutations [12], fusions [13], and amplifications [14] in genes coding for chromatin regulators, histones, or histone variants have been documented to lead to cancer-related chromatin alterations.

The role of alterations of chromatin landscape in tumor initiation and progression has been extensively studied in many cancer types. For instance, in neuroblastoma, analysis of superenhancers, characterized by high level of H3K27ac, has identified two specific subgroups of cells with distinct identities and revealed a specific superenhancer of MYC, a well-known oncogene, responsible for high expression of this TF in the absence of gene amplification [15]. In gliomas, loss of H3K27me3 has been shown to be associated with aberrant expression of several oncogenes and TFs, while H3K27me3 gain was demonstrated to cause silencing of certain oncosuppressors [16]. The enzyme EZH2, a member of the PRC2 complex that deposits H3K27me3, has been shown to be deregulated in many cancers; its increased expression was negatively associated with patient survival [17, 18]. These and other studies of chromatin changes in cancer led to the identification and use in clinics of several epigenetic drugs [19]. Further research attempting to decipher the role of epigenetic landscape in carcinogenesis is ongoing.

Analysis of transcription factor binding in cancer using ChIP-seq also provided important insights. For example, T-cell acute lymphocytic leukemia recurrently exhibits activating mutation of NOTCH1. NOTCH1 has been shown to bind an enhancer downstream of MYC gene, driving high expression of this infamous oncogene [20]. CTCF binding has also been intensively studied in cancer. Indeed, CTCF, through its binding in insulator regions, participates in the creation of boundaries of Topologically Associating Domains (TADs). Several studies have demonstrated that alteration of CTCF binding can disrupt TADs, resulting in ectopic enhancer-promoter interactions [21].

Overall, ChIP-seq provided important information on chromatin changes and transcription factor binding in numerous cancer

studies and allowed for the discovery of several new therapeutic targets. This technology has been also applied to study different aspects of Ewing sarcoma (EwS) detailed in the next section.

**1.3 ChIP-seq
Contributed to
Important
Breakthrough in Ewing
Sarcoma**

Ewing sarcoma is characterized by the expression of a fusion gene *EWSR1-FLI1* (EF1) or, more rarely, a fusion of *EWSR1* and another member of the ETS TF family. The EF1 fusion gene, resulting from the t(11,22) chromosomal translocation, codes for a TF that acts as a major driver of oncogenic transformation through activation and repression of thousands of target genes [22]. ChIP-seq for this transcription factor has been performed in various models of ES cells; several strategies have been applied for EF1 immunoprecipitation: use of an anti-FLAG antibody to pull-down ectopically expressed flagged version of EF1 or employment of an antibody targeting endogenous FLI1 as well as the fusion protein [23]. ChIP-seq experiments for EF1 provided crucial information about the mechanisms of action and the roles of this TF. In particular, it has been shown that EF1 binds in the proximal promoter and intronic regions of different specific targets, such as *CDKN1A*, *TNC*, *TGFBR2*, *MYC*, *AURKA* and *AURKB*, *CCND1*, *IGFBP3*, *STYXLI*, *GLI1*, and *EZH2* [24]. EF1 also regulates a large number of genes through binding to distant enhancers located in intergenic regions [25].

ChIP-seq experiments in Ewing sarcoma have also shown that EF1 binds to GGAA microsatellites, preferentially constituted of tracts of 9 repeats or more [25]. However, in promoters or enhancer-like regions, binding of EF1 does not have the same impact on gene modulation. In the promoter, the number of repeats correlates with EF1 binding and gene activation while in enhancer-like regions, the number of repeats only correlates with EF1 binding affinity but not with gene activation. In addition, the number of repeats has not been shown to be associated with gene repression [26].

ChIP-seq analysis of histone modification profiles surrounding EF1 binding sites identified that GGAA repeats bound by EF1 exhibit characteristics of enhancers with an enrichment in H3K4me1 and H3K27ac, as it can be observed, for instance, in a region 180 kb upstream of the *CALCB* gene, highly overexpressed in ES [27]. Interestingly, analysis of these histone marks confirmed that EF1 can act as a pioneer transcription factor as it is able to create de novo enhancers through accumulation of H3K4me1 and H3K27ac [28, 29]. Recently, Lin et al. performed H3K27ac in order to analyze superenhancers (SEs) in EwS. Their results show that in ES cells SEs are significantly enriched in active EF1 binding motifs [30]. They also showed that EF1 regulates the expression of the TF MEIS1 through binding to an SE. This TF plays a pro-survival role in EwS by cooperating with EF1 to regulate the expression of another SE-associated gene, *APCDD1*.

In contrast to the activatory function of EF1 associated with GGAA microsatellite binding, EF1 was shown to repress genes when binding to conserved enhancers exhibiting canonical ETS motifs. The proposed repression mechanism was based on the displacement of more active wild-type ETS TFs, such as ELF1 and GABPA, and consecutive decrease in p300 recruitment and H3K27ac deposition [29]. EF1 is also demonstrated to be able to regulate cellular plasticity, mainly by repressing genes involved in cytoskeletal reorganization [31]. This repression was shown to result from shared genomic binding regions between EF1 and the transcription factor MRTFB disrupting the MRTFB/YAP-1/TEAD transcriptional module. This disruption leads to the suppression of the Rho-actin pathway and results in the repression of a large number of structural and regulatory cytoskeletal genes [32].

Johnson et al. performed knockdown/rescue experiments for EF1 in A673 Ewing sarcoma cells, where the rescue was done with a wild-type (WT) or mutant forms of EF1. ChIP-seq experiments targeting the WT or mutated EF1 showed that a part of EWS located within the amino acids 1–82 and 246–264 of the EWS portion of EF1 was necessary for DNA binding of the fusion protein [33]. Before these ChIP-seq experiments were performed, it was assumed that the DNA-binding property of EF1 was solely carried by the FLI1 DNA binding domain while the EWS part was necessary for transcriptional activity and oncogenic transformation.

The interactions between EF1 and some other transcription factors were also investigated using the ChIP-seq technique. Bilke et al. showed that EF1 proteins have a tendency to form clusters, for example, with 15 binding locations near the *DLGAP1* gene [34]. Furthermore, more than 50% of the binding sites of E2F3, another transcription factor, were shown to overlap with EF1 binding, mainly in proximal promoters of genes upregulated by EF1, leading to the hypothesis that EF1 could recruit E2F3 on the promoter of its target genes.

ChIP-seq has been also used to decipher mechanisms by which the anticancer drug trabectedin represses EF1 activity. Recent work has shown that trabectedin induces a relocalization of EF1 within the nucleus as well as an eviction of the SWI/SNF chromatin-remodeling complex, leading to an increased accumulation in H3K27me3 and H3K9me3 histones marks. This increase was preferentially observed around GGAA microsatellites and EF1 target genes [35].

In this chapter, we provide a framework to analyze ChIP-seq experiments (*see* Fig. 1 for an overview of a usual ChIP-seq analysis pipeline). We describe how to go from the raw ChIP-seq reads to the identification of binding sites (peak calling) and show how to accomplish downstream analysis steps, some of which have been already performed in Ewing sarcoma and allowed for gaining knowledge on the genetics and epigenetics of this disease.

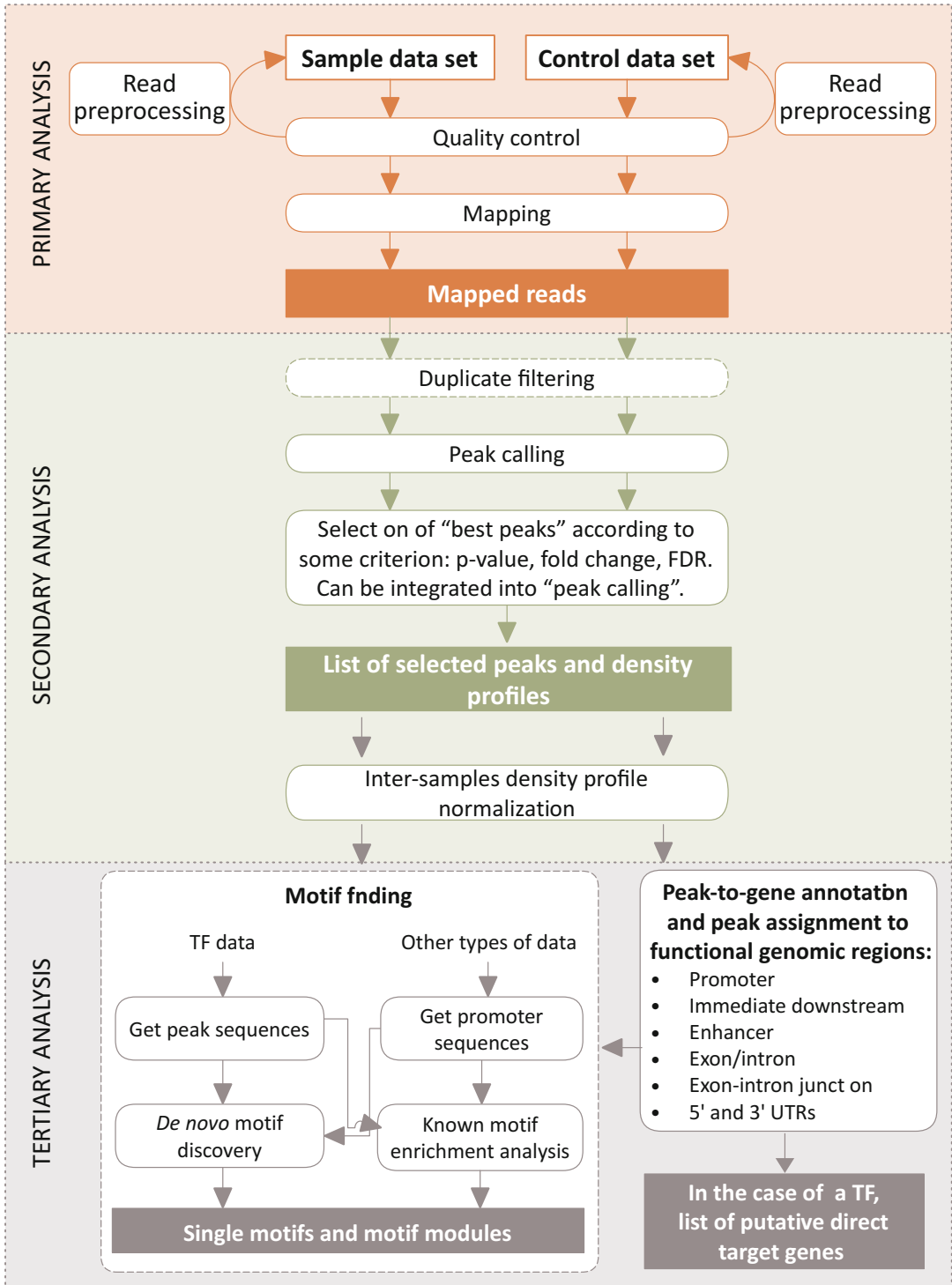


Fig. 1 Outline of a typical ChIP-seq analysis. Primary analysis includes read preprocessing and mapping; low-quality reads can be filtered out based on mapping quality scores. Secondary analysis consists in peak calling and signal postprocessing; duplicate reads can be kept or be filtered out. Our example of tertiary analyses includes de novo motif discovery and/or motif enrichment analysis, functional annotation of the peaks, and identification of putative targets of TFs

2 Materials

2.1 Computing Requirements

The workflow can be performed on a standard UNIX or MacOS desktop or laptop with a minimum of 8Gb of RAM. More details on computing requirements can be found in **Note 1**.

2.2 Software Installation

1. Quality control with FASTQC: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
2. Reads preprocessing: *see Note 2* for a nonexhaustive list of tools that can be used to remove low-quality bases prior to read mapping and perform other preprocessing steps.
3. R Statistical Programming Language [36]: <https://www.r-project.org>.
4. R package libraries (Bioconductor):
 - (a) rtracklayer (<https://bioconductor.org/packages/release/bioc/html/rtracklayer.html>).
 - (b) plyr (<https://cran.r-project.org/web/packages/plyr/>).
5. HMCAN [37]: <https://github.com/BoevaLab/HMCAN>
6. BWA [38]: <http://bio-bwa.sourceforge.net/>
7. Samtools [39]: <http://samtools.sourceforge.net/>
8. CHIPIN: <https://github.com/BoevaLab/CHIPIN> (additional R libraries are required for CHIPIN installation).
9. LILY [15]: <http://boevalab.inf.ethz.ch/LILY/>
10. Integrative Genomics Viewer (IGV) [40]: <https://software.broadinstitute.org/software/igv/>
11. ChomHMM [41]: <http://compbio.mit.edu/ChromHMM/>
12. UCSC tools:
 - (a) liftOver (http://hgdownload.cse.ucsc.edu/admin/exe/linux.x86_64/liftOver),
 - (b) wigToBigWig (http://hgdownload.cse.ucsc.edu/admin/exe/linux.x86_64/wigToBigWig),
 - (c) bedGraphToBigWig: (http://hgdownload.cse.ucsc.edu/admin/exe/linux.x86_64/bedGraphToBigWig).
 - (d) chain file (<http://hgdownload.cse.ucsc.edu/goldenPath/hg18/liftOver/>).

2.3 Data

1. ChIP-seq reads (sample files), usually in .FASTQ format *see Note 3*.
2. Blacklist file [42] (auxiliary file) *see Note 4*.
3. Reference genome build in .FASTA format as one file (for human, usually hg19.fa or hg38.fa, referred as REF.fa in this

- protocol) and as separate files for each chromosome (<https://hgdownload.soe.ucsc.edu/downloads.html>) (auxiliary files).
4. A file with chromosome lengths, REF.sizes, containing two columns: chromosome name and size in base pairs (<https://hgdownload.soe.ucsc.edu/downloads.html>) (auxiliary file).
 5. A file containing GC content and mappability scores for large regions of DNA in CNP format (included with HMCAN) (auxiliary file).

3 Methods

3.1 *Sample Collection and Processing*

To obtain a strong ChIP-seq signal as a result of peak calling, DNA samples provided for sequencing should be of high quality. To ensure sample quality, one should pay attention to the collection of samples and the chromatin immunoprecipitation (ChIP) step. ChIP can be performed either using commercially available kits (Diagenode, Abcam, and Active motif) or using well-described noncommercial protocols [43, 44]. The purpose of this chapter is to emphasize several critical steps of the ChIP library preparation essential to obtain good quality data.

Chromatin shearing: Chromatin should be sheared in ~150–300 bp fragments before performing the ChIP. This size corresponds to mono- or dinucleosome chromatin fragments, allowing for a high resolution for the detection of binding sites. The distribution of the fragment sizes should not vary a lot between the samples. Indeed, the shape of the fragment size distribution determines the resolution of binding sites. Ensuring similar peak resolution is important to be able to compare several samples.

Antibody: The choice of the antibody is one of the most important factors affecting the quality of ChIP-seq data [45]. Antibodies have to offer both high sensitivity and high specificity. Indeed, the more sensitive the antibody is, the less background noise will be obtained in the final result. The use of highly specific antibodies allows for being confident that the detected peaks correspond to binding sites of the protein of interest. Unfortunately, commercially available “ChIP grade” or even “ChIP-seq grade” antibodies are not all appropriate to obtain high-quality results as it has been shown by several studies [46, 47]. In addition, the specificity may vary between lots for the same antibody. Finally, the choice of monoclonal antibody, which may reduce the background noise, or polyclonal one, can also be important as some epitopes may be masked on some loci depending on the chromatin context and protein–protein interactions. Thus, it is essential to carefully choose and test the antibodies, especially when no prior test information is available [48].

Tagged protein: Epitope-tagged (e.g., GFP, FLAG-tag, and HA-tag) experiments can be used in cells that do not express the target protein (or in which the expression of the target has been knocked out and a tagged version of this protein has been reintroduced). This approach can be useful to discriminate multiple similar TF paralogs or isoforms or to immunoprecipitate protein for which no ChIP-seq grade antibody is available. In the case of Ewing sarcoma, it can be used to specifically target tagged-version of the fusion protein EWS/FLI1 in cells that also express FLI1 and/or EWS [49, 50].

Sequencing library: Even if several methods exist to estimate the fragment size distribution from single-ends data, paired-end reads allow to obtain the exact size of each fragment, therefore increasing the accuracy of the obtained results [37, 51]. In addition, paired-end reads improve the library complexity and increase mapping efficiency, especially, in repetitive regions [52].

3.2 Quality Control (QC) and Read Filtering

As the first step after sequencing, one should ensure the good quality of reads. In this protocol, we assume that the ChIP-seq has been performed in paired-end mode and corresponding read files are named READ1.fastq and READ2.fastq. Quality controls can be performed with the FASTQC tool using the following command line:

```
$ fastqc READ1.fastq
$ fastqc READ2.fastq
```

Details about the output and the different metrics investigated by the FASTQC tool can be found on its website (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Particular attention should be paid to the base quality calls and the number of duplicates, which reflect the overall success of sequencing and the complexity of the initial DNA library. One should also check for a potential presence of over-represented k-mers as it can reveal a bad trimming of the sequencing adapters.

At this step, no preprocessing or read filtering are usually necessary since low-quality reads will be filtered during and after the mapping step. However, if the reads are of really poor quality, one could preprocess the reads using one of the tools cited in **Note 1**. Examples of how to use such tools can be found in other reviews [53, 54].

3.3 Mapping Using BWA

In this protocol, we show how to map reads using the BWA method but other mappers can also be considered (*see Note 5*). As variant calling is seldom performed on ChIP-seq data, the choice of the read mapping tool is relatively unimportant.

1. [Optionally] Before starting, a BWA index of the reference genome required to map the reads should be created. This

step should only be performed once for each reference genome used. Process as follows:

```
$ bwa index REF.fa
```

This step will output 5 files next to the REF.fa file: REF.fa.ann, REF.fa.pac, REF.fa.amb, REF.fa.bwt, [REF.fa.sa](#)

2. Mapping the reads to the indexed genome may then be performed using the command “**bwa mem**”. In most cases, default parameters can be used. However, to speed up the calculations, several threads can be used for the mapping using the “-t” option (8 in the following example). Default output is in the uncompressed .SAM format. It is useful to convert it directly to the binary format .BAM to save disk space. This can be done by directly piping the result of “**bwa mem**” to “**samtools view**” as follows:

```
$ bwa mem -t 8 <PATH_TO_BWA_INDEXES>/hg38.fa <PATH_TO_FASTQ>/
READ1.fastq <PATH_TO_FASTQ>/READ2.fastq | samtools view -b >
<PATH_TO_OUPUT_DIR>/RESULT.bam
```

The -b option in “samtools” makes the output to be written in the .BAM format. It is also possible to increase the number of threads using the -@ option. The final output will be written in <PATH_TO_OUPUT_DIR>/RESULT.bam

3. The .BAM files resulting from the previous command line contain all mapped reads including low-quality mappings. These low-quality mappings consist of reads aligned to multiple genomic locations and reads containing many low-quality bases or mismatches. This filtering of low-quality mappings can be done using the tool “**samtools view**.” Additionally, a filtering of so-called duplicate reads is often performed. Duplicate reads constitute read pairs mapping to the exactly same genomic locations. Unfortunately, it is usually not possible to discriminate between “natural” duplicates and duplicates that arise from multiple PCR products of the same DNA template molecule. Keeping duplicates can lead to false positive peak calls, but the downside of duplicate removal is the inability to call peaks in the regions of the genomic amplification in cancer. To remove duplicate reads, one can apply “**samtools markdup**.” The following command lines exemplify these steps:

```
$ samtools view -u -q 20 <PATH_TO_OUPUT_DIR>/RESULT.bam >
<PATH_TO_OUPUT_DIR>/RESULT.q20.bam
```

```
$ samtools sort -@ 8 -m 5G -n -o <PATH_TO_OUPUT_DIR>/RESULT.
q20.nsrt.bam
<PATH_TO_OUPUT_DIR>/RESULT.q20.bam
```

```
$ samtools fixmate -@ 8 -m <PATH_TO_OUPUT_DIR>/RESULT.q20.
nsrt.bam
```

```

<PATH_TO_OUPUT_DIR>/RESULT.q20.nsrt.fixmate.bam

$ samtools sort -@ 8 -m 5G -o
<PATH_TO_OUPUT_DIR>/RESULT.q20.nsrt.fixmate.coordsrt.bam
<PATH_TO_OUPUT_DIR>/RESULT.q20.nsrt.fixmate.bam

$ samtools markdup -@ 8 -r
<PATH_TO_OUPUT_DIR>/RESULT.q20.nsrt.fixmate.coordsrt.bam
<PATH_TO_OUPUT_DIR>/RESULT.q20.noDup.bam

```

Note that, using these command lines, all intermediate files will be kept, which can drastically increase the disk space used.

Line 1: “-q 20” allows “**samtools view**” to write to the output only the mappings with quality of at least 20. “-u” allows for writing of uncompressed .BAM file saving time on compression/decompression. To write a compressed .BAM file, one can use “-b” with “-t” option setting the number of threads used for the compression.

Line 2: In order to use “**samtools markdup**” to remove the duplicates, “**samtools fixmate**” should be used to fill in mate coordinates. To do so, the .BAM file should be name-sorted (“-n”). The “-o” option specifies the output file. The “-m” option is used to specify the amount of memory per thread to be used for the sorting process. In this example, 8 threads are used for a total amount of RAM of 40G.

Line 3: “**samtools fixmate**” is then performed. The -m option is used to add the mate score tags that will be use by “**samtools markdup**” to keep the best reads.

Line 4: The .BAM file needs to be be coordinate-sorted before removing the duplicates (default behavior of “**samtools sort**”).

Line 5: “**samtools markdup**” can be used to only mark duplicates or to remove them; for the latter, one should use the “-r” option. “**samtools markdup**” can only be applied to a coordinate-sorted .BAM file.

In order to be visualized using any genome browser, for instance IGV, .BAM files have to be indexed using “**samtools index**”:

```
$ samtools index <PATH_TO_OUPUT_DIR>/RESULT.q20.noDup.bam
```

3.4 Peak Calling Using HMcAn

The next analysis step is the detection of binding sites based on mapped reads (peak calling). In this protocol, we recommend to use the HMcAn method [37], which was specifically developed for the analysis of cancer ChIP-seq data. HMcAn corrects the ChIP-seq signal for the GC-content bias that results from PRC amplification of DNA libraries and removes the bias arising from copy number alterations inherent to cancer genomes. Both IP and Input mapped reads (filtered .BAM files obtained at the previous

step) should be provided to HMCAN to get the best peak calling accuracy; however, the tool can technically run without a control sample. The following command lines assume that the command “**hmcAN**” is in your PATH and the .BAM files are named IP.bam and Input.bam:

```
$ hmcAN <PATH_TO_BAM>/IP.bam <PATH_TO_BAM>/Input.bam config.txt
<PATH_TO_OUPUT_DIR>/NAME > NAME.log 2> NAME.err.log
```

<PATH_TO_OUPUT_DIR>/NAME: Prefix to be used for the output files, for example, name of the sample.

All information sent to standard output or standard error will be written in NAME.log and NAME.err.log respectively.

The configuration file (config.txt) should look like below:
format bam.

```
GCIndex <PATH>/GC_profile_25KbWindow_Mapp50_hg19.cnp
genomePath <PATH>/hg19/chromosomes/
pairedEnds True
smallBinLength 50
largeBinLength 25000
pvalueThreshold 0.05
mergeDistance 3000
GCmergeDistance 1000
iterationThreshold 1
finalThreshold 0.1
maxIter 20
PosteriorProb 0.75
PrintWig True
PrintPosterior False
blackListFile <PATH>/hg19-blacklist.bed
CallPeaks True
PrintBedGraph False
RemoveDuplicates False
calculateEmpiricalPvalue True
```

HMCAN will output three main files: one for narrow peaks corresponding to individual binding sites, one file with larger regions overall enriched in signal, and the whole genome signal density profile (as a .WIG file). Output files in the plain text .WIG format are usually compressed into the binary .BIGWIG format example:

```
$ wigToBigWig NAME_narrowPeaks.wig hg19.size NAME_narrowPeaks.
bw
```

3.5 *Between-Sample Normalization of ChIP-seq Data*

Commonly, normalization of ChIP-seq signal is performed using the total number of fragments mapped per sample. However, this method does not take into account variation in the signal-to-noise ratio observed in different samples due to natural variability in antibody efficiency. To overcome this issue when comparing signals in several conditions, a number of methods have been developed to perform between-sample normalization [15, 55–58]. These methods are based, for example, on a weighted regression approach or linear MA-fit using peaks common to all conditions.

In this protocol, we provide an example of normalization implemented in the R package ChIPIN (<https://github.com/BoevaLab/CHIPIN/>). Using ChIPIN, the user can apply either a quantile normalization or a normalization based on a linear regression with nonzero intercept. The method assumes that regulatory regions of genes that have constant expression levels between conditions (further called “constant genes”) have matching ChIP-seq intensities. The user can apply ChIPIN when (1) the above assumption is supposed to be true and (2) the user can provide either gene expression profiles for each sample (RNA-seq or microarray) or a list of “constant genes” in .BED format.

Further, we assume that the user has installed the ChIPIN package and has to normalize ChIP-seq signal for three samples S1, S2, and S3 (.BIGWIG files); here we use “H3K27ac” as the experiment’s name. To estimate “constant genes,” we assume we have raw read count RNA-seq data (RNAseq_RAW.tsv). Then, we launch **R**, define paths to sample files and proceed to the normalization of the data:

```
$ R
> pathToGeneExpressionFile = "<PATH_TO_INTPUT_DIR>/RNAseq_RAW.tsv"
> pathToBigWigFiles = paste0("<PATH_TO_INTPUT_DIR>/", c("S1.H3K27ac.bw", "S2.H3K27ac.bw", "S3.H3K27ac.bw"))
> outputFolder = "<PATH_TO_OUPUT_DIR>"
> histoneMarkName = "H3K27Ac"
> sampleName = <NAME>
> CHIPIN_normalize(path_to_bw=pathToBigWigFiles, type_norm="quantile", raw_read_count=pathToGeneExpressionFile, sample_name=sampleName, output_dir=outputFolder, organism="hg19", expression_plot=TRUE, compute_stat=TRUE, percentage=0.1, nGroup=20, histone_mark=histoneMarkName)
```

Several parameters can be modified, for instance beforeRegionStartLength, afterRegionStartLength, regionBodyLength, binSize are parameters of the “computeMatrix” function from the “deepTools” package called by ChIPIN. Modifying it allows the user to give more or less weight to regions upstream and downstream of genes for performing normalization.

ChIPIN creates several output files:

- NAME_constant_genes.bed: contains a list of regions representing “constant genes” used for normalization;
- NAME_CPMmeanVSsd.pdf: biplot showing the distribution of average expression values and standard deviation of expression across samples (“constant genes” are highlighted in red);
- normalized .BIGWIG files;
- StatsAfter.txt and StatsBefore.txt: text files with statistics reflecting the success of the normalization procedure;
- After_Normalization.pdf and Before_Normalization.pdf: files that allow for visualizing the success of the normalization by showing density profiles around transcription start sites of “constant genes” before and after normalization.
- If the expression_plot parameter is set to TRUE, a .pdf file containing the density profile around the TSS of low, medium and highly expressed genes is created for each sample. This plot is useful to verify the efficiency of the antibody used. For example, it is known that highly expressed genes should show higher density of the histone mark H3K27Ac than lowly expressed ones.

The normalized .BIGWIG density files can then be used for any subsequent analysis that involves a comparison of signal intensities between samples.

3.6 Subsequent Analysis

At this point, the ChIP-seq data have been processed to call peaks and create density profiles; densities profiles can be normalized across samples if needed. Depending on the questions asked, different types of subsequent analysis can be performed. We provide several examples of downstream analysis in this section.

3.6.1 Motif Analysis

Motif enrichment analysis provides numerous biological insights [59]. Applied to a ChIP-seq dataset generated for a TF of interest, it allows for identifying yet unknown TF binding motifs using de novo motif discovery and detecting TFs that cooperate or compete with the TF of interest by scanning peaks for known motifs. In the context of Ewing sarcoma, motif analysis allowed for discovery of two types of binding motifs of EFl (classic ETS family motif and GGAA microsatellite repeat [25]). It also provided insights on EFl binding partners [34].

Motif discovery performed on regions carrying a specific histone mark allows for identifying TFs with an activator or repressive function in a given cell type. For example, it enables discovery of TFs that bind active enhancers and promoters and represent major transcriptional regulators [15].

Numerous tools have been designed to perform motif analysis. Among the most widely used tools, we would name ChIPmunk [60], MEME and MEME-ChIP [61] for de novo motif discovery, i-cisTarget [62], AME and CentriMO from the MEME toolkit [61] for known motif enrichment, and Homer [63] and RSAT [64] for both analysis types. More detail is covered in this review [59].

3.6.2 Annotation of Binding Sites

Binding sites of TFs or regions enriched in histone modifications detected with ChIP-seq usually have to be assigned to genes they potentially regulate. The easiest, although not so accurate way to assign regulatory regions to genes is by choosing the closest transcription start site. This can be done manually using home-made scripts or by using published tools such as ChIPseeker [65], annotatePeaks [66] or ChIPpeakAnno [67]. A more sophisticated method, TIP [68], proposes a probabilistic model that quantitatively measures the regulatory relationships between TFs and target genes. In addition, the commonly used GREAT tool tries to predict functions of noncoding regions by analyzing the annotation of nearby genes [69]. GREAT is easy to use via the website version (<http://great.stanford.edu/public/html/>); the user has to choose a genome assembly and upload both test and background regions. If no background regions are available, it is possible to use the whole genome as background. The association rules can also be modified according to the user's preferences.

3.6.3 Super Enhancer (SE) Calling

SEs are regions of the genome containing multiple enhancers that are bound by a set of transcription factors that determine cell identity. SEs can be identified using ChIP-seq experiments targeting the H3K27ac histone mark [70], mediator protein MED1 [71] or bromodomain protein BRD4 [72, 73]. Here, we will show how to identify SEs using the LILY tool based on H3K27ac profiles [15]. LILY replicates the ROSE algorithm originally developed by the Young lab [71], but in addition to ROSE, the LILY method includes explicit correction for copy number variation inherent to cancer samples. To run LILY based on the HMCAN output, run the following in the command line:

```
$ cat PATHTO_LILY/runLILY.R | R --slave --args SAMPLE
OUTPUT 12500 2500 REF_refseq.ucsc REF.size
```

SAMPLE: prefix for the sample of interest containing the path to the HMCAN output for the corresponding sample.

OUTPUT: path where the results of LILY will be written.

12,500: peak stitching distance (bp).

2500: distance around gene TSSs to annotate gene promoters (bp).

REF.refseq.ucsc: file with transcriptome information.

REF.size: file with chromosome lengths used to convert .WIG to .BIGWIG files.

The output of LILY is a .BED file “SAMPLE.scores.bed”, where SEs are sorted according to their strengths. Ultimately, SE regions can be used to investigate Core Regulatory Circuits using COLTRON [74] or CRCmapper [75] tools.

3.6.4 Conversion Between Genome Builds

In certain situations, one may have to compare ChIP-seq results obtained with one genome build (for example, human genome build hg38) to results obtained with a different build (for example hg18 or hg19). Indeed, the read files (.FASTQ or .BAM) are not always available and thus it may not be possible to realign the reads on the genome build of interest and redo the peak calling. In this case, it is possible to convert genomic coordinates from one build to another using the LiftOver tool. This can be done online (<http://genome-euro.ucsc.edu/cgi-bin/hgLiftOver>) or in the command line as follows:

```
$ liftOver PEAKS.hg18.bed hg18ToHg19.over.chain.gz PEAKS.hg19.
bed unlifted.bed
```

This command converts the coordinates of ChIP-seq peaks obtained on hg18 (.BED file PEAKS.hg18.bed) to coordinates on hg19 (.BED file PEAKS.hg19.bed). The positions that fail to be transformed are stored in the file “unlifted.bed”.

It is also possible to convert Wig/BigWig files from one build to another even though it is a rather complicated procedure, which we do not recommend. In this case, the .BIGWIG file should first be converted into the .BEDGRAPH format using the **bigwigToBedGraph** command, then lifted as previously described to the new genome build and ultimately converted back from .BEDGRAPH to .BIGWIG using **bedGraphToBigWig**. If the density profile is in .WIG format, it should first be converted to .BIGWIG using **wigToBigWig**.

```
$ wigToBigWig PEAKS.hg18.wig hg18.size PEAKS.hg18.bw
$ bigwigToBedGraph PEAKS.hg18.bw PEAKS.hg18.bedGraph
$ liftOver PEAKS.hg18.bedGraph hg18ToHg19.over.chain.gz PEAKS.
hg19.bed unlifted.bed
$ bedGraphToBigWig PEAKS.hg19.bed hg19.size PEAKS.hg19.bw
```

4 Notes

1. *Computing requirements*: The running time can be decreased by using more powerful configurations or working on a high-powered server, especially when dealing with a large number of samples. One should also be aware of the disk space available on the computer for data analysis as the .BAM files generated at the mapping step can be relatively large; moreover, keeping all intermediate analysis files can significantly increase space requirements. Once the peak calling has been performed, all subsequent analysis steps can be performed at a standard desktop or laptop computer. We recommend the Linux operating system to facilitate the use of command line tools and R.
2. *Nonexhaustive list of reads preprocessing tools*: fastx_clipper (http://hannonlab.cshl.edu/fastx_toolkit/), Btrim [76] (<http://graphics.med.yale.edu/trim/>), Trimmomatic [77] (<http://www.usadellab.org/cms/?page=trimmomatic>), AlienTrimmer [78] (<ftp://ftp.pasteur.fr/pub/gensoft/projects/AlienTrimmer/>), Cutadapt (<https://pypi.org/project/cutadapt/>), and AdapterRemoval [79] (<https://github.com/MikkelSchubert/adapterremoval>).
3. *ChIP-seq reads*: Paired-end reads are preferred to single end reads. Each ChIP-seq sample should be paired with a matched control sample: a whole cell extract sample (WCE, or “input”) or a mock ChIP reaction background sample (IgG control).
4. *ENCODE blacklist*: This list represents a set of regions in the human and other genomes that exhibit anomalous, unstructured, or high signal in all sequencing experiments independently of cell line or experiment (<https://github.com/Boyle-Lab/Blacklist/tree/master/lists>).
5. *Nonexhaustive list of read mappers*: BWA [38] (<http://bio-bwa.sourceforge.net/>), bowtie2 [80] (<http://bowtie-bio.sourceforge.net/bowtie2/>), STAR [81] (<https://github.com/alexdobin/STAR>), and NovoAlign (<http://www.novocraft.com/products/novoalign/>).

Acknowledgments

This work was funded through institutional support from Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, ATIP-Avenir, the ARC Foundation (ARC-RAC16002KSA-R15093KS), the SIRIC CARPEM (No. INCA-DGOS-INSERM_12561), and the “Who Am I?” Laboratory of Excellence ANR-11-LABX-0071, funded by the French Government through its Investissement d’Avenir program, operated by the French National Research Agency (ANR-11-IDEX-0005-02).

References

1. Audia JE, Campbell RM (2016) Histone modifications and cancer. *Cold Spring Harb Perspect Biol* 8:a019521
2. Cosgrove MS, Wolberger C (2005) How does the histone code work? *Biochem Cell Biol* 83:468–476
3. Suganuma T, Workman JL (2011) Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* 80:473–499
4. Josling GA, Selvarajah SA, Petter M et al (2012) The role of bromodomain proteins in regulating gene expression. *Genes (Basel)* 3:320–343
5. Connelly KE, Weaver TM, Alpsoy A et al (2019) Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association. *Nucleic Acids Res* 47:2289–2305
6. Reiter F, Wienerroither S, Stark A (2017) Combinatorial function of transcription factors and cofactors. *Curr Opin Genet Dev* 43:73–81
7. Groth A, Rocha W, Verreault A et al (2007) Chromatin challenges during DNA replication and repair. *Cell* 128:721–733
8. Gilmour DS, Lis JT (1984) Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proc Natl Acad Sci U S A* 81:4275–4279
9. Gilmour DS, Lis JT (1985) In vivo interactions of RNA polymerase II with genes of *Drosophila melanogaster*. *Mol Cell Biol* 5:2009–2018
10. Solomon MJ, Larsen PL, Varshavsky A (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell* 53:937–947
11. Beischlag TV, Prefontaine GG, Hankinson O (2018) ChIP-re-ChIP: co-occupancy analysis by sequential chromatin immunoprecipitation. In: Visa N, Jordán-Pla A (eds) *Chromatin immunoprecipitation: methods and protocols*. Springer, New York, NY, pp 103–112
12. Zhang J, Dominguez-Sola D, Hussein S et al (2015) Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med* 21:1190–1198
13. Alekseyenko AA, Walsh EM, Wang X et al (2015) The oncogenic BRD4-NUT chromatin regulator drives aberrant transcription within large topological domains. *Genes Dev* 29:1507–1523
14. Rui L, Emre NCT, Kruhlik MJ et al (2010) Cooperative epigenetic modulation by cancer amplicon genes. *Cancer Cell* 18:590–605
15. Boeva V, Louis-Brennetot C, Peltier A et al (2017) Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat Genet* 49:1408–1413
16. Funato K, Major T, Lewis PW et al (2014) Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* 346:1529–1533
17. Drelon C, Berthon A, Mathieu M et al (2016) EZH2 is overexpressed in adrenocortical carcinoma and is associated with disease progression. *Hum Mol Genet* 25(13):2789–2800
18. Kim KH, Roberts CWM (2016) Targeting EZH2 in cancer. *Nat Med* 22:128–134
19. Kelly AD, Issa J-PJ (2017) The promise of epigenetic therapy: reprogramming the cancer epigenome. *Curr Opin Genet Dev* 42:68–77
20. Yashiro-Ohtani Y, Wang H, Zang C et al (2014) Long-range enhancer activity determines Myc sensitivity to notch inhibitors in T cell leukemia. *Proc Natl Acad Sci U S A* 111: E4946–E4953
21. Lupiáñez DG, Kraft K, Heinrich V et al (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161:1012–1025
22. Braun BS, Frieden R, Lessnick SL et al (1995) Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis. *Mol Cell Biol* 15:4623–4630
23. Sand LGL, Szuhai K, Hogendoorn PCW (2015) Sequencing overview of Ewing sarcoma: a journey across genomic, Epigenomic and transcriptomic landscapes. *Int J Mol Sci* 16:16176–16215
24. Kovar H (2010) Downstream EWS/FLI1—upstream Ewing's sarcoma. *Genome Med* 2:8
25. Guillon N, Tirode F, Boeva V et al (2009) The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PLoS One* 4:e4932
26. Johnson KM, Taslim C, Saund RS et al (2017) Identification of two types of GGAA-microsatellites and their roles in EWS/FLI binding and gene regulation in Ewing sarcoma. *PLoS One* 12:e0186275
27. Dallmayer M, Li J, Ohmura S et al (2019) Targeting the CALCB/RAMP1 axis inhibits growth of Ewing sarcoma. *Cell Death Dis* 10:1–13
28. Grünewald TGP, Bernard V, Gilardi-Hebenstreit P et al (2015) Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene *EGR2* via a GGAA microsatellite. *Nat Genet* 47:1073–1078

29. Riggi N, Knoechel B, Gillespie SM et al (2014) EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* 26:668–681
30. Lin L, Huang M, Shi X et al (2019) Super-enhancer-associated MEIS1 promotes transcriptional dysregulation in Ewing sarcoma in co-operation with EWS-FLI1. *Nucleic Acids Res* 47:1255–1267
31. Chaturvedi A, Hoffman LM, Jensen CC et al (2014) Molecular dissection of the mechanism by which EWS/FLI expression compromises actin cytoskeletal integrity and cell adhesion in Ewing sarcoma. *Mol Biol Cell* 25:2695–2709
32. Katschnig AM, Kauer MO, Schwentner R et al (2017) EWS-FLI1 perturbs MRTFB/YAP-1/TEAD target gene regulation inhibiting cytoskeletal autoregulatory feedback in Ewing sarcoma. *Oncogene* 36:5995–6005
33. Johnson KM, Mahler NR, Saund RS et al (2017) Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. *Proc Natl Acad Sci U S A* 114:9870–9875
34. Bilke S, Schwentner R, Yang F et al (2013) Oncogenic ETS fusions deregulate E2F3 target genes in Ewing sarcoma and prostate cancer. *Genome Res* 23:1797–1809
35. Harlow ML, Chasse MH, Boguslawski EA et al (2019) Trabectedin inhibits EWS-FLI1 and evicts SWI/SNF from chromatin in a schedule-dependent manner. *Clin Cancer Res* 25:3417–3429
36. R Development Core Team (2010) a language and environment for statistical computing: reference index, R Foundation for Statistical Computing, Vienna
37. Ashoor H, Hérault A, Kamoun A et al (2013) HMCAn: a method for detecting chromatin modifications in cancer samples using ChIP-seq data. *Bioinformatics* 29:2979–2986
38. Li H, Durbin R (2009) Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25:1754–1760
39. Li H, Handsaker B, Wysoker A et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079
40. Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genomics viewer. *Nat Biotechnol* 29:24–26
41. Ernst J, Kellis M (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 9:215–216
42. Amemiya HM, Kundaje A, Boyle AP (2019) The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep* 9:1–5
43. Rodríguez-Ubreva J, Ballestar E (2014) Chromatin immunoprecipitation. In: Stockert JC, Espada J, Blázquez-Castro A (eds) *Functional analysis of DNA and chromatin*. Humana Press, Totowa, NJ, pp 309–318
44. Wiehle L, Breiling A (2016) Chromatin immunoprecipitation. In: Lanzuolo C, Bodega B (eds) *Polycomb group proteins: methods and protocols*. Springer, New York, NY, pp 7–21
45. Wardle FC, Tan H (2015) A ChIP on the shoulder? Chromatin immunoprecipitation and validation strategies for ChIP antibodies. *F1000Res* 4:235
46. Egelhofer TA, Minoda A, Klugman S et al (2011) An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* 18:91–93
47. Goens G, Rusu D, Bultot L et al (2009) Characterization and quality control of antibodies used in ChIP assays. *Methods Mol Biol* 567:27–43
48. Landt SG, Marinov GK, Kundaje A et al (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* 22:1813–1831
49. Kinsey M, Smith R, Iyer AK et al (2009) EWS/FLI and its downstream target NROB1 interact directly to modulate transcription and oncogenesis in Ewing’s sarcoma. *Cancer Res* 69:9047–9055
50. Shimizu R, Tanaka M, Tsutsumi S et al (2018) EWS-FLI1 regulates a transcriptional program in cooperation with Foxq1 in mouse Ewing sarcoma. *Cancer Sci* 109:2907–2918
51. Zhang Y, Liu T, Meyer CA et al (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9:R137
52. Chen Y, Negre N, Li Q et al (2012) Systematic evaluation of factors influencing ChIP-seq fidelity. *Nat Methods* 9:609–614
53. Buisine N, Kerdivel G, Sachs LM (2018) De novo transcriptomic approach to study thyroid hormone receptor action in non-mammalian models. In: Plateroti M, Samarut J (eds) *Thyroid hormone nuclear receptor: methods and protocols*. Springer, New York, NY, pp 265–285
54. Davis-Turak J, Courtney SM, Hazard ES et al (2017) Genomics pipelines and data integration: challenges and opportunities in the research setting. *Expert Rev Mol Diagn* 17:225–237
55. Liang K, Keleş S (2012) Normalization of ChIP-seq data with control. *BMC Bioinformatics* 13:199

56. Nair NU, Sahu AD, Bucher P et al (2012) ChIPnorm: a statistical method for normalizing and identifying differential regions in histone modification ChIP-seq libraries. *PLoS One* 7:e39573
57. Taslim C, Wu J, Yan P et al (2009) Comparative study on ChIP-seq data: normalization and binding pattern characterization. *Bioinformatics* 25:2334–2340
58. Shao Z, Zhang Y, Yuan G-C et al (2012) MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol* 13:R16
59. Boeva V (2016) Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in eukaryotic cells. *Front Genet* 7:24
60. Kulakovskiy IV, Boeva VA, Favorov AV et al (2010) Deep and wide digging for binding motifs in ChIP-Seq data. *Bioinformatics* 26:2622–2623
61. McLeay RC, Bailey TL (2010) Motif enrichment analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11:165
62. Herrmann C, Van de Sande B, Potier D et al (2012) I-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. *Nucleic Acids Res* 40:e114
63. Heinz S, Benner C, Spann N et al (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38:576–589
64. Thomas-Chollier M, Sand O, Turatsinze J-V et al (2008) RSAT: regulatory sequence analysis tools. *Nucleic Acids Res* 36:W119–W127
65. Yu G, Wang L-G, He Q-Y (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31:2382–2383
66. Chen X, Xu H, Yuan P et al (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133:1106–1117
67. Zhu LJ, Gazin C, Lawson ND et al (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11:237
68. Cheng C, Min R, Gerstein M (2011) TIP: a probabilistic method for identifying transcription factor target genes from ChIP-seq binding profiles. *Bioinformatics* 27:3221–3227
69. McLean CY, Bristol D, Hiller M et al (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 28:495–501
70. Hnisz D, Abraham BJ, Lee TI et al (2013) Super-enhancers in the control of cell identity and disease. *Cell* 155:934–947
71. Whyte WA, Orlando DA, Hnisz D et al (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153:307–319
72. Lovén J, Hoke HA, Lin CY et al (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153:320–334
73. Chapuy B, McKeown MR, Lin CY et al (2013) Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* 24:777–790
74. Lin CY, Erkek S, Tong Y et al (2016) Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature* 530:57–62
75. Saint-André V, Federation AJ, Lin CY et al (2016) Models of human core transcriptional regulatory circuitries. *Genome Res* 26:385–396
76. Kong Y (2011) Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics* 98:152–153
77. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120
78. Criscuolo A, Brisse S (2013) AlienTrimmer: a tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. *Genomics* 102:500–506
79. Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* 9:88
80. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359
81. Dobin A, Davis CA, Schlesinger F et al (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21



Epigenetic Analysis in Ewing Sarcoma

Jeremy M. Simon and Nicholas C. Gomez

Abstract

Ewing sarcoma is a highly malignant tumor characterized by a chromosomal translocation that modifies the activity of an ETS family transcription factor. The most prevalent translocation product, EWSR1-FLI1, exploits a permissive and unique chromatin environment of stem cells, and transforms them into an oncogenic state through alterations to gene expression and gene regulatory programs. Though the transformation ability of, and subsequent reliance on EWSR1-FLI1 had been previously described, the advent of genome-wide sequencing technologies allowed for the specific identification of genomic loci and genes targeted by EWSR1-FLI1. Furthermore, the characterization of the chromatin environment in these, and other, cell types could not have been accomplished without the computational and statistical methods that enable large-scale data analysis. Here, we outline in detail the tools and steps needed to analyze genome-wide transcription factor binding and histone modification data (chromatin immunoprecipitation, ChIP-seq), as well as chromatin accessibility data (assay for transposase-accessible chromatin, ATAC-seq) from Ewing sarcoma cells. Our protocol includes a compilation of data quality control metrics, trimming of adapter sequences, reference genome alignment, identification of enriched sites (“peaks”) and motifs, as well as annotation and visualization, using real-world data. These steps should provide a platform on which molecular biologists can build their own analytical pipelines to aid in data processing, analysis, and interpretation.

Key words Next-generation sequencing, ChIP-seq, ATAC-Seq, EWSR1-FLI1, H3K27ac, Epigenetics, Bioinformatics, Ewing sarcoma

1 Introduction

Ewing sarcoma is a tumor of the bone and soft tissue that primarily affects children and young adults. It is characterized by chromosomal translocations that involve members of the TET family and the ETS family of transcription factors; 80–85% of tumors exhibit an in-frame fusion of EWSR1 and FLI1 (hereafter EWSR1-FLI1) [1, 2]. Translocations with other ETS genes are detected in most of the remaining tumors, yielding similarly functioning fusion proteins [3, 4]. EWSR1-FLI1 exhibits altered binding properties when compared to its DNA-binding parental protein FLI1 despite identical ETS-family DNA binding domains [5]. Instead of localizing to

canonical ETS motifs [6], EWSR1-FLI1 is driven to a certain class of microsatellite repeats characterized by a multimerization of the GGAA core of the ETS sequence motif [7–9]. Which specific microsatellites and other GGAA-containing sequences in the genome EWSR1-FLI1 localizes to is further dependent on the underlying organization of chromatin, both in tumor cells as well as the putative cell of origin, the mesenchymal stem cell [10].

In eukaryotes, DNA is organized into chromatin through tight but dynamic interactions with histones, together forming a structure known as the nucleosome [11]. Histones can be posttranslationally modified, particularly on their flexible tails [12], remodeled by incorporating noncanonical proteins [13, 14], or repositioned or displaced during DNA-templated processes such as transcription [15–17]. This reorganization is often a necessary precursor to the binding of transcription factors, which typically can only bind to their cognate sites in accessible regions of chromatin [18]. What makes Ewing sarcoma etiology both unique and interesting is that stem cells, particularly mesenchymal stem cells, exhibit chromatin accessibility over repetitive elements including GGAA-containing microsatellites [9, 10, 19]. What this suggests is that upon translocation, EWSR1-FLI1 can coopt the already permissive chromatin environment of stem cells, recognize and bind to GGAA-containing repetitive elements, and then reinforce their role as enhancers, driving oncogenic gene expression [5, 9, 10]. Therefore, chromatin accessibility at critical sites specific to stem cells can facilitate EWSR1-FLI1 targeting [10] and influence downstream gene expression and cancer progression [7, 20–23].

Studying genome-wide chromatin accessibility, transcription factor binding, and localization of histone modifications in depth in multiple cell types relied heavily on the use of high-throughput DNA sequencing technologies [24, 25]. The advent of high-throughput sequencing enabled the development and wide adoption of assays such as chromatin immunoprecipitation (ChIP-seq) [26–28], formaldehyde-assisted isolation of regulatory elements (FAIRE-seq) [29, 30], DNase hypersensitivity (DNase-seq) [31–33], and assay for transposase-accessible chromatin (ATAC-seq) [34, 35]. Each of these experiments yields a billion or more sequenced nucleotides, the fragments of which then need to be assigned to their source coordinates in the reference genome. Then, in order to tell which sequenced fragments represent true transcription factor binding sites, or accessible chromatin regions, or loci with modified histones, an algorithm must separate true enrichment from background signal [36, 37], while minimizing false discoveries. This is broadly true whether studying chromatin or transcription factors from Ewing sarcoma cells or another biological system.

Every one of these steps requires intense computation; however, molecular biologists interested in analyzing their own data can do so by utilizing the example protocol we detail below. We stress

that each of the outlined steps are flexible and modular such that other algorithms can be easily substituted if a given user or lab prefers, and that the best methodology will depend on the underlying dataset and hypotheses.

2 Materials/Applications

This chapter requires the installation of specific applications and functions to work. Due to the specifics and complexities of different operating systems this chapter will not discuss how to install each package but instead the user is referred to the extensive documentation that each of the respective applications provides. The required applications are listed in Tables 1 and 2.

3 Methods

This chapter will utilize the publicly available ATAC-seq and ChIP-seq data generated in from the study performed by Riggi et al. [9]. Specifically, we will process and analyze the data generated from SK-N-MCs, a Ewing sarcoma cell line. The first step is to download the data from the NCBI Gene Expression Omnibus (GEO) using SRA Toolkit. The raw reads will be trimmed of sequencing adapters and processed through FASTQC for quality control. Trimmed reads are then aligned to the human genome (hg38) and regions of enrichment are identified by calling peaks. Peaks are then scanned for statistically significant enrichment of transcription factor motifs and associated with genomic features such as promoters, introns, and exons. We then create a heatmap displaying ATAC and H3K27ac signal at EWSR1-FLI1 binding sites to demonstrate a simple option for integrating different types data sets.

To perform the following commands please use a command-line UNIX/LINUX interface, Terminal for Mac OS X, or Cygwin for Windows.

3.1 Data Acquisition and Preprocessing

1. Download SK-N-MC Data from GEO using SRA toolkit (*see Note 1*).

```
#ATAC-Seq
fastq-dump --split-files SRX718183
#ChIP-seq H3K27ac
fastq-dump --split-files SRX718118
#ChIP-Seq FLI1
fastq-dump --split-files SRX718117
```

Table 1
Required Programs/Applications

Software	Ver.	Website	References
SRA toolkit	2.10.0	https://github.com/ncbi/sra-tools	
Skewer	0.2.2	https://github.com/relipmoc/skewer	[38]
FASTQC	0.11.8	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	[39]
Bowtie2	2.3.5.1	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml	[40]
Samtools	1.9	http://www.htslib.org/	[41]
Samblaster	0.1.24	https://github.com/GregoryFaust/samblaster	[42]
Homer	4.11	http://homer.ucsd.edu/homer/index.html	[43]
R	3.5.1	https://www.r-project.org/	[44]
R studio	1.0.153	https://rstudio.com/	[45]
MACS2	2.2.5	https://github.com/taoliu/MACS	[37]
deepTools	3.3.1	https://deeptools.readthedocs.io/en/develop/	[46]
IGV	2.7.2	http://software.broadinstitute.org/software/igv/	[47]

Table 2
Required Bioconductor Packages—available from [Bioconductor.org](https://www.bioconductor.org)

Package	Ver.	Website	References
Bioconductor	3.8	https://www.bioconductor.org/install/	[48]
ATACseqQC	1.6.4	https://www.bioconductor.org/packages/release/bioc/html/ATACseqQC.html	[49]
Rsamtools	1.34	https://www.bioconductor.org/packages/release/bioc/html/Rsamtools.html	[50]
BSgenome. Hsapiens. UCSC.hg38	1.4.1	https://www.bioconductor.org/packages/release/data/annotation/html/BSgenome.Hsapiens.UCSC.hg38.html	[51]
GenomicRanges	1.34.0	https://www.bioconductor.org/packages/release/bioc/html/GenomicRanges.html	[52]
ChIPpeakAnno	3.16.0	https://www.bioconductor.org/packages/release/bioc/html/ChIPpeakAnno.html	[53]
Rtracklayer	1.42.1	https://www.bioconductor.org/packages/release/bioc/html/rtracklayer.html	[54]

2. Trim Adapter from sequencing reads using Skewer [Optional but recommended] (*see Note 2*).
3. Skewer requires a FASTA file of known adapter sequences. We can create this file from the known adapter sequences obtained from the Illumina support site (*see Note 3*).
4. Run Skewer using 6 threads (-t 6), give the output a new base name (-o SKNMC_ATAC), run in the paired end mode (-m pe), and use the adapter FASTA reference file (-x).

```
skewer -t 6 -o SKNMC_ATAC -m pe -x Skewer_Nextera_adapter_
sequences.fa SRX718183_1.fastq SRX718183_2.fastq
```

5. Trim adapters from H3K27ac ChIP Seq (*see Note 4*).

```
skewer -t 6 -o SKNMC_H3K27ac -m tail -x Skewer_Truseq_
adapters.fa SRX718118_1.fastq
```

6. Trim adapters from FLI1 ChIP Seq.

```
skewer -t 6 -o SKNMC_FLI1 -m tail -x Skewer_Truseq_adapters.fa
SRX718117_1.fastq
```

7. Quality Control using FASTQC [Optional] (*see Note 5*). Each FASTQ file will be processed independently through the FASTQC software (*see Note 6*).

```
#ATAC-Seq
fastqc --noextract --outdir FASTQC_ATAC_1 SKNMC_ATAC-trimmed-
pair1.fastq
fastqc --noextract --outdir FASTQC_ATAC_2 SKNMC_ATAC-trimmed-
pair2.fastq

#H3K27ac ChIP-Seq
fastqc --noextract --outdir FASTQC_H3K27ac SKNMC_H3K27ac-
trimmed.fastq

#FLI1 ChIP-Seq
fastqc --noextract --outdir FASTQC_FLI1 SKNMC_FLI1-trimmed.
fastq
```

3.2 Alignment of Data

1. Align to Human Genome (GRCh38): use bowtie2 to align to the human genome (hg38) (*see Note 7*) setting the options to very-sensitive, maximum fragment length of 2000 (-X 2000) and number of threads to 8 (-p).

```
#ATAC-Seq
bowtie2 -p 8 --very-sensitive -X 2000 --rg-id SKNMC_ATAC -x
Bowtie2Index/genome -1 SKNMC_ATAC-trimmed-pair1.fastq -2
SKNMC_ATAC-trimmed-pair2.fastq -S SKNMC_ATAC.sam

#H3K27ac ChIP-Seq
bowtie2 -p 8 --very-sensitive --rg-id SKNMC_H3K27ac -x Bow-
```

```
tie2Index/genome -U SKNMC_H3K27ac-trimmed.fastq -S
SKNMC_H3K27ac.sam
```

```
#FLI1 ChIP-Seq
```

```
bowtie2 -p 8 --very-sensitive --rg-id SKNMC_FLI1 -x Bowtie2Index/genome -U SKNMC_FLI1-trimmed.fastq -S SKNMC_FLI1.sam
```

2. Sort and remove duplicates from the datasets (*see Note 8*). The ATAC-seq data needs to be first sorted by read-pair and then finally stored as compressed version and indexed (*see Note 9*).

```
#ATAC-Seq
```

```
#Sort by read pair in order to remove duplicates
```

```
samtools sort -n -o SKNMC_ATAC.sortedName.sam SKNMC_ATAC.sam
```

```
#Remove duplicates
```

```
samblaster -i SKNMC_ATAC.sortedName.sam -o
```

```
SKNMC_ATAC.sortedName.DupsRemoved.sam -r
```

```
#Convert to bam and sort by coordinate
```

```
samtools view -bS SKNMC_ATAC.sortedName.DupsRemoved.sam | samtools sort -o SKNMC_ATAC.DupsRemoved.sorted.bam
```

```
#Create the index
```

```
samtools index SKNMC_ATAC.DupsRemoved.sorted.bam
```

```
#H3K27ac ChIP-Seq
```

```
#Sort by coordinate
```

```
samtools view -bS SKNMC_H3K27ac.sam | samtools sort -o SKNMC_H3K27ac.sorted.bam
```

```
#Remove duplicates
```

```
samtools rmdup -s SKNMC_H3K27ac.sorted.bam
```

```
SKNMC_H3K27ac.DupsRemoved.sorted.bam
```

```
#Create the index
```

```
samtools index SKNMC_H3K27ac.DupsRemoved.sorted.bam
```

```
#FLI1 ChIP-Seq
```

```
samtools view -bS SKNMC_FLI1.sam | samtools sort -o SKNMC_FLI1.sorted.bam
```

```
#Remove duplicates
```

```
samtools rmdup -s SKNMC_FLI1.sorted.bam SKNMC_FLI1.DupsRemoved.sorted.bam
```

```
#Create the index
```

```
samtools index SKNMC_FLI1.DupsRemoved.sorted.bam
```

3. Shift ATAC-seq reads to account for transposition event (*see Note 10*).

```
Rscript Shift_Bam_Files.R SKNMC_ATAC.DupsRemoved.sorted.bam
shifted/
```

3.3 Peak Calling

1. The reads are now aligned to the human genome and duplicate reads have been removed.
2. We can now call peaks on the data sets to identify regions of significant signal enrichment.

```
#ATAC
macs2 callpeak -f BAMPE -t shifted/SKNMC_ATAC.DupsRemoved.
sorted_shifted.bam -g hs --outdir ATAC_MACS2_Peaks -n SKNMC_
ATAC
```

```
#H3K27ac
macs2 callpeak -t SKNMC_H3K27ac.sorted.bam -g hs --outdir
H3K27ac_MACS2_Peaks -n SKNMC_H3K27ac
```

```
#FLI1
macs2 callpeak -t SKNMC_FLI1.sorted.bam -g hs --outdir FLI1_
MACS2_Peaks -n SKNMC_FLI1
```

MACS2 outputs:

- (a) an excel readable file with information about the peak calls,
- (b) a `.narrowPeak` which is a BED 6+ 4 which can be loaded into IGV or the UCSC Genome Browser for visualization (below).
- (c) a `.summits` file which contains the bp of the summit of each peak and.
- (d) a `model.R` file which will create an image of the model for your data.

3.4 Visualization

There are many ways to visualize the data but common tools include the Integrative Genomics Viewer (IGV) or UCSC Genome Browser [55]. The sorted bam files from above can be loaded directly into IGV in order to visualize the alignments (Fig. 1) (*see Note 11*).

3.5 CREATE HEATMAPS

A common analysis is to create heatmaps of various genomic signals at different loci. In this section we will take the EWSR1-FLI1 peaks called from above and plot the ATAC and H3K27ac signal at these regions. This analysis is highly flexible in terms of data sets, genomic regions, and even visualization options. Therefore, the processing steps should be edited to fit the specific needs of the project.

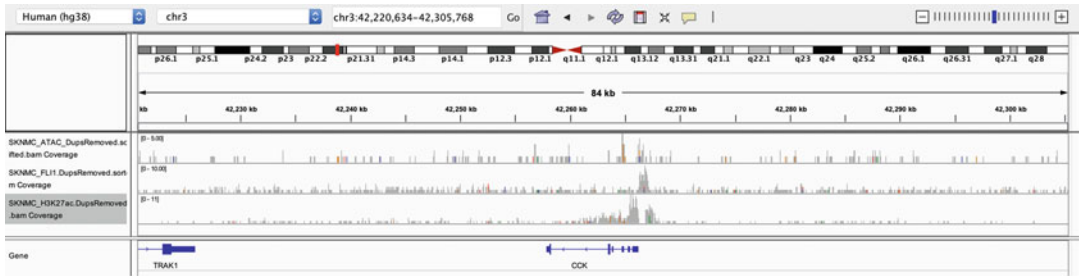


Figure 1 - Simon and Gomez

Fig. 1 IGV Snapshot of Genomic Data. A Browser shot demonstrating ATAC-seq signal (top), FLI1 ChIP-seq (middle), and H3K27ac ChIP-seq data at the Cholecystokinin (CCK) locus. We can see ATAC signal and FLI1 signal at the transcription start site as well as H3K27ac signal flanking the FLI1 binding site

We recommend using deepTools which is a suite of python tools that, among many other useful functions, creates and visualizes heatmaps.

1. deepTools works primarily with bigWig files which are used to display dense continuous data such as alignments along the genome. Therefore, the bam files will need to be first converted into bigWig (bw) files. -b specifies the input file to be converted, -o is the new bw being created, -of is the format requested (bigwig), -bs is the bin size of the file which we set to 1 to give single-bp resolution, and -p is the number of threads to use (*see Note 12*).

```
#Convert ATAC to bigWig
bamCoverage -b shifted/SKNMC_ATAC_DupsRemoved.sorted_shifted.bam -o SKNMC_ATAC_DupsRemoved.sorted_shifted.bw -of bigwig -bs 1 -p 8
```

```
#Convert H3K27ac to bigWig
bamCoverage -b SKNMC_H3K27ac.DupsRemoved.sorted.bam -o SKNMC_H3K27ac.DupsRemoved.sorted.bw -of bigwig -bs 1 -p 8
```

```
#Convert FLI1 to bigWig
bamCoverage -b SKNMC_FLI1.DupsRemoved.sorted.bam -o SKNMC_FLI1.DupsRemoved.sorted.bw -of bigwig -bs 1 -p 8
```

2. In order to compute the matrix for the heatmap we need to give deepTools a bedfile of the regions we are interested in. In this case, we will modify the FLI1 narrowPeak file into the proper format by using the cut command (*see Note 13*).

```
cut -f 1-6 FLI1_MACS2_Peaks/SKNMC_FLI1_peaks.narrowPeak > SKNMC_FLI1_peaks.bed
```

- Now we can run the `computeMatrix` reference-point function from the `deepTools` suite to create a matrix of scores per genomic region of interest. We indicate the signal file(s) using `-S` and the regions of interest `-R`, how far upstream and downstream in basepairs from the reference point to plot (`-a`, `-b`), the name of the output (`-o`), the point of reference in the genomic region (`--referencePoint`), how to sort the regions in the matrix (`--sortRegions`, and `--sortUsing`), and finally the number of threads (`-p`).

```
computeMatrix reference-point -S SKNMC_FLII1.
DupsRemoved.sorted.bw SKNMC_ATAC_DupsRemoved.
sorted_shifted.bw SKNMC_H3K27ac.DupsRemoved.
sorted.bw -R SKNMC_FLII1_peaks.bed -a 2000 -b 2000
-o SKNMC_signal_at_FLII1_Peaks.txt.gz --referen-
cePoint center --sortRegions descend --sortUsing
mean -p 8.
```

- Using the output matrix, `SKNMC_signal_at_FLII1_Peaks.txt.gz`, we can create a heatmap with the `plotHeatmap` function.

```
plotHeatmap -m SKNMC_signal_at_FLII1_Peaks.txt.gz -o
SKNMC_signal_at_FLII1_peaks.pdf --refPointLabel Peak_Center
--sortUsingSamples 1
```

- The resulting pdf plot has our three heatmaps of FLII, ATAC, and H3K27ac signal at FLII peak sites ± 2 kb from the center of the FLII peak (Fig. 2).

3.6 Transcription Factor Motif Enrichment

- We can look for the presence of statistically enriched transcription factor motifs within our peaks using HOMER (*see Note 14*).

```
#ATAC Motif Finding
findMotifsGenome.pl ATAC_MACS2_Peaks/SKNMC_ATAC_peaks.narrow
Peak WholeGenomeFasta/genome.fa ATAC_hg38_HOMER -p 8
```

```
#FLII Motif Finding
findMotifsGenome.pl FLII_MACS2_Peaks/SKNMC_FLII1_peaks.narrow
Peak WholeGenomeFasta/genome.fa FLII_hg38_HOMER -p 8
```

- The output will be an HTML file in which you can read with any browser. The results for FLII demonstrate a statistically significant enrichment of the repetitive GGAA motif and predict the EWSR1-FLII transcription factor as expected (Fig. 3).

3.7 Genomic Feature Association of Peaks

One common, and useful analysis, is to associate peak calls with specific genomic features. In order to perform these next steps, we will utilize R, R Studio, and various Bioconductor R Packages (*see Note 15*).

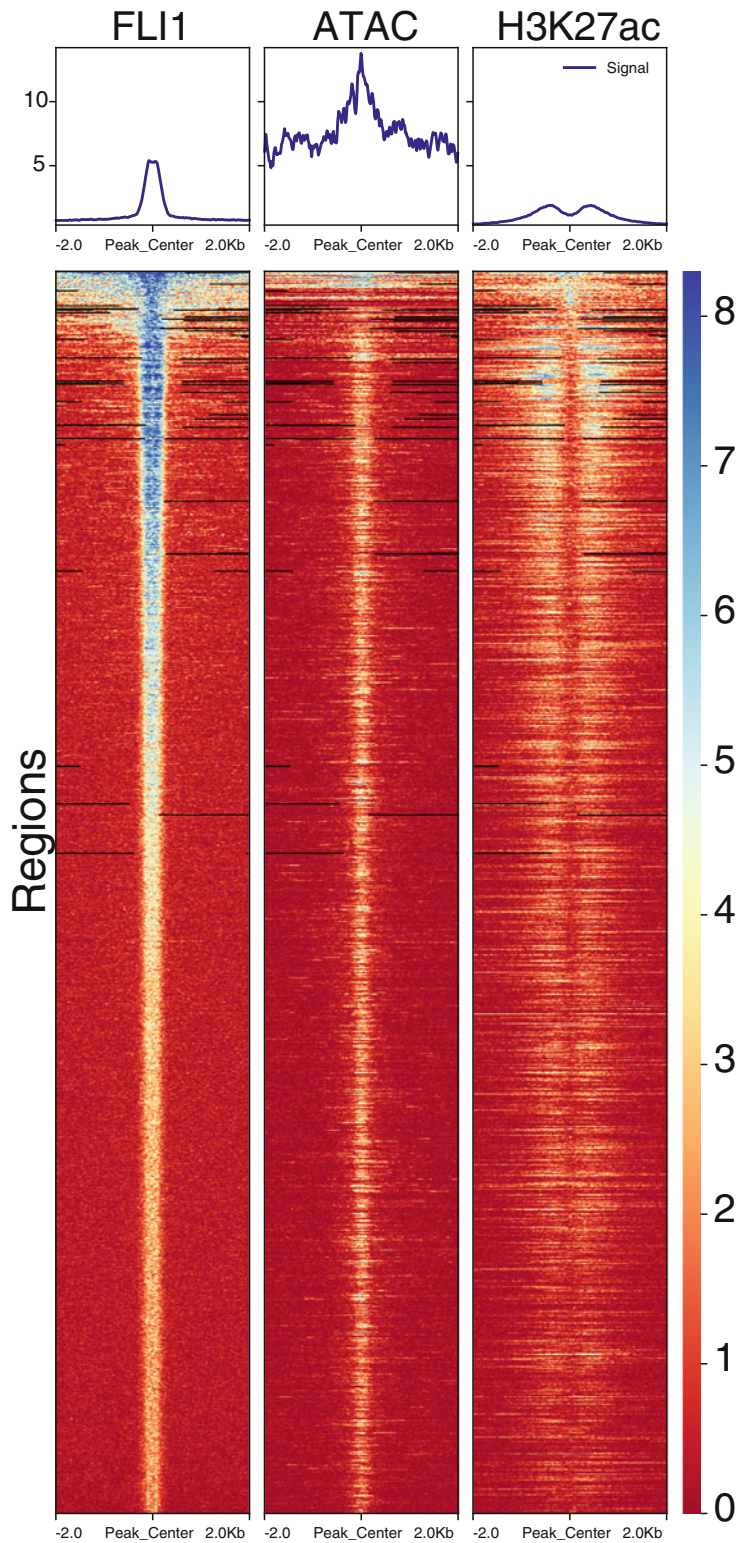


Fig. 2 Heatmap produced from deepTools. Visualization of FLI1 signal (left), ATAC-seq (middle), and H3K27ac (right) at FLI1 peak sites ± 2000 bp from the peak center. The data is sorted in descending order by mean FLI1 signal across the window. The scale goes from low (red) to high (blue) signal. We identify coenrichment of FLI1 and ATAC-signal followed by flanking H3K27ac signal consistent with previous reports

Homer *de novo* Motif Results FLI1_hg38_HOMER

[Known Motif Enrichment Results](#)

[Gene Ontology Enrichment Results](#)

If Homer is having trouble matching a motif to a known motif, try copy/pasting the matrix file into [STAMP](#)

More information on motif finding results: [HOMER | Description of Results | Tips](#)

Total target sequences = 4502

Total background sequences = 34154

* - possible false positive





Rank	Motif	P-value	log P-value	% of Targets	% of Background	STD(Bg STD)	Best Match/Details	Motif File
1		1e-1532	-3.528e+03	32.03%	1.26%	42.3bp (140.6bp)	EWSR1-FLI1/MA0149.1/Jaspar(0.814) More Information Similar Motifs Found	motif file (matrix)
2		1e-789	-1.818e+03	30.61%	3.93%	40.8bp (86.3bp)	Elk4(ETS)/Hela-Elk4-ChIP-Seq(GSE31477)/Homer(0.979) More Information Similar Motifs Found	motif file (matrix)
3		1e-303	-6.993e+02	4.40%	0.05%	34.0bp (102.3bp)	GFX(?) / Promoter/Homer(0.981) More Information Similar Motifs Found	motif file (matrix)
4		1e-146	-3.377e+02	5.49%	0.60%	58.3bp (120.1bp)	RME1(MacIsaac)/Yeast(0.684) More Information Similar Motifs Found	motif file (matrix)

Fig. 3 Screenshot of HOMER output. HOMER outputs an html file that can be read by any browser. It contains the *de novo* identified motifs and the putative transcription factor that binds that site. It displays *p*-values which are recommended to be very significant ($<1e-50$). Consistent with previous studies we find that FLI1 binding sites significantly enriched for GGAA repetitive sequences

1. Open R studio and install that latest version of Bioconductor by entering the following code within R.

```
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install()
```

2. Install the required Bioconductor packages.

```
BiocManager::install(c("ATACseqQC", "Rsamtools", "BSgenome.Hsapiens.UCSC.hg38", "GenomicRanges", "ChIPpeakAnno", "rtracklayer"))
```

3. Load the needed packages and database that contains the annotation of the features.

```
library(GenomicRanges)
library(ChIPpeakAnno)
library(TxDb.Hsapiens.UCSC.hg38.knownGene)
library(rtracklayer)
```

4. Import the narrowPeak files that resulted from MACS2 as GenomicRange objects. This allows them to be easily manipulated within R.

```
SKNMC_ATAC_Peaks =
import("SKNMC_ATAC_peaks.narrowPeak", format="narrowPeak")
SKNMC_FLI1_Peaks =
import("SKNMC_FLI1_peaks.narrowPeak", format="narrowPeak")
```

```
SKNMC_H3K27ac_Peaks =
import("SKNMC_H3K27ac_peaks.narrowPeak", format="narrowPeak")
```

5. Assign the peaks to genomic features using the assignChromosomeRegion function from the ChIPSeqAnno package.

```
ATAC_CR<-assignChromosomeRegion(SKNMC_ATAC_Peaks, nucleotide-
Level=FALSE,
precedence=c("Promoters", "immediateDownstream",
"fiveUTRs", "threeUTRs",
"Exons", "Introns"),
TxDb=TxDb.Hsapiens.UCSC.hg38.knownGene)
```

```
H3K27ac_CR<-assignChromosomeRegion(SKNMC_H3K27ac_Peaks,
nucleotideLevel=FALSE,
precedence=c("Promoters", "immediateDownstream",
"fiveUTRs", "threeUTRs",
"Exons", "Introns"),
TxDb=TxDb.Hsapiens.UCSC.hg38.knownGene)
```

```
FLI1_CR<-assignChromosomeRegion(SKNMC_FLI1_Peaks, nucleotide-
Level=FALSE,
precedence=c("Promoters", "immediateDownstream",
"fiveUTRs", "threeUTRs",
"Exons", "Introns"),
TxDb=TxDb.Hsapiens.UCSC.hg38.knownGene)
```

6. Now create barplots of the peak associations with the genomic features.

```
#ATAC Barplot
barplot(ATAC_CR$percentage,ylim=c(0,50),names.arg=c("Promo-
ter","Downstream","5'UTR","3'UTR","Exons","Introns","Inter-
genic"),col="darkred",main="ATAC Peak Associations")
```

```
#H3K27ac Barplot
barplot(H3K27ac_CR$percentage,ylim=c(0,50),names.arg=c("Pro-
moter","Downstream","5'UTR","3'UTR","Exons","Introns","Inter-
genic"),col="darkgreen",main="H3K27ac Peak Associations")
```

```
#FLI1 Barplot
barplot(FLI1_CR$percentage,ylim=c(0,50),names.arg=c("Promo-
ter","Downstream","5'UTR","3'UTR","Exons","Introns","Inter-
genic"),col="darkblue",main="FLI1 Peak Associations")
```

7. We can put each of these plots on the same page and export the file as a .pdf for viewing (Fig. 4).

```
#Create the pdf file name and dimensions
pdf("Genomic_Association_of_Peaks.pdf",width = 6, height = 12)
```

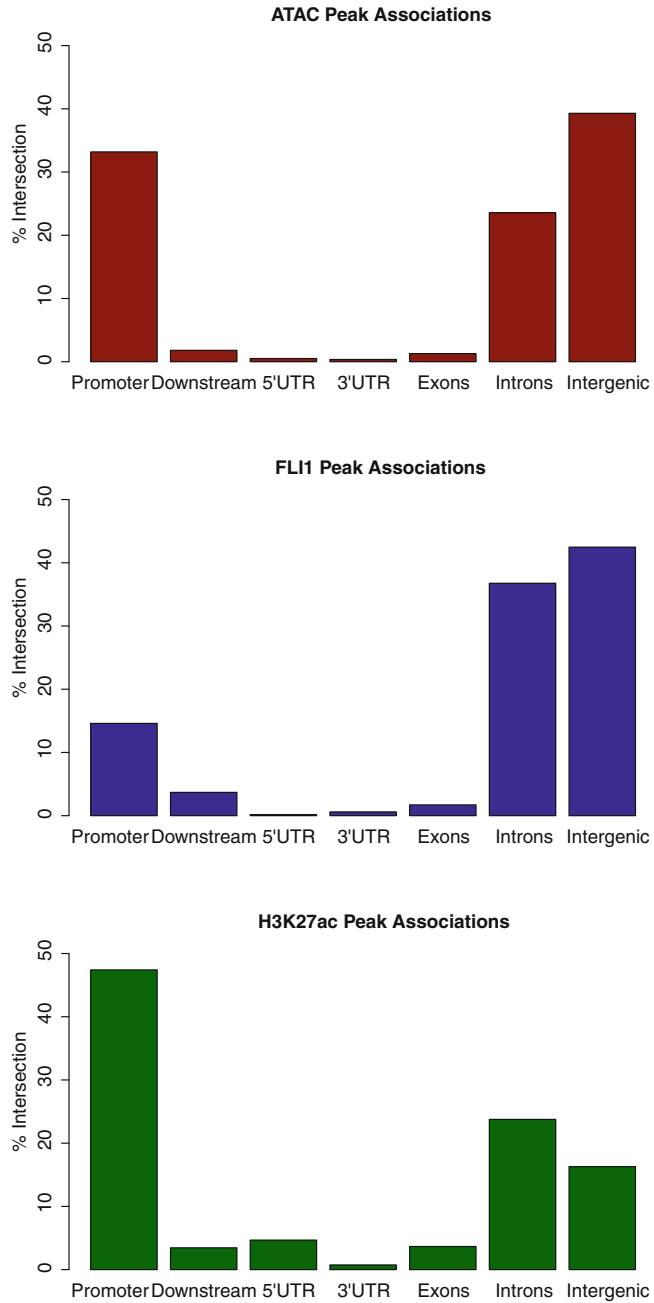


Fig. 4 Genomic Associations of Peaks. Peaks were associated with specific genomic features (*x*-axis) using the Bioconductor package ChIPpeakAnno. The percentage of ATAC-Seq (top), FLI1 (middle), and H3K27ac (bottom) peaks within each feature is plotted (*y*-axis)

```

#Create a new graphic that is a 1 x 3 grid
par(mfrow=c(3,1))
#plot the data
barplot(ATAC_CR$percentage,ylim=c(0,50),names.arg=c("Promo-
ter","Downstream","5'UTR","3'UTR","Exons","Introns","Inter-
genic"),col="darkred",main="ATAC Peak Associations",ylab="%
Intersection")

barplot(FLI1_CR$percentage,ylim=c(0,50),names.arg=c("Promo-
ter","Downstream","5'UTR","3'UTR","Exons","Introns","Inter-
genic"),col="darkblue",main="FLI1 Peak Associations",ylab="%
Intersection")

barplot(H3K27ac_CR$percentage,ylim=c(0,50),names.arg=c("Pro-
moter","Downstream","5'UTR","3'UTR","Exons","Introns","Inter-
genic"),col="darkgreen",main="H3K27ac Peak Associations",
ylab="% Intersection")
dev.off()

```

4 Notes

1. The fastq-dump application allows for the extraction of raw fastq reads from SRA files. As ATAC data sets are typically paired-end sequencing experiments, the `--split-files` command extracts the paired reads into separate files. If the SRA file does not contain paired-end reads `--split-files` has no effect. Therefore, as a precaution, it is recommended to include this option even if the data is not expected to be paired-end.
2. ATAC-Seq and other specific data sets enrich for small target fragments (<140 bp). When read length exceeds DNA insert size, the sequenced read may read through the DNA fragment and into the 3' sequencing adapter. This can result in spurious alignments or cause reads to not align to the reference genome. Therefore, it is advised to identify and trim the sequencing adapters prior to alignment.
3. Many ATAC-Seq protocols, including the example dataset, use the Nextera DNA Sample Prep Kit (Illumina). We create a FASTA file with the known adapter sequences that were obtained from the Illumina support site. This file can be tuned to your specific set of adapters by simply editing the FASTA file. The files used for this analysis are also available on the Github repository (https://github.com/gomeznick86/Ewing_Book_Chapter).
4. Skewer can also search single-end reads by changing the mode using the `-m`. We use the default option by searching the tail of the read for the adapter, `-m tail`. Many single-end protocols use the TruSeq Illumina kit. We include these adapter sequence in a separate file as they differ from those in the Nextera kit.

5. FASTQC will create an html file of basic quality control metrics for your sample. It can be useful for identifying problems such as read quality and GC biases. Please see the guide online for more information.
6. FASTQC is unable to make the output directories. Therefore they need to be created prior to running FASTQC. You can make a new directory from the command line by using the `mkdir` command, for example as follows:

```
mkdir FASTQ_OUTPUT
```

7. Bowtie2 requires a genome index in order to run. This can either be created on your own or you can use a prebuilt index. This chapter will use a prebuilt *H. sapiens* UCSC hg38 index that was obtained from the Illumina iGenomes support site. This file can be downloaded from (https://support.illumina.com/sequencing/sequencing_software/igenome.html).
8. The method for removing duplicate reads depends on the method of sequencing. For paired-end data, we use `sambalster`, whereas with single-end data, we use `samtools rmdup`. See the documentation for each method for further details.
9. Bam files are binary versions of sam files which allow for easier long-term storage and are a useful format for many different downstream applications. They can be sorted by coordinate or read-pair depending on the options selected. In our experience most applications use coordinate sorted bam files.
10. Due to the nature of the transposase, the ATAC-seq reads need to be shifted (+4) or (-5) bp depending on the strand. In this example, we take advantage of a few R Packages in order to shift the reads properly. This process took ~45 min using 1 core but can take a few hours depending on the size of the input file. The R script is available from the github page (https://github.com/gomeznick86/Ewing_Book_Chapter). Please install the required Bioconductor packages (`Rsamtools`, `ATACseqQC`, and `Bsgenome.Hsapiens.UCSC.hg38`) in order to run successfully.
11. IGV is convenient due to its ease of use and drag and drop properties. Take note that IGV defaults to the hg19 genome. Ensure that you select the correct genome build for your data, in this chapter we need to select hg38. IGV can also handle many other data sets including the narrowPeak format, which is the format used to annotate peak coordinates.
12. Ensure that the bam index (.bai) is in the same directory as the .bam file in order to prevent an indexing error from occurring.

13. cut is a command that is included in UNIX that allows for the extraction of specific fields. In this case we are using it to extract the first 6 columns (-f 1-6) and saving the output as a .bed file.
14. HOMER is able to perform de novo as well as known motif enrichment. It can also look for motif enrichment between two data sets. Please see HOMER documentation for more information. Also note hg38 is not compatible with HOMER by default. In order to overcome this, we can instead provide HOMER with a hg38 whole-genome FASTA file. This file is also included in the Illumina iGenomes referenced in **Note 7**. Make sure to change the paths of the narrowPeak and genome.fa file to the locations of the files on your own device.
15. Bioconductor is an open source community that provides tools for the analysis and comprehension of genomic data. It has an active community in terms of development, maintenance, as well as troubleshooting. This is a great resource if specific questions arise during analysis of data.

Acknowledgments

We would like to thank Y Yang and L Whitehouse for constructive comments. J.M.S. was supported by The Eunice Kennedy Shriver National Institute of Child Health and Human Development (U54HD079124) and NINDS (P30NS045892). N.C.G. holds a Postdoctoral Enrichment Program Award from the Burroughs Wellcome Fund and is supported by a NIH Postdoctoral Ruth L. Kirschstein National Research Service Award F32CA221353.

References

1. Grünewald TGP, Cidre-Aranaz F, Surdez D et al (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4:5
2. Delattre O, Zucman J, Plougastel B et al (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 359:162–165
3. Sorensen PH, Lessnick SL, Lopez-Terrada D et al (1994) A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor. *ERG*, *Nature Genetics* 6:146–151
4. Jeon IS, Davis JN, Braun BS et al (1995) A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* 10:1229–1234
5. Patel M, Simon JM, Iglesia MD et al (2012) Tumor-specific retargeting of an oncogenic transcription factor chimera results in dysregulation of chromatin and transcription. *Genome Res* 22:259–270
6. Wei GH, Badis G, Berger MF et al (2010) Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J* 29:2147–2160
7. Gangwal K, Close D, Enriquez CA et al (2010) Emergent properties of EWS/FLI1 regulation via GGAA microsatellites in Ewing's sarcoma. *Genes Cancer* 1:177–187
8. Guillon N, Tirode F, Boeva V et al (2009) The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PLoS One* 4:e4932
9. Riggi N, Knoechel B, Gillespie SM et al (2014) EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or

- repress enhancer elements in Ewing sarcoma. *Cancer Cell* 26:668–681
10. Gomez NC, Hepperla AJ, Dumitru R et al (2016) Widespread chromatin accessibility at repetitive elements links stem cells with human cancer. *Cell Rep* 17:1607–1620
 11. Luger K, Mäder AW, Richmond RK et al (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251–260
 12. Strahl BD, Allis CD (2000) The language of covalent histone modifications. *Nature* 403:41–45
 13. Szenker E, Ray-Gallet D, Almouzni G (2011) The double face of the histone variant H3.3. *Cell Res* 21:421–434
 14. Weber CM, Henikoff S (2014) Histone variants: dynamic punctuation in transcription. *Genes Dev* 28:672–682
 15. Orphanides G, LeRoy G, Chang CH et al (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92:105–116
 16. Vignali M, Hassan AH, Neely KE et al (2000) ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* 20:1899–1910
 17. Fry CJ, Peterson CL (2001) Chromatin remodeling enzymes: who's on first? *Curr Biol* 11:R185–R197
 18. Zaret KS, Carroll JS (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25:2227–2241
 19. Riggi N, Suva ML, Suva D et al (2008) EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res* 68:2176–2185
 20. Cidre-Aranaz F, Grünewald TGP, Surdez D et al (2017) EWS-FLI1-mediated suppression of the RAS-antagonist Sprouty 1 (SPRY1) confers aggressiveness to Ewing sarcoma. *Oncogene* 36:766–776
 21. Kinsey M, Smith R, Lessnick SL (2006) NR0B1 is required for the oncogenic phenotype mediated by EWS/FLI in Ewing's sarcoma. *Mol Cancer Res* 4:851–859
 22. García-Aragoncillo E, Carrillo J, Lalli E et al (2008) DAX1, a direct target of EWS/FLI1 oncoprotein, is a principal regulator of cell-cycle progression in Ewing's tumor cells. *Oncogene* 27:6034–6043
 23. Gangwal K, Lessnick SL (2008) Microsatellites are EWS/FLI response elements: genomic "junk" is EWS/FLI's treasure. *Cell Cycle* 7:3127–3132
 24. E.P. Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74
 25. Thurman RE, Rynes E, Humbert R et al (2012) The accessible chromatin landscape of the human genome. *Nature* 489:75–82
 26. Johnson DS, Mortazavi A, Myers RM et al (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316:1497–1502
 27. Barski A, Cuddapah S, Cui K et al (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837
 28. Mikkelsen TS, Ku M, Jaffe DB et al (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448:553–560
 29. Giresi PG, Lieb JD (2009) Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). *Methods* 48:233–239
 30. Simon JM, Giresi PG, Davis IJ et al (2012) Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat Protoc* 7:256–267
 31. Boyle AP, Davis S, Shulha HP et al (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132:311–322
 32. Song L, Crawford GE (2010) DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor Protocols* 2010:pdb.prot5384
 33. Crawford GE, Holt IE, Whittle J et al (2006) Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Res* 16:123–131
 34. Buenrostro JD, Giresi PG, Zaba LC et al (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10:1213–1218
 35. Buenrostro JD, Wu B, Chang HY et al (2015) ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* 109:21.29.1–21.29.9
 36. Luger K, Richmond TJ (1998) The histone tails of the nucleosome. *Curr Opin Genet Dev* 8:140–146
 37. Zhang Y, Liu T, Meyer CA et al (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9:R137
 38. Jiang H, Lei R, Ding S-W et al (2014) Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15:182–112
 39. S. Andrews (2012). FastQC: a quality control tool for high throughput sequence data.

- <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
40. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* 9:357–359
 41. Li H, Handsaker B, Wysoker A et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079
 42. Faust GG, Hall IM (2014) SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* 30:2503–2505
 43. Heinz S, Benner C, Spann N et al (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38:576–589
 44. R.C.T. (2019). R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>
 45. R.T. (2015). RStudio: Integrated Development Environment for R. <http://www.rstudio.com/>
 46. Ramirez F, Ryan DP, Grüning B et al (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44:W160–W165
 47. Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genomics viewer. *Nat Biotechnol* 29:24–26
 48. Huber W, Carey VJ, Gentleman R et al (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* 12:115–121
 49. Ou J, Liu H, Yu J et al (2018) ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics* 19:169–113
 50. M. Morgan, H. Pagès, V. Obenchain, et al. (2019)., Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import, <http://bioconductor.org/packages/Rsamtools>
 51. T.B.D. Team (2015), BSgenome.Hsapiens. UCSC.hg38: Full genome sequences for *Homo sapiens* (UCSC version hg38)
 52. Lawrence M, Huber W, Pagès H et al (2013) Software for computing and annotating genomic ranges. *PLoS Comput Biol* 9:e1003118
 53. Zhu LJ, Gazin C, Lawson ND et al (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11:237–210
 54. Lawrence M, Gentleman R, Carey V (2009) rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics* 25:1841–1842
 55. Kent WJ, Sugnet CW, Furey TS et al (2002) The human genome browser at UCSC. *Genome Res* 12:996–1006



Systems Biology Analysis for Ewing Sarcoma

Marianyela Petrizzelli, Jane Merlevede, and Andrei Zinovyev

Abstract

Ewing sarcoma (EwS) is a highly aggressive pediatric bone cancer that is defined by a somatic fusion between the *EWSR1* gene and an ETS family member, most frequently the *FLI1* gene, leading to expression of a chimeric transcription factor EWSR1-FLI1. Otherwise, EwS is one of the most genetically stable cancers. The situation when the major cancer driver is well known looks like a unique opportunity for applying the systems biology approach in order to understand the EwS mechanisms as well as to uncover some general mechanistic principles of carcinogenesis. A number of studies have been performed revealing the direct and indirect effects of EWSR1-FLI1 on multiple aspects of cellular life. Nevertheless, the emerging picture of the oncogene action appears to be highly complex and systemic, with multiple reciprocal influences between the immediate consequences of the driver mutation and intracellular and intercellular molecular mechanisms, including regulation of transcription, epigenome, and tumoral micro-environment. In this chapter, we present an overview of existing molecular profiling resources available for EwS tumors and cell lines and provide an online comprehensive catalogue of publicly available omics and other datasets. We further highlight the systems biology studies of EwS, involving mathematical modeling of networks and integration of molecular data. We conclude that despite the seeming simplicity, a lot has yet to be understood on the systems-wide mechanisms connecting the driver mutation and the major cellular phenotypes of this pediatric cancer. Overall, this chapter can serve as a guide for a systems biology researcher to start working on EwS.

Key words Ewing sarcoma, EWSR1-FLI1, Cancer systems biology, Omics data, Network, Mathematical modeling, Data integration

1 Introduction

Ewing sarcoma (EwS) is a rare aggressive bone and soft tissue cancer, with the peak at 15 years old in age distribution. Its worldwide incidence rate is about 2–3 cases per million children, being lower in African population compared to European population. The 5-years survival is currently 70–80% for the cases with localized tumors and drops to 30% for those with metastases [1].

From the point of view of cancer genomics, EwS represents one of the most genetically stable cancers. Together with rhabdoid tumors, EwS was ranked as the cancer having the lowest somatic

mutation frequency among 27 analyzed cancer types [2]. Therefore, EwS is also relatively homogeneous at the genomic level. At the same time, the major genetic cancer driver event in EwS is well established: it is a balanced chromosomal translocation leading to the fusion of a member of the FET gene family with an ETS transcription factor. In 85% of the cases, this leads to appearance of a chimeric transcription factor EWSR1-FLI1, which activity leads to the widespread changes of the cellular molecular profiles and phenotypes. In the following we will refer to the systemic properties of EWSR1-FLI1 gene and protein, unless explicitly specified, since other translocation types generally have similar characteristics and are rare.

EWSR1 gene encodes a multifunctional protein that is involved in various cellular processes, including gene expression, cell signaling, RNA processing and transport. The protein includes an N-terminal prion-like low-complexity domain PrLD and a C-terminal RNA-binding domain (the latter is commonly missing in the fusion protein). Characterizing the precise normal biological function of EWSR1 protein is difficult due to its potency to interact with many other proteins, which is usually articulated as that EWSR1 connects together several important biological functions such as transcription and splicing [3]. From the network biology point of view, EWSR1 frequently plays the role of a major hub in the protein-protein interaction (PPI) networks of proteins associated to diseases [4]. Genomic translocations involving *EWSR1* are not exclusively attributed to EwS; they serve driver mutations in a broad variety of mesenchymal lesions which includes Ewing's sarcoma/peripheral neuroectodermal tumor, desmoplastic small round cell tumor, clear cell sarcoma, angiomatoid fibrous histiocytoma, extraskeletal myxoid chondrosarcoma, and a subset of myxoid liposarcoma [5].

FLI1 is a member of the large E26 transformation-specific (ETS) transcription factor family characterized by a specific binding DNA motif with consensus sequence CAGGAAG. It is normally implicated in the development of different animal tissues. For example, it transcriptionally regulates genes that drive normal hematopoiesis and vasculogenesis [6].

The fusion between two normal genes *EWSR1* and *FLI1* leads to the new properties of the resulting protein, EWSR1-FLI1, the most remarkable of which is its ability to bind microsatellite sequences containing exact GGAA repeats, which are rather abundant in the human genome and are not functional in healthy cells [7, 8]. Upon EWSR1-FLI1 binding, the microsatellites become potent enhancers of genes located sometimes at hundreds of thousands of base pairs away. This new property has a profound effect, with activation of EWSR1-FLI1 leading to the drastic rewiring of gene expression and epigenetic reprogramming. These changes appear to be detrimental to cells and lead to apoptosis for

the majority of human cell types, but can be tolerated in few, including pediatric mesenchymal stem cells (MSCs) or neural crest stem cells (NCSC) [1].

Importantly, the process of EwS tumorigenesis can be recapitulated to some extent in several inducible cell line models, in which the expression of EWSR1-FLI1 can be modulated through doxycycline-controlled short hairpin RNA (shRNA) or application of siRNA-based oncogene knockdown. Reducing the expression of EWSR1-FLI1 in EwS cells to some minimal level (e.g., 20% from its initial concentration) leads to the drastic reduction of proliferation and changes in the cell morphology toward MSC-like phenotype [9]. This process can be reversed by reactivating the chimeric oncogene [10]. Further, reduction of the EWSR1-FLI1 expression leads to apoptosis, showing addiction of the EwS cells to the oncogene activity. Inducible systems have been extensively used to study the role of other genetic and epigenetic actors and the downstream mechanisms of the EWSR1-FLI1 action. The development of genetically engineered mouse models of EwS has not been very successful so far, probably due to the distribution of GGAA microsatellites in the mouse genome being different from human [11].

Understanding the mechanisms of adult cancers is usually heavily complicated by genomic complexity and genetic heterogeneity of the tumors, which might be a result of long (sometimes, decades-long) history of their evolution. In adults, the tumorigenic process is usually a direct and indirect consequence of a combinatorial action of several driver genes, with importance of usually vaguely determined temporal sequence of events. EwS looks a much simpler case and hence it seems, at first glance, a promising target for application of the systems biology approach in order to unravel this particular mechanism of carcinogenesis.

A systems biology approach is here understood as studying a biological phenomenon by combining a collection of system-wide molecular information: in particular, by applying multiple perturbations or profiling series of tumor samples, and recapitulating the available data in the form of mathematical models that reflect some aspects of the cellular life.

Indeed, EwS is a cancer type which is relatively well characterized at the molecular level (see description of available molecular data in the corresponding sections of this chapter and some examples in Table 1).

Understanding EwS and other pediatric cancers through the analysis of “big” omics data, by applying mathematical modeling and, more recently, machine learning approaches stimulated several interdisciplinary consortia to launch large-scale European projects with this motivation in mind. The ambition of these projects is to unravel the mystery of pediatric (sometimes called embryonal) tumors, including EwS. The driving idea is that many embryonal tumors exploit the same type of fragilities such as the regulation of

Table 1
Exemplary omics datasets generated in EwS studies

Year	Experimental system and data	Dataset characteristics	Publication, ID
<i>Genomics</i>			
2012	aCGH profiling of EwS	67 EwS tumors and 16 EwS cell lines	[21], GSE20355
2014	WGS + targeted sequencing	112 (WGS) and 199 (TS) tumors	[26], ICGC
2014	WES, WGS, SNP array for EwS tumors and cell lines	WES 96 tumors and 11 cell lines, WGS of 7 pairs, SNP array of 28 pairs	[28], PedcBio-Portal
2014	Sequencing of EwS tumors and cell lines	65 tumors and 36 cell lines, 6 paired WGS, 79 TS, 6 SNP arrays	[27], dbGaP phs000768.v1.p1
<i>Epigenomics</i>			
2014	Epigenome of primary EwS and cell lines, MSC cells	ChIP-seq for 4 histone modifications, 5 transcription factors	[57], GSE61944
2015	“Epigenome map” of A673 cell line	ChIP-Seq profiles with active and inhibited EWSR1-FLI1	[58], https://tinyurl.com/r6ddvpb
2017	DNA methylation sequencing of EwS tumor and cell lines	140 EwS tumor samples, 16 EwS cell lines, and 32 primary MSCs	[55], GSE88826
2020	ChIP-Seq time series of A-673/TR/shEF	ChIP-Seq of FLI1 (6 time points) and H7K27ac (2 conditions)	[29], GSE129155
<i>Transcriptomics</i>			
2012	French CIT program EwS tumor Cohort	117 samples, Affymetrix HG-U133Plus2	[22], GSE34620
2013	EWSR1-FLI1 silencing	Time series, 13 time points	[10]
2014	ICGC transcriptomic dataset	57 EwS tumors profiled with RNA-Seq	[26]
<i>Proteomics</i>			
2016	RPPA of 18 EwS cell lines resistant to targeted therapy	218 proteins in RPPA panels	[64], GSE78124
2018	Proteomic profiling of 2 EwS cell lines	2336 and 847 proteins of which 543 and 259 secreted proteins	[60], PXD007909
<i>Drug and gene screens</i>			
2012	GDSC project	Drug sensitivity for 21 EwS cell lines	[73]
2017	siRNA-based screening of EwS	7000 genes, with low/high EWSR1-FLI1	[75]
2018	Determining druggable EWSR1-FLI1 interactome	3325 experimental compounds in the inducible cell line	[77]

(continued)

Table 1
(continued)

Year	Experimental system and data	Dataset characteristics	Publication, ID
<i>miRNAome</i>			
2012	mRNA/miRNA EwS profiling	39 EwS tumors	[84], GSE37371
2016	miRNA profiling of EwS tumors	20 tumors and 4 MSCs	[44], GSE80201
<i>Single cell</i>			
2020	Single cell RNA-Seq of EwS inducible cell line and PDXs	7 time points (383 cells), 5 PDX (8431 cells)	[29], GSE130019, GSE130023, GSE130024

The complete catalogue is available at https://github.com/sysbio-curie/EwingSarcoma_Omics_Atlas

cell cycle or apoptosis and the blockage of normal developmental and differentiation programs [12], which enable the appearance of these cancers early in human life, without accumulation of large number of mutations.

Examples of such large projects are the FP6 project “European Embryonal Tumour Pipeline” (EETP), aimed at generating omics data for EwS, and the FP7 project “ASSET: Analysing and Striking the Sensitivities of Embryonal Tumours” which collected efforts of 14 partner institutions <https://cordis.europa.eu/project/id/259348>. The “European Network for Cancer research in Children and Adolescents” (ENCCA) facilitated and structured networking activities for prioritization of, access to and clinical research on innovative, biologically targeted drugs for the treatment of childhood cancers. Currently, in the Horizon 2020 program, iPædiatricCancer (iPC) consortium (21 partner institutions, including major cancer research centers in Europe and the USA, IBM Research and Barcelona Supercomputer Center) works on integrating large-scale omics data on several pediatric cancers and making them available, using computational clouds, for machine learning and modeling-based analysis (<https://ipc-project.eu/>). The above-mentioned European consortia represent truly interdisciplinary team, putting in close collaboration experts in cancer biology, computer science and computational biology. The ambitious long-term goal of such efforts is to provide individualized diagnostic and prognostic tools for ongoing and future clinical trials focused on pediatric tumors such as EURO EWING, MOSCATO 02, MAPPYACTS.

The purpose of this chapter is to (1) outline the available omics datasets suitable for application of systems biology and machine learning-based approaches and (2) highlight several applications of

the systems biology approach to study EwS. Having these objectives in mind, we catalogued the EwS-related datasets from various sources, with the intention to update this catalogue online in the future. We collected and harmonized some of the dataset collections for immediate application of various computational analyses. Therefore, this chapter serves as a guide and a resource for the future systems biology studies of EwS.

2 Omics and Large-Scale Perturbation Studies of EwS

2.1 Sources of Public Omics Datasets

Considering its rareness, EwS is a pediatric cancer type which is relatively well studied at the molecular level. For example, International Cancer Genome Consortium (ICGC) included moderate size collection of EwS donor profiles as a multiomics dataset available through the corresponding data portal. Otherwise, the molecular profiles of Ewing tumors or model systems are scattered across data repositories and publications. For example, querying “Ewing” in ArrayExpress returns 135 datasets. EwS molecular profiling is part of data resources specialized in pediatric cancers such as Kids First initiative (<https://kidsfirstdrc.org/>), which also provides an associated computational cloud Cavatica (<https://cavatica.squareSPACE.com/>). Some EwS-related datasets can be obtained through a specialized instance of cBioPortal (focused on adult cancers), named PedcBioPortal which has been developed for the Childhood Cancer Genomics program (pediatric cancer-specialized data resource, supporting the curation and pancancer integration of public pediatric cancer genomics datasets, <http://pedcbiportal.org>).

In this section, we made a systematic effort to identify EwS-related molecular data across multiple sources, having in mind their potential usefulness in the future systems biology or machine learning-based studies. We focused on molecular data generated either as a result of profiling of tumor cohorts or the results of chemical or genetic or epigenetic perturbations applied to model systems (mainly cell lines). This effort resulted in a catalogue which is available as an online table at https://github.com/sysbio-curie/EwingSarcoma_Omics_Atlas. Some exemplary datasets from this catalogue are listed in Table 1.

Previously, an effort to summarize available large-scale EwS-related omics data has been undertaken in a review that we strongly recommend [13]. However, in [13] the effort was to summarize the biological insights obtained from the sequencing studies of EwS, while here we focus on the data itself and its possible reuse in a systems biology-oriented approach. We consider all data generation efforts for EwS, not only limited to sequencing, and mention more recent (e.g., single cell) studies as well as the types of data not previously summarized (such as proteomics and metabolomics).

2.2 Genomic Studies of EwS

The genome of EwS patients has been studied for decades. After the identification of the genomic fusion as initiating event, the focus of the studies has shifted towards the identification of secondary events. The initial efforts were drawn towards the identification of chromosome copy number variations (CNV) through microarray-based comparative genomic hybridization (aCGH) approaches [14–18]. Later, next-generation sequencing (NGS) allowed for characterizing the genome of EwS in a more precise way.

2.2.1 Microarray-Based Comparative Genomic Hybridization Studies

Ferreira et al. [19] studied 25 patients, 23 primary tumors obtained before any treatment and relapse for two of them. The authors observed a median number of three aberrations per case, with 21 tumors (84%) showing at least one DNA copy number aberrations (CNA). The most frequent gains were entire chromosomes 8 (56%), 12 (20%), 18 (12%), 20 (12%) and the short arm of chromosome 5 (5p) (20%). The most frequent losses involved entire chromosomes 10 (16%) and 19 (16%), and partial regions of chromosome arms 16q (16%) and 7q (25%). They computed the fraction of the genome affected (FGA) by this instability. The median of FGA for all samples was 6%.

An unsupervised clustering analysis was performed, using the smallest overlapping regions of imbalance (SORI) as variable [19]. Two subgroups were identified: one genomically unstable characterized by a high number of aberrations with a median number of 8 SORI per case (range: 4–18) and one genomically stable, with a median number of one SORI per case (range: 0–3). Trisomy 8, the most common secondary aberration in EwS, was equally found in both groups.

Based on the 20 patients with tumor samples at diagnosis for whom clinical and follow-up data were available, the genomically stable group showed a higher tendency to achieve complete remission during or after treatment than the genomically unstable group (100 vs. 62%). In addition, genomically unstable group was more refractory to chemotherapy and thus associated to a poor prognosis.

Savola et al. [20] studied 31 samples by aCGH: 23 were primary tumors, two recurrences, and six metastatic tumors. They confirmed most of the findings previously obtained by conventional CGH and array CGH studies [19] about the most recurrent CNVs. They also observed a significantly better prognosis for primary tumors with three or less CNVs than for tumors with higher number of CNVs both in terms of event-free and overall survival.

Finally, Mackintosh et al. [21] screened 67 tumors from untreated patients and 16 cell lines for CNAs by aCGH. They confirmed most of the previous findings: most frequent gains comprised the entire chromosome 8, and the chromosome arms 1q and 12p, while most frequent losses were located in 3p, 9p, 16q, and

17p. In their cohort, twice bigger than previously discussed cohorts, 1qG, detected in 31% of tumors, was the CNA with the highest clinical impact, associated with relapse and poor overall and disease-free survival. It was also confirmed that the number of CNAs drives the clinical outcome of patients.

Altogether, several seminal studies of EwS genome involving relatively small cohort sizes (between 30 and 70 tumors) converged to similar conclusions about small but clinically relevant genomic instability in EwS.

2.2.2 Genomewide Association Studies (GWAS) of EwS

Two large-scale GWAS studies of EwS have been published so far in order to find genetic determinants for predisposition to this cancer type. One of the major questions addressed by these studies concerns the geographical differences in the occurrence frequencies of EwS, considering the fact that this cancer is not considered to be highly heritable. In one of the studies, genotypes of 401 French individuals with EwS, 684 unaffected French individuals and 3668 unaffected individuals of European descent and living in the USA have been analyzed [22]. In a more recent analysis, 733 EwS cases and 1346 unaffected individuals of European ancestry were genotyped [23]. Both studies showed consistent results pointing to six susceptibility regions with the effect sizes larger than observed in the majority of cancers. Interestingly, the exact mechanism underlying one of the loci located close to *EGR2* gene was revealed [24]. EwS cell proliferation was shown to depend on the activity of EGR2. It appeared that a single SNP can increase the number of consecutive GGAA motifs in a genomic region located near *EGR2* with epigenetic characteristics of an active regulatory element, and thus increase the EWSR1-FLI1-dependent enhancer activity. Similar mechanism explained clinically relevant upregulated activity of *MYBL2*, a potent regulator of cell proliferation and cell survival in some EwS tumors and not the others, thus illustrating a possibility of cooperation of cancer drivers with regulatory germline variants, even if some of them were not yet identified in case-control GWAS [25].

2.2.3 Sequencing EwS Genome, WES and WGS Approaches

With the outbreak of NGS, more precise characterization of both coding and noncoding parts of the genome became possible. Several simultaneously finalized large-scale studies have investigated the EwS tumor genomes. Tirode et al. [26] performed whole-genome sequencing (WGS) of 112 tumors, and showed that EwS tumors have rather stable genomes, with a median number of somatic single-nucleotide variants (SNV) of 319, of somatic coding mutations of 10 and of somatic structural variations (SV) of 7 per tumor. Somatic CNAs were frequent, as previously demonstrated. The most frequently mutated genes were *STAG2* (in 19 cases, 17% of the patients), *TP53* (8 cases), and *EZH2* (3 cases). Interestingly, a

significantly greater number of SVs was observed in *STAG2* mutated cases but *STAG2* mutation was not associated with the number of SNVs and indels. The authors also identified in the one hand, mutual exclusivity between *STAG2* mutation and *CDKN2A* deletion and, on the other hand, a significant coassociation between *STAG2* and *TP53* mutations. Finally, linking these recurrently mutated genes to clinical features showed that patients with *STAG2* or *TP53* mutations had a significantly lower probability of survival, patients with neither *STAG2* nor *TP53* mutations had the highest probability of survival, and patients with both genes mutated had the worst outcome. A last track investigated by the authors was that subclonal *STAG2* mutations may expand at relapse.

Two other studies reported similar findings [27, 28]. A limitation of these studies in the use of their data is the variability of the sequencing strategy used on their cohorts and the limited number of normal-tumor available pairs. In Crompton et al., the authors performed WES of tumor-normal pairs from 26 patients, tumors from 66 patients, and 11 cell lines. They also performed WGS on 7 paired samples, genotyping array (SNP array) on 29 samples, and transcriptome sequencing (RNA-Seq) on 30 EwS samples. In Brohl et al., 101 samples were investigated: 65 tumors (including 13 normal-tumor pairs) and 36 cell lines [27]. Only six patients were investigated using paired WGS and 80 samples (including cell lines) were subjected to targeted sequencing, with a panel based on the mutations detected by WGS. In addition, RNA-Seq was performed on 30 samples.

In the end, the three independent studies made similar conclusions, using independent cohorts and different sequencing strategies, which reinforced their findings.

Another major dataset of EwS WGS data should be released during this year by the Kids First project, with the sequencing of around 1000 individuals and 400 families with an EwS case (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001228.v1.p1).

2.3 Studies of EwS Transcriptome

2.3.1 Expression of Coding Genes in EwS

EwS is relatively well characterized at the transcriptomic level. By browsing public repositories, we catalogued more than 1300 distinct transcriptomic profiles obtained at bulk level, among which approximately 600 were profiles of tumoral samples and 600 were profiles of cell lines under different conditions and perturbations. The absolute majority of transcriptomic profiles have been generated using various microarray platforms, with Affymetrix HG-U133Plus2 chip being the most popular (40% of profiles). Among experimental systems used for gene expression profiling, cell lines A-673 (Cellosaurus ID: CVCL_0080) and SK-N-MC (Cellosaurus ID: CVCL_0530) appear to be the most popular. Ironically, both these cell lines are marked as problematic because

of confusion with their cancer of origin: A-673 was initially thought to be a rhabdomyosarcoma, SK-N-MC a neuroblastoma cell line. Both of them are included in the Cancer Cell Line Encyclopedia (CCLE) and are well characterized at genomic and epigenomic levels (including gene expression, DNA methylation and chromatin mark profiles). For example, A-673 cell line was transformed to inducible cell line system, where the expression of EWSR1-FLI1 transcript can be modulated, and time series of transcriptomic changes measured upon inhibition or reactivation have been produced [10]. To this collection, one can add a rich dataset of single cell RNA-Seq profiles (9000 profiles) recently published and comprising time-resolved measurements of the inducible A-673-derived cell line upon activation of the oncogene, cells in several patient-derived xenografts (PDXs) and in xenograft-inducible cellular systems [29]. Another recently published large single cell dataset contains 9783 scRNA-Seq profiles for three EwS cell lines CHLA9, CHLA10, and TC71 [30].

Even if few hundreds of EwS tumors have been profiled at gene expression level, their simultaneous comparative study is complicated by the fact that different transcriptomic platforms have been used in these studies: therefore, one has to deal with laboratory-specific and platform-specific effects. The largest cohort of 117 EwS tumors has been profiled in the frame of the French Cartes d'Identité des Tumeurs (CIT) from the Ligue Nationale Contre le Cancer (<http://cit.ligue-cancer.net>) and made publicly available [22]. Two sets of samples for 44 EwS tumors have been profiled at the level of gene expression in order to characterize the inflammatory response of the tumoral tissues compared to cell lines [31]. More recently, 85 EwS tumors from two sample sources were profiled in order to construct and validate transcriptomic signatures of survival for EwS in [32]. International Cancer Genome Consortium (ICGC) openly provides 57 processed bulk RNA-Seq profiles of EwS (as a part of BOCA-FR dataset), released as a part of a large-scale EwS genomic study [26]. Few smaller datasets with the profiles of EwS tumors have been released in the public domain [9, 33–35].

We collected most of the identified transcriptomic datasets in a form suitable for further computational analysis in a unified collection of data matrices available at https://github.com/sysbio-curie/EwingSarcoma_Omics_Atlas.

2.3.2 Expression of Noncoding Parts of the Genome: *microRNA* and *lncRNA*

Deregulations in gene expression have also been studied in non-coding elements of EwS cells, like miRNA and lncRNA. Some of these datasets are listed in Table 1. Sand et al. in [13] reviewed, among others, performed expression studies on noncoding elements. We briefly recapitulate here the main studies and discuss more recent works.

One focus is to identify miRNAs deregulated in EwS cells compared with control cells. Mostly, limited number of cell lines were used to highlight the function of one specific miRNA, as in [36–38] and a series of papers published by the group of Tsumura [39–43]. For example, in the mentioned studies, they took advantage of one microarray experiment in 5 human EwS cell lines and human mesenchymal stem cells to show the role of let-7a, miR-16, and miR-29b in the cell cycle [39], miR-138 on EwS cell proliferation, invasion, and migration [43], miR-301a on cell proliferation [40], miR-20b in EwS cell proliferation [41] and miR-181c in apoptosis [42]. A recent study by Parafioriti et al. [44] used larger sample size of tumor samples, 20 patients affected by primary untreated tumors and normal MSCs from 4 healthy donors. The miRNAs microarray analysis of 954 miRNAs showed 58 significantly differentially expressed in EwS samples compared to MSCs, with 36 being up- and 22 being down-regulated. They suggest to consider BCL-2 as a novel biomarker for EwS. This dataset was then used in the paper by Liu Y [45], together with a gene expression dataset, to demonstrate the potential use of miR-21/CD166 as diagnostic markers and therapeutic targets for this disease.

Another focus is to identify miRNAs deregulated between different stages/characteristics of the disease. The aim is to either identify prognostic miRNAs or to compare the effects of the founding translocation type on the miRNome landscape. In such studies, sample sizes are usually bigger and tumor samples are preferred to cell lines. In the work published by Nakatani et al. [46], the authors studied 34 primary tumors to identify prognostic miRNAs, related to treatment response and outcome, comparing 21 tumors from patients who had an early tumor relapse with 13 tumors from patients who never recurred. They concluded that miR-34a expression was a strong predictor of outcome in EwS. In the work by Karnuth et al. [47], the cohort was composed of 40 EwS biopsies with different translocation types, six EwS cell lines and mesenchymal stem cells from six healthy donors. Of the 35 differentially expressed microRNAs between tumors and controls (over 377 investigated), 19 were higher and 16 lower expressed in EwS. miR-31 was the most differentially expressed microRNA, with lowest expression in mesenchymal stem cells. It was described as a potential tumor suppressor in EwS with influence on proliferation and invasiveness. In addition, no significantly differentially expressed microRNAs were detected between EwS samples with EWSR1-FLI or EWSR1-ERG translocations.

Finally, we mention two studies that differ from the main topics previously described. Teicher et al. [48] screened 63 human adult and pediatric sarcoma cell lines including 23 EwS with 100 FDA approved and 345 investigational agents. Both microRNA expression and gene expression were measured. The drug and compound

response, gene expression and microRNA expression data are publicly available at <http://sarcoma.cancer.gov>. De Feo et al. [49] applied an innovative approach, looking for the first time at the exosome of CD99, one of the hallmark surface molecules of EwS [50]. The cells are prone to differentiate toward neural lineage if deprived of CD99 [50]. Using relatively small sample size, three CD99-positive exosomes and four CD99-negative exosomes, the authors could decipher the repertoire in these 2 types of exosomes and identify miR-199a-3p as contributing to EwS aggressiveness.

As miRNA, lncRNAs are attractive as tissue-specific biomarkers. Marques Howarth et al. [51] performed RNA-Seq to look for novel transcripts regulated by EWSR1-FLI1 (six tumor samples). They compared pediatric human mesenchymal progenitor cells (pMPCs) expressing EWSR1-FLI1 with control pMPCs and identified 157 genes with higher expression in cells expressing EWSR1-FLI1, while only 13 genes had reduced expression. They focused primarily on genes not previously established as EWSR1-FLI1 targets, identifying 16 candidate genes, of which 15 were protein coding genes and a single lncRNA of unknown function, *EWSATI*. They showed that *EWSATI* is a lncRNA specifically upregulated as a consequence of the oncogenic fusion.

2.4 Characterizing the EwS's Epigenome

As EwS, apart from EWSR1-FLI1 fusion, rarely shows recurrent mutations, improvements in the knowledge of its epigenomic landscape can provide novel breakthrough. In this sense, DNA methylation profiling provides a valuable approach to study the states of EwS cells. Some efforts have been drawn in this direction to boost the understanding of the interplay between DNA methylation and the pathogenesis of EwS and, in general, the carcinogenesis of human malignancies.

Indeed, methylation profiling has been carried out on a relatively large set of cohorts such as on 52 EwS tumors, three cell lines and eight MSC using bead chip methylation [52]; on 69 EwS tumors by Illumina GoldenGate Methylation Cancer Panel I microarray [53]; on 15 EwS tumors, seven cell lines, ten healthy tissues and four human MSC by Infinium Human Methylation 450 K [54] and on 140 EwS tumors, 16 EwS cell lines, 32 MSCs by reduced representation bisulfite sequencing (RRBS) [55]. Sheffield et al. [55] defined a DNA methylation signature of EwS samples that resulted in an accuracy close to 100% for distinguishing EwS samples from various other tumor types. Differences between tumors were further assessed by comparing aggregate DNA methylation profiles and the methylation-based Inference of Regulatory Activity score. A continuous disease spectrum underlying EwS and between mesenchymal and stem cell signatures was identified. This is another way to say that Ewing tumors are not characterized by well-defined molecular subtypes that can be defined at the level of DNA methylation.

Since EWSR1-FLI1 is a transcription factor, an important part of the EwS omics atlas is constituted by a collection of ChIP-Seq and ATAC-Seq profiles. Historically, characterizing the binding affinity of EWSR1-FLI1 to GGAA microsatellites was a result of introducing genome-wide ChiP-on-chip technology which was later replaced by ChIP-Seq [7, 8]. EwS model systems such as A-673 cell line, its inducible modifications and related constructs based on MSC cells are relatively well-characterized in terms of genome-wide ChIP-Seq profiles for EWSR1-FLI1 itself (sometimes using antibodies against FLI1) and some other important co-factors and transcription factors (such as *MYBL2* or *E2F3*) [25, 56–58]. ChIP-Seq profiles are generated in various conditions (e.g., with activated or knocked-down oncogene), and include recently published time-resolved profiles [29].

ChIP-Seq profiling was used to characterize the state of chromatin modifications and their dependence on the activity of EWSR1-FLI1, in inducible model systems [57, 58]. In [58], an impressive effort was undertaken to chart the “epigenome map” (collection of RNA-Seq, ATAC-Seq, DNA methylation, ChiP-Seq profiles including histone mark modifications) of the A-673-derived inducible model system. The complete dataset was made public, easily available online (<http://www.medical-epigenomics.org/papers/tomazou2015/>) and was reused in a number of EwS studies.

These datasets present a unique and unprecedented opportunity to use computational and machine learning methods for the quantification of EwS heterogeneity between and within tumors at epigenetic level.

2.5 Proteomic Studies

The EWSR1-FLI1 fusion is a key driver in EwS oncogenesis. As such, downstream effectors and target proteins of EWSR1-FLI1 are likely implicated in disease pathogenesis and are thus of interest to the discovery of new biomarkers and therapeutic targets but also to the identification of protein interactions and signaling pathway partners playing a key role in the onset and progression of cancer hallmarks.

Proteomic profiling has been performed mostly on EwS cell lines: in 293, and A-673 cell lines [59] and in TC32J and CHLA10 under serum-starved conditions [60].

Few publications investigate the consequences of EWSR1-FLI1 modulation on EwS proteomic profiles. Madoz-Gurpide et al. profiled and compared the proteomic expression of TC-71 EwS cell lines relative to an EWSR1-FLI1 knockdown TC-71 cell line variant, using 2D-DIGE [61]. Franzetti et al. measured the proteomic expression of A-673 and SK-N-MC cell lines prior and after EWSR1-FLI1 knockdown [62].

Proteomic profiling has also been used to compare EwS profiles of patients from different prognosis groups, to identify synergistic

drug combinations that improve clinical efficacy and to elucidate the mechanism of acquired drug resistance. For instance, Kikuta et al. examined the proteomic profile of 8 biopsy samples from EwS patients (with good and bad retrospective prognosis) using two-dimensional difference gel electrophoresis [63].

Lamhamedi-Cherradi et al. generated 37 EwS cell lines resistant to IGF-1R- or mTOR-targeted therapy [64]. In this study, reverse-phase protein lysate arrays (RPPAs) revealed proteomic changes linked to IGF-1R/mTOR resistance, and selected proteins were validated in cell-based assays, xenografts, and within human clinical samples.

Puerto-Camacho et al. [65] investigated the therapeutic value of ENG targeting, a core receptor of the TGF β family, through characterization of ENG, sENG, and MMP14 expression by flow cytometry analysis in a panel of ten EwS cell lines and by immunohistochemistry (IHC) analysis in a set of three EwS xenografts, nine PDX models, and 43 FFPE patient samples, assaying them for the efficacy of targeted antibody therapy.

In a systems biology perspective, a new method for target discovery which can be used as surrogate tool for the analysis of the proteome has been proposed in [66]. The approach consisted in the analysis of whole cellular transcriptomes by RNA-Seq to identify candidate cell proteins. As a proof of concept, the method was applied on three EwS cell lines (A-673, TC-32, and TTC-466) and two MSC lines, and revealed a set of candidate target proteins differentially expressed in tumor cells.

2.6 Metabolomic Studies

EwS is a unique model system to study metabolic alterations caused by the oncogene and to increase the understanding of metabolic reprogramming in general, in particular the metabolic switch from oxidative to glycolytic metabolism (Warburg effect).

To our knowledge, only two studies have characterized EwS cell lines' metabolome. Jonker characterized the metabolome of A-673, SK-N-MC, and A673-C1 doxycycline inducible cell lines under different conditions [67], in EWSR1-FLI1^{HIGH} and EWSR1-FLI1^{LOW} conditions, and time-resolved metabolomics time series after inhibition of the oncogene. Metabolic analysis identified twenty-four commonly changed metabolites in different pathways, implicated in such processes as energy metabolism, the tryptophan pathway, N-glycosylation, fatty acid synthesis and glutathione metabolism. Consistently with the Warburg effect, a partial reversion from glycolysis to ATP generation through oxidative phosphorylation was observed upon EWSR1-FLI1 inhibition.

Tanner et al. characterized the metabolome of A673 cell lines after EWSR1-FLI1 silencing by shRNA and on a control knock-down [68]. In agreement with the previous study, this metabolomic analysis revealed distinct separation of metabolic profiles in EwS-knockdown versus control-knockdown cells. Metabolites in

several metabolic pathways were altered and phosphoglycerate dehydrogenase was found to be highly expressed in EwS and correlated with worse patient survival.

Characterization of the metabolome remains costly. Other strategies have been proposed in the literature to study EwS cell metabolism through exploration of targeting metabolic dependencies. Using such approach, Dasgupta et al. studied the metabolic dependencies in A673, TC-71, MHC EwS cell lines and two non-malignant cells [69]; Sen et al. in SK-N-MC, TC-32, HCT116, and HEK-293T [70]; and Issaq et al. in TC71, EW8, and 5838 cell lines [71].

2.7 Systematic Perturbation Studies, Drug and Gene Invalidation Screenings

EwS cell lines were extensively used in screenings in which either gene functions were systematically invalidated or when the cells were exposed to a set of drugs. In particular, a set of EwS cells participated in the large-scale drug sensitivity profiling projects such as screening the Cancer Cell Line Encyclopedia (CCLE) or Genomics of Drug Sensitivity in Cancer (GDSC) [72, 73]. Both screenings resulted in predicting potential drugs for EwS among the main reported results. Olaparib (PARP1 inhibitor) was suggested in the GDSC paper, and irinotecan (topoisomerase I inhibitor) appeared to be efficient for killing EwS cell lines with elevated SLFN11 expression, in the CCLE study. Both drugs were tested in clinical trials of EwS although with limited success. The sensitivity of EwS cell lines to olaparib was particularly surprising because it was thought to be efficient in cancers with increased BRCA1-dependent genomic instability while EwS cell lines usually have robust expression and no mutations in *BRCA1*. One of the possible explanations was that PARP1 appears to be a direct target of EWSR1-FLI1 and at the same time acts as its transcriptional co-factor. This creates a positive feedback loop, on which EwS cells depend, and its disruption leads to inactivation of the driver oncogene [74].

Since 2012, several more focused efforts have been made to screen vulnerabilities of EwS cells. To give some examples, siRNA-based screening in [75] targeted around 7000 genes in the low and high EWSR1-FLI1 activity conditions, highlighting the particular and clinically relevant sensitivity to *LRWD1* gene. The largest, to our knowledge, drug screening tested more than 300,000 compounds and highlighted the proteasome addiction of EwS cells [76]. The druggable interactome of EWS-FLI1 was nicely charted in [77] based on a screening by 3325 compounds in two EWSR1-FLI1 activity conditions in the form of a hallmark-like image. ‘Apoptosis’, ‘Translation’, ‘Histone deacetylation’, ‘Transcription regulation’, ‘Topoisomerase activity’ and ‘Microtubule organization’ appeared to be the ‘hallmarks’ of EwS druggability.

Combinatorial screening identified some synergistic drug effects in EwS context, in particular, the synergy between

PKC412 and IGF1R inhibitors [78]. This result is of particular interest in the light of that most of IGF1R inhibitors used alone lead to drug resistance in treating EwS. Finally, the novel genome-wide CRISPR-Cas9-based screening identified druggable dependencies in EwS cells having wild-type *TP53*, which is the representative genetic background for the majority of EwS tumors [79].

Quite interestingly, some of the EwS screenings were preceded by *in silico* predictions. Thus, the Connectivity Map database was used in order to identify those drugs whose transcriptomic signature would have a potential to “reverse” the signature of EWSR1-FLI1 [80]. Two drugs, auranofin, a thioredoxin reductase inhibitor, and ganetespib, an HSP90 inhibitor, were predicted to have anti-cancer activities *in silico* and were confirmed active across a panel of genetically diverse EwS cells. Moreover, their combined effect appeared to be synergistic.

Besides simple viability screens, EwS cells were subject to siRNA-based High-Content Screening. 672 EwS-relevant genes were invalidated followed by microscopy imaging which allowed quantifying not only the number of cells, but also distinguishing mitotic and apoptotic cells, as well as distribution of cells in different cell cycle phases, in a fully automated fashion [81].

3 Computational Systems Biology Studies of EwS

Characterizing EwS at multiple levels of molecular description allows combining different types of data in order to either validate the conclusions made in one particular dataset or apply joint integrative data analysis, finding biological signals emergent across several levels of molecular description. This is the purpose of multi-omics and integrative data analysis in cancer systems biology which becomes a major tool in deciphering the complexity of cancer disease [82].

3.1 Multi-omics EwS Datasets

Some large-scale efforts such as The Cancer Genome Atlas (TCGA) systematically collect multi-level molecular description of adult cancers, and were already subject to multi-omics data analysis [82]. Nevertheless, multi-omics datasets of pediatric cancers are still rare and contain only few samples with matched molecular profiles, e.g. the one generated for medulloblastoma [83]. In the case of EwS, several datasets exist where multiple-level of omics profiling were combined for a sufficient number of samples (e.g. more than 20). First of all, such a dataset is publicly available as a part of International Cancer Genome Consortium database (ICGC), under the reference BOCA-FR. This dataset contains molecular characterization of 98 tumor samples, of which 43 samples combine three levels of molecular description: genome (CNVs, somatic mutations, including structural somatic mutations),

transcriptome (profiled by RNA-Seq) and DNA methylation (profiled by reduced representation bisulfite sequencing).

Other examples of multi-omics tumor description in EwS include joint profiling of mRNA and miRNA expression such as in the study [84] (39 EwS tumors) or systematic characterization of cell lines available in the standard collections such as NCI Sarcoma Cell Line panel (23 EwS cell lines) or Cancer Cell Line Encyclopedia (12 EwS cell lines) [72].

Processed versions of some of the datasets for the multi-omics analyses in EwS are available as a part of the data repository associated with this chapter, https://github.com/sysbio-curie/EwingSarcoma_Omics_Atlas, and will be updated in the future.

3.2 Integrative Biology Studies Combining Several Data Types and Sources

Several studies have been performed combining several sources of data in order to address various biological questions, such as clarifying the origin of EwS cells. For example, in the study [9] integrative analysis of transcriptomic data from EwS inducible cell lines, EwS tumors, transcriptomes of MSC and other cell types, provided arguments in favor of mesenchymal stem cells as potential precursors of EwS cells. Indeed, knocking-down the EWSR1-FLI1 expression in EwS cell lines resulted in convergence of the tumor gene expression profiles to that of mesenchymal stem cells. At the same time, another integration of transcriptomic data of various origins showed that the expression profile of EwS is more similar to that of neural crest stem cells than other cell types such as mesenchymal stem cells [85]. Ten years later, the origin of EwS cells is still disputed between mesenchymal stem cells and neural crest stem cells [1]. In a recent study, more than 40,000 publicly available RNA-Seq profiles (including 260 EwS) have been re-analyzed that defined a set of ‘EwS-like’ transcriptomes corresponding to the neural crest cells, induced pluripotent stem cells, human embryonic stem cells, MSCs. The manifold learning-based analysis of EwS transcriptomes in the context of EWS-like profiles showed that the EwS transcriptome can be placed at an intermediate position on the developmental trajectory connecting pluripotent, neuroectodermal, and mesodermal cell states [30]. We believe that multi-level data integration (rather than comparing only gene expression profiles) might provide the right metric to resolve the several decades-old question on the origin of EwS cells, or suggest a new concept.

Another important question is characterizing the transcriptional response for the induction of EWSR1-FLI1, including up- and down-regulated genes. In 2008, a meta-analysis of 13 different model systems and transcriptomic profiles has been performed in order to define the consensus or core ‘transcriptomic signature’ of EWSR1-FLI1 (503 up- and 293 down-regulated genes) [86]. In another study, a molecular function map of EwS was built based on the joint analysis of EwS cell lines and tumors together with MSC

transcriptomes. This analysis highlighted distinct clusters of activities for EWSR1-FLI1 regulated genes in EwS and revealed another definition of the transcriptional EWSR1-FLI1 signature (367 up- and 252 down-regulated genes) [87]. However, both analyses were not able to distinguish direct and indirect downstream effects of EWSR1-FLI1 activity, that requires using more molecular description layers, specific mathematical modeling or single cell approaches as described below.

Few other examples of integration of several data types are represented in the genomic landscape of EwS papers described above, as well as several joint studies of coding and non-coding gene expression [45, 84].

3.3 Assessing EwS Tumor Composition Via Mathematical Modeling

Mathematical modeling of the bulk tumor transcriptome as a complex mixture of various cell types can shed light on the composition of tumor microenvironment via applying so-called computational deconvolution tools [88]. To our knowledge, there exists only one study where the state-of-the-art deconvolution method, CIBERSORT, has been applied in order to determine the relative fraction of 22 immune cell types using 197 microarray expression profiles for EwS tumors [89]. From this analysis, it followed that the most abundant type of immune cells present in the microenvironment of EwS tumors are immunosuppressive M2 type macrophages, and that increased number of neutrophils, albeit a low number, was associated to poor survival (although with a border-line statistical significance). A minority of EwS tumors appeared to be in the “hot” state, with dominating T-cells populations. Interestingly, this study pointed to a link between hypoxia and the immunogenic status of the EwS tumor, with high hypoxia been associated with the “cold” state, characterized by decreased infiltration of T-cells.

Deconvolution of DNA methylation profiles can also serve the purpose of quantifying immune tumoral composition (e.g., via MeDeCom tool [90]) but has not been applied to EwS so far. Nevertheless, some deconvolution-related approaches have been used to estimate the levels of the within-sample heterogeneity (WSH) of Ewing tumors [55]. In a recent study, the RRBS profiles of EwS were used to benchmark six different WSH measures, concluding that different WSH measures may be more suitable to quantify different aspects of WSH measured through DNA methylation (cell type heterogeneity, DNA methylation erosion, cellular contamination or allele-specific methylation) [91].

3.4 Network Modeling Approaches

The wide-spread action of EWSR1-FLI1 is caused by its properties as a potent transcription factor and as a protein able to interact with many other proteins. This oncogene is frequently referred to as a “network hub” regulating various biological mechanisms such as splicing [92]. The pleiotropic effect of EWSR1-FLI1 appeared a

difficult case from the point of view of mechanistical modeling, since the perturbation caused by its activity is distributed across major cellular functions, and, therefore, it is difficult to define the potential model borders. Some of the published studies aimed at understanding the networks transducing the immediate action of EWSR1-FLI1 downstream to the major cellular phenotypes.

For example, transcriptome dynamics upon inhibition and reactivation of EWSR1-FLI1 in the A673 cell line transformed into inducible system was used to define a list of candidate genes connecting the oncogene with apoptosis and cell cycle phenotypes [10]. The selection of the genes was achieved using a model-based approach, assuming sigmoid-like response in the gene expression, which appeared to be more sensitive than the standard fold change-based approaches and more adapted to the temporal nature of the data.

As a result, a complex influence network of downstream action of EWSR1-FLI1 has been inferred, using literature knowledge about biological interactions (*see* Fig. 1a and the interactive version at <https://navicell.curie.fr/navicell/maps/ewing/master/> created using the online network visualization platform NaviCell [93]). A limited number of genes was further selected for data-based network inference in order to validate part of the mechanistic connections between the network members. A complete perturbation-response matrix has been experimentally constructed for 11 selected genes (FOXO1A, IER3, CFLAR and their known regulators), by systematically knocking down them one by one using specific siRNAs. The response was quantified by qRT-PCR for each gene in the set. The perturbation-response matrix was binarized and analyzed together with the influence network in order to distinguish direct and indirect effects of each perturbation.

The network analysis validated a number of direct interactions between EWSR1 and FLI1 and its targets and discovered few new ones, in particular, a member of E3 ubiquitin-protein ligase complex CUL1, indicating a link between the oncogene and the protein turnover regulation in the context of EwS. The hypothesis that *CUL1* is a direct target of the oncogene was further supported by ChIP-Seq data analysis. More generally, the reconstructed network can serve as a basis for the mechanistic modeling of the EWSR1-FLI1 action, for example, using probabilistic Boolean modeling approach [94].

Another study applied mathematical modeling to better understand the functional synergy between the action of EWSR1-FLI1 and the cellular E2F dependent gene regulatory network which is the central part of the cell cycle progression mechanism [95]. The model focused on explaining the observation that knocking down EWSR1-FLI1 is accompanied by loss of *E2F3*/pRB (activator complex) and gain of *E2F4*/p130 (inhibitory complex) occupancy at E2F target promoters. The originality of the approach was in that

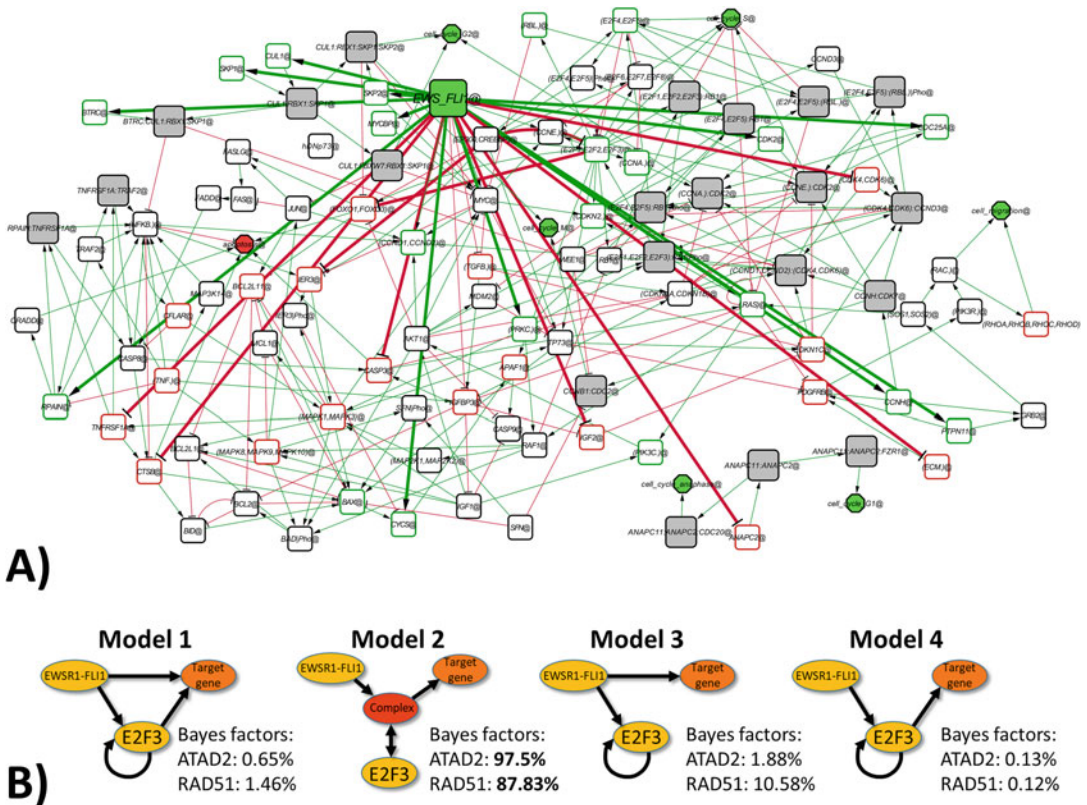


Fig. 1 Examples of mathematical modeling of EwS networks. **(a)** Reverse-engineered network of the downstream effect of EWSR1-FLI1 leading to proliferation and cell cycle phenotypes. Thick edges designate regulations inferred from transcriptome time series and siRNA/RT-QPCR data, green for activation and red for inhibition effect. White nodes are simple proteins or related groups, grey nodes are protein complexes, green pentagons are cellular phenotypes. The network image is adapted from [10]. **(b)** Use of mathematical modeling and model selection in order to test a biological hypothesis on the interplay between EWSR1-FLI1, the E2F3 transcription factor and target genes. Model 1 assumes that the EWSR1-FLI1 and E2F3 proteins independently target genes including E2F3 itself. Model 2 postulates that target gene transcription depends on the co-binding of EWSR1-FLI1 and E2F3 proteins as a complex or separately but in interdependence. Model 3 presumes EWSR1-FLI1 protein activates transcription of target genes alone without a contribution of E2F3. Finally, model 4 supposes that the EWSR1-FLI1 protein first activates the transcription of E2F3, and E2F3 protein subsequently activates transcription of target genes. (The images are adapted from [95])

four alternative mathematical models based on the standard formalism of chemical kinetics were suggested to explain this phenomenon (see Fig. 1b). The parameters of all four models were fit to the expression dynamics of four genes (*EWSR1-FLI1*, *E2F3*, and two E2F target genes *ATAD2* and *RAD51*) measured by qRT-PCR in 14 time points following the knock-down of *EWSR1-FLI1*. The Bayesian model selection approach was used to rank the models accordingly to their ability to explain the data. One of the four models was way more probable accordingly to this analysis (Fig. 1b). It predicted the synergy either through physical and/or

functional interaction between EWSR1-FLI1 and an E2F3 complex as a necessary prerequisite for combinatorial promoter binding and activation. This study provided an excellent example for the power of systems biology in the study of complex gene regulatory mechanisms that are otherwise difficult to assess experimentally.

4 Single Cell Studies of EwS

From the point of view of systems biology, emergence of technologies allowing to study biological systems at single cell level dramatically improves our understanding of the mechanisms of tumorigenesis, genetic and epigenetic tumoral heterogeneity connected to resistance to treatment. Each cell represents a possible state of a biological system under specific (even though not completely characterized) conditions. Therefore, with an advent of the single cell technologies, the amount of data available for reverse-engineering the biological mechanisms dramatically increased in the last 5 years.

The first recently published single-cell study of EwS provided a valuable resource comprising several single cell datasets [29]. First, the doxycycline-inducible system based on the A-673 cell line was profiled in 7 time points at single cell level using C1 single-cell system (Fluidigm) (383 cells at all time points), tracing the induction of the EWSR1-FLI1 from the meta-stable state where its expression was the lowest to the meta-stable state where the expression of EWSR1-FLI1 was high. RNA velocity-based analysis combined with pseudo-time quantification showed a picture of relatively rapid transition of Ewing cells between two meta-stable states EWSR1-FLI1^{LOW} and EWSR1-FLI1^{HIGH} (see Fig. 2a). For each individual cell, the duration of the transition appeared to be much shorter than the total duration of the experiment (15 days) which allowed quantification of the RNA velocity vectors. In each metastable state, EwS cells can proliferate: in the EWSR1-FLI1^{LOW} state, the proliferation appears to be possible as soon as 2 days after the doxycycline was removed from the system. In the EWSR1-FLI1^{HIGH} state, non-proliferating cells were rare. Two cellular trajectory types described the heterogeneity of the transition between two states. In the first scenario, followed by the majority of EwS cells, the activation of proliferation approximately coincided with the full activation of the oncogene or even preceded it, in the second, the activation of proliferation was delayed after the oncogene activation (Fig. 2a).

The cell line dataset was jointly analyzed with several other scRNA-Seq datasets, including EwS PDXs (142 cells) and a xenograft implanted with the inducible cell line system (215 cells). Independent Component Analysis (ICA) was applied in order to identify distinct sources of gene expression heterogeneity [96],

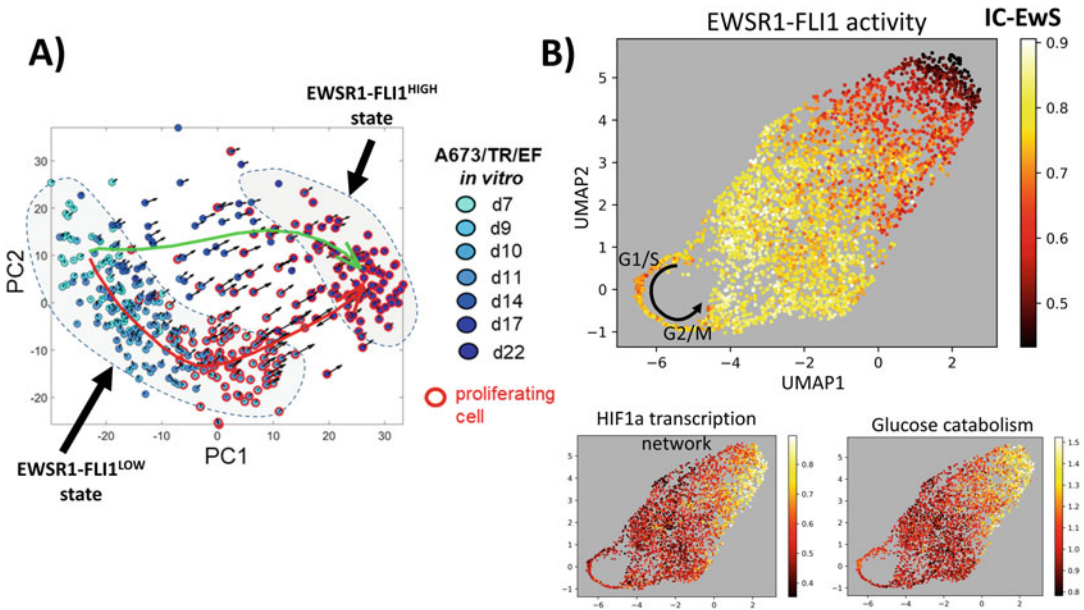


Fig. 2 Single cell study of EwS. (a) RNA velocity plot produced for the inducible cell line system. Each arrow shows a potential direction of the change of the transcriptome for a given cell. (The image is adapted from [29]). Two branches of pseudo-time (shown by green and red curves) recapitulate two types of the transcriptional dynamics after induction of the oncogene. (b) Visualization of intratumoral heterogeneity in an EwS patient-derived xenograft (data from [29]). Each point represents a single-cell tumoral transcriptome. The large panel shows cell heterogeneity caused by variability of the estimated EWSR1-FLI1 activity. The ring-like structure is formed by cells in the proliferative state. Small panels show two transcriptomic scores connected to hypoxia and metabolic heterogeneity

which allowed to distinguish the proper activating transcriptional program of EWSR1-FLI1, called IC-EwS (220 genes), from seemingly indirect and non-specific to EwS effects such as the transcriptional program of G1/S or G2/M phases or response to hypoxia. The results of the functional enrichment analysis of the identified independent components can be browsed online at http://bioinfo-out.curie.fr/projects/sitcon/mosaic/toppgene_analysis/.

The IC-EwS signature was validated by time-resolved bulk FLI1 ChIP-Seq measurements in the same time points as used for the transcriptomic profiling, and H3K27ac histone mark modification profiling in EWSR1-FLI1^{HIGH} and EWSR1-FLI1^{LOW} conditions. The conclusion was that IC-EwS is strongly enriched with direct targets of EWSR1-FLI1. The downregulation program of EWSR1-FLI1 was recapitulated in a component called IC-ECM (for extracellular matrix) but appeared to be non-specific to EwS tumors and less directly associated with the direct action of the oncogene.

Moreover, the study provided data on the intratumoral heterogeneity (ITH) at single cell level in 5 EwS PDXs profiled rather 10× Genomics (8431 cells in total). Some of the transcriptional

programs identified from the inducible cell line analysis, were shown to significantly contribute to EwS ITH, including those connected to proliferation, oxidative phosphorylation, splicing, hypoxia and IC-EwS (representing a surrogate measure of EWSR1-FLI1) itself. It appeared that a well-defined intermediate level of EWSR1-FLI1 activity was associated with cells in the proliferative state. Below and, more intriguingly, above this range, almost no proliferative cells were observed. Among the cells with low estimated activity of EWSR1-FLI1, there existed a sub-population characterized by increased hypoxia signaling and increased expression of genes involved in glycolysis. Those cells having the maximum estimated activity of EWSR1-FLI1 were also characterized by increased hypoxia, after regressing out the dominant signal connected with expression of EWSR1-FLI1 direct targets.

Three EwS cell lines (CHLA9, CHLA10, TC71) have been recently sequenced at single cell level using $10\times$ Genomics (9783 cells in total), and these data were analyzed together with the above described PDX profiles with the purpose to demonstrate the existence of mesodermal-like cell subpopulations [30]. In another, yet unpublished study, the authors performed single cell profiling of three EwS cell lines in the conditions with and without knocking down EWSR1-FLI1 using siRNA [97]. The dataset contained close to one thousand viable cells profiled using C1 single-cell system (Fluidigm). The authors reported that they identified existence of two rare subpopulations in EwS cells: dormant-like and neural stem-like in the EWSR1-FLI1^{HIGH} state, with distinct population dynamics after knocking-down the driver oncogene. It was suggested that the existence of these rare states can provide a survival mechanism upon the stress caused by the inhibition of EWSR1-FLI1.

Overall, it seems that single cell studies (not only of transcriptome but also other modalities such as scATAC-Seq) can provide insights on the structure of intratumoral heterogeneity and shed light on the mechanisms connecting EWSR1-FLI1 activity and major cellular phenotypes such as proliferation. Ongoing single cell profiling of EwS tumors as a part of EU Horizon-2020 iPae-diaticCure project should provide new insights in the mechanisms of interaction of EwS cells with the major actors of tumoral micro-environment, including immune cells.

5 Conclusion

Independently on its clinical significance, EwS is an outstanding cancer type in the light of cancer systems biology for several reasons. It is one of the most genetically stable and homogeneous cancer, lacking clear relation to normal tissues and characterized by

a single known cancer driver event. There exist multiple evidences that EWSR1-FLI1 blocks some normal developmental cellular trajectories. EwS was relatively well characterized at the molecular level in the last 15 years, partially thanks to the European-level collaborative efforts, bringing together computational and cancer biologists. Despite this, researchers from both fields still seem to be far from having a mechanistically complete picture of the connection between the fact of appearance of the particular genomic fusion causing EwS and the downstream shifts in the functioning of the major cellular mechanisms at multiple levels.

The reason for this appears to be the systemic action of the chimeric oncogene, such that its complexity evades simple intuition and the usual reductionist approach. It seems that having so many molecular clues in hand, we must have already reverse-engineered the “EWSR1-FLI1 pathway”, if the action of EWSR1-FLI1 could be reduced to dysregulation of a small number of key regulators of cellular life. The biological reality, however, seems to be more complex than this simplistic representation. This complexity is reflected in that the molecular studies of EwS, besides genomic ones, do not seem to converge to a limited set of “principal molecular players” involved in this disease.

A recent study, based on single-cell analysis of the well-studied inducible cellular EwS system, resulted in a definition of the proper transcriptional signature of EWSR1-FLI1, designated as IC-EwS, statistically independent from its indirect downstream effects on cell cycle, organization of extracellular matrix, regulation of RNA splicing, etc. [29]. This signature appeared to be more specific to the EwS tumors than any other previously suggested signature. EWSR1-FLI1 as a transcriptional activator binds to the repetitive sequences more or less randomly distributed across the genome. As one would expect from this, the IC-EwS signature is not enriched with any particular biological process or function. Despite this, the perturbation caused by EWSR1-FLI1 fusion and expression seems to have profound and consistent changes, collectively pushing the EwS cell towards the cancerous phenotype. This apparent discrepancy between a disorganized nature of the perturbation and the well-defined malignant outcome can be called the EwS enigma, both from the biological and the systems theory points of view.

In order to resolve this enigma, we might need to understand some yet unknown principles of cellular adaptation and selection, at epigenetic level, to the drastic changes in the topology of the connections between the modules of the global molecular network. These principles should be properly formalized in the language of mathematics, because standard descriptions, such as chemical kinetics or logical formalisms, appear to be poorly suited to this situation. In order to advance in understanding cancer in general and rationalizing its treatment, the genesis of EwS might serve us as an important prototypical real-life scenario for application of the systems biology approach.

Acknowledgements

This work was supported by the European Union's Horizon 2020 program (grant 826121, iPC project) and the Agence Nationale de la Recherche as part of the "Investissements d'avenir" program (reference ANR-19-P3IA-0001; PRAIRIE 3IA Institute).

References

1. Grunewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, Sorensen PH, Delattre O, Dirksen U (2018) Ewing sarcoma. *Nat Rev Dis Primers* 4(1):5
2. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH, Pugh TJ, Stransky N, Helman E, Kim J, Sougnez C, Ambrogio L, Nickerson E, Shefler E, Cortes ML, Auclair D, Saksena G, Voet D, Noble M, DiCara D, Lin P, Lichtenstein L, Heiman DI, Fennell T, Imielinski M, Hernandez B, Hodis E, Baca S, Dulak AM, Lohr J, Landau DA, Wu CJ, Melendez-Zajgla J, Hidalgo-Miranda A, Koren A, SA MC, Mora J, Crompton B, Onofrio R, Parkin M, Winckler W, Ardlie K, Gabriel SB, CWM R, Biegel JA, Stegmaier K, Bass AJ, Garraway LA, Meyerson M, Golub TR, Gordenin DA, Sunyaev S, Lander ES, Getz G (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499(7457):214–218
3. Boulay G, Sandoval GJ, Riggi N, Iyer S, Buisson R, Naigles B, Awad ME, Rengarajan S, Volorio A, McBride MJ, Broye LC, Zou L, Stamenkovic I, Kadoch C, Rivera MN (2017) Cancer-specific retargeting of BAF complexes by a prion-like domain. *Cell* 171(1):163–178.e19
4. Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S, Albalá JS, Lim J, Fraughton C, Llamosas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S, Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP, Vidal M (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437(7062):1173–1178
5. Romeo S, Dei Tos AP (2010) Soft tissue tumors associated with EWSR1 translocation. *Virchows Arch* 456(2):219–234
6. Li Y, Luo H, Liu T, Zacksenhaus E, Ben-David Y (2015) The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene* 34(16):2022–2031
7. Gangwal K, Sankar S, Hollenhorst PC, Kinsey M, Haroldsen SC, Shah AA, Boucher KM, Watkins WS, Jorde LB, Graves BJ, Lessnick SL (2008) Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proc Natl Acad Sci U S A* 105(29):10149–10154
8. Guillon N, Tirode F, Boeva V, Zynovyev A, Barillot E, Delattre O (2009) The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PLoS One* 4(3):e4932
9. Tirode F, Laud-Duval K, Prieur A, Delorme B, Charbord P, Delattre O (2007) Mesenchymal stem cell features of Ewing tumors. *Cancer Cell* 11(5):421–429
10. Stoll G, Surdez D, Tirode F, Laud K, Barillot E, Zynovyev A, Delattre O (2013) Systems biology of Ewing sarcoma: a network model of EWS-FLI1 effect on proliferation and apoptosis. *Nucleic Acids Res* 41(19):8853–8871
11. Kovar H, Amatruda J, Brunet E, Burdach S, Cidre-Aranaz F, de Alava E, Dirksen U, van der Ent W, Grohar P, Grunewald TGP, Helman L, Houghton P, Iljin K, Korsching E, Ladanyi M, Lawlor E, Lessnick S, Ludwig J, Meltzer P, Metzler M, Mora J, Moriggl R, Nakamura T, Papamarkou T, Radic Sarikas B, Redini F, Richter GHS, Rossig C, Schadler K, Schafer BW, Scotlandi K, Sheffield NC, Shelat A, Snaar-Jagalska E, Sorensen P, Stegmaier K, Stewart E, Sweet-Cordero A, Szuhai K, Tirado OM, Tirode F, Toretsky J, Tsaoufou K, Uren A, Zynovyev A, Delattre O (2016) The second European interdisciplinary Ewing sarcoma research summit—A joint effort to deconstructing the multiple layers of a complex disease. *Oncotarget* 7(8):8613–8624

12. Maguire LH, Thomas AR, Goldstein AM (2015) Tumors of the neural crest: common themes in development and cancer. *Dev Dyn* 244(3):311–322
13. Sand LGL, Szuhai K, Hogendoorn PCW (2015) Sequencing overview of Ewing sarcoma: A journey across genomic, Epigenomic and transcriptomic landscapes. *Int J Mol Sci* 16 (7):16176–16215
14. Amiel A, Ohali A, Fejgin M, Sardos-Albertini F, Bouaron N, Cohen I, Yaniv I, Zaizov R, Avigad S (2003) Molecular cytogenetic parameters in Ewing sarcoma. *Cancer Genet Cytogenet* 140 (2):107–112
15. Armengol G, Tarkkanen M, Virolainen M, Forus A, Valle J, Bohling T, Asko-Seljavaara S, Blomqvist C, Elomaa I, Karaharju E et al (1997) Recurrent gains of 1q, 8 and 12 in the Ewing family of tumours by comparative genomic hybridization. *Br J Cancer* 75 (10):1403–1409
16. Brisset S, Schleiermacher G, Peter M, Mairal A, Oberlin O, Delattre O, Aurias A (2001) Cgh analysis of secondary genetic changes in Ewing tumors: correlation with metastatic disease in a series of 43 cases. *Cancer Genet Cytogenet* 130 (1):57–61
17. Ozaki T, Paulussen M, Poremba C, Brinkschmidt C, Rerin J, Ahrens S, Homann C, Hillmann A, Wai D, Schaefer KL et al (2001) Genetic imbalances revealed by comparative genomic hybridization in Ewing tumors. *Genes Chromosom Cancer* 32 (2):164–171
18. Tarkkanen M, Kiuru-Kuhlefelt S, Blomqvist C, Armengol G, Bohling T, Ekfors T, Virolainen M, Lindholm P, Monge O, Picci P et al (1999) Clinical correlations of genetic changes by comparative genomic hybridization in Ewing sarcoma and related tumors. *Cancer Genet Cytogenet* 114(1):35–41
19. Ferreira B, Alonso J, Carrillo J, Acquadro F, Largo C, Suela J, Teixeira M, Cerveira N, Molares A, Gomez-Lopez G et al (2008) Array cgh and gene-expression profiling reveals distinct genomic in-stability patterns associated with dna repair and cell-cycle checkpoint pathways in Ewing's sarcoma. *Oncogene* 27 (14):2084–2090
20. Savola S, Klami A, Tripathi A, Niini T, Serra M, Picci P, Kaski S, Zambelli D, Scotlandi K, Knuutila S (2009) Combined use of expression and cgh arrays pinpoints novel candidate genes in Ewing sarcoma family of tumors. *BMC Cancer* 9(1):17
21. Mackintosh C, Ordóñez JL, García-Domínguez DJ, Sevillano V, Llombart-Bosch A, Szuhai K, Scotlandi K, Alberghini M, Sciort R, Sinnaeve F, Hogendoorn PCW, Picci P, Knuutila S, Dirksen U, Debiec-Rychter M, Schaefer KL, de Alava E (2012) 1q gain and CDT2 overexpression underlie an aggressive and highly proliferative form of Ewing sarcoma. *Oncogene* 31(10):1287–1298
22. Postel-Vinay S, Veron AS, Tirode F, Pierron G, Reynaud S, Kovar H, Oberlin O, Lapouble E, Ballet S, Lucchesi C, Kontny U, Gonzalez-Neira A, Picci P, Alonso J, Patino-Garcia A, de Paillerets BB, Laud K, Dina C, Froguel P, Clavel-Chapelon F, Doz F, Michon J, Chanock SJ, Thomas G, Cox DG, Delattre O (2012) Common variants near TARDBP and EGR2 are associated with susceptibility to Ewing sarcoma. *Nat Genet* 44(3):323–327
23. Machiela MJ, Grunewald TGP, Surdez D, Reynaud S, Mirabeau O, Karlins E, Rubio RA, Zaidi S, Grossetete-Lalami S, Ballet S, Lapouble E, Laurence V, Michon J, Pierron G, Kovar H, Gaspar N, Kontny U, Gonzalez-Neira A, Picci P, Alonso J, Patino-Garcia A, Corradini N, Berard PM, Freedman ND, Rothman N, Dagnall CL, Burdett L, Jones K, Manning M, Wyatt K, Zhou W, Yeager M, Cox DG, Hoover RN, Khan J, Armstrong GT, Leisenring WM, Bhatia S, Robison LL, Kulozik AE, Kriebel J, Meitinger T, Metzler M, Hartmann W, Strauch K, Kirchner T, Dirksen U, Morton LM, Mirabello L, Tucker MA, Tirode F, Chanock SJ, Delattre O (2018) Genome-wide association study identifies multiple new loci associated with Ewing sarcoma susceptibility. *Nat Commun* 9(1):3184
24. Grunewald TGP, Bernard V, Gilardi-Hebenstreit P, Raynal V, Surdez D, Aynaud MM, Mirabeau O, Cidre-Aranaz F, Tirode F, Zaidi S, Perot G, Jonker AH, Lucchesi C, Le Deley MC, Oberlin O, Marec-Berard P, Veron AS, Reynaud S, Lapouble E, Boeva V, Rio Frio T, Alonso J, Bhatia S, Pierron G, Cancel-Tassin G, Cussenot O, Cox DG, Morton LM, Machiela MJ, Chanock SJ, Charnay P, Delattre O (2015) Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite. *Nat Genet* 47 (9):1073–1078
25. Musa J, Cidre-Aranaz F, Aynaud MM, Orth MF, Knott MML, Mirabeau O, Mazor G, Varon M, Holting TLB, Grossetete S, Gartlgruber M, Surdez D, Gerke JS, Ohmura S, Marchetto A, Dallmayer M, Baldauf MC, Stein S, Sannino G, Li J, Romero-Perez L, Westermann F, Hartmann W, Dirksen U, Gylmek M, Anderson ND, Shlien A, Rotblat B, Kirchner T, Delattre O, TGP G (2019) Cooperation of cancer drivers

- with regulatory germline variants shapes clinical outcomes. *Nat Commun* 10(1):1–10
26. Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, Bahrami A, Zhang Z, Lapouble E, Reynaud S, Rusch M et al (2014) Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of *stag2* and *tp53* mutations. *Cancer Discov* 4(11):1342–1353
 27. Brohl AS, Solomon DA, Chang W, Wang J, Song Y, Sindiri S, Patidar R, Hurd L, Chen L, Shern JF et al (2014) The genomic landscape of the Ewing sarcoma family of tumors reveals recurrent *stag2* mutation. *PLoS Genet* 10(7):e1004475
 28. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, Kiezun A, Carter SL, Shukla SA, Mehta SS et al (2014) The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov* 4(11):1326–1341
 29. Aynaud MM, Mirabeau O, Gruel N, Grossetete S, Boeva V, Durand S, Surdez D, Saulnier O, Za di S, Gribkova S, Fouche A, Kairov U, Raynal V, Tirode F, Grunewald TG, Bohec M, Baulande S, Janoueix-Lerosey I, Vert JP, Barillot E, Delattre O, Zinovyev A (2020) Transcriptional programs define intratumoral heterogeneity of Ewing sarcoma at single-cell resolution. *Cell Rep* 30(6):1767–1779.e6
 30. Miller HE, Gorthi A, Bassani N, Lawrence LA, Iskra BS, Bishop AJR (2020) Reconstruction of Ewing sarcoma developmental context from mass-scale transcriptomics reveals characteristics of EWSR1-FLI1 permissibility. *Cancers* 12(4):948
 31. Savola S, Klami A, Myllykangas S, Manara C, Scotlandi K, Picci P, Knuutila S, Vakkila J (2011) High expression of complement component 5 (C5) at tumor site associates with superior survival in Ewing's sarcoma family of tumour patients. *ISRN Oncol* 2011:1–10
 32. Volchenboum SL, Andrade J, Huang L, Bar-kauskas DA, Krailo M, Womer RB, Ranft A, Potratz J, Dirksen U, Triche TJ, Lawlor ER (2015) Gene expression profiling of Ewing sarcoma tumors reveals the prognostic importance of tumor-stromal interactions: A report from the children's oncology group. *J Pathol Clin Res* 1(2):83–94
 33. Schaefer KL, Eisenacher M, Braun Y, Brachwitz K, Wai DH, Dirksen U, Lanvers-Kaminsky C, Juergens H, Herrero D, Stegmaier S, Koscielniak E, Eggert A, Nathrath M, Gosheger G, Schneider DT, Bury C, Diallo-Danebrock R, Ottaviano L, Gabbert HE, Poremba C (2008) Microarray analysis of Ewing's sarcoma family of tumours reveals characteristic gene expression signatures associated with metastasis and resistance to chemotherapy. *Eur J Cancer* 44(5):699–709
 34. Scotlandi K, Remondini D, Castellani G, Manara MC, Nardi F, Cantiani L, Francesconi M, Mercuri M, Caccari AM, Serra M, Knuutila S, Picci P (2009) Overcoming resistance to conventional drugs in Ewing sarcoma and identification of molecular predictors of outcome. *J Clin Oncol* 27(13):2209–2216
 35. Svoboda LK, Harris A, Bailey NJ, Schwentner R, Tomazou E, von Levetzow C, Magnuson B, Ljung-man M, Kovar H, Lawlor ER (2014) Overexpression of HOX genes is prevalent in Ewing sarcoma and is associated with altered epigenetic regulation of developmental transcription programs. *Epigenetics* 9(12):1613–1625
 36. Ban J, Jug G, Mestdagh P, Schwentner R, Kauer M, Aryee DN, Schaefer KL, Nakatani F, Scotlandi K, Reiter M et al (2011) Hsa-mir-145 is the top *ews-flil*-repressed microRNA involved in a positive feedback loop in Ewing's sarcoma. *Oncogene* 30(18):2173–2180
 37. De Vito C, Riggi N, Suva ML, Janiszewska M, Horlbeck J, Baumer K, Provero P, Stamenkovic I (2011) *Let-7a* is a direct *ews-flil* target implicated in Ewing's sarcoma development. *PLoS One* 6(8):e23592
 38. McKinsey E, Parrish J, Irwin A, Niemeyer B, Kern H, Birks D, Jedlicka P (2011) A novel onco-genic mechanism in Ewing sarcoma involving *igf* pathway targeting by *ews/il*-regulated microRNAs. *Oncogene* 30(49):4910–4920
 39. Kawano M, Tanaka K, Itonaga I, Iwasaki T, Tsumura H (2015) *C-myc* represses tumor-suppressive microRNAs, *let-7a*, *mir-16* and *mir-29b*, and induces cyclin d2-mediated cell proliferation in Ewing's sarcoma cell line. *PLoS One* 10(9):e0138560
 40. Kawano M, Tanaka K, Itonaga I, Iwasaki T, Tsumura H (2016) MicroRNA-301a promotes cell proliferation via *pten* targeting in Ewing's sarcoma cells. *Int J Oncol* 48(4):1531–1540
 41. Kawano M, Tanaka K, Itonaga I, Iwasaki T, Tsumura H (2017) MicroRNA-20b promotes cell proliferation via targeting of *tgf-receptor ii* and upregulates *myc* expression in Ewing's sarcoma cells. *Int J Oncol* 51(6):1842–1850
 42. Kawano M, Tanaka K, Itonaga I, Iwasaki T, Tsumura H (2018) MicroRNA-181c prevents apoptosis by targeting of *fas* receptor in Ewing's sarcoma cells. *Cancer Cell Int* 18(1):37
 43. Tanaka K, Kawano M, Itonaga I, Iwasaki T, Miyazaki M, Ikeda S, Tsumura H (2016)

- Tumor suppressive microRNA-138 inhibits metastatic potential via the targeting of focal adhesion kinase in Ewing's sarcoma cells. *Int J Oncol* 48(3):1135–1144
44. Parafioriti A, Bason C, Armiraglio E, Calciano L, Daolio PA, Berardocco M, Di Bernardo A, Colosimo A, Luksch R, Berardi AC (2016) Ewing's sarcoma: an analysis of miRNA expression profiles and target genes in paraffin-embedded primary tumor tissue. *Int J Mol Sci* 17(5):656
 45. Liu Y, Chen G, Liu H, Li Z, Yang Q, Gu X, Du Z, Zhang G, Wang J (2019) Integrated bioinformatics analysis of miRNA expression in Ewing sarcoma and potential regulatory effects of mir-21 via targeting *alcam/cd166*. *Artif Cells Nanomed Biotechnol* 47(1):2114–2122
 46. Nakatani F, Ferracin M, Manara MC, Ventura S, Del Monaco V, Ferrari S, Alberghini M, Grilli A, Knuutila S, Schaefer KL et al (2012) Mir-34a predicts survival of Ewing's sarcoma patients and directly influences cell chemo-sensitivity and malignancy. *J Pathol* 226(5):796–805
 47. Karnuth B, Dedy N, Spieker T, Lawlor ER, Gattenlohner S, Ranft A, Dirksen U, Jurgens H, Brauninger A (2014) Differentially expressed miRNAs in Ewing sarcoma compared to mesenchymal stem cells: low mir-31 expression with effects on proliferation and invasion. *PLoS One* 9(3):e93067
 48. Teicher BA, Polley E, Kunkel M, Evans D, Silvers T, Delosh R, Laudeman J, Ogle C, Reinhart R, Selby M et al (2015) Sarcoma cell line screen of oncology drugs and investigational agents identifies patterns associated with gene and microRNA expression. *Mol Cancer Ther* 14(11):2452–2462
 49. De Feo A, Sciandra M, Ferracin M, Felicetti F, Astol A, Pignochino Y, Picci P, Care A, Scottlandi K (2019) Exosomes from cd99-deprived Ewing sarcoma cells reverse tumor malignancy by inhibiting cell migration and promoting neural differentiation. *Cell Death Dis* 10(7):1–15
 50. Rocchi A, Manara MC, Sciandra M, Zambelli D, Nardi F, Nicoletti G, Garofalo C, Meschini S, Astol A, Colombo MP et al (2010) Cd99 inhibits neural differentiation of human Ewing sarcoma cells and thereby contributes to oncogenesis. *J Clin Invest* 120(3):668–680
 51. Howarth MM, Simpson D, Ngok SP, Nieves B, Chen R, Siprashvili Z, Vaka D, Breese MR, Crompton BD, Alexe G et al (2014) Long noncoding rna *ewsat1*-mediated gene repression facilitates Ewing sarcoma oncogenesis. *J Clin Invest* 124(12):5275–5290
 52. Patel N, Black J, Chen X, Marcondes AM, Grady WM, Lawlor ER, Borinstein SC (2012) DNA methylation and gene expression profiling of Ewing sarcoma primary tumors reveal genes that are potential targets of epigenetic inactivation. DOI <https://doi.org/10.1155/2012/498472>, <https://www.hindawi.com/journals/sarcoma/2012/498472/>, iSSN: 1357-714X library catalog: www.hindawi.com Pages: e498472 Publisher: Hindawi volume: 2012
 53. Park HR, Jung WW, Kim HS, Park YK (2014) Microarray-based DNA methylation study of Ewing's sarcoma of the bone. *Oncol Lett* 8(4):1613–1617
 54. Huertas-Mart Nez J, Court F, Rello-Varona S, Herrero-Mart ND, Almacellas-Rabaiget O, Sainz-Jaspeado M, Garcia-Monclus S, Lagares-Tena L, Buj R, Hontecillas-Prieto L, Sastre A, Azorin D, Sanjuan X, Lopez-Aleman R, Moran S, Roma J, Gallego S, Mora J et al (2017) DNA methylation profiling identifies PTRF/Cavin-1 as a novel tumor suppressor in Ewing sarcoma when co-expressed with caveolin-1. *Cancer Lett* 386:196–207
 55. Sheffield NC, Pierron G, Klughammer J, Datlinger P, Schonegger A, Schuster M, Hadler J, Surdez D, Guillemot D, Lapouble E, Freneaux P, Champigneulle J, Bouvier R, Walder D, Ambros IM, Hutter C, Sorz E, Amaral AT, de Alava E, Schallmoser K, Strunk D, Rinner B, Liegl-Atzwanger B, Huppertz B, Leithner A, de Pinieux G, Terrier P, Laurence V, Michon J, Ladenstein R, Holter W, Windhager R, Dirksen U, Ambros PF, Delattre O, Kovar H, Bock C, Tomazou EM (2017) DNA methylation heterogeneity defines a disease spectrum in Ewing sarcoma. *Nat Med* 23(3):386–395
 56. Bilke S, Schwentner R, Yang F, Kauer M, Jug G, Walker RL, Davis S, Zhu YJ, Pineda M, Meltzer PS, Kovar H (2013) Oncogenic ETS fusions deregulate E2F3 target genes in Ewing sarcoma and prostate cancer. *Genome Res* 23(11):1797–1809
 57. Riggi N, Knoechel B, Gillespie SM, Rheinbay E, Boulay G, Suva ML, Rossetti NE, Boonseng WE, Oksuz O, Cook EB, Formey A, Patel A, Gymrek M, Thapar V, Deshpande V, Ting DT, Hornicek FJ, Nielsen GP, Stamenkovic I, Aryee MJ, Bernstein BE, Rivera MN (2014) EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* 26(5):668–681
 58. Tomazou EM, Sheffield NC, Schmidl C, Schuster M, Schonegger A, Datlinger P,

- Kubicek S, Bock C, Kovar H (2015) Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS-FLI1. *Cell Rep* 10(7):1082–1095
59. Elzi DJ, Song M, Hakala K, Weintraub ST, Shiiro Y (2014) Proteomic analysis of the EWS-FLI-1 interactome reveals the role of the lysosome in EWS-FLI-1 turnover. *J Proteome Res* 13(8):3783–3791
 60. Hawkins AG, Basur V, da Veiga LF, Pedersen E, Sperring C, Nesvizhskii AI, Lawlor ER (2018) The Ewing sarcoma secretome and its response to activation of wnt/beta-catenin signaling. *Mol Cell Proteomics* 17(5):901–912
 61. Madoz-Gurpide J, Herrero-Martin D, Gomez-Lopez G, Hontecillas-Prieto L, Biscuola M, Chamizo C, Garcia-Dominguez D, Marcilla D, Amaral AT, Ordóñez JL, Ed A (2016) Proteomic profiling of Ewing sarcoma reveals a role for TRAF6 in proliferation and ribonucleoproteins/RNA processing. *J Proteomics Bioinform* 9(6):1–10
 62. Franzetti GA, Laud-Duval K, van der Ent W, Brisac A, Irondelle M, Aubert S, Dirksen U, Bouvier C, de Pinieux G, Snaar-Jagalska E, Chavrier P, Delattre O (2017) Cell-to-cell heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in Ewing sarcoma cells. *Oncogene* 36(25):3505–3514
 63. Kikuta K, Tochigi N, Shimoda T, Yabe H, Morioka H, Toyama Y, Hosono A, Beppu Y, Kawai A, Hirohashi S, Kondo T (2009) Nucleophosmin as a candidate prognostic biomarker of Ewing's sarcoma revealed by proteomics. *Clin Cancer Res* 15(8):2885–2894
 64. Lamhamedi-Cherradi SE, Menegaz BA, Ramamoorthy V, Vishwamitra D, Wang Y, Maywald RL, Buford AS, Fokt I, Skora S, Wang J, Naing A, Lazar AJ, Rohren EM, Daw NC, Subbiah V, Benjamin RS, Ratan R, Priebe W, Mikos AG, Amin HM, Ludwig JA (2016) IGF-1R and mTOR blockade: novel resistance mechanisms and synergistic drug combinations for Ewing sarcoma. *J Natl Cancer Inst* 108(12):djw182
 65. Puerto-Camacho P, Amaral AT, Lamhamedi-Cherradi SE, Menegaz BA, Castillo-Ecija H, Ordóñez JL, Domínguez S, Jordan-Perez C, Diaz-Martin J, Romero-Pérez L, Lopez-Alvarez M, Civantos-Jubera G, Robles-Frías MJ, Biscuola M, Ferrer C, Mora J, Cuglievan B, Schadler K, Seifert O, Kontermann R, Pfizenmaier K, Simón L, Fabre M, Carcaboso AM, Ludwig JA, Ed A (2019) Preclinical efficacy of endoglin-targeting antibody–drug conjugates for the treatment of Ewing sarcoma. *Clin Cancer Res* 25(7):2228–2240
 66. Town J, Pais H, Harrison S, Stead LF, Bataille C, Bunjobpol W, Zhang J, Rabbitts TH (2016) Exploring the surfaceome of Ewing sarcoma identifies a new and unique therapeutic target. *Proc Natl Acad Sci* 113(13):3603–3608
 67. Jonker A (2014) Synthetic Lethality and Metabolism in Ewing Sarcoma: Knowledge Through Silence. PhD thesis, Paris 11. <http://www.theses.fr/2014PA11T039>
 68. Tanner JM, Bensard C, Wei P, Krahn NM, Schell JC, Gardiner J, Schiman J, Lessnick SL, Rutter J (2017) EWS/FLI is a master regulator of metabolic reprogramming in Ewing sarcoma. *Mol Cancer Res* 15(11):1517–1530
 69. Dasgupta A, Trucco M, Rainusso N, Bernardi RJ, Shuck R, Kurenbekova L, Loeb DM, Yustein JT (2017) Metabolic modulation of Ewing sarcoma cells inhibits tumor growth and stem cell properties. *Oncotarget* 8(44):77292–77308
 70. Sen N, Cross AM, Lorenzi PL, Khan J, Gryder BE, Kim S, Caplen NJ (2018) EWS-FLI1 reprograms the metabolism of Ewing sarcoma cells via positive regulation of glutamine import and serine-glycine biosynthesis. *Mol Carcinog* 57(10):1342–1357
 71. Issaq SH, Mendoza A, Fox SD, Helman LJ (2019) Glutamine synthetase is necessary for sarcoma adaptation to glutamine deprivation and tumor growth. *Oncogenesis* 8(3):1–12
 72. Barretina J, Caponigro G, Stransky N, Venkatesan K, AA M, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jane-Valbuena J, FA M, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P, de Silva M, Jagtap K, Jones MD, Wang L, Hatton C, Palesscandolo E, Gupta S, Mahan S, Sounguez C, Onofrio RC, Liefeld T, MacConaill L, Winckler W, Reich M, Li N, Mesirov JP, Gabriel SB, Getz G, Ardlie K, Chan V, Myer VE, Weber BL, Porter J, LA G (2012) The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483(7391):603–607
 73. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, Greninger P, Thompson IR, Luo X, Soares J, Liu Q, Iorio F, Surdez D, Chen L, Milano RJ, Bignell GR, Tam AT, Davies H, Stevenson JA, Barthorpe S, Lutz SR, Kogera F, Lawrence K, McLaren-Douglas A, Mitropoulos X, Mironenko T, Thi H, Richardson L, Zhou W, Jewitt F, Zhang T, O'Brien P, Boisvert JL, Price S, Hur W, Yang W, Deng X, Butler A, Choi HG, Chang JW, Baselga J, Stamenkovic I, Engelman JA, Sharma SV, Delattre O, Saez-

- Rodriguez J, Gray NS, Settleman J, Futreal PA, Haber DA, Stratton MR, Ramaswamy S, McDermott U, Benes CH (2012) Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 483 (7391):570–575
74. Brenner JC, Feng FY, Han S, Patel S, Goyal SV, Bou-Maroun LM, Liu M, Lonigro R, Prensner JR, Tomlins SA, Chinnaiyan AM (2012) PARP-1 inhibition as a targeted strategy to treat Ewing's sarcoma. *Cancer Res* 72 (7):1608–1613
 75. He T, Surdez D, Rantala JK, Haapa-Paananen S, Ban J, Kauer M, Tomazou E, Fey V, Alonso J, Kovar H, Delattre O, Iljin K (2017) High-throughput RNAi screen in Ewing sarcoma cells identifies leucine rich repeats and WD repeat domain containing 1 (LRWD1) as a regulator of EWS-FLI1 driven cell viability. *Gene* 596:137–146
 76. Shukla N, Somwar R, Smith RS, Ambati S, Munoz S, Merchant M, D'Arcy P, Wang X, Kobos R, Antczak C, Bhinder B, Shum D, Radu C, Yang G, Taylor BS, Ng CK, Weigelt B, Khodos I, De Stanchina E, Reis-Filho JS, Ouerfelli O, Linder S, Djaballah H, Ladanyi M (2016) Proteasome addiction defined in Ewing sarcoma is effectively targeted by a novel class of 19S proteasome inhibitors. *Cancer Res* 76(15):4525–4534
 77. Tsafou K, Katschnig AM, Radic-Sarikas B, Mutz CN, Iljin K, Schwentner R, Kauer MO, Muhlbacher K, Aryee DN, Westergaard D, Haapa-Paananen S, Fey V, Superti-Furga G, Toretzky J, Brunak S, Kovar H (2018) Identifying the druggable interactome of EWS-FLI1 reveals MCL-1 dependent differential sensitivities of Ewing sarcoma cells to apoptosis inducers. *Oncotarget* 9(57):31018–31031
 78. Branka RS, Kalliopi PT, Kristina BE, Theodore P, Huber KV, Cornelia M, Jerzy AT, Keiryn LB, Jesper VO, Sren B, Heinrich K, Giulio SF (2017) Combinatorial drug screening identifies Ewing sarcoma-specific sensitivities. *Mol Cancer Ther* 16 (1):88–101
 79. Stolte B, Iniguez AB, Dharia NV, Robichaud AL, Conway AS, Morgan AM, Alexe G, Schauer NJ, Liu X, Bird GH, Tsherniak A, Vazquez F, Buhrlage SJ, Walensky LD, Stegmaier K (2018) Genome-scale CRISPR-Cas9 screen identifies druggable dependencies in TP53 wild-type Ewing sarcoma. *J Exp Med* 215(8):2137–2155
 80. Passetto ZY, Chen B, Alturkmani H, Hyter S, Flynn CA, Baltezor M, Ma Y, Rosenthal HG, Neville KA, Weir SJ, Butte AJ, Godwin AK (2017) In silico and in vitro drug screening identifies new therapeutic approaches for Ewing sarcoma. *Oncotarget* 8(3):4079–4095
 81. Pauwels E, Surdez D, Stoll G, Lescure A, Del Nery E, Delattre O, Stoven V (2012) A probabilistic model for cell population phenotyping using HCS data. *PLoS One* 7(8):1–12
 82. Chakraborty S, Hosen MI, Ahmed M, Shekhar HU (2018) Onco-multi-OMICS approach: A new frontier in cancer research. *Biomed Res Int* 2018:9836256
 83. Forget A, Martignetti L, Puget S, Calzone L, Brabetz S, Picard D, Montagud A, Liva S, Sta A, Dingli F, Arras G, Rivera J, Loew D, Besnard A, Lacombe J, Pages M, Varlet P, Dufour C, Yu H, Mercier AL, Indersie E, Chivet A, Leboucher S, Sieber L, Beccaria K, Gombert M, Meyer FD, Qin N, Bartl J, Chavez L, Okonechnikov K, Sharma T, Thatikonda V, Bourdeaut F, Pouppnot C, Ramaswamy V, Korshunov A, Borkhardt A, Reifenberger G, Pouillet P, Taylor MD, Kool M, Pster SM, Kawauchi D, Barillot E, Remke M, Ayrault O (2018) Aberrant ERBB4-SRC signaling as a Hallmark of group 4 Medulloblastoma revealed by integrative Phosphoproteomic Profiling. *Cancer Cell* 34 (3):379–395.e7
 84. Martignetti L, Laud-Duval K, Tirode F, Pierron G, Reynaud S, Barillot E, Delattre O, Zinovyev A (2012) Antagonism pattern detection between MicroRNA and target expression in Ewing's sarcoma. *PLoS One* 7(7):e41770
 85. von Levetzow C, Jiang X, Gweye Y, von Levetzow G, Hung L, Cooper A, Hsu JHR, Lawlor ER (2011) Modeling initiation of Ewing sarcoma in human neural crest cells. *PLoS One* 6(4):e19305
 86. Hancock JD, Lessnick SL (2008) A transcriptional profiling meta-analysis reveals a core EWS-FLI gene expression signature. *Cell Cycle* 7(2):250–256
 87. Kauer M, Ban J, Koer R, Walker B, Davis S, Meltzer P, Kovar H (2009) A molecular function map of Ewing's sarcoma. *PLoS One* 4(4): e5415
 88. Avila Cobos F, Vandesompele J, Mestdagh P, De Preter K (2018) Computational deconvolution of transcriptomics data from mixed cell populations. *Bioinformatics* 34 (11):1969–1979
 89. Stahl D, Gentles AJ, Thiele R, Gutgemann I (2019) Prognostic profiling of the immune cell microenvironment in Ewing's sarcoma family of tumors. *Onco Targets Ther* 8(12): e1674113
 90. Lutsik P, Slawski M, Gasparoni G, Vedenev N, Hein M, Walter J (2017) MeDeCom:

- discovery and quantification of latent components of heterogeneous methylomes. *Genome Biol* 18(1):55
91. Scherer M, Nebel A, Franke A, Walter J, Lengauer T, Bock C, Muller F, List M (2020) Quantitative comparison of within-sample heterogeneity scores for DNA methylation data. *Nucleic Acids Res* 48(8):e46
 92. Selvanathan SP, Graham GT, Erkizan HV, Dirksen U, Natarajan TG, Dakic A, Yu S, Liu X, Paulsen MT, Ljungman ME, Wu CH, Lawlor ER, Uren A, Toretsky JA (2015) Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing. *Proc Natl Acad Sci U S A* 112(11):E1307–E1316
 93. Bonnet E, Viara E, Kuperstein I, Calzone L, Cohen DPA, Barillot E, Zinovyev A (2015) NaviCell web service for network-based data visualization. *Nucleic Acids Res* 43(W1):W560–W565
 94. Stoll G, Caron B, Viara E, Dugourd A, Zinovyev A, Naldi A, Kroemer G, Barillot E, Calzone L (2017) MaBoSS 2.0: an environment for stochastic Boolean modeling. *Bioinformatics* 33(14):2226–2228
 95. Schwentner R, Papamarkou T, Kauer MO, Stathopoulos V, Yang F, Bilke S, Meltzer PS, Girolami M, Kovar H (2015) EWS-FLI1 employs an E2F switch to drive target gene expression. *Nucleic Acids Res* 43(5):2780–2789
 96. Sompairac N, Nazarov PV, Czerwinska U, Cantini L, Biton A, Molkenov A, Zhumadilov Z, Barillot E, Radvanyi F, Gorban A, Kairov U, Zinovyev A (2019) Independent component analysis for unraveling the complexity of cancer omics datasets. *Int J Mol Sci* 20(18):4414
 97. Khoogar R, Lawlor ER, Chen Y, Ignatius M, Kitagawa K, H-M Huang T, Houghton PJ (2019) Single-cell RNA profiling identifies diverse cellular responses to EWSR1-FLI1 Down-regulation in Ewing sarcoma. *bioRxiv* 750539

INDEX

A

- Anchored multiplex PCR (AMP)..... 106
- Antibodies16–18, 21, 22, 24, 51, 55, 56, 58, 152, 185, 187, 198, 235, 240, 246, 253, 266–268, 272, 273, 277, 278, 315, 316
- ArrayExpress 260, 308
- Assay for transposase-accessible chromatin (ATAC-seq)..... 286–292, 294, 295, 297, 299, 315

B

- BCOR..... 51, 55, 58, 59, 66, 67, 69, 78–80, 105
- Binding sites266, 268, 269, 272, 274, 276, 279, 286, 287, 295
- Bioinformatics 12, 78, 260, 271
- Biopsies 27, 28, 30, 32, 33, 36, 39–45, 50–52, 54, 56, 61, 68, 74, 88, 99, 228, 313, 316

C

- Cancer genomics 259, 260, 262, 263, 303, 308
- Cancer systems biology..... 318, 325
- cBioPortal260–263, 308
- Cell death151, 152, 156, 157, 252
- Cell division 151, 152
- Cell-free circulating DNA40, 41, 43–45
- Cell viability 29, 146, 152, 159, 160, 186
- ChIP-seq..... vi, 12, 265–281, 286–290, 292, 315
- CIC 66, 67, 69, 77, 78, 105
- Colorimetry 16, 159–165
- Core biopsies 50, 51, 56, 61
- Cryoprotective solutions 30

D

- Databasesvi, 260–263, 295, 318
- Data integration 319
- Decalcification 50–52, 55, 61, 73
- Differential diagnosis 50–52, 55, 58, 65, 66, 85, 105, 106
- DNA5, 6, 8, 11, 12, 15, 27–29, 32–34, 36, 40–45, 52, 70, 73, 90, 92, 96, 100, 101, 106, 111–113, 121, 123, 125–131, 133–135, 140–145, 147–149, 152, 195, 237, 239, 252, 260–262, 265–267, 269, 272–274, 285, 286, 295, 304, 309, 312, 313, 315, 319, 320

- Double-quenched probes40–42
- Droplet digital PCR (ddPCR) 40–42, 44
- Drug screening 159–162, 164, 317

E

- Electrophoresis 18, 20, 22, 34, 88, 92, 96, 99, 121, 127, 128, 142–145, 148, 316
- Embryonic superficial zone (eSZ)..... 184, 185
- Enhancers 4, 134, 139, 140, 266, 268, 269, 278–280, 286, 304, 310
- Epigenetics224, 259, 266, 269, 304, 305, 310, 315, 321, 326
- Ewing sarcoma (EwS) 3–12, 15–24, 27–37, 39–45, 49–62, 65–80, 85–101, 105–115, 139–149, 151–157, 159–178, 183–188, 191–199, 201–213, 215, 220, 224, 226–227, 230, 233, 235–239, 243–254, 259–263, 265–281, 285–300, 303–327
- EWSR1-ERG..... 22, 39, 41, 56, 74, 75, 86, 87, 91, 94–98, 100, 101
- EWSR1-FLI1 4, 5, 17, 22, 39, 41, 56, 57, 65, 72, 74, 86, 87, 91, 94–98, 100, 101, 120, 139, 140, 268, 285–287, 290, 292, 310, 315, 317, 318, 324, 326

F

- FFPE samples 27–37, 108, 109
- Fresh tissues28, 29, 35, 68, 73
- Frozen section 50
- Frozen tissues 5, 28–30
- FUS 66, 67, 69, 72, 76, 86, 105

G

- Gel 15, 16, 18, 20, 21, 23, 88, 92, 96, 97, 99, 101, 121, 126–128, 134, 135, 142–145, 148, 168–170, 173, 178, 316
- Gene expression 12, 29, 119–121, 125, 127, 132, 133, 136, 139, 151, 224, 237, 260, 262, 266, 277, 286, 304, 311–314, 319–321
- Gene Expression Omnibus (GEO) 260, 287
- Genome-wide association studies (GWAS) 310
- Germline variants 7, 8, 10, 310
- Grossing 50, 51, 53

H

H3K27ac 266, 268, 269, 277–279,
 287–290, 292–294, 296–298, 324
 Hemocytometers 153, 155, 157, 160, 161,
 173, 185, 203
 Histone modifications 265, 268, 279, 286
 Histopathology 50, 228
 Housekeeping protein 16, 21

I

ImageJ 170, 172, 177, 178
 Immunocompromised mice 191, 192, 217,
 220, 223
 Immunohistochemistry (IHC) 30, 31, 36,
 50–52, 54, 55, 72, 75, 198, 226, 228, 233, 316
 Inducible vector 132
 Invasion 167–178, 186, 202, 313

L

Liquid biopsies 27, 39
 Live imaging 243
 Luciferase assays 140, 142, 149

M

Mathematical modeling 305, 320, 321
 Metastasis 167, 186, 188, 201, 202, 207,
 209–213, 230, 244, 248
 Microsatellites 5, 8, 140, 268, 269, 278,
 286, 304, 305, 315
 Migration 167–178, 244, 250, 313
 Molecular diagnosis 88
 Motif analysis 278, 279
 Mouse models v, 39, 183–188,
 215, 224, 305
 Mouse surgery 209, 212

N

Neoadjuvant therapy 7, 50–52, 60
 Networks 4, 120, 222, 266,
 304, 307,
 320, 321, 323, 326
 Next-generation sequencing (NGS) 8, 10–12, 33,
 34, 110, 259, 260, 262, 263, 265–281, 309, 310
 Nucleic acids v, 27–37, 40, 52, 61,
 106–109, 114, 230

O

Omics data 307, 308, 318
 OnkoKB 262
 Orthotopic injection 218–219
 Overexpression 78, 120, 183, 198

P

Patient-derived xenograft (PDX) 29, 212,
 215, 216, 220, 223–240, 316, 325
 Pediatric sarcoma 24, 313
 Precision diagnostics 105
 Preclinical testing 220
 Proliferation 5, 151, 152, 156, 159, 160, 168,
 176, 177, 233, 244, 248, 251, 254, 305, 310,
 313, 321, 325
 Promoters 120, 140, 141, 266, 268, 269,
 278, 279, 287, 296, 298, 321, 323
 Protein expression 15
 Protein ladder 16, 20

R

Regulatory DNA 140
 Regulatory sequences 139–149
 Reporter assays 139–149
 Reporting template 59, 60
 Repository 119, 133, 175, 260–263, 295,
 308, 311, 319
 Resazurin 152, 159–165
 Reverse transcriptase-PCR (RT-PCR) 75, 78,
 80, 85–102, 106
 RIPA 17, 18
 RNA 5, 15, 27–30, 32–36, 39, 52, 66,
 74, 78, 88–90, 92–94, 97, 99–101, 105–115,
 132, 136, 185, 195, 207, 227, 236, 237, 260,
 265, 266, 304, 305, 319, 321, 326

S

Senescence 151, 152, 183
 Serum markers 39
 Small round cell tumors (SRCT) 65, 66, 69, 72,
 73, 76, 85
 SNPs 310, 311
 Soft tissues 50, 53, 54, 56, 60, 62, 76, 78,
 105, 106, 183, 185, 201, 243, 285, 303
 Solid biopsies v
 Somatic mutations 4, 6, 8, 12, 139, 262, 303, 318
 Subcutaneous injection 188, 192, 194, 195

T

Targeted RNA-seq 106
 The Cancer Genome Atlas (TCGA) Research
 Network 262
 Tissue preservation 30, 35
 Tissues 5, 6, 8, 11, 15, 17, 27–37, 49,
 51, 52, 55–57, 60–62, 66–70, 72, 73, 89, 92, 99,
 106, 151, 191, 192, 195, 198, 203, 207, 208,
 210, 212, 213, 216, 220, 221, 225, 227, 228,
 230, 233, 235, 237–240, 304, 312, 313, 325

Transillumination 16
 Translocations 4, 16, 22, 65, 66, 71, 72,
 74, 76, 86, 88, 120, 176, 268, 285, 286, 304, 313
 Transwell 167, 169, 170, 172–175
 Trypan blue 151–157, 160, 161, 163, 192

W

Western blot 15–17, 20, 23

Wound healing assays 167–169, 171–172, 176

X

Xenograft models 191, 192, 215, 223

Z

Zebrafish vi, 243–254