



**HAL**  
open science

## Unraveling Ewing Sarcoma Tumorigenesis Originating from Patient-Derived Mesenchymal Stem Cells

Anna Sole, Sandrine Grossetête, Maxime Heintzé, Loelia Babin, Sakina Zaïdi, Patrick Revy, Benjamin Renouf, Anne de Cian, Carine Giovannangeli, Cécile Pierre-Eugène, et al.

► **To cite this version:**

Anna Sole, Sandrine Grossetête, Maxime Heintzé, Loelia Babin, Sakina Zaïdi, et al.. Unraveling Ewing Sarcoma Tumorigenesis Originating from Patient-Derived Mesenchymal Stem Cells. *Cancer Research*, 2021, 81 (19), pp.4994-5006. 10.1158/0008-5472.CAN-20-3837 . mnhn-03873854

**HAL Id: mnhn-03873854**

**<https://mnhn.hal.science/mnhn-03873854v1>**

Submitted on 27 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Unraveling Ewing sarcoma tumorigenesis originating from patient-derived**  
2 **Mesenchymal Stem Cells**

3 Anna Sole<sup>1</sup>, Sandrine Grossetête<sup>2</sup>, Maxime Heintzé<sup>1</sup>, Loelia Babin<sup>1</sup>, Sakina Zaidi<sup>2</sup>, Patrick  
4 Revy<sup>1</sup>, Benjamin Renouf<sup>3</sup>, Anne De Cian<sup>3</sup>, Carine Giovannangeli<sup>3</sup>, Cécile Pierre-Eugène<sup>2</sup>,  
5 Isabelle Janoueix-Lerosey<sup>2</sup>, Lucile Couronné<sup>4,5</sup>, Sophie Kaltenbach<sup>5,6</sup>, Mark Tomishima<sup>7,8</sup>,  
6 Maria Jasin<sup>7</sup>, Thomas G. P. Grünwald<sup>9,10,11,12</sup>, Olivier Delattre<sup>2</sup>, Didier Surdez<sup>2,13\*†</sup> and Erika  
7 Brunet<sup>1\*†</sup>

8

9 <sup>1</sup> Institut Imagine, INSERM UMR1163, «Laboratory of Genome Dynamics in the Immune  
10 System », Équipe Labellisée La Ligue Nationale Contre Le Cancer, University of Paris, Paris,  
11 France

12 <sup>2</sup> INSERM U830, Équipe Labellisée LNCC « Genetics and Biology of Pediatric Cancers »,  
13 PSL Research University, SIREDO Oncology Centre, Institut Curie Research Centre, Paris,  
14 France

15 <sup>3</sup> Muséum National d'Histoire Naturelle, Inserm U1154, CNRS UMR 7196, Sorbonne  
16 Universités, Paris, France

17 <sup>4</sup> Institut Imagine, INSERM UMR1163, «Laboratory of Onco-Hematology, Laboratory of  
18 Cellular and Molecular Mechanisms of Hematological Disorders and Therapeutical  
19 Implications», University of Paris, Paris, France

20 <sup>5</sup> Laboratory of Onco-Hematology, Hôpital Necker - Enfants Malades, Assistance Publique -  
21 Hôpitaux de Paris (APHP), Paris, France

22 <sup>6</sup> INSERM U1151, Institut Necker-Enfants Malades, Paris, France

23 <sup>7</sup> Developmental Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY  
24 10065, USA

25 <sup>8</sup> BlueRock Therapeutics, New York, NY 10016, USA

26 <sup>9</sup> Division of Translational Pediatric Sarcoma Research, German Cancer Research Center  
27 (DKFZ), Heidelberg, Germany

28 <sup>10</sup> Hopp-Children's Cancer Center (KiTZ), Heidelberg, Germany

29 <sup>11</sup> Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany

30 <sup>12</sup> German Cancer Consortium (DKTK), Partner Site Heidelberg, Heidelberg, Germany

31 <sup>13</sup> Balgrist University Hospital, University of Zurich, Zurich, Switzerland

32 † co-corresponding authors

33

34

35 \* correspondence to:

36 [erika.brunet@inserm.fr](mailto:erika.brunet@inserm.fr) (E.B)

37 [didier.surdez@balgrist.ch](mailto:didier.surdez@balgrist.ch) (D.S)

38

39 The authors declare no potential conflicts of interest.

40

41

42 **Funding:**

43 This work was supported by the Institut National du Cancer (PLBIO16-291), by La Ligue  
44 Nationale Contre le Cancer (P.R/E.B. and O.D/D.S teams: Équipes Labellisées), the  
45 Fondation ARC (L.B.); the AIDA association (M.H.), by grants from the Institut Curie; the  
46 INSERM; the Canceropôle Ile-de-France; the projet de Recherche ‘Enfants, Adolescents et  
47 Cancer’; the Agence Nationale de la Recherche (ANR-10-EQPX-03, Institut Curie  
48 Génomique d’Excellence (ICGex) and the société française de lutte contre les cancers de  
49 l’enfant et de l’adolescent. This project also received support from European funding: ERA-  
50 NET TRANSCAN JTC 2014 (TRAN201501238) and TRANSCAN JTC 2017  
51 (TRANS201801292). D.S. is supported by the Institut Curie–SIRIC (Site de Recherche  
52 Intégrée en Cancérologie) program. T.G.P.G. is supported by grants from the Gert and  
53 Susanna Mayer foundation, the Barbara and Wilfried Mohr foundation, and the SMARCB1  
54 association. We are indebted to the following associations for providing essential support:  
55 L’Etoile de Martin, Aida, la Course de l’Espoir, M la vie avec Lisa, ADAM Couleur Jade,  
56 Dans les pas du Géant, Courir pour Mathieu, Marabout de Ficelle, Olivier Chape, Les  
57 Bagouzamanon, Enfants et Santé, and les Amis de Claire.

58

59

60 **ABSTRACT**

61 Ewing sarcoma (EwS) is characterized by pathognomonic translocations fusing most  
62 frequently *EWSR1* with *FLI1*. Moreover, around 30% of EwS tumors also display genetic  
63 alterations in *STAG2*, *TP53*, or *CDKN2A (SPC)*. Numerous attempts to develop relevant EwS  
64 models from primary human cells have been poorly successful. By engineering the  
65 t(11;22)(q24;q12) translocation together with a combination of *SPC* mutations, we report here  
66 successful generation of a wide collection of immortalized cells tolerating EWSR1-FLI1  
67 expression (EWIma cells) from primary mesenchymal stem cells (MSCs) derived from a EwS  
68 patient. Our results revealed that *SPC* alterations strongly favor EwS oncogenicity.  
69 Importantly, xenograft experiments with independent EWIma cells induced tumors and  
70 metastases in mice which displayed *bona fide* EwS features. EWIma cells present balanced  
71 but also more complex translocation profiles mimicking chromoplexy, which is frequently  
72 observed in EwS and other cancers. Collectively, our results demonstrate that bone marrow  
73 derived MSCs are a source of origin for EwS and also provide original experimental models  
74 to investigate Ewing sarcomagenesis.

75

76 **SIGNIFICANCE**

77 We demonstrate that Ewing sarcoma can originate from human bone marrow derived MSCs  
78 and that additional recurrent mutations support the EwS translocation mediated  
79 transformation.

## 80 INTRODUCTION

81 Ewing sarcoma (EwS) is the second most frequent bone or soft-tissue cancer of  
82 children, adolescents, and young adults. It is characterized by a chromosomal translocation  
83 between *EWSR1* and members of the *ETS* (E26 transforming-specific) family of transcription  
84 factors, most frequently with *FLI1* (t(11;22)(q24;q12)) (1). EWSR1-FLI1 exerts a strong  
85 oncogenic role but also cytotoxic effects when expressed in various primary cells (2,3). Only  
86 few additional recurrent genetic alterations are observed in EwS, primarily including  
87 inactivating mutations of *STAG2* (~15–20%) and *TP53* (~5–10%), as well as *CDKN2A*  
88 deletions (~9–22%) (4-6). Co-occurrence of *STAG2* and *TP53* mutations was reported to be  
89 associated with poor outcome (6). However, the pro-oncogenic role of these additional  
90 mutations in EwS origin and progression remained elusive.

91 Despite numerous efforts to generate murine EwS models, none of them faithfully  
92 recapitulated phenotypic, transcriptomic and epigenetic features of EwS (7-11). This may be  
93 partially explained by the poor conservation of *cis*-regulatory enhancers containing GGAA-  
94 microsatellites (mSats) that are uniquely bound by EWSR1-FLI1, and that appear critical for  
95 Ewing sarcomagenesis (12). However, several factors such as the exact nature of the cell(s) of  
96 origin, the timing and the (co)-occurrence of oncogenic events involved in Ewing  
97 sarcomagenesis are still poorly characterized (12). Although EwS histogenesis has been a  
98 long-lasting debate, experimental evidence has converged on either a neural crest origin, as  
99 neural crest-derived cells appeared to be permissive to EWSR1-FLI1 expression (13), or a  
100 mesenchymal origin, as for instance EWSR1-FLI1 inhibition in EwS cells induced features of  
101 mesenchymal stem cells (MSCs) (14). However, none of the attempts succeeded to model  
102 *bona fide* EwS tumors *in vivo* from any types of primary human cells.

103 To better mimic the pathophysiological context of EwS, we and others engineered the  
104 *EWSR1-FLI1* translocation using genome editing technologies in human stem cells (15,16).

105 These approaches lead to formation of the specific t(11;22)(q24;q12) translocation, starting  
106 from two double-strand breaks (17), one in *EWSR1* and the other in *FLII*. In contrast to  
107 models with ectopic EWSR1-FLI1 expression, the precise generation of the translocation at  
108 the endogenous loci enables faithful and ‘natural’ oncogene regulation, and reproduces  
109 heterozygosity at the *EWSR1* locus resulting from the translocation of one allele. However,  
110 the specific isolation of immortalized/transformed *EWSR1-FLII* translocated clones remained  
111 unsuccessful (15,16). To overcome this issue, we reasoned that the genetic background of the  
112 starting cells could be of strong relevance. Indeed, the incidence of EwS is much higher in  
113 Europeans or European-Americans than in Africans or Afro-Americans (18-20). Genome-  
114 wide association studies (GWAS) suggest that a specific genetic germline background may be  
115 more permissive to *EWSR1-ETS* translocation and may favor EWSR1-ETS activity (21), as  
116 reported for the *EGR2* susceptibility locus (22). In addition, we reasoned that additional  
117 recurrent mutations identified in this cancer may also contribute to Ewing sarcomagenesis.  
118 Here, we describe a model generated from primary human MSCs of a European EwS patient  
119 by introducing the t(11;22)(q24;q12) translocation and additional alterations in *STAG2*, *TP53*  
120 and *CDKN2A*. Strikingly, this model displays molecular and phenotype features of EwS  
121 tumor, including expression of EwS-associated markers (including membranous CD99), and  
122 efficiently generated EwS tumors and metastases in immunodeficient mice. The *bona fide*  
123 genetically engineered EwS model generated in this study provides novel insights in Ewing  
124 sarcomagenesis and highlights the role of additional somatic mutations in this transformation.

125

## 126 **MATERIALS AND METHODS**

### 127 **Primary cell and cell line culture**

128 MSC<sup>Pat</sup> cells (MSC7-BJ), human primary BMSCs, derived from bone marrow aspirates  
129 previously described in (23) and hMPCs described in (24) were cultured in  $\alpha$ MEM

130 supplemented with 10% MSC-FBS (12662029; Life Technologies), Glutamine 10mM (Life  
131 Technologies) and 2 ng/mL Recombinant Human FGF (233-FB-025; R&D Systems). For  
132 these primary cells, written informed consent was obtained according to the Declaration of  
133 Helsinki and studies were approved by the ethics committees of the contributing institutions.  
134 A673 were cultured in DMEM medium supplemented with 10% FBS (Life Technologies) and  
135 cultured at 37 °C in a humidified atmosphere with 5% CO<sub>2</sub> and 20% O<sub>2</sub>. MSC<sup>Pat</sup> and hMPCs  
136 were cultured in hypoxic-like conditions (3% O<sub>2</sub>). Culture cells are tested monthly for  
137 mycoplasma contamination using the mycoplasma treatment kit (Myco-1&2 set A8360.0010,  
138 VWR). Cells used in the study are fully described in Supplementary Materials.

139

#### 140 **CRISPR/Cas9 transfections**

141 Cells were transfected by 4D Nucleofector Amaxa technology (Lonza) using the cell line  
142 nucleofector (solution P1, FF-104) with 1 µg of plasmid pCAS9-GFP (44719; Addgene) and  
143 1 µg of each plasmid (MLM3636, 43860; Addgene) encoding for the different gRNAs (2 µg  
144 for gRNA<sup>EWS</sup>). For Cas9/gRNA RNP complexes, MSC<sup>Pat</sup> and hMPCs were transfected  
145 directly with the different combination of gRNAs and Cas9-GFP protein (ratio 2:1). gRNA  
146 sequences are listed in the Supplementary material.

147

#### 148 **PCR-based translocation detection and frequency**

149 For detection of translocations from bulk cells, DNA (E.Z.N.A. Tissue DNA Kit, Omega Bio-  
150 Tek, GA, USA) was amplified by PCR or Nested PCR, 6 to 8 days post-transfection. Serial  
151 dilutions of DNA enable the assessment of translocation frequency as previously described in  
152 (25). Primer sequences are reported in Supplementary Material.

153

#### 154 **Soft-agar colony formation assays**

155 A first agar layer was placed in 10 cm plates at 0.8% (w/v) of low melting temperature  
156 agarose (50101; Life) in  $\alpha$ MEM-10%FBS. Once solidified, a second layer of 0.48% agar  
157 containing  $4 \times 10^4$  cells was added. The plates were maintained at 4 °C for 5 min and 10mL of  
158 fresh culture medium was subsequently deposited as a top layer. The plates were incubated in  
159 hypoxia conditions and colonies were isolated after 3 to 4 weeks and further analyzed.  
160 Counting of colonies was performed using ImageJ/Fiji software and T-test analysis was  
161 applied.

162

### 163 **Cell proliferation and siRNA assay**

164 hMPCs were seeded in 6-well plates (40.000 or 60.000 cells/well) and maintained in hypoxic  
165 like or standard conditions for proliferation assay. Cell growth was monitored and analyzed  
166 by the IncuCyte Live Cell Analysis system (IncuCyte S3, Essen Bioscience) every 24 hours  
167 for 4 to 10 days. For knock-down experiments, 40nM of siRNA against *EWSRI-FLII* (5'-  
168 GGC AGC AGA ACC CUU CUU A-3') (Eurofins) (or control siRNA (D-001810-01-50;  
169 Dharmacon) was transfected using Lipofectamine RNAiMAX Transfection Reagent  
170 (ThermoFisher Scientific) following manufacturer's instructions. Cell growth was monitored  
171 and analyzed by the IncuCyte system every 4 hours during 3 days. All experiments were  
172 independently performed in triplicate, and T-test analysis was applied. SiRNA sequences are  
173 reported in Supplementary Material.

174

### 175 **Fluorescence *In Situ* Hybridization (FISH), multicolor FISH analysis and conventional** 176 **cytogenetic analysis**

177 Fresh cells with few passages were harvested after 1-6 hours with 10 $\mu$ l/mL of KaryoMAX  
178 colcemid (Gibco) treatment, resuspended in 0.075M KCl at 37 °C for 30 minutes and fixed in  
179 methanol/acetic acid (3:1). Cells were dropped onto glass slides and dried. FISH was



180 performed on the metaphases using EWSR1 and FLI1 probes (LPS007, Cytocell) to detect the  
181 t(11;22) chromosomal translocation. Cell images were captured with the Zeiss Spinning Disk  
182 Confocal microscopy 63x. Alternatively, for multicolor FISH imaging metaphase spreads  
183 were stained with 24XCyte, Multicolor Painting mFISH Probe Kit (MetaSystems), which was  
184 prepared following supplier's instructions. Metaphases were imaged using a ZEISS  
185 AxioImager.Z2 microscope and the Metafer automated capture system (MetaSystems).  
186 Karyotyping was performed using Isis software (MetaSystems). For conventional karyotypes,  
187 metaphase spreads, R-banded chromosomes were analyzed by standard procedures.

188

### 189 **Flow cytometry**

190 Immunostaining of cells for CD99 marker was performed by incubating cells with FITC  
191 Mouse Anti-Human CD99 (BD Pharmingen, 555688) or FITC Mouse IgG2a,  $\kappa$  Isotype Ctrl  
192 Antibody (Biolegend, 400208) for 30 minutes at 4 °C prior to flow cytometry analysis based  
193 on SS-A/FS-H gating on alive cells (BD FACS Aria II- BD bioscience, and FlowJo software).  
194 Cells labelled with unspecific FITC Mouse IgG2a were used as negative cells for CD99  
195 expression.

196

### 197 **Western blotting**

198 Whole-cell extracts were prepared with protein lysis buffer (50mM Tris-HCl at pH 7.4, 1%  
199 Triton X-100, 0.1% SDS, 150mM NaCl, 1mM EDTA, and 1mM DTT), with addition of  
200 cocktail protease inhibitor tablets (Complete, Roche). Membranes were stained with FLI1  
201 (ab133485; Abcam) (used to detect the EWSR1-FLI1 translocation), STAG2 (sc81852; Santa  
202 Cruz), p53 (sc126; Santa Cruz) or p16 (554079; BD Pharmingen) antibodies. ACTIN  
203 (sc1616; Santa Cruz) and VINCULIN (sc73614; Santa Cruz) antibodies were used as loading

204 controls. Membranes were visualized with Odyssey CLx Imaging System (LI-COR  
205 Biosciences).

206

### 207 **qRT-PCR**

208 RNA was extracted with the RNeasy Plus Mini Kit (Qiagen) and reverse-transcribed using the  
209 High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). qRT-PCRs were  
210 performed using PowerSYBR green Mastermix (Applied Biosystems). Oligonucleotides were  
211 purchased from MWG Eurofins Genomics (Oligonucleotides and Primers, see Supplementary  
212 Materials). Reactions were run on CFX384 Touch Real-Time PCR instrument (Bio-Rad) and  
213 analyzed using the CFX Manager Software (Bio-Rad). Relative expression level was assessed  
214 with the ddCT method using *RPLP0* as a housekeeping gene. Primer sequences are reported  
215 in Supplementary Material.

216

### 217 **Telomerase Repeat Amplification Protocol (TRAP)**

218 Fresh cells were resuspended in 100 $\mu$ L of CHAPS Lysis Buffer and TRAP assay was  
219 performed following the manufacturer's instructions of TRAPeze Telomerase Detection Kit  
220 (S7700, Millipore). PCR products were run on a TBE/acrylamide:bisacrylamide (19:1) gel,  
221 stained with SYBR Gold Nucleic Acid Gel Stain (Invitrogen) and visualized with a FLA-  
222 3000 Phosphorimager (Fujifilm).

223

### 224 **Telomere Restriction Fragment (TRF)**

225 DNA was isolated and digested overnight with HinfI and RsaI enzymes. DNA samples were  
226 run in a 0.7% agarose gel overnight and transferred to a membrane. The [ $\gamma^{32}$ P]ATP-labeled  
227 telomere probe (CCCTAAA)<sub>4</sub> was subsequently hybridized by using the Easy Hyb reagent

228 (Roche). Membranes were exposed to Phosphorimager screens and screens were scanned with  
229 a FLA-3000 Phosphorimager (Fujifilm).

230

### 231 **SNP array**

232 Infinium Core-24 Chip (Illumina Inc. San Diego, USA) containing more than 300.000 SNPs  
233 that were hybridized with genomic DNA.

234

### 235 **RNA-seq and bioinformatics analysis**

236 Sequencing was carried out using 2x100 cycles (paired-end reads 100 nucleotides) for all  
237 samples on Illumina HiSeq2500 or NovaSeq6000 instruments. Reads were aligned with  
238 STAR 2.5.3 (Supplementary Reference 47) to the human genome (GRCh37/hg19 version).  
239 We used the count matrix generated by STAR with the human gene annotation v19 of  
240 GENCODE as reference. DESEQ2 (Supplementary Reference 48) was used to normalize data  
241 and performed differential analysis with the Wald test. The p-value was adjusted using the  
242 Benjamini-Hochberg method. We considered a gene expressed if the normalized expression  
243 was higher than 10. STAR-Fusion v1.4.0 (Supplementary Reference 49) was applied to  
244 predict fusion transcripts. No statistical methods were used to predetermine sample size.  
245 RNAseq experiments were performed in duplicates or triplicates.

246

### 247 **ChIP-Seq**

248 Chromatin Immunoprecipitation (ChIP) experiments were performed following manufacturer  
249 instructions using iDeal ChIP-seq kit for transcription factors and for histones (Diagenode)  
250 with respectively rabbit polyclonal anti-FLI1 antibody (ab15289, Abcam), rabbit polyclonal  
251 anti-H3K4me3 (C15410003, Diagenode) and rabbit polyclonal anti-H3K27ac (ab4729,  
252 Abcam). For ChIP sequencing, libraries were generated using TruSeq ChIP library

253 preparation kit (Illumina) and sequenced on Illumina HiSeq 2500 (single end, 100 bp). Reads  
254 were aligned to human reference genome (GRCh37/hg19) with bowtie2 2.2.9 (Supplementary  
255 Reference 50). Peaks were called with MACS2 2.1.1 (Supplementary Reference 51) with the  
256 option narrow for FLI1 ChIP-seq and broad for H3K27ac histone mark. ChIP-seq data were  
257 normalized according to their respective input DNA sample. The ChIP-seq signal tracks were  
258 generated by macs2 with bdgcmp option (and m FE to compute fold enrichment between the  
259 ChIP and the control). Then, we run bedGraphToBigWig to convert the file to a binary format  
260 (BigWig). To define enhancers and super-enhancers, we used the ROSE tool (Supplementary  
261 Reference 52) with the parameter -t 2500 in order to exclude H3K27ac peaks which overlap  
262 the theoretical TSS (<2.5kb) regions. Here defined enhancers were then stitched and therefore  
263 some enhancers or super-enhancers can contain active promoters. We annotate them by  
264 associating the closest expressed genes. Control-FREEC (29) on input DNA was used to  
265 determine Copy Number Variants in EWIma1 (MSC<sup>Pat</sup> was used as reference).  
266 Primer sequences are reported in Supplementary Material.

267

## 268 **Mice**

269 Animal care and use for this study were performed in accordance with the recommendations  
270 of the European Community (2010/63/UE) for the care and use of laboratory animals.  
271 Experimental procedures were specifically approved by the ethics committee of the Institut  
272 Curie CEEA-IC #118 (Authorization APAFIS#11206-2017090816044613-v2 given by  
273 National Authority) in compliance with the international guidelines. The tumorigenic and  
274 metastatic potential of hMPC clones (4-7 mice) or EWIma1, 5 and 7 (2-4 mice per clone)  
275 cells was investigated by injecting 1 million cells in an orthotopic intra-osseous model as  
276 described previously (Supplementary Reference 53). Seven-week-old NSG (NOD.Cg-  
277 Prkdcscid IL2rgtm1Wjl/SzJ) female mice were purchased from Charles Rivers (France). Mice

278 were anesthetized by inhalation of a combination isoflurane/air (1.5%, 1 L/min) and followed  
279 up for tumor growth. The tumor volume was calculated by using the formula  $L \times (l^2)/2$ ,  
280 where L and l represent respectively the longest and the smallest perpendicular diameter.  
281 Tumor samples were fixed for 24 hours in AFA solution and processed for paraffin  
282 embedding and sectioning.

283

#### 284 **Immunohistochemistry**

285 Xenograft sections (4 $\mu$ m) were cut and stained with hematoxylin and eosin.  
286 Immunohistochemistry was performed using the following antibodies and dilutions: FLI1  
287 1:50 (Abcam, ab15289), STAG2 1:25 (Santa Cruz, sc81852), Ki67 1:500 (Abcam, Ab15580),  
288 cleaved CASP3 1:250 (Cell Signaling, #9661) and CD99 1:1 ready-to-use (Agilent, IS057).

289

#### 290 **Data availability**

291 ChIP-seq and RNA-seq data have been deposited in NCBI's Gene Expression Omnibus under  
292 accession number: GSE150783.

293 Raw data are available at Mendeley. Reserved DOI: doi:10.17632/fx29by5k43.1

294 <https://data.mendeley.com/datasets/fx29by5k43/draft?a=63bd5c8f-f13b-4100-b9bf-988ddb0c9131>

295

296

297

298 **RESULTS**

299 **Generation of stable cell lines expressing EWSR1-FLI1 (EWIma cells) starting from**  
300 **wild type MSCs of an EwS patient**

301 Aiming at recapitulating the t(11;22)(q24;q12) *in situ* and knowing that the genetic  
302 background is a potential factor of incidence in Ewing sarcoma (18), we derived primary  
303 normal MSCs of a patient (MSC<sup>Pat</sup>) (23) who was affected by an EwS of the ulna. The EwS  
304 tumor from this patient exhibited four large deletions (chr3q; chr9p comprising *CDKN2A*;  
305 chr16q; and chr17p comprising *TP53*), one gain (chr1q), and an isodisomy (chr5p) at time of  
306 diagnosis (**Supplementary Fig. 1 and 2A**). This tumor also exhibited altered expression of  
307 the 3' end of *STAG2* (4.90 ratio of average coverage of [Exons 26–30]/[Exons 1–25],  
308 compared to 1.62 for MSC<sup>Pat</sup> cells suggesting the existence of 3' truncated transcripts)  
309 (**Supplementary Fig. 2A**), however, without detectable genetic alteration in the coding  
310 region of this gene. MSC<sup>Pat</sup> were derived from a bone marrow aspirate at the time of diagnosis  
311 and exhibit no copy number alterations (**Supplementary Fig. 1**). Genomic analyses of these  
312 cells confirmed the absence of mutations or alterations in oncogene or tumor suppressor  
313 genes. As comparison, we also used mesenchymal stem/precursor cells (hMPCs) derived from  
314 human embryonic stem cells which are proficient for multi-lineage differentiation (fat,  
315 cartilage, bone, and skeletal muscle)(24). We noticed a growth advantage of hypoxia on cell  
316 morphology and proliferation of hMPCs (**Supplementary Fig. 2B and C**). This is consistent  
317 with the observation that the center of solid tumors and the niche of MSCs are mostly hypoxic  
318 environments and that hypoxia enhances growth of MSCs (also reported in 25). Based on  
319 these results, we cultured MSC<sup>Pat</sup> and hMPC cells in hypoxic conditions (3% O<sub>2</sub>) throughout  
320 this study and used the CRISPR-Cas9 technology to engineer the t(11;22)(q24;q12)  
321 translocation (26)(**Fig. 1A**). Upon transfection of Cas9, gRNA<sup>EWSR1</sup> and gRNA<sup>FLI1</sup> (*EF*)  
322 coding plasmids in MSC<sup>Pat</sup>, *EWSR1-FLI1* translocation positive cells were readily detected at

323 12 and up to 61 day post-transfection (**Supplementary Fig. 2D**), although at a lower  
324 frequency for the latter time point. However, we could not recover any viable clones in these  
325 conditions.

326 We hypothesized that additional somatic mutations found in EwS may facilitate the  
327 transformation potential of EWSR1-FLI1 fusion protein. We focused on the three most  
328 recurrently mutated genes identified in EwS: *STAG2* (*S*), *TP53* (*P*) and *CDKN2A* (*C*),  
329 knowing that their expression is also altered in the tumor of origin of this EwS patient  
330 (**Supplementary Fig. 2A**). Using CRISPR-Cas9 and gRNA plasmid transfection, we  
331 simultaneously induced the translocation with *EF* gRNAs with *SPC* gRNAs in MSC<sup>Pat</sup> cells  
332 (**Fig. 1A**), and could recover numerous clones with *SPC* mutations. One clone (hereafter  
333 termed EWIma1), among hundred isolated clones, displayed morphologic changes with  
334 rounder cells, indicative of a gradual acquisition of the classical small-round-cell morphology  
335 of EwS cells (2-3 weeks after its isolation, **Supplementary Fig. 2E**). Three additional  
336 independent experiments using the same approach allowed recovering two additional EWIma  
337 clones (termed EWIma1\* and EWIma1<sup>#</sup>). SNP arrays of EWIma1 cells did not show any of  
338 the copy number changes found in the patient EwS cells, excluding a hypothetical initial  
339 contamination of the MSC<sup>Pat</sup> with patient tumor cells (**Supplementary Fig. 1**). In contrast to  
340 the original MSC<sup>Pat</sup>, but similarly to the prototypic A673 EwS cell line, EWIma1 cells stably  
341 expressed EWSR1-FLI1 oncoprotein (**Supplementary Fig. 2F**). As expected, these cells did  
342 not express *STAG2* nor *p16* and expressed a truncated form of *p53* (**Supplementary Fig.**  
343 **2G**). Inhibition of EWSR1-FLI1 by RNA interference led to a significant decrease of their  
344 proliferation (**Fig. 1B**) and reverted EWIma1 cells to a more mesenchymal spread-like  
345 morphology, as previously described in EwS cell lines (14)(**Supplementary Fig. 2H and 2I**).

346 Karyotypes and PCR analysis of EWIma1 cells revealed three additional translocations  
347 involving chromosomes 9, 11, 13, and 17 (**Fig. 1C and 1D, Supplementary Fig. 2J, 2K and**

348 **2L** and **Supplementary Table 1**). Most of the breakpoint junctions corresponded to the  
349 CRISPR/Cas9 targeted loci or to a predicted off-target site of gRNA<sup>CDKN2A</sup> located on chr13  
350 (in the promoter region of *USPL1*) (**Supplementary Fig. 2K**). Only t(9;17) breakpoints,  
351 which involve the centromere of chr17, could not be amplified by PCR but probably  
352 implicates the *CDKN2A* on chr9 as only one allele of *CDKN2A* can be amplified. An  
353 additional duplication of chr20q, potentially involved in a derivative chromosome 20, was  
354 detected in 73% (16/22) of the metaphases (**Supplementary Fig. 2L** and **Supplementary**  
355 **Table 1**). Similarly to EWIma1, we identified translocations involving chromosomes 9, 11  
356 and 13, 17 in EWIma1<sup>\*</sup> and EWIma1<sup>#</sup> (**Supplementary Fig. 2J**). Strikingly, these particular  
357 chromosomal rearrangements are reminiscent of concomitant intricate genetic events  
358 previously described as chromoplexy, which is characterized by a sudden burst of complex,  
359 loop-like rearrangements (27), found in more than 35% of EwS tumors (28).

360 To further characterize the EWIma1 model and to assess whether it faithfully  
361 recapitulates EwS properties, we performed ChIP-seq experiments against FLI1 and H3K27ac  
362 in MSC<sup>Pat</sup> and EWIma1 cells. Chromatin patterns for these marks at known EWSR1-FLI1  
363 targets genes in EWIma1 cells strongly resembled those of established A673 EwS cells and  
364 noticeably differed from the one of MSC<sup>Pat</sup> (**Fig. 1E** and **1F**).

365 Importantly, the canonical *EWSR1-FLI1* GGAA-mSat and ETS binding motifs (29)  
366 were identified as the first and second motifs among all known transcription factor motifs in  
367 EWIma1 FLI1 ChIP-seq data (**Fig. 1G**). Conversely, ETS site was the most prominent  
368 identified motif in ChIP-seq peaks in MSC<sup>Pat</sup> which express FLI1 (**Fig. 1G, Supplementary**  
369 **Fig. 3A**). In EWIma1 cells, 16,338 specific EWSR1-FLI1 peaks were identified (1,741 at  
370 GGAA-mSats and 14,597 at ETS binding sites). These specific EWSR1-FLI1 peaks were  
371 highly similar to those identified in the A673 cells (**Fig. 1H**). Strikingly, peaks at GGAA-  
372 mSat regions were completely absent in MSC<sup>Pat</sup>. Using H3K27ac ChIP-seq, we further



373 identified 8,685 regions that were specific for EWIma1 compared to MSC<sup>Pat</sup> (**Fig. 1H**).  
374 Again, these regions were highly conserved in A673. Furthermore, we performed ROSE  
375 analysis in EWIma1 and MSC<sup>Pat</sup> cells and identified super-enhancers (SEs) associated to  
376 known EwS-specific genes such as *BCL11B*, *CCND1*, *GLG1*, *NKX2-2*, and *SOX6* (30-32) as  
377 top hits in EWIma1 cells (**Supplementary Fig. 3B**).

378

379 ***De novo* Ewing sarcomagenesis models display heterogeneous morphologies and**  
380 **immortalization patterns, simple and chromoplectic-like phenotypes and are favored by**  
381 **additional mutations.**

382 Aiming at obtaining a broad collection of EWIma models by increasing the  
383 translocation frequency, we transfected MSC<sup>Pat</sup> with the ribonucleic protein (RNP)/Cas9  
384 complexes (replacing the above plasmid based approach) (33). Very few small colonies grew  
385 in agar after transfection with gRNA<sup>*EWSR1*</sup> and gRNA<sup>*FLI1*</sup> (*EF*) or together with gRNA<sup>*TP53*</sup>  
386 (*EF+P*). In contrast, additional combinations with gRNA<sup>*CDKN2A*</sup> (*EF+PC*) and even more  
387 strikingly with gRNA<sup>*STAG2*</sup> (*EF+SPC*) significantly increased the size of these soft agar grown  
388 colonies (**Fig. 2A** and **2B**). Using *EF+SPC* conditions, we reached a translocation frequency  
389 of  $1.6 \times 10^{-3}$  (**Fig. 2C**), representing a ~30-fold improvement as compared to plasmid based  
390 approach (**Supplementary Fig. 2D**). Remarkably, *EF+SPC* combination allowed for the  
391 identification of 116 PCR positive clones for *EWSR1-FLI1* out of a total of 274 isolated  
392 clones from these agar plates (**Supplementary Fig. 3C**). We further kept in culture 30 clones  
393 (**Supplementary Table 2**). Of these, we first randomly selected 13 clones (EWIma2 to  
394 EWIma14) for further molecular and cellular characterization, and confirmed expression of  
395 EWSR1-FLI1 fusion protein (**Fig. 2D**) and mutation in *SPC* genes for all of them (**Table 1**).  
396 Using flow cytometry, we also confirmed higher CD99 expression levels in EWIma clones as  
397 compared to MSC<sup>Pat</sup> mutated in *SPC* (**Supplementary Fig. 3D**). Most EWIma clones

398 displayed classical EwS cell morphology (**Fig. 2E, Supplementary Fig. 3E**). Using TRAP  
399 assay, most EWIma clones showed strong telomerase activity and grew past 100 days, in  
400 agreement with a full immortalized phenotype (**Fig. 2F**). Interestingly, we observed a more  
401 mesenchymal intermediate morphology in some clones (EWIma 11, 12, but also for two  
402 additional clones named EWIma 30 and 31) (**Fig. 2E, Supplementary Fig. 3E**). These  
403 intermediate clones appeared to express lower levels of EWSR1-FLI1 transcript  
404 (**Supplementary Table 2**) and protein (**Fig. 2D**), as compared to EWIma models displaying a  
405 clear EwS morphology, and also exhibited a lower telomerase activity (**Fig. 2F**). Karyotype  
406 analysis of immortalized EWIma models revealed that 7 out of 13 analyzed clones (EWIma 2  
407 to 4 and 6 to 8, 12) displayed a rather simple and stable karyotype with t(11;22)(q24;q12)  
408 (**Fig. 2G and Supplementary Fig. 4A**). A few additional somatic alterations could be also  
409 detected by simple karyotype, including a loss of 16q (EWIma2), which is also recurrently  
410 identified in EwS tumors (**Supplementary Fig. 4A**). EWIma1, 1<sup>\*</sup>, 1<sup>#</sup>, 5, 11 and 14 displayed  
411 chromoplectic-like translocation patterns (**Fig. 1C, 1D, 2H, 2I, Supplementary Fig. 4A**).  
412 This result is strikingly representative of the recently reported 30-40% incidence of  
413 chromoplexy in EwS tumors (28). Whereas the chromoplectic-like pattern of experimental  
414 independent EWIma1, EWIma1<sup>\*</sup> and EWIma1<sup>#</sup> clones is almost identical, (**Supplementary**  
415 **Fig. 2J**), differences in the translocation patterns were observed in the other complex models.  
416 For instance, EWIma5 and EWIma14 showed a chromoplectic-like karyotype with t(11;22)  
417 but also the additional translocation t(9;22) (between *CDKN2A* and *EWSR1*), addition of  
418 chr17p (probably from the *TP53* gene DNA break as only one allele can be detected) and loss  
419 of *STAG2* in all metaphases (**Fig. 2H and 2I, Supplementary Fig. 4A and 4B**). Altogether,  
420 immortalized EWIma1 cells display stable EWSR1-FLI1 expression and faithfully  
421 recapitulate EwS cells features.

422

423 **Engineering of t(11;22)(q24;q12) positive cells from multipotent mesenchymal precursor**  
424 **cells**

425 Since MSC<sup>Pat</sup> cells are extremely limited resources to generate these EwS models, we  
426 attempted to reproduce these results using published multipotent mesenchymal precursor cells  
427 (hMPCs)(24). An initial translocation frequency above 10<sup>-3</sup> independently of the presence or  
428 absence of *SPC* gRNA could be achieved in these cells (**Supplementary Fig. 5A**). As  
429 reported for MSC<sup>Pat</sup>, transfection of the unique pair of gRNA<sup>EWS</sup> and gRNA<sup>FLII</sup> in hMPCs  
430 resulted in a progressive loss of the t(11;22)(q24;q12) after two weeks of culture  
431 (**Supplementary Fig. 5B**). However, addition of *SPC* gRNAs increased the proliferation rate  
432 of bulk *EF* gRNAs transfected hMPCs cells (**Supplementary Fig. 5C**) and led to longer  
433 detection (up to 27 days) of *EF* fusion transcript (**Supplementary Fig. 5B**). Numerous clones  
434 grown in soft agar could be isolated from gRNA *EF*+*SPC* hMPC transfected cells and 0.9%  
435 contained the EwS translocation (3/336 clones). A similar frequency (0.7%, 3/408 isolated  
436 clones) of *EWSR1-FLII* positive clones was obtained if only additional gRNAs targeting  
437 *TP53* and *CDKN2A* (but not *STAG2*) were used. Small colonies were obtained in soft agar  
438 when gRNA<sup>EWS</sup> and gRNA<sup>FLII</sup> were transfected alone, but they could ultimately not be  
439 recovered after isolation.

440 We further analyzed *EWSR1-FLII* translocated clones, with and without *STAG2*  
441 mutations (**Supplementary Fig. 5D to 5K**). As in EwS cells, we could detect the *EWSR1-*  
442 *FLII* fusion transcript in all clones, with *STAG2* WT clone 1 exhibiting a very low level of  
443 transcript (**Supplementary Fig. 5F and 5G**). Expression of *EWSR1-FLII* was detected in 3  
444 clones for which we collected sufficient protein extract, confirming the low level of fusion  
445 protein in *STAG2* WT clone 1 (**Supplementary Fig. 5H**). *TP53* was mutated in these clones  
446 which expressed a truncated p53 protein (**Supplementary Fig. 5E and 5I**). *CDKN2A*  
447 mutations were also present in all four clones (**Supplementary Fig. 5E**). While no p16

448 expression was detected for the *STAG2* knockout (KO) clones, the *STAG2* WT clone 1  
449 showed p16 expression (related to the induced mutation that leads to a late stop codon)  
450 (**Supplementary Fig. 5E and 5I**). All clones expressed the EwS-specific cell surface marker  
451 CD99 (**Supplementary Fig. 5J**). However, and in contrast to EWIma clones obtained from  
452 MSC<sup>Pat</sup>, none of these *EWSR1-FLI1* translocated clones was fully immortalized *in vitro* and  
453 they all stopped growing after 2–3 months in culture (after agar selection). During that time,  
454 we observed a progressive telomere shortening associated with a weak telomerase activity,  
455 factors known to be deleterious for long-term culture (**Supplementary Fig. 5K**).

456

#### 457 **Transcriptomic analysis of t(11;22)(q24;q12) engineered mesenchymal precursor** 458 **models revealed a palette of EWSR1-FLI1 activation signature in these cells**

459 To further characterize the different models generated so far, we performed  
460 transcriptomic analyses in hMPCs, MSC<sup>Pat</sup> cells, and the derived models containing either the  
461 *EWSR1-FLI1* translocation and/or *SPC* mutations. These results were compared to  
462 transcriptomic profiles of 31 established EwS cell lines. Principal component analysis (PCA)  
463 revealed a striking weight of the first component (PC1) as compared to the other ones (**Fig.**  
464 **3A**). PC1 was clearly associated with EWSR1-FLI1 expression (**Fig. 3B**). Importantly, most  
465 EWIma1 cells clustered close to a collection of EwS cells but far apart from their parental  
466 MSC<sup>Pat</sup>. Conversely, MSC<sup>Pat</sup> with or without *SPC* mutations co-localized on PC1 axis and  
467 were moderately segregated by PC2 and PC3. Short-term silencing of EWSR1-FLI1 in  
468 EWIma1 cells reverted their PC1 component (**Fig. 2B and Supplementary Fig. 2H, 2I and**  
469 **6A**). Interestingly, EWIma 11, 12, 30 and 31 models (termed hereafter EWIma<sup>low</sup> models),  
470 which displayed a more mesenchymal intermediate morphology (**Fig. 2E and**  
471 **Supplementary Fig. 3E**) and a moderate telomerase activity (**Fig. 2F**) were clearly  
472 segregated from the other EWIma models (EWIma<sup>high</sup>, defined hereafter as all EWIma models

473 except EWIma<sup>low</sup>) on the PC1 axis (**Fig. 3B**). Similarly, using an unsupervised hierarchical  
474 clustering analysis (HCA), all EWIma<sup>high</sup> *in vitro* clones emerged distinctly of the original  
475 MSC<sup>Pat</sup>, hMPCs and EWIma<sup>low</sup>, from a slightly distant branch compared to EwS patient cells  
476 (**Fig. 3C**). In contrast, engineered clones obtained from hMPCs cells clustered in between  
477 EwS cell lines and their parental cells (**Fig. 3C**). Quite remarkably, EWSR1-FLI1 expression  
478 appeared as the key factor driving the segregation of mesenchymal cells from EwS  
479 populations in both PCA and HCA. MSC<sup>Pat</sup> carrying SPC alterations or hMPC-derived  
480 models displaying lower EWSR1-FLI1 levels (in particular clones 1 and 2 from hMPCs WT  
481 STAG2) (**Fig. 3B** and **3C**), clustered close to WT MSC<sup>Pat</sup> or hMPCs (**Fig. 3C**). To measure  
482 PC1 activity and evaluate the transition state of a broader collection of EWIma models  
483 without having to perform RNA-seq for each clone, we wondered if a small panel of known  
484 EWSR1-FLI1 transcriptional activated (*EGR2*, *NKX2-2*, *PRKCB*) (21,26,27) and repressed  
485 (*TNC*, *DKK1*, *IGFBP3*) target genes could be used as a surrogate marker using an RT-QPCR  
486 approach. For this, we defined an “EWINGness” score as the sum of log<sub>2</sub>FC (*EGR2* + *NKX2-*  
487 *2* + *PRKCB*) - (*TNC* + *DKK1* + *IGFBP3*). Quite remarkably, a strong correlation (R<sup>2</sup>=0.93)  
488 between PC1 and EWINGness scores was observed in MSC<sup>Pat</sup> and EWIma models  
489 (**Supplementary Fig. 6B**). A similar observation was made when all data were considered  
490 (R<sup>2</sup>=0.85) and not surprisingly, this correlation was poor among EwS cell lines (R<sup>2</sup>=0.08)  
491 (**Supplementary Fig. 6B**). Using this approach, we were able to evaluate the “EWINGness”  
492 of 30 EWIma clones (**Supplementary Table 2**) emphasizing a broad palette of EWSR1-FLI1  
493 activation signature in these models (**Supplementary Fig. 6C**). Indeed, clones that were  
494 negative for the EWSR1-FLI1 translocation but mutated for *SPC* typically displayed a  
495 EWINGness score below 25. Independently of an analysis on their morphological aspect,  
496 EWIma<sup>low</sup> model had an intermediate EWINGness score which was comparable to EWIma1  
497 cells silenced for EWSR1-FLI1. Similarly, hMPC-derived models, which were not fully

498 transformed, clearly scored in the intermediate 25-50 window. All other EWIma models  
499 presenting an EWSR1-FLI1 translocation had a score > 50, which was also observed in a  
500 panel of 22 EwS cell lines (**Supplementary Fig. 6C**). We further explored the transcriptomic  
501 signature between MSC<sup>Pat</sup> and EWIma models using gene-set enrichment analysis (GSEA).  
502 Heatmap of the top50 features for each phenotype highlighted again the intermediate  
503 signature of EWIma<sup>low</sup> models (**Supplementary Fig. 6D**). These last were removed from  
504 further GSEA to identify gene set signatures correlated with MSC<sup>Pat</sup> or EWIma<sup>high</sup> models.  
505 Quite remarkably, published EWSR1-FLI1 activation signatures ranked among the top50  
506 EWIma correlated signatures (among 18580 investigated signatures) (**Supplementary Table**  
507 **3**). Similarly, four E2F family member signatures were also identified in this top50 (**Fig. 3D,**  
508 **Supplementary Table 3**). Among those and of particular interest, E2F3 was previously  
509 shown to co-localize with EWSR1-FLI1 and to participate in the deregulation of cell cycle  
510 control of EwS (34). Conversely, signatures associated with a mesenchymal state were  
511 strongly enriched in MSC<sup>Pat</sup> GSEA analysis (**Fig. 3D, Supplementary Table 3**).

512

### 513 **EWIma cells display tumorigenic and metastatic properties in mice**

514 Ultimately, to evaluate the ability of the above described models to give rise to tumors  
515 *in vivo*, we performed orthotopic intra-femoral xenograft experiments. When injecting four  
516 hMPC derived models (*STAG2* KO clone 1, n=4; *STAG2* KO clone 2, n=4; *STAG2* WT  
517 clone1 hMPC, n=4; *STAG2* WT clone2 1, n=3), no tumors were detected in any of these mice  
518 after 5.5 months. Remarkably, when injecting two chromoplectic-like EWIma models  
519 (EWIma1, n=4 and EWIma5, n=2) and one simple rearranged model (EWIma7, n=4), all  
520 mice developed tumors that reached ethical endpoint criteria within 44-60 days (EWIma1),  
521 84-90 days (EWIma7) and 90 days (EWIma5) post injection. These tumors were positive for  
522 t(11;22;)(q24;q12) and expressed EWSR1-FLI1 (**Fig. 4A**). Necropsy revealed distant

523 metastases to the lungs and liver with all 3 models. Histological analysis evidenced a typical  
524 EwS small-round-cell morphology at primary and metastatic sites (**Fig. 4B, Supplementary**  
525 **Fig. 7 A and 7B**). Immunohistochemistry (IHC) experiments confirmed a strong and  
526 homogeneous CD99 membrane staining, a nuclear FLI1 pattern and absence of nuclear  
527 STAG2 expression (**Fig. 4B, Supplementary Fig. 7A and 7B**). These tumors were highly  
528 proliferative and displayed a non-apoptotic pattern as revealed respectively by Ki67 and  
529 cleaved caspase 3 staining (**Fig. 4B, Supplementary Fig. 7A and 7B**). Finally, we profiled  
530 these EWIma tumors using RNA-seq. Notably, using unsupervised hierarchical clustering  
531 analysis (HCA), transformed EWIma1, 5 and 7 tumors clustered together with EwS cells,  
532 whereas their *in vitro* respective models emerged from a more distant branch (**Fig. 3C**).  
533 Similarly, these tumors were clearly more left shifted on PC1 axis as compared to their  
534 respective *in vitro* counterparts (**Fig. 3B**). All together, these EWIma models likely represent  
535 a novel and large panel of *de novo* generated EwS cellular models with immortalized and  
536 transforming properties.

537

538 **DISCUSSION**

539 In the present work, we successfully and efficiently generated EWSR1-FLI1  
540 transformed cells starting from “normal” non-cancerous MSCs of a EwS patient. These  
541 models (EWIma) faithfully recapitulated *bona fide* EwS characteristics, including cell  
542 morphology, transcriptomic, epigenetic, metastatic and plasticity aspects that have been  
543 previously reported in EwS cells lines and tumors (12). In addition, we defined here a  
544 EWINGness score as a simple surrogate marker to evaluate transformation potential of  
545 mesenchymal stem cells towards Ewing sarcoma. Our results further support that these  
546 MSC<sup>Pat</sup> are permissive to *EWSR1-FLI1* expression under its *EWSR1* endogenous promoter  
547 and ultimately leads to their transformation *in vivo*. This demonstrates that Ewing sarcoma  
548 can originate from human bone marrow-derived MSC as previously anticipated but never  
549 demonstrated so far (8, 9, 11, 14). Whereas this cell is the only permissive one remains to be  
550 formally elucidated. For instance, repeating our experimental approach in neural-crest-derived  
551 or other stem/progenitor populations would be complementary to define if various cells of  
552 origin in EwS exist (13,35,36). The EwS tumor generated in this study combines endogenous  
553 *EWSR1-FLI1* translocation together with most recurrent mutations found in EwS. How  
554 *STAG2*, *TP53*, and *CDKN2A* alterations specifically contribute to transformation in our model  
555 remains to be clarified in future studies. However, even if the *SPC* mutations appear to confer  
556 a growth advantage to mesenchymal stem cells (**Supplementary Fig. 5C**), they do not appear  
557 to confer a “primed” Ewing transcriptomic signature (MSC<sup>Pat</sup> vs MSC<sup>Pat</sup><sub>SPC</sub>), which is  
558 clearly mediated by the EWSR1-FLI1 transcription factor in our MSC<sup>Pat</sup> or hMPC derived  
559 models (**Fig. 3B** and **3C**). *STAG2* and *TP53* mutations can co-occur in EwS at diagnosis and  
560 appear to define an aggressive subtype (6). We also recently demonstrated that *STAG2* loss-  
561 of-function (LOF) mutations reduced the cis-mediated activity of EWSR1-FLI1 (37). In that  
562 respect, we can speculate that *STAG2* LOF in EWIma models may attenuate the known



563 EWSR1-FLI1 toxicity and therefore favors the emergence of these clones. We also showed  
564 that STAG2 LOF increased migratory properties of EwS cells, including in EWIma1 cells  
565 (37), which was also previously reported at the clinical level to be associated with metastasis  
566 (5) and poor outcome (4). In possible support of this notion, orthotopically engrafted  
567 EWIma1, 5 and 7 cells also grew at distant sites such as in the lungs and liver. In addition, the  
568 original EwS tumor cells of this patient displayed two chromosomal deletions containing  
569 *CDKN2A* and *TP53* at the time of the diagnosis, and gene expression data showed a transcript  
570 alteration of *STAG2* (**Supplementary Fig. 2A**). These findings raise the possibility that the  
571 simultaneous alteration of p16, p53 and *STAG2* expression had a direct “boosting” effect on  
572 Ewing sarcomagenesis in this particular patient. Whereas individual or combined *SPC*  
573 mutations in this particular patient tumor or more generally in EwS are concomitant to the  
574 translocation in EwS tumors or appear as secondary events remains to be elucidated, and both  
575 scenarios may occur. Since *SPC* mutations are absent from many EwS tumors at the time of  
576 diagnosis, it is likely that other combinations of more private mutations together with the  
577 pathognomic translocation may also allow to successfully transform MSC into faithful EwS  
578 models. Indeed, on average, ten coding variants per tumor were detected in EwS tumors at the  
579 time of diagnosis (6) and 120 unique genes were involved in chromoplectic breakpoints in  
580 Ewing sarcoma (28). Finally, the time scale and the *in vitro* aspects of our approach may also  
581 explain why we only transformed few MSC<sup>Pat</sup> with the *EF+SPC* cocktail as several years  
582 and/or microenvironmental factors may be necessary to fully transform a cell of origin. For  
583 instance, using clock-like mutation signatures in primary and relapse EwS tumors, it was  
584 estimated that the cell that would give rise to the relapse existed years before diagnosis (28).  
585 The over-proportional weight of the first PCA component and the GSEA signatures (**Fig. 3A**  
586 and **3D**) highlighted that the EWSR1-FLI1 transcriptional signature is the predominant  
587 feature in our model. Notably, the EWSR1-FLI1 binding pattern at both GGAA-mSats and

588 canonical ETS-like binding sites in EWIma1 were strikingly overlapping with that of the  
589 established A673 EwS cell line but highly divergent from the FLI1 binding pattern observed  
590 in MSC<sup>Pat</sup> (**Fig. 1G** and **1H**). Acquisition of well-known SEs, reminiscent of a specific EwS  
591 identity, clearly demonstrates that EWIma1 cells also display *bona fide* (neo)-enhancer  
592 properties for EWSR1-FLI1 (**Fig. 1H**). Hierarchical clustering showed that EWIma tumors  
593 cluster within a large panel of EwS cell lines. Yet, PC2 which mostly discriminated hMPC  
594 and MSC<sup>Pat</sup> derived models, also slightly segregated EWIma1 from EwS cell lines (**Fig. 3B**).  
595 Interestingly, all EWIma tumors display additional ‘EWINGness’ (similar PC1 values to EwS  
596 cell lines) as compared to their respective EWIma *in vitro* models (**Fig. 3B** and **3C**).  
597 Exogenous signaling from the microenvironment may account for this difference but remains  
598 to be determined.

599         In this study, we suggested the existence of a ‘permissive milieu’ that could alleviate  
600 the potential toxicity of EWSR1-FLI1 expression while favoring its appropriate  
601 regulation. We recently demonstrated using GWAS that at least 6 loci were significantly  
602 associated with Ewing sarcoma (21). Future experiments, using MSCs with different  
603 genotypes at susceptibility loci will enable to more precisely decipher the key genetic  
604 elements that are required for EWSR1-FLI1-induced transformation. Besides, recent single  
605 cell RNA-seq study of EwS patient-derived xenograft (PDX) tumors highlighted a window in  
606 which EwS cells can proliferate (23). Low levels of EWSR1-FLI1 were associated with  
607 mesenchymal and apoptotic phenotypes (14, 38), while EwS cells displaying very high  
608 EWSR1-FLI1 activity led to absence of proliferation and HIF1 $\alpha$  pathway activation (23). In  
609 that respect, the collection of EWSR1-FLI1 positive clones generated in this study display a  
610 broad and heterogeneous range of EWINGness scores, possibly recapitulating various levels  
611 of EWSR1-FLI1 transcriptional activity. However, although all EWIma models generated  
612 from MSC<sup>Pat</sup> carried EWSR1-FLI1 translocation and *SPC* mutations, not all displayed fully

613 immortalized patterns indicating that additional factors (e.g. stemness, cell cycle, oxphos  
614 status...) may contribute to the transformation of the EwS cell of origin. Additional  
615 investigation with these valuable models will be necessary to answer these key questions.

616 Here, we also attempted but did not succeed to engineer a transformed EwS model  
617 with hMPC cells. Whereas hMPC cells display a *bona fide* multipotent differentiation  
618 potential (24), we anticipated that MSC derived from healthy adolescent bone marrow (match  
619 of MSC<sup>Pat</sup> conditions) would have been a better control. However, we could not collect such  
620 controls due to the very limited occurrence of such pediatric samples.

621 Besides modelling Ewing sarcomagenesis, we genetically engineered cells with  
622 different karyotypes, including some that are reminiscent of chromoplexy (EWIma 1, 5 and  
623 14), which has been described in ~35% of EwS tumors. Chromoplexy comprises multiple  
624 chromosomal translocations that reshuffles chromosomes in a new and scrambled  
625 configuration instead of creating simple reciprocal translocations. Recently, chromoplexy has  
626 been described in 17.8% of 2,648 whole-cancer genomes from 38 tumor types (39,40) and  
627 plausibly as the source of their oncogenic transformation. However, if the exact mechanism of  
628 chromoplexy remains to be fully elucidated, modelling this event, as made possible with our  
629 gRNA cocktail approach, is of major interest for cancer research. Indeed, in our EWIma  
630 models, most chromoplectic-like rearrangements were proved to be initiated at gRNA target  
631 sites (and off-target sites for gRNA<sup>CDKN2A</sup>), indicating that they originated from a single burst  
632 in MSC<sup>Pat</sup> at the time of *EWSRI-FLII* translocation formation as suggested by genomic data  
633 on EwS tumors (28). In addition, genomic regions implicated in chromoplexy are often found  
634 in early replicating regions, rich in expressed genes (27, 28, 41). Remarkably, all *loci*  
635 implicated in chromoplectic-like events in EWIma1 cells (including the intergenic off-target  
636 site of gRNA<sup>CDKN2A</sup>) were located within early replicating domains of the human MSC  
637 genome (**Supplementary Fig. 7C**)(42). In that respect, our approach may represent an

638 attractive model to investigate how ‘normal’ cells adapt to such a catastrophic burst of  
639 rearrangements. In addition to chromoplexy, chromosomal alterations that have been  
640 observed in EwS are also present in some of our EWIma models. For instance, a duplication  
641 of chr20q is observed in EWIma1. In EwS, trisomy or focal amplifications of chr20 have been  
642 described in up to 15% of these tumors (43, 44). Similarly, deletion of chr16q observed in  
643 EwS tumors (43,44) is particularly obvious in EWIma2 cells (**Supplementary Fig. 4A**).

644         Recent studies allowed to reconstruct clonal and temporal evolution of tumors using  
645 mutational signatures, multiple spatio-temporal tumor sampling and/or single cell sequencing  
646 approaches. These top-down approaches allow to speculate about the timing of genetic lesions  
647 in the cell of origin without, however, achieving this original stage (40, 45). Our bottom-up  
648 strategy is very complementary to these approaches that were also used in EwS to speculate  
649 about the timing of the translocation and of the additional alterations. In addition, since EwS  
650 genetic susceptibility loci have been identified (21), it would be interesting to expand this  
651 collection when starting from additional untransformed cell of origin collections (possibly  
652 derived from EwS and non-EwS patients) which may ultimately allow to determine how  
653 eQTL related genes affect Ewing sarcomagenesis. Combining top-down and bottom-up  
654 strategies may ultimately allow to answer these complex questions, especially in sarcoma  
655 where many new entities are presumed to be driven by candidate gene fusion oncogenes (46).

656         In conclusion, this work demonstrates that EwS can originate from bone marrow  
657 derived mesenchymal stem cells. It further provides evidence of the necessity to reach a  
658 minimal level of EWSR1-FLI1-mediated transcriptional activity, within a defined genomic  
659 susceptibility context, to achieve full immortalization and transformation of this cell of origin.  
660 All together, we successfully bypassed here the challenge of modeling EwS *ab initio*. Our  
661 model mimics a rather aggressive form of EwS with *SPC* mutations displaying single  
662 balanced EWSR1-FL1 translocation but also chromoplectic-like events and transforming

663 properties in mice. Our successful approach to generate *bona fide* EwS cells opens broad  
664 avenues to gain important insights into Ewing sarcomagenesis but also into mechanisms  
665 related to chromoplexy. More generally, this transposable approach shall allow to investigate  
666 sarcomagenesis in the highly heterogeneous family of sarcoma tumors.

667

668 **Author contributions:**

669 A.S., M.H., L.B., S.Z., P.R., B.R., C.P.E., T.G.P.G. and D.S. carried out experiments. S.G.  
670 performed the bioinformatics analyses. S.G. analyzed processed sequencing data. A.D.C. and  
671 C.G. produced and provided key reagents. L.C. and S.K. analyzed karyotypes of EWIma  
672 clones. G.P and F.T. provided patient sequencing data. M.T., I.J.L and M.J. provided cells  
673 and valuable expertise. O.D., D.S. and E.B. originally conceived the project, designed the  
674 experiments and supervised the work. A.S., D.S. and E.B wrote the paper with the help of  
675 O.D., M.J., and T.G.P.G. who provided valuable feedback, and all authors reviewed and  
676 agreed on the final manuscript.

677

678 **Acknowledgements**

679 We thank the iPS (Nathalie Lefort), Imagine cytometry (Olivier Pellé), and microscopy  
680 (Meriem Garfa-Traore) platforms for help with SNP karyotyping, cell analysis and image  
681 analysis, Lina El Kassar (CECS/I-STEM), Dr. Chloé Lescale and Dr. Ludovic Deriano  
682 (Institut Pasteur) for help with the multicolor FISH analysis and Dr. Marion Piganeau for  
683 technical support and helpful discussions. We thank Rosalie Borry and each member of the  
684 ‘Genome Dynamics in the Immune System’ lab for scientific discussions and technical  
685 support. We thank E. Lapouble, G. Pierron, C. Thirant, R. Leclerc, V. Raynal, S. Baulande, P.  
686 Legoix, C. Kamoun, M-M. Aynaud and E. Barillot and all members of the Genetics and  
687 Biology of Pediatric Cancers laboratory for helpful discussions and/or technical or

688 bioinformatics assistance. We thank H. Kovar for providing Ewing sarcoma cell lines and  
689 COG Childhood Cancer Repository for the CHLA cell line.

690  
691 **References**

- 692 1. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, *et al.* Gene fusion  
693 with an ETS DNA-binding domain caused by chromosome translocation in human  
694 tumours. *Nature* **1992**;359:162-5
- 695 2. Lessnick SL, Dacwag CS, Golub TR. The Ewing's sarcoma oncoprotein EWS/FLI  
696 induces a p53-dependent growth arrest in primary human fibroblasts. *Cancer Cell*  
697 **2002**;1:393-401
- 698 3. Sohn EJ, Li H, Reidy K, Beers LF, Christensen BL, Lee SB. EWS/FLI1 oncogene  
699 activates caspase 3 transcription and triggers apoptosis in vivo. *Cancer Res*  
700 **2010**;70:1154-63
- 701 4. Brohl AS, Solomon DA, Chang W, Wang J, Song Y, Sindiri S, *et al.* The genomic  
702 landscape of the Ewing Sarcoma family of tumors reveals recurrent STAG2 mutation.  
703 *PLoS Genet* **2014**;10:e1004475
- 704 5. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, *et al.*  
705 The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov* **2014**;4:1326-41
- 706 6. Tirode F, Surdez D, Ma X, Parker M, Le Deley MC, Bahrami A, *et al.* Genomic  
707 landscape of Ewing sarcoma defines an aggressive subtype with co-association of  
708 STAG2 and TP53 mutations. *Cancer Discov* **2014**;4:1342-53
- 709 7. Minas TZ, Surdez D, Javaheri T, Tanaka M, Howarth M, Kang HJ, *et al.* Combined  
710 experience of six independent laboratories attempting to create an Ewing sarcoma  
711 mouse model. *Oncotarget* **2017**;8:34141-63
- 712 8. Riggi N, Cironi L, Provero P, Suva ML, Kaloulis K, Garcia-Echeverria C, *et al.*  
713 Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal  
714 progenitor cells. *Cancer Res* **2005**;65:11459-68
- 715 9. Castellero-Trejo Y, Eliazar S, Xiang L, Richardson JA, Ilaria RL, Jr. Expression of the  
716 EWS/FLI-1 oncogene in murine primary bone-derived cells Results in EWS/FLI-1-  
717 dependent, ewing sarcoma-like tumors. *Cancer Res* **2005**;65:8698-705
- 718 10. El Beaino M, Liu J, Wasylishen AR, Pourebrahim R, Migut A, Bessellieu BJ, *et al.*  
719 Loss of Stag2 cooperates with EWS-FLI1 to transform murine Mesenchymal stem  
720 cells. *BMC Cancer* **2020**;20:3
- 721 11. Riggi N, Suva ML, De Vito C, Provero P, Stehle JC, Baumer K, *et al.* EWS-FLI-1  
722 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell  
723 reprogramming toward Ewing sarcoma cancer stem cells. *Genes Dev* **2010**;24:916-32
- 724 12. Grünewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Alava E, Kovar H, *et al.*  
725 Ewing sarcoma. *Nat Rev Dis Primers* **2018**;4:5
- 726 13. von Levetzow C, Jiang X, Gwye Y, von Levetzow G, Hung L, Cooper A, *et al.*  
727 Modeling initiation of Ewing sarcoma in human neural crest cells. *PLoS One*  
728 **2011**;6:e19305
- 729 14. Tirode F, Laud-Duval K, Prieur A, Delorme B, Charbord P, Delattre O. Mesenchymal  
730 stem cell features of Ewing tumors. *Cancer Cell* **2007**;11:421-9
- 731 15. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, *et al.*  
732 Cancer translocations in human cells induced by zinc finger and TALE nucleases.  
733 *Genome research* **2013**;23:1182-93

- 734 16. Torres-Ruiz R, Martinez-Lage M, Martin MC, Garcia A, Bueno C, Castano J, *et al.*  
735 Efficient Recreation of t(11;22) EWSR1-FLI1(+) in Human Stem Cells Using  
736 CRISPR/Cas9. *Stem Cell Reports* **2017**;8:1408-20
- 737 17. Brunet E, Simsek D, Tomishima M, DeKolver R, Choi VM, Gregory P, *et al.*  
738 Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl*  
739 *Acad Sci U S A* **2009**;106:10620-5
- 740 18. Fraumeni JF, Jr., Glass AG. Rarity of Ewing's sarcoma among U.S. Negro children.  
741 *Lancet* **1970**;1:366-7
- 742 19. Jawad MU, Cheung MC, Min ES, Schneiderbauer MM, Koniaris LG, Scully SP.  
743 Ewing sarcoma demonstrates racial disparities in incidence-related and sex-related  
744 differences in outcome: an analysis of 1631 cases from the SEER database, 1973-  
745 2005. *Cancer* **2009**;115:3526-36
- 746 20. Randall RL, Lessnick SL, Jones KB, Gouw LG, Cummings JE, Cannon-Albright L, *et*  
747 *al.* Is There a Predisposition Gene for Ewing's Sarcoma? *J Oncol* **2010**;2010:397632
- 748 21. Machiela MJ, Grünewald TGP, Surdez D, Reynaud S, Mirabeau O, Karlins E, *et al.*  
749 Genome-wide association study identifies multiple new loci associated with Ewing  
750 sarcoma susceptibility. *Nat Commun* **2018**;9:3184
- 751 22. Grünewald TG, Bernard V, Gilardi-Hebenstreit P, Raynal V, Surdez D, Aynaud MM,  
752 *et al.* Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2  
753 via a GGAA microsatellite. *Nat Genet* **2015**;47:1073-8
- 754 23. Aynaud MM, Mirabeau O, Gruel N, Grossetete S, Boeva V, Durand S, *et al.*  
755 Transcriptional Programs Define Intratumoral Heterogeneity of Ewing Sarcoma at  
756 Single-Cell Resolution. *Cell Rep* **2020**;30:1767-79.e6
- 757 24. Barberi T, Willis LM, Socci ND, Studer L. Derivation of multipotent mesenchymal  
758 precursors from human embryonic stem cells. *PLoS Med* **2005**;2:e161
- 759 25. Fehrer C, Brunauer R, Laschober G, Unterluggauer H, Reitingner S, Kloss F, *et al.*  
760 Reduced oxygen tension attenuates differentiation capacity of human mesenchymal  
761 stem cells and prolongs their lifespan. *Aging Cell* **2007**;6:745-57
- 762 26. Renouf B, Piganeau M, Ghezraoui H, Jasin M, Brunet E. Creating Cancer  
763 Translocations in Human Cells Using Cas9 DSBs and nCas9 Paired Nicks. *Methods in*  
764 *enzymology* **2014**;546:251-71
- 765 27. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, *et al.*  
766 Punctuated evolution of prostate cancer genomes. *Cell* **2013**;153:666-77
- 767 28. Anderson ND, de Borja R, Young MD, Fuligni F, Rosic A, Roberts ND, *et al.*  
768 Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors.  
769 *Science* **2018**;361
- 770 29. Boeva V, Surdez D, Guillon N, Tirode F, Fejes AP, Delattre O, *et al.* De novo motif  
771 identification improves the accuracy of predicting transcription factor binding sites in  
772 ChIP-Seq data analysis. *Nucleic Acids Res* **2010**;38:e126
- 773 30. Kennedy AL, Vallurupalli M, Chen L, Crompton B, Cowley G, Vazquez F, *et al.*  
774 Functional, chemical genomic, and super-enhancer screening identify sensitivity to  
775 cyclin D1/CDK4 pathway inhibition in Ewing sarcoma. *Oncotarget* **2015**;6:30178-93
- 776 31. Baldauf MC, Orth MF, Dallmayer M, Marchetto A, Gerke JS, Rubio RA, *et al.* Robust  
777 diagnosis of Ewing sarcoma by immunohistochemical detection of super-enhancer-  
778 driven EWSR1-ETS targets. *Oncotarget* **2018**;9:1587-601
- 779 32. Marchetto A, Ohmura S, Orth MF, Knott MML, Colombo MV, Arrigoni C, *et al.*  
780 Oncogenic hijacking of a developmental transcription factor evokes vulnerability  
781 toward oxidative stress in Ewing sarcoma. *Nat Commun* **2020**;11:2423

- 782 33. Hendel A, Bak RO, Clark JT, Kennedy AB, Ryan DE, Roy S, *et al.* Chemically  
783 modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells.  
784 *Nature biotechnology* **2015**;33:985-9
- 785 34. Bilke S, Schwentner R, Yang F, Kauer M, Jug G, Walker RL, *et al.* Oncogenic ETS  
786 fusions deregulate E2F3 target genes in Ewing sarcoma and prostate cancer. *Genome*  
787 *research* **2013**;23:1797-809
- 788 35. Staeger MS, Hutter C, Neumann I, Foja S, Hattenhorst UE, Hansen G, *et al.* DNA  
789 microarrays reveal relationship of Ewing family tumors to both endothelial and fetal  
790 neural crest-derived cells and define novel targets. *Cancer Res* **2004**;64:8213-21
- 791 36. Tanaka M, Yamazaki Y, Kanno Y, Igarashi K, Aisaki K, Kanno J, *et al.* Ewing's  
792 sarcoma precursors are highly enriched in embryonic osteochondrogenic progenitors.  
793 *The Journal of clinical investigation* **2014**;124:3061-74
- 794 37. Surdez D, Zaidi S, Grossetête S, Laud-Duval K, Ferre AS, Mous L, *et al.* STAG2  
795 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and  
796 EWSR1-FLI1 activity in Ewing sarcoma. *Cancer Cell* **2021**
- 797 38. Stoll G, Surdez D, Tirode F, Laud K, Barillot E, Zinovyev A, *et al.* Systems biology  
798 of Ewing sarcoma: a network model of EWS-FLI1 effect on proliferation and  
799 apoptosis. *Nucleic Acids Res* **2013**;41:8853-71
- 800 39. Campbell PJ, Getz G, Korbel JO. Pan-cancer analysis of whole genomes. *Nature*  
801 **2020**;578:82-93
- 802 40. Gerstung M, Jolly C, Leshchiner I, D'Antonio SC, Gonzalez S, Rosebrock D, *et al.* The  
803 evolutionary history of 2,658 cancers. *Nature* **2020**;578:122-8
- 804 41. Lee JJ, Park S, Park H, Kim S, Lee J, Lee J, *et al.* Tracing Oncogene Rearrangements  
805 in the Mutational History of Lung Adenocarcinoma. *Cell* **2019**;177:1842-57.e21
- 806 42. Rivera-Mulia JC, Buckley Q, Sasaki T, Zimmerman J, Didier RA, Natorf K, *et al.*  
807 Dynamic changes in replication timing and gene expression during lineage  
808 specification of human pluripotent stem cells. *Genome research* **2015**;25:1091-103
- 809 43. Jahromi MS, Jones KB, Schiffman JD. Copy Number Alterations and Methylation in  
810 Ewing's Sarcoma. *Sarcoma* **2011**;2011:362173
- 811 44. Ozaki T, Paulussen M, Poremba C, Brinkschmidt C, Rerimoli J, Ahrens S, *et al.* Genetic  
812 imbalances revealed by comparative genomic hybridization in Ewing tumors. *Genes*  
813 *Chromosomes Cancer* **2001**;32:164-71
- 814 45. Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S, *et al.*  
815 Clock-like mutational processes in human somatic cells. *Nat Genet* **2015**;47:1402-7
- 816 46. Watson S, Perrin V, Guillemot D, Reynaud S, Coindre JM, Karanian M, *et al.*  
817 Transcriptomic definition of molecular subgroups of small round cell sarcomas. *The*  
818 *Journal of pathology* **2018**;245:29-40
- 819
- 820



821 **FIGURE LEGENDS**

822

823 **Figure 1. Generation of EWIma1 cells derived from MSC<sup>Pat</sup> cells, recapitulating**  
824 **molecular and epigenetic features of EwS.**

825 **A-** CRISPR/Cas9 based strategy to obtain *EWSR1-FLI1* translocated clones from MSCs with  
826 or without inducing *STAG2*, *TP53* and *CDKN2A* additional mutations. EF: gRNA<sup>EWS</sup> and  
827 gRNA<sup>FLI1</sup>. EF+SPC: gRNA<sup>EWS</sup>, gRNA<sup>FLI1</sup>, gRNA<sup>TP53</sup>, gRNA<sup>CDKN2A</sup> and gRNA<sup>STAG2</sup>.

828 **B-** Cumulative cell counts over time in EWIma1 cells transfected with siRNA targeting  
829 *EWSR1-FLI1* (si-EF1) compared to wild type (WT) and siRNA control (si-CTL). Top,  
830 western blot against *EWSR1-FLI1*, shown at 4 days post transfection. Results represent the  
831 mean  $\pm$  SD from three independent experiments. \* p<0.05.

832 **C-** Representative image of spectral karyotype (SKY multi-colored fluorescent FISH  
833 analysis) obtained from EWIma1 cells with a chromoplectic like pattern (reciprocal  
834 translocations t(11;22)(q24;q12), t(13;17), t(17;9)) and derivative chromosome 11 (der11) of  
835 t(11;13)). N= 22. See also Supplementary Table 1.

836 **D-** Schematic circos plot of the main rearrangements seen in EWIma1 cells (dot lines:  
837 rearrangements not found in all cells).

838 **E-** Integrative Genomics Viewer representation for *FLI1* and H3K27ac ChIP-seq profiles at  
839 *DKK1* locus showing the disappearance of super-enhancer in EWIma1, and A673 cells  
840 compared to MSC<sup>Pat</sup>. Super-enhancers are framed in red.

841 **F-** Integrative Genomics Viewer representation for *FLI1* and H3K27ac ChIP-seq profiles at  
842 *PRKCB*, *CCND1* and *NKX2-2* loci showing the appearance of super-enhancers in EWIma1,  
843 and A673 cells compared to MSC<sup>Pat</sup>. Super-enhancers are framed in red.

844 **G-** Top two motifs predicted by ChIPMunk corresponding to known motifs in Jaspar database  
845 identified in EWIma1 and MSC<sup>Pat</sup> *FLI1* ChIP-seq data.

846 **H-** Left: Heatmap representation of *FLI1* ChIP-seq peaks ranked by intensity at GGAA  
847 microsatellite (GGAAm) or ETS sites in EWIma1 only or MSC<sup>Pat</sup> common sites. Right:  
848 Heatmap representation of H3K27ac ChIP-seq peaks sorted by ROSE SE rank in EWIma1  
849 and MSC<sup>Pat</sup> specific and common sites. A673 data are shown as control. Read density is  
850 displayed within a 5kb (H3K27ac or *FLI1*) window around peak center and color scale  
851 intensities are shown in normalized coverage (scale is shown on the right of each panel).

852

853 **Figure 2. Molecular characterization of a collection of EWIma cells derived from**  
854 **MSC<sup>Pat</sup> cells.**

855 **A-** Representative images of agar colony formation assays using combinations of gRNA<sup>EWS</sup>  
856 (*E*), gRNA<sup>FLI1</sup> (*F*), gRNA<sup>TP53</sup> (*P*), gRNA<sup>CDKN2A</sup> (*C*), and gRNA<sup>STAG2</sup> (*S*).

857 **B-** Colony size quantification mean  $\pm$  SD, p value (\* p= 0.0186; \*\* p=0.0029;  
858 \*\*\*\*p<0.0001); n>600 colonies per condition.

859 **C-** Nested PCR to detect the translocated chromosome derivative 22 (der22) on serial  
860 dilutions from a DNA pool of MSC<sup>Pat</sup> cells transfected with *EF+SPC* gRNAs (from 50 to 1.6  
861 ng in triplicates). Translocation frequency (f) is calculated as described in (26) using the  
862 assumption that a human diploid cell contains  $\sim$ 6 pg of DNA.

863 **D-** Western blot against *EWSR1-FLI1* in a panel of EWIma clones compared to positive  
864 (A673) and negative (MSC<sup>Pat</sup>) controls. Vinculin is used as loading control.

865 **E-** Representative images of cellular morphology for the negative *EWSR1-FLI1* translocated  
866 MSC-SPC-#5 clone (with typical MSC morphology), for a positive *EWSR1-FLI1*  
867 translocated EWIma11 (intermediate EwS morphology) and EWIma7 (classical EwS  
868 morphology) clones. Scale bar: 800  $\mu$ m.

869 **F-** Telomerase Repeat Amplification Protocol (TRAP) assay showing telomerase activity in  
870 the different EWIma clones compared to A673 Ewing cells (A), MSC-SPC-#5 (CTL) and  
871 MSC<sup>Pat</sup> cells.

872 **G-** Karyotype analysis of EWIma7.

873 **H-** Karyotype analysis of EWIma5.

874 **I-** Karyotype analysis of EWIma14.

875

876

877

878 **Figure 3. EWSR1-FLI1 transcriptional signature is predominant in the *de novo* models.**

879 **A-** Histogram for weight percentage of top 10 dimensions of principal component analysis  
880 (PCA).

881 **B-** PCA representation for parental and hMPC or MSC<sup>Pat</sup> derived models (including *in vitro*,  
882 *in vivo* and EWSR1-FLI1 silenced EWIma1 models) compared to a collection of EwS cell  
883 lines. Left: PC1 vs PC2 and right: PC1 vs PC3.

884 **C-** Unsupervised Hierarchical clustering and heatmap based on top 1% inter-quartile range  
885 gene expression values. EWIma<sup>low</sup> models (EWIma12, 11, 31 and 30) are framed in red.  
886 EWIma orthotopic tumors (EWIma1, 5 and 7) are framed in brown. EwS cell lines are framed  
887 in black. R<sub>n</sub>: biological replicates numbering.

888 **D-** GSEA enrichment plots from top signatures upon MSC<sup>Pat</sup> versus EWIma models  
889 comparison. Top MSC<sup>Pat</sup> correlated signatures included commonly down-regulated genes in  
890 mesenchymal progenitors upon EWSR1-FLI1 and EWSR1-ERG expression (Myagawa  
891 targets of EWSR1-ETS fusions DN) and the hallmark epithelial to mesenchymal transition  
892 gene sets. Top EWIma correlated signatures included the cell cycle independent EWSR1-  
893 FLI1 activation signature (IC\_EWS from (23)) and up-regulated genes in embryonic  
894 fibroblasts upon serum stimulation and E2F3 knockdown (KONG\_E2F3\_TARGETS). In this  
895 analysis, EWIma<sup>low</sup> models were excluded.

896

897 **Figure 4. *In vivo* tumors obtained from engineered EWIma7 cells**

898 **A-** Western blot against EWSR1-FLI1 in EWIma 5, 7 and 1 tumor cells and parental EWIma1  
899 cells.

900 **B-** Histology of EWIma7 tumors (representative images). Top panel row: left, H&E staining  
901 in tumors at primary orthotopic implantation site (scale bar 200 μm) and anti-CD99 IHC  
902 staining at lung metastatic sites (scale bar, middle panel: 1 mm, right panel: 50 μm). Middle  
903 row left panel: magnification of H&E in the primary orthotopic tumor displaying a classical  
904 small round tumor cell morphology (scale bar: 50 μm). Additional IHC stainings against  
905 CD99, FLI1, STAG2, Ki67 and cleaved CASP3 (cIcASP3) in primary orthotopic tumors are  
906 shown in middle and bottom panel rows (scale bar: 50 μm).

907

908 **Table 1: Representative panel of genetic and morphologic features of EWIma model.**

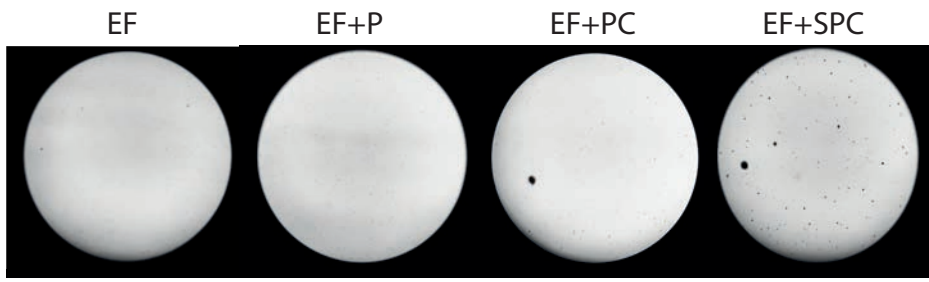
909 Sequences of *STAG2*, *TP53* and *CDKN2A* mutations and cellular morphology (+ classical  
910 EwS morphology, +/- intermediate EwS morphology or - MSC morphology) are indicated.  
911 Genomic sequences of der11 and der22 at breakpoints are indicated. MSC-SPC-#5 clone does  
912 not contain the EWSR1-FLI1 translocation and is used as negative control. EWINGness score  
913 for each EWIma model is displayed in the last column (see also **Supplementary Table 2**).  
914 “No seq” means that no sequence was PCR amplified for these clones probably due to large  
915 DNA deletions or presence of translocation implicated the gene.

916

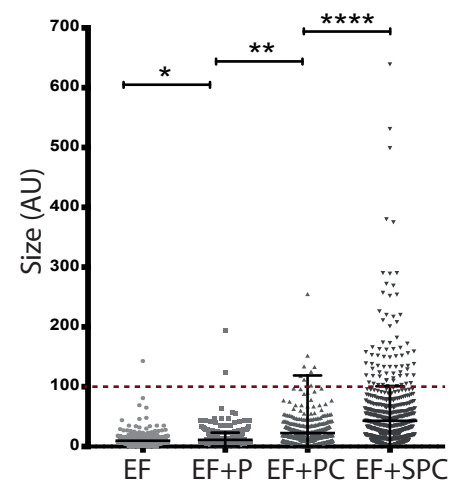


**Figure 2**

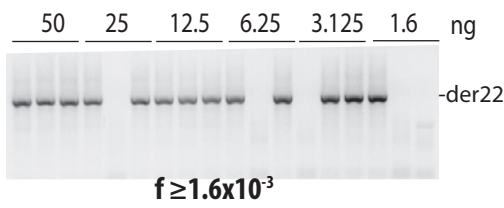
**A**



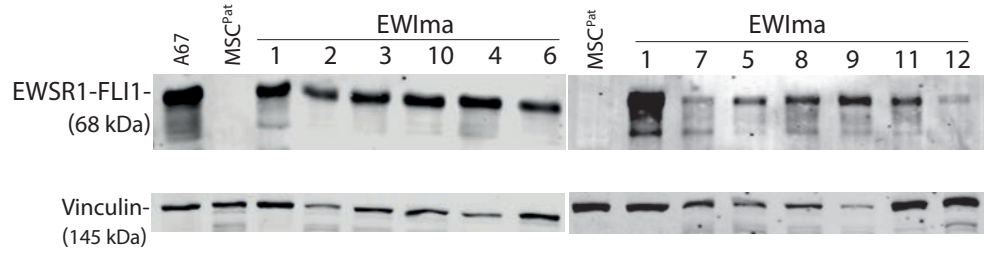
**B**



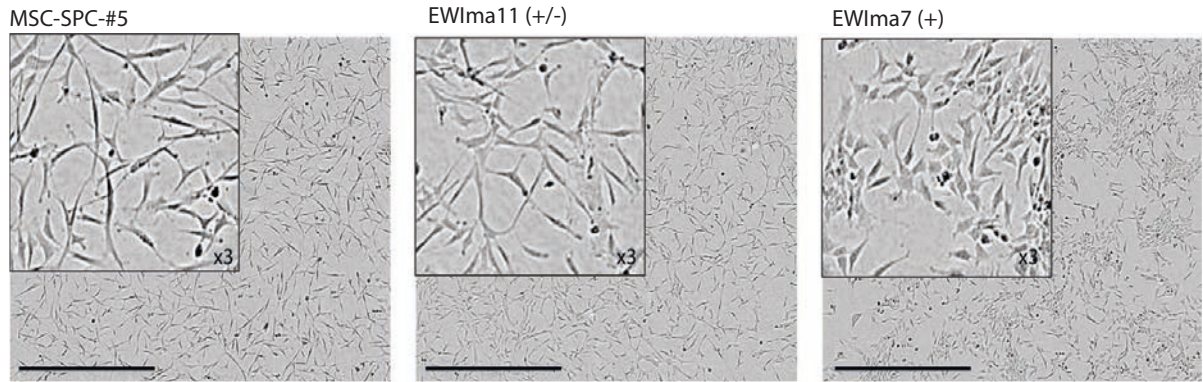
**C**



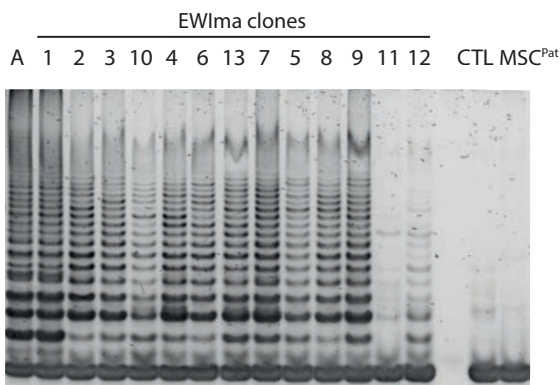
**D**



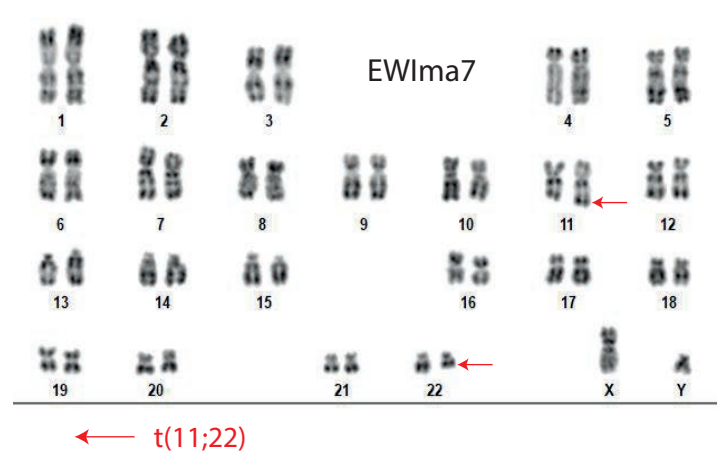
**E**



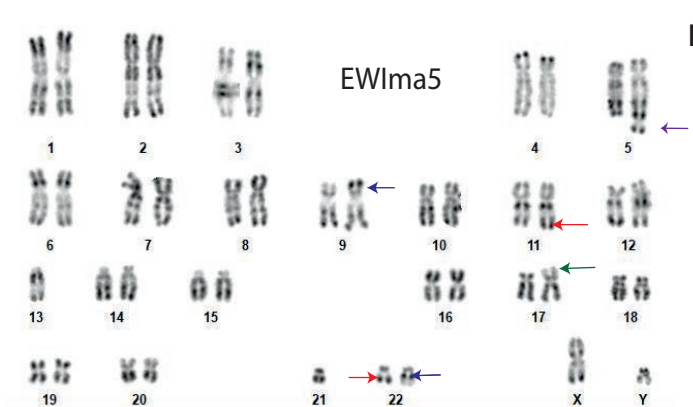
**F**



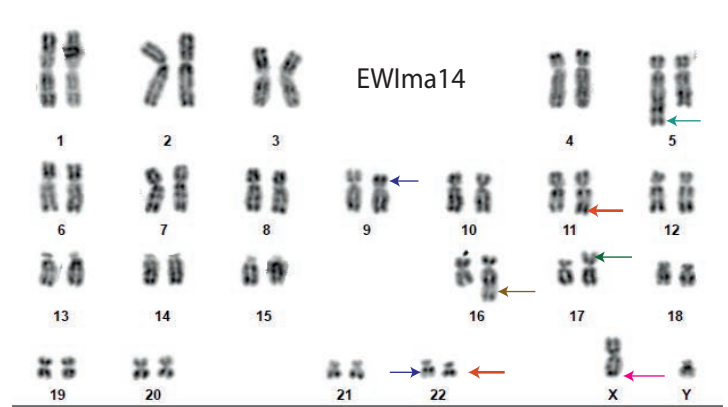
**G**



**H**



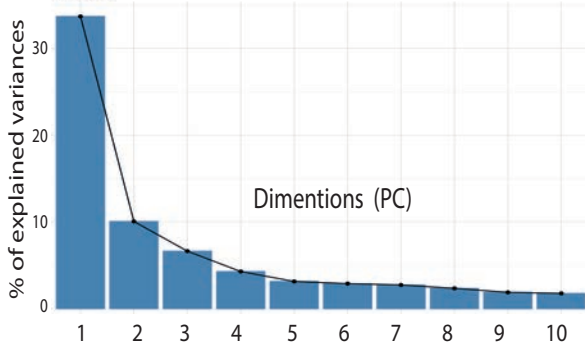
**I**



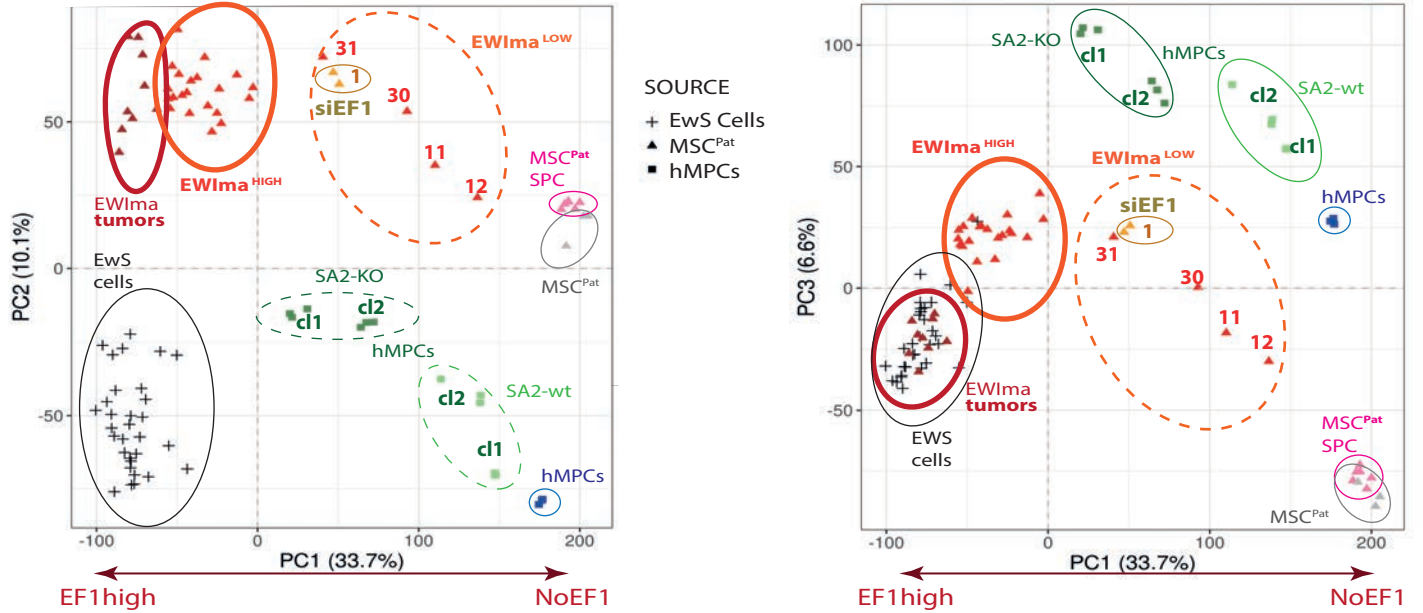
← t(11;22)  
 ← t(9;22)  
 ← gain 17  
 ← gain of 5q (70%)  
 ← t(11;22)  
 ← t(9;22)  
 ← gain 17  
 ← gain 16q (90%)  
 ← gain 5q (90%)  
 ← teloX (STAG2-)

**Figure 3**

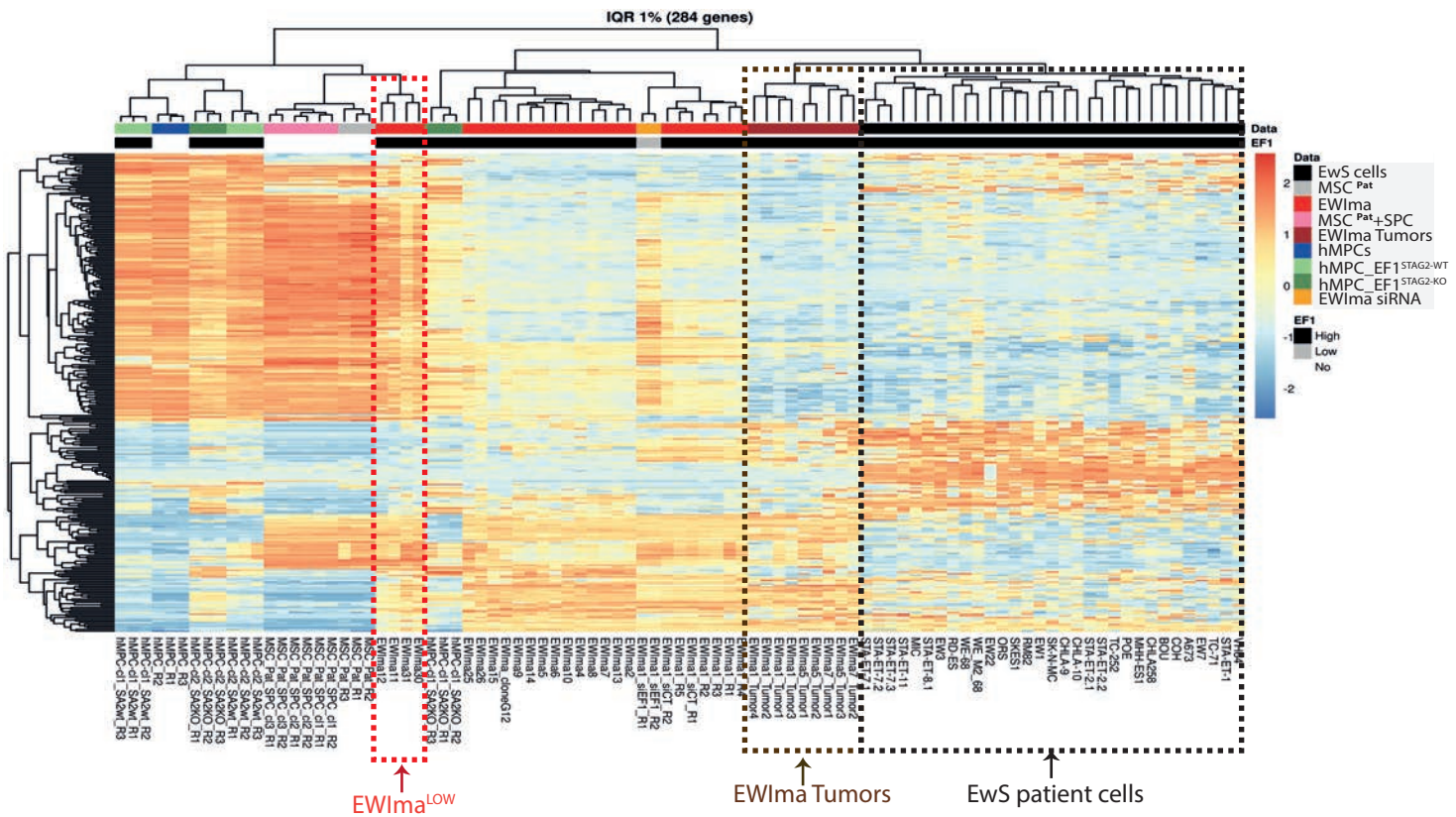
**A**



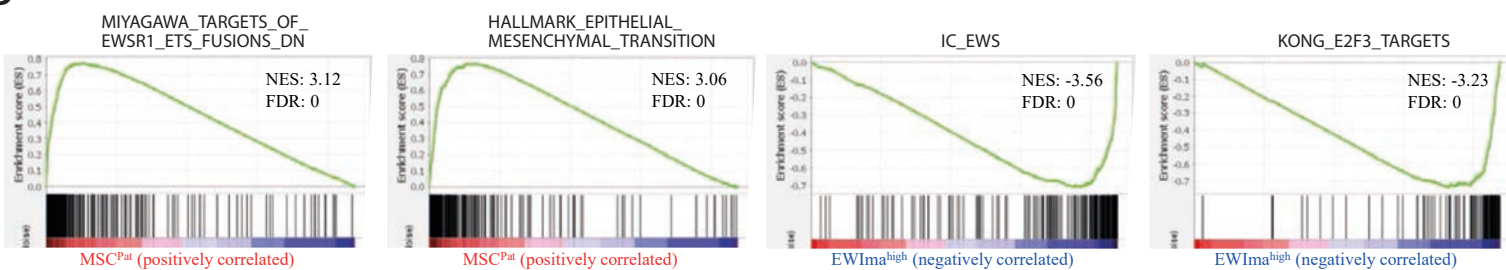
**B**



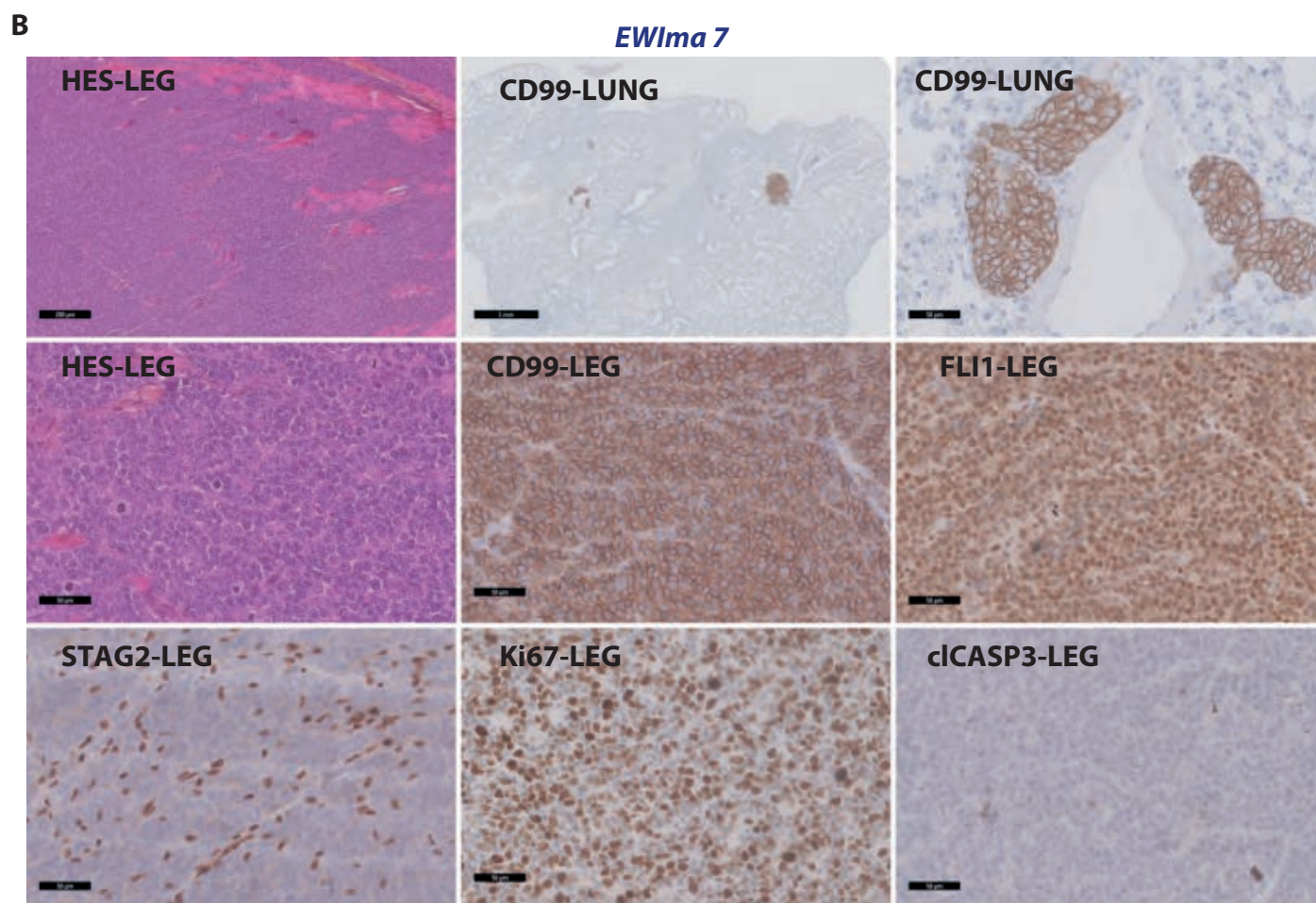
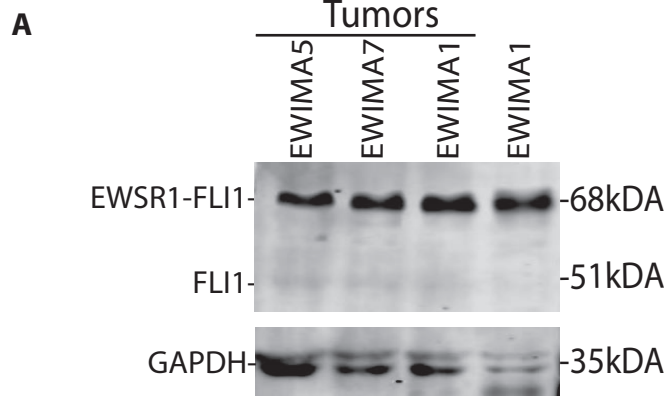
**C**



**D**



**Figure 4**



**Table 1**

Name	STAG2	TP53	CDKN2A	Mophology	Der11	Der22	Score*
EWIma2	c.661_662insT	c.831delG	c.512_513insG / c.512_514insTT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	70,0
EWIma3	c.661_662insT	no seq	c.512_513insG / c.512_514insTT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	66,6
EWIma13	c.661_662insT	c.831delG	c.512_513insG / c.512_514insTT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	66,5
EWIma10	c.661_662insT	c.831delG / c.817_838del	c.513delT / c.512_513del	+	Del204 - Del24	Del70 - Del15	66,2
EWIma14	c.661_662insT	c.811_830del	c.512_513insT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	65,4
EWIma4	c.661_662insT	c.831delG	c.512_513insG	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	65,4
EWIma6	c.661_662insT	c.831delG / c.832_833insCC	c.512delG / c.512_513insT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	64,5
EWIma7	c.661_662insT	c.831_832insG / c.831_837del	c.512_513insT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	61,4
EWIma9	c.661_662insT	c.831delG	c.512_513insT / c.511_512del	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	60,8
EWIma5	no seq	c.811_830del	c.512_513insT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	60,4
EWIma8	c.661_662insT	c.831_832insG / c.831_837del	c.512_513insT	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	60,1
EWIma1	c.661_662insT	c.831delG	c.511_519del	+	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	58,3
EWIma11	c.661_662insT	c.832_833insC / c.831delG + c.834C>T	c.512_515del / c.512_513insT	+/-	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	40,9
EWIma12	c.661_662insT	c.832_833insC / c.831delG + c.834C>T	c.512_513insG	+/-	TCCAGCTA-CTTCACAC	TTTCCTAT-TAAACATCT	33,5
MSC-SPC#5	c.660_662del	c.831_836del / c.815_836del	c.509_516del / c.512_513insT	-	-	-	0,0