# Disruption of Microhomology-mediated End-joining in Ewing Sarcoma

Shuhei Asada[1#, *], Guangli Zhu[1,2#], Jithma Prasad Abeykoon[1,3], Yutaro Tanaka[4,5], Huy Nguyen[1,6], Yuna Hirohashi[1], Divya R. Iyer[1], Nicholas William Ashton[1], Sirisha Mukkavalli[1], Lige Jiang[1], Miles Del Busso[3], Judith Jebastin Thangaiah[7], Steven I. Robinson[8], Kalindi Parmar[1,6], Eliezer M. Van Allen[5,9,10,11], Riaz Gillani[4,5,9,12,] Geoffrey I. Shapiro[10], Alan D. D'Andrea[1, 6*]

1. Division of Radiation and Genome Stability, Department of Radiation Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA.
2. Department of Radiation Oncology, First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China
3. Division of Hematology, Department of Internal Medicine, Mayo Clinic, MN, USA
4. Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA.
5. Cancer Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA.
6. Center for DNA Damage and Repair, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA.
7. Department of Pathology, Mayo Clinic, MN, USA
8. Department of Oncology, Mayo Clinic, MN, USA
9. Harvard Medical School, Boston, MA, USA.
10. Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA.
11. Center for Cancer Genomics, Dana-Farber Cancer Institute, Boston, MA, USA.
12. Boston Children's Hospital, Boston, MA, USA.
# These authors contributed equally to this work.

*Corresponding authors: Shuhei Asada (shuhei_asada@dfci.harvard.edu) and Alan D. D'Andrea (alan_dandrea@dfci.harvard.edu)

Keywords: Ewing's Sarcoma, DNA double-strand break, Alternative end-joining, TMEJ, DNA-PKc, CDK12

**SUMMARY**

Ewing sarcoma (EwS) is a group of bone and soft tissue cancers in children and young adults. The EWS-FLI1 protein, the oncogenic driver in EwS, is a fusion of the FET protein, EWSR1, at its amino terminus, with the carboxy terminus of the transcription factor, FLI1. Since EwS cells have pronounced sensitivity to radiation and chemotherapy-induced DNA damage, the role of EWS-FLI1 in DNA repair is likely. Here, we demonstrate that EWS-FLI1 causes a defect in microhomology-mediated end-joining (MMEJ) repair. EWSR1 is a splicing factor that promotes the faithful splicing of the *POLQ* pre-mRNA, required for the expression of POLΘ, a critical protein in the MMEJ pathway. Expression of EWS-FLI1, or loss of EWSR1, causes exon 25 skipping of the *POLQ* transcript, decreased POLΘ protein expression, impaired MMEJ, and cellular sensitivity to inhibitors of NHEJ repair and HR repair, through the mechanism of synthetic lethality. Knockdown of EWS-FLI1 expression restores POLΘ mitotic foci and increases MMEJ activity. Based on this mechanism, inhibitors of NHEJ or HR may provide a targeted therapy for patients with Ewing sarcoma.

**Highlights:**

-Ewing sarcoma tumors have a deficiency in POLθ expression and a corresponding loss of MMEJ activity

-EWSR1 is a splicing factor that interacts with other splicing factors such as FUBP1 and KHSRP/FUBP2 to accurately splice the *POLQ* mRNA.

-The EWS-FLI1 fusion oncoprotein, or loss of EWSR1, causes a splicing defect, leading to exon25 skipping of the *POLQ* pre-mRNA and loss of POLθ expression

-The MMEJ deficiency of EwS cells results in cellular sensitivity to inhibitors of Homologous Recombination or Non-Homologous End-Joining

-Exon 25 skipping of *POLQ* mRNA is a predictive biomarker for HR inhibitors in human cancers

67 **INTRODUCTION**

68

69 Ewing sarcoma (EwS) is a group of bone and soft tissue cancers in children and young
70 adults[1,2]. EwS tumor cells have a chromosome 11;22 translocation, resulting in expression of
71 the EWS-FLI1 fusion oncoprotein. Their elevated sensitivity to ionizing radiation and DNA-
72 damaging chemotherapy, as well as their relatively low level of somatic mutations, suggest that
73 EwS cells have an underlying defect in an error-prone DNA repair (DDR) pathway. Despite this
74 high drug sensitivity, novel treatments for EwS are required, especially for patients with relapsed
75 and/or refractory disease, who have a low 5-year overall survival rate of only 10%.

76 Recent studies have identified some abnormalities in DDR in Ewing sarcoma cells.
77 Expression of EWS-FLI1 fusion protein promotes R-loop formation and impaired BRCA1
78 recruitment to chromatin, suggesting a defect in the homologous-recombination repair pathway
79 (HRD) [3]. Despite this preclinical observation, PARP inhibition, an approved treatment for HR-
80 deficient tumors, was ineffective in the treatment of EwS patients[4]. Other cellular defects
81 resulting from EWS-FLI1 expression may indirectly affect DDR, such as defects in BAF
82 complex-mediated transcription [5,6], condensate generation [7-9], and mRNA splicing [10]. However,
83 the mechanism by which EWS-FLI1 specifically influences DDR pathways is unknown, and
84 other more profound DNA repair defects, beyond the HRD phenotype, may exist.

85 DNA double-strand breaks (DSBs) are the most critical forms of DNA damage in cells.
86 There are three main pathways for repairing DSBs - non-homologous end-joining (NHEJ),
87 homologous recombination (HR), and microhomology-mediated end-joining (MMEJ). For a
88 cancer cell to survive and proliferate, DSBs must be repaired, and the cell must rely on at least
89 one of these pathways. DNA polymerase theta (POLΘ), an error-prone polymerase encoded by
90 the *POLQ* gene, is required for the MMEJ pathway. Overexpression of *POLQ* confers

91 chemoresistance, whereas depletion of POLΘ results in cancer predisposition and
92 hypersensitivity to irradiation and chemotherapy [11,12]. *To date, there are no reported human*
93 *cancers with defects in the MMEJ pathway.*
94 Here, we demonstrate that EwS cells have an underlying defect in MMEJ. Expression of
95 EWS-FLI1, or loss of EWSR1, in these cells promotes exon skipping of the *POLQ* pre-mRNA,
96 generating a premature termination codon and a profound reduction of POLΘ protein expression.
97 EWS-FLI1 expressing cells exhibit a phenotype like *POLQ* deficient cells, including
98 hypersensitivity to topoisomerase inhibitors, irradiation, and synthetic lethality with loss of the
99 NHEJ or HR pathways. Knockout of the EWS-FLI1 oncoprotein in EwS cells results in reduced
100 exon skipping, re-expression of the POLΘ protein, and a rescue of MMEJ activity. Accordingly,
101 the DDR defect in EwS tumors could be exploited for therapeutic benefit by inhibiting another
102 DSB repair pathway.
103
104
105

106 **RESULTS**

107 ***EWSR1* knock-out cells phenocopy MMEJ-deficient cells.**

108 Previous CRISPR screens and Depmap analyses have demonstrated that genes in the
109 MMEJ pathway are synthetic lethal in cells with an underlying defect in an HR gene [13,14],
110 Interestingly, in these screens, knockout of the *EWSR1* gene was highly similar to knockout of
111 known MMEJ genes, including *POLQ* [15], *XRCC1* [16], *APEX2* [17], and *RHNO1* [18], in the killing of
112 HR-deficient tumor cells (**Extended Data Fig. 1A)**. To investigate the function of EWSR1 in
113 DNA repair, we initially generated an *EWSR1* knock-out in RPE *TP53-/-* cells (**Fig. 1A**). *EWSR1*
114 knockout resulted in hypersensitivity to ionizing irradiation, topoisomerase inhibitors, and ATR
115 inhibitors, but not to the PARP inhibitor, suggesting a loss of MMEJ activity (**Extended Data Fig.**
116 **1 B-F**). The DDR defect observed in *EWSR1* knockout cells was similar to the defect observed in
117 *POLQ* knockout cells [11]. Consistently, *EWSR1* is mapped closely to *POLQ* in the previous repair-
118 seq, indicating a functional relationship between these genes in DNA damage repair [19].

119 Next, the impact of *EWSR1* loss on the activity of DNA double-strand break repair
120 pathways in U2OS cells was evaluated, using established Isce-I-based reporter systems [20,21].
121 *EWSR1* knockout mildly decreased HR activity (**Extended Data Fig. 1G**); however, the *EWSR1*-
122 deficient cells were not sensitive to the PARP inhibitor, Olaparib. Interestingly, *EWSR1* knockout
123 significantly decreased MMEJ activity (**Fig. 1B**). To further confirm a MMEJ reduction upon
124 *EWSR1* depletion, a novel Cas9-based MMEJ reporter system was generated (**Extended Data Fig.**
125 **1H**). Again, *EWSR1* knockout decreased MMEJ activity in this Cas9-based reporter cells (**Fig.**
126 **1C**). Consistent with this result, *EWSR1*-deficient cells were resistant to the *POLQ* inhibitor,
127 Novobiocin [22](**Extended Data Fig. 1I**).

128 Recent studies suggest that *EWSR1* knockdown is synthetic lethal with BRCA2 loss [13,14].
129 To validate these results, depletion of EWSR1 in BRCA2-proficient and BRCA2-deficient DLD1
130 cells was evaluated. Like *POLQ* loss, EWSR1 loss profoundly decreased colony-forming activity
131 in the *BRCA2*-deficient DLD1 cells (**Fig. 1D**). EWSR1 loss also efficiently decreased clonogenic
132 activity of the *BRCA1*-mutated breast cancer cell lines, MDA-MB-468 and HCC1937 cells
133 (**Extended Data Fig1. J, K**). To further confirm the synthetic lethality of EWSR1 and BRCA1/2,
134 we generated RPE *TP53-/-* cells stably expressing Cas9 (**Extended Data Fig. 1L**). These cells
135 were subsequently lentivirally-transduced with a sgRNA targeting control, *POLQ,* or *EWSR1* (co-
136 expressing GFP) together with another sgRNA targeting control, *BRCA1* or *BRCA2* (co-expressing
137 tRFP657), and changes in the frequency of GFP/tRFP657 cells were monitored. Again, like *POLQ*
138 loss, *EWSR1* loss resulted in synthetic lethality with *BRCA1/2* loss in this system. Taken together,
139 *EWSR1* functions as a MMEJ gene.
140
141 **EWSR1 promotes MMEJ by maintaining POLΘ expression.**
142 Since EWSR1 loss caused a reduction of MMEJ activity, we examined the impact of
143 EWSR1 loss on protein expression of known MMEJ-related proteins. To monitor the POLΘ
144 protein expression, RPE *TP53-/-* cells expressing 3xFLAG-tagged *POLQ* were generated.
145 Interestingly, *EWSR1* knockout resulted in decreased POLΘ expression (**Fig. 1E, Extended Data**
146 **Fig. 2A**). Increased radiosensitivity of the *EWSR1* knockout cells was observed (**Extended Data**
147 **Fig. 2B**). To further evaluate the direct and acute effect of EWSR1 loss on POLΘ expression, we
148 tagged endogenous EWSR1 protein with the mutant cytosolic prolyl isomerase $FKBP12^{F36V}$ in
149 endogenously 3xFLAG-tagged *POLQ* expressing HEK293T cells [23,24]. In this targeted protein
150 degradation system, the small bifunctional molecule $dTAG^{v1}$ bridges $FKBP12^{F36V}$ with the E3

151 ubiquitin ligase VHL, resulting in substrate ubiquitination and proteasomal degradation. Treatment
152 of cells harboring the FKBP12$^{F36V}$-tagged EWSR1 with dTAG$^{v1}$ resulted in rapid degradation of
153 EWSR1 protein (**Extended Data Fig. 2C-F**). Interestingly, this acute depletion of EWSR1 rapidly
154 reduced the expression of POLΘ but did not affect the expression of other known MMEJ-related
155 proteins, including APEX2, HMCES, FEN1, LIG3, PARP1, and XRCC1 nor HR-related proteins
156 such as BRCA1 and BRCA2 (**Fig 1F, Extended Data Fig. 2G).** Since EWSR1, as well as FUS
157 and TAF15, belongs to the FET family of proteins [25], we speculated that POLΘ loss and MMEJ
158 deficiency might be a common phenotype of the loss of FET family genes. However, FUS or
159 TAF15 loss did not affect the protein level of POLΘ or MMEJ activity (**Extended Data 2H, I**).
160 Consistently, *EWSR1* knockout reduced the POLΘ foci assembly upon irradiation-induced DNA
161 damage (**Fig. 1G**).

162 The functional significance of the loss of POLΘ on MMEJ activity in *EWSR1*-depleted
163 cells was next examined. Transduction with the WT *POLQ* cDNA rescued MMEJ activity in
164 *EWSR1* knockout cells (**Fig 1H**). Moreover, *POLQ* cDNA transduction corrected the
165 hypersensitivity to the topoisomerase inhibitors, indicating that loss of POLΘ protein is
166 responsible for these phenotypes in *EWSR1* knockout cells (**Fig 1I, Extended Data Fig 2J**). Like
167 POLΘ loss, EWSR1 loss increased unrepaired DNA damage and micronuclei in BRCA-deficient
168 cells (**Extended Data Fig 2K, L**). Collectively, EWSR1 promotes MMEJ via the maintenance of
169 POLΘ protein expression.

170

171 **EWSR1 directly binds to the *POLQ* transcript and prevents *POLQ* exon skipping.**

172

173 To investigate the cause of the reduced POLΘ protein expression following loss of
174 EWSR1, we evaluated the stability of POLΘ protein in the presence or absence of EWSR1. U2OS

175 MMEJ-reporter cells were transduced with the *POLQ* cDNA, followed by *EWSR1* knockout by
176 the CRISPR-Cas9 system. However, the protein level of POLΘ expressed by *POLQ* cDNA did
177 not change after *EWSR1* depletion (**Extended Data Fig. 3A**). Furthermore, a proteasome inhibitor
178 MG132, or the autophagy inhibitor chloroquine, did not rescue endogenous POLΘ expression in
179 *EWSR1*-depleted cell (**Extended Data Fig. 3B**), suggesting a pre-translational mechanism.

180 To assess the overall influence of EWSR1 loss on transcript levels, including the POLQ
181 transcripts, we next performed an RNA-seq analysis on RPE1 *TP53-/-* cells and RPE1 *TP53-/-*
182 *EWSR1* knockout cells. Alternative splicing events were analyzed by the MISO framework [26] as
183 well as the rMATs framework[27]. Based on this analysis, various alternative splicing events,
184 including exon skipping events, were observed in *EWSR1* knock-out cells (**Extended Data Fig.**
185 **3C-G)**. Among DNA repair genes with exon skipping events, we identified several exon skipping
186 events of the *POLQ* transcript, including exon 7, exon 21 and exon 25, as top hits (**Fig. 2A**).
187 Among these exon skipping events within the *POLQ* transcript, the skipping of exon 25 was the
188 most prominent (**Fig. 2B)**. Of note, *POLQ* exon 25 is located between long introns, intron 24 and
189 intron 25. Furthermore, we re-analyzed previously published RNA-seq datasets using si-control
190 and si-EWSR1 cells, and again the *POLQ* exon 25 skipping is one of the most significant events
191 (**Extended Data Fig. 3H, I**). We validated the results in multiple *EWSR1* knockout clones by RT-
192 PCR (**Extended Data Fig. 3J**).

193 To evaluate the impact of exon-skipping events on POLΘ protein expression, we
194 transduced sgRNAs targeting individual exons. Targeting *POLQ* exon 7, exon 21, or exon 25 by
195 CRISPR-Cas9-mediated gene knockout induced profound protein loss of POLΘ as well as a
196 reduction of MMEJ activity (**Fig.2C, D**). Moreover, a cycloheximide chase assay revealed that a
197 shorter isoform of the POLΘ protein, expressed from a *POLQ* cDNA lacking exon 25, is more

198 unstable than the full-length POLΘ protein (**Fig. 2E**). Exon 25 skipping causes premature
199 termination of POLΘ protein synthesis, preventing the complete expression of the "Fingers" and
200 "Palm" of the encoded polymerase. Structural modeling by AlphaFold 3 demonstrated that this
201 truncation disrupts the structural integrity of this domain, likely explaining the instability
202 compared to the full-length protein (**Extended Data Fig. 3K**). The *POLQ* cDNA lacking exon 25
203 or exon7 failed to rescue the MMEJ-deficiencies in *POLQ* knockout cells (**Fig. 2F**).

204 To further investigate the biological significance of *POLQ* exon 25 skipping in the MMEJ
205 activity, we designed antisense oligonucleotides (ASOs) that target exon 25 and induce exon 25
206 skipping (**Extended Data Fig. 4A**). *POLQ* exon 25 skipping decreased POLΘ protein expression
207 and MMEJ activity (**Extended Data Fig. 4B, C**). Moreover, a mNeongreen-based minigene
208 reporter, as previously described [28], was generated in which the mNeongreen is split into two exons
209 with synthetic genes derived from the partial intron 24/25 and exon 25 of *POLQ* inserted in
210 between (**Extended Data Fig. 4D**). The minigene reporter recapitulated the *POLQ* exon 25
211 skipping driven by depletion of EWSR1, indicating a direct involvement of EWSR1 in the
212 reduction of *POLQ* exon 25 skipping by binding the *POLQ* pre-mRNA near the exon 25 (**Fig. 2G,**
213 **Extended Data Fig. 4E**). Using an RNA immunoprecipitation protocol [29], we determined whether
214 EWSR1 interacts with the *POLQ* pre-mRNA, using HEK293T cells harboring the FKBP12$^{F36V}$-
215 tagged EWSR1 (**Extended Data Fig. 4F**). Indeed, EWSR1 efficiently co-immunoprecipitated
216 with the *POLQ* pre-mRNA near exon 7 and exon 25, as compared to the *GAPDH* control
217 (**Extended Data Fig. 4G, H**). Taken together, EWSR1 binds directly to *POLQ* pre-mRNA and
218 prevents exon skipping of the *POLQ* transcript, leading to maintenance of POLΘ protein levels
219 and MMEJ activity.

220 Given that one allele of *EWSR1* is lost in EwS cells, we speculated that these phenomena
221 would also be observed in EwS cells, including patient samples. A higher level of exon 7 and
222 exon 25 skipping and decreased expression of POLΘ protein were evident in EwS cell lines,
223 compared to the osteosarcoma cell line U2OS (**Fig. 2H, I**). Importantly, patient-derived EwS
224 tumor samples exhibited a higher frequency of *POLQ* exon 25 skipping than primary tumors from
225 other pediatric solid tumors, including Neuroblastoma, Osteosarcoma, and Wilms tumor (**Fig. 2J**).
226 Moreover, *POLQ* mRNA levels in Ewing sarcoma were comparable to or lower than other
227 pediatric solid tumors (**Fig. 2K**).
228 We also evaluated single nucleotide variants (SNVs), the mutational signature, and small
229 deletions in patient-derived EwS tumors. Consistent with previous studies [30], the EwS tumors have
230 a low SNV frequency compared to Neuroblastoma, consistent with a MMEJ deficiency in the
231 tumors (**Fig. 2L**). Although the level of HRD hallmarks, ID-6 and signature 3, were overall low,
232 which is consistent with a previous study [31], Ewing sarcoma tumors showed an even lower level
233 of signature 3 or ID-6, compared to Neuroblastomas, again confirming that EwS is an HR-
234 proficient tumor **(Extended Data Fig. 5A-F**). In contrast, the indels observed in the EwS samples,
235 compared with neuroblastoma samples, have a lower frequency of deletions with >3bp micro-
236 homologies and templated insertions, consistent with a low level of *POLQ*-mediated MMEJ in
237 EwS [32] **(Fig. 2M, Extended Data Fig. 5G-I**). To validate our findings in the primary patient
238 samples, we measured the degree of microhomology at breakpoints of the *EWSR1* fusion in EwS
239 tumor samples from two different EwS patient cohorts. Interestingly, the degree of microhomology
240 at canonical translocation breakpoints was very low in the EwS genomes, consistent with a cellular
241 defect in the MMEJ pathway and a compensatory increase in NHEJ (**Extended Data Fig. 5J, K**).

242 Collectively, EwS tumor cells have a bona fide defect in MMEJ, which is probably due to increased
243 *POLQ* exon 25 skipping.
244
245 **The Prion-like domain of EWSR1 is required for the faithful splicing of the *POLQ***
246 **transcript.**
247 To identify the functional domain of EWSR1 required for accurate *POLQ* splicing, we next
248 generated EWSR1 mutants lacking individual domains (**Fig. 3A**). Deletion of the prion-like
249 domain (M1) or the C-terminus RRG2 domain (M6) resulted in failure to rescue the *POLQ* exon
250 skipping and POLΘ protein expression level (**Fig. 3B, C, Extended Data Fig. 6A**). Consistent
251 with these results, MMEJ activity was not restored in *EWSR1*-depleted cells by these two mutant
252 forms of EWSR1 (**Extended Data Fig. 6B**). In addition, these mutant proteins were unable to
253 rescue the cellular hypersensitivity to topoisomerase inhibitors or an ATR inhibitor. The mutants
254 also failed to rescue the synthetic lethality resulting from BRCA2 loss in *EWSR1*-deficient cells
255 (**Extended Data Fig. 6C-G**).
256 A nuclear localization signal (NLS) is located at the RRG2 domain, and its loss is likely to
257 destabilize the EWSR1 protein. The NLS mutant protein failed to rescue *POLQ* exon skipping and
258 failed to restore POLΘ protein expression. To further evaluate the prion-like domain, the domain
259 was subdivided into five parts, and mutant proteins lacking the individual regions were generated
260 (**Extended Data Fig. 6H**). Interestingly, all mutant forms were able to rescue the POLΘ protein
261 level, albeit with some minor differences, suggesting that the domains have functional
262 redundancies (**Extended Data Fig. 6I)**. Since the prion-like domain of EWSR1 is critical for the
263 formation of condensates of EWSR1 [33], we infer that the formation of EWSR1 condensates may
264 play a role in maintaining the faithful splicing of POLΘ. To test this hypothesis, we generated

265 multiple mutant forms of EWSR1 in which tyrosine residues of the prion-like domain were
266 substituted with serine residues, as previously described[33]. Strikingly, an EWSR1-7YS mutant, in
267 which tyrosine residues were substituted by serine residues, failed to rescue the *POLQ* exon
268 skipping and POLΘ protein expression level in *EWSR1*-depleted cells (**Fig. 3D, E, Extended Data**
269 **Fig. 6J, K**). Furthermore, an EWSR1-37YS mutant, in which all thirty-seven tyrosine residues in
270 the prion-like domain were replaced with serine residues, and known to be defective in condensate
271 formation, abolished the ability of EWSR1 to prevent exon skipping of the *POLQ* transcript,
272 further suggesting that condensate formation is critical for faithful splicing. Consistent with these
273 findings, the EWSR1-7YS and EWSR1-37YS mutants failed to rescue MMEJ activity in EWSR1-
274 depleted cells (**Fig. 3F**). Taken together, the prion-like domain of EWSR1 is required for the
275 faithful splicing of the *POLQ* transcript and for the expression of the POLΘ protein, suggesting
276 that condensate formation of EWSR1 is also required for these activities as previously described
277 [5,33].
278
279
280 **The prion-like domain of EWSR1 is critical for interaction with spliceosomes including**
281 **FUBP1 and KHSRP.**
282 The condensate-defective EWSR1 mutant may have lost the ability to bind the *POLQ*-pre
283 mRNA. However, RNA-immunoprecipitation revealed that the EWSR1-7YS and EWSR1-37YS
284 mutants can still bind to the *POLQ* pre-mRNA (**Extended Data Fig 4. F-I**). Alternatively, the
285 condensate-defective EWSR1 mutant may be competent for *POLQ* mRNA binding but defective
286 in its interaction with RNA-binding proteins in spliceosomes. To test this hypothesis, we
287 performed an IP/mass spectrometry analysis using HEK293T cells with a biallelic FKBP12-
288 V36F knock-in at the N-terminus of *EWSR1* (**Fig. 3G**). The cDNAs encoding the control V5-

289 tagged *EWSR1*-WT, or the *EWSR1* 37YS mutant, were transduced into the cells and, following
290 dTAG$^{v1}$ treatment, cellular proteins were immunoprecipitated with the anti-V5 antibody. A
291 previous study demonstrated that the prion-like domain of EWS-FLI1 interacts with BAF
292 complex subunit components in condensates [5]. Accordingly, wildtype EWSR1, but not the
293 EWSR1 37YS mutant protein, preferentially bound to BAF subunits, including ARID1A,
294 SMARCA2, SMARCA4, and SMARCC1 (**Fig. 3G, I, Extended Data Fig. 7A**). In addition to
295 these BAF subunits, additional RNA binding proteins were identified as binding partners of
296 wild-type EWSR1. Intriguingly, the interactors contain prion-like domains, further indicating a
297 functional interaction of EWSR1 with these proteins in spliceosomes (**Fig. 3H**).
298 Since exon skipping in *EWSR1* KO cells is typically observed for exons that are flanked
299 by longer introns (**Extended Data Fig. 7B**), we focused on FUBP1, among the prion-like
300 domain-dependent EWSR1 binding partners. FUBP1 is a spliceosome known to mediate the
301 splicing of long introns [34] (**Fig. 3H, I**). In addition, KHSRP/FUBP2, known to have a similar
302 function in splicing, was further evaluated. Depletion of FUBP1 or KHSRP alone cause only a
303 minor reduction in POLΘ protein expression or MMEJ activity. However, dual depletion of
304 FUBP1 and KHSRP increased *POLQ* exon25 skipping, reduced the POLΘ protein level, and
305 reduced the cellular MMEJ activity, indicating a functional redundancy between FUBP1 and
306 KHSRP (**Fig. 3 J, K, Extended Data Fig. 7C).** A FUBP1-16YS mutant, with tyrosine residues
307 of its C-terminus region substituted with serine residues, failed to bind to wildtype EWSR1 and
308 failed to rescue the phenotype of FUBP1/KHSRP DKO cells, suggesting that tyrosine residues of
309 EWSR1 and FUBP1 are critical for the protein-protein interaction and for *POLQ* exon25
310 inclusion (**Fig. 3 L-Q, Extended Data Fig. 7D)**. Similarly, a KHSRP-17YS mutant failed to
311 rescue the phenotype of FUBP1/KHSRP DKO cells **(Extended Data Fig. 7E-H)**. Taken

312 together, the protein-protein interaction between EWSR1 and FUBP1 or KHSRP via the prion-
313 like domain is critical for the faithful splicing of the *POLQ* pre-mRNA.
314
315
316 **The EWS-FLI1 oncoprotein inhibits formation of POLQ mitotic foci at DSB loci and**
317 **disrupts MMEJ.**
318 The chromosomal translocation of 11;22 leads not only to the loss of one wildtype EWSR1
319 allele but also to the generation of an oncogenic fusion protein EWS-FLI1 in EwS cells (**Extended**
320 **Data Fig. 8A**). Moreover, EWS-FLI1 is known to heterodimerize with EWSR1 and to disrupt its
321 activity [5,35,36]. We therefore hypothesized that this oncoprotein may directly or indirectly interfere
322 with MMEJ activity. To test this hypothesis, the EWS-FLI1 oncoprotein was lentivirally expressed
323 in U2OS-based MMEJ reporter cells and evaluated for MMEJ activity. Interestingly, EWS-FLI1
324 diminished the MMEJ activity in U2OS MMEJ reporter cells (**Fig.4A, Extended Data Fig. 8B**).
325 Consistent with these results, EwS cell lines exhibited resistance to POLΘ inhibitors, Novobiocin
326 and ART558 [37], compared with osteosarcoma cells U2OS (**Extended Data Fig. 8C, D)**. The
327 influence of various mutant forms of the EWS-FLI1 fusion oncoprotein on MMEJ activity was
328 tested (**Fig. 4B**). Expression of the SYGQ1-FLI1, SYGQ2-FLI1, ΔSYGQ1, or ΔSYGQ2 mutant
329 by cDNA-mediated transfection decreased the MMEJ activity in the U2OS reporter cells,
330 indicating functional redundancies in the prion-like domain of EWS-FLI1. Interestingly, the
331 EWS-FLI1 37YS mutant, which failed to disrupt MMEJ activity (**Fig. 4B, Extended Data Fig.**
332 **8B**), is defective in condensate formation and in binding to wild-type EWSR1 protein [5,36].
333 Moreover, expression of EWS-FLI1, but not the EWS-FLI1 37YS mutant, conferred
334 hypersensitivity to topoisomerase inhibitors (**Extended Data Fig. 8E, F**).

335 To determine the effect of the EWS-FLI1 oncoprotein on MMEJ activity in Ewing sarcoma
336 cells, we transduced an Isce-I-based MMEJ reporter into the A673, RH1, and RD-ES EwS lines.
337 As expected, depletion of EWS-FLI1 from these cells by sgFLI1, as confirmed by western blot,
338 restored MMEJ activity (**Fig. 4C, Extended Data Fig. 8G**). These EwS lines, with EWS-FLI1
339 depletion, exhibited resistance to the POLΘ inhibitor, ART558 (**Extended Data Fig. 8H),**
340 consistent with their restoration of MMEJ following EWS-FLI1 depletion. The minigene reporter
341 assay showed that expression of EWS-FLI1 but not the EWS-FLI1 37YS mutant increased the
342 *POLQ* exon 25 skipping, whereas depletion of EWS-FLI1 in EwS cells ameliorated the *POLQ*
343 exon 25 skipping (**Fig. 4D, E)**. Consistent with a previous study [5], EWS-FLI1 preferentially bound
344 to EWSR1 compared with the EWS-FLI1 37YS mutant, suggesting an inhibitory effect of EWS-
345 FLI1 on EWSR1-mediated splicing (**Fig. 4F**). Given that R-loop accumulation driven by the EWS-
346 FLI1 expression inhibits the recruitment of BRCA1 to the DSB loci, we asked whether R-loop
347 accumulation may be involved in the MMEJ reduction in Ewing sarcoma. Expression of wildtype
348 RNaseH1 or mutant RNaseH1 did not affect EWS-FLI1-mediated MMEJ reduction (**Extended**
349 **Data Fig. 8I, J**).
350 Recent studies indicate that DSB repair by POLΘ and MMEJ occurs during mitosis of the
351 cell cycle [38]. POLΘ is phosphorylated by PLK1 and assembles with TOPBP1, along with Rhino
352 and the 9-1-1 complex, into POLΘ foci [18,39]. To assess the influence of EWS-FLI1 expression on
353 the assembly of DSB loci, we engineered endogenously V5-tagged POLΘ expressing MMEJ
354 reporter U2OS cells by CRISPR/Cas9-mediated knock-in. After generation of adenoviral I-SceI-
355 induced DSBs, we performed chromatin-immunoprecipitation (ChIP) with an anti-V5 antibody,
356 followed by qPCR analysis. Overexpression of EWS-FLI1 in these cells disrupted the assembly
357 of POLΘ with sites of DNA double-strand breaks **(Fig. 4G).** Consistent with this result, MMEJ

358 activity and the assembly of POLΘ foci assembly, as measured by immunofluorescence, was
359 decreased by the overexpression of EWS-FLI1 (**Fig. 4H**). Moreover, EWS-FLI1 overexpression
360 in these U2OS cells also increased γH2AX foci, consistent with the loss of MMEJ activity in
361 mitosis. EWS-FLI1 overexpression can therefore reduce POLΘ-mediated DSB repair in mitosis,
362 at a time when POLΘ-mediated MMEJ is highly active [18,39]. EWS-FLI1 was also depleted in
363 endogenously V5-tagged POLΘ expressing A673 cells. As predicted, the loss of EWS-FLI1
364 restored POLΘ foci assembly (**Fig. 4I).** Collectively, the EWS-FLI1 oncoprotein can reduce the
365 level of POLΘ mitotic foci, leading to a functional impairment of MMEJ in EwS cells.

366
367
368 **Inhibition of NHEJ or HR efficiently kills the MMEJ-deficient Ewing sarcoma cells via**
369 **synthetic lethality.**

370 Since inhibiting two of three DSB repair pathways is synthetically lethal in cells [40,41], we
371 asked whether inhibition of NHEJ or HR would provide a viable therapeutic modality in MMEJ-
372 deficient Ewing sarcoma cells. Previous studies have shown that inhibitors of NHEJ are
373 synthetic lethal with MMEJ loss [40,41]. Consistent with our study, EwS cells are sensitive to DNA-
374 PK inhibitors, known to inhibit NHEJ and enhance the activity of cytotoxic agents in these cells
375 [42]. We confirmed that EwS cell lines are sensitive to the DNA-PKci (**Fig. 5A**) and have an
376 elevated baseline level of activated pDNA-PKc (**Fig. 5B**). EwS cells exhibited high phospho-
377 RPA, a replication stress marker, and were hypersensitive to the topoisomerase inhibitors,
378 topotecan, and etoposide (**Extended Data Fig. 8A, 9A, B**). Since etoposide is used in dose-
379 intensive chemotherapy of Ewing sarcoma, we asked whether inhibition of NHEJ could increase
380 the therapeutic efficacy of etoposide in EWS cells. Strikingly, the DNA-PK inhibitor, Peposertib,
381 enhanced the activity of etoposide in two EWS cell lines (**Fig. 5C, D)**. Moreover, the EwS cell

382 line A673, grown as a xenograft, was also sensitive to peposertib, *in vivo*, and the drug enhanced
383 the activity of etoposide (**Fig. 5E, F, Extended Data Fig. 9C, D**). Consistent with the enhanced
384 sensitivity to peposertib, immunohistochemistry analysis of the EwS primary patient samples
385 revealed an elevated baseline level of pDNA-PKc compared with Osteosarcoma patient samples
386 (**Fig. 5G**).
387 On the other hand, EWS-FLI1- but not EWS-FLI1-YS mutant-expressing cells are more
388 dependent on BRCA2, confirming that the HR pathway is also a rational target in MMEJ-
389 deficient EwS cells (**Fig. 5H**). Previous studies have also suggested that EwS cells are sensitive
390 to HR inhibitors, perhaps through a mechanism of synthetic lethality. CDK12 is a known
391 activator of HR repair, based on its ability to promote the transcription of long HR genes [43]. A
392 targeted inhibitor of CDK12 is known to enhance the killing of EwS cells [44]. Consistent with this
393 report, EwS cell lines are sensitive to the CDK12 inhibitors THZ531 and MFH290 (**Fig 5 I, J**).
394 Depletion of the EWS-FLI1 oncoprotein in the EwS cells resulted in a restoration of MMEJ
395 activity and acquisition of CDK12i resistance (**Fig 5K, L, Extended Data Fig. 9E, F**). Taken
396 together, the killing of the EwS cells by this HR inhibitor results from synthetic lethality between
397 the HR pathway and the MMEJ pathway.
398
399
400

401 **DISCUSSION**

402 Despite the high sensitivity to DNA-damaging agents, EwS cells, expressing the EWS-
403 FLI1 fusion oncoprotein, have a remarkably low somatic mutation burden [30,45,46]. This
404 observation suggested that EwS cells are deficient in an error-prone DNA repair mechanism,
405 such as MMEJ. Indeed, CRISPR screens confirmed that the knockout of the EWSR1 gene
406 functionally resembles the knockout of known MMEJ genes, such as *POLQ*, *XRCC1*, *APEX2*,
407 and *RHNO1*, suggesting that EWSR1 protein cooperates with these other encoded proteins in the
408 MMEJ pathway. Moreover, EWSR1-deficient cancer cell lines and EwS tumors have a low ID6
409 mutational signature, indicating a low level of microhomology-mediated repair of double-strand
410 breaks. Since EWSR1 has a known function as a mRNA splicing factor [35,47-49], we hypothesized
411 that EWSR1 may regulate MMEJ indirectly through its splicing of the mRNA encoding a MMEJ
412 protein. EWSR1 is indeed required for the accurate splicing of the long mRNA encoding the
413 critical MMEJ polymerase, POLΘ, although it is not required for splicing of mRNAs encoding
414 other MMEJ proteins. Disruption of EWSR1 activity in EwS cells results in skipping of multiple
415 exons in the *POLQ* pre-mRNA, especially exon 25, leading to decreased POLΘ expression in
416 EwS cells and to a strong MMEJ defect.

417 In contrast to previous reports, our study indicates that EwS tumors are competent for
418 high-fidelity DNA repair pathways, such as the homologous recombination pathway, and the
419 tumors are PARP inhibitor-resistant. Other recent studies indicate that EwS cells are dependent
420 on the HR genes, BRCA1, BARD1, and PALB2 [50,51] for survival. Previous studies have also
421 demonstrated that HR deficiency is synthetic lethal in cells with a MMEJ deficiency [22,52,53].
422 Accordingly, the strong MMEJ defect in EwS cells could result in an increased reliance on HR
423 repair.

424 The chromosome 11/22 translocation in EwS cells and the generation of the EWS-FLI1
425 fusion oncoprotein appears to account, at least in part, for the cellular defect in EWSR1 splicing
426 activity. First, since one allele of EWSR1 is disrupted by the chromosome translocation, the
427 EwS cells already have a 50% reduction in EWSR1 protein expression. Second, the EWS-FLI1
428 oncoprotein heterodimerizes with EWSR1 and exerts a dominant negative effect on EWSR1
429 function [5]. Our studies indicate that loss of EWSR1 function can account for the decreased
430 expression of POLΘ protein, the elevated exon 25 skipping of the *POLQ* pre-mRNA, and the
431 defective MMEJ activity in EwS cells. Enhanced expression of the EWSR1 protein, or acute
432 knockdown of the EWS-FLI1 oncoprotein, can reverse these defects and restore MMEJ activity,
433 thereby reducing the cellular reliance on other DSB repair pathways, such as HR and NHEJ.

434 Splicing of specific mRNAs is regulated, at least in part, by RNA Binding Proteins
435 (RBPs) which function in membrane-less condensates. Condensates known as nuclear speckles
436 are required for splicing activity[54]. Many RNPs, including the EWSR1 protein, contain a prion-
437 like domain (PrLD), allowing the proteins to assemble with other RBPs contained in
438 condensates. We reasoned that EWSR1 might interact with other RBPs and function
439 cooperatively as a spliceosome for the *POLQ* pre-mRNA. Moreover, disruption of the PrLD of
440 EWSR1, through the generation of multiple Tyrosine to Serine (YS) mutations, disrupts the
441 assembly of functional spliceosomes. Indeed, the EWSR 37YS mutant, known to be defective in
442 condensate assembly [5], also fails to correct exon 25 skipping of the *POLQ* mRNA, and fails to
443 restore POLΘ protein expression. IP/Mass Spectrometry identified specific RBPs that interact
444 with EWSR1. Interestingly, multiple RBPs containing prion-like domains were identified in the
445 immune complexes, and these proteins did not coimmunoprecipitate with the EWSR1 37YS
446 mutant protein. Components of SWI/SNF complexes, including subunits of the cBAF complex

447 with prion-like domains, were also found to interact with EWSR1 and EWS-FLI1, consistent
448 with a previous report [5]. Unlike EWSR1, other FET proteins containing PrLD domains, such as
449 FUS and TAF15, were not required for accurate *POLQ* splicing, suggesting that EWSR1 forms
450 distinct RBP complexes, perhaps functioning in distinct condensates. Taken together, EWSR1
451 may act as a "hub" for spliceosomes, using its specific prion-like domain for binding to a distinct
452 set of RBPs and for splicing of the *POLQ* mRNA and probably other mRNAs. Expression of
453 EWS-FLI1 disrupts this splicing mechanism.

454 Some RNA binding proteins, including FUBP1 and KHSRP/FUBP2, specifically interact
455 with wildtype EWSR1, but not with the EWSR1 37YS mutant. FUBP1 and KHSRP have high
456 similarity in their protein sequence and exhibit functional redundancy in their splicing activity [55].
457 FUBP1 and KHSRP are multifunctional RNA binding proteins and are known to redundantly
458 regulate the splicing of the SMN2 mRNA [55]. A carboxy-terminal, tyrosine-rich region is present
459 in both FUBP1 and KHSRP/FUBP2 [56]. However, the biological functions of the tyrosine-rich
460 regions of FUBP1 and KHSRP have remained unclear. Since the tyrosine residues of the prion-
461 like domain are critical for liquid-liquid phase separation[57,58] and condensate formation, a YS
462 mutant of FUBP1 or KHSRP, may fail to form condensates with EWSR1 and may lead to
463 aberrant splicing events, such as *POLQ* exon skipping. FUBP1 is crucial for splicing long
464 introns [34] such as the long intron 24/25 in the *POLQ* pre-mRNA. Therefore, a condensate
465 formed by EWSR1-FUBP1/KHSRP may be required for the accurate splicing of exons trapped
466 between long introns. Interestingly, the dual knockout of FUBP1 and KHSRP, like the knockout
467 of EWSR1, results in exon 25 skipping of the *POLQ* mRNA and a reduction in POLΘ protein
468 expression and MMEJ. Re-expression of the wildtype FUBP1, but not of the FUBP1-YS
469 mutant, rescued these defects in the dual knockout cells. The RBPs, EWSR1, FUBP1, and

470 KHSRP therefore appear to functionally interact via their PrLD domains and to cooperate in a
471 EWSR1-initiated splicing pathway for the *POLQ* mRNA (**Extended Data Fig. 9G**).

472 Given that overexpression of EWS-FLI1 oncoprotein induces extensive DNA damage[59],
473 we reasoned that the EWS-FLI1 fusion oncoprotein could impair MMEJ activity in EwS cells.
474 Indeed, the acute degradation of the EWS-FLI1 oncoprotein rapidly rescued MMEJ, suggesting
475 an acute upregulation of POLΘ expression and POLΘ mitotic foci. While the EWS-FLI1
476 protein forms condensate with EWSR1, it disrupts the function of EWSR1, reduces POLΘ
477 expression, and inhibits MMEJ. The EWS-FLI1-37YS mutant fails to form condensates with
478 wild-type EWSR and fails to reduce MMEJ. While knockout of EWS-FLI1 restores MMEJ, loss
479 of the oncoprotein is also detrimental to EwS cell growth and results in a reduction in the
480 transcription of EWS-FLI1 target genes which contributes to EwS cell growth.

481 Previous studies on the pathogenesis of Ewing Sarcoma have focused primarily on the
482 transcriptional activity of the EWS-FLI1 oncoprotein, driven by the carboxy-terminal FLI1
483 transcriptional domain. Indeed, this oncoprotein upregulates the transcription of genes which in
484 turn promote R-loops and the neoplastic growth of the EwS tumor. Our study also provides new
485 insights into the pathogenesis of the cancer. While the chromosome 11/22 translocation results
486 in the formation of the EWS-FLI1 oncoprotein, it also simultaneously inactivates the EWSR1
487 *tumor suppressor* function. One allele of this EWSR1 tumor suppressor gene is inactivated by
488 the translocation itself, and the remaining EWSR1 suppressor function is inactivated by the
489 binding of the fusion oncoprotein to the remaining EWSR1 protein. Due to the loss of the
490 splicing activity of EWSR1, the tumor cells exhibit the characteristic MMEJ defect- namely,
491 radiation sensitivity, a defect in microhomology-mediated DSB repair, and a hypo-mutable
492 phenotype. The splicing defect may also result in the inactivation of other tumor suppressor

493 genes and a more direct role in the process of oncogenesis. For instance, the EWS-FLI1 fusion
494 protein may have additional direct oncogenic functions beyond its ability to promote
495 transcription, R-loops, and replication stress. Specifically, the EWS-FLI1 fusion can disrupt
496 PLK1-activated POLΘ mitotic foci [39] which are known sites of TOPBP1-mediated DSB repair.
497 Accordingly, the EWS-FLI1 oncoprotein may further block the repair of DSBs in mitosis,
498 promote the formation of micronuclei, and enhance genomic instability.

499 Taken together, *Ewing sarcoma is the first known cancer with loss of MMEJ activity.* The
500 impairment of the MMEJ pathway in EwS prompted us to target NHEJ and HR, the two DNA
501 repair pathways known to be synthetic lethal with MMEJ loss. Indeed, EwS cells are sensitive to
502 the NHEJ inhibitor, peposertib, as a monotherapy. Peposertib is also synergistic with the
503 genotoxic drug, Etoposide, in killing EwS cells. Accordingly, the use of this drug combination
504 could reduce the need for high doses of Etoposide for EWS patients and thereby reduce its long-
505 term toxicity, such as therapy-related leukemia[60]. EWS cells, due to their MMEJ defect, are also
506 sensitive to the HR inhibitor, THZ531, a CDK12 inhibitor. This inhibitor is one of the first
507 known inhibitors of the HR pathway. Exon 25 skipping of the *POLQ* mRNA in fresh tumor
508 samples may provide a biomarker for identifying tumors with an underlying MMEJ defect which
509 may in turn respond to targeted NHEJ inhibitors or HR inhibitors through the mechanism of
510 synthetic lethality.

511

512 **ACKNOWLEDGMENTS**

513 We thank members of the D'Andrea laboratory for their helpful suggestions and comments. We
514 thank Harvard Medical School and the Dana-Farber/Harvard Cancer Center in Boston, MA, for
515 the use of the Animal, Flowcytometry, and Taplin Mass Spectrometry core facilities. This work
516 was supported by U.S. National Institutes of Health grants R01HL052725 and R01CA296618
517 (A.D.D), the Breast Cancer Research Foundation (A.D.D.), the Ludwig Center at Harvard
518 (A.D.D.), the David Liposarcoma Research Initiative at Harvard (A.D.D.), and the Smith Family
519 Foundation (A.D.D.), and the Alex's Lemonade Stand Foundation (S.A.). This work was also
520 supported by U.S. National Institutes of Health grants P50 CA168504 (A.D.D, G.I.S.), P50
521 CA240243 (A.D.D., G.I.S.), the SENSHIN Medical Research Foundation (S.A.), the Japan
522 Society for the Promotion of Science Overseas Research Fellowships (S.A.), NCI K08 Career
523 Development Award (R.G.), the DoD PRCRP Fellow Career Development Award (R.G.), the
524 ASCO Career Development Award (R.G.), the Rally Foundation Career Development Award
525 (R.G.) and the Boston Children's Hospital TRP Mentored Career Development Award (R.G.).
526 Graphical abstract (Extended Fig, 9G) was created using BioRender.com.

527

528 **AUTHOR CONTRIBUTIONS**

529 S.A and A.D.D. conceived the study. S.A., G.Z., J.P.A., Y.H., D.R.I., S.M. and L.J. designed and
530 performed biological experiments and analyzed the data. S.A. and Y.H designed and performed
531 animal experiments. Y.T., H.N. and R.G. performed computational analysis. N.W.A. performed
532 AlphaFold analysis. M.D.B., J.J.T and S.I.R. performed IHC analysis using patient samples.
533 K.P., E.M.V.A. and G.I.S. provided scientific inputs. S.A. and A.D.D. wrote the manuscript.

534

535 **FIGURE LEGENDS**

536 **Figure 1: *EWSR1* knock-out cells display MMEJ-deficient phenotypes through POLΘ**
537 **protein reduction.**

538 (A) Immunoblotting analysis for the detection of EWSR1 in RPE p53-/- clones edited with
539 CRISPR-Cas9.

540 (B, C) U2OS I-SceI-based (B, n=9) or Cas9-based (C, n=4) or were transduced lentivirally
541 transduced with control sgRNA or sgRNA targeting *EWSR1* or *POLQ* (co-expressing puromycin
542 resistance gene). Cells were treated with puromycin for 4 days, and then subjected to MMEJ
543 reporter analysis.

544 (D) Colony formation assay data of DLD1 *BRCA2* wild-type and *BRCA2* knockout cells
545 transduced with control, sg*POLQ* or sg*EWSR1* (n=3, left). Shown are representative images of the
546 colonies (right).

547 (E) RPE *TP53-/-* 3xFLAG-*POLQ* knock-in cells were lentivirally transduced with Cas9 and a
548 sgRNA targeting control (NT), *EWSR1* or *POLQ* (co-expressing puromycin resistance gene). After
549 puromycin selection, cells were subjected to immunoblotting analysis.

550 (F) HEK293T 3xFLAG-*POLQ* single knock-in cells and HEK293T 3xFLAG-*POLQ*/2xHA-
551 FKBP12 V36F *EWSR1* double knock-in cells are treated with DMSO or $dTAG^{V1}$ for indicated
552 time. Cell lysates were extracted and followed by immunoblotting analysis.

553 (G) Immunofluorescence analysis showing DNA damage-induced the foci formation of V5-tagged
554 endogenous POLΘ was decreased by EWSR1 depletion (scale bars = 2μm).

555 (H) U2OS MMEJ reporter cells (I-SceI based) were transduced with a control vector or POLQ
556 cDNA (co-expressing blasticidin resistance gene). After blasticidin selection, cells were
557 subsequently transduced with Cas9 and a sgRNA targeting control or EWSR1 (co-expressing

558 puromycin resistance gene). After puromycin selection, cells were subjected to MMEJ reporter
559 analysis (n=4).
560 (I) RPE *p53*-/- cells and RPE *p53*-/- *EWSR1* knockout cells were transduced with control, wildtype
561 *EWSR1* or *POLQ* cDNA. Colony formation assays were performed in the presence of Topotecan
562 (n=3).
563
564
565 **Figure 2: Disruption of MMEJ activity in Ewing sarcoma is driven by *POLQ* exon skipping.**
566 (A) Volcano plots showing alternatively spliced skipped exon (SE) in RPE *p53*-/- *EWSR1* knock-
567 out cells compared with RPE *p53*-/- cells defined by deep RNA-seq analysis.
568 (B) Sashimi plots of the alternative splicing events in control and *EWSR1* knock-out RPE *p53*-/-
569 cells: the region between exon24 and exon26 of the *POLQ* transcript.
570 (C) U2OS endogenous V5 tagged- POLΘ expressing cells were transduced with Cas9 and sgRNA
571 targeting the indicated exon of the *POLQ* gene. Cell lysates were extracted and followed by
572 immunoblotting analysis.
573 (D) U2OS MMEJ ver.2 (Cas9-based) cells were lentivirally transduced with control sgRNA or
574 sgRNA targeting different *POLQ* exons (co-expressing puromycin resistance gene). Cells were
575 treated with puromycin for 4 days, and then subjected to MMEJ reporter analysis (n=4).
576 (E) *POLQ* mRNA without exon 25 produces shorter POLΘ which is unstable. Blot showing
577 HEK293T cells transfected with FLAG-*POLQ* cDNA containing *POLQ* full-length CDS or CDS
578 lacking exon25. Cells were treated with CHX for 3–9 hr.
579 (F) HEK293T *POLQ* knockout cells were transfected with MMEJ reporter (ver.2) and sgRNA
580 targeting the reporter together with control vector or indicated *POLQ* cDNA. 48h after
581 transfection, cells were subjected to FACS analysis to measure the MMEJ activity (n=3).

582 (G) Schematic representation of the minigene reporter construct (left). K562 cells carrying the
583 minigene reporter were lentivirally transduced with Cas9 and sgRNA targeting the indicated genes.
584 Shown are relative frequencies of mNeonGreen positive fraction within mScarlet positive cells
585 (right, n=3).
586 (H) RT-PCR showing the exon 7 and exon 25 skippings of *POLQ* transcript in Ewing sarcoma cell
587 lines.
588 (I) Immunoblotting analysis showing the lower levels of POLΘ expression in Ewing sarcoma cell
589 lines compared to U2OS cells.
590 (J) rMATS analysis showing *POLQ* exon skipping events in primary samples from pediatric cancer
591 patients. n = 498 (EWS = 44, NBL = 131, OS = 88, WT = 110, RMS = 125).
592 (K) RNA-seq data from the patient sample shows the *POLQ* mRNA expression is significantly
593 lower in Ewing sarcoma patients compared with RMS/WT. n = 832 (EWS = 83, NBL = 195, OS
594 = 182, WT = 162, RMS = 210).
595 (L, M) Low somatic SNV and MMEJ deletion mutation burden in Ewing sarcoma tumors (ES; n
596 = 23) as compared to neuroblastoma tumors (NB; n = 47). Short somatic mutations were
597 determined from the whole genome sequencing data of tumors and matched normals. Each dot in
598 K represents the total number of somatic SNVs for each tumor. Each dot in L represents the
599 proportion of 3+-bp deletions with 2+-bp microhomology among all indels of each tumor.
600
601 **Figure 3: Formation of EWSR1 condensates through the prion-like domain is critical for**
602 **faithful splicing of *POLQ* transcript via interacting with FUBP1 and KHSRP/FUBP2.**
603 **(A)** Schematic representation of the individual domains of EWSR1 and the deletion mutants used
604 in Fig.3B, C.

605 (B, D) RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells were
606 retrovirally transduced with empty vector, wildtype EWSR1, or indicated EWSR1 mutants. After
607 transduction, cells were treated with blasticidin for four days. Then, cells were treated with DMSO
608 or 1μM $dTAG^{v1}$ for 24hr. After $dTAG^{v1}$ treatment, cell lysates were subjected to immunoblot
609 analysis.
610 (C, E) RT-PCR showing the *POLQ* exon 25 skipping in the indicated cells.
611 (F) U2OS MMEJ ver.2 (Cas9-based) cells were lentivirally transduced with control sgRNA or
612 sgRNA targeting *EWSR1* together with empty vector wild-type EWSR1 or the indicated EWSR1
613 mutants by retroviral transduction. Cells were treated with puromycin and blasticidin for 4days,
614 and then subjected to MMEJ reporter analysis (n=4).
615 (G) Volcano plots comparing detected protein levels by IP-MS between EWSR1 wildtype and the
616 EWSR1 37YS mutant.
617 (H) A ven diagram showing the frequency of all the genes and positive hits with a prion-like
618 domain identified by PLAAC.
619 (I) Immunoprecipitation followed by immunoblotting analysis showing wildtype EWSR1
620 preferentially bound to FUBP1, KHSRP, SMARCA4 and SMARCC1 compared to the EWSR1
621 mutants.
622 (J) FUBP1/KHSRP DKO cells exhibited POLΘ protein reduction (upper) and increased *POLQ*
623 exon 25 skipping (lower).
624 (K) MMEJ activity was profoundly decreased in FUBP1/KHSRP DKO cells measured by U2OS
625 MMEJ ver.2 cells (n=4).
626 (L) Schematic representation of wildtype and YS-mutated EWSR1 or FUBP1.

627 (M) Immunoblotting showing wildtype FUBP1 preferentially bound to EWSR1 compared to the
628 FUBP1 16YS mutant.
629 (N) U2OS MMEJ ver.2 cells were lentivirally and retrovirally transduced with indicated sgRNA(s)
630 (co-expressing puromycin resistance gene) and control or FUBP1 cDNA (co-expressing
631 blastidcidin resistance gene). Cells were treated with puromycin and blasticidin for 4days, and then
632 subjected to MMEJ reporter analysis (n=4).
633 (O) RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells were
634 lentivirally transduced with Cas9 and a sgRNA targeting control (NT) or sgRNAs targeting FUBP1
635 and KHSRP together with retrovirally transduced with a control vector, wildtype FUBP1 or the
636 FUBP1 16YS mutant. After puromycin and blasticidin selection, cells were subjected to
637 immunoblotting analysis.
638 (P) RT-PCR showing the *POLQ* exon25 skipping in the indicated cells.
639
640 **Figure 4: Expression of EWS-FLI1 perturbs POLΘ foci assembly, resulting in impairment**
641 **of MMEJ.**
642 (A) I-SceI-based MMEJ reporter cells were lentivirally transduced with an empty vector, wild-
643 type *EWS-FLI1* or the FKBP12 V36F-tagged *EWS-FLI1*. Cells were treated with 1uM dTAG$^{V1}$
644 for 18hr, and then adenovirally transduced with I-sceI. 48hr after I-sceI transduction, the
645 frequencies of GFP positive cells were assessed by FACS (n=6).
646 (B) Schematic representation of the EWS-FLI and mutant forms of EWS-FLI1 used in MMEJ
647 reporter assay (left). I-SceI-based MMEJ reporter cells were lentivirally transduced with an empty
648 vector, wild-type *EWS-FLI1* or indicated mutant forms of *EWS-FLI1* (co-expressing puromycin

649 resistance gene). After puromycin selection, cells were subjected to MMEJ reporter assay (right,
650 n=6).
651 (C) I-SceI based MMEJ reporter was lentivirally transduced with U2OS, A673, RH1 and RD-ES
652 cell lines. Cells were subsequently transduced with Cas9 together with control sgRNA or sgRNA
653 targeting FLI1. Cells were treated with puromycin for 4days, and then subjected to MMEJ reporter
654 analysis (n=3).
655 (D) EWS-FLI1 but not the EWS-FLI1 37YS mutant increased *POLQ* exon 25 skipping measured
656 by K562 minigene reporter cells (n=6).
657 (E) A673 and TC32 cells carrying the minigene reporter showing depletion of EWS-FLI1
658 decreased *POLQ* exon 25 skipping (n=3).
659 (F) Immunoprecipitation followed by immunoblotting analysis showing wildtype EWS-FLI11
660 preferentially bound to EWSR1 compared to the EWS-FLI1 37YS mutant.
661 (G) ChIP-qPCR assay showing expression of EWS-FLI1 perturbed POLΘ assembly to the DSB
662 loci induced by Isce-I (n=2).
663 (H) Immunofluorescence analysis showing endogenous V5-tagged POLΘ foci formation was
664 perturbed by EWS-FLI1 expression both asynchronous and mitotic arrested cells. Cells were
665 treated with 2Gy irradiation and then incubated for 4hr (scale bars = 2μm).
666 (I) Immunofluorescence analysis showing endogenous V5-tagged POLΘ foci formation was
667 restored by EWS-FLI1 depletion in V5-tagged-*POLQ* knock-in A673 cells. Cells were treated with
668 2Gy irradiation and then incubated for 4hr (scale bars = 2μm).
669
670 **Figure 5: Inhibition of NHEJ or HR efficiently eliminates Ewing sarcoma cells.**

671 (A) The CTG assay showing that compared with U2OS, Ewing sarcoma cells showed
672 hypersensitivity to Peposertib.
673 (B) Immunoblot analysis showing phospho-DNAPK, DNAPK and GAPDH protein level in U2OS
674 and Ewing sarcoma cell lines.
675 (C, D) The CTG assay showing that a combination of Peposertib and Etoposide showed a strong
676 synergistic effect on Ewing sarcoma cell lines A673 (C) and TC32 (D).
677 (E, F) Xenograft model using TC32 showing therapeutic efficacy of a Peposertib and Etoposide.
678 NSG mice (n = 10 mice/group) bearing luciferized TC32 cells were treated with 100 mg/kg of
679 Peposertib (orally), 8 mg/kg etoposide (via IP), or both for 2 weeks. Tumor sizes were measured
680 every 2-3 days (E) and also monitored by bioluminescence images (F).
681 (G) Immunohistochemistry analysis from the patient samples of EwS (n=8) showing higher
682 pDNA-PKc activities compared with that of OS (n=9).
683 (H) Colony formation assay data of RPE *p53*-/- cells transduced with a control vector, EWS-FLI1
684 or EWS-FLI1 37YS together with or without BRCA2 depletion (n=3, top) and the representative
685 colony images (bottom).
686 (I, J) The CTG assay showing that compared with U2OS, Ewing sarcoma cells showed
687 hypersensitivity to CDK12 inhibitors, THZ531 (I) and MFH290 (J).
688 (K, L) The CTG assay showing that depletion of EWS-FLI1 by sg*FLI*1 in TC32 cells led to
689 acquisition of resistance to CDK12 inhibitors, THZ531 (K) and MFH290 (L).

690 **Extended Data Figure 1: *EWSR1* loss is synthetic lethal with BRCA1/2 deficiency.**

691 (A) A Venn diagram showing overlap between genes identified in the previously reported CRISPR
692 screens to identify synthetic *BRCA2* lethality and genes whose dependence is positively correlated
693 with *POLQ* in the Depmap analysis.

694 (B-F, I) Colony formation assay data of RPE *p53*-/- cells and RPE *p53*-/- *EWSR1* knock-out clones
695 treated with X-ray (B), VE822 (C), Topotecan (D), Etoposide (E), Olaparib (F) and Novobiocin (I
696 4) (n=4).

697 (G) U2OS DR-GFP reporter cells (I-SceI based) were transduced lentivirally transduced with
698 control sgRNA or sgRNA targeting EWSR1, BRCA1 or BRCA2 (co-expressing puromycin
699 resistance gene). Cells were treated with puromycin for 4days, and then subjected to reporter
700 analysis (n=3).

701 (H) Schematic representation of the Cas9-based MMEJ reporter assay (ver.2).

702 (J, K) Colony formation assay data of MDA-MB-436 (J, n=3) and HCC1937 (K, n=3) cells and
703 the cells complemented with BRCA1 cDNA transduced with control, sg*POLQ* or sg*EWSR1* (left).
704 Shown are representative images of the colonies (right).

705 (L) RPE *p53*-/- cells were lentivirally transduced Cas9 (expressing balasticidin resistance gene).
706 After balasticicin selection, cells were lentivirally transduced with a sgRNA targeting control,
707 *POLQ* or *EWSR1* (co-expressing GFP) together with with a sgRNA targeting control, *BRCA1* or
708 *BRCA2* (co-expressing tRFP657). 72hr after the sgRNA transduction, the frequency of the
709 GFP/tRFP657 cells was traced by FACS (n=3).

710

711 **Extended Data Figure 2: Subacute depletion of EWSR1 results in POLΘ protein loss.**

712 **(A)** RPE *TP53-/-* 3xFLAG-POLQ knock-in cells were transduced with Cas9 and a sgRNA
713 targeting *EWSR1* by electroporation. Single clones are isolated 7days after electroporation. Shown
714 are immunoblotting analyses for FLAF- POLΘ, EWSR1 and GAPDH in control cells and *EWSR1*
715 knockout clones.
716 (B) Colony formation assay data of RPE *p53*-/- 3xFLAG-*POLQ* knock-in cells and RPE *p53*-/-
717 *EWSR1* knock-out 3xFLAG-*POLQ* knock-in clones treated with X-ray (n=3).
718 (C) RPE *p53*-/- 2xHA-FKBP12 V36F *EWSR1* cells were treated with DMSO, 1μM $dTAG^{13}$ or
719 1μM $dTAG^{v1}$ for 24hr. The cell lysates were extracted and then analyzed by immunoblotting.
720 $dTAG^{v1}$ but not $dTAG^{13}$ efficiently degraded endogenous EWSR1 protein.
721 (D) RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells are treated
722 with 1μM $dTAG^{V1}$ for the indicated time. The cell lysates were extracted and then analyzed by
723 immunoblotting.
724 (E) RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells are treated
725 with $dTAG^{V1}$ at the indicated concentration for 24hr. The cell lysates were extracted and then
726 analyzed by immunoblotting.
727 (F) RPE *p53*-/- 3xFLAG-*POLQ* single knock-in cells and RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-
728 FKBP12 V36F *EWSR1* double knock-in cells are treated with $dTAG^{V1}$ for indicated time. The cell
729 lysates were extracted and then analyzed by immunoblotting.
730 (G) Immunoblotting analyses showed that BRCA1, BRCA2 and other MMEJ proteins, except for
731 POLΘ, were not affected by *EWSR1* knockout in RPE *p53*-/- cells.
732 (H) RPE *p53*-/- 3xFLAG-*POLQ* knock-in cells were lentivirally transduced with control sgRNA
733 or sgRNA targeting *EWSR1, FUS* or *TAF15* (co-expressing puromycin resistance gene). Cells

734 were treated with puromycin for 4days, and then the cell lysates were extracted and then analyzed
735 by immunoblotting.
736 (I) K562 MMEJ ver.2 (Cas9-based) cells were lentivirally transduced with control sgRNA or
737 sgRNA targeting *EWSR1, FUS* or *TAF15* (co-expressing puromycin resistance gene). Cells were
738 treated with puromycin for 4days, and then cells were subjected to MMEJ reporter analysis (n=3).
739 (J) RPE *p53*-/- cells and RPE *p53*-/- *EWSR1* knockout cells were transduced with control, wildtype
740 *EWSR1* or *POLQ* cDNA. Colony formation assays were performed in the presence of Etoposide
741 (n=3).
742 (K) MDA-MB-436 cells and the cells complemented with BRCA1 were transduced with Cas9 and
743 a sgRNA targeting control or *EWSR1*. Cells were subjected to comet assay to measure unrepaired
744 DNA damage.
745 (L) RPE *p53*-/- cells and RPE *p53*-/- *EWSR1* knockout cells were transduced with Cas9 and a
746 sgRNA targeting control or *BRCA2*. Cells were treated with or without synchronization by 8μM
747 RO3306. 1 hour after removal of RO3306, cells were treated with 2 Gy irradiation and then
748 incubated for 24 hours. The frequency of micronuclei positive cells is shown.
749
750 **Extended Data Figure 3: EWSR1 prevents multiple exon skipping events of the *POLQ***
751 **transcript.**
752 (A) U2OS MMEJ reporter cells (I-SceI based) were transduced with *POLQ* cDNA (co-expressing
753 blasticidin resistance gene). After blasticidin selection, cells were subsequently transduced with
754 Cas9 and a sgRNA targeting control or *EWSR1* (co-expressing puromycin resistance gene). After
755 puromycin selection, cell lysates were extracted and then analyzed by immunoblotting. See also
756 Fig.2C.

757 (B) RPE *p53*-/- 3xFLAG-*POLQ* knock-in cells were transduced with Cas9 and a control sgRNA
758 or sgRNA targeting *EWSR1* by electroporation. 72hr after electroporation, cells were treated with
759 MG132 or Chloroquine for 4hr. After the treatment, cell lysates were extracted and then analyzed
760 by immunoblotting.
761 (C) A pie chart summarizing alternative splicing event types in RPE *p53*-/- *EWSR1* knock-out cells
762 obtained by MISO analysis.
763 (D-G) Volcano plots showing retained introns (RI) (D), alternatively spliced 5' splice cites (A5SS)
764 (E), alternatively spliced 3' splice cites (A3SS) (F), mutually exclusive exons (MXE) (G) in RPE
765 *p53*-/- *EWSR1* knock-out cells compared with RPE *p53*-/- cells.
766 (H) Volcano plots showing alternatively spliced skipped exon (SE) in HEK293T si*EWSR1*
767 transduced cells compared with control cells defined by RNA-seq analysis (GSE154944).
768 (I) Sashimi plots of the alternative splicing events in control and *EWSR1* knock-out RPE *p53*-/-
769 cells: the region between exon 23 and exon 26 of *POLQ* transcript.
770 (J) RT-PCR showing the exon 7, 21, 22 and 25 skipping of *POLQ* transcript in *EWSR1* knock-out
771 cells derived from RPE *p53*-/- (top) and from RPE *p53*-/-3xFLAG-*POLQ* knock-in cells (bottom).
772 (K) The schematics illustrates the location of protein domains with the primary sequence POLΘ.
773 Numbers refer to amino acid residues. Exo = exonuclease, T = thumb, P = palm, F = fingers. The
774 ribbon illustrations are AlphaFold 3-generated predictions of the WT and truncated form of the
775 POLΘ polymerase domain. The two views represent 180º rotation around the y-axis. Colors
776 illustrate the sub-domain architecture.
777
778 **Extended Data Figure 4: Exon 25 skipping of the *POLQ* transcript decreased POLΘ protein,**
779 **resulting in profound MMEJ deficiencies.**

780 (A) Schematics of the *POLQ* exon 25, intron 24/25 and the design of ASO and sgRNAs targeting
781 the splice sites.
782 (B) Immunoblotting (upper) and RT-PCR (lower) analysis showing that transduction of two ASOs
783 efficiently induced exon 25 skipping and POLΘ protein loss in RPE *p53*-/- 3xFLAG-*POLQ* knock-
784 in cells.
785 (C) U2OS MMEJ ver.2 (Cas9-based) were transduced control or two ASOs targeting *POLQ* exon
786 25 splice sites. 48h after ASO transduction, cells were subjected to MMEJ reporter analysis (n=4).
787 (D) Schematic representation of the endogenous *POLQ* exon 25 and intron 24/25 and the minigene
788 constructs.
789 (E) HEK293T 2xHA-FKBP12 V36F *EWSR1* dknock-in cells were transfected with the indicated
790 minigene construct and then treated with DMSO or $dTAG^{V1}$. 48h after transfection, cells were
791 subjected to assess the endogenous- and minigene- derived *POLQ* exon 25 skipping (left) and
792 FACS analysis to measure the reporter activity (right).
793 (F-H) Binding of EWSR1 to *POLQ* pre-mRNA evaluated using the RNA immuno-precipitation
794 (RIP) assay, performed using an anti-EWS antibody and an anti-HA antibody (N=2). The assay
795 was performed in HEK293T 3xFLAG-POLQ/2xHA-FKBP12 V36F *EWSR1* double knock-in cells
796 in the presence of DMSO or $dTAG^{v1}$. qPCR was performed with primers targeting *POLQ* pre-
797 mRNA around exon 7 and 25, *POLQ* mature mRNA and *GAPDH* (bottom). Shown were
798 immunoblotting for cell lysates with or without immunoprecipitation (F), RIP-qPCR assay for the
799 samples with EWS-IP (G) and HA-IP (H).
800 (I) Binding of multiple EWSR1 mutants to *POLQ* pre-mRNA evaluated using the RNA immuno-
801 precipitation (RIP) assay, performed using anti-EWS antibody. The assay was performed in
802 HEK293T 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells transduced with

803 a control vector, EWSR1 WT, EWSR1 7YS mutant or EWSR1 37YS mutant in the presence of
804 DMSO or dTAG$^{v1}$. qPCR was performed with primers targeting *POLQ* pre-mRNA around exon
805 7 and 25 and *GAPDH* (bottom).
806
807 **Extended Data Figure 5: Samples from Ewing sarcoma patients show no HRD genomic**
808 **signatures.**
809 (A-C) Contribution of COSMIC indel (ID)-signature (A) and stacked bar plots of the numbers of
810 indels attributed to COSMIC ID signatures for Ewing sarcoma tumors (ES) (in B), and
811 neuroblastoma tumors (NB) (in C). Short somatic mutations were determined from whole genome
812 sequencing data of tumors and matched normals using GATK4, and SigProfilerAssignment was
813 used to estimate the COSMIC ID signatures from the mutations. Each bar contains relative
814 numbers of indels attributed to specific ID signatures, which are shown in different colors. There
815 are no ES tumors with MMEJ-related ID6 indels (B). In contrast, there are several NB tumors with
816 ID6 indels (C).
817 (D-F) Contribution of COSMIC SNV-signature 3 (D) and stacked bar plots of the number of SNVs
818 attributed to COSMIC SNV signatures for Ewing sarcoma tumors (ES) (in E), and neuroblastoma
819 tumors (NB) (in F). Short somatic mutations were determined from whole genome sequencing
820 data of tumors and matched normals using GATK4, and SigProfilerAssignment was used to
821 estimate the COSMIC SNV signatures from the mutations. Each bar represents estimated numbers
822 of SNVs attributed to specific SNV signatures, which are shown in different colors. There are no
823 ES tumors with MMEJ-related Signature 3 SNVs (E). In contrast, there are several NB tumors
824 with Signature 3 SNVs (F).

825 (G-H) Distribution of insertions and deletions by size in Ewing sarcoma tumors (ES; n = 23), and
826 neuroblastoma tumors (NB; n = 47). Short somatic mutations were determined from the whole
827 genome sequencing data of tumors and matched normals. Each bar represents the proportion of
828 deletions (G) or insertions (H) of a particular size among all deletions or all insertions, respectively,
829 of each cohort.
830 (I) Low templated insertions in Ewing sarcoma tumors (ES; n = 23) as compared to neuroblastoma
831 tumors (NB; n = 47).
832 (J, K) The bar graphs show the length of microhomology at canonical translocation breakpoints in
833 Ewing sarcoma samples from GMKF (J, n=46) and St. Jude (K, N=43) cohorts. All translocations
834 were called using the Manta structural variant caller. Canonical translocations are defined as those
835 impacting one of the following genes: EWSR1, FLI1, ETV1, FEV, FUS, ERG.
836
837 **Extended Data Figure 6: The prion-like domain of EWSR1 is critical for the faithful splicing**
838 **of *POLQ* transcript.**
839 (A) RPE *p53*-/- 3xFLAG-*POLQ* knock-in cells were retrovirally transduced with empty vector,
840 wildtype *EWSR1*, or indicated *EWSR1* mutants (co-expressing blasticidin resistance gene)
841 together with lentivirally Cas9 and a control sgRNA or sgRNA targeting *EWSR1* (co-expressing
842 puromycin resistance gene). Cells were treated with puromycin and blasticidin for 4days, and
843 then cell lysates were extracted and then analyzed by immunoblotting.
844 (B) U2OS MMEJ ver.2 (Cas9-based) cells were lentivirally transduced with control sgRNA or
845 sgRNA targeting *EWSR1* together with empty vector wild-type EWSR1 or the indicated EWSR1
846 mutants by retroviral transduction. Cells were treated with puromycin and blasticidin for 4days
847 and then subjected to MMEJ reporter analysis (n=4).

848 (C-E) Colony formation assay data of RPE *p53*-/- cells and RPE *p53*-/- *EWSR1* knock-out clones
849 transduced with an empty vector, wild-type *EWSR1* or the indicated *EWSR1* mutants by
850 retroviral transduction in the presence of 10nM Topotecan (C), 62.5nM Etoposide (D), or 100nM
851 VE822 (E), respectively. The data is normalized to the DMSO control (n=3).
852 (F, G) K562 cells were lentivirally transduced with sgRNAs targeting control or *EWSR1* (co-
853 expressing puromycin resistance gene) and control or *BRCA2* (co-expressing blasticidin resistance
854 gene). Cells were treated with puromycin and blasticidin for 4days, and then retrovirally
855 transduced with indicated constructs (co-expressing GFP). 72hr after retroviral transduction, the
856 change of frequency of GFP positive fraction was traced by FACS (n=3).
857 (H) Schematic representation of the deletion mutants within the prion-like domain used in Fig. S6I.
858 (I) RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells were
859 retrovirally transduced with empty vector, wildtype EWSR1, or indicated EWSR1 mutants. After
860 transduction, cells were treated with blasticidin for four days. Then, cells were treated with DMSO
861 or 1μM $dTAG^{v1}$ for 24hr. After $dTAG^{v1}$ treatment, cell lysates were subjected to immunoblot
862 analysis.
863 (J) Wildtype EWSR1 but not the EWSR1 7YS or 37YS mutant rescued POLΘ protein expression
864 in RPE *p53*-/- 3xFLAG-*POLQ* knock-in cells.
865 (K) Wildtype EWSR1 but not the EWSR1 7YS or 37YS mutant prevented *POLQ* exon 25 skipping
866 measured by K562 minigene reporter cells (n=3).
867
868 **Extended Data Figure 7: Exon skipping in *EWSR1* KO cells was typically observed for exons**
869 **that are flanked by longer introns.**

870 (A) Volcano plots comparing detected protein levels by IP-MS between EWSR1 wildtype and
871 control.
872 (B) The length of introns surrounding included, excluded or background exons in *EWSR1*
873 knockout cells identified by RNA-seq analysis.
874 (C) Dual depletion of FUBP1 and KHSRP increased *POLQ* exon 25 skipping measured by K562
875 minigene reporter cells (n=3).
876 (D) The FUBP1 16YS mutant failed to rescue *POLQ* exon 25 skipping induced by FUBP1/KHSRP
877 dual depletion measured by K562 minigene reporter cells (n=3).
878 (E) RPE *p53*-/- 3xFLAG-*POLQ*/2xHA-FKBP12 V36F *EWSR1* double knock-in cells were
879 lentivirally transduced with Cas9 and a sgRNA targeting control (NT) or sgRNAs targeting
880 FUBP1 and KHSRP together with retrovirally transduced with a control vector, wildtype
881 KHSRP or the KHSRP 16YS mutant. After puromycin and blasticidin selection, cells were
882 subjected to immunoblotting analysis.
883 (F) RT-PCR showing the *POLQ* exon25 skipping in the indicated cells.
884 (G) U2OS MMEJ ver.2 (Cas9-based) cells were lentivirally transduced with indicated sgRNA(s)
885 (co-expressing puromycin resistance gene). Cells were treated with puromycin for 4days, and
886 then subjected to MMEJ reporter analysis (n=4).
887 (H) The KHSRP 17YS mutant failed to rescue *POLQ* exon 25 skipping induced by
888 FUBP1/KHSRP dual depletion measured by K562 minigene reporter cells (n=3).
889
890 **Extended Data Figure 8: Depletion of EWS-FLI1 ameliorates MMEJ activity in Ewing**
891 **sarcoma cells.**

892 (A) The cell lysates were extracted from an Osteosarcoma cell line U2OS and Ewing Sarcoma cell
893 lines (A673, TC32, TC71, RH1, and RD-ES, and then analyzed by immunoblotting.
894 (B) Cas9-based MMEJ reporter cells were lentivirally transduced with an empty vector, wild-type
895 *EWS-FLI1* or indicated mutant forms of *EWS-FLI1* (co-expressing puromycin resistance gene).
896 After puromycin selection, cells were subjected to MMEJ reporter assay (n=3).
897 (C, D) the CellTiter-Glo® (CTG) assay showing Ewing sarcoma cells showed resistance to POLΘ
898 inhibitors, Novobiocin (C) and ART558 (D).
899 (E, F) Colony formation assay data of RPE *p53*-/- cells lentivirally transduced with control vector,
900 EWS-FLI WT or the EWS-FLI1 37YS mutant treated with Etoposide (E) and Topotecan (F) (n=2).
901 (G) The cell lysates used in Fig.4C were extracted and then analyzed by immunoblotting. The
902 black and red arrows showing EWS-FLI1 and FLI1, respectively.
903 (H) The CTG assay shows that depletion of EWS-FLI1 in A673 cells sensitized to ART558.
904 (I) U2OS I-SceI-based MMEJ reporter cells were transduced with control or EWS-FLI1 together
905 with an empty vector, RNaseH1 WT, the RNaseH1 DN, or the RNaseH1 WKKD mutant. Cells
906 were treated with puromycin and blasticidin for 4days and then subjected to MMEJ reporter
907 analysis (n=6).
908 (J) A673 carrying the Isce-I-based MMEJ reporter were transduced with an empty vector,
909 RNaseH1 WT, the RNaseH1 DN, or the RNaseH1 WKKD mutant. Cells were treated with
910 blasticidin for 4days and then subjected to MMEJ reporter analysis (n=3).
911
912 **Extended Data Figure 9: Combination of Peposertib and Etoposide efficiently killed Ewing**
913 **sarcoma cells in vivo.**

914 (A, B) The CTG assay showing that compared with U2OS, Ewing sarcoma cells showed
915 hypersensitivity to Topotecan (A) and Etoposide (B).
916 (C) Overall survivals showing a therapeutic efficacy of a combination of Etoposide and Peposertib
917 in a TC32 xenograft model (n=10, each group).
918 (D) The changes of the body weight of the treated mice.
919 (E, F) The CTG assay showing that depletion of EWS-FLI1 in A673 cells led to acquisition of
920 resistance to CDK12 inhibitors, THZ531 (E) and MFH290 (F).
921 (G) The schematic representation of the findings of this study.

922 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

923 **Cell cultures**

924 RPE *p53*-/- cells generated previously[61] were grown in DMEM/F12 containing GlutaMAX
925 (Gibco) and supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich). HEK293T
926 (ATCC, CRL-3216) cells were grown in DMEM (Gibco) supplemented with 10%FBS. TC32 and
927 TC71 cell lines were obtained from children cancer repository and grown in RPMI (Gibco)
928 supplemented with 10%FBS. RH-1 (DSMZ, ACC493) cell grown in RPMI 1640 supplemented
929 with 10%FBS, 1% P/S. RD-ES (ATCC, HTB-166) cells were grown in RPMI 1640 supplemented
930 with 20%FBS, 1% P/S. SK-ES1 (ATCC, HTB-86) and SK-PN-DW (ATCC, CRL-2139) cells
931 were cultured in DMEM (Gibco) supplemented with 10%FBS. U2OS parental (ATCC, HTB-96)
932 and carrying a DNA repair template cells were grown in DMEM (Gibco) supplemented with
933 10%FBS. The BRCA1-deficient cell lines HCC1937 (ATCC, CRL-2336) and HCC1954 (ATCC,
934 CRL-2338) and BRCA1-complemented cells were grown in RPMI (Gibco) supplemented with
935 10%FBS. RH-1 (DSMZ, ACC493). K562 (ATCC, CCL-243) cells were grown in RPMI 1640
936 supplemented with 20%FBS, 1% P/S. MDA-MB-436 and MDA-MB-436+BRCA1 cells were a
937 kind gift from Dr. Geoffrey Shapiro (Dana Farber Cancer Institute) and grown in RPMI (GIBCO)
938 supplemented with 10% fetal bovine serum. DLD1 BRCA2 -/- and DLD1 Parental were obtained
939 from Horizon discovery, Waterbeach, UK. U2OS cells with HR reporter (DR-GFP) was a kind
940 gift from Dr. Jeremy Stark (City of Hope).

941 To establish the U2OS-MMEJ reporter cells, U2OS cells were lentivirally transduced with the
942 Isce-I-based (ver.1) or Cas9-based (ver.2) MMEJ reporter and then incubated with puromycin.
943 After selection with puromycin, a single clone was isolated and subsequently the clones were
944 validated.

945 All cells were cultured in an incubator maintained at 37°C, 5% $CO_2$ and a relative humidity of
946 95%.
947
948 **Compounds and antibodies**
949 Chemical compounds, including Olaparib (#S1060), Novobiocin (#S2492), Etoposide (#S1225),
950 VE822 (#S7102), Ro-3306 (#S7747), M3814 (#S8586) and THZ0531 (#$6595) were purchased
951 from Selleckchem. ART558 (#HY-141520) and MFH-290 (#HY-153244) were purchased from
952 MedChemExpress. Nocodazole (#1404) was purchased from Sigma Aldrich. Chemicals were
953 dissolved in DMSO and kept in small aliquots at −20°C or 80°C. The key sources of antibodies
954 are: POLθ (Cell Signaling Technology, #64708), γH2AX (Millipore #05–636), LIG3 (BD
955 Biosciences, #611876), HMCES (Santa Cruz Biotechnology, SC-514238), HMCES (Atlas
956 Antibodies #HPA044968), APEX2 (Cell Signaling Technology, #74728), PARP1(Cell Signaling
957 Technology, #9542), CIP2A (Abcam, ab128179), DNA-PK (Cell Signaling Technology, #12311),
958 Phospho-DNA-PKcs (Ser2056) (Cell Signaling Technology, #68716), TOPBP1 (Abcam, ab2402),
959 XRCC1 (Cell Signaling Technology, #76998), FEN1 (Cell Signaling Technology, #82354),
960 EWSR1 (Cell Signaling Technology, #11910 for WB), EWSR1 (Santa Cruz Biotechnology, sc-
961 28327 for WB), EWSR1 (Life Technologies, A300417A for IP and RIP), BRCA1 (Merck
962 Millipore, OP92), BRCA2 (Cell Signaling Technology, #10741), FLI1 (Abcam, ab133485), FUS
963 (Cell Signaling Technology, #67840), TAF15(Cell Signaling Technology, #28409), Vinculin (Cell
964 Signaling Technology, #4650), FLAG (Sigma, F1804), HA (Cell Signaling Technology, #3724
965 for WB), HA (Roche, #11867431001 for IP and RIP), V5 (Cell Signaling Technology, #13202 for
966 WB and IF), V5(Abcam, ab15828 for ChIP), V5(Abcam, ab27671 for IP and WB), GAPDH (Cell
967 Signaling Technology, #5174), Histone H3 (Cell Signaling Technology, #9715), FUBP1 (Abcam

968 ab181111), KHSRP (Cell Signaling Technology, #13398), SMARCA4/BRG1 (Cell Signaling
969 Technology, #3508), SMARCC1 (Cell Signaling Technology, #11956), MYC (Santa Cruz
970 Biotechnology, SC-40).
971
972 **Generation of CRISPR engineered cell lines**
973 For generating *EWSR1* knock-out clones in HEK293T and RPE *p53*-/- cells, a ribonucleoprotein
974 (RNP) complex was formed by Alt-R™ S.p. HiFi Cas9 and CRISPR-Cas9 sgRNA targeting
975 *EWSR1* (Integrated DNA Technologies). For generating endogenously 3xFLAG-tagged *POLQ*
976 knock-in HEK293T and RPE *p53*-/- cells, 3xFLAG-tagged *POLQ* HEK293T *EWSR1*-FKBP12
977 V36F double knock-in cells or V5-tagged *POLQ* knock-in U2OS and A673 cells, the Cas9/sgRNA
978 RNP complex together with a template was delivered. Cells were resuspended in SE (U2OS cells),
979 SF (HEK293T and A673 cells) or P3 (RPE *p53*-/- cells) Nucleofector Solution with supplement
980 (Lonza) and then mixed with the Cas9/sgRNA RNP complex and Alt-R™ Cas9 Electroporation
981 Enhancer (Integrated DNA Technologies), respectively. The Cas9/sgRNA RNP complex was
982 delivered to the cells by 4D-Nucleofector™ (Lonza). Media was changed the following day, and
983 then 2 days after cells were seeded in a 96-well plate to isolate single clones. Individual clones
984 were tested for genomic editing analyses using immunoblotting and genomic PCR with subsequent
985 Sanger sequencing.
986
987 **Plasmids**
988 psPAX2 and pMD2.G (Addgene #12260 and #12259) were gifts from Dr. Didier Trono. pLV-
989 EF1a-IRES-Blast (pLV-Blast) (Addgene #85133) was a gift from Dr. Tobias Meyer.[62]

990 pCDH-puro-EWS-FLI1 was a gift from Jialiang Wang (Addgene plasmid # 102813). pLenti
991 CMV Puro LUC (w168-1) was a gift from Eric Campeau & Paul Kaufman (Addgene plasmid #
992 17477). pCVL.SFFV.d14GFP.EF1a.HA.NLS.Sce(opt).T2A.TagBFP was a gift from Andrew
993 Scharenberg (Addgene plasmid # 32627). EWS-FLI1 37YS mutant was cloned into pCDH-puro
994 vector. EWSR1, FUBP1, KHSRP, RNaseH1 and its mutant forms were cloned into pMYs-IRES-
995 Blastcidin vector, respectively. sgRNAs were cloned into a lentiCRISPR-v2 construct as
996 previously described[63]. Of note, sgEWSR1 #2, sgFUBP1 #2 or sgKHSRP #2 target genomic
997 DNA but not cDNA, so that endogenous knockout and ectopic cDNA expression are possible
998 simultaneously. All the sgRNA sequences used in this study were provided in Supplemental
999 Table 1.
1000
1001 **Virus generation**
1002 Lentivirus was generated by co-transfecting lentiviral constructs with packaging vectors, psPAX2
1003 and pMD2.G, into HEK293T cells using CalPhos Mammalian Transfection Kit (Takara Bio). Two
1004 days after transfection, supernatant containing virus was harvested and filtered through 0.45 μm
1005 filter. Viral transductions were performed in the presence of 8 μg/ml (for RPE *p53*-/-, MDA-MB
1006 436, DLD1, HCC1937, and HCC1954 cells) or 4 μg/ml (for K562 cells) polybrene (Sigma-
1007 Aldrich). For all virus generations and transductions media supplemented with heat inactivated
1008 FBS was used. Retrovirus was generated by co-transfecting retroviral constructs with packaging
1009 vectors, and, into HEK293T cells using CalPhos Mammalian Transfection Kit (Takara Bio). Two
1010 days after transfection, supernatant containing virus was harvested and filtered through 0.45 μm
1011 filter. Viral transductions were performed in the presence of 8 μg/mL polybrene (Sigma-Aldrich).
1012 Puromycin dihydrochloride (Sigma-Aldrich) or Blasticidin S HCl (Thermo Fisher Scientific) was
1013 used after infection of retrovirus or lentivirus.

1014

1015 **Colony formation assay**

1016 Cells were seeded in 6-well plates at 500-10,000 cells/well. The day after seeding, cells were
1017 exposed to each drug or x-ray. X-ray irradiation was performed using RS-2000 Irradiator (Rad
1018 Source Technologies). After incubation, cells were washed with PBS and then fixed in a mixture
1019 of 50% methanol, 10% acetic acid, and 40% water for 20 minutes. The fixed cells were stained
1020 with 0.5% crystal violet (Sigma-Aldrich) in 20% methanol for 2–6 h and then washed with water
1021 twice. The plates were imaged on GE Amersham Imager 600 using colorimetric transillumination
1022 setting. The area fraction of the colonies in each well was estimated using ImageJ.

1023

1024 **CellTiter-Glo assay**

1025 To test drug sensitivity, U2OS or Ewing sarcoma cells were seeded in 96-well plates at 1000–3000
1026 cells/well. The day after seeding, drug was added to the wells, and then the plates were incubated
1027 for 3-5 days. After incubation, ATP in each well was quantified using CellTiter-Glo Luminescent
1028 Cell Viability Assay (Promega). CellTiter-Glo assays were performed in technical duplicate if not
1029 otherwise specified in a figure legend.

1030

1031 **Immunofluorescence**

1032 Mitotic synchronization was achieved using Ro-3306 (8 μM, 16 hours) followed by 30min release
1033 in fresh medium. Then cells were incubated with 100ng/ml nocodazole and subsequently
1034 irradiated. 4h after irradiation or in the absence of treatment, cells were fixed with 4%
1035 paraformaldehyde for 15 minutes at room temperature. Cells were then permeabilized with 0.3%
1036 Triton X-100 for 10 minutes on ice, followed by blocking with 3% non-fat milk for 1 hour at room

1037 temperature. The slides were stained with primary antibodies at 4°C overnight. Afterward, they
1038 were stained with secondary antibodies for 1 hour at room temperature. The slides were scanned
1039 using fluorescence microscope. At least 100 cells were counted for each sample. Foci
1040 quantification was performed using software CellProfiler version 4.2.6.
1041
1042 **Immunoblotting**
1043 Whole cell lysate was prepared by lysing cells in ice-cold RIPA buffer (Cell Signaling
1044 Technology) supplemented with Protease/Phosphatase inhibitor Cocktail (Cell Signaling
1045 Technology). Cell lysate was cleared by centrifuging the samples at 12000 rpm, 4°C for 20
1046 minutes. Supernatant was recovered and protein concentration was estimated using BCA assay kit
1047 (Thermo Fisher Scientific). Protein samples were boiled in 2X Laemmli sample buffer (Bio-rad)
1048 supplemented with 5% 2-mercaptoethanol or NuPAGE LDS sample buffer (Thermo Fisher
1049 Scientific) supplemented with NuPAGE sample reducing agent (Thermo Fisher Scientific).
1050 Samples were run in NuPAGE 4-12% Bis-Tris gels (Thermo Fisher Scientific) using NuPAGE
1051 MES SDS Running Buffer (Thermo Fisher Scientific) or NuPAGE 3-8% Tris-Acetate Gel
1052 (Thermo Fisher Scientific) using NuPAGE Tris-Acetate SDS Running Buffer (Thermo Fisher
1053 Scientific). After electrophoresis, samples were transferred onto a nitrocellulose membrane
1054 (Amersham). Membranes were blocked for 1 h with Blocking Buffer, which was prepared by
1055 diluting Fish Serum Blocking Buffer (Thermo Fisher Scientific) with Tris-buffered saline
1056 containing 0.1% Tween20 (TBST). Membranes were incubated with primary antibodies diluted in
1057 Blocking Buffer overnight at 4°C. Membranes were washed with TBST three times, 5 minutes
1058 each, incubated with secondary antibodies diluted in blocking buffer for 1h, which either
1059 horseradish peroxidase (HRP)-conjugated for chemiluminescent detection or fluorophore-

1060 conjugated for visualization using the LI-COR imaging system. After secondary antibody
1061 incubation, membranes were washed with TBST and imaged using chemiluminescence setting on
1062 GE Amersham Imager 600 or the LI-COR imaging system.
1063
1064 **DNA repair template assay**
1065 DR-GFP and EJ5-GFP cells were kind gifts from Dr. Jeremy Stark. MMEJ-GFP cells (ver.1 and
1066 ver.2) were generated by lentivirally transducing a cassette into U2OS or Ewing sarcoma cells
1067 using a pLV-EF1a-IRES-puro vector (addgene #85132) as previously described [20]. U2OS cells
1068 carrying a DNA repair template reporter (DR-GFP, EJ5-GFP, and MMEJ-GFP) or Ewing sarcoma
1069 cells carrying a MMEJ reporter were lentivirally transduced with Cas9 and sgRNA (co-expressing
1070 puromycin resistance gene). After puromycin or blasticidin selection, 40,000 DNA repair template
1071 reporter cells were seeded in 12-well plates and adenovirally (for U2OS) or lentivirally (co-
1072 expressing BFP, for Ewing sarcoma cells) transduced with Isce-I the following day. 48h after Isce-
1073 I transduction, GFP signals were analyzed by CytoFLEX (Beckman). The signals were normalized
1074 to control cells. For U2OS Cas9-based MMEJ reporter assay, the cells were subsequently
1075 transduced with Cas9 and sgRNA targeting the reporter (co expressing BFP). 72h after
1076 Cas9/sgRNA transduction, GFP signals within BFP positive fraction were analyzed by CytoFLEX
1077 (Beckman). For Cas9-based MMEJ reporter assay using HEK293T *POLQ* WT or knockout cells,
1078 the MMEJ reporter, Cas9 and sgRNA targeting the reporter (co expressing BFP) together with a
1079 control vector, *POLQ* WT cDNA or *POLQ* mutant cDNA were transduced using Lipofectamine
1080 LTX with Plus reagent (Thermo Fisher Scientific). After 72 h incubation, GFP signals within BFP
1081 positive fraction were analyzed by CytoFLEX (Beckman).
1082

1083 **Comet assay**

1084 The alkaline comet assays were performed to detect both single and double-stranded DNA breaks.
1085 1000 cells, suspended in PBS, were mixed with 50μl melted low melting agarose and pipetted onto
1086 the Cometslides (R&D Systems). Once the agarose solidified, the slides were immersed in a lysis
1087 solution for 18 hours to facilitate cell lysis. Following lysis, the slides were incubated in an alkaline
1088 unwinding solution for 1 hour to denature the DNA. Electrophoresis was performed at 21 V for 45
1089 minutes in an alkaline electrophoresis solution. The slides were stained with SYBR green solution,
1090 scanned by fluorescence microscope, and analyzed using CometScore 2.0 software.

1091

1092 **Chromatin immunoprecipitation (ChIP) assay**

1093 U2OS biallelic endogenous V5-*POLQ* expressing cells carrying an Isce-I based MMEJ reporter
1094 were adenovirally transuduced with Isce-I in the presence or absence of 100ng/ml nocodazole.
1095 12h after Isce-I trunscduction, chromatin immunoprecipitation was performed using SimpleChIP
1096 Enzymatic Chromatin IP Kit (Magnetic Beads) (Cell Signaling Technology). Protein-DNA
1097 cross-linking was performed with 1% formaldehyde and glycine solution. Chromatin was
1098 fragmented with Micrococcal Nuclease. The digested chromatin was incubated at 4°C with each
1099 antibody for 4 h. After the antibody reaction, ChIP-Grade Protein G Magnetic Beads were added
1100 to the samples, and the samples were incubated at 4°C for 1–2 h. For crosslinking reversal, each
1101 sample was incubated at 65°C with Proteinase-K overnight. After DNA purification,
1102 immunoprecipitated DNA was analyzed by qPCR using a primer set targeting the integrated
1103 MMEJ reporter locus. ChIP-qPCR was performed in technical triplicate. All the primer
1104 sequences for ChIP-qPCR used in this study were provided in Supplemental Table 1.

1105

1106 **Reverse transcription PCR (RT-PCR)**

1107 Total RNA was extracted using RNeasy Mini Kit (QIAGEN). Complementary DNA synthesis
1108 was performed using SuperScript IV First-Strand cDNA Synthesis Reaction (Thermo Fisher
1109 Scientific). The PCR products were separated in 2% agarose gel stained with SYBR Safe
1110 (Thermo Fisher Scientific) and imaged using ChemiDoc MP Imaging System (Bio-rad). All the
1111 primer sequences for RT-PCR used in this study were provided in Supplemental Table 1.

1112

1113 **ASO transfection**

1114 ASO was designed to target near the splice sites of the *POLQ* exon 25. The ASO was synthesized
1115 using 2′-O-methoxyethyl-phosphorothioated bases by Integrated DNA Technologies (IDT). Prior
1116 to ASO transfection, cells were seeded in a 6-well plate and cultured overnight. ASO was
1117 transfected at a final concentration of 0.1 μM using Lipofectamine LTX with Plus reagent (Thermo
1118 Fisher Scientific). After 48 h incubation, cells were harvested and analyzed by RT-PCR and
1119 western blotting. All the sequences for ASO used in this study were provided in Supplemental
1120 Table 1.

1121

1122 **Splicing minigene reporter assay**

1123 A mini-gene reporter carrying mNeongreen split by synthetic intronic sequence derived from
1124 *POLQ* intron 24/25 and exon 25 was generated as previously described [28]. The synthetic intron
1125 was made up of 250 bp of the start of the *POLQ* intron 24, 250 bp or 1000bp of the end of the
1126 *POLQ* intron 24, the *POLQ* exon 25, 250 bp or 1000bp of the start of the *POLQ* intron 25m and
1127 250 bp of the end of the *POLQ* intron 25. The synthetic introns and exon so generated were
1128 placed between the sequences encoding the N-terminus and C-terminus of split mNeongreen.

1129 The mScarlet-P2A-split mNeongreen (N-terminus)-synthetic introns/exon-split mNeongreen (C
1130 terminus) construct was integrated into a pLV-EF1a-IRES-Neo vector (Addgene #85139) to
1131 generate the mini-gene reporter. K562 cells or Ewing sarcoma cell lines were lentivirally
1132 transduced with the mini-gene reporter. 48hr after lentiviral transduction, the cells were enriched
1133 by using G418 (0.8mg/ul) for seven days. After antibiotic selection, mScarlet and mNeongreen
1134 signals in the reporter cells were analyzed by CytoFLEX (Beckman). For the flow cytometry
1135 analyses, mScarlet positive fractions within living cells were gated. Thereafter, the frequency of
1136 the mNeongreen positive cells within mScarlet expressing cells was analyzed by using Flowjo
1137 ver.10.9.0.
1138
1139 **Mass spectrometry analysis**
1140 HEK293T EWSR1-FKBP12 V36F knock-in cells were seeded into 15cm dishes and cultured
1141 overnight. The cells were transfected with pMYS-IB empty vector (EV) or pMYs-IB-EWSR1
1142 wildtype or pMYs-IB-EWSR1 37YS plasmid. Cells were transfected with 30 µg of vector /dish
1143 using Lipofectamine LTX with Plus reagent. 16h after transfection, media was changed with 1uM
1144 dTAG$^{v1}$. The cells were harvested 40 h after transfection and then were lysed with 40 mM Tris-
1145 HCl pH 8 buffer supplemented with 0.5% NP-40, 200 mM NaCl, 2 mM EDTA,
1146 Protease/Phosphatase inhibitor Cocktail (Cell Signaling Technology), and MG-132. The lysate
1147 was centrifuged, and the pellet was resuspended with 20 mM Tris-HCl pH 7.5 supplemented with
1148 100 mM KCl, 2 mM MgCl2, 1 mM CaCl2, and Protease/Phosphatase inhibitor Cocktail and treated
1149 with Micrococcal Nuclease (Thermo Fisher Scientific) for approximately 30 minutes, followed by
1150 addition of 5 mM EGTA (final concentration) to stop digestion. After three pulses of sonication
1151 (30 s each), samples were centrifuged to collect the supernatants as chromatin fractions.

1152 Immunoprecipitation was performed using Anti-FLAG M2 antibody (Sigma-Aldrich) conjugated
1153 with Dynabeads Protein G for Immunoprecipitation (Thermo Fisher Scientific). Beads were
1154 washed with TGN-150 wash buffer containing 20 mM Tris-HCl pH 7.6, 150 mM NaCl, 3 mM
1155 MgCl2, 10% glycerol, and 0.01% NP-40. Beads were then boiled at 70□C for 20 minutes in 1:1
1156 diluted NuPAGE LDS Sample buffer (4X) (Thermo Fisher Scientific) to elute the proteins. The
1157 eluted proteins were applied to SDS-PAGE, and Coomassie stained gel band samples were used
1158 for the analyses. Two independent biological replicate samples were analyzed by mass
1159 spectrometry.
1160 The gel bands were cut into approximately 1 mm3 pieces. Gel pieces were then subjected to a
1161 modified in-gel trypsin digestion procedure. Gel pieces were washed and dehydrated with
1162 acetonitrile for 10 minutes and completely dried in a SpeedVac (Thermo Fisher Scientific). The
1163 gel pieces were rehydrated with 50 mM ammonium bicarbonate containing 12.5 µg/mL modified
1164 sequencing-grade trypsin (Promega) at 4°C. After rehydration, the excess trypsin solution was
1165 removed and replaced with 50 mM ammonium bicarbonate solution. Samples were then incubated
1166 at 37°C overnight. Peptides were later extracted with a solution containing 50% acetonitrile and
1167 1% formic acid and dried in a SpeedVac. The samples were reconstituted in 5–10 µL of HPLC
1168 solvent A (2.5% acetonitrile, 0.1% formic acid). A nano-scale reverse-phase HPLC capillary
1169 column was created by packing 2.6 µm C18 spherical silica beads into a fused silica capillary (100
1170 µm inner diameter x ~30 cm length) with a flame-drawn tip. After equilibrating the column each
1171 sample was loaded via an EASY-nLC (Thermo Fisher Scientific). A gradient was formed, and
1172 peptides were eluted with increasing concentrations of solvent B (90% acetonitrile, 0.1% formic
1173 acid). As peptides eluted, they were subjected to electrospray ionization and then entered into a
1174 Orbitrap Exploris480 mass spectrometer (Thermo Fisher Scientific). Peptides were detected,

1175 isolated, and fragmented to produce a tandem mass spectrum of specific fragment ions for each
1176 peptide. Peptide sequences were determined by matching protein databases with the acquired
1177 fragmentation pattern by the software program, Sequest (Thermo Fisher Scientific). All databases
1178 include a reversed version of all the sequences and the data was filtered to between a one and two
1179 percent peptide false discovery rate. Perseus (version 2.0.11.0) [64] was used to preprocess the
1180 protein data. The protein intensities were normalized by the median and log2-transformed. Proteins
1181 with only one intensity value or with a sum of all total peptide counts of less than 5 were filtered
1182 out. For each sample, missing log2-intensity values were imputed using the normal distribution
1183 with width of 0.3 and down shift of 1.8. Histograms of log2-instensity values were inspected to
1184 confirm that the imputed values were at the lower end. Finally, limma analysis (version 3.54.2) [65]
1185 was performed on the log2-instensity values to determine enriched and depleted proteins. Volcano
1186 plots were generated using ggplot2 and ggrepel R packages.
1187
1188 **RNA immuno-precipitation (RIP) assay**
1189 Protocol for RIP assay was adapted from Nicholson-Shaw et al., 2022[29]. HEK293T EWSR1-
1190 FKBP12 V36F knock-in cells were cultured in 15 cm dishes and 2-3 plates were used per cell line
1191 for each assay. 16hr after transfection, media was changed with 1uM dTAG$^{v1}$. Cells were harvested
1192 40hr after transfection, washed with ice-cold PBS 2X times, cross-linked with 0.2% PFA for 15
1193 min. and then quenched by adding glycine to a final concentration of 125mM. Cells were lysed in
1194 iCLIP buffer (50mM Tris-HCl pH 7.4, 100mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium
1195 deoxycholate) supplemented with protease, phosphatase and RNase inhibitor cocktail. The
1196 samples were rotated at 4°C for 10 min. and then sonicated on 'light' setting twice. The lysate was
1197 incubated with Protein A beads on a rotor for 10 min. at 4°C to pre-clear the extract. The lysate

1198 was centrifuged at max speed for 10 min. and the supernatant was recovered and incubated for 2
1199 hrs on a rotor at 4°C with Protein A beads coupled to EWSR1 Ab (Life Technologies). A part of
1200 the lysate (0.4-0.6%) was kept aside as input. After 2 hrs the beads were washed twice with iCLIP
1201 buffer. The input and the IP samples were all subjected to TURBO DNase treatment (Thermo
1202 Fisher Scientific) for 30 min. at 37°C to get rid of the genomic DNA and the cross-links were
1203 reversed by digesting with Proteinase K (NEB) for 30 min. at 37°C and further digestion and
1204 denaturation of proteins was achieved by adding urea to a final concentration of 2.5M and
1205 incubating the samples at 37°C for additional 20 min. After digestion of the DNA and protein from
1206 the samples, the RNA was purified by adding Trizol and processing the samples through Directzol
1207 RNA mini-prep kit (Zymo Research). The eluted RNA from the input and IP samples were
1208 converted to cDNA using the High-capacity RNA-to-cDNA kit (Thermo Fisher Scientific). The
1209 reverse transcription was performed with and without adding the reverse transcriptase enzyme.
1210 qPCR was set up with primers targeting *POLQ* pre-mRNA, *POLQ* mature mRNA, and GAPDH
1211 in technical replicates for each assay. Ct values were normalized to that from the control samples. .
1212 All the sequences for RIP-qPCR used in this study were provided in Supplemental Table 1.
1213
1214 **Next generation sequencing**
1215 Library preparation and sequencing reactions were conducted at GENEWIZ, Lin/Azenta US, Inc.
1216 Deep RNA-seq (300 million reads/sample, N = 3 for each cell line) was performed for RPE *p53-*
1217 */-* (parental control) and *p53-/-EWSR1* knock-out cells. RNA samples were prepared using RNeasy
1218 Mini Kit and RNase-Free DNase Set (QIAGEN). RNA samples were quantified using Qubit 2.0
1219 Fluorometer, and the RNA integrity was checked with 4200 TapeStation. NEBNext Ultra II RNA
1220 Library Prep Kit for Illumina was used following the manufacturer's recommendations (New

1221 England BioLabs). Briefly, mRNAs were initially enriched with Oligo d(T) beads. Enriched
1222 mRNAs were fragmented for 15 minutes at 94°C. First strand and second strand cDNA were
1223 subsequently synthesized. cDNA fragments were end repaired and adenylated at 3´ends, and
1224 universal adapters were ligated to cDNA fragments, followed by index addition and library
1225 enrichment by PCR with limited cycles. The sequencing libraries were validated using Agilent
1226 TapeStation and quantified using Qubit 2.0 Fluorometer as well as by quantitative PCR (KAPA
1227 Biosystems). The sequencing libraries were clustered on flowcell lanes. After clustering, the
1228 flowcell was loaded on the Illumina NovaSeq instrument according to manufacturer's instructions.
1229 The samples were sequenced using a 2x150 Paired End configuration. Image analysis and base
1230 calling were conducted by the Control Software (NCS). Raw sequence data (.bcl files) generated
1231 from the Illumina instrument was converted into fastq files and de-multiplexed using Illumina's
1232 bcl2fastq (version 2.17) software. One mismatch was allowed for index sequence identification.
1233 Several bioinformatics analyses were performed on the RNA-seq data to identify differentially
1234 expressed genes, differentially used exons, and differentially spliced events between RPE *p53*-/-
1235 control cells and *p53-/-EWSR1* knock-out cells.
1236 For differential gene expression analysis, Kallisto (version 0.48.0) [66] was used to map the raw
1237 RNA-seq reads against the GENCODE human reference gene annotation (Release 44) [67] and
1238 generate raw transcript abundance estimates, and the *tximport* Bioconductor R package [68] was used
1239 to summarize the transcript abundance estimates to produce gene-level estimated counts.
1240 Differentially expressed genes between control and *EWSR1* knock-out cell samples (N = 3) were
1241 determined using both edgeR (version 3.40.2) [69] and DESeq2 (version 1.38.3) [70] Bioconductor
1242 packages. The high concordance between the estimated $\log_2$ foldchange of the two algorithms was
1243 verified.

1244 For both differential exon usage and differential splicing analyses, STAR (version 2.7.10b) [71] was
1245 used to map RNA-seq reads and generate BAM files. Specifically, the 2-pass STAR was used to
1246 map the raw RNA-seq reads against the GRCh38/hg38 human reference genome, with the
1247 GENCODE human reference gene annotation (Release 44) [67]. The first round of the 2-pass STAR
1248 was run to identify splice junctions, and more confidence junctions then were filtered by requiring
1249 them to have at least 5 total mapped reads. The second round of the 2-pass STAR was run using
1250 the filtered junctions to generate BAM files.
1251 For differential exon usage analysis, DEXSeq (version 1.46.0) [72] was used. In brief, a list of unique
1252 exonic regions was initially generated using the Python script *dexseq_prepare_annotation.py.* This
1253 script "collapsed" exonic regions from different transcripts to be used as exon counting bins. Next,
1254 the *dexseq_count.py* Python script was used to count reads for each sample. Finally, differentially
1255 used exons between control and *EWSR1* knock-out cells (N = 3) were determined using DEXSeq.
1256 For the differential splicing analysis, MISO (version 0.5.4) [26] and other tools were used as
1257 discussed below. The exon-centric splicing event annotations were generated using the *rnaseqlib*
1258 Python package (http://github.com/yarden/rnaseqlib). These annotations were derived
1259 considering all transcripts from Ensembl, UCSC RefSeq and NCBI RefSeq gene annotations,
1260 which were obtained from https://hgdownload.soe.ucsc.edu/goldenPath/hg38/database, in files
1261 *knownGene.txt*, *refGene.txt*, and *ncbiRefSeq.txt*. Each annotation event belongs to one of these 5
1262 types: skipped exons (SE), alternative 3′ splice sites (A3SS), alternative 5′ splice sites (A5SS),
1263 mutually exclusive exons (MXE), and retained introns (RI). These annotated events were linked
1264 to representative overlapping exons or introns of the GENCODE human reference gene
1265 annotation (Release 44) using the intersect function of the *bedtools* toolset (version 2.31.0). The
1266 representative exons and introns were prioritized with a high overlapping percentage (95%),

1267 followed by good transcript quality information including whether they were annotated as
1268 canonical and basic, and with better transcript support level. MISO was performed to quantify
1269 the isoforms and compute percent-spliced-in values (PSIs) of each alternative splicing event
1270 from the sorted BAM of each sample. Lower quality events were filtered out as follows. For each
1271 replicate sample pair, the *compare_miso* program was used to compare the results of the control
1272 and the *EWSR1* knock-out samples. The *filter_events* program was then used to filter the
1273 comparison results, by requiring each sample to have at least 35 total isoform-identifying reads.
1274 Finally, *limma* analysis (version 3.54.2) was performed on the PSI values of the filtered events to
1275 determine differential splicing events between control and *EWSR1* knock-out cells (N = 3).
1276 Significant differential alternative splicing events were called if they had *FDR* smaller than 0.05
1277 and absolute delta PSI of at least 0.05. Events that had significant FDR (<0.05) but small
1278 absolute delta PSI (<0.05) were likely to be artifacts and thus were excluded from *limma* results.
1279 Volcano plots were generated using ggplot2.
1280 To have more confidence in the MISO results, we also independently analyzed differential
1281 splicing using rMATS (version v4.2.0) [27]. Since SE events were the most frequent event type,
1282 we focused on analyzing the concordance of SE events between MISO and rMATS. First,
1283 rMATS was run using all sorted BAMs as inputs and the GENCODE human reference gene
1284 annotation (Release 44). Lower quality events were filtered out by requiring each sample to
1285 have at least 35 total isoform-identifying reads. Then, the shared events that were highly similar
1286 between MISO and rMATS were identified if their target regions overlapped by at least 95%.
1287 Since one MISO event could be linked to multiple slightly different rMATS events, for each
1288 MISO event, a representative rMATS event was chosen that had the largest absolute difference
1289 in inclusion level. Finally, we assessed the concordance between MISO and rMATS by

1290 comparing their inclusion differences (delta PSI values and differences in inclusion levels,
1291 respectively) from these shared events. We observed that a great majority of significant MISO
1292 events had a shared rMATS event. Furthermore, we observed that the inclusion differences of
1293 MISO and rMATS were positively correlated with $R^2$ of 0.32. The scatterplot comparing the
1294 inclusion differences of the two algorithms was generated using ggplot2. Sashimi plots were
1295 generated using the *rmats2sashimiplot* program (https://github.com/Xinglab/rmats2sashimiplot/).
1296

1297 **AlphaFold analysis**

1298 AlphaFold3 and the AlphaFold Server [73] were used to model full-length (POLQ aa 1824-2579)
1299 and truncated (POLQ aa 1824-2390) POLQ polymerases domains in complex with two
1300 complementary 9 bp DNA molecules (5'-CCAATGACA-3' and 5'-TGTCATTGG-3'). Models
1301 were visualized in Pymol V3.0.4 (Schrödinger) using the top-ranking predictions per seed. The
1302 full-length model was compared with a published crystal structure of the POLQ polymerase
1303 domain (PDB 6XBU) [74] to ensure its accuracy.
1304

1305 **Mouse xenograft study**

1306 All animal experiments were conducted in accordance with Institutional Animal Care and Use
1307 Committee-approved protocol (#08-036) at Dana-Farber Cancer Institute. TC32 cells were
1308 lentivirally transduced with luciferase by using a pLenti CMV Puro LUC (w168-1) vector. Two
1309 million TC32 cells expressing luciferase were subcutaneously injected into 7-week-old male and
1310 female NOD.Cg-*Prkdc*$^{scid}$ Il2rg$^{tm1Wjl}$/SzJ mice. After confirming tumor engraftment, mice were
1311 treated with Peposertib (100 mg/kg) or vehicle control once a day via oral administration together
1312 with Etoposide (8mg/kg) or vehicle control via intraperitoneal injection (IP injection) twice a week

1313 for 2 weeks. Tumors were measured every 2 to 3 days using an electronic caliper, and tumor
1314 volumes were calculated by using the formula L×W×W/2. For immunohistochemistry analysis,
1315 tumors were excised and fixed with formaldehyde. Mice with a tumor of more than 20 mm in
1316 length or width were euthanized.
1317

1318 **Data acquiring and processing of transcriptomic sequencing from patient tumor samples**

1319 Patient tumor sample bulk transcriptomic sequencing data were obtained from St. Jude's Cloud [75].
1320 For expression analysis, we used the mapped expression counts data; for the exon splicing data,
1321 we used the BAM files provided. For the expression counts data, all of the samples were uniformly
1322 processed and mapped to the same reference genome as part of the St. Jude's Cloud resource, and
1323 as such, no correction between samples or cancer types was performed. To evaluate the *POLQ*
1324 expression, we first computed the TPM (transcript-per-million) counts for all samples. In this
1325 computation, we omitted genes annotated as pseudogenes in GENCODE v.31, and used the
1326 median gene length as the gene length for each gene symbol.
1327

1328 **Computation of exon skipping rates in primary patient samples**

1329 To compute the proportions of reads with exons skipped in the sequencing reads data obtained
1330 from St. Jude's Cloud [76], rMATS-turbo, a computational tool to quantify alternative splicing
1331 patterns from RNA sequencing data, was applied. To compute the proportion of reads that each
1332 exon was skipped in, the rMATS-turbo results inclusive of supporting reads from counted from
1333 both junction and exon reads were used. Of note, for each individual exon, any sample that had no
1334 supporting reads for both the inclusive and skipped cases was omitted. All comparisons between
1335 distributions of skipped read proportions were performed using Mann-Whitney U tests.

1336

1337 **Mutational Analysis of Tumors**

1338 Mutational analysis was performed on the whole genome sequencing (WGS) data of Ewing
1339 sarcoma and neuroblastoma tumors obtained from St. Jude Cloud (https://www.stjude.cloud/).
1340 The samples include 23 Ewing sarcoma (ES) and 48 neuroblastoma (NB) tumor and matched-
1341 normal pairs. Somatic variants were determined as follows. First, for each cohort, short variants
1342 were called from the HG38 Genomic Variant Call Format (gVCF) data using the GATK joint
1343 genotyping workflow (version 4.2.4). Second, bcftools (version 1.14) was used to filter out low-
1344 quality variants with these hard filters: QD < 7.5, MQ < 59.9, -g3, and -G10. Only biallelic SNPs
1345 and INDELs were selected. Third, cohort-level singleton variants were selected if there was only
1346 one sample with read depth (DP) between 17 and 300, and variant allele fraction (VAF) of at
1347 least 0.1. And finally, a variant was considered somatic if both the tumor and matched-normal
1348 had SOR <= 3, MQ > 59.9, and DP > 17, and either, tumor VAF >= 0.175 and matched normal
1349 VAF = 0, or tumor VAF <= 0.825 and matched normal VAF = 1.
1350 Profiles of SNV mutations and INDELs were generated using SigProfilerMatrixGenerator
1351 (version 1.2.13; https://github.com/AlexandrovLab/SigProfilerMatrixGenerator). Mutational
1352 signature analysis was performed using SigProfilerAssignment (version 0.0.21;
1353 https://github.com/AlexandrovLab/SigProfilerAssignment). This involved fitting the mutational
1354 profiles to the COSMIC (https://cancer.sanger.ac.uk/signatures/) SNV signatures (SBS version 2)
1355 and INDEL signatures (ID version 3.3).

1356 The MMEJ signature was characterized by the presence of microhomologies at the
1357 boundaries of deletions and the presence of nearby template sequences in insertions, or templated
1358 insertions [77,78]. The presence of microhomologies in deletions was also independently determined

1359 using the method described by Taheri-Ghahfarokhi *et al* [77,78]. This method essentially finds the
1360 longest possible complementary subsequence between one end of the deleted sequence and the
1361 flanking sequence at the other end. In brief, for each deletion, 50-bp flanking sequences around
1362 the deleted sequence were first retrieved from the reference human genome (HG38) using
1363 bedtools (version 2.27.1; https://bedtools.readthedocs.io/). The deletion was then repeatedly
1364 shifted leftward by one base whenever its rightmost base matched the reference base
1365 immediately to its left. Finally, the microhomology sequence was the longest matching
1366 subsequence between the deleted sequence and the right flanking sequence starting from their
1367 leftmost bases. Deletions with length of at least 3bp with GATK ReadPosRankSum of greater
1368 than -5 and with microhomology of at least 2 bp were designated as MMEJ events [77]. For
1369 comparison between ES and NB, tumors with at least 15 deletions and the proportions of MMEJ
1370 deletions among all indels were used. Local templated insertions were characterized by the
1371 presence of subsequences in the inserted sequence that were identical to the flanking sequences
1372 (templated inserts) [15,79,80]. The templated inserts were detected by aligning the inserted sequence
1373 with the 50bp-flanking sequences around the insertion site using the Smith–Waterman local
1374 alignment algorithm [81]. The analysis was restricted to insertions with GATK ReadPosRankSum
1375 of greater than -5 and with length of at least 5bp. The insertions with templates adjacent to the
1376 insertion position were excluded. The remaining insertions with templates covering at least 60%
1377 of the inserted sequence were designated as MMEJ events. For comparison between ES and NB,
1378 tumors with at least 2 insertions, and the proportions of MMEJ templated insertions among all
1379 insertions were used.
1380
1381 **Immunohistochemistry (IHC)**

1382 Primary patient samples for IHC analysis were obtained from the Mayo Clinic (IRB approval #24-
1383 007702). IHC was performed using a Parhelia autostainer. Slides were deparaffinized using Slide
1384 Brite xylene substitute (23401-01) followed by rehydration in ethyl alcohol. Endogenous blocking
1385 was performed using methanol hydrogen peroxide solution. Slides were heated for 30 minutes in
1386 citrate solution for epitope retrieval. Slides were incubated for 30minutes in anti-DNAPKcs
1387 phosphor S2056 antibody at 1:150 followed by washing in wash buffer (Dako S3006) and
1388 inbubation for 30minutes in Mach 3 rabbit HRP (Buocare Medical RP531L). Primary staining was
1389 performed by incubation in DAB (Dako K3468) for 10minutes. Counterstaing was performed
1390 using hematoxylin (Sigma-Aldrich MHS16) and bluing buffer (Dako C5702). Slides were then
1391 dehydrated in ethyl alcohol before coverslipping using Permount mounting mediun (Fisher
1392 Scientific SP15-100). Slides were digitally imaged using a Motic Easyscan Pro 6.
1393 Histological image analysis was performed using QuPath (version 0.5.1). Whole-slide images
1394 (WSIs) were imported into QuPath, and automated color deconvolution was applied to separate
1395 the hematoxylin and DAB (3,3′-diaminobenzidine) staining channels. The mean optica l density
1396 (OD) of the DAB-positive areas was measured to assess staining intensity. For patients with
1397 multiple tissue sections, and the mean OD across all sections for the same patient was used to
1398 represent the patient's staining intensity.
1399
1400 **Statistical analysis**
1401 Data were analyzed and visualized using GraphPad Prism (Version 10.2.2, GraphPad Software,
1402 LLC). For comparisons between two groups, either a two-tailed Student's t-test or a Mann-
1403 Whitney U test was used. For comparisons involving more than two groups, one-way analysis of

variance (ANOVA) was performed. For *in vivo* survival analysis study, a log-rank test was performed. A p-value of less than 0.05 ($p < 0.05$) was considered statistically significant.

1415

1416 **REFERENCES**


1417 1 Riggi, N., Suva, M. L. & Stamenkovic, I. Ewing's Sarcoma. *N Engl J Med* **384**, 154-164
1418 (2021). https://doi.org/10.1056/NEJMra2028910
1419 2 Grunewald, T. G. P. *et al.* Ewing sarcoma. *Nat Rev Dis Primers* **4**, 5 (2018).
1420 https://doi.org/10.1038/s41572-018-0003-x
1421 3 Gorthi, A. *et al.* EWS-FLI1 increases transcription to cause R-loops and block BRCA1
1422 repair in Ewing sarcoma. *Nature* **555**, 387-391 (2018).
1423 https://doi.org/10.1038/nature25748
1424 4 Choy, E. *et al.* Phase II study of olaparib in patients with refractory Ewing sarcoma
1425 following failure of standard chemotherapy. *BMC Cancer* **14**, 813 (2014).
1426 https://doi.org/10.1186/1471-2407-14-813
1427 5 Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like
1428 Domain. *Cell* **171**, 163-178 e119 (2017). https://doi.org/10.1016/j.cell.2017.07.036
1429 6 Davis, R. B., Supakar, A., Ranganath, A. K., Moosa, M. M. & Banerjee, P. R.
1430 Heterotypic interactions can drive selective co-condensation of prion-like low-complexity
1431 domains of FET proteins and mammalian SWI/SNF complex. *Nat Commun* **15**, 1168
1432 (2024). https://doi.org/10.1038/s41467-024-44945-5
1433 7 Chong, S. *et al.* Tuning levels of low-complexity domain interactions to modulate
1434 endogenous oncogenic transcription. *Mol Cell* **82**, 2084-2097 e2085 (2022).
1435 https://doi.org/10.1016/j.molcel.2022.04.007
1436 8 Zuo, L. *et al.* Loci-specific phase separation of FET fusion oncoproteins promotes gene
1437 transcription. *Nat Commun* **12**, 1491 (2021). https://doi.org/10.1038/s41467-021-21690-7
1438 9 Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that
1439 control gene transcription. *Science* **361** (2018). https://doi.org/10.1126/science.aar2555
1440 10 Selvanathan, S. P. *et al.* EWS-FLI1 modulated alternative splicing of ARID1A reveals
1441 novel oncogenic function through the BAF complex. *Nucleic Acids Res* **47**, 9619-9636
1442 (2019). https://doi.org/10.1093/nar/gkz699
1443 11 Wang, Z. *et al.* DNA polymerase theta (POLQ) is important for repair of DNA double-
1444 strand breaks caused by fork collapse. *J Biol Chem* **294**, 3909-3919 (2019).
1445 https://doi.org/10.1074/jbc.RA118.005188
1446 12 Goff, J. P. *et al.* Lack of DNA polymerase theta (POLQ) radiosensitizes bone marrow
1447 stromal cells in vitro and increases reticulocyte micronuclei after total-body irradiation.
1448 *Radiat Res* **172**, 165-174 (2009). https://doi.org/10.1667/RR1598.1
1449 13 Mengwasser, K. E. *et al.* Genetic Screens Reveal FEN1 and APEX2 as BRCA2 Synthetic
1450 Lethal Targets. *Mol Cell* **73**, 885-899 e886 (2019).
1451 https://doi.org/10.1016/j.molcel.2018.12.008
1452 14 Adam, S. *et al.* The CIP2A-TOPBP1 axis safeguards chromosome stability and is a
1453 synthetic lethal target for BRCA-mutated cancer. *Nat Cancer* **2**, 1357-1371 (2021).
1454 https://doi.org/10.1038/s43018-021-00266-w
1455 15 Chan, S. H., Yu, A. M. & McVey, M. Dual roles for DNA polymerase theta in alternative
1456 end-joining repair of double-strand breaks in Drosophila. *PLoS Genet* **6**, e1001005
1457 (2010). https://doi.org/10.1371/journal.pgen.1001005
1458 16 Della-Maria, J. *et al.* Human Mre11/human Rad50/Nbs1 and DNA ligase
1459 IIIalpha/XRCC1 protein complexes act together in an alternative nonhomologous end

1460 joining pathway. *J Biol Chem* **286**, 33845-33853 (2011).
1461 https://doi.org/10.1074/jbc.M111.274159
1462 17 Fleury, H. *et al.* The APE2 nuclease is essential for DNA double-strand break repair by
1463 microhomology-mediated end joining. *Mol Cell* **83**, 1429-1445 e1428 (2023).
1464 https://doi.org/10.1016/j.molcel.2023.03.017
1465 18 Brambati, A. *et al.* RHINO directs MMEJ to repair DNA breaks in mitosis. *Science* **381**,
1466 653-660 (2023). https://doi.org/10.1126/science.adh3694
1467 19 Hussmann, J. A. *et al.* Mapping the genetic landscape of DNA double-strand break
1468 repair. *Cell* **184**, 5653-5669 e5625 (2021). https://doi.org/10.1016/j.cell.2021.10.002
1469 20 Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous
1470 Recombination share the initial end resection step to repair DNA double-strand breaks in
1471 mammalian cells. *Proc Natl Acad Sci U S A* **110**, 7720-7725 (2013).
1472 https://doi.org/10.1073/pnas.1213431110
1473 21 Pierce, A. J., Johnson, R. D., Thompson, L. H. & Jasin, M. XRCC3 promotes homology-
1474 directed repair of DNA damage in mammalian cells. *Genes Dev* **13**, 2633-2638 (1999).
1475 https://doi.org/10.1101/gad.13.20.2633
1476 22 Zhou, J. *et al.* A first-in-class Polymerase Theta Inhibitor selectively targets
1477 Homologous-Recombination-Deficient Tumors. *Nat Cancer* **2**, 598-610 (2021).
1478 https://doi.org/10.1038/s43018-021-00203-x
1479 23 Nabet, B. *et al.* The dTAG system for immediate and target-specific protein degradation.
1480 *Nat Chem Biol* **14**, 431-441 (2018). https://doi.org/10.1038/s41589-018-0021-8
1481 24 Nabet, B. *et al.* Rapid and direct control of target protein levels with VHL-recruiting
1482 dTAG molecules. *Nat Commun* **11**, 4687 (2020). https://doi.org/10.1038/s41467-020-
1483 18377-w
1484 25 Johnson, C. N. *et al.* Insights into Molecular Diversity within the FET Family:
1485 Unraveling Phase Separation of the N-Terminal Low Complexity Domain from RNA-
1486 Binding Protein EWS. *bioRxiv* (2023). https://doi.org/10.1101/2023.10.27.564484
1487 26 Katz, Y., Wang, E. T., Airoldi, E. M. & Burge, C. B. Analysis and design of RNA
1488 sequencing experiments for identifying isoform regulation. *Nat Methods* **7**, 1009-1015
1489 (2010). https://doi.org/10.1038/nmeth.1528
1490 27 Shen, S. *et al.* rMATS: robust and flexible detection of differential alternative splicing
1491 from replicate RNA-Seq data. *Proc Natl Acad Sci U S A* **111**, E5593-5601 (2014).
1492 https://doi.org/10.1073/pnas.1419161111
1493 28 Harada, N. *et al.* The splicing factor CCAR1 regulates the Fanconi anemia/BRCA
1494 pathway. *Mol Cell* **84**, 2618-2633 e2610 (2024).
1495 https://doi.org/10.1016/j.molcel.2024.06.031
1496 29 Nicholson-Shaw, A. L., Kofman, E. R., Yeo, G. W. & Pasquinelli, A. E. Nuclear and
1497 cytoplasmic poly(A) binding proteins (PABPs) favor distinct transcripts and isoforms.
1498 *Nucleic Acids Res* **50**, 4685-4702 (2022). https://doi.org/10.1093/nar/gkac263
1499 30 Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-
1500 associated genes. *Nature* **499**, 214-218 (2013). https://doi.org/10.1038/nature12213
1501 31 Thatikonda, V. *et al.* Comprehensive analysis of mutational signatures reveals distinct
1502 patterns and molecular processes across 27 pediatric cancers. *Nat Cancer* **4**, 276-289
1503 (2023). https://doi.org/10.1038/s43018-022-00509-4

32 Schimmel, J., Kool, H., van Schendel, R. & Tijsterman, M. Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells. *EMBO J* **36**, 3634-3649 (2017). https://doi.org/10.15252/embj.201796948

33 Johnson, C. N. *et al.* Insights into Molecular Diversity within the FUS/EWS/TAF15 Protein Family: Unraveling Phase Separation of the N-Terminal Low-Complexity Domain from RNA-Binding Protein EWS. *J Am Chem Soc* **146**, 8071-8085 (2024). https://doi.org/10.1021/jacs.3c12034

34 Ebersberger, S. *et al.* FUBP1 is a general splicing factor facilitating 3' splice site recognition and splicing of long introns. *Mol Cell* **83**, 2653-2672 e2615 (2023). https://doi.org/10.1016/j.molcel.2023.07.002

35 Graham, G. T. *et al.* Comprehensive profiling of mRNA splicing indicates that GC content signals altered cassette exon inclusion in Ewing sarcoma. *NAR Cancer* **4**, zcab052 (2022). https://doi.org/10.1093/narcan/zcab052

36 Selvanathan, S. P. *et al.* Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing. *Proc Natl Acad Sci U S A* **112**, E1307-1316 (2015). https://doi.org/10.1073/pnas.1500536112

37 Zatreanu, D. *et al.* Poltheta inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. *Nat Commun* **12**, 3636 (2021). https://doi.org/10.1038/s41467-021-23463-8

38 Llorens-Agost, M. *et al.* POLtheta-mediated end joining is restricted by RAD52 and BRCA2 until the onset of mitosis. *Nat Cell Biol* **23**, 1095-1104 (2021). https://doi.org/10.1038/s41556-021-00764-0

39 Gelot, C. *et al.* Poltheta is phosphorylated by PLK1 to repair double-strand breaks in mitosis. *Nature* **621**, 415-422 (2023). https://doi.org/10.1038/s41586-023-06506-6

40 Kumar, R. J. *et al.* Dual inhibition of DNA-PK and DNA polymerase theta overcomes radiation resistance induced by p53 deficiency. *NAR Cancer* **2**, zcaa038 (2020). https://doi.org/10.1093/narcan/zcaa038

41 Patterson-Fortin, J. *et al.* Targeting DNA Repair with Combined Inhibition of NHEJ and MMEJ Induces Synthetic Lethality in TP53-Mutant Cancers. *Cancer Res* **82**, 3815-3829 (2022). https://doi.org/10.1158/0008-5472.CAN-22-1124

42 Collins, V. J. *et al.* Enhancing standard of care chemotherapy efficacy using DNA-dependent protein kinase (DNA-PK) inhibition in preclinical models of Ewing sarcoma. *Mol Cancer Ther* (2024). https://doi.org/10.1158/1535-7163.MCT-23-0641

43 Dubbury, S. J., Boutz, P. L. & Sharp, P. A. CDK12 regulates DNA repair genes by suppressing intronic polyadenylation. *Nature* **564**, 141-145 (2018). https://doi.org/10.1038/s41586-018-0758-y

44 Iniguez, A. B. *et al.* EWS/FLI Confers Tumor Cell Synthetic Lethality to CDK12 Inhibition in Ewing Sarcoma. *Cancer Cell* **33**, 202-216 e206 (2018). https://doi.org/10.1016/j.ccell.2017.12.009

45 Gounder, M. M. *et al.* Clinical genomic profiling in the management of patients with soft tissue and bone sarcoma. *Nat Commun* **13**, 3406 (2022). https://doi.org/10.1038/s41467-022-30496-0

46 Tirode, F. *et al.* Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov* **4**, 1342-1353 (2014). https://doi.org/10.1158/2159-8290.CD-14-0622

47 Dutertre, M. *et al.* Cotranscriptional exon skipping in the genotoxic stress response. *Nat Struct Mol Biol* **17**, 1358-1366 (2010). https://doi.org/10.1038/nsmb.1912

48 Paronetto, M. P., Minana, B. & Valcarcel, J. The Ewing sarcoma protein regulates DNA damage-induced alternative splicing. *Mol Cell* **43**, 353-368 (2011). https://doi.org/10.1016/j.molcel.2011.05.035

49 Verdile, V. *et al.* EWS splicing regulation contributes to balancing Foxp1 isoforms required for neuronal differentiation. *Nucleic Acids Res* **50**, 3362-3378 (2022). https://doi.org/10.1093/nar/gkac154

50 Maurer, L. M. *et al.* BRCA1-associated RING domain-1 (BARD1) loss and GBP1 expression enhance sensitivity to DNA damage in Ewing sarcoma. *Cancer Res Commun* **2**, 220-232 (2022). https://doi.org/10.1158/2767-9764.crc-21-0047

51 Menon, S. *et al.* FET fusion oncoproteins disrupt physiologic DNA repair networks and induce ATR synthetic lethality in cancer. *Res Sq* (2023). https://doi.org/10.21203/rs.3.rs-2869150/v1

52 Ceccaldi, R. *et al.* Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair. *Nature* **518**, 258-262 (2015). https://doi.org/10.1038/nature14184

53 Mateos-Gomez, P. A. *et al.* Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature* **518**, 254-257 (2015). https://doi.org/10.1038/nature14157

54 Bhat, P. *et al.* Genome organization around nuclear speckles drives mRNA splicing efficiency. *Nature* **629**, 1165-1173 (2024). https://doi.org/10.1038/s41586-024-07429-6

55 Wang, J., Schultz, P. G. & Johnson, K. A. Mechanistic studies of a small-molecule modulator of SMN2 splicing. *Proc Natl Acad Sci U S A* **115**, E4604-E4612 (2018). https://doi.org/10.1073/pnas.1800260115

56 Duncan, R., Collins, I., Tomonaga, T., Zhang, T. & Levens, D. A unique transactivation sequence motif is found in the carboxyl-terminal domain of the single-strand-binding protein FBP. *Mol Cell Biol* **16**, 2274-2282 (1996). https://doi.org/10.1128/MCB.16.5.2274

57 Farag, M. *et al.* Condensates formed by prion-like low-complexity domains have small-world network structures and interfaces defined by expanded conformations. *Nat Commun* **13**, 7722 (2022). https://doi.org/10.1038/s41467-022-35370-7

58 Wang, J. *et al.* A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688-699 e616 (2018). https://doi.org/10.1016/j.cell.2018.06.006

59 Seong, B. K. A. *et al.* TRIM8 modulates the EWS/FLI oncoprotein to promote survival in Ewing sarcoma. *Cancer Cell* **39**, 1262-1278 e1267 (2021). https://doi.org/10.1016/j.ccell.2021.07.003

60 Chaudhary, U. B. & Haldas, J. R. Long-term complications of chemotherapy for germ cell tumours. *Drugs* **63**, 1565-1577 (2003). https://doi.org/10.2165/00003495-200363150-00004

61 Lim, K. S. *et al.* USP1 Is Required for Replication Fork Protection in BRCA1-Deficient Tumors. *Mol Cell* **72**, 925-941 e924 (2018). https://doi.org/10.1016/j.molcel.2018.10.045

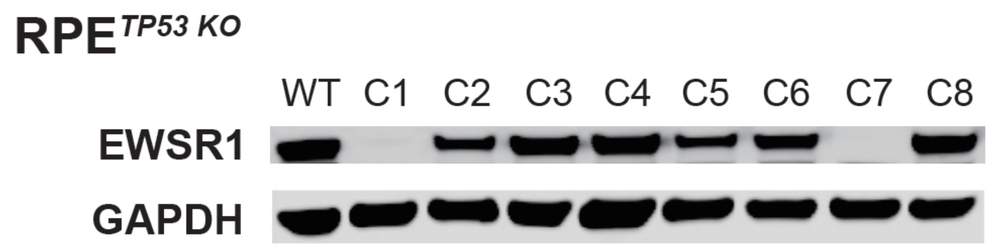
62 Hayer, A. *et al.* Engulfed cadherin fingers are polarized junctional structures between collectively migrating endothelial cells. *Nat Cell Biol* **18**, 1311-1323 (2016). https://doi.org/10.1038/ncb3438

63 Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783-784 (2014). https://doi.org/10.1038/nmeth.3047
64 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016). https://doi.org/10.1038/nmeth.3901
65 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015). https://doi.org/10.1093/nar/gkv007
66 Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525-527 (2016). https://doi.org/10.1038/nbt.3519
67 Frankish, A. *et al.* Gencode 2021. *Nucleic Acids Res* **49**, D916-D923 (2021). https://doi.org/10.1093/nar/gkaa1087
68 Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015). https://doi.org/10.12688/f1000research.7563.2
69 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010). https://doi.org/10.1093/bioinformatics/btp616
70 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014). https://doi.org/10.1186/s13059-014-0550-8
71 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013). https://doi.org/10.1093/bioinformatics/bts635
72 Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res* **22**, 2008-2017 (2012). https://doi.org/10.1101/gr.133744.111
73 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024). https://doi.org/10.1038/s41586-024-07487-w
74 Chandramouly, G. *et al.* Poltheta reverse transcribes RNA and promotes RNA-templated DNA repair. *Sci Adv* **7** (2021). https://doi.org/10.1126/sciadv.abf1771
75 McLeod, C. *et al.* St. Jude Cloud: A Pediatric Cancer Genomic Data-Sharing Ecosystem. *Cancer Discov* **11**, 1082-1099 (2021). https://doi.org/10.1158/2159-8290.CD-20-1230
76 Wang, Y. *et al.* rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. *Nat Protoc* **19**, 1083-1104 (2024). https://doi.org/10.1038/s41596-023-00944-2
77 Tobalina, L., Armenia, J., Irving, E., O'Connor, M. J. & Forment, J. V. A meta-analysis of reversion mutations in BRCA genes identifies signatures of DNA end-joining repair mechanisms driving therapy resistance. *Ann Oncol* **32**, 103-112 (2021). https://doi.org/10.1016/j.annonc.2020.10.470
78 Taheri-Ghahfarokhi, A. *et al.* Decoding non-random mutational signatures at Cas9 targeted sites. *Nucleic Acids Res* **46**, 8417-8434 (2018). https://doi.org/10.1093/nar/gky653
79 van Schendel, R., Roerink, S. F., Portegijs, V., van den Heuvel, S. & Tijsterman, M. Polymerase Theta is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. *Nat Commun* **6**, 7394 (2015). https://doi.org/10.1038/ncomms8394
80 Yu, A. M. & McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res* **38**, 5706-5717 (2010). https://doi.org/10.1093/nar/gkq379

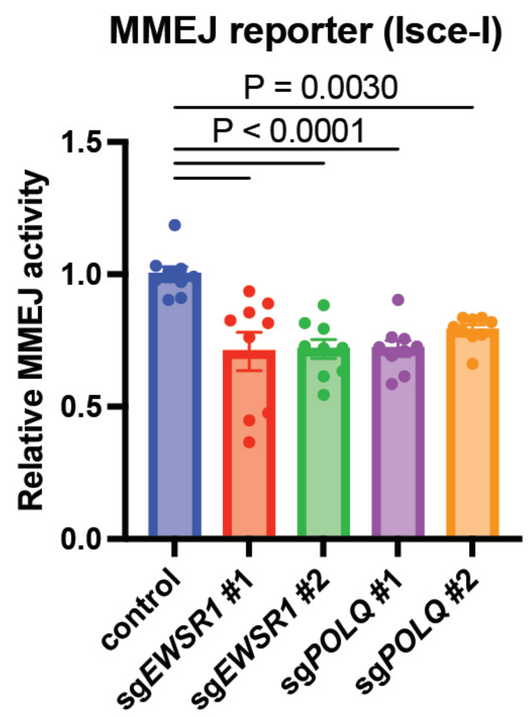
81 Smith, T. F. & Waterman, M. S. Identification of common molecular subsequences. *J Mol Biol* **147**, 195-197 (1981). https://doi.org/10.1016/0022-2836(81)90087-5

# Figure 1

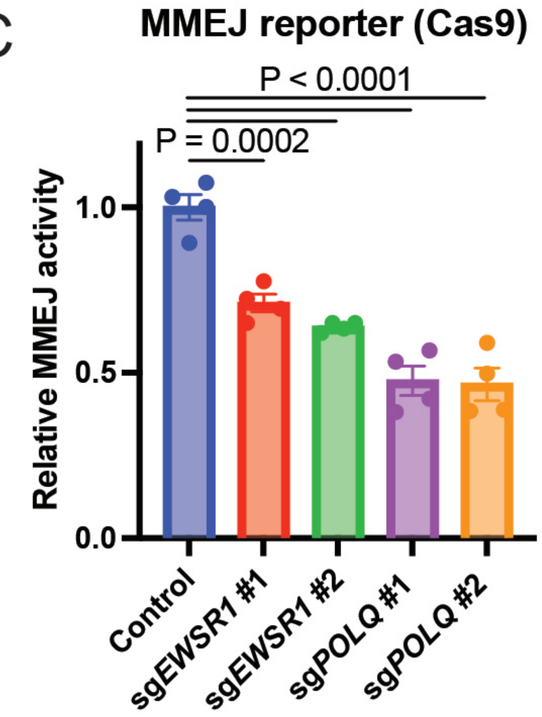
A



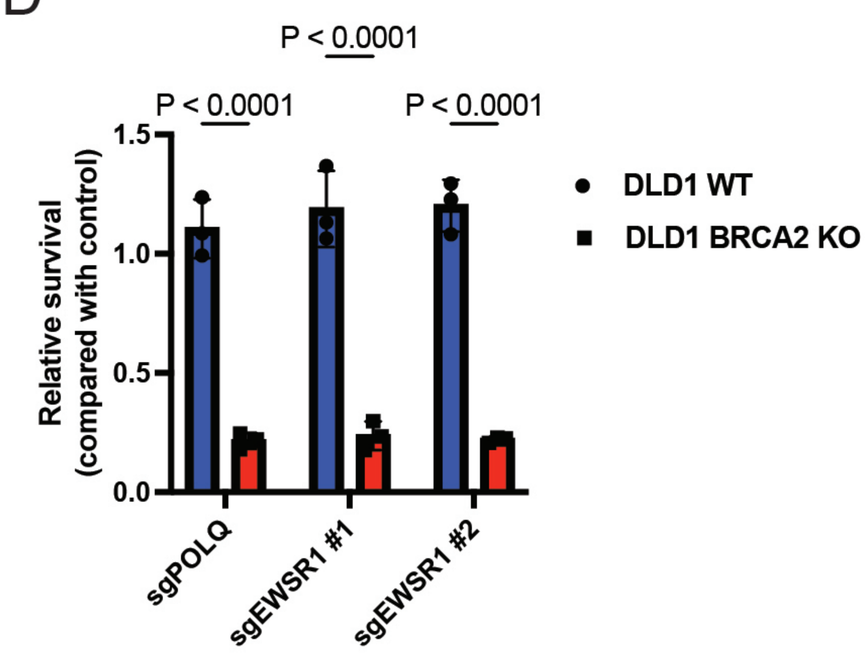

B

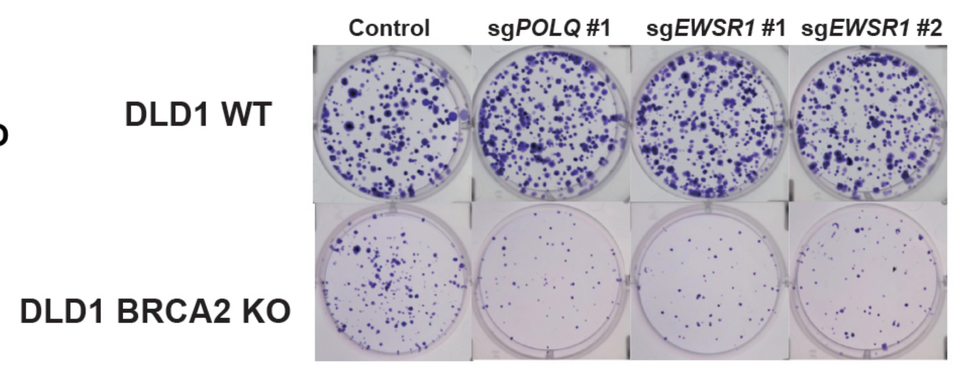



C

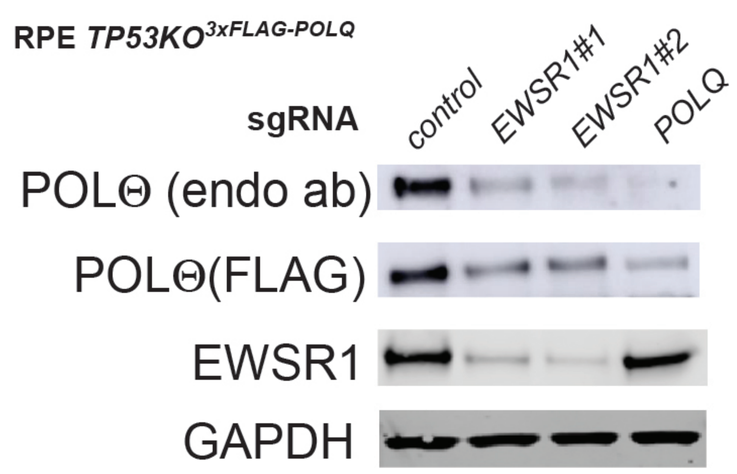



D

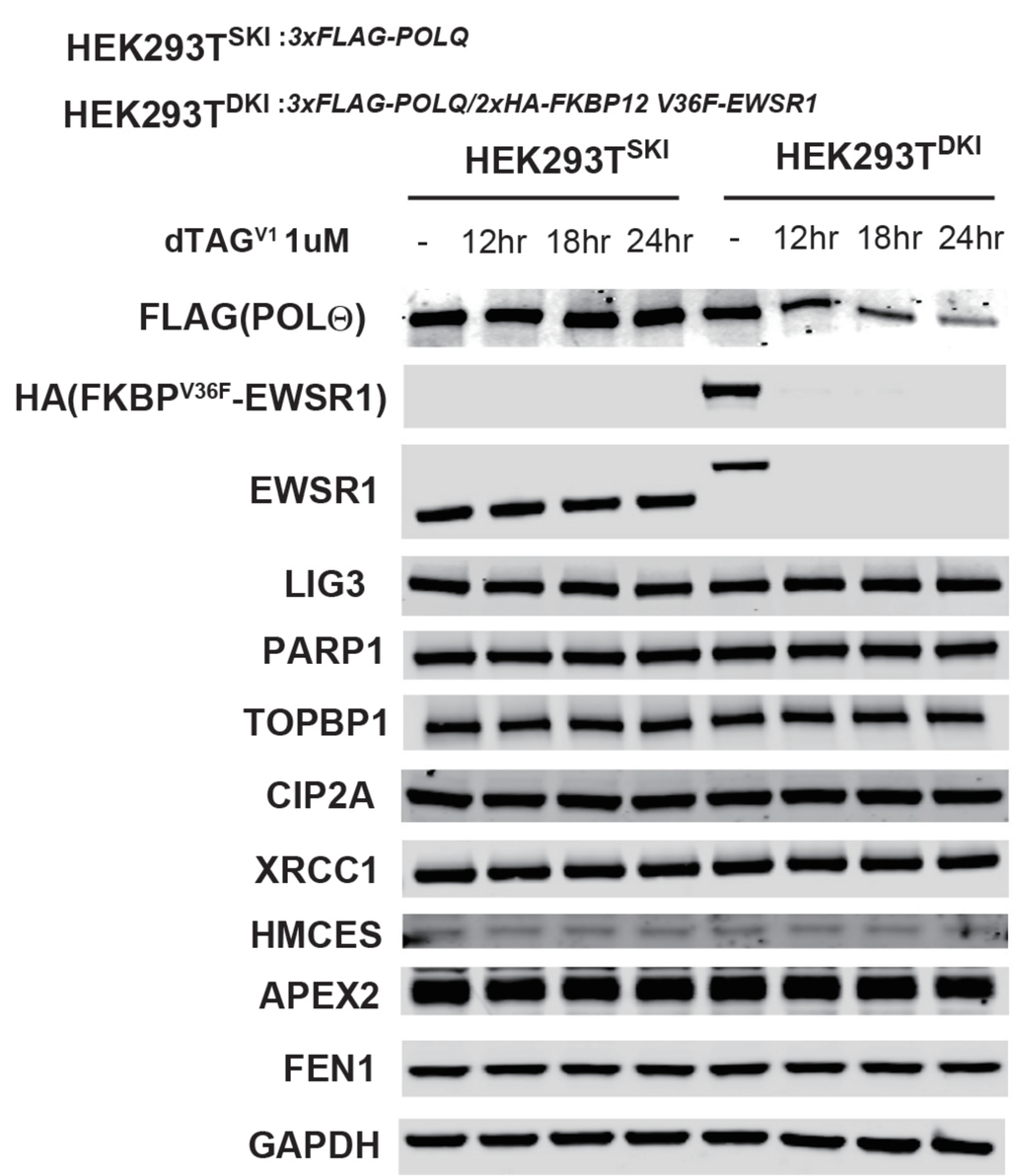


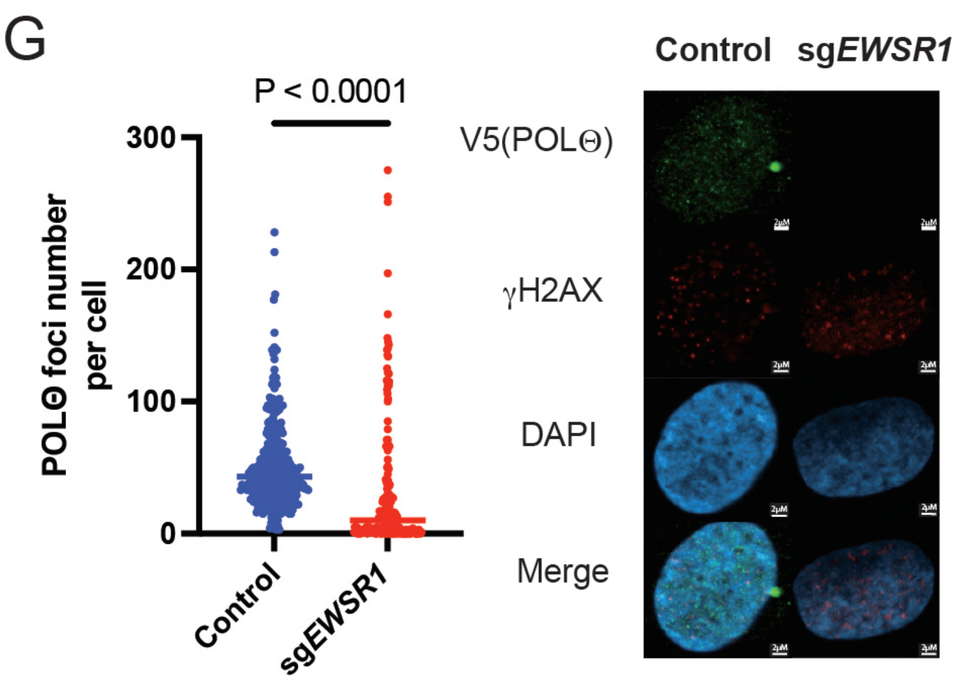

E



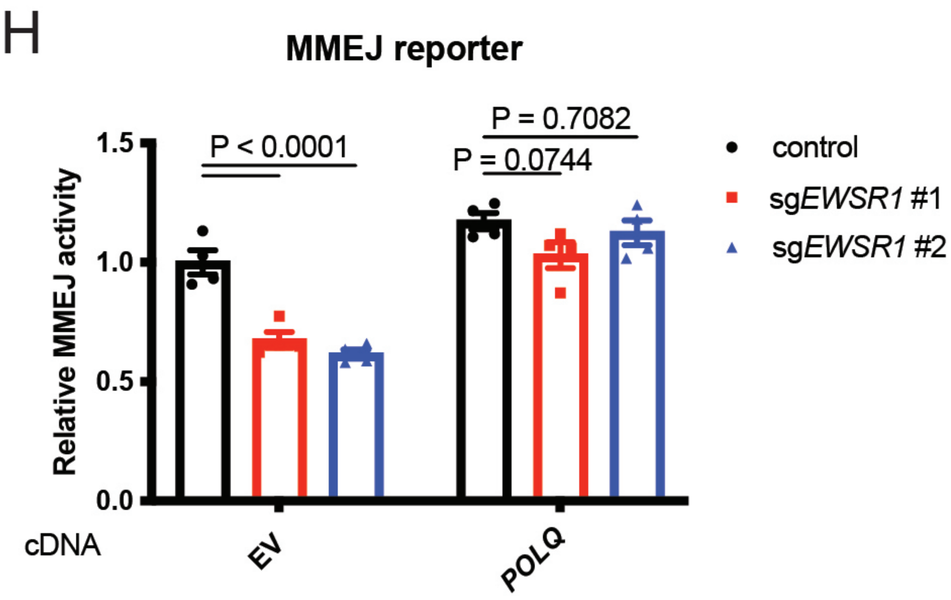

F



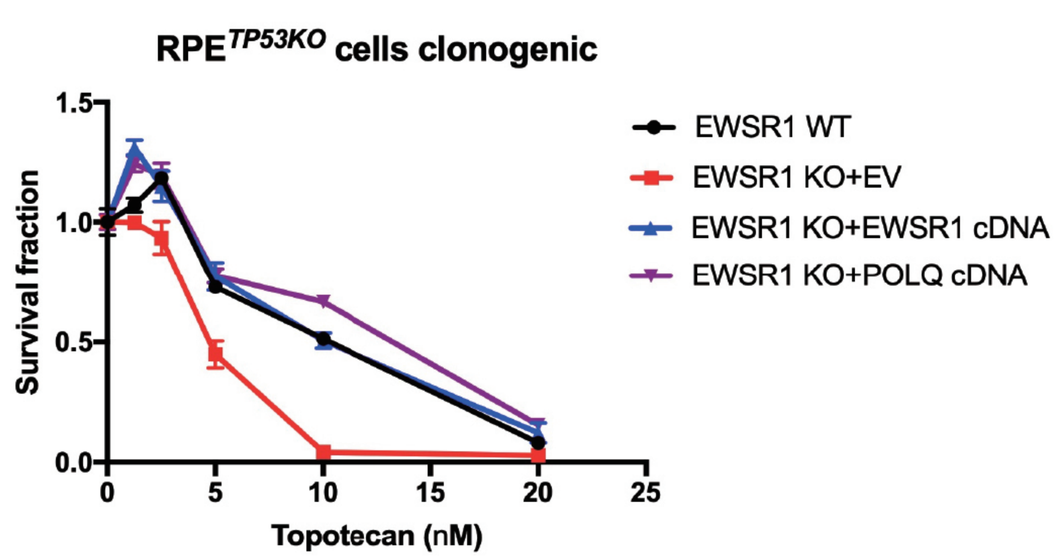

G



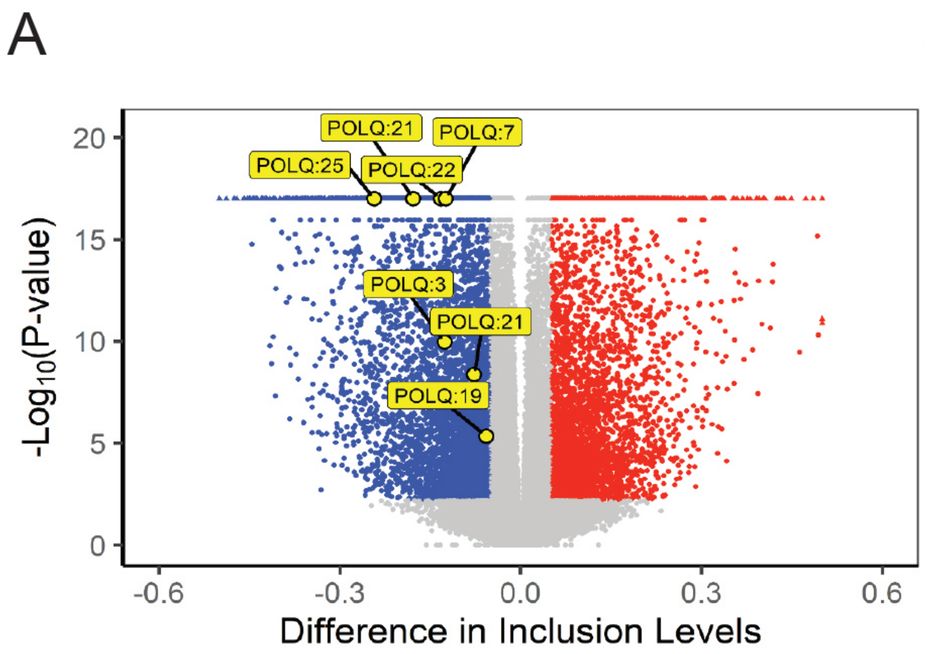

H

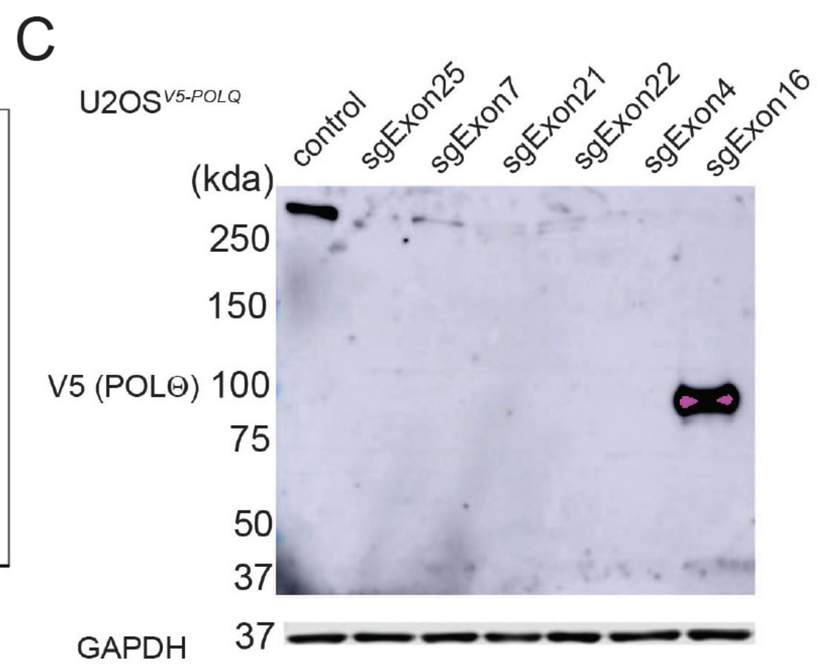


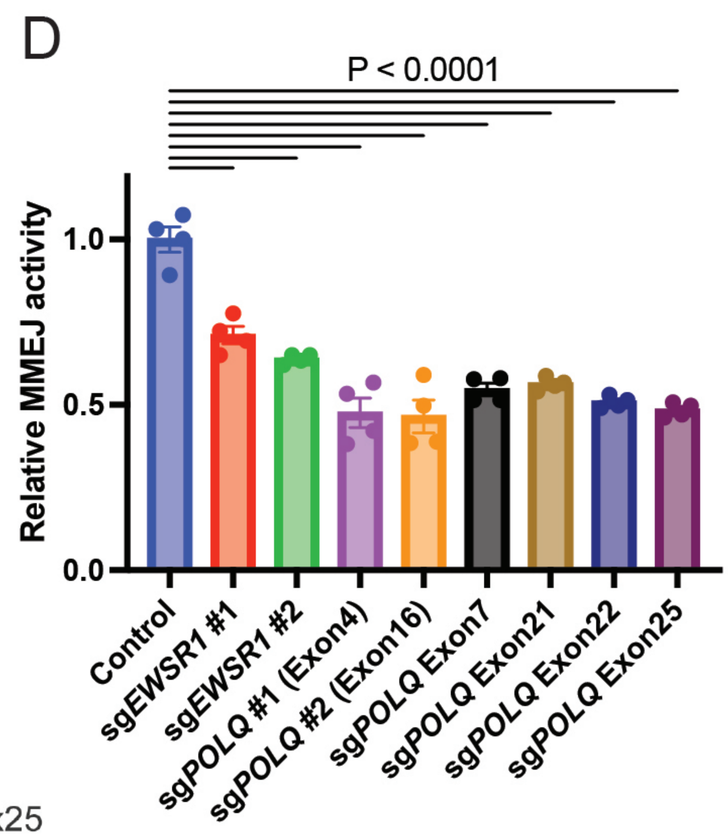

I

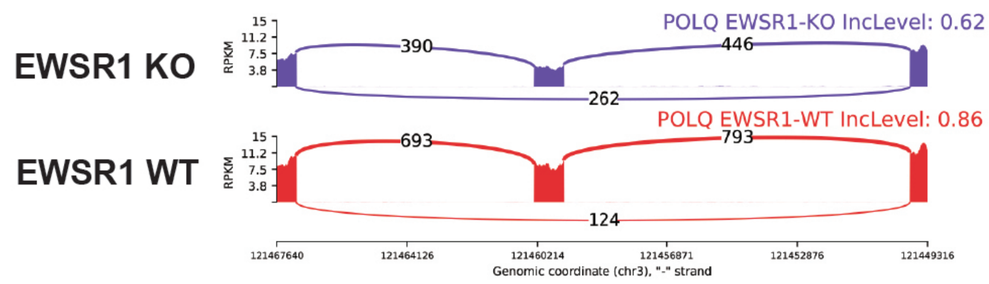


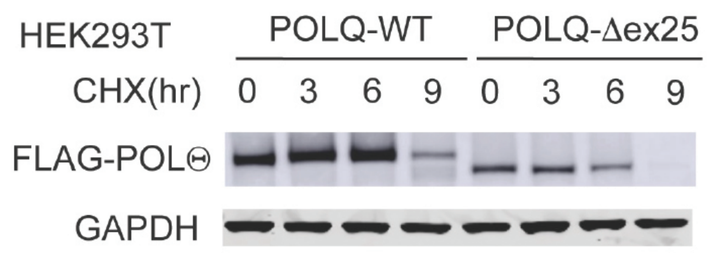
# Figure 2

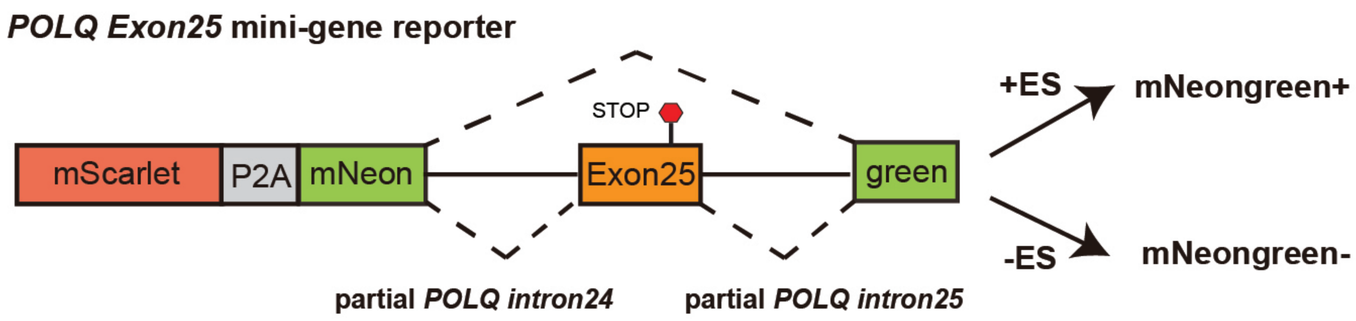

A
POLQ:21
POLQ:7
POLQ:25
POLQ:22
POLQ:3
POLQ:21
POLQ:19
-Log10(P-value)
Difference in Inclusion Levels

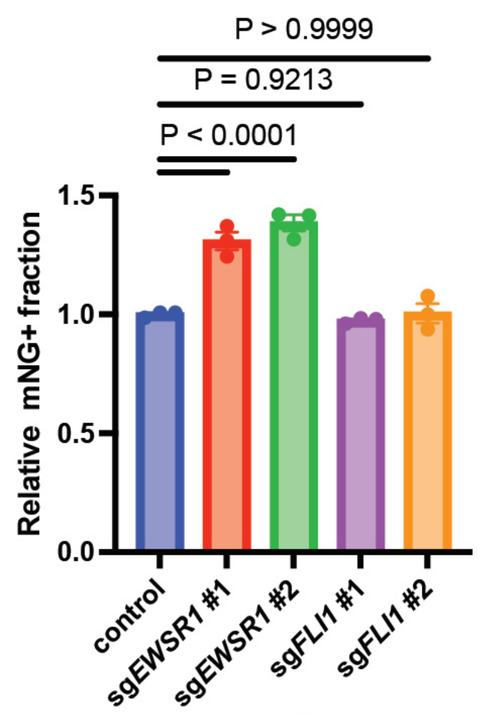

C
U2OS V5-POLQ
control
sgExon25
sgExon7
sgExon21
sgExon22
sgExon4
sgExon16
(kda)
250
150
V5 (POLΘ) 100
75
50
37
GAPDH 37

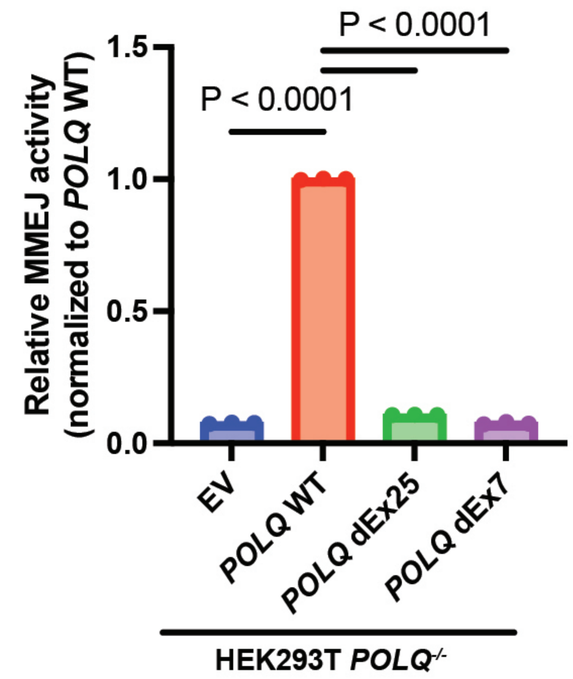

D
P < 0.0001
Relative MMEJ activity
Control
sgEWSR1 #1
sgEWSR1 #2
sgPOLQ #1 (Exon4)
sgPOLQ #2 (Exon16)
sgPOLQ Exon7
sgPOLQ Exon21
sgPOLQ Exon22
sgPOLQ Exon25

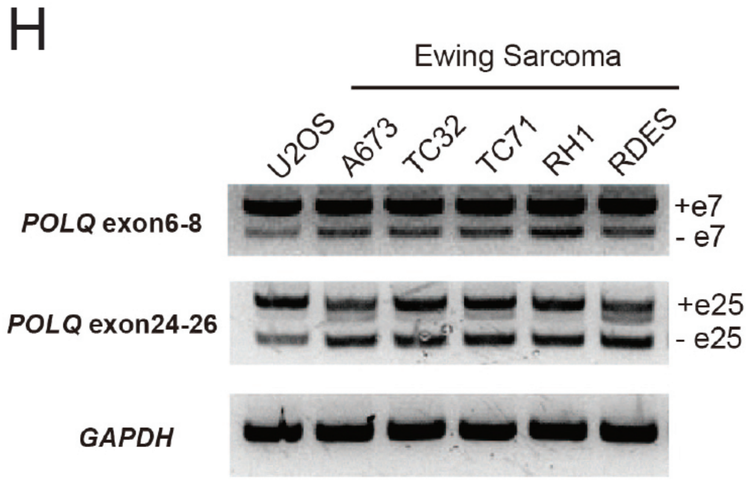

B
EWSR1 KO
POLQ EWSR1-KO IncLevel: 0.62
390
446
262
EWSR1 WT
POLQ EWSR1-WT IncLevel: 0.86
693
793
124
Genomic coordinate (chr3), "-" strand

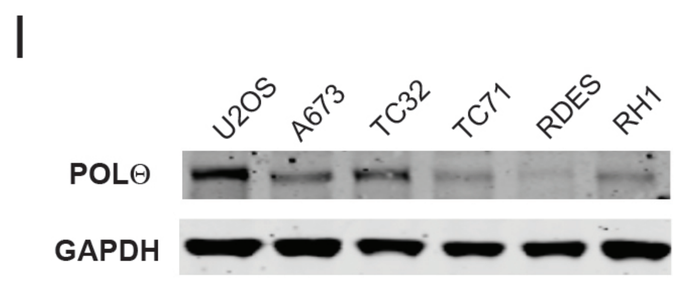

E
HEK293T
POLQ-WT
POLQ-Δex25
CHX(hr) 0 3 6 9 0 3 6 9
FLAG-POLΘ
GAPDH

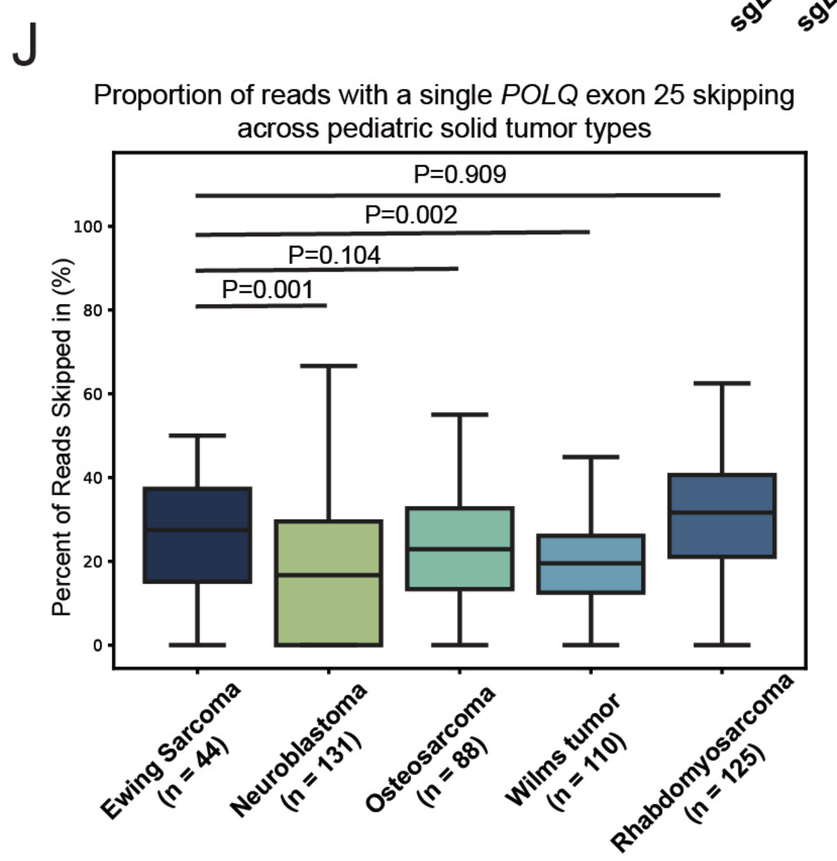

G
POLQ Exon25 mini-gene reporter
STOP
mScarlet
P2A
mNeon
Exon25
green
+ES
mNeongreen+
-ES
mNeongreen-
partial POLQ intron24
partial POLQ intron25

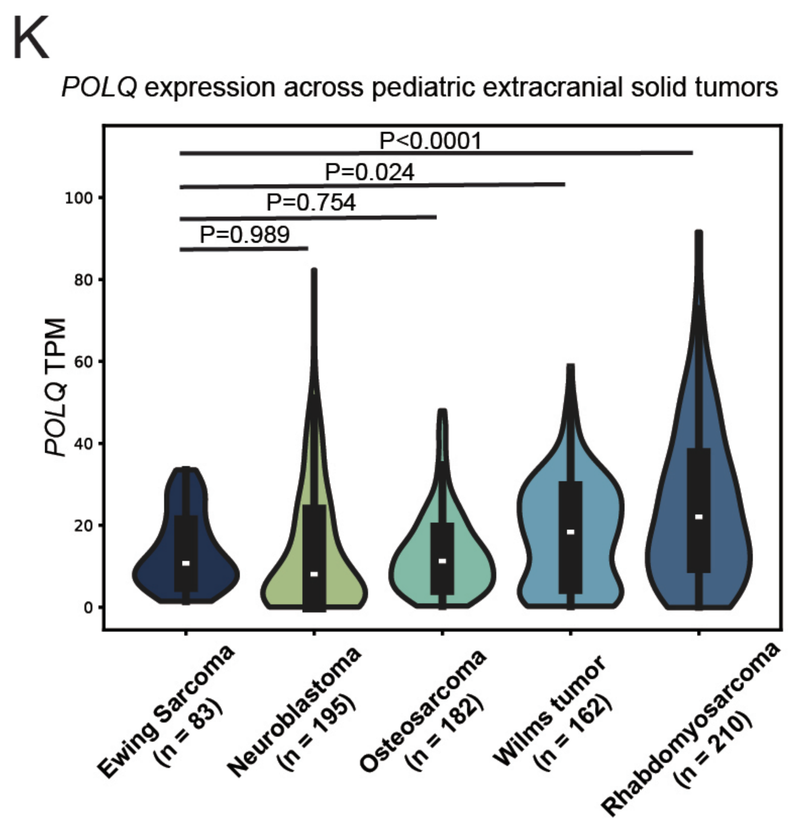

P > 0.9999
P = 0.9213
P < 0.0001
Relative mNG+ fraction
control
sgEWSR1 #1
sgEWSR1 #2
sgFLI1 #1
sgFLI1 #2

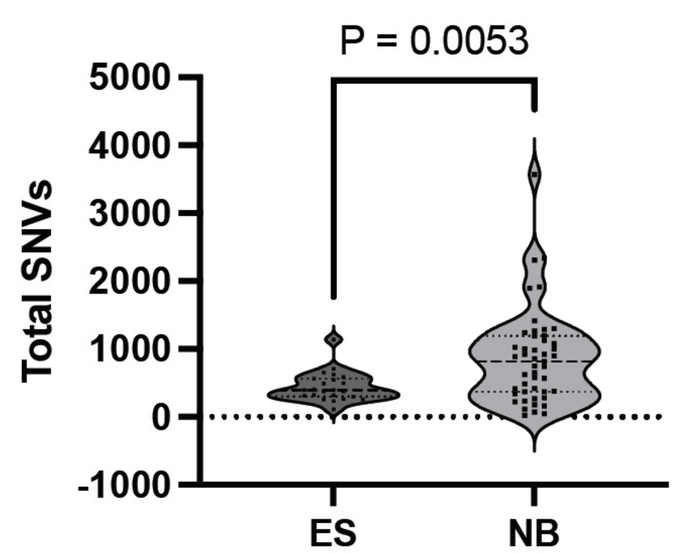

F
P < 0.0001
P < 0.0001
Relative MMEJ activity (normalized to POLQ WT)
EV
POLQ WT
POLQ dEx25
POLQ dEx7
HEK293T POLQ-/-

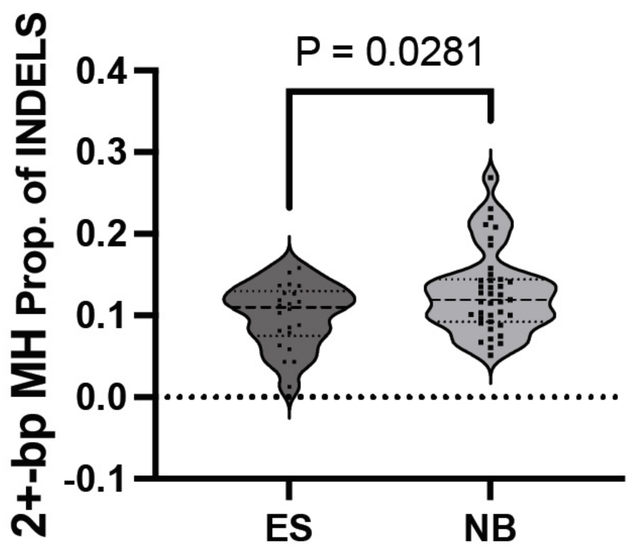

H
Ewing Sarcoma
U2OS
A673
TC32
TC71
RH1
RDES
POLQ exon6-8
+e7
- e7
POLQ exon24-26
+e25
- e25
GAPDH


I
U2OS
A673
TC32
TC71
RDES
RH1
POLΘ
GAPDH

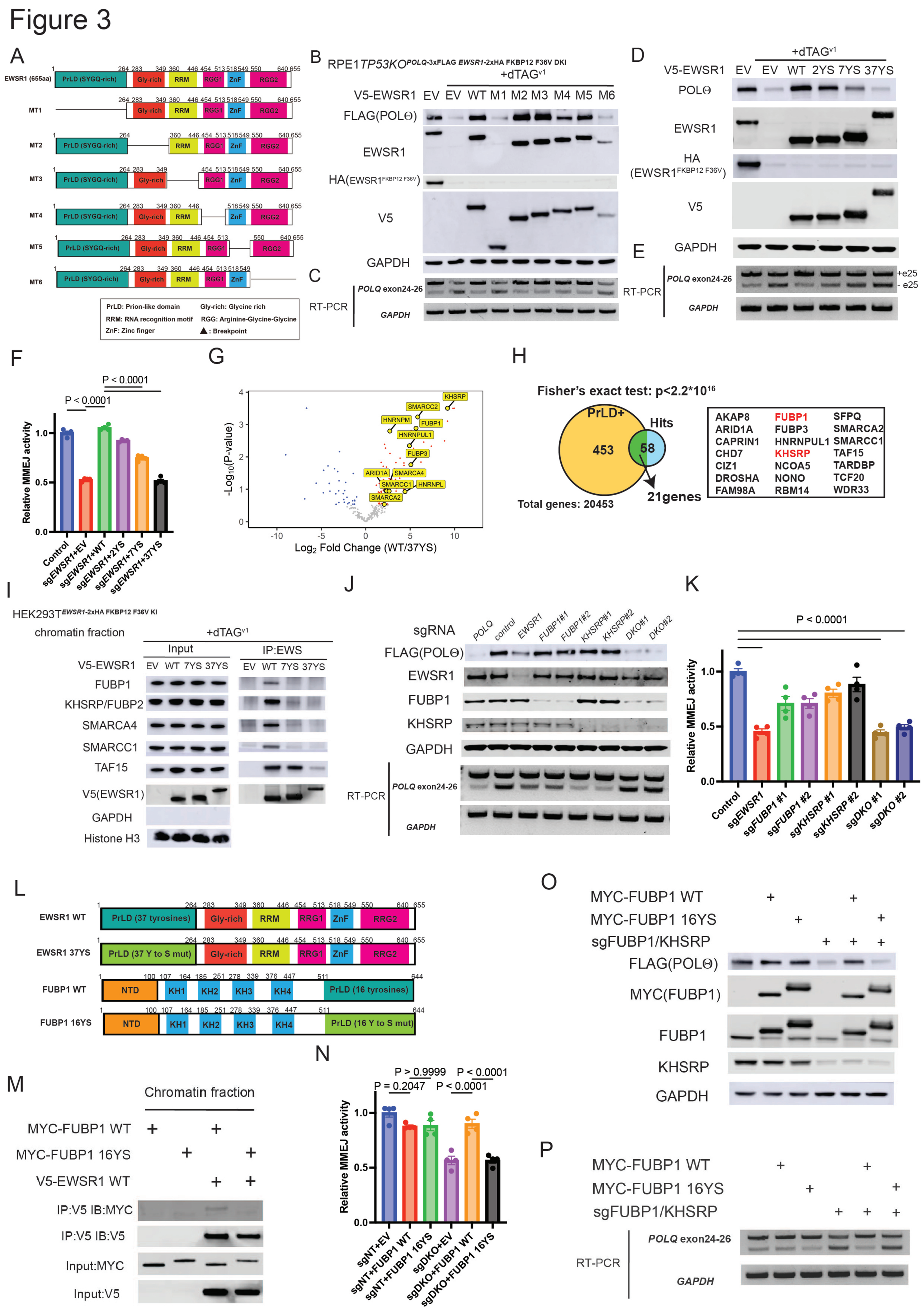

J
Proportion of reads with a single POLQ exon 25 skipping across pediatric solid tumor types
P=0.909
P=0.002
P=0.104
P=0.001
Percent of Reads Skipped in (%)
Ewing Sarcoma (n = 44)
Neuroblastoma (n = 131)
Osteosarcoma (n = 88)
Wilms tumor (n = 110)
Rhabdomyosarcoma (n = 125)


K
POLQ expression across pediatric extracranial solid tumors
P<0.0001
P=0.024
P=0.754
P=0.989
POLQ TPM
Ewing Sarcoma (n = 83)
Neuroblastoma (n = 195)
Osteosarcoma (n = 182)
Wilms tumor (n = 162)
Rhabdomyosarcoma (n = 210)

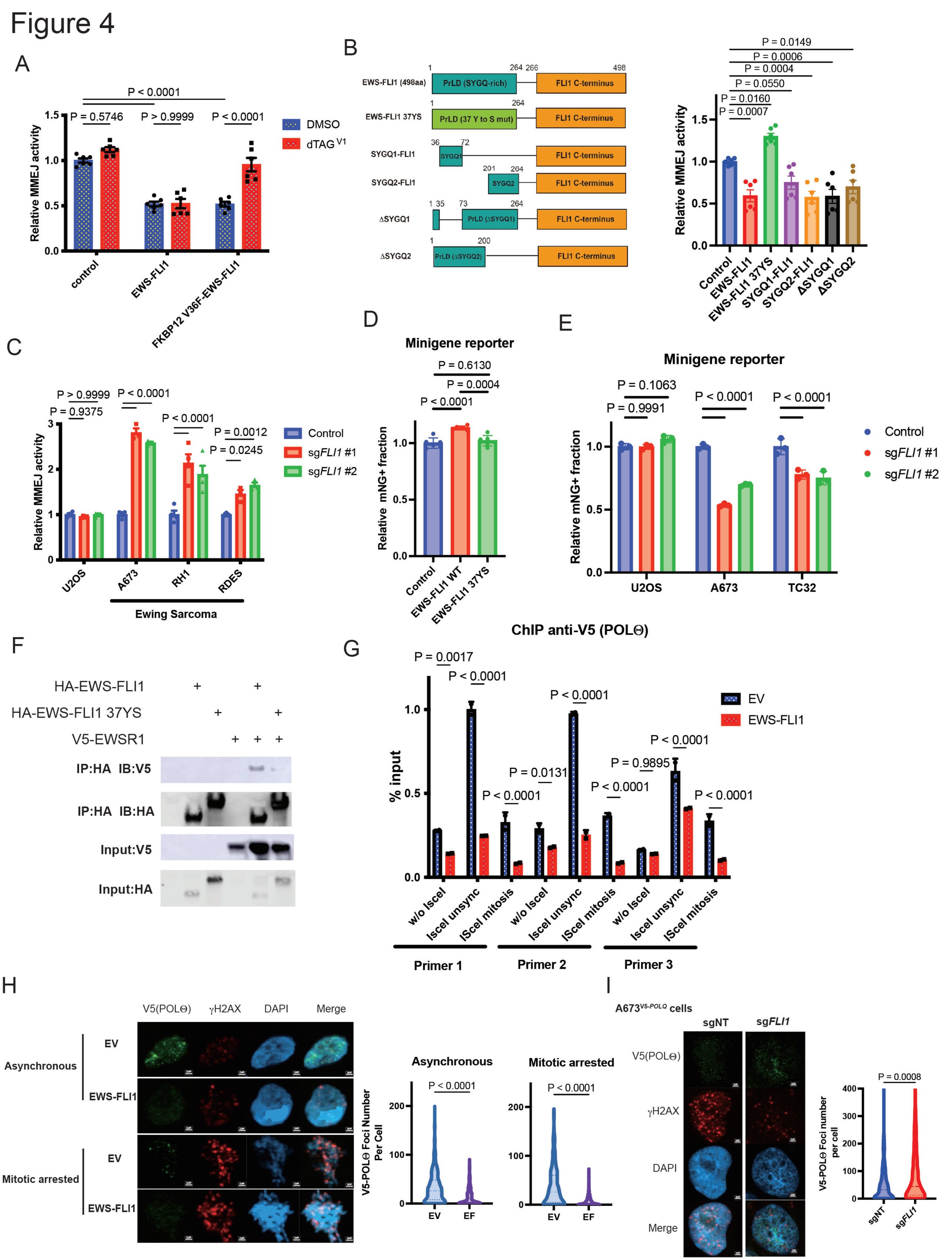

L
P = 0.0053
Total SNVs
ES
NB


M
P = 0.0281
2+-bp MH Prop. of INDELS
ES
NB

# Figure 3

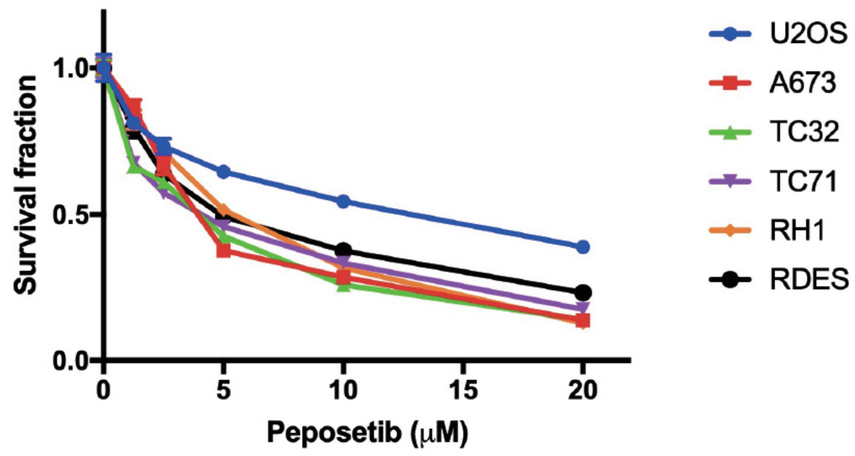

A
EWSR1 (655aa)
MT1
MT2
MT3
MT4
MT5
MT6
PrLD (SYGQ-rich)
Gly-rich
RRM
RGG1
ZnF
RGG2
PrLD: Prion-like domain
Gly-rich: Glycine rich
RRM: RNA recognition motif
RGG: Arginine-Glycine-Glycine
ZnF: Zinc finger
▲: Breakpoint
B
RPE1 TP53KO POLQ-3xFLAG EWSR1-2xHA FKBP12 F36V DKI
+dTAGv1
V5-EWSR1 EV EV WT M1 M2 M3 M4 M5 M6
FLAG(POLΘ)
EWSR1
HA(EWSR1FKBP12 F36V)
V5
GAPDH
C
RT-PCR
POLQ exon24-26
GAPDH
D
+dTAGv1
V5-EWSR1 EV EV WT 2YS 7YS 37YS
POLΘ
EWSR1
HA (EWSR1FKBP12 F36V)
V5
GAPDH
E
RT-PCR
POLQ exon24-26
+e25
- e25
GAPDH
F
P < 0.0001
P < 0.0001
Relative MMEJ activity
Control
sgEWSR1+EV
sgEWSR1+WT
sgEWSR1+2YS
sgEWSR1+7YS
sgEWSR1+37YS
G
-Log10(P-value)
Log2 Fold Change (WT/37YS)
KHSRP
SMARCC2
HNRNPM
FUBP1
HNRNPUL1
FUBP3
ARID1A
SMARCA4
SMARCC1
HNRNPL
SMARCA2
H
Fisher's exact test: p<2.2*10^16
PrLD+
Hits
453
58
21genes
Total genes: 20453
AKAP8 FUBP1 SFPQ
ARID1A FUBP3 SMARCA2
CAPRIN1 HNRNPUL1 SMARCC1
CHD7 KHSRP TAF15
CIZ1 NCOA5 TARDBP
DROSHA NONO TCF20
FAM98A RBM14 WDR33
I
HEK293T EWSR1-2xHA FKBP12 F36V KI
chromatin fraction
+dTAGv1
Input
IP:EWS
V5-EWSR1 EV WT 7YS 37YS EV WT 7YS 37YS
FUBP1
KHSRP/FUBP2
SMARCA4
SMARCC1
TAF15
V5(EWSR1)
GAPDH
Histone H3
J
sgRNA POLQ control EWSR1 FUBP1#1 FUBP1#2 KHSRP#1 KHSRP#2 DKO#1 DKO#2
FLAG(POLΘ)
EWSR1
FUBP1
KHSRP
GAPDH
RT-PCR
POLQ exon24-26
GAPDH
K
P < 0.0001
Relative MMEJ activity
Control
sgEWSR1
sgFUBP1 #1
sgFUBP1 #2
sgKHSRP #1
sgKHSRP #2
sgDKO #1
sgDKO #2
L
EWSR1 WT
EWSR1 37YS
PrLD (37 tyrosines)
PrLD (37 Y to S mut)
Gly-rich
RRM
RRG1
ZnF
RRG2
FUBP1 WT
FUBP1 16YS
NTD
KH1
KH2
KH3
KH4
PrLD (16 tyrosines)
PrLD (16 Y to S mut)
M
Chromatin fraction
MYC-FUBP1 WT
MYC-FUBP1 16YS
V5-EWSR1 WT
IP:V5 IB:MYC
IP:V5 IB:V5
Input:MYC
Input:V5
N
P > 0.9999
P < 0.0001
P = 0.2047
P < 0.0001
Relative MMEJ activity
sgNT+EV
sgNT+FUBP1 WT
sgNT+FUBP1 16YS
sgDKO+EV
sgDKO+FUBP1 WT
sgDKO+FUBP1 16YS
O
MYC-FUBP1 WT
MYC-FUBP1 16YS
sgFUBP1/KHSRP
FLAG(POLΘ)
MYC(FUBP1)
FUBP1
KHSRP
GAPDH
P
MYC-FUBP1 WT
MYC-FUBP1 16YS
sgFUBP1/KHSRP
RT-PCR
POLQ exon24-26
GAPDH

# Figure 4

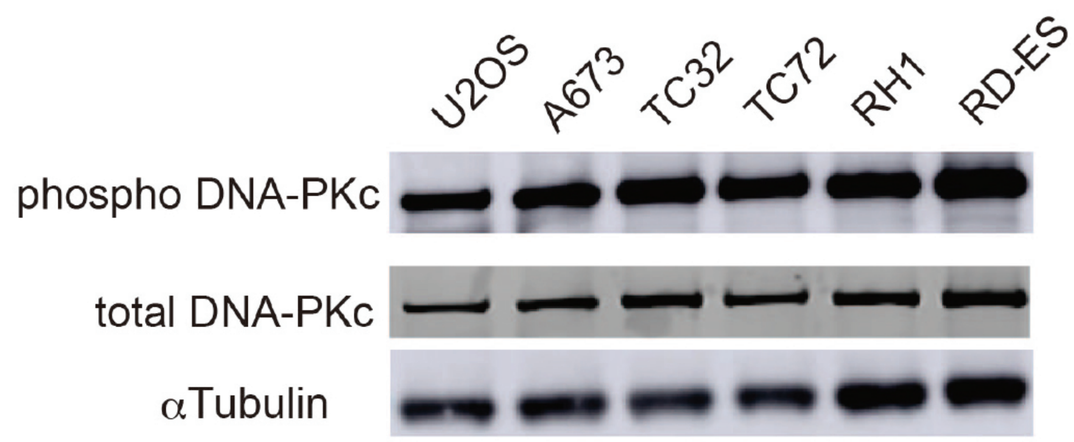

A
P < 0.0001
P = 0.5746
P > 0.9999
P < 0.0001
Relative MMEJ activity
DMSO
dTAG V1
control
EWS-FLI1
FKBP12 V36F-EWS-FLI1
B
EWS-FLI1 (498aa)
PrLD (SYGQ-rich)
FLI1 C-terminus
EWS-FLI1 37YS
PrLD (37 Y to S mut)
SYGQ1-FLI1
SYGQ1
SYGQ2-FLI1
SYGQ2
ΔSYGQ1
PrLD (ΔSYGQ1)
ΔSYGQ2
PrLD (ΔSYGQ2)
P = 0.0149
P = 0.0006
P = 0.0004
P = 0.0550
P = 0.0160
P = 0.0007
Control
EWS-FLI1
EWS-FLI1 37YS
SYGQ1-FLI1
SYGQ2-FLI1
ΔSYGQ1
ΔSYGQ2
C
P > 0.9999
P = 0.9375
P < 0.0001
P < 0.0001
P = 0.0012
P = 0.0245
Control
sgFLI1 #1
sgFLI1 #2
U2OS
A673
RH1
RDES
Ewing Sarcoma
D
Minigene reporter
P = 0.6130
P = 0.0004
P < 0.0001
Relative mNG+ fraction
Control
EWS-FLI1 WT
EWS-FLI1 37YS
E
Minigene reporter
P = 0.1063
P = 0.9991
P < 0.0001
P < 0.0001
U2OS
A673
TC32
F
HA-EWS-FLI1
HA-EWS-FLI1 37YS
V5-EWSR1
IP:HA IB:V5
IP:HA IB:HA
Input:V5
Input:HA
G
ChIP anti-V5 (POLΘ)
P = 0.0017
P < 0.0001
P < 0.0001
P = 0.0131
P < 0.0001
P < 0.0001
P = 0.9895
P < 0.0001
P < 0.0001
% input
EV
EWS-FLI1
w/o IScel
IScel unsync
IScel mitosis
Primer 1
Primer 2
Primer 3
H
V5(POLΘ)
γH2AX
DAPI
Merge
Asynchronous
Mitotic arrested
EV
EWS-FLI1
V5-POLΘ Foci Number Per Cell
P < 0.0001
P < 0.0001
EV
EF
I
A673V5-POLQ cells
sgNT
sgFLI1
V5(POLΘ)
γH2AX
DAPI
Merge
P = 0.0008
V5-POLΘ Foci number per cell
sgNT
sgFLI1

# Figure 5

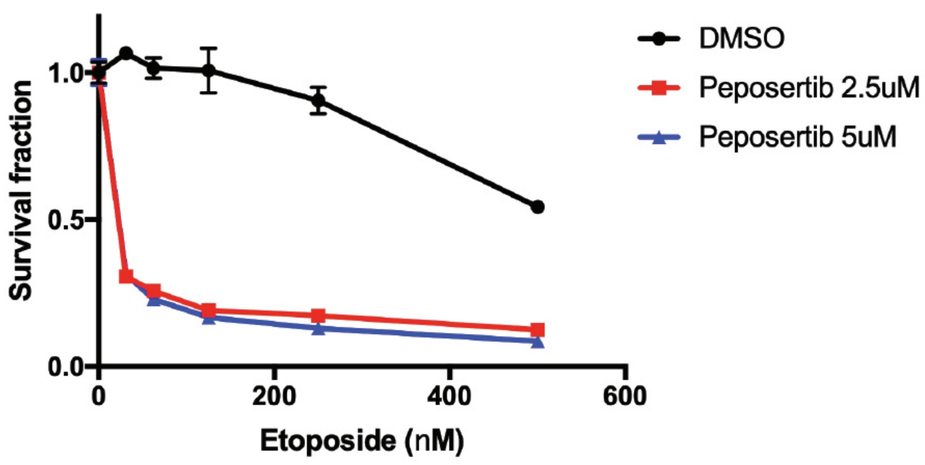
A



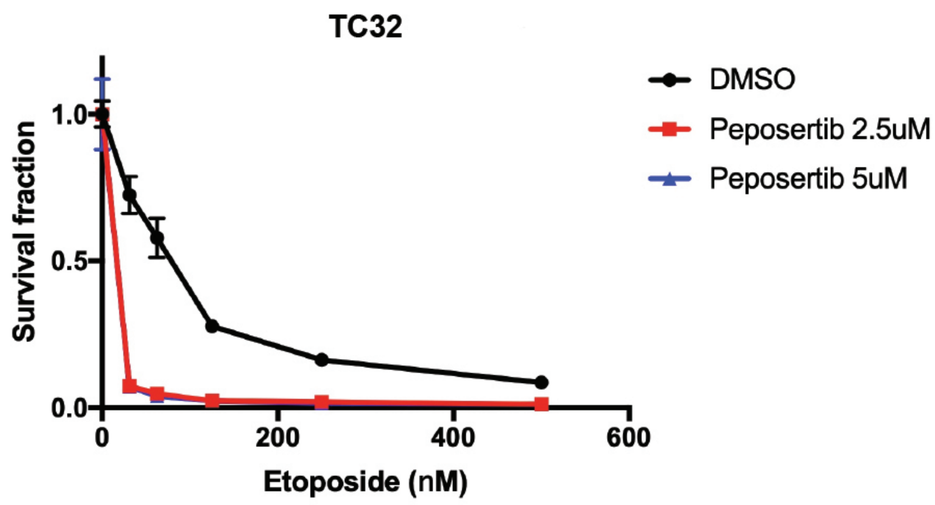

B



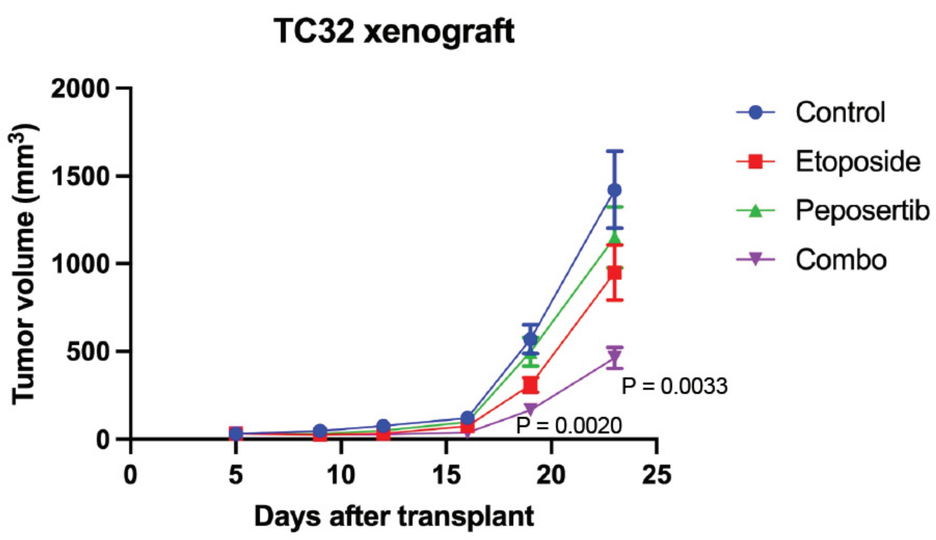

C



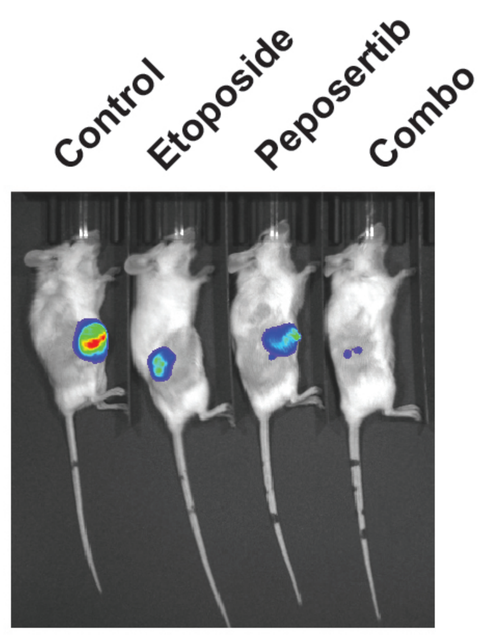

D



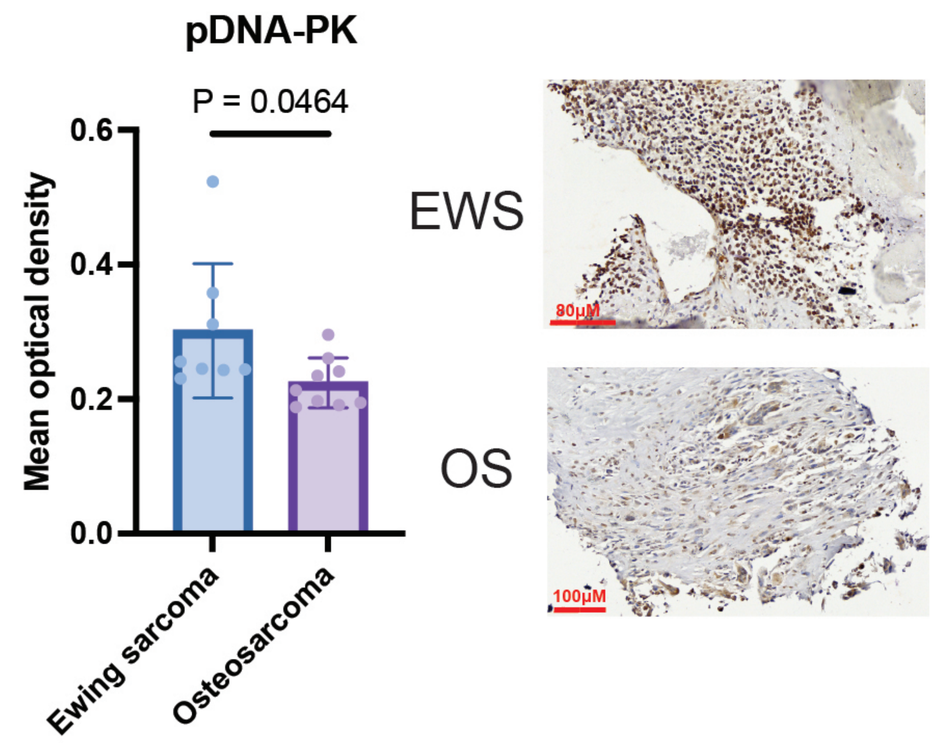

E



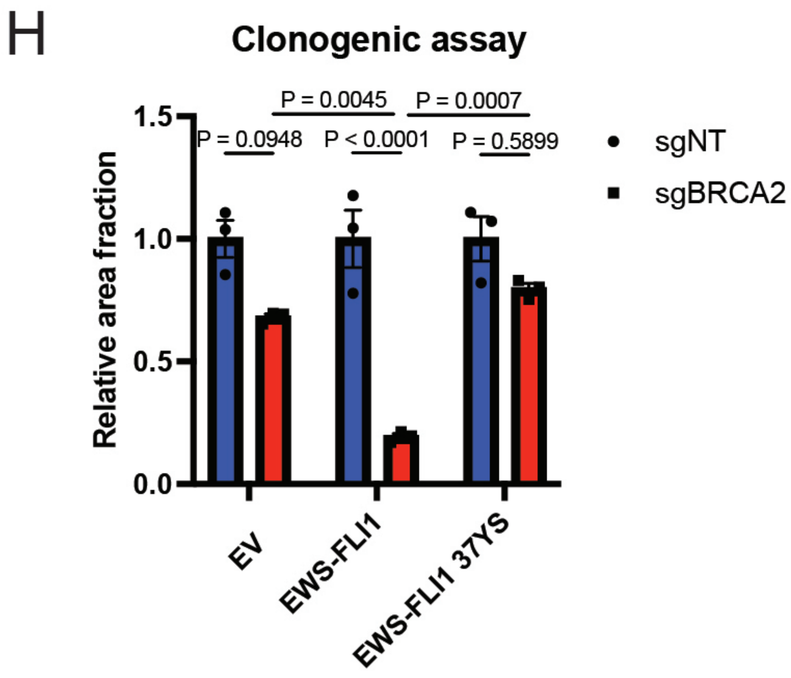

F

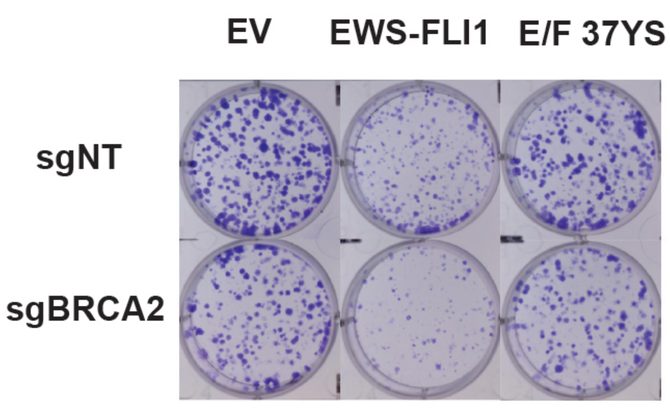



G

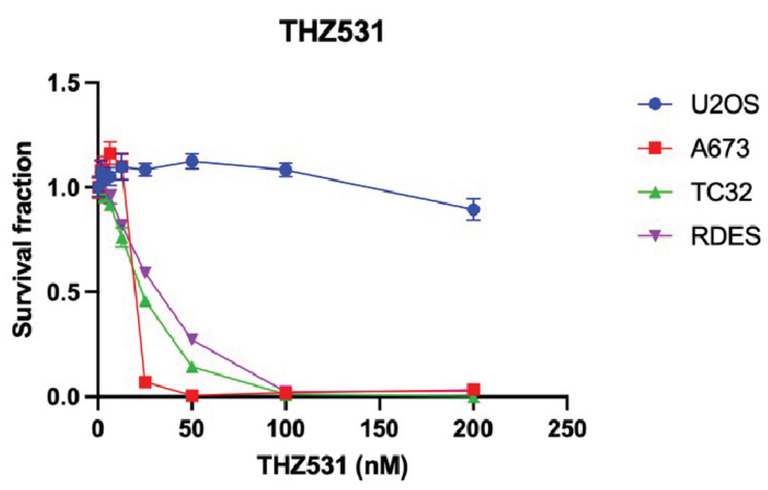



H

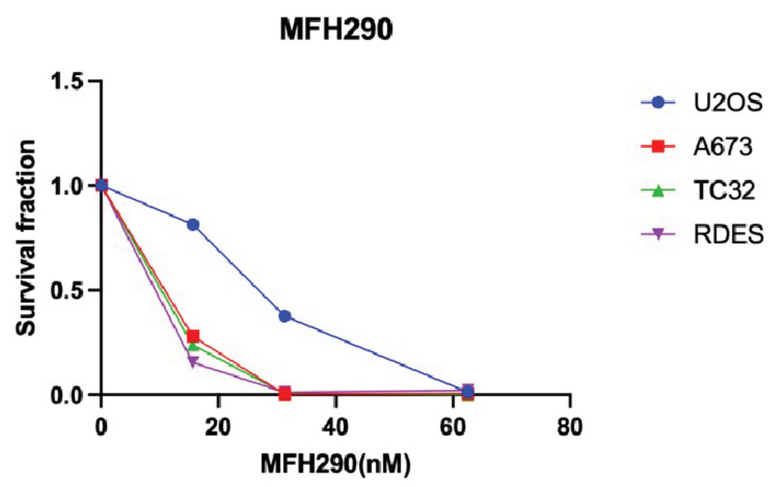


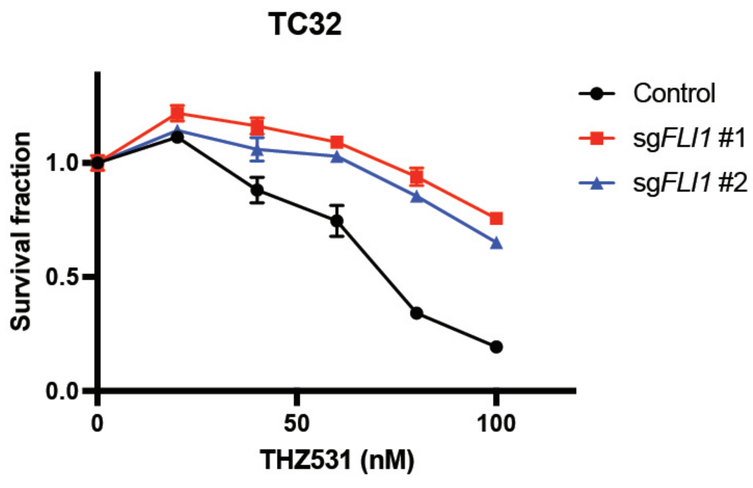

I

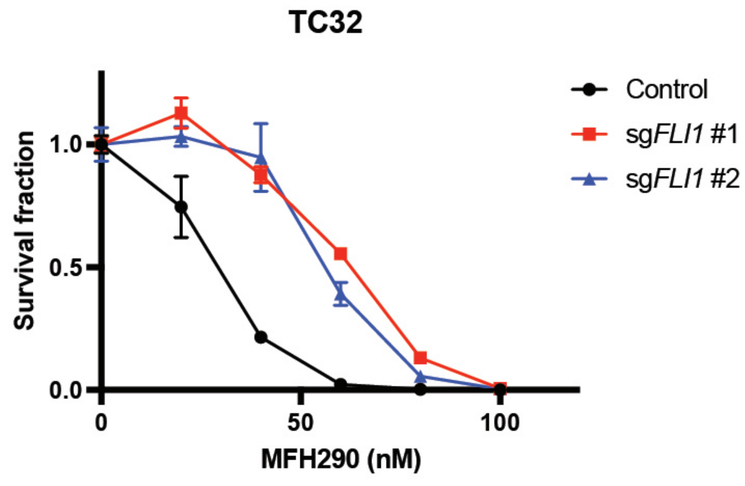



J




K




L