



HAL
open science

STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma

Didier Surdez, Sakina Zaidi, Sandrine Grossetête, Karine Laud-Duval, Anna Sole Ferre, Lieke Mous, Thomas Vourc'H, Franck Tirode, Gaelle Pierron, Virginie Raynal, et al.

► To cite this version:

Didier Surdez, Sakina Zaidi, Sandrine Grossetête, Karine Laud-Duval, Anna Sole Ferre, et al.. STAG2 mutations alter CTCF-anchored loop extrusion, reduce cis-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma. *Cancer Cell*, 2021, 39 (6), pp.810-826.e9. 10.1016/j.ccell.2021.04.001 . hal-03882743

HAL Id: hal-03882743

<https://hal.science/hal-03882743>

Submitted on 16 Jun 2023

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **STAG2 mutations alter CTCF-anchored loop extrusion,**
2 **reduce cis-regulatory interactions and EWSR1-FLI1 activity**
3 **in Ewing sarcoma**

4 Didier Surdez^{1,7*}, Sakina Zaidi^{1,9}, Sandrine Grossetête^{1,9}, Karine Laud-Duval¹, Anna
5 Sole Ferre², Lieke Mous^{1,7}, Thomas Vourc'h^{3,8}, Franck Tirode⁴, Gaelle Pierron⁵, Virginie
6 Raynal^{1,6}, Sylvain Baulande⁶, Erika Brunet², Véronique Hill¹, Olivier Delattre^{1,5,10*}

7 ¹ INSERM U830, Équipe Labellisée LNCC, Diversity and Plasticity of Childhood Tumors
8 Lab, PSL Research University, SIREDO Oncology Centre, Institut Curie Research
9 Centre, 75005 Paris, France.

10 ² INSERM UMR 1163, Laboratory of Genome Dynamics in the Immune System, Equipe
11 Labellisée Ligue contre le Cancer and Université de Paris, Imagine Institute, 75005
12 Paris, France.

13 ³ UMR 168, Biology Inspired Physics at Mesoscales, PSL Research University, Institut
14 Curie Research Centre, 75005 Paris, France.

15 ⁴ Univ Lyon, Université Claude Bernard Lyon 1, CNRS 5286, INSERM U1052, Cancer
16 Research Center of Lyon, 69008 Lyon, France

17 ⁵ Unité de Génétique Somatique, Service d'oncogénétique, Institut Curie, Centre
18 Hospitalier, 75005 Paris, France.

19 ⁶ Institut Curie Genomics of Excellence (ICGex) Platform, PSL Université, Institut Curie
20 Research Centre, 75005 Paris, France

21 ⁷ Present address: Balgrist University Hospital, University of Zurich, Zurich, Switzerland

22 ⁸ Present address: Université Clermont Auvergne, CNRS, Sigma Clermont, UMR 6602
23 Institut Pascal, 63000 Clermont-Ferrand, France

24 ⁹ These authors contributed equally to this work

25 ¹⁰ Lead Contact

SUMMARY

***STAG2*, a cohesin family gene, is among the most recurrently mutated genes in cancer. *STAG2* loss-of-function (LOF) is associated with aggressive behavior in Ewing sarcoma, a childhood cancer driven by aberrant transcription induced by the *EWSR1-FLI1* fusion oncogene. Here, using isogenic Ewing cells, we show that while *STAG2* LOF profoundly changes the transcriptome, it does not significantly impact *EWSR1-FLI1*, CTCF-cohesin or acetylated H3K27 DNA binding patterns. In contrast, it strongly alters the anchored dynamic loop extrusion process at boundary CTCF sites and dramatically decreases promoter-enhancer interactions, particularly affecting the expression of genes regulated by *EWSR1-FLI1* through binding GGAA microsatellite elements. Down-modulation of cis-mediated *EWSR1-FLI1* activity, observed in *STAG2*-LOF conditions, is associated with enhanced migration and invasion properties of Ewing cells previously observed in *EWSR1-FLI1*^{low} cells. Our study illuminates a process whereby *STAG2*-LOF fine-tunes the activity of an oncogenic transcription factor through altered CTCF-anchored loop extrusion and cis-mediated enhancer mechanisms.**

43 INTRODUCTION

44 Ewing sarcoma is an aggressive bone cancer mostly observed in adolescent and young
45 adults (Grünewald et al., 2018). This cancer is characterized by fusions between
46 *EWSR1* and *ETS* transcription factor family members, most frequently *FLI1* (Delattre et
47 al., 1992; Grünewald et al., 2018). This oncoprotein behaves as a pioneer transcription
48 factor, generating neo-enhancers through binding to GGAA microsatellites (Boulay et
49 al., 2017; Gangwal et al., 2008; Guillon et al., 2009; Sheffield et al., 2017; Tomazou et
50 al., 2015). *STAG2* mutation in Ewing sarcoma is the most frequent secondary genetic
51 alteration (15-21%) in an otherwise stable genome (Brohl et al., 2014; Crompton et al.,
52 2014; Solomon et al., 2011; Tirode et al., 2014). It is associated with poor prognosis and
53 metastasis (Crompton et al., 2014; Tirode et al., 2014). Furthermore, subclonal *STAG2*
54 mutations detected in tumors at diagnosis are preferentially expanded in relapsed
55 tumors suggesting their positive selection during cancer progression and treatment
56 (Crompton et al., 2014; Tirode et al., 2014).

57 *STAG2* is an integral member of the cohesin complex which is essential to hold sister
58 chromatids together during mitosis and to shape the three-dimensional genome
59 structure through its association with CTCF at the boundaries of topologically
60 associating domains (TAD) (Bintu et al., 2018; Dixon et al., 2012; Downen et al., 2014;
61 Michaelis et al., 1997; Nora et al., 2012; Rao et al., 2014; Wendt et al., 2008). Two
62 CTCF molecules binding at convergent sites and interacting with the cohesin complex
63 allow for the generation of a chromatin loop in which gene regulation processes
64 preferentially occur (Beagrie et al., 2017; Guo et al., 2015; Rao et al., 2014, 2017; Tang
65 et al., 2015). Key components of the CTCF/cohesin complex are necessary for the
66 maintenance of chromatin loop structures (Haarhuis et al., 2017; Nora et al., 2017; Rao
67 et al., 2017; Schwarzer et al., 2017; Wutz et al., 2017) and are typically identified
68 through a dot at corner peaks on Hi-C contact maps. These loops are likely generated
69 in a dynamic process called chromatin extrusion (Davidson et al., 2019; Fudenberg et
70 al., 2016; Hassler et al., 2018; Hsieh et al., 2020; Kim et al., 2019; Krietenstein et al.,
71 2020; Nasmyth, 2001; Nuebler et al., 2018; Sanborn et al., 2015; Vian et al., 2018).
72 Chromatin extrusion by the condensin complex has been recently visualized in yeast
73 (Ganji et al., 2018). ATP and NIPBL-MAU2 are essential factors for the extrusion of

74 chromatin loops by the cohesin complex *in vitro* (Davidson et al., 2019; Kim et al.,
75 2019). *In vivo*, it is thought that the dynamic extrusion process is reflected by
76 architectural stripes detected in Hi-C and Micro-C experiments (Hsieh et al., 2020;
77 Krietenstein et al., 2020; Vian et al., 2018). Current understanding of the exact
78 mechanisms describing the interplay between the CTCF/cohesin complex and the
79 chromatin during loop extrusion is however still incomplete. Similarly, the role of
80 *STAG2*, which is frequently altered in human cancer mainly through loss-of-function
81 (LOF) mutations, is still poorly understood (Bailey et al., 2018; Hill et al., 2016;
82 Lawrence et al., 2014; Romero-Pérez et al., 2019). Since *STAG2* is located on the X
83 chromosome, inactivating mutations of a single allele are sufficient for complete LOF
84 (Romero-Pérez et al., 2019). Inactivating mutations of the *STAG1* paralog are much
85 less frequent in cancer. *STAG1* and *STAG2* LOF mutations have been shown to be
86 synthetic lethal (Benedetti et al., 2017; van der Lelij et al., 2017).

87 Here, we addressed whether *STAG2* LOF mutations alter transcriptome, epigenome
88 and chromatin topology of Ewing sarcoma cellular models and investigated the
89 mechanisms by which this mutation could contribute to increased aggressiveness of this
90 cancer.

91 **RESULTS**

92 ***STAG2* knock-out profoundly alters the transcriptomic landscape**

93
94 To decipher the oncogenic mechanisms related to *STAG2* LOF in Ewing sarcoma, we
95 used a CRISPR/Cas9 approach with two different sgRNAs targeting *STAG2* (SA2m#1
96 and SA2m#2). We generated three knock-out (KO) isogenic pairs derived from A673
97 (A673^{SA2m#1}) and TC71 (TC71^{SA2m#1}, TC71^{SA2m#2}), two *STAG2* wild type (WT) Ewing
98 sarcoma cell lines (Figure 1A). Absence of *STAG2* protein expression was confirmed in
99 each of these clones (Figure 1A). Proliferation rate of *STAG2*-WT and -KO cells was
100 similar (Figure S1A-D). RNA-seq comparing paired *STAG2* proficient and deficient lines
101 highlighted a broad transcriptional modulation (Figure 1B, Table S1). Altogether, these
102 three isogenic clones define a set of 546 *STAG2*-modulated genes, 204 being
103 commonly up-regulated and 198 being commonly down-regulated genes in *STAG2*-KO
104

105 cells. To validate the specificity of our findings, we used a CRISPR/Cas9-based
106 approach to correct the *STAG2* mutation and generated a line (A673^{SA2r}) with rescued
107 *STAG2* expression (Figure 1A). Expression profiling in this rescue line showed highly
108 significant reversion of both the *STAG2*-KO-associated down- and up-regulated effects
109 (Figure 1C and Table S1). We also performed short term knock-down experiments with
110 two different siRNAs (siSA2#6, siSA2#8) at three different time points (24, 48, 72hrs) in
111 A673 and TC71. To confirm that the *STAG2*-regulated signature was not limited to A673
112 and TC71 isogenic clones, we also knocked-down *STAG2* in three additional Ewing cell
113 lines (EW1, CHLA-10 and CHLA-258) (Figure 1A) and further validated RNA-seq results
114 for some genes using RT-QPCR (Figure S1E-J). As shown in Figure 1D and S2A, all
115 these experiments indicated that the sets of *STAG2*-modulated genes defined with
116 isogenic clones were regulated as soon as 24-48hrs post siRNA transfection. At 72 hrs,
117 most of the gene expression changes detected in the isogenic clones were observed,
118 with only minor variations between siRNAs or cell lines and with a lower dynamic range
119 of modulation than in stable KO experiments (Figure S2A). We can therefore conclude
120 that this set of genes represents a robust signature of *STAG2*-KO Ewing cells,
121 modulated at short-term and which hence accounts for early, possibly direct,
122 transcriptomic consequences of *STAG2* inactivation.

123 124 **Functional analysis reveals major impact of *STAG2* inactivation on *EWSR1-FLI1*-** 125 **induced genes**

126 The functional aspect of *STAG2* proficient and deficient Ewing sarcoma was
127 investigated by Gene Set Enrichment Analysis (GSEA) and using the DoRothEA
128 curated transcription factor/target gene set database (Garcia-Alonso et al., 2019;
129 Mootha et al., 2003). Taking advantage of additional Ewing sarcoma cell line and tumor
130 RNA-seq data, we performed this analysis in four independent Ewing sarcoma datasets:
131 i) *STAG2*-WT parental cells vs *STAG2*-KO isogenic cells vs, ii) *STAG2*-WT si-control vs
132 *STAG2* knock-down transfected cells iii) *STAG2*-WT vs *STAG2*-mutated Ewing cell
133 lines, iv) *STAG2*-WT vs *STAG2*-mutated Ewing tumors. Strikingly, when investigating
134 18,889 signatures ranked by average Normalized Enrichment Score (NES), several of
135 the top20 signatures enriched in *STAG2* proficient condition were *EWSR1-FLI1*-

136 regulated gene signatures. (Table 1 and S2). Ranking first in GSEA and DoRothEA
137 analyses (Table 1 and S3), IC-EWS is a recently described signature (Aynaud et al.,
138 2020) that was defined based on independent component analysis of single cell RNA-
139 seq experiments upon induction of *EWSR1-FLI1* in Ewing cells. This signature is
140 exquisitely specific for Ewing sarcoma, enriched in genes modulated by EWSR1-FLI1
141 activity on GGAA microsatellite sequences and mostly devoid of cell cycle genes which
142 are frequently confounding factors in such GSEA analyses. Apart from a borderline-
143 significant signature observed at 24hrs post transfection in A673 cells knock-down for
144 *STAG2*, all other comparisons yielded strongly significant GSEA results with the IC-
145 EWS gene set (Figure S2B-G). Direct comparison also showed that IC-EWS distinguish
146 *STAG2*-WT and -mutated cell lines and tumors (Figure S2H, I). Further evidences for
147 increased activity of EWSR1-FLI1 in *STAG2* proficient conditions are provided by Riggi
148 and Miyagawa data sets highlighting genes that are up-regulated upon ectopic
149 expression of *EWSR1-FLI1* in mesenchymal/progenitor cells. Beyond EWSR1-FLI1-
150 related signatures, other gene sets provided much less consistent information across
151 datasets and hence appeared less meaningful (Table 1). One gene set each of EGF or
152 TGF β signaling, and one gene set each of MYC, P53 and NF κ B targets were ranked
153 among the first twenty gene sets (Table 1 and S3). It is noteworthy that the P53 gene
154 set is particularly enriched in analyses performed in cell lines and in tumors, an
155 observation which may be linked to the frequent association of *STAG2* and *TP53*
156 mutations in Ewing cell lines and tumors (Tirode et al., 2014). In contrast to down-
157 regulated genes in *STAG2*-LOF systems, analyses of up-regulated gene sets did not
158 provide obvious illuminating information (Table 1 and S3). NES scores were usually
159 weaker and the highest scores observed in the tumor comparison were not strongly
160 supported by the cell line systems. We may nevertheless note the YAP/TAZ pathway
161 which has been recently shown to counteract EWSR1-FLI1 activity (Katschnig et al.,
162 2017; Rodríguez-Núñez et al., 2020) (Table 1). Altogether, these analyses showed that
163 *STAG2* inactivation has a major impact on EWSR1-FLI1 up-regulated genes and hence
164 suggested that *STAG2* may modulate EWSR1-FLI1 transcriptional effects.

166 **Binding patterns at H3K27ac, EWSR1-FLI1, CTCF and cohesin sites are mostly**
167 **unmodulated upon STAG2 knock out**

168 Our primary hypothesis was that the inactivation of *STAG2* may impair EWSR1-FLI1
169 binding and/or accessibility to chromatin. We hence performed western blot and ChIP-
170 seq experiments against plausible causative factors in our isogenic models. A
171 decreased expression of EWSR1-FLI1 which could account for its decreased activity
172 was neither observed in isogenic nor in knocked-down cells. The opposite, a slight
173 increase of EWSR1-FLI1 protein expression, could be observed in TC71 *STAG2*-KO
174 and A673 *STAG2*-KD cells (Figure 1A). Similarly, no consistent variation of the level of
175 H3K27ac was observed (Figure 1A). ChIP-seq analyses of EWSR1-FLI1 and H3K27ac
176 also did not provide an explanation for decreased EWSR1-FLI1 activity (Figure 2A-B,
177 S3A-C). A slight increase of EWSR1-FLI1 binding can be noted in TC71-KO and A673-
178 KD cells (Figure S3B, C), which may possibly reflect the increased expression of
179 EWSR1-FLI1 mentioned above but this cannot account for the paradoxical decreased
180 activity of this protein. As *STAG2* is a member of the cohesin complex, we also
181 investigated the expression and binding patterns of CTCF and of other subunits of the
182 cohesin complex. None of the expression or binding patterns were altered upon
183 *STAG2*-KO, apart from a slight, possibly compensatory, increase of *STAG1* expression
184 and binding to DNA in *STAG2*-KO cells (Figure 1A, 2C, S3A-C). As a prototypic locus,
185 we used the *DKK2* gene which is a well-known EWSR1-FLI1 target (Kauer et al., 2009;
186 Miyagawa et al., 2009; Riggi et al., 2008), which displays EWSR1-FLI1-bound GGAA
187 microsatellites as plausible cis-regulatory enhancer elements and for which all genomic
188 features can be displayed in a single panel. Expression of *DKK2* is strongly decreased
189 in *STAG2*-KO cells and restored in *STAG2*-rescued cells (Figure 2D). Figure 2E and S3
190 illustrate the global conservation of CTCF, cohesin, EWSR1-FLI1 and enhancer
191 H3K27ac marks at the *DKK2* locus in *STAG2*-KO cells. Genome wide, we also noticed
192 that *STAG1*-specific- or *STAG2*-specific-cohesin binding sites were rare (Figure 2C,
193 S3A-C) as compared to recently published data (Kojic et al., 2018; Viny et al., 2019).
194 Altogether, protein expression and ChIP-seq profiles provide no consistent explanation
195 for the decreased EWSR1-FLI1 signature observed in *STAG2* deficient Ewing sarcoma
196 cells.

197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227

CTCF HiChIP highlights a STAG2 dependent anchored extrusion mechanism

We hypothesized that the loss of *STAG2* may alter cis-mediated enhancer activity through changes in CTCF-cohesin loop domains. Three-dimensional genome conformation results from the sum of multiple interaction types: A-B domains and CTCF/cohesin loops as well as promoter/enhancer or polycomb complex hubs, phase separated domains, and transcription factors contracted loci (Hnisz et al., 2017; Hsieh et al., 2020; Krietenstein et al., 2020; Merkmenschlager and Nora, 2016; Rowley and Corces, 2018; Stadhouders et al., 2019). Hi-C and Micro-C experiments capture all of these different types of interaction but only allow to infer their exact origin based on concomitant ChIP-seq data. Importantly, *STAG2* has been reported to bridge CTCF and RAD21 (one of the three core subunits of the cohesin ring) through direct protein-protein interactions (Xiao et al., 2011; Zhang et al., 2013; Li et al., 2020). Aiming at specifically investigating CTCF/cohesin interactions, we performed CTCF HiChIP (Mumbach et al., 2016) and generated for all isogenic models a high coverage, 5 kb resolution interaction map using HiC-Pro (Servant et al., 2015) (Table S4). At low (250kb bins) resolution, no obvious change of the global “plaid” pattern was observed between the different isogenic cells (Figure S4A). At high resolution (5kb bins), the size and positions of the loop domains were mostly unaffected (Figure S4B). However, a striking difference could be observed throughout the genome at “stripes” that reveal high interaction frequencies between a single CTCF locus and contiguous loci (Figure 3A-H, S4B, C). Such stripes have recently been reported to occur at particular loci with super-enhancer (SE) features and have also been predicted by computer simulations (Fudenberg et al., 2016; Vian et al., 2018). They are suggested to reflect the dynamic process of CTCF/cohesin loop domain generation whereby one cohesin complex is arrested at a first CTCF site whereas the second CTCF is sliding concomitantly with chromatin (CTCF-anchored extrusion model, Figure 3H). As shown for the prototypic *DKK2* locus and at a representative broader region on chromosome 10, a decreased signal at such “anchored extrusion stripes” was observed in *STAG2*-KO cells (Figure 3B, S4B, C) as compared to *STAG2*-WT or -rescued cells (Figure 3A, C and S4B, C). CTCF HiChIP experiments were also conducted at 72h following transfection with two different siRNAs

228 (siSA#6 and siSA#8) in A673 cells and confirmed the stripe reduction (Figure 3D-F and
229 Figure S4B). To further investigate this observation in a cellular context where only
230 STAG2 is expressed, we also generated Ewing cells with a KO of *STAG1*, the paralog
231 of *STAG2* (Figure 1A). In contrast to *STAG2*-KO cells, stripes were even more intense
232 in *STAG1*-KO cells compared to WT (Figure 3G and S4B).

233

234 **Genome-wide loop detection highlights the role of STAG2 in the anchored** 235 **extrusion process**

236 In order to quantify and expand these observations at the genome-wide level, we
237 developed an algorithm (Tweed) for the detection of stripes and the resulting loop
238 domains in simple or interlaced loop regions (Figure 4A, S5A-H, STAR Methods). The
239 vast majority of detected stripes were flanked by convergent CTCF sites (Figure S5E)
240 and were intensely decorated by cohesin members therefore demonstrating efficient
241 identification of loop domains (Figure S5F, G). Because loop formation in the anchored
242 extrusion model can emerge from either left or right CTCF boundaries, we also
243 distinguished hereafter X and Y stripes (Figure 3H, S5H, S6A, B). Three major findings
244 could thus be highlighted: first, 18,774 (A673^{WT}) and 14,263 (TC71^{WT}) loops were
245 detected throughout the genome (Figure S6C-E). Here, these loops are detected
246 through stripe structures and not through corner peak dots as usually reported in Hi-C
247 contact maps data. These results are however consistent with numbers previously
248 reported in mammalian genomes, thus indicating that anchored extrusion probably
249 represents a general mechanism of loop formation (Figure 3H). Second, most loops
250 showed equivalent X and Y stripe intensities, indicating an extrusion process with
251 balanced anchorage at either CTCF loop boundaries (Figure 4B). In the few loops
252 presenting skewed X or Y stripe patterns, the anchored CTCF boundary was
253 predominantly decorated by cohesin members (Figure 4C, S6A, B). Third, and most
254 importantly, anchored loop extrusions were strongly reduced, genome-wide, in *STAG2*-
255 KO cells and significantly reverted upon *STAG2* restoration (Figure 4D, S6E). In
256 contrast, these stripes were strongly reinforced in *STAG1*-KO as compared to parental
257 WT cells (Figure 4D). In *STAG2*-KO cells, only very few stripes demonstrated increased
258 signals (Figure 4D and S6E). Moreover, outlier stripes were not consistent between the

259 different *STAG2*-KO models. Similarly, only outlier and non-consistent stripes
260 demonstrated decreased signal in *STAG1*-KO cells (Figure 4D and S6E). Decreased
261 stripes were also readily observed in *STAG2*-knocked-down cells as compared to cells
262 transfected with a control siRNA (Figure 3D-F, Figure 4D). Altogether, these
263 experiments indicated that *STAG2* LOF impacts the loop extrusion process genome-
264 wide and that this phenotype is an early, likely direct consequence of the inactivation of
265 *STAG2*.

266 267 **STAG2 dependent anchorage is associated with decreased cis-promoter- 268 enhancer interactions within loops**

269 Abundant evidence shows that CTCF/cohesin chromatin loops are preferential
270 structures allowing for promoter-enhancer interactions (Dixon et al., 2012; Downen et al.,
271 2014; Guo et al., 2015; Hnisz et al., 2016; Lupiáñez et al., 2015; Narendra et al., 2015;
272 Phillips-Cremins et al., 2013; Rao et al., 2014; Splinter et al., 2006; Tang et al., 2015;
273 Vian et al., 2018; Weintraub et al., 2017). In particular, recent data indicates that loop
274 extrusion dynamically juxtaposes elements necessary for antibody class switching in B-
275 cells (Zhang et al., 2019). We therefore speculated that *STAG2* LOF could impair cis-
276 gene activation. To explore this mechanism, we performed H3K27ac HiChIP in *STAG2*-
277 proficient and -deficient cells. At the prototypical *DKK2* locus, we observed a global loss
278 of interactions in *STAG2*-KO cells that was reverted upon *STAG2* re-expression (Figure
279 S7A). Similarly, other classical EWSR1-FLI1 target genes displayed loss of interaction
280 patterns (Figure S7B). To extend our findings genome wide, we developed a
281 bioinformatic pipeline to detect promoter-enhancer (pe) and enhancer-enhancer (ee)
282 interaction chains. Starting from 5kb bins overlapping H3K4me3 promoter peaks, the
283 strongest H3K27ac HiChIP chain-interactions were reported until the twentieth
284 interaction sites. To validate that this H3K27ac interaction-based, gene expression
285 agnostic, approach, accurately identifies regulatory regions, we isolated all genes that
286 include an EWSR1-FLI1-bound GGAA microsatellite sequence within their regulatory
287 chain. A total of 2331 and 1625 genes were retrieved in A673 and TC71, respectively
288 (Table S5). Functional analysis of this set of genes confirmed a very significant
289 enrichment in EWSR1-FLI1 up-regulated genes (Table S6). We then investigated

290 whether promoter-enhancer interactions were modified in *STAG2*-KO cells. H3K27ac
291 ChIP-seq signal at enhancers in these chains remained mostly unaffected upon
292 *STAG2*-KO or -rescue (Figure 5A-B), emphasizing our first observation that H3K27ac
293 signal is largely conserved upon *STAG2* LOF (Figure 1A, 2A, S3A-C). In contrast,
294 H3K27ac HiChIP demonstrated a genome-wide global loss of interactions within these
295 chains in *STAG2*-KO cells (Figure 5C-D). These interactions were significantly restored
296 in *STAG2*-rescued isogenic cells (Figure 5C). Both loss and restoration effects were
297 similar at EWSR1-FLI1-bound and -unbound enhancers (Figure S8A-D). Since SEs
298 consist of clusters of enhancers (Hnisz et al., 2013; Whyte et al., 2013), we also
299 investigated the impact of *STAG2* LOF at these loci and observed a strong reduction of
300 intra-SE H3K27ac interactions that was rescued in *STAG2* rescued cells (Figure 5E).
301 We then studied the relationship between chromatin loops and loss of H3K27ac
302 promoter-enhancer interactions. We plotted the density of interactions depending upon
303 the distance between cis-interacting pairs on the genome (Figure 5F). Patterns of CTCF
304 and of H3K27ac interactions were highly similar in both *STAG2*-WT and -KO conditions
305 (Figure 5F) with a noticeable decrease of interaction read frequency in a window
306 comprised between 20kb to around 500kb in *STAG2*-KO cells (Figure 5F zoom) which
307 was reverted upon *STAG2* re-expression. Quite strikingly, this distance corresponds to
308 the size of most chromatin loops (Figure 5F). Though correlative, these data therefore
309 strongly suggest that altered loop extrusion as a result of *STAG2* LOF impairs efficient
310 formation of promoter-enhancer interactions.

311

312 **EWSR1-FLI1-induced genes are particularly impacted by loosened cis-regulatory** 313 **interactions**

314 We next wondered why loosened chromatin interactions that occur genome-wide within
315 loop-congruent distances may particularly impact EWSR1-FLI1-induced genes.
316 Previous publications have shown that EWSR1-FLI1 binding occurs predominantly at
317 enhancers and is particularly enriched at SE (Baldauf et al., 2018; Boulay et al., 2018;
318 Kennedy et al., 2015; Riggi et al., 2014; Sheffield et al., 2017; Tomazou et al., 2015). In
319 particular, EWSR1-FLI1 transforms otherwise quiescent GGAA-microsatellites into
320 active “neo-enhancers”. We therefore concentrated our analyses on such sequences.

321 EWSR1-FLI1-bound GGAA microsatellite sequences, similarly to other types of
322 enhancers, super-enhancers and TSS of expressed genes were predominantly located
323 within chromatin loops (Figure 6A, B). In *STAG2*-WT cells, the number of interaction-
324 pairs at EWSR1-FLI1-bound enhancers, whether GGAA microsatellites or single ETS
325 sites, was significantly higher than at other enhancer sites in both A673 and TC71 cells
326 (Figure 6C, D). We next investigated the degree of loosened interactions at genes
327 regulated by *STAG2* LOF and noted that, while a decrease of interactions in *STAG2*-KO
328 cells was observed for all genes, it was significantly greater for down-regulated genes
329 than on un- or up-regulated genes in A673 and TC71 cells (Figure 6E, F). In addition,
330 genes containing an EWSR1-FLI1-bound GGAA microsatellite in their regulatory chain
331 were more frequently down-regulated by *STAG2* LOF compared to other genes (Chi-
332 square, $p=1.4 \times 10^{-5}$ in A673 cells, $p=2.2 \times 10^{-16}$ in TC71 cells). Altogether these data
333 show that, though the impact of *STAG2* LOF on promoter-enhancer interactions is
334 genome-wide, EWSR1-FLI1-activated genes, and particularly those regulated by
335 microsatellite sequences, are predominantly impacted likely due to the abundance of
336 interactions at these sequences.

337

338 ***STAG2* mutation promotes migration and invasive properties of Ewing cancer** 339 **cells**

340 Previous reports have indicated that, in addition to decreased proliferation, the knock-
341 down of *EWSR1-FLI1* in Ewing cells partly reverts the phenotype of Ewing cells into that
342 of mesenchymal stem cells, the likely cell-of-origin (Tirode et al., 2007) with
343 morphological changes such as cell flattening, increased cell matrix adhesion and
344 increased migration ability (Chaturvedi et al., 2014; Franzetti et al., 2017; Pedersen et
345 al., 2016). As mentioned above, proliferation rate of *STAG2*-KO cells was not obviously
346 and consistently altered and was highly similar to that of parental cells (Figure S1A-D).
347 This is consistent with the observation that IC-EWS, which is mostly devoid of cell cycle
348 genes, is ranked first compared to other EWSR1-FLI1-activation signatures which do
349 include such cell cycle genes (Table 1 and S2). An increase of paxillin associated stress
350 fiber adhesion foci was observed in the A673-derived clone, with *STAG2*-rescue (Figure
351 7A, B). Using soft agar assay, clonogenicity of A673^{SA2m#1} cells was increased as

352 compared to A673^{WT} and A673^{SA2r} cells, however this was not observed in TC71
353 isogenic models (Figure S1D). A more striking effect was observed on migration.
354 Wound healing assays documented a strongly decreased healing time in the *STAG2*-
355 KO A673 clone as compared to parental (Figure 7C, D) but could not be conducted in
356 TC71 cells due to massive detachment of the plate at confluency of parental and
357 isogenic TC71 cells. As the wound healing phenotype was poorly rescued in A673^{SA2r}
358 (Figure 7D), we decided to investigate other Ewing cell lines. We first observed that
359 *STAG2*-KD in A673 cells with two different siRNAs fully reproduced the increased
360 migration observed in the *STAG2*-KO clone (Figure 7E). This early impact on wound
361 healing was further observed using the CHLA-258 Ewing sarcoma cell line (Figure 7F).
362 We also took advantage of a recently engineered human MSC-derived cell line which
363 presents a CRISPR/Cas9 induced t(11;22) translocation leading to an *EWSR1-FLI1*
364 fusion and a deletion of *STAG2* (Sole et al, manuscript under review). This cell line was
365 further engineered to express a DOX-inducible *STAG2* transgene. Induction of *STAG2*
366 expression led to decreased migration that was not observed in control cells (Figure 7G
367 and H). We also explored the invasion ability of *STAG2*-WT and -LOF cells using the
368 transwell assay. Again, *STAG2*-KO or short term -knocked down A673 cells
369 demonstrated increased invasion with rescue in A673^{SA2r} cells (Figure 7I-K). Similarly,
370 CHLA-258 cells showed increased invasion with two different siRNAs (Figure 7L). In
371 addition, collagen embedded aggregates of A673^{SA2m#1} cells generated anisotropic
372 elongated 3D structures (Figure 7M) and displayed (collective) migration prone
373 properties at one leading edge of these structures (Figure 7M, N Supplemental Video 1,
374 2). These properties were not observed in A673^{WT} and A673^{SA2r} cells, which displayed
375 classical spherical 3D structures. Based on migratory properties of *EWSR1-FLI1*^{low}
376 (Chaturvedi et al., 2012; Franzetti et al., 2017) and *STAG2* LOF cells (this work), we
377 questioned whether the *STAG2*-LOF migratory effect was dependent or independent of
378 the decreased *EWSR1-FLI1* activity. We therefore asked if over-expression of *EWSR1*-
379 *FLI1* in *STAG2*-KO cells may counteract their migratory properties. Strikingly,
380 A673^{SA2m#1} cells transduced with an *EWSR1-FLI1* expression lentiviral vector displayed
381 a drastic reduction of their wound invasion properties as compared to empty control
382 transduced cells (Figure 7O). This demonstrates that the *STAG2*-LOF-induced

383 migratory effect is dependent on EWSR1-FLI1 activity. Finally, using a signature of 39
384 genes decreased in A673 and TC71 *STAG2*-LOF cells and displaying EWSR1-FLI1-
385 bound microsatellite sequences in their promoter enhancer chains, we show that a
386 lower expression of these genes is associated with adverse prognosis and with the
387 presence of metastasis (Figure S8E-G). This provides additional evidence that *STAG2*
388 LOF is associated with metastasis in Ewing sarcoma and that this effect is mediated, at
389 least in part, by EWSR1-FLI1. Altogether, these data show that Ewing cells with *STAG2*
390 LOF recapitulate certain phenotypic characteristics of EWSR1-FLI1^{low} cells and
391 particularly their migration phenotype. Importantly, knock-down experiments indicate
392 that this phenotype is an early consequence of *STAG2* LOF.

393

394

395 **DISCUSSION**

396

397 The hypothesis that cohesin dynamically extrudes DNA loops in an ATP-dependent
398 manner has recently been demonstrated *in vitro* (Davidson et al., 2019; Kim et al.,
399 2019). The detection of architectural stripes, either by Hi-C (Fudenberg et al., 2016;
400 Vian et al., 2018), by Micro-C (Hsieh et al., 2020; Krietenstein et al., 2020) or by
401 enriching of the DNA fragments associated with CTCF (this work) strongly supports the
402 hypothesis that such a dynamic process takes place *in vivo* in the context of chromatin
403 and that these stripes, frequently referred as extrusion stripes or extrusion-associated
404 stripes, constitute direct hints of this process. One process of loop extrusion may rely on
405 an initial anchoring of DNA on one of the CTCF loop-flanking site and then progression
406 of the loop by the sliding of DNA until it reaches a conversely oriented CTCF site as
407 indicated in the model shown in Figure 3H. Our comprehensive analysis confirms that
408 these stripes delineate loops flanked by CTCF sites in convergent orientations. It further
409 indicates that most loops can initiate from both flanking sides. However, in a minority of
410 cases, loops preferentially start from one of the flanking CTCF. In these cases, the loop
411 initiator CTCF site is much more intensely bound by core cohesin members suggesting
412 that the presence of these core members is necessary for the initiation of the loop
413 extrusion process, a result which is fully consistent with the *in vitro* data of Davidson

414 and colleagues. Interestingly, the ratio of stripes (reflecting the extrusion process)
415 versus corner peaks (corresponding to fully extruded loops) is highly heterogeneous
416 across the genome suggesting a broad variability in the dynamics of loop extrusion.
417 Our data show that STAG2 and STAG1 demonstrate strikingly different roles in this
418 CTCF-anchored loop extrusion mechanism. Whereas cells expressing only STAG1
419 exhibit a strongly decreased intensity of stripes, cells expressing only STAG2
420 demonstrate dramatically increased intensity of these stripes. This strongly suggests
421 that STAG2 enhances the anchored loop extrusion process while STAG1 is much less
422 efficient in this mechanism, one hypothesis being that STAG1 may promote other types
423 of extrusion mechanisms. While this mechanism was not investigated here, recent
424 papers also indicate that STAG1 and STAG2 have specific roles in genome
425 organization (Arruda et al., 2020; Casa et al., 2020). Altogether, our data are consistent
426 with STAG1 and STAG2 having distinct but complementary functions in the topological
427 shaping of the genome which may account for their synthetic lethality (Benedetti et al.,
428 2017; van der Lelij et al., 2017; Liu et al., 2018). *STAG1* and *STAG2* expression levels
429 across various cancers and normal tissues are quite heterogeneous (Romero-Pérez et
430 al., 2019). It can therefore be anticipated that the ratio between STAG1 and STAG2 at
431 loop boundaries may allow for the relaxation or stabilization of this anchorage and
432 consequently fine tunes the level of enhancer activation within the chromatin loop. In
433 line with the above hypothesis, the increased STAG1/STAG2 ratio in the A673^{SA2r} as
434 compared to A673 parental cells may account for the highly significant but partial
435 rescues of the transcriptional-, stripe- and enhancer interaction- profiles observed in this
436 study. Other cancers, such as urothelial carcinoma, glioblastoma or myeloid leukemia
437 display frequent mutations of *STAG2* (Romero-Pérez et al., 2019). Interestingly,
438 applying our Tweed algorithm to recently published CTCF HiChIP data (GSE111537) in
439 *STAG2*-WT and -KO AML cell models (Smith et al., 2020) also showed a significant
440 decrease of stripe pattern in *STAG2* LOF cells (P-values < 2.2x 10⁻¹⁶). This thus
441 indicates that altered chromatin extrusion upon *STAG2* LOF is not limited to Ewing
442 sarcoma and is also observed in other malignancies.
443 The folding of DNA into chromatin loops is a critical determinant of the function of the
444 genome and particularly of gene regulation. Vian and colleagues recently reported that

445 79% of stripe domains were associated with active enhancers (Vian et al., 2018). We
446 confirm this observation and further show that EWSR1-FLI1-bound microsatellites,
447 which act as neo-enhancers in Ewing sarcoma, are also strongly enriched in such
448 domains. In Ewing cells and in agreement with these observations, the loss of extrusion
449 stripes caused by *STAG2* LOF is associated with important changes in the gene
450 expression pattern, particularly on EWSR1-FLI1 regulated genes. This is not associated
451 with impaired access to DNA of transcription factors or histone modifying enzymes in
452 *STAG2*-KO cells. In contrast, we observed a dramatic decrease of H3K27ac
453 interactions which is predominantly observed within a distance range that corresponds
454 to loop size. Mapping H3K27ac chain interactions from promoter sites enables definition
455 of putative regulatory chains and further identifies those containing EWSR1-FLI1-bound
456 GGAA microsatellites as highly specific for EWSR1-FLI1-regulated genes. Our agnostic
457 regulatory chain analysis now enables a more direct identification of genes regulated by
458 EWSR1-FLI1 binding on GGAA-microsatellites. The latter sequences are enriched in
459 H3K27Ac levels as compared to other enhancers which may explain why they are
460 particularly sensitive to *STAG2* LOF. Interestingly, recent studies in mouse embryos
461 indicates that *STAG2* LOF alters transcription of particular tissue-specific genes (De
462 Koninck et al., 2020). Another publication identifies a specific *STAG2*-*RUNX2* interplay
463 in gene regulation in hematopoiesis (Ochi et al., 2020). Together with our observations
464 in Ewing cells, this strongly supports the hypothesis that the decreased enhancer
465 interactions genome-wide associated with *STAG2* LOF have cell-type-specific
466 consequences depending upon which master transcription factors are active in these
467 cells.

468 Ewing cells demonstrate plasticity between EWSR1-FLI1^{high} and EWSR1-FLI1^{low} states
469 and latter cells are a plausible cause of metastasis (Aynaud et al., 2020; Franzetti et al.,
470 2017). Variation of transcript or protein levels are not the only mechanisms to account
471 for heterogeneous EWSR1-FLI1 activity. Indeed, Wnt/beta-catenin activity was reported
472 to account for metastasis in Ewing tumors through antagonizing EWSR1-FLI1 activity
473 (Pedersen et al., 2016). A similar observation has been made for the YAP/TAZ signaling
474 pathway (Katschnig et al., 2017; Rodríguez-Núñez et al., 2020). Altogether these data
475 indicate that EWSR1-FLI1 activity can vary dynamically, with medium to high activity

476 cells being highly proliferative whereas low activity cells have more limited proliferative
477 ability but have increased mesenchymal properties including propensity to migrate and
478 invade. Our data show that *STAG2* LOF constitutes an alternative mechanism to
479 moderate EWSR1-FLI1 activity and hence potentially increase the invasive and
480 metastatic potential of Ewing cells. Our data also strongly support the hypothesis that
481 *STAG2* LOF effects on migration and invasion are dependent upon a lower activity of
482 EWSR1-FLI1. Indeed, we show not only that these effects can be reverted by
483 increasing EWSR1-FLI1 expression but that a signature, containing the subset of genes
484 down-regulated in *STAG2*-KO cells and containing microsatellite EWSR1-FLI1 binding
485 sites in their regulatory chains, is of strong prognostic significance in Ewing sarcoma.
486 *STAG2* LOF may hence induce a shift of the spectrum of EWSR1-FLI1 activity,
487 increasing the proportion of cells with a low activity, yet keeping the appropriate window
488 of activity for proliferation. *STAG2* LOF Ewing cells might thus combine aggressive
489 features of EWSR1-FLI1^{-high} and ^{-low} cells. Though our data highlight a critical role of
490 *STAG2* LOF in moderating EWSR1-FLI1 activity, it is likely that this alteration may also
491 contribute additional oncogenic functions. In that respect, it is noteworthy that a
492 YAP/TAZ gene set significantly shows up in *STAG2*-LOF up-regulated genes,
493 suggesting that YAP/TAZ activity may also contribute to lower EWSR1-FLI1 activity.
494 *DKK2* and other WNT antagonists (*KREMEN1*, *FRZB*) and agonists (*LRP5*, *FZD8*) are
495 respectively down- and up-regulated upon *STAG2* LOF in most data sets, indicating that
496 Wnt signaling activation may also possibly contribute to cell migration. Other
497 transcription factors (MYC, NF κ B) or signaling pathways (EMT transition, TGF β and
498 EGF) are impacted upon *STAG2* LOF and further studies will be needed to precisely
499 investigate their potential role in *STAG2* LOF associated oncogenesis.

500 In conclusion, this study shows that *STAG2* promotes a dynamic anchored loop
501 extrusion process which likely favors the establishment of cis-interactions between
502 promoters and enhancers, hence unravelling a fundamental role of *STAG2* in gene
503 regulation. In Ewing sarcoma, this process enhances the oncogenic role of EWSR1-
504 FLI1 by the paradoxical moderation of its transcriptional activity.

505

506 **ACKNOWLEDGMENTS**

507 This work was supported by grants from the Institut Curie: the INSERM; the
508 Canceropôle Ile-de-France; the Ligue Nationale Contre le Cancer (Equipe labellisée)
509 and projet de Recherche “Enfants, Adolescents et Cancer”; the Institut National du
510 Cancer (PLBIO16-291), the Fondation ARC, the Agence Nationale de la Recherche
511 (ANR-10-EQPX-03, Institut Curie Génomique d’Excellence (ICGex) and the Société
512 Française de lutte contre les Cancers et les leucémies de l’Enfant et de l’adolescent
513 (SFCE). This project also received support from European funding: ERA-NET
514 TRANSCAN JTC-2011 (01KT1310), ERA-NET TRANSCAN JTC 2014
515 (TRAN201501238) and TRANSCAN JTC 2017 (TRANS201801292), EEC (HEALTH-
516 F2-2013-602856), H2020-IMI2-JTI-201 5-07 (116064 – ITCC P4), H2020-SC1-DTH-
517 2018-1 (SEP-210506374 - iPC). We are indebted to the following associations for
518 providing essential supports : L’Etoile de Martin, la Course de l’Espoir, M la vie avec
519 Lisa, ADAM, Couleur Jade, Dans les pas du Géant, Courir pour Mathieu, Marabout de
520 Ficelle, Olivier Chape, Les Bagouzamanon, l’association Hubert Gouin, Enfance et
521 Cancer, and les Amis de Claire. D.S. is supported by the Institut Curie–SIRIC (Site de
522 Recherche Intégrée en Cancérologie) program. We thank N. Servant, J. Waterfall, A.
523 Zinovyev, MM Aynaud, N. Gruel, V. Berdah, A. Fu, E. Lapouble, E. Tomazou and all
524 members of the Genetics and Biology of Pediatric Cancers laboratory for helpful
525 discussions. We thank H. Kovar and Childhood Cancer Repository for providing Ewing
526 cell lines and C. Pierre Eugene for her experimental assistance as well as P. Legoix, O.
527 Mirabeau, C. Kamoun and E. Barillot for bioinformatic or technical assistance. Finally,
528 we thank J. Waterfall for his diligent proofreading of this paper.

529

530 **AUTHOR CONTRIBUTIONS**

531 D.S., S.Z, K.L-D and L.M. carried out experiments. S.G. and V.H. performed the
532 bioinformatics analyses. D.S., S.G. and V.H. analyzed processed sequencing data. G.P
533 and F.T. provided patient sequencing data. E.B and A.S.F provided EWIma1 models
534 and contributed valuable expertise. S.G implemented the Tweed algorithm and V.H the
535 enhancer interaction algorithm. T.V. analyzed spheroid invasion assays. V.R performed
536 with S.B. sequencing experiments. D.S. originally conceived the project and designed

537 the experiments. D.S. and O.D. supervised the work overall and wrote the paper, all
538 authors reviewed the manuscript.

539 540 **DECLARATION OF INTERESTS**

541 The authors declare no competing interests

542 543 **FIGURE LEGENDS**

544 545 **Figure 1. *STAG2* knock out profoundly alters transcriptomic landscape**

546 (A) Representative western blotting in cellular extracts from isogenic *STAG2* KO
547 (generated with two independent sgRNAs: SA2m#1 and SA2m#2), *STAG1* KO (sgRNA:
548 SA1m#1), *STAG2*-rescued (sgRNA: SA2r) and *STAG2* knock-down (KD) at 48 hours
549 (generated with two independent siRNAs: siSA2#6 and siSA2#8) Ewing sarcoma cells.
550 Color code for sgRNA isogenic models is indicated for each model and kept identical
551 throughout the manuscript. (B) Scaled Venn diagram for modulated genes between
552 *STAG2*-WT and -KO conditions (n=3), total modulated genes for each condition
553 represent the sum of intra-circle numbers, universe includes expressed genes
554 (n=13780), *P* value for intersection was calculated with SuperExact test. (C) Box plots of
555 log₂ fold change for up-, un- and down-regulated genes in *STAG2*-KO and *STAG2*-
556 rescued cells as compared to A673 or TC71 parental cells (n=3 for each model),
557 number of genes is indicated for each category, *P* values: two-tailed paired Wilcoxon
558 test. Box represents the central 50% of data points (interquartile range). Upper and
559 lower whiskers represent the largest and smallest observed values within 1.5 times the
560 interquartile range from the ends of the box. (D) Heat map for core set of commonly up-
561 (left panel) and down- (right panel) regulated genes identified in B, C for *STAG2*-KO
562 and -KD Ewing cell lines. Time after siRNA transfection is indicated at the top, sgRNA
563 and siRNA identifiers at the bottom.

564 See also Figure S1, S2, Table S1.

565 566 **Figure 2. Binding pattern at H3K27ac, EWSR1-FLI1, CTCF and cohesin sites is** 567 **mostly unaffected upon *STAG2* knock out**

568 (A-C) Heatmap representation of ChIP-seq data in *STAG2*-WT or -KO A673 cells for (A)
569 H3K27ac, ranked by ROSE algorithm, (B) EWSR1-FLI1, ranked by peak intensity in
570 GGAA microsatellite (μ sat) and non μ sat sites and (C) CTCF/cohesin members, ranked
571 by *STAG2* peak intensity. Read density is displayed within a 20 kb (H3K27ac) or 4 kb
572 (other marks) window around peak center and color scale intensities are shown in
573 normalized coverage (scale is shown on the bottom of each panel). Presence (black) or
574 absence (grey) of CTCF peaks is indicated on the left. (D) Dot plot for *DKK2* expression
575 ($n=3$) for each model. (E) ChIP-seq binding profiles at the *DKK2* locus for CTCF,
576 EWSR1-FLI1, cohesin members, H3K27ac and H3K4me3 histone marks in *STAG2*-WT
577 or -KO A673 cells. Number of consecutive GGAA repetitions are shown in grey.
578 Promoter-enhancer H3K27ac HiChIP inferred chain at *DKK2* locus is shown.
579 See also Figure S3.

580

581 **Figure 3. High-resolution CTCF HiChIP data at the *DKK2* locus**

582 (A) *STAG2*-WT ($n=5$), (B) *STAG2*-KO ($n=5$), (C) *STAG2*-rescued ($n=3$), (D) *STAG2*-KD
583 72 hrs after siRNA transfection with siCT ($n=2$), (E) siSA2#6 ($n=2$), (F) siSA2#8 ($n=2$)
584 and (G) *STAG1*-KO ($n=2$) A673 cells. Each panel is scaled to total CTCF HiChIP valid
585 pairs at a 5kb/bin resolution. CTCF ChIP-seq profiles are shown above each Juicebox
586 print screens HiChIP data. Dashed box highlights the region with the first CTCF loop
587 associated stripe. (H) scheme illustrating the process of loop extrusion: **1**, upon cohesin
588 loading on chromatin, extrusion can occur possibly in any direction allowing ultimately
589 interactions between CTCF and cohesin. **2**, in this model, chromatin is then anchored
590 by the CTCF/cohesin complex at the left CTCF site and extrusion allows for the
591 formation of an **3**, intermediate- and a **4**, fully- extruded loop. On the right, schemes
592 illustrating CTCF-HiChIP data with loop initiation from the left CTCF/cohesin site and
593 resulting in a X stripe pattern. **4'**, an extrusion process anchored at right CTCF/cohesin
594 site results in a Y stripe pattern.

595 See also Figure S4, Table S4.

596

597 **Figure 4. Genome-wide loop detection highlights the role of *STAG2* in the** 598 **anchored extrusion process**

599 (A) Loop domains prediction at *DKK2* locus in A673-WT CTCF HiChIP data. (B) Density
600 distribution of X and Y stripes for each loop across the genome according to Tweed
601 prediction in CTCF HiChIP data from A673 and TC71 *STAG2*-WT cells. Top scheme:
602 most loops present a symmetric distribution centered on zero indicating a 50% X- and
603 50% Y-stripe based anchorage. Unbalanced CTCF/cohesin signal at one of the two
604 anchorage sites results in asymmetric stripe patterns. (C) Global analysis for the
605 presence of one, two, three or four cohesin peak(s) associated with CTCF peaks in
606 Tweed-detected loop domains presenting predominant X or Y stripes (anchored at
607 either left or right CTCF site) in A673 (n=5) and TC71 (n=2) cells. (D) Box plot of X- or
608 Y- stripe ratios in CTCF HiChIP data between A673^{siCT} and *STAG2*-KD (A673^{siSA2#6},
609 A673^{siSA2#8}) at 72h, *STAG1/2*-KO models (A673^{SA1m#1}, A673^{SA2m#1}, TC71^{SA2m#2}) and
610 parental -WT (A673, TC71) or *STAG2* rescued (A673^{SA2r}) cells, A673 n=18,774; TC71
611 n=14,263. Right scheme: positive-, negative-ratio values indicate a gain or a loss of
612 stripes, respectively, Two-tailed paired t-test on normalized coverage (log10) between
613 parental/siCT and KO/KD or *STAG2*-KO and -rescued models are all highly significant
614 ($P < 2.2 \times 10^{-16}$). Box represents the central 50% of data points (interquartile range).
615 Upper and lower whiskers represent the largest and smallest observed values within 1.5
616 times the interquartile range from the ends of the box.
617 See also Figures S5, S6, Table S4.

618

619 **Figure 5. *STAG2* mutation is globally associated with decreased cis-promoter-**
620 **enhancer and enhancer-enhancer interactions within loops**

621 (A-D) Color coded box plot of comparative analysis of (A-B) H3K27ac ChIP-seq peak
622 intensities and (C-D) H3K27ac HiChIP interactions along promoter-enhancer chains
623 between *STAG2*-WT and -KO as well as between *STAG2*-KO and -rescued cells, *P*
624 values: two-tailed Wilcoxon test. P: promoter, E: enhancer, position in the chain is
625 shown for rank 1 to 5. (E) Comparative analysis of intra super-enhancer interactions in
626 H3K27ac HiChIP data between *STAG2*-WT and -KO or -rescue cells, *P* values: two-
627 tailed t-test. Box represents the central 50% of data points (interquartile range). Upper
628 and lower whiskers represent the largest and smallest observed values within 1.5 times
629 the interquartile range from the ends of the box. (n) for each condition is indicated (F)

630 Top, curve of cumulative percentage of loop presence upon genomic distance (log10
631 scale) in A673 and TC71 (*STAG2*-WT). Loop size threshold: 75% of cumulative loops
632 corresponding to 595kb for A673 (red line) and 340kb (green line) for TC71. Numbers of
633 loop and median loop size for each cell line is indicated. Bottom, percentage of cis-
634 interaction read-pairs upon genomic distance between *STAG2*-WT, -KO and -rescue
635 conditions in CTCF- and H3K27ac HiChIP data. A threshold of 20kb used for H3K27ac
636 chain detection is displayed (blue dashed line). Right, zoom in CTCF- and H3K27ac-
637 HiChIP plot flanking lower and top threshold (highlighted in grey).
638 See also Figure S7, S8, Tables S4, S5, S6.

639

640 **Figure 6. *STAG2* induced genes and EWSR1-FLI1 bound elements are highly**
641 **dependent on cis-regulatory interactions**

642 (A, B) Percentage of depicted categories inside loops in (A) A673 and (B) TC71
643 (*STAG2*-WT). Chi-square test between observed and expected percentage was
644 performed for each category. All adjusted p-values (Bonferroni) were significant. (C-D)
645 Normalized numbers of H3K27ac HiChIP interactions at EWSR1-FLI1 (microsatellite,
646 msat or single GGAA ETS sites) bound or unbound enhancers in (C) A673 and (D)
647 TC71 (*STAG2*-WT), P values: two-tailed paired Wilcoxon test. Log2 FC normalized
648 H3K27ac HiChIP interaction ratio between promoter and first enhancer for down-, un- or
649 up-regulated genes upon *STAG2*-KO in (E) A673 and (F) TC71 (*STAG2*-WT), Box
650 represents the central 50% of data points (interquartile range). Upper and lower
651 whiskers represent the largest and smallest observed values within 1.5 times the
652 interquartile range from the ends of the box. P values: two-tailed t-test.

653

654 **Figure 7. *STAG2* mutation promotes cell migration**

655 (A, B) Immunofluorescence with phalloidin (red), paxillin (green) and merged pictures in
656 A673 and A673^{SA2m#1} cells and (B), box plot of paxillin associated focal adhesion loci
657 (FA) per cell in A673, A673^{SA2m#1} and A673^{SA2r}, P value: two-tailed t-test. (C)
658 representative wound healing assay picture at initial and 24h time points for A673 and
659 A673^{SA2m#1} cells, computed migration front line is shown (yellow). (D-H) Quantification
660 of wound surface (cell free) along the assay for: (D) A673, A673^{SA2m#1} and A673^{SA2r}

661 cells; (E) A673-, (F) CHLA-258 transfected cells with siCT or siSA2#6 and siSA2#8, (G)
662 *STAG2*- or (H) empty vector control- inducible EWlma1 cells (MSC cells engineered to
663 present *EWSR1-FLI1* translocation and *STAG2* mutation) in the presence or absence of
664 doxycycline to induce *STAG2* expression, *P* value: two-tailed t-test. * *P*<0.05 (I)
665 macroscopic and magnified picture (10X) of a representative membrane from a Boyden
666 chamber migration assay with A673 (n=3) and A673^{SA2m#1} (n=3) cells. (J-L) Box plots of
667 cells per field in A673, A673^{SA2m#1} and A673^{SA2r} cells at 48h in Boyden chamber
668 migration assay (J). Box plots of cells per field in (K) A673-, (L) CHLA-258- cells upon
669 siCT, siSA2#6 and siSA2#8 transfection for 48hrs and plated 24hrs in Boyden chamber
670 migration assay. *P* value: two-tailed t-test, Box represents the central 50% of data
671 points (interquartile range). Upper and lower whiskers represent the largest and
672 smallest observed values within 1.5 times the interquartile range from the ends of the
673 box. (M, N) Invasion assay, (M) representative pictures at indicated time post-collagen
674 embedding and (N) eccentricity and area measurements across time of spheroids
675 generated with A673, A673^{SA2m#1} and A673^{SA2r} cells. (O), Left, quantification of wound
676 surface along the assay for A673, A673^{SA2m#1} and A673^{SA2m#1} transduced with either
677 empty (pCDH1, yellow) or *EWSR1-FLI1* (pCDH1-EF1, green) cells. *P* value: two-tailed
678 t-test. * *P*<0.05, right, a representative western blot for *EWSR1-FLI1* (EF1), *STAG2* and
679 beta-actin (β -act).

680 See also Figure S8, Supplemental Videos 1, 2

681

Table 1. Top20 gene sets enriched in STAG2 proficient or deficient Ewing

682

sarcoma data sets. See also Figure S2, Tables S2, S3.

enriched in	RANK	Gene sets (18 889 gene sets from MSigDB or published data)	NES					pVal				FDR						
			1	2	3	4	AVG NES	1	2	3	4	1	2	3	4			
			siSA2	Isogenic	Cell Lines	Tumors		•: pVal <0.01				•: FDR <0.2						
STAG2 proficient	1	IC_EWS	2.0	1.7	2.2	2.1	2.00	•	•	•	•	•	•	•	•	•	•	•
	2	KERLEY_RESPONSE_TO_CISPLATIN_UP	1.8	1.8	2.0	2.0	1.90	•	•	•	•	•	•	•	•	•	•	•
	3	VERHAAK_GLIOMASTOMA_MESENCHYMAL	1.9	1.9	1.8	1.7	1.80	•	•	•	•	•	•	•	•	•	•	•
	4	HALLMARK_TNFA_SIGNALING_VIA_NFKB	1.4	1.5	1.7	2.6	1.80	•	•	•	•	•	•	•	•	•	•	•
	5	GALINDO_IMMUNE_RESPONSE_TO_ENTEROTOXIN	1.7	1.6	1.7	2.1	1.79	•	•	•	•	•	•	•	•	•	•	•
	6	JOHNSTONE_PARVB_TARGETS_3_UP	1.8	1.3	1.8	2.2	1.77	•	•	•	•	•	•	•	•	•	•	•
	7	NAGASHIMA_NRG1_SIGNALING_UP	1.4	1.6	1.7	2.3	1.77	•	•	•	•	•	•	•	•	•	•	•
	8	RIGGI_EWING_SARCOMA_PROGENITOR_UP	1.9	1.5	1.9	1.8	1.76	•	•	•	•	•	•	•	•	•	•	•
	9	KINSEY_TARGETS_OF_EWSR1_FLI1_FUSION_DN	1.8	1.4	1.6	2.1	1.73	•	•	•	•	•	•	•	•	•	•	•
	10	NAGASHIMA_EGF_SIGNALING_UP	1.1	1.5	1.9	2.4	1.72	•	•	•	•	•	•	•	•	•	•	•
	11	ACOSTA_PROLIFERATION_INDEPENDENT_MYC_TARGETS_DN	1.7	1.5	1.8	1.8	1.71	•	•	•	•	•	•	•	•	•	•	•
	12	AMIT_DELAYED_EARLY_GENES	1.4	1.4	1.8	2.2	1.69	•	•	•	•	•	•	•	•	•	•	•
	13	FISCHER_DIRECT_P53_TARGETS_META_ANALYSIS	1.3	1.6	1.8	2.0	1.69	•	•	•	•	•	•	•	•	•	•	•
	14	REACTOME_FCER1_MEDIATED_CA_2_MOBILIZATION	1.3	1.7	1.2	2.3	1.64	•	•	•	•	•	•	•	•	•	•	•
	15	MIYAGAWA_TARGETS_OF_EWSR1_ETS_FUSIONS_UP	1.8	1.4	1.5	1.9	1.63	•	•	•	•	•	•	•	•	•	•	•
	16	HUANG_DASATINIB_RESISTANCE_UP	1.7	1.6	1.7	1.6	1.63	•	•	•	•	•	•	•	•	•	•	•
	17	KAUER_EWS-FLI1_DOWN	1.9	1.6	1.5	1.5	1.63	•	•	•	•	•	•	•	•	•	•	•
	18	PLASARI_TGFB1_TARGETS_1HR_UP	1.6	1.5	1.7	1.7	1.63	•	•	•	•	•	•	•	•	•	•	•
	19	GSE21546_UNSTIM_VS_ANTI_CD3_STIM_ELK1_KO_DP_THYMOCYTES_UP	1.1	1.2	2.0	2.2	1.62	•	•	•	•	•	•	•	•	•	•	•
	20	SMIRNOV_RESPONSE_TO_IR_2HR_UP	1.4	1.8	1.5	1.8	1.61	•	•	•	•	•	•	•	•	•	•	•
STAG2 deficient	1	CHR1Q44	-1.7	-1.0	-2.2	-2.1	-1.75	•	•	•	•	•	•	•	•	•	•	•
	2	PYEON_CANCER_HEAD_AND_NECK_VS_CERVICAL_UP	-1.2	-1.3	-2.0	-2.1	-1.65	•	•	•	•	•	•	•	•	•	•	•
	3	KEGG_ASTHMA	-1.3	-1.7	-1.7	-1.8	-1.61	•	•	•	•	•	•	•	•	•	•	•
	4	VERHAAK_GLIOMASTOMA_PRONEURAL	-1.4	-1.5	-1.9	-1.6	-1.60	•	•	•	•	•	•	•	•	•	•	•
	5	VILLANUEVA_LIVER_CANCER_KRT19_UP	-0.9	-0.8	-2.0	-2.4	-1.54	•	•	•	•	•	•	•	•	•	•	•
	6	PYEON_HPV_POSITIVE_TUMORS_UP	-0.9	-1.4	-1.6	-2.2	-1.51	•	•	•	•	•	•	•	•	•	•	•
	7	KUNINGER_IGF1_VS_PDGF_B_TARGETS_UP	-0.8	-1.3	-1.3	-2.5	-1.49	•	•	•	•	•	•	•	•	•	•	•
	8	TOYOTA_TARGETS_OF_MIR34B_AND_MIR34C	-1.0	-1.1	-1.6	-2.2	-1.48	•	•	•	•	•	•	•	•	•	•	•
	9	CHR2P25	-1.5	-1.8	-1.4	-1.1	-1.47	•	•	•	•	•	•	•	•	•	•	•
	10	RICKMAN_TUMOR_DIFFERENTIATED_WELL_VS_POORLY_UP	-1.2	-1.0	-1.9	-1.7	-1.47	•	•	•	•	•	•	•	•	•	•	•
	11	MIR4800_5P	-1.2	-1.6	-1.8	-1.2	-1.45	•	•	•	•	•	•	•	•	•	•	•
	12	GO_OLFACTORY_RECEPTOR_ACTIVITY	-1.6	-1.6	-1.3	-1.3	-1.45	•	•	•	•	•	•	•	•	•	•	•
	13	GO_CHROMATIN_REMODELING	-1.6	-0.9	-1.4	-1.9	-1.44	•	•	•	•	•	•	•	•	•	•	•
	14	MEISSNER_BRAIN_HCP_WITH_H3_UNMETHYLATED	-1.3	-1.7	-1.4	-1.4	-1.44	•	•	•	•	•	•	•	•	•	•	•
	15	RODRIGUEZ-NUNEZ_YAP+TAZ_UP	-1.2	-1.9	-1.6	-1.1	-1.44	•	•	•	•	•	•	•	•	•	•	•
	16	CHR15Q25	-1.1	-1.4	-1.5	-1.8	-1.43	•	•	•	•	•	•	•	•	•	•	•
	17	CHR20Q12	-1.8	-1.1	-1.4	-1.4	-1.41	•	•	•	•	•	•	•	•	•	•	•
	18	GO_G0_TO_G1_TRANSITION	-1.3	-1.2	-1.5	-1.6	-1.41	•	•	•	•	•	•	•	•	•	•	•
	19	MITSIADES_RESPONSE_TO_APLIDIN_DN	-0.8	-0.9	-2.0	-1.9	-1.41	•	•	•	•	•	•	•	•	•	•	•
	20	REACTOME_TRANSCRIPTIONAL_REGULATION_BY_E2F6	-1.3	-1.2	-1.4	-1.8	-1.41	•	•	•	•	•	•	•	•	•	•	•

683

684

685

686

687

688

689

690

1, Ewing sarcoma cell lines (A673, CHLA-10, CHLA-258, EW1 and TC71), siCT (n=10) vs siSA2#6/8 (n=20) at 72h. 2, Ewing sarcoma isogenic lines, A673WT, TC71WT (n=6) vs A673SA2m#1, TC71SA2m#1, TC71SA2m#2 (n=9). 3, Panel of Ewing sarcoma cell lines, STAG2-WT (n=14) vs STAG2-mutated (n=15). 4, Panel of Ewing sarcoma tumors, STAG2-WT (n=32) vs STAG2-mutated (n=13). NES, Normalized Enrichment Score; pVal, p value; FDR, false discovery rate.

691 **STAR ★ METHODS**

692 **RESOURCE AVAILABILITY**
693

694
695 **LEAD CONTACT**

696 Further information and requests for resources and reagents should be directed to and
697 will be fulfilled by the Lead Contact Olivier Delattre (olivier.delattre@curie.fr).

698
699 **MATERIALS AVAILABILITY**

700 Availability of the isogenic *STAG1* and *STAG2* models generated in this study is
701 subjected to a Material Transfer Agreement.

702
703 **DATA AND CODE AVAILABILITY**

704 All data reported in this study are available at the Gene Expression Omnibus (GEO)
705 under accessions GSE133228, GSE34620 and at the European Genome-phenome
706 Archive (EGA) under accession EGAS00001003333. All other data, custom code and
707 materials are available from the corresponding authors upon reasonable request.

708
709 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

710
711 **Tumor Cell Lines**

712 The Ewing sarcoma A673 cell line was obtained from the American Type Culture
713 Collection (ATCC) and the Ewing sarcoma TC71 cell line was obtained from the
714 German Collection of Microorganisms and Cell Cultures (DSMZ). EW1 cell line was
715 obtained from International Agency for Research on Cancer (IARC), CHLA-10 and
716 CHLA-258 cell lines were obtained from Childhood Cancer Repository (COG
717 Repository). EWIma1 cells were generated in Erika Brunet laboratory by Anna Sole
718 Ferre at Imagine Institute in Paris (Sole et al. under review). EWIma1 STAG2
719 (EWIma1^{SA2}) and empty (EWIma1^{Empty}) tetracycline inducible model cells were
720 generated by transfecting EWIma1 cells with respectively empty pCW57-GFP-2A MCS
721 vector (Addgene) or pCW57-GFP-2A-MCS-STAG2 vector. Transfection was performed
722 with the Amaxa Nucleofector I Device (Lonza) according to the manufacturer's protocol

723 and cells were selected with puromycin at 1µg/mL. Cells were cultured in (A673) DMEM
724 (GE Healthcare, SH30022.01) or (TC71 and EW1) RPMI (GE Healthcare, SH30027.01,)
725 supplemented with 10% fetal bovine serum (GE Helthcare, SV30160.03). CHLA-10 and
726 CHLA-258 were cultured in IMDM (Gibco, 12440-53) supplemented with 20% fetal
727 bovine serum and 1X Insuline-Transferrin-selenium (Gibco, 41400-045). EWIma1^{SA2}
728 and EWIma1^{Empty} were cultured in Alpha-MEM (Gibco, 22561021) supplemented with
729 10% of MSC qualified FBS (Gibco, 12662029), 1% L-glutamine (Gibco, 25030024) and
730 1% Penicillin-streptomycine (Gibco, 15140122). When available, STR profiling proved
731 each cell line matched with the reference profile provided by ATCC and DSMZ,
732 respectively. Cells were routinely tested negative for mycoplasma contamination by
733 qPCR (VenorGeM qEP (11-9250, Minerva Biolabs).

734

735 **Patient samples**

736

737 Samples were stored in a tumor bank at the Institut Curie. The study was approved by
738 the Institutional Review Board of the Institut Curie (Paris, France) and by the regional
739 ethics committees (Comité de Protection de Personnes) from Kremlin Bicêtre (Project
740 n°99-25, June 9th 1999) and Ile-de-France I (GenEwing n° IC 2009-02). Written
741 informed consent was obtained. Most patients were treated according to Euro-Ewing 99
742 or EuroEwing 2012 protocols.

743

744 **METHOD DETAILS**

745

746 **CRISPR/Cas9 based genome-editing**

747 sgRNA guides were designed using Crispor tool (<http://crispor.tefor.net/>).

748 Sense and antisense primers matching sgRNA regions in *STAG1* (SA1m#1: forward, 5'-
749 ACACCGATGTGCCGAGTACACCAAGG-3'; reverse, 5'-
750 AAAACCTTGGTGTACTCGGCACATCG-3'), *STAG2* (SA2m#1: forward, 5'-
751 ACACCGCGACATACAAGCACCCCTGGCG-3'; reverse, 5'-
752 AAAACGCCAGGGTGCTTGTATGTGCGCG-3' and SA2m#2: forward, 5'-
753 ACACCGATTTGACATACAAGCACCCG-3'; reverse, 5'-

754 AAAACGGGTGCTTGTATGTGCGAAATCG-3') and *STAG2* mutated (SA2r: forward, 5'-
755 ACACCGTTTTCGACATACAAGCACCCCTG-3'; reverse, 5'-
756 AAAACAGGGTGCTTGTATGTGCGAAACG-3')

757 Loci were annealed and cloned into MLM3636 vector. MLM3636 was a gift from Keith
758 Joung (Addgene plasmid # 43860; <http://n2t.net/addgene:43860>;
759 RRID:Addgene_43860). Phosphorylation and annealing was performed by mixing 1 µL
760 of each sgRNA at 100 µM (Eurofin), 1 µL of 10X T4 ligation buffer (B0202S, New
761 England Biolabs), 1 µL of PNK (EK0031, Fermentas), 6 µL of water and using the
762 following program: 37 °C for 30 min; 95 °C for 5 min; ramp down to 25 °C at 5 °C/min. 2
763 µL of 1:200 diluted sgRNA were cloned in 100ng of MLM3636 with 2 µL of 10X Tango
764 buffer (BY5, Fermentas), 1 µL of DTT (D0632, Sigma) and ATP (U120D, Promega)
765 each at 10mM, 1 µL of BsmB1 (R05805, New England Biolabs), 0.5 µL of T4 DNA
766 Ligase (M0202S, New England Biolabs) and 10.5 µL of water. The cloning was
767 performed using the following program: 6 cycles of 37°C for 5 min, 21°C for 5
768 min. Transformation was performed with 5 ng of DNA using Stellar™ competent cells
769 (636763, Ozyme) according to manufacturer recommendations.

770 Cells were transfected using Amaxa Nucleofector™ (Lonza) following manufacturer
771 instruction. Briefly, one million of cells were co-transfected with 2 µg of MLM3636-
772 SA1m#1 or MLM3636-SA2m#1 or MLM3636-SA2m#2 plasmid and 2 µg of pCas9_GFP
773 plasmid (gift from Kiran Musunuru, Addgene plasmid # 44719 ;
774 <http://n2t.net/addgene:44719> ; RRID: Addgene_44719) and 1 µg of pCDH1-CMV-MCS-
775 EF1-Puro plasmid (System Biosciences, CA) providing puromycin resistance in V
776 Nucleofector™ solution using T020 program. 24 hours after transfection, the cells are
777 selected during 2 days with 0.5 µg/mL or 1 µg/mL of puromycin for TC71 and A673 cell
778 lines respectively. Transfected cells were then cloned and screened for knockout using
779 Western blot to assess *STAG1* and *STAG2* expression. For rescue experiment,
780 A673SA2m#1 cells were co-transfected with 2 µg of MLM3636-SA2r, 2 µg of single
781 stranded phosphorotioate modified oligonucleotide template (SA2t: 5'-
782 ctTCTTACAGGATTGTCTGACTCACAAGTCAGAGCATTTCGACATACAAGCACGCTA
783 GCAGGTCGGTATTTAGAAATATTTTCTGCATAttg -3', nucleotides with
784 phosphorotioate modification are shown in lower case), 2 µg of pCas9_GFP plasmid

785 and 1 µg of pCDH1-CMV-MCS-EF1-Puro plasmid as described above and isolated as
786 for the other isogenic models. All isogenic models were validated using Sanger
787 sequencing. DNA was extracted from clones of interest using QiAamp®DNA mini kit
788 (51304, Qiagen) according to the manufacturer's instructions. PCR reaction was
789 performed to amplify sgRNA targeting region in *STAG1* or *STAG2* loci (1619-*STAG1*:
790 forward, 5'-CAGCATCCTCAAGGCTGTGA-3'; reverse, 5'-
791 TCAGTGGAAGTGAAGAAGCTCT-3' and 1460-*STAG2*: forward, 5'-
792 AGAGCTGAAGTGTTTCAGAGGT-3'; reverse, 5'-AGGAATTCGCAGGAGGGATG-3')
793 with 200 ng of DNA, 2.5 µL of 10X TP Gold (4311806, Thermo Fisher Scientific), 1.5 µL
794 of MgCl₂ 25mM, 2 µL dNTP 2.5mM (U1420, Promega), 1.5 µL forward and reverse
795 primers (10 µM) and 0.12 µL Amplitaq Gold®. PCR was performed using the following
796 program: 95 °C for 12 min, 35 cycles of (94 °C for 15 s, 57°C for 30 s and 72°C for 2
797 min) and 72 °C for 7 min. PCR products were purified (Nucleofast 96 PCR plate
798 743100.50, Macherey Nagel) and Sanger sequenced validated using their respective
799 forward and reverse primers.

800

801 **siRNA**

802 Cells were transfected using lipofectamine RNAiMAX (Invitrogen, ref 13778030) with
803 siCT (ON-TARGET plus Non-targeting Control siRNA #1, Dharmacon, ref D-001810-01-
804 50, or All stars negative Control siRNA, ref 1027281, Qiagen), siSA2#6 (siRNA Human
805 *STAG2*, Dharmacon, ref J-021351-06-0050), siSA2#8 (siRNA Human *STAG2*,
806 Dharmacon , ref J-021351-08-0050), *siEWSR1-FLI1* fusion type 1 (7/6) for A673, TC71
807 and CHLA-10 or type 2 (7/5) for EW1 (Qiagen, custom order). For a 6 well plate, siRNA
808 transfection mix was prepared by adding 3µl of Lipofectamine RNAiMAX to 230µl of
809 Opti-Mem (Thermo Fisher, ref 31985062) and combining it for 12 to 20 minutes with 233
810 µl of Opti-Mem and siRNA mix. This mix was then added to 1.9 mL of respective cell
811 media. All experiments were conducted without antibiotics and scale up when
812 necessary. The cells were harvested after 24, 48 or 72 hours post transfection.

813

814 **Cell infection**

815 Lentivirus was produced in HEK293T cells (from ATCC) as previously described
816 (Surdez et al., 2012) with pCDH1 empty control or pCDH1-EWSR1-FLI1 vectors
817 (Guillon et al., 2009). A673^{SA2m#1} cells were infected with a multiplicity of infection
818 ranging from 20 to 50. 24h post infection, cells were selected with puromycin for 2
819 weeks using 1 µg/ml puromycin (ant-pr-1; InvivoGen) before starting experiments.

820

821 **RNA extraction, cDNA and RT-QPCR**

822 RNA was extracted with the RNeasy Plus Mini Kit (Qiagen, ref 74134) and reverse-
823 transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied
824 Biosystems, ref 4368814). RT-QPCRs were performed using Power SYBR Green PCR
825 Master Mix (Applied Biosystems, ref 4367659). Oligonucleotides were purchased from
826 MWG Eurofins Genomics and listed in oligonucleotide section of the key resource table.
827 Reactions were run on an CFX384 Touch Real-Time PCR Detection System instrument
828 (Bio-Rad) and analyzed using the Bio-Rad CFX Manager 3.1 software.

829

830 **Immunoblotting**

831 Cells were trypsinized, counted, washed with ice-cold PBS and lysed in Laemmli buffer
832 (50 mM Tris-HCL, 2.5 mM EDTA, 2.55 mM EGTA, 2% SDS 20%, 5% Glycerol, 1%
833 Bromophenol blue, protease inhibitor cocktail tablets and 2 mM DL-Dithiothreitol
834 solution) at 10 million cells/ml. Protein lysates were sonicated and denatured at 95 °C
835 for 5 min and electrophorated on 4-15% Mini-PROTEAN®TGXTM gels (456-1084, BIO-
836 RAD), transferred onto nitrocellulose membranes (1704159, BIO-RAD). Membranes
837 were incubated overnight at 4°C with mouse anti-STAG2 (1:1,000, Santa Cruz
838 Biotechnology, sc-81852), goat anti-STAG1 (1:5,000, ab4457, Abcam), rabbit anti-H3
839 (1:50,000, ab1791, Abcam), mouse anti-β-Actin (1:20,000, A5316, Sigma-Aldrich),
840 rabbit anti-FLI1 antibody (1:1,000, ab133485, Abcam) and rabbit anti-H3K27ac
841 (1:1,000, ab4729, Abcam). Then membranes were incubated 1h at room temperature
842 with respective anti-rabbit, anti-mouse immunoglobulin G horseradish peroxidase (HRP)
843 coupled secondary antibody (1:3,000, NA934 or NXA931, respectively; GE Healthcare)
844 or anti-goat IgG-HRP (1: 10,000, SC-2354, Santa Cruz). Proteins were visualized using

845 SuperSignal™ West Pico Plus (34580, Thermo Scientific) and ChemiDoc™ Imaging
846 System (BIO-RAD).

847

848 **RNA-seq**

849 RNA was extracted from independent experiments using RNeasy Plus Mini Kit (Qiagen,
850 ref 74134) according to manufacturer recommendations. RNA sequencing libraries were
851 prepared from 1 µg of total RNA using the TruSeq Stranded mRNA Library preparation
852 kit (Illumina, ref 20020594) following manufacturer instructions. Sequencing was carried
853 out using 2x100 cycles (paired-end reads 100 nucleotides) for all samples on Illumina
854 HiSeq2500 or NovaSeq6000 instruments. Reads were aligned with STAR 2.5.3 (Dobin
855 et al., 2013) to the human genome (GRCh37/hg19 version). We used the count matrix
856 generated by STAR using the human gene annotation v19 of GENCODE. DESEQ2
857 1.20.0 (Love et al., 2014) was used to normalize data and performed differential
858 analysis with the Wald test. The p-value was adjusted using Benjamini-Hochberg. For
859 differential analysis, we used an adjusted p-value < 0.01 and $|\log_2(\text{fold change})| > 1$.
860 We considered a gene expressed if the normalized expression is higher than 10.

861

862 **Tumors and cell lines expression data**

863 RNA-seq dataset from Ewing sarcoma patients were previously published
864 (EGAS00001003333). Microarray data were profiled using HG-U133-Plus2 arrays
865 (Affymetrix) and processed as previously published (Postel-Vinay et al., 2012; Surdez et
866 al., 2012). For Microarray analyses, Ewing tumors samples from GSE34620 and from
867 this study (GSE133228) were used. RNA-seq of this study (GSE133228) were
868 generated from previously described Ewing cell lines (Batra et al., 2004; Dauphinot et
869 al., 2001; Kovar et al., 1997; Tirode et al., 2014). *STAG2* mutational status of tumors
870 (Tirode et al., 2014) and cell lines (Brohl et al., 2014; Crompton et al., 2014; Tirode et
871 al., 2014) was previously determined.

872

873 **ChIP-seq**

874 Chromatin Immunoprecipitation (ChIP) experiments were performed following
875 manufacturer instructions using iDeal ChIP-seq kit for transcription factors and for

876 histones (Diagenode) with respectively rabbit polyclonal anti-FLI1 antibody (ab15289,
877 Abcam), rabbit polyclonal anti-RAD21 (ab992, Abcam), goat polyclonal anti-STAG1
878 (ab4457, Abcam), goat polyclonal anti-STAG2 (ab4463, Abcam), anti-CTCF (provided
879 in iDeal ChIP-seq kit Diagenode), rabbit polyclonal anti-SMC1 (A300-055A, Bethyl
880 laboratories), rabbit polyclonal anti-H3K4me3 (C15410003, Diagenode) and rabbit
881 polyclonal anti-H3K27ac (ab4729, Abcam). Briefly, Ewing cell lines were fixed for 10
882 minutes with 1% of methanol-free formaldehyde (28908, Thermo-Scientific). Chromatin
883 was sonicated (Bioruptor, Diagenode) for 20 cycles (30-sec on, 30-sec off) set at
884 position “high” to generate DNA fragments with an average size around 150-300pb. For
885 ChIP sequencing, libraries were generated using TruSeq ChIP library preparation kit
886 (Illumina) and sequenced on Illumina HiSeq 2500 or NovaSeq 6000 (single end, 100
887 bp). Reads were aligned to human reference genome (GRCh37/hg19) with bowtie2
888 2.2.9 (Langmead and Salzberg, 2012). Uninformative reads (multimapped reads,
889 duplicated reads and reads with low mapping score) were filtered out with samtools 1.3
890 (Li et al., 2009). Peaks were called with MACS2 2.1.1 (Zhang et al., 2008) with the
891 option narrow for FLI1, CTCF and cohesin members ChIP-seq and broad for histone
892 marks. Genomic regions containing four or more consecutive GGAA sequences were
893 considered as microsatellites (msat) and divided into FLI1-bound and -unbound
894 categories. For each cell line, ChIP-seq were normalized according to their respective
895 input DNA sample. The ChIP-seq signal tracks were generated by macs2 with bdgcmp
896 option (and -m FE to compute fold enrichment between the ChIP and the control).
897 Then, we run bedGraphToBigWig to convert the file to a binary format (BigWig). To
898 identify super-enhancers, we apply ROSE 0.1 algorithm on the H3K27ac peaks (Whyte
899 et al., 2013). For the heatmap, the region was binned (50 bp/bin) around the reference
900 peaks and the normalized coverage was computed for each bin.

901

902 **HiChIP**

903 Several adaptations were made and are detailed below in an overview of the otherwise
904 original protocol described by Mumbach and colleagues (Mumbach et al., 2016).
905 Experiments were performed as independent biological replicates. Cells were plated 2
906 days before starting the experiment to reach a 90-95% confluence at the time of fixation

907 in 1% of methanol-free formaldehyde (28908, Thermo-Scientific) in freshly prepared
908 respective cell culture media (20ml/T150 flask). Cells were fixed at room temperature
909 for 10 minutes under gentle shaking platform (50mvt/min). 2 mL of glycine solution (2M)
910 were added and cells were incubated for 5 additional minutes at room temperature
911 under gentle shaking platform. The supernatant was removed and cells were washed 3
912 times with PBS at room temperature. Cells were scraped vigorously with ice cold PBS
913 supplemented with protease inhibitor cocktail tablets (11836145001, Roche), and
914 flushed five time through a syringe with a 21 gauge needle (301155, BD Microlance). In
915 situ contact libraries were performed starting from 15 million nuclei digested overnight at
916 37°C with Mbo1 (R0147M, New England Biolabs). After proximity ligation (4 h at room
917 temperature), the nuclear pellet was sonicated and the chromatin immunoprecipitation
918 step was performed using the iDeal ChIP-seq kit for transcription factors (Diagenode)
919 according to the supplier's recommendation with some modification. Nuclei were
920 resuspended in IL1b and IL2 buffer following Diagenode protocol and all centrifugations
921 were performed at 4°C for 5 min at 1,950 RCF for these steps. The chromatin was
922 sonicated (5 million nuclei per tube) with Bioruptor pico (Diagenode) for 10 cycles (30-
923 sec on, 30-sec off) set at position "high". The tubes were pooled and the chromatin was
924 clarified by centrifugation at 4°C for 10 min at 16,000 RCF. Sonicated chromatin from 11
925 and 3 million nuclei were used respectively for CTCF and H3K27ac immunoprecipitation
926 step using the equivalent of 3 ChIP reactions pooled in one tube for each HiChIP
927 reaction (final volume was 1050 µl/tube). Immunoprecipitation was then carried out
928 following Diagenode kit instructions by multiplying all reagents by a factor 3 until end of
929 elution step (50µl).

930 Biotin capture was performed as previously described (Mumbach et al., 2016) and
931 library for illumina sequencing were prepared using 10 ng of chromatin and 0.5 µL of
932 TN5 (15028211, Illumina). A first PCR in a final volume of 50µl with 5 cycles was
933 performed (72 °C for 5 min, 98 °C for 1 min, then 5 cycles at 98 °C for 15 s, 63 °C for 30
934 s, and 72 °C for 1 min). To determine how many additional PCR cycles were required
935 for optimal library preparation, a QPCR with: 5µL of the first PCR product, 1 µL of
936 Nextera Ad1_noMX and Nextera Ad2.X (each at 1.25 µM), 5 µl of phusion HF 2X
937 (M0531S, New England Biolabs), 0,75 µL Evagreen® 20X (31000, Biotium) and 2,25 µL

938 of water was runned (same program as first PCR but with 30 cycles). The optimal
939 number of additional cycle was determined for each library by setting a threshold just
940 before reaching the end of exponential amplification step. The 45µl left from the first
941 PCR were further amplified using additional PCR cycles as determined above. Size
942 selection was performed using Ampure XP beads (A63881, Beckman Coulter) to
943 capture fragments greater than 300 bp. Libraries were quantified and analyzed using
944 Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and LabChIP (Perkin Elmer). A
945 first validation of HiChIP experiments was performed using 150 bp paired-end
946 sequencing on MiSeq-microV2-300-PE150 (Illumina). Deep 75 bp paired-end
947 sequencing was then performed on HiSeq 2500 or NovaSeq 6000 systems (Illumina).
948 HiC-Pro 2.10.1 (Servant et al., 2015) pipeline was used to map reads against the
949 human genome (GRCh37/hg19 version), assign reads to restriction fragments, remove
950 duplicate pairs, filter out invalid interaction products and generate a 5kb interaction
951 matrix of valid pairs (read pairs mapping on two different restriction fragments, see
952 Servant et al., 2015 for more details). All HiChIP data were ultimately normalized to the
953 number of valid cis-interactions for subsequent analyses. HiC-Pro was also used to
954 generate “.hic” file for data visualization in Juicebox software (Durand et al., 2016; Rao
955 et al., 2014).

956

957 **Loop prediction by Tweed**

958 To detect loops from CTCF HiChIP data, we developed a new algorithm called Tweed
959 (Sup. Figure 5). Loops can be defined by the presence of CTCF sites at both ends, by
960 increased signal as compared to background and by the presence of a corner peak at
961 the intersection of a X and a Y stripe. First, we identified all CTCF peaks using the
962 MACS algorithm on ChIP-seq data. All candidate stripes starting on a CTCF site and
963 ending on all possible CTCF sites within a limit of 2.5 Mb were then considered. We
964 excluded from the analysis a 75 kb (15 bins) region close to the diagonal where the
965 signal is difficult to interpret. We then used two different approaches. The first one was
966 based on the expected increased signal of stripes (enrichment method). This method
967 compared the total intensity of each candidate stripe with that of a control region defined
968 by the 10 flanking parallel lines of identical size (the control region was defined

969 upstream or downstream the candidate X and Y stripes, respectively). The signal
970 intensity of a stripe was then normalized along its entire length. For each control
971 regions, we also calculated the median value of the normalized intensity (10 lines). For
972 each stripe, we defined a control value corresponding to the median values of all control
973 regions of equivalent size. This last step allowed to avoid local bias and not to be too
974 stringent in the definition of stripes. The intensity of each candidate stripe of a given
975 length was then compared with this value. The intersection of X and Y stripes was
976 computed based on their coordinates. A loop was defined when both X and Y values
977 were > 2 or when either of the values was > 3 .

978 The second approach (second derivative method) was based on the detection of
979 decreased interaction intensity at the end of the stripe (frequently defined as the corner
980 peak). This inflection point can be detected by a negative value of the second derivative
981 of the function defined by the cumulative curve of interaction counts along the stripe. If
982 the value was negative on the corner, or lower than -10 in the ± 1 adjacent bin to the
983 corner, a candidate loop was predicted. Only loops predicted by both methods and with
984 mean intensity higher than 2 reads/bin were kept. For adjacent loops (i.e. differing only
985 by 1 bin at their CTCF boundary sites), the loop with the highest coverage intensity was
986 kept. At this point, most predicted loops presented at their boundaries a strong
987 prevalence of cohesin members and convergent CTCF motif orientations as determined
988 with the FIMO tool from MEME suite (Grant et al., 2011) using the CTCF motif
989 (MA0139.1 from Jaspar database) (Sup. Figure 5e-g). Final filtering for presence of
990 convergent CTCF motif orientations and of at least one member of the cohesin complex
991 at each CTCF boundary sites was therefore used for final loop selection. At the end,
992 18,774 (A673WT) and 14,263 (TC71WT) loops were detected by our Tweed algorithm.
993 Final loops were classified in two categories: balanced intensity of X and Y stripes or
994 asymmetric stripes with predominant X or Y pattern (defined by a \log_2 X/Y stripe
995 coverage ratio higher or lower than the mean value of all stripes intensity $\pm 2SD$,
996 respectively). To compare loop coverage for each loop detected in *STAG2* WT
997 conditions (Fig. 4D), we computed the stripe coverage and normalized by the total
998 number of cis-interactions as well as by the loop size. We then assessed the difference

999 between *STAG1/2* -WT and -KO conditions (paired T-test on
1000 $\log_{10}(\text{normalised_coverage})$).

1001

1002 **Interaction chains**

1003 H3K27ac HiChIP were processed with HiC-Pro (v2.10.1) using at least two replicates
1004 for each experiment with a bin resolution of 5kb and all analyses were performed using
1005 valid pairs (Sup. Table S4). Chains started from H3K4me3 peaks overlapping TSS (V19
1006 genome annotation of the hg19 mapping assembly) of expressed genes in A673^{WT} or
1007 A673^{SA2m#1} conditions (for A673 chains) or TC71^{WT} or TC71^{SA2m#2} conditions (for TC71
1008 chains). Each 5kb promoter bin (BIN-P) overlapped with at least one of these promoter
1009 regions (i.e. several BIN-P overlapping H3K4me3 peaks were allowed). Starting from a
1010 single BIN-P, the first enhancer (BIN-E1) element of the promoter enhancer chain was
1011 identified as bin displaying overlap with H3K27ac peaks (in respectively A673^{WT} and
1012 TC71^{WT} ChIP-seq data) and displaying the strongest interaction (greater than 4 reads)
1013 located at least 20,000 (4 bin gap) away from BIN-P in H3K27ac HiChIP matrix of
1014 respectively A673^{WT} or TC71^{WT} data. A recursive algorithm following these rules
1015 allowed to construct promoter enhancer chains up to the 20th enhancer BIN (chains:
1016 BIN-P linked to BIN-E1 up to BIN-E20). Promoter enhancer chains were assigned to
1017 respectively 10,716 and 10,708 genes in A673^{WT} and TC71^{WT}, representing respectively
1018 74% and 72% of the population of expressed genes in A673 and TC71. Promoter
1019 enhancer chains containing GGAA microsatellite-bound EWSR1-FLI1 peaks at any BIN-
1020 P or BIN-E position within their chains are summarized in Table S5. Interactions
1021 (separated by at least 20kbp and displaying at least 4 reads) localized within super
1022 enhancer (SE) regions were defined in parental *STAG2*-WT cells (634 and 725 SEs
1023 were respectively considered in A673^{WT} and TC71^{WT}). So defined interacting bins were
1024 compared to H3K27ac HiChIP data of isogenic *STAG2*-mutated or -rescue conditions.

1025

1026 **Soft agar colony formation assay**

1027 A first agar layer was placed in 6 well cell culture plates at 0.8% (w/v) of low melting
1028 point agarose (16520; Invitrogen) in respective cell culture media. Once solidified, a
1029 second layer of 0.4% agar was added, containing respectively 1000 or 2000 cells for

1030 TC71 and A673 lines. The plates were maintained at 4°C for 5 minutes and 1 mL of
1031 fresh medium was subsequently deposited as a top layer. The plates were incubated at
1032 37°C at 5% CO₂ and colonies (>50 µm diameter) were counted 2 weeks post seeding
1033 using the FIJI Particle Analysis tool.
1034

1035 **Actin cytoskeleton and focal adhesion**

1036 A673 (2.5X10⁴ cells) and A673SA2m#1 (2X10⁴ cells) were plated in 24-well plate on
1037 poly-D-Lysine coated coverslips (Corning, #354086). After 24 hours, the cells were fixed
1038 with 4% paraformaldehyde, permeabilized in 0.2% Triton X-100/1XPBS and blocked in
1039 1% BSA/1XPBS solution. Cells were then incubated with phalloidin-
1040 tetramethylrhodamine B isothiocyanate (1:100, Sigma, #P1951) and paxillin antibody
1041 (1:1000, BD biosciences, #610051) diluted in 0.5% BSA/1X PBS for 1 hour, washed
1042 with 0.5% BSA/1X PBS and incubated with Alexa Fluor 488-conjugated donkey anti-
1043 mouse (1:100, Invitrogen, #A21202) for 45 min. Slides were mounted using Prolong
1044 Gold Antifade reagent (Thermofisher, #P36930). Images were acquired with an upright
1045 widefield Apotome microscope (Zeiss) equipped with a Coolsnap HQ2 camera through
1046 a x63 NA 1.4 oil-immersion objective lens. Paxillin focal adhesion foci associated with
1047 phalloidin stress fibers were counted in individual cells for each condition.
1048

1049 **Wound healing migration assay**

1050 Migratory capacity of the cells was examined by seeding respectively 7X10⁴ A673,
1051 6X10⁴ A673SA2m#1, 7X10⁴ A673SA2r, 7-8X10⁴ EWIma1^{SA2} or EWIma1^{Empty} and
1052 9X10⁴ CHLA-258 9X10⁴ cells in each well of a cell culture-insert (Ibidi, #80209). At cell
1053 confluence, insert was carefully removed using tweezers and cell migration into the
1054 wound (cell-free 500 µm gap) was monitored using time-lapse microscopy (IncuCyte
1055 Live Cell Analysis Systems, 4x or 10x objective lens, Sartorius) with an interval of 2
1056 hours during 48 hours. The wound surface was measured for each time point using the
1057 MRI Wound Healing Tool ([http://dev.mri.cnrs.fr/projects/imagej-
1058 macos/wiki/Wound_Healing_Tool](http://dev.mri.cnrs.fr/projects/imagej-macos/wiki/Wound_Healing_Tool)) in ImageJ.
1059

1060 **Spheroid invasion assay**

1061 Spheroids were formed in 96-wells Ultra-Low Attachment microplates (Corning, 7007)
1062 by seeding 2000 cells in 50 μ l of cell culture media. After 3 days of aggregation,
1063 spheroids were embedded in 50 μ l of Collagen I matrix (5mg/ml) (Corning, 354249) and
1064 incubated in a cell culture incubator (37°C, 5% CO₂). Time-lapse microscopy (IncuCyte
1065 Live Cell Analysis Systems, 4X objective lens, Sartorius) was started 2h post
1066 embedding for 7 days using an interval of 1 hour between each acquisition.

1067

1068 **Image analysis**

1069 To analyze the shape and size of cellular aggregates, raw images are binarized through
1070 a custom-made MATLAB code. Briefly, the code detects the contours of the aggregate
1071 with a standard-deviation filter. Contour are then filled to define the shape of the
1072 aggregate, whose geometric properties are assessed through MATLAB's regionprops
1073 function. In particular, we record for each frame the area and the eccentricity, defined as
1074 the ratio of the distance between the two foci of the equivalent ellipse and the length of
1075 the long axis. This value is equal to 1 for a straight line and 0 for a perfect circle.

1076

1077 **Velocity measurements:**

1078 Particle image velocimetry (PIV) was used to obtain the velocity fields around the
1079 contour of the aggregates. The PIV analysis was performed using the Matpiv package
1080 in MATLAB (<https://www.mn.uio.no/math/english/people/aca/jks/matpiv/>). A three-pass
1081 computation using a final window of 64 x 64 pixels (195 x 195 μ m) was used, with a .75
1082 overlap. Aberrant vectors were detected and removed from the analysis when their
1083 magnitude exceeded the local median value by three times the standard deviation. The
1084 time interval between consecutive images was 4 hours. For clarity, we only draw the
1085 velocity arrows located at less than 100 μ m from the edge of the aggregate.

1086

1087 **Boyden chambers assay**

1088 A673 and CHLA-258 cells were transfected 48hrs prior the assay. The following
1089 numbers of trypsinized cells were seeded in the upper compartment of a Boyden
1090 chamber (353182, BD Biosciences): 5X10⁴ A673, 5X10⁴ A673SA2m#1, 5x10⁴
1091 A673SA2r cells and 7.5X10⁴ A673 or CHLA-258 transfected cells (siCT, siSA#6 or

1092 siSA#8) in respectively their respective media containing 0.5% FBS. The lower well
1093 chamber was in contact with respective media containing 10% FBS. After 24 hours
1094 (siRNA) or 48 hours (*STAG2* isogenic models), the remaining cells present on the upper
1095 membrane of the Boyden chamber were carefully wiped out with a cotton swab and the
1096 chamber was then stained with crystal violet solution. Cell count was performed from 10
1097 non-overlapping pictures acquired with a microscope (10X objective lens).

1098

1099 **GSEA and DoRothEA Analyses**

1100 GSEA was performed using default parameters and the MSigDB H, C2, C3, C4, C5, C6,
1101 C7 (V7.1.symbols) gene sets collections and published gene sets (listed in Table S2).
1102 The analysis presented in table S6 was generated with the web based Investigate Gene
1103 Sets tool (<https://www.gsea-msigdb.org/gsea/msigdb/annotate.jsp>) using H, C2, C3, C4,
1104 C5, C6, C7 gene sets and default parameters except for max gene (n=500) per gene
1105 set. Due to restriction of maximum input genes (n<1995) of this tool, only genes
1106 displaying expression levels >250 TPM were selected for this analysis (1918 genes
1107 among 2331 in A673 and 1332 genes among 1625 in TC71). DoRothEA analyses were
1108 performed using default parameters and restricting results to A and B confidence level
1109 transcription factors.

1110

1111 **QUANTIFICATION AND STATISTICAL ANALYSIS**

1112 The tests used for statistical analyses are described in the legends of each concerned
1113 figure and have been performed using R v3.4. Symbols for significance are described in
1114 the legends of each concerned figure. Experimental group, n represent the number of
1115 subjects within each group

1119 REFERENCES

- 1120 Arruda, N.L., Carico, Z.M., Justice, M., Liu, Y.F., Zhou, J., Stefan, H.C., and Downen,
1121 J.M. (2020). Distinct and overlapping roles of STAG1 and STAG2 in cohesin localization
1122 and gene expression in embryonic stem cells. *Epigenetics Chromatin* *13*, 32.
- 1123 Aynaud, M.-M., Mirabeau, O., Gruel, N., Grossetête, S., Boeva, V., Durand, S., Surdez,
1124 D., Saulnier, O., Zaïdi, S., Gribkova, S., et al. (2020). Transcriptional Programs Define
1125 Intratumoral Heterogeneity of Ewing Sarcoma at Single-Cell Resolution. *Cell Rep* *30*,
1126 1767-1779.e6.
- 1127 Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe,
1128 A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., et al. (2018). Comprehensive
1129 Characterization of Cancer Driver Genes and Mutations. *Cell* *173*, 371-385.e18.
- 1130 Baldauf, M.C., Orth, M.F., Dallmayer, M., Marchetto, A., Gerke, J.S., Rubio, R.A., Kiran,
1131 M.M., Musa, J., Knott, M.M.L., Ohmura, S., et al. (2018). Robust diagnosis of Ewing
1132 sarcoma by immunohistochemical detection of super-enhancer-driven EWSR1-ETS
1133 targets. *Oncotarget* *9*, 1587–1601.
- 1134 Batra, S., Reynolds, C.P., and Maurer, B.J. (2004). Fenretinide cytotoxicity for Ewing’s
1135 sarcoma and primitive neuroectodermal tumor cell lines is decreased by hypoxia and
1136 synergistically enhanced by ceramide modulators. *Cancer Res.* *64*, 5415–5424.
- 1137 Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C.A., Chotalia, M., Xie, S.Q.,
1138 Barbieri, M., de Santiago, I., Lavitas, L.-M., Branco, M.R., et al. (2017). Complex multi-
1139 enhancer contacts captured by genome architecture mapping. *Nature* *543*, 519–524.
- 1140 Benedetti, L., Cereda, M., Monteverde, L., Desai, N., and Ciccarelli, F.D. (2017).
1141 Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog
1142 STAG1. *Oncotarget* *8*, 37619–37632.
- 1143 Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S.,
1144 Yamaya, K., Boettiger, A.N., and Zhuang, X. (2018). Super-resolution chromatin tracing
1145 reveals domains and cooperative interactions in single cells. *Science* *362*.
- 1146 Boulay, G., Sandoval, G.J., Riggi, N., Iyer, S., Buisson, R., Naigles, B., Awad, M.E.,
1147 Rengarajan, S., Volorio, A., McBride, M.J., et al. (2017). Cancer-Specific Retargeting of
1148 BAF Complexes by a Prion-like Domain. *Cell* *171*, 163-178.e19.
- 1149 Boulay, G., Volorio, A., Iyer, S., Broye, L.C., Stamenkovic, I., Riggi, N., and Rivera, M.N.
1150 (2018). Epigenome editing of microsatellite repeats defines tumor-specific enhancer
1151 functions and dependencies. *Genes Dev.* *32*, 1008–1019.
- 1152 Brohl, A.S., Solomon, D.A., Chang, W., Wang, J., Song, Y., Sindiri, S., Patidar, R.,
1153 Hurd, L., Chen, L., Shern, J.F., et al. (2014). The genomic landscape of the Ewing
1154 sarcoma family of tumors reveals recurrent STAG2 mutation. *PLoS Genet* *10*,
1155 e1004475.

1156 Casa, V., Moronta Gines, M., Gade Gusmao, E., Slotman, J.A., Zirkel, A., Josipovic, N.,
1157 Oole, E., van IJcken, W.F.J., Houtsmuller, A.B., Papantonis, A., et al. (2020).
1158 Redundant and specific roles of cohesin STAG subunits in chromatin looping and
1159 transcriptional control. *Genome Res.* *30*, 515–527.

1160 Chaturvedi, A., Hoffman, L.M., Welm, A.L., Lessnick, S.L., and Beckerle, M.C. (2012).
1161 The EWS/FLI oncogene drives changes in cellular morphology, adhesion, and migration
1162 in Ewing sarcoma. *Genes Cancer* *3*, 102–116.

1163 Crompton, B.D., Stewart, C., Taylor-Weiner, A., Alexe, G., Kurek, K.C., Calicchio, M.L.,
1164 Kiezun, A., Carter, S.L., Shukla, S.A., Mehta, S.S., et al. (2014). The genomic
1165 landscape of pediatric Ewing sarcoma. *Cancer Discov* *4*, 1326–1341.

1166 Dauphinot, L., De Oliveira, C., Melot, T., Sevenet, N., Thomas, V., Weissman, B.E., and
1167 Delattre, O. (2001). Analysis of the expression of cell cycle regulators in Ewing cell
1168 lines: EWS-FLI-1 modulates p57KIP2 and c-Myc expression. *Oncogene* *20*, 3258–3265.

1169 Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.-M. (2019). DNA
1170 loop extrusion by human cohesin. *Science*.

1171 De Koninck, M., Lapi, E., Badía-Careaga, C., Cossío, I., Giménez-Llorente, D.,
1172 Rodríguez-Corsino, M., Andrada, E., Hidalgo, A., Manzanares, M., Real, F.X., et al.
1173 (2020). Essential Roles of Cohesin STAG2 in Mouse Embryonic Development and Adult
1174 Tissue Homeostasis. *Cell Rep* *32*, 108014.

1175 Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H.,
1176 Joubert, I., de Jong, P., Rouleau, G., et al. (1992). Gene fusion with an ETS DNA-
1177 binding domain caused by chromosome translocation in human tumours. *Nature* *359*,
1178 162–165.

1179 Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B.
1180 (2012). Topological domains in mammalian genomes identified by analysis of chromatin
1181 interactions. *Nature* *485*, 376–380.

1182 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,
1183 Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner.
1184 *Bioinformatics* *29*, 15–21.

1185 Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub,
1186 A.S., Schuijers, J., Lee, T.I., Zhao, K., et al. (2014). Control of Cell Identity Genes
1187 Occurs in Insulated Neighborhoods in Mammalian Chromosomes. *Cell* *159*, 374–387.

1188 Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S.,
1189 and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact
1190 Maps with Unlimited Zoom. *Cell Syst* *3*, 99–101.

1191 Franzetti, G.-A., Laud-Duval, K., van der Ent, W., Brisac, A., Irondelle, M., Aubert, S.,
1192 Dirksen, U., Bouvier, C., de Pinieux, G., Snaar-Jagalska, E., et al. (2017). Cell-to-cell

1193 heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in
1194 Ewing sarcoma cells. *Oncogene* 36, 3505–3514.

1195 Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A.
1196 (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep* 15, 2038–
1197 2049.

1198 Gangwal, K., Sankar, S., Hollenhorst, P.C., Kinsey, M., Haroldsen, S.C., Shah, A.A.,
1199 Boucher, K.M., Watkins, W.S., Jorde, L.B., Graves, B.J., et al. (2008). Microsatellites as
1200 EWS/FLI response elements in Ewing’s sarcoma. *Proceedings of the National Academy*
1201 *of Sciences of the United States of America* 105, 10149–10154.

1202 Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C.
1203 (2018). Real-time imaging of DNA loop extrusion by condensin. *Science* 360, 102–105.

1204 Garcia-Alonso, L., Holland, C.H., Ibrahim, M.M., Turei, D., and Saez-Rodriguez, J.
1205 (2019). Benchmark and integration of resources for the estimation of human
1206 transcription factor activities. *Genome Res* 29, 1363–1375.

1207 Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a
1208 given motif. *Bioinformatics* 27, 1017–1018.

1209 Grünewald, T.G.P., Cidre-Aranaz, F., Surdez, D., Tomazou, E.M., de Álava, E., Kovar,
1210 H., Sorensen, P.H., Delattre, O., and Dirksen, U. (2018). Ewing sarcoma. *Nat Rev Dis*
1211 *Primers* 4, 5.

1212 Guillon, N., Tirode, F., Boeva, V., Zynovyev, A., Barillot, E., and Delattre, O. (2009). The
1213 oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with
1214 potential transcriptional activation function. *PLoS ONE* 4, e4932.

1215 Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y.,
1216 Tang, Y., et al. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and
1217 Enhancer/Promoter Function. *Cell* 162, 900–910.

1218 Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yáñez-Cuna, J.O., Amendola, M.,
1219 van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., et al.
1220 (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell*
1221 169, 693-707.e14.

1222 Hassler, M., Shaltiel, I.A., and Haering, C.H. (2018). Towards a Unified Model of SMC
1223 Complex Function. *Curr. Biol.* 28, R1266–R1281.

1224 Hill, V.K., Kim, J.-S., and Waldman, T. (2016). Cohesin mutations in human cancer.
1225 *Biochim. Biophys. Acta* 1866, 1–11.

1226 Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A.,
1227 and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease.
1228 *Cell* 155, 934–947.

- 1229 Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.-L., Bak, R.O., Li, C.H., Goldmann, J.,
1230 Lajoie, B.R., Fan, Z.P., Sigova, A.A., et al. (2016). Activation of proto-oncogenes by
1231 disruption of chromosome neighborhoods. *Science* 351, 1454–1458.
- 1232 Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A
1233 Phase Separation Model for Transcriptional Control. *Cell* 169, 13–23.
- 1234 Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R.,
1235 and Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked
1236 Mammalian Chromatin Folding. *Mol Cell* 78, 539-553.e8.
- 1237 Katschnig, A.M., Kauer, M.O., Schwentner, R., Tomazou, E.M., Mutz, C.N., Linder, M.,
1238 Sibilica, M., Alonso, J., Aryee, D.N.T., and Kovar, H. (2017). EWS-FLI1 perturbs
1239 MRTFB/YAP-1/TEAD target gene regulation inhibiting cytoskeletal autoregulatory
1240 feedback in Ewing sarcoma. *Oncogene* 36, 5995–6005.
- 1241 Kauer, M., Ban, J., Kofler, R., Walker, B., Davis, S., Meltzer, P., and Kovar, H. (2009). A
1242 molecular function map of Ewing’s sarcoma. *PLoS ONE* 4, e5415.
- 1243 Kennedy, A.L., Vallurupalli, M., Chen, L., Crompton, B., Cowley, G., Vazquez, F., Weir,
1244 B.A., Tsherniak, A., Parasuraman, S., Kim, S., et al. (2015). Functional, chemical
1245 genomic, and super-enhancer screening identify sensitivity to cyclin D1/CDK4 pathway
1246 inhibition in Ewing sarcoma. *Oncotarget* 6, 30178–30193.
- 1247 Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., and Yu, H. (2019). Human cohesin
1248 compacts DNA by loop extrusion. *Science* 366, 1345–1349.
- 1249 Kojic, A., Cuadrado, A., De Koninck, M., Giménez-Llorente, D., Rodríguez-Corsino, M.,
1250 Gómez-López, G., Le Dily, F., Marti-Renom, M.A., and Losada, A. (2018). Distinct roles
1251 of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nat. Struct. Mol.*
1252 *Biol.* 25, 496–504.
- 1253 Kovar, H., Jug, G., Aryee, D.N., Zoubek, A., Ambros, P., Gruber, B., Windhager, R., and
1254 Gadner, H. (1997). Among genes involved in the RB dependent cell cycle regulatory
1255 cascade, the p16 tumor suppressor gene is frequently lost in the Ewing family of
1256 tumors. *Oncogene* 15, 2225-32.
- 1257 Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.-H.S.,
1258 Parsi, K.M., Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural Details of
1259 Mammalian Chromosome Architecture. *Mol Cell* 78, 554-565.e7.
- 1260 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2.
1261 *Nat. Methods* 9, 357–359.
- 1262 Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub,
1263 T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and
1264 saturation analysis of cancer genes across 21 tumour types. *Nature* 505, 495–501.

1265 van der Lelij, P., Lieb, S., Jude, J., Wutz, G., Santos, C.P., Falkenberg, K., Schlattl, A.,
1266 Ban, J., Schwentner, R., Hoffmann, T., et al. (2017). Synthetic lethality between the
1267 cohesin subunits STAG1 and STAG2 in diverse cancer contexts. *Elife* 6.

1268 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.,
1269 Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009).
1270 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.

1271 Li, Y., Haarhuis, J.H.I., Sedeño Cacciatore, Á., Oldenkamp, R., van Ruiten, M.S.,
1272 Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D., et al. (2020). The
1273 structural basis for cohesin-CTCF-anchored loops. *Nature* 578, 472–476.

1274 Liu, Y., Xu, H., Van der Jeught, K., Li, Y., Liu, S., Zhang, L., Fang, Y., Zhang, X.,
1275 Radovich, M., Schneider, B.P., et al. (2018). Somatic mutation of the cohesin complex
1276 subunit confers therapeutic vulnerabilities in cancer. *J. Clin. Invest.* 128, 2951–2965.

1277 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
1278 dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.

1279 Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D.,
1280 Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin
1281 domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012–
1282 1025.

1283 Merkschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and
1284 Transcriptional Gene Regulation. *Annu Rev Genomics Hum Genet* 17, 17–43.

1285 Michaelis, C., Ciosk, R., and Nasmyth, K. (1997). Cohesins: chromosomal proteins that
1286 prevent premature separation of sister chromatids. *Cell* 91, 35–45.

1287 Miyagawa, Y., Okita, H., Itagaki, M., Toyoda, M., Katagiri, Y.U., Fujimoto, J., Hata, J.,
1288 Umezawa, A., and Kiyokawa, N. (2009). EWS/ETS regulates the expression of the
1289 Dickkopf family in Ewing family tumor cells. *PLoS One* 4, e4634.

1290 Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J.,
1291 Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1alpha-
1292 responsive genes involved in oxidative phosphorylation are coordinately downregulated
1293 in human diabetes. *Nat. Genet.* 34, 267–273.

1294 Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and
1295 Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome
1296 architecture. *Nat. Methods* 13, 919–922.

1297 Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and
1298 Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the
1299 Hox clusters during differentiation. *Science* 347, 1017–1021.

1300 Nasmyth, K. (2001). Disseminating the genome: joining, resolving, and separating sister
1301 chromatids during mitosis and meiosis. *Annu. Rev. Genet.* *35*, 673–745.

1302 Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T.,
1303 van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the
1304 regulatory landscape of the X-inactivation centre. *Nature* *485*, 381–385.

1305 Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N.,
1306 Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF
1307 Decouples Local Insulation of Chromosome Domains from Genomic
1308 Compartmentalization. *Cell* *169*, 930-944.e22.

1309 Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N., and Mirny, L.A. (2018).
1310 Chromatin organization by an interplay of loop extrusion and compartmental
1311 segregation. *Proc Natl Acad Sci U S A* *115*, E6697–E6706.

1312 Ochi, Y., Kon, A., Sakata, T., Nakagawa, M.M., Nakazawa, N., Kakuta, M., Kataoka, K.,
1313 Koseki, H., Nakayama, M., Morishita, D., et al. (2020). Combined Cohesin-RUNX1
1314 Deficiency Synergistically Perturbs Chromatin Looping and Causes Myelodysplastic
1315 Syndromes. *Cancer Discov* *10*, 836–853.

1316 Pedersen, E.A., Menon, R., Bailey, K.M., Thomas, D.G., Van Noord, R.A., Tran, J.,
1317 Wang, H., Qu, P.P., Hoering, A., Fearon, E.R., et al. (2016). Activation of Wnt/ β -Catenin
1318 in Ewing Sarcoma Cells Antagonizes EWS/ETS Function and Promotes Phenotypic
1319 Transition to More Metastatic Cell States. *Cancer Res.* *76*, 5040–5053.

1320 Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell,
1321 J.S.K., Ong, C.-T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein
1322 subclasses shape 3D organization of genomes during lineage commitment. *Cell* *153*,
1323 1281–1295.

1324 Postel-Vinay, S., Véron, A.S., Tirode, F., Pierron, G., Reynaud, S., Kovar, H., Oberlin,
1325 O., Lapouble, E., Ballet, S., Lucchesi, C., et al. (2012). Common variants near TARDBP
1326 and EGR2 are associated with susceptibility to Ewing sarcoma. *Nature Genetics* *44*,
1327 323–327.

1328 Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson,
1329 J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the
1330 human genome at kilobase resolution reveals principles of chromatin looping. *Cell* *159*,
1331 1665–1680.

1332 Rao, S.S.P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-
1333 Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al.
1334 (2017). Cohesin Loss Eliminates All Loop Domains. *Cell* *171*, 305-320.e24.

1335 Riggi, N., Suva, M.L., Suva, D., Cironi, L., Provero, P., Tercier, S., Joseph, J.M., Stehle,
1336 J.C., Baumer, K., Kindler, V., et al. (2008). EWS-FLI-1 expression triggers a Ewing's

1337 sarcoma initiation program in primary human mesenchymal stem cells. *Cancer Res* *68*,
1338 2176–2185.

1339 Riggi, N., Knoechel, B., Gillespie, S.M., Rheinbay, E., Boulay, G., Suvà, M.L., Rossetti,
1340 N.E., Boonseng, W.E., Oksuz, O., Cook, E.B., et al. (2014). EWS-FLI1 utilizes divergent
1341 chromatin remodeling mechanisms to directly activate or repress enhancer elements in
1342 Ewing sarcoma. *Cancer Cell* *26*, 668–681.

1343 Rodríguez-Núñez, P., Romero-Pérez, L., Amaral, A.T., Puerto-Camacho, P., Jordán, C.,
1344 Marcilla, D., Grünewald, T.G., Alonso, J., de Alava, E., and Díaz-Martín, J. (2020).
1345 Hippo pathway effectors YAP1/TAZ induce an EWS-FLI1-opposing gene signature and
1346 associate with disease progression in Ewing sarcoma. *J. Pathol.* *250*, 374–386.

1347 Romero-Pérez, L., Surdez, D., Brunet, E., Delattre, O., and Grünewald, T.G.P. (2019).
1348 STAG Mutations in Cancer. *Trends in Cancer*.

1349 Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome
1350 architecture. *Nat. Rev. Genet.* *19*, 789–800.

1351 Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I.,
1352 Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion
1353 explains key features of loop and domain formation in wild-type and engineered
1354 genomes. *Proc. Natl. Acad. Sci. U.S.A.* *112*, E6456-6465.

1355 Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie,
1356 Y., Fonseca, N.A., Huber, W., Haering, C., Mirny, L., et al. (2017). Two independent
1357 modes of chromatin organization revealed by cohesin removal. *Nature* *551*, 51–56.

1358 Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E.,
1359 Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C
1360 data processing. *Genome Biol.* *16*, 259.

1361 Sheffield, N.C., Pierron, G., Klughammer, J., Datlinger, P., Schönegger, A., Schuster,
1362 M., Hadler, J., Surdez, D., Guillemot, D., Lapouble, E., et al. (2017). DNA methylation
1363 heterogeneity defines a disease spectrum in Ewing sarcoma. *Nat. Med.* *23*, 386–395.

1364 Smith, J.S., Lappin, K.M., Craig, S.G., Liberante, F.G., Crean, C.M., McDade, S.S.,
1365 Thompson, A., Mills, K.I., and Savage, K.I. (2020). Chronic loss of STAG2 leads to
1366 altered chromatin structure contributing to de-regulated transcription in AML. *J Transl*
1367 *Med* *18*, 339.

1368 Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Elkahloun, A.G., Harris, B.T.,
1369 Toretsky, J.A., Rosenberg, S.A., Shukla, N., Ladanyi, M., et al. (2011). Mutational
1370 Inactivation of STAG2 Causes Aneuploidy in Human Cancer. *Science* *333*, 1039–1043.

1371 Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N., and
1372 de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone
1373 modification in the beta-globin locus. *Genes Dev.* *20*, 2349–2354.

- 1374 Stadhouders, R., Fillion, G.J., and Graf, T. (2019). Transcription factors and 3D genome
1375 conformation in cell-fate decisions. *Nature* 569, 345–354.
- 1376 Surdez, D., Benetkiewicz, M., Perrin, V., Han, Z.-Y., Pierron, G., Ballet, S., Lamoureux,
1377 F., Rédini, F., Decouvelaere, A.-V., Daudigeos-Dubus, E., et al. (2012). Targeting the
1378 EWSR1-FLI1 Oncogene-Induced Protein Kinase PKC- β Abolishes Ewing Sarcoma
1379 Growth. *Cancer Res* 72, 4494–4503.
- 1380 Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska,
1381 A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-Mediated Human 3D Genome
1382 Architecture Reveals Chromatin Topology for Transcription. *Cell* 163, 1611–1627.
- 1383 Tirode, F., Laud-Duval, K., Prieur, A., Delorme, B., Charbord, P., and Delattre, O.
1384 (2007). Mesenchymal stem cell features of Ewing tumors. *Cancer Cell* 11, 421–429.
- 1385 Tirode, F., Surdez, D., Ma, X., Parker, M., Le Deley, M.C., Bahrami, A., Zhang, Z.,
1386 Lapouble, E., Grossetête-Lalami, S., Rusch, M., et al. (2014). Genomic landscape of
1387 Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53
1388 mutations. *Cancer Discov* 4, 1342–1353.
- 1389 Tomazou, E.M., Sheffield, N.C., Schmidl, C., Schuster, M., Schönegger, A., Datlinger,
1390 P., Kubicek, S., Bock, C., and Kovar, H. (2015). Epigenome mapping reveals distinct
1391 modes of gene regulation and widespread enhancer reprogramming by the oncogenic
1392 fusion protein EWS-FLI1. *Cell Rep* 10, 1082–1095.
- 1393 Vian, L., Pękowska, A., Rao, S.S.P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L.,
1394 Huang, S.-C., El Khattabi, L., Dose, M., Pruett, N., et al. (2018). The Energetics and
1395 Physiological Impact of Cohesin Extrusion. *Cell* 175, 292–294.
- 1396 Viny, A.D., Bowman, R.L., Liu, Y., Lavallée, V.-P., Eisman, S.E., Xiao, W., Durham,
1397 B.H., Navitski, A., Park, J., Braunstein, S., et al. (2019). Cohesin Members Stag1 and
1398 Stag2 Display Distinct Roles in Chromatin Accessibility and Topological Control of HSC
1399 Self-Renewal and Differentiation. *Cell Stem Cell* 25, 682-696.e8.
- 1400 Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S.,
1401 Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 Is a Structural
1402 Regulator of Enhancer-Promoter Loops. *Cell* 171, 1573-1588.e28.
- 1403 Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S.,
1404 Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional
1405 insulation by CCCTC-binding factor. *Nature* 451, 796–801.
- 1406 Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl,
1407 P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator
1408 establish super-enhancers at key cell identity genes. *Cell* 153, 307–319.
- 1409 Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W.,
1410 Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., et al. (2017). Topologically

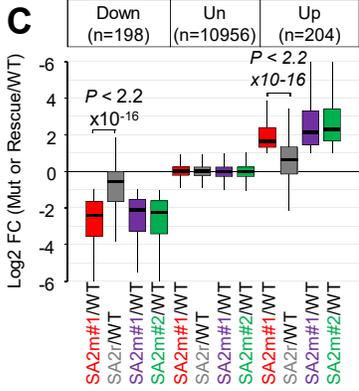
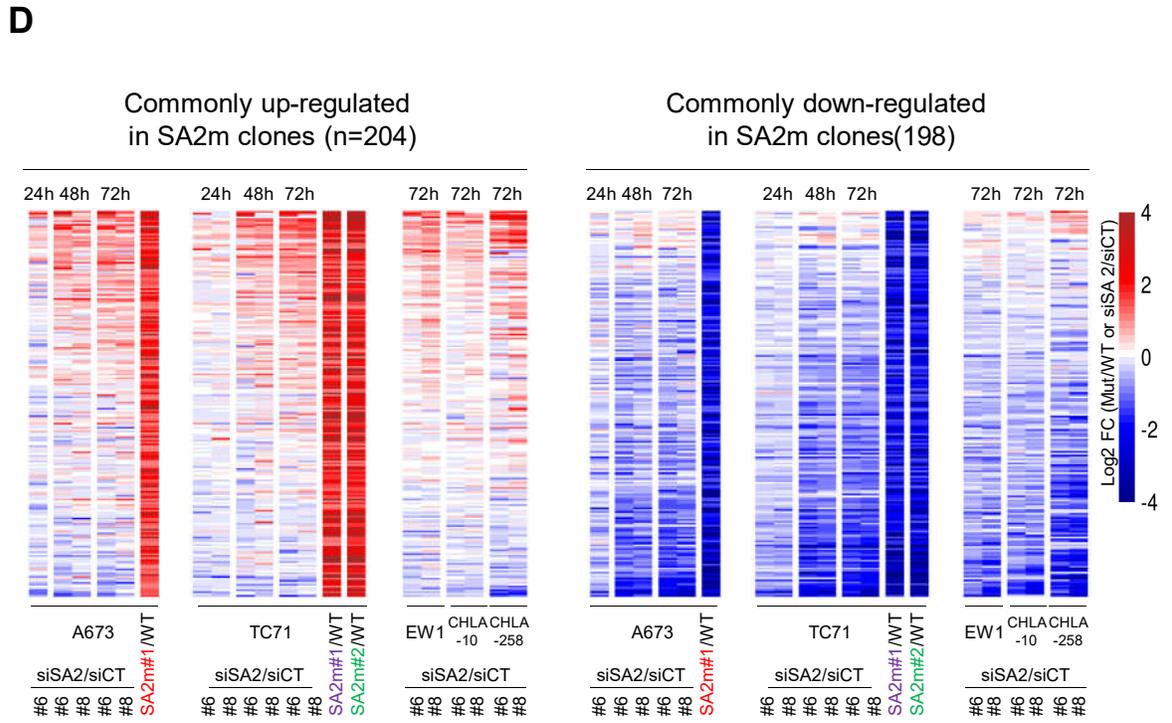
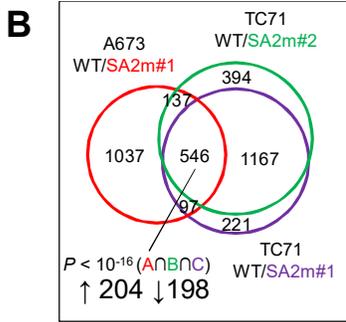
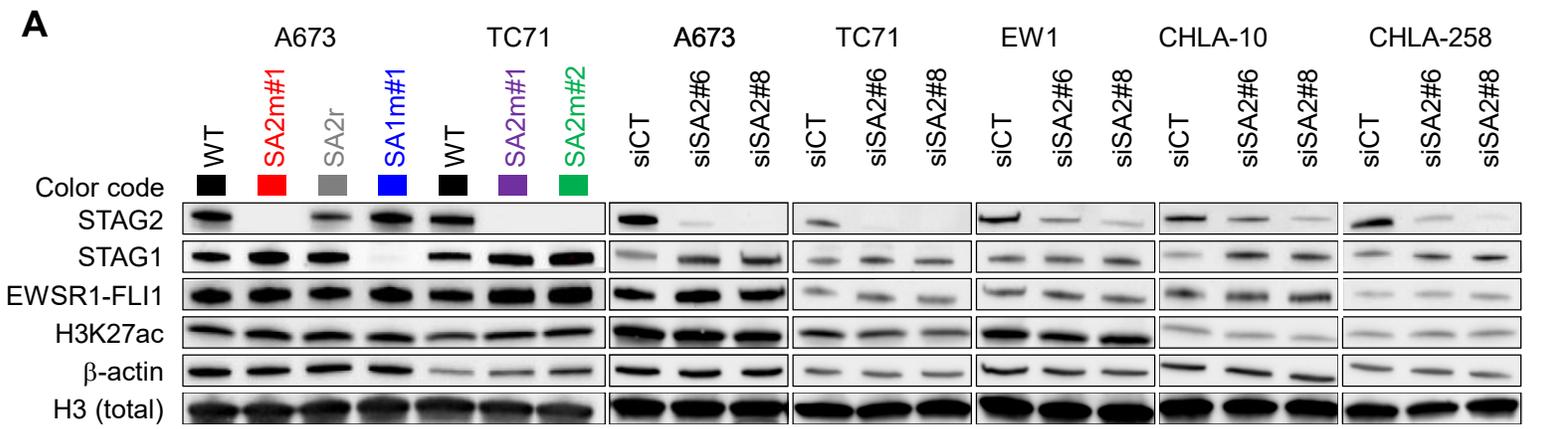
1411 associating domains and chromatin loops depend on cohesin and are regulated by
1412 CTCF, WAPL, and PDS5 proteins. *EMBO J.* 36, 3573–3599.

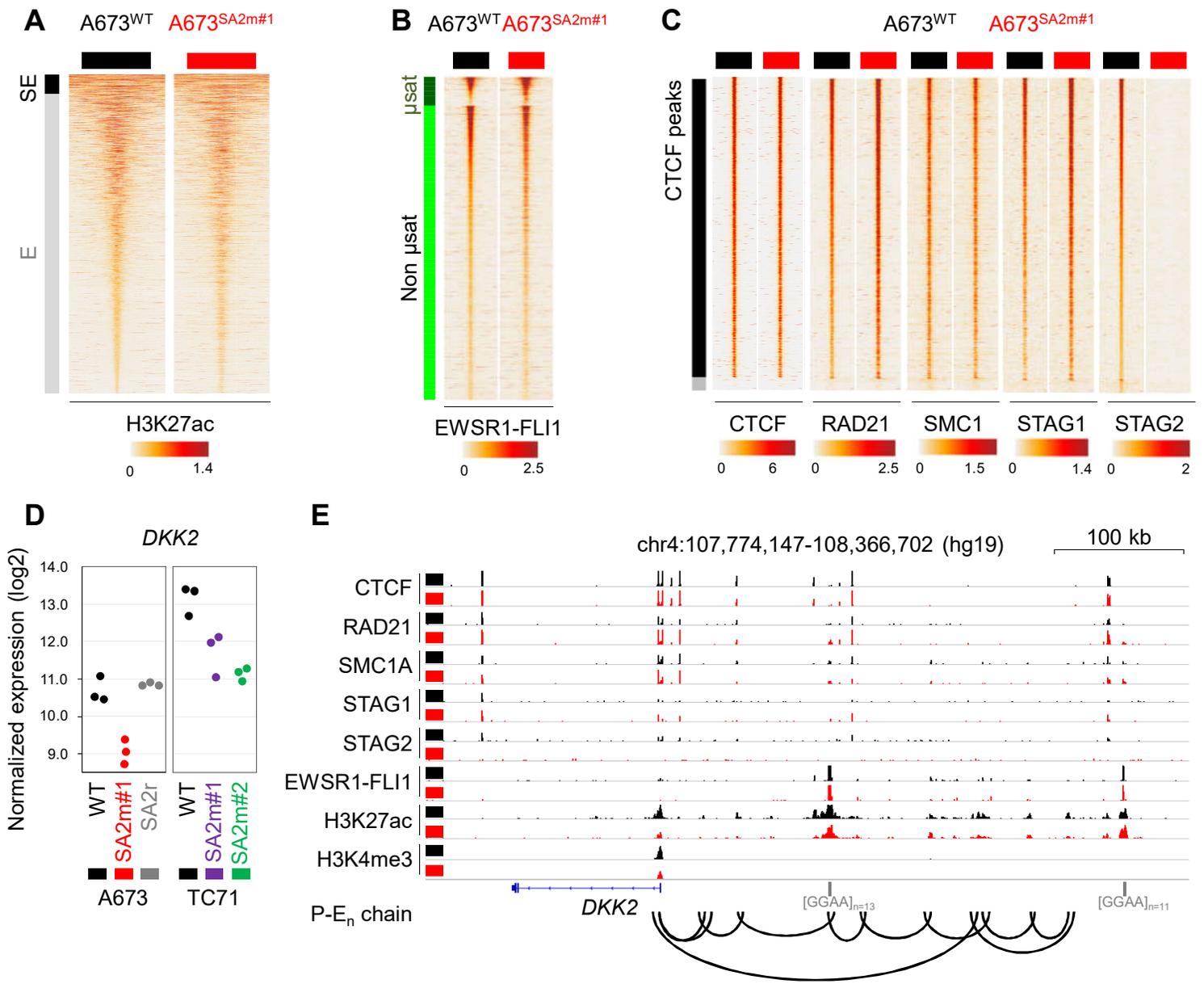
1413 Xiao, T., Wallace, J., and Felsenfeld, G. (2011). Specific Sites in the C Terminus of
1414 CTCF Interact with the SA2 Subunit of the Cohesin Complex and Are Required for
1415 Cohesin-Dependent Insulation Activity [\[2\]](#). *Mol Cell Biol* 31, 2174–2183.

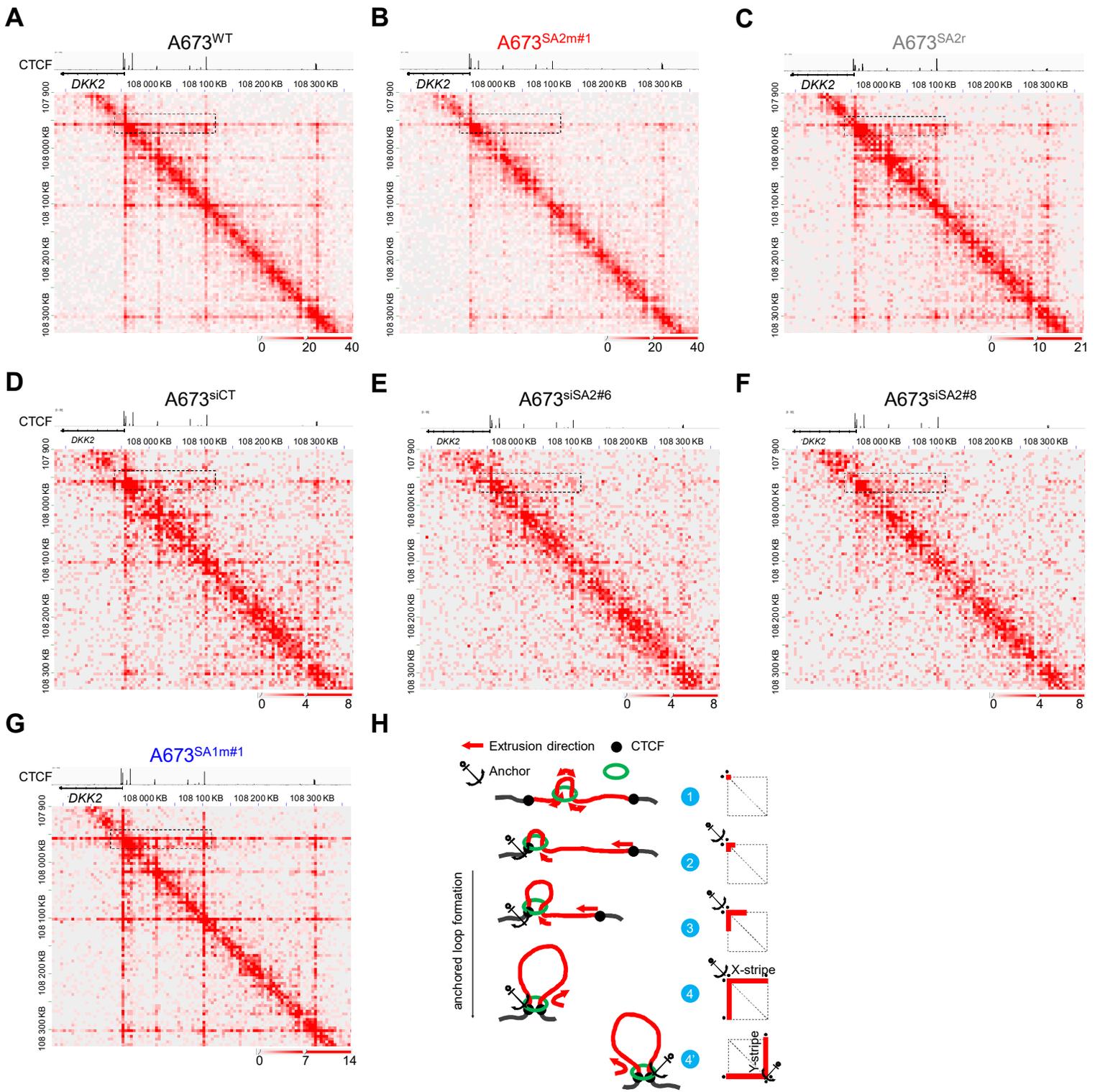
1416 Zhang, N., Jiang, Y., Mao, Q., Demeler, B., Tao, Y.J., and Pati, D. (2013).
1417 Characterization of the interaction between the cohesin subunits Rad21 and SA1/2.
1418 *PLoS ONE* 8, e69458.

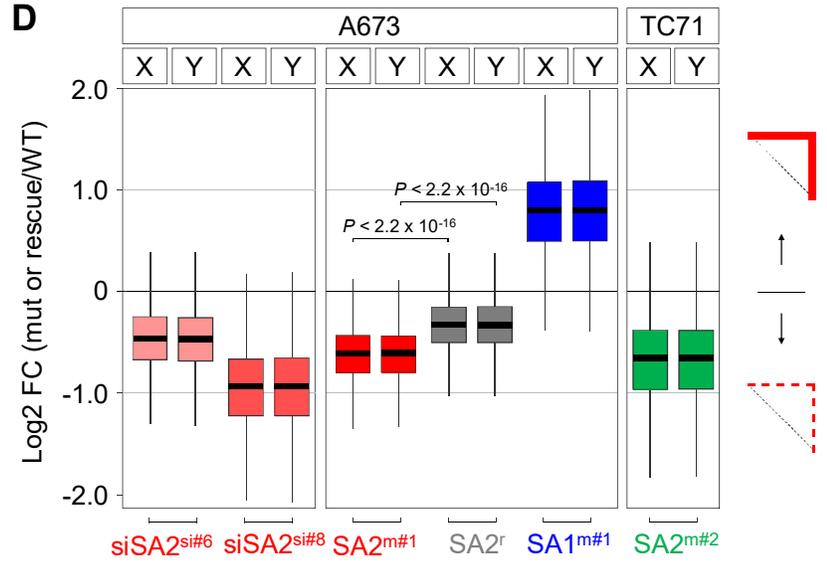
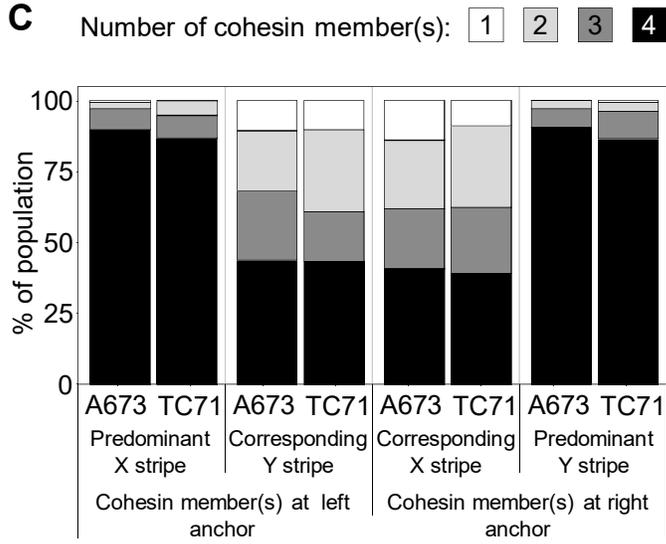
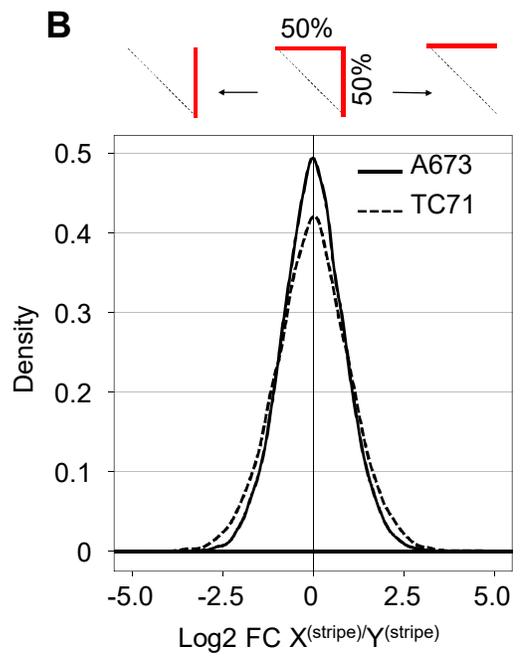
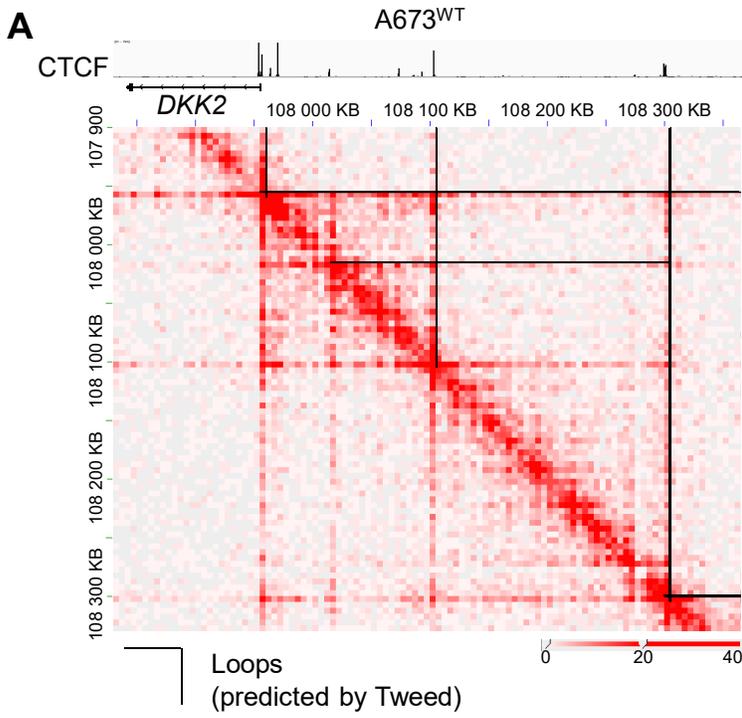
1419 Zhang, X., Zhang, Y., Ba, Z., Kyritsis, N., Casellas, R., and Alt, F.W. (2019).
1420 Fundamental roles of chromatin loop extrusion in antibody class switching. *Nature* 575,
1421 385–389.

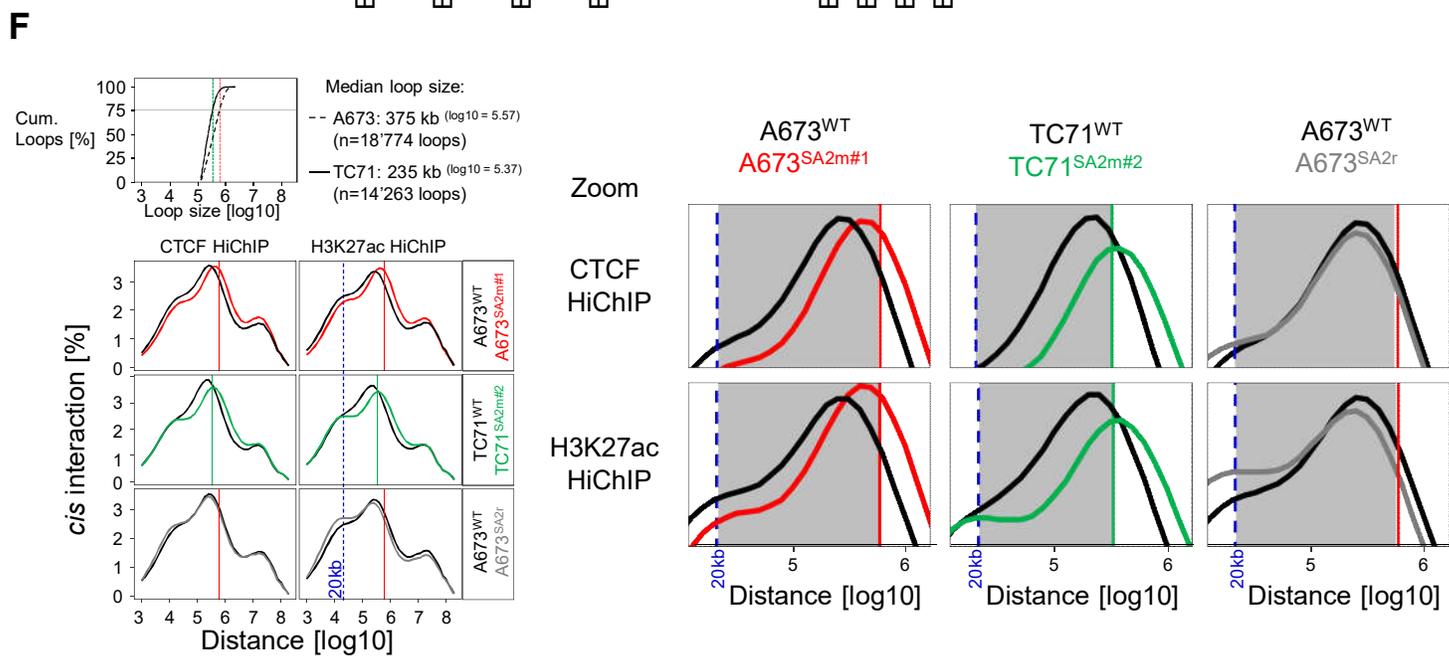
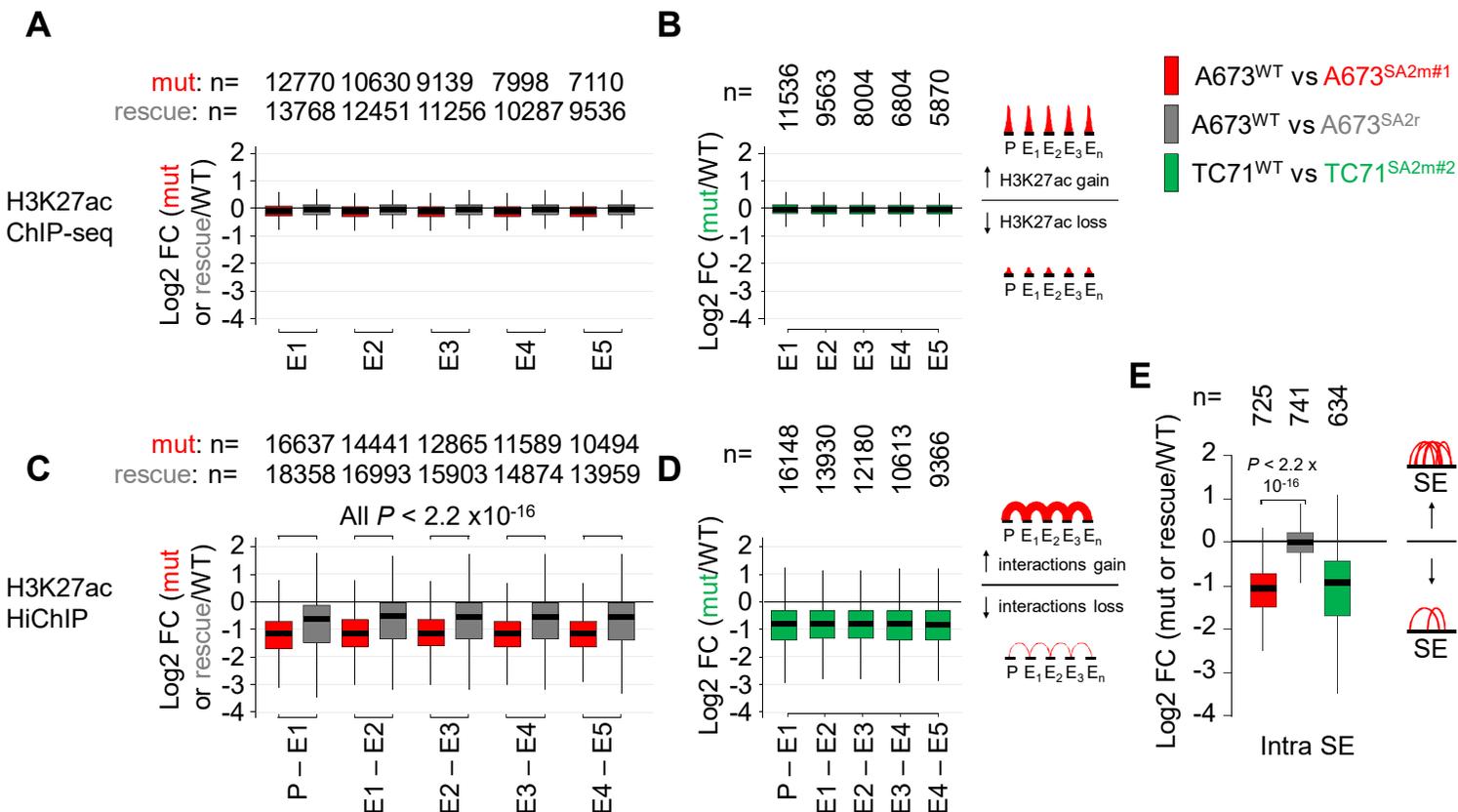
1422 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E.,
1423 Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of
1424 ChIP-Seq (MACS). *Genome Biol.* 9, R137.

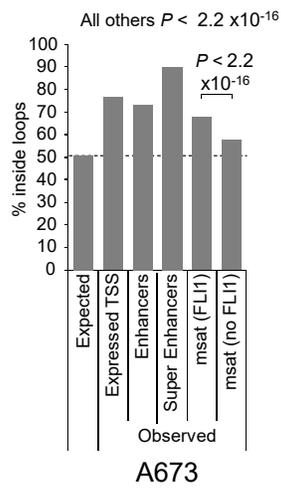
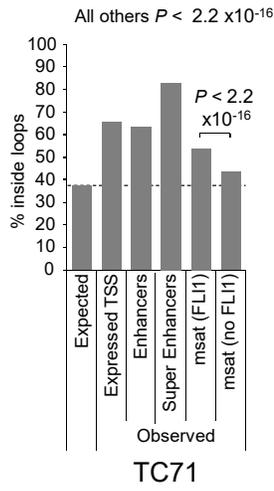
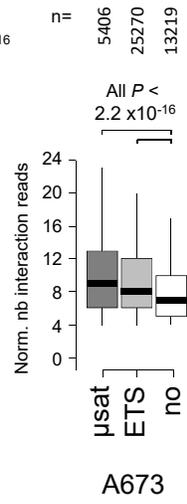
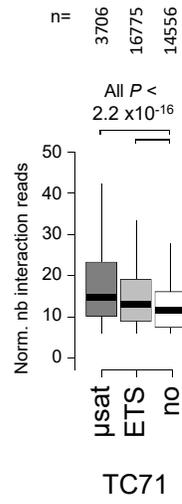
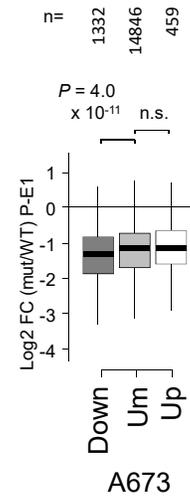
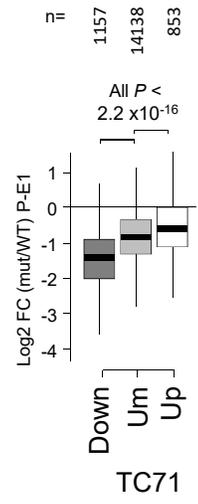


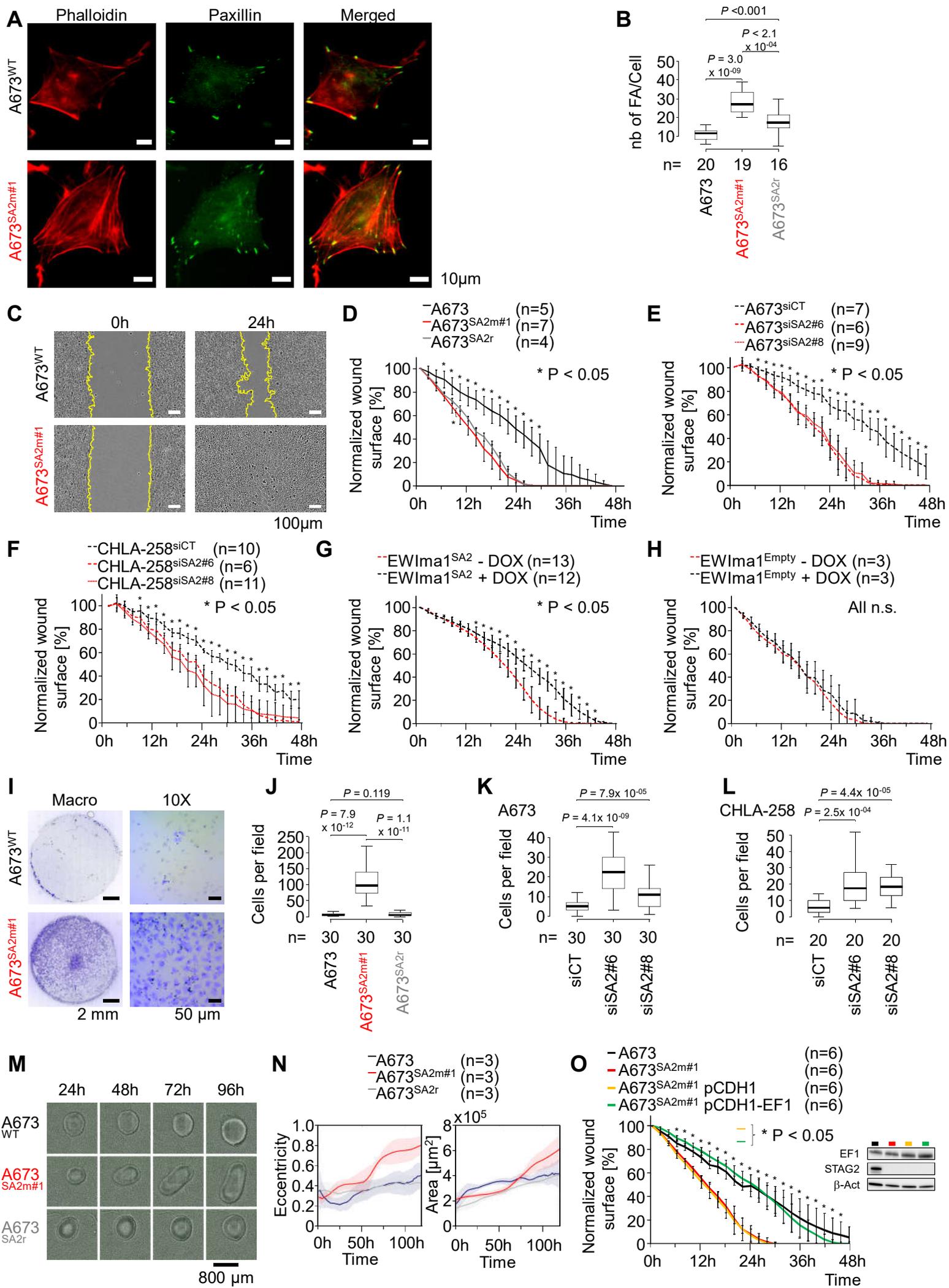








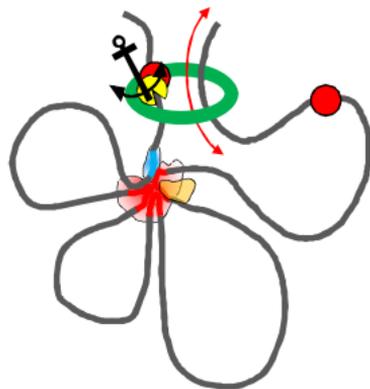
A**B****C****D****E****F**



STAG2^{WT}

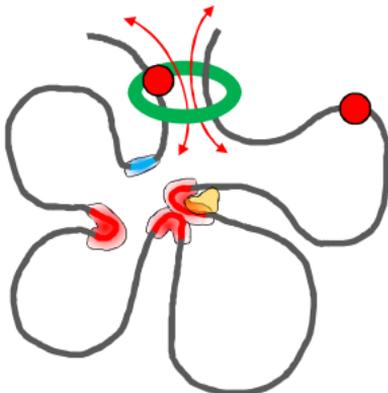
CTCF/cohesin-STAG2 favors:

- anchored chromatin extrusion
- promoter-enhancers and intra-enhancer interactions



STAG2 Mutated(KO)

- ↓ anchored chromatin extrusion
- ↓ cis-enhancer activity



in Ewing sarcoma

