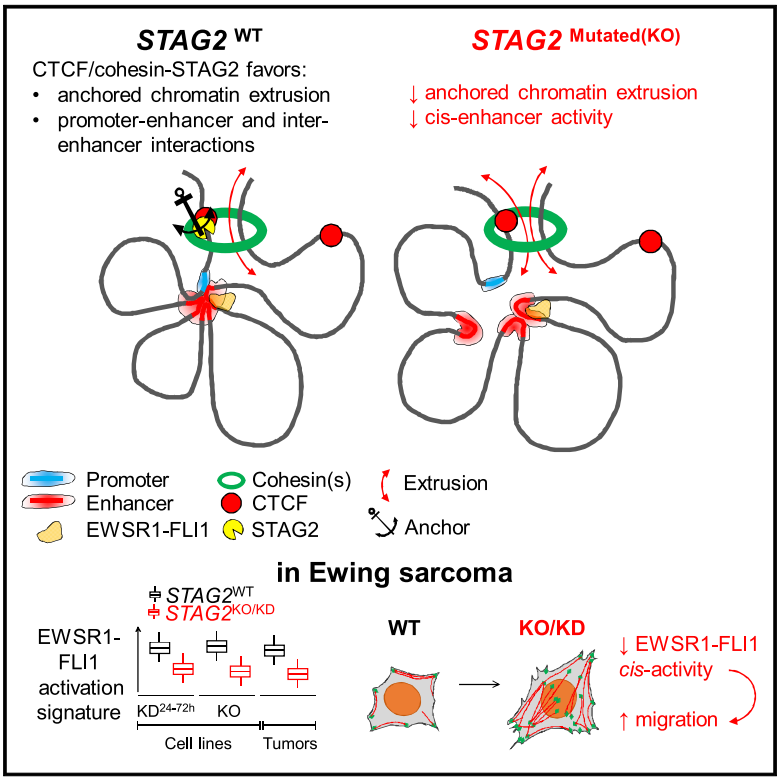
Cancer Cell

Article

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

Graphical abstract



Highlights

- STAG2 promotes anchored chromatin extrusion at CTCF loop boundaries
- *STAG2* LOF reduces *cis*-mediated EWSR1-FLI1 transcriptional activity
- *STAG2* LOF increases mesenchymal features and migratory potential of Ewing cells
- *STAG2* LOF signature is associated with adverse prognosis in Ewing sarcoma

Authors

Didier Surdez, Sakina Zaidi, Sandrine Grossetête, ..., Erika Brunet, Véronique Hill, Olivier Delattre

Correspondence

didier.surdez@curie.fr (D.S.), olivier.delattre@curie.fr (O.D.)

In brief

Surdez et al. find that STAG2, an integral member of the cohesin complex, promotes anchored chromatin loop extrusion and favors promoter-enhancer interactions. In Ewing sarcoma, *STAG2* mutations do not affect EWSR1-FLI1 expression level but reduce its *cis*-mediated enhancer activity, increasing mesenchymal features of Ewing sarcoma cells.

Surdez et al., 2021, Cancer Cell *39*, 1–17
June 14, 2021
https://doi.org/10.1016/j.ccell.2021.04.001

CellPress

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

Cancer Cell

CellPress

Article

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

Didier Surdez,[1,7,*] Sakina Zaidi,[1,9] Sandrine Grossetête,[1,9] Karine Laud-Duval,[1] Anna Sole Ferre,[2] Lieke Mous,[1,7] Thomas Vourc'h,[3,8] Franck Tirode,[4] Gaelle Pierron,[5] Virginie Raynal,[1,6] Sylvain Baulande,[6] Erika Brunet,[2] Véronique Hill,[1] and Olivier Delattre[1,5,10,*]

[1]INSERM U830, Équipe Labellisée LNCC, Diversity and Plasticity of Childhood Tumors Lab, PSL Research University, SIREDO Oncology Centre, Institut Curie Research Centre, 75005 Paris, France
[2]INSERM UMR 1163, Laboratory of Genome Dynamics in the Immune System, Equipe Labellisée Ligue contre le Cancer and Université de Paris, Imagine Institute, 75005 Paris, France
[3]UMR 168, Biology Inspired Physics at Mesoscales, PSL Research University, Institut Curie Research Centre, 75005 Paris, France
[4]Univ Lyon, Université Claude Bernard Lyon 1, CNRS 5286, INSERM U1052, Cancer Research Center of Lyon, 69008 Lyon, France
[5]Unité de Génétique Somatique, Service d'oncogénétique, Institut Curie, Centre Hospitalier, 75005 Paris, France
[6]Institut Curie Genomics of Excellence (ICGex) Platform, PSL Université, Institut Curie Research Centre, 75005 Paris, France
[7]Present address: Balgrist University Hospital, University of Zurich, Zurich, Switzerland
[8]Present address: Université Clermont Auvergne, CNRS, Sigma Clermont, UMR 6602 Institut Pascal, 63000 Clermont-Ferrand, France
[9]These authors contributed equally
[10]Lead contact
*Correspondence: didier.surdez@curie.fr (D.S.), olivier.delattre@curie.fr (O.D.)
https://doi.org/10.1016/j.ccell.2021.04.001

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 at GGAA microsatellite neo-enhancers. 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.

INTRODUCTION

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

STAG2 is an integral member of the cohesin complex, which is essential to hold sister chromatids together during mitosis and to shape the three-dimensional (3D) genome structure through its association with CTCF at the boundaries of topologically associating domains (Bintu et al., 2018; Dixon et al., 2012; Dowen et al., 2014; Michaelis et al., 1997; Nora et al., 2012; Rao et al., 2014; Wendt et al., 2008). Two CTCF molecules binding at convergent sites and interacting with the cohesin complex allow for the generation of a chromatin loop in which gene regulation processes preferentially occur (Beagrie et al., 2017; Guo et al., 2015; Rao et al., 2014, 2017; Tang et al., 2015). Key components of the CTCF/cohesin complex are necessary for the maintenance of chromatin loop structures (Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017; Wutz et al.,

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

2017) and are typically identified through a dot at corner peaks on Hi-C contact maps. These loops are likely generated in a dynamic process called chromatin extrusion (Davidson et al., 2019; Fudenberg et al., 2016; Hassler et al., 2018; Hsieh et al., 2020; Kim et al., 2019; Krietenstein et al., 2020; Nasmyth, 2001; Nuebler et al., 2018; Sanborn et al., 2015; Vian et al., 2018). Chromatin extrusion by the condensin complex has been recently visualized in yeast (Ganji et al., 2018). ATP and NIPBL-MAU2 are essential factors for the extrusion of chromatin loops by the cohesin complex *in vitro* (Davidson et al., 2019; Kim et al., 2019). *In vivo*, it is thought that the dynamic extrusion process is reflected by architectural stripes detected in Hi-C and Micro-C experiments (Hsieh et al., 2020; Krietenstein et al., 2020; Vian et al., 2018). Current understanding of the exact mechanisms describing the interplay between the CTCF/cohesin complex and the chromatin during loop extrusion is, however, incomplete. Similarly, the role of *STAG2*, which is frequently altered in human cancer mainly through loss-of-function (LOF) mutations, remains poorly understood (Bailey et al., 2018; Hill et al., 2016; Lawrence et al., 2014; Romero-Pérez et al., 2019). Since *STAG2* is located on the X chromosome, inactivating mutations of a single allele are sufficient for complete LOF (Romero-Pérez et al., 2019). Inactivating mutations of the *STAG1* paralog are much less frequent in cancer. *STAG1* and *STAG2* LOF mutations have been shown to be synthetic lethal (Benedetti et al., 2017; van der Lelij et al., 2017).

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

RESULTS

STAG2 knockout profoundly alters the transcriptomic landscape

To decipher the oncogenic mechanisms related to *STAG2* LOF in Ewing sarcoma, we used a CRISPR-Cas9 approach with two different sgRNAs targeting *STAG2* (SA2m#1 and SA2m#2). We generated three knockout (KO) isogenic pairs derived from A673 ($A673^{SA2m\#1}$) and TC71 ($TC71^{SA2m\#1}$, $TC71^{SA2m\#2}$), two *STAG2* wild-type (WT) Ewing sarcoma cell lines (Figure 1A). Absence of STAG2 protein expression was confirmed in each of these clones (Figure 1A). Proliferation rate of *STAG2* WT and KO cells was similar (Figures S1A–S1D). RNA sequencing (RNA-seq) comparing paired *STAG2* proficient and deficient lines highlighted a broad transcriptional modulation (Figure 1B; Table S1). Altogether, these three isogenic clones define a set of 546 *STAG2*-modulated genes, 204 being commonly upregulated and 198 being commonly downregulated genes in *STAG2* KO cells. To validate the specificity of our findings, we used a CRISPR-Cas9-based approach to correct the *STAG2* mutation and generated a line ($A673^{SA2r}$) with rescued *STAG2* expression (Figure 1A). Expression profiling in this rescue line showed highly significant reversion of both the *STAG2* KO-associated down- and upregulated effects (Figure 1C; Table S1). We also performed short-term knockdown (KD) experiments with two different small interfering RNAs (siRNAs) (siSA2#6 and siSA2#8) at three different time points (24, 48, and 72 h) in A673 and TC71. To confirm that the *STAG2*-regulated signature was not limited to A673 and TC71 isogenic clones, we also knocked down *STAG2* in three additional *STAG2* WT Ewing cell lines (EW1, CHLA-10, and CHLA-258) (Figure 1A) and further validated RNA-seq results for some genes using qRT-PCR (Figures S1E–S1J). As shown in Figures 1D and S2A, all these experiments indicated that the sets of STAG2-modulated genes defined with isogenic clones were regulated as soon as 24–48 h after siRNA transfection. At 72 h, most of the gene expression changes detected in the isogenic clones were observed, with only minor variations between siRNAs or cell lines and with a lower dynamic range of modulation than in stable KO experiments (Figure S2A). We can therefore conclude that this set of genes represents a robust signature of *STAG2* KO Ewing cells, modulated at short term and which hence accounts for early, possibly direct, transcriptomic consequences of *STAG2* inactivation.

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

The functional aspect of *STAG2*-proficient and -deficient Ewing sarcoma was investigated by Gene Set Enrichment Analysis (GSEA) and the DoRothEA-curated transcription factor/target gene set database (Garcia-Alonso et al., 2019; Mootha et al., 2003). Taking advantage of additional Ewing sarcoma cell line and tumor RNA-seq data, we performed this analysis in four independent Ewing sarcoma datasets: (1) *STAG2* WT parental cells versus *STAG2* KO isogenic cells, (2) *STAG2* WT si-control versus *STAG2* KD transfected cells, (3) *STAG2* WT versus *STAG2* mutated Ewing cell lines, and (4) *STAG2* WT versus *STAG2* mutated Ewing tumors. Strikingly, when investigating 18,889 signatures ranked by average normalized enrichment score (NES), several of the top 20 signatures enriched in *STAG2* proficient condition were EWSR1-FLI1-regulated gene signatures (Tables 1 and S2). Ranking first in GSEA and DoRothEA analyses (Tables 1 and S3), IC-EWS is a recently described signature (Aynaud et al., 2020) that was defined based on independent component analysis of single-cell RNA-seq experiments upon induction of *EWSR1-FLI1* in Ewing cells. This signature is exquisitely specific for Ewing sarcoma, enriched in genes modulated by EWSR1-FLI1 activity on GGAA microsatellite sequences and mostly devoid of cell-cycle genes, which are frequently confounding factors in such GSEA analyses. Apart from a borderline significant signature observed at 24 h post transfection in A673 cells knocked down for *STAG2*, all other comparisons yielded strongly significant GSEA results with the IC-EWS gene set (Figures S2B–S2G). Direct comparison also showed that IC-EWS distinguish *STAG2* WT and mutated cell lines and tumors (Figures S2H and S2I). Further evidences for increased activity of EWSR1-FLI1 in *STAG2*-proficient conditions are provided by Riggi and Miyagawa datasets highlighting genes that are upregulated upon ectopic expression of *EWSR1-FLI1* in mesenchymal/progenitor cells. Beyond EWSR1-FLI1-related signatures, other gene sets provided much less consistent information across datasets and hence appeared less meaningful (Table 1). One gene set each of epidermal growth factor (EGF) or transforming growth factor β (TGF-β) signaling, and one gene set each of MYC, P53, and nuclear factor κB (NF-κB) targets were ranked among the first 20 gene sets

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

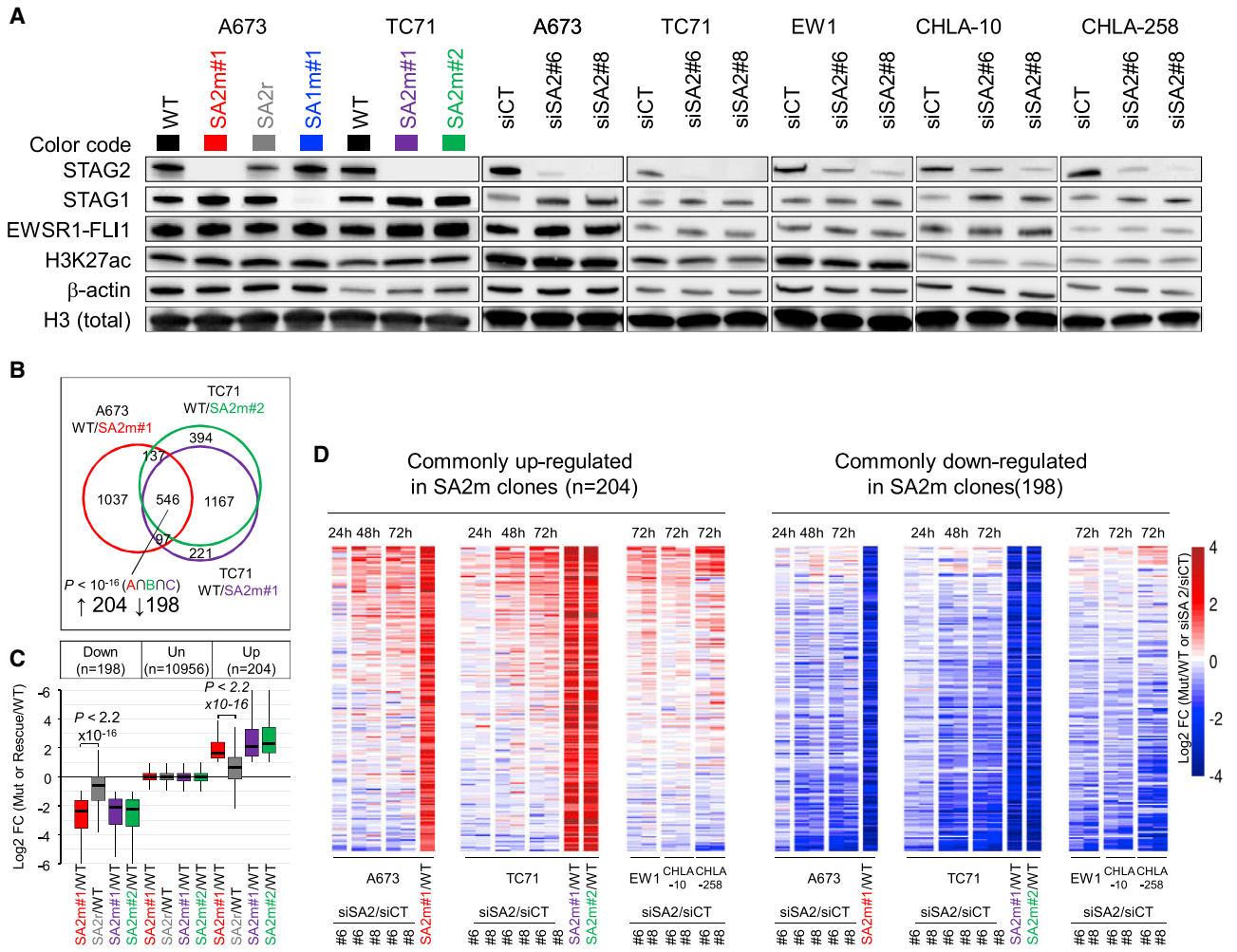


Figure 1. *STAG2* knockout profoundly alters the transcriptomic landscape

(A) Representative western blotting in cellular extracts from isogenic *STAG2* knockout (KO) (generated with two independent sgRNAs: SA2m#1 and SA2m#2), *STAG1* KO (sgRNA: SA1m#1), *STAG2* rescued (sgRNA: SA2r) and *STAG2* knockdown (KD) at 48 h (generated with two independent siRNAs: siSA2#6 and siSA2#8) Ewing sarcoma cells. Color code for sgRNA isogenic models is indicated for each model and kept identical throughout the article.

(B) Scaled Venn diagram for modulated genes between *STAG2* WT and KO conditions (n = 3); total modulated genes for each condition represent the sum of intra-circle numbers; universe includes expressed genes (n = 13,780). p value for intersection was calculated with the SuperExact test.

(C) Boxplots of log2 fold change for up-, un-, and downregulated genes in *STAG2* KO and *STAG2* rescued cells as compared with A673 or TC71 parental cells (n = 3 for each model); number of genes is indicated for each category. p values: two-tailed paired Wilcoxon test. Boxes represent the central 50% of data points (interquartile range). Upper and lower whiskers represent the largest and smallest observed values within 1.5 times the interquartile range from the ends of the box.

(D) Heatmap for the core set of commonly upregulated (left panel) and downregulated (right panel) genes identified in (B and C) for *STAG2* KO and KD Ewing cell lines. Time after siRNA transfection is indicated at the top, sgRNA and siRNA identifiers at the bottom.

See also Figures S1 and S2, Table S1.

(Tables 1 and S3). It is noteworthy that the P53 gene set is particularly enriched in analyses performed in cell lines and in tumors, an observation that may be linked to the frequent association of *STAG2* and *TP53* mutations in Ewing cell lines and tumors (Tirode et al., 2014). In contrast to downregulated genes in *STAG2* LOF systems, analyses of upregulated gene sets did not provide obvious illuminating information (Tables 1 and S3). NES scores were usually weaker and the highest scores observed in the tumor comparison were not strongly supported by the cell line systems. We may nevertheless note the YAP/TAZ pathway, which has been recently shown to counteract EWSR1-FLI1 activity (Katschnig et al., 2017; Rodríguez-Núñez et al., 2020) (Table 1). Altogether, these analyses showed that *STAG2* inactivation has a major impact on EWSR1-FLI1 upregulated genes and hence suggest that STAG2 may modulate EWSR1-FLI1 transcriptional effects.

Binding patterns at H3K27ac, EWSR1-FLI1, CTCF, and cohesin sites are mostly unmodulated upon *STAG2* KO

Our primary hypothesis was that the inactivation of *STAG2* may impair EWSR1-FLI1 binding and/or accessibility to chromatin. Hence, we performed western blot and chromatin

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

Table 1. Top 20 gene sets enriched in *STAG2* proficient or deficient Ewing sarcoma datasets

			NES					p value				FDR			
			1	2	3	4		1	2	3	4	1	2	3	4
Enriched in	Rank	Gene sets (18,889 gene sets from MSigDB or published data)	siSA2	Isogenic	Cell lines	Tumors	**AVG NES**	●: $p < 0.01$				●: 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_GLIOBLASTOMA_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_FLII_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_FCERI_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-FLI_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_GLIOBLASTOMA_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_PDGFB_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**		●	●				●	

(*Continued on next page*)

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

Table 1. ***Continued***

			NES					p value				FDR			
			1	2	3	4		1	2	3	4	1	2	3	4
Enriched in	Rank	Gene sets (18,889 gene sets from MSigDB or published data)	siSA2	Isogenic	Cell lines	Tumors	**AVG NES**	●: p < 0.01				●: FDR <0.2			
	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**				●				●

see also Figure S2, Tables S2 and S3.

(1) Ewing sarcoma cell lines (A673, CHLA-10, CHLA-258, EW1, and TC71), siCT (n = 10) versus siSA2#6/8 (n = 20) at 72 h. (2) Ewing sarcoma isogenic lines, $A673^{WT}$, $TC71^{WT}$ (n = 6) versus $A673^{SA2m\#1}$, $TC71^{SA2m\#1}$, and $TC71^{SA2m\#2}$ (n = 9). (3) Panel of Ewing sarcoma cell lines, *STAG2* WT (n = 15) versus *STAG2* mutated (n = 15). (4) Panel of Ewing sarcoma tumors, *STAG2* WT (n = 32) versus *STAG2* mutated (n = 13). NES, normalized enrichment score. FDR, false discovery rate.

immunoprecipitation sequencing (ChIP-seq) experiments against plausible causative factors in our isogenic models. A decreased expression of EWSR1-FLI1, which could account for its decreased activity, was not observed in isogenic or in knocked down cells. The opposite, a slight increase of EWSR1-FLI1 protein expression, could be observed in TC71 *STAG2* KO and A673 *STAG2* KD cells (Figure 1A). Similarly, no consistent variation of the level of H3K27ac was observed (Figure 1A). ChIP-seq analyses of EWSR1-FLI1 and H3K27ac also did not provide an explanation for decreased EWSR1-FLI1 activity (Figures 2A, 2B and S3A–S3C). A slight increase of EWSR1-FLI1 binding can be noted in TC71 KO and A673 KD cells (Figures S3B and S3C), which may possibly reflect the increased expression of EWSR1-FLI1 mentioned above but this cannot account for the paradoxical decreased activity of this protein. As STAG2 is a member of the cohesin complex, we also investigated the expression and binding patterns of CTCF and of other subunits of the cohesin complex. None of the expression or binding patterns were altered upon *STAG2* KO, apart from a slight, possibly compensatory, increase of STAG1 expression and binding to DNA in *STAG2* KO cells (Figures 1A, 2C and S3A–S3C). As a prototypic locus, we used the *DKK2* gene, which is a well-known EWSR1-FLI1 target (Kauer et al., 2009; Miyagawa et al., 2009; Riggi et al., 2008), which displays EWSR1-FLI1-bound GGAA microsatellites as plausible *cis*-regulatory enhancer elements and for which all genomic features can be displayed in a single panel. Expression of *DKK2* is strongly decreased in *STAG2* KO cells and restored in *STAG2* rescued cells (Figure 2D). Figure 2E and S3 illustrate the global conservation of CTCF, cohesin, EWSR1-FLI1, and enhancer H3K27ac marks at the *DKK2* locus in *STAG2* KO cells. Genome wide, we also noticed that *STAG1*-specific- or *STAG2*-specific cohesin binding sites were rare (Figures 2C and S3A–S3C) compared with recently published data (Kojic et al., 2018; Viny et al., 2019). Altogether, protein expression and ChIP-seq profiles provide no consistent explanation for the decreased EWSR1-FLI1 signature observed in *STAG2*-deficient Ewing sarcoma cells.

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; Merkenschlager 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 (250 kb bins) resolution, no obvious change of the global ''plaid'' pattern was observed between the different isogenic cells (Figure S4A). At high resolution (5 kb bins), the size and positions of the loop domains were mostly unaffected (Figure S4B). However, a striking difference

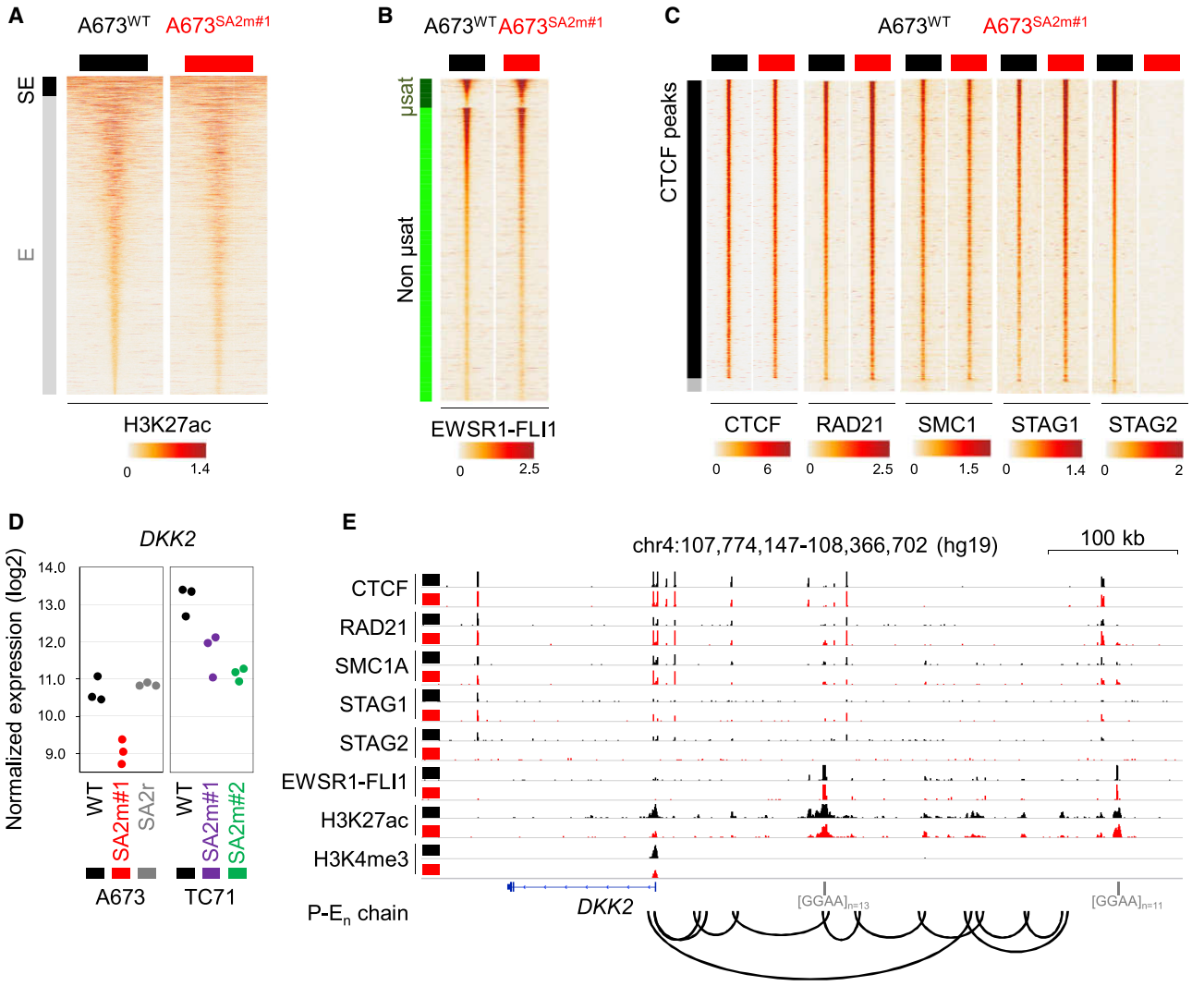


Figure 2. Binding pattern at H3K27ac, EWSR1-FLI1, CTCF, and cohesin sites is mostly unaffected upon *STAG2* KO

(A–C) Heatmap representation of ChIP-seq data in *STAG2* WT or KO A673 cells for (A) H3K27ac, ranked by ROSE algorithm at super-enhancer (SE) and enhancer (E) sites (B) EWSR1-FLI1, ranked by peak intensity at GGAA microsatellite (µsat) and non-µsat sites, and (C) CTCF/cohesin members, ranked by STAG2 peak intensity. Read density is displayed within a 20 kb (H3K27ac) or 4 kb (other marks) window around the peak center, and color scale intensities are shown in normalized coverage (the scale is shown on the bottom of each panel). Presence (black) or absence (gray) of CTCF peaks is indicated on the left.

(D) Dot plot for *DKK2* expression (n = 3) for each model.

(E) ChIP-seq binding profiles at the *DKK2* locus for CTCF, EWSR1-FLI1, cohesin members, H3K27ac, and H3K4me3 histone marks in *STAG2* WT or KO A673 cells. Numbers of consecutive GGAA repetitions are shown in gray. Promoter-enhancer H3K27ac HiChIP-inferred chain at the *DKK2* locus is shown.

See also Figure S3.

could be observed throughout the genome at "stripes" that reveal high interaction frequencies between a single CTCF locus and contiguous loci (Figures 3A–3H, S4B, and S4C). 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 (Figures 3B, S4B, and S4C) as compared with *STAG2* WT or rescued cells (Figures 3A, 3C, S4B, and S4C). CTCF HiChIP experiments were also conducted at 72 h after transfection with two different siRNAs (siSA#6 and siSA#8) in A673 cells and confirmed the stripe reduction (Figures 3D–3F and S4B). To further investigate this observation in a cellular context where only STAG2 is expressed, we also generated Ewing cells with a KO of *STAG1*, the paralog of *STAG2* (Figure 1A). In contrast to *STAG2* KO cells, stripes were even more intense in *STAG1* KO cells compared with WT (Figures 3G and S4B).

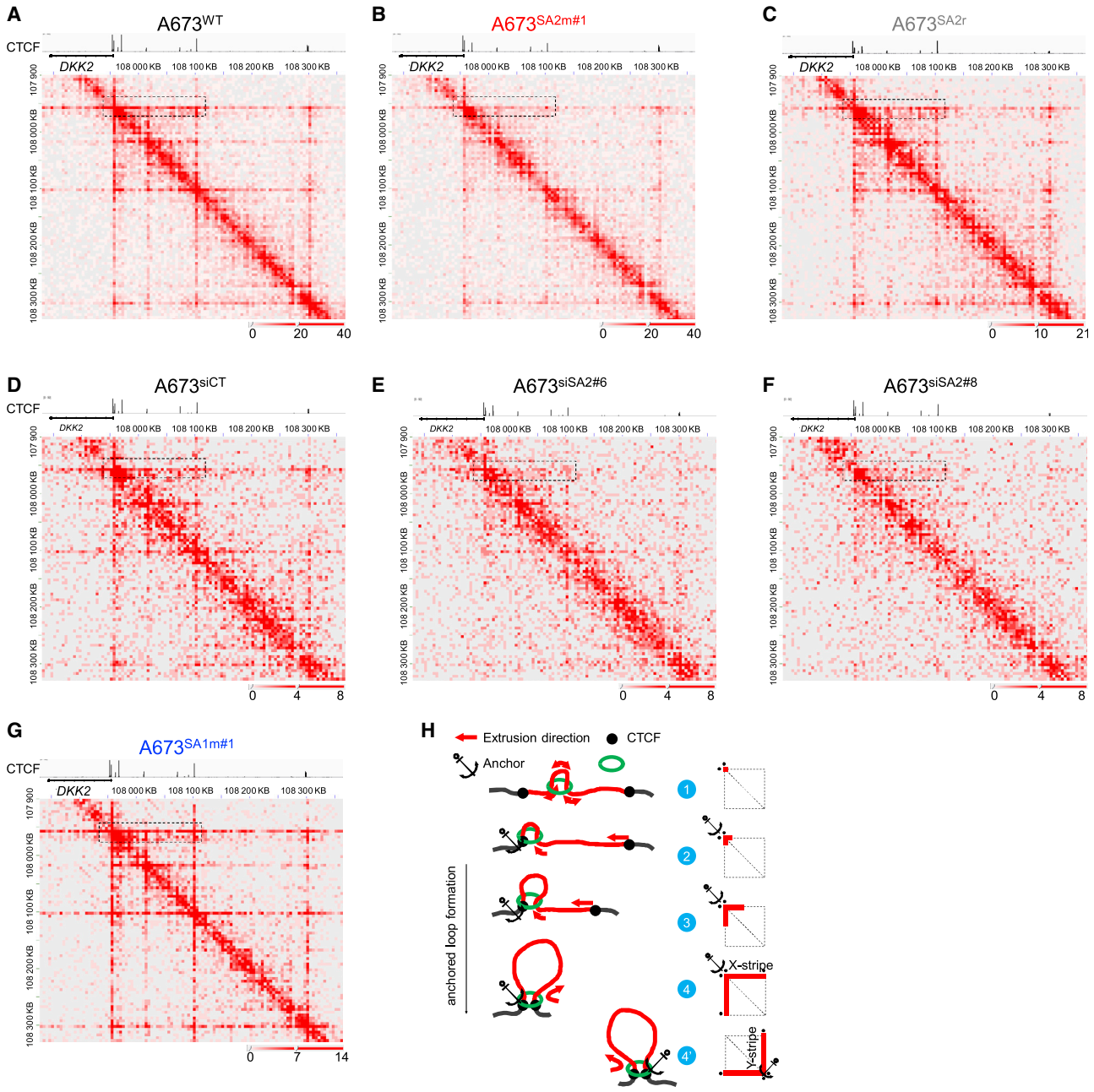


Figure 3. High-resolution CTCF HiChIP data at the *DKK2* locus

(A–G) (A) *STAG2* WT (n = 5), (B) *STAG2* KO (n = 5), (C) *STAG2* rescued (n = 3), (D) *STAG2* KD cells 72 h after siRNA transfection with siCT (n = 2), (E) siSA2#6 (n = 2), (F) siSA2#8 (n = 2), and (G) *STAG1* KO (n = 2) A673 cells. Each panel is scaled to total CTCF HiChIP valid pairs at a 5 kb/bin resolution. CTCF ChIP-seq profiles are shown above each Juicebox print screens HiChIP data. Dashed boxes highlight the region with the first CTCF loop-associated stripe.

(H) Scheme illustrating the process of loop extrusion: (1) upon cohesin loading on chromatin, extrusion can occur possibly in any direction allowing ultimately interactions between CTCF and cohesin. (2) In this model, chromatin is then anchored by the CTCF/cohesin complex at the left CTCF site and extrusion allows for the formation of (3) an intermediate and (4) a fully extruded loop. On the right, schemes illustrating CTCF HiChIP data with loop initiation from the left CTCF/cohesin site, resulting in an X stripe pattern. (4′) An extrusion process anchored at the right CTCF/cohesin site results in a Y stripe pattern.

See also Figure S4, Table S4.

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

To quantify and expand these observations at the genome-wide level, we developed an algorithm (Tweed) for the detection of stripes and the resulting loop domains in simple or interlaced loop regions (Figures 4A and S5A–S5H; STAR methods). The vast majority of detected stripes were flanked by convergent CTCF sites (Figure S5E) and were intensely decorated by cohesin members therefore demonstrating efficient identification of loop domains (Figures S5F and S5G). Because loop formation

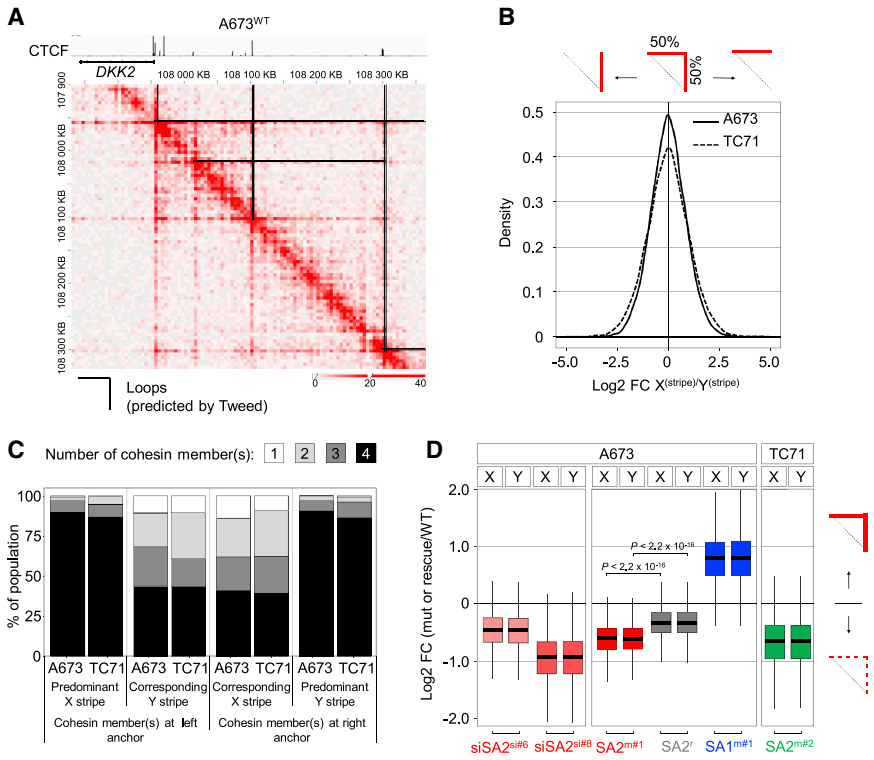


Figure 4. Genome-wide loop detection highlights the role of STAG2 in the anchored extrusion process

(A) Loop domains prediction at the *DKK2* locus in A673 WT CTCF HiChIP data.

(B) Density distribution of X and Y stripes for each loop across the genome according to Tweed prediction in CTCF HiChIP data from A673 and TC71 *STAG2* WT cells. Top scheme: most loops present a symmetric distribution centered on zero, indicating a 50% X stripe- and 50% Y stripe-based anchorage. Unbalanced CTCF/cohesin signal at one of the two anchorage sites results in asymmetric stripe patterns.

(C) Global analysis for the presence of one, two, three, or four cohesin peak(s) associated with CTCF peaks in Tweed-detected loop domains presenting predominant X or Y stripes (anchored at either left or right CTCF sites) in A673 (n = 5) and TC71 (n = 2) cells.

(D) Boxplots of X or Y stripe ratios in CTCF HiChIP data between $A673^{siCT}$ and *STAG2* KD ($A673^{siSA2\#6}$, $A673^{siSA2\#8}$) at 72 h, *STAG1/2* KO models ($A673^{SA1m\#1}$, $A673^{SA2m\#1}$, $TC71^{SA2m\#2}$), and parental WT (A673, TC71) or *STAG2* rescued ($A673^{SA2r}$) cells: A673 n = 18,774; TC71 n = 14,263. Right scheme: positive- and negative-ratio values indicate a gain or a loss of stripes, respectively. Two-tailed paired t test on normalized coverage (log10) between parental/siCT and KO/KD or *STAG2* KO and rescued models are all highly significant ($p < 2.2 \times 10^{-16}$). Boxes represent the central 50% of data points (interquartile range). Upper and lower whiskers represent the largest and smallest observed values within 1.5 times the interquartile range from the ends of the box.

See also Figures S5 and S6, Table S4.

in the anchored extrusion model can emerge from either left or right CTCF boundaries, we hereafter also distinguish X and Y stripes (Figures 3H, S5H, S6A, and S6B). Three major findings could thus be highlighted: first, 18,774 ($A673^{WT}$) and 14,263 ($TC71^{WT}$) loops were detected throughout the genome (Figures S6C–S6E). Here, these loops are detected through stripe structures and not through corner peak dots as usually reported in Hi-C contact maps data. These results are, however, consistent with numbers previously reported in mammalian genomes, thus indicating that anchored extrusion probably represents a general mechanism of loop formation (Figure 3H). Second, most loops showed equivalent X and Y stripe intensities, indicating an extrusion process with balanced anchorage at either CTCF loop boundary (Figure 4B). In the few loops presenting skewed X or Y stripe patterns, the anchored CTCF boundary was predominantly decorated by cohesin members (Figures 4C, S6A, and S6B). Third, and most importantly, anchored loop extrusions were strongly reduced, genome wide, in *STAG2* KO cells and significantly reverted upon *STAG2* restoration (Figures 4D and S6E). In contrast, these stripes were strongly reinforced in *STAG1* KO as compared with parental WT cells (Figure 4D). In *STAG2* KO cells, only very few stripes demonstrated increased signals (Figures 4D and S6E). Moreover, outlier stripes were not consistent between the different *STAG2* KO models. Similarly, only outlier and non-consistent stripes demonstrated decreased signal in *STAG1* KO cells (Figure 4D and S6E). Decreased stripes were also readily observed in *STAG2* knocked down cells as compared with cells transfected with a control siRNA (Figures 3D–3F and 4D). Altogether, these experiments indicated that *STAG2* LOF impacts the loop extrusion process genome wide and that this phenotype is an early, likely direct consequence of the inactivation of *STAG2*.

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

Abundant evidence shows that CTCF/cohesin chromatin loops are preferential structures allowing for promoter-enhancer interactions (Dixon et al., 2012; Dowen et al., 2014; Guo et al., 2015; Hnisz et al., 2016; Lupiáñez et al., 2015; Narendra et al., 2015; Phillips-Cremins et al., 2013; Rao et al., 2014; Splinter et al., 2006; Tang et al., 2015; Vian et al., 2018; Weintraub et al., 2017). In particular, recent data indicate that loop extrusion dynamically juxtaposes elements necessary for antibody class switching in B cells (Zhang et al., 2019). We therefore speculated that *STAG2* LOF could impair *cis*-gene activation. To explore this mechanism, we performed H3K27ac HiChIP in *STAG2*-proficient and -deficient cells. At the prototypical *DKK2* locus, we observed a global loss of interactions in *STAG2* KO cells that was reverted upon STAG2 re-expression (Figure S7A). Similarly, other classical EWSR1-FLI1 target genes displayed loss of interaction patterns (Figure S7B). To extend our findings genome wide, we developed a bioinformatic pipeline to detect promoter-enhancer and enhancer-enhancer interaction chains. Starting from 5 kb bins overlapping H3K4me3 promoter peaks, the strongest H3K27ac HiChIP chain interactions were reported

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

until the 20th interaction sites. To validate that this H3K27ac interaction-based, gene expression agnostic, approach accurately identifies regulatory regions, we isolated all genes that include an EWSR1-FLI1-bound GGAA microsatellite sequence within their regulatory chain. A total of 2,331 and 1,625 genes were retrieved in A673 and TC71, respectively (Table S5). Functional analysis of this set of genes confirmed a very significant enrichment in EWSR1-FLI1 upregulated genes (Table S6). We then investigated whether promoter-enhancer interactions were modified in *STAG2* KO cells. H3K27ac ChIP-seq signal at enhancers in these chains remained mostly unaffected upon *STAG2* KO or *STAG2* rescued cells (Figures 5A and 5B), emphasizing our first observation that H3K27ac signal is largely conserved upon *STAG2* LOF (Figures 1A, 2A, and S3A–S3C). In contrast, H3K27ac HiChIP demonstrated a genome-wide global loss of interactions within these chains in *STAG2* KO cells (Figures 5C and 5D). These interactions were significantly restored in *STAG2* rescued isogenic cells (Figure 5C). Both loss and restoration effects were similar at EWSR1-FLI1-bound and -unbound enhancers (Figures S8A–S8D). Since SEs consist of clusters of enhancers (Hnisz et al., 2013; Whyte et al., 2013), we also investigated the impact of *STAG2* LOF at these loci and observed a strong reduction of intra-SE H3K27ac interactions that was rescued in *STAG2* rescued cells (Figure 5E). We then studied the relationship between chromatin loops and loss of H3K27ac promoter-enhancer interactions. We plotted the density of interactions depending upon the distance between *cis*-interacting pairs on the genome (Figure 5F). Patterns of CTCF and of H3K27ac interactions were highly similar in both *STAG2* WT and KO conditions (Figure 5F) with a noticeable decrease of interaction read frequency in a window comprised between 20 and around 500 kb in *STAG2* KO cells (Figure 5F zoom), which was reverted upon *STAG2* re-expression. Quite strikingly, this distance corresponds to the size of most chromatin loops (Figure 5F). Although correlative, these data therefore strongly suggest that altered loop extrusion as a result of *STAG2* LOF impairs efficient formation of promoter-enhancer interactions.

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

We next wondered why loosened chromatin interactions that occur genome wide within loop-congruent distances may particularly impact EWSR1-FLI1-induced genes. Previous publications have shown that EWSR1-FLI1 binding occurs predominantly at enhancers and is particularly enriched at SE (Baldauf et al., 2018; Boulay et al., 2018; Kennedy et al., 2015; Riggi et al., 2014; Sheffield et al., 2017; Tomazou et al., 2015). In particular, EWSR1-FLI1 transforms otherwise quiescent GGAA microsatellites into active "neo-enhancers." We therefore concentrated our analyses on such sequences. EWSR1-FLI1-bound GGAA microsatellite sequences, similarly to other types of enhancers, SEs and transcription start site (TSS) of expressed genes were predominantly located within chromatin loops (Figures 6A and 6B). In *STAG2* WT cells, the number of interaction pairs at EWSR1-FLI1-bound enhancers, whether GGAA microsatellites or single ETS sites, was significantly higher than at other enhancer sites in both A673 and TC71 cells (Figures 6C and 6D). We next investigated the degree of loosened interactions at genes regulated by *STAG2* LOF and noted that, while a decrease of interactions in *STAG2* KO cells was observed for all genes, it was significantly greater for downregulated genes than for un- or upregulated genes in A673 and TC71 cells (Figures 6E and 6F). In addition, genes containing an EWSR1-FLI1-bound GGAA microsatellite in their regulatory chain were more frequently downregulated by *STAG2* LOF compared with other genes (chi-square, $p = 1.4 \times 10^{-5}$ in A673 cells, $p = 2.2 \times 10^{-16}$ in TC71 cells). Altogether these data show that, although the impact of *STAG2* LOF on promoter-enhancer interactions is genome wide, EWSR1-FLI1-activated genes, and particularly those regulated by microsatellite sequences, are predominantly impacted likely due to the abundance of interactions at these sequences.

STAG2 mutation promotes migration and invasive properties of Ewing cancer cells

Previous reports have indicated that, in addition to decreased proliferation, the KD of *EWSR1-FLI1* in Ewing cells partly reverts the phenotype of Ewing cells into that of mesenchymal stem cells, the likely cell-of-origin (Tirode et al., 2007) with morphological changes, such as cell flattening, increased cell matrix adhesion, and increased migration ability (Chaturvedi et al., 2012; Franzetti et al., 2017; Pedersen et al., 2016). As mentioned above, proliferation rate of *STAG2 KO* cells was not obviously and consistently altered and was highly similar to that of parental cells (Figures S1A–S1D). This is consistent with the observation that IC-EWS, which is mostly devoid of cell-cycle genes, is ranked first compared with other EWSR1-FLI1 activation signatures that do include such cell-cycle genes (Tables 1 and S2). An increase of paxillin-associated stress fiber adhesion foci was observed in the A673-derived clone, with *STAG2* rescue (Figures 7A and 7B). Using soft agar assay, clonogenicity of $A673^{SA2m\#1}$ cells was increased as compared with $A673^{WT}$ and $A673^{SA2r}$ cells. However, this was not observed in TC71 isogenic models (Figure S1D). A more striking effect was observed on migration. Wound healing assays documented a strongly decreased healing time in the *STAG2* KO A673 clone as compared with parental cells (Figures 7C and 7D), but could not be conducted in TC71 cells due to massive detachment of the plate at confluency of parental and isogenic TC71 cells. As the wound healing phenotype was poorly rescued in $A673^{SA2r}$ (Figure 7D), we decided to investigate other Ewing cell lines. We first observed that *STAG2* KD in A673 cells with two different siRNAs fully reproduced the increased migration observed in the *STAG2* KO clone (Figure 7E). This early impact on wound healing was further observed using the CHLA-258 Ewing sarcoma cell line (Figure 7F). We also took advantage of a recently engineered human MSC-derived cell line that presents a CRISPR-Cas9-induced t(11;22) translocation leading to an *EWSR1-FLI1* fusion and a deletion of *STAG2* (A.S., unpublished data). This cell line was further engineered to express a DOX-inducible *STAG2* transgene. Induction of *STAG2* expression led to decreased migration that was not observed in control cells (Figures 7G and 7H). We also explored the invasion ability of *STAG2* WT and LOF cells using the transwell assay. Again, the fact that *STAG2* KO or short-term *STAG2* knocked down A673 cells demonstrated increased invasion with rescue in $A673^{SA2r}$ cells (Figures 7I–7K). Similarly, CHLA-258 cells showed increased invasion with two different siRNAs (Figure 7L). In addition, collagen-embedded aggregates of $A673^{SA2m\#1}$ cells generated anisotropic elongated 3D structures (Figure 7M) and displayed (collective)

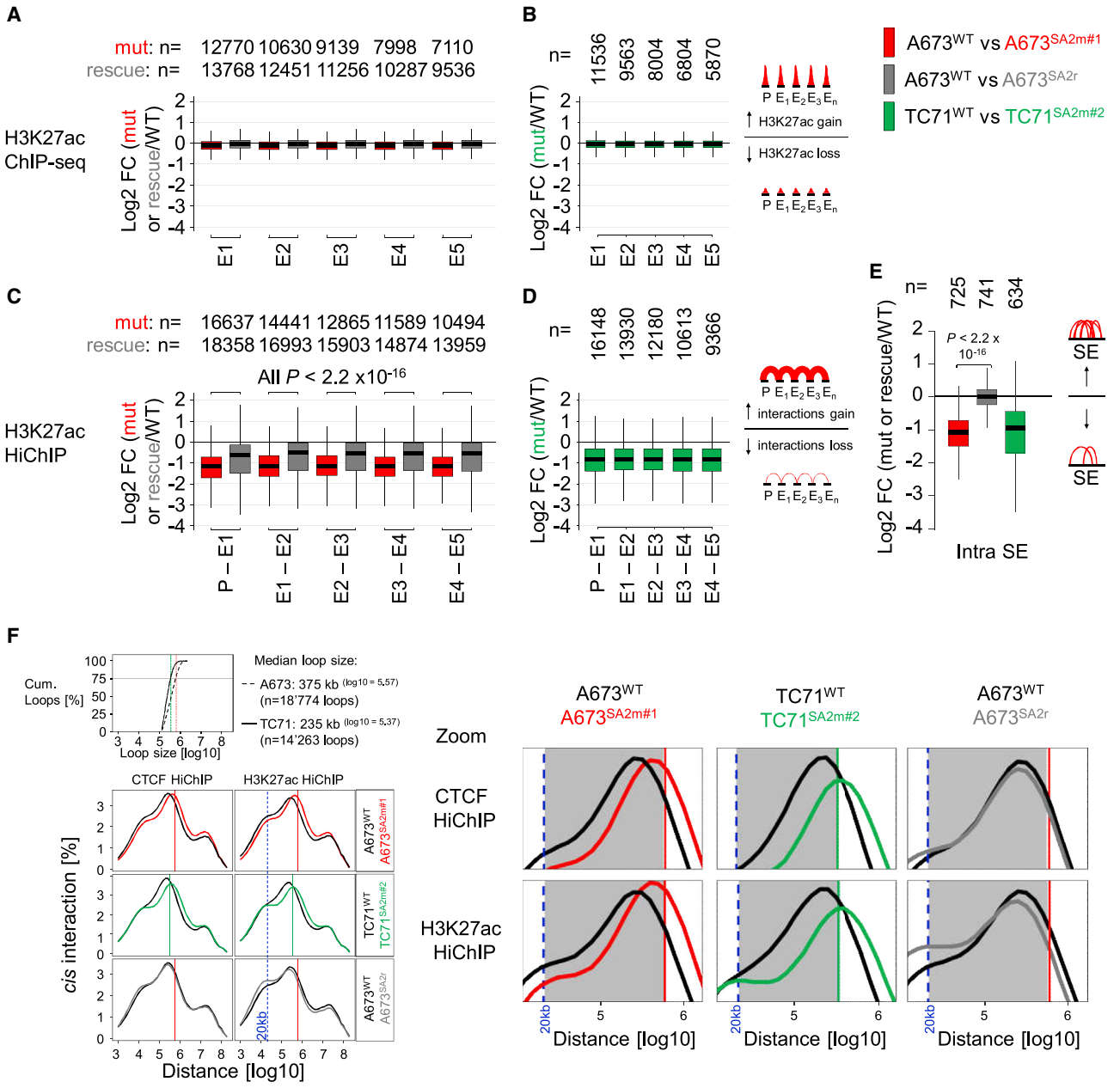


Figure 5. *STAG2* mutation is globally associated with decreased *cis*-promoter-enhancer and enhancer-enhancer interactions within loops

(A–D) Color-coded boxplots of comparative analysis of (A–B) H3K27ac ChIP-seq peak intensities and (C–D) H3K27ac HiChIP interactions along promoter-enhancer chains between *STAG2* WT and KO as well as between *STAG2* KO and rescued cells. p values: two-tailed Wilcoxon test. P (promoter) and E (enhancer) positions in the chain are shown for ranks 1 to 5.

(E) Color-coded boxplots of comparative analysis of intra super-enhancer interactions in H3K27ac HiChIP data between *STAG2* WT and KO or rescued cells. p values: two-tailed t test. (A–E) Boxes represent the central 50% of data points (interquartile range). Upper and lower whiskers represent the largest and smallest observed values within 1.5 times the interquartile range from the ends of the box. The n value for each condition is indicated.

(F) Top, curve of cumulative percentage of loop presence upon genomic distance (log10 scale) in A673 and TC71 (*STAG2* WT). Loop size threshold: 75% of cumulative loops corresponding to 595 kb for A673 (red line) and 340 kb (green line) for TC71. Numbers of loop and median loop size for each cell line is indicated. Bottom, percentage of *cis*-interaction read pairs upon genomic distance between *STAG2* WT, KO, and rescued conditions in CTCF and H3K27ac HiChIP data. A threshold of 20 kb used for H3K27ac chain detection is displayed (blue dashed line). Right, zoom in CTCF and H3K27ac HiChIP plots flanking lower and top thresholds (highlighted in gray).

See also Figures S7 and S8, Tables S4, S5, and S6.

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

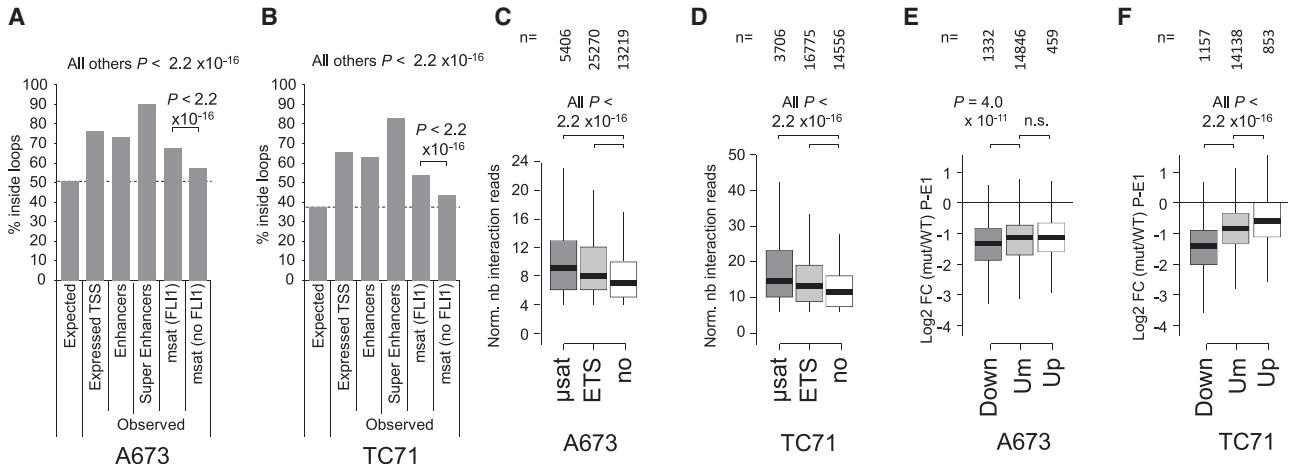


Figure 6. STAG2-induced genes and EWSR1-FLI1 bound elements are highly dependent on *cis*-regulatory interactions

(A and B) Percentage of depicted categories inside loops in (A) A673 and (B) TC71 (*STAG2* WT). Chi-square test between observed and expected percentage was performed for each category. All adjusted p values (Bonferroni) were significant.

(C and D) Normalized numbers of H3K27ac HiChIP interactions at EWSR1-FLI1 (microsatellite, µsat, or single GGAA ETS sites) bound or unbound enhancers in (C) A673 and (D) TC71 (*STAG2* WT). p values: two-tailed paired Wilcoxon test.

(E and F) Log2 fold change (FC) normalized H3K27ac HiChIP interaction ratio between promoter and first enhancer for down-, un-, or upregulated genes upon *STAG2* KO in (E) A673 and (F) TC71 (*STAG2* WT). Boxes represent the central 50% of data points (interquartile range). Upper and lower whiskers represent the largest and smallest observed values within 1.5 times the interquartile range from the ends of the box. p values: two-tailed t test.

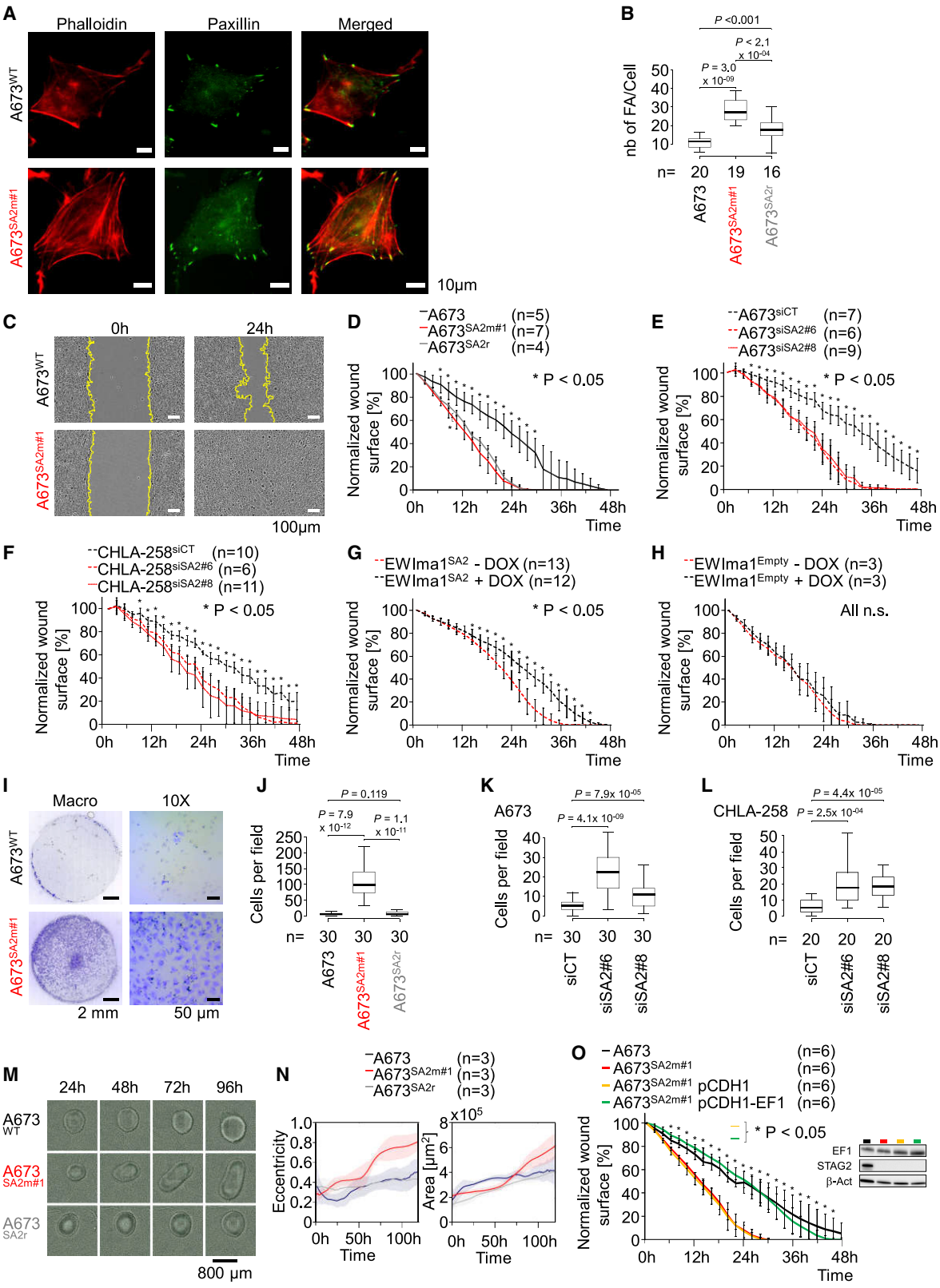
migration prone properties at one leading edge of these structures (Figures 7M and 7N; Videos S1 and S2). These properties were not observed in A673^{WT} and A673^{SA2r} cells, which displayed classical spherical 3D structures. Based on the migratory properties of *EWSR1-FLI1*^{low} (Chaturvedi et al., 2012; Franzetti et al., 2017) and *STAG2* LOF cells (this work), we questioned whether the *STAG2* LOF migratory effect was dependent or independent of the decreased EWSR1-FLI1 activity. We therefore asked if overexpression of EWSR1-FLI1 in *STAG2* KO cells may counteract their migratory properties. Strikingly, A673$^{SA2m\#1}$ cells transduced with an EWSR1-FLI1 expression lentiviral vector displayed a drastic reduction of their wound invasion properties as compared with empty control transduced cells (Figure 7O). This demonstrates that the *STAG2* LOF-induced migratory effect is dependent on EWSR1-FLI1 activity. Finally, using a signature of 39 genes decreased in A673 and TC71 *STAG2* LOF cells and displaying EWSR1-FLI1-bound microsatellite sequences in their promoter-enhancer chains, we show that a lower expression of these genes is associated with adverse prognosis and with the presence of metastasis (Figures S8E–S8G). This provides additional evidence that *STAG2* LOF is associated with metastasis in Ewing sarcoma and that this effect is mediated, at least in part, by EWSR1-FLI1. Altogether, these data show that Ewing cells with *STAG2* LOF recapitulate certain phenotypic characteristics of EWSR1-FLI1^{low} cells and particularly their migration phenotype. Importantly, KD experiments indicate that this phenotype is an early consequence of *STAG2* LOF.

DISCUSSION

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

Our data show that STAG2 and STAG1 demonstrate strikingly different roles in this CTCF-anchored loop extrusion mechanism. Whereas cells expressing only STAG1 exhibit a strongly decreased intensity of stripes, cells expressing only STAG2 demonstrate dramatically increased intensity of these stripes. This strongly suggests that STAG2 enhances the anchored loop extrusion process, while STAG1 is much less efficient in this mechanism, one hypothesis being that STAG1 may promote other types of extrusion mechanisms. While this mechanism was not investigated here, recent papers also indicate that STAG1 and STAG2 have specific roles in genome organization (Arruda

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001



(legend on next page)

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

et al., 2020; Casa et al., 2020). Altogether, our data are consistent with STAG1 and STAG2 having distinct but complementary functions in the topological shaping of the genome which may account for their synthetic lethality (Benedetti et al., 2017; van der Lelij et al., 2017; Liu et al., 2018). *STAG1* and *STAG2* expression levels across various cancers and normal tissues are quite heterogeneous (Romero-Pérez et al., 2019). It can therefore be anticipated that the ratio between STAG1 and STAG2 at loop boundaries may allow for the relaxation or stabilization of this anchorage and consequently fine-tunes the level of enhancer activation within the chromatin loop. In line with the above hypothesis, the increased STAG1/STAG2 ratio in the A673^{SA2r} as compared with A673 parental cells may account for the highly significant but partial rescues of the transcriptional, stripe, and enhancer interaction profiles observed in this study. Other cancers, such as urothelial carcinoma, glioblastoma, or myeloid leukemia display frequent mutations of *STAG2* (Romero-Pérez et al., 2019). Interestingly, applying our Tweed algorithm to recently published CTCF HiChIP data (GSE111537) in *STAG2* WT and KO acute myeloid leukaemia cell models (Smith et al., 2020) also showed a significant decrease of stripe pattern in *STAG2* LOF cells ($p < 2.2 \times 10^{-16}$). This thus indicates that altered chromatin extrusion upon *STAG2* LOF is not limited to Ewing sarcoma and is also observed in other malignancies.

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

Ewing cells demonstrate plasticity between EWSR1-FLI1^{high} and EWSR1-FLI1^{low} states and latter cells are a plausible cause of metastasis (Aynaud et al., 2020; Franzetti et al., 2017). Variations of transcript or protein levels are not the only mechanisms to account for heterogeneous EWSR1-FLI1 activity. Indeed, Wnt/beta-catenin activity was reported to account for metastasis in Ewing tumors through antagonizing EWSR1-FLI1 activity (Pedersen et al., 2016). A similar observation has been made for the YAP/TAZ signaling pathway (Katschnig et al., 2017; Rodríguez-Núñez et al., 2020). Altogether these data indicate that EWSR1-FLI1 activity can vary dynamically, with medium- to high-activity cells being highly proliferative, whereas low activity cells have more limited proliferative ability but have increased mesenchymal properties, including propensity to migrate and invade. Our data show that *STAG2* LOF constitutes an alternative mechanism to moderate EWSR1-FLI1 activity and hence potentially increase the invasive and metastatic potential of Ewing cells. Our data also strongly support the hypothesis that *STAG2* LOF effects on migration and invasion are dependent upon a lower activity of EWSR1-FLI1. Indeed, we show not only that these effects can be reverted by increasing EWSR1-FLI1 expression but that a signature, containing the subset of genes downregulated in *STAG2* KO cells and containing microsatellite EWSR1-FLI1 binding sites in their regulatory chains, is of strong prognostic significance in Ewing sarcoma. *STAG2* LOF may hence induce a shift of the spectrum of EWSR1-FLI1 activity, increasing the proportion of

Figure 7. *STAG2* mutation promotes cell migration

(A and B) (A) Immunofluorescence with phalloidin (red), paxillin (green), and merged pictures in A673 and A673$^{SA2m\#1}$ cells, and (B) boxplots of paxillin-associated focal adhesion loci (FA) per cell in A673, A673$^{SA2m\#1}$, and A673^{SA2r}. p value: two-tailed t test. Boxes represent the central 50% of data points (interquartile range). Upper and lower whiskers represent the largest and smallest observed values within 1.5 times the interquartile range from the ends of the box.

(C) Representative wound healing assay picture at initial and 24 h time points for A673 and A673$^{SA2m\#1}$ cells, computed migration front line is shown (yellow).

(D–H) Quantification of wound surface (cell free) along the assay for: (D) A673, A673$^{SA2m\#1}$, and A673^{SA2r} cells, (E) A673, (F) CHLA-258-transfected cells with siCT or siSA2#6 and siSA2#8, (G) *STAG2*, or (H) empty vector control-inducible EWIma1 cells (MSCs engineered to present *EWSR1-FLI1* translocation and *STAG2* mutation) in the presence or absence of doxycycline to induce *STAG2* expression. Mean ± SD, *p < 0.05: two-tailed t test.

(I) Macroscopic and magnified picture (10×) of a representative membrane from a Boyden chamber migration assay with A673 (n = 3) and A673$^{SA2m\#1}$ (n = 3) cells.

(J–L) Boxplots of cells per field in A673, A673$^{SA2m\#1}$, and A673^{SA2r} cells at 48 h in a Boyden chamber migration assay (J). Boxplots of cells per field in (K) A673 and (L) CHLA-258 cells upon siCT, siSA2#6, and siSA2#8 transfection for 48 h and plated 24 h in a Boyden chamber migration assay. p value: two-tailed t test. Boxes represent the central 50% of data points (interquartile range). Upper and lower whiskers represent the largest and smallest observed values within 1.5 times the interquartile range from the ends of the box.

(M and N) Invasion assay. (M) Representative pictures at the indicated time after collagen embedding and (N) eccentricity and area measurements across time of spheroids generated with A673, A673$^{SA2m\#1}$, and A673^{SA2r} cells. Mean ± SD.

(O) Left, quantification of wound surface along the assay for A673, A673$^{SA2m\#1}$, and A673$^{SA2m\#1}$ transduced with either empty (pCDH1, yellow) or EWSR1-FLI1 (pCDH1-EF1, green) cells. Mean ± SD, p value: two-tailed t test. *p < 0.05, right, a representative western blot for EWSR1-FLI1 (EF1), STAG2, and beta-actin (β-act).

See also Figure S8, Videos S1 and S2.

cells with a low activity, yet keeping the appropriate window of activity for proliferation. *STAG2* LOF Ewing cells might thus combine aggressive features of EWSR1-FLI1^{-high} and EWSR1-FLI1^{-low} cells. Although our data highlight a critical role of *STAG2* LOF in moderating EWSR1-FLI1 activity, it is likely that this alteration may also contribute additional oncogenic functions. In that respect, it is noteworthy that a YAP/TAZ gene set significantly shows up in *STAG2* LOF upregulated genes, suggesting that YAP/TAZ activity may also contribute to lower EWSR1-FLI1 activity. *DKK2* and other WNT antagonists (*KREMEN1*, *FRZB*) and agonists (*LRP5*, *FZD8*) are, respectively, down- and upregulated upon *STAG2* LOF in most datasets, indicating that Wnt signaling activation may also possibly contribute to cell migration. Other transcription factors (MYC, NF-κB) or signaling pathways (EMT transition, TGF-β, and EGF) are impacted upon *STAG2* LOF and further studies will be needed to precisely investigate their potential role in *STAG2* LOF-associated oncogenesis.

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

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Tumor cell lines
  - Patient samples
- METHOD DETAILS
  - CRISPR/Cas9 based genome-editing
  - siRNA
  - Cell infection
  - RNA extraction, cDNA and RT-QPCR
  - Immunoblotting
  - RNA-seq
  - Tumors and cell lines expression data
  - ChIP-seq
  - HiChIP
  - Loop prediction by Tweed
  - Interaction chains
  - Soft agar colony formation assay
  - Actin cytoskeleton and focal adhesion
  - Wound healing migration assay
  - Spheroid invasion assay
  - Image analysis
  - Velocity measurements:
  - Boyden chambers assay
  - GSEA and DoRothEA analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.ccell.2021.04.001.

ACKNOWLEDGMENTS

This work was supported by grants from the Institut Curie: the INSERM; the Cancéropôle Ile-de-France; the Ligue Nationale Contre le Cancer (Equipe labellisée), and projet de Recherche "Enfants, Adolescents et Cancer"; the Institut National du Cancer (PLBIO16-291), the Fondation ARC, the Agence Nationale de la Recherche (ANR-10-EQPX-03, Institut Curie Génomique d'Excellence (ICGex), and the Société Française de lutte contre les Cancers et les leucémies de l'Enfant et de l'adolescent (SFCE). This project also received support from European funding: ERA-NET TRANSCAN JTC-2011 (01KT1310), ERA-NET TRANSCAN JTC 2014 (TRAN201501238), and TRANSCAN JTC 2017 (TRANS201801292), EEC (HEALTH-F2-2013-602856), H2020-IMI2-JTI-201 5-07 (116064-ITCC P4), H2020-SC1-DTH-2018-1 (SEP-210 506374-iPC). We are indebted to the following associations for providing essential support: L'Etoile de Martin, la Course de l'Espoir, M la vie avec Lisa, ADAM, Couleur Jade, Dans les pas du Géant, Courir pour Mathieu, Marabout de Ficelle, Olivier Chape, Les Bagouzamanon, l'association Hubert Gouin, Enfance et Cancer, and les Amis de Claire. D.S. is supported by the Institut Curie-SIRIC (Site de Recherche Intégrée en Cancérologie) program. We thank N. Servant, J. Waterfall, A. Zinovyev, M.M. Aynaud, N. Gruel, V. Berdah, A. Fu, E. Lapouble, E. Tomazou, and all members of the Genetics and Biology of Pediatric Cancers laboratory for helpful discussions. We thank H. Kovar and the Childhood Cancer Repository for providing Ewing cell lines and C. Pierre Eugene for her experimental assistance as well as P. Legoix, O. Mirabeau, C. Kamoun, and E. Barillot for bioinformatic or technical assistance. Finally, we thank J. Waterfall for his diligent proofreading of this paper.

AUTHOR CONTRIBUTIONS

D.S., S.Z., K.L.-D., and L.M. carried out the experiments. S.G. and V.H. performed the bioinformatics analyses. D.S., S.G., and V.H. analyzed processed sequencing data. G.P and F.T. provided patient sequencing data. E.B. and A.S.F. provided EWIma1 models and contributed valuable expertise. S.G. implemented the Tweed algorithm and V.H. the enhancer interaction algorithm. T.V. analyzed spheroid invasion assays. V.R performed with S.B. sequencing experiments. D.S. originally conceived the project and designed the experiments. D.S. and O.D. supervised the work overall and wrote the paper, all authors reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 2, 2019
Revised: August 31, 2020
Accepted: April 2, 2021
Published: April 29, 2021

REFERENCES

Arruda, N.L., Carico, Z.M., Justice, M., Liu, Y.F., Zhou, J., Stefan, H.C., and Dowen, J.M. (2020). Distinct and overlapping roles of STAG1 and STAG2 in cohesin localization and gene expression in embryonic stem cells. Epigenetics Chromatin *13*, 32.

Aynaud, M.-M., Mirabeau, O., Gruel, N., Grossetête, S., Boeva, V., Durand, S., Surdez, D., Saulnier, O., Zaïdi, S., Gribkova, S., et al. (2020). Transcriptional programs define intratumoral heterogeneity of Ewing sarcoma at single-cell resolution. Cell Rep. *30*, 1767–1779.e6.

Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., et al. (2018). Comprehensive characterization of cancer driver genes and mutations. Cell *173*, 371–385.e18.

Baldauf, M.C., Orth, M.F., Dallmayer, M., Marchetto, A., Gerke, J.S., Rubio, R.A., Kiran, M.M., Musa, J., Knott, M.M.L., Ohmura, S., et al. (2018). Robust diagnosis of Ewing sarcoma by immunohistochemical detection of super-enhancer-driven EWSR1-ETS targets. Oncotarget *9*, 1587–1601.

Batra, S., Reynolds, C.P., and Maurer, B.J. (2004). Fenretinide cytotoxicity for Ewing's sarcoma and primitive neuroectodermal tumor cell lines is decreased by hypoxia and synergistically enhanced by ceramide modulators. Cancer Res. *64*, 5415–5424.

Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C.A., Chotalia, M., Xie, S.Q., Barbieri, M., de Santiago, I., Lavitas, L.-M., Branco, M.R., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. Nature *543*, 519–524.

Benedetti, L., Cereda, M., Monteverde, L., Desai, N., and Ciccarelli, F.D. (2017). Synthetic lethal interaction between the tumour suppressor STAG2 and its paralog STAG1. Oncotarget *8*, 37619–37632.

Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N., and Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science *362*, eaau1783.

Boulay, G., Sandoval, G.J., Riggi, N., Iyer, S., Buisson, R., Naigles, B., Awad, M.E., Rengarajan, S., Volorio, A., McBride, M.J., et al. (2017). Cancer-specific retargeting of BAF complexes by a prion-like domain. Cell *171*, 163–178.e19.

Boulay, G., Volorio, A., Iyer, S., Broye, L.C., Stamenkovic, I., Riggi, N., and Rivera, M.N. (2018). Epigenome editing of microsatellite repeats defines tumor-specific enhancer functions and dependencies. Genes Dev. *32*, 1008–1019.

Brohl, A.S., Solomon, D.A., Chang, W., Wang, J., Song, Y., Sindiri, S., Patidar, R., Hurd, L., Chen, L., Shern, J.F., et al. (2014). The genomic landscape of the Ewing sarcoma family of tumors reveals recurrent STAG2 mutation. PLoS Genet. *10*, e1004475.

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

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

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

Dauphinot, L., De Oliveira, C., Melot, T., Sevenet, N., Thomas, V., Weissman, B.E., and Delattre, O. (2001). Analysis of the expression of cell cycle regulators in Ewing cell lines: EWS-FLI-1 modulates p57KIP2and c-Myc expression. Oncogene *20*, 3258–3265.

Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.-M. (2019). DNA loop extrusion by human cohesin. Science *366*, 1338–1345.

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

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

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

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., et al. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374–387.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101.

Franzetti, G.-A., Laud-Duval, K., van der Ent, W., Brisac, A., Irondelle, M., Aubert, S., Dirksen, U., Bouvier, C., de Pinieux, G., Snaar-Jagalska, E., et al. (2017). Cell-to-cell heterogeneity of EWSR1-FLI1 activity determines proliferation/migration choices in Ewing sarcoma cells. Oncogene *36*, 3505–3514.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of chromosomal domains by loop extrusion. Cell Rep. *15*, 2038–2049.

Gangwal, K., Sankar, S., Hollenhorst, P.C., Kinsey, M., Haroldsen, S.C., Shah, A.A., Boucher, K.M., Watkins, W.S., Jorde, L.B., Graves, B.J., et al. (2008). Microsatellites as EWS/FLI response elements in Ewing's sarcoma. Proc. Natl. Acad. Sci. U S A *105*, 10149–10154.

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

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

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

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

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

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., et al. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. Cell *162*, 900–910.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yáñez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., et al. (2017). The cohesin release factor WAPL restricts chromatin loop extension. Cell *169*, 693–707.e14.

Hassler, M., Shaltiel, I.A., and Haering, C.H. (2018). Towards a unified model of SMC complex function. Curr. Biol. *28*, R1266–R1281.

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

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

Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.-L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., et al. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science *351*, 1454–1458.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A phase separation model for transcriptional control. Cell *169*, 13–23.

Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D landscape of transcription-linked mammalian chromatin folding. Mol. Cell *78*, 539–553.e8.

Katschnig, A.M., Kauer, M.O., Schwentner, R., Tomazou, E.M., Mutz, C.N., Linder, M., Sibilia, M., Alonso, J., Aryee, D.N.T., and Kovar, H. (2017). EWS-FLI1 perturbs MRTFB/YAP-1/TEAD target gene regulation inhibiting cytoskeletal autoregulatory feedback in Ewing sarcoma. Oncogene *36*, 5995–6005.

Kauer, M., Ban, J., Kofler, R., Walker, B., Davis, S., Meltzer, P., and Kovar, H. (2009). A molecular function map of Ewing's sarcoma. PLoS One *4*, e5415.

Kennedy, A.L., Vallurupalli, M., Chen, L., Crompton, B., Cowley, G., Vazquez, F., Weir, B.A., Tsherniak, A., Parasuraman, S., Kim, S., et al. (2015). Functional, chemical genomic, and super-enhancer screening identify sensitivity to cyclin D1/CDK4 pathway inhibition in Ewing sarcoma. Oncotarget *6*, 30178–30193.

Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., and Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. Science *366*, 1345–1349.

Kojic, A., Cuadrado, A., De Koninck, M., Giménez-Llorente, D., Rodríguez-Corsino, M., Gómez-López, G., Le Dily, F., Marti-Renom, M.A., and Losada, A. (2018). Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. Nat. Struct. Mol. Biol. *25*, 496–504.

De Koninck, M., Lapi, E., Badía-Careaga, C., Cossío, I., Giménez-Llorente, D., Rodríguez-Corsino, M., Andrada, E., Hidalgo, A., Manzanares, M., Real, F.X., et al. (2020). Essential roles of cohesin STAG2 in mouse embryonic development and adult tissue homeostasis. Cell Rep. *32*, 108014.

Kovar, H., Jug, G., Aryee, D.N., Zoubek, A., Ambros, P., Gruber, B., Windhager, R., and Gadner, H. (1997). Among genes involved in the RB dependent cell cycle regulatory cascade, the p16 tumor suppressor gene is frequently lost in the Ewing family of tumors. Oncogene *15*, 2225–2232.

Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.-H.S., Parsi, K.M., Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural details of mammalian chromosome architecture. Mol. Cell *78*, 554–565.e7.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

van der Lelij, P., Lieb, S., Jude, J., Wutz, G., Santos, C.P., Falkenberg, K., Schlattl, A., Ban, J., Schwentner, R., Hoffmann, T., et al. (2017). Synthetic lethality between the cohesin subunits STAG1 and STAG2 in diverse cancer contexts. eLife *6*, e26980.

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

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

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

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

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

Merkenschlager, M., and Nora, E.P. (2016). CTCF and cohesin in genome folding and transcriptional gene regulation. Annu. Rev. Genomics Hum. Genet. *17*, 17–43.

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

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

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

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

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

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

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

Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell *169*, 930–944.e22.

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

Ochi, Y., Kon, A., Sakata, T., Nakagawa, M.M., Nakazawa, N., Kakuta, M., Kataoka, K., Koseki, H., Nakayama, M., Morishita, D., et al. (2020). Combined cohesin-RUNX1 deficiency synergistically perturbs chromatin looping and causes myelodysplastic syndromes. Cancer Discov. *10*, 836–853.

Pedersen, E.A., Menon, R., Bailey, K.M., Thomas, D.G., Van Noord, R.A., Tran, J., Wang, H., Qu, P.P., Hoering, A., Fearon, E.R., et al. (2016). Activation of Wnt/β-catenin in Ewing sarcoma cells antagonizes EWS/ETS function and promotes phenotypic transition to more metastatic cell states. Cancer Res. *76*, 5040–5053.

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

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

Prieur, A., Tirode, F., Cohen, P., and Delattre, O. (2004). EWS/FLI-1 silencing and gene profiling of Ewing cells reveal downstream oncogenic pathways and a crucial role for repression of insulin-like growth factor binding protein 3. Mol. Cell. Biol. *24*, 7275–7283.

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

Rao, S.S.P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al. (2017). Cohesin loss eliminates all loop domains. Cell *171*, 305–320.e24.

Riggi, N., Suva, M.L., Suva, D., Cironi, L., Provero, P., Tercier, S., Joseph, J.M., Stehle, J.C., Baumer, K., Kindler, V., et al. (2008). EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. Cancer Res. *68*, 2176–2185.

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

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

Romero-Pérez, L., Surdez, D., Brunet, E., Delattre, O., and Grünewald, T.G.P. (2019). STAG mutations in cancer. Trends Cancer *5*, 506–520.

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

Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. U S A *112*, E6456–E6465.

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

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

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

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

Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Elkahloun, A.G., Harris, B.T., Toretsky, J.A., Rosenberg, S.A., Shukla, N., Ladanyi, M., et al. (2011). Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science *333*, 1039–1043.

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

Stadhouders, R., Filion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. Nature *569*, 345–354.

Surdez, D., Benetkiewicz, M., Perrin, V., Han, Z.-Y., Pierron, G., Ballet, S., Lamoureux, F., Rédini, F., Decouvelaere, A.-V., Daudigeos-Dubus, E., et al. (2012). Targeting the EWSR1-FLI1 oncogene-induced protein kinase PKC-β abolishes Ewing sarcoma growth. Cancer Res. *72*, 4494–4503.

Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. Cell *163*, 1611–1627.

Tirode, F., Laud-Duval, K., Prieur, A., Delorme, B., Charbord, P., and Delattre, O. (2007). Mesenchymal stem cell features of Ewing tumors. Cancer Cell *11*, 421–429.

Tirode, F., Surdez, D., Ma, X., Parker, M., Le Deley, M.C., Bahrami, A., Zhang, Z., Lapouble, E., Grossetête-Lalami, S., Rusch, M., et al. (2014). Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. Cancer Discov. *4*, 1342–1353.

Tomazou, E.M., Sheffield, N.C., Schmidl, C., Schuster, M., Schönegger, A., Datlinger, P., Kubicek, S., Bock, C., and Kovar, H. (2015). Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS-FLI1. Cell Rep. *10*, 1082–1095.

Vian, L., Pękowska, A., Rao, S.S.P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., Huang, S.-C., El Khattabi, L., Dose, M., Pruett, N., et al. (2018). The energetics and physiological impact of cohesin extrusion. Cell *175*, 292–294.

Viny, A.D., Bowman, R.L., Liu, Y., Lavallée, V.-P., Eisman, S.E., Xiao, W., Durham, B.H., Navitski, A., Park, J., Braunstein, S., et al. (2019). Cohesin members Stag1 and Stag2 display distinct roles in chromatin accessibility and topological control of HSC self-renewal and differentiation. Cell Stem Cell *25*, 682–696.e8.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 is a structural regulator of enhancer-promoter loops. Cell *171*, 1573–1588.e28.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796–801.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., et al. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J. *36*, 3573–3599.

Xiao, T., Wallace, J., and Felsenfeld, G. (2011). Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity ▽. Mol. Cell Biol. *31*, 2174–2183.

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

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

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

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-STAG1	Abcam	Cat# ab4457; RRID: AB_2286589
anti-STAG2	Santa Cruz	Cat# sc-81852; RRID: AB_2199948
anti-STAG2	Abcam	Cat# ab4463; RRID: AB_304471
anti-FLI1	Abcam	Cat# ab133485; RRID: AB_2722650
anti-H3K27ac	Abcam	Cat# ab4729; RRID: AB_2118291
anti-H3	Abcam	Cat# ab1791; RRID: AB_302613
anti-beta-Actin	anti-beta-Actin	Cat# A5316; RRID: AB_476743
anti-rabbit-IgG-HRP	GE Healthcare	Cat# NA934; RRID: AB_772206
anti-mouse-IgG-HRP	GE Healthcare	Cat# NXA931; RRID: AB_772209
anti-FLI1	Abcam	Cat# ab15289; RRID: AB_301825
anti-RAD21	Abcam	Cat# ab992; RRID: AB_2176601
anti-CTCF	Diagenode	Cat# C15410210; RRID: AB_2753160
anti-SMC1	Bethyl laboratories	Cat# A300-055A; RRID: AB_2192467
anti-H3K4me3	Diagenode	Cat# pAb-003-050; RRID: AB_2616052
Anti-paxillin	BD biosciences	Cat# 610051; RRID: AB_397463
Phalloidin-tetramethylrhodamine B isothiocyanate	Sigma	Cat# P1951; RRID: AB_2315148
anti-mouse-IgG-Alexa Fluor 488	Invitrogen	Cat# A-21202; RRID: AB_141607
Bacterial and virus strains		
MLM3636 vector	Addgene	Addgene_43860
pCas9_GFP	Addgene	Addgene_44719
pCDH1-CMV-MCS-EF1-Puro	System Biosciences, CA	CD510B-1
Stellar™ competent cells	Ozyme	636763
Biological samples		
Frozen human tumor samples	Curie Institute, UGS	N/A
Chemicals, peptides, and recombinant proteins		
Alpha-MEM	Gibco	22561021
DMEM	GE Healthcare	SH30022.01
IMDM	Gibco	12440-53
RPMI	GE Healthcare	SH30027.01
fetal bovine serum	GE Helthcare	SV30160.03
MSC qualified fetal bovine serum	Gibco	2662029
Insuline-Transferrin-selenium 1X	Gibco	41400-045
10X T4 ligation buffer	New England Biolabs	B0202S
PNK	Fermentas	EK0031
10X Tango buffer	Fermentas	BY5
DTT	Sigma	D0632
ATP	Promega	U120D
low melting point agarose	Invitrogen	16520
Ultra-Low Attachment microplates	Corning	7007
Collagen I, High Concentration, Rat Tail,	Corning	354249
BsmB1	New England Biolabs	R05805
T4 DNA Ligase	New England Biolabs	M0202S
4-15% Mini-PROTEAN®TGX gels	BIO-RAD	456-1084

(*Continued on next page*)

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
nitrocellulose membranes	BIO-RAD	1704159
SuperSignal™ West Pico Plus	Thermo Scientific	34580
methanol-free formaldehyde	Thermo-Scientific	28908
protease inhibitor cocktail	Roche	11836145001
Mbo1	New England Biolabs	R0147M
TN5	Illumina	15028211
phusion HF 2X	New England Biolabs	M0531S
Evagreen 20X	Biotium	31000
Ampure XP beads	Beckman Coulter	A63881
Critical commercial assays		
mycoplasma detection kit	Minerva Biolabs	11-9250
QiAamp®DNA mini kit	Qiagen	51304
Nucleofast 96 PCR plate	Macherey Nagel	743100.50
RNeasy plus mini kit	Qiagen	74134
Lipofectamine RNAiMAX	Invitrogen	13778030
cDNA Reverse Transcription Kit	Applied Biosystems	4368814
Power SYBR Green PCR Master	Applied Biosystems	4367659
TruSeq Stranded mRNA Library preparation kit	Illumina	20020594
iDeal ChIP-seq kit for transcription factors	Diagenode	C01010170
iDeal ChIP-seq kit for for histones	Diagenode	C01010171
TruSeq ChIP library preparation kit	Illumina	IP-202-1012
Deposited data		
Ewing sarcoma RNAseq, ChiP-Seq, HiChIP, Affymetrix arrays	This paper	GSE133228
Ewing sarcoma transcriptome	Tirode et al., 2014	EGAS00001003333
Ewing sarcoma Affymetrix arrays	Postel-Vinay et al., 2012	GSE34620
Experimental models: cell lines		
A673	ATCC	ATCC CRL-1598
BOU	Institut Curie	N/A
CHLA-10	COG Repository	N/A
CHLA-258	COG Repository	N/A
COH	Institut Curie	N/A
EW1	IARC	N/A
EW16	IARC	N/A
EW22	Institut Curie	N/A
EW3	IARC	N/A
EW7	IARC	N/A
EWIma1^{SA2}	A.S., unpublished data	N/A
EWIma1^{Empty}	A.S., unpublished data	N/A
MHH-ES1	DSMZ	N/A
MIC	Institut Curie	N/A
ORS	Institut Curie	N/A
POE	Institut Curie	N/A
RM82	Kovar et al., 1997	N/A
SK-ES-1	ATCC	N/A
SK-N-MC	ATCC	N/A
STA-ET-1	Kovar et al., 1997	N/A
STA-ET-11	Kovar et al., 1997	N/A
STA-ET-2.1	Kovar et al., 1997	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
STA-ET-2.2	Kovar et al., 1997	N/A
STA-ET-7.1	Kovar et al., 1997	N/A
STA-ET-7.2	Kovar et al., 1997	N/A
STA-ET-7.3	Kovar et al., 1997	N/A
STA-ET-8.1	Kovar et al., 1997	N/A
TC-252	Kovar et al., 1997	N/A
TC-71	DSMZ	ACC 516
VH64	Kovar et al., 1997	N/A
WE-68	Kovar et al., 1997	N/A
WE-M2-68	Kovar et al., 1997	N/A
Oligonucleotides		
sgRNA primer SA1m#1: forward, 5′-ACA CCGATGTGCCGAGTACACCAAGG-3′	This paper	N/A
sgRNA primer SA1m#1: reverse, 5′-AAA ACCTTGGTGTACTCGGCACATCG-3′	This paper	N/A
sgRNA primer SA2m#1: forward, 5′-ACA CCGCGACATACAAGCACCCTGGCG-3′	This paper	N/A
sgRNA primer SA2m#1: reverse, 5′-AAA ACGCCAGGGTGCTTGTATGTCGCG-3′	This paper	N/A
sgRNA primer SA2m#2: forward, 5′-ACA CCGATTTCGACATACAAGCACCCG-3′	This paper	N/A
sgRNA primer SA2m#2: reverse, 5′-AAA ACGGGTGCTTGTATGTCGAAATCG-3′	This paper	N/A
sgRNA primer SA2r: forward, 5′-ACA CCGTTTCGACATACAAGCACCCTG-3′	This paper	N/A
sgRNA primer SA2r: reverse, 5′-AAA ACAGGGTGCTTGTATGTCGAAACG-3′	This paper	N/A
oligonucleotide template SA2t: 5′- ctT CTTACAGGATTGTCTGACTCACAAGT CAGAGCATTTCGACATACAAGCACG CTAGCAGGTCGGTATTTAGAAATA TTTTCTGCATAttg -3′, nucleotides with phosphorotioate modification are shown in lower case	This paper	N/A
All stars negative Control siRNA (proprietary target sequence, not disclosed)	Qiagen	1027281
siRNA targeting *EWSR1-FLI1* fusion (Type 1) (DNA template provided for ordering) 5′-AAGGCAGCAGAACCCTTCTTA-3	Prieur et al., 2004	N/A
siRNA targeting *EWSR1-FLI1* fusion (Type 2)5′- GGCAGCAGAGUUCACUGCUd CdG-3	This paper	N/A
ON-TARGETplus Non-targeting Control siRNA #1 (proprietary target sequence, not disclosed)	Dharmacon	D-001810-01-50
siRNA targeting *STAG2* (siSA2#6) 5′-GUAGAUGAUUGGAUAGAAU-3’	Dharmacon	ON-Targert J021351-06
siRNA targeting *STAG2* (siSA2#8) 5′-CCACUGAUGUCUUACCGAA-3’	Dharmacon	ON-Targert J021351-08
PCR *STAG1* (for SA1m#1 CRIPSR/Cas9 validation): forward, 5′-AGCATCCTCAAGGCTGTGA-3′	This paper	N/A

(*Continued on next page*)

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PCR *STAG1* (for SA1m#1 CRIPSR/Cas9 validation) : reverse, 5′-TCAGTGGAAGTGAAGAAGCTCT-3′	This paper	N/A
PCR *STAG2* (for SA2m#1, SA2m#2 and SA2r CRIPSR/Cas9 validation): forward, 5′-AGAGCTGAAGTGTTCAGAGGT-3′	This paper	N/A
PCR *STAG2*: (for SA2m#1, SA2m#2 and SA2r CRIPSR/Cas9 validation): reverse, 5′-AGGAATTCGCAGGAGGGATG-3′)	This paper	N/A
RT-QPCR primer for *CAV2* :forward, 5'-TCAACTCGCATCTCAAGCTG-3'	This paper	N/A
RT-QPCR primer for *CAV2* :reverse, 5'-GATTTCAAAGAGGGCATGGC-3'	This paper	N/A
RT-QPCR primer for *DKK2* :forward, 5'-CCAGTACCCGCTGCAATAA-3'	This paper	N/A
RT-QPCR primer for *DKK2* :reverse, 5'-ATGACCGTGGTTTCGATCTC-3'	This paper	N/A
RT-QPCR primer for *EGR2* :forward, 5'-TTGACCAGATGAACGGAGTG-3'	This paper	N/A
RT-QPCR primer for *EGR2* :reverse, 5'-GCCCATGTAAGTGAAGGTCTG-3'	This paper	N/A
RT-QPCR primer for *NR0B1* :forward, 5'-ATGCTGGAGTCTGAACATCAG-3'	This paper	N/A
RT-QPCR primer for NR0B1 :reverse, 5'-TGAGTATTTGCTGAGTTCCCC-3'	This paper	N/A
RT-QPCR primer for *RPLP0* :forward, 5'-GAAACTCTGCATTCTCGCTTC-3'	Surdez et al. (2012)	N/A
RT-QPCR primer for *RPLP0* :reverse, 5'-GGTGTAATCCGTCTCCACAG-3'	Surdez et al. (2012)	N/A
RT-QPCR primer for *STAG2* :forward, 5'-TCAGTCGTAGAGATCCAGAGG-3'	This paper	N/A
RT-QPCR primer for *STAG2* :reverse, 5'-TCCCACATGCTATCCACAAG-3'	This paper	N/A
Software and algorithms		
GSEAv4.0.1		N/A
R v3.4		https://cran.r-project.org
Bio-Rad CFX Manager 3.1 software	Bio-Rad	N/A
bowtie 2 2.2.9	Langmead and Salzberg (2012)	N/A
samtools 1.3	Li et al. (2009)	N/A
MACS2 2.1.1	Zhang et al. (2008)	N/A
ROSE 0.1	Whyte et al. (2013)	N/A
HiC-Pro 2.10.1	Servant et al. (2015)	N/A

RESOURCE AVAILABILITY

Lead contact

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

Materials availability

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

Data and code availability

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

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Tumor cell lines

The Ewing sarcoma A673 cell line was obtained from the American Type Culture Collection (ATCC) and the Ewing sarcoma TC71 cell line was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). EW1 cell line was obtained from International Agency for Research on Cancer (IARC), CHLA-10 and CHLA-258 cell lines were obtained from Childhood Cancer Repository (COG Repository). EWIma1 cells were generated in Erika Brunet laboratory by Anna Sole Ferre at Imagine Institute in Paris (A.S., unpublished data). EWIma1 STAG2 (EWIma1^{SA2}) and empty (EWIma1^{Empty}) tetracycline inducible model cells were generated by transfecting EWIma1 cells with respectively empty pCW57-GFP-2A MCS vector (Addgene) or pCW57-GFP-2A-MCS-STAG2 vector. Transfection was performed with the Amaxa Nucleofector I Device (Lonza) according to the manufacturer's protocol and cells were selected with puromycin at 1μg/mL. Cells were cultured in (A673) DMEM (GE Healthcare, SH30022.01) or (TC71 and EW1) RPMI (GE Healthcare, SH30027.01,) supplemented with 10% fetal bovine serum (GE Helthcare, SV30160.03). CHLA-10 and CHLA-258 were cultured in IMDM (Gibco, 12440-53) supplemented with 20% fetal bovine serum and 1X Insuline-Transferrin-selenium (Gibco, 41400-045). EWIma1^{SA2} and EWIma1^{Empty} were cultured in Alpha-MEM (Gibco, 22561021) supplemented with 10% of MSC qualified FBS (Gibco, 12662029), 1% L-glutamine (Gibco, 25030024) and 1% Penicillin-streptomycine (Gibco, 15140122). When available, STR profiling proved each cell line matched with the reference profile provided by ATCC and DSMZ, respectively. Cells were routinely tested negative for mycoplasma contamination by qPCR (VenorGeM qEP (11-9250, Minerva Biolabs).

Patient samples

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

METHOD DETAILS

CRISPR/Cas9 based genome-editing

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

Sense and antisense primers matching sgRNA regions in *STAG1* (SA1m#1: forward, 5′-ACACCGATGTGCCGAGTACACCAAGG-3′; reverse, 5′-AAAACCTTGGTGTACTCGGCACATCG-3′), *STAG2* (SA2m#1: forward, 5′-ACACCGCGACATACAAGCACCCTGGCG-3′; reverse, 5′-AAAACGCCAGGGTGCTTGTATGTCGCG-3′ and SA2m#2: forward, 5′-ACACCGATTTCGACATACAAGCACCCG-3′; reverse, 5′-AAAACGGGTGCTTGTATGTCGAAATCG-3′) and *STAG2* mutated (SA2r: forward, 5′-ACACCGTTTCGACATACAAG CACCCTG-3′; reverse, 5′-AAAACAGGGTGCTTGTATGTCGAAACG-3′)

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

Cells were transfected using Amaxa Nucleofector™ (Lonza) following manufacturer instruction. Briefly, one million of cells were co-transfected with 2 μg of MLM3636-SA1m#1 or MLM3636-SA2m#1 or MLM3636-SA2m#2 plasmid and 2 μg of pCas9_GFP plasmid (gift from Kiran Musunuru, Addgene plasmid # 44719; http://n2t.net/addgene:44719; RRID: Addgene_44719) and 1 μg of pCDH1-CMV-MCS-EF1-Puro plasmid (System Biosciences, CA) providing puromycin resistance in V Nucleofector ™solution using T020 program. 24 hours after transfection, the cells are selected during 2 days with 0.5 μg/mL or 1 μg/mL of puromycin for TC71 and A673 cell lines respectively. Transfected cells were then cloned and screened for knockout using Western blot to assess STAG1 and STAG2 expression. For rescue experiment, A673SA2m#1 cells were co-transfected with 2 μg of MLM3636-SA2r, 2 μg of single stranded phosphorotioate modified oligonucleotide template (SA2t: 5′- ctTCTTACAGGATTGTCTGACTCACAAGTCAGAGCATTTC GACATACAAGCACGCTAGCAGGTCGGTATTTAGAAATATTTTCTGCATAttg -3′, nucleotides with phosphorotioate modification are shown in lower case), 2 μg of pCas9_GFP plasmid and 1 μg of pCDH1-CMV-MCS-EF1-Puro plasmid as described above and isolated as for the other isogenic models. All isogenic models were validated using Sanger sequencing. DNA was extracted from clones of interest using QiAamp®DNA mini kit (51304, Qiagen) according to the manufacturer's instructions. PCR reaction was performed to amplify sgRNA targeting region in *STAG1* or *STAG2* loci (1619-*STAG1*: forward, 5′-CAGCATCCTCAAGGCTGTGA-3′;

reverse, 5′-TCAGTGGAAGTGAAGAAGCTCT-3′ and 1460-*STAG2*: forward, 5′-AGAGCTGAAGTGTTCAGAGGT-3′; reverse, 5′-AGGAATTCGCAGGAGGGATG-3′) with 200 ng of DNA, 2.5 μL of 10X TP Gold (4311806, Thermo Fisher Scientific), 1.5 μL of MgCl2 25mM, 2 μL dNTP 2.5mM (U1420, Promega), 1.5 μL forward and reverse primers (10 μM) and 0.12 μL Amplitaq Gold®. PCR was performed using the following program: 95°C for 12 min, 35 cycles of (94°C for 15 s, 57°C for 30 s and 72°C for 2 min) and 72°C for 7 min. PCR products were purified (Nucleofast 96 PCR plate 743100.50, Macherey Nagel) and Sanger sequenced validated using their respective forward and reverse primers.

siRNA

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

Cell infection

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

RNA extraction, cDNA and RT-QPCR

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

Immunoblotting

Cells were trypsinized, counted, washed with ice-cold PBS and lysed in Laemmli buffer (50 mM Tris-HCL, 2.5 mM EDTA, 2.55 mM EGTA, 2% SDS 20%, 5% Glycerol, 1% Bromophenol blue, protease inhibitor cocktail tablets and 2 mM DL-Dithiothreitol solution) at 10 million cells/ml. Protein lysates were sonicated and denatured at 95°C for 5 min and electrophorated on 4-15% Mini-PROTEAN®TGXTM gels (456-1084, BIO-RAD), transferred onto nitrocellulose membranes (1704159, BIO-RAD). Membranes were incubated overnight at 4°C with mouse anti-STAG2 (1:1,000, Santa Cruz Biotechnology, sc-81852), goat anti-STAG1 (1:5,000, ab4457, Abcam), rabbit anti-H3 (1:50,000, ab1791, Abcam), mouse anti-β-Actin (1:20,000, A5316, Sigma-Aldrich), rabbit anti-FLI1 antibody (1:1,000, ab133485, Abcam) and rabbit anti-H3K27ac (1:1,000, ab4729, Abcam). Then membranes were incubated 1h at room temperature with respective anti-rabbit, anti-mouse immunoglobulin G horseradish peroxidase (HRP) coupled secondary antibody (1:3,000, NA934 or NXA931, respectively; GE Healthcare) or anti-goat IgG-HRP (1: 10,000, SC-2354, Santa Cruz). Proteins were visualized using SuperSignal™ West Pico Plus (34580, Thermo Scientific) and ChemiDoc™ Imaging System (BIO-RAD).

RNA-seq

RNA was extracted from independent experiments using RNeasy Plus Mini Kit (Qiagen, ref 74134) according to manufacturer recommendations. RNA sequencing libraries were prepared from 1 μg of total RNA using the TruSeq Stranded mRNA Library preparation kit (Illumina, ref 20020594) following manufacturer instructions. Sequencing was carried out using 2x100 cycles (paired-end reads 100 nucleotides) for all samples on Illumina HiSeq2500 or NovaSeq6000 instruments. Reads were aligned with STAR 2.5.3 (Dobin et al., 2013) to the human genome (GRCh37/hg19 version). We used the count matrix generated by STAR using the human gene annotation v19 of GENCODE. DESEQ2 1.20.0 (Love et al., 2014) was used to normalize data and performed differential analysis with the Wald test. The p-value was adjusted using Benjamini-Hochberg. For differential analysis, we used an adjusted p-value < 0.01 and $|\log2(\text{fold change})| > 1$. We considered a gene expressed if the normalized expression is higher than 10.

Tumors and cell lines expression data

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

ChIP-seq

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

HiChIP

Several adaptations were made and are detailed below in an overview of the otherwise original protocol described by Mumbach and colleagues (Mumbach et al., 2016). Experiments were performed as independent biological replicates. Cells were plated 2 days before starting the experiment to reach a 90-95% confluence at the time of fixation in 1% of methanol-free formaldehyde (28908, Thermo-Scientific) in freshly prepared respective cell culture media (20ml/T150 flask). Cells were fixed at room temperature for 10 minutes under gentle shaking platform (50mvt/min). 2 mL of glycine solution (2M) were added and cells were incubated for 5 additional minutes at room temperature under gentle shaking platform. The supernatant was removed and cells were washed 3 times with PBS at room temperature. Cells were scraped vigorously with ice cold PBS supplemented with protease inhibitor cocktail tablets (11836145001, Roche), and flushed five time through a syringe with a 21 gauge needle (301155, BD Microlance). *In situ* contact libraries were performed starting from 15 million nuclei digested overnight at 37°C with Mbo1 (R0147M, New England Biolabs). After proximity ligation (4 h at room temperature), the nuclear pellet was sonicated and the chromatin immunoprecipitation step was performed using the iDeal ChIP-seq kit for transcription factors (Diagenode) according to the supplier's recommendation with some modification. Nuclei were resuspended in IL1b and IL2 buffer following Diagenode protocol and all centrifugations were performed at 4°C for 5 min at 1,950 RCF for these steps. The chromatin was sonicated (5 million nuclei per tube) with Bioruptor pico (Diagenode) for 10 cycles (30-sec on, 30-sec off) set at position "high". The tubes were pooled and the chromatin was clarified by centrifugation at 4°C for 10 min at 16,000 RCF. Sonicated chromatin from 11 and 3 million nuclei were used respectively for CTCF and H3K27ac immunoprecipitation step using the equivalent of 3 ChIP reactions pooled in one tube for each HiChIP reaction (final volume was 1050 µl/tube). Immunoprecipitation was then carried out following Diagenode kit instructions by multiplying all reagents by a factor 3 until end of elution step (50µl).

Biotin capture was performed as previously described (Mumbach et al., 2016) and library for illumina sequencing were prepared using 10 ng of chromatin and 0.5 µL of TN5 (15028211, Illumina). A first PCR in a final volume of 50µl with 5 cycles was performed (72°C for 5 min, 98°C for 1 min, then 5 cycles at 98°C for 15 s, 63°C for 30 s, and 72°C for 1 min). To determine how many additional PCR cycles were required for optimal library preparation, a QPCR with: 5µL of the first PCR product, 1 µL of Nextera Ad1_noMX and Nextera Ad2.X (each at 1.25 µM), 5 µl of phusion HF 2X (M0531S, New England Biolabs), 0,75 µL Evagreen® 20X (31000, Biotium) and 2,25 µL of water was runned (same program as first PCR but with 30 cycles). The optimal number of additional cycle was determined for each library by setting a threshold just before reaching the end of exponential amplification step. The 45µl left from the first PCR were further amplified using additional PCR cycles as determined above. Size selection was performed using Ampure XP beads (A63881, Beckman Coulter) to capture fragments greater than 300 bp. Libraries were quantified and analyzed using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and LabChIP (Perkin Elmer). A first validation of HiChIP experiments was performed using 150 bp paired-end sequencing on MiSeq-microV2-300-PE150 (Illumina). Deep 75 bp paired-end sequencing was then performed on HiSeq 2500 or NovaSeq 6000 systems (Illumina).

HiC-Pro 2.10.1 (Servant et al., 2015) pipeline was used to map reads against the human genome (GRCh37/hg19 version), assign reads to restriction fragments, remove duplicate pairs, filter out invalid interaction products and generate a 5kb interaction matrix of valid pairs (read pairs mapping on two different restriction fragments, see Servant et al., 2015 for more details). All HiChIP data were ultimately normalized to the number of valid cis-interactions for subsequent analyses. HiC-Pro was also used to generate ".hic" file for data visualization in Juicebox software (Durand et al., 2016; Rao et al., 2014).

Please cite this article in press as: Surdez et al., *STAG2* mutations alter CTCF-anchored loop extrusion, reduce *cis*-regulatory interactions and EWSR1-FLI1 activity in Ewing sarcoma, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.04.001

Loop prediction by Tweed

To detect loops from CTCF HiChIP data, we developed a new algorithm called Tweed (Figure S5). Loops can be defined by the presence of CTCF sites at both ends, by increased signal as compared to background and by the presence of a corner peak at the intersection of a X and a Y stripe. First, we identified all CTCF peaks using the MACS algorithm on ChIP-seq data. All candidate stripes starting on a CTCF site and ending on all possible CTCF sites within a limit of 2.5 Mb were then considered. We excluded from the analysis a 75 kb (15 bins) region close to the diagonal where the signal is difficult to interpret. We then used two different approaches. The first one was based on the expected increased signal of stripes (enrichment method). This method compared the total intensity of each candidate stripe with that of a control region defined by the 10 flanking parallel lines of identical size (the control region was defined upstream or downstream the candidate X and Y stripes, respectively). The signal intensity of a stripe was then normalized along its entire length. For each control regions, we also calculated the median value of the normalized intensity (10 lines). For each stripe, we defined a control value corresponding to the median values of all control regions of equivalent size. This last step allowed to avoid local bias and not to be too stringent in the definition of stripes. The intensity of each candidate stripe of a given length was then compared with this value. The intersection of X and Y stripes was computed based on their coordinates. A loop was defined when both X and Y values were > 2 or when either of the values was > 3.

The second approach (second derivative method) was based on the detection of decreased interaction intensity at the end of the stripe (frequently defined as the corner peak). This inflection point can be detected by a negative value of the second derivative of the function defined by the cumulative curve of interaction counts along the stripe. If the value was negative on the corner, or lower than -10 in the +/- 1 adjacent bin to the corner, a candidate loop was predicted. Only loops predicted by both methods and with mean intensity higher than 2 reads/bin were kept. For adjacent loops (i.e. differing only by 1 bin at their CTCF boundary sites), the loop with the highest coverage intensity was kept. At this point, most predicted loops presented at their boundaries a strong prevalence of cohesin members and convergent CTCF motif orientations as determined with the FIMO tool from MEME suite (Grant et al., 2011) using the CTCF motif (MA0139.1 from Jaspar database) (Figures S5E–S5G). Final filtering for presence of convergent CTCF motif orientations and of at least one member of the cohesin complex at each CTCF boundary sites was therefore used for final loop selection. At the end, 18,774 (A673WT) and 14,263 (TC71WT) loops were detected by our Tweed algorithm. Final loops were classified in two categories: balanced intensity of X and Y stripes or asymmetric stripes with predominant X or Y pattern (defined by a log2 X/Y stripe coverage ratio higher or lower than the mean value of all stripes intensity +/- 2SD, respectively). To compare loop coverage for each loop detected in *STAG2* WT conditions (Figure 4D), we computed the stripe coverage and normalized by the total number of cis-interactions as well as by the loop size. We then assessed the difference between *STAG1/2* -WT and -KO conditions (paired T-test on log10(normalised_coverage)).

Interaction chains

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

Soft agar colony formation assay

A first agar layer was placed in 6 well cell culture plates at 0.8% (w/v) of low melting point agarose (16520; Invitrogen) in respective cell culture media. Once solidified, a second layer of 0.4% agar was added, containing respectively 1000 or 2000 cells for TC71 and A673 lines. The plates were maintained at 4°C for 5 minutes and 1 mL of fresh medium was subsequently deposited as a top layer. The plates were incubated at 37°C at 5% CO2 and colonies (>50 μm diameter) were counted 2 weeks post seeding using the FIJI Particle Analysis tool.

Actin cytoskeleton and focal adhesion

A673 (2.5X104 cells) and A673SA2m#1 (2X104 cells) were plated in 24-well plate on poly-D-Lysine coated coverslips (Corning, #354086). After 24 hours, the cells were fixed with 4% paraformaldehyde, permeabilized in 0.2% Triton X-100/1XPBS and blocked in 1% BSA/1XPBS solution. Cells were then incubated with phalloidin-tetramethylrhodamine B isothiocyanate (1:100, Sigma, #P1951) and paxillin antibody (1:1000, BD biosciences, #610051) diluted in 0.5% BSA/1X PBS for 1 hour, washed with 0.5%

BSA/1X PBS and incubated with Alexa Fluor 488-conjugated donkey anti-mouse (1:100, Invitrogen, #A21202) for 45 min. Slides were mounted using Prolong Gold Antifade reagent (Thermofisher, #P36930). Images were acquired with an upright widefield Apotome microscope (Zeiss) equipped with a Coolsnap HQ2 camera through a x63 NA 1.4 oil-immersion objective lens. Paxillin focal adhesion foci associated with phalloidin stress fibers were counted in individual cells for each condition.

Wound healing migration assay

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

Spheroid invasion assay

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

Image analysis

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

Velocity measurements:

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

Boyden chambers assay

A673 and CHLA-258 cells were transfected 48hrs prior the assay. The following numbers of trypsinized cells were seeded in the upper compartment of a Boyden chamber (353182, BD Biosciences): 5X104 A673, 5X104 A673SA2m#1, 5x104 A673SA2r cells and 7.5X104 A673 or CHLA-258 transfected cells (siCT, siSA#6 or siSA#8) in respectively their respective media containing 0.5% FBS. The lower well chamber was in contact with respective media containing 10% FBS. After 24 hours (siRNA) or 48 hours (*STAG2* isogenic models), the remaining cells present on the upper membrane of the Boyden chamber were carefully wiped out with a cotton swab and the chamber was then stained with crystal violet solution. Cell count was performed from 10 non-overlapping pictures acquired with a microscope (10X objective lens).

GSEA and DoRothEA analyses

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

QUANTIFICATION AND STATISTICAL ANALYSIS

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