www.sciencemag.org/content/361/6405/eaam8419/suppl/DC1

Supplementary Materials for

Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors

Nathaniel D. Anderson, Richard de Borja*, Matthew D. Young*, Fabio Fuligni*, Andrej Rosic, Nicola D. Roberts, Simon Hajjar, Mehdi Layeghifard, Ana Novokmet, Paul E. Kowalski, Matthew Anaka, Scott Davidson, Mehdi Zarrei, Badr Id Said, L. Christine Schreiner, Remi Marchand, Joseph Sitter, Nalan Gokgoz, Ledia Brunga, Garrett T. Graham, Anthony Fullam, Nischalan Pillay, Jeffrey A. Toretsky, Akihiko Yoshida, Tatsuhiro Shibata, Markus Metzler, Gino R. Somers, Stephen W. Scherer, Adrienne M. Flanagan, Peter J. Campbell, Joshua D. Schiffman, Mary Shago, Ludmil B. Alexandrov, Jay S. Wunder, Irene L. Andrulis, David Malkin†, Sam Behjati†, Adam Shlien†

*These authors contributed equally to this work.
†Corresponding author. Email: adam.shlien@sickkids.ca (A.S.); sb31@sanger.ac.uk (S.B.); david.malkin@sickkids.ca (D.M.)

Published 31 August 2018, *Science* **361**, eaam8419 (2018)
DOI: 10.1126/science.aam8419

This PDF file includes:

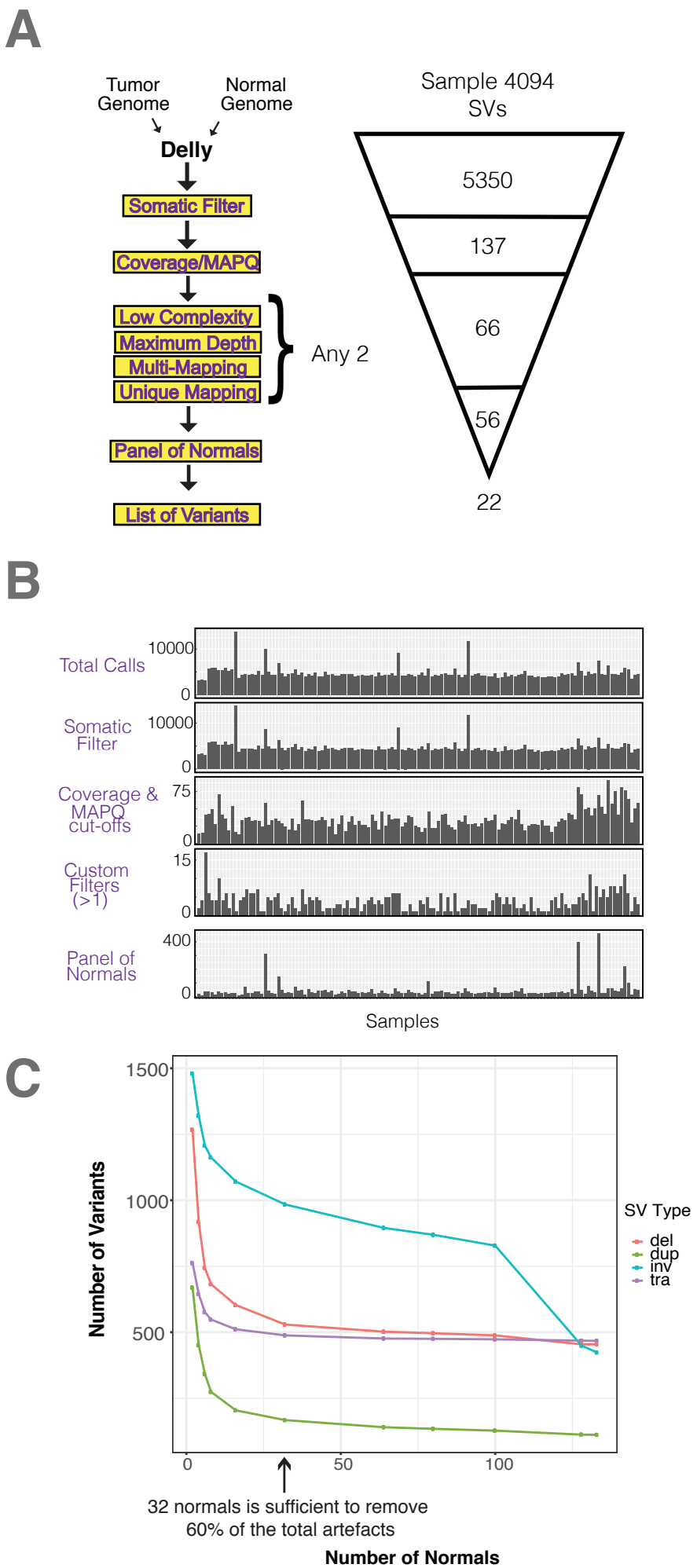
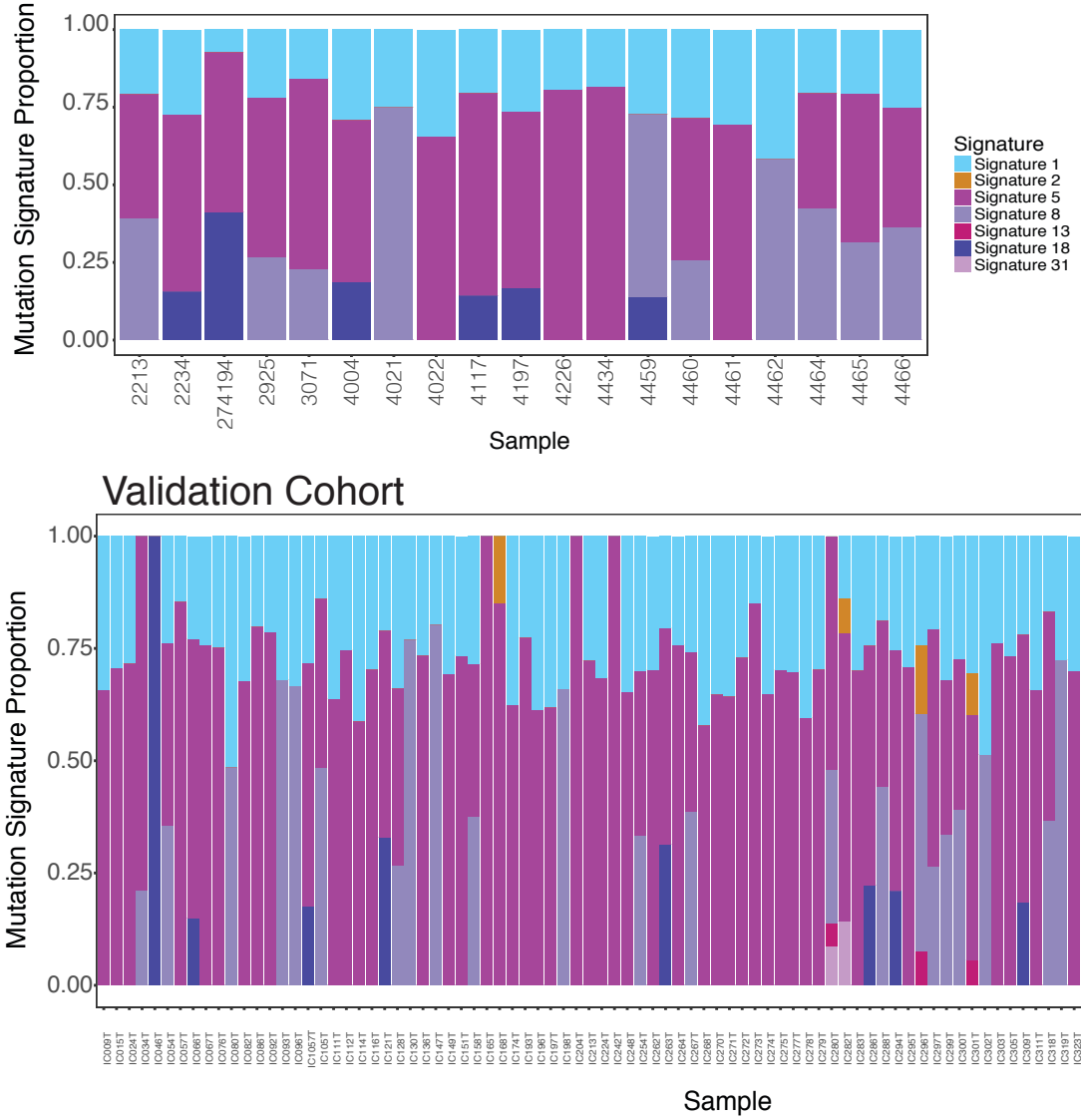

A
Tumor Genome
Normal Genome
Delly
Somatic Filter
Coverage/MAPQ
Low Complexity
Maximum Depth
Multi-Mapping
Unique Mapping
Any 2
Panel of Normals
List of Variants
Sample 4094 SVs
5350
137
66
56
22
B
Total Calls
Somatic Filter
Coverage & MAPQ cut-offs
Custom Filters (>1)
Panel of Normals
10000
0
75
15
400
Samples
C
Number of Variants
1500
1000
500
0
50
100
SV Type
del
dup
inv
tra
32 normals is sufficient to remove 60% of the total artefacts
Number of Normals

Figure S1

Fig. S1. Informatics pipelines to detect somatic substitutions and rearrangements. **(A)** Rearrangement pipeline overview schematic illustrates the use of the detection tool, Delly, and the custom filters applied to its raw output. A representative sample is illustrated showing the number of variants retained after each filtering step. **(B)** First panel shows number of rearrangements in raw output. Subsequent panels show number of variants removed after each step. **(C)** Power of panel-of-normals: dose-response curves for rearrangements. With an increase in the number of normal samples used in your panel, the power to remove artefacts increases.

A


Toronto Cohort
Mutation Signature Proportion
1.00
0.75
0.50
0.25
0.00
Signature
Signature 1
Signature 2
Signature 5
Signature 8
Signature 13
Signature 18
Signature 31
Sample
Validation Cohort
Mutation Signature Proportion
Sample

B

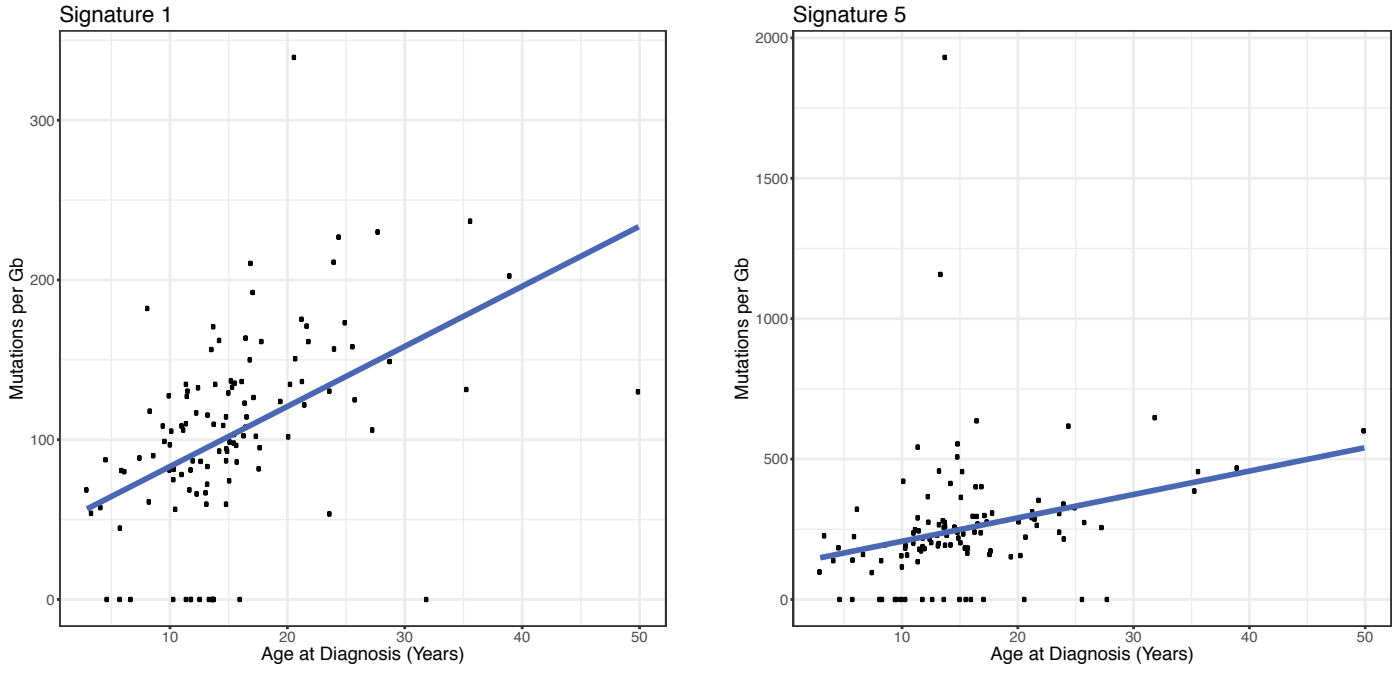
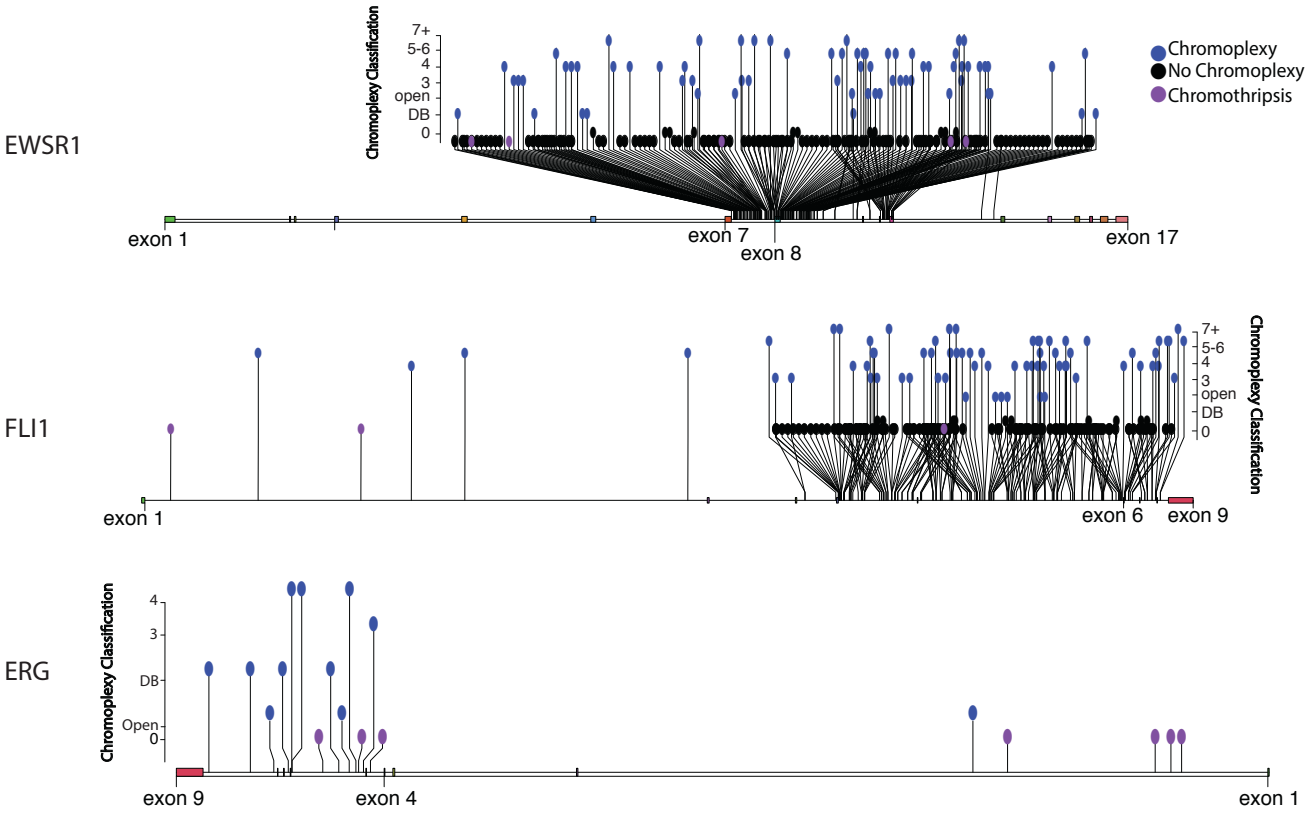

Signature 1
Mutations per Gb
Age at Diagnosis (Years)
Signature 5
Mutations per Gb
Age at Diagnosis (Years)

Figure S2

Fig S2. Mutation Signatures in Ewing sarcoma. (A) Stacked bar charts illustrate the sample-by-sample breakdown of the proportion of mutation signatures in each tumor for the Toronto and Validation cohort. Signatures for diagnostic tumors with a matched relapsed or metastatic tumor (n=4) can be seen in Fig 5A. COSMIC Signatures 1 and 5 are the most common (purple and sky blue). **(B)** Clock-like mutation signatures (COSMIC 1 and 5) have been found to be associated with the age of diagnosis in other cancers. Scatter plot illustrates positive correlation between mutation burden and age at diagnosis in ES patients for COSMIC signatures 1 and 5.

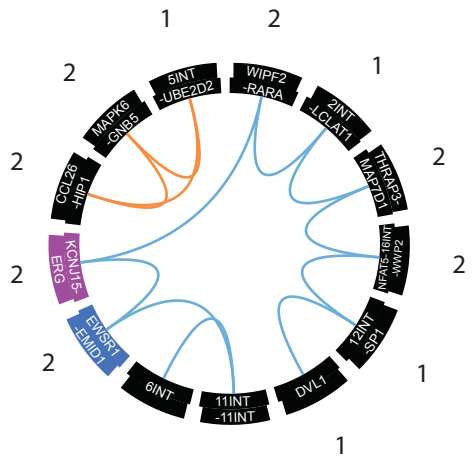
A



B

	FLI1	ERG
Age at Diagnosis (months)	199.67	155.52
Sex (M:F)	19:16	7:4
Proportion of Genic Hits	56.20%	57.88%
Average # of Rearrangements	13	18

C



Sample 4004
18 genes disrupted

D

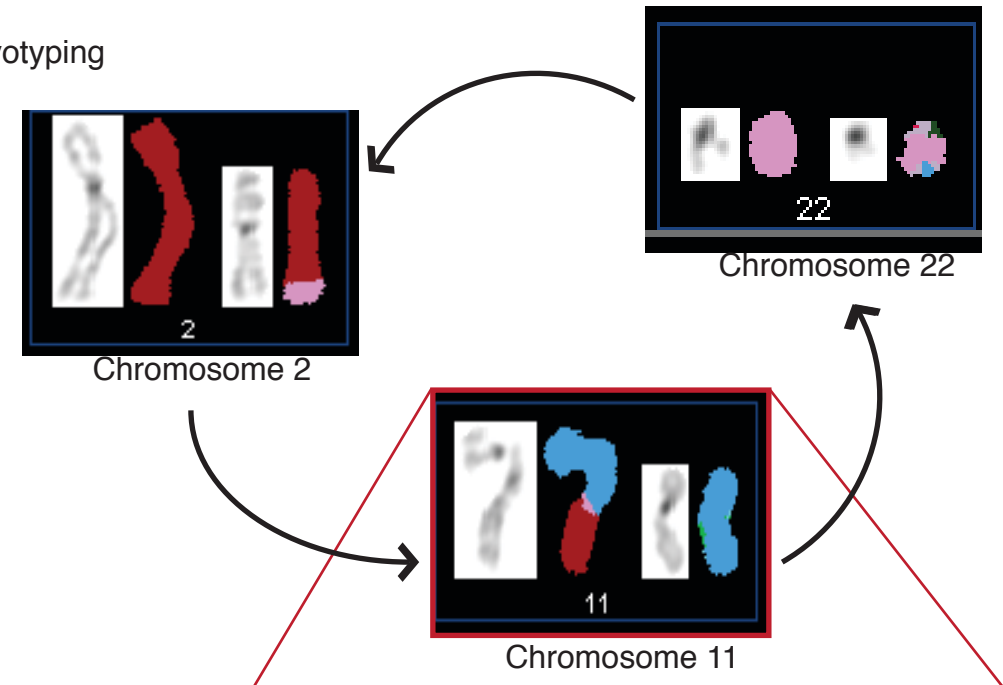
EWS ↓ ETS →

RINL	MAP4K5	ATP6V1C2	CCDC117	TTC28	RARA	MYO1B	RRM2	OSM	HIP1
SEC14L2	AASDHPPT	GACAT3	RASL10A	PPP4R1L	KCNJ15	DEPDC5	ROCK2	**AP1B1**	MAP7D1
ERH	THOC5	CYS1	PHF21B	PCM1	EYS	TMEM52	DDX1	**ETS1**	SP1
TMC2	MEGF11	GREB1	WDR63	CRLF3	MACROD2	CDKAL1	LOC101926966	USP37	LCLAT1
EXT2	ERCC1	ZNRF3	NOS1AP	MAN1C1	C11orf49	NARS2	CAMSAP1	COL11A1	NFAT5
PTPRK	RSF1	LO101927356	NME7	IFI16	MCHR1	SGK1	TRAF2	LRRK1	WIPF2
WDR7	**GAS2L1**	GPC5	EML6	ATF6	SLC25A17	LINC01505	FAM129A	NUP210L	CAMTA1
EMID1	MAPK1P1L	GMDS	SLC25A25-AS1	ETV3L	ARID1B	ZNF462	COL4A4	CUBN	LOC102477328
KREMEN1	CDH11	MN1	CALN1	CCL26	SBF2	YAP1	EGFR	NFATc2	OPCML
NF2	LOXL3	CHEK2	SLC5A1	MAPK6	CXCR5	FAM167a-AS1	PGAM1	THRAP3	PDE7B
NFS1	C2orf48	MIAT	HPSE2	GNB5	HNRNPL	ZZEF1	TMEM135	DVL1	ZNF143
EP400	HPCAL1	MIATNB	LINC01426	WWP2	FAM118B	PTCD3	ADAT1	UBE2D2	SCUBE3

Figure S3

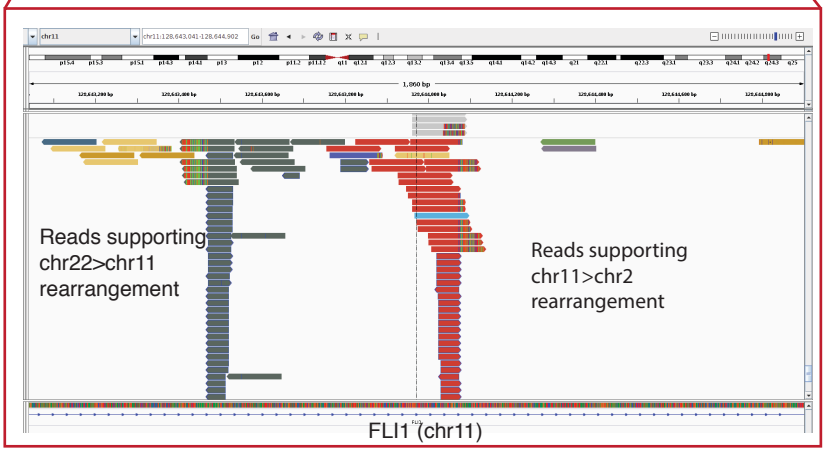
Fig. S3. Characterizing chromoplectic breakpoints in Ewing sarcoma. (A) Breakpoint distribution for the *EWSR1*, *FLI1* and *ERG* genes are shown. We observed no difference in intronic breaks for chromoplectic and non-chromoplectic *EWSR1* breaks. Non-canonical *FLI1* breaks (before exon 3) are exclusively chromoplectic. All ERG-rearranged Ewing sarcoma tumors exhibited patterns of chromoplexy. The color and height of lolliplot balls refer to complex rearrangements and their classification (i.e. number of rearrangements in a loop, open loops or the existence of deletion bridges (DB)). **(B)** Table overview comparing chromoplexy in EWS-FLI1 and EWS-ERG driven tumors. **(C)** Circos web illustrating the Ewing sarcoma sample with 2 looped rearrangement clusters affecting 18 genes (n=18). **(D)** Flow diagram of all the unique genes (n=120) affected by chromoplexy in Ewing sarcoma. Established cancer-associated genes are highlighted (red) and recurrently affected genes are bolded.

A Spectral Karyotyping



B

Custom-Capture Sequencing



C

Sample	Nextera	Matched Genome	Chromoplexy	Fusion	Other Rearrangements
2213	Yes	Yes	Yes	EWS-FLI1	EWS-SCUBE3, SCUBE3-FLI1
2226	Yes	Yes	Yes	EWS-FLI1	FLI1-CXCR5
2234	Yes	Yes	Yes	EWS-FLI1	EWSR1-chr8, FLI1-chr6, chr6 deletion
2925	Yes	Yes	Yes	EWS-FLI1	EWS-chr5, chr5-FLI1, FLI1-chr11
4117	Yes	Yes	Yes	EWS-FLI1	EWS-chr7, FLI1-CUBN
4311	Yes	Yes	Yes	EWS-FLI1	EWS-UNC5D, UNC5D-FLI1
PD9684 (3071)	Yes	Yes	Yes	EWS-ERG	EWS-CAMTA1
PD9668 (2498)	Yes	No	No	EWS-FLI1	NA
PD9675 (2301)	Yes	No	Yes	EWS-FLI1	EWS-NCOR2, FLI1-NCOR2
PD9679 (2331)	Yes	No	No	EWS-FLI1	NA
PD9681 (2134)	Yes	No	Yes	EWS-FLI1	chr22-FLI1
PD9682 (2187)	Yes	No	Yes	EWS-FLI1	EWS-chr2, FLI1-chr2
PD9691 (3935)	Yes	No	Yes	EWS-ERG	EWS-GRAMD4
PD9692 (3938)	Yes	No	No	EWS-FLI1	NA

Figure S4

Fig. S4. Validating chromoplexy using cytogenetics and targeted sequencing. (A) Spectral Karyotyping (SKY, multi-colored fluorescent *in situ* hybridization) identifies 24 human chromosomes in a single hybridization and stains each one for a different color. Derivative chromosomes for a three-way rearrangement between chromosomes 22, 11 and 2 are depicted. **(B)** Visual flow diagram of a three-way rearrangement between chromosomes 22, 11 and 2 is shown on the left. Nextera custom capture sequencing was applied to this (and many other) samples to confirm complex rearrangements detected by whole-genome sequencing. We achieved coverage of target regions between 900-1000X. An IGV (integrated genomics viewer) screenshot of FLI1 shows reads translocated to chromosome 22 (green) and chromosome 2 (orange) (right). **(C)** Validation table provides an overview of all samples that have undergone Nextera custom capture sequencing and other platforms that have also performed on these samples. Many rearrangements not included on this panel were also detected, proving chromoplexy occurs in *cis* (on the same strand, otherwise they would not be sequenced).

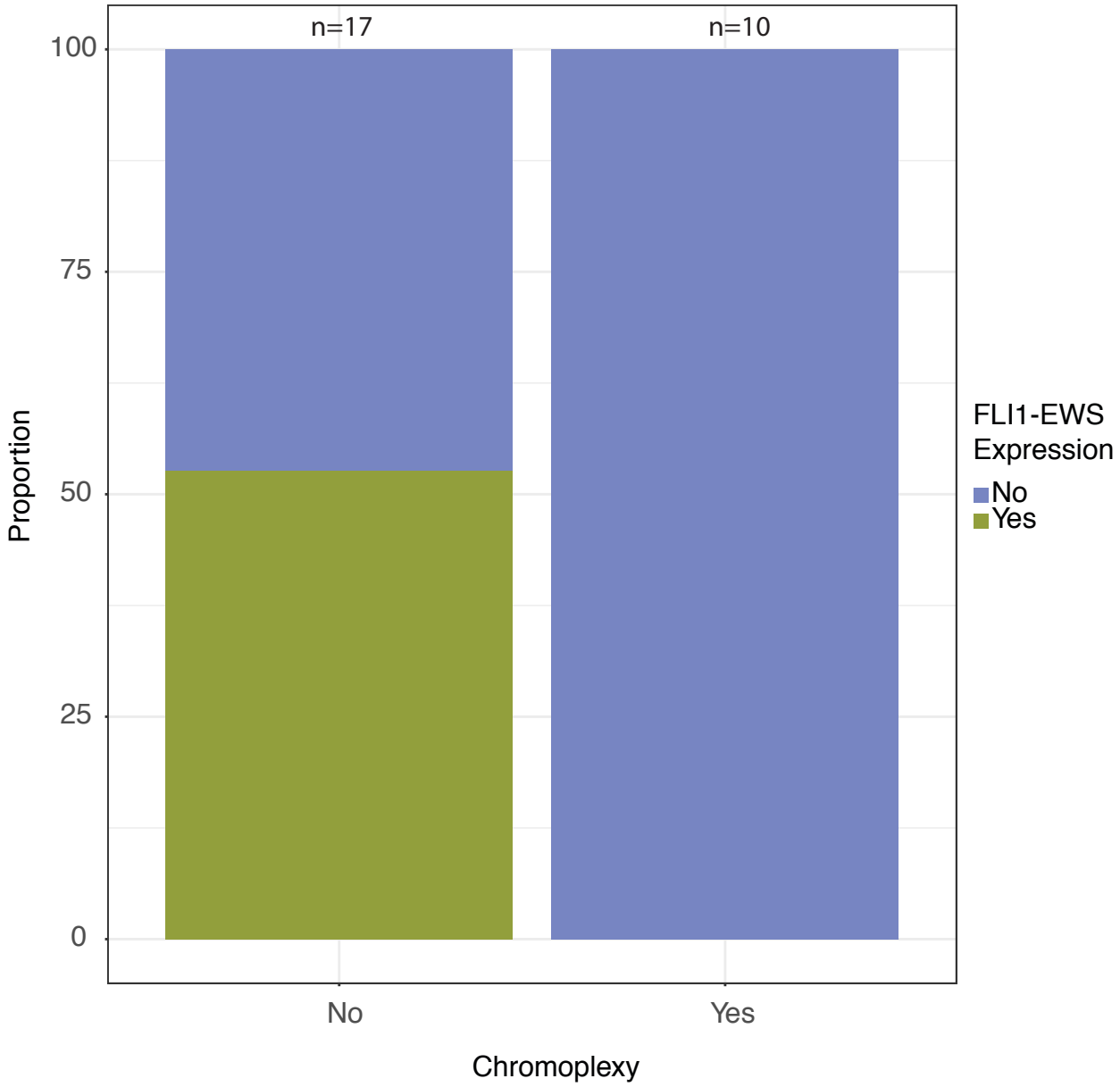


Fig. S5. *FLI1-EWSR1* expression in Ewing sarcoma primary tumors from RNA-seq. In the genome, the presence of chromoplexy and the reciprocal fusion were mutually exclusive. As shown here, in the transcriptome, when chromoplexy occurs, the *FLI1-EWSR1* reciprocal fusion is not expressed. Conversely, when chromoplexy does not occur, the *FLI1-EWSR1* fusion is expressed 52% of the time.

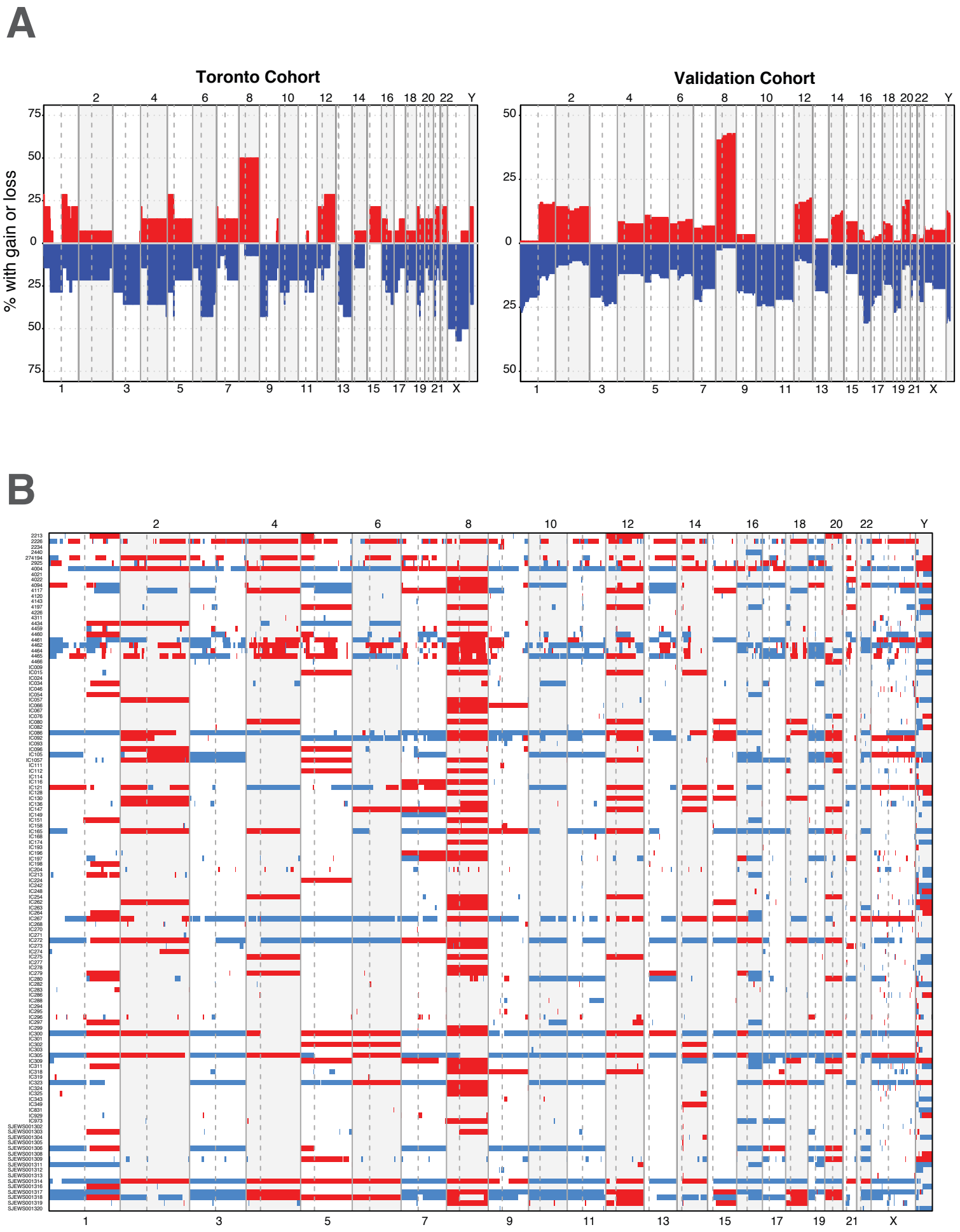


Figure S6

Figure S6. Global Copy-number alterations in ES. (A) Copy-number profiles for Toronto (left) and Tirode, Surdez *et al.* (right) demonstrate common copy number alterations in both cohorts**. (B)** Copy-number heatmap provides sample-by-sample breakdown of arm-length and focal copy-number alterations (CNAs) in Ewing sarcoma (ES) using a low threshold cut-off (+/- 0.2). For a full catalogue of CNAs in ES, see Tirode et al 2014 *Cancer Discovery*.

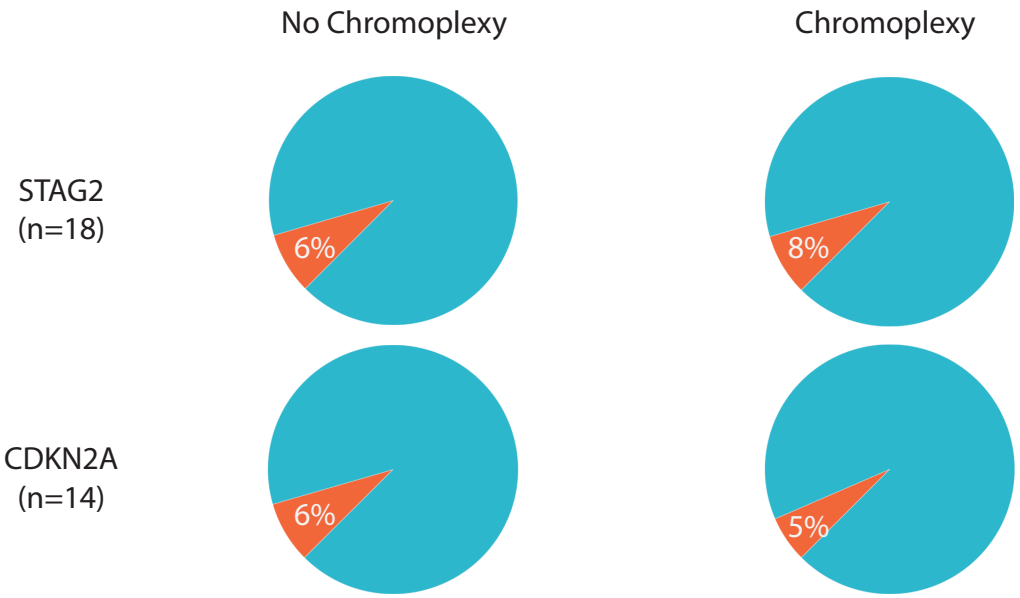


Fig. S7. Distribution of *STAG2* and *CDKN2A* mutations between simple and chromoplexy ES. Pie charts illustrate no difference in the prevalence of somatic *STAG2* mutations or *CDKN2A* deletions in Ewing sarcomas with and without chromoplexy.

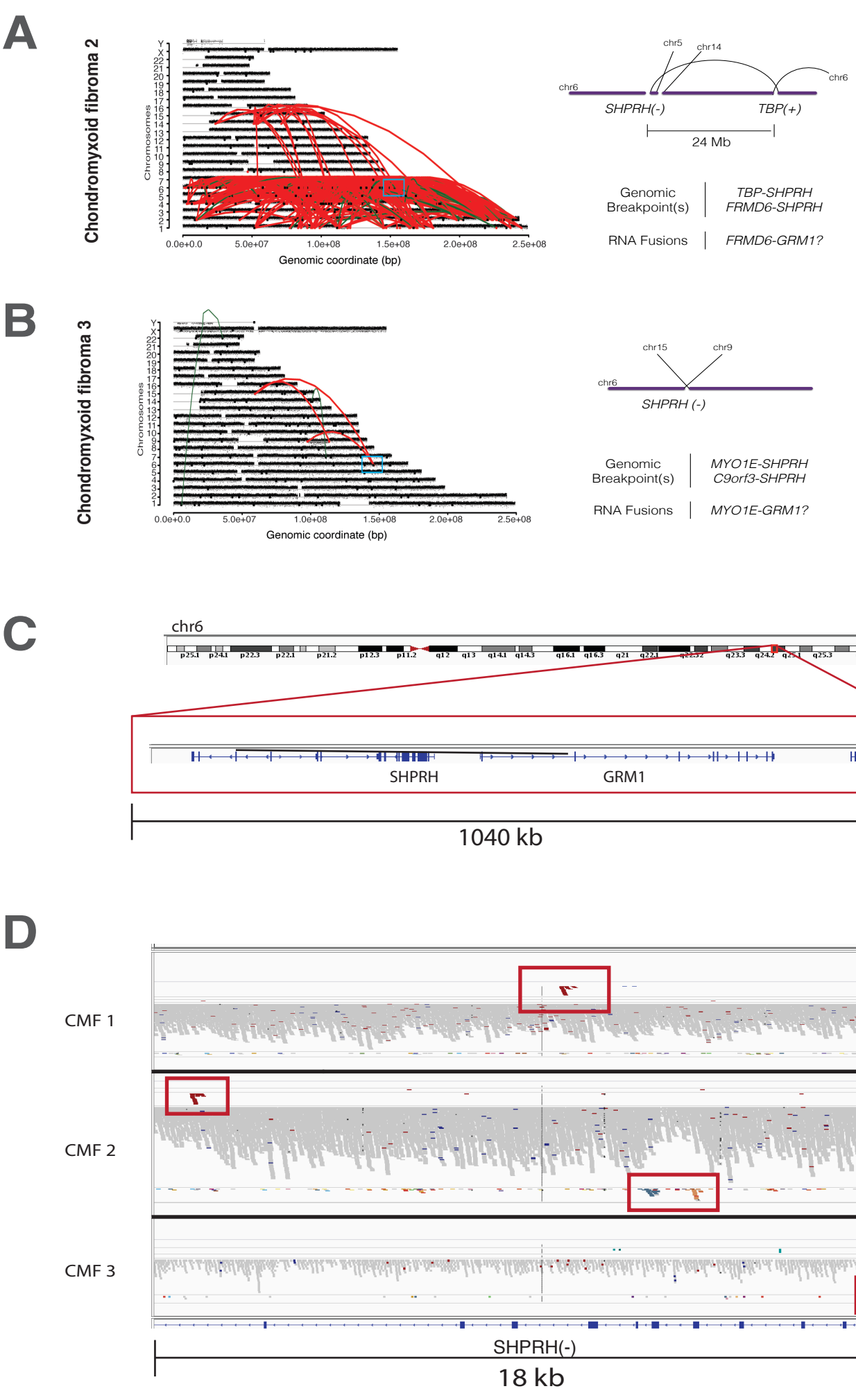

Genomic Complexity
Canonical Fusion
A
Chondromyxoid fibroma 2
Chromosomes
Genomic coordinate (bp)
0.0e+0.0
5.0e+07
1.0e+08
1.5e+08
2.0e+08
2.5e+08
chr5
chr14
chr6
chr6
SHPRH(-)
TBP(+)
24 Mb
Genomic Breakpoint(s)
TBP-SHPRH
FRMD6-SHPRH
RNA Fusions
FRMD6-GRM1?
B
Chondromyxoid fibroma 3
Chromosomes
Genomic coordinate (bp)
chr15
chr9
chr6
SHPRH (-)
Genomic Breakpoint(s)
MYO1E-SHPRH
C9orf3-SHPRH
RNA Fusions
MYO1E-GRM1?
C
chr6
SHPRH
GRM1
1040 kb
D
CMF 1
CMF 2
CMF 3
SHPRH(-)
18 kb

Figure S8

Fig. S8. Genomic Complexity of CMFs. (A) Chondromyxoid fibroma with chromothripsis. Over 500 rearrangements can be detected by genome sequencing in this sample, with three *SHPRH* breakpoints. A *FRMD6-GRM1* fusion is predicted based on the genomic breakpoint (exon 1) of *FRMD6,* as indicated by the question mark (?). **(B)** Chrondromyxoid fibroma with chromoplexy. A simple three-way rearrangement loop involving a deletion bridge on chromosome 9 was observed, predicted to generate the *MYO1E-GRM1* fusion. **(C)** Genomic rearrangements generating the canonical fusion in this cancer type do not directly affect the *GRM1* gene, but rather the *SHPRH* gene, which is the immediate 5' gene (directly upstream), as depicted in this figure. **(D)** In our three samples with genomic complexity generating *GRM1* fusions, the breakpoints appear to cluster within an 18-kb region within the *SHPRH* gene.

Enlarged Figure 3A for readability

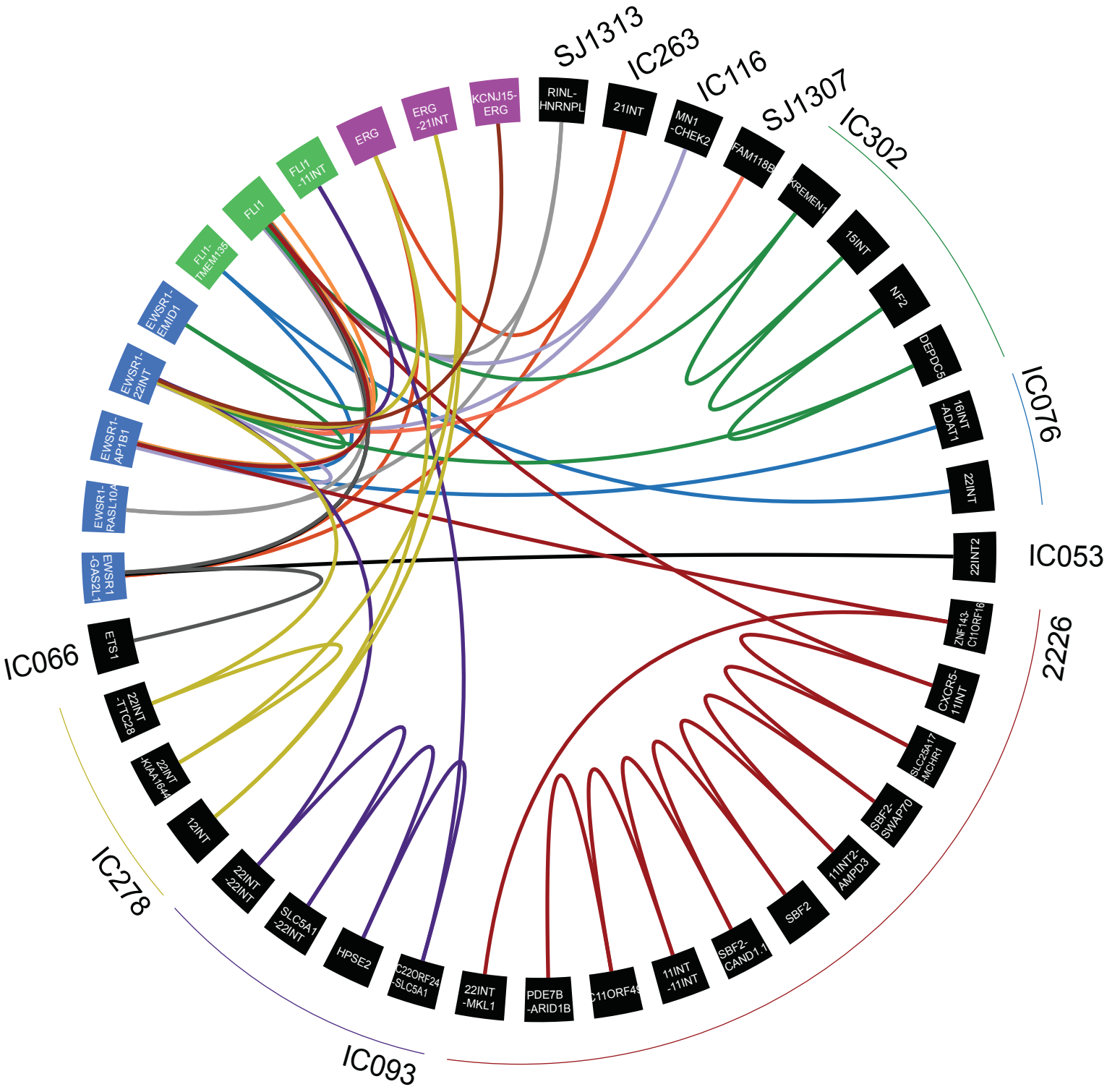
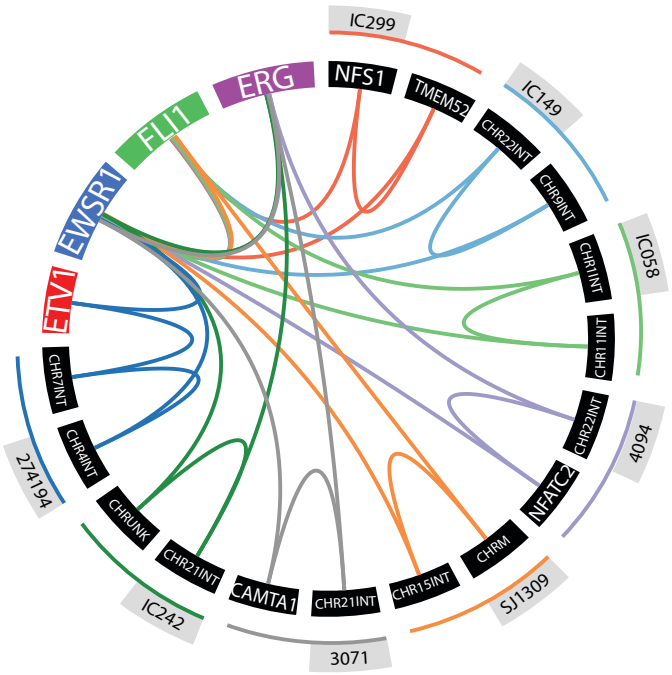
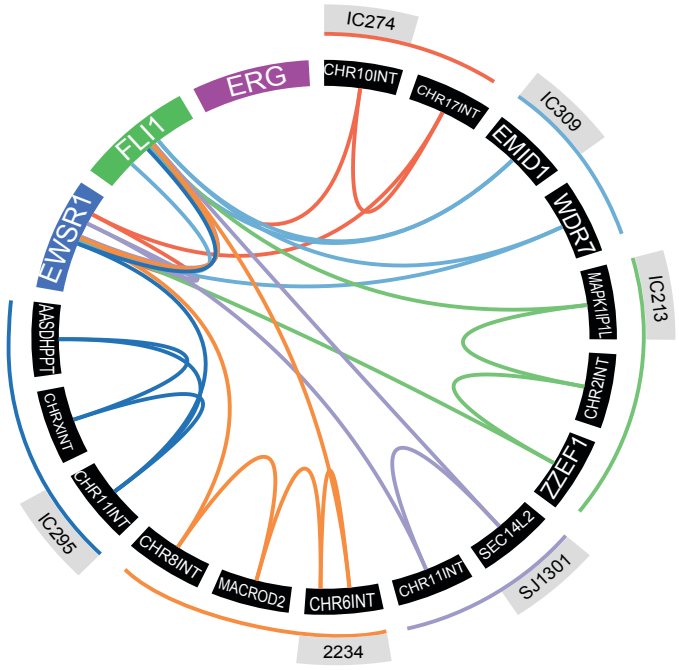


Fig. S9. Enlarged Figure 3A for Readability. Circos web is enlarged “deletion-bridge” (DB) chromoplexy in order to identify genes/loci disrupted by these rearranegments. DB chromoplexy occurs in a loop-structure, but the breakpoints are linked due to the intervening loss of DNA (illustrated in fig.S11)

Four-part (n=8)



Five or six-part (n=6)



>Seven-part (n=5)

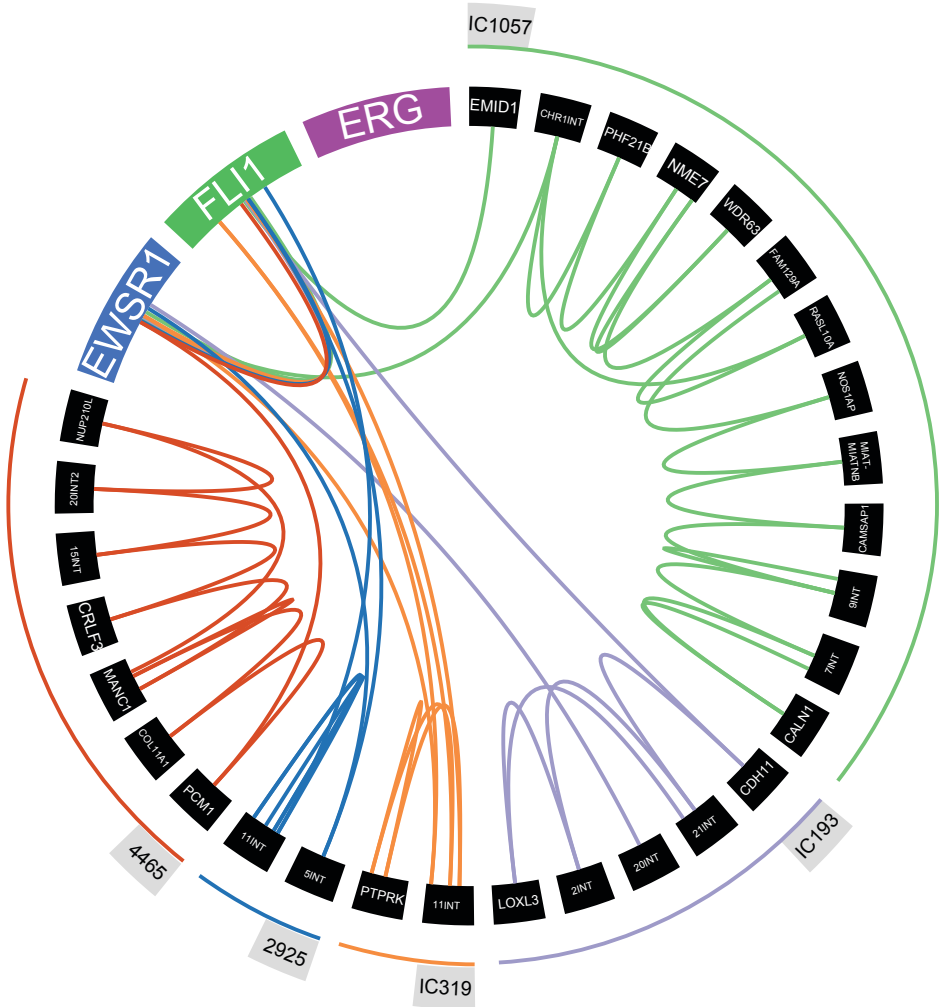
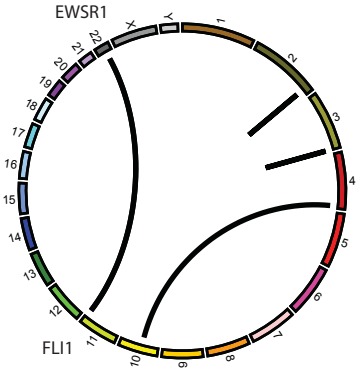


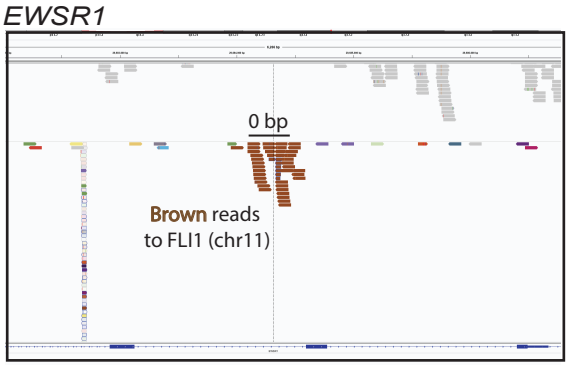
Figure S10

Figure S10. Consequences of Chromoplexy. (A) Circos webs illustrate the chromoplexy fusion loop patterns in ES, from four-part to greater than seven-part. *EWSR1*, *FLI1* and *ERG* are central to all loops. Other complex cases include open and non-EWS loops (n=5, each). One exceptionally rare case of an *EWSR1-ETV1* fusion (red, four-part loop) was detected at the core of a chromoplexy fusion loop.

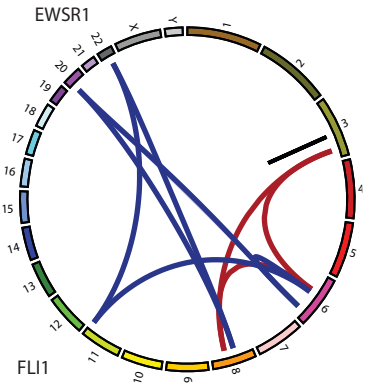
A

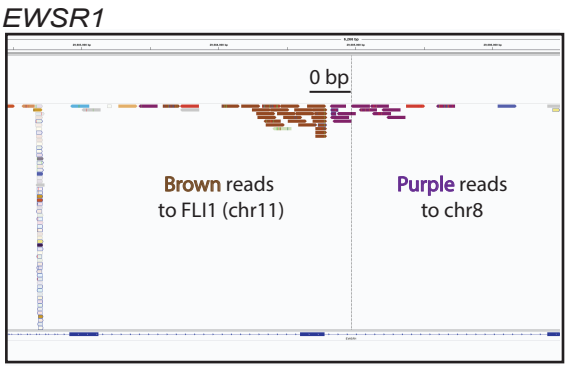
Simple Reciprocal



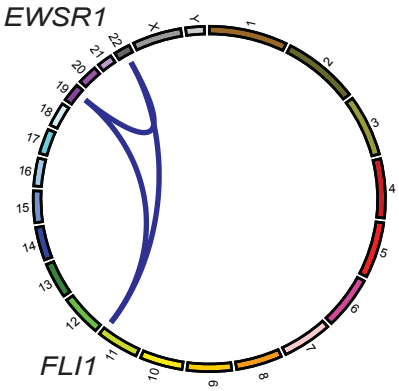


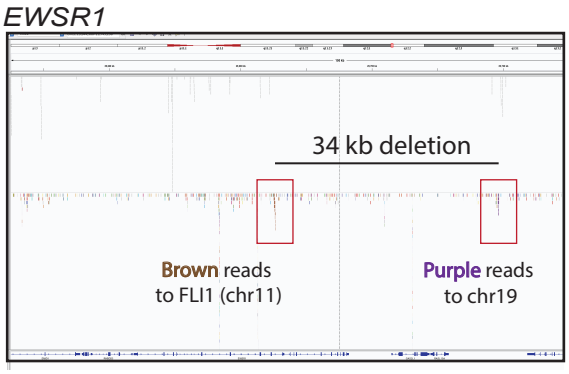
Chromoplexy





DB Chromoplexy





B

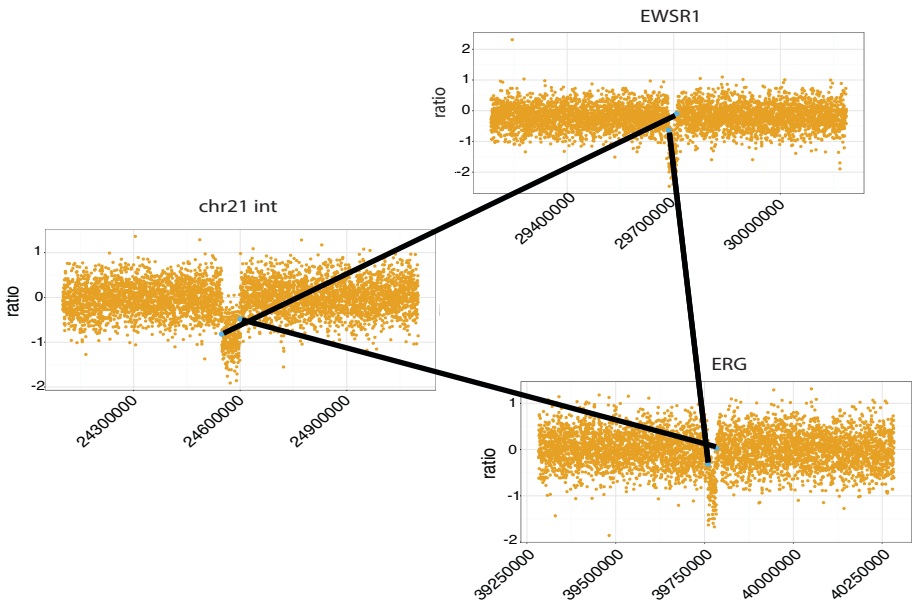


Figure S11

Fig. S11. Deletion Bridges are common in Ewing sarcoma. (A) Representative circos plots and integrated genomics viewer (IGV) screenshots illustrate how breakpoint distances are calculated. Simple reciprocal translocations (top) have no complexity and the breakpoints are right next to each other (>100bp apart). Chromoplectic rearrangements (middle) look similar to reciprocal translocations with respect to breakpoint distances, however the forward and reverse reads go to different chromosomes (brown reads to chr11 vs purple reads to chr8). Lastly, deletion-bridge (DB) chromoplexy (bottom) are further apart as a DNA deletion spans the two breakpoints with the forward and reverse reads going to different chromosomes (characteristic of chromoplexy). **(B)** Example of consistent deletion bridges across all chromoplectic loci. Loss of genetic material at chromoplectic loci joins adjacent breakpoints.

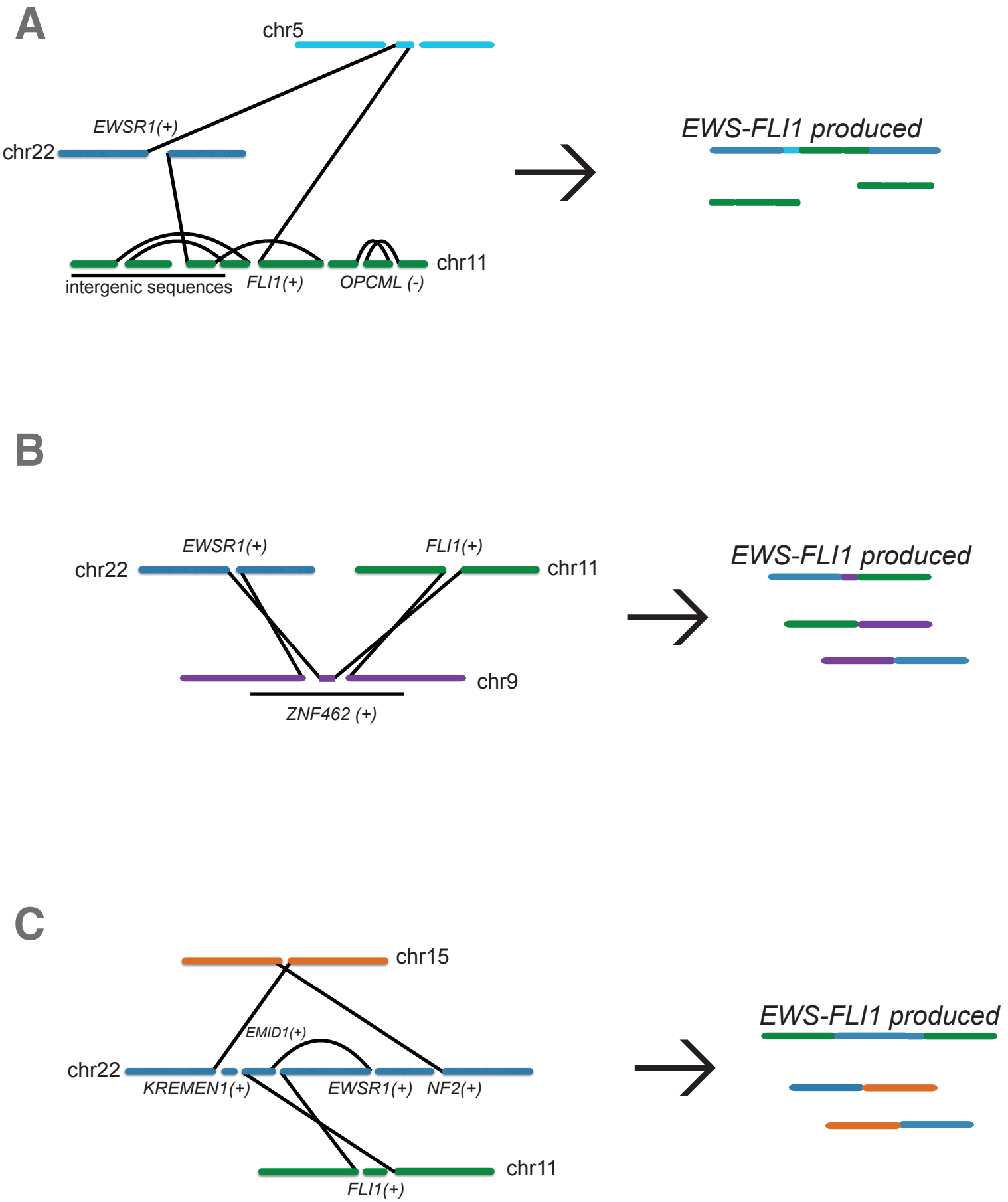

A
chr5
EWSR1(+)
chr22
chr11
intergenic sequences
FLI1(+)
OPCML (-)
EWS-FLI1 produced
B
EWSR1(+)
FLI1(+)
chr22
chr11
chr9
ZNF462 (+)
EWS-FLI1 produced
C
chr15
EMID1(+)
chr22
KREMEN1(+)
EWSR1(+)
NF2(+)
chr11
FLI1(+)
EWS-FLI1 produced

Figure S12

Fig. S12. Chromoplexy acts as a molecular scaffold to produce the EWS-FLI1 driver. Two rearrangements and a third locus are required to bring together *EWSR1* and *FLI1*. In these cases, no direct *EWSR1-FLI1* translocation can be detected using genome sequencing. However, a diagnostic fusion is detected at the transcriptional level. Most of these scaffolding events would be missed by cytogenetics, as an *EWSR1* break-apart FISH would show that these are *EWSR1*-rearranged tumors and G-band karyotyping would show the classical t(11;22) karyotype. **(A)** The 5' portion of *EWSR1* is translocated to chromosome 5 and similarly, the 3' portion of FLI1 is translocated to the same chromosome and position. A functional *EWSR1-FLI1* driver is produced transcriptionally. **(B)** In a double reciprocal translocation event, the 5' of *EWSR1* and the 3' of FLI1 are both translocated to the *ZNF462* gene on chromosome 9 and an *EWSR1-FLI1* transcript is detected at the RNA level. **(C)** An inversion between two adjacent genes, *EWSR1* and *EMID1,* and an *EMID1-FLI1* translocation can be observed in the genome and a canonical *EWSR1-FLI1* fusion transcript is detected in the RNA.

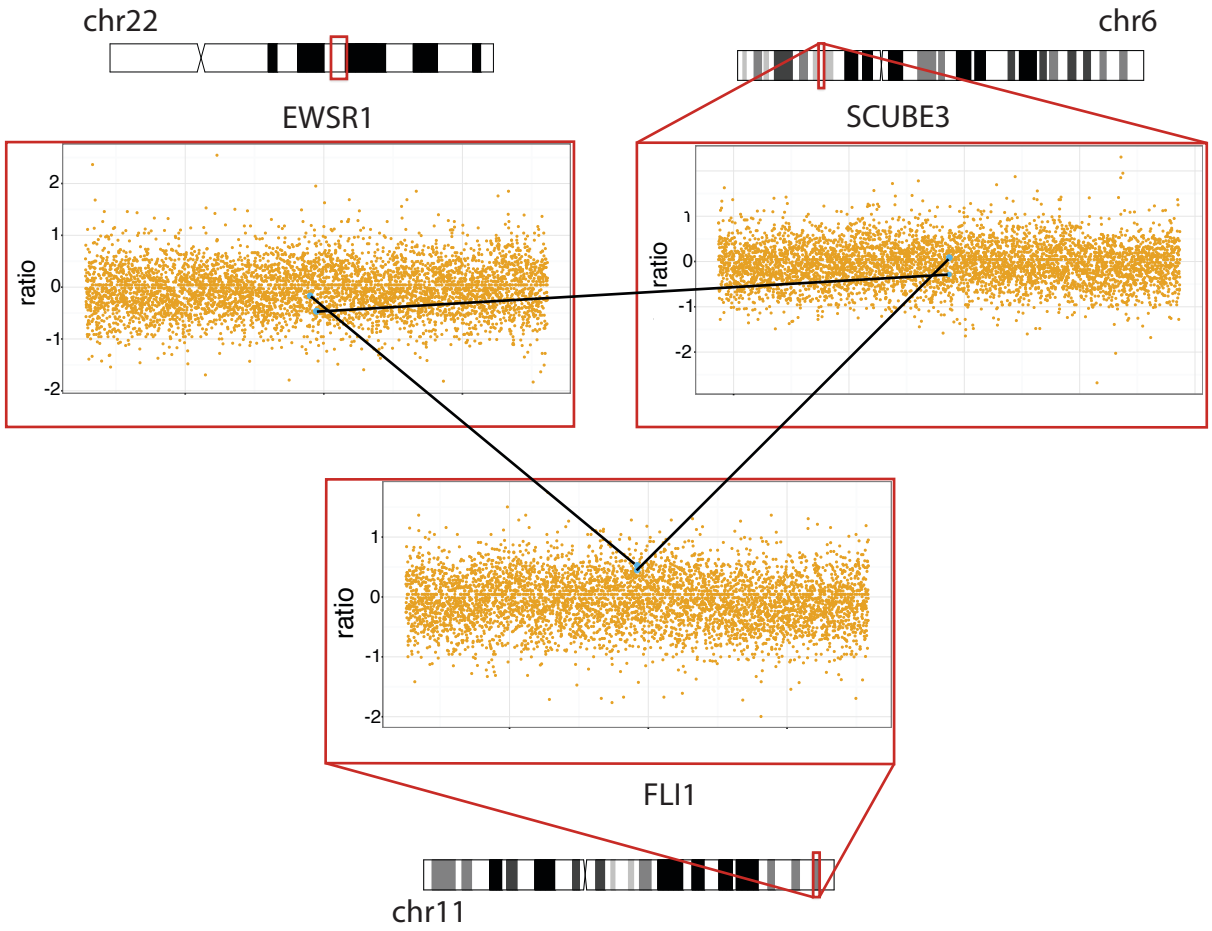
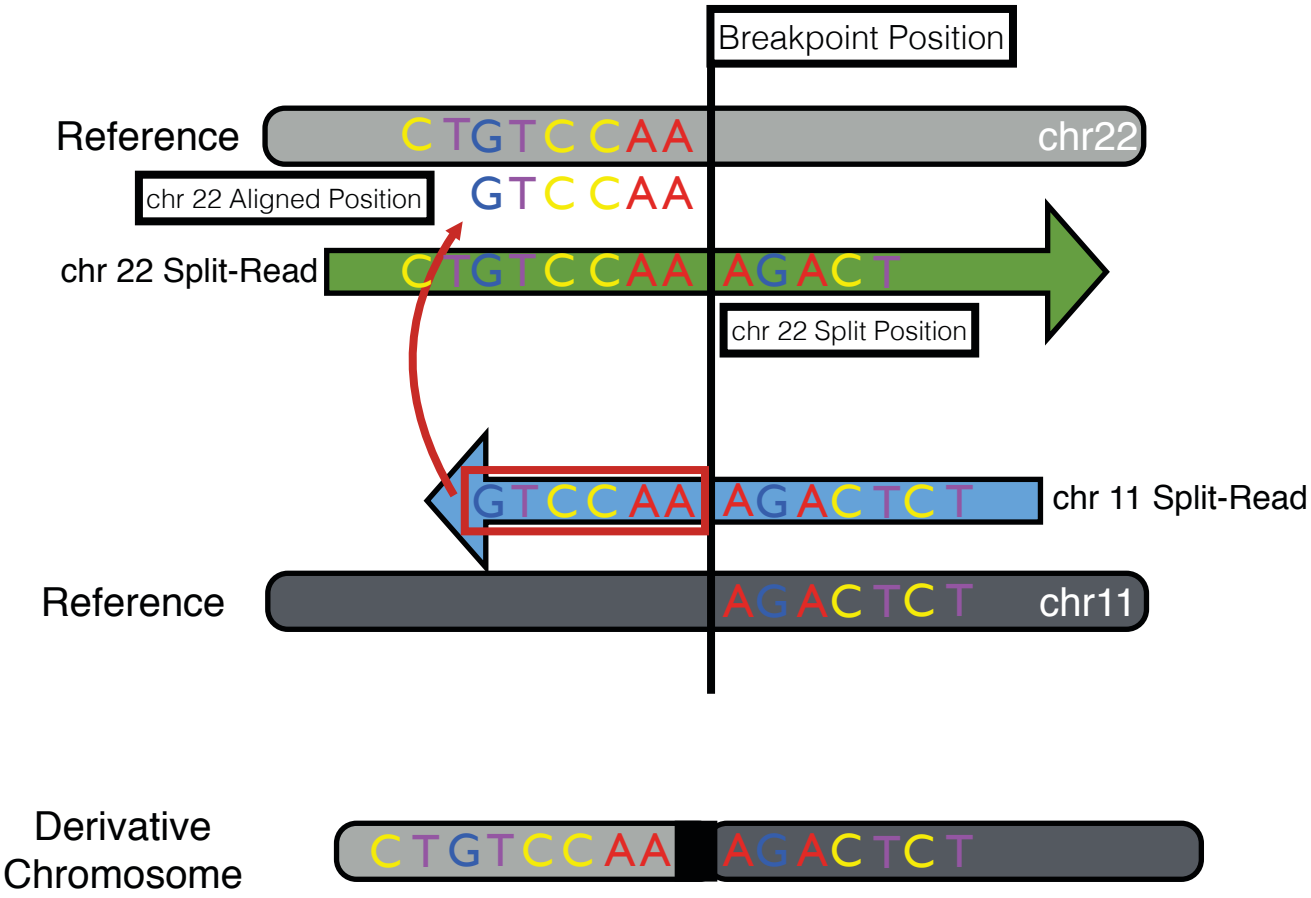


Fig. S13. Copy-number state suggests a one-off burst of rearrangements. A focused view (1MB window) of chromoplexy illustrates that there is copy-neutrality at chromoplectic breakpoint loci suggesting the mechanism of chromoplexy is a one-off mutation burst. A three-way rearrangement pattern is depicted.

A



B

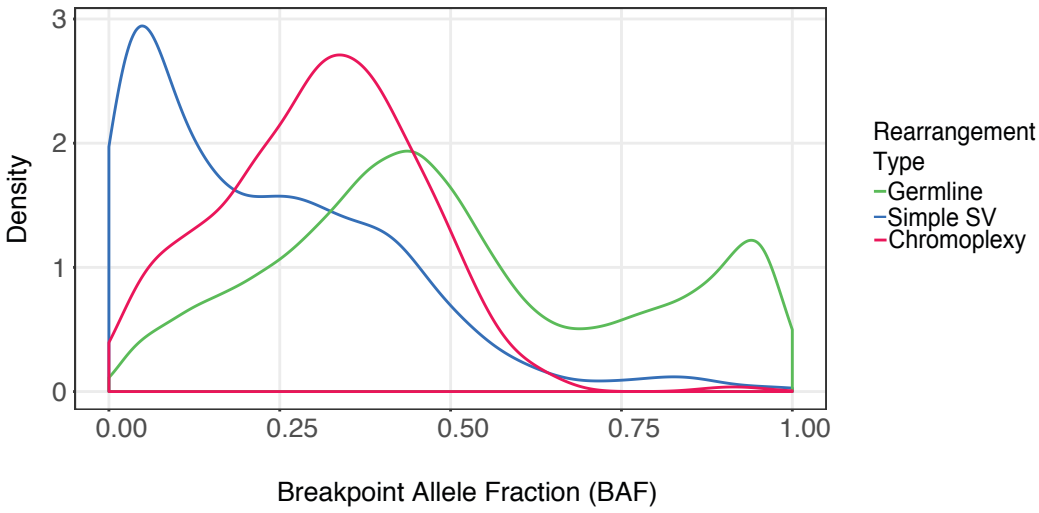


Figure S14

Fig. S14. Calculating and evaluating the breakpoint allele fraction (BAF). The breakpoint allele fraction (BAF) is the proportion of reads supporting the breakpoint of a rearrangement compared to the local coverage and is analogous to a variant allele fraction (VAF). A higher BAF signifies an early rearrangement. **(A)** A schematic representation of how the BAF is calculated. Breakpoint-containing reads map to the reference (chr22), but contain nucleotide bases that map elsewhere in the genome (chr11). The position where this transition occurs is the split position. Our tool maps split reads to the other chromosome. Split-reads from the other side of the rearrangement (chr11) are mapped to side 1 (chr22). The aligned position is represented by read alignments. The BAF is calculated by dividing the number of split-reads (split and aligned) by the total coverage. **(B)** BAFs demonstrate that chromoplectic rearrangements (pink) arise earlier than simple structural variants (blue) given the higher breakpoint allele frequency (BAF). As a control, germline structural variants (green) are plotted and exhibit a bimodal distribution at 0.5 and 1.0, as expected.

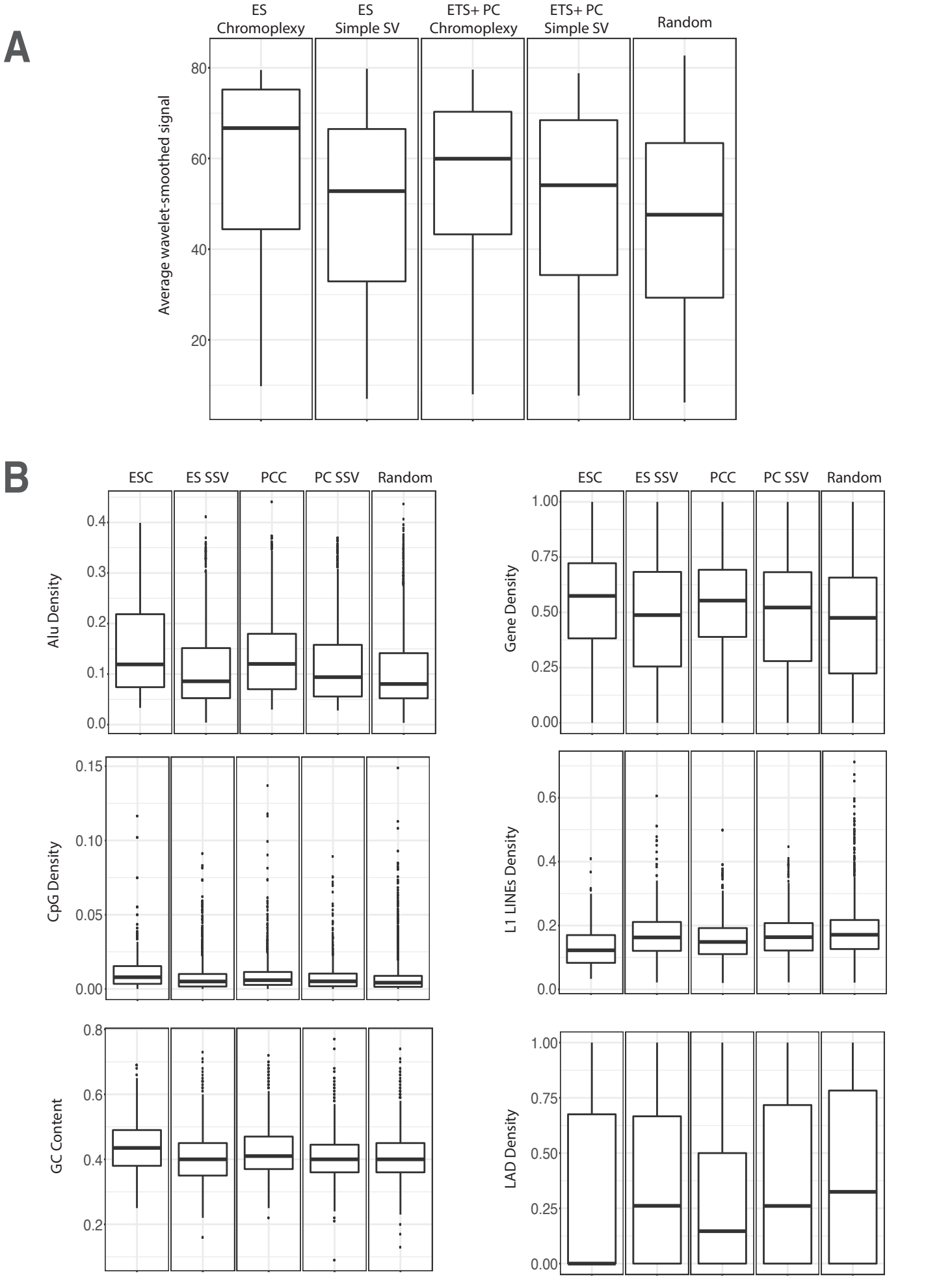

A
ES Chromoplexy
ES Simple SV
ETS+ PC Chromoplexy
ETS+ PC Simple SV
Random
Average wavelet-smoothed signal
80
60
40
20
B
ESC
ES SSV
PCC
PC SSV
Random
Alu Density
Gene Density
CpG Density
L1 LINEs Density
GC Content
LAD Density

Figure S15

Fig. S15. Chromoplexy rearrangements are consistent with early replication timing. (A) Genomic property analysis demonstrates that chromoplexy in Ewing sarcoma is strongly associated replication timing. The average wavelet-smoothed signal (WSS) shows chromoplectic rearrangements have a higher average WSS than simple structural variants or random points. This suggests chromoplexy is associated with early replication timing. This is also true for ETS+ Prostate cancer **(B)** Other features of the genome also support the association of Ewing chromoplexy (ESC) with early replication timing. Ewing chromoplexy rearrangements have higher GC content, Alu Density, gene density, and CpG island density. Other features such as L1 LINEs and lamina-associated domains (LADs) were depleted around chromoplexy regions. These features supporting early replication timing have a Cohen's d equal to or greater than 0.3 and a corrected p-value less than 0.5. This is not observed in Ewing simple structural variants (ES SSV). Other features consistent with early replication timing in ETS+ Prostate Cancer (PCC) are high Alu density and depletions in LADs and L1 LINEs. All feature density metrics are in 1MB sliding windows centered at 1kb bins, except for GC content which was calculated in a 1 kb bin (see table S4).

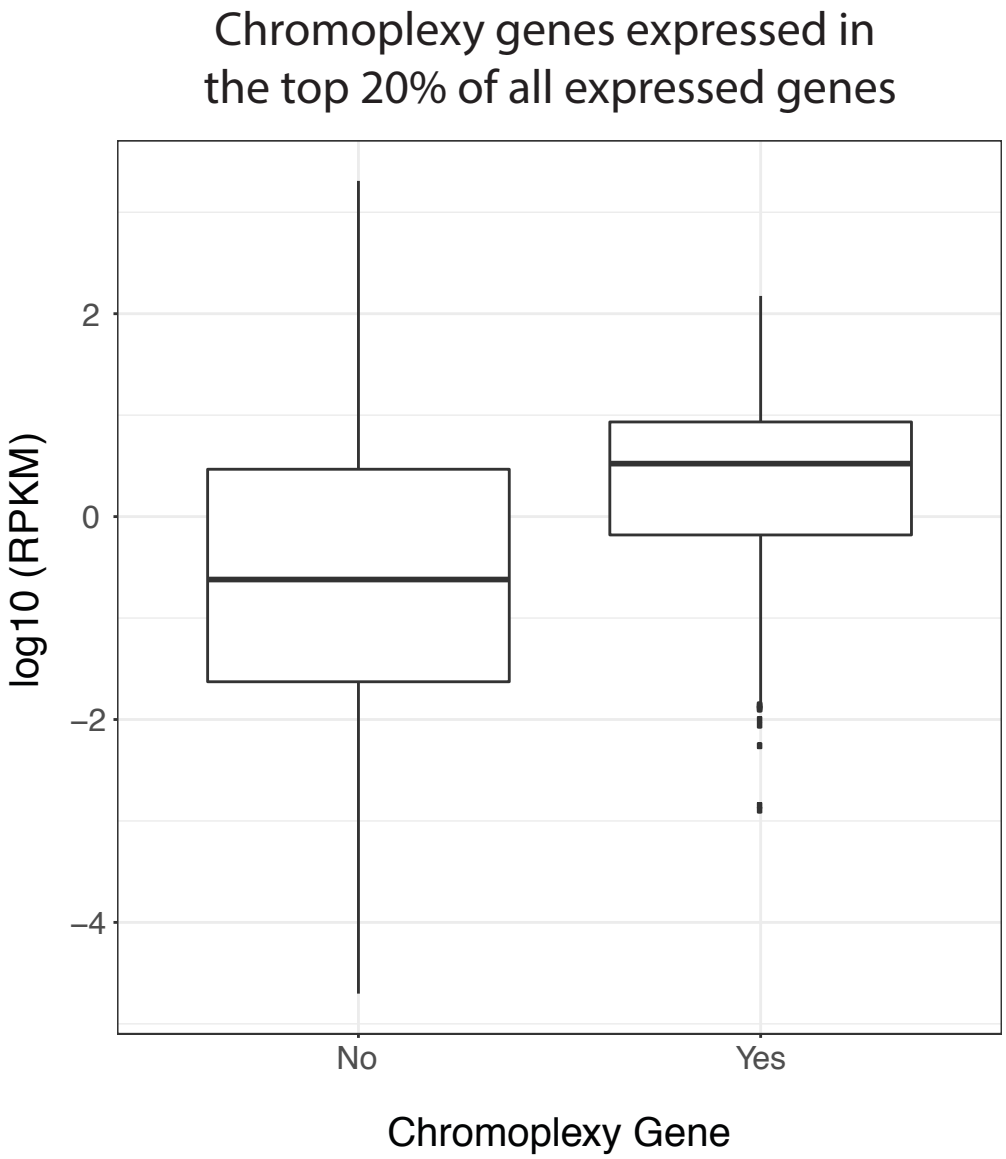


Fig. S16. Global chromoplexy gene expression. Boxplot depicts global transcriptional activity (read per kilobase per million: RPKM) for chromoplexy genes and other expressed genes in the transcriptome. Chromoplexy genes are expressed in the top 20% of the most expressed genes in the transcriptome across the cohort.

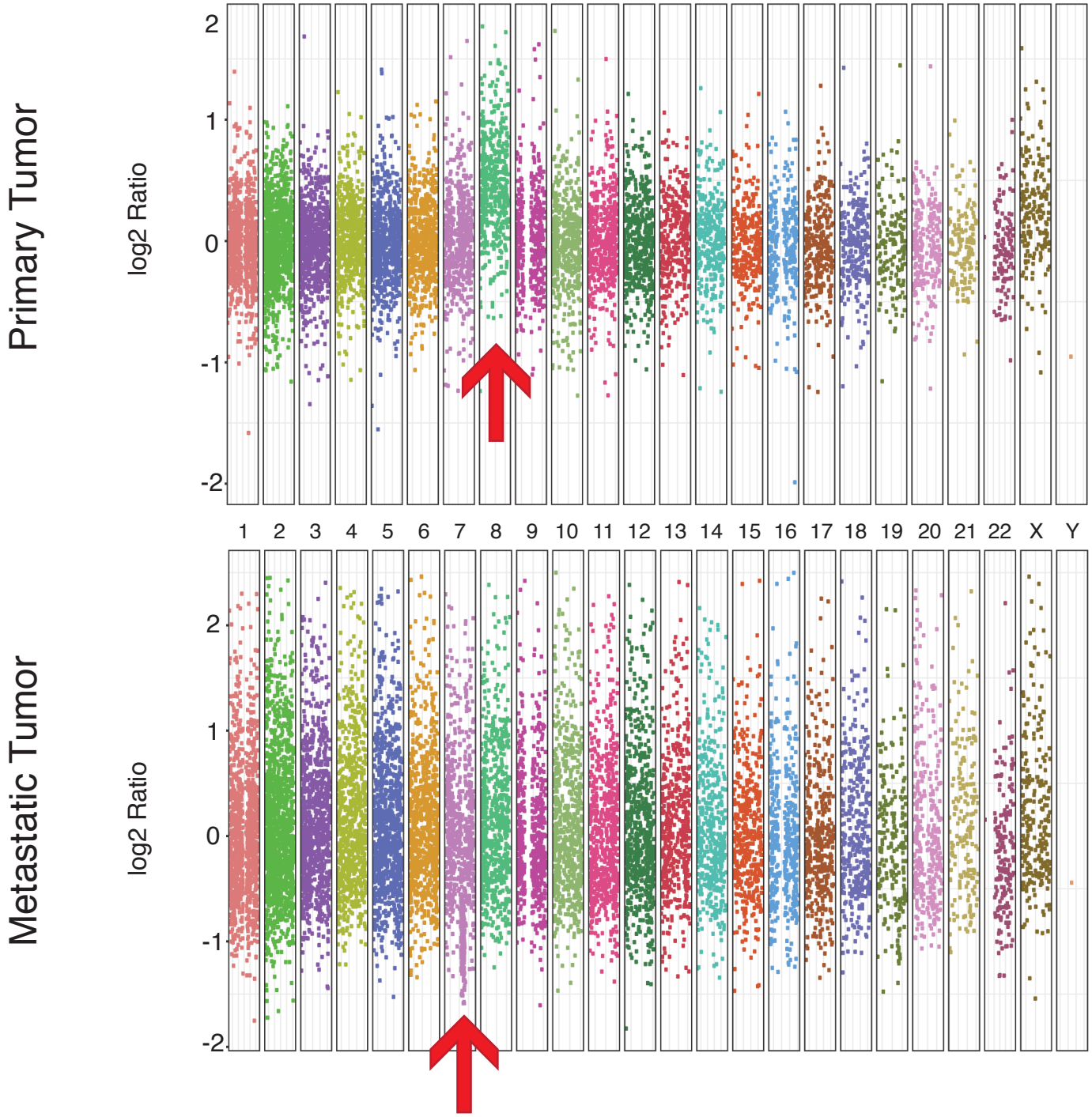


Fig. S17. Molecular Inversion Probe (MIP) assay. MIP copy number was determined for three primary-metastatic ES pairs. Data are shown for one pair in **(A)** and **(B).** Large copy-number alterations are unique to the biopsy or metastatic tumor (red arrow) and are not shared, supporting a model early divergence and parallel evolution between primary and metastatic tumors in Ewing sarcoma.

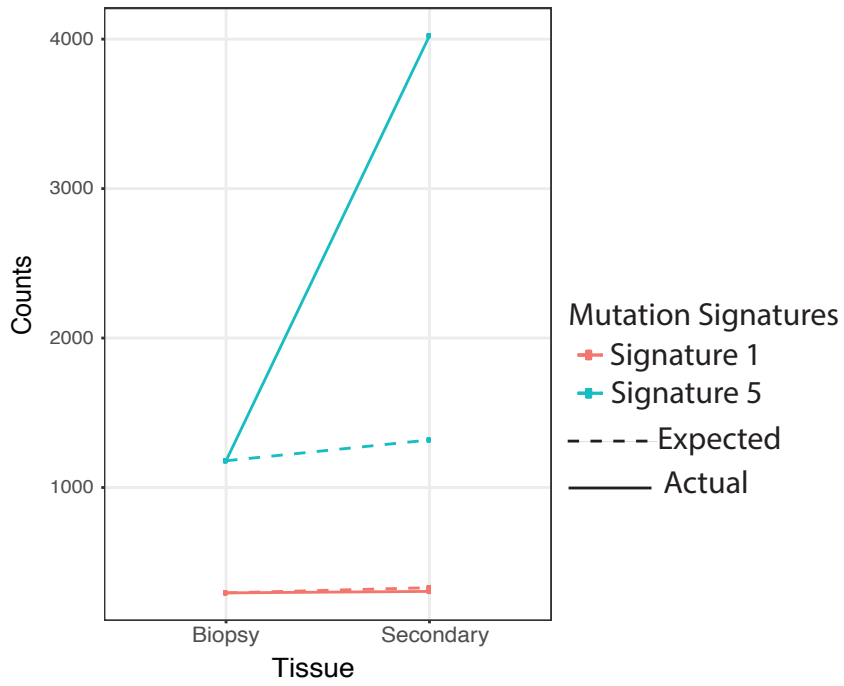


Fig. S18. Validation of COSMIC Signature 1 consistency. We calculated the Signature 1 mutation rate in an ES primary tumor and predicted the number of Signature 1 mutations that should arise in a relapse tumor 1.2 years later. Having sequenced the relapse material, we then calculated the observed Signature 1 mutation accumulation. The line graphs illustrate the consistency of Signature 1 between predicted (dotted line) and actual (solid line) values of primary and relapse tumors. We repeated the analysis for Signature 5 and did not see a steady mutation rate.

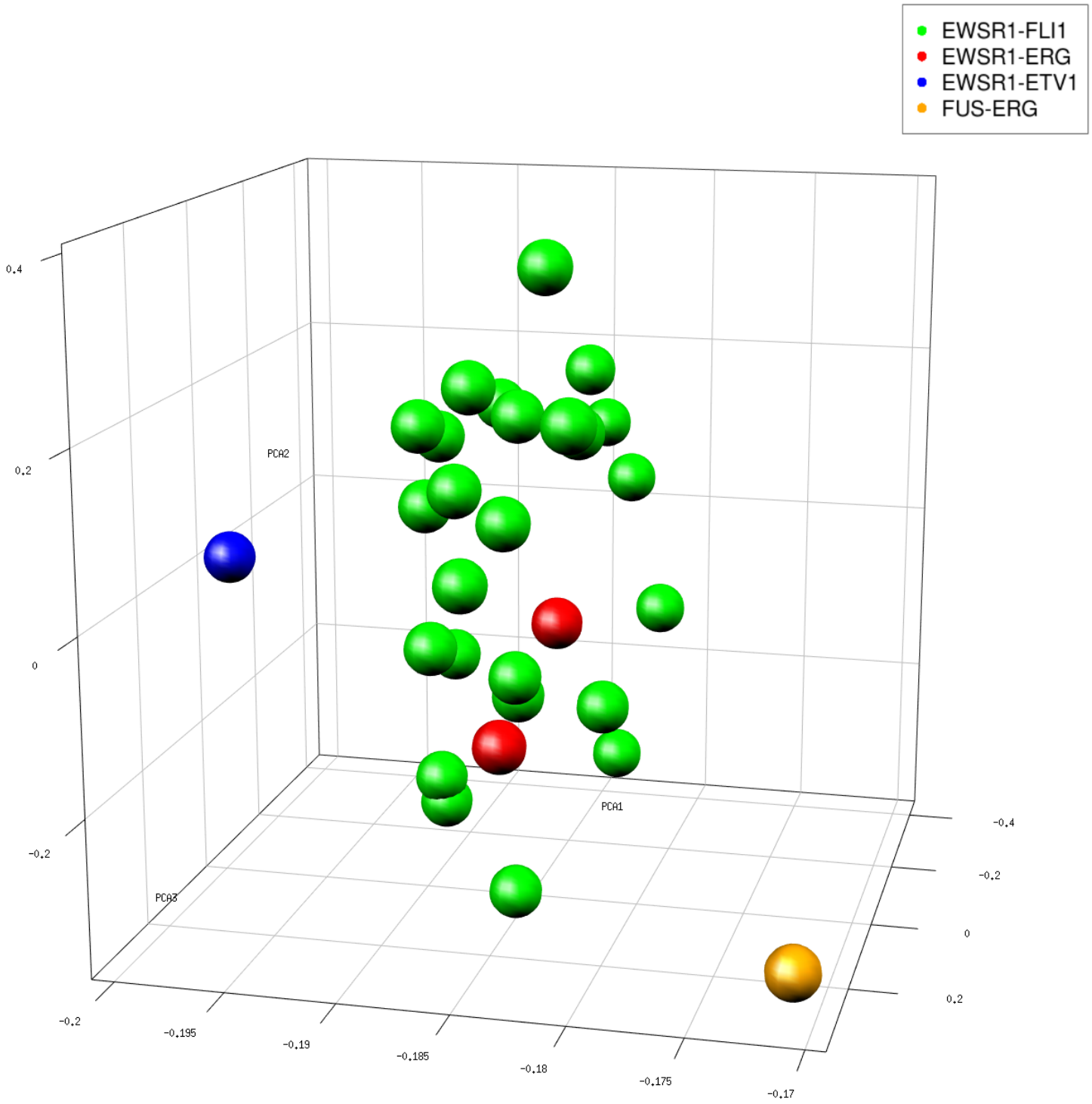


Fig. S19. Principal Component Analysis (PCA) of Ewing sarcoma tumors. PCA of Ewing sarcoma gene expression demonstrates that *EWSR1-FLI1* (green) and *EWSR1-ERG* (red) driven tumors cluster together by gene expression. Non-canonical *FUS-ERG* (gold) and *EWSR1-ETV1*(blue) fusion-driven tumors cluster separately. The *FUS-ERG* tumor is not *EWSR1*-rearranged and was omitted from this study.

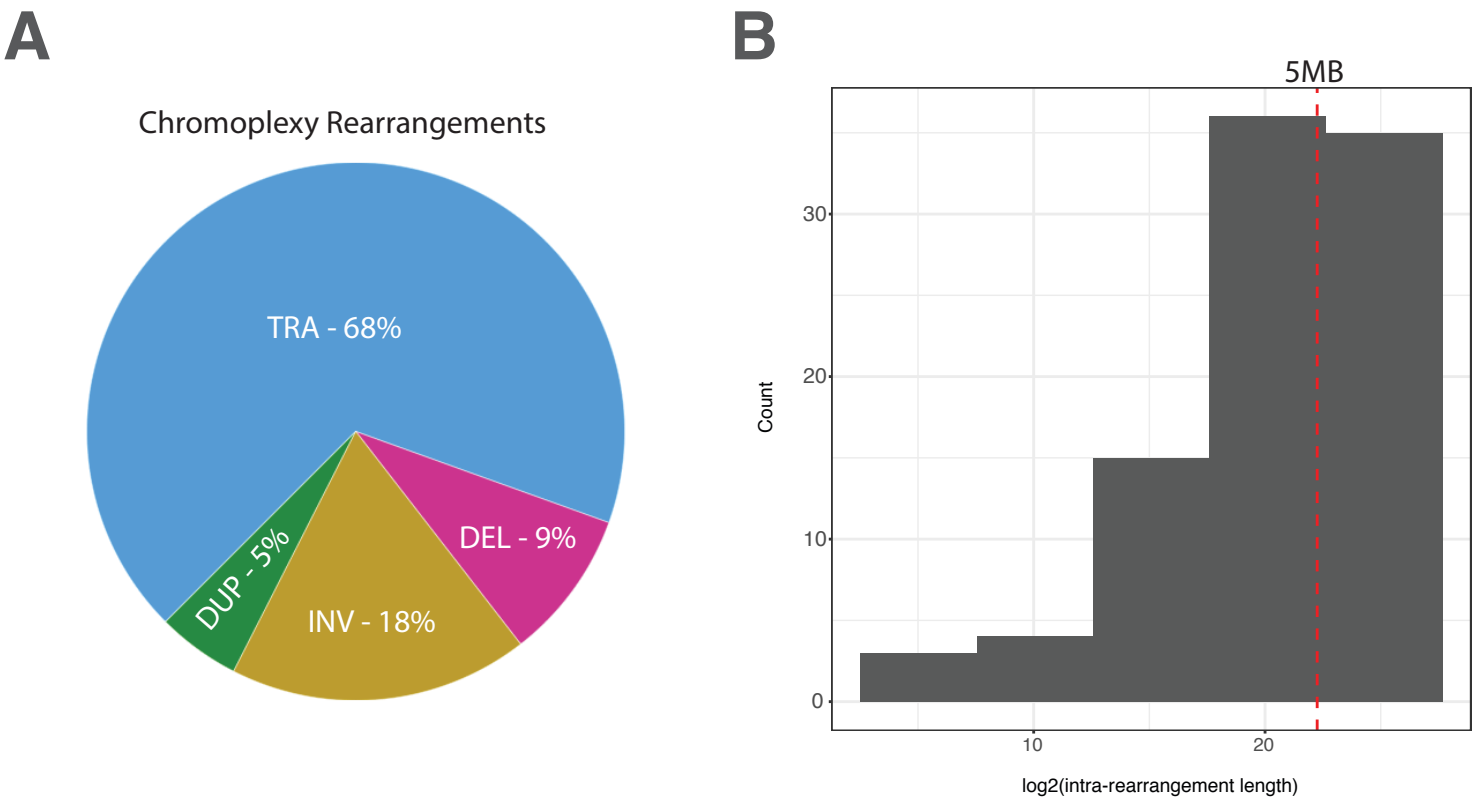
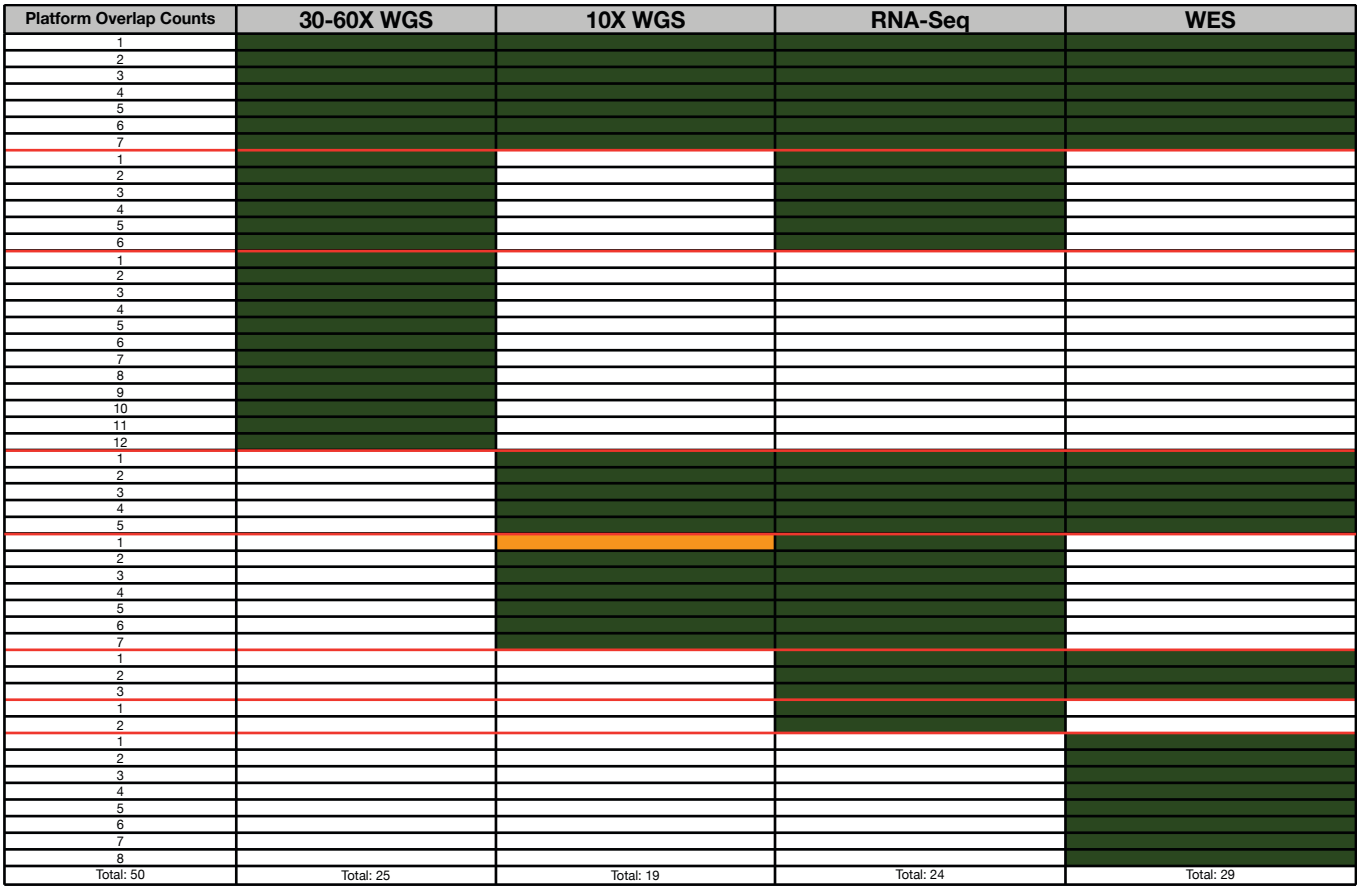


Fig. S20. Prevalence of sub-microscopic rearrangements in ES chromoplexy. (A) Pie chart characterizies the rearrangement subtypes of chromoplexy rearrangements (mainly translocations). (B) 32% of chromoplexy rearrangements are intra-chromosomal with a quarter of them being sub-microscopic (>5MB). These small intrachromosomal rearrangements (to the left of the red dotted line) would be missed by routine cytogenetics.

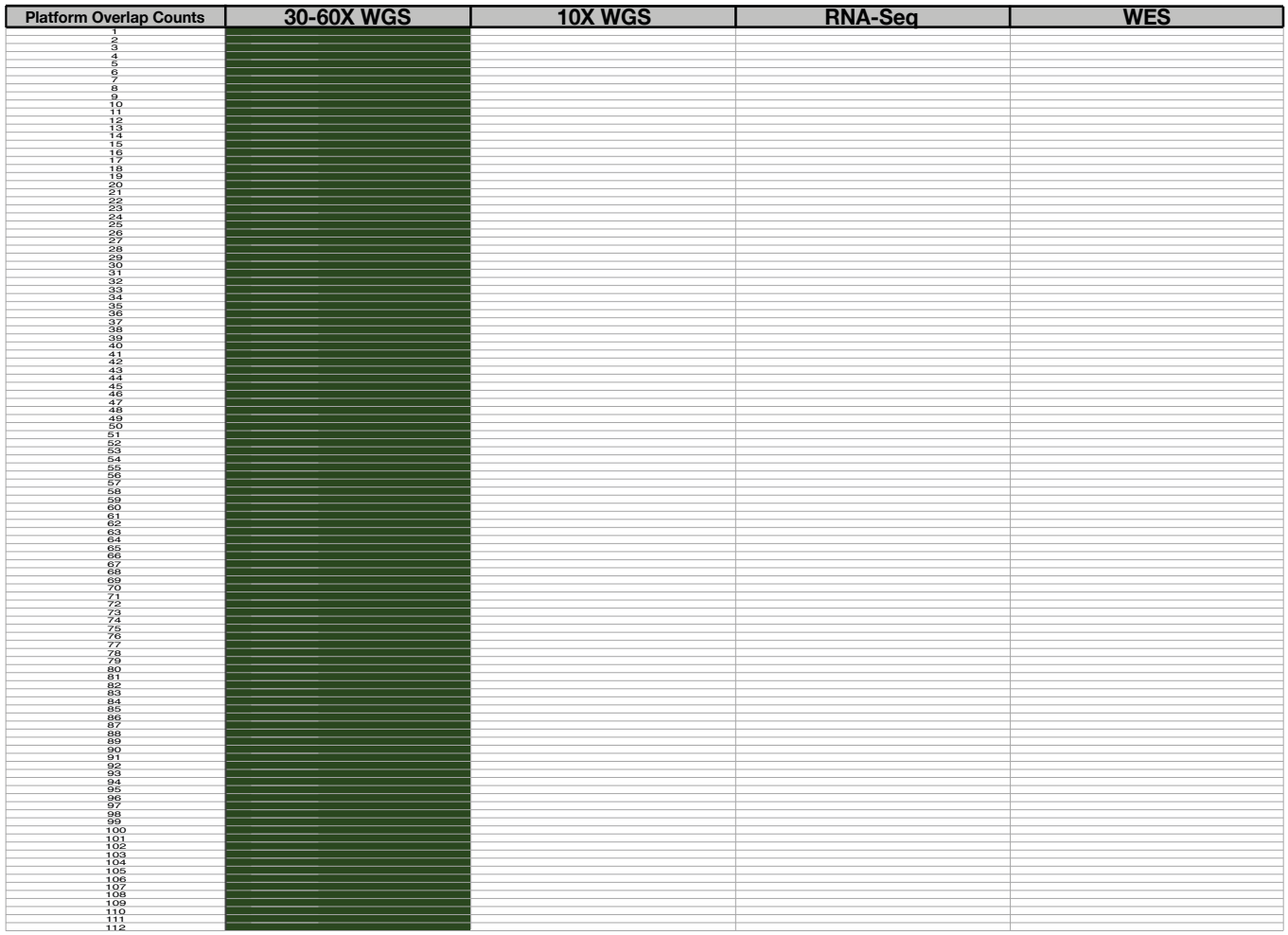
Table S1. Ewing sarcoma sequencing. We performed whole-genome sequencing (WGS) to a depth of either 30-60x (high coverage) or ~10x (rearrangement screen) for 37 Ewing sarcoma tumors. This study focused on the analysis of the high coverage genomes. One rearrangement screen was included as it had unexpectedly high coverage (~20x, orange). For all high coverage genomes, greater than 96% of the reads successfully aligned to the reference genome. To identify transcriptional consequences of mutations, we performed RNA-sequencing on 24 tumors, of which 10 exhibited patterns of chromoplexy. Whole-exome sequencing was also performed on select samples, as indicated. The average coverage of the WES samples was 85x. Finally, we obtained raw sequencing reads from 112 Ewing sarcoma genomes reported by Tirode, Durdez et al. 2014 *Cancer Discovery* to validate our findings.

Toronto Cohort



Platform Overlap Counts	30-60X WGS	10X WGS	RNA-Seq	WES
1				
2				
3				
4				
5				
6				
7				
1				
2				
3				
4				
5				
6				
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
1				
2				
3				
4				
5				
1				
2				
3				
4				
5				
6				
7				
1				
2				
3				
1				
2				
1				
2				
3				
4				
5				
6				
7				
8				
Total: 50	Total: 25	Total: 19	Total: 24	Total: 29

Validation Cohort



Platform Overlap Counts	30-60X WGS	10X WGS	RNA-Seq	WES
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				
101				
102				
103				
104				
105				
106				
107				
108				
109				
110				
111				
112				

Table S2. Disease-defining fusions in sarcomas. We performed low-pass or high-depth whole genome sequencing on 13 sarcomas with known disease-defining fusions. RNA was available for two samples (PD7514 and PD7515).

Tumor type	Canonical fusion	Number of fusion positive tumors screened	Tumor IDs	Rearrangement bursts with fusion	Tumor IDs
Phosphaturic mesenchymal tumor	FN1-FGFR1	1	PD13472	1	PD13472
Chondromyxoid fibroma	Variable-GRM1 fusion	4	PD7514a PD7515a PD10079 PD10080	3	PD7514 PD7515 PD10079
Synovial sarcoma	SS18-SSX	2	PD19375 PD19398	1	PD19398
Solitary fibrous tumor	NAB2-STAT6	6	PD20639 PD20640 PD20667	0	-

Table S3. Gene Set Enrichment Analysis (GSEA). A GSEA of hallmark pathways shows 30/50 pathways dysregulated in chromoplexy ES transcriptomes, which are listed here.

Gene Set Name	Description	Adjusted p-value
ALLOGRAFT REJECTION	*Genes up-regulated during transplant rejection.*	7.00E-24
INFLAMMATORY RESPONSE	*Genes defining inflammatory response.*	6.89E-18
INTERFERON GAMMA RESPONSE	*Genes up-regulated in response to IFNG [GeneID=3458].*	6.89E-18
COMPLEMENT	*Genes encoding components of the complement system, which is part of the innate immune system.*	1.50E-14
KRAS SIGNALING UP	*Genes up-regulated by KRAS activation.*	1.75E-13
IL6-JAK-STAT3 SIGNALING	*Genes up-regulated by IL6 [GeneID=3569] via STAT3 [GeneID=6774], e.g., during acute phase response.*	1.83E-13
TNFA SIGNALING VIA NFKB	*Genes regulated by NF-kB in response to TNF [GeneID=7124].*	1.92E-12
IL2-STAT5 SIGNALING	*Genes up-regulated by STAT5 in response to IL2 stimulation.*	1.98E-11
UV RESPONSE DN	*Genes down-regulated in response to ultraviolet (UV) radiation.*	2.23E-09
ADIPOGENESIS	*Genes up-regulated during adipocyte differentiation (adipogenes*	1.46E-08
APICAL JUNCTION	*Genes encoding components of apical junction complex.*	1.14E-07
HYPOXIA	*Genes up-regulated in response to low oxygen levels (hypoxia).*	8.20E-07
APOPTOSIS	*Genes mediating programmed cell death (apoptosis) by activation of caspases.*	4.97E-06
EPITHELIAL MESENCHYMAL TRANSITION	*Genes defining epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis.*	5.43E-06
ESTROGEN RESPONSE LATE	Genes defining late response to estrogen.	5.43E-06
MYOGENESIS	Genes involved in development of skeletal muscle (myogenesis)	5.43E-06
HEDGEHOG SIGNALING	Genes up-regulated by activation of hedgehog signaling.	2.64E-05
XENOBIOTIC METABOLISM	Genes encoding proteins involved in processing of drugs and other xenobiotics.	3.28E-05
COAGULATION	Genes encoding components of blood coagulation system; also up-regulated in platelets.	7.17E-05
CHOLESTEROL HOMEOSTASIS	Genes involved in cholesterol homeostasis.	9.22E-05
PI3K-AKT MTOR SIGNALING	Genes up-regulated by activation of the PI3K/AKT/mTOR pathway.	0.000621
ESTROGEN RESPONSE EARLY	Genes defining early response to estrogen.	0.00088
FATTY ACID METABOLISM	Genes encoding proteins involved in metabolism of fatty acids.	0.00101
REACTIVE OXIGEN SPECIES PATHWAY	Genes up-regulated by reactive oxigen species (ROS).	0.00138
ANDROGEN RESPONSE	Genes defining response to androgens.	0.00329
KRAS SIGNALING DN	Genes down-regulated by KRAS activation.	0.00384
MTORC1 SIGNALING	Genes up-regulated through activation of mTORC1 complex.	0.00384
UV RESPONSE UP	Genes up-regulated in response to ultraviolet (UV) radiation.	0.00491
INTERFERON ALPHA RESPONSE	Genes up-regulated in response to alpha interferon proteins.	0.0156
BILE ACID METABOLISM	Genes involve in metabolism of bile acids and salts.	0.025

Table S4. Genomic Properties Table. 38 genomic properties are described in this table, including the method (metric of analysis), source and references for the analysis.

Genomic Property	Method	Source	Reference
Centromere	log2 Distance of Breakpoints to Centromeres	UCSC GB hg19 gap track	doi:10.1093/nar/gkt1168
Telomere	log2 Distance of Breakpoints to Tentromeres	UCSC GB hg19 gap track	doi:10.1093/nar/gkt1168
GC content	100 bp bins across the genome	Hg19 FASTA	NA
Sequence complexity	Assigns complexity score to a sequence (high score = low complexity, low score = high complexity)	Hg19 FASTA	NA
H3K9me3	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K36me3	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H4K20me1	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K79me2	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K4me1	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K27ac	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
DNase	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K9ac	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K4me3	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K4me2	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H2A.Z	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
H3K27me3	p-value signal tracks from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
RNA expression	Imputed logRPKM track from RNA-seq data from ChromImpute. Raw data in 25bp bins.	ROADMAP (1,2)	doi:10.1038/nature14248
DNA methylation	Imputed fractional methylation track from DNAMethylSBS data from ChromImpute. Raw data is fraction methylated at particular positions	ROADMAP (1,2)	doi:10.1038/nature14248
Replication time	Average wavelet smoothed signal (WSS) from from 3 cell lines: NHEK (normal skin, ectoderm), GM12878 (normal blood, mesoderm), and MR90 (normal lung, endoderm).	Repli-Seq data available via UCSC	doi:10.1073/pnas.0912402107
Gene Density	log2(distance to),density in sliding window (3)	GENCODE v19	doi:10.1101/gr.135350.111
Lamina associated domains	log2(distance to),density in sliding window (3)	Tig3ET normal human embryonic lung fibroblasts	doi:10.1038/nature06947
CpG islands	log2(distance to),density in sliding window (3)	UCSC GB cpgislands track	doi:10.1093/nar/gkt1168
Direct repeats	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
G-quadruplex	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
Cruciform inverted repeats	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
Triplex mirror repeats	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
Short tandem repeats	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
Z-DNA motifs	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
ALU repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
MIR repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
L1 repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
L2 repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
LTR repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
DNA repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
Simple repeats	log2(distance to),density in sliding window (3)	Repeatmasker 20140131	http://www.repeatmasker.org
G4L1	log2(distance to),density in sliding window (3)	non-B DB v2	doi:10.1093/nar/gks955
Nucleosome occupancy	Nucleosome density signal from K562 and GM12878 cell lines	ENCODE, MNase experiment	https://www.encodeproject.org/experiments/ENCSR000CXQ/
Recombination Rate	Value at nearest point	HAPMAP Phase II	doi:10.1038/nature06258

Notes:
1. ROADMAP tracks used for Ewing's sarcoma: E025, E107, E108, E129 (average 1kb bins)
2. ROADMAP tracks used for Prostate Cancer: E028, E065, E076, E079, E094, E096, E098, E109, E126, E127 (average 1 kb bins)
3. Sizes of sliding windows analyzed (centered around 1kb bin): 1kb, 10kb, 100kb, 1MB

Table S5. Clinical Information. Detailed clinical information for the ES Toronto cohort is described in this table.

Sample	Age at Diagnosis	Sex	Age at Death	Primary Site (+ Mets at Dx)	Relapse	Platform
2849	14 yrs, 1 mth	F		left proximal humerus	NO	WXS
2226	16 yrs	M		ethmoid sinus	NO	WXS, WGS, WTS
2234	14 yrs, 3 mths	F		left distal femur	NO	WXS, WGS, WTS
2283	11 yrs, 10 mths	M		left scapula	NO	WXS
2498	8 yrs, 4 mths	F		Unknown	NO	WXS
3080	15 yrs, 9 mths	M	17 yrs, 10 mths	left superior pubic ramus	YES	WXS, WTS
3553	12 yrs, 6 mths	M		right iliac (+ right lung)	NO	WXS
3943	10 yrs, 7 mths	F		right distal tibia	NO	WXS
2213	10 yrs	M	13 yrs, 8 mths	right distal femur (+ left femur, bilateral lung mets)	YES	WXS, WGS, WTS
2440	8 yrs, 4 mths	M		right ischium (+ bilateral tibia)	NO	WXS, WGS, WTS
2244	9 yrs, 10 mths	M		apex of right thorax	NO	WXS
2253	17 yrs, 3 mths	M	19 yrs, 9 mths	right distal fibula (+ multiple bone, BM)	YES	WXS
2289	15 yrs, 8 mths	F		3rd left rib	YES	WXS
2301	12 yrs, 8 mths	F	13 yrs, 4 mths	left 6th rib (+ BM, multiple bone, lung)	N/A - died while on therapy	WXS
2560	4 yrs, 3 mths	M		right 3rd & 4th ribs	NO	WXS
3336	9 yrs, 7 mths	M		left proximal fibula	NO	WXS, WTS
3480	7 yrs, 3 mths	M		right proximal humerus	NO	WXS
3938	12 yrs, 9 mths	M		right mid humerus	NO	WXS
2331	13 yrs, 9 mths	M	15 yrs, 3 mths	left tibia (+ bilateral lung)	YES	WXS, WTS
2561	13 yrs, 2 mths	M	16 yrs, 2 mths	right buttock/lateral iliac crest (+ bilateral lung)	YES	WXS
2134	14 yrs, 11 mths	F		left distal femur	NO - Developed AML secondary to treatment	WXS, WTS
2187	13 yrs, 4 mths	F		right maxillary sinus	YES	WXS
2925	16 yrs, 7 mths	F		right proximal fibula	NO	WXS, WGS, WTS
3935	15 yrs, 4 mths	M	16 yrs, 8 mths	left calcaneus (+ lung, multiple bone)	YES	WXS
3071	16 yrs, 6 mths	M	Deceased-no date	pelvis (+ lung)	N/A - DOD	WXS, WGS, WTS
3073	15 yrs, 11 mths	F		scapula	NO	WXS
3075	16 yrs, 6 mths	F	18 yrs, 4 mths	pelvis (+ BM)	N/A - DOD	WXS, WTS
3077	17 yrs, 5 mths	M		tibia	NO	WXS, WTS
3131	19 yrs, 4 mths	F		quadriceps	NO	WXS, WTS
1514	2 yrs, 8 mths	F	3 yrs, 2 mths	skull (invading dura and left frontal lobe)	N/A - progressed; DOD	WXS
4021	9 yrs, 7 mths	M		nasopharynx	NO	WGS, WTS
4311	10 yrs, 6 mths	M		right iliac/sacroiliac joint (+right femur, left humerus, right radius, skull, L2, T6, right lung)	YES	WGS, WTS
4094	10 yrs, 2 mths	F		left 7th rib	YES	WGS
4004	13 yr, 7 mths	M		posterior left 4th rib	NO	WGS
4022	16 yrs, 11 mths	M		left hemi pelvis (multiple bone + BM)	N/A - refractory disease	WGS, WTS
4117	13 yrs, 3 mths	M		pelvis (+bilateral lung)	NO	WGS
4120	4 yrs, 8 mths	F		right proximal ulna	NO	WGS
4197	16 yrs, 2 mths	F		anterior left 4th rib	NO	WGS
4226	12 yrs, 4 mths	F		left medial thigh	NO	WGS
4434	14 yrs, 10 mths	M		Unknown	NO	WGS, WTS
4436	15 yrs, 8 mths	F		Unknown	NO	WTS
4438	2 yrs, 10 mths	M	3 yrs, 6 mths	Unknown	NO	WTS
4458	36 years, 7 months	M	38 years, 5 months	proximal tibia	YES (Lung)	WGS, WTS
4459	17 years, 1 month	M		proximal humerus	NO	WGS, WTS
4460	24 years	F		proximal femur	NO	WGS, WTS
4461	21 years, 4 months	M		scapula	NO	WGS, WTS
4462	25 years, 7 months	F		ilium	YES (bone- skull; lung; brain)	WGS, WTS
4463	28 years, 10 months	M	30 years, 10 months	distal femur (primary M1)	NO	WGS
4464	17 years, 8 months	M		acetabulum	NO	WGS, WTS
4465	21 years, 6 months	F	22 years, 8 months	proximal femur	YES (bone - spine; pelvis; contralateral femur)	WGS, WTS
4466	21 years, 8 months	F	23 years, 4 months	ilium	YES (Lung)	WGS, WTS

References and Notes

1. P. J. Stephens, C. D. Greenman, B. Fu, F. Yang, G. R. Bignell, L. J. Mudie, E. D. Pleasance, K. W. Lau, D. Beare, L. A. Stebbings, S. McLaren, M.-L. Lin, D. J. McBride, I. Varela, S. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A. P. Butler, J. W. Teague, M. A. Quail, J. Burton, H. Swerdlow, N. P. Carter, L. A. Morsberger, C. Iacobuzio-Donahue, G. A. Follows, A. R. Green, A. M. Flanagan, M. R. Stratton, P. A. Futreal, P. J. Campbell, Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011). doi:10.1016/j.cell.2010.11.055 Medline

2. S. C. Baca, D. Prandi, M. S. Lawrence, J. M. Mosquera, A. Romanel, Y. Drier, K. Park, N. Kitabayashi, T. Y. MacDonald, M. Ghandi, E. Van Allen, G. V. Kryukov, A. Sboner, J.-P. Theurillat, T. D. Soong, E. Nickerson, D. Auclair, A. Tewari, H. Beltran, R. C. Onofrio, G. Boysen, C. Guiducci, C. E. Barbieri, K. Cibulskis, A. Sivachenko, S. L. Carter, G. Saksena, D. Voet, A. H. Ramos, W. Winckler, M. Cipicchio, K. Ardlie, P. W. Kantoff, M. F. Berger, S. B. Gabriel, T. R. Golub, M. Meyerson, E. S. Lander, O. Elemento, G. Getz, F. Demichelis, M. A. Rubin, L. A. Garraway, Punctuated evolution of prostate cancer genomes. *Cell* **153**, 666–677 (2013). doi:10.1016/j.cell.2013.03.021 Medline

3. F. Mitelman, B. Johansson, F. Mertens, The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **7**, 233–245 (2007). doi:10.1038/nrc2091 Medline

4. E. Papaemmanuil, I. Rapado, Y. Li, N. E. Potter, D. C. Wedge, J. Tubio, L. B. Alexandrov, P. Van Loo, S. L. Cooke, J. Marshall, I. Martincorena, J. Hinton, G. Gundem, F. W. van Delft, S. Nik-Zainal, D. R. Jones, M. Ramakrishna, I. Titley, L. Stebbings, C. Leroy, A. Menzies, J. Gamble, B. Robinson, L. Mudie, K. Raine, S. O'Meara, J. W. Teague, A. P. Butler, G. Cazzaniga, A. Biondi, J. Zuna, H. Kempski, M. Muschen, A. M. Ford, M. R. Stratton, M. Greaves, P. J. Campbell, RAG-mediated recombination is the predominant driver of oncogenic rearrangement in *ETV6-RUNX1* acute lymphoblastic leukemia. *Nat. Genet.* **46**, 116–125 (2014). doi:10.1038/ng.2874 Medline

5. S. Sankar, S. L. Lessnick, Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet.* **204**, 351–365 (2011). doi:10.1016/j.cancergen.2011.07.008 Medline

6. H. Kovar, Blocking the road, stopping the engine or killing the driver? Advances in targeting EWS/FLI-1 fusion in Ewing sarcoma as novel therapy. *Expert Opin. Ther. Targets* **18**, 1315–1328 (2014). doi:10.1517/14728222.2014.947963 Medline

7. A. S. Brohl, D. A. Solomon, W. Chang, J. Wang, Y. Song, S. Sindiri, R. Patidar, L. Hurd, L. Chen, J. F. Shern, H. Liao, X. Wen, J. Gerard, J.-S. Kim, J. A. Lopez Guerrero, I. Machado, D. H. Wai, P. Picci, T. Triche, A. E. Horvai, M. Miettinen, J. S. Wei, D. Catchpool, A. Llombart-Bosch, T. Waldman, J. Khan, The genomic landscape of the Ewing sarcoma family of tumors reveals recurrent *STAG2* mutation. *PLOS Genet.* **10**, e1004475 (2014). doi:10.1371/journal.pgen.1004475 Medline

8. B. D. Crompton, C. Stewart, A. Taylor-Weiner, G. Alexe, K. C. Kurek, M. L. Calicchio, A. Kiezun, S. L. Carter, S. A. Shukla, S. S. Mehta, A. R. Thorner, C. de Torres, C. Lavarino, M. Suñol, A. McKenna, A. Sivachenko, K. Cibulskis, M. S. Lawrence, P. Stojanov, M. Rosenberg, L. Ambrogio, D. Auclair, S. Seepo, B. Blumenstiel, M. DeFelice, I. Imaz-Rosshandler, A. Schwarz-Cruz Y Celis, M. N. Rivera, C. Rodriguez-Galindo, M. D.

Fleming, T. R. Golub, G. Getz, J. Mora, K. Stegmaier, The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov.* **4**, 1326–1341 (2014). doi:10.1158/2159-8290.CD-13-1037 Medline

9. M. S. Lawrence, P. Stojanov, P. Polak, G. V. Kryukov, K. Cibulskis, A. Sivachenko, S. L. Carter, C. Stewart, C. H. Mermel, S. A. Roberts, A. Kiezun, P. S. Hammerman, A. McKenna, Y. Drier, L. Zou, A. H. Ramos, T. J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M. L. Cortés, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D. I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A. M. Dulak, J. Lohr, D.-A. Landau, C. J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S. A. McCarroll, J. Mora, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S. B. Gabriel, C. W. M. Roberts, J. A. Biegel, K. Stegmaier, A. J. Bass, L. A. Garraway, M. Meyerson, T. R. Golub, D. A. Gordenin, S. Sunyaev, E. S. Lander, G. Getz, Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013). doi:10.1038/nature12213 Medline

10. F. Tirode, D. Surdez, X. Ma, M. Parker, M. C. Le Deley, A. Bahrami, Z. Zhang, E. Lapouble, S. Grossetête-Lalami, M. Rusch, S. Reynaud, T. Rio-Frio, E. Hedlund, G. Wu, X. Chen, G. Pierron, O. Oberlin, S. Zaidi, G. Lemmon, P. Gupta, B. Vadodaria, J. Easton, M. Gut, L. Ding, E. R. Mardis, R. K. Wilson, S. Shurtleff, V. Laurence, J. Michon, P. Marec-Bérard, I. Gut, J. Downing, M. Dyer, J. Zhang, O. Delattre; St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project and the International Cancer Genome Consortium, Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of *STAG2* and *TP53* mutations. *Cancer Discov.* **4**, 1342–1353 (2014). doi:10.1158/2159-8290.CD-14-0622 Medline

11. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. J. R. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A.-L. Børresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjörd, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, N. Jäger, D. T. W. Jones, D. Jones, S. Knappskog, M. Kool, S. R. Lakhani, C. López-Otín, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. J. Tutt, R. Valdés-Mas, M. M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates, J. Zucman-Rossi, P. A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, M. R. Stratton; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). doi:10.1038/nature12477 Medline

12. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, P. J. Campbell, M. R. Stratton, Deciphering signatures of mutational processes operative in human cancer. *Cell Reports* **3**, 246–259 (2013). doi:10.1016/j.celrep.2012.12.008 Medline

13. L. B. Alexandrov, P. H. Jones, D. C. Wedge, J. E. Sale, P. J. Campbell, S. Nik-Zainal, M. R. Stratton, Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015). doi:10.1038/ng.3441 Medline

14. A. Zoubek, C. Pfleiderer, M. Salzer-Kuntschik, G. Amann, R. Windhager, F. M. Fink, E. Koscielniak, O. Delattre, S. Strehl, P. F. Ambros, H. Gadner, H. Kovar, Variability of EWS chimaeric transcripts in Ewing tumours: A comparison of clinical and molecular data. *Br. J. Cancer* **70**, 908–913 (1994). doi:10.1038/bjc.1994.419 Medline

15. C. M. Hattinger, U. Pötschger, M. Tarkkanen, J. Squire, M. Zielenska, S. Kiuru-Kuhlefelt, L. Kager, P. Thorner, S. Knuutila, F. K. Niggli, P. F. Ambros, H. Gadner, D. R. Betts, Prognostic impact of chromosomal aberrations in Ewing tumours. *Br. J. Cancer* **86**, 1763–1769 (2002). doi:10.1038/sj.bjc.6600332 Medline

16. P. M. Neilsen, K. I. Pishas, D. F. Callen, D. M. Thomas, Targeting the p53 pathway in Ewing sarcoma. *Sarcoma* **2011**, 746939 (2011). doi:10.1155/2011/746939 Medline

17. C. S. Cooper, R. Eeles, D. C. Wedge, P. Van Loo, G. Gundem, L. B. Alexandrov, B. Kremeyer, A. Butler, A. G. Lynch, N. Camacho, C. E. Massie, J. Kay, H. J. Luxton, S. Edwards, Z. Kote-Jarai, N. Dennis, S. Merson, D. Leongamornlert, J. Zamora, C. Corbishley, S. Thomas, S. Nik-Zainal, M. Ramakrishna, S. O'Meara, L. Matthews, J. Clark, R. Hurst, R. Mithen, R. G. Bristow, P. C. Boutros, M. Fraser, S. Cooke, K. Raine, D. Jones, A. Menzies, L. Stebbings, J. Hinton, J. Teague, S. McLaren, L. Mudie, C. Hardy, E. Anderson, O. Joseph, V. Goody, B. Robinson, M. Maddison, S. Gamble, C. Greenman, D. Berney, S. Hazell, N. Livni, C. Fisher, C. Ogden, P. Kumar, A. Thompson, C. Woodhouse, D. Nicol, E. Mayer, T. Dudderidge, N. C. Shah, V. Gnanapragasam, T. Voet, P. Campbell, A. Futreal, D. Easton, A. Y. Warren, C. S. Foster, M. R. Stratton, H. C. Whitaker, U. McDermott, D. S. Brewer, D. E. Neal; ICGC Prostate Group, Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat. Genet.* **47**, 367–372 (2015). doi:10.1038/ng.3221 Medline

18. A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015). doi:10.1016/j.cels.2015.12.004 Medline

19. K. H. Nord, H. Lilljebjörn, F. Vezzi, J. Nilsson, L. Magnusson, J. Tayebwa, D. de Jong, J. V. M. G. Bovée, P. C. W. Hogendoorn, K. Szuhai, *GRM1* is upregulated through gene fusion and promoter swapping in chondromyxoid fibroma. *Nat. Genet.* **46**, 474–477 (2014). doi:10.1038/ng.2927 Medline

20. A. Malhotra, M. Lindberg, G. G. Faust, M. L. Leibowitz, R. A. Clark, R. M. Layer, A. R. Quinlan, I. M. Hall, Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. *Genome Res.* **23**, 762–776 (2013). doi:10.1101/gr.143677.112 Medline

21. J. O. Korbel, P. J. Campbell, Criteria for inference of chromothripsis in cancer genomes. *Cell* **152**, 1226–1236 (2013). doi:10.1016/j.cell.2013.02.023 Medline

22. R. S. Hansen, S. Thomas, R. Sandstrom, T. K. Canfield, R. E. Thurman, M. Weaver, M. O. Dorschner, S. M. Gartler, J. A. Stamatoyannopoulos, Sequencing newly replicated DNA

reveals widespread plasticity in human replication timing. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 139–144 (2010). doi:10.1073/pnas.0912402107 Medline

23. J. Sima, D. M. Gilbert, Complex correlations: Replication timing and mutational landscapes during cancer and genome evolution. *Curr. Opin. Genet. Dev.* **25**, 93–100 (2014). doi:10.1016/j.gde.2013.11.022 Medline

24. J. Merlevede, N. Droin, T. Qin, K. Meldi, K. Yoshida, M. Morabito, E. Chautard, D. Auboeuf, P. Fenaux, T. Braun, R. Itzykson, S. de Botton, B. Quesnel, T. Commes, E. Jourdan, W. Vainchenker, O. Bernard, N. Pata-Merci, S. Solier, V. Gayevskiy, M. E. Dinger, M. J. Cowley, D. Selimoglu-Buet, V. Meyer, F. Artiguenave, J.-F. Deleuze, C. Preudhomme, M. R. Stratton, L. B. Alexandrov, E. Padron, S. Ogawa, S. Koscielny, M. Figueroa, E. Solary, Mutation allele burden remains unchanged in chronic myelomonocytic leukaemia responding to hypomethylating agents. *Nat. Commun.* **7**, 10767 (2016). doi:10.1038/ncomms10767 Medline

25. J. W. Gray, Evidence emerges for early metastasis and parallel evolution of primary and metastatic tumors. *Cancer Cell* **4**, 4–6 (2003). doi:10.1016/S1535-6108(03)00167-3 Medline

26. J. F. Brasme, M. Chalumeau, O. Oberlin, D. Valteau-Couanet, N. Gaspar, Time to diagnosis of Ewing tumors in children and adolescents is not associated with metastasis or survival: A prospective multicenter study of 436 patients. *J. Clin. Oncol.* **32**, 1935–1940 (2014). doi:10.1200/JCO.2013.53.8058 Medline

27. P. Roberts, S. A. Burchill, S. Brownhill, C. J. Cullinane, C. Johnston, M. J. Griffiths, D. J. McMullan, N. P. Bown, S. P. Morris, I. J. Lewis, Ploidy and karyotype complexity are powerful prognostic indicators in the Ewing's sarcoma family of tumors: A study by the United Kingdom Cancer Cytogenetics and the Children's Cancer and Leukaemia Group. *Genes Chromosomes Cancer* **47**, 207–220 (2008). doi:10.1002/gcc.20523 Medline

28. K. Cibulskis, M. S. Lawrence, S. L. Carter, A. Sivachenko, D. Jaffe, C. Sougnez, S. Gabriel, M. Meyerson, E. S. Lander, G. Getz, Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013). doi:10.1038/nbt.2514 Medline

29. T. Rausch, T. Zichner, A. Schlattl, A. M. Stütz, V. Benes, J. O. Korbel, DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012). doi:10.1093/bioinformatics/bts378 Medline

30. D. W. Craig, S. Nasser, R. Corbett, S. K. Chan, L. Murray, C. Legendre, W. Tembe, J. Adkins, N. Kim, S. Wong, A. Baker, D. Enriquez, S. Pond, E. Pleasance, A. J. Mungall, R. A. Moore, T. McDaniel, Y. Ma, S. J. M. Jones, M. A. Marra, J. D. Carpten, W. S. Liang, A somatic reference standard for cancer genome sequencing. *Sci. Rep.* **6**, 24607 (2016). doi:10.1038/srep24607 Medline

31. R. Xi, A. G. Hadjipanayis, L. J. Luquette, T.-M. Kim, E. Lee, J. Zhang, M. D. Johnson, D. M. Muzny, D. A. Wheeler, R. A. Gibbs, R. Kucherlapati, P. J. Park, Copy number variation detection in whole-genome sequencing data using the Bayesian information criterion. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1128–E1136 (2011). doi:10.1073/pnas.1110574108 Medline

32. A. Shlien, B. B. Campbell, R. de Borja, L. B. Alexandrov, D. Merico, D. Wedge, P. Van Loo, P. S. Tarpey, P. Coupland, S. Behjati, A. Pollett, T. Lipman, A. Heidari, S. Deshmukh, N. Avitzur, B. Meier, M. Gerstung, Y. Hong, D. M. Merino, M. Ramakrishna, M. Remke, R. Arnold, G. B. Panigrahi, N. P. Thakkar, K. P. Hodel, E. E. Henninger, A. Y. Göksenin, D. Bakry, G. S. Charames, H. Druker, J. Lerner-Ellis, M. Mistry, R. Dvir, R. Grant, R. Elhasid, R. Farah, G. P. Taylor, P. C. Nathan, S. Alexander, S. Ben-Shachar, S. C. Ling, S. Gallinger, S. Constantini, P. Dirks, A. Huang, S. W. Scherer, R. G. Grundy, C. Durno, M. Aronson, A. Gartner, M. S. Meyn, M. D. Taylor, Z. F. Pursell, C. E. Pearson, D. Malkin, P. A. Futreal, M. R. Stratton, E. Bouffet, C. Hawkins, P. J. Campbell, U. Tabori; Biallelic Mismatch Repair Deficiency Consortium, Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nat. Genet.* **47**, 257–262 (2015). doi:10.1038/ng.3202 Medline

33. H. Li, Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**, 2843–2851 (2014). doi:10.1093/bioinformatics/btu356 Medline

34. A. Morgulis, E. M. Gertz, A. A. Schäffer, R. Agarwala, A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *J. Comput. Biol.* **13**, 1028–1040 (2006). doi:10.1089/cmb.2006.13.1028 Medline

35. P. Rice, I. Longden, A. Bleasby, EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000). doi:10.1016/S0168-9525(00)02024-2 Medline

36. M. Zarrei, J. R. MacDonald, D. Merico, S. W. Scherer, A copy number variation map of the human genome. *Nat. Rev. Genet.* **16**, 172–183 (2015). doi:10.1038/nrg3871 Medline

37. J. Cohen, The statistical power of abnormal-social psychological research: A review. *J. Abnorm. Soc. Psychol.* **65**, 145–153 (1962). doi:10.1037/h0045186 Medline

38. J. Cohen, A power primer. *Psychol. Bull.* **112**, 155–159 (1992). doi:10.1037/0033-2909.112.1.155 Medline

39. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015). doi:10.1093/bioinformatics/btu638 Medline

40. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi:10.1093/bioinformatics/btp616 Medline

41. D. J. McCarthy, Y. Chen, G. K. Smyth, Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012). doi:10.1093/nar/gks042 Medline

42. M. S. Jahromi, A. R. Putnam, C. Druzgal, J. Wright, H. Spraker-Perlman, M. Kinsey, H. Zhou, K. M. Boucher, R. L. Randall, K. B. Jones, D. Lucas, A. Rosenberg, D. Thomas, S. L. Lessnick, J. D. Schiffman, Molecular inversion probe analysis detects novel copy number alterations in Ewing sarcoma. *Cancer Genet.* **205**, 391–404 (2012). doi:10.1016/j.cancergen.2012.05.012 Medline

43. C. Pilati, J. Shinde, L. B. Alexandrov, G. Assié, T. André, Z. Hélias-Rodzewicz, R. Ducoudray, D. Le Corre, J. Zucman-Rossi, J.-F. Emile, J. Bertherat, E. Letouzé, P. Laurent-Puig, Mutational signature analysis identifies *MUTYH* deficiency in colorectal cancers and adrenocortical carcinomas. *J. Pathol.* **242**, 10–15 (2017). doi:10.1002/path.4880 Medline