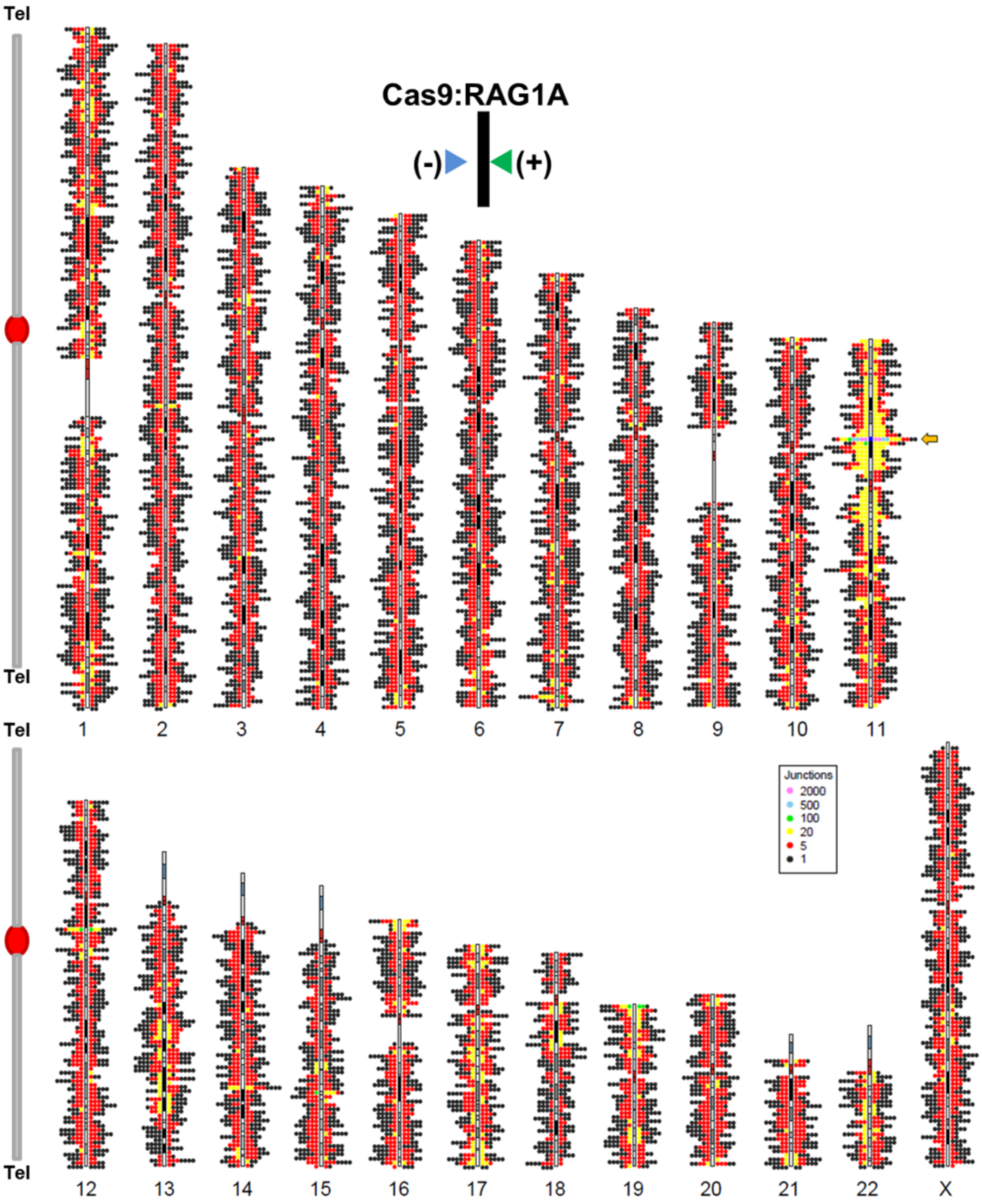
Supplementary Information

Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases

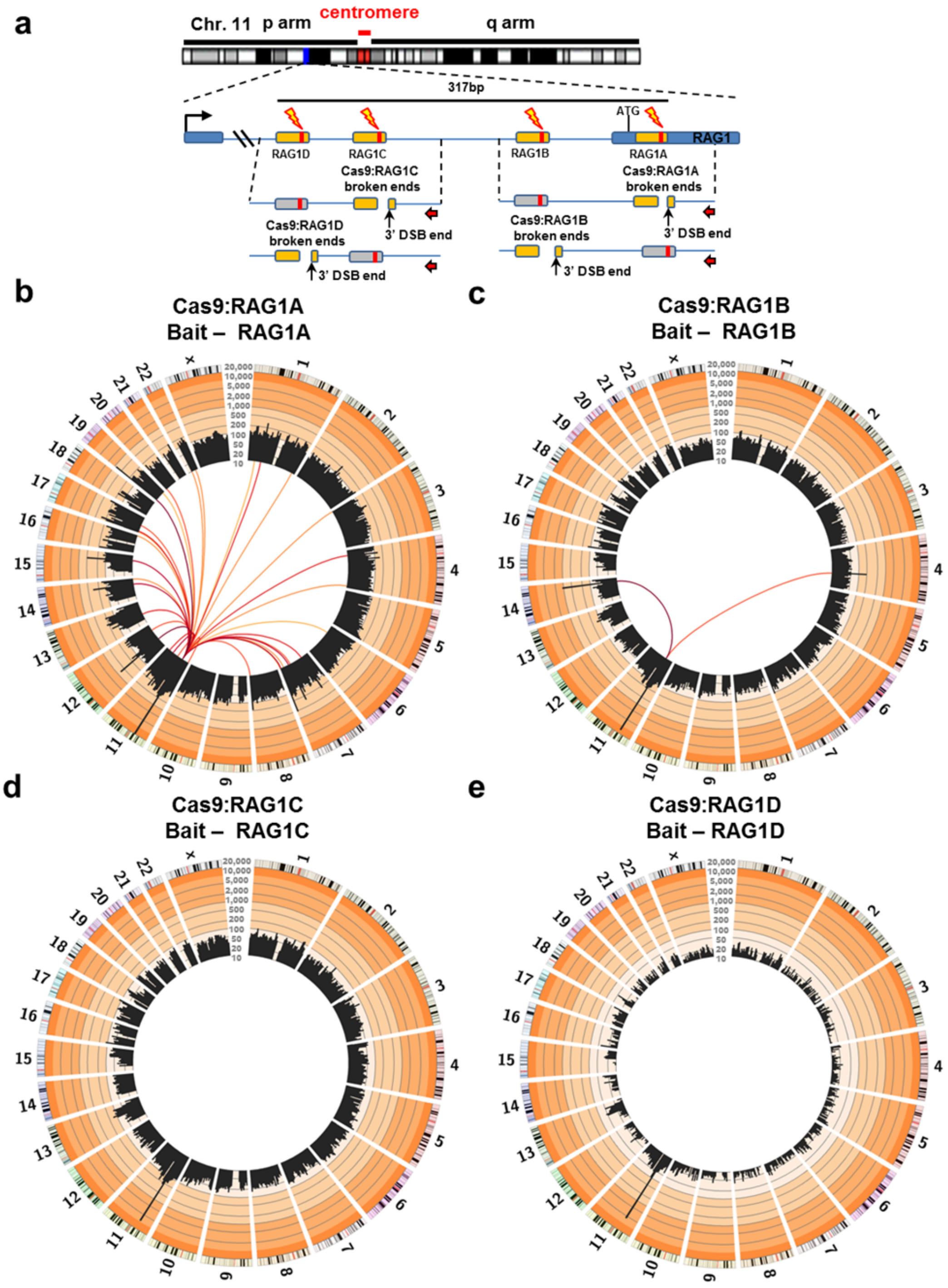
Richard L. Frock[1,2,3,4], Jiazhi Hu[1,2,3,4], Robin M. Meyers[1,2,3], Yu-Jui Ho[1,2,3], Erina Kii[1,2,3] &
Frederick W. Alt[1,2,3,5]

[1]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, Massachusetts, USA
[2]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA
[3]Howard Hughes Medical Institute, Boston, Massachusetts, USA
[4]Equal Contribution
[5]Correspondence should be addressed to alt@enders.tch.harvard.edu



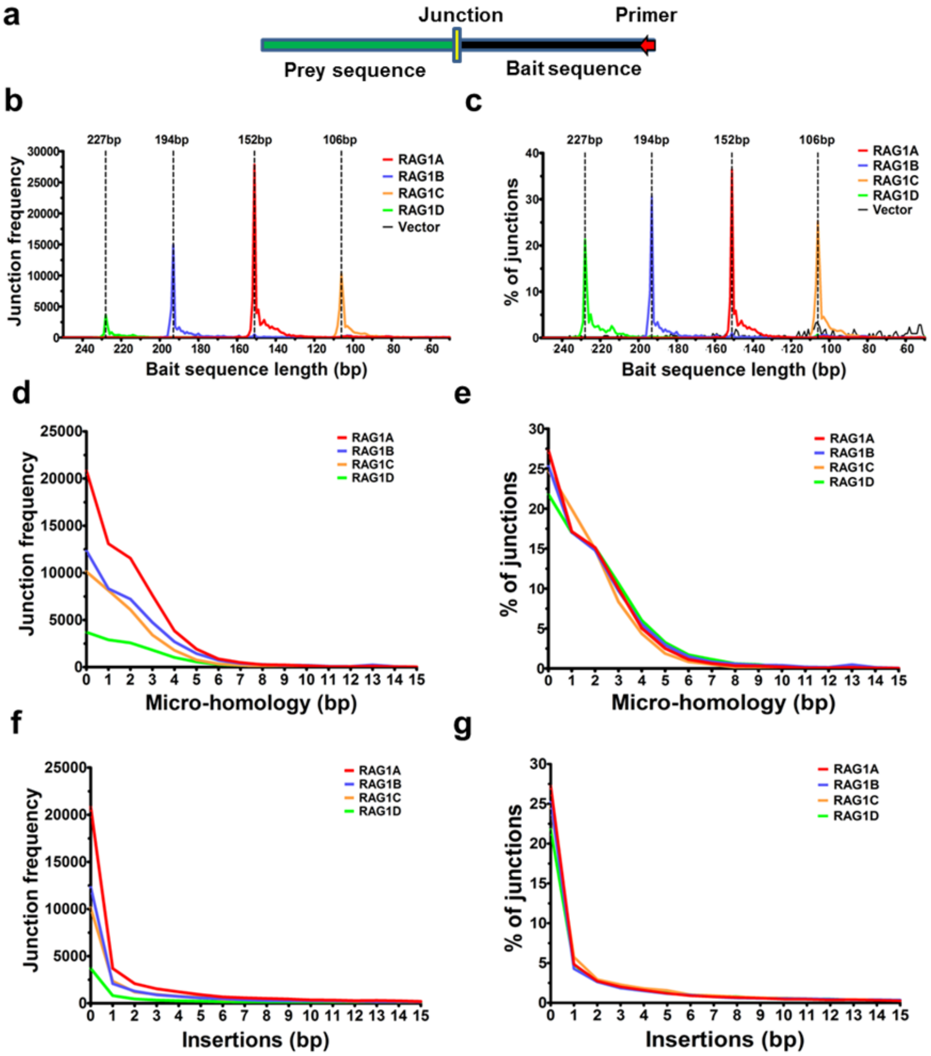
Supplementary Figure 1. Human genome dot plot of Cas9:RAG1A prey junctions using LAM-PCR HTGTS. Joins to the (-) and (+) chromosomal orientations are displayed on the left and right sides of each chromosome respectively. Colored dots represent prey junction frequency within a 1.5Mb binned region. The following colors indicate the following frequencies: black = 1, red = 5, yellow =20, green = 100 blue = 500 pink =2,000. Barren areas of variable length occurred in and around centromeres or rDNA genes, consistent with mapping difficulties in these highly repetitive regions[47]. Total junctions listed: 61,700; N=2.

Supplementary Figure 2

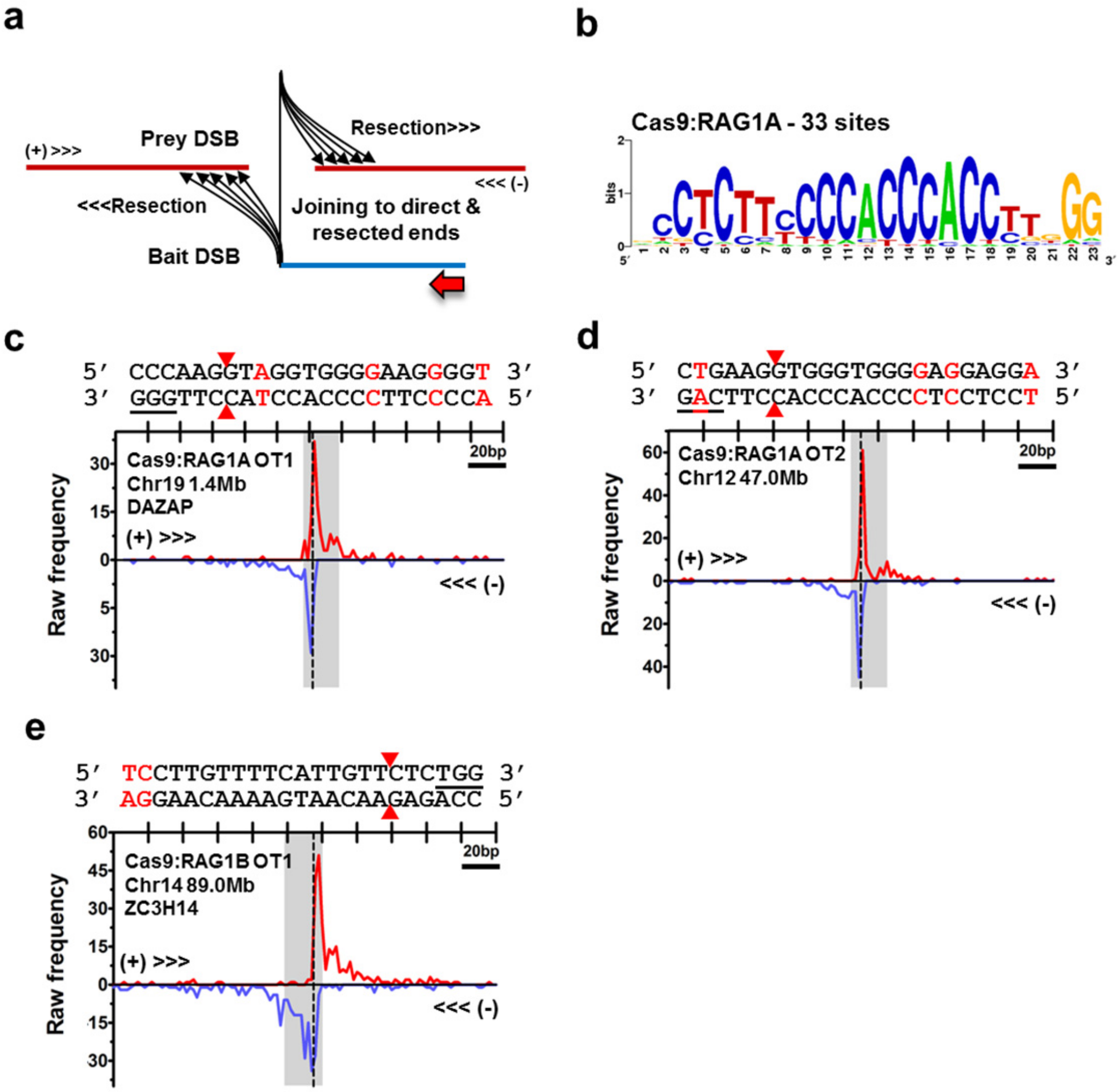


Supplementary Figure 2. HTGTS libraries from Cas9:RAG1A-D. (a) RAG1 locus indicating RAG1B, RAG1C, and RAG1D targeted sites. Red arrow indicates cloning primer. (b-e) Circos plots show (b) Cas9:RAG1A, (c) Cas9:RAG1B, (d) Cas9:RAG1C, and (e) Cas9:RAG1D HTGTS libraries. No recurrent hotspots were detected for RAG1C and D. Biological replicates: N=6 each. Circos plots are presented as similarly as described in Fig 1 (see legend); junction scale: 10-20,000. See Supplementary Table 1 for total library sizes.

Supplementary Figure 3

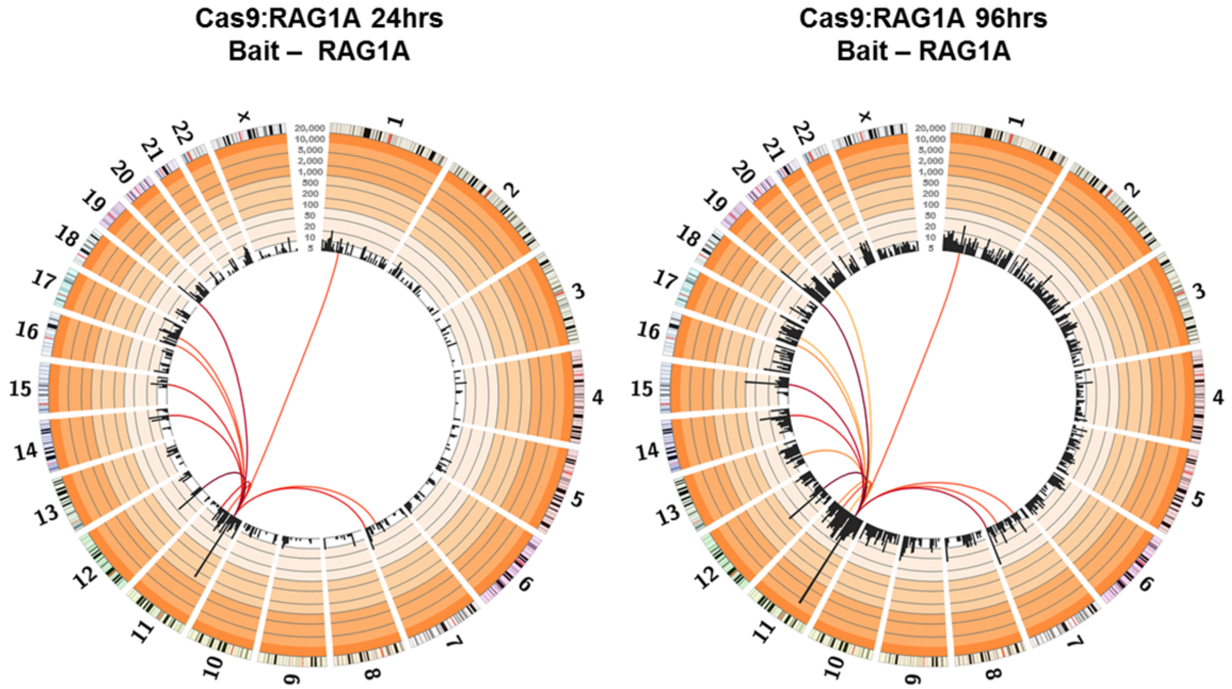


Supplementary Figure 3. HTGTS library junction compositions. (a) Filtered DNA sequence reads are composed of (from right to left) the bait primer (red arrow), the bait sequence (black rectangle) leading up to the junction (yellow box), followed by the aligned prey sequence (green rectangle). (b,c) Bait sequence junction distribution corresponds to predicted break-site position for Cas9:RAG1A-D libraries displayed as either (b) total junctions or (c) percent junctions (black dashed lines; N=6 each). Due to the constraints of sequencing length for Miseq and, correspondingly, primer positions from the bait DSB, bait junctions predominately display short resections (up to 25bp for most junctions) with some displaying longer resections leading up to the primer position; bait resections beyond the primer position are not detected with this method. However, this is not true for prey junctions, and for those which map to the break-site (See Supplementary Figure 7), we observe that >99% of resection events are within 15kb with an average of around 2-3 kb while less than 1% extends for up to 250-300 kb. (d-g) Junctions from HTGTS libraries are primarily composed of short micro-homologies and direct joins with a smaller proportion of junctions containing base pair insertions of variable lengths.



Supplementary Figure 4. Hotspots are related to recurrent off-target nuclease activity. (a) Diagram of bait 3′ DSB ends translocating to a separate broken end resulting in direct and resected joins in both orientations. (b) Guide RNA sequence preference extrapolated from RAG1A OT sites. (c-e) Prey joining distribution to select OT sites from (c,d) Cas9:RAG1A or (e) Cas9:RAG1B bait DSBs. Plus and minus orientation junction enrichment are indicated by red and blue curves respectively. Dashed lines indicate predicted off-target cleavage site. Shaded area represents off-target sequence listed above each plot; red text indicates mismatch from targeted sequence and underlined nucleotides indicate the PAM.

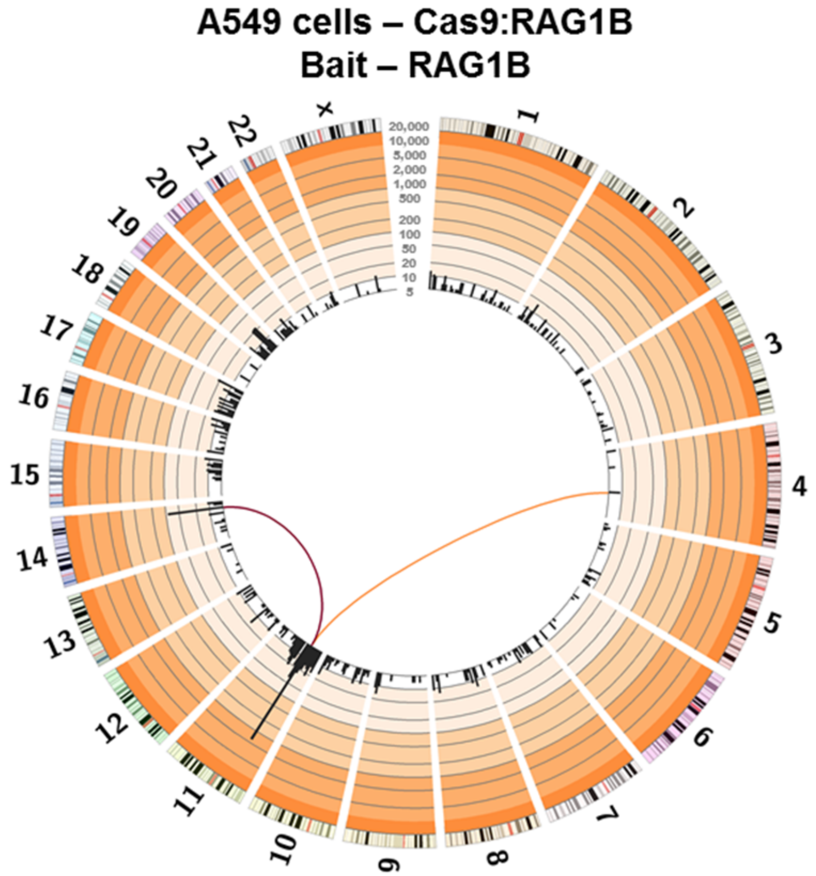
Supplementary Figure 5


Cas9:RAG1A 24hrs
Bait – RAG1A

Cas9:RAG1A 96hrs
Bait – RAG1A

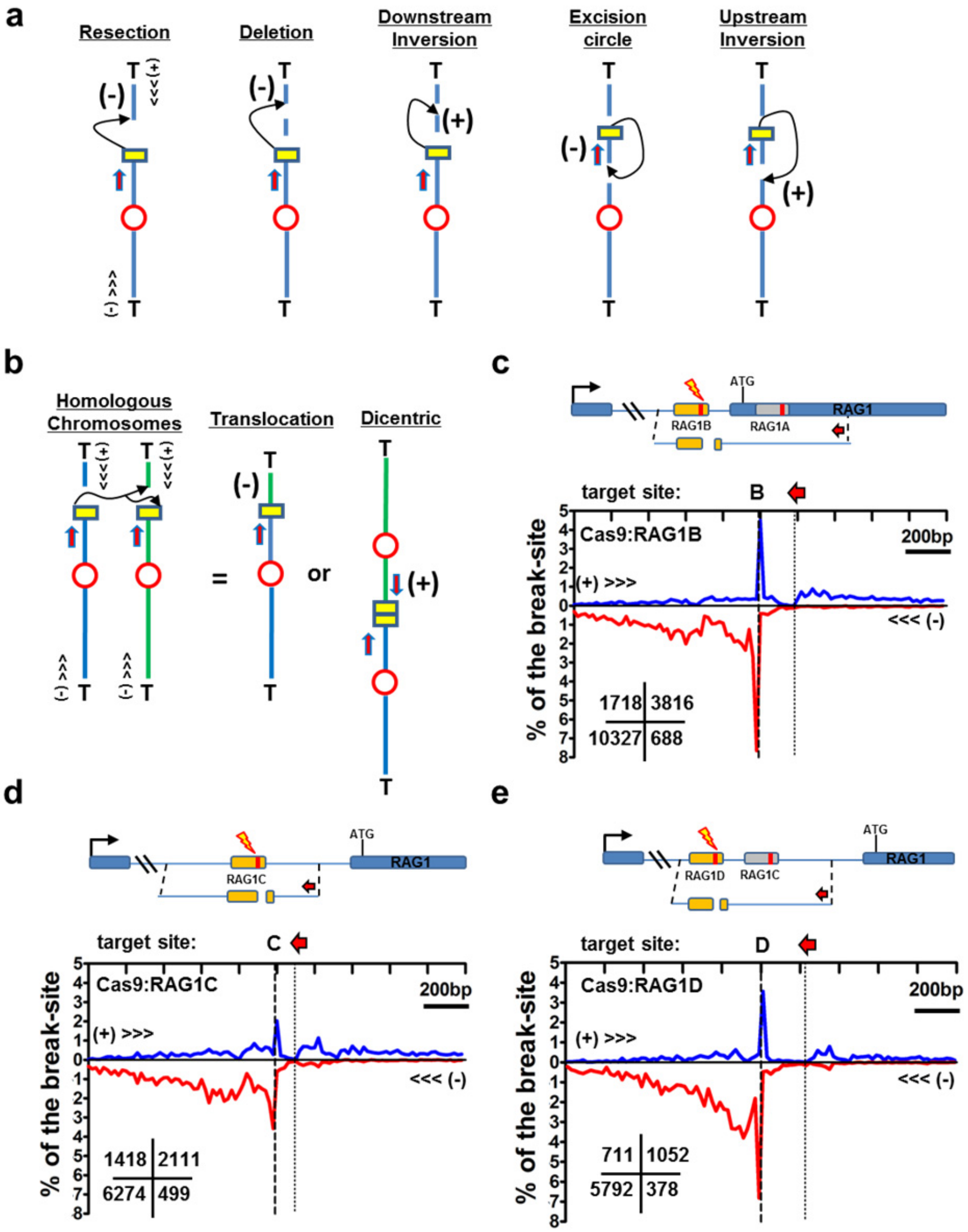
Supplementary Figure 5. Off-target detection with varied culturing times post-transfection. Cas9:RAG1A-transfected 293T cells were cultured for 24hrs and 96hrs prior to DNA collection. HTGTS libraries indicate identification of similar off-target sites with fewer off-targets detected at 24hrs. Library replicates: 24hrs: N=2; 96hrs: N=2. Circos plots are similarly presented as in Fig. 1 (see legend). Junction scale range: 5-20,000.

Supplementary Figure 6


A549 cells – Cas9:RAG1B
Bait – RAG1B

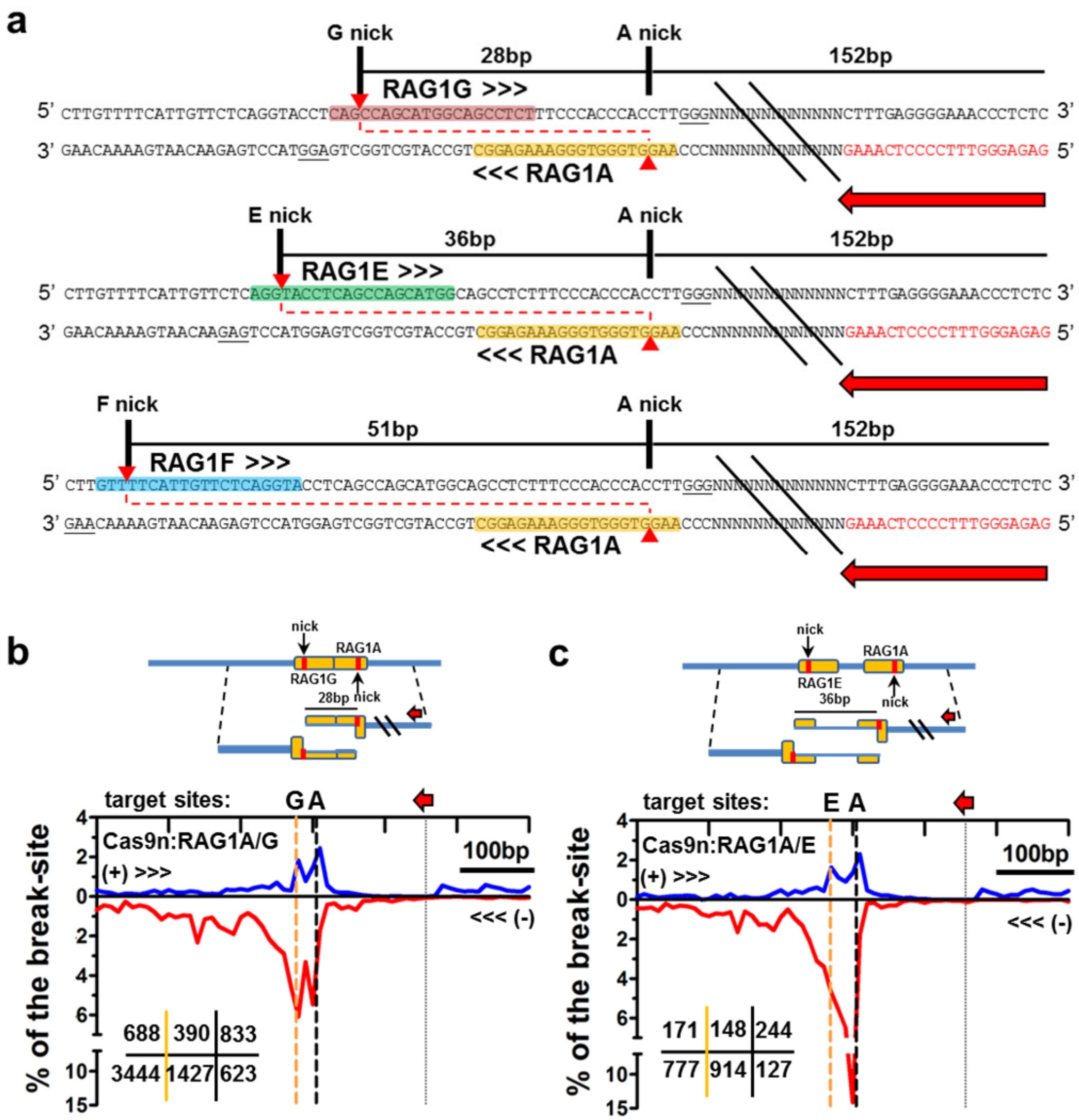
Supplementary Figure 6. Cas9:RAG1B bait HTGTS library from A549 cells. Circos plot with red to yellow lines (high to low hotspot enrichment) linking the RAG1B bait to the RAG1B off-target hotspots (N=2). Circos plots are presented as similarly described in Fig. 1 (see legend). Junction scale range: 5-20,000.

Supplementary Figure 7



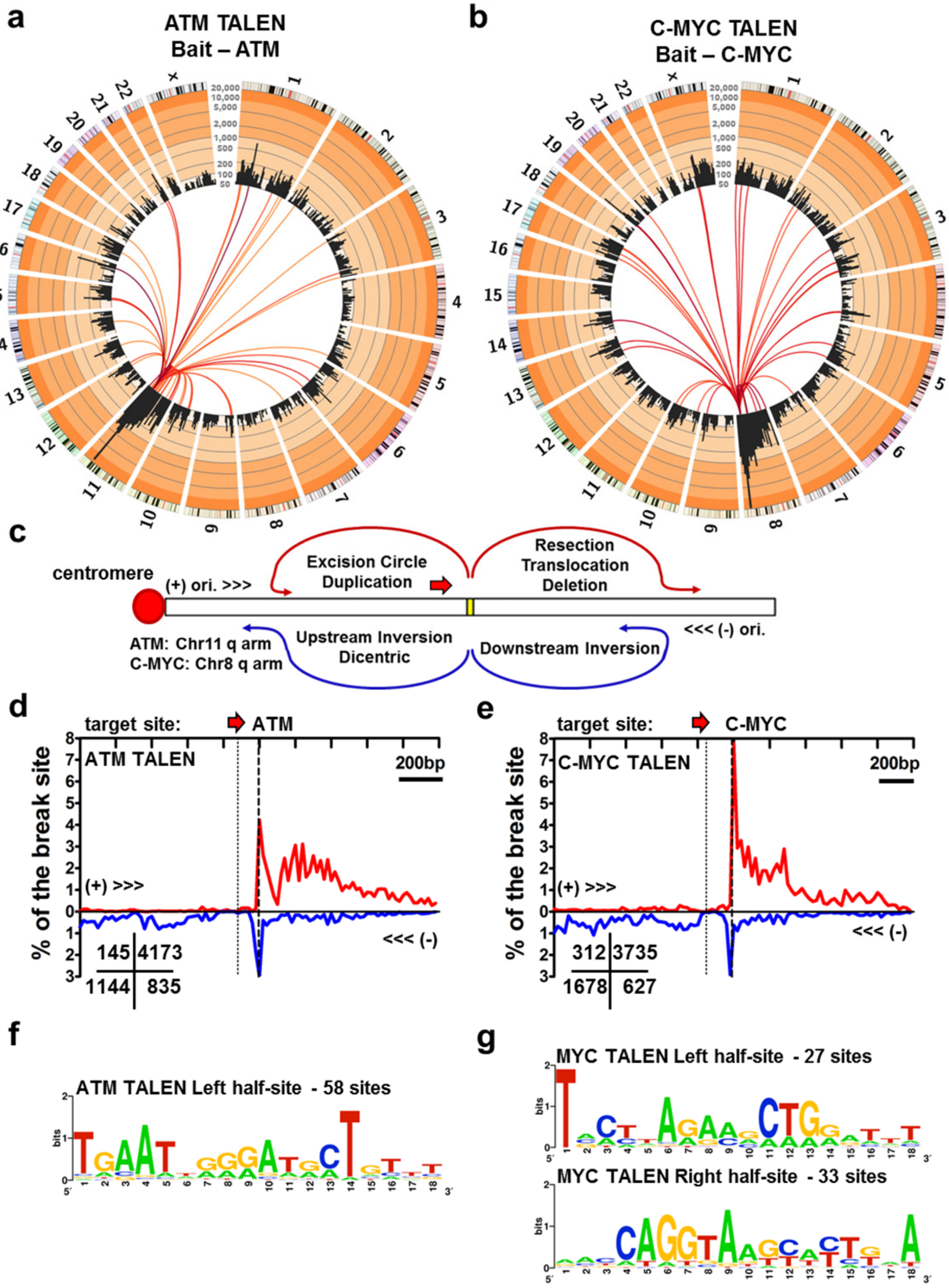
Supplementary Figure 7. Prey joining outcomes at the bait break-site. (a) From a single bait DSB, nearby prey junctions are primarily derived from rejoining the original, but resected, DSB. Joining to broken ends of nearby spontaneous DSBs can result in deletions, inversions and excision circles depending on the position of the 2nd DSB relative to the bait and which broken end of the 2nd DSB is joined. (b) Targeting endogenous loci typically include the homologous chromosome and the homologous target site. Thus, joins to the homologous chromosome can occur which can lead to the formation of dicentric chromosomes. (c-e) Break-site prey junction enrichment for (c) Cas9:RAG1B, (d) Cas9:RAG1C, and (e) Cas9:RAG1D (1kb regions flanking the break-site). Orange boxes indicate Cas9 targeted site for the respective gRNA. Red bars indicate cleavage site. Total junctions for each quadrant are listed (N=6 for each).

Supplementary Figure 8



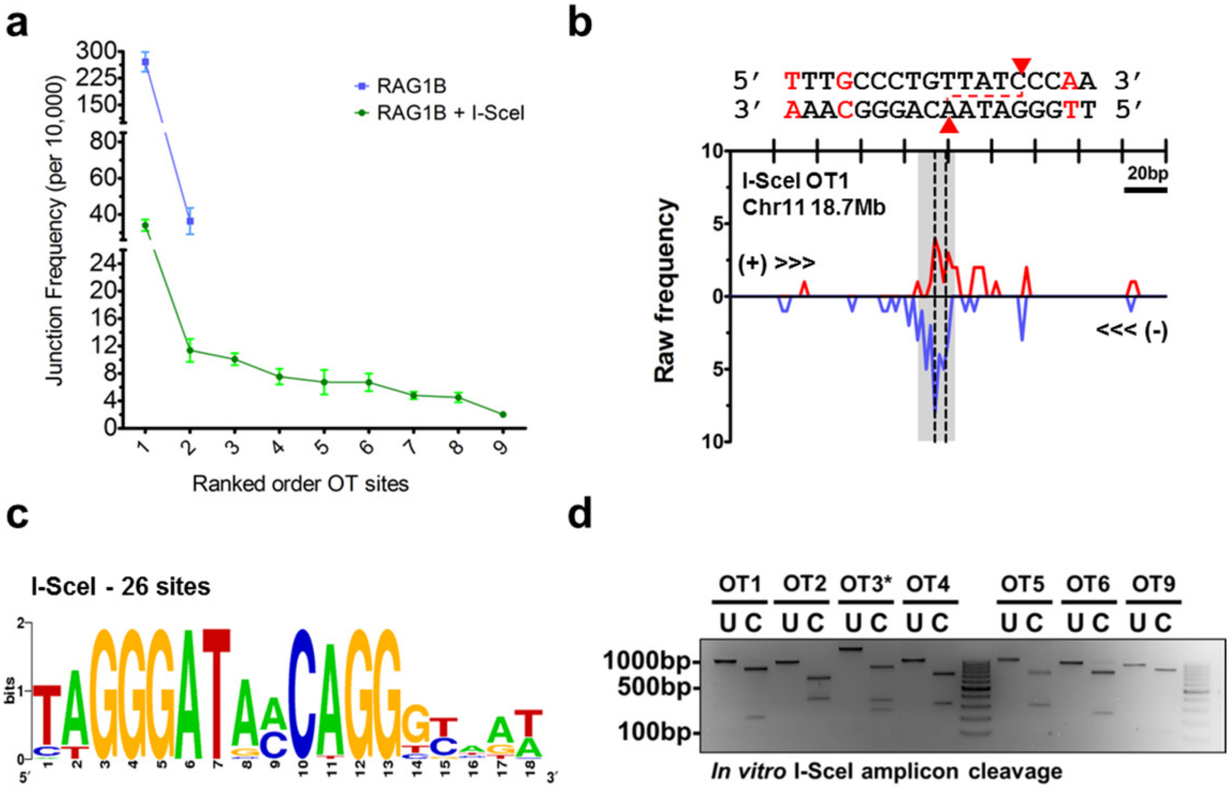
Supplementary Figure 8. Cas9n:RAG1A/G, A/E, A/F paired nickase targets. (a) Sequences specific for RAG1A, G, E, and F gRNAs (yellow, red, green and blue highlight respectively) with variable length 5' overhangs (dashed lines connecting red arrowheads; PAM is underlined). (b,c) Prey junction distribution at the break-site for (b) Cas9n:RAG1A/G (N=6) and (c) Cas9n:RAG1A/E (N=2). Orange dashed lines indicate the (b) RAG1G or (c) RAG1E site relative to the RAG1A site (black dashed line). Break-site prey junction totals within 0.3kb are listed for each section including prey junctions between off-set nick sites. Preliminary data of RAG1A single nicking experiments do not yield high numbers of HTGTS junctions and no HTGTS hotspots of any kind. However, we detect small numbers of junctions at the break-site, which are 2-fold greater than vector control, indicating that nicks indeed can be converted into DSBs. In these controls, there appears to be few if any translocations to the homologous chromosome. Future studies using a separate bait-site (using a different nuclease class) to capture single nick DSBs (co-expression assay) will yield more information into the mechanisms associated with Cas9n-related single nick-driven DSBs.

Supplementary Figure 9



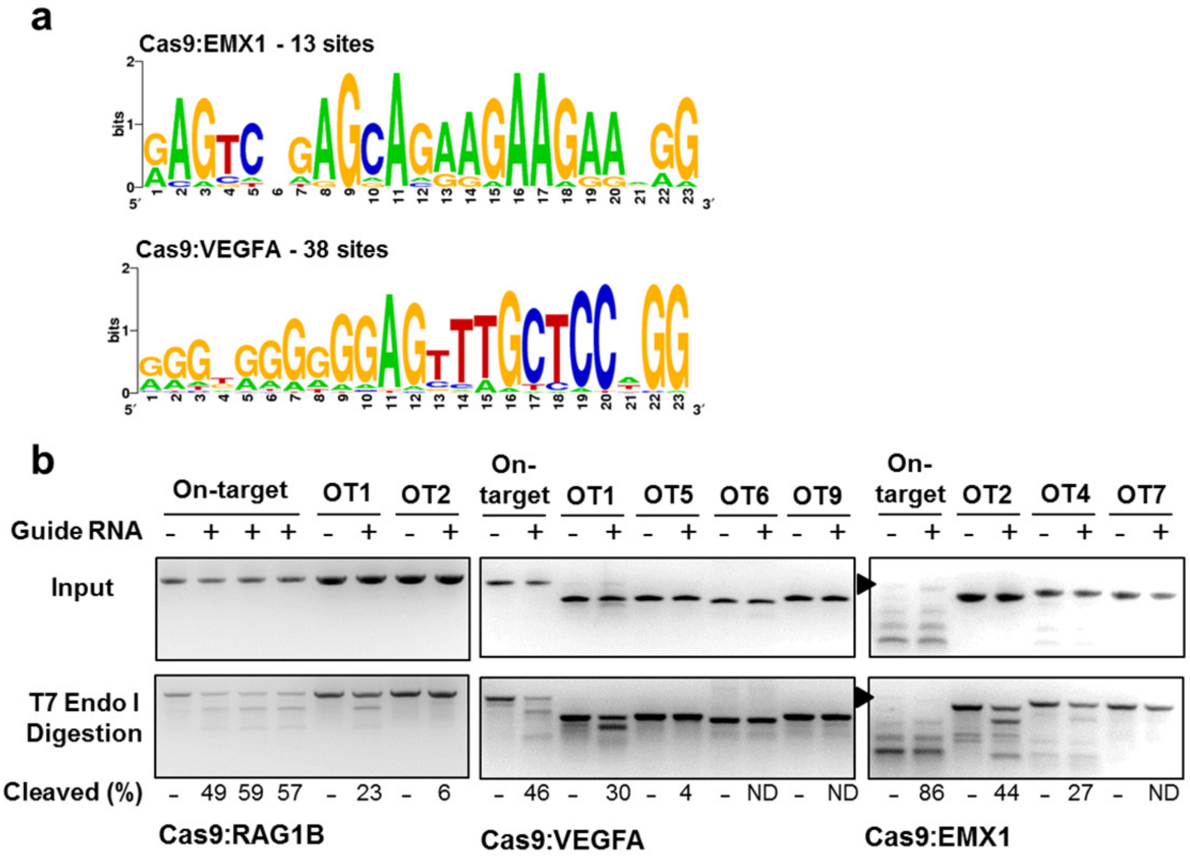
Supplementary Figure 9. ATM and c-MYC TALEN HTGTS libraries. (a,b) Circos plots of genome-wide prey junctions cloned from the (a) ATM TALEN (left circos) and the (b) c-MYC TALEN (right circos) on chromosomes 11 and 8 respectively. Circos plots are similarly presented as in Fig. 1 (see legend). Junction scale range: 50-20,000. Interior colored lines linking the bait to prey hotspots (gradient: dark red to yellow = high to low enrichment) display the top ten percent of off-target sites identified for each TALEN library; N=3 each. (c) Cloning from 5' bait DSBs in TALEN libraries and outcomes associated with local joining (see Supplementary Figure 7a,b). (d,e) Relative to the primer (red arrow), downstream prey junctions beyond the break-site (middle dashed line) predominantly represent resections and deletions. Dicentric-forming upstream inversions at the break-site are enriched. Numbers in the lower left corner are total frequencies for each quadrant (N=3 each). (f) ATM and (g) C-MYC TALEN sequence preference derived from OT half-sites (Supplementary Tables 4 & 5). No ATM TALEN right half-site analysis was performed since only two sites were described.

Supplementary Figure 10



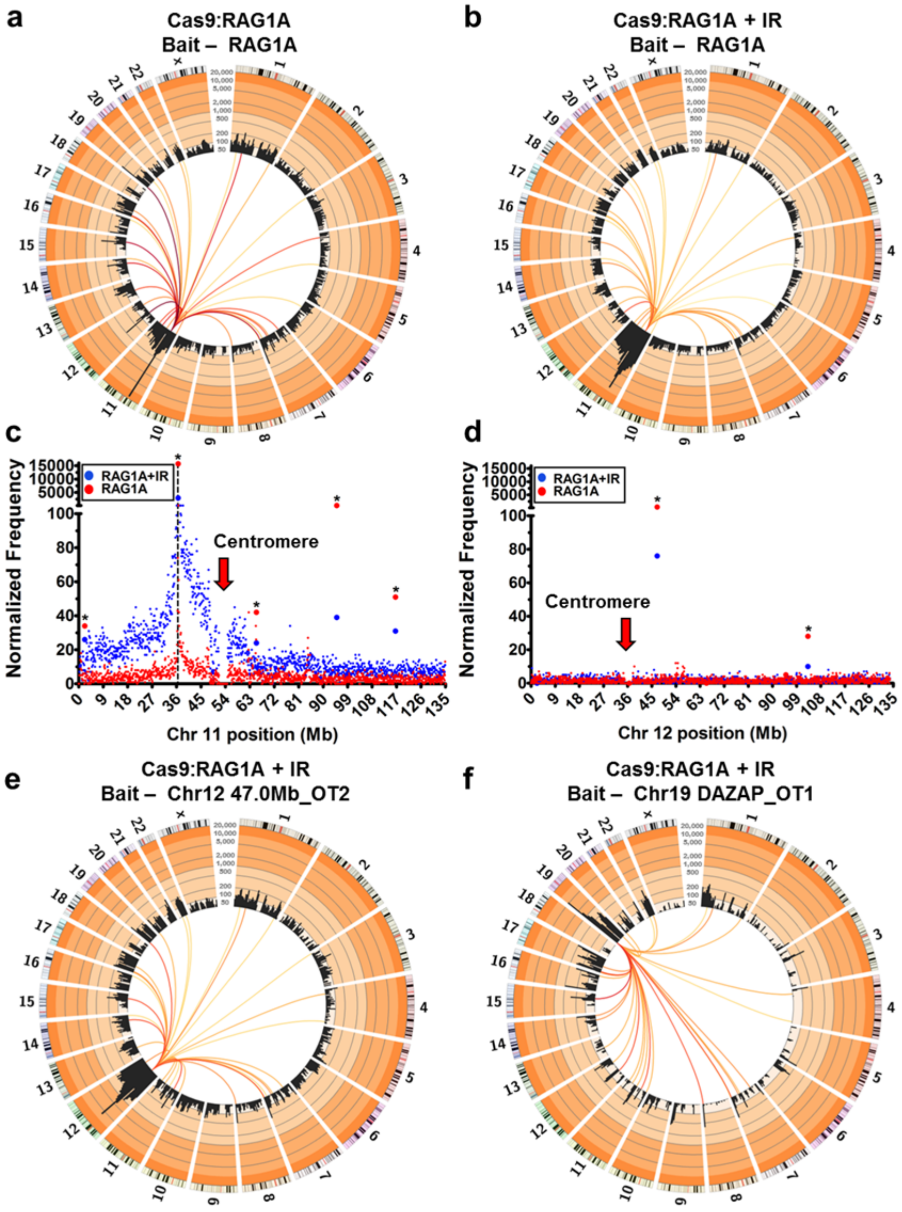
Supplementary Figure 10. Analysis of I-SceI OT sites recovered by Cas9:RAG1B bait DSBs. (a) Relative junction frequency of RAG1B (N=6; blue) and (b) I-SceI (N=3; green) OTs. Error bars are displayed as S.E.M.. (b) Select I-SceI OT site showing focal prey junction enrichment. Plus and minus orientation junction enrichment are indicated by red and blue curves respectively. Vertical dashed lines indicate predicted off-target cleavage site. Shaded area represents off-target sequence listed above the plot; red text indicates mismatch from targeted sequence. (c) Nuclease sequence preference derived from both human and mouse[23] I-SceI OT sites. (d) I-SceI off-target sites were PCR amplified and either mock digested (U) or digested with I-SceI (C) with most sites efficiently cleaved. Asterisk indicates site with 2 identical off-target sequences.

Supplementary Figure 11



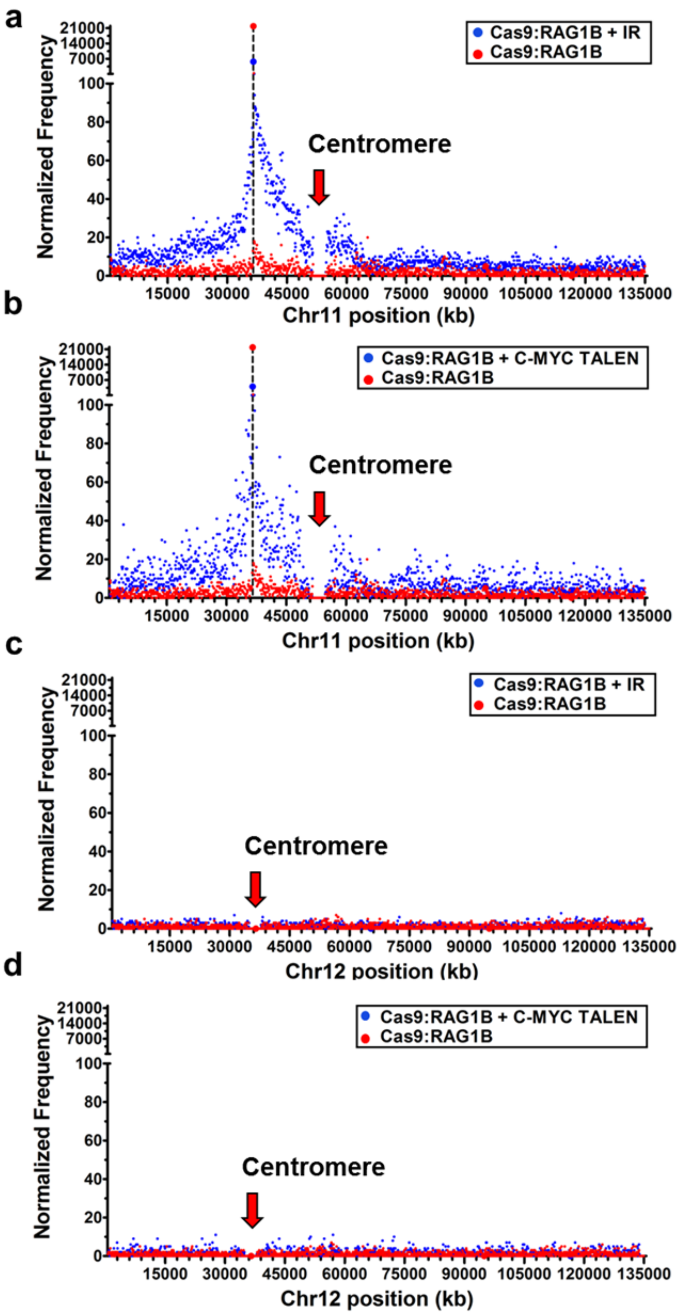
Supplementary Figure 11. Sequence analysis and in vitro cleavage assay for Cas9:EMX1 and Cas9:VEGFA. (a) Nuclease sequence preference derived from Cas9:EMX1 and Cas9:VEGFA OT sites. (b) T7 Endonuclease I cleavage assay of RAG1B, VEGFA, and EMX1 gRNA on- and off-target sites. 293T genomic DNA sources were from non-transfected, Cas9:RAG1B alone, or in combination with VEGFA or EMX1 gRNAs. For Cas9:EMX on-target, black triangles indicate correct germline band. See online methods for measuring amplicon cleavage.

Supplementary Figure 12



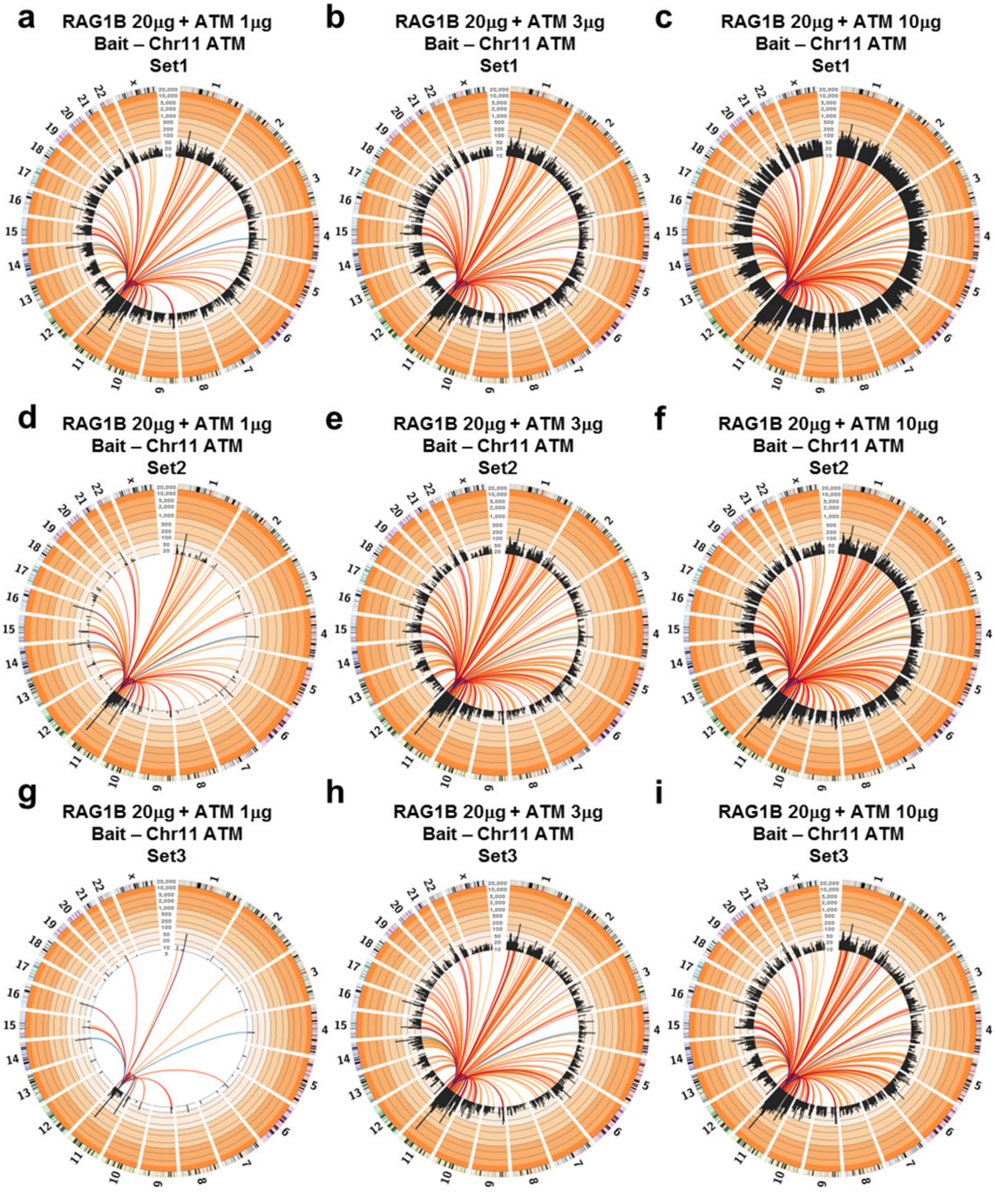
Supplementary Figure 12. IR-treatment increases break-site chromosome junction enrichment. (a,b) Circos plot comparison of (a) Cas9:RAG1A and (b) Cas9:RAG1A + IR libraries normalized to 75,000 junctions (N=6 and N=3 respectively). Note the change in dark red colored lines to lighter orange/yellow colored lines indicating a decrease in hotspot junction enrichment. (c,d) Dot plot of Cas9:RAG1A (red) and Cas9:RAG1A + IR (blue) (c) across the break-site chromosome 11 and (d) across a similar sized chromosome 12. Dot scale: 100kb. Enlarged dots with asterisks signify RAG1A OT sites or the on-target site (at the dashed line). (e,f) Circos plots of Cas9:RAG1A + IR OT bait HTGTS libraries from (e) chromosome 12 and (f) chromosome 19 (N=3 each). Maximum ranges for OT sites were derived from RAG1A chromosome 12 and chromosome 19 bait OT libraries respectively. Circos plots are similarly presented as in Fig. 1. Junction scale: 50-20,000.

Supplementary Figure 13



Supplementary Figure 13. Cas9:RAG1B bait wide-spread, low level DSB activity detection assay. (a-d) Dot plot of break-site chromosome junctions from (a,b) RAG1B (N=6), (a) RAG1B + IR (N=3), and (b) RAG1B + C-MYC TALEN (N=3) HTGTS libraries using RAG1B as bait. A similar sized chromosome (12) is also displayed for (c,d) RAG1B, (c) RAG1B + IR, and (d) RAG1B + C-MYC TALEN. Dot scale: 100kb. Dashed line indicates break-site with enlarged dots representing the 100kb region of the break-site. For all HTGTS libraries shown in the figure, junctions within C-MYC OT sites (200 chromosome 11 sites, mean size/site: 990bp) were removed prior to normalization at 59,830 junctions.

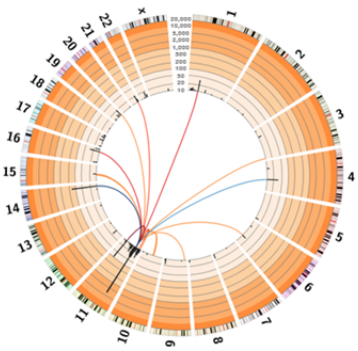
Supplementary Figure 14



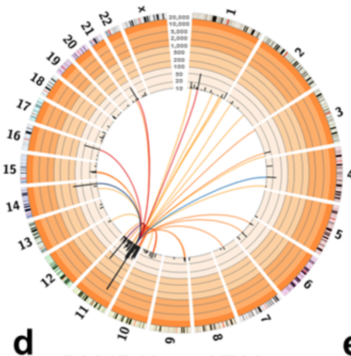
Supplementary Figure 14. ATM TALEN bait HTGTS libraries with increasing concentrations of transfected ATM TALEN plasmids. (a-i) Circos plots displaying additional replicates of titrated ATM TALEN bait plasmids at (a,d,g) 1μg, (b,e,h) 3μg, and (c,f,i) 10μg each along with 20μg of Cas9:RAG1B prey plasmid as (a-c) the 1st set, (d-f) 2nd , and (g-i) 3rd replicate sets. The blue colored lines connect the ATM bait to RAG1B OT prey hotspots, and the red to yellow colored lines connect ATM bait to ATM OT prey hotspots. Circos plots are presented as similarly described in Fig. 1. Junction scale: set1:10-20,000; set 2: 20-20,000; set 3: 5- or 10-20,000.

Supplementary Figure 15

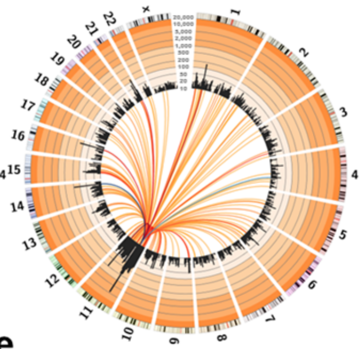
a RAG1B 20μg + ATM 1μg
Bait – Chr11 RAG1B
Set2



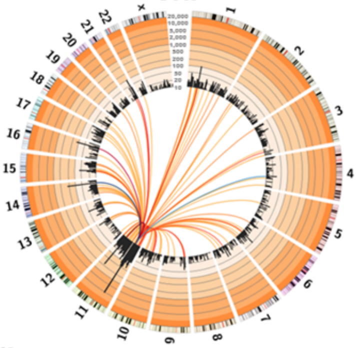
b RAG1B 20μg + ATM 3μg
Bait – Chr11 RAG1B
Set2



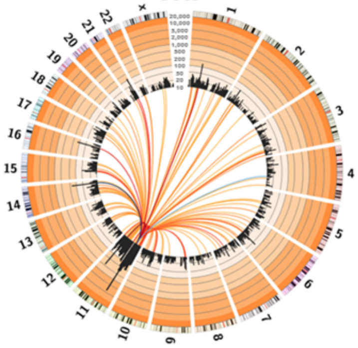
c RAG1B 20μg + ATM 10μg
Bait – Chr11 RAG1B
Set2



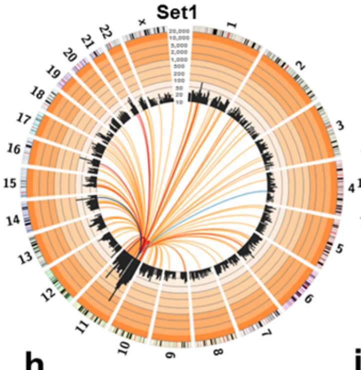
d RAG1B 20μg + ATM 3μg
Bait – Chr11 RAG1B
Set3



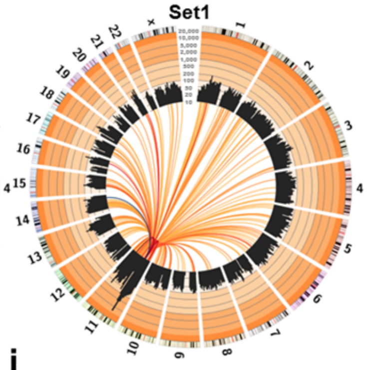
e RAG1B 20μg + ATM 10μg
Bait – Chr11 RAG1B
Set3



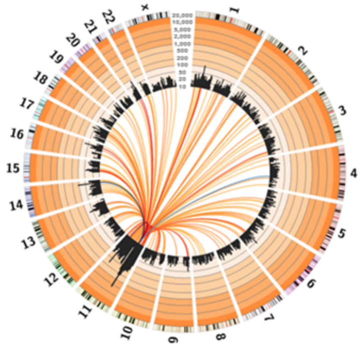
f RAG1B 20μg + ATM 20μg
Bait – Chr11 RAG1B
Set1



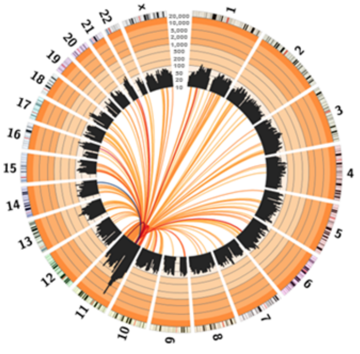
g RAG1B 20μg + ATM 100μg
Bait – Chr11 RAG1B
Set1



h RAG1B 20μg + ATM 20μg
Bait – Chr11 RAG1B
Set2

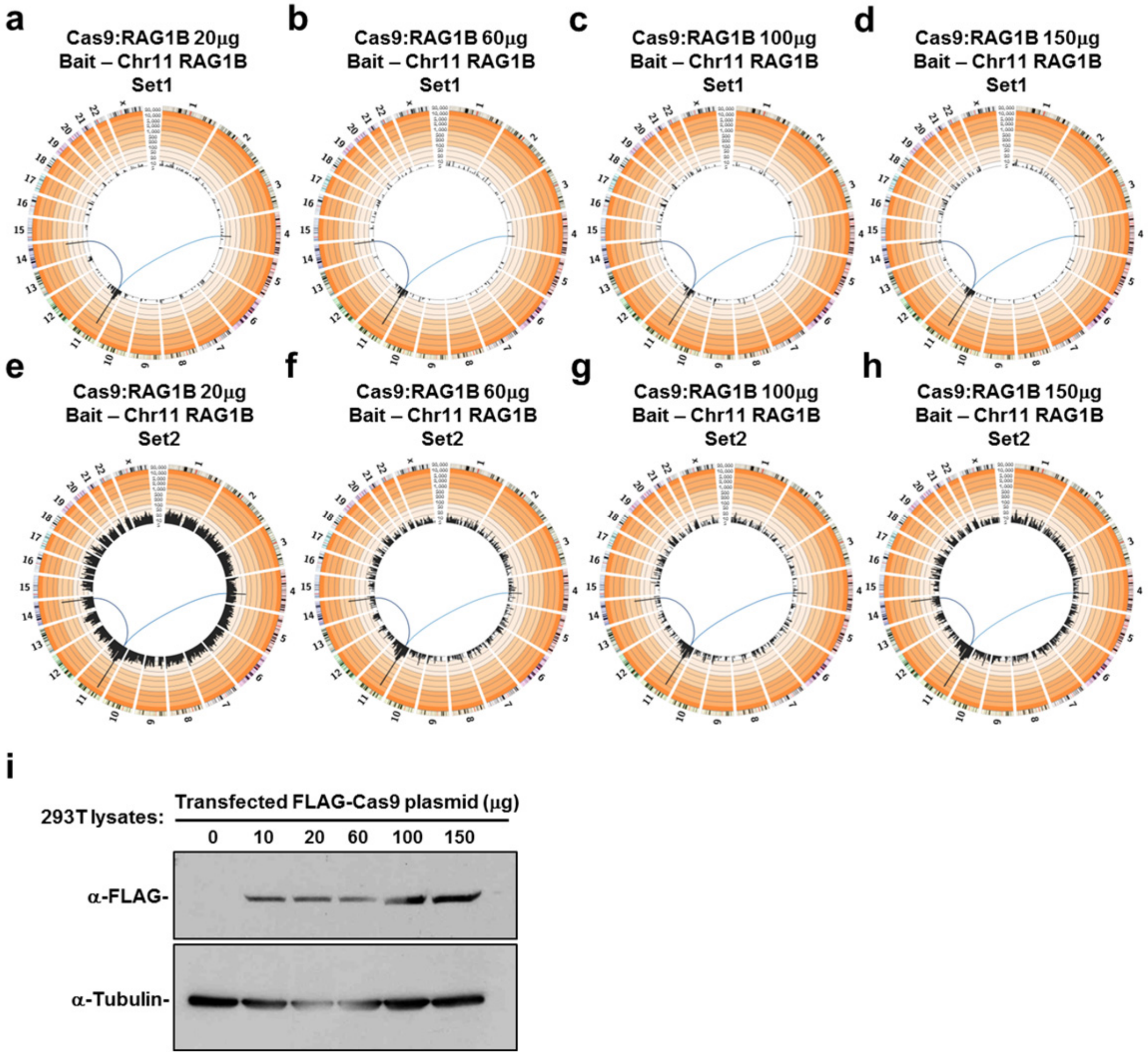


i RAG1B 20μg + ATM 100μg
Bait – Chr11 RAG1B
Set2



Supplementary Figure 15. RAG1B bait ATM TALEN titration HTGTS libraries. Circos plots showing additional biological replicate sets of ATM titration RAG1B bait libraries for (a-e) 1-10μg of transfected ATM TALEN plasmids from Figure 5, and two replicate sets of RAG1B bait libraries with 20 and 100μg of transfected ATM TALEN plasmids. Blue colored lines link the RAG1B bait to RAG1B OT prey hotspots, and red to yellow colored lines link the RAG1B bait to ATM TALEN prey hotspots. Circos plots are presented as similarly described as in Fig. 1 (see legend). Junction scale: 10-20,000.

Supplementary Figure 16


a Cas9:RAG1B 20μg Bait – Chr11 RAG1B Set1

b Cas9:RAG1B 60μg Bait – Chr11 RAG1B Set1

c Cas9:RAG1B 100μg Bait – Chr11 RAG1B Set1

d Cas9:RAG1B 150μg Bait – Chr11 RAG1B Set1

e Cas9:RAG1B 20μg Bait – Chr11 RAG1B Set2

f Cas9:RAG1B 60μg Bait – Chr11 RAG1B Set2

g Cas9:RAG1B 100μg Bait – Chr11 RAG1B Set2

h Cas9:RAG1B 150μg Bait – Chr11 RAG1B Set2

i

Supplementary Figure 16. Increased amounts of Cas9:RAG1B plasmid does not greatly increase non-specific DSB activity. (a-h) Increased amounts of transfected Cas9:RAG1B plasmid DNA at (a,e) 20μg, (b,f) 60μg, (c,g) 100μg, and (d,h) 150μg Cas9:RAG1B plasmid for (a-d) set1 and (e-h) set 2 HTGTS libraries. (i) Western analysis probing for FLAG-tagged Cas9 and a loading control Tubulin from 293T extracts transfected with increased amounts of Cas9:RAG1B plasmid DNA. Circos plots presented similarly as in Fig. 1 (see legend). Junction scale: 5-20,000.

Supplementary Table 1. Cas9 and TALEN target sites used to generate HTGTS libraries.

Bait: Cas9 endonuclease

Exp	Nuclease/Locus	GFP	Total DNA	Method	Library	Total Bait aligned	Filtered Junctions
#1	Cas9:RAG1A	81%	100μg	EM-PCR	1	1,385,548	5,373
#2	Cas9:RAG1A	45%	100μg	EM-PCR	1	1,188,505	2,846
#3	Cas9:RAG1A	76%	100μg	EM-PCR	1	837,126	2,836
#4	Cas9:RAG1A	89%	100μg	EM-PCR	1	1,123,900	3,619
#5	Cas9:RAG1A	90%	50μg	LAM-PCR	1	1,523,373	33,122
#6	Cas9:RAG1A	91%	50μg	LAM-PCR	1	1,640,590	42,911
TOTAL			**500μg**		**6**	**7,699,042**	**90,707**
#1	Cas9:RAG1A – 24hrs	42%	50μg	LAM-PCR	1	1,407,635	2,501
#2	Cas9:RAG1A – 24hrs	46%	50μg	LAM-PCR	1	646,784	1,898
TOTAL			**100μg**		**2**	**2,054,419**	**4,399**
#1	Cas9:RAG1A – 96hrs	62%	50μg	LAM-PCR	1	1,551,789	9,550
#2	Cas9:RAG1A – 96hrs	72%	50μg	LAM-PCR	1	811,246	4,045
TOTAL			**100μg**		**2**	**2,363,035**	**13,595**
#1	Cas9:RAG1A+IR	86%	50μg	LAM-PCR	1	777,941	62,344
#2	Cas9:RAG1A+IR	85%	50μg	LAM-PCR	1	949,984	83,302
#3	Cas9:RAG1A+IR	86%	50μg	LAM-PCR	1	765,342	60,525
TOTAL			**150μg**		**3**	**2,493,267**	**206,171**
#1	Cas9:RAG1A OT Chr19	81%	50μg	LAM-PCR	1	1,433,690	3,483
#2	Cas9:RAG1A OT Chr19	90%	50μg	LAM-PCR	1	1,343,499	10,878
#3	Cas9:RAG1A OT Chr19	91%	50μg	LAM-PCR	1	1,275,773	8,771
TOTAL			**150μg**		**3**	**4,052,962**	**23,132**
#1	Cas9:RAG1 A OT Chr19+IR	86%	50μg	LAM-PCR	1	921,997	15,934
#2	Cas9:RAG1 A OT Chr19+IR	85%	50μg	LAM-PCR	1	1,214,518	22,406
#3	Cas9:RAG1 A OT Chr19+IR	86%	50μg	LAM-PCR	1	1,357,751	21,071
TOTAL			**150μg**		**3**	**3,494,266**	**59,411**
#1	Cas9:RAG1A OT Chr12	81%	50μg	LAM-PCR	1	1,252,516	746
#2	Cas9:RAG1A OT Chr12	90%	50μg	LAM-PCR	1	1,235,424	13,702
#3	Cas9:RAG1A OT Chr12	91%	50μg	LAM-PCR	1	1,134,490	2,711
TOTAL			**150μg**		**3**	**3,622,430**	**17,158**
#1	Cas9:RAG1A OT Chr12+IR	86%	50μg	LAM-PCR	1	1,183,240	25,223
#2	Cas9:RAG1A OT Chr12+IR	85%	50μg	LAM-PCR	1	1,105,648	29,601
#3	Cas9:RAG1A OT Chr12+IR	86%	50μg	LAM-PCR	1	1,358,359	32,528
TOTAL			**150μg**		**3**	**3,647,247**	**87,352**
#1	Cas9:RAG1A OT Chr7	90%	50μg	LAM-PCR	1	1,324,361	3,077
#2	Cas9:RAG1A OT Chr7	80%	50μg	LAM-PCR	1	581,954	1,389
#3	Cas9:RAG1A OT Chr7	87%	50μg	LAM-PCR	1	1,422,336	3,777
TOTAL			**150μg**		**3**	**3,328,651**	**8,243**
#1	Cas9:RAG1B	83%	100μg	EM-PCR	1	1,241,696	4,801
#2	Cas9:RAG1B	48%	100μg	EM-PCR	1	1,235,718	2,044
#3	Cas9:RAG1B	66%	100μg	EM-PCR	1	798,151	1,693
#4	Cas9:RAG1B	87%	100μg	EM-PCR	1	883,814	1,554
#5	Cas9:RAG1B	90%	50μg	LAM-PCR	1	1,493,069	21,720
#6	Cas9:RAG1B	90%	50μg	LAM-PCR	1	1,478,753	28,036
TOTAL			**500μg**		**6**	**7,131,201**	**59,848**
#1	Cas9:RAG1B – A549	38%	200μg	LAM-PCR	2	2,308,874	2,489
#2	Cas9:RAG1B – A549	51%	100μg	LAM-PCR	1	1,351,727	2,540
TOTAL			**300μg**		**3**	**3,660,601**	**5,029**
#1	Cas9:RAG1B 60μg	68%	50μg	LAM-PCR	1	891,902	11,593
#2	Cas9:RAG1B 60μg	66%	50μg	LAM-PCR	1	552,083	4,767
TOTAL			**100μg**		**2**	**1,443,985**	**16,360**

#1	Cas9:RAG1B 100μg	65%	50μg	LAM-PCR	1	895,365	9,819
#2	Cas9:RAG1B 100μg	60%	50μg	LAM-PCR	1	1,113,910	5,350
TOTAL			**100μg**		**2**	**2,009,275**	**15,169**
#1	Cas9:RAG1B 150μg	59%	50μg	LAM-PCR	1	1,064,148	14,551
#2	Cas9:RAG1B 150μg	52%	50μg	LAM-PCR	1	1,104,544	5,840
TOTAL			**100μg**		**2**	**2,168,692**	**20,391**
#1	Cas9:RAG1B+IR	94%	50μg	LAM-PCR	1	1,110,892	47,639
#2	Cas9:RAG1B+IR	96%	50μg	LAM-PCR	1	1,140,162	39,555
#3	Cas9:RAG1B+IR	94%	50μg	LAM-PCR	1	1,188,273	46,621
TOTAL			**150μg**		**3**	**3,439,327**	**133,815**
#1	Cas9:RAG1B +IScel	57%	50μg	LAM-PCR	1	988,054	15,978
#2	Cas9:RAG1B +IScel	54%	50μg	LAM-PCR	1	1,163,969	10,676
#3	Cas9:RAG1B +IScel	58%	50μg	LAM-PCR	1	976,679	9,900
TOTAL			**150μg**		**3**	**3,128,702**	**36,554**
#1	Cas9:RAG1B +Cas9:EMX1	92%	50μg	LAM-PCR	1	1,358,175	19,474
#2	Cas9:RAG1B +Cas9:EMX1	87%	50μg	LAM-PCR	1	940,919	16,730
#3	Cas9:RAG1B +Cas9:EMX1	88%	50μg	LAM-PCR	1	699,736	714
TOTAL			**150μg**		**3**	**2,998,830**	**36,918**
#1	Cas9:RAG1B +Cas9:VEGFA	88%	50μg	LAM-PCR	1	1,265,815	1,996
#2	Cas9:RAG1B +Cas9:VEGFA	92%	50μg	LAM-PCR	1	769,647	14,352
#3	Cas9:RAG1B +Cas9:VEGFA	89%	50μg	LAM-PCR	1	712,697	12,103
TOTAL			**150μg**		**3**	**2,748,159**	**28,451**
#1	Cas9:RAG1B +ATM TALEN	88%	50μg	LAM-PCR	1	1,429,068	22,426
#2	Cas9:RAG1B +ATM TALEN	77%	50μg	LAM-PCR	1	823,007	24,473
#3	Cas9:RAG1B +ATM TALEN	78%	50μg	LAM-PCR	1	588,161	10,200
TOTAL			**150μg**		**3**	**2,840,236**	**57,099**
#1	Cas9:RAG1B +MYC TALEN	88%	50μg	LAM-PCR	1	1,723,351	40,995
#2	Cas9:RAG1B +MYC TALEN	77%	50μg	LAM-PCR	1	1,006,825	34,443
#3	Cas9:RAG1B +MYC TALEN	75%	50μg	LAM-PCR	1	1,027,535	35,514
TOTAL			**150μg**		**3**	**3,757,711**	**110,952**
#1	Cas9:RAG1C	85%	100μg	EM-PCR	1	1,434,562	1,878
#2	Cas9:RAG1C	51%	100μg	EM-PCR	1	1,238,636	1,512
#3	Cas9:RAG1C	58%	100μg	EM-PCR	1	743,963	407
#4	Cas9:RAG1C	89%	100μg	EM-PCR	1	814,688	644
#5	Cas9:RAG1C	85%	50μg	LAM-PCR	1	1,712,220	20,213
#6	Cas9:RAG1C	85%	50μg	LAM-PCR	1	1,920,004	25,046
TOTAL			**500μg**		**6**	**7,864,073**	**49,700**
#1	Cas9:RAG1D	72%	100μg	EM-PCR	1	1,498,394	729
#2	Cas9:RAG1D	56%	100μg	EM-PCR	1	1,458,917	784
#3	Cas9:RAG1D	64%	100μg	EM-PCR	1	822,563	132
#4	Cas9:RAG1D	89%	100μg	EM-PCR	1	825,052	184
#5	Cas9:RAG1D	87%	50μg	LAM-PCR	1	1,875,731	9,826
#6	Cas9:RAG1D	87%	50μg	LAM-PCR	1	1,595,863	9,516
TOTAL			**500μg**		**6**	**8,076,520**	**21,171**
#1	Cas9:RAG1B +ATM TALEN 100μg	75%	50μg	LAM-PCR	1	1,757,242	54,960
#2	Cas9:RAG1B +ATM TALEN 100μg	77%	50μg	LAM-PCR	1	1,685,816	63,678
TOTAL			**100μg**		**2**	**3,443,058**	**118,638**
#1	Cas9:RAG1B +ATM TALEN 10μg	96%	50μg	LAM-PCR	1	655,488	20,436
#2	Cas9:RAG1B +ATM TALEN 10μg	86%	50μg	LAM-PCR	1	1,476,240	48,912
	Cas9:RAG1B +ATM TALEN 10μg	87%	50μg	LAM-PCR	1	1,485,080	18,331
TOTAL			**150μg**		**3**	**3,626,808**	**87,679**
#1	Cas9:RAG1B +ATM TALEN 3μg	94%	50μg	LAM-PCR	1	760,330	8,122
#2	Cas9:RAG1B +ATM TALEN 3μg	87%	50μg	LAM-PCR	1	1,461,868	41,457
	Cas9:RAG1B +ATM TALEN 3μg	83%	50μg	LAM-PCR	1	1,293,894	17,930
TOTAL			**150μg**		**3**	**3,516,092**	**67,509**

#1	Cas9:RAG1B +ATM TALEN 1μg	92%	50μg	LAM-PCR	1	678,188	7,418
#2	Cas9:RAG1B +ATM TALEN 1μg	77%	50μg	LAM-PCR	1	1,399,326	24,120
TOTAL			**100μg**		**2**	**2,077,514**	**31,538**

Bait: Cas9 D10A nickase

Exp	Nuclease/Locus	GFP	Total DNA	Method	Library	Total Bait aligned	Filtered Junctions
#1	Cas9n:RAG1A/G	86%	100μg	EM-PCR	1	1,225,465	2,496
#2	Cas9n:RAG1A/G	88%	100μg	EM-PCR	1	1,489,939	2,018
#3	Cas9n:RAG1A/G	86%	100μg	EM-PCR	1	1,514,264	1,203
#4	Cas9n:RAG1A/G	91%	50μg	LAM-PCR	1	1,272,060	13,937
#5	Cas9n:RAG1A/G	90%	50μg	LAM-PCR	1	1,310,951	13,246
#6	Cas9n:RAG1A/G	92%	50μg	LAM-PCR	1	1,131,288	12,794
TOTAL			**450μg**		**6**	**7,943,967**	**45,694**
#1	Cas9n:RAG1A/E	89%	100μg	EM-PCR	1	1,409,177	3,826
#2	Cas9n:RAG1A/E	86%	50μg	LAM-PCR	1	1,078,141	13,796
TOTAL			**150μg**		**2**	**2,487318**	**17,622**
#1	Cas9:RAG1A/F	87%	100μg	EM-PCR	1	1,555,687	1,788
#2	Cas9:RAG1A/F	87%	50μg	LAM-PCR	1	1,196,599	10,502
TOTAL			**150μg**		**2**	**2,752,286**	**12,290**

Bait: TALEN endonuclease

Exp	Nuclease/Locus	GFP	Total DNA	Method	Library	Total Bait aligned	Filtered Junctions
#1	ATM TALEN	85%	50μg	LAM-PCR	1	1,629,352	22,844
#2	ATM TALEN	86%	50μg	LAM-PCR	1	440,065	10,760
#3	ATM TALEN	85%	50μg	LAM-PCR	1	1,058,993	35,553
TOTAL			**150μg**		**3**	**3,128,410**	**69,157**
#1	c-MYC TALEN	87%	50μg	LAM-PCR	1	957,292	62,473
#2	c-MYC TALEN	97%	50μg	LAM-PCR	1	1,043,966	24,886
#3	c-MYC TALEN	95%	50μg	LAM-PCR	1	1,130,927	20,993
TOTAL			**150μg**		**3**	**3,132,185**	**108,352**
#1	ATM 1μg +C9B:RAG1B 20μg	92%	50μg	LAM-PCR	1	1,288,833	2,227
#2	ATM 1μg +C9B:RAG1B 20μg	81%	50μg	LAM-PCR	1	1,816,729	22,633
#3	ATM 1μg +C9B:RAG1B 20μg	77%	50μg	LAM-PCR	1	1,659,566	15,185
TOTAL			**150μg**		**3**	**4,765,128**	**40,045**
#1	ATM 3μg +C9B:RAG1B 20μg	94%	50μg	LAM-PCR	1	1,262,584	21,345
#2	ATM 3μg +C9B:RAG1B 20μg	83%	50μg	LAM-PCR	1	1,580,141	28,890
#3	ATM 3μg +C9B:RAG1B 20μg	87%	50μg	LAM-PCR	1	1,575,902	37,100
TOTAL			**150μg**		**3**	**4,418,627**	**87,335**
#1	ATM 10μg +C9B:RAG1B 20μg	96%	50μg	LAM-PCR	1	1,104,338	25,359
#2	ATM 10μg +C9B:RAG1B 20μg	87%	50μg	LAM-PCR	1	1,403,545	67,220
#3	ATM 10μg +C9B:RAG1B 20μg	86%	50μg	LAM-PCR	1	1,463,665	59,393
TOTAL			**150μg**		**3**	**3,971,548**	**151,972**

Supplementary Table 2. Estimated HTGTS background for various libraries

Target DNA	Vector control	Vector control	Cas9:RAG1A	Cas9:RAG1B	Cas9:RAG1C	Cas9:RAG1D
HTGTS cloning method	EM-PCR RAG1A/B	EM-PCR RAG1C/D	EM-PCR RAG1A/B	EM-PCR RAG1A/B	EM-PCR RAG1C/D	EM-PCR RAG1C/D
Input DNA	100μg	100μg	100μg	100μg	100μg	100μg
Control DNA	NA	NA	Mouse	Mouse	Mouse	Mouse
Target Junctions	40	38	3,600	1,541	636	182
Control Junctions	-	-	23	19	5	5
Background	**NA**	**NA**	**0.64%**	**1.23%**	**0.79%**	**2.75%**

Target DNA	Vector control	Vector control	Cas9:RAG1A	Cas9:RAG1B	Cas9:RAG1C	Cas9:RAG1D
HTGTS cloning method	LAM-PCR RAG1A/B	LAM-PCR RAG1C/D	LAM-PCR RAG1A/B	LAM-PCR RAG1A/B	LAM-PCR RAG1C/D	LAM-PCR RAG1C/D
Input DNA	50μg	50μg	50μg	50μg	50μg	50μg
Control DNA	NA	NA	Mouse	Mouse	Mouse	Mouse
Target Junctions	2,346	134	26,508	16,516	20,247	7,500
Control Junctions	-	-	191	113	206	94
Background	**NA**	**NA**	**0.72%**	**0.68%**	**1.02%**	**1.25%**

Supplementary Table 3. Cas9:RAG1A off-target site mismatches and translocations

Chromo-some	Nuclease/Locus	Off-target site sequence (NGG)	Mis-match	Cas9:RAG1A Frequency	Cas9n:RAG1A Frequency	OT1 Bait Freq.	OT2 Bait Freq.	OT17 Bait Freq.
Chr11	**Cas9:RAG1A**	**GCCTCTTTCCCACCCACCTT GGGNNN**	0	NA	NA	177.59	220.90	23.53
Chr19	***DAZAP1 OT1***	ACCCCTTCCCCACCTACCTT GGGTCG	4	77.18***	0.18	NA	89.65	51.47
Chr12	***47.0Mb OT2***	TCCTCCTCCCCACCCACCTT CAGACT	4	74.07***	0.26	121.84	NA	36.01
Chr7	155.9Mb OT3	CCCTCCTTCCCACCCACTTT GGGTGA	3	40.46***	0.09	105.13	130.56	172.11
Chr15	SIN3A OT4	TCCTCTCTCCCACCCACCTC CGGCTC	3	40.19***	0.20	137.20	174.13	37.91
Chr11	94.9Mb OT5	ACGTCTTCCCCACCCACCTG GGGCCT	4	24.84***	-	7.33	19.21	13.55
Chr14	78.5Mb OT6	ACTTCCTTCCCACCCACCTT CAGCCA	4	23.98***	0.09	54.95	88.34	34.48
Chr1	CACHD1 OT7	TCCTCTTCTCCACCCACCTC TGGTTT	4	22.62***	-	33.59	63.81	25.28
Chr17	6.8Mb OT8	GTCTCTTTCTCACCCACTTT GGGTTG	3	19.14***	-	68.77	78.62	29.64
Chr7	TARP OT9	CACCCTTTCCCATCCACCTT TGGATA	4	17.57***	-	23.86	29.82	57.59
Chr7	112.9Mb OT10	GCCTCTTCCACACCCACCCT GGGCCC	3	14.90***	-	33.84	88.24	270.44
Chr13	RB1 OT11	GCCTCTTTACCACCTCACCT TGGGCA	5	8.03	-	20.78	39.41	24.20
Chr4	RBPJ OT12	TTCTCTTCCCCACCCACCTT TGAGCT	4	7.87	-	8.05	18.23	5.45
Chr11	116.5Mb OT13	TCCTCTTCTCCACCCACCAT AGGGTG	4	7.69	-	11.99	6.36	13.53
Chr20	ATRN OT14	GTCTCTTTCCCATCCACCTT TGATAA	3	6.01	-	12.18	20.92	20.96
Chr11	PCNXL3 OT15	GCCCCTACCCCACCCACCTG GGTGGA	5	5.62	-	5.76	1.46	0.88
Chr17	FBXL20 OT16	ACCTCTTATCCACCCACCTT GGCCTC	4	4.49	0.09	38.21	44.59	28.79
Chr7	***103.6Mb OT17***	GCCTCTCCCCACCCACCCTT GGCTTG	7	4.25	-	7.68	4.41	NA
Chr8	143.6Mb OT18	ACCTCCTTCCCGCCCACCTG GGGCTC	4	3.43	-	39.66	34.66	39.52
Chr7	AGK OT19	ACCCCTTCCCCACTCACCTC CGGGAT	5	3.12	-	0.69	1.95	12.90
Chr12	103.2Mb OT20	GCTTCTTCCCCACCCACACT TGGTGG	4	2.82	0.1	3.90	26.07	21.15
Chr1	SLC35F3 OT21	CACCCTTCCCCACCCACCCT GGGACC	5	2.27	-	4.51	3.43	7.65
Chr11	2.2Mb OT22	GCCCCTCCCCCACCCACCCT GGGGGA	4	1.95	-	4.19	10.35	12.90
Chr22	SUN2/DNAL3 OT23	TGCTCTTCCCCACCCACCCA AGGCCT	5	1.82	-	8.13	4.42	4.37
Chr1	33.6Mb OT24	ACCTCCATCCTCCCCACCTT GGGTCT	5	1.54	-	1.37	0.73	1.08
Chr14	XRCC3 OT25	GGCTCCTCCTCACCCACCTC TGGGTC	5	1.43	0.09	11.80	3.43	5.88
Chr17	LOC284009 OT26	GACTCTTTTCCACCCACCCT AGGGGA	3	1.40	-	24.14	14.50	9.58
Chr3	0.4Mb OT27	ACCTCTTCCCTCCCCACCTT TGGAGG	4	1.33	-	0.69	1.72	8.50
Chr4	181.5Mb OT28	ACCTCTTTCCATCCCACCTT GGGAAA	3	1.14	-	2.67	1.72	7.85
Chr6	40.0Mb OT29	GTGTATTCCCCACCCACCTT GGGAAG	4	0.77	-	-	-	-
Chr16	FBXO31 OT30	GCCCCTTCCCCACCCACCCT GTGGAA	4	0.74	-	7.49	6.91	6.33
Chr19	CYTH2 OT31	ACCTTGTCCCCACCCACCAG GGGATC	6	0.61	-	8.75	1.47	-
Chr22	CECR5 OT32	GCCCCTTACCCATCCACCCT GGGTGC	4	0.60	-	0.61	3.92	1.97
Chr7	VIPR2 OT33	GCCCCTGCCCCACCCACCTG TGGAGC	4	0.55	-	1.99	1.72	11.35

Mismatches to targeted Cas9:RAG1A and PAM are listed in red. Average junction frequencies (per 10,000 unique junctions) for off-target sites are listed for the Cas9:RAG1A (N=6) and the Cas9n:RAG1 combinations (total N=10; A/G (N=6), A/E (N=2), A/F (N=2)). Two-way ANOVA indicates an extremely significant interaction between Cas9:RAG1A and Cas9n:RAG1 A/G, A/E, A/F with the off-targets listed ($p<0.0001$). Post-tests indicate significance for each site ($p<0.001$ = ***). Additional Cas9:RAG1A OT bait site libraries with average on-target and OT junction frequencies also listed (N=3 each).

Supplementary Table 4. ATM TALEN off-target site sequences

Chromosome	Nuclease/ Locus	Left half-site	Spacer	Right half-site	Mis-match	RAG1B Bait Freq.	ATM Bait Freq.
Chr11	**ATM TALEN**	**TGAATTGGGATGCTGTTT (+)**	**18**	**TTTATTTTACTGTCTTTA (+)**	**NA**	**189.32**	**NA**
Chr16	HYDIN OT	TGGATAGGGATGATGTGC (+) TGAATAGAACTGCTTTTC (−)	17	–	5,7	**43.83**	**61.03**
Chr1	JAK1 OT	TAAATAGGGATTCTGAGC (+) TGAAATGGGGTTTTGATT (−)	15	–	6,5	**43.25**	**100.42**
Chr21	BRWD1-AS1 OT	TGAATAGGAACAAAGATT (+) CGAATCGGGAAGATATTC (−)	14	–	7,6	**38.43**	**31.99**
Chr15	FRMD5 OT	TCAATTGAAAAGCTGAGG (+) CGAATAGGAATGCTCTGC (−)	17	–	7,6	**17.88**	**27.53**
Chr21	MIR5009 OT	TGAATAGAAATGCTGTCA (+) AAAATTGGGATGATTTTC (−)	14	–	5,5	**14.52**	**11.22**
Chr10	65.5Mb OT	TTGAATAGAATGATGTAT (+) TGAATGAGAATGCTCATT (−)	18	–	7,5	**12.48**	**15.25**
Chr6	TSPYL4 OT	TGAACAAGGATGCTGCAT (+) TCACTTAAAATACTGTTT (−)	18	–	5,6	**12.32**	**26.15**
Chr3	176.1Mb OT	TAGATAGTGATACTGTTT (+) TAAATAGTGATGCTGTGG (−)	15	–	5,5	**11.75**	**19.26**
Chr9	7.7Mb OT	TAAATTAGGAAGCTGAGG (+) TGAATCAGGAAACTGTCC (−)	23	–	6,6	**12.25**	**13.29**
Chr9	13.4Mb OT	TGACTTCAGATGCTGCCT (+) TGAATTGGAAAGCAGTAG (−)	15	–	5,5	**10.17**	**14.19**
Chr15	37.9Mb OT	TGGACTAGAGTGCTGGTT (+) TTAGTTGGGATACTGCTT (−)	18	–	5,4	**8.25**	**14.38**
Chr1	22.9Mb OT	GGAATCAGGATACTCCTC (+) TGGATTAGAATGATCTAC (−)	17	–	7,7	**7.70**	**11.68**
Chr1	180.0Mb OT	TGAATGGAGATGATTACA (+) TGAATAGGCACGCTGTTC (−)	16	–	7,4	**8.18**	**15.30**
Chr6	LRFN2 OT	TGATTTGAGATGCTCTTA (+) TGCATAGGGGAGATGTTT (−)	14	–	4,5	**6.10**	**12.37**
Chr6	160.5Mb OT	TGAAGCAGGATACTGCAT (+) TGCATAAAGATGCTATTG (−)	18	–	6,6	**6.49**	**12.42**
Chr2	40.0Mb OT	TGCATAGTGAAGCTGCTT (+) TGAATAGGGAACCTATAG (−)	17	–	5,6	**6.53**	**7.65**
Chr15	FBN1 OT	TGAGTGAGGGTTCTAGTT (−)	16	ACTATTTTACTGTCTTTC (−)	7,3	**6.08**	**3.11**
Chr13	PCDH9 OT	TGAGATGAGAGACTCACT (+) TGAATTAAGATACTGTTT (−)	16	–	8,3	**6.37**	**4.28**
Chr18	20.3Mb OT	TGCATTGAAAAACTATCT (+) TAAATAAGAAAGCTTTCT (−)	17	–	7,7	**6.41**	**4.24**
Chr14	GMFB OT	TGAAAGGAAATACTGTCT (+) TGAAATAGGATATTATAT (−)	15	–	6,6	**4.82**	**4.72**
Chr2	55.8Mb OT	TAAATTGAAAGGCTGTTG (+) GGCATTGCGTTGAATTGT (−)	17	–	5,8	**4.63**	**3.87**
Chr10	SH3PXD2A OT	GGAAATGGAAACCTGTTT (+) TGAAATGAGATGATTTAT (−)	13	–	5,5	**6.28**	**18.27**
Chr12	MLXIP OT	TGAAGTGGGGTGCTGCCT (+) TTGATTGGGATTATGTTT (−)	17	–	4,4	**5.24**	**8.65**
Chr19	ZNF814 OT	TGAATGAGGCTGCTCTTT (+) TGAATGTGGAGAATGTGG (−)	15	–	4,7	**4.93**	**8.47**
Chr3	WNT7A OT	TGAACAGGGATGCTTTGT (+) TAAATTGAGATATTGCTT (−)	15	–	4,5	**3.71**	**7.96**
Chr12	FBXO21 OT	TACATTAGAATGATGTCC (+) TGAATGTGAATGCTAATG (−)	17	–	7,6	**4.26**	**4.31**
Chr15	TCF12 OT	TTAATTGGGACCCTGCCT (+) TGCATCTGGCTAATGTTT (−)	17	–	5,6	**4.50**	**5.06**
Chr22	PI4KA OT	TAAATGAAGCAACTGTTT (+)	18	GACACTCTACTGTCTTCA (+)	7,6	**4.41**	**5.21**
Chr7	ABCB1 OT	TGAATATGTATACTGCTT (+) TGAATTGGAAGACTATTG (−)	21	–	5,5	**4.50**	**7.10**
Chr15	82.1Mb OT	AGAATGCTGTTGCTCTTT (+) CCAATGGGGAGAATGTTT (−)	16	–	6,6	**3.97**	**6.49**

Left and Right TALEN half-sites are reported in the same chromosomal orientation except for loci displaying the same half-site for both sides with the 2nd site reported in the (-) orientation. Average frequencies (per 10,000) are listed for RAG1B and ATM bait libraries.

Supplementary Table 5. C-MYC TALEN off-target site sequences

Chromosome	Nuclease/ Locus	Left half-site	Spacer	Right half-site	Mismatch	RAG1B Bait Freq.	C-MYC Bait Freq.
Chr8	**C-MYC TALEN**	**TGCTTAGACGCTGGATTT (+)**	**16**	**AACCAGGTAAGCACCGAA (+)**	**NA**	**26.10**	**NA**
Chr19	TIMM44 OT	TACTTAGAACCTGGGTTT (-)	18	GGCCAGGGTGGCATCTGA (-)	5,8	**17.23**	**25.11**
Chr11	CUL5 OT	-	16	TATTAGTTAAGTACTTGT (+) CCACAGGTAAGTACTTAA (-)	9,6	**23.12**	**3.60**
Chr3	SEMA5b OT	-	20	ACCCAGGCACACACAGAA (+) CCTCAGGTATGCATTGAA (-)	5,6	**9.23**	**13.27**
Chr14	SLC35F4 OT	TACTCAGAAAATGAATTA (-)	17	AACCAGGTAATCTTTGTA (-)	7,5	**8.41**	**22.76**
Chr19	4.5Mb OT	-	22	TTACAGGAGTGCACTGCC (+) AACACGAGAATCGCTTGA (-)	9,8	**8.07**	**20.79**
Chr21	25.3Mb OT	TCCCGAGAAGCTGGAATT (-)	18	AACCAGGTAATTTTTGTA (-)	5,6	**4.52**	**2.40**
Chr8	1.6Mb OT	TGATTAAACCCTGATTCT (+) TCCTTGAACACTGGCTTT (-)	16	-	7,5	**5.34**	**33.84**
Chr1	PHC2 OT	-	15	AAGCAGGTAAGGACCTTA (+) AACCAGGGAAGTATCCAA (-)	4,4	**4.01**	**6.26**
Chr3	CCDC14 OT	-	15	GATCAGGTAGGCATTCAA (+) AGTCAGGTAAGCACACCA (-)	6,5	**6.90**	**5.84**
Chr8	LRRC69 OT	TCCCAAGAAGCTGGGATT (-)	18	AACCAGGTAATTTTTGTA (-)	6,7	**5.81**	**66.63**
Chr5	RhoBTB3 OT	-	18	CATCTGGTGAGTGCTGAA (+) AATGAGGCAGGCACATAA (-)	7,6	**4.34**	**3.09**
Chr13	39.6Mb OT	TACTAAAAGACAAGGTTT (-)	17	AACCAGGTAGGGAACTAT (-)	8,5	**3.09**	**7.99**
Chr4	43.3Mb OT	-	10	ATGCTTGTAATCCCAGCA (+) AACCAGGTACTTTCTTAA (-)	7,5	**3.78**	**4.04**
Chr17	14.6Mb OT	TACTCATGAGCTAAATAT (-)	17	TACCAGGTATGTCTTGAA (-)	8,6	**4.29**	**5.72**
Chr13	74.7Mb OT	-	17	GCCCAAGTAAGCATCCAA (+) GCGCAGGTAAGCATCTAA (-)	5,5	**3.71**	**6.58**
Chr3	165.2Mb OT	TGCTAAGAAGCTGGACTC (+) TGCTTAGAAGCAGGTCTT (-)	15	-	4,4	**3.56**	**14.73**
Chr10	FAM171A OT	TACATAAAAGCAAGATTC (+)	17	GGGAAGGTAAACATTGAA (+)	7,7	**3.09**	**5.43**
Chr8	PBK OT	TAATAAGACACAAGATAC (+) TGCTTCTCTGCTGAATTA (-)	16	-	8,6	**4.07**	**31.67**
Chr5	108.7Mb OT	TAAATAGAAGAAGGCTTG (+) TGATTAGAACATAGATCC (-)	14	-	8,7	**3.42**	**11.77**
Chr5	SH3PXD2B OT	TCCATAGACACTGCATCT (+)	16	AACCAGTCAAGCACTTTT (+)	5,6	**3.19**	**10.95**
Chr10	101.5Mb OT	TCCTGAGTAGCTGGATTA (+)	17	GCCCAGGTAATTTTTGTA (+)	5,8	**3.64**	**5.83**
Chr18	SS18 OT	TACCTAGAAAAAGAATGT (-)	17	CAGCAAGTATGCACTGGA (-)	8,7	**3.64**	**3.60**
Chr21	30.3Mb OT	TGTATAGAAGCTGGGTGT (+) TCCCAAAGCACTAGGATT (-)	16	-	5,9	**3.12**	**4.58**
Chr3	FAM208A OT	TAATGAAGCCCTGAACAT (+)	13	GTCCAGTTAAGCTCTTTA (-)	9,7	**3.73**	**13.00**
Chr10	DNA2 OT	TCCCAAGAAGCTGGGTTT (-)	18	ACCCAGCTAATTTTTGTA (-)	5,8	**2.82**	**3.76**
Chr3	MRPS22 OT	TAATGAAGCCCTGAACAT (+)	14	GTCCAGTTAAGCTCTTTA (+)	9,7	**2.63**	**4.55**
Chr17	69.9Mb OT	-	22	AATATGGTAAACTTTGAA (+) AATCAGGTAAGTATTTCA (-)	7,6	**4.21**	**4.66**
ChrX	PGK1 OT	TGCTTAAAAAAAGAATAT (+) TACCTAGAAGCTGATGGA (-)	16	-	7,8	**2.96**	**14.29**
ChrX	122.9Mb OT	TGCTTTGGCCCTGTGAGT (-)	21	GAGATAGGAAGCACTTAA (-)	7,8	**2.85**	**5.19**
Chr17	HS3ST3A1 OT	TCTCGAGAAGCTGAAATT (+)	20	GCCCAGGTAATTTTTGTA (+)	6,8	**3.37**	**4.85**

Left and Right TALEN half-sites are reported in the same chromosomal orientation except for loci displaying the same half-site for both sides with the 2nd site reported in the (-) orientation. Average frequencies (per 10,000) are listed for RAG1B and C-MYC bait libraries.

Supplementary Table 6. Cas9:RAG1B and I-SceI off-target site mismatches and translocations

Bait	Chromosome	Nuclease/Locus	Off-target site sequence	Mismatch	Frequency (per 10,000)
Cas9:RAG1B	**Chr11**	**Cas9:RAG1B**	**GACTTGTTTTCATTGTTCTC AGGNNN**	**0**	**NA**
“	Chr14	ZC3H14 OT1	TCCTTGTTTTCATTGTTCTC TGGTGG	2	**289.64**
“	Chr4	120.5Mb OT2	CATTTGTTTTCATTGTTCTC TGGCTG	2	**37.42**
Cas9:RAG1B	**NA**	**I-SceI**	**NNN TAGGGATAACAGGGTAAT NNN**	**0**	**NA**
“	Chr11	18.7Mb OT1	GTC TTGGGATAACAGGGCAAA GCA	3	**28.85**
“	Chr8	31.4Mb OT2	ATT TTGGGATAACAGGGCAAT ACT	2	**8.45**
“	Chr11	29.5Mb OT3*	TTG TAGGGATACCAGGTTTAT TTC	3	**6.97**
“	Chr1	CACNA1E OT4	GGC TAGGGATACCAGGTCAAA CAA	4	**5.20**
“	Chr11	GLYAT OT5	CAC TAGGGATAACAGGCTATT CGG	2	**5.17**
“	Chr20	11.4Mb OT6	TAC TAGGGATACCAGGGTCAT TCA	2	**4.86**
“	Chr9	24.8Mb OT7	TGC TAGGGATAACAGGTTGAA GGT	3	**3.64**
“	Chr6	CCND3 OT8	CAG TAGGGATAACAGGGCTGT TGA	3	**3.54**
“	Chr2	119.9Mb OT9	CAC TAGGGATGCCAGGGTGAA CAA	4	**1.4**

Mismatches to target sequences are listed in red. Average frequencies are listed (Cas9:RAG1B N=6; I-SceI N=3). Asterisk (*) indicates locus with 2 identical off-target sequences.

Supplementary Table 7. Cas9:EMX1 and Cas9:VEGFA off-target site mismatches and translocations

Bait	Chromosome	Nuclease/Locus	Off-target site sequence	Mismatch	Frequency (per 10,000)
Cas9:RAG1B	**Chr2**	**Cas9:EMX1**	**GAGTCCGAGCAGAAGAAGAA GGGNNN**	**0**	**136.25**
“	Chr15	MFAP1 OT1	GAGTCTAAGCAGAAGAAGAA GAGAGC	3	**189.39**
“	Chr5	HCN1 OT2**	GAGTTAGAGCAGAAGAAGAA AGGCAT	2	**120.13**
“	Chr5	SEMA5A OT3	AAGTCTGAGCACAAGAAGAA TGGTGA	3	**46.42**
“	Chr2	219.8Mb OT4	GAGGCCGAGCAGAAGAAAGA CGGCGA	3	**30.31**
“	Chr11	HSD17B12 OT5	AAGCCCGAGCAAAGGAAGAA AGGAGA	4	**28.48**
“	Chr8	128.8Mb OT6	GAGTCCTAGCAGGAGAAGAA GAGGCA	3	**16.35**
“	ChrX	53.4Mb OT7	GAGTCCGGGAAGGAGAAGAA AGGCTC	3	**8.11**
“	Chr5	DPYSL3 OT8	GAGCCGGAGCAGAAGAAGGA GGGAGG	3	**6.21**
“	Chr11	MTA2 OT9	GAATCCAAGCAGAAGAAGAG AAGGAG	4	**5.95**
“	Chr6	WASF1 OT10	AAGTCAGAGCAGAAAAAGAG AGGACA	4	**5.41**
“	Chr6	9.1Mb OT11	ACGTCTGAGCAGAAGAAGAA TGGACA	3	**1.34**
“	Chr1	TCEA3 OT12	AAGTCCGAGGAGAGGAAGAA AGGGTT	3	**1.14**
“	Chr13	27.7Mb OT13	GAGTAGGAGCAGGAGAAGAA GGAGGA	4	**1.08**
Cas9:RAG1B	**Chr6**	**Cas9:VEGFA**	**GGGTGGGGGGAGTTTGCTCC TGGNNN**	**0**	**190.26**
“	Chr15	IGDCC3 OT1**	GGATGGAGGGAGTTTGCTCC TGGGGT	2	**176.22**
“	Chr17	KRT42P OT2	TAGTGGAGGGAGCTTGCTCC TGGCTG	4	**165.88**
“	Chr1	99.3Mb OT3**	GGGGAGGGGAAGTTTGCTCC TGGCAT	3	**136.19**
“	Chr12	LINC01257 OT4**	GGGAGGGTGGAGTTTGCTCC TGGGGA	3	**72.27**
“	Chr12	CACNA2D4 OT5**	CGGGGGAGGGAGTTTGCTCC TGGGGA	3	**66.03**
“	Chr22	PVALB OT6	GGGTGGGGGGAGTTTGCCCC AGGCCA	1	**49.04**
“	Chr6	14.3Mb OT7	GTGGGGGTAGAGTTTGCTCC AGGTGT	4	**38.70**
“	Chr22	19.6Mb OT8	GAGGGGGAGCAGTTTGCTCC AGGTGA	4	**33.47**
“	Chr5	32.9Mb OT9	GCGTGGGGGGTGTTTGCTCC CGGGCA	2	**19.33**
“	Chr17	47.3Mb OT10	CTGGTGGGGGAGCTTGCTCC AGGGAA	5	**18.04**
“	Chr11	67.5Mb OT11	AGGAAGGAGGAGTTAGCTCC TGGGGG	5	**16.66**
“	Chr5	MAP3K1 OT12	GGTGGGGGTGGGTTTGCTCC TGGTAT	4	**15.86**
“	Chr11	DSCAML1 OT13	GGGCAAGGGGAGGTTGCTCC TGGAGA	4	**15.63**
“	Chr3	128.2Mb OT14	AGGTGGTGGGAGCTTGTTCC TGGCTT	4	**15.12**
“	Chr1	PCNXL2 OT15	GGAGGAGGGGAGTCTGCTCC AGGTTT	4	**14.02**
“	Chr11	KIAA1549L OT16	AGCTGAGGGGAGCTTGCTCT GGGCTG	5	**12.31**
“	Chr4	TRMT44 OT17	GAGTGGGTGGAGTTTGCTAC AGGCAG	3	**12.12**
“	Chr13	ATP8A2 OT18	GGTTGAGGGGAGTCTGCTCC AGGCTT	3	**11.80**
“	Chr1	TRIM62 OT19	GGGTGGGTGGAGTTTGCTAC TGGCAT	2	**10.41**
“	Chr20	56.1Mb OT20	AGGGAGGAGGAATTTGCTCC AGGAGT	5	**9.96**
“	Chr17	32.9Mb OT21	GGGGGTGGGGACTTTGCTCC AGGGCC	3	**8.65**
“	Chr22	SUN2 OT22	GGGCAGAGGGAGTTAGCACC GGGCGT	5	**7.95**
“	Chr16	8.7Mb OT23	AAGTAAGGGAAGTTTGCTCC TGGTCC	5	**6.11**
“	Chr11	3.4Mb OT24	AGGAAGGAGGAGTTAGCTCC TGGGGG	5	**5.69**
“	Chr11	71.4Mb OT25	AGGAAGGAGGAGTTAGCTCC TGGGGG	5	**5.52**
“	Chr10	124.7Mb OT26	AGCTGGAGGGAGTTTGCCCC AGGTGA	4	**5.03**
“	Chr5	7.0Mb OT27	GAGGGTGGGGAGTTTACTCC TGGAAG	4	**4.72**
“	Chr20	C20orf62 OT28	GGGTGAGGGGAATAAACTCC AGGGTG	5	**3.80**
“	Chr22	38.2Mb OT29	AGGTCGGGGGAGTTAGATCC CGGGGT	4	**3.66**
“	Chr11	47.2Mb OT30	GGGGAGGGGGAGGGTGCTCC AGGCAG	3	**3.47**
“	ChrX	19.1Mb OT31	GGGAGGGGAGAGTTTGTTCC AGGAAA	3	**3.38**
“	Chr21	RUNX1 OT32	AAGTGGGAAGAGTTTGTTCC AGGCTC	5	**2.96**
“	Chr12	C12orf39 OT33	GGGAGGGGGCAGGTTGCTCC AGGATA	3	**2.58**
“	Chr11	NAV2 OT34	AGGTAGGAGAAGCTTGCTCC TGAGAT	6	**2.45**
“	Chr3	LINC00885 OT35	GGTGGGGGAGAGCTAGCTCC GGGAGG	5	**2.41**
“	Chr8	28.4Mb OT36	AAGTGGGAGGAGACTGCTCC AGGTAG	5	**1.80**
“	Chr10	6.7Mb OT37	AAATGGGGGGAGTTTGCCCC CCGGAG	5	**1.52**
“	Chr19	18.5Mb OT38	GGGGGCGGGGAGGTTGCCCC GGGGAA	4	**1.52**

Mismatches to target sequences are listed in RED. Average frequencies are listed (Cas9:EMX1 N=3; Cas9:VEGFA N=3). Double Asterisks (**) represent sites previously identified[15].

Supplementary Table 8. Translocation cloning and off-target oligos

Name	Sequence5'→3' **	Purpose
I-SceI off-target (OT) oligos		
Chr11_18.7Mb_F	AGC CCC GTC TTC TCT GAA TG	1kb amplicon; OT site 800bp / 200bp
Chr11_18.7Mb_R	CGT TGA GCT GCT TTT TCC TC	
Chr8_31.4Mb_F	GCA ACA CGG TGT GGT ATT TC	980bp amplicon OT site 600bp / 380bp
Chr8_31.4Mb_R	TGA ACC AAT CCT GAC ATT GC	
Chr11_29.5Mb_F	TTT GGA AAC AAG CCC AGT TG	1,419bp amplicon; 2 OT sites 250bp / 800bp / 350bp
Chr11_29.5Mb_R	TGA CCC CAG ATT CCT TCT TC	
CACNA1E_F	GGT GGT TCC TGA GTG TTT CC	995bp amplicon; OT site 300bp / 695bp
CACNA1E_R	TGG CCA AAA GTC ATG AAG TG	
Chr20_11.4Mb_F	GGT TGC CAT TGT GAT TCC TC	1kb amplicon; OT site 700bp / 300bp
Chr20_11.4Mb_R	GGC ATA AAC CCA CAA AAA GG	
GLYAT_F	CCT TAG CCC ATG GAT TCT ACC	964bp amplicon; OT site 220bp / 744bp
GLYAT_R	TCA GTT TAC CCC AAC CAA GC	
Chr2_119.9Mb_F	TGC TGT GGC TTG AAT GTC TC	900bp amplicon; OT site 800bp / 100bp
Chr2_119.9Mb_R	ACT TCT GAG GGG CCT TTG TC	
Oligos for T7 Endo I		
RAG1_F	CCC CCT GGA AGA CTG CTT TA	690bp amplicon; 470bp / 220bp Cas9:RAG1B site
RAG1_R	AGG ACT GCT GGA GAT TGC TC	
RAG1B_OT1_F	TTG GTT GCC CAT CTT ATT CC	700bp amplicon; OT site 220bp / 480bp
RAG1B_OT1_R	AGC TAA ACC TGC CTG CAG AA	
RAG1B_OT2_F	AAT AAC TTG CAG CCA TTC CA	700bp amplicon; OT site 180bp / 520bp
RAG1B_OT2_R	TAA ATT GCC CAT GAT TGC AC	
EM-PCR adapter oligos		
Universal oligo	/5phos/ TAA CCA GCC C /3InvdT/	Suppress adapter-ended amplification
Priming oligo	GTA ATA CGA CTC ACT ATA GGG CAC GCG TGG TCG ACG GCC CGG GCT GGT TAT	Contains adapter primer
RAG1 A/B		
Bio RAG1A/B - F1	/5BiosG/ AGG ACT GCT GGA GAT TGC TC	Bio-PCRI
RAG1A/B - F2 I5 M00B	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TGA GAG GGT TTC CCC TCA AAG	Miseq tail
RAG1A/B - F2 I5 M02B	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTG GAG AGG GTT TCC CCT CAA AG	Miseq tail multiplex
RAG1A/B - F2 I5 M03B	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCA TGA GAG GGT TTC CCC TCA AAG	Miseq tail multiplex
RAG1A/B - F2 I5 M04B	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TGC TCG AGA GGG TTT CCC CTC AAA G	Miseq tail multiplex

RAG1A/B - F2 I5 M05B	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TGC CAT GAG AGG GTT TCC CCT CAA AG	Miseq tail multiplex
RAG1A/B - F2 I5 M06B	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAT ATC GGA GAG GGT TTC CCC TCA AAG	Miseq tail multiplex
AP2 I7	TCT CGG CAT TCC TGC TGA ACC GCT CTT CCG ATC TAC TAT AGG GCA CGC GTG GT	Adapter primer; Miseq tail reverse complement
P5	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT	Miseq tail PCRIII
P7	CAA GCA GAA GAC GGC ATA CGA GAT CGG TCT CGG CAT TCC TGC TGA ACC	Miseq tail PCRIII reverse complement
RAG1 C/D		
Bio RAG1C/D - F1	/5BiosG/ CCT GAG AAC AAT GAA AAC AAG TC	Bio-PCRI
RAG1C/D - F2 I5 M00C	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAC CAA TAT CAA TAT CCC ACT GAT G	Miseq tail
RAG1C/D - F2 I5 M02C	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCG ACC AAT ATC AAT ATC CCA CTG ATG	Miseq tail multiplex
RAG1C/D - F2 I5 M04C	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCA CGA CCA ATA TCA ATA TCC CAC TGA TG	Miseq tail multiplex
RAG1C/D - F2 I5 M06C	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TGA CTC GAC CAA TAT CAA TAT CCC ACT GAT G	Miseq tail multiplex
Bridge adapter oligos (LAM-PCR)		
ssA-upper-TN	GCG ACT ATA GGG CAC GCG TGG TNN NNN-/3AmMO/	Common bridge oligo
ssA-lower	/5Phos/ CCA CGC GTG CCC TAT AGT CGC-/3AmMO/	To link A-tailed DNA
ssA-lower-A	/5Phos/ ACC ACG CGT GCC CTA TAG TCG C-/3AmMO/	To link G-,C-, T-tailed or blunt DNA
Cas9:RAG1A – OT		
Bio DAZAP - F1A	/5BiosG/ TTC CCT GTA ACT TGG GAT GG	Bio-PCRI
DAZAP - F2A I5 M00F	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTG CCC AGA AAT CAG AAC AAC	Miseq tail
DAZAP - F2A I5 M02F	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAG TGC CCA GAA ATC AGA ACA AC	Miseq tail multiplex
DAZAP - F2A I5 M03F	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTC ATG CCC AGA AAT CAG AAC AAC	Miseq tail multiplex
DAZAP - F2A I5 M04F	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TGA TCT GCC CAG AAA TCA GAA CAA C	Miseq tail multiplex
DAZAP - F2A I5 M05F	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCT GAA TGC CCA GAA ATC AGA ACA AC	Miseq tail multiplex
Bio 12_47 + F1	/5BiosG/ GGA CAA CAA CCC CCA GTT AG	Bio-PCRI
12_47 + F2A I5 M00G	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTG TGG TAA AGG AGA CAA TGC T	Miseq tail
12_47 + F2A I5 M02G	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCT TGT GGT AAA GGA GAC AAT GCT	Miseq tail multiplex
12_47 + F2A I5 M03G	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TGA CTG TGG TAA AGG AGA CAA TGC T	Miseq tail multiplex
12_47 + F2A I5 M04G	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAC GTT GTG GTA AAG GAG ACA ATG CT	Miseq tail multiplex
12_47 + F2A I5 M05G	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTG ACA TGT GGT AAA GGA GAC AAT GCT	Miseq tail multiplex
TALEN sites		
Bio ATM + F1	/5BiosG/ CTG CTG CCG TCA ACT AGA AC	Bio-PCRI
ATM + F2 I5 M00H	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAT TTA AGC GCC TGA TTC GAG	Miseq tail multiplex
ATM + F2 I5 M02H	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCG ATT TAA GCG CCT GAT TCG AG	Miseq tail multiplex

ATM + F2 I5 M03H	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTC AAT TTA AGC GCC TGA TTC GAG	Miseq tail multiplex
ATM + F2 I5 M04H	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAT CTA TTT AAG CGC CTG ATT CGA G	Miseq tail multiplex
ATM + F2 I5 M05H	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCC GCG ATT TAA GCG CCT GAT TCG AG	Miseq tail multiplex
ATM + F2 I5 M05H2	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTG TAC ATT TAA GCG CCT GAT TCG AG	Miseq tail multiplex
Bio MYC + F1	/5BiosG/ CGA AAC TTT GCC CAT AGC AG	Bio-PCRI
MYC + F2 I5 M00I	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCT TAC AAC ACC CGA GCA AGG	Miseq tail multiplex
MYC + F2 I5 M03I	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TAG CCT TAC AAC ACC CGA GCA AGG	Miseq tail multiplex
MYC + F2 I5 M04I	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TCT CGC TTA CAA CAC CCG AGC AAG G	Miseq tail multiplex
MYC + F2 I5 M05I	TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TTG TAT CTT ACA ACA CCC GAG CAA GG	Miseq tail multiplex
**Color code: Green = adapter sequence complementarity; Blue = multiplex sequence; Red = Terminal ends for MiSeq; Orange = Internal MiSeq priming sequence; /5phos/ = 5' phosphorylation; /3InvdT/ = 3' inverted dT; /5BiosG/ = 5' Biotin; /3AmMO/ = 3' Amino modified.		

Supplementary References

47. Schueler, M.G. & Sullivan, B.A. Structural and functional dynamics of human centromeric chromatin. *Annu. Rev. Genomics Hum. Genet.* **7**, 301-313 (2006).