eLIFE
elifesciences.org

Figures and figure supplements

APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition

Sandra R Richardson, et al.

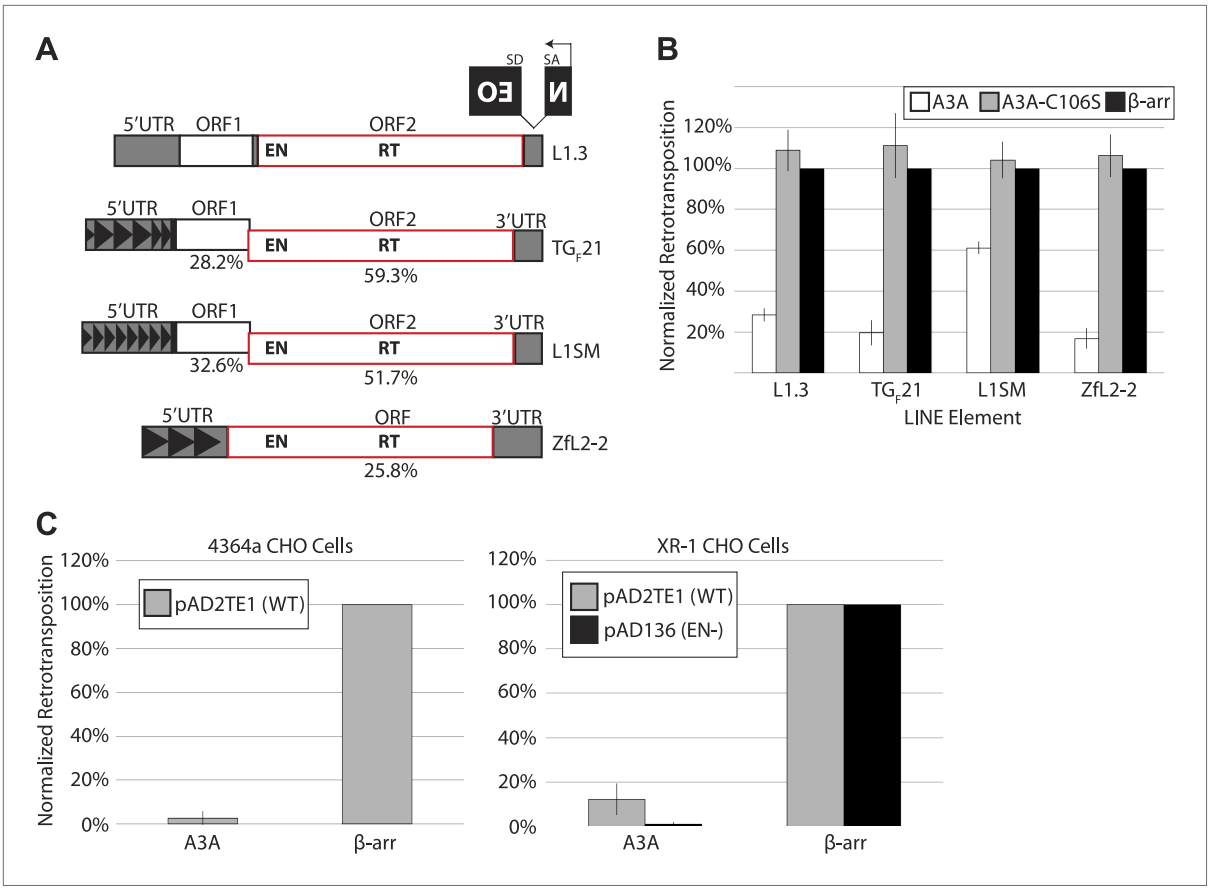


Figure 1. A3A suppresses retrotransposition of distinct LINEs, but does not specifically inhibit L1 endonuclease activity. (**A**) *LINE schematics*: L1.3, TG_F21, and L1SM encode two ORFs (ORF1 and ORF2, outlined in black and red, respectively). ZfL2-2 contains one ORF (outlined in red). Indicated are the endonuclease (EN) and reverse transcriptase (RT) domains in the respective ORFs. All elements are tagged in their 3′UTRs with the *mneoI* (NEO) indicator cassette. The percent nucleotide identity of TG_F21, L1SM, and ZfL2-2 to human L1.3 is indicated below each schematic. Black arrowheads indicate repeated monomeric sequences in the 5′ UTRs of TG_F21, L1SM, and ZfL2-2. (**B**) *A3A inhibits retrotransposition of distinct LINEs*: shown is the effect of wild-type A3A (white bars), deaminase-deficient A3A_C106S (gray bars), and β-arrestin (β-arr) control (black bars) on LINE retrotransposition. The x-axis indicates the LINE element. The y-axis indicates percent retrotransposition; in each case the β-arr control is set to 100%. Data were normalized using a circular *NEO* expression cassette as detailed in ***Figure 1—figure supplement 1A***. Data are expressed as the mean percent retrotransposition derived from three independent experiments consisting of two technical replicates each, with error bars representing the standard deviation among all six technical replicates. (**C**) *Retrotransposition assays in 4364a (wild-type CHO cells (left)) and XR-1 (XRCC4-deficient CHO cells (right))*: the x-axis indicates experiments conducted with the A3A or β-arr expression vector. The y-axis indicates percent retrotransposition; for each reaction the β-arr control is set to 100%. Gray bars indicate retrotransposition of wild-type L1 (pAD2TE1) and black bars indicate the L1 EN mutant (pAD136). Data were normalized using the circular *NEO* control (***Figure 1—figure supplement 1A***). Data are shown as the mean percent retrotransposition derived from three independent experiments consisting of two or three technical replicates each, with error bars representing the standard deviation among all eight technical replicates.
DOI: 10.7554/eLife.02008.003

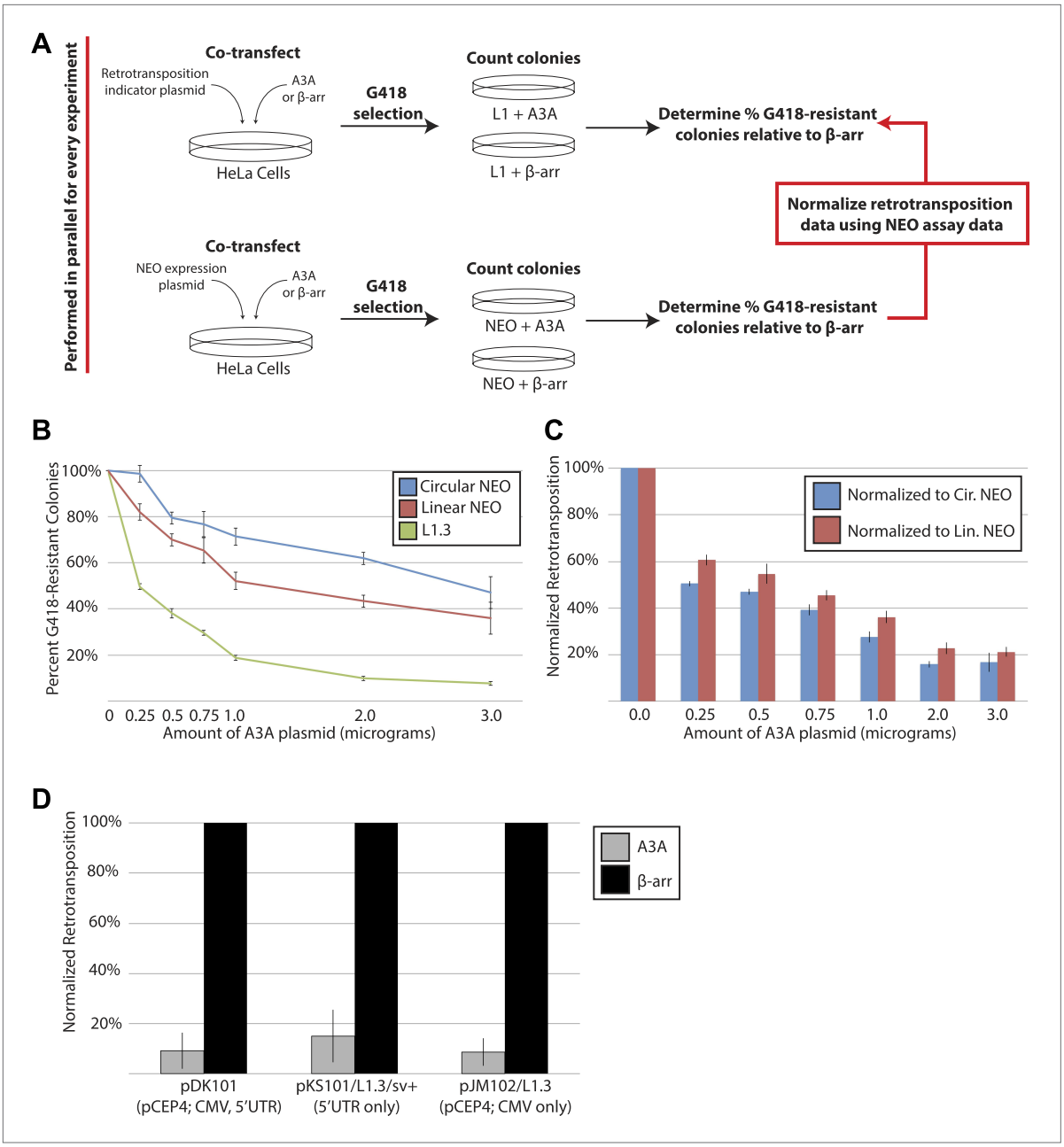


Figure 1—figure supplement 1. Additional Control Experiments.

DOI: 10.7554/eLife.02008.004

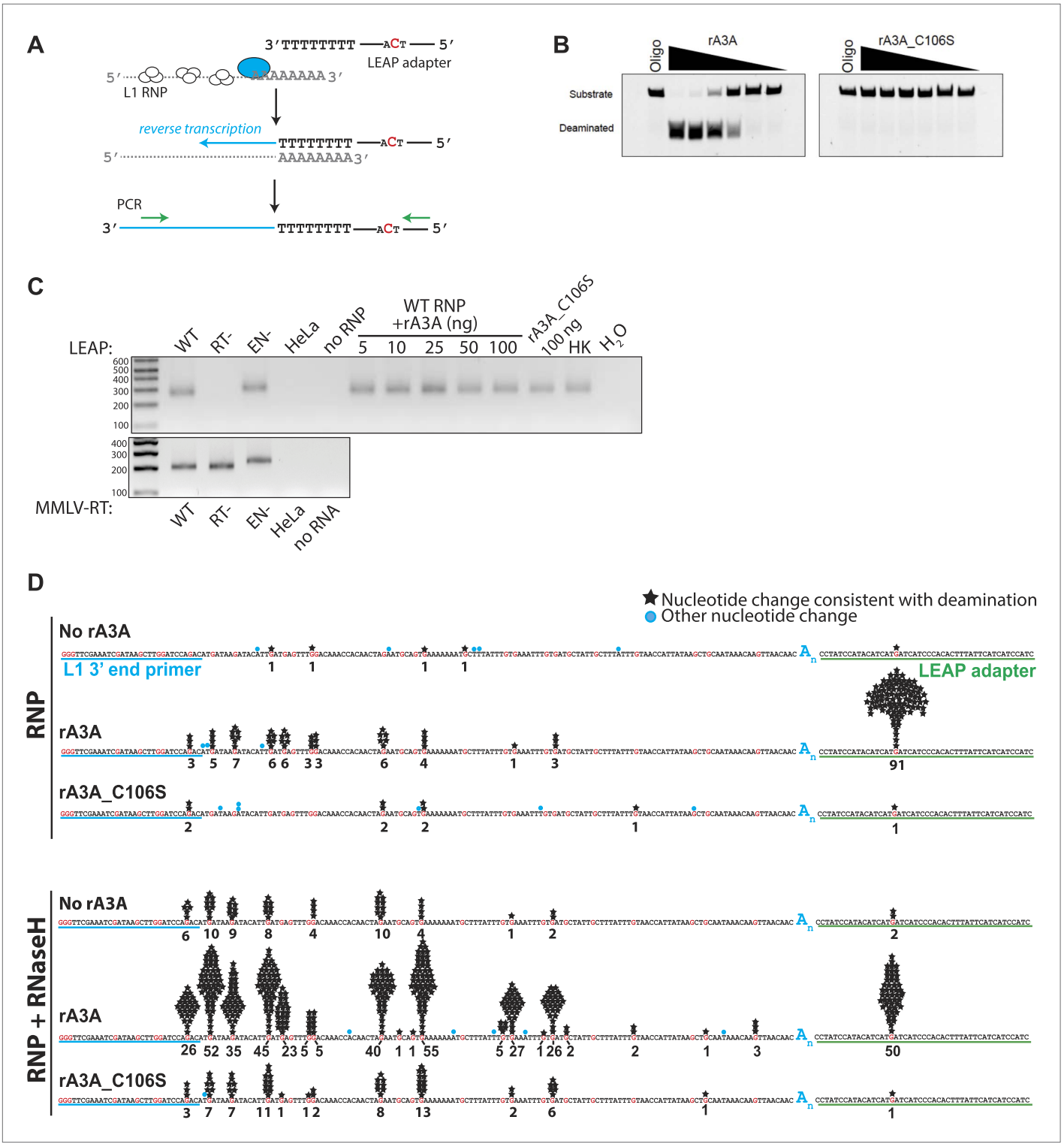


Figure 2. Recombinant A3A (rA3A) deaminates L1 cDNAs in vitro. (**A**) *LEAP assay rationale*: L1 RNP preparations consisting of the L1 RNA (gray), L1 ORF1p (white ovals), and L1 ORF2p (blue oval) are incubated with a 3′ RACE primer consisting of a unique adapter sequence that contains a single cytidine (red) followed by an oligo dT sequence (black lettering). After reverse transcription (blue arrow), the resultant L1 cDNAs (blue line) are PCR amplified using primers specific to the engineered L1 and the unique adapter sequence (green arrows). (**B**) *Recombinant A3A has deaminase activity in vitro*: twofold serial dilutions (500 ng–15.62 ng) of WT rA3A (left panel) or deaminase-deficient rA3A_C106S (right panel) were incubated with a fluorescein isothiocyanate (FITC) labeled single-strand DNA oligonucleotide containing a single cytidine residue. The products were treated with recombinant

Figure 2. Continued on next page

Figure 2. Continued

uracil DNA glycosylase (UDG) and NaOH and then were resolved by gel electrophoresis. A control reaction was included without recombinant protein (marked Oligo). (**C**) *Recombinant A3A does not inhibit L1 RT activity*: control LEAP reactions with RNP preparations from HeLa cells transfected with WT (pDK101), RT- (pDK135), or EN- (pJJH230A/L1.3) human L1s (upper gel). HeLa indicates untransfected HeLa cells; no RNP indicates control reactions lacking RNPs. Increasing amounts of rA3A (ng) did not significantly affect LEAP activity. Samples containing a deaminase-deficient rA3A_C106S, a heat killed rA3A (HK), or without LEAP products (H_2O) served as controls. MMLV RT reactions (lower gel) confirm the integrity of purified RNA isolated from RNP preps used in the LEAP assay. Notably, the increased size of the EN- RNP RT products is due to a higher molecular weight product generated from pJJH230A/L1.3, which contains an *mblastI* indicator cassette instead of an *mneoI* indicator cassette. Size standards (bp) are indicated at the left of the gel. (**D**) *Sequence characterization of LEAP Products*: shown is the (+) strand sequence of the LEAP product. Guanosine nucleotides are indicated in red. Black stars and numbers indicate the frequency of G-to-A mutations (corresponding to C-to-U mutations in the minus (−) strand L1 cDNA) that occurred on (+) strand L1 cDNA. Blue circles indicate other nucleotide changes. The blue A_n indicates the LEAP product poly (**A**) tail. Blue underlining indicates the L1 3′ end PCR primer. Green underlining indicates the LEAP adapter (5np1) sequence. *Top panel*: LEAP products generated under the following conditions: no rA3A protein, 100 ng of wild-type rA3A, 100 ng of deaminase-deficient rA3A_C106S. *Bottom panel*: LEAP products generated under the following conditions in the presence of RNase H: no rA3A protein, 100 ng of wild-type rA3A, 100 ng of deaminase-deficient rA3A_C106S. One hundred products were characterized for each condition.

DOI: 10.7554/eLife.02008.005

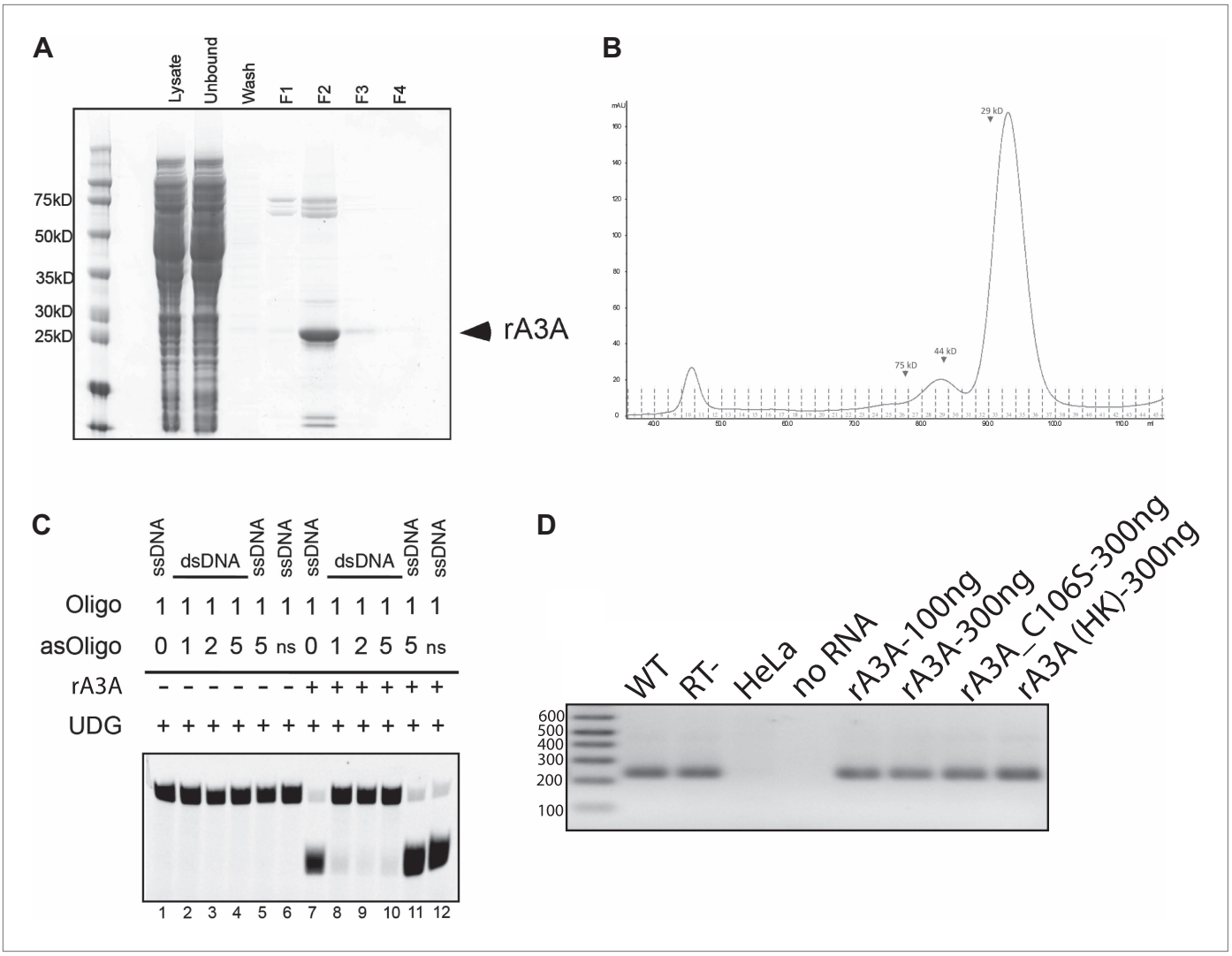
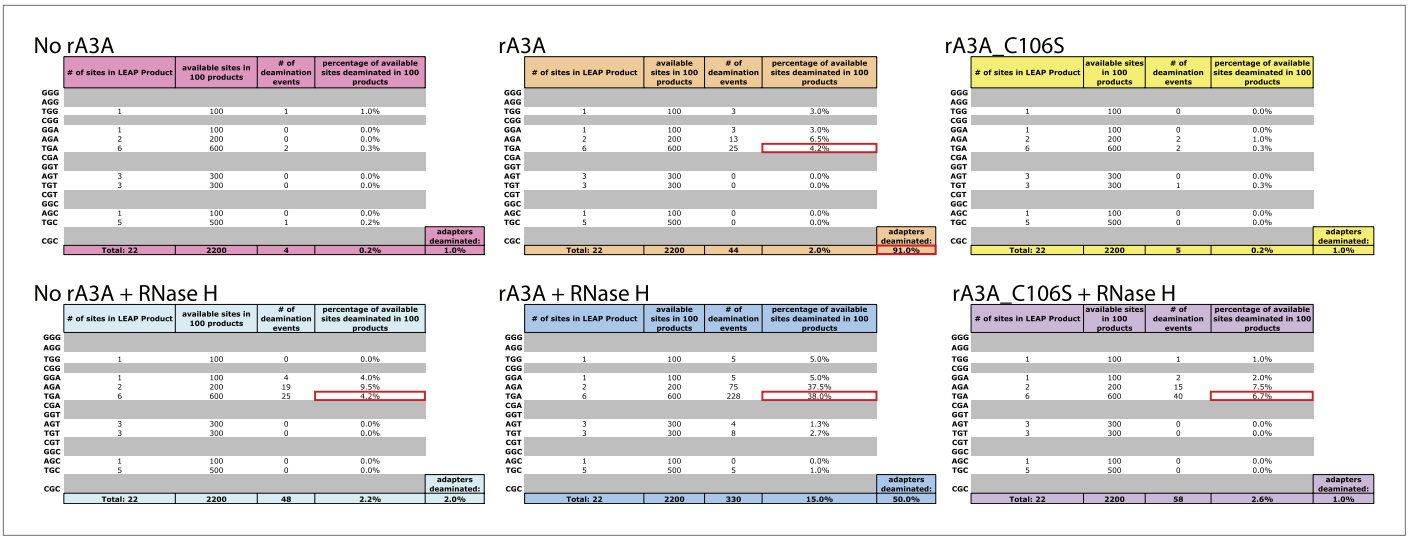


Figure 2—figure supplement 1. Control Experiments with Recombinant A3A.

DOI: 10.7554/eLife.02008.006



No rA3A	# of sites in LEAP Product	available sites in 100 products	# of deamination events	percentage of available sites deaminated in 100 products	
GGG					
AGG					
TGG	1	100	1	1.0%	
CGG					
GGA	1	100	0	0.0%	
AGA	2	200	0	0.0%	
TGA	6	600	2	0.3%	
CGA					
GGT					
AGT	3	300	0	0.0%	
TGT	3	300	0	0.0%	
CGT					
GGC					
AGC	1	100	0	0.0%	
TGC	5	500	1	0.2%	
CGC					adapters deaminated:
	Total: 22	2200	4	0.2%	1.0%

rA3A	# of sites in LEAP Product	available sites in 100 products	# of deamination events	percentage of available sites deaminated in 100 products	
GGG					
AGG					
TGG	1	100	3	3.0%	
CGG					
GGA	1	100	3	3.0%	
AGA	2	200	13	6.5%	
TGA	6	600	25	4.2%	
CGA					
GGT					
AGT	3	300	0	0.0%	
TGT	3	300	0	0.0%	
CGT					
GGC					
AGC	1	100	0	0.0%	
TGC	5	500	0	0.0%	
CGC					adapters deaminated:
	Total: 22	2200	44	2.0%	91.0%

rA3A_C106S	# of sites in LEAP Product	available sites in 100 products	# of deamination events	percentage of available sites deaminated in 100 products	
GGG					
AGG					
TGG	1	100	0	0.0%	
CGG					
GGA	1	100	0	0.0%	
AGA	2	200	2	1.0%	
TGA	6	600	2	0.3%	
CGA					
GGT					
AGT	3	300	0	0.0%	
TGT	3	300	1	0.3%	
CGT					
GGC					
AGC	1	100	0	0.0%	
TGC	5	500	0	0.0%	
CGC					adapters deaminated:
	Total: 22	2200	5	0.2%	1.0%

No rA3A + RNase H	# of sites in LEAP Product	available sites in 100 products	# of deamination events	percentage of available sites deaminated in 100 products	
GGG					
AGG					
TGG	1	100	0	0.0%	
CGG					
GGA	1	100	4	4.0%	
AGA	2	200	19	9.5%	
TGA	6	600	25	4.2%	
CGA					
GGT					
AGT	3	300	0	0.0%	
TGT	3	300	0	0.0%	
CGT					
GGC					
AGC	1	100	0	0.0%	
TGC	5	500	0	0.0%	
CGC					adapters deaminated:
	Total: 22	2200	48	2.2%	2.0%

rA3A + RNase H	# of sites in LEAP Product	available sites in 100 products	# of deamination events	percentage of available sites deaminated in 100 products	
GGG					
AGG					
TGG	1	100	5	5.0%	
CGG					
GGA	1	100	5	5.0%	
AGA	2	200	75	37.5%	
TGA	6	600	228	38.0%	
CGA					
GGT					
AGT	3	300	4	1.3%	
TGT	3	300	8	2.7%	
CGT					
GGC					
AGC	1	100	0	0.0%	
TGC	5	500	5	1.0%	
CGC					adapters deaminated:
	Total: 22	2200	330	15.0%	50.0%

rA3A_C106S + RNase H	# of sites in LEAP Product	available sites in 100 products	# of deamination events	percentage of available sites deaminated in 100 products	
GGG					
AGG					
TGG	1	100	1	1.0%	
CGG					
GGA	1	100	2	2.0%	
AGA	2	200	15	7.5%	
TGA	6	600	40	6.7%	
CGA					
GGT					
AGT	3	300	0	0.0%	
TGT	3	300	0	0.0%	
CGT					
GGC					
AGC	1	100	0	0.0%	
TGC	5	500	0	0.0%	
CGC					adapters deaminated:
	Total: 22	2200	58	2.6%	1.0%

Figure 2—figure supplement 2. A3A Deamination Events in LEAP products.
DOI: 10.7554/eLife.02008.007

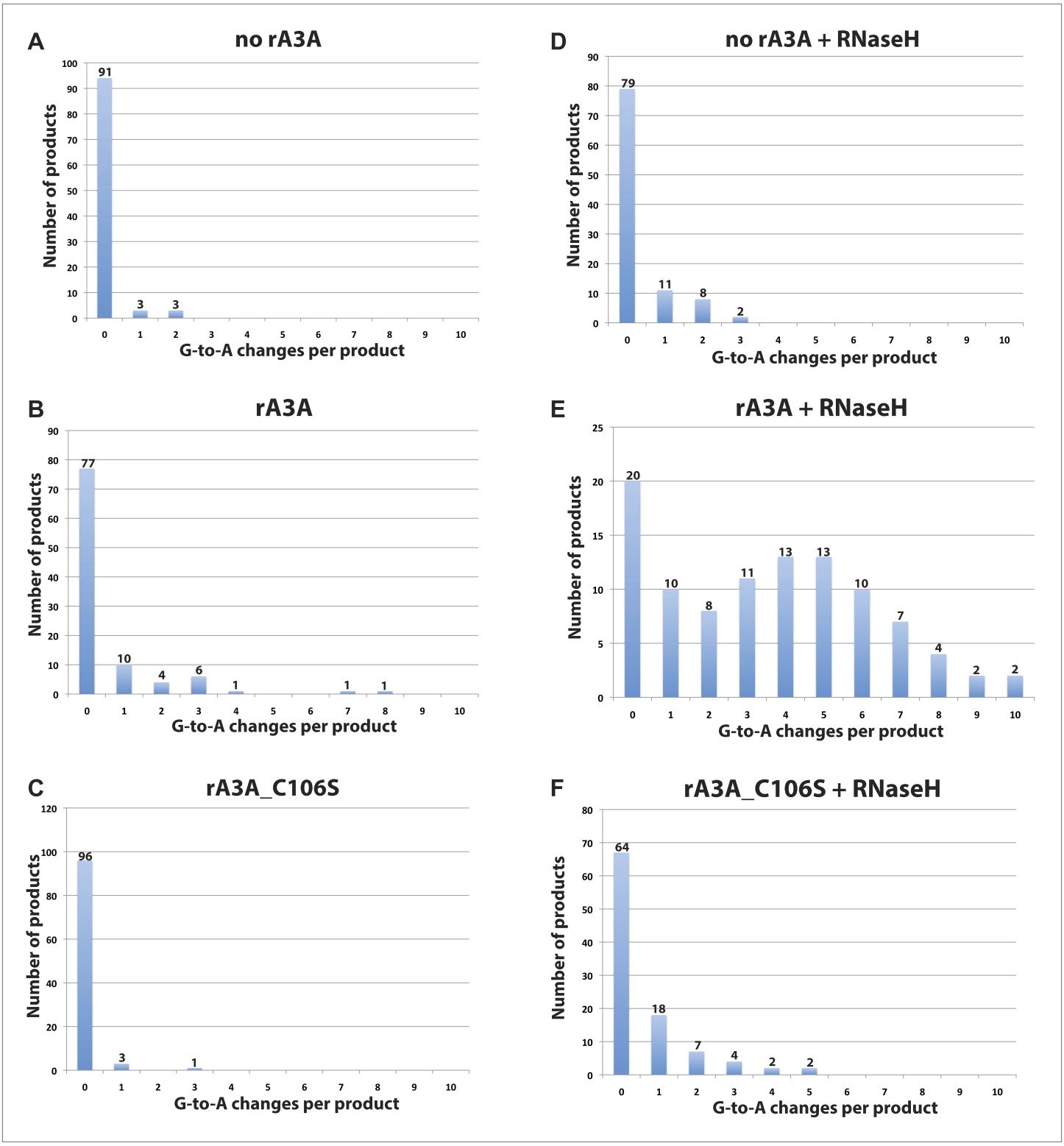


Figure 2—figure supplement 3. Distribution of deamination events per LEAP product.
DOI: 10.7554/eLife.02008.008

Sample no.	polyA length	Total GtoA	Total Other	G29	G34	G39	G47	G50	G56	G57	G72	G76	G79	G81	G99	G101	G108	G110	G113	G128	G144	G155	Other	Adapter (TCA)
A3A+RNaseH 1	56	0																						X
A3A+RNaseH 2	68	2			X									X										X
A3A+RNaseH 3	101	8		X	X	X	X	X			X			X				X						X
A3A+RNaseH 4	62	2					X			X														X
A3A+RNaseH 5	85	7			X		X	X	X	X	X			X										
A3A+RNaseH 6	77	9		X	X	X	X	X			X			X	X	X								X
A3A+RNaseH 7	12-C-41	6		X	X	X	X	X						X										X
A3A+RNaseH 8	63	0	1																				A104T	
A3A+RNaseH 9	44	1									X													
A3A+RNaseH 10	91	3	1	X		X										X							T148C	
A3A+RNaseH 11	84	5			X		X	X						X				X						X
A3A+RNaseH 12	94	3			X						X			X										X
A3A+RNaseH 13	61	6			X	X	X				X	X		X										X
A3A+RNaseH 14	38	3									X			X				X						
A3A+RNaseH 15	88	0																						
A3A+RNaseH 16	92	4			X		X				X			X										X
A3A+RNaseH 17	86	7		X	X		X				X			X		X		X						X
A3A+RNaseH 18	48	5			X		X	X						X				X						X
A3A+RNaseH 19	85	6		X	X	X	X							X				X						X
A3A+RNaseH 20	85	2			X								X											X
A3A+RNaseH 21	84	0																						
A3A+RNaseH 22	47	0																						
A3A+RNaseH 23	94	9			X	X	X	X			X			X	X	X	X							X
A3A+RNaseH 24	48	10		X	X	X	X	X			X			X		X		X				X		
A3A+RNaseH 25	61	0																						X
A3A+RNaseH 26	57	1												X										
A3A+RNaseH 27	55	7		X	X	X	X							X		X		X						X
A3A+RNaseH 28	39-GATGGCTAATGAAT-12	7			X	X	X	X								X		X	X					X
A3A+RNaseH 29	45	6			X		X	X			X			X				X						X
A3A+RNaseH 30	50	5			X			X			X			X		X								X
A3A+RNaseH 31	46-GACGATC-15	3			X	X								X										
A3A+RNaseH 32	81	6			X	X	X	X			X			X										X
A3A+RNaseH 33	51	0																						
A3A+RNaseH 34	84	0	1																				A88G	
A3A+RNaseH 35	60	1									X													
A3A+RNaseH 36	99	5		X			X				X			X		X								
A3A+RNaseH 37	79	8			X	X	X	X			X			X		X		X						X
A3A+RNaseH 38	96	3					X				X			X										X
A3A+RNaseH 39	31	10		X	X	X	X	X			X			X		X		X		X				X
A3A+RNaseH 40	31	0																						X
A3A+RNaseH 41	58	6		X		X	X	X						X				X						X
A3A+RNaseH 42	93	4		X	X	X					X													X
A3A+RNaseH 43	30	6		X	X	X	X				X			X										
A3A+RNaseH 44	72	5			X		X							X		X		X						X
A3A+RNaseH 45	54	3			X									X		X								
A3A+RNaseH 46	60	0																						X
A3A+RNaseH 47	58	0																						
A3A+RNaseH 48	59	0																						X
A3A+RNaseH 49	62	1			X																			
A3A+RNaseH 50	31	5			X		X	X						X				X						X
A3A+RNaseH 51	42	4		X	X		X									X								
A3A+RNaseH 52	50	5			X		X		X					X		X								X
A3A+RNaseH 53	48	0																						
A3A+RNaseH 54	12	0																						X
A3A+RNaseH 55	91	4				X	X							X						X				
A3A+RNaseH 56	55	0																						
A3A+RNaseH 57	68	1									X													
A3A+RNaseH 58	51	2			X	X																		X
A3A+RNaseH 59	62	5		X	X		X				X			X										
A3A+RNaseH 60	50	0																						X
A3A+RNaseH 61	49	0																						
A3A+RNaseH 62	51	1									X													
A3A+RNaseH 63	79	4		X	X	X					X													X
A3A+RNaseH 64	57	6		X	X	X	X							X				X						
A3A+RNaseH 65	52	4					X				X			X					X					X
A3A+RNaseH 66	81	3			X					X				X										
A3A+RNaseH 67	93	2									X					X								X
A3A+RNaseH 68	58	3				X	X	X																X
A3A+RNaseH 69	54	4			X			X			X			X										X
A3A+RNaseH 70	39	1														X								
A3A+RNaseH 71	107	6		X	X	X			X					X		X								
A3A+RNaseH 72	37	1												X										
A3A+RNaseH 73	62	4			X	X					X			X										
A3A+RNaseH 74	63	7		X	X	X					X			X		X		X						X
A3A+RNaseH 75	40	0																						
A3A+RNaseH 76	76	5					X				X				X			X			X			X
A3A+RNaseH 77	77	6			X		X							X		X		X				X		
A3A+RNaseH 78	100	0																						
A3A+RNaseH 79	50	4			X			X								X		X						X
A3A+RNaseH 80	60	1									X													X
A3A+RNaseH 81	38	0																						
A3A+RNaseH 82	43	1			X																			X
A3A+RNaseH 83	59	5				X	X							X				X				X		X
A3A+RNaseH 84	54	4					X							X	X	X								
A3A+RNaseH 85	28	3		X	X									X										X
A3A+RNaseH 86	69	5		X		X	X				X			X										
A3A+RNaseH 87	55	7	1	X	X	X	X				X			X				X					A64T	X
A3A+RNaseH 88	32	5		X		X	X							X				X						X
A3A+RNaseH 89	55	2								X	X													X
A3A+RNaseH 90	69	4		X		X	X							X										
A3A+RNaseH 91	64	8			X	X	X		X		X			X		X		X						
A3A+RNaseH 92	62	4				X	X	X			X													
A3A+RNaseH 93	104	2				X	X																	
A3A+RNaseH 94	60	7	1		X	X		X	X	X				X		X							T97A	
A3A+RNaseH 95	53	3			X										X	X								
A3A+RNaseH 96	44	2									X			X										
A3A+RNaseH 97	56	5			X		X	X			X			X										
A3A+RNaseH 98	53	3		X	X									X										
A3A+RNaseH 99	47	4			X			X								X		X						
A3A+RNaseH 100	38	8		X	X	X	X				X			X		X		X						
			TOTAL per position:	26	52	35	45	23	5	5	40	1	1	55	5	27	1	26	2	2	1	3		
				AGA	TGA	AGA	TGA	TGA	TGG	GGA	AGA	TGC	AGT	TGA	TGT	TGA	TGT	TGA	TGC	TGT	TGC	AGT		50

Figure 2—figure supplement 4. Summary of LEAP products generated in the presence of rA3A and RNase H.
DOI: 10.7554/eLife.02008.009

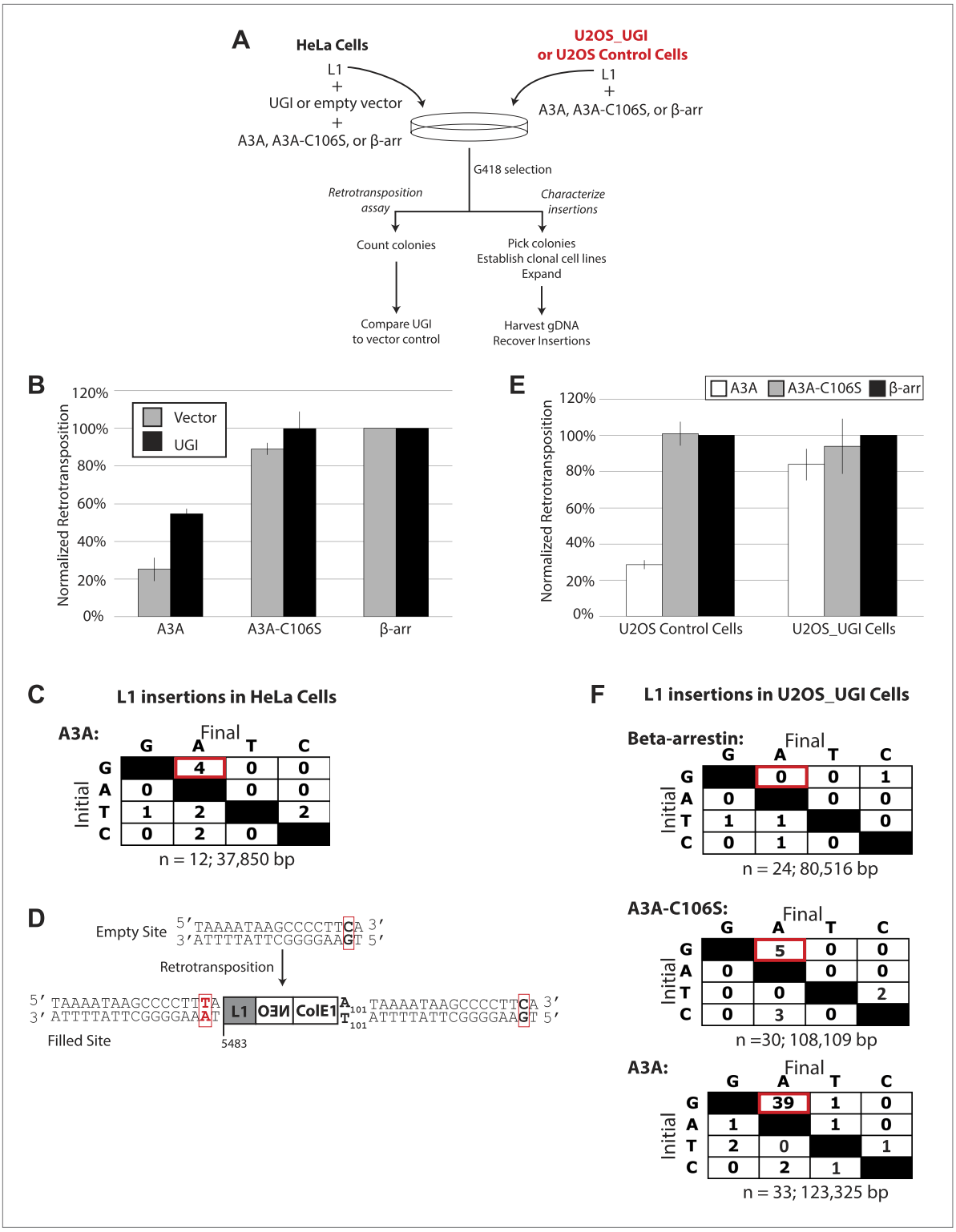


Figure 3. UGI expression alleviates A3A-mediated inhibition of L1 retrotransposition and allows detection of deaminated L1 retrotransposition events. (**A**) *Experimental strategy*: cells (HeLa, upper left, or U2OS, upper right) were co-transfected with the indicated expression plasmids and assayed for L1 retrotransposition (bottom left branch) or insertion analysis (bottom right branch). (**B**) *UGI expression alleviates A3A-mediated retrotransposition inhibition*: the x-axis indicates the co-expression vector. The y-axis depicts the efficiency of L1 retrotransposition. Shown are results of experiments in the presence (black bars) or absence (gray bars) of UGI. The results were normalized as in ***Figure 1—figure supplement 1A***. Data are expressed as the mean percent retrotransposition derived from three independent experiments consisting of two technical replicates each, with error bars representing

Figure 3. Continued on next page

Figure 3. Continued

the standard deviation among all six technical replicates. (**C**) *DNA sequencing results*: the left column indicates the engineered L1 sequence. The top row indicates sequence changes observed in recovered L1 insertions; n indicates the number of characterized retrotransposition events. The total retrotransposed sequence observed (in bp) includes the L1 sequence and the *mneoI/ColE1 cassette*. (**D**) *An L1 insertion harboring mismatched TSDs*: The pre- (empty) and post- (filled) L1 integration sites are shown. The truncation point (bp 5483) and structure of the engineered retrotransposed L1 are indicated. Deamination of single-strand genomic DNA in the pre-integration site (red rectangle) leads to inexact TSDs in the post-integration site (red lettering/rectangles). (**E**) *Stable UGI expression alleviates A3A-mediated L1 inhibition*: The x-axis indicates the U2OS cell line. The y-axis depicts the efficiency of L1 retrotransposition. Shown are the effects of wild-type A3A (white bars), A3A-C106S (gray bars), and β-arrestin (β-arr, black bars) control on L1 retrotransposition. Data were normalized to controls conducted with a circular *NEO* expression vector (***Figure 3—figure supplement 1 B–C***). Data are expressed as the mean percent retrotransposition derived from three independent experiments consisting of two technical replicates each, with error bars representing the standard deviation among all six technical replicates. (**F**) *DNA sequencing results*: The left column indicates the engineered L1 sequence. The top row indicates sequence changes observed in recovered L1 insertions; n indicates the numbers of characterized retrotransposition events. The total retrotransposed sequence observed (in bp) includes the L1 sequence and the *mneoI/ColE1 cassette*.

DOI: 10.7554/eLife.02008.010

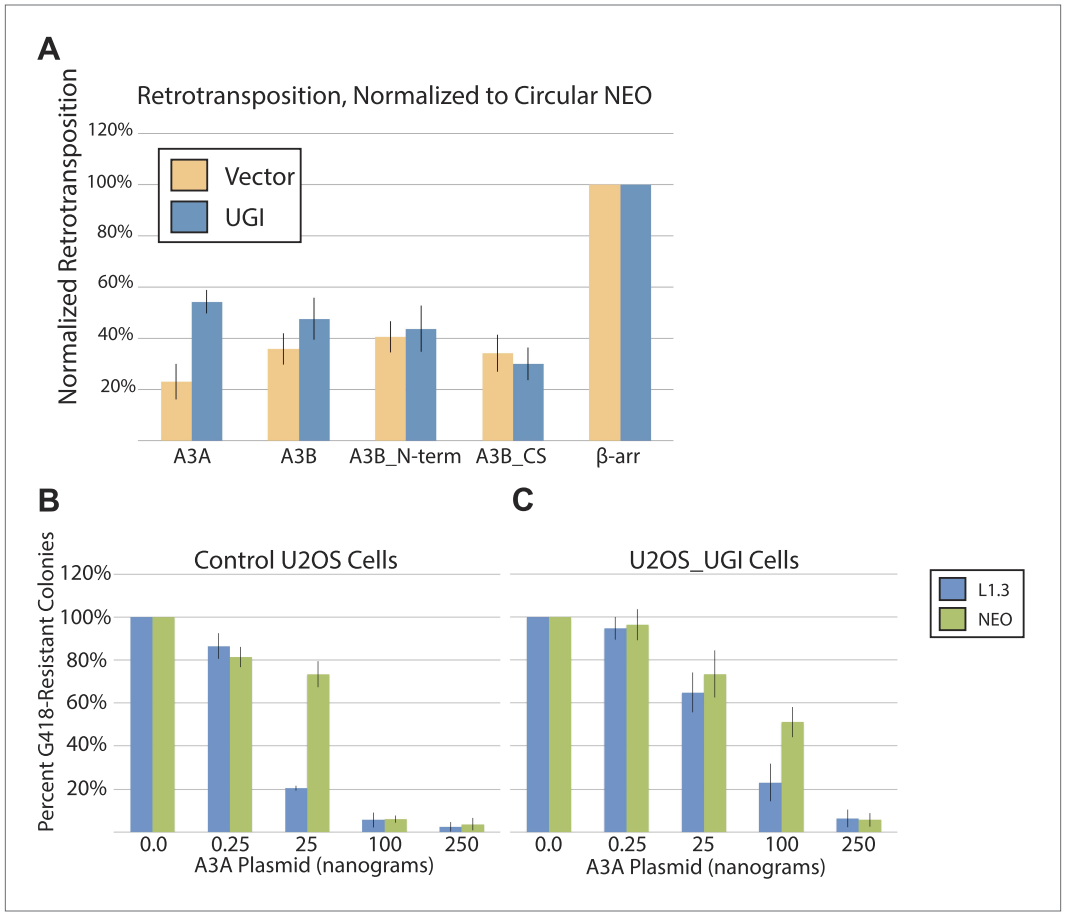


Figure 3—figure supplement 1. Additional Control Experiments.

DOI: 10.7554/eLife.02008.011

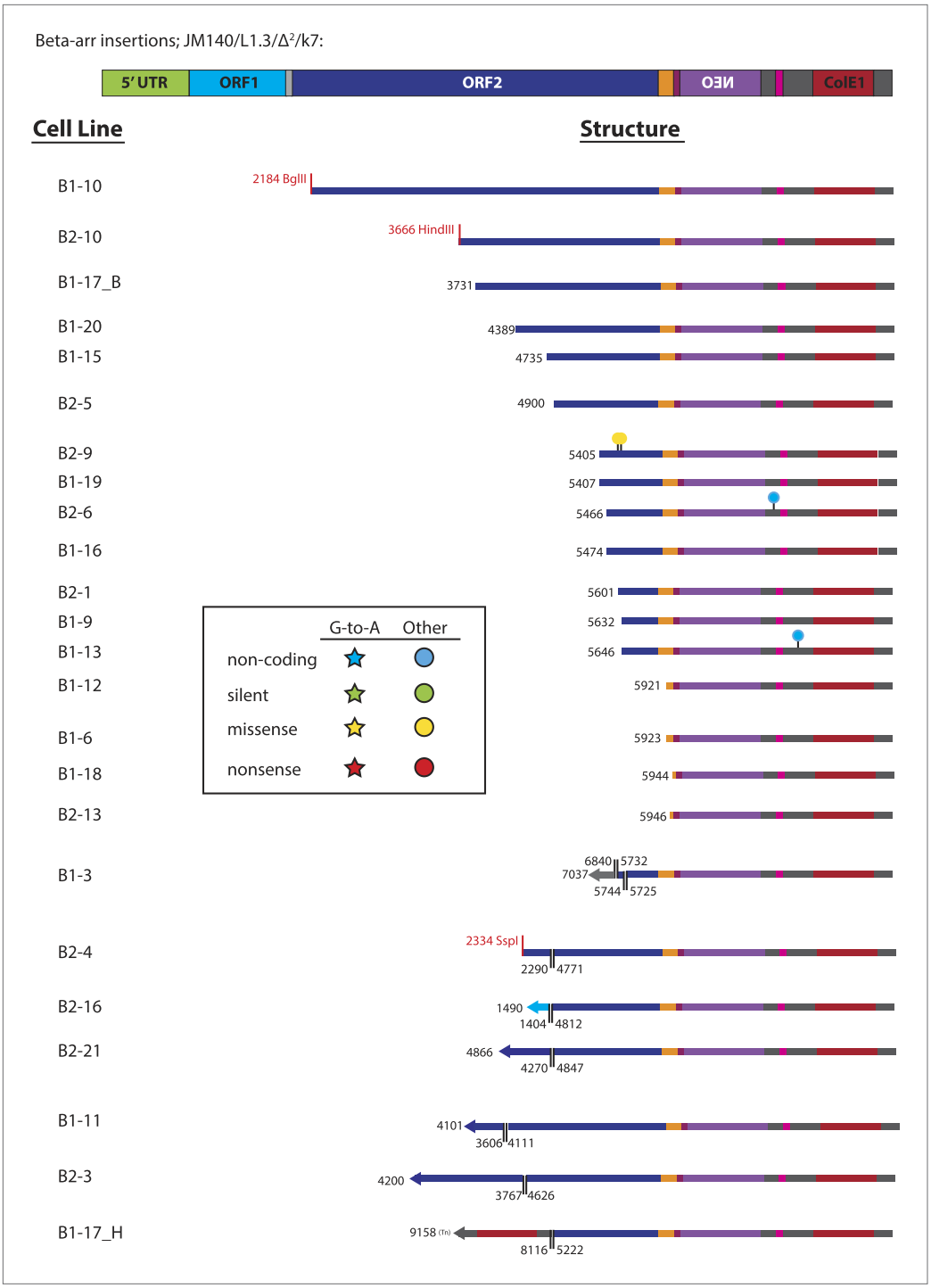


Figure 3—figure supplement 2. Summary of L1 retrotransposition events generated in U2OS_UGI cells in the presence of β-arrestin.

DOI: 10.7554/eLife.02008.012

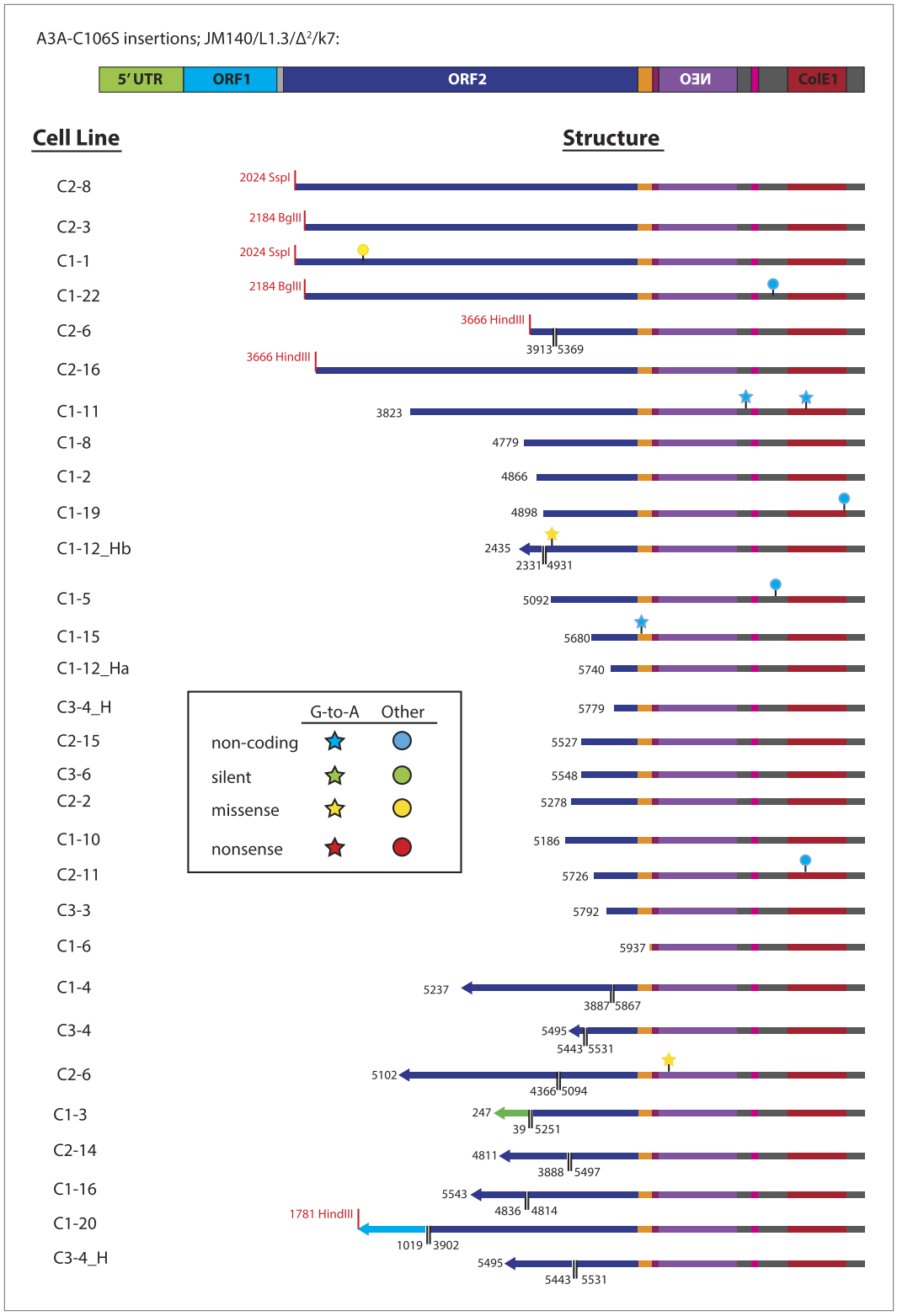


Figure 3—figure supplement 3. Summary of L1 retrotransposition events generated in U2OS_UGI cells in the presence of A3A_C106S.

DOI: 10.7554/eLife.02008.013

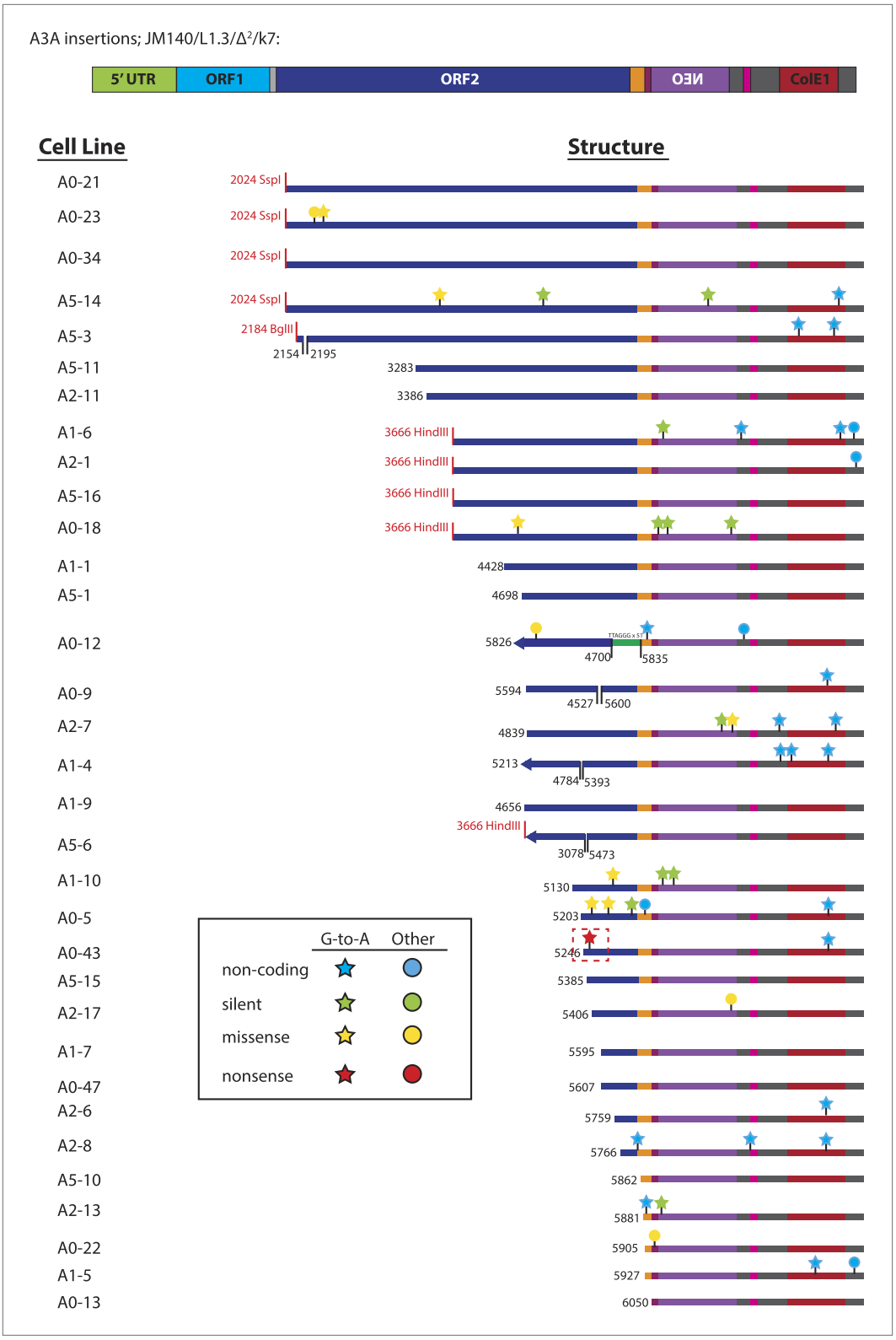


Figure 3—figure supplement 4. Summary of L1 retrotransposition events generated in U2OS_UGI cells in the presence of A3A.

DOI: 10.7554/eLife.02008.014

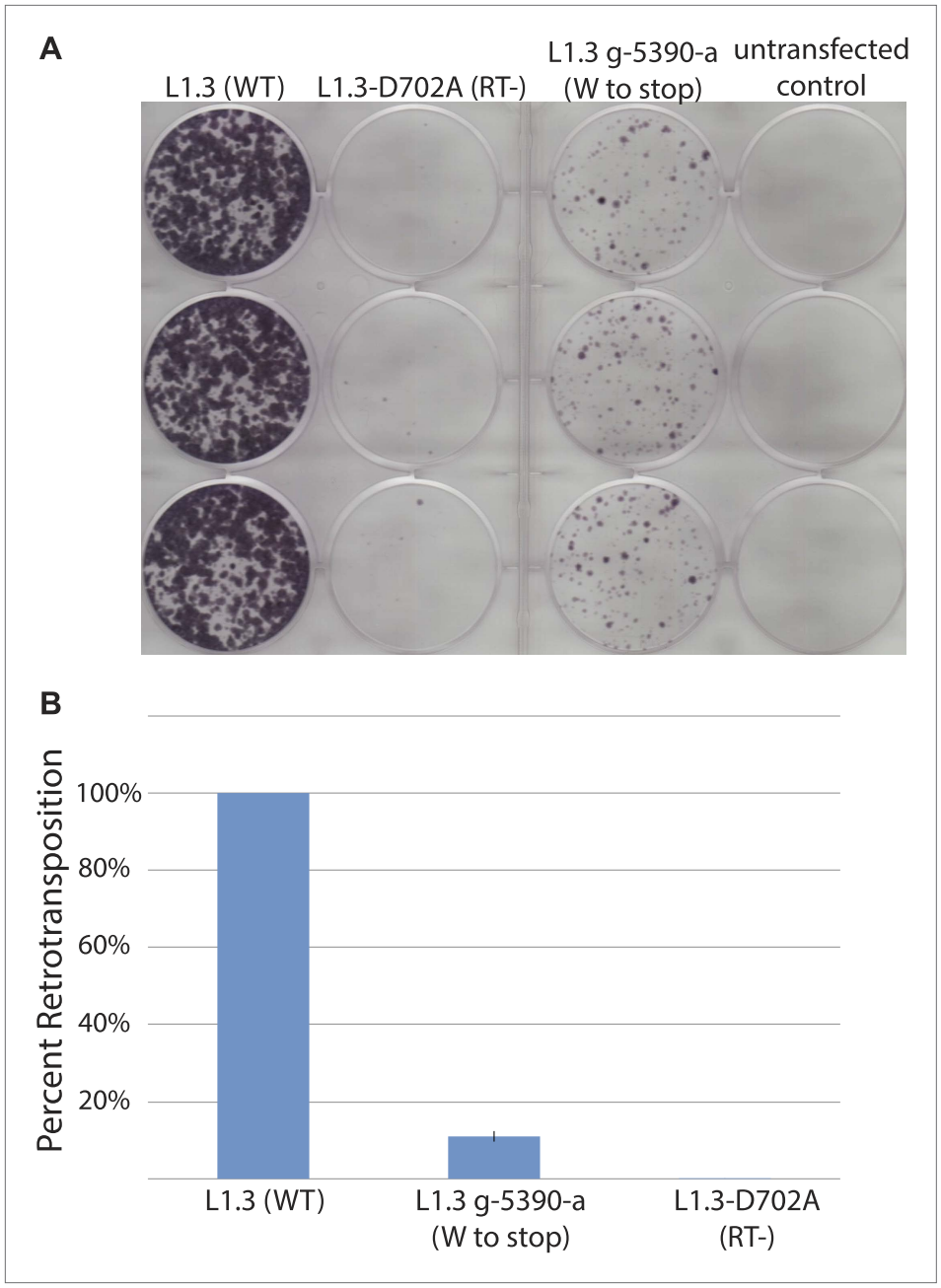


Figure 3—figure supplement 5. The L1.3 G-5390-A mutation decreases L1 retrotransposition efficiency.
DOI: 10.7554/eLife.02008.015

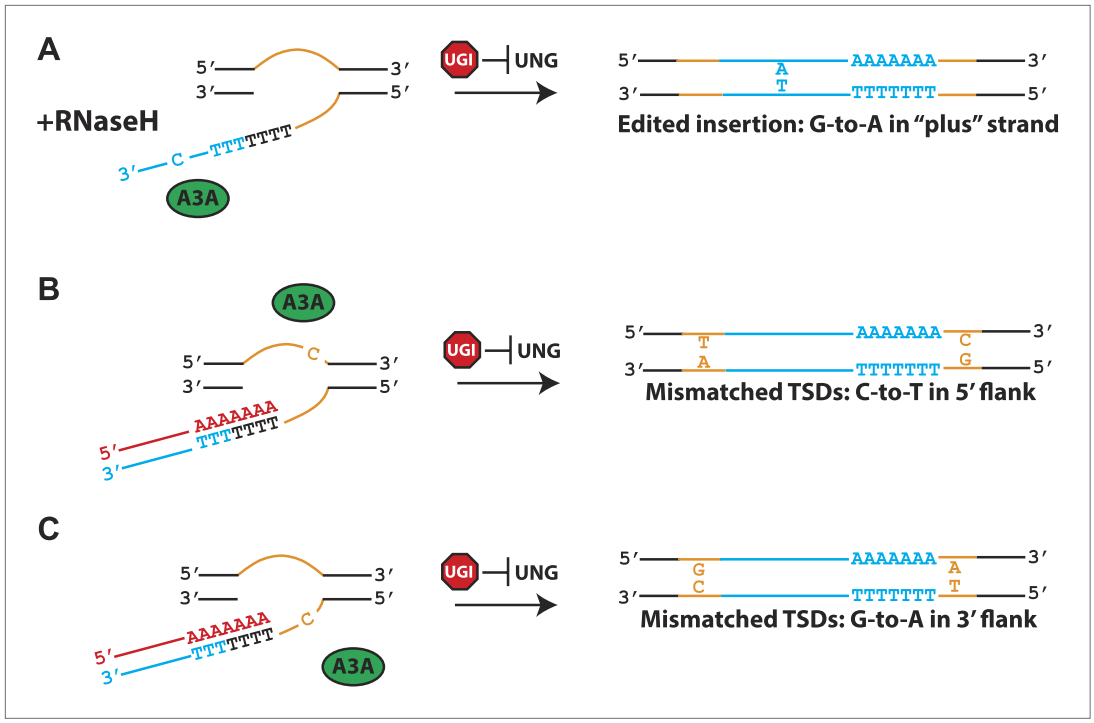


Figure 4. A model for A3A-mediated inhibition of L1 retrotransposition. The left side of the Figure shows L1 integration reactions that occur via TPRT. The right side of the Figure shows the predicted structures of the resultant L1 retrotransposition events. Shown are transiently exposed single-strand genomic DNA regions that ultimately give rise to target site duplications (TSDs: orange lines), the L1 RNA (red line), the L1 cDNA (blue line), and the A3A protein (green oval). UGI expression (red stop sign) can inhibit UNG activity. (**A**) *Deamination of the L1 cDNA*: in the presence of cellular RNase H, A3A-mediated deamination of the L1 (−) strand cDNA (blue lettering) in the presence of UGI leads to C-to-T mutations on the L1 non-coding strand and G-to-A mutations on the L1 (+) coding strand. (**B**) *Deamination of the 5′ flanking top-strand genomic DNA*: in the presence of UGI, deamination of the transiently exposed single-strand 5′ flanking genomic DNA (top orange line) during TPRT results in C-to-T changes in the 5′ TSD relative to the 3′ TSD. (**C**) *Deamination of the 3′ flanking bottom-strand genomic DNA*: in the presence of UGI, deamination of transiently exposed single-strand 3′ flanking genomic DNA during TPRT, in principle, is predicted to result in a G-to-A change in the 3′ TSD relative to the 5′ TSD.

DOI: 10.7554/eLife.02008.016