



Group	U3RU5	HIV-SD	SAPPT
Anova	0.0097	< 0.0001	< 0.0001
PGK vs MutSD	ns	0.001	ns
PGK vs Mut1_13	ns	ns	< 0.0001
PGK vs Mut14-15	ns	ns	ns
MutSD vs Mut1_13	ns	0.01	< 0.0001
MutSD vs Mut14-15	ns	< 0.0001	0.01
Mut1_13 vs Mut14-15	0.05	0.001	< 0.0001

Supplemental Figure 1. Gene expression analyses of specific LV backbone positions on canonical and recoded vectors.

RT-qPCR results on transcription levels of different LV backbone portions performed on RNA extracted from JY cells transduced with SIN.LV.PGK and the recoded vectors MutSD, Mut 1_13 and Mut 14_15.

Each graph shows the ΔC_t obtained with the probe sets U3RU5 HIV-SD and SA-PPT (indicated) normalized against GFP. The table below shows the comparative results of the ANOVA analysis.

Supplemental Table 1

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
CD34	DWLUSA	5' cell_exon - LV.exon_2	ex_3	6	43446456	-	<i>TJAP1</i>	TTGGTA
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_8	2	71597013	+	<i>ZNF638</i>	AGGCTT
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_8	3	63819875	-	<i>THOC7</i>	CATCTG
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_6	4	190878553	+	<i>FRG1</i>	AGGGGA
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	7	35912276	+	<i>SEPT7</i>	AGGTGG
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_2	7	74047983	+	<i>AY927540</i>	AGGTCT
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_11	11	17336933	+	<i>NUCB2</i>	AGGTTG
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_7	18	33270976	+	<i>GALNT1</i>	AGATTT
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_25	18	39661087	+	<i>PIK3C3</i>	AGTACT
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_5	19	13051057	+	<i>CALR</i>	AGGATG
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_3	20	32659872	+	<i>RALY</i>	AGGTGG
CD34	UPLVSD	LV.exon_1 - 3' cell_exon	ex_3	22	42071200	-	<i>NHP2L1</i>	GGCTGC
CD34	UPLVSD	LV.exon_1 - 3' cell_intron	intr_6	17	62147124	+	<i>ERN1</i>	AGGTCT
CD34	UPLVSD	LV.exon_1 - LV.exon_2 - 3' cell_exon	ex_10	2	114709297	+	<i>ACTR3</i>	AGAATA
CD34	UPLVSD	LV.exon_1 - LV.exon_2 - 3' cell_exon	ex_8	2	174946789	-	<i>OLA1</i>	ATCAAC
CD34	UPLVSD	LV.exon_1 - LV.exon_2 - 3' cell_exon	ex_5	5	154339540	+	<i>MRPL22</i>	AGGTTC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_3	1	183775619	-	<i>RGL1</i>	GGGGTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_8	2	85770157	-	<i>MAT2A</i>	CAGGTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_2	6	31590678	-	<i>PRRC2A</i>	CTGGTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_8	7	6624891	-	<i>ZDHHC4</i>	CAGGTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_2	7	100473333	-	<i>SRRT</i>	CAGGTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_19	9	131348127	-	<i>SPTAN1</i>	GAGGAC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_5	10	98704002	-	<i>LCOR</i>	AAGGTT
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_10	11	32990689	-	<i>QSER1</i>	AAGGTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_16	15	67995746	-	<i>MAP2K5</i>	GAGGTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	16	2807946	-	<i>SRRM2</i>	CAGGTA

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_5	16	15771806	-	<i>NDE1</i>	GCG GTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_2	16	29828613	-	<i>C16orf53</i>	CAG GTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_23	16	30522249	-	<i>ITGAL</i>	GGGG AC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_6	16	50788335	-	<i>CYLD</i>	CAG GTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_13	19	10793272	-	<i>ILF3</i>	GAG GAC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_14	19	11114073	-	<i>SMARCA4</i>	GAG GTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	19	52085135	-	<i>ZNF175</i>	AAG G TG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_3	1	32377297	+	<i>PTP4A2</i>	ACCTTC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_1	1	45987281	+	<i>PRDX1</i>	GCCTTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_2	3	141811903	+	<i>TFDP2</i>	ACCTTT
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_7	7	37272700	+	<i>ELMO1</i>	ACCTTC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_3	7	130630479	+	<i>FLJ43663</i>	ACCTCA
JY	DWLVSAs	5' cell_exon - LV.exon_1 - LV.exon_2	ex_3	7	130630479	+	<i>FLJ43663</i>	ACCTCA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_12	10	75138668	+	<i>ANXA7</i>	ACCTCAC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_1	11	47869400	+	<i>NUP160</i>	ACCAAT
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_2	12	56579939	+	<i>SMARCC2</i>	ACCATC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_28	12	133233722	+	<i>POLE</i>	ACCTGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_28	12	133233722	+	<i>POLE</i>	ACCTGT
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_24	17	1565252	+	<i>PRPF8</i>	TCCACC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_6	17	5377969	+	<i>DERL2</i>	TCCTTC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_1	17	76834669	+	<i>USP36</i>	ACCTGG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_3	17	79813018	+	<i>P4HB</i>	ACCTTC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	17	79813018	+	<i>P4HB</i>	ACCTTC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	19	2046185	+	<i>MKMK2</i>	ACCTTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_6	19	3981337	+	<i>EEF2</i>	ACCTTC
JY	DWLVSAs	5' cell_exon-LV.exon_1a - LV.exon_2	ex_6	19	3981337	+	<i>EEF2</i>	ACCTTC
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_5	19	40587726	+	<i>ZNF780A</i>	ACCTGG

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	19	45922356	+	<i>ERCC1</i>	ACCACA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	X	20213183	+	<i>RPS6KA3</i>	ACCATA
JY	DWLUSA	5' cell_exon - LV.exon_1 - LV.exon_2	ex_4	17	80529746	-	<i>FOXK2</i>	CAGGTA
JY	DWcrypSA	5' cell_intr - LV.exon_1	intr_2	9	37251841	-	<i>ZCCHC7</i>	CTGTGC
JY	DWcrypSA	5' cell_intr - LV.exon_1	intr_7	2	53928525	+	<i>GPR75-ASB3</i>	AGCTGA
JY	DWcrypSA	5' interg - LV.exon_1	interg	X	74965307	+	<i>None</i>	TCCTGG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_1	1	160681472	-	<i>CD48</i>	ACCTTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_2	4	184577034	-	<i>RWDD4A</i>	ACCCTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_5	12	7344296	+	<i>PEX5</i>	CAGGTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	16	1779579	+	<i>MAPK8IP3</i>	GAGGTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	16	67116211	+	<i>CBFB</i>	GAGGTG
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_10	19	10748991	+	<i>SLC44A2</i>	CTGGTA
JY	DWcrypSA	5' cell_exon - LV.exon_1	ex_4	21	17909759	+	<i>LINC00478</i>	CAGGTA
JY	DWLUSA	5' cell_exon - LV.exon_2	ex_2	2	208425883	-	<i>CREB1</i>	ACAGGT
JY	DWLUSA	5' cell_exon - LV.exon_2	ex_1	10	98031109	-	<i>BLNK</i>	ACCTCA
JY	DWLUSA	5' cell_exon - LV.exon_2	ex_41	X	153580253	+	<i>FLNA</i>	CCTGGC
JY	DWLUSA	5' cell_exon - LV.exon_3	ex_4	16	24981828	-	<i>ARHGAP17</i>	ACCCCA
JY	DWLUSA	5' cell_exon - LV.exon_4	ex_1	20	61847484	-	<i>YTHDF1</i>	ACAGCA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_4	1	44280605	-	<i>ST3GAL3</i>	ACTGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_5	2	28210954	-	<i>BRE</i>	CAGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_3	2	152267852	-	<i>RIF1</i>	AAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_33	6	33648457	-	<i>ITPR3</i>	CAGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_9	6	43573056	-	<i>POLH</i>	GATGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_2	11	66888829	-	<i>KDM2A</i>	TTGGTT
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_11	11	66995634	-	<i>KDM2A</i>	TGGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_3	14	103894777	-	<i>MARK3</i>	AAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_4	16	89628799	-	<i>RPL13</i>	AACGTA

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_11	19	14512572	-	CD97	CAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_6	20	2903931	-	PTPRA	AAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_2	22	50810579	-	SAPS2	AAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_1	22	50946874	-	NCAPH2	GAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_1	1	3563972	+	WDR8	ACCTGC
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_1	1	160681472	+	CD48	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_1	1	236433175	+	ERO1LB	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_3	2	55445008	+	C2orf63	CAAACC
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_22	2	55871772	+	PNPT1	ACCTGT
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_7	3	47755905	+	SMARCC1	ACCTTC
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_5	7	72903524	+	BAZ1B	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_1	8	117738255	+	EIF3H	ACCTTC
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_22	15	51778270	+	DMXL2	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_5	15	77344727	+	TSPAN3	ACCTGA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_12	17	2075952	+	SMG6	ACCTTA
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_12	17	2075952	+	SMG6	ACCTTA
JY	DWcrypSA	5' cell_exon-LV.exon_1a	ex_8	19	36386882	+	NFKBID	ACCTGG
JY	DWLVSA	5' cell_exon-LV.exon_1a - LV.exon_2	ex_11	12	110600870	-	IFT81	GAGGTA
JY	DWLVSA	5' cell_exon-LV.exon_1a - LV.exon_2	ex_4	12	113698297	-	TPCN1	CAGGTG
JY	DWLVSA	5' cell_exon-LV.exon_1a - LV.exon_2	ex_9	19	6745005	-	TRIP10	AAGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_15	1	161882142	-	ATF6	TCAATT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	1	186349016	-	C1orf27	CAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	2	65283662	-	CEP68	TAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	3	47437707	-	PTPN23	CAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	3	113604431	-	GRAMD1C	AAGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	4	40198920	-	RHOH	CAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	4	44682859	-	GUF1	CCGGTA

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	4	87593696	-	<i>PTPN13</i>	AAG G TA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_5	5	60817269	-	<i>ZSWIM6</i>	AAG G TG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	5	149915447	-	<i>NDST1</i>	AT G TG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_26	5	169230210	-	<i>DOCK2</i>	G C G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	8	6590171	-	<i>AGPAT5</i>	CC A G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_7	8	27255305	-	<i>PTK2B</i>	CG G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	8	145515556	-	<i>HSF1</i>	CC G G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_10	9	139313558	-	<i>PMPCA</i>	AC G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	9	139983457	-	<i>MAN1B1</i>	C A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_13	10	21971186	-	<i>MLLT10</i>	AC G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	10	75506762	-	<i>SEC24C</i>	A A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_10	11	250865	-	<i>PSMD13</i>	G A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	11	67197066	-	<i>RPS6KB2</i>	A A G G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	11	72527872	-	<i>ATG16L2</i>	CT G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	11	121175202	-	<i>SC5DL</i>	AT G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_5	14	60758172	-	<i>PPM1A</i>	ACT G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	14	64375945	-	<i>SYNE2</i>	A A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	14	96994343	-	<i>PAPOLA</i>	G A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	15	76146828	-	<i>UBE2Q2</i>	TT G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	16	27414541	-	<i>IL21R</i>	AT G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_19	17	7407110	-	<i>POLR2A</i>	ATT G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	17	25877734	-	<i>KSR1</i>	C A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	17	79852462	-	<i>ANAPC11</i>	ACT G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	19	1244184	-	<i>ATP5D</i>	GG G G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	19	7070664	-	<i>ZNF557</i>	C A G G T A
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_7	19	13051701	-	<i>CALR</i>	C A G G T G
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	19	17671274	-	<i>GLT25D1</i>	CT G G T A

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_7	19	17877619	-	<i>FCHO1</i>	AAGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_5	19	49841286	-	<i>CD37</i>	CAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	19	50185516	-	<i>PRMT1</i>	ACGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	20	35812776	-	<i>RPN2</i>	AAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	20	35993680	-	<i>SRC</i>	CAGGTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	22	3981485	-	<i>TAB1</i>	TGGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_5	22	50633469	-	<i>TRABD</i>	CAGGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	1	1756836	+	<i>GNB1</i>	ACTCGA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_10	2	48059488	+	<i>FBXO11</i>	ACCATG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_11	2	160731955	+	<i>LY75-CD302</i>	ACCTTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	2	230744698	+	<i>TRIP12</i>	ACCTTC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	3	49129830	+	<i>QRICH1</i>	ACCAGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_16	4	68510416	+	<i>UBA6</i>	ACCATT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_9	4	106307724	+	<i>PPA2</i>	ACCGAT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_8	6	30675161	+	<i>MDC1</i>	ACCCTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	8	80954855	+	<i>TPD52</i>	GCCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	8	145979068	+	<i>ZNF251</i>	ACCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	9	114213705	+	<i>KIAA0368</i>	ACCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_7	9	132689475	+	<i>FNBP1</i>	CATTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	9	132691846	+	<i>FNBP1</i>	ACCTGG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	9	138758302	+	<i>CAMSAP1</i>	ACCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	9	139292751	+	<i>SNAPC4</i>	ACCTGC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	10	101933965	+	<i>ERLIN1</i>	ACCTGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	10	104465103	+	<i>ARL3</i>	ACCTGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	11	46842726	+	<i>CKAP5</i>	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	11	62356503	+	<i>TUT1</i>	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	intr_2	11	110150054	+	<i>RDX</i>	TATGGT

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_8	12	53589158	+	<i>ITGB7</i>	ACATTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	12	56578628	+	<i>SMARCC2</i>	ACCTGG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	12	57029966	+	<i>BAZ2A</i>	ACCCAT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	12	109072048	+	<i>CORO1C</i>	ACCATG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	13	95243107	+	<i>TGDS</i>	ACCTAC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	14	91875003	+	<i>CCDC88C</i>	ACCTGG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	15	59204762	+	<i>SLTM</i>	ACTTGG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	15	62325640	+	<i>VPS13C</i>	ACCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_8	15	76566753	+	<i>ETFA</i>	ACTGGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	16	786270	+	<i>NARFL</i>	ATTTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	16	11824501	+	<i>TXNDC11</i>	ACCTCG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	16	74640555	+	<i>GLG1</i>	ACCTCC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_19	16	89845209	+	<i>FANCA</i>	ACCTCT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	16	89880933	+	<i>FANCA</i>	TAAATG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_6	17	7504727	+	<i>FXR2</i>	ACCTCT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_12	17	37421650	+	<i>FBXL20</i>	ACCAAT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	17	38020373	+	<i>IKZF3</i>	ACCTTC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	17	42400652	+	<i>SLC25A39</i>	ACCGCT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_2	17	76704279	+	<i>CYTH1</i>	ACCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	17	80583265	+	<i>WDR45L</i>	ACCTTT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	19	2085174	+	<i>MOB3A</i>	ACCGTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_8	19	2123829	+	<i>AP3D1</i>	ACCTGG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_4	19	2132470	+	<i>AP3D1</i>	ACCAGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	19	2137010	+	<i>AP3D1</i>	ACCTTA
JY	DWLUSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_3	19	2137010	+	<i>AP3D1</i>	ACCTTA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	19	2520606	+	<i>GNG7</i>	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_7	19	2761458	+	<i>SGTA</i>	ACCATG

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	19	3061158	+	AES	ACCTGT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_11	19	3977817	+	EEF2	ACCTCC
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_18	19	10465188	+	TYK2	ACCTCG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	19	10490291	+	TYK2	ACCTGA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_9	19	46133209	+	EML2	ACCTTG
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_3	19	52588040	+	ZNF841	ACCTGA
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_1	22	50733148	+	PLXNB2	ACCCCT
JY	DWcrypSA	5' cell_exon-LV.exon_1b	ex_14	X	153592390	+	FLNA	ACCCCTG
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_2	11	58352357	-	ZFP91-CNTF	AAGGTA
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_13	19	4433536	-	CHAF1A	CAGGTG
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_11	19	6750077	-	TRIP10	GAGGTC
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_5	19	34859607	-	GPI	CAGGTA
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_4	X	123159768	-	STAG2	AAAGTA
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_2	13	113219425	+	TUBGCP3	ACCTGT
JY	DWLVSA	5' cell_exon-LV.exon_1b - LV.exon_2	ex_6	17	17782941	+	TOM1L2	ACCTGC
JY	DWLVSA	5' cell_intr - LV.exon_2	intr_1	16	28220309	+	XPO6	CAATGG
JY	DWcrypSA	5' cell_intr-LV.exon_1a	intr_30	15	43831216	+	HISPPD2A	GCCTGA
JY	DWcrypSA	5' cell_intr-LV.exon_1a	intr_19	16	2228662	+	CASKIN1	GGGCAA
JY	DWcrypSA	5' cell_intr-LV.exon_1a	intr_2	17	28088168	+	SSH2	ACAGCT
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_1	4	83957184	-	COPS4	TAGGTA
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_4	17	5114426	-	LOC100130950	CAAGTG
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_1	17	5200809	-	RABEP1	CAGGTC
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_1	6	29526882	+	UBD	ACCCCT
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_4	12	116518434	+	MED13L	ACCTTC
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_6	15	70366842	+	TLE3	ACGGTA
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_1	16	28220308	+	XPO6	ACAATG
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_1	16	67750541	+	GFOD2	ACCTTC

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	5' cell_intr-LV.exon_1b	intr_2	19	2609424	+	<i>GNG7</i>	ACCCTG
JY	DWLUSA	5' interg - LV.exon_2	interg	7	14079487	+	<i>None</i>	TCTCTC
JY	DWcrypSA	5' interg-LV.exon_1a	interg	5	4242959	+	<i>None</i>	AGAAAA
JY	DWcrypSA	5' interg-LV.exon_1a	interg	14	106329408	+	<i>None</i>	ACCTGA
JY	DWcrypSA	5' interg-LV.exon_1b	interg	1	9690282	-	<i>None</i>	CAGGTG
JY	DWcrypSA	5' interg-LV.exon_1b	interg	11	6767580	+	<i>None</i>	ACCTGT
JY	DWLUSA	5' interg-LV.exon_1b - LV.exon_2	interg	2	119620369	+	<i>None</i>	GAAAAA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_3	1	6253114	-	<i>RPL22</i>	CTCCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_14	2	148693267	-	<i>ORC4L</i>	AGCCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_7	3	189702379	-	<i>LEPREL1</i>	ATTCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	10	97023820	-	<i>PDLIM1</i>	CTCCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_7	12	57058584	-	<i>PTGES3</i>	ATCCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	16	2013257	-	<i>RPS2</i>	TTCCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_17	16	3808049	-	<i>CREBBP</i>	GTCCTT
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_3	16	29572972	-	<i>LOC440354</i>	TCTCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_9	17	17158259	-	<i>COPS3</i>	GGTCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_11	17	45219802	-	<i>CDC27</i>	CTCCTG
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_9	18	28722041	-	<i>DSC1</i>	GGATTT
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_2	21	26802321	-	<i>LINC00158</i>	TTTCTA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	1	6931817	+	<i>CAMTA1</i>	AGAGAT
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	1	8029405	+	<i>PARK7</i>	AGGGAC
JY	UPLVSD	LV.exon_1 - 3' cell_exon	3' UTR	1	53459285	+	<i>SCP2</i>	AAAATA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_2	2	61118818	+	<i>REL</i>	AGGTGC
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	3	157920861	+	<i>RSRC1</i>	AGGTCC
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_5	4	39776454	+	<i>UBE2K</i>	AGGGCA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_5	4	41946810	+	<i>TMEM33</i>	AGGCAA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	3' UTR	4	140311455	+	<i>NAA15</i>	AGGTTT

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_6	5	113822747	+	KCNN2	AGGTAA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_23	10	91505656	+	KIF20B	AGGAGC
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_13	11	252505	+	PSMD13	AGATCA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_3	16	3117378	+	IL32	AGGTCC
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_5	16	89629292	+	RPL13	AGGTCT
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_5	17	2224600	+	SRR	AGTCCA
JY	UPLVSD	LV.exon_1 - 3' cell_exon	ex_4	X	154456666	+	VBP1	AGGAGT
JY	UPLVSD	LV.exon_1 - 3' cell_intron	intr_2	6	41840054	-	USP49	CAGCTA
JY	UPLVSD	LV.exon_1 - 3' cell_intron	intr_2	16	10892858	-	FAM18A	TCCCTA
JY	UPLVSD	LV.exon_1 - 3' cell_intron	intr_3	6	124969798	+	NKAIN2	AGTTCA
JY	UPLVSD	LV.exon_1 - 3' intergenic region	interg	6	48211380	-	None	AGTCTT
JY	UPLVSD	LV.exon_1 - 3' intergenic region	interg	9	72844305	-	None	GAACTG
JY	UPLVSD	LV.exon_1 - 3' intergenic region	interg	13	47666452	-	None	CCACTC
JY	UPLVSD	LV.exon_1 - 3' intergenic region	interg	3	142712839	+	None	AGCACG
JY	UPLVSD	LV.exon_1 - LV.exon_2 - 3' cell_exon	ex_13	3	32568400	-	DYNC1L1	GGCTA
JY	UPLVSD	LV.exon_1 - LV.exon_2 - 3' cell_exon	ex_3	17	47700238	-	SPOP	ACCCTG
JY	UPLVSD	LV.exon_1 - LV.exon_2 - 3' intergenic region	interg	7	10620311	-	None	AACTG
JY	UPcrypSD	LV.exon_2 - 3' cell_exon	ex_4	9	19093301	-	HAUS6	ACCCTA
JY	UPcrypSD	LV.exon_2 - 3' cell_exon	ex_17	16	56518779	-	BBS2	CAACTG
JY	UPcrypSD	LV.exon_2 - 3' cell_exon	ex_7	1	43166452	+	YBX1	AGGGGC
JY	UPcrypSD	LV.exon_2 - 3' cell_exon	ex_7	9	136217857	+	RPL7A	AGGGAA
JY	UPcrypSD	LV.exon_2 - 3' intergenic region	interg	7	45492743	+	None	AGGAAA
JY	DWcrypSA	LV.exon_3 - 3' cell exon	ex_60	1	19455568	-	UBR4	AGCTA
JY	DWcrypSA	LV.exon_3 - 3' cell exon	ex_10	1	85122167	-	SSX2IP	AAACTA
JY	DWcrypSA	LV.exon_3 - 3' cell exon	ex_9	2	17902528	-	SMC6	CCGCTT
JY	DWcrypSA	LV.exon_3 - 3' cell exon	ex_57	2	61459667	-	USP34	CCACTG
JY	DWcrypSA	LV.exon_3 - 3' cell exon	ex_14	3	195785503	-	TFRC	CACCTG

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	4	100477377	-	<i>RG9MTD2</i>	AAACTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_2	6	31760897	-	<i>VARS</i>	AGCCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_7	6	43533473	-	<i>XPO5</i>	GCCCTA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_11	6	146256577	-	<i>SHPRH</i>	GATCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_13	7	124481232	-	<i>POT1</i>	TGCCTA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_11	11	800666	-	<i>PIDD</i>	CACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_15	11	28045397	-	<i>KIF18A</i>	GAACTT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	11	65652657	-	<i>FIBP</i>	AGTCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	12	15097773	-	<i>ARHGDI1B</i>	CACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_2	12	46355640	-	<i>SCAF11</i>	CACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_7	12	100482873	-	<i>UHRF1BP1L</i>	GCTCTA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	15	65308865	-	<i>MTFMT</i>	GGGCTA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_6	17	1442246	-	<i>PITPNA</i>	CACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_2	17	2381179	-	<i>METTL16</i>	AGGCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_16	17	18152197	-	<i>FLII</i>	AACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_35	17	67246762	-	<i>ABCA5</i>	CATCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_48	17	76476868	-	<i>DNAH17</i>	TACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	17	80020849	-	<i>DUS1L</i>	AAACTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_13	17	80047048	-	<i>FASN</i>	CACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_13	19	4690969	-	<i>DPP9</i>	CATCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	19	6387570	-	<i>GTF2F1</i>	AACCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_1	19	18426875	-	<i>LSM4</i>	AAGCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	X	70329240	-	<i>IL2RG</i>	TTCCTT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_3	1	218478370	+	<i>RRP15</i>	AGCGTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_2	6	30876811	+	<i>GTF2H4</i>	AGGTGA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_2	6	31583298	+	<i>AIF1</i>	AGGAGG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_15	6	117890783	+	<i>DCBLD1</i>	AGGTTA

Cell Type	Primer set	Fusion Structure	Target	Chr	Chr Position	Str	Gene	Junction Seq
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	11	763344	+	TALDO1	AGGGAG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_3	11	118239346	+	UBE4A	AGATGA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_3	13	100172290	+	TM9SF2	AGGTTT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_36	14	102482193	+	DYNC1H1	AGATCC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_8	14	104139351	+	KLC1	AGGTTT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_14	16	722084	+	RHOT2	AGCCTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_16	16	2225302	+	TRAF7	AGGTGT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_13	16	2647124	+	PDPK1	AGGGTT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	16	24566936	+	RBBP6	AGACTG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	16	30093805	+	PPP4C	AGGTGT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_2	17	5211989	+	RABEP1	AGTTTC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_3	17	7245267	+	ACAP1	AGCTCC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	17	7604059	+	WRAP53	AGGTCC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	17	61869772	+	DDX42	AGCTAT
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	19	2245445	+	SF3A2	AGGGCC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_9	19	17273160	+	MYO9B	AGGCCA
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_5	19	50295192	+	AP2A1	AGGGAC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_16	19	54656604	+	CNOT3	AGGCAG
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_4	19	59058743	+	TRIM28	AGGGCC
JY	DWcrypSA	LV.exon_3 - 3'cell exon	ex_11	21	45736127	+	PFKL	AGACTC
JY	DWcrypSA	LV.intr_3 - 3'cell exon	intr_10	11	64963203	+	CAPN1	AGGTAA

Supplemental Table 1: List of aberrant transcripts identified by cLAM PCR

Cell type: indicates the cell type (cord blood CD34+ or JY) in which the aberrant transcript was retrieved.

Primer Set: indicates the primer set used (see figure).

Enz: Enzyme used in the cLAM-PCR reaction.

Fusion Structure: depicts the structure of the aberrant transcript. The '-' sign indicates splicing between the two surrounding elements. LV.exon_1 is defined by splice sites SA1 and SD1, LV.exon_2 is defined by splice sites SA2 and (SD4). Genomic regions can be exons, introns, or intergenic regions.

Target: indicates the genomic element targeted by the aberrant splicing event. UCSC genome browser and RefSeq database were used to determine the exact splicing target. Exons are referred to the targeted gene (see *Gene* column), when applicable.

Chr: Chromosome number

Chr Position: Chromosomal position of the nucleotide targeted by the splicing event corresponds to the sequence identify breakpoint between LV and human genome

Str: Orientation of the aberrant transcript with respect to the human genome (+, first chromosomal strand; -, second strand).

Gene: RefSeq symbol of the targeted gene. 'None' is used when sequences from intergenic regions are mapped.

Human genome GRCh37/hg19 assembly was used to map the aberrant transcripts.

Junction Seq: sequence of the exonic/intronic junction of the genomic portions of our fusion transcript. Highlighted in red are the canonical dinucleotide splice signal found in intron (GT, in case of donor splice site; AG, in case of acceptor splice site). In blue are highlighted the non-canonical splice signal.

Supplemental Table 2

SYSTEM	GENE CLASSES	COUNT	p VALUE	FOLD CHANGE
MF	N-methyltransferase activity	16	2.40E-06	4.3
MF	protein methyltransferase activity	15	6.00E-06	4.2
BP	protein amino acid alkylolation	13	8.90E-06	4.7
BP	protein amino acid methylation	13	8.90E-06	4.7
MF	histone methyltransferase activity	13	9.50E-06	4.7
BP	biopolymer methylation	15	2.60E-04	3.1
MF	protein-lysine N-methyltransferase activity	10	2.70E-04	4.4
MF	lysine N-methyltransferase activity	10	2.70E-04	4.4
MF	histone-lysine N-methyltransferase activity	10	2.70E-04	4.4
MF	arginine N-methyltransferase activity	4	9.90E-03	8.1
MF	protein-arginine N-methyltransferase activity	4	9.90E-03	8.1
BP	sulfur amino acid biosynthetic process	5	1.80E-02	4.7
BP	methionine biosynthetic process	4	2.10E-02	6.3
BP	central nervous system projection neuron axonogenesis	4	4.70E-02	4.7

Supplemental Table 2: Overrepresentation analysis of the gene targeted by LV integration sites in CD34+ and JY cell types.

Go analyses performed on cellular genes targeted by LV integration in CD34+ and JY cells transduced with SINLV.PGK. Gene classes significantly overrepresented are indicated. Column System; MF, Molecular Function; BP, Biological Process. Count: number of genes identified in the dataset belonging to each specific class. P value: P values < 0.05 are shown. Significant p values after Bonferroni correction are highlighted. Fold Change: Fold increase over the predicted random distribution.

Supplemental Table 3

Assay	PGK	Mut_SD	Mut 1_13	Mut 14_15
Vector Titer (TU/ml)	$2.3 \cdot 10^7$	$2.1 \cdot 10^6$	$2.7 \cdot 10^6$	$2.5 \cdot 10^7$
p24 (ng/ml)	330	460	160	270
Infectivity (TU/ng p24)	$7 \cdot 10^4$	$5 \cdot 10^3$	$1.6 \cdot 10^4$	$9.4 \cdot 10^4$
Single Copy MFI	7400± 400	7900± 300	7500± 160	7900± 300

Supplemental Table 3: Titer and Infectivity of the recoded vector.

Vector Titer: The titer is defined as number of transducing units per milliliter (TU/ml) of vector preparation.

It has been performed by end-point titration on 293T cell line.

p24: measure of the Gag capsid protein p24, it represents the total content of vector particles in the vector stock. It has been estimated by immune-capture ELISA assay.

Infectivity: transducing activity of the vector preparation *per* unit of physical particle.

Single Copy MFI: Mean Fluorescence Intensity of GFP transgene in 293T cells transduced @ single vector copy.

Supplemental Table 4

Assay	Forward Primer	Reverse Primer	Probe/ Context sequence
U3RU5	GACCCACCTCCCAACCC	TGTCTCTCTCTCCACCTTCTTCTTC	CAGGCCCGAAGGAATA
LV-FUSION	CAGGACTCGGCTTGCTGAA	GGTTGGGAGGTGGGTCAG	TCGCCGCCCTCGCC
U3RU5	NA	NA	GCTTCAAGTAGTGTGTGCCCGTCTG
SA-PPT	NA	NA	TCTCGACGGTATCGGTTAACTTTTA
HIV-SD	TACTGACGCTCTCGCACC	TCTCGACGCAGGACTCG	ATCTCTCTCCTTCTAGCCTC
B2M	NA	NA	AGTGGGATCGAGACATGTAAGCAGC
GFP	CAGCTCGCCGACCACTA	GGGCCGTCGCCGAT	CCAGCAGAACACCCCC

Supplemental Table 4: Sequences of the Primers and Probes used for the RT qPCR analyses