

Supplementary Information Online

Supplementary Table 1 online

Proviral integration site sequences detected by LAM-PCR.

Supplementary Table 2 online

Statistical analysis of common integration sites (CIS).

Legends

Supplementary Table 1 online Proviral integration site sequences detected by LAM-PCR. The genomic part of each sequenced retroviral integration site (RIS) was aligned to the human genome database using University of California, Santa Cruz BLAT search tools (<http://genome.ucsc.edu/>). The RefSeq gene next to the RIS is listed. Integrations that could not be mapped to the human genome (too short, no definite or multiple hits) are not shown. Pat., defines if the RIS is derived from a post transplantation sample of patient 1-5, the pre-engraftment sample from patient 6 or the normal donor (nd) control transduction sample; 'Source', analyzed cell fraction; 'Time', months post transplantation; 'Length', size of the LAM-PCR amplicon without linker- and LTR- sequences. 'Chromosome', exact location of the RIS in a chromosome; 'Identity', RIS was detected in different fractions or at different time points; 5' of TSS, RIS is located upstream of the RefSeq gene; 3' of TSS, RIS is located in the RefSeq gene; Intron, intron in which RIS is located; E, exon in which RIS is located; 3' of RefSeq gene, RIS is located downstream of the RefSeq gene; 'UCSC Locus', integration locus of the RIS according to the UCSC database; 'CpG island', reveals if a CpG island is located within 1kb up- or downstream of the RIS; 'Repetitive Elements', reveals if the RIS is located in a repetitive element and defines it; TSS, transcription start site; bp, base pairs; CD3, purified T cells; CD13, purified granulocyte/monocyte progenitor cells; CD15, purified granulocytes; CD34, hematopoietic stem- / progenitor cells; nd, normal donor; pre, pre-transplantation; mon, months post transplantation.

Supplementary Table 2 online Statistical analysis of common integration sites (CIS) occurrence. Computer simulations to compare the likelihood of random CIS occurrence on the haploid size of the human genome (3.12×10^9) with the number of CIS detected in our patients. CIS were defined as viral integration events located less than 30 kb apart from each other, independently of whether the loci were in genes or outside of gene coding regions. * p-Value, quota per 10000 simulation runs, in which the number of CIS was at least as high as in the random sampling. SD, standard deviation; MV, Median, Min, Max, mean value, median, minimum and maximum number of expected CIS according to simulations, respectively; CIS, indicates the number of common integration sites detected; Post, engrafted patient T cells; Pre-P6, patient 6 pre-transplant CD34+ cells; Pre-nd, healthy normal donor transduced CD34+ cells.

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
1	CD13	14	90	10q24.32	SUFU			100568	7		104028919	no	no
1	CD13	14	113	12p13.33	PRKWNK1		25138				707855	no	no
1	CD13	14	60	14q21.3	C14orf138					21567	48543827	no	no
1	CD13	14	53	3p21.32	ZNF445		33693				44513447	no	no
1	CD13	14	23	7p15.2	SCAP2					113304	26335658	no	no
						CD3, 20 mon							
1	CD13	14	72	8p21.2	ADAM28		14861				24158693	no	LINE
1	CD13	14	42	9q22.32	PTPDC1					48065	92261610	no	SINE
1	CD3	20	63	10p11.23	MAP3K8					36329	30791096	no	no
1	CD3	20	26	10q22.1	CDH23			332278	25		72833572	no	no
1	CD3	20	86	11p15.4	HBE1		9333				5265381	no	no
1	CD3	20	127	11q14.3	MTNR1B		144365				92246577	no	LINE
1	CD3	20	125	12p11.22	MLSTD1			17400	1		29285265	no	SINE
1	CD3	20	197	12p13.2	STYK1			4222	1		10713684	no	LINE
1	CD3	20	147	12p13.2	ETV6			22799	1		11716854	no	no
1	CD3	20	178	12p13.31	TAPBPL			508	1		6432020	no	no
1	CD3	20	62	12q14.1	USP15			11519	1		60951973	no	no
1	CD3	20	123	12q14.3	HMGA2			56207	3		64560665	no	no
1	CD3	20	35	14q24.2	PCNX			91078	7		69455241	no	SINE
1	CD3	20	49	14q24.3	SPTLC2			62045	8		76011106	no	SINE
1	CD3	20	138	16p11.2	SPN		831				29711385	no	SINE
1	CD3	20	103	16p11.2	HUMMLC2B			2215	5		30424397	yes	SINE
1	CD3	20	86	16p13.3	CLDN9		6321				3056170	no	SINE
1	CD3	20	122	16q22.1	TK2			32978	7		66327756	no	no
1	CD3	20	113	16q22.1	SIAT4B			7384	1		70241788	no	SINE
1	CD3	20	27	17p13.1	MYH10		4137				8738739	no	LINE
1	CD3	20	385	17p13.1	GAS7			76329	1		10226105	no	no
1	CD3	20	180	17q25.3	MSF					70513	76164310	no	no
1	CD3	20	88	18p11.31	MYOM1					6636	3050170	no	SINE
1	CD3	20	59	18q11.2	HRH4					57111	20369018	no	no
1	CD3	20	138	18q12.2	BRUNOL4		1125701				34523292	no	no
1	CD3	20	20	18q21.33	BCL2					71499	58868049	no	no
1	CD3	20	181	19p13.2	MGC33407					28288	8640465	no	no
1	CD3	20	323	19q13.31	KCNN4			2388	1		48974861	no	no
1	CD3	20	71	1p13.3	VAV3			123456	2		107738756	no	LINE
1	CD3	20	86	1p21.2	CGI-30		85904				101040158	no	no
1	CD3	20	89	1p31.3	APG4C			12107	1		62631827	no	no
1	CD3	20	220	1p32.1	FLJ10986			401952	12		59534509	no	no
1	CD3	20	100	1p36.33	GNB1		2152				1772944	no	no
1	CD3	20	40	1q21.3	TSRC1					5581	147755571	no	no
1	CD3	20	121	1q32.2	CD34			20481	4		205147833	no	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
1	CD3	20	146	1q42.3	TOMM20					81283	232232526	no	no
1	CD3	20	95	1q44	SMYD3			311339	5		243216096	no	no
1	CD3	20	85	20q11.21	HM13			53199	10		30900440	no	no
1	CD3	20	335	20q11.21	ID1					10057	30949370	no	SINE
1	CD3	20	117	20q11.21	BCL2L1			48716	1		31006940	no	no
1	CD3	20	170	21q22.11	FLJ46020			135	E1		34241761	no	no
1	CD3	20	77	21q22.3	HSF2BP			870	1		43934706	yes	no
1	CD3	20	67	22q13.1	ADSL		82614				38902988	no	no
1	CD3	20	128	22q13.31	NUP50					36318	43893589	no	LINE
1	CD3	20	101	2p13.3	AAK1		1974				69847357	yes	no
1	CD3	20	24	2p15	TMEM17					22554	62679742	no	no
1	CD3	20	60	2p15	UGP2			14040	1		64057810	no	no
1	CD3	20	105	2q22.3	ZFH1B		987891				146476576	no	LINE
1	CD3	20	41	2q32.3	STK17B			12924	2		197225894	no	no
1	CD3	20	214	2q33.1	SF3B1					98055	198361188	no	no
1	CD3	20	75	3p22.2	MOBP					118638	39647085	no	SINE
1	CD3	20	92	3p25.1	ANKRD28		4035				15817372	no	no
1	CD3	20	33	3p26.1	EDEM1					151397	5388039	no	LINE
1	CD3	20	114	3q26.32	TBL1XR1		146029				178381982	no	no
1	CD3	20	72	4p15.32	PROM1			44695	4		15783754	no	no
1	CD3	20	100	4p16.1	MIST			13786	1		10423586	no	no
1	CD3	20	95	4q31.21	INPP4B		164178				144094337	no	no
1	CD3	20	48	5q11.2	FLJ11795		211607				55640303	no	no
1	CD3	20	138	5q31.2	SIL1			103617	3		138506582	no	LINE
1	CD3	20	177	5q35.1	STK10			80674	5		171515344	no	no
1	CD3	20	142	6p12.1	ELOVL5			52353	1		53208426	no	no
1	CD3	20	27	6p21.2	PIM1					12408	37202465	no	no
1	CD3	20	142	6p21.32	HLA-DMB		2988				32958725	no	DNA
1	CD3	20	114	6p22.3	JARID2			50043	1		15404549	no	no
1	CD3	20	83	7q31.2	HIC					392365	114606124	no	no
1	CD3	20	117	7q33	UNQ1940		52428				134901974	no	no
1	CD3	20	35	8q21.3	DKFZp762O076		1193				92010824	no	LTR
1	CD3	20	196	9p21.3	MLLT3			220914	4		20391536	no	LINE
1	CD3	20	94	9q33.2	CEP110			86342	40		119312961	no	no
1	CD3	20	88	9q34.11	DNM1			43698	13		126385430	no	no
1	CD3	20	33	Xq26.2	MASK		5528				129857422	no	no
1	CD3	20	75	Xq26.2	HPRT1		2728				132297136	no	DNA
1	CD3	14	137	10p15.1	COPEB		139166				3920621	no	no
1	CD3	14	254	13q12.3	FLT1			71382	8		26795850	no	SINE
1	CD3	14	89	13q33.1	BIVM			2110	1		101151510	yes	no
1	CD3	14	39	17q21.2	STAT3			67245	15		40846433	no	SINE

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
1	CD3	14	42	19q13.12	GPR43		13316				40619141	no	no
1	CD3	14	236	1p36.13	LOC255104		479				19596364	yes	no
1	CD3	14	58	20q13.12	PRKCBP1		58400				46729296	no	no
1	CD3	14	166	20q13.2	ZNF217		167737				53052795	no	SINE
1	CD3	14	88	22q13.1	PLA2G6			4041	1		36816761	no	SINE
1	CD3	14	248	2p25.1	LOC339789		157595				8647883	no	no
						CD3, 20 mon							
1	CD3	14	56	3p26.1	ITPR1			278126	27		4788262	no	no
1	CD3	14	30	5q13.2	CENPH			600	1		68501859	yes	no
1	CD3	14	46	5q23.2	CSNK1G3					441018	123468374	no	no
1	CD3	14	276	6q22.32	NCOA7			18732	1		126111416	no	no
1	CD3	14	164	7p14.1	CDC2L5		2336				39728553	no	no
1	CD3	14	424	Xq28	IRAK1					4016	151739850	no	LINE
1	CD3	20	334	10p11.23	MAP3K8					39821	30794588	no	no
1	CD3	20	96	10q22.1	KIAA1274					8614	71681422	no	LINE
1	CD3	20	145	10q22.3	DC-TM4F2					9278	81953239	no	no
1	CD3	20	127	10q25.2	PDCD4		6796				112289387	no	no
1	CD3	20	83	11q13.2	RPS6KB2		685				66970662	yes	SINE
1	CD3	20	62	11q22.3	PDGFD			91121	1		103481485	no	no
1	CD3	20	71	12q23.3	MGC40368			15085	3		105214133	no	no
1	CD3	20	43	12q24.23	PRKAB1			594	1		118518308	yes	no
1	CD3	20	43	12q24.31	ATP6V0A2			30672	9		122580636	no	no
1	CD3	20	96	13q14.11	DNAJD1		236				41395103	yes	SINE
1	CD3	20	24	14q22.1	C14orf29		12595				49316321	no	no
1	CD3	20	365	14q22.1	PTGER2		3556				50767595	no	no
1	CD3	20	199	15q13.3	C15orf16					48671	29442714	no	no
1	CD3	20	71	15q21.3	TCF12			195074	3		55121963	no	SINE
1	CD3	20	84	16p11.2	ITGAL		13146				30506985	no	SINE
1	CD3	20	40	16p12.3	LOC57149			7148	2		20886127	no	no
1	CD3	20	60	16p13.3	KCTD5		356				2732063	yes	no
1	CD3	20	44	17q21.33	WFIKKNRP		20086				49367392	no	DNA
1	CD3	20	109	17q23.2	BCAS3			186680	6		59416334	no	DNA
1	CD3	20	24	17q24.3	MAP2K6					215577	68352178	no	no
1	CD3	20	32	18p11.21	C18orf9					66947	12595685	no	no
1	CD3	20	85	18p11.23	PTPRM			4313	1		7562130	no	no
1	CD3	20	56	18q12.2	C18orf37			1940	1		31328002	no	LINE
1	CD3	20	173	18q21.33	BCL2			13603	1		58953151	no	no
1	CD3	20	92	19p13.2	TSPAN16			3619	6		11295053	no	SINE
1	CD3	20	375	19p13.2	HSPC023					9299	13759572	no	no
1	CD3	20	117	19p13.3	CNN2					620	990683	yes	no
1	CD3	20	55	19q13.12	FLJ38451		1459				40144503	no	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
1	CD3	20	33	1p31.2	IL23R					5606	67101170	no	LINE
1	CD3	20	26	1p36.11	SOC		1306				26239117	no	no
1	CD3	20	47	1q32.1	PTPRC			22094	1		195918576	no	DNA
1	CD3	20	345	1q32.1	CACNA1S					12033	198284866	no	no
1	CD3	20	349	1q32.2	CD34			7313	1		205160901	no	no
1	CD3	20	173	20q13.2	BCAS1					8358	53237581	no	no
1	CD3	20	60	20q13.2	BCAS1					2307	53243632	no	no
1	CD3	20	60	21q21.1	USP25		289248				15735119	no	no
1	CD3	20	22	22q13.31	C22orf9			5907	1		43907309	no	no
1	CD3	20	109	2p21	ZFP36L2					12254	43412987	no	no
1	CD3	20	39	2p23.3	UBXD4			39458	4		24177369	no	LINE
1	CD3	20	151	3p26.2	CRBN		126522				3322912	no	DNA
1	CD3	20	50	4p15.2	DHX15			2052	1		24335008	no	no
1	CD3	20	82	4q21.21	RASGEF1B		35216				82886672	no	LTR
1	CD3	20	118	6p21.32	BRD2			838	1		32984197	yes	no
1	CD3	20	144	7q31.1	HIC			52790	2		114169502	no	no
1	CD3	20	78	9q21.12	ALDH1A1		39383				71064471	no	no
1	CD3	20	65	9q22.32	C9orf3			77593	5		92907996	no	no
1	CD3	20	47	9q34.3	UBADC1		895				134293406	yes	LTR
2	CD15	9	29	11p11.2	MGC4707		216				46922387	yes	no
2	CD15	9	137	12q13.13	NFE2		6633				52982444	no	no
2	CD3	13	689	16q22.3	FLJ34389			209	E1		74514143	yes	no
2	CD3	13	210	19p13.3	LRG1		844				4491880	no	no
2	CD3	13	380	19q13.31	MGC 51082		57				48815903	yes	no
2	CD3	13	151	19q13.32	BCL3		7654				49936227	no	SINE
2	CD3	13	231	19q13.32	BCL3		3255				49937616	no	SINE
2	CD3	13	64	1q32.1	SNRPE			752	2		201008929	yes	no
2	CD3	13	23	1q42.2	TARBP1		44226				231743344	no	no
2	CD3	13	220	20p12.2	JAG1					97385	10515949	no	DNA
2	CD3	13	194	20q11.21	BCL2L1		149				31055805	yes	no
2	CD3	13	33	20q11.21	SPAG4L					12725	32274533	no	no
2	CD3	13	36	4p12	TEC			21779	1		48165577	no	no
2	CD3	13	41	4p15.2	RBPSUH		31087				26041128	no	LTR
2	CD3	13	113	5p14.3	CDH18		554571				20581361	no	no
2	CD3	13	133	6q23.2	AKAP7			53375	4		131500406	no	LINE
2	CD3	13	115	6q24.3	STXBP5		304382				147201063	no	no
2	CD3	13	110	6q25.2	PIP3-E			2635	1		154706255	no	no
2	CD3	13	226	7p21.3	NDUFA4			434	1		10723716	yes	no
2	CD3	13	188	7q21.13	CLDN12			944	1		89645769	no	no
2	CD3	13	188	8p21.2	PTK2B			37753	1		27242717	no	LINE
2	CD3	13	241	Xp22.31	TBL1X		146248				8698293	no	LTR

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
2	CD3	9	33	11q24.2	SIAT4C			38213	1		125801961	no	no
2	CD3	9	25	12p13.33	NINJ2			64984	1		578032	no	no
2	CD3	9	57	12q21.33	ATP2B1		292170				88844482	no	no
2	CD3	9	113	14q21.3	ARF6	CD3, 23 mon				75646	48427418	no	no
2	CD3	9	72	19q13.31	KCNN4	CD3, 23 mon		3011	1		48974238	no	SINE
2	CD3	9	37	20p12.3	KIAA1434		42641				5629313	no	SINE
2	CD3	9	111	21q21.2	FLJ42200	CD3, 23 mon	89004				24633921	no	LTR
2	CD3	9	85	2p15	B3GNT1	CD3, 23 mon	14710				62383086	no	SINE
2	CD3	9	22	2q37.3	LRRFIP1		3273				238884406	no	no
2	CD3	9	40	6p25.1	RPP40		32806				4980227	no	no
2	CD3	9	37	9p24.2	RFX3	CD3, 23 mon; CD19, 19 mon	4036				3520019	no	no
2	CD3	13	155	10q23.33	CYP2C8			31618	9		96462151	no	no
2	CD3	13	271	11p15.1	NAV2			296649	12		19995870	no	LINE
2	CD3	13	36	15q21.2	MGC35274			6876	1		49739496	no	LINE
2	CD3	13	172	17p13.1	LOC92162		15703				7943247	no	no
2	CD3	13	115	17p13.3	SERPINF1			2375	1		1874225	no	no
2	CD3	13	32	17q11.2	POLDIP2			585	1		26829545	yes	no
2	CD3	13	33	1q32.2	TRAF3IP3			455	1		207009368	no	no
2	CD3	13	189	20p11.23	SNX5			299	1		17943855	yes	no
2	CD3	13	57	2p15	B3GNT1		14191				62383605	no	no
2	CD3	13	23	2q14.3	PROC		8997				128260288	yes	no
2	CD3	13	97	2q37.3	LRRFIP1	CD3, 23 mon; CD15, 19 mon	3205				238884474	no	no
2	CD3	13	185	4p12	TEC			21977	1		48165379	no	LINE
2	CD3	13	20	4q21.3	AFF1		45568				88340964	no	no
2	CD3	13	165	4q31.3	LRBA		14614				152528890	no	no
2	CD3	13	88	8q12.1	NSMAF		192900				59815247	no	no
2	CD3	13	62	Xp22.22	EGFL6		99939				12849228	no	no
3	CD3	18	36	10p12.1	BAMBI					253612	29229486	no	no
3	CD3	18	45	10p15.1	COPEB	CD3, 17 mon and 30 mon	25924				3807379	no	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
3	CD3	18	75	10q26.3	MGMT					101922	131141690	yes	no
3	CD3	18	135	11p15.1	PIK3C2A			13591	1		17142006	no	no
3	CD3	18	133	13q14.2	VDRIP		56317				46523558	no	LINE
3	CD3	18	115	15q26.2	MCTP2		93786				92477477	no	LINE
3	CD3	18	62	17q23.3	PECAM1		22518				62959880	no	no
3	CD3	18	36	1p34.2	HIVEP3			34652	1		41762803	no	LINE
3	CD3	18	134	1q42.3	TOMM20					130904	232182905	no	no
3	CD3	18	57	22q13.31	CELSR1		56433				45201741	no	no
						CD3, 17 mon							
3	CD3	18	142	2p14	PPP3R1		1127				68455312	yes	no
3	CD3	18	24	4q31.1	MAML3			60909	1		141591951	no	no
3	CD3	18	49	5p13.2	BRIX			9652	1		34961634	yes	no
3	CD3	18	22	6q21	SCML4		51285				108190310	no	no
3	CD3	18	155	7q11.21	LOC285908		97046				65334091	no	LINE
						CD3, 30 mon							
3	CD3	17	34	11p15.4	MRPL17		537				6669420	yes	DNA
						CD3, 18 mon and 30 mon							
3	CD3	17	24	14q32.33	CRIP1					931	103927729	yes	no
3	CD3	17	82	15q25.1	IL16			115285	15		79306304	no	no
3	CD3	17	223	17p11.2	NT5M		773				17406442	yes	LINE
3	CD3	17	36	19p13.12	HSH2D			6675	1		16112513	no	SINE
3	CD3	17	165	20q13.12	PRKCBP1			479	1		46670417	no	no
3	CD3	17	84	21q22.12	RUNX1			23000	1		35318991	yes	no
3	CD3	17	129	2p13.3	MAD					7954	70148093	no	SINE
3	CD3	17	34	2p13.3	FLJ20558					7936	70343590	yes	no
3	CD3	17	160	2q35	FLJ10116			9496	1		217071361	no	LINE
3	CD3	17	53	2q37.3	COL6A3					32838		no	no
3	CD3	17	106	3p13	FLJ10539			79629	10		72738870	no	no
3	CD3	17	76	3q27.1	KLHL6			1602	1		184592719	no	no
3	CD3	17	196	5p13.1	FYB			42264	2		39222904	no	no
3	CD3	17	73	6p24.1	EDN1					139283	12544044	no	LINE
3	CD3	17	69	6q24.1	C6orf80		2851				139072376	no	no
3	CD3	17	21	7q21.11	GNAI1					117074	79577850	no	no
3	CD3	17	101	9p24.3	SMARCA2			147013	25		2152355	no	no
3	CD3	30	88	11p15.4	OR52N2		10486				5795389	no	LINE
3	CD3	30	249	11q21	FN5			11339	1		92953328	no	LTR
3	CD3	30	41	13q31.3	C13orf25		97633				89600442	no	LTR
3	CD3	30	51	15q12	GABRA5		181716				24540998	no	LINE
3	CD3	30	79	15q21.2	MAPK6			2679	1		50030182	no	no
3	CD3	30	166	15q26.3	ARRDC4					22315	96269143	no	LINE

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
3	CD3	30	195	15q26.3	IGF1R		49029				96879000	no	no
3	CD3	30	31	17q21.33	MGC15396		613				49260514	yes	no
3	CD3	30	62	17q23.2	MSI2			117632	2		55926201	no	no
3	CD3	30	64	18q11.2	C18orf17			21056	1		19845780	yes	DNA
3	CD3	30	246	19p13.3	GNA15		3828				3083402	no	SINE
3	CD3	30	27	1p36.22	SSB1					58108	9197242	no	SINE
3	CD3	30	27	20p12.2	PLCB4			67130	1		9139062	no	no
3	CD3	30	63	2p13.3	PCBP1					26871	70317737	no	no
3	CD3	30	32	2p21	PRKCE			205320	2		46058897	no	DNA
3	CD3	30	59	7q35	CNTNAP2			542491	1		145748061	no	no
						CD3,							
3	CD3	30	112	8q11.23	LOC115294	18 mon	35924				52859681	no	no
3	CD3	30	46	9q33.2	OR1J1					12302	120602980	no	no
3	CD3	18	149	12q24.31	P2RX4		12472				120047145	no	no
3	CD3	18	43	14q24.1	FLJ39779			913	1		67252318	yes	no
3	CD3	18	226	22q12.3	NCF4			1914	1		35501998	no	no
3	CD3	18	25	3p21.2	GLYCTK		7167				52273174	no	no
3	CD3	18	38	4q31.21	GYPE		36472				145440815	no	no
3	CD3	17	300	11q21	FN5			11636	1		92953031	no	LINE
3	CD3	17	142	4p16.3	RGS12			70026	3		3417277	no	SINE
3	CD3	17	91	4q31.1	MAML3			60822	1		141592038	no	no
3	CD3	17	76	8q11.23	LOC115294		35852				52859609	no	no
3	CD3	17	59	9q21.2	GNAQ		1563				76104874	yes	simple repeat
4	CD3	12	207	11q13.2	ADRBK1			8291	1		66817744	no	no
4	CD3	12	214	11q13.4	FLJ20625			3316	1		71537314	no	LINE
4	CD3	12	221	11q13.5	GARP			5212	2		76101991	no	no
4	CD3	12	51	12p12.1	BCAT1					93387	24762175	no	no
4	CD3	12	26	17q23.2	TMEM49			123347	10		58382797	no	SINE
4	CD3	12	76	18q21.33	PLEKHE1			237631	12		58769334	no	no
4	CD3	12	85	1p34.3	MACF1			114665	1		39075077	no	SINE
4	CD3	12	71	1q21.2	TSRC1					6239	147756234	no	SINE
4	CD3	12	21	20q13.2	ATP9A			64925	4		51005405	no	no
4	CD3	12	161	2q37.1	MGC35154					16370	232677901	no	no
4	CD3	12	42	3q23	RASA2			43858	4		142570693	no	no
4	CD3	12	35	4q35.1	ACSL1		29450				186472589	no	no
4	CD3	12	47	6p24.2	NEDD9		150399				11491283	no	SINE
4	CD3	12	163	6p25.2	SERPINB9					10488	2822017	no	no
4	CD3	12	75	8q24.12	ENPP2					4923	120520991	no	LINE
4	CD3	12	41	8q24.12	HAS2					229101	122353027	no	no
4	CD3	12	210	9p21.2	CCDC2			1297	1		26947707	yes	no
4	CD3	12	172	Xp11.4	CASK			301658	11		40511548	no	LTR

Supplementary Table 1

Pat.	Source	Time	Length	Chromosome	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
5	CD3	9	114	10p15.1	KLF6		195383				3976838	no	no
5	CD3	9	173	10p15.1	IL2RA			11647	1		6096631	no	no
5	CD3	9	102	11p15.4	PPFIBP2		1486				7497876	yes	DNA
5	CD3	9	159	12p11.22	MLSTD1			16713	1		29284578	no	no
5	CD3	9	97	12q23.3	DIP13B		3889				104136364	no	LINE
5	CD3	9	81	12q24.12	LNK			28678	2		110284383	no	LINE
5	CD3	9	234	13q12.11	LATS2			15610	E2		19418044	no	no
5	CD3	9	197	13q12.3	LGR8		149954				29961725	no	no
5	CD3	9	218	13q14.11	COG6			15421	4		38043235	no	no
5	CD3	9	123	13q32.1	TGDS		268				92946752	yes	no
5	CD3	9	24	14q23.2	HIF1A		23941				60128339	no	LTR
5	CD3	9	109	14q32.11	C14orf143					20479	88231429	no	no
5	CD3	9	527	16p12.2	PRKCB1			115757	2		23929632	no	LINE
5	CD3	9	98	16q22.1	NQO1		817				69537397	no	SINE
5	CD3	9	28	19p13.3	MRPL54		401				3713271	yes	simple repeat
5	CD3	9	22	1q32.1	BTG2					12684	200579668	no	SINE
5	CD3	9	135	20q13.12	PRKCBP1		5009				46675905	no	LINE
5	CD3	9	88	20q13.13	CEBPB					67916	49562550	no	LINE
5	CD3	9	61	2p11.2	POLR1A					27937	86200983	no	no
5	CD3	9	44	2p22.2	CYP1B1		20738				38298565	no	no
5	CD3	9	85	2q11.2	INPP4A			17431	1		98700244	no	SINE
5	CD3	9	35	2q37.1	PTMA		45893				232729882	yes	no
5	CD3	9	141	3p26.2	CRBN		12795				3209185	no	no
5	CD3	9	68	3q13.2	CD200					32144	113434701	no	SINE
5	CD3	9	33	3q26.32	TBL1XR1		214036				178449989	no	no
5	CD3	9	51	3q29	HES1					67752	195245061	no	LINE
5	CD3	9	75	4q21.23	MGC11324			13777	2		84929825	no	SINE
5	CD3	9	83	6p13.12	ERCC4		494685				13485777	no	SINE
5	CD3	9	231	6p24.2	NEDD9		157120				11498004	no	no
5	CD3	9	224	6p25.2	LOC340156			25385	2		2670768	no	low complexity
5	CD3	9	46	6q21	C6orf51		17858				111307475	no	no
5	CD3	9	47	6q24.2	UTRN			52651	1		144646186	no	no
5	CD3	9	93	7q31.31	ANKRD7					781347	118218587	no	no
5	CD3	9	25	8q23.3	CSMD3		299836				114705661	no	LINE
5	CD3	9	22	9q21.12	C9orf77			14100	1		69968866	no	no
5	CD3	9	27	9q22.33	COL15A1					6488	97219384	no	no
6	CD34	pre	31	10p12.31	PLXDC2		34582				20074796	no	LINE
6	CD34	pre	406	10p12.33	ST8SIA6			25044	2		17475216	no	no
6	CD34	pre	219	10p13	FAM107B			97531	2		14662096	no	no
6	CD34	pre	94	10p15.1	KLF6		22692				3804147	no	SINE
6	CD34	pre	133	10q22.1	PRG1		27823				70164608	no	SINE

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
6	CD34	pre	169	10q24.1	ENTPD1			77562	2		97258065	no	LTR
6	CD34	pre	59	10q24.1	PIK3AP1			32295	2		98112571	no	no
6	CD34	pre	122	11p13	LDLRAD3			146231	3		36076152	no	no
6	CD34	pre	235	11p13	FLJ14213			34322	4		36396186	no	DNA
6	CD34	pre	32	11p14.3	LUZP2		182002				25068491	no	SINE
6	CD34	pre	92	11q12.2	MS4A7		1034				59920284	no	SINE
6	CD34	pre	69	11q13.4	FLJ20625			3546	1		71537084	no	no
6	CD34	pre	41	11q13.4	FLJ20625			1756	1		71538874	no	SINE
6	CD34	pre	133	11q13.4	TNFRSF19L			8474	1		72822187	no	no
6	CD34	pre	191	11q13.4	PGM2L1		729				73836539	yes	no
6	CD34	pre	90	11q22.3	ACAT1			826	1		107530733	yes	no
6	CD34	pre	33	12p12.1	SSPN					41369	26318503	no	no
6	CD34	pre	42	12p13.31	NANOG			4996	2		7838258	no	no
6	CD34	pre	90	12q13.2	MLC1SA		436				54832166	yes	no
6	CD34	pre	34	12q14.3	C12orf31			149	1		64810651	yes	no
6	CD34	pre	41	12q15	MDM1		32221				67044600	no	no
6	CD34	pre	244	12q24.31	FBXL10			24866	3		120338293	no	SINE
6	CD34	pre	50	12q24.31	MORN3			6105	1		120499761	no	LINE
6	CD34	pre	62	13q14.3	FLJ11712		14348				49267580	no	no
6	CD34	pre	21	13q21.32	PCDH9		344719				65947183	no	LTR
6	CD34	pre	42	13q32.1	TGDS			711	1		92945801	yes	LINE
6	CD34	pre	57	13q32.1	ABCC4			28995	1		93622689	no	no
6	CD34	pre	72	13q32.3	PHGDHL1			69245	2		97620409	no	no
6	CD34	pre	114	14q23.3	FUT8		132809				63735072	no	DNA
6	CD34	pre	58	14q24.2	SFRS5			322	1		68224192	yes	no
6	CD34	pre	39	14q31.2	SEL1L		1928909				81919083	no	no
6	CD34	pre	417	14q32.11	C14orf143					74234	88177674	no	no
6	CD34	pre	110	14q32.32	MARK3			624	E1		101842436	yes	no
6	CD34	pre	112	15q14	SPRED1		139901				36120671	no	simple repeat
6	CD34	pre	59	15q22.31	DAPK2			60089	2		61994249	no	SINE
6	CD34	pre	85	15q23	IQCH					19873	65528985	yes	no
6	CD34	pre	31	15q24.2	SCAMP5			125990	1		73016308	no	LINE
6	CD34	pre	396	15q26.1	VPS33B		535				89296116	yes	LINE
6	CD34	pre	21	16p13.13	KIAA0350			191362	21		11196236	no	no
6	CD34	pre	71	17p11.2	PIGL		388				16320714	yes	no
6	CD34	pre	118	17p13.2	UNQ5783					42015	5296687	yes	no
6	CD34	pre	188	17p13.2	MIS12			368	1		5591180	yes	SINE
6	CD34	pre	91	17p13.3	SMG6			192719	13		2220891	no	no
6	CD34	pre	67	17q11.2	KSR1			101461	3		26046061	no	LTR
6	CD34	pre	62	17q11.2	NF1/EVI2B			214943/403	35/1		29782623	no	no
6	CD34	pre	86	17q12	MGC19764		316				33715326	yes	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromosome	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
6	CD34	pre	62	17q12	CCL23		17386				34507941	no	SINE
6	CD34	pre	65	17q21.31	LOC641522		10547				45142506	no	no
6	CD34	pre	36	17q21.33	TOB1		3375				49419426	yes	no
6	CD34	pre	44	17q21.33	FLJ42842					92721	49987291	no	LTR
6	CD34	pre	195	17q23.1	PCTP					36533	54365905	no	SINE
6	CD34	pre	40	17q25.1	GRB2			60283	3		73938574	no	no
6	CD34	pre	36	18p11.31	DLGAP1			184903	4		3650393	no	LINE
6	CD34	pre	96	18q11.2	HRH4					60662	20372569	no	SINE
6	CD34	pre	80	18q12.3	PIK3C3		2156236				35630950	no	LTR
6	CD34	pre	181	18q12.3	SETBP1					41367	40938127	no	no
6	CD34	pre	57	19p13.11	ZNF429			906	1		21481209	no	LTR
6	CD34	pre	347	19p13.12	AKAP8					22168	15303168	yes	no
6	CD34	pre	146	19p13.12	H2H2D			16648	4		16122486	no	SINE
6	CD34	pre	84	19p13.2	TRMT1					773	13075943	yes	SINE
6	CD34	pre	127	19p13.3	SCAMP4					7189	1884200	no	SINE
6	CD34	pre	49	19q13.32	IGFL2					31484	51387869	no	no
6	CD34	pre	64	19q13.33	PRR62					1133	54787136	yes	no
6	CD34	pre	147	1p13.2	PHTF1			1106	2		113598977	no	no
6	CD34	pre	405	1p13.3	CD53			6264	1		110720764	no	SINE
6	CD34	pre	121	1p22.3	COL24A1			414057	58		85630657	no	SINE
6	CD34	pre	83	1p34.3	CSF3R		354				36376525	no	no
6	CD34	pre	37	1p35.3	PTAFR					22232	28058320	no	no
6	CD34	pre	74	1p35.3	MED18			233	1		28260470	yes	no
6	CD34	pre	27	1q23.1	SPTA1		6338				155879618	no	no
6	CD34	pre	21	1q42.3	IRF2BP2					7675	231819372	yes	no
6	CD34	pre	69	20p11.23	C20orf74		24461				20712584	no	no
6	CD34	pre	70	21q11.2	SAMSN1		91320				14931855	no	no
6	CD34	pre	47	21q21.1	C21orf37		468010				17265069	no	DNA
6	CD34	pre	77	22q11.21	ATP6V1E1			226	1		16485916	yes	no
6	CD34	pre	114	22q11.23	PIWIL3			45521	14		23449716	no	LTR
6	CD34	pre	87	22q11.23	LRP5L		53116				24136194	no	no
6	CD34	pre	50	22q13.1	MKL1			22985	5		39072316	no	no
6	CD34	pre	116	22q13.2	CYB5R3		695				41277041	yes	SINE
6	CD34	pre	84	22q31.2	RGS18			2870	3		189418721	no	no
6	CD34	pre	46	2p16.1	BCL11A					85361	60567475	yes	no
6	CD34	pre	90	2p21	HAAO		136974				43131224	no	no
6	CD34	pre	38	2p22.2	FAM82A		128608				38023533	no	SINE
6	CD34	pre	35	2p25.3	ALLC					398364	3713132	no	LINE
6	CD34	pre	69	2q11.2	MGAT4A			23008	2		98945981	no	no
6	CD34	pre	91	2q11.2	AFF3			333044	7		100010401	no	no
6	CD34	pre	75	2q22.1	LRP1B			994837	7		142105202	no	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
6	CD34	pre	38	2q32.1	CALCRL			24985	1		188490536	no	no
6	CD34	pre	249	3p14.3	ARHGEF3			26282	1		56767019	no	no
6	CD34	pre	31	3p24.2	THRB			133295	1		24377983	no	no
6	CD34	pre	120	3p26.1	EDEM1		645				5203788	yes	no
6	CD34	pre	88	3q25.1	SUCNR1			2923	1		152915269	no	LINE
					MCF2L2/			172423/	15/				
6	CD34	pre	150	3q27.1	B3GNT5			2387	1		184294345	yes	no
6	CD34	pre	266	3q29	LOC348840		16836				199151047	no	no
6	CD34	pre	420	4p12	CEP135			14567	6		56745146	no	no
6	CD34	pre	98	4p14	WDR19		574				39080630	yes	no
6	CD34	pre	43	4p16.1	KIAA0232					7695	6958127	no	no
6	CD34	pre	46	4q31.1	MAML3			418736	2		141234124	yes	no
6	CD34	pre	50	4q31.3	RNF175			10527	2		155248	no	DNA
6	CD34	pre	117	5p15.5	BASP1					850598	18180277	no	no
6	CD34	pre	64	5q13.3	IQGAP2			55069	1		75838291	no	no
6	CD34	pre	30	5q15	GLRX		13871				95246325	no	no
6	CD34	pre	34	5q33.1	IL17B		50884				148838232	no	no
6	CD34	pre	333	5q33.1	ANAXA6					5926	150502852	no	no
6	CD34	pre	106	6p12.2	PAQR8		19800				52253981	no	no
6	CD34	pre	61	6p21.33	TRIM26			12023	2		30275307	no	no
6	CD34	pre	74	6p21.33	HLA-E					22520	30589774	no	SINE
6	CD34	pre	135	6p21.33	TUBB		800				30793511	yes	no
6	CD34	pre	43	6p21.33	IER3		41683				30860169	no	LINE
6	CD34	pre	108	6p22.1	ZNF323			16051	3		28413900	no	SINE
6	CD34	pre	52	6p22.2	HIST1H2BE					282	26292719	yes	no
6	CD34	pre	79	6p24.1	HIVEP1		2465				12118092	no	no
6	CD34	pre	124	6p24.3	CAGE1			62256	13		7272382	no	SINE
6	CD34	pre	76	6q21	FLJ37396					33325	109671112	yes	LTR
6	CD34	pre	52	7p12.3	TNS3			45660	3		47274319	no	no
6	CD34	pre	22	7p13	CCM2		5223				44775344	no	LTR
6	CD34	pre	141	7p15.2	NFE2L		53190				25880942	no	LINE
6	CD34	pre	191	7p22.1	FOXK1			63447	2		4529715	yes	SINE
6	CD34	pre	32	7p22.1	FBXL18			61673	4		5236047	no	no
6	CD34	pre	308	7p22.1	TRIAD3		3595				5569225	no	no
6	CD34	pre	270	7p22.3	MAD1L1			123134	11		1893787	no	no
6	CD34	pre	172	7q36.1	XRCC2			292	1		151764699	yes	no
6	CD34	pre	44	8q21.11	MSC					104987	72698769	no	no
6	CD34	pre	79	8q24.13	TRIB1					27503	126434734	no	no
6	CD34	pre	102	9q21.11	PRKACG		5217				67091344	no	DNA
6	CD34	pre	45	9q22.23	TGFBR1			4796	1		97252046	no	no
6	CD34	pre	261	9q34.11	AK1					1183	126003531	no	LINE

Supplementary Table 1

Pat.	Source	Time	Length	Chromosome	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
6	CD34	pre	303	9q34.11	FNBP1		2155				128183673	no	no
6	CD34	pre	137	9q34.2	PXRA		40551				132692251	no	no
6	CD34	pre	96	Xp22.22	ARHGAP6			291954	1		10753293	no	no
6	CD34	pre	116	Xq11.2	LOC139886		295				61438297	yes	no
6	CD34	pre	80	Xq23	ACSL4		107087				107847677	no	no
6	CD34	pre	118	Xq23	RP13-360B22.2		42648				107967666	no	no
6	CD34	pre	57	Xq26.1	XPNPEP2					6535	127615683	no	no
nd	CD34	pre	68	10p15.1	ASB13		17849				5730396	yes	no
nd	CD34	pre	22	10q11.21	HNRPF			643	1		43187977	yes	no
nd	CD34	pre	40	10q11.21	HNRPF		10998				43199618	no	SINE
nd	CD34	pre	47	10q23.31	IFIT3			473	1		90752803	no	no
nd	CD34	pre	30	10q24.2	PI4KII			14051	2		99079081	no	LINE
nd	CD34	pre	75	10q25.2	GPAM		16622				113624727	no	SINE
nd	CD34	pre	100	10q25.2	TCF7L2			6436	2		114381234	no	no
nd	CD34	pre	73	10q26.2	ADAM12			344054	14		127307370	no	no
nd	CD34	pre	55	11p11.2	NR1H3					271	47254976	yes	no
nd	CD34	pre	198	11q12.2	CD6			20603	1		60535137	no	SINE
nd	CD34	pre	731	11q13.1	FKSG44					6881	64963234	no	no
nd	CD34	pre	85	11q13.4	ARHGEF17			5340	1		72751317	no	no
nd	CD34	pre	32	11q14.1	GAB2			68543	1		77786531	no	LINE
nd	CD34	pre	161	11q24.1	KIAA1959			15484	1		122079580	no	no
nd	CD34	pre	20	12p13.2	APOLD1		29139				12800743	no	LINE
nd	CD34	pre	98	12p13.31	HOM-TES-103			847	1		6534643	yes	no
nd	CD34	pre	152	12p13.32	HRMT1L3			32171	1		3503189	no	LTR
nd	CD34	pre	103	12q13.2	RPS26		661				54721532	yes	SINE
nd	CD34	pre	69	12q13.3	KIF5A			2457	1		56232571	no	no
nd	CD34	pre	104	12q24.31	DYNLL1			25190	1		119344803	no	SINE
nd	CD34	pre	22	13q12.3	HMGB1					64687	28767310	no	no
nd	CD34	pre	401	13q14.13	LCP1			11583	1		44542812	no	no
nd	CD34	pre	120	13q22.3	MYCBP2		1897				75701075	yes	LINE
nd	CD34	pre	48	13q32.3	TM9SF2		2215				97849514	yes	no
nd	CD34	pre	191	14q22.1	DDHD1		84496				51694534	no	no
nd	CD34	pre	181	14q32.12	RIN3			55959	2		91026125	no	no
nd	CD34	pre	48	15q14	FLJ20582		50249				33046759	no	no
nd	CD34	pre	23	15q14	FSIP1			87027	10		37704060	no	no
nd	CD34	pre	35	15q15.2	TGM5			16446	5		41258665	no	SINE
nd	CD34	pre	45	15q22.31	SMAD6		58164				64652324	no	no
nd	CD34	pre	177	15q24.2	MDS018					15509	73074392	no	LTR
nd	CD34	pre	160	16q13	GPR97			19469	11		57497037	no	no
nd	CD34	pre	93	16q22.1	VAC14			105734	17		70505509	no	no
nd	CD34	pre	51	16q24.3	GAS8		677				89830510	yes	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromosome	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
nd	CD34	pre	94	17p12	ADORA2B					6852	16086628	no	SINE
nd	CD34	pre	21	17p13.1	FLJ32734					22222	8811596	no	SINE
nd	CD34	pre	74	17q12	ACACA			117957	23		35793652	no	SINE
nd	CD34	pre	52	17q21.2	HSD17B1			3051	E6		41080200	yes	no
nd	CD34	pre	102	17q21.2	COASY		641				41086724	yes	SINE
nd	CD34	pre	128	17q21.32	MYL4		4830				45756236	no	no
nd	CD34	pre	39	17q21.32	SCAP1			1743	1		46980447	yes	no
nd	CD34	pre	43	17q25.3	FLJ23058		59482				80071994	yes	no
nd	CD34	pre	152	18p11.22	NDUFV2		189755				8902970	no	no
nd	CD34	pre	65	18q12.1	MAPRE2			2268	1		30875646	no	no
nd	CD34	pre	178	18q22.1	C18orf4					54106	63268687	no	no
nd	CD34	pre	115	18q22.3	FBXO15			64312	8		69899606	no	no
nd	CD34	pre	186	19p13.11	PGPEP1					8877	18347033	no	SINE
nd	CD34	pre	95	19p13.2	LASS4		15645				8164612	no	SINE
nd	CD34	pre	156	19p13.3	RFX2			43959	1		6017595	no	no
nd	CD34	pre	362	19q13.41	SIGLEC6		16390				56743266	no	LTR
nd	CD34	pre	51	1p13.1	VANGL1			2859	1		115543263	no	no
nd	CD34	pre	75	1p13.3	GPSM2			29304	10		108795743	no	LINE
nd	CD34	pre	33	1p22.2	HFU1			10912	3		91327555	no	SINE
nd	CD34	pre	52	1p34.2	CAP1		2035				39917687	yes	no
nd	CD34	pre	326	1p36.12	KIAA0478			2712	1		22250537	no	no
nd	CD34	pre	68	1q22	MTX1			1053	1		152396335	yes	no
nd	CD34	pre	26	1q32.2	CD34					10854	205132729	no	no
nd	CD34	pre	25	1q42.3	TOMM20					112304	232201505	no	no
nd	CD34	pre	88	20p11.23	SLC24A3			342996	3		19531286	no	no
nd	CD34	pre	24	20q11.21	BCL2L1			16376	1		31039280	no	no
nd	CD34	pre	220	20q13.2	ZNF217		124662				53009720	no	no
nd	CD34	pre	143	21q22.11	IFNAR2			6997	1		33529624	no	no
nd	CD34	pre	220	21q22.11	GART			298	1		33834514	yes	no
nd	CD34	pre	224	21q22.11	C21orf51		1200				34666995	yes	no
nd	CD34	pre	78	21q22.2	HMG1		13843				39655286	no	no
nd	CD34	pre	32	21q22.3	ITGB2			703	1		45196303	no	no
nd	CD34	pre	184	21q22.3	C21orf56		1189				46461692	no	no
nd	CD34	pre	48	22q12.1	CRYBA4					1755	25352941	no	LTR
nd	CD34	pre	99	22q13.1	FLJ20232		559				38140764	yes	no
nd	CD34	pre	66	22q13.31	CELSR1		66851				45212159	no	LINE
nd	CD34	pre	71	2p11.2	CAPG			4496	1		85607518	no	no
nd	CD34	pre	347	2p11.2	CAPG			3717	1		85608297	no	SINE
nd	CD34	pre	86	2p14	AFTIPHILIN					14358	64809028	no	no
nd	CD34	pre	29	2p14	PLEK		13410				68553548	no	LINE
nd	CD34	pre	42	2q11.2	IL1RL1			24994	1		102574356	no	no

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
nd	CD34	pre	136	2q12.3	LIMS1			846	1		108827013	no	no
nd	CD34	pre	28	2q14.3	TSN					398529	123019103	no	LINE
nd	CD34	pre	32	2q31.1	GPR155			4628	1		175549724	no	no
nd	CD34	pre	128	3p13	FOXP1			191411	2		71362106	no	no
nd	CD34	pre	109	3q25.1	MBNL1		22299				153284439	no	no
nd	CD34	pre	133	4p15.32	CD38		12188				15518626	no	no
nd	CD34	pre	64	4p16.1	MIST		20483				10457855	no	LTR
nd	CD34	pre	33	4p16.2	C17		14584				5100323	no	no
nd	CD34	pre	136	4q12	POLR2B		243				57760408	yes	no
nd	CD34	pre	126	4q13.3	SLC4A4			68023	4		72739091	no	no
nd	CD34	pre	70	5p13.2	FLJ23577			879	1		35664369	yes	no
nd	CD34	pre	20	5q31.2	LOC340061					679	138883372	no	no
nd	CD34	pre	64	5q33.2	LARP		29004				154091968	no	no
nd	CD34	pre	38	5q33.3	FLJ31951		3864				158619593	no	no
nd	CD34	pre	26	5q35.3	RNF130			23017	1		179585754	no	no
nd	CD34	pre	42	6p12.2	C6orf33			28678	1		52302459	no	no
nd	CD34	pre	159	6p25.1	PECI		39572				4120402	no	LTR
nd	CD34	pre	297	6p25.1	LY86					5279	6605492	no	no
nd	CD34	pre	63	6q27	CCR6					7122	167469039	no	no
nd	CD34	pre	76	7p14.1	AMPH					32991	38131096	no	no
nd	CD34	pre	67	7p21.3	ARL4A					29169	12504065	no	no
nd	CD34	pre	145	7q21.2	CDK6			180522	4		91894719	no	no
nd	CD34	pre	54	7q31.1	DLD		13374				107078768	no	SINE
nd	CD34	pre	42	7q32.1	SND1			310014	15		127156729	no	no
nd	CD34	pre	151	7q33	SLC35B4		1547				133419476	no	DNA
nd	CD34	pre	51	7q33	TIF1			36976	1		137596680	no	SINE
nd	CD34	pre	121	8q11.23	PCMTD1		101830				52925587	no	LINE
nd	CD34	pre	44	8q21.13	PKIA		390822				79087476	no	no
nd	CD34	pre	50	8q22.1	RAD54B			37404	3		95406489	no	no
nd	CD34	pre	89	8q24.12	DEPDC6					14839	121034183	no	no
nd	CD34	pre	64	9q33.1	TLR4			519	1		115843148	no	SINE
nd	CD34	pre	436	9q34.11	GPR107			58082	14		128250331	no	no
nd	CD34	pre	47	9q34.3	EGFL7		13346				134901324	no	no
nd	CD34	pre	59	Xp22.33	FLJ11323			6495	2		91666	no	LTR
nd	CD34	pre	79	Xq12	OPHN1			23	E1		66520415	yes	no
nd	CD34	pre	90	Xq21.1	ITM2A					88305	77283452	no	no
nd	CD34	pre	153	Xq25	SH2D1A			6143	1		122184092	no	LINE
nd	CD34	pre	56	10q22.1	KIAA1279					32727	70154064	no	no
nd	CD34	pre	103	10q25.2	VTI1A			2550	1		113873893	no	SINE
nd	CD34	pre	40	11p15.4	CARS			35214	9		3007766	no	LTR
nd	CD34	pre	132	11p15.4	C11orf56		1588				6221772	no	LINE

Supplementary Table 1

Pat.	Source	Time	Length	Chromo- some	RefSeq Gene	Identity	5' of TSS [bp]	3' of TSS [bp]	Intron	3' of RefSeq Gene [bp]	UCSC Locus	CpG Island (1kb)	Repetitive Elements
nd	CD34	pre	94	11q14.2	PRSS23		9711				86228160	no	no
nd	CD34	pre	160	11q22.3	CASP4					348277	104003035	no	no
nd	CD34	pre	77	11q23.3	ARHGEF12			468	1		1197460066	no	no
nd	CD34	pre	96	12q13.12	SPATS2			55548	2		48103108	no	no
nd	CD34	pre	58	14q32.2	AK7			14657	3		94863146	no	SINE
nd	CD34	pre	254	15q23	UACA					26497	68636215	no	no
nd	CD34	pre	137	17p13.3	GARNL4			109045	2		3015368	no	no
nd	CD34	pre	83	17q11.2	CRLF3			980	1		29296274	no	no
nd	CD34	pre	50	17q21.32	MYL4		1573				45759493	no	no
nd	CD34	pre	50	17q23.3	MRC2			25197	1		61203330	no	
nd	CD34	pre	59	19p13.13	RFX1		1251				13979348	no	no
nd	CD34	pre	49	1p21.2	DPH5		654				100954908	no	no
nd	CD34	pre	142	1p34.1	C1orf164			13126	2		44297380	no	no
nd	CD34	pre	166	1q42.12	PARP1		246				223570668	no	no
nd	CD34	pre	107	1q44	SMYD3			242488	4		243284947	no	no
nd	CD34	pre	39	20p12.1	RRBP1			16650	1		17641278	no	no
nd	CD34	pre	113	2p13.1	DGUOK					71565	74232191	no	no
nd	CD34	pre	91	2p23.2	ALK			471326	4		29647641	no	SINE
nd	CD34	pre	80	2q34	ZNFN1A2					48447	214025439	no	no
nd	CD34	pre	52	3p25.2	IQSEC1					19093	12895281	no	LTR
nd	CD34	pre	53	6p25.2	WRNIP1			3886	2		2714551	no	no
nd	CD34	pre	87	7p22.3	MAFK		20533				1294173	no	SINE
nd	CD34	pre	166	Xp22.2	TBL1X			87546	2		8932090	no	no
nd	CD34	pre	24	Xq21.1	ATP7A			43324	1		75965397	no	no

Supplementary Table 2

Sample	Detected	Expected					p-Value*
	CIS	MV	Median	Min	Max	SD	
Post	15	0.90	1	0	7	0.95	0
Pre P6	1	0.17	0	0	4	0.42	0.16
Pre nd	3	0.19	0	0	3	0.44	0.001