Supplemental Material

Figure S1. Optimization of Alignment Parameters.
Using 5 randomly chosen human-baboon orthologous pairs, we tested a range gap-opening and gap-extension parameters. For each parameter, we considered the effect on sequence similarity, sequence similarity counting each indel as one mismatch, the number of mismatches, and the number of indels. We also considered the correspondence of gaps and the position of young retrotransposition events. Parameters were favored (gap-opening penalty of –50 and gap-extension penalty of –1) which simultaneously reduced the number of indels and mismatches, and which treated new retrotransposition events as a single event. Suboptimal alignment regions were analyzed separately.

Figure S2. Sequence and Alignment Properties.

A. Human-baboon sequence alignments
B. Human-chimpanzee sequence alignments

We obtained 42 qualified human-baboon and 51 human-chimpanzee orthologous pairs. For each alignment (labeled by baboon or chimpanzee reference accession), a number of mutational properties such as Kimura distance, large indel frequency (count per 10 kb) and retrotransposition frequency (count per 10 kb). Other sequence properties such as human %GC, human %SINEs, human % LINE baboon/chimpanzee GC%, baboon/chimpanzee SINEs%, baboon/chimpanzee LINEs% were also compared. Results varied within the expected levels of human genome variation. All Y-axis are in a logarithm scale with the base = 2.

Table S3. Non human primate genomic clones' map location with respect to the human genome.
We obtained 51 chimpanzee, 42 baboon and 9 lemur genomic clones. Within 102 non-human primate genomic clones, 99 of the sequences mapped to phylogenetic group chromosome 7.

Figure S4. Single Nucleotide Variation (500 bp non-overlapping windows).
A scatter plot of genetic distances determined from non-overlapping 500 bp sliding windows for human-chimpanzee, human-baboon and human-lemur sequence alignments. These were plotted against human divergence times of 5.5, 25 and 55 million years ago for chimpanzee, baboon and lemur alignments respectively. Suboptimal alignments were excluded. The means and their standard deviations were shown.

Table S5. The Number and Length Distribution of Large Insertion/Deletion Events.
The number of all insertion/deletion events and their length distribution (the number of bases assigned as an insertion to the species based on the alignments) is summarized. All large (>100 bp in length) insertion/deletion events were further classified into one of three categories1) indels that were characteristic of a retrotransposition event (see text), 2) RAID (Repeat associated insertion/deletions) in which a repeat was identified at one or both gap junctions and 3) NRAID (Non-repeat associated insertion/deletions). Significant differences in the number of counts between species were tested by χ^2 test assuming similar distributions. Only the numbers of lineage specific retrotransposition events were significantly different for each species comparison. No significant difference was observed in the number of RAID or NRAID.

Figure S6. Length Distribution of Indels Shorter than 100bp in Human-Chimpanzee, Human-Baboon and Human-Lemur Sequence Alignment.
More than 80% of all indels are equal to or less than 15 bp in length but contribute to less than 3.6% of the overall length variation within an alignment. Length (bp) is in a logarithm scale with the base = 2.

Figure S7. Length Distribution of NRAID (>100 bp).
Length distribution was plotted for NRAID (non repeat associated indel) for human-baboon, human-chimpanzee and human-lemur genomic sequences alignments. Peaks are observed between 125-250 bp for in human-baboon and human-lemur genomic sequence alignments.

Table S8. Junction Analysis of Large RAID (Repeat Associated Indel) (>100 bp).
RAID (>=100bp) which associated with repeats but could not be classified as retrotransposition are categorized according to its 3', 5' boundary repeat class. R1, the 5' boundary of the sequence alignment gap; R2, the 3' boundary repeat gap; none, no repeat was found on the boundary but repeat(s) may reside within the insertion/deletion. Others include unclassified, small RNA, satellites, simple repeats, low complexity repeats. Assignment to species based on location of insertion sequence.

Table S9. Genome Size Difference and Repeat Class.
Based on aligned orthologous sequence, (9 loci for human-lemur and 8 loci for baboon-lemur sequence alignments) the difference (bp) and relative difference in bp and in counts is shown for each repeat class. More evolutionarily recent repeat classes, SINE/Alu and LINE/L1, repeats are dramatically decreased within the lemur genome as compared with human and baboon. Older repeat elements that integrated before the divergence of these two lineages (L2, L3, MIR and DNA) show very little difference in overall frequency.

Figure S1. Optimization of alignment gap penalty parameters.

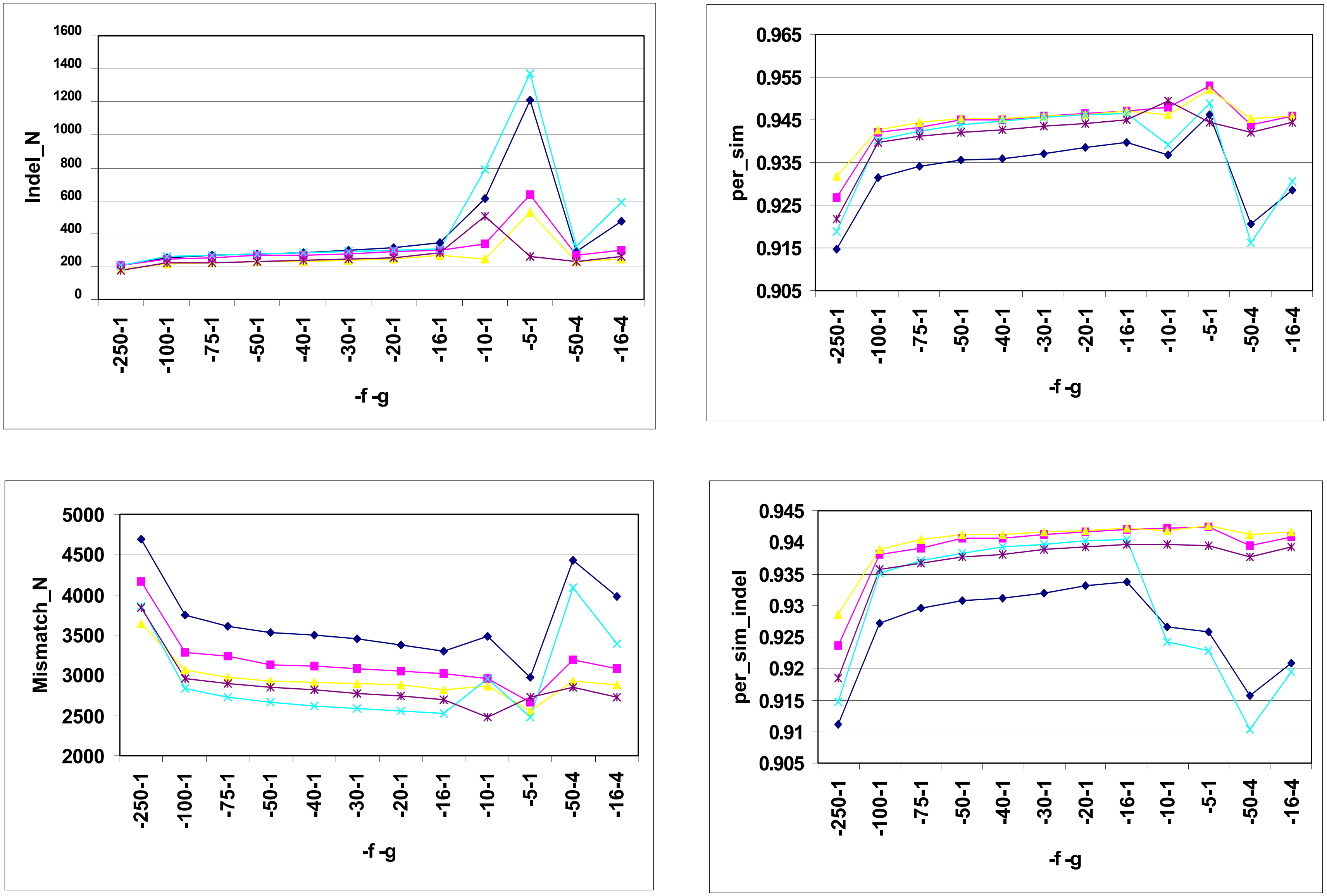


Figure S2A. Human-baboon sequence alignments.

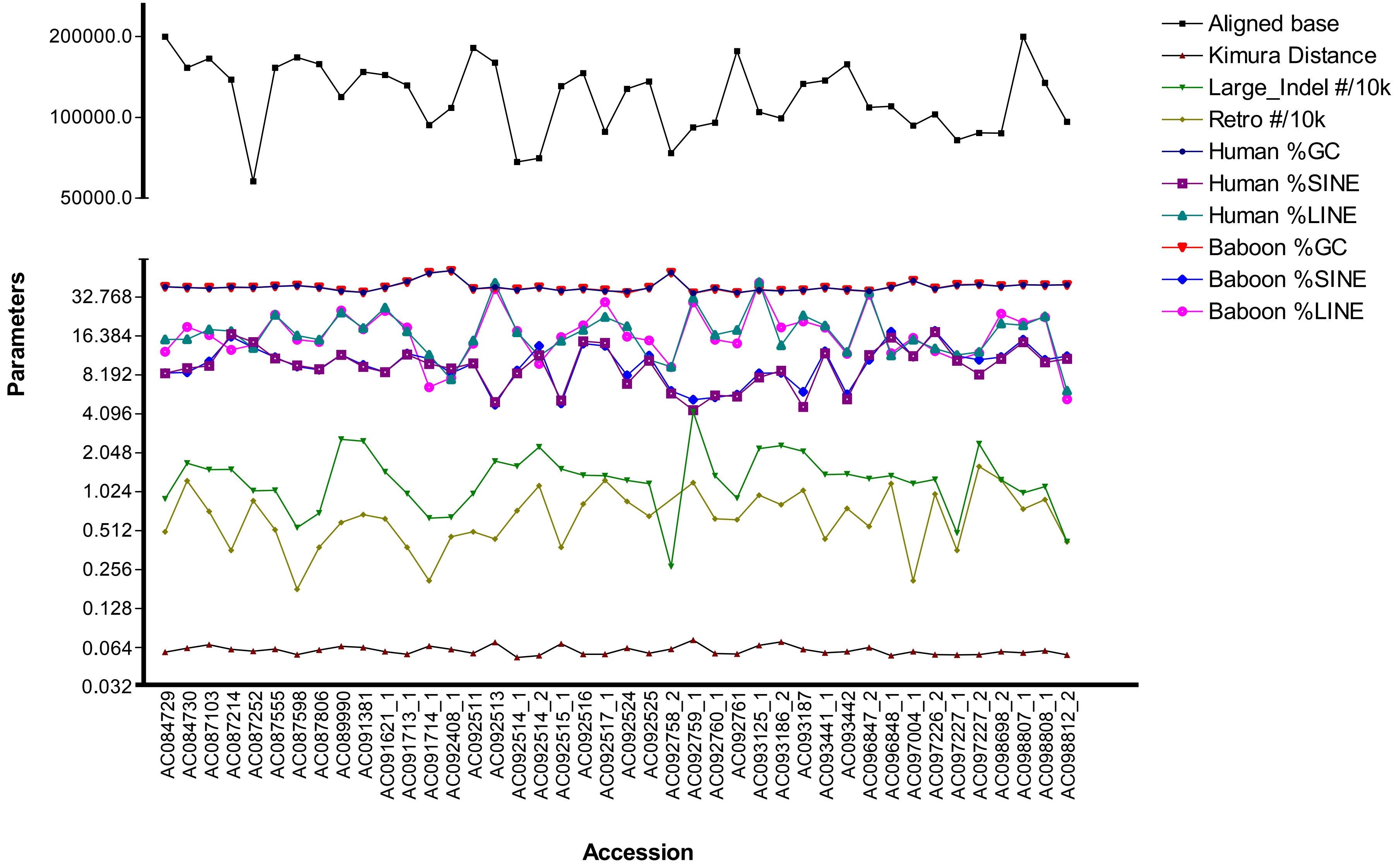


Figure S2B. Human-chimpanzee sequence alignments.

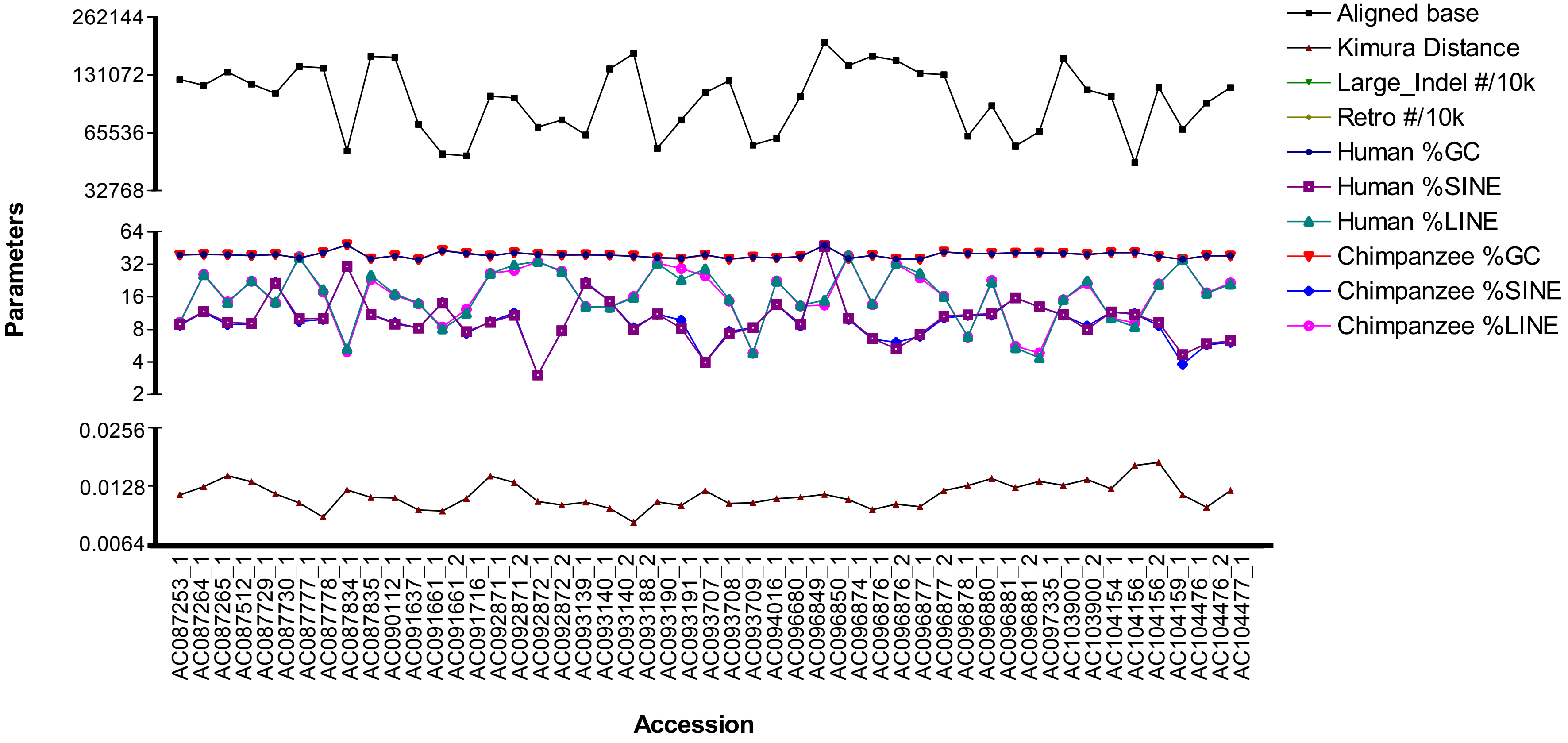


Table S3. All non human primate accession's map location with respect to the human genome

Species	Accession	Human Chr	HG_pos_B	HG_pos_E
Chimpanzee	AC087253_1	7	125456204	125581162
Chimpanzee	AC087264_1	7	125577822	125694350
Chimpanzee	AC087265_1	7	125327806	125464454
Chimpanzee	AC087512_1	7	125212754	125330783
Chimpanzee	AC087729_1	7	125696086	125805674
Chimpanzee	AC087730_1	7	126606480	126752809
Chimpanzee	AC087777_1	7	126427359	126572051
Chimpanzee	AC087778_1	7	70800000	70853035
Chimpanzee	AC087834_1	7	126699881	126869696
Chimpanzee	AC087835_1	7	126891582	127054765
Chimpanzee	AC090112_1	7	126846442	126919261
Chimpanzee	AC091637_1	7	126565474	126616364
Chimpanzee	AC091661_1	7	33906503	33956460
Chimpanzee	AC091661_2	7	33802397	33904458
Chimpanzee	AC091716_1	7	32523503	32633017
Chimpanzee	AC092871_1	7	98179517	98249871
Chimpanzee	AC092871_2	7	98077231	98154568
Chimpanzee	AC092872_1	7	97576866	97641328
Chimpanzee	AC092872_2	7	97641272	97782804
Chimpanzee	AC093139_1	7	97769610	97940065
Chimpanzee	AC093140_1	7	100411668	100466372
Chimpanzee	AC093140_2	7	100107108	100195267
Chimpanzee	AC093188_2	7	96615694	96728807
Chimpanzee	AC093190_1	7	98603581	98727830
Chimpanzee	AC093191_1	7	100653552	100711148
Chimpanzee	AC093707_1	7	100437252	100499014
Chimpanzee	AC093708_1	7	101345614	101448275
Chimpanzee	AC093709_1	7	79957136	80153435
Chimpanzee	AC094016_1	7	100494703	100642766
Chimpanzee	AC096680_1	7	99356639	99525029
Chimpanzee	AC096849_2	7	98756046	98932615
Chimpanzee	AC096850_1	7	98877265	99023391
Chimpanzee	AC096874_1	7	31087040	31219632
Chimpanzee	AC096876_1	7	30916193	30979930
Chimpanzee	AC096876_2	7	31018861	31110370
Chimpanzee	AC096877_1	7	30860451	30916635
Chimpanzee	AC096877_3	7	30793652	30860374
Chimpanzee	AC096878_1	7	30524531	30684518
Chimpanzee	AC096880_1	7	34045551	34157278
Chimpanzee	AC096881_1	7	30608131	30709889
Chimpanzee	AC096881_2	7	30710302	30757419
Chimpanzee	AC097335_1	7	126695306	126809671
Chimpanzee	AC103900_1	7	99279053	99348906
Chimpanzee	AC103900_2	7	99349278	99446988
Chimpanzee	AC104154_1	7	96109879	96224851

Chimpanzee	AC104156_1	7	95800051	95859495
Chimpanzee	AC104156_2	7	95881017	95958347
Chimpanzee	AC104159_1	7	99525274	99666958
Chimpanzee	AC104476_1	7	96194376	96270254
Chimpanzee	AC104476_2	7	96270281	96364226
Chimpanzee	AC104477_1	7	95912624	96071593
Baboon	AC084729	7	125443396	125661497
Baboon	AC084730	7	125284551	125450778
Baboon	AC087103	7	127074330	127250078
Baboon	AC087214	7	125626957	125781423
Baboon	AC087252	7	126252597	126312406
Baboon	AC087555	7	126516263	126673383
Baboon	AC087598	7	126400349	126570209
Baboon	AC087806	7	126924597	127087544
Baboon	AC089990	7	126671704	126799648
Baboon	AC091381	7	126771442	126935734
Baboon	AC091621	7	33738973	33889790
Baboon	AC091713	7	31956301	32099986
Baboon	AC091714	7	33202196	33304830
Baboon	AC092408	7	33392338	33503141
Baboon	AC092511	7	99431955	99624848
Baboon	AC092513	7	98089875	98265147
Baboon	AC092514_1	7	97819684	97889551
Baboon	AC092514_2	7	97745407	97818933
Baboon	AC092515	7	98232863	98367358
Baboon	AC092516	7	97366768	97527099
Baboon	AC092517	7	97339115	97429747
Baboon	AC092524	7	98605746	98751994
Baboon	AC092525	7	101577265	101716069
Baboon	AC092758	7	33619251	33693937
Baboon	AC092759	7	98866894	98975929
Baboon	AC092760	7	99325764	99431353
Baboon	AC092761	7	98265141	98455626
Baboon	AC093125	7	101140404	101253233
Baboon	AC093186	7	98987815	99090637
Baboon	AC093187	7	98393159	98549580
Baboon	AC093441	7	100801942	100943386
Baboon	AC093442	7	99098700	99276803
Baboon	AC096847	7	100471556	100587130
Baboon	AC096848	7	97635586	97747528
Baboon	AC097004	7	32015656	32114475
Baboon	AC097226	7	29858815	29964715
Baboon	AC097227_1	7	30176487	30259843
Baboon	AC097227_2	7	30083356	30175967
Baboon	AC098698	7	30409831	30499997
Baboon	AC098807	7	31200256	31407650
Baboon	AC098808	7	30500266	30639035
Baboon	AC098812	7	30756181	30854421
Lemur	AC123543_1	7	126735098	126829353

Lemur	AC123543_2	7	126848422	126961833
Lemur	AC123544_1	7	126461921	126656554
Lemur	AC123970_1	7	125516578	125565453
Lemur	AC123971_1	7	127055958	127101317
Lemur	AC123971_2	7	126987165	127049740
Lemur	AC118574_1	11	132765214	132841665
Lemur	AC118569_1	19	14331781	14437483
Lemur	AC118570_2	19	21399214	21545653

Fig. S4.

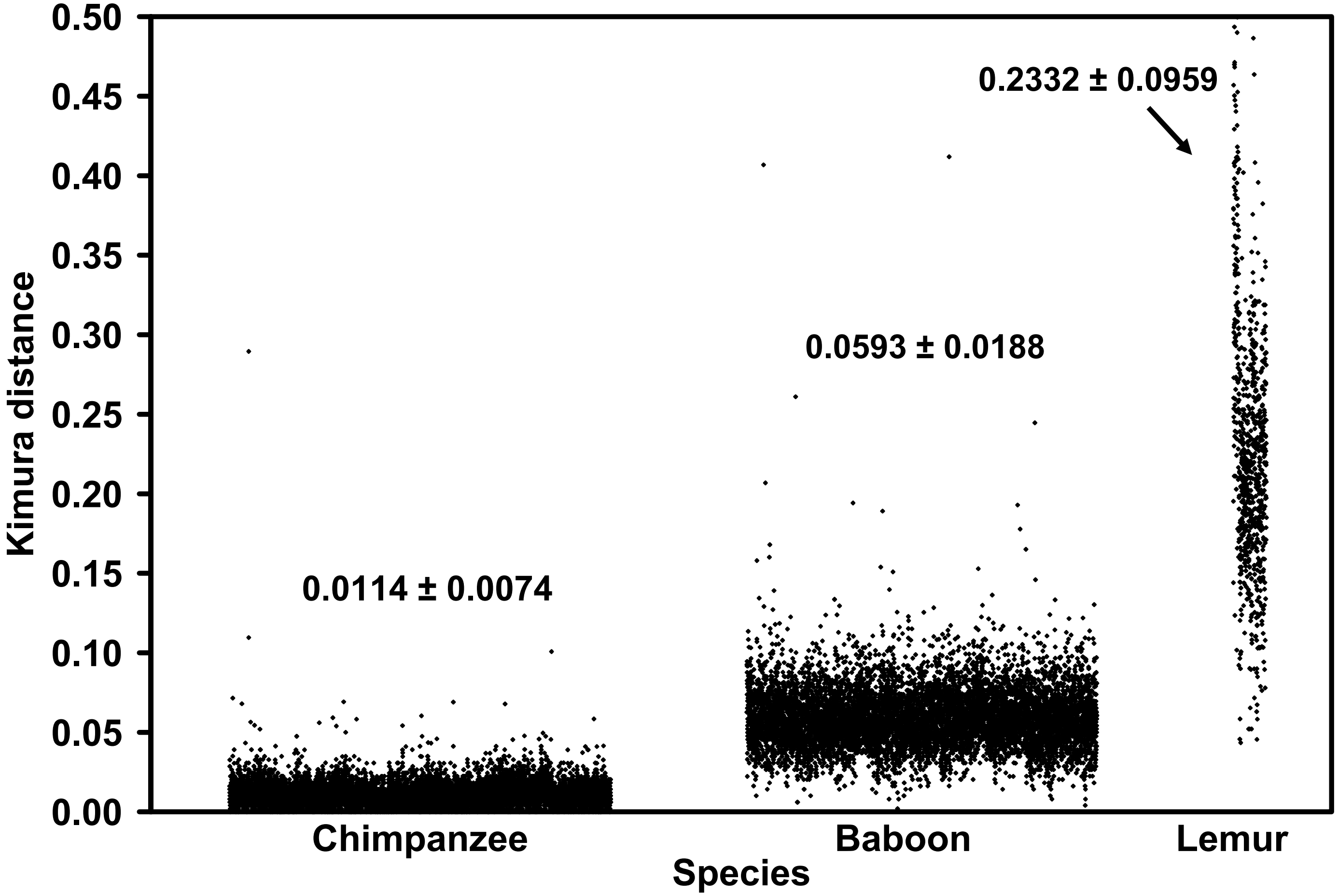

0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Kimura distance
0.0114 ± 0.0074
0.0593 ± 0.0188
0.2332 ± 0.0959
Chimpanzee
Baboon
Lemur
Species

Table S5. Count of the Total Number of Insertion/Deletion Events and Their Length Distribution.

Human Chimpanzee		Both	Human	Chimp
All_Indel*	COUNT	7102	3465	3637
	TOTAL_LEN	228029	161289	66740
	AVE_LEN	32	47	18
Indel*	COUNT	88	56	33
	TOTAL_LEN	115104	87751	27659
	AVE_LEN	1308	1567	838
Retro*	COUNT	49	34	16
	TOTAL_LEN	51207	43673	7840
	AVE_LEN	1045	1285	490
RAID	COUNT	31	19	12
	TOTAL_LEN	60286	42588	17698
	AVE_LEN	1945	2241	1475
NRAID	COUNT	8	3	5
	TOTAL_LEN	3611	1490	2121
	AVE_LEN	451	497	424
Human Baboon		Both	Human	Baboon
All_Indel***	COUNT	25522	13139	12383
	TOTAL_LEN	706522	367128	339394
	AVE_LEN	28	28	27
Indel	COUNT	774	369	405
	TOTAL_LEN	1576883	771653	805230
	AVE_LEN	2037	2091	1988
Retro**	COUNT	364	153	211
	TOTAL_LEN	246402	130505	115897
	AVE_LEN	677	853	549
RAID	COUNT	365	195	170
	TOTAL_LEN	1314364	634338	680026
	AVE_LEN	3601	3253	4000
NRAID	COUNT	45	21	24
	TOTAL_LEN	16117	6810	9307
	AVE_LEN	358	324	388

Figure S6.

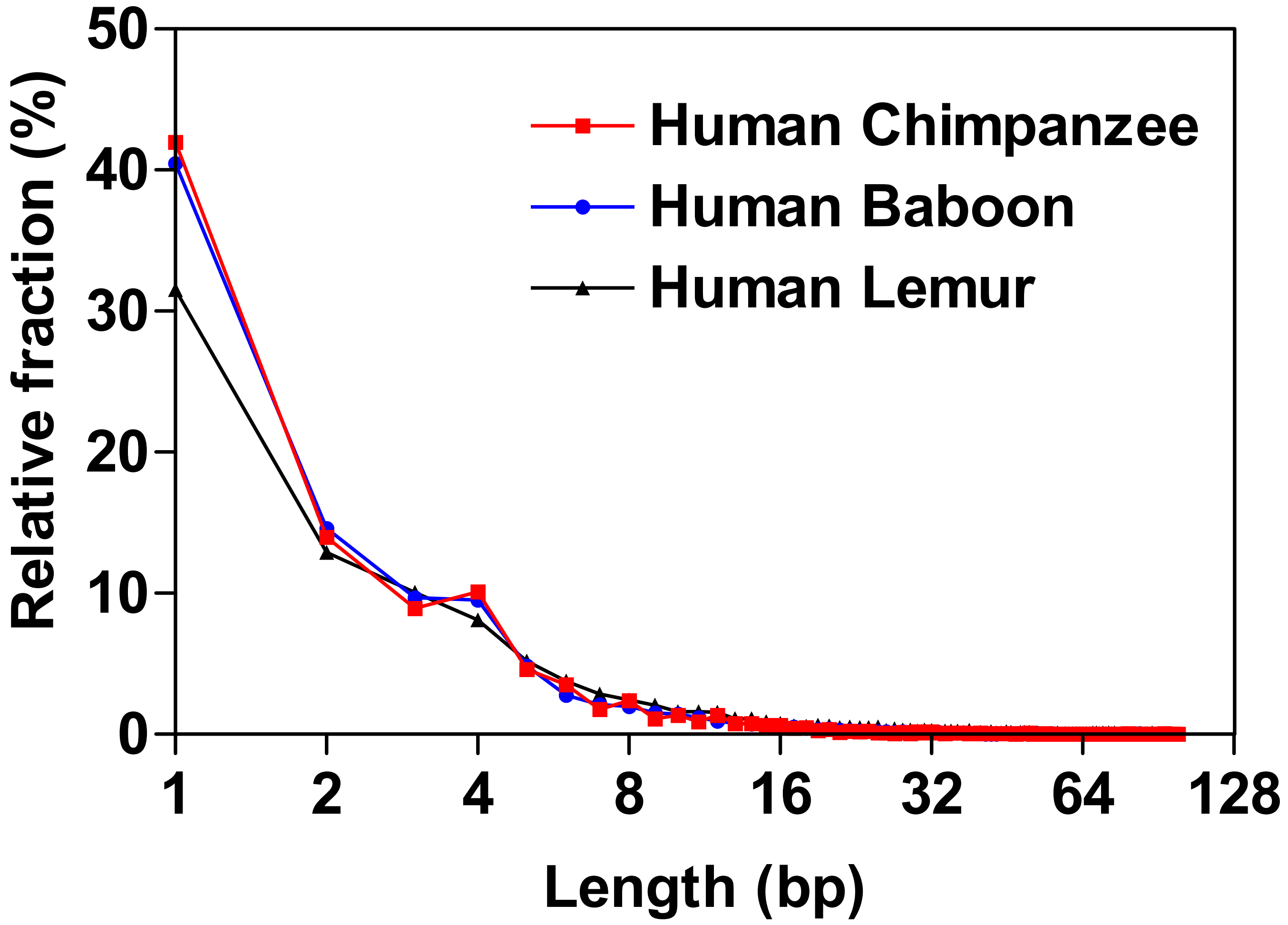
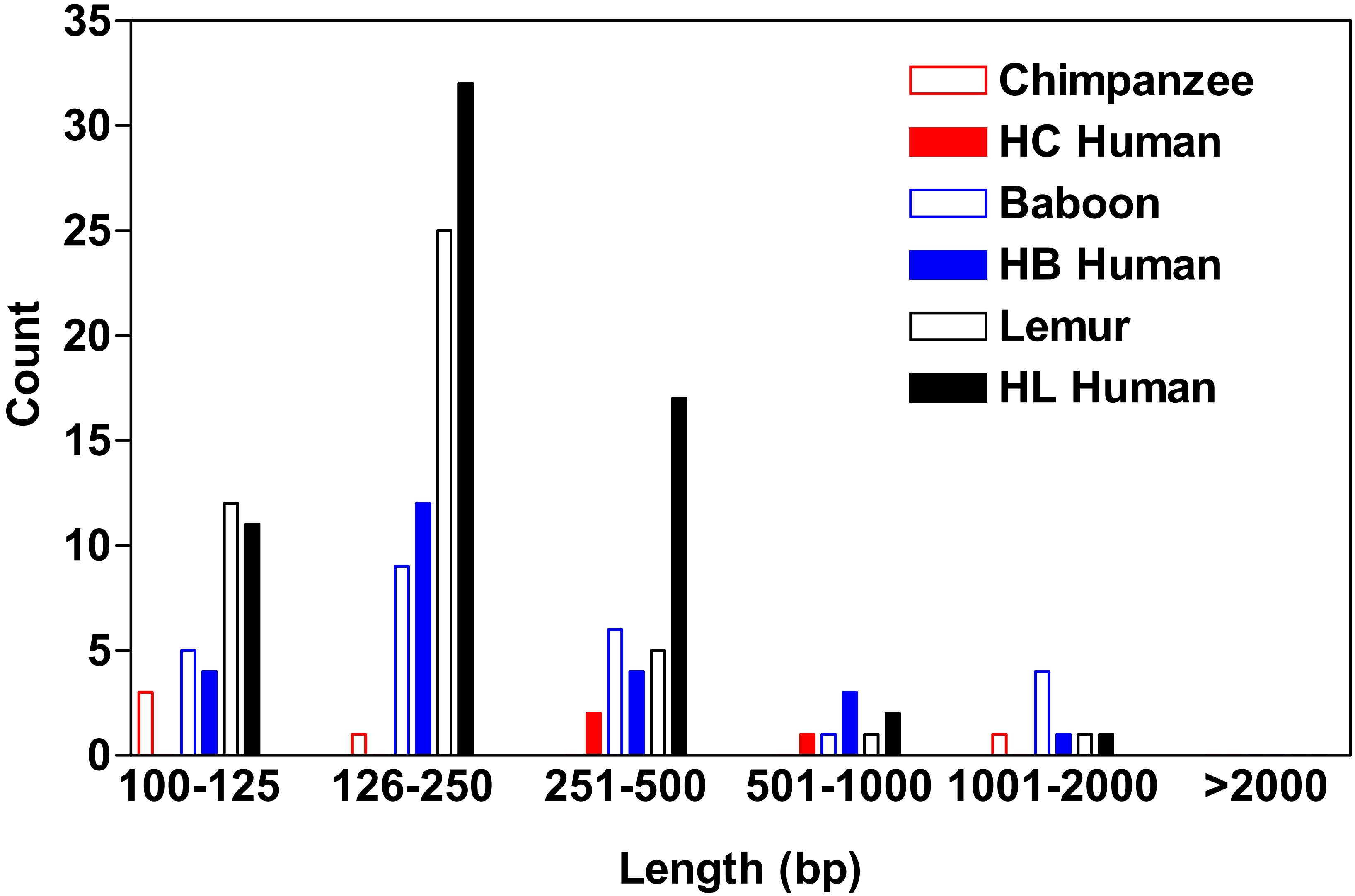

50
40
30
20
10
0
Relative fraction (%)
1
2
4
8
16
32
64
128
Length (bp)
Human Chimpanzee
Human Baboon
Human Lemur

Figure S7.



Count
0
5
10
15
20
25
30
35
Chimpanzee
HC Human
Baboon
HB Human
Lemur
HL Human
100-125
126-250
251-500
501-1000
1001-2000
>2000
Length (bp)

Table S8.

		R2 (3’)						
	R1 (5’)	**LINE**	**SINE**	**LTR**	**DNA**	**None**	**Others**	**subtotal**
Baboon	**LINE**	8	0	1	1	7	0	17
	SINE	4	10	1	0	4	0	19
	LTR	0	0	6	0	4	0	10
	DNA	0	0	0	0	3	0	3
	None	6	5	1	2	9	0	23
	Others	0	0	0	0	0	0	0
	subtotal	18	15	9	3	27	0	72
Human	**LINE**	14	2	0	0	11	0	27
	SINE	0	11	1	0	11	1	24
	LTR	0	0	1	0	3	0	4
	DNA	2	0	0	1	0	0	3
	none	4	5	2	0	16	0	27
	other	0	0	0	0	0	0	0
	subtotal	20	18	4	1	41	1	85

Table S9.

A. Human –Lemur sequence alignments

	Human_BP	Lemur_BP	Dif_BP	Dif_BP%	Human_N	Lemur_N	Dif_N
SINEs:	**145065**	**51861**	**93204**	**10.08%**	**608**	**277**	**331**
ALUs	123290	30756	92534	10.01%	456	124	332
MIRs	21775	21105	670	0.07%	152	153	-1
LINEs:	**162507**	**114362**	**48145**	**5.21%**	**255**	**239**	**16**
LINE1	128504	82207	46297	5.01%	139	128	11
LINE2	30996	29690	1306	0.14%	105	103	2
L3/CR1	3007	2465	542	0.06%	11	8	3
LTR	**32121**	**22249**	**9872**	**1.07%**	**76**	**63**	**13**
MaLRs	20619	15282	5337	0.58%	52	44	8
ERVL	4813	3269	1544	0.17%	12	10	2
ERV_classI	6689	3698	2991	0.32%	12	9	3
ERV_classII	0	0	0	0.00%	0	0	0
DNA	**21922**	**21857**	**65**	**0.01%**	**89**	**98**	**-9**
MER1_type	12059	13422	-1363	-0.15%	55	67	-12
MER2_type	6650	3762	2888	0.31%	18	12	6
Unclassified:	255	348	-93	-0.01%	1	1	0
Total IR	*358657*	*206004*	*152653*	*16.51%*	*1013*	*659*	*354*
Small	773	240	533	0.06%	4	3	1
Satellites:	0	0	0	0.00%	0	0	0
Simple	7578	2756	4822	0.52%	108	61	47
Low	4243	3238	1005	0.11%	95	91	4
Total Other	*12594*	*6234*	*6360*	*0.69%*	*207*	*155*	*52*
Grand Total	374742	216540	159013	17.20%	1220	814	406

B. Baboon-Lemur sequence alignments

	Baboon_BP	Lemur_BP	Dif_BP	Dif_BP%	Baboon_N	Lemur_N	Dif_N
SINEs:	**83116**	**42268**	**40848**	**5.17%**	**373**	**238**	**135**
ALUs	58949	17358	41591	5.26%	215	69	146
MIRs	24167	24910	-743	-0.09%	158	169	-11
LINEs:	**132710**	**93831**	**38879**	**4.92%**	**223**	**207**	**16**
LINE1	99962	64883	35079	4.44%	106	102	4
LINE2	28725	25268	3457	0.44%	103	94	9
L3/CR1	4023	3680	343	0.04%	14	11	3
LTR	**33299**	**21687**	**11612**	**1.47%**	**74**	**60**	**14**
MaLRs	17820	12264	5556	0.70%	41	35	6
ERVL	6887	5769	1118	0.14%	18	16	2
ERV_classI	7523	3654	3869	0.49%	14	9	5
ERV_classII	1069	0	1069	0.14%	1	0	1
DNA	**19885**	**22428**	**-2543**	**-0.32%**	**82**	**90**	**-8**
MER1_type	9054	11294	-2240	-0.28%	46	54	-8
MER2_type	8200	5552	2648	0.34%	18	14	4
Unclassified:	351	348	3	0.00%	1	1	0
Total IR	*269361*	*180562*	*91750*	*11.61%*	*735*	*574*	*161*
Small	103	186	-83	-0.01%	1	2	-1
Satellites:	0	0	0	0.00%	0	0	0
Simple	4378	2232	2146	0.27%	104	49	55
Low	4414	4187	227	0.03%	113	120	-7
Total Other	*8895*	*6605*	*2290*	*0.29%*	*218*	*171*	*47*
Grand Total	278145	187084	94040	11.90%	953	745	208