

Figure S1

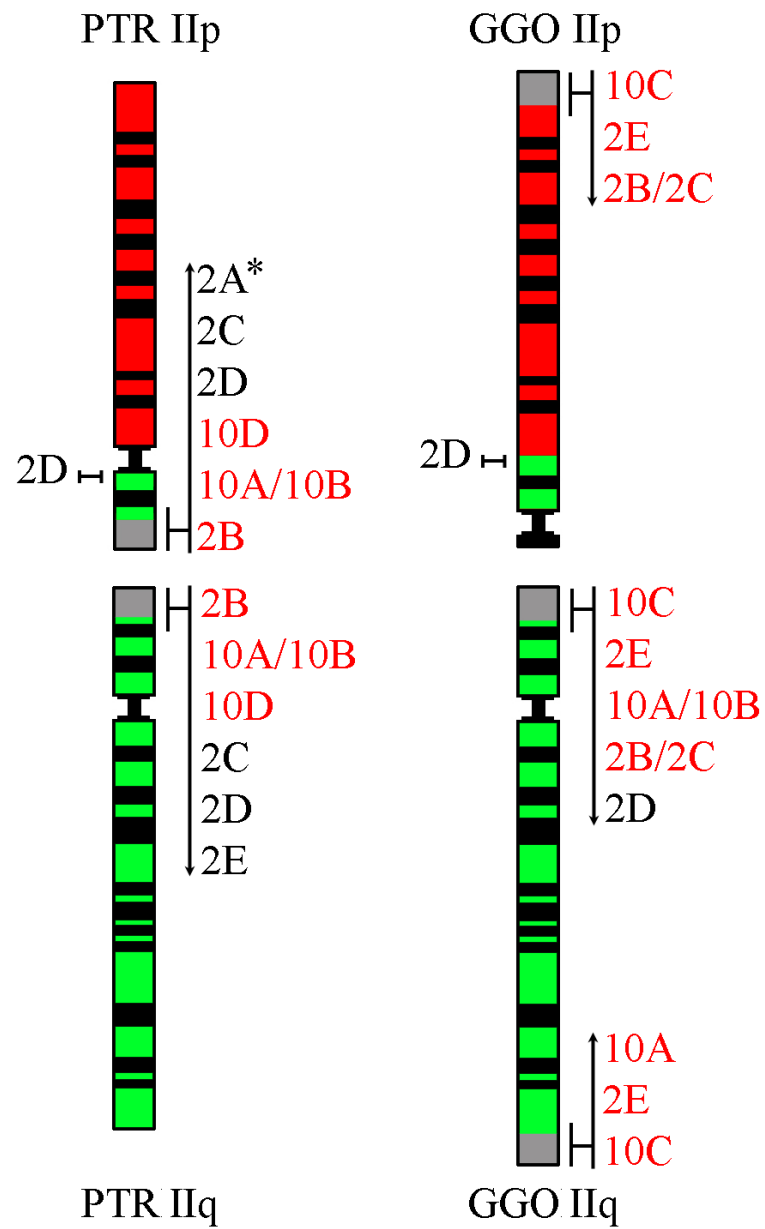


Figure S2

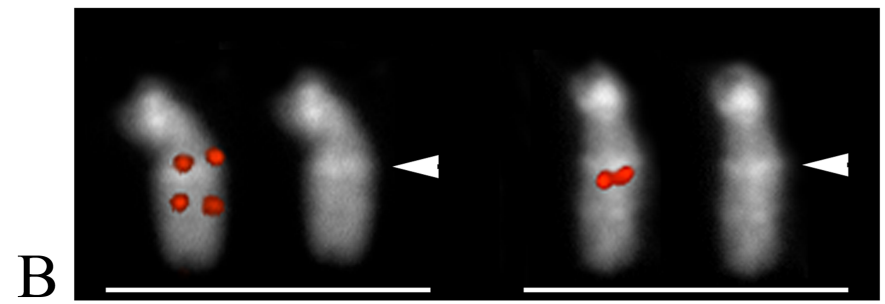
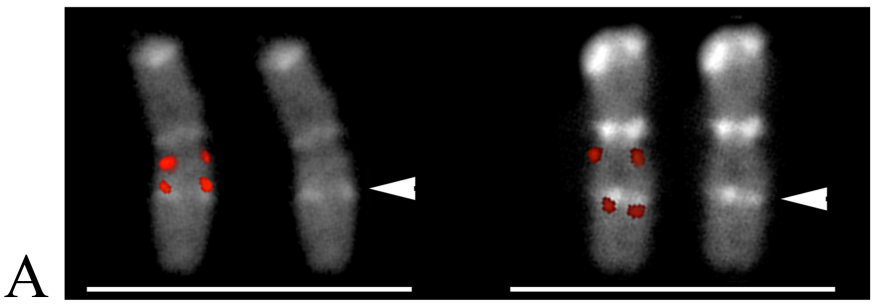
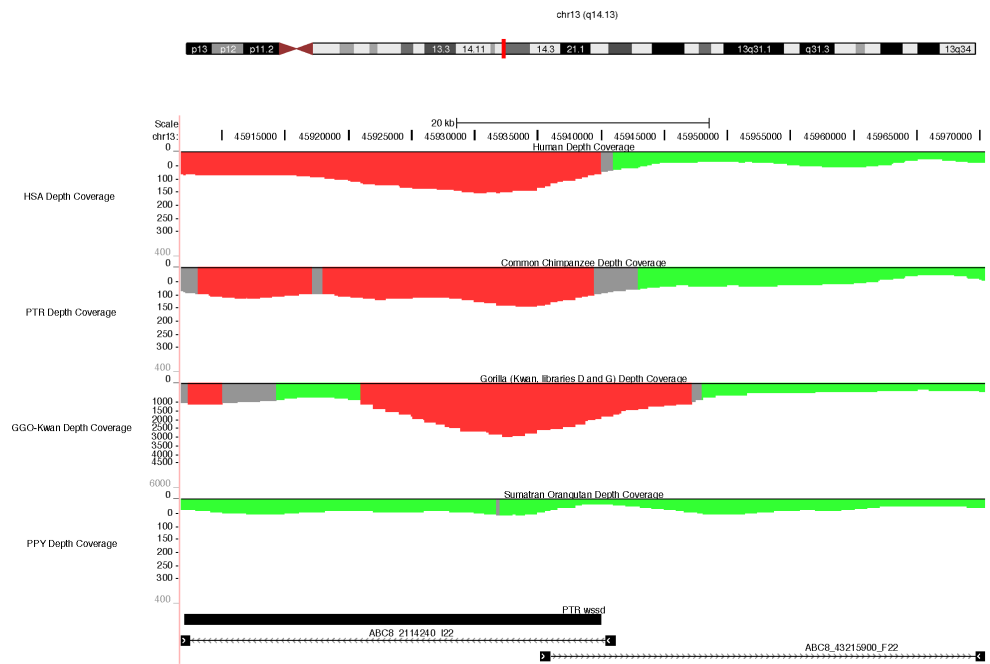
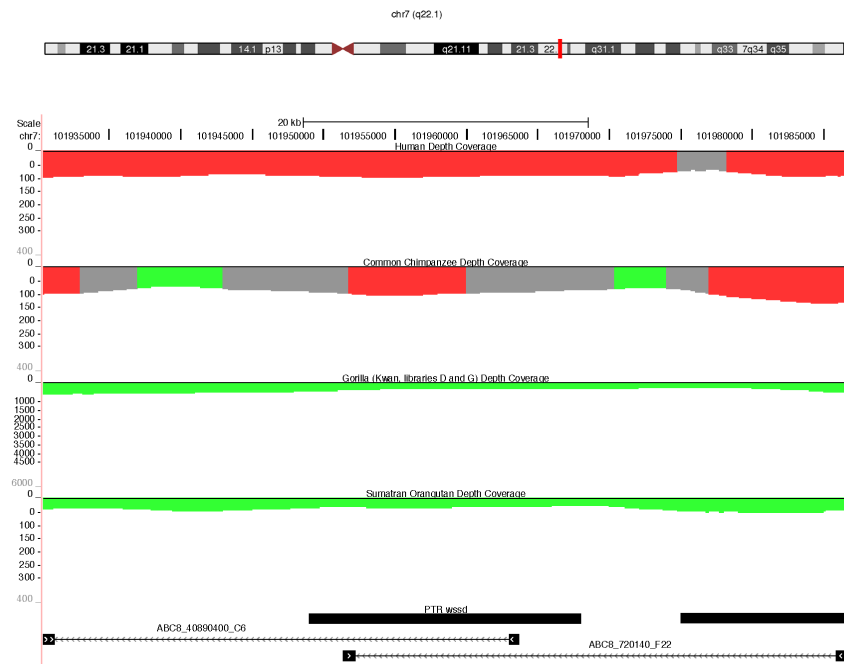


Figure S3

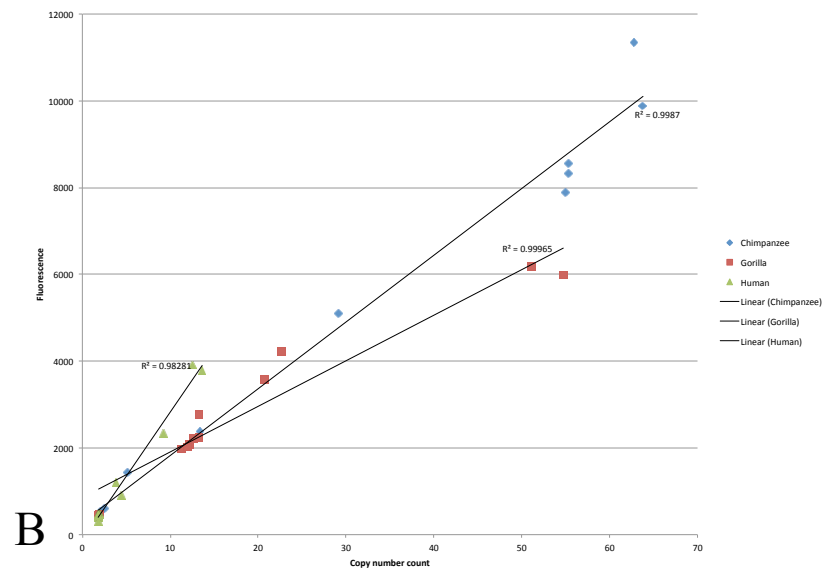
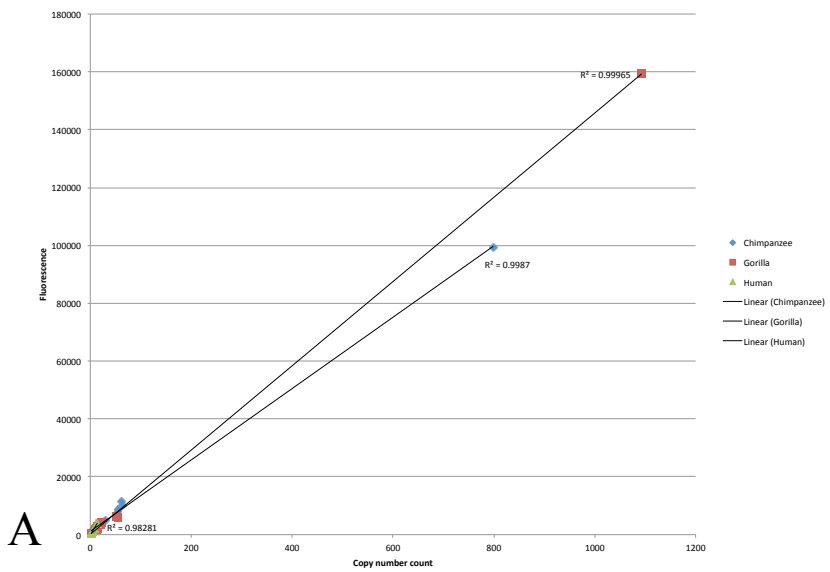
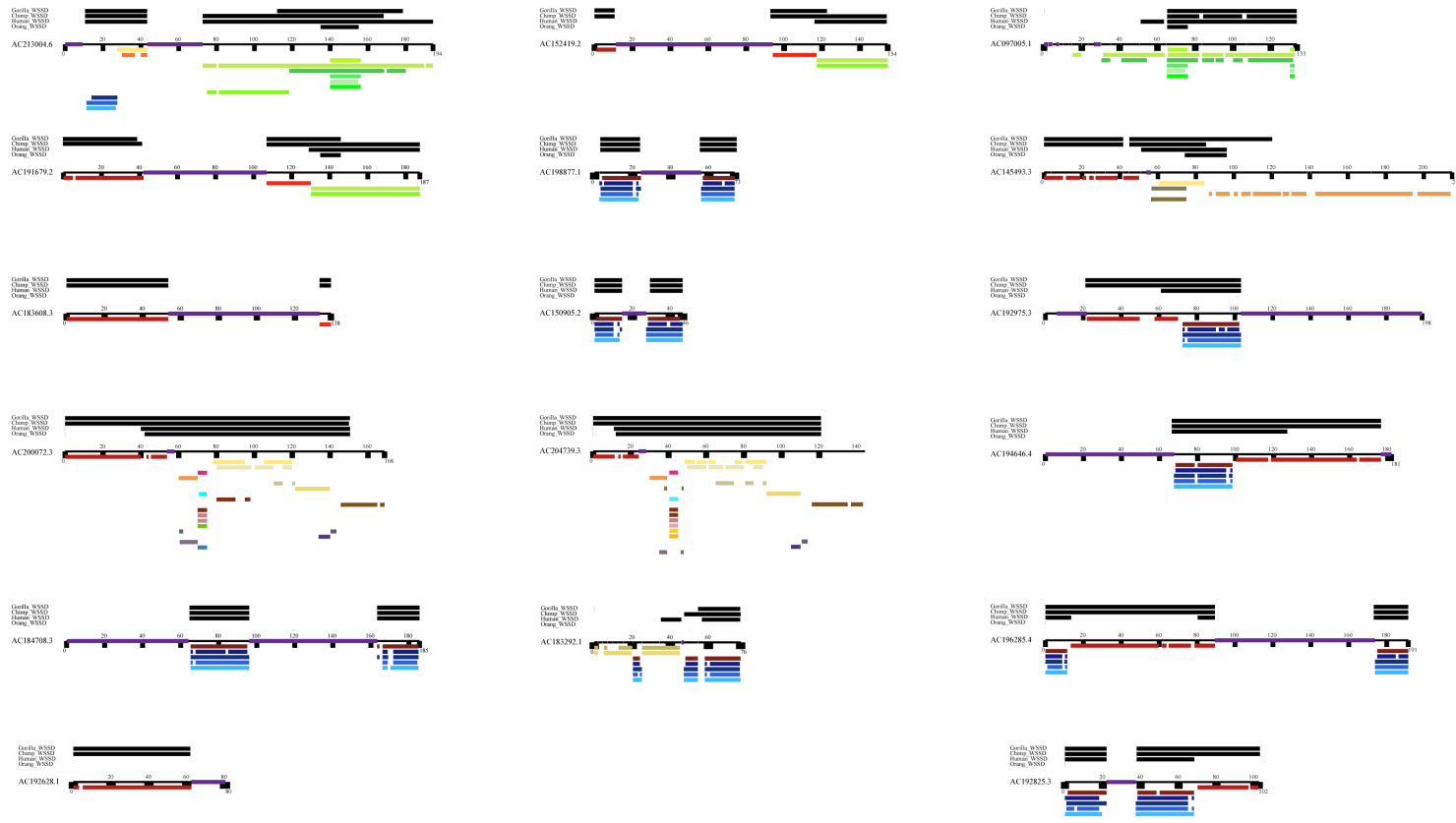


Figure S4



PTR specific regions

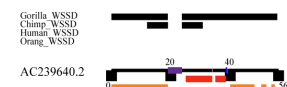
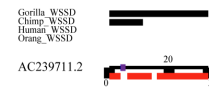
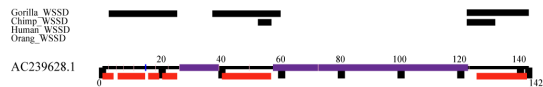
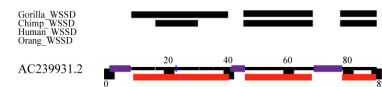
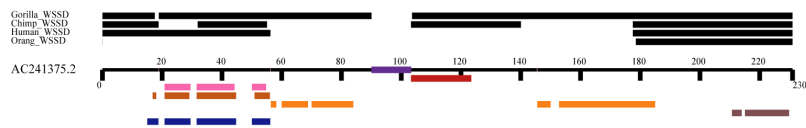
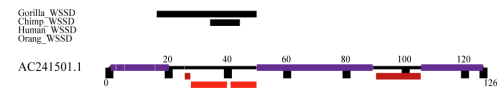
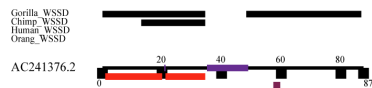
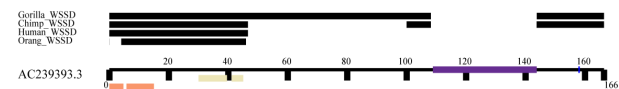
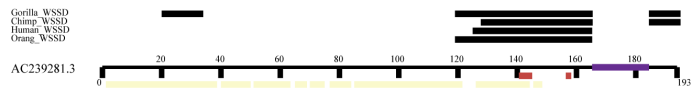
- chr1:76977-100983
- chr1:78739-100983
- chr1:225578-247298
- chr1:698480-729985
- chr1:795156-804654
- chr1:120372053-120405147
- chr1:142602189-142648762
- chr1:220405325-220467093
- chr1:239525339-239565523
- chr2:43799262-43803855
- chr2:10620864-10621910
- chr5:39824507-39829506
- chr5:151437589-151442629
- chr5:151568491-151573091
- chr6:105007438-105012021
- chr6:105045396-105049999
- chr7:30253096-30258095
- chr7:43774170-43854629
- chr7:66185868-66202117
- chr7:74767488-74782267
- chr7:75801303-75817475
- chr7:76346757-76453622
- chr7:101575631-101591846
- chr7:101708105-102022998
- chr8:5-10917
- chr8:151487-159266
- chr8:170359-176239
- chr9:119898066-119903096
- chr13:107960815-107965854
- chr16:21052-140599
- chr16:8881171-88821485
- chr18:55222147-55226748
- chr18:55590898-55594894
- chr19:130718-148638
- chr19:212828-240466

PTR/GGO shared regions

- chr1:239525339-239565523
- chr2:113970467-114032187
- chr9:2285-111502
- chr9:66942239-66987465
- chr10:19441552-19528768
- chr10:19541022-19610790
- chr9:67180348-67245795
- chr9:68144022-68217981

• 1Mb apart from telomere
 • 1Mb apart from centromere

Figure S5 A



PTR/GGO shared regions



GGO specific regions

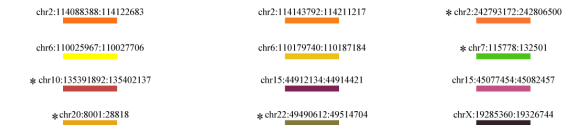


Figure S5 B

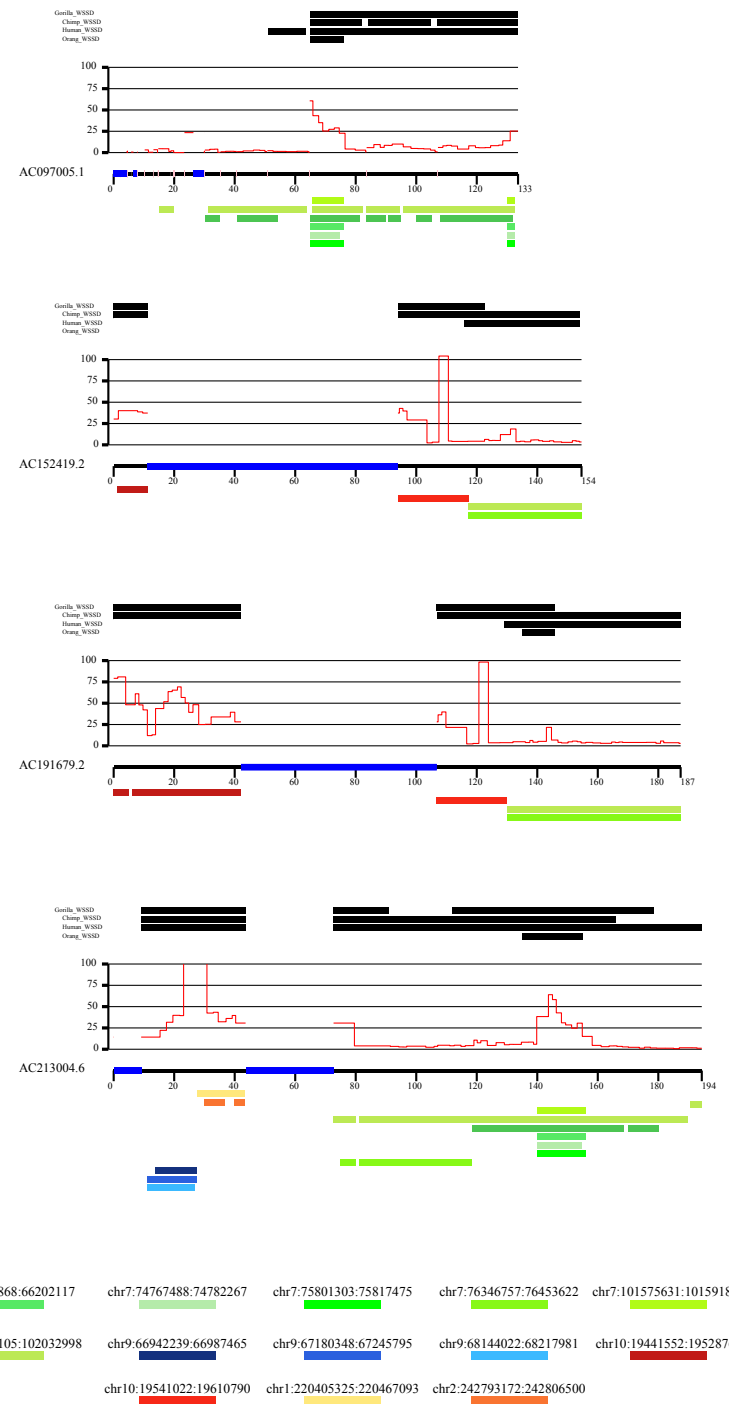


Figure S6

Table S1. Mapping of probes on chromosomes 2 and 10

	Fosmid Clone	Position NCBI35	PTR IIq_parm	PTR IIq_qarm	PTR IIp_parm	PTR IIp_Cenparm	PTR IIp_qarm	GGO IIq_parm	GGO IIq_qarm	GGO IIp_parm	GGO IIp_interstitial	GGO IIp_qarm
	ABC8 41147200 H23	chr2:113840263-113878692	No	No	Yes	Yes*	No	No	No	No	Yes	No
2A	ABC8 42163800 E20#	chr2:113884171-113915881	No	No	Yes	No	No	No	No	No	No	No
	ABC8 41115100 I17#	chr2:113894854-113929406	No	No	Subter	No	No	No	No	No	No	No
	ABC8 43008200 P2#	chr2:113919580-113958771	No	No	Subter	No	No	No	No	No	No	No
2B	ABC8 40925900 F12#	chr2:113978394-114020431	Cap	No	Cap	No	No	Subcap	No	No	No	Subcap (het)
	ABC8 41134100 D22#	chr2:114021524-114059681	Cap	No	Cap	No	No	Subcap	No	No	No	Subcap (het)
2C	ABC8 40903600 D20@	chr2:114080059-114122590	Subter	No	Subter	No	No	Subcap (weak)	No	No	No	Subcap (het)
2D	ABC8 660422 B2\$	chr2:114111934-114153971	Yes	No	Yes	Yes*	No	Yes (weak)	No	No	Yes	No
2E	ABC8 2126940 M24	chr2:114149372-114184630	Yes	No	No	No	No	Subcap	Subcap	No	No	Subcap
	ABC8 738840 B20	chr2:114196346-114231473	Yes	No	No	No	No	Subcap (het)	Subcap (het)	No	No	Subcap
	ABC8 45707700 L24	chr2:114223493-114258048	Yes	No	No	No	No	Yes	No	No	No	No
	ABC8 41123600 B12	chr2:114267007-114303469	Yes	No	No	No	No	Yes	No	No	No	No
	ABC8 43746700 B12	chr10:19359773-19395890	No	No	No	No	No	No	No	No	No	No
	ABC8 708622 I7	chr10:19399597-19435714	No	No	No	No	No	No	No	No	No	No
10A	ABC8 41110300 L12	chr10:19427867-19467099	Subcap	No	Subcap	No	No	No	No	No	No	No
	ABC8 757522 C14	chr10:19428257-19466794	Subcap	No	Subcap	No	No	Subcap	No	No	No	No
	ABC8 661222 G22	chr10:19465621-19502910	Subcap	No	Subcap	No	No	No	Subcap (het)	No	No	No
10B	ABC8 40859600 G4	chr10:19496838-19535777	Subcap	No	Subcap	No	No	Subcap (het)	Subcap	No	No	No
10C	ABC8 40868200 C16	chr10:19530349-19564732	No	No	No	No	No	Cap	Cap	No	No	Cap
10D	ABC8 41045500 J9	chr10:19581695-19620805	Subcap	No	Subcap	No	No	No	No	No	No	No
	ABC8 2134240 I12	chr10:19615532-19651206	No	No	No	No	No	No	No	No	No	No
	ABC8 42229400 M23	chr10:19655864-19700390	No	No	No	No	No	No	No	No	No	No

Note. * Shared with human; # Duplicated on chr 9 in human and chimpanzee; \$ Identifies chr 2q inv bp; Subter, light up tel of chr with no caps; Subcap, light up tel of chr with no caps; @ Duplicated on chr 2q in human, in bold clones identifying the CAP in chimpanzee or gorilla.

Table S2. Definition of interstitial heterochromatic integration point in chimpanzee

Clone	Position	Chimpanzee mapping
RP11-1134K14	chr7:101,890,344-102,048,014	M. Gross 2006*
ABC8 40893700 L2	chr7:101906133-101936332	7qprox
ABC8 40890400 C6	chr7:101930395-101963680	7q11.2; 7qprox
ABC8 720140 F22	chr7:101951324-101986595	7q11.2; 7qdist
ABC8 42230900 F1	chr7:101969183-102013237	7q11.2; 7qdist
ABC8 41051200 J2	chr7:101787281-101830912	7p; 7q11.2; 7qdist
ABC8 42164700 J9	chr7:101862516-101906594	7p; 7q11.2; 7qdist
RP11-189B4	chr13:45,867,742-45,984,270	M. Gross 2006*
ABC8 2112940 H11	chr13:45872397-45910777	13qprx; 13qdup
ABC8 2114240 I22	chr13:45906763-45941160	13qprx ; 13qdup
ABC8 43215900 F22	chr13:45935232-45970371	13qdistal
WIBR2-2936M02 G248P89166G1	chr13:45,966,323-46,009,057	13qdistal

Note. * Position as reported in Gross et al. 2006.

Table S3. Copy number estimates of gorilla and chimpanzee hyperexpanded loci and proximal regions

Genomic regions	Probe	Chr	Start	End	Window size (bp)	HUMAN			CHIMPANZEE			GORILLA			NEANDERTHAL##	
						Genomic amount (bp)	(n=175)	Fluor Value	Genomic amount (bp)	(n=12)	Fluor Value	Genomic amount (bp)	(n=4)	Fluor Value	Genomic amount (bp)	(n=1)
2A	abc8_43008200_p2	chr2	113919580	113958771	39191	488712	12	3907	522416	13	2377	76031	2	465	548674	14
2B	abc8_40925900_f12	chr2	113978394	114020431	42037	390103	9	2346	33545526	798	99316	529666	13	2222	420370	10
	abc8_41134100_d22	chr2	114021524	114059681	38157	520080	14	3777	2099398	55	7885	428885	11	1980	534198	14
2C	abc8_40903600_d20	chr2	114080059	114122590	42531	188838	4	905	221161	5	1436	519304	12	2078	212655	5
2D	abc8_660422_b2	chr2	114111934	114153971	42037	160161	4	1206	106774	3	602	500240	12	2013	168148	4
2E	abc8_2126940_m24	chr2	114149372	114184630	35258	68753	2	479	64522	2	438	1928965	55	5976	68753	2
	abc8_738840_b20	chr2	114196346	114231473	35127	68146	2	459	66741	2	484	1795341	51	6194	68146	2
10A	abc8_41110300_l12	chr10	19427867	19467099	39232	71010	2	421	2167568	55	8564	522570	13	2252	71010	2
	abc8_757522_c14	chr10	19428257	19466794	38537	69752	2	399	2129169	55	8332	513313	13	2761	69752	2
	abc8_661222_g22	chr10	19465621	19502910	37289	69358	2	456	2378292	64	9873	774120	21	3563	69358	2
10B	abc8_40859600_g4	chr10	19496838	19535777	38939	70090	2	309	2444590	63	11354	880021	23	4237	38939	1
10C	abc8_40868200_c16	chr10	19530349	19564732	34383	61889	2	460	71860	2	539	37546236	1092	159360	61889	2
10D	abc8_41045500_j9	chr10	19581695	19620805	39110	70398	2	436	1139274	29	5116	72354	2	445	70398	2
						2297290#		0.9981*	46971586#		0.9987*	46078105#		0.99965*	2402290	

Note. * Correlation (r^2) values between copy number and fluorescence value; # Total genomic amount; ## Reich D. et al. 2010

Table S5. Pairwise distance calculation

Region: chr2 113996786 114006404															
	Start	End	Sequence size	1	2	3	4	5	6	7	8	9	10	11	
[1] #AC183292.1 rc	60717	66802	6085	1											
[2] #hg17 chr9	72761	83240	10479	2	0.02										
[3] #ponAbe2 chr9	63019186	63029250	10064	3	0.037	0.033									
[4] #hg17 chr9 rc	68182581	68192206	9625	4	0.019	0.013	0.029								
[5] #hg17 chr2	113996786	114006404	9618	5	0.021	0.009	0.034	0.012							
[6] #AC192825.3	12002	18076	6074	6	0.015	0.02	0.037	0.02	0.022						
[7] #AC192825.3	59453	67944	8491	7	0.014	0.017	0.034	0.016	0.017	0.013					
[8] #AC184708.3 rc	163323	177841	14518	8	0.014	0.019	0.036	0.019	0.02	0.007	0.011				
[9] #AC184708.3 rc	64064	73947	9883	9	0.012	0.018	0.036	0.017	0.019	0.007	0.01	0.006			
[10] #AC150905.2	4537	14509	9972	10	0.011	0.019	0.036	0.018	0.02	0.012	0.012	0.011	0.009		
[11] #AC198877.1	14340	20411	6071	11	0.011	0.019	0.037	0.018	0.02	0.012	0.013	0.011	0.009	0.002	

Region: chr2 114016544 114026608																	
	Start	End	Sequence size	1	2	3	4	5	6	7	8	9	10	11	12	13	14
[1] #AC150905.2	27266	37264	9998	1													
[2] #AC198877.1	55930	65925	9995	2	0.004												
[3] #AC183292.1	47063	54412	7349	3	0.009	0.009											
[4] #AC184708.3 rc	86024	96028	10004	4	0.006	0.006	0.007										
[5] #AC196285.4	173832	183829	9997	5	0.008	0.007	0.009	0.007									
[6] #AC192825.3	37402	47391	9989	6	0.008	0.008	0.009	0.007	0.009								
[7] #AC194646.4	67566	77574	10008	7	0.015	0.015	0.016	0.014	0.015	0.013							
[8] #AC192975.3 rc	93195	103219	10024	8	0.013	0.013	0.013	0.012	0.013	0.011	0.015						
[9] #hg17 chr2	114016544	114026608	10064	10	0.018	0.018	0.018	0.016	0.018	0.015	0.019	0.017					
[10] #hg17 chr9	52553	62603	10050	11	0.018	0.018	0.018	0.016	0.018	0.016	0.018	0.017	0.005				
[11] #hg17 chr9	66947388	66957451	10063	12	0.018	0.018	0.018	0.016	0.018	0.016	0.02	0.017	0.014	0.014			
[12] #hg17 chr9	67185486	67195407	9921	13	0.018	0.018	0.017	0.016	0.018	0.016	0.019	0.017	0.013	0.014	0.006		
[13] #hg17 chr9 rc	68202409	68212474	10065	14	0.018	0.018	0.018	0.016	0.018	0.015	0.019	0.017	0.011	0.011	0.012	0.011	
[14] #ponAbe2 chrUn	15170555	15180669	10114	15	0.033	0.034	0.034	0.032	0.034	0.031	0.034	0.033	0.027	0.028	0.028	0.028	0.025

Region: chr10 19564292 19576051													
	Start	End	Sequence size	1	2	3	4	5	6	7	8	9	10
[1] #AC241376.2	21136	28864	7728	1									
[2] #AC239711.2 rc	1	3712	3711	2	0.008								
[3] #AC239931.2 rc	15929	22179	6250	3	0.02	0.018							
[4] #AC239628.1 rc	10585	14268	3683	4	0.019	0.017	0.021						
[5] #AC239931.2	77917	81190	3273	5	0.018	0.016	0.005	0.02					
[6] #hg17 chr10 rc	19564292	19576051	11759	6	0.025	0.023	0.024	0.027	0.022				
[7] #AC152419.2	93953	103951	9998	7	0.015	0.013	0.014	0.017	0.012	0.016			
[8] #AC183608.3	132940	138895	5955	8	0.014	0.012	0.014	0.016	0.012	0.015	0		
[9] #ponAbe2 chr10	20026147	20030619	4472	9	0.045	0.043	0.045	0.047	0.042	0.035	0.037	0.037	
[10] #AC191679.2	106779	115272	8493	10	0.014	0.012	0.014	0.016	0.012	0.015	0	0	0.037

Region: chr10 19541142 19549347														
	Start	End	Sequence size	1	2	3	4	5	6	7	8	9	10	11
[1] #AC241376.2	10749	20725	9976	1										
[2] #AC239628.1	134030	142675	8645	2	0.002									
[3] #AC239711.2 rc	5002	14078	9076	3	0.002	0.002								
[4] #AC239598.1	62011	72009	9998	4	0.002	0.002	0.002							
[5] #AC239931.2	58442	68453	10011	5	0.018	0.019	0.017	0.019						
[6] #ponAbe2 chr10	19994255	20002519	8264	6	0.049	0.049	0.047	0.049	0.053					
[7] #hg17 chr10 rc	19541143	19549347	8204	7	0.027	0.027	0.025	0.027	0.03	0.048				
[8] #AC239931.2 rc	81283	84647	3364	8	0.014	0.014	0.012	0.014	0.015	0.048	0.026			
[9] #AC241501.1 rc	27610	30977	3367	9	0.018	0.018	0.016	0.018	0.02	0.053	0.031	0.015		
[10] #AC241375.2 rc	123748	127093	3345	10	0.011	0.011	0.009	0.011	0.015	0.047	0.024	0.01	0.015	
[11] #AC239628.1 rc	2148	10447	8299	11	0.006	0.005	0.006	0.006	0.023	0.053	0.031	0.018	0.022	0.015

Figure S1. Telomere and StSat cohybridization experiment. Results of a bicolor FISH experiment on chimpanzee (PTR) and gorilla (GGO) metaphase chromosomes are shown. The cohybridization at chromosome tips demonstrates the distal localization of the commercial telomeric PNA (Peptide Nucleic Acid) probe (green signals) to the StSat (red signals).

Figure S2. Schematic organization of subtelomeric region in IIp and IIq in chimpanzee and gorilla. Subtelomeric caps have been reported in gray; regions being part of the cap and showing multiple chromosomal signals are highlighted (red letters) compared to regions displaying single signal or not mapping to the subtelomeric regions neither in chimpanzee nor in gorilla (black letters).

Figure S3. Definition of the integration point on chromosomes VII and XIII in chimpanzee. Chromosomal ideograms are shown (top) with the corresponding sequence read-depth of highlighted regions in human, chimpanzee, gorilla, and orangutan based on human genome sequence coordinates (hg17, NCBI Build 35). Regions of excess read-depth (red) and WSSD positive regions (black horizontal bars) are indicated with respect to the end sequence placement of fosmid clones predicted to flank the integration point. Lower panels show FISH experiments using the fosmid clones above for (A) chromosomes VII and (B) chromosome XIII; arrows denote the location of interstitial heterochromatin.

Figure S4. Comparison of copy number versus fluorescence in human, chimpanzee, and gorilla. Linear distributions for copy number counts and fluorescence intensity in human, chimpanzee, and gorilla with (A) and without (B) hyperexpanded regions are shown. The results suggest that copy number may be accurately estimated by sequence read-depth analyses even for repeats of high copy number.

Figure S5. Nonhuman primate BAC clone duplication analysis. Sequenced (A) chimpanzee (PTR) and (B) gorilla (GGO) BAC clone inserts containing StSat and mapping to chromosome VII were analyzed for segmental duplications. Segmental

duplications detected by WSSD are annotated in each species (black bars above line), StSat (blue), and human segmental duplications based on BLAST sequence similarity searches (whole-genome assembly comparison method) (lower colored bars) and classified as to whether they were shared, chimpanzee (PTR) or gorilla specific (GGO).

Figure S6. WSSD and BLAST analysis of chimpanzee BAC clones containing regions homologous to human chromosomes 7 and 10. Segmental duplications detected by WSSD are annotated in each species (black bars above line), StSat location in the clones is shown (blue), and human segmental duplications based on BLAST sequence similarity searches (whole-genome assembly comparison method) are reported (lower colored bars). Predicted copy number is reported in the graph by red lines and highlights the transition point from euchromatin to the interstitial heterochromatin in the chimpanzee genome. Human chromosomes 7 and 10 correspond to phylogenetic group VII and X.

Table S1. Mapping of probes on chromosomes 2 and 10. FISH results obtained using human fosmid on chimpanzee (PTR) and gorilla (GGO) chromosomes.

Table S2. Definition of interstitial heterochromatic integration point in chimpanzee. Clone mapping on chimpanzee of human clones flanking interstitial heterochromatic blocks on chromosomes VII and XIII.

Table S3. Copy number estimates of gorilla and chimpanzee hyperexpanded loci and proximal regions. Quantitative analysis by fluorescence estimates for clones mapping on hyperexpanded and proximal regions on human chromosomes 2 and 10.

Table S4. BLAST analysis using pCht sequences on chimpanzee and gorilla BAC clones. Sequence analysis on chimpanzee and gorilla and exact localization of pCht (X74281.1) in PTR and GGO BAC clones has been reported.

Table S5. Pairwise distance calculation. Pairwise genetic distance matrix of the chr2_113996786_114006404, chr2_114016544_114026608, chr10_19564292_19576051 and chr10_19541142_19549347 regions using human, chimpanzee, gorilla and orangutan sequences. Human genome reference: hg17, May 2004, NCBI Build 35.