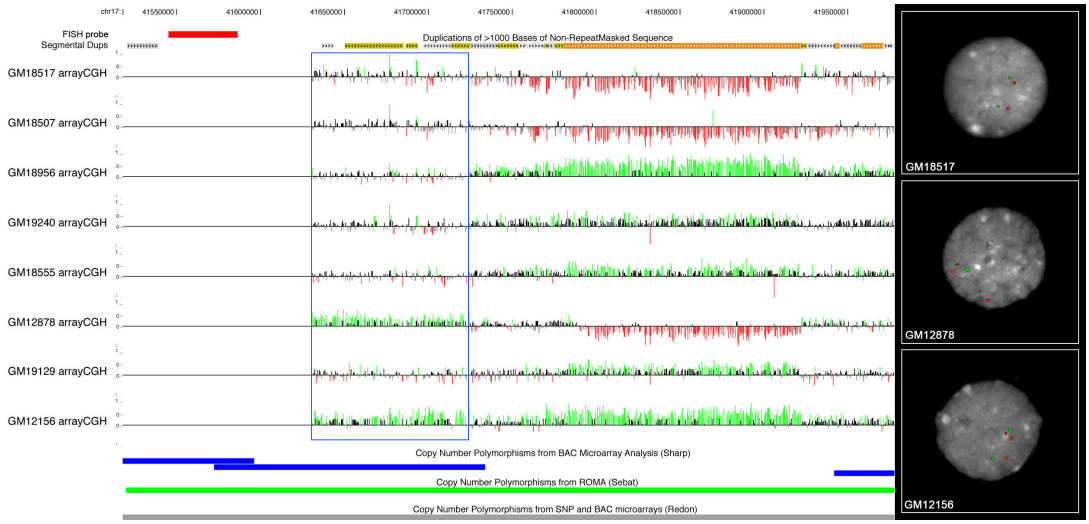
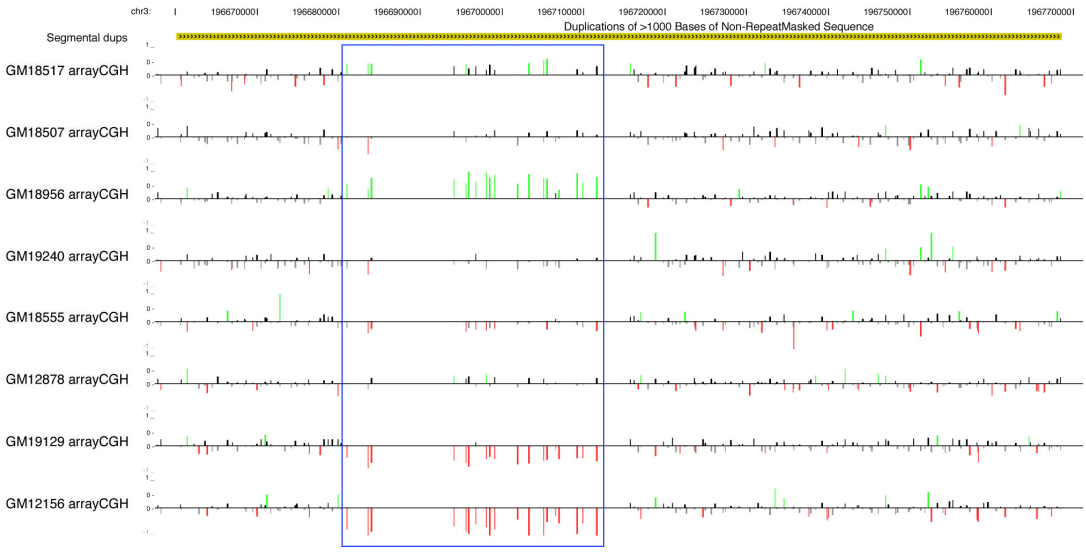
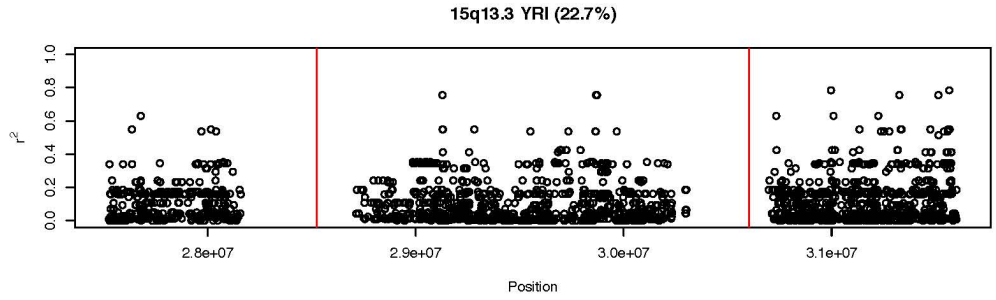
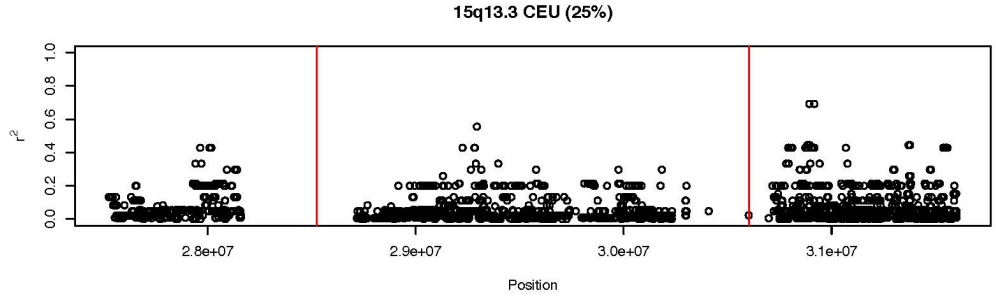


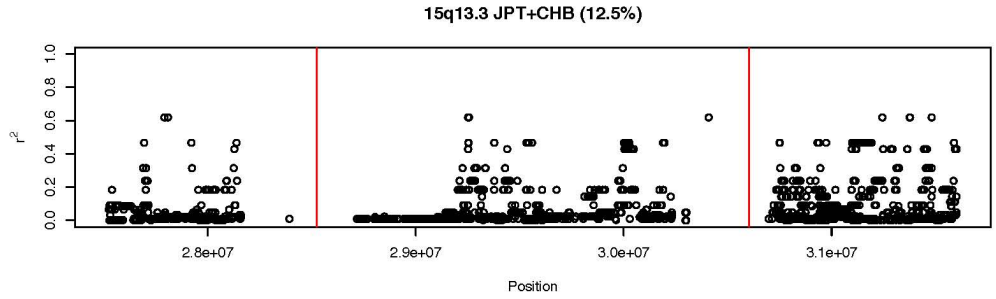
Supplementary Figure 1: Oligonucleotide arrayCGH results for the 17q21.31 inversion distal breakpoint region. Individuals GM12156 (heterozygous H1/H2) and GM12878 (homozygous H1/H1) show a 75-kbp extended duplication (in the blue rectangle) compared to the other individuals. On the right, interphase FISH validation using probe G248P85784G12 (red) mapping in the polymorphic region (chr17:41545746-41586094) and probe G248P8967C6 (green) mapping in the single copy region (chr17:39477456-39515250).

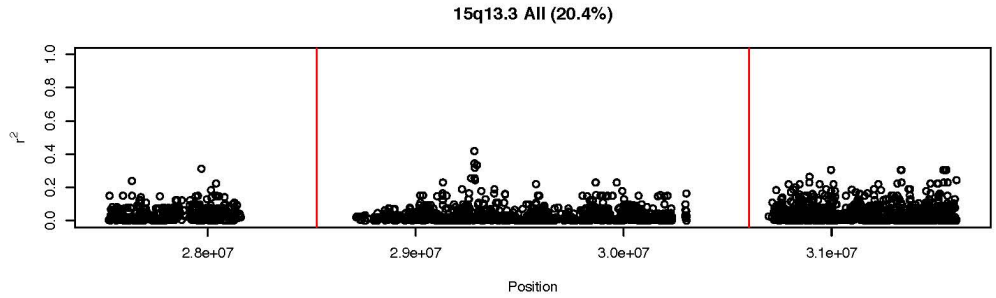


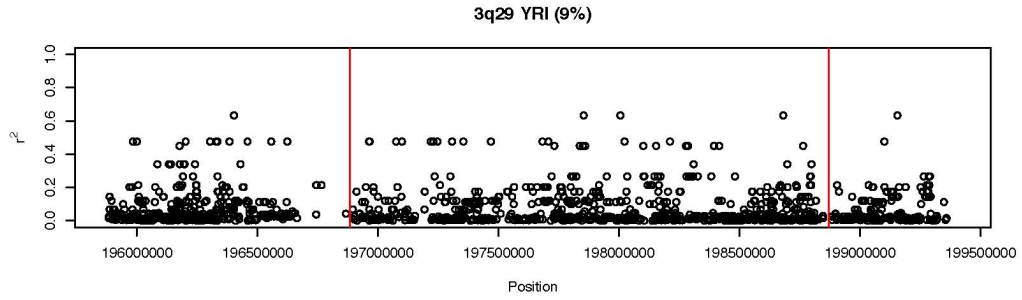
Supplementary Figure 2: Oligonucleotide arrayCGH results show that individuals GM19129 and GM12156, both carriers for the inverted allele, show a 25-kbp copy-number loss within the duplicated sequence at proximal breakpoint of the 3q29 inversion.

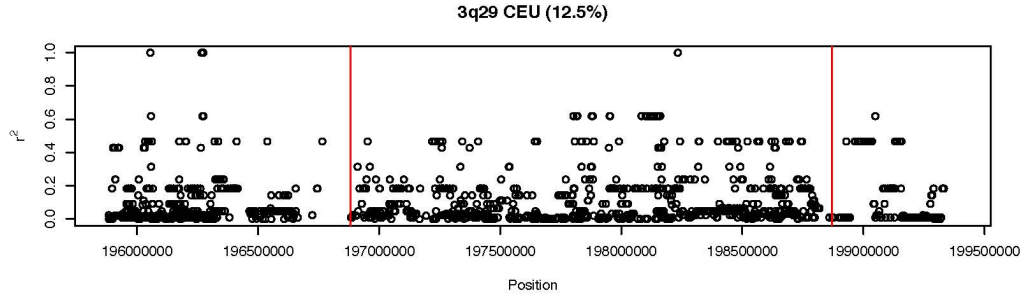

15q13.3 YRI (22.7%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

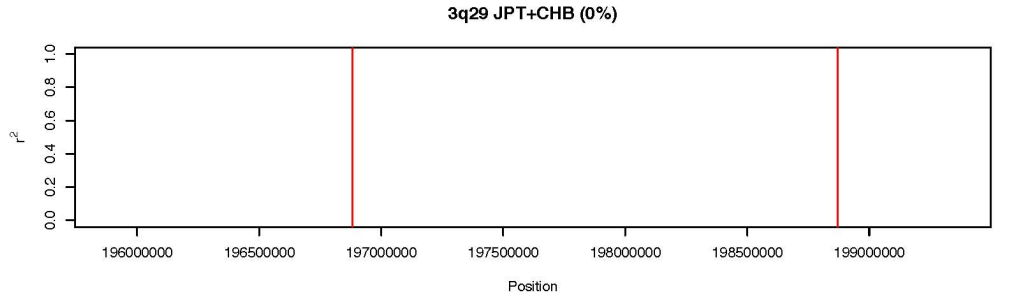

15q13.3 CEU (25%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

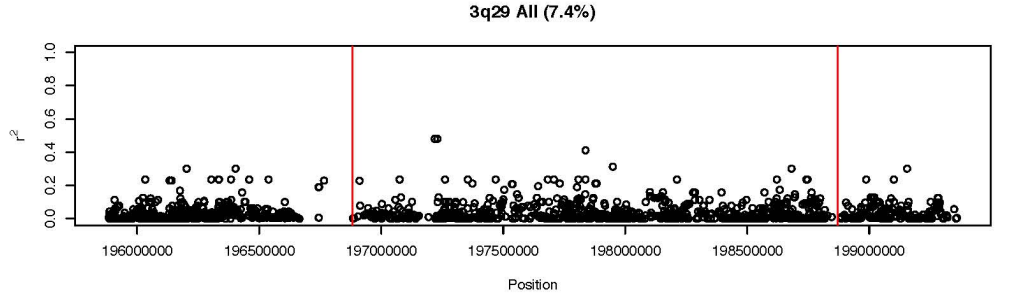

15q13.3 JPT+CHB (12.5%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

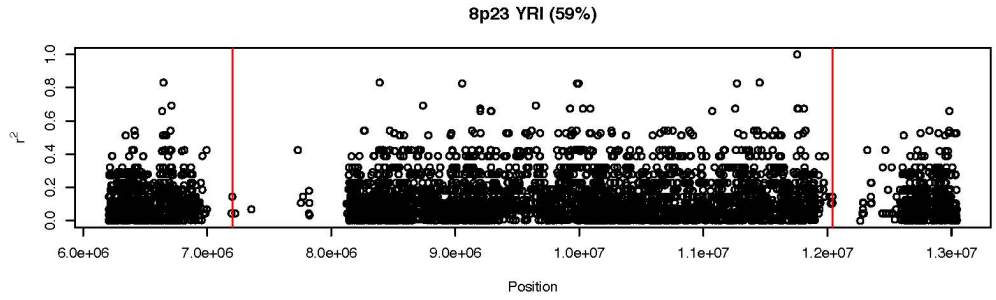

15q13.3 All (20.4%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

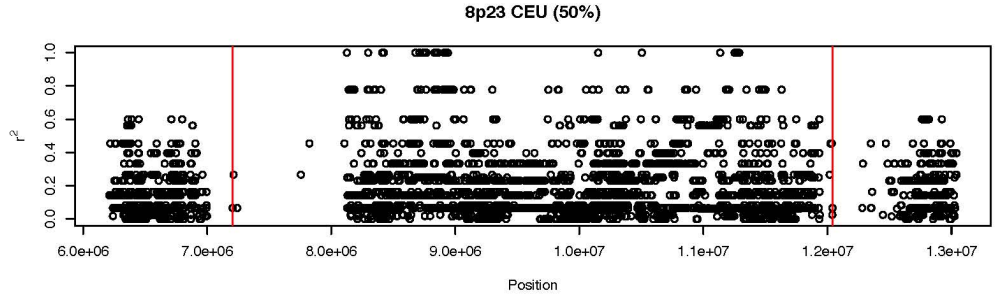

3q29 YRI (9%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
199500000
Position

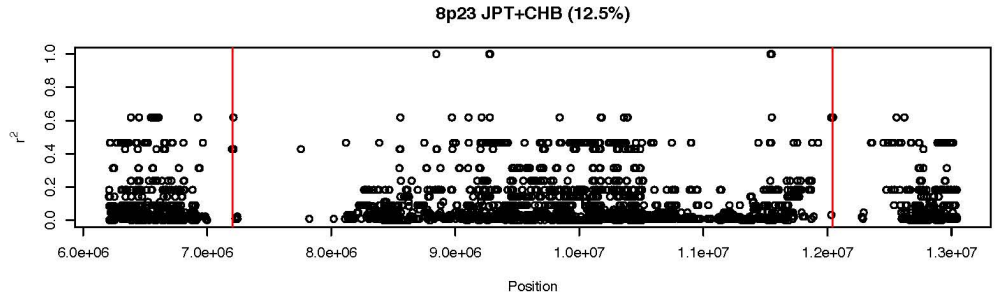

3q29 CEU (12.5%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
199500000
Position

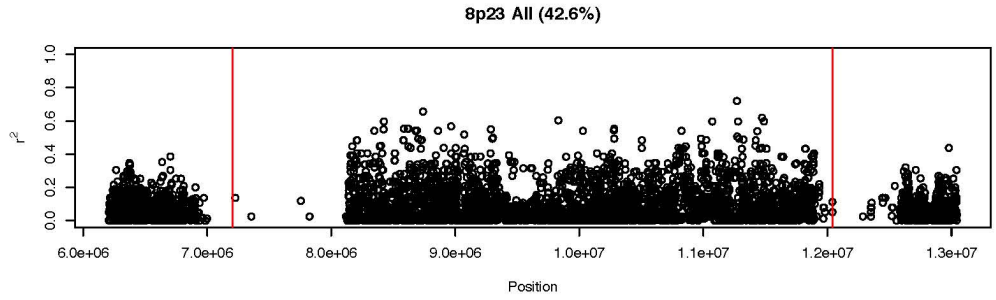

3q29 JPT+CHB (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
Position

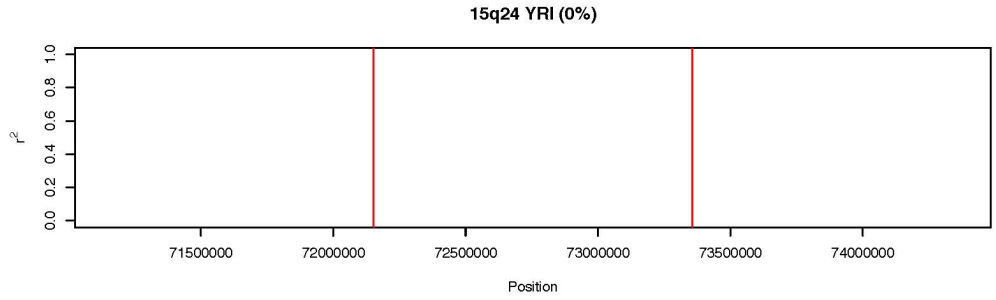

3q29 All (7.4%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
Position

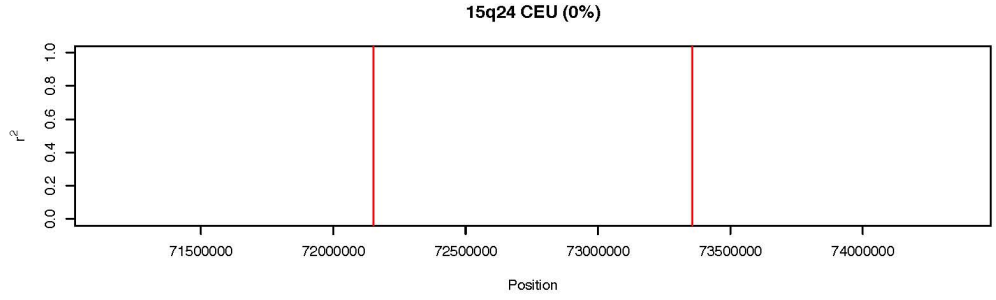

8p23 YRI (59%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

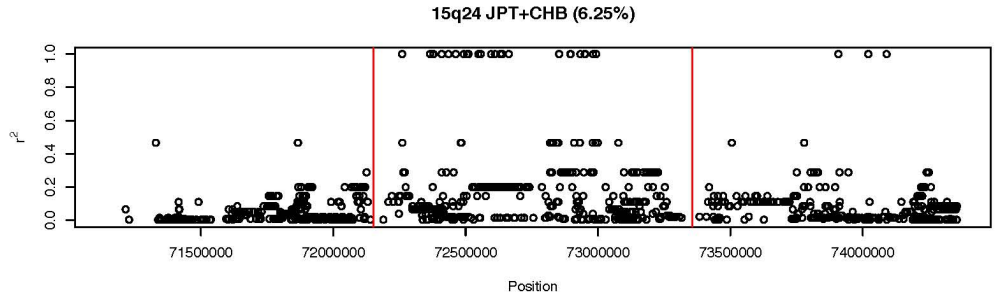

8p23 CEU (50%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

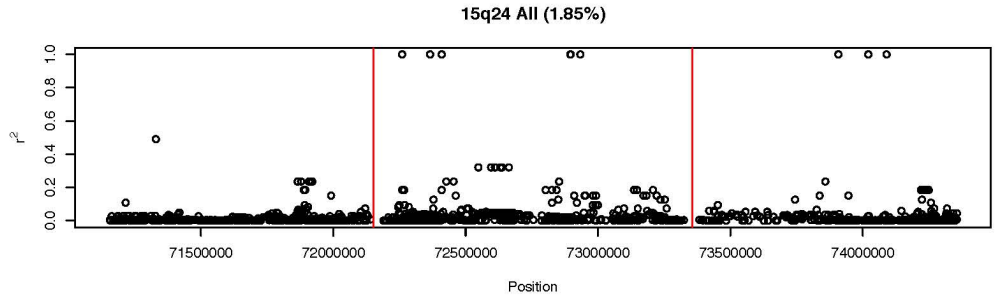

8p23 JPT+CHB (12.5%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

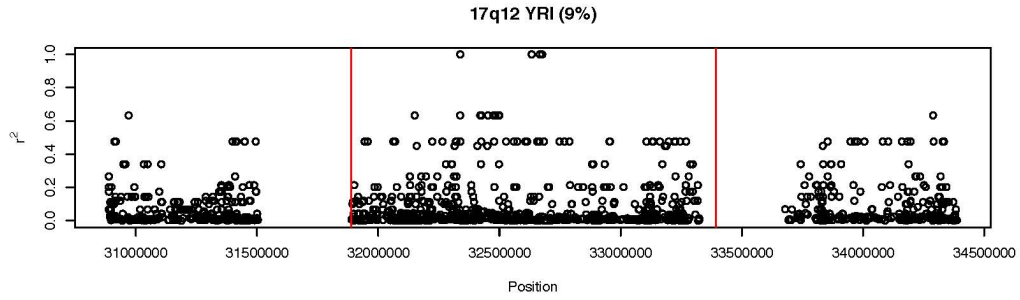

8p23 All (42.6%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

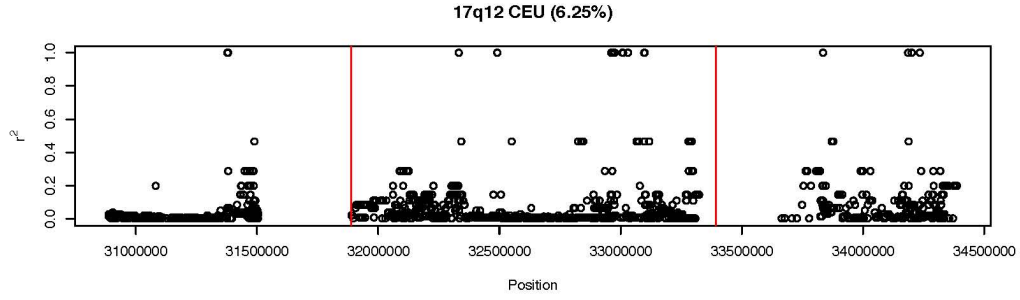

15q24 YRI (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

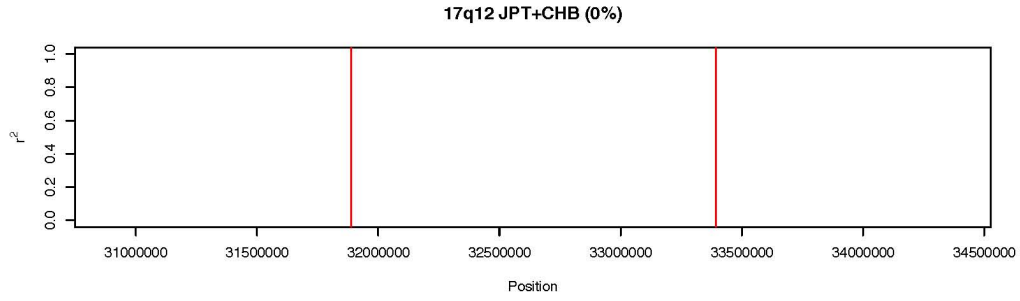

15q24 CEU (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

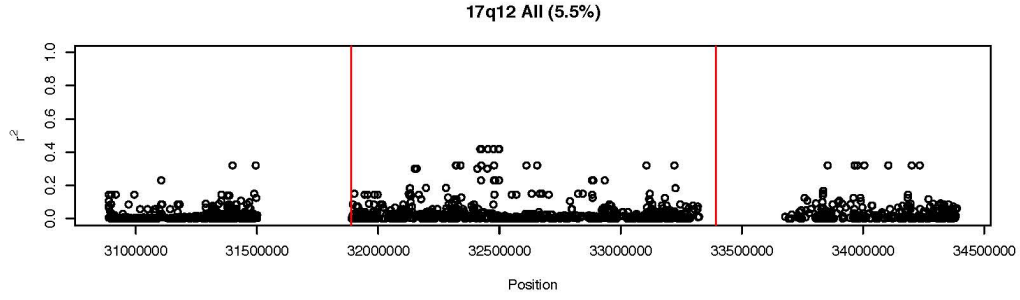

15q24 JPT+CHB (6.25%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

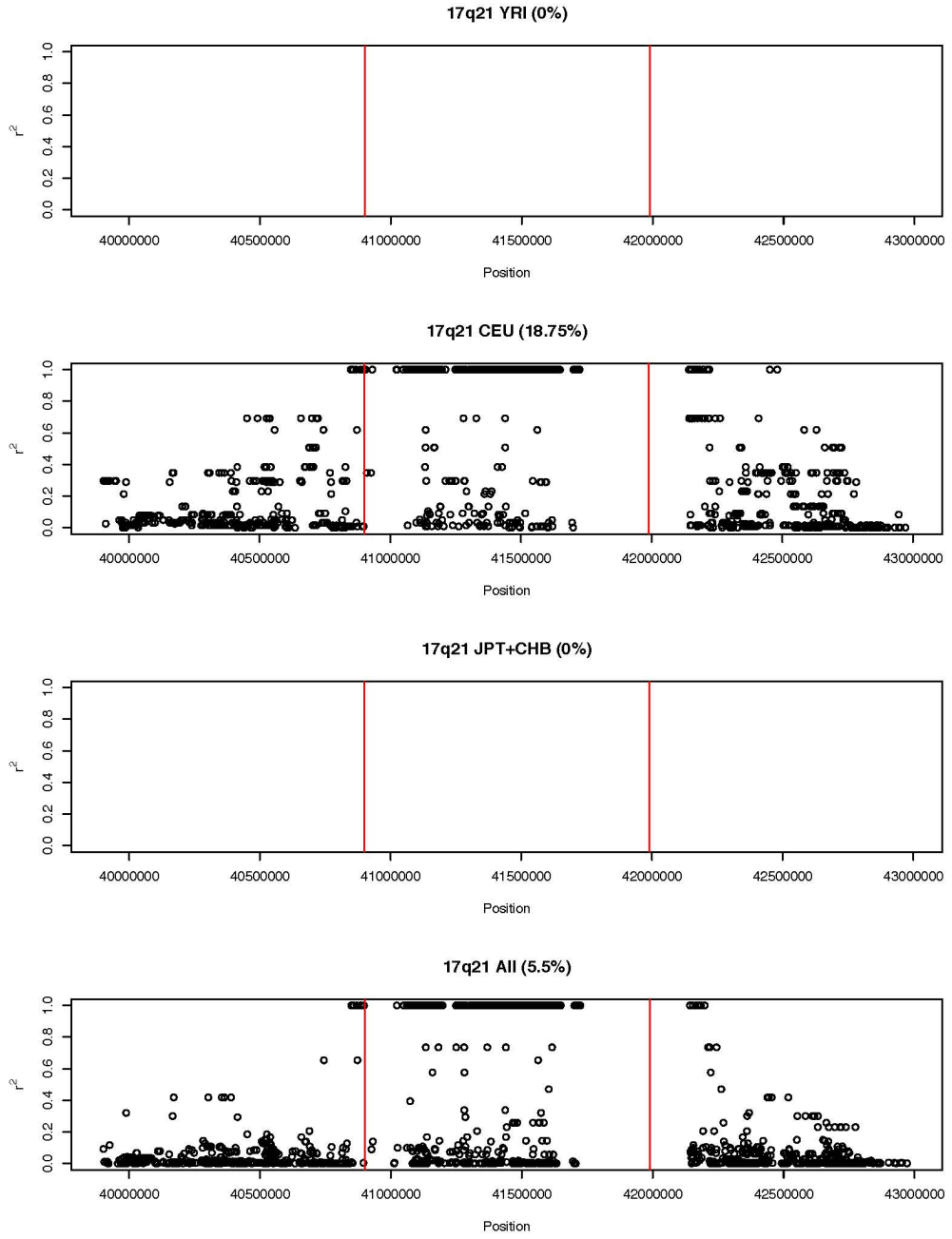

15q24 All (1.85%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position


17q12 YRI (9%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
31000000
31500000
32000000
32500000
33000000
33500000
34000000
34500000
Position

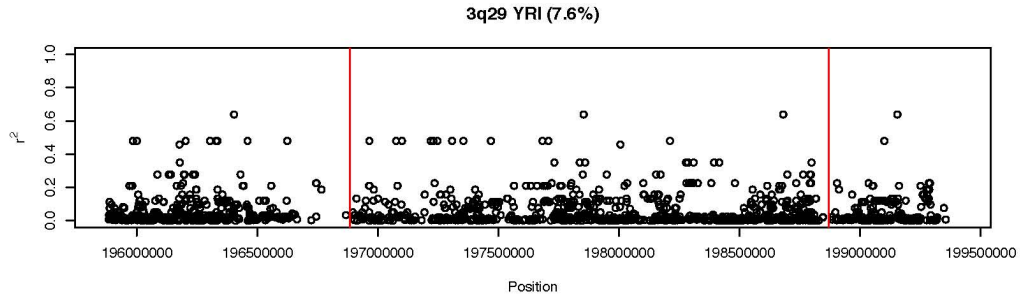

17q12 CEU (6.25%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
31000000
31500000
32000000
32500000
33000000
33500000
34000000
34500000
Position

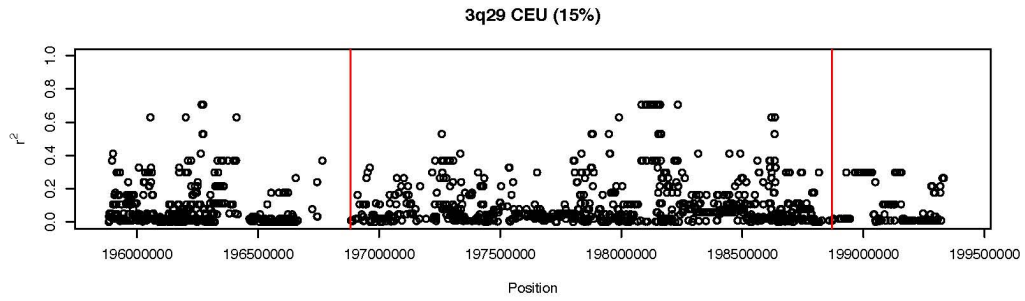

17q12 JPT+CHB (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
31000000
31500000
32000000
32500000
33000000
33500000
34000000
34500000
Position

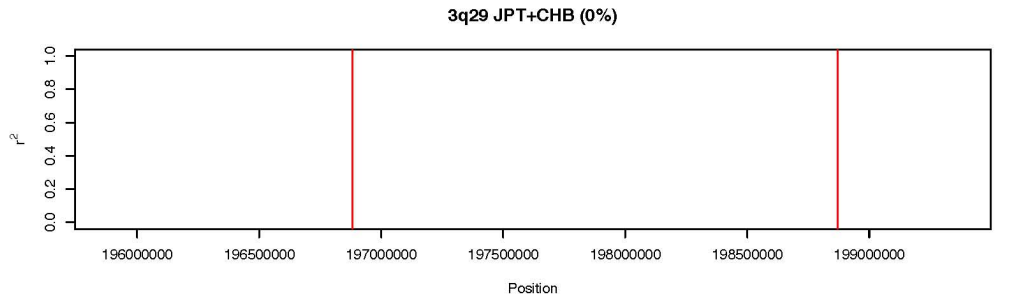

17q12 All (5.5%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
31000000
31500000
32000000
32500000
33000000
33500000
34000000
34500000
Position

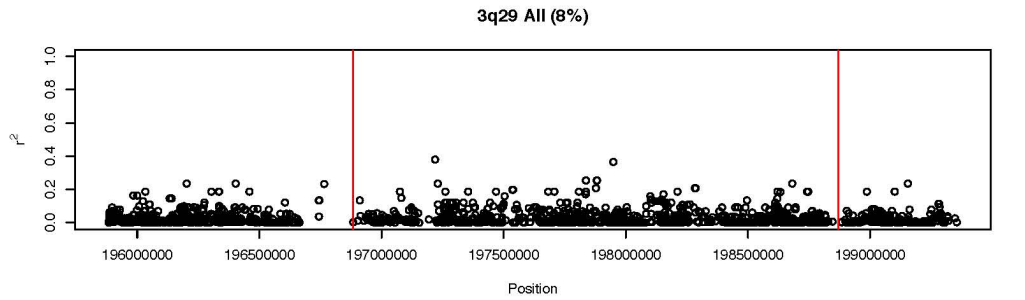


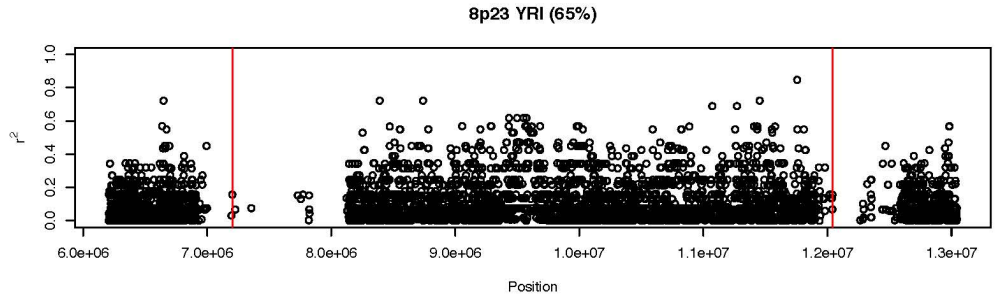
Supplementary Figure 3: Pair-wise r^2 between HapMap SNP and inversion genotypes. The approximate inversion breakpoints are denoted by the red vertical lines. Data are presented from the three HapMap continental groupings individually, as well as for all 27 samples pooled. The frequency of the inversion allele in each sample set is given in parenthesis. For analysis, double heterozygotes were assumed to have minor alleles in a cis configuration (see Materials and Methods).

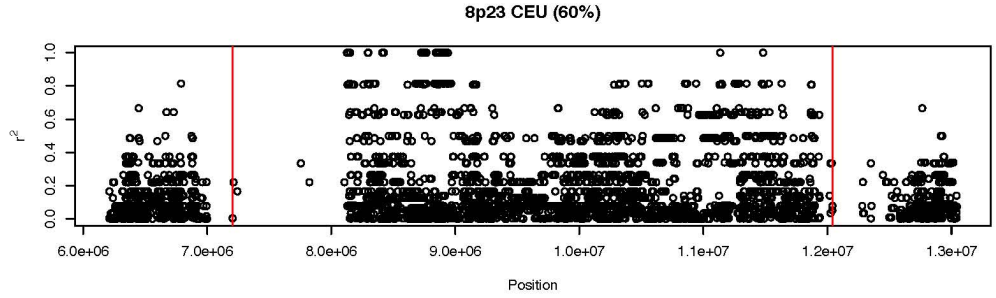

3q29 YRI (7.6%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
199500000
Position

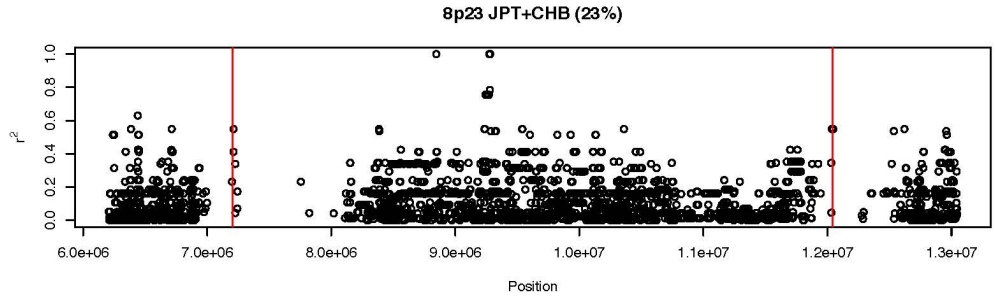

3q29 CEU (15%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
199500000
Position

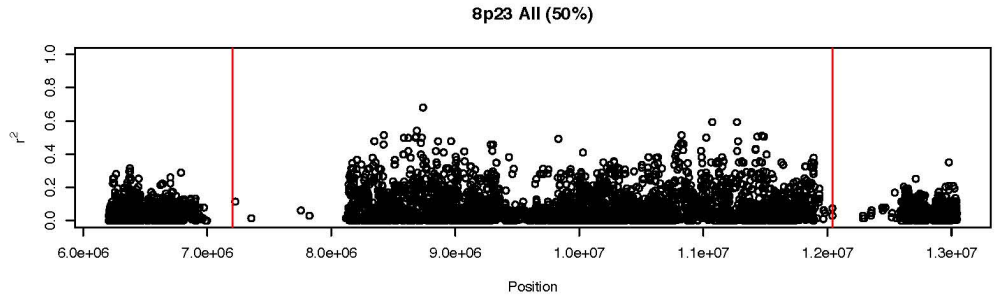

3q29 JPT+CHB (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
Position

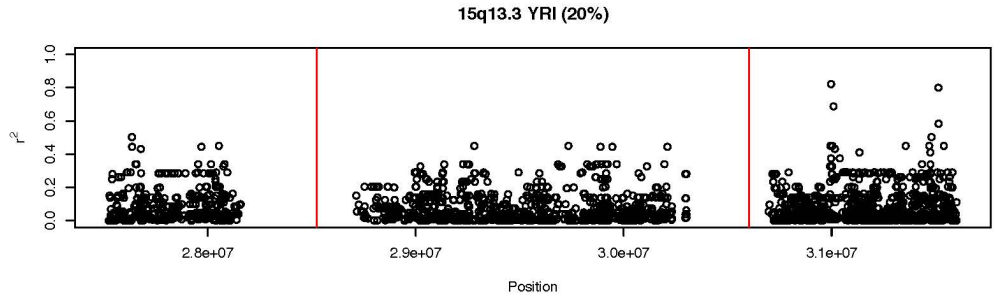

3q29 All (8%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
196000000
196500000
197000000
197500000
198000000
198500000
199000000
Position

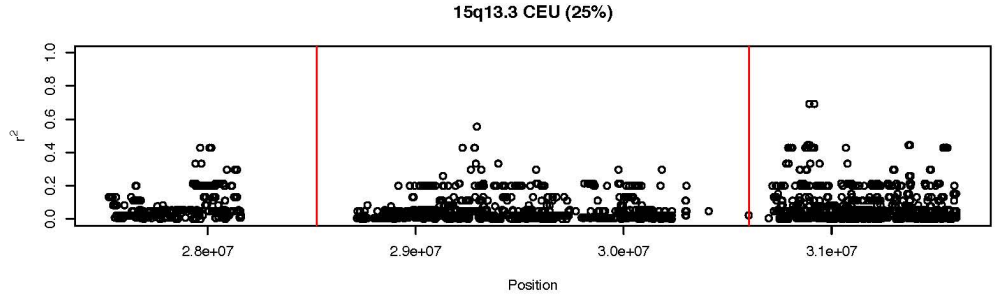

8p23 YRI (65%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

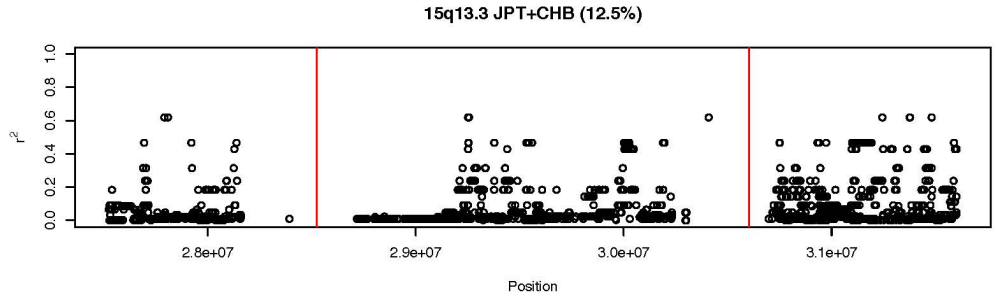

8p23 CEU (60%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

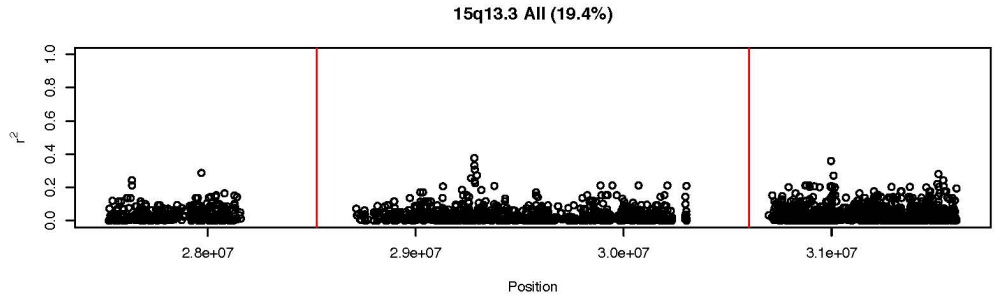

8p23 JPT+CHB (23%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

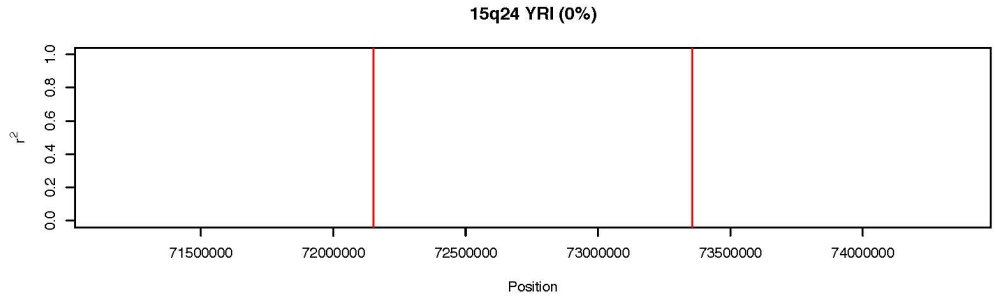

8p23 All (50%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
6.0e+06
7.0e+06
8.0e+06
9.0e+06
1.0e+07
1.1e+07
1.2e+07
1.3e+07
Position

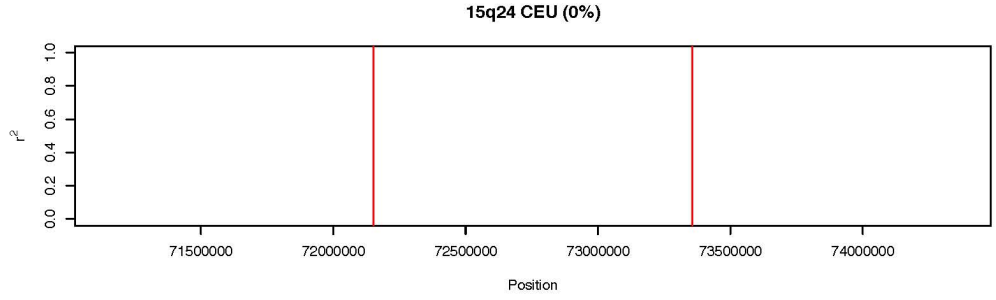

15q13.3 YRI (20%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

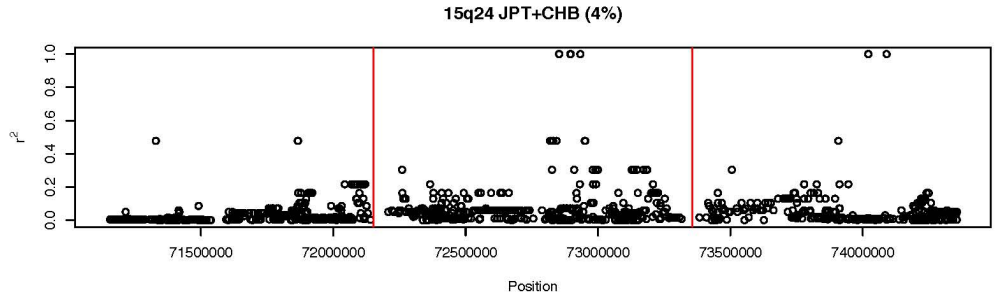

15q13.3 CEU (25%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

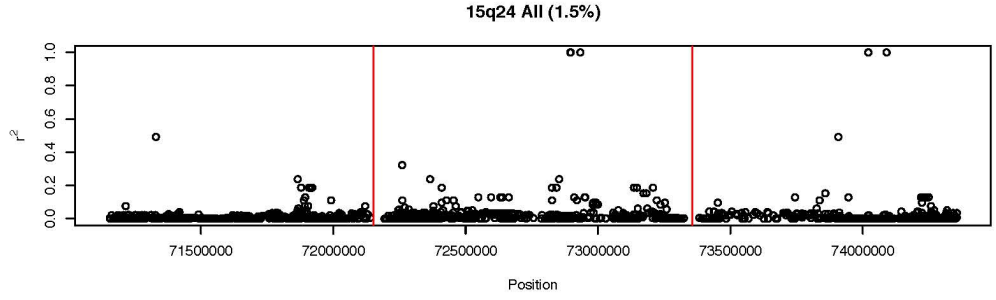

15q13.3 JPT+CHB (12.5%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position

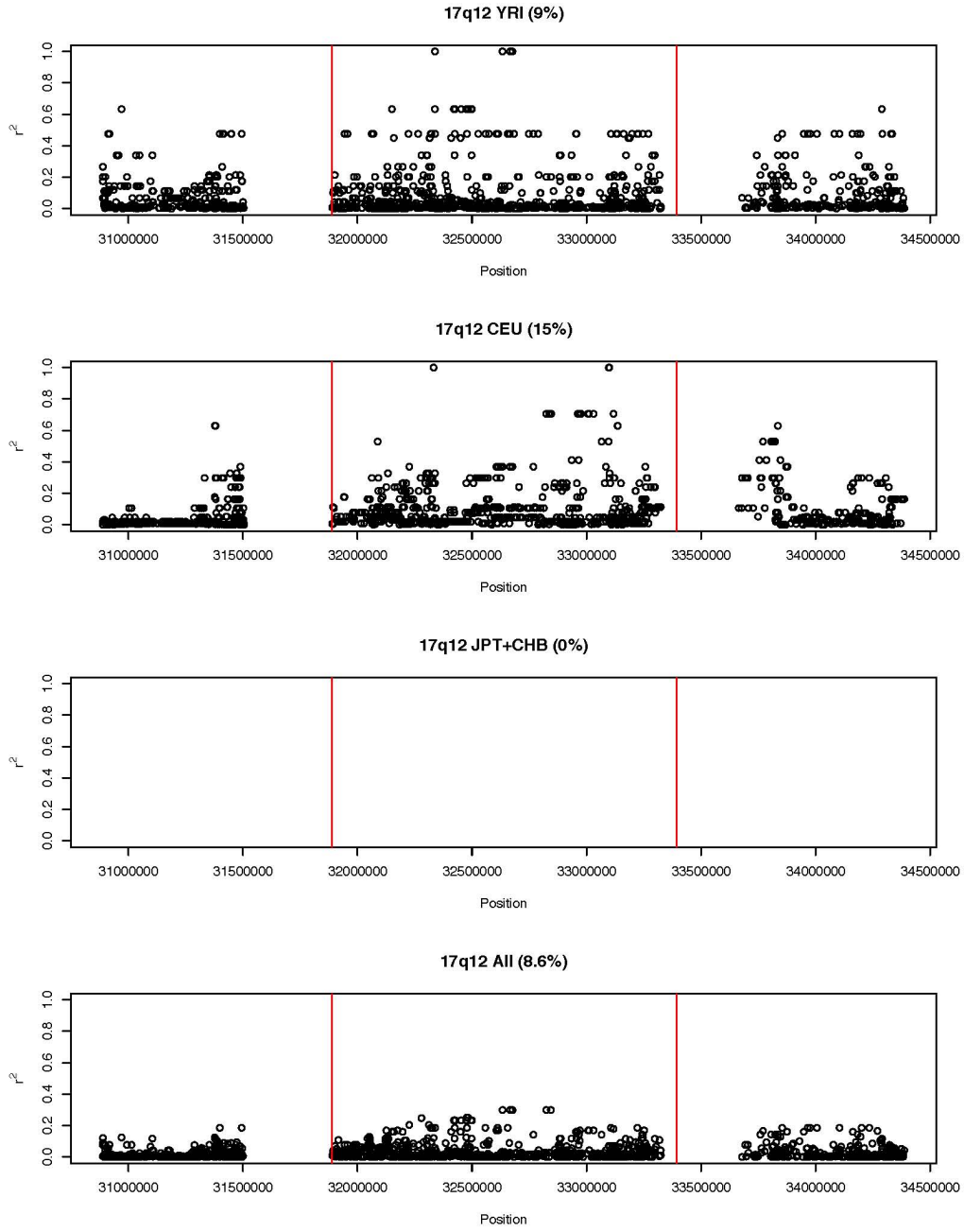

15q13.3 All (19.4%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
2.8e+07
2.9e+07
3.0e+07
3.1e+07
Position


15q24 YRI (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

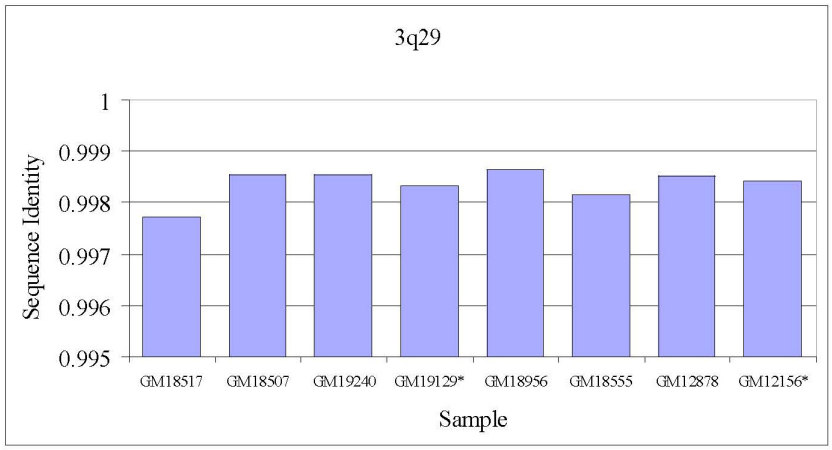

15q24 CEU (0%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

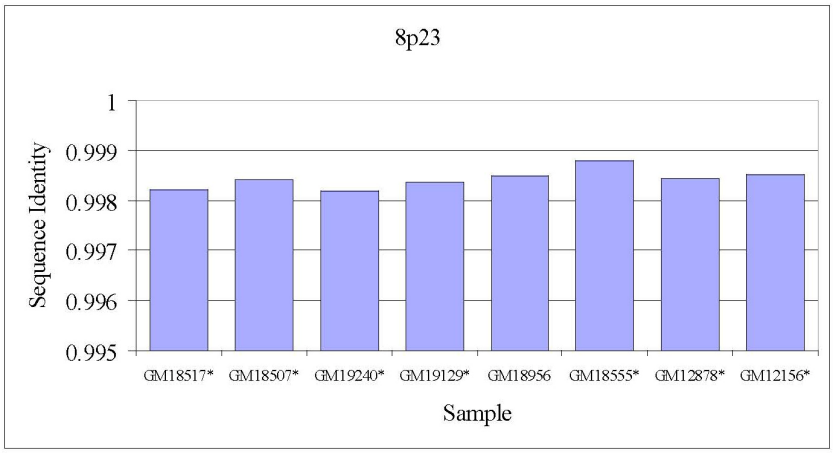

15q24 JPT+CHB (4%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

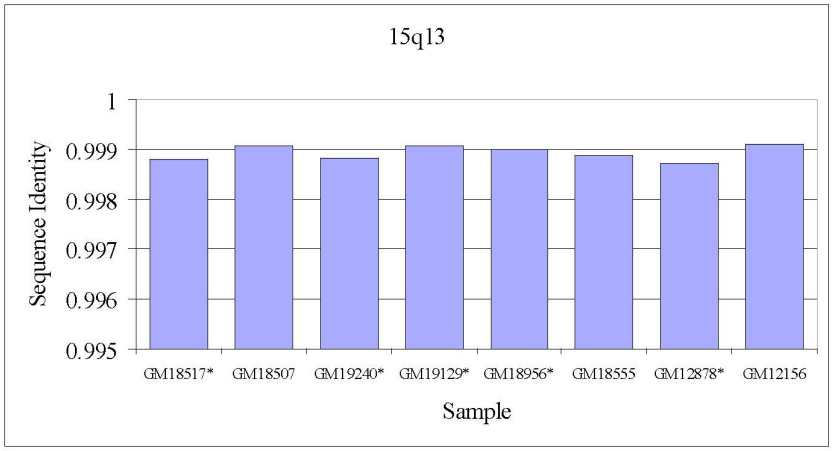

15q24 All (1.5%)
r2
0.0
0.2
0.4
0.6
0.8
1.0
71500000
72000000
72500000
73000000
73500000
74000000
Position

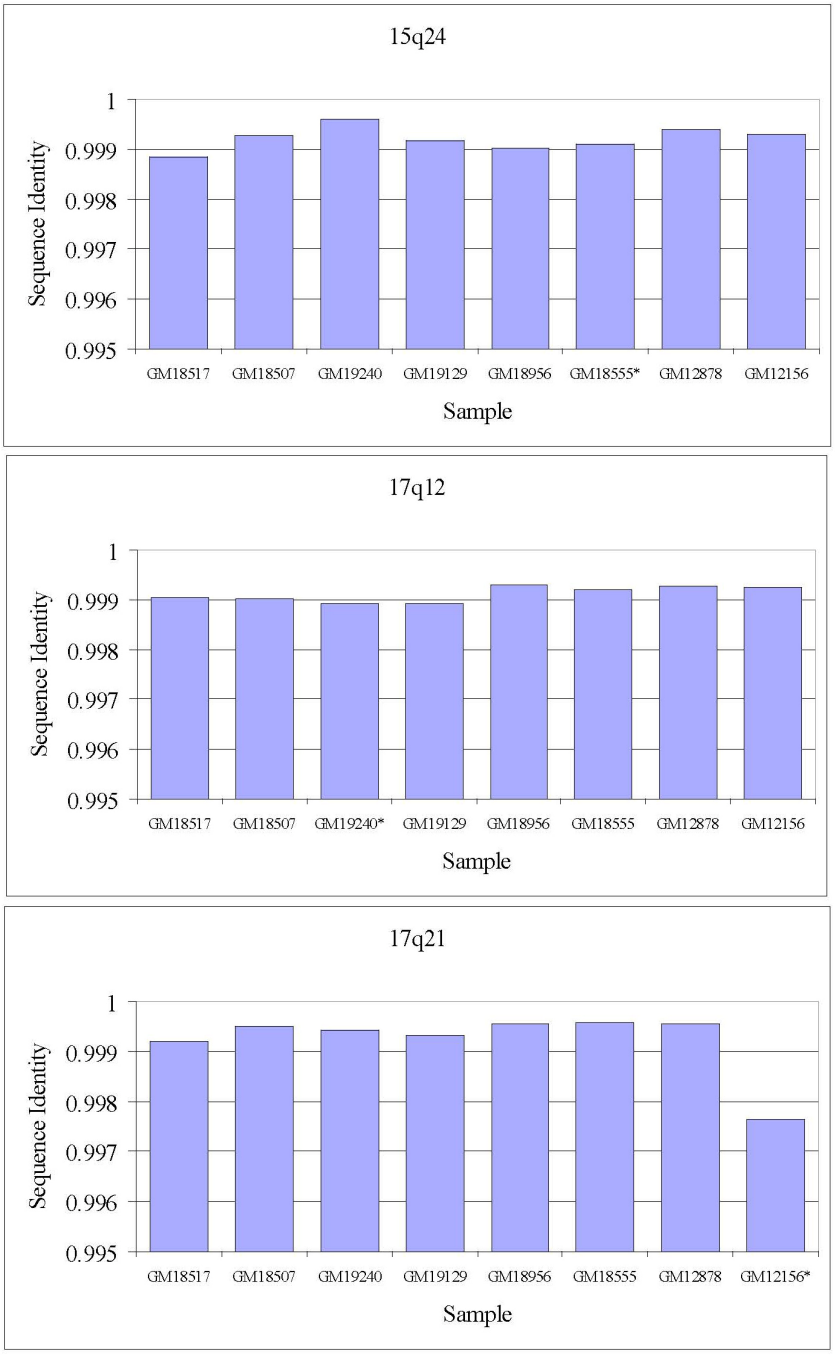


Supplementary Figure 4: R^2 values as in Figure S1, but including genotype results from the samples predicted to carry inversions (see Table 3).


3q29
Sequence Identity
1
0.999
0.998
0.997
0.996
0.995
GM18517
GM18507
GM19240
GM19129*
GM18956
GM18555
GM12878
GM12156*
Sample

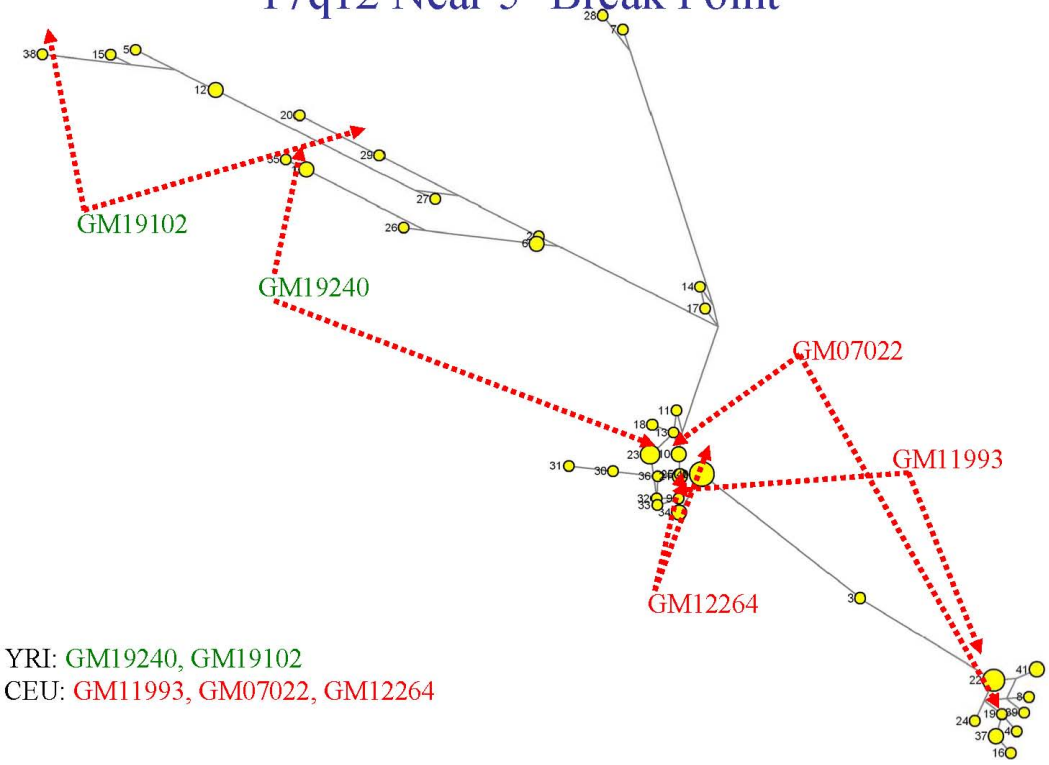

8p23
Sequence Identity
1
0.999
0.998
0.997
0.996
0.995
GM18517*
GM18507*
GM19240*
GM19129*
GM18956
GM18555*
GM12878*
GM12156*
Sample


15q13
Sequence Identity
1
0.999
0.998
0.997
0.996
0.995
GM18517*
GM18507
GM19240*
GM19129*
GM18956*
GM18555
GM12878*
GM12156
Sample



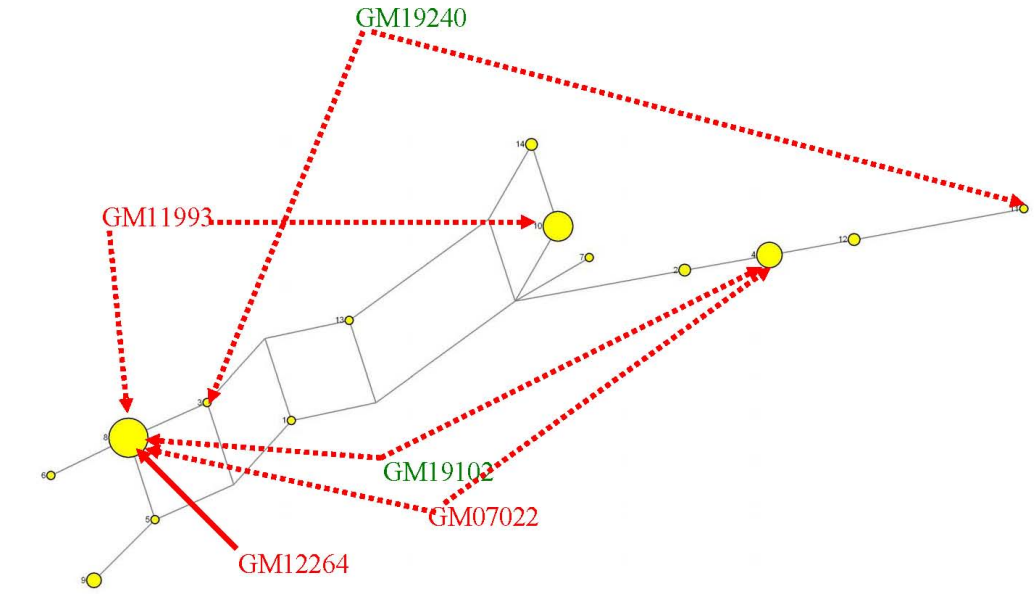
Supplementary Figure 5: Sequence divergence of inversion intervals. The sequence identity of eight HapMap individuals relative to the reference genome assembly (build 35) is shown for the unique portion of each inversion interval. Sequence identity was calculated based on all uniquely mapping fosmid end sequences from each sample. Only mapped positions having a sequence quality of at least Q20 were considered. Individual samples carrying an inversion are indicated by an asterisk.

17q12 Near 5' Break Point

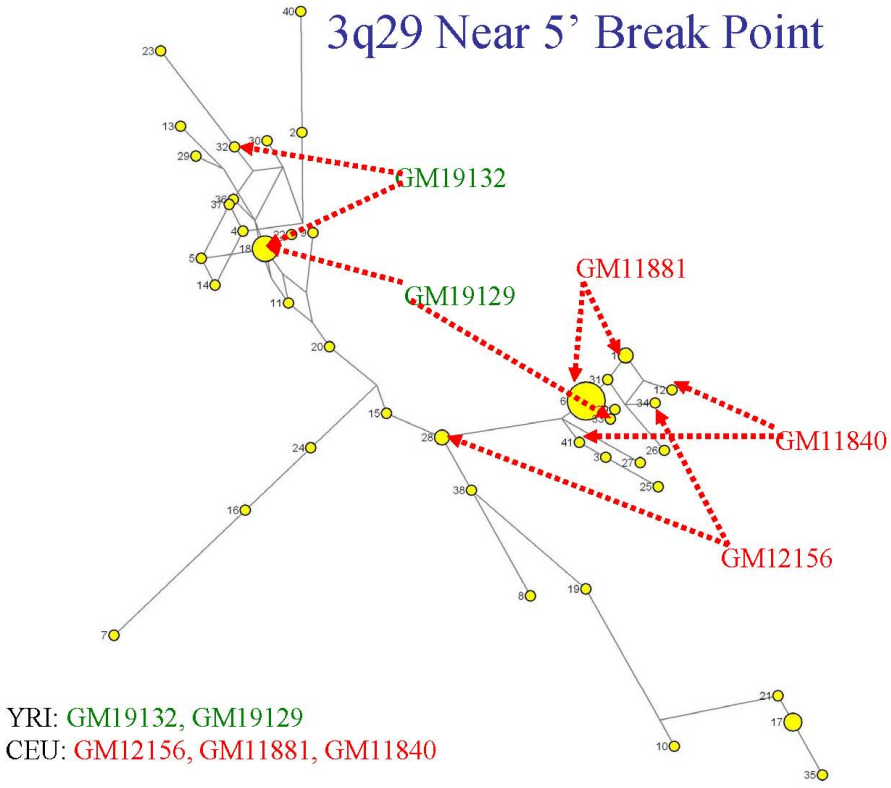


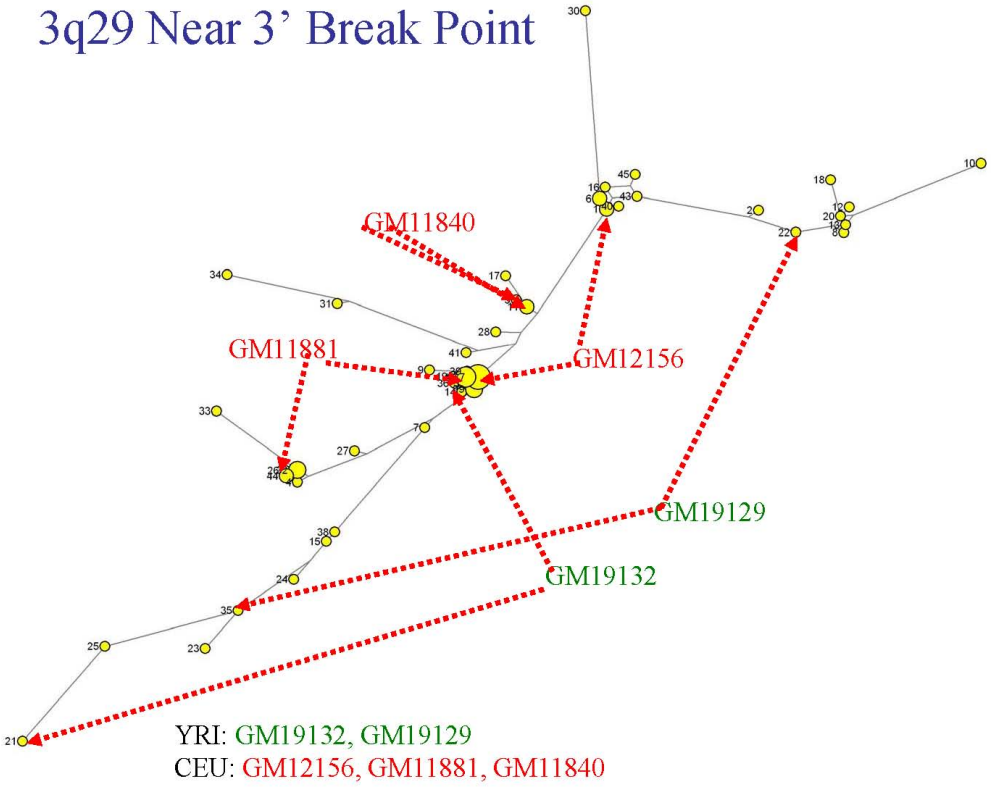
YRI: GM19240, GM19102
CEU: GM11993, GM07022, GM12264

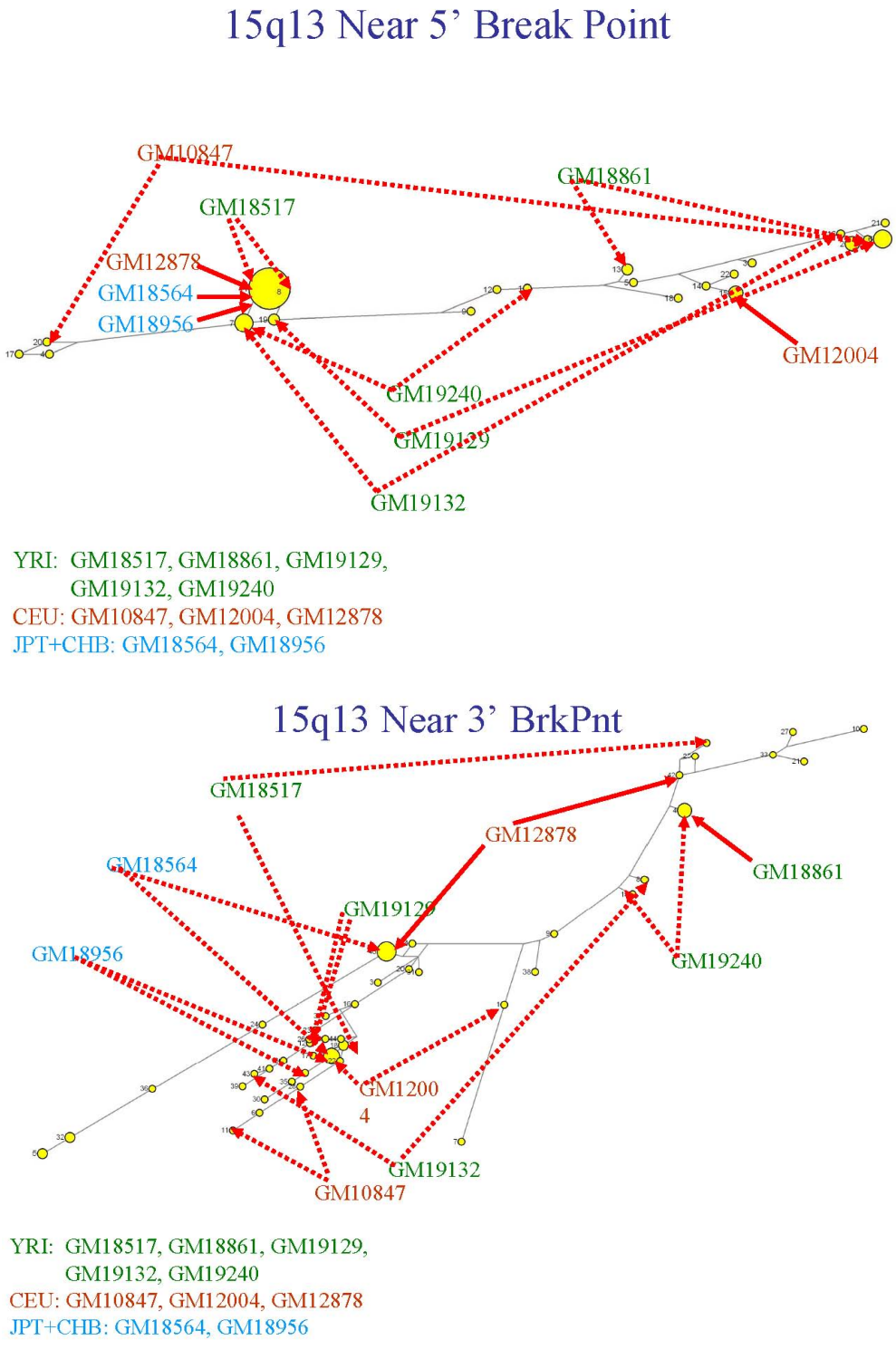
17q12 Near 3' Break Point



YRI: GM19240, GM19102
CEU: GM11993, GM07022, GM12264

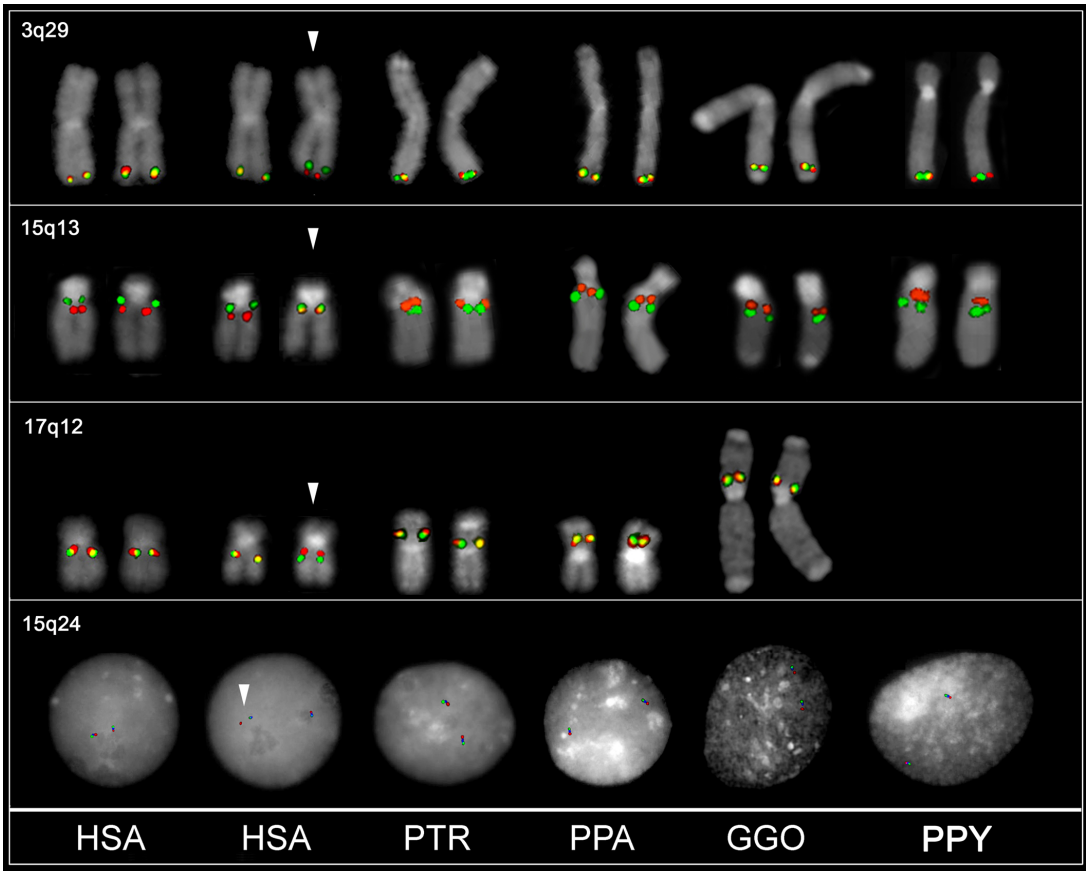

3q29 Near 5' Break Point
GM19132
GM19129
GM11881
GM11840
GM12156
YRI: GM19132, GM19129
CEU: GM12156, GM11881, GM11840


3q29 Near 3' Break Point
GM11840
GM11881
GM12156
GM19129
GM19132
YRI: GM19132, GM19129
CEU: GM12156, GM11881, GM11840

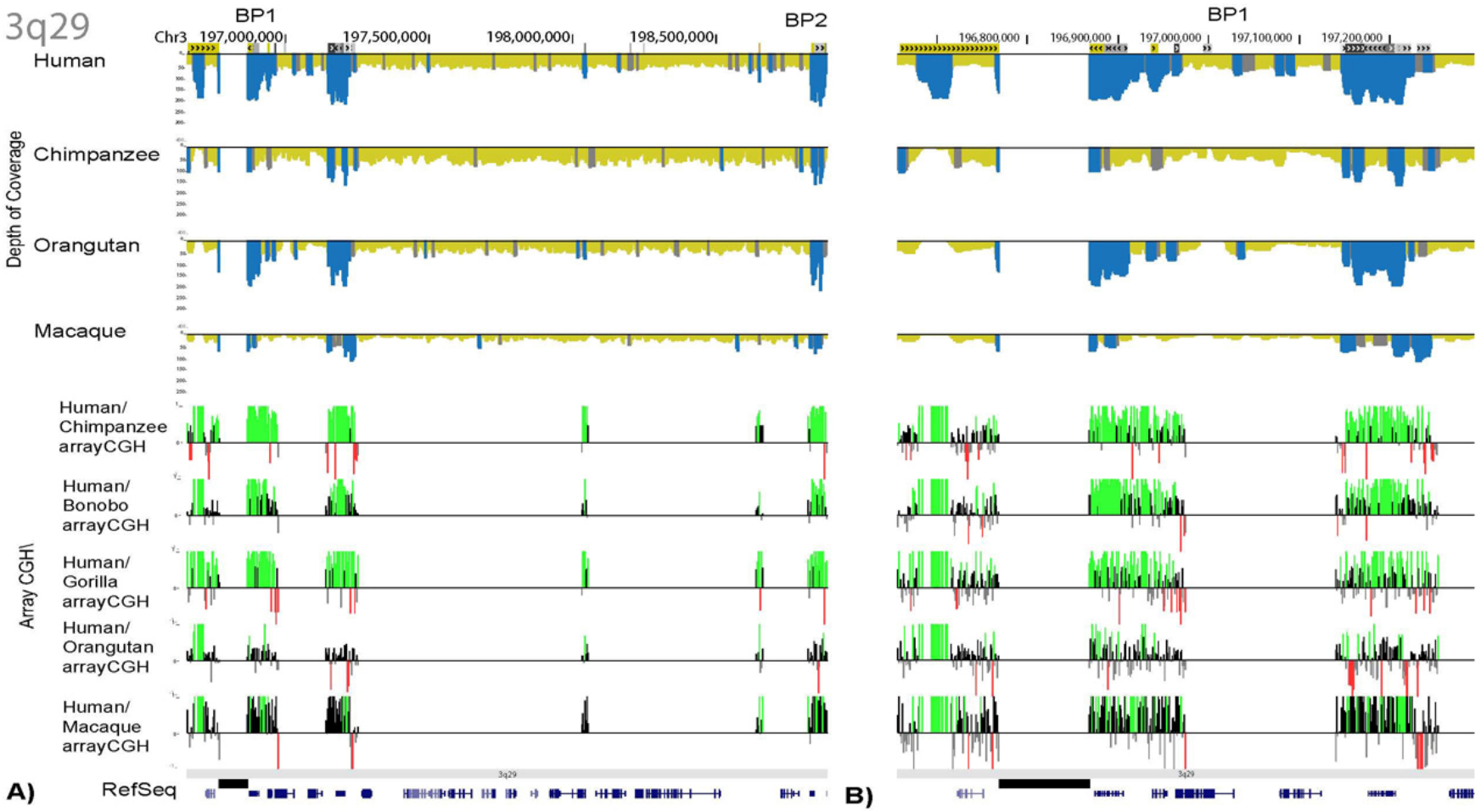


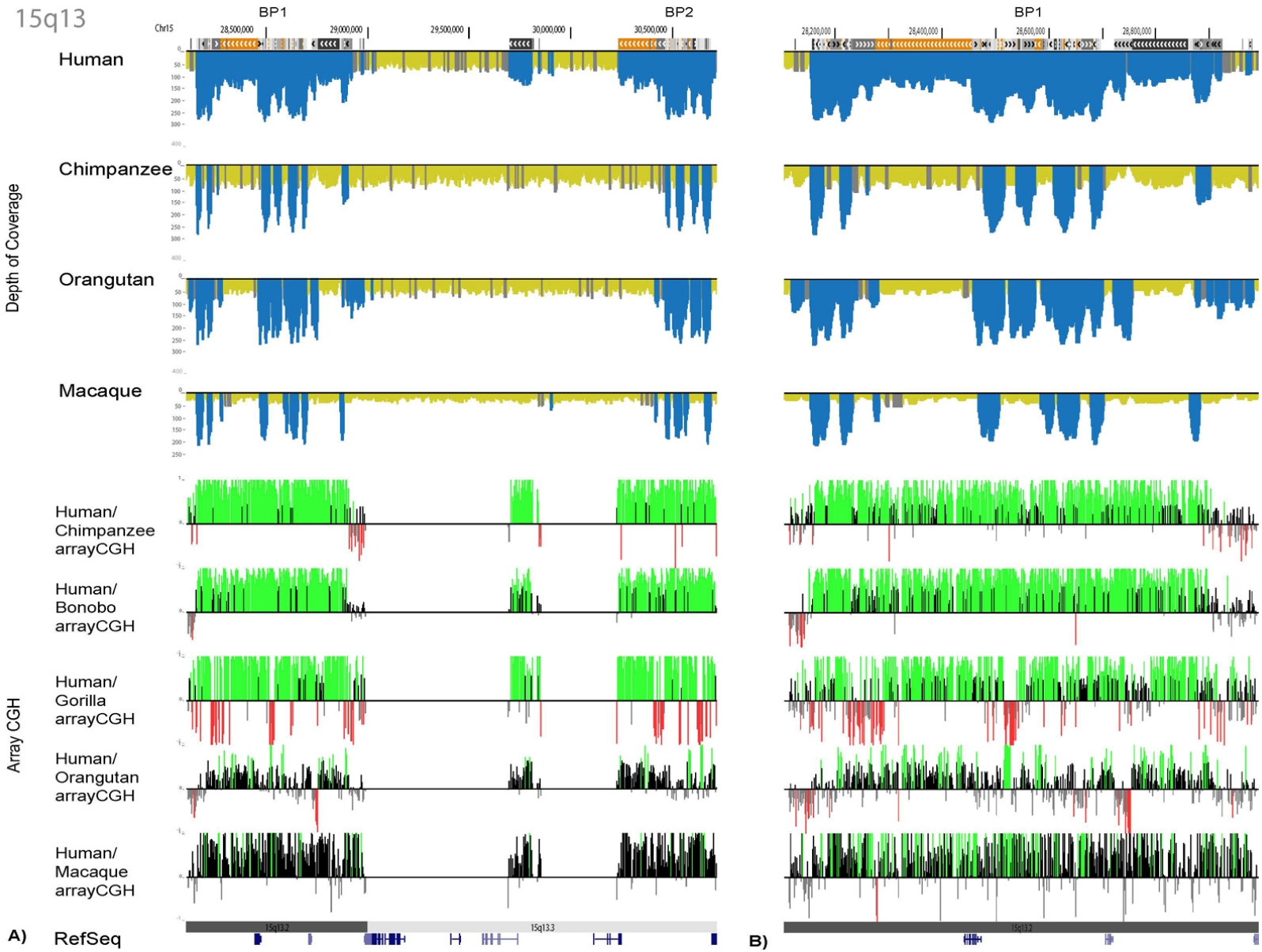
Supplementary Figure 6: Median-joining haplotype networks constructed from HapMap phased haplotypes. Networks were constructed separately for each breakpoint for the 3q29, 17q12 and 15q13.3 inversions (Table S6). Haplotype groups are denoted by yellow circles. The size of the circles is proportional to the number of chromosomes carrying the represented haplotype. The

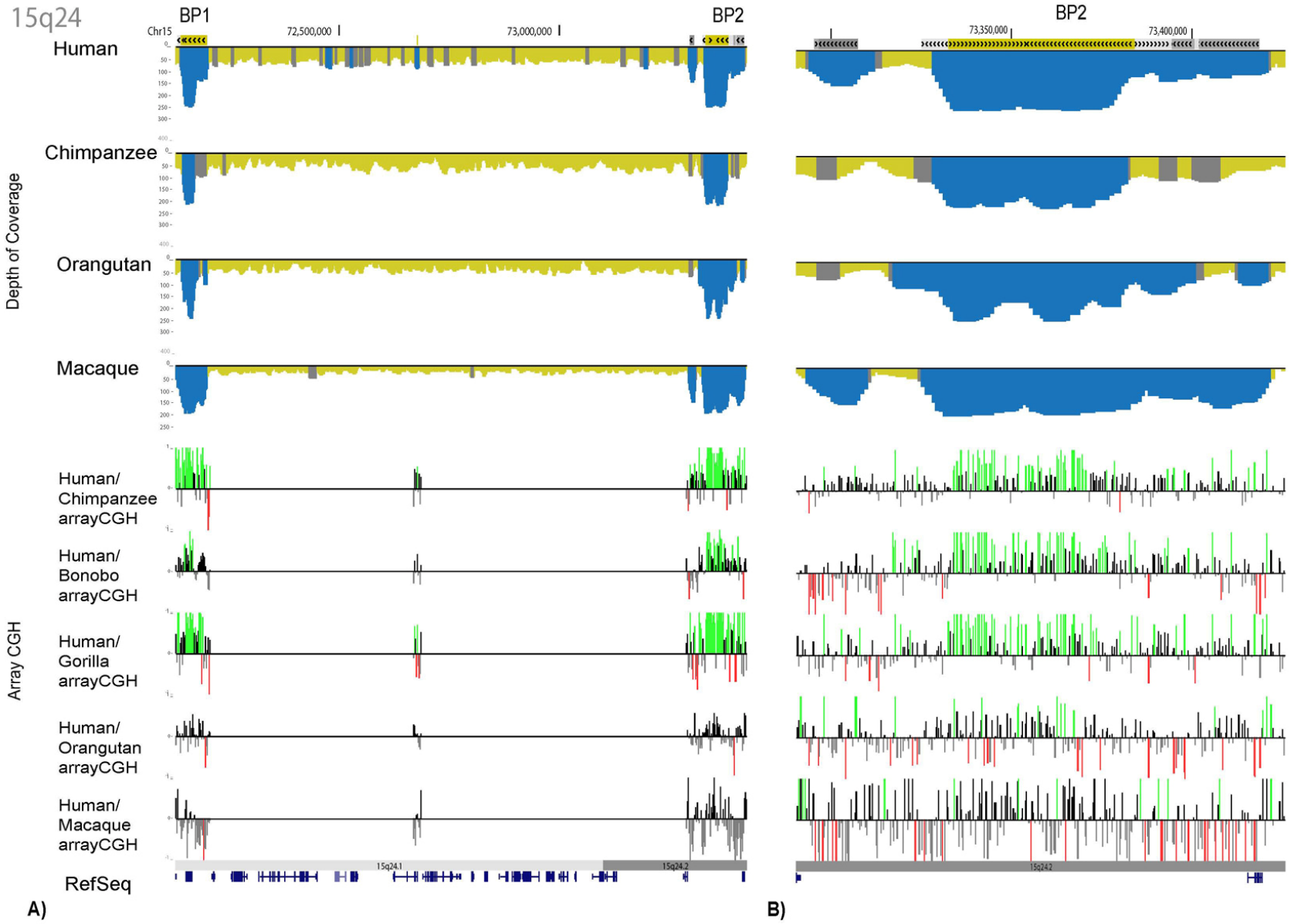
distance between haplotype groups is proportional to the number of SNP mutational events separating each haplotype. The haplotypes carried by individuals known to also carry an inversion allele are indicated by dashed and solid lines. For heterozygotes (dashed lines), only one of the two indicated haplotypes corresponds to the inversion. For homozygotes (solid lines), all indicated haplotypes must be present in the inverted orientation. Configurations where it is not possible to define closely related haplotype groups, which account for all inversion chromosomes, suggest recurrent mutation or substantial exchange between inverted and non-inverted haplotypes.

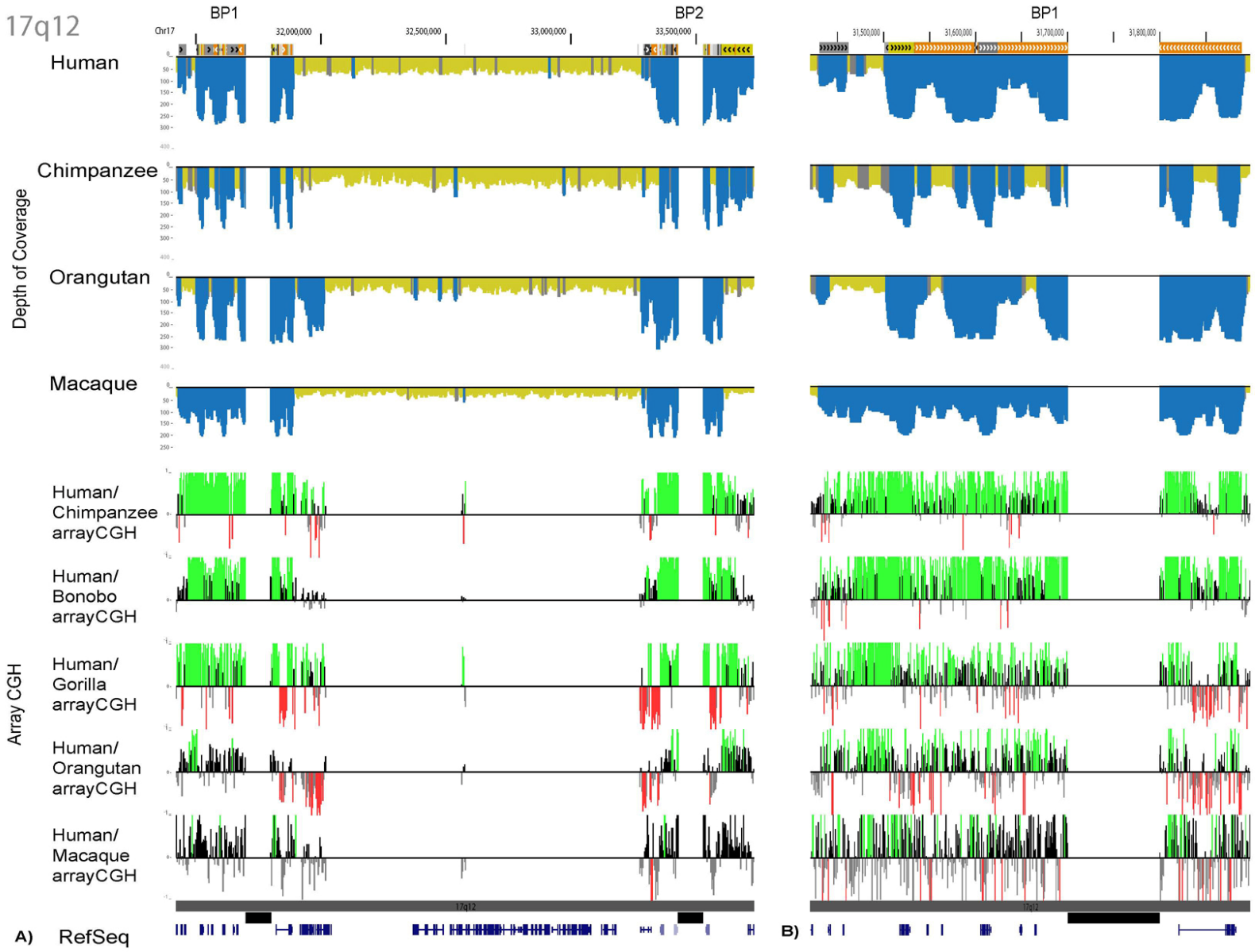


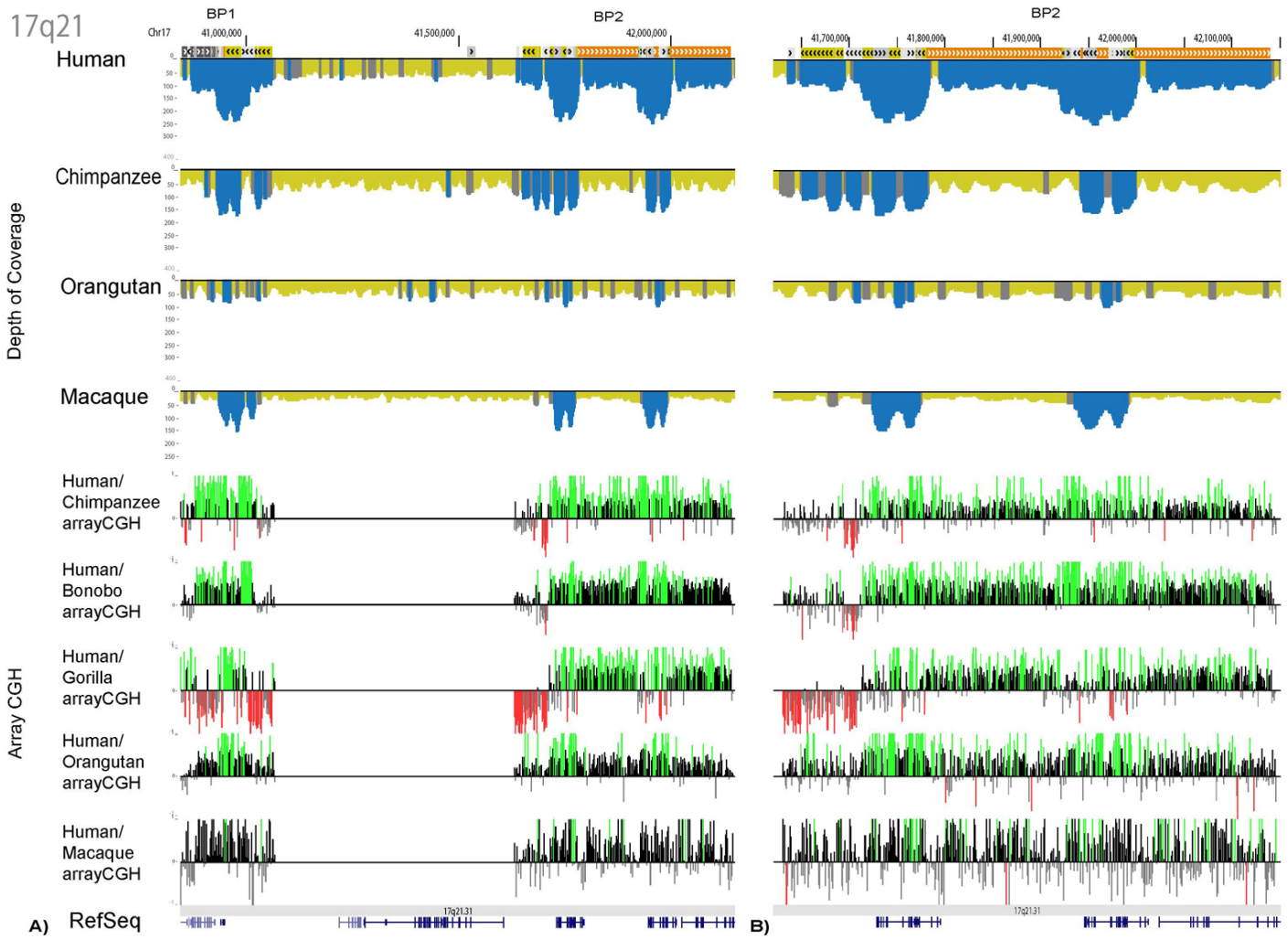
Supplementary Figure 7: Evolutionary origin of 3q29, 15q13.3, 17q12 and 15q24 inversions. None of these individuals were found to be inverted, suggesting that the direct (H1) configuration represents the ancestral state. Arrows indicate inverted chromosomes. Metaphase FISH-based assay could not be used to resolve the 17q12 inversion in PPY as the fosmid probe WIBR2-1900P13 maps within a larger evolutionary paracentric inversion in orangutan (2).


3q29
BP1
BP2
Chr3 197,000,000
197,500,000
198,000,000
198,500,000
Human
Chimpanzee
Orangutan
Macaque
Depth of Coverage
Human/ Chimpanzee arrayCGH
Human/ Bonobo arrayCGH
Human/ Gorilla arrayCGH
Human/ Orangutan arrayCGH
Human/ Macaque arrayCGH
Array CGH\
RefSeq
A)
BP1
196,800,000
196,900,000
197,000,000
197,100,000
197,200,000
B)

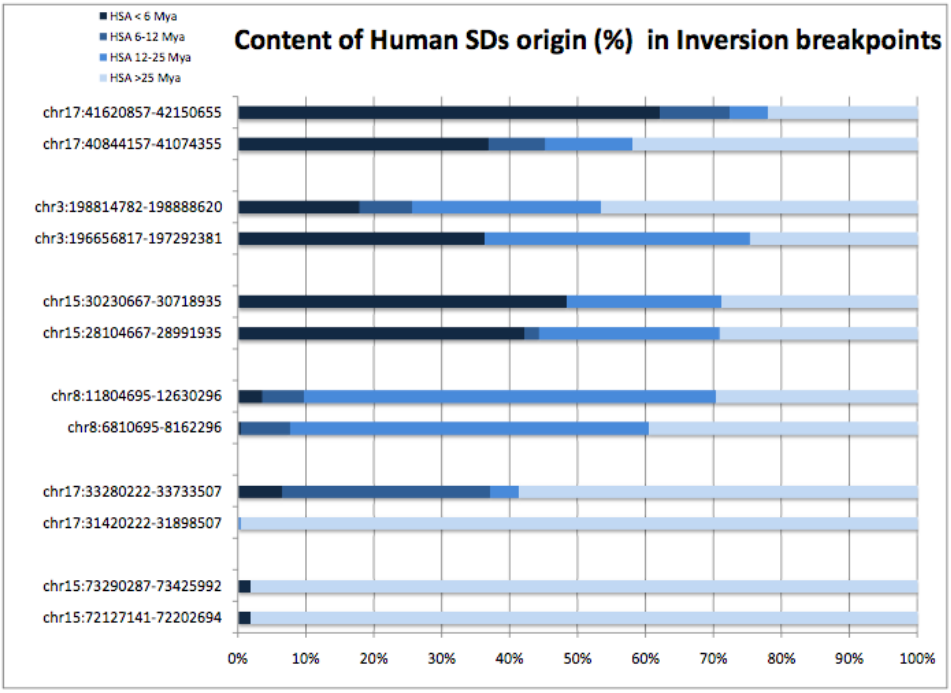

15q13
BP1
BP2
Chr15
28,500,000
29,000,000
29,500,000
30,000,000
30,500,000
Human
Chimpanzee
Orangutan
Macaque
Depth of Coverage
Human/
Chimpanzee
arrayCGH
Human/
Bonobo
arrayCGH
Human/
Gorilla
arrayCGH
Human/
Orangutan
arrayCGH
Human/
Macaque
arrayCGH
Array CGH
A)
RefSeq
15q13.2
15q13.3
BP1
28,200,000
28,400,000
28,600,000
28,800,000
B)
15q13.2


15q24
BP1
BP2
Chr15
72,500,000
73,000,000
73,350,000
73,400,000
Human
Chimpanzee
Orangutan
Macaque
Depth of Coverage
Human/
Chimpanzee
arrayCGH
Human/
Bonobo
arrayCGH
Human/
Gorilla
arrayCGH
Human/
Orangutan
arrayCGH
Human/
Macaque
arrayCGH
Array CGH
15q24.1
15q24.2
RefSeq
A)
B)

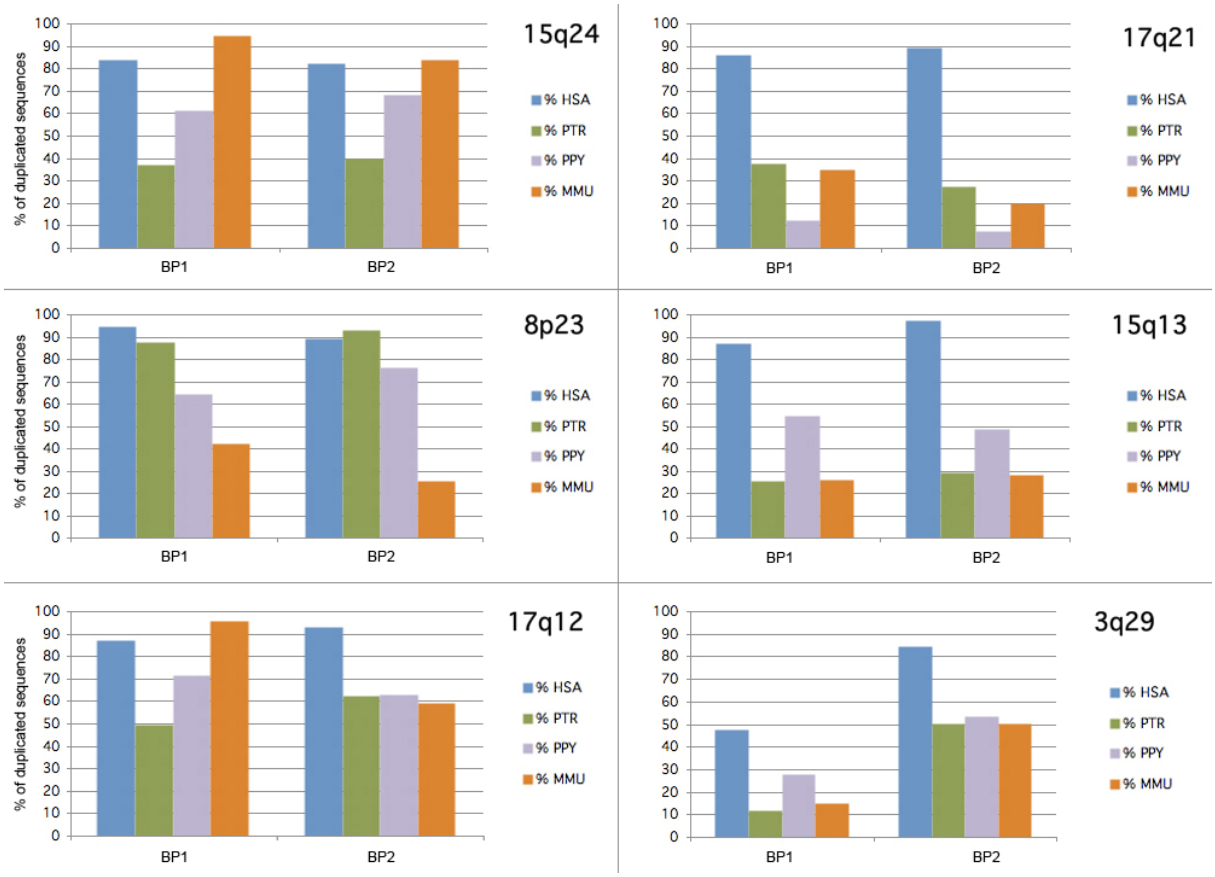

17q12
BP1
BP2
Chr17
32,000,000
32,500,000
33,000,000
33,500,000
31,500,000
31,600,000
31,700,000
31,800,000
Depth of Coverage
Human
Chimpanzee
Orangutan
Macaque
Array CGH
Human/ Chimpanzee arrayCGH
Human/ Bonobo arrayCGH
Human/ Gorilla arrayCGH
Human/ Orangutan arrayCGH
Human/ Macaque arrayCGH
A)
RefSeq
B)



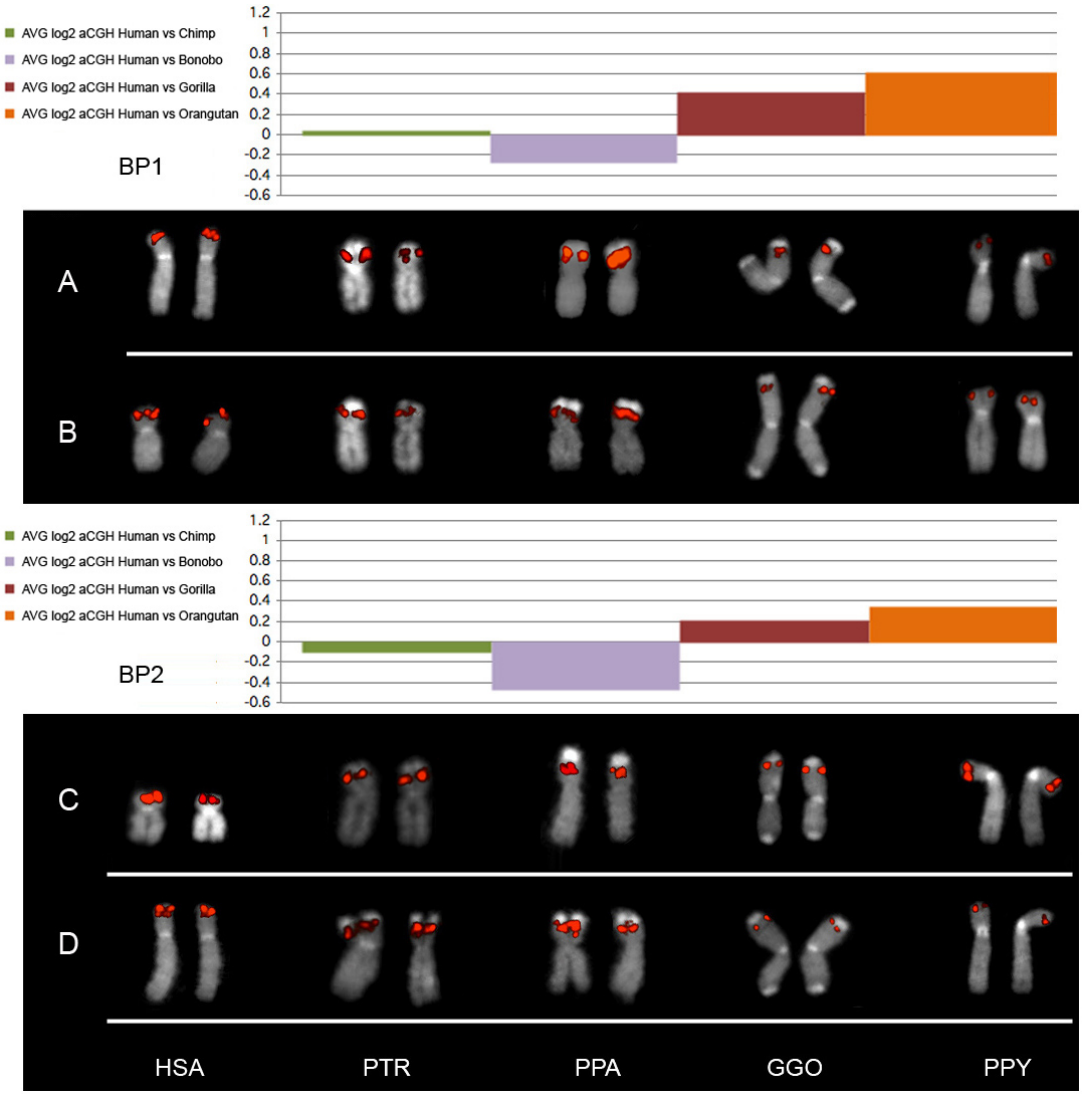
Supplementary Figure 8: (A) Comparative segmental duplication analysis of the 3q29, 15q13.3, 17q12, 15q24 and 17q21.31 inversion regions. Regions of excess WGS depth of coverage (duplicated regions) are shown in blue. ArrayCGH experiments using human as a reference are shown for chimpanzee, bonobo, gorilla, orangutan and macaque. (B) Enlargement of the distal or proximal breakpoints is shown.



Supplementary Figure 9: Estimated age of human segmental duplications at the breakpoints of the inversions is shown as a percentage of the total. Human segmental duplications were categorized into different evolutionary timepoints based on a parsimonious four-way comparison of macaque, orangutan, chimpanzee and human segmental duplications. A human-specific segmental duplication was set to younger than 6 million years (Myr); 6–12 Myr for duplications only shared by human and chimpanzee; 12–25 Myr for those shared by human, chimpanzee and orangutan; and finally, a duplication shared by all four species (human, chimpanzee, orangutan and macaque) was set to be older than 25 Myr.



Supplementary Figure 10: Primate duplication content. The percentage of sequence detected as duplicated at each inversion breakpoint (BP1 and BP2) is shown for each breakpoint region.



Supplementary Figure 11: A bonobo expansion of 8p23 segmental duplications. The average $\log_2$ ratio by arrayCGH for segmental duplications at the 8p23 inversion breakpoint is compared against FISH using clones ABC13-48886900D12 (A), ABC13-48857200H3 (B), WIBR2-1720G15 (C) and WIBR2-1318J03 (D). Note, both arrayCGH and FISH predict excess copy number of segmental duplications in bonobo when compared to human.

Supplementary References

1. Bailey, J.A., Yavor, A.M., Massa, H.F., Trask, B.J. and Eichler, E.E. (2001) Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res,* **11,** 1005-17.
2. Cardone, M.F., Jiang, Z., D'Addabbo, P., Archidiacono, N., Rocchi, M., Eichler, E.E. and Ventura, M. (2008) Hominoid chromosomal rearrangements on 17q map to complex regions of segmental duplication. *Genome Biol,* **9,** R28.