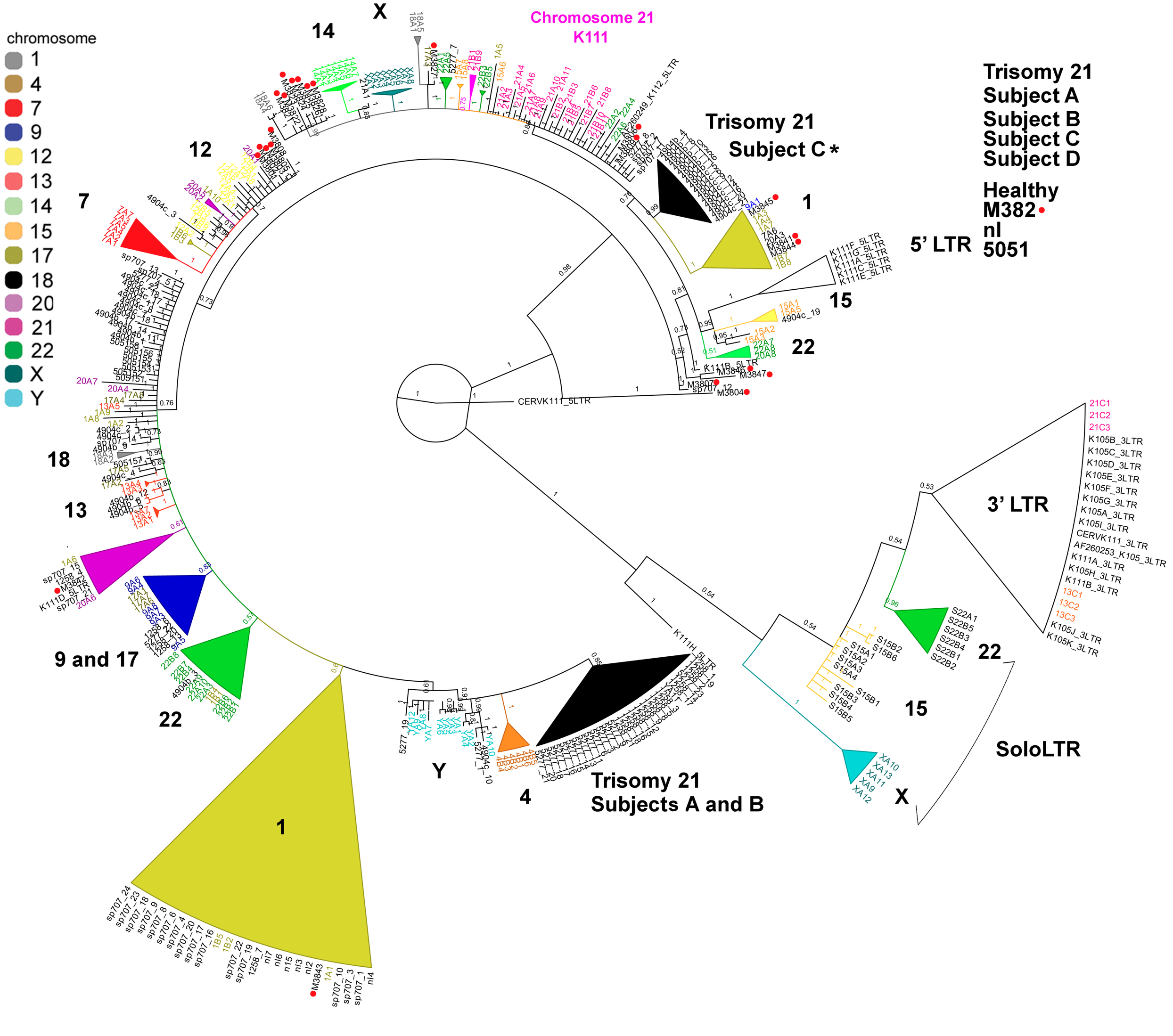
chromosome
1
4
7
9
12
13
14
15
17
18
20
21
22
X
Y
Chromosome 21
K111
Trisomy 21
Subject C *
Trisomy 21
Subject A
Subject B
Subject C
Subject D
Healthy
M382
nl
5051
5' LTR
3' LTR
SoloLTR
Trisomy 21
Subjects A and B
CERVK111_5LTR
K111H_5LTR
K111F_5LTR
K111G_5LTR
K111A_5LTR
K111C_5LTR
K111E_5LTR
K111B_5LTR
21C1
21C2
21C3
K105B_3LTR
K105C_3LTR
K105D_3LTR
K105E_3LTR
K105F_3LTR
K105G_3LTR
K105A_3LTR
K105I_3LTR
CERVK111_3LTR
AF260253_K105_3LTR
K111A_3LTR
K105H_3LTR
K111B_3LTR
13C1
13C2
13C3
K105J_3LTR
K105K_3LTR
14
X
12
7
18
13
9 and 17
22
1
4
Y
15
22
15
X

Figure S11. The pericentric K111 marker shows loss and apparent recombination in trisomy 21. Bayesian inference tree of the K111 5′ LTR insertions amplified from each individual human chromosome together with total K111 proviruses amplified in healthy individuals and subjects with trisomy 21. Sequence labels are colored to indicate from which human chromosome they arise. Each chromosome was assigned a color depicted on the legend at left. Note that each color tends to cluster to specific evolutionary branches, indicating that individual K111s often spread within an individual chromosome (black bold numbers or triangle colored branches). Posterior probability values greater than 70 are shown. K111 sequences from healthy individuals (M382, nl, and 5051) and DNA from individuals with trisomy 21 (legend on the right; Subject A 1258, Subject B 5277, Subject C 4904, and Subject D sp707) distribute along the tree clustering close to the K111 sequences specific to each chromosome. As an example, K111 sequences from a healthy individual (M386) are indicated (red dots). Several K111 sequences in trisomy 21 subjects A, B, and C cluster to novel branches (black triangle branches), likely the result of homologous recombination, and K111 sequences are substantially reduced in number in Chromosome 21 in these individuals with Down Syndrome (sequences in magenta). The tree was generated using Bayesian inference with four independent chains run for at least 10,000,000 generations until sufficient trees were sampled to generate >99% credibility. 5′ LTRs, 3′ LTRs and soloLTR lineages are shown along with the chimpanzee LTRs (CERV-K111). Nucleotide sequence substitutions that are specific for the K111 group of sequences found in each chromosome were used to generate the tree (Contreras-Galindo et al. 2013).