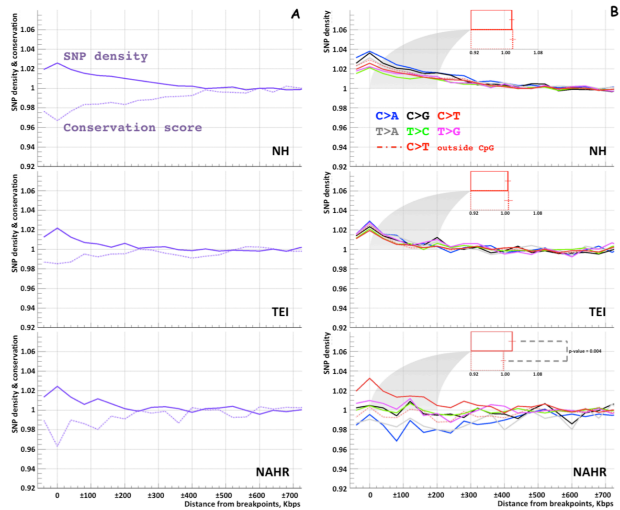
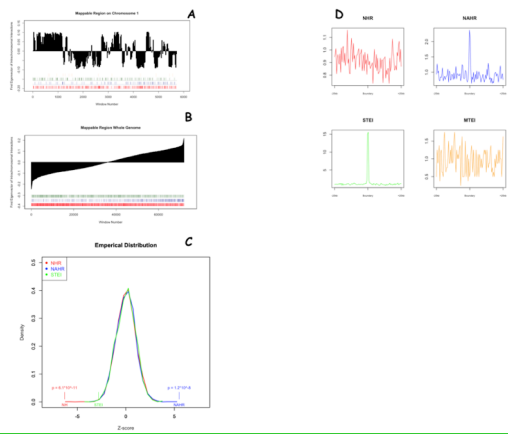


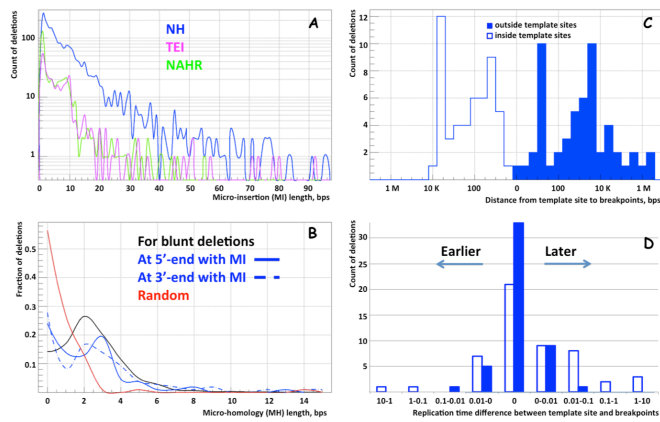
**Figure 1.** Deriving confident set of breakpoints. A) Conceptual steps for the derivation. Breakpoints from local target assembly are filtered by mapping reads to putative junctions. B) Null model for breakpoint filtering. C) Comparison of different breakpoint sets. Note, the pilot set {REF} was included in the derivation as one of the call sets. Integrated set {REF} was bias towards large non-repetitive deletions for the purpose of reliable genotyping, resulting in mobile element insertions being strongly under represented.



**Figure 2.** Co-aggregation of SNPs and deletion breakpoints found in the analyzed samples. A) Normalized SNP densities increased while conservation decreased in 400 kbps regions around breakpoints of each class. B) Densities increase for substitutions of all types around NH and TEI breakpoints. Increase of C>T substitutions around NAHR breakpoints is explained by enrichment of CpG motifs.



**Figure 3.** Analysis of NAHR. A) breakpoint incidence on Chromosome1 with first eigenvector of Hi-C data attached. B) Sorted Hi-C first eigenvector on the whole genome and incidences of breakpoint. C) Empirical distribution given by permutation shows the depletion of NH and enrichment of NAHR in open chromatin is statistically significant.



**Figure 4.** Analysis of micro-insertions (MI) at deletion junctions. A) Most of MIs are up to 10 bps in length and can be explained by mismatch/indel close to breakpoint leading to unambiguous alignment. Larger MIs are typically found for NH deletions. B) Length of micro-homology (MH) at deletion junction. For deletions with MIs and identifies template site MH are calculate for 5'-ends/3'-ends of the deletion and the template site (**Fig. S3**). C) The distribution of the nearest distance between template site and either of the breakpoints exhibits distinct peaks in the ranges 10-100 bps (**proximal sites**) and 2-7 kbps (**distal sites**). D) The different in replication time between template site and breakpoints reveals later replication of template sites when template site is outside the deletion (p-value < 0.03 by binomial test). The effect is even more significant (p-value < 0.01) when excluding difference of up to 0.01 as such small values are comparable to measurement error.