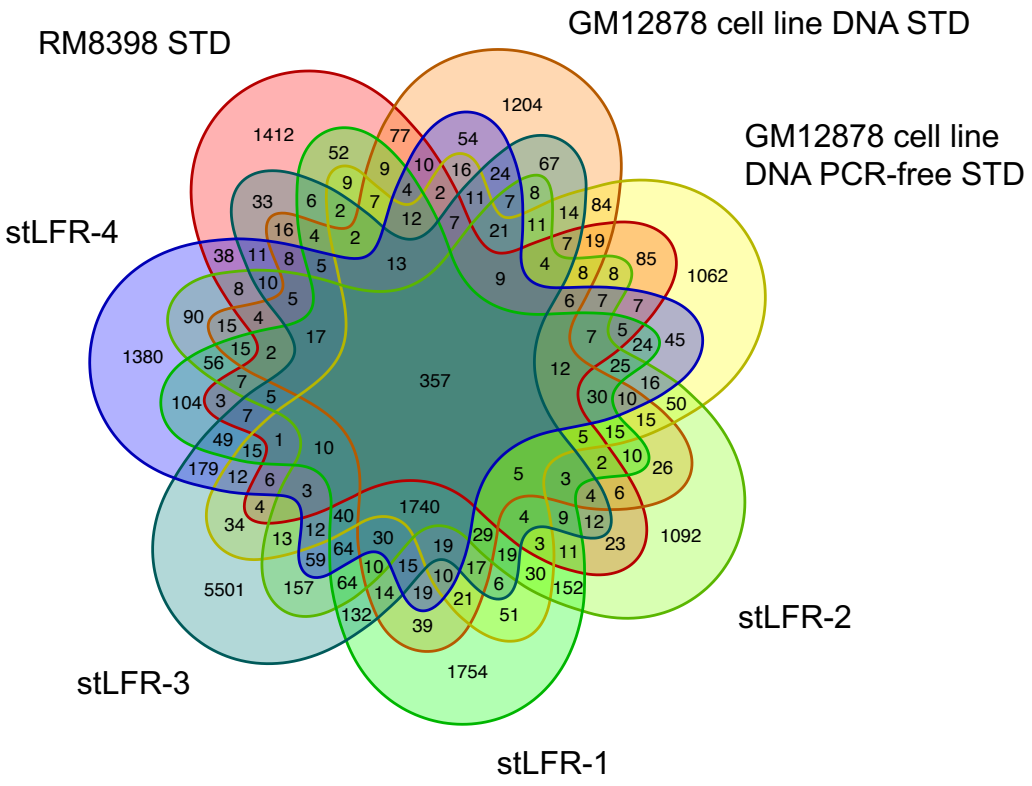
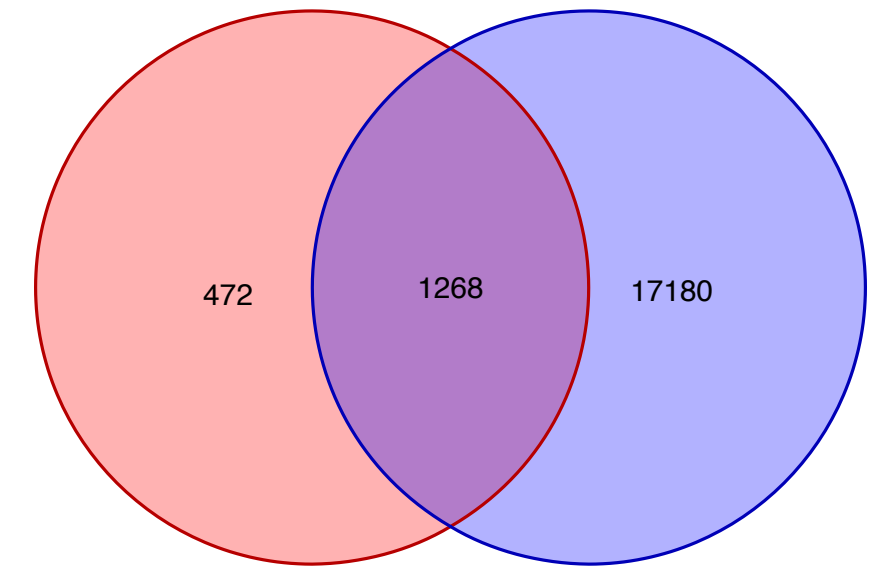
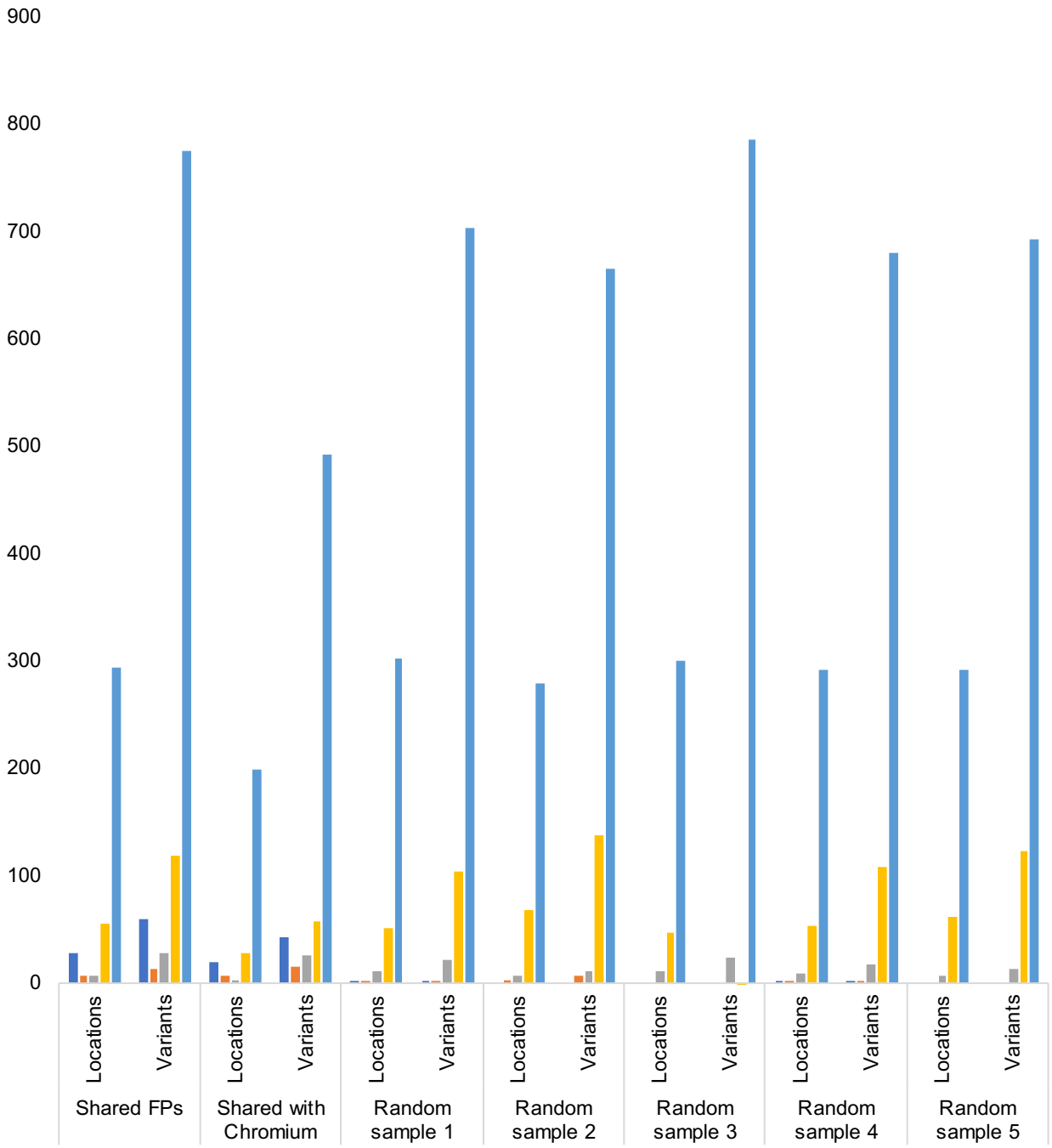
A



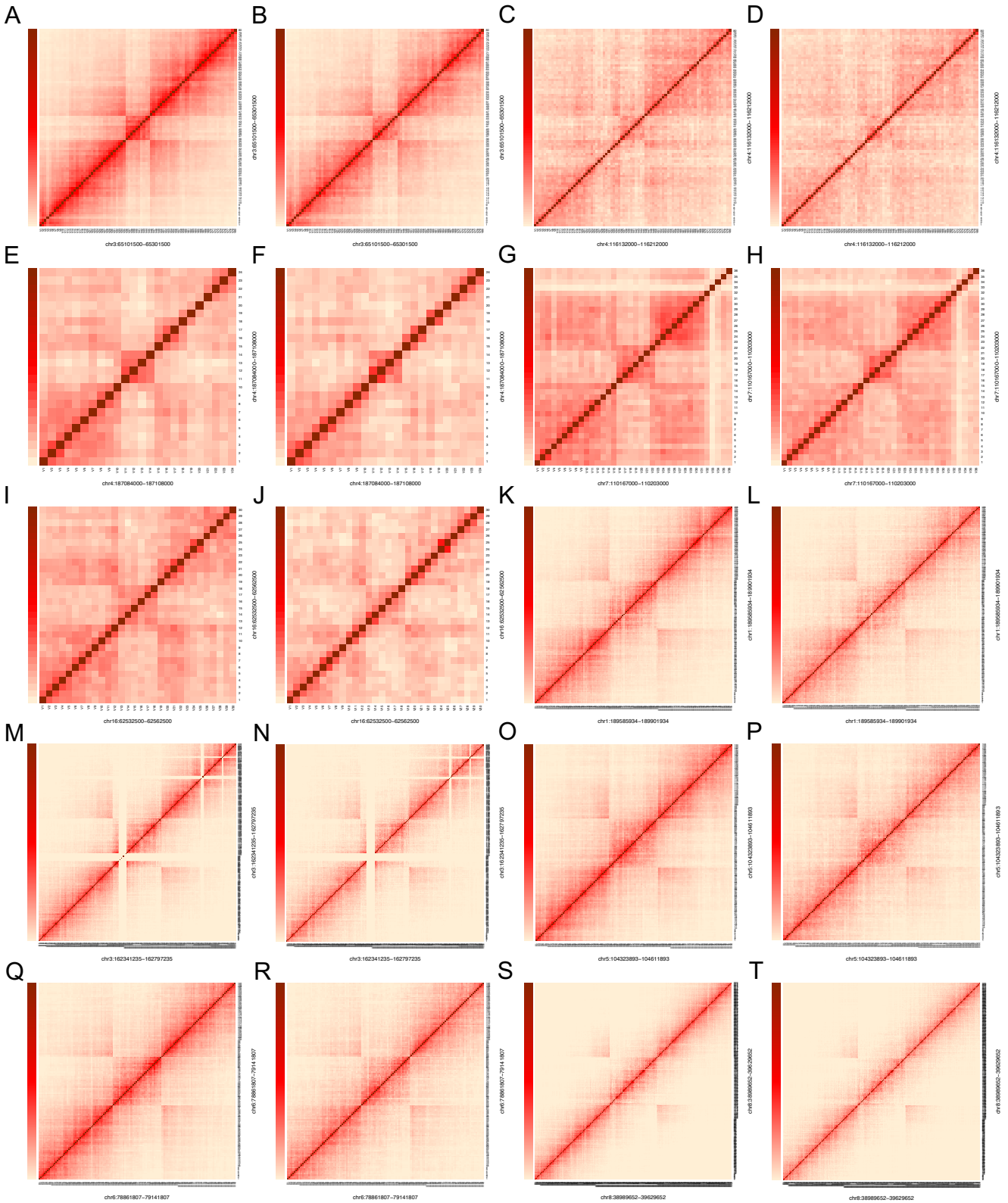
B



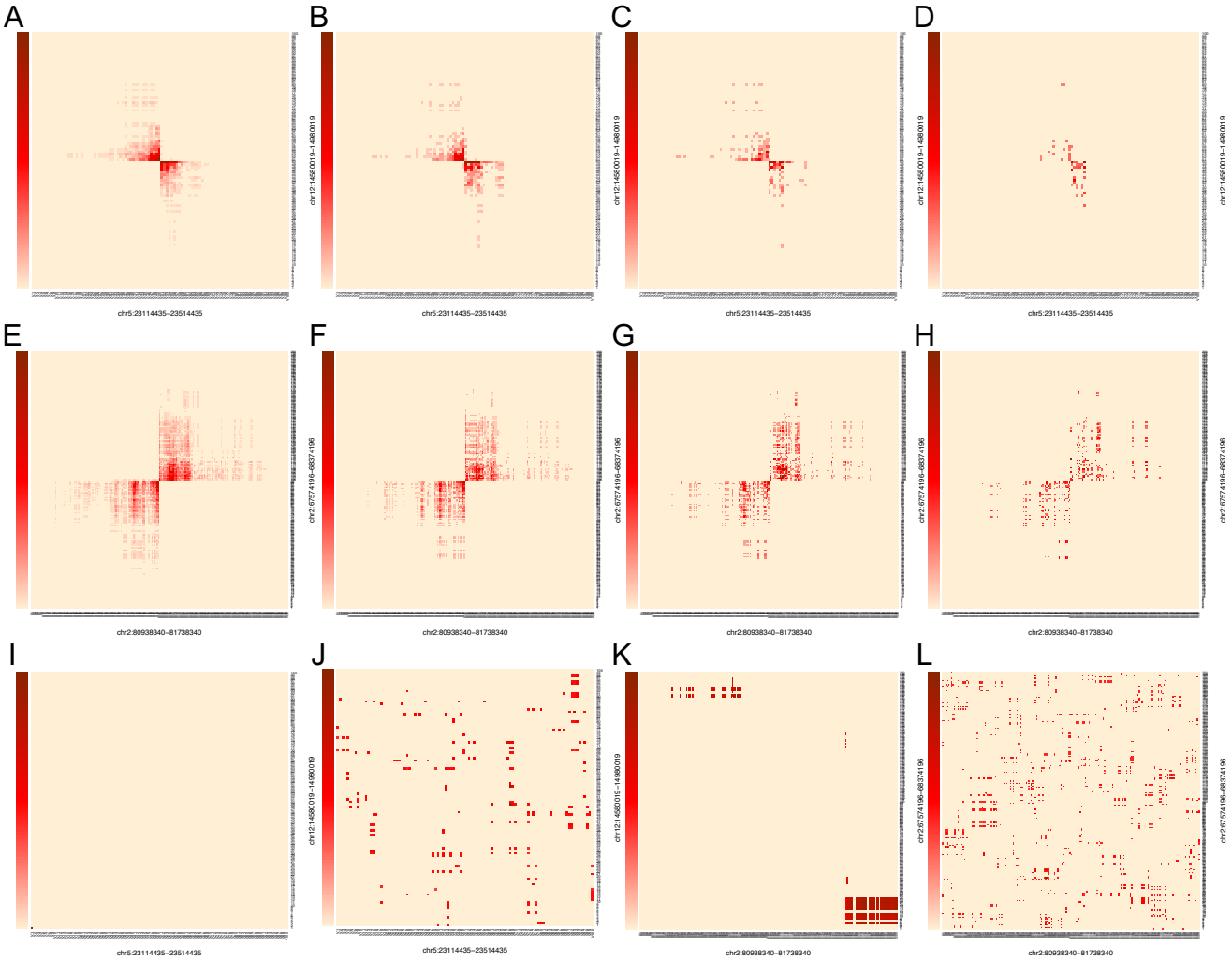
Supplemental Fig S1. Overlap of FPs between libraries. (A) The FPs from each stLFR library and the BGISEQ-500 standard libraries from GM12878 DNA and GIAB reference material were plotted in a Venn diagram. 1,740 FPs are shared between the four stLFR libraries and the two standard libraries made from cell line DNA, but not found in the standard library made from the GIAB reference material. (B) The overlap of stLFR library FPs and Chromium library FPs shows that 1,268 are shared between the two different technologies that both used DNA isolated from GM12878 as opposed to the GIAB reference material for NA12878. 472 FPs are unique to stLFR libraries.



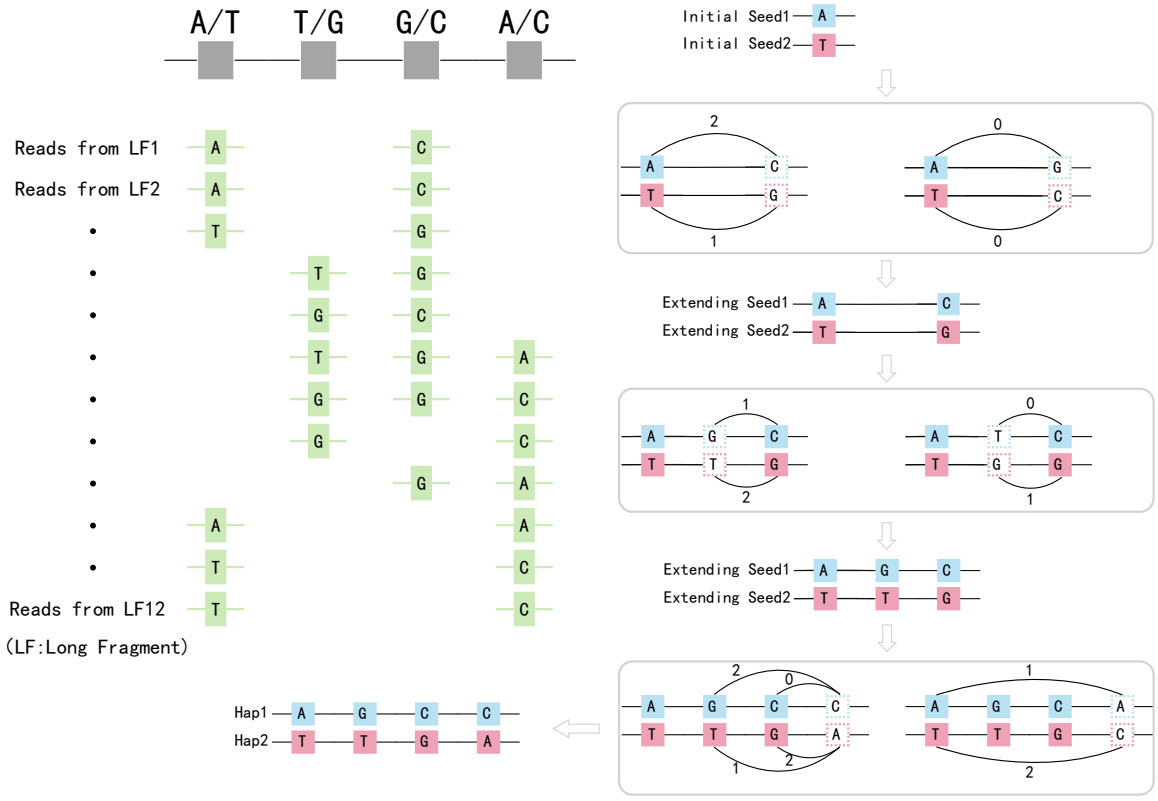
Supplemental Fig S2. Shared false positive variant distribution. The genomic distance separating 1,758 shared FP variants were summed within consecutive bins of 100 bp (dark blue), 1,000 bp (orange), 10,000 bp (grey), 100,000 bp (yellow), and 1 million bp (light blue). 5 sets of 1,758 randomly selected variants from the stLFR-1 library were also plotted. For each sample the total number of locations or the total number of variants are plotted. Only bins or the variants within bins where 2 or more variants are found are summed.



Supplemental Fig S3. NA12878 deletion detection using barcode sharing heatmaps. Detection of deletions in the stLFR-1 library at chr3:65189000-65213999 using 230 Gb (A) or 100 Gb (B), chr4:116167000-116176999 using 230 Gb (C) or 100 Gb (D), chr4:187094000-187097999 230 Gb (E) or 100 Gb (F), chr7:110182000-110187999 230 Gb (G) or 100 Gb (H), chr16:62545000-62549999 230 Gb (I) or 100 Gb (J), chr1:189704509-189783359 230 Gb (K) or 100 Gb (L), chr3:162512134-162569235 230 Gb (M) or 100 Gb (N), chr5:104432113-104467893 230 Gb (O) or 100 Gb (P), chr6:78967194-79001807, and chr8:39232074-39309652 230 Gb (S) or 100 Gb (T) of read data.



Supplemental Fig S4. Translocation and inversion detection with stLFR. A patient cell line and cell line GM20759 harboring a translocation between chromosome 5 and 12 and an inversion on chromosome 2, respectively, were analyzed with stLFR. For each library the total sequence coverage was downsampled to investigate the detection ability at lower coverages. The translocation between chromosome 12 and 5 was easily detected at total sequence coverages of 40 Gb (A), 20 Gb (B), 10 Gb (C), and even 5 Gb (D). The inversion in GM20759 was also easily detected at total sequence coverages of 46 Gb (E), 20 Gb (F), 10 Gb (G), and 5 Gb (H). In addition, we investigated these regions in the GM12878 cell line which is not known to harbor either of these SVs. The translocation between chromosome 5 and 12 was not evident in either the stLFR library from 1 ng with 230 Gb of coverage (I) or the 10 ng library with 126 Gb of coverage (J). The transversion also was not found in the stLFR-1 (K) or stLFR-4 library (L) either.



Supplemental Fig S5. LongHap phasing. A full description of the phasing algorithm applied with LongHap can be found in the Methods and Materials.