Haplotype-resolved Sequencing the Major Histocompatibility Complex (MHC) Locus using HLS-CATCH™ Target Enrichment and TELL-Seq™ Linked Read Sequencing

sage science

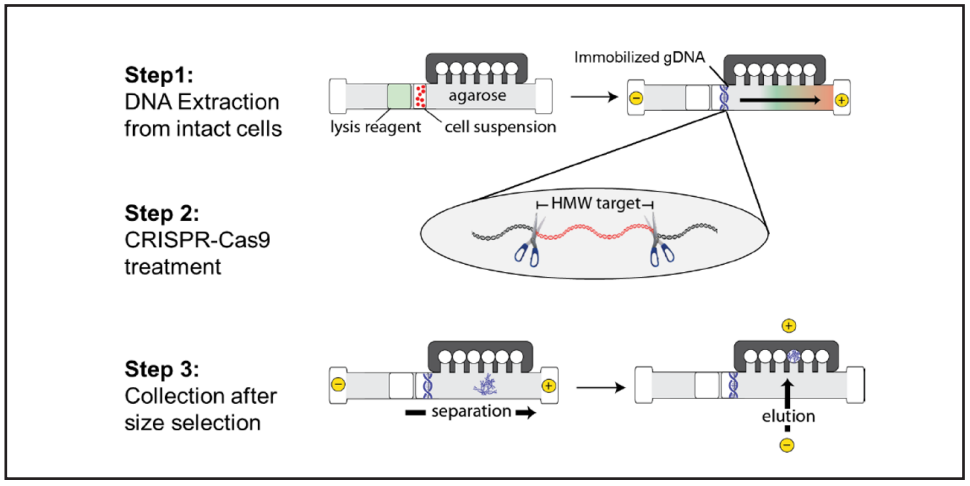
Application Note: HLS2

The MHC locus was enriched using Cas9 digestion of restriction sites at 400kb intervals using the HLS-CATCH process. Phasing analysis was undertaken by TELL-Seq linked read analysis from Universal Sequencing.

Chris Boles[1], Tom Chen[2], and Peter Chang[2]
[1]Sage Science, Inc. [2]Universal Sequencing Technology Corporation

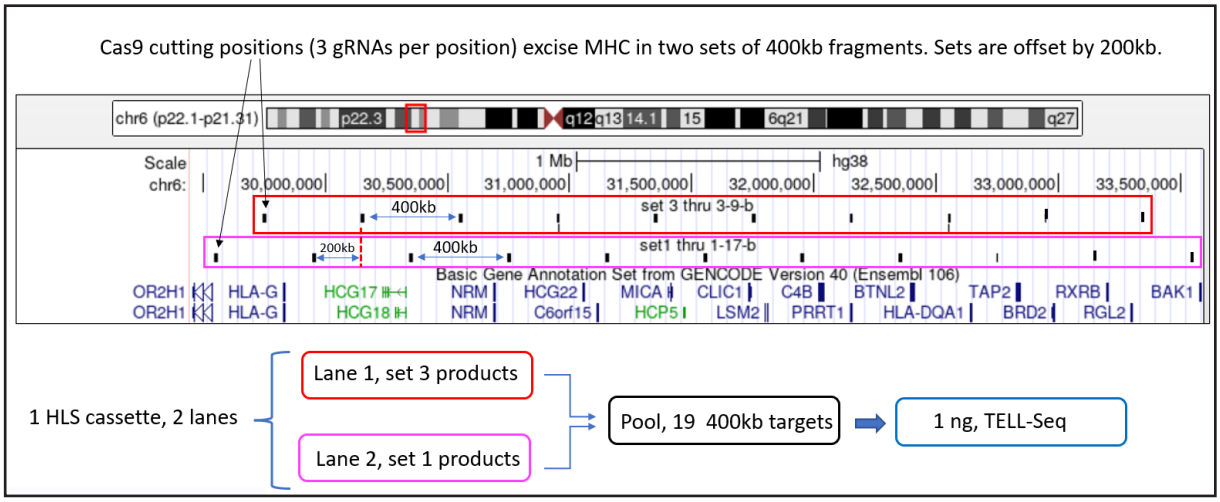
The HLS-CATCH Method

Long-range sequencing information is required for haplotype phasing and detection of large structural variants. HLS-CATCH is a method that purifies high molecular weight genomic regions with targeted Cas9 restriction, illustrated below.



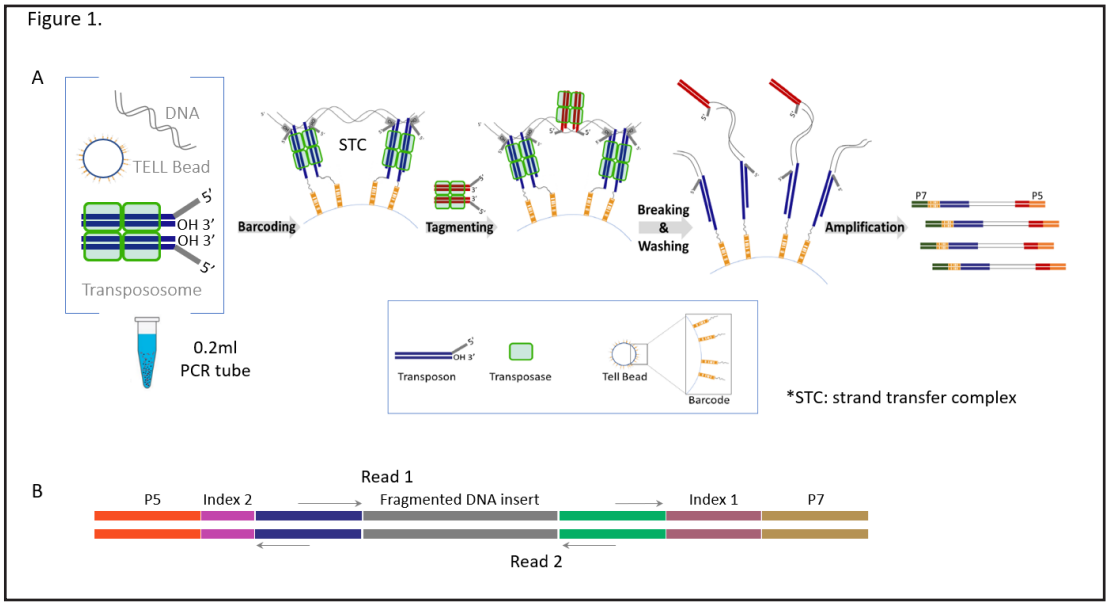
MHC Guide RNA design and pooling strategy

Cells from a lymphblastoid cell line (Ashkenazi son, HG002, Coriell Instiitute) were processed using the HLS-CATCH method. To purify the MHC locus (~4MB), guideRNAs (gRNA) were designed with three guides per cutting position which each position ~400kb apart to fragment the entire locus into ~10 fragments. A second set of gRNAs were designed in the same manner but off-set by a distance of 200kb relative to the first set (below). These were run in two separate reactions on different lanes on a gel cassette on the HLS2 platform. The 400kb target fractions were pooled and then subject to TELL-Seq linked read sequencing.



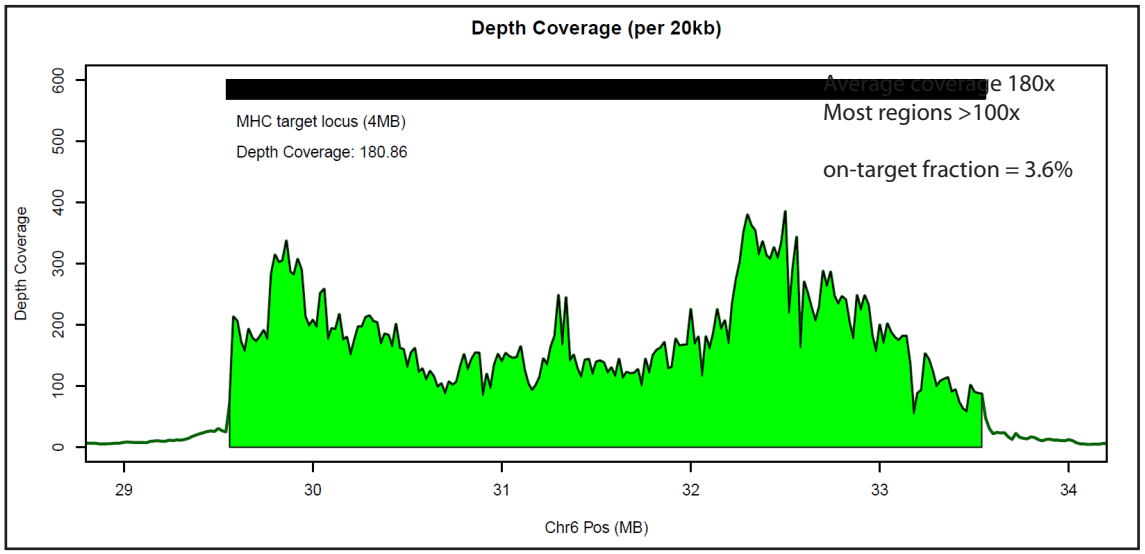
The TELL-Seq Linked Read Library Construction

TELL-Seq linked-read read libraries are constructed in a single tube with a 3 hour workflow and requires 1 ng of DNA input. For phasing analysis, linked read sequencing benefits from HMW input. The transposase-based Tell-Seq method is outlined below:



Preliminary Results

Prelimiary results showed sufficient coverage and indicates that this approach can provide useful phasing data. Some of these metrics are summarized below.



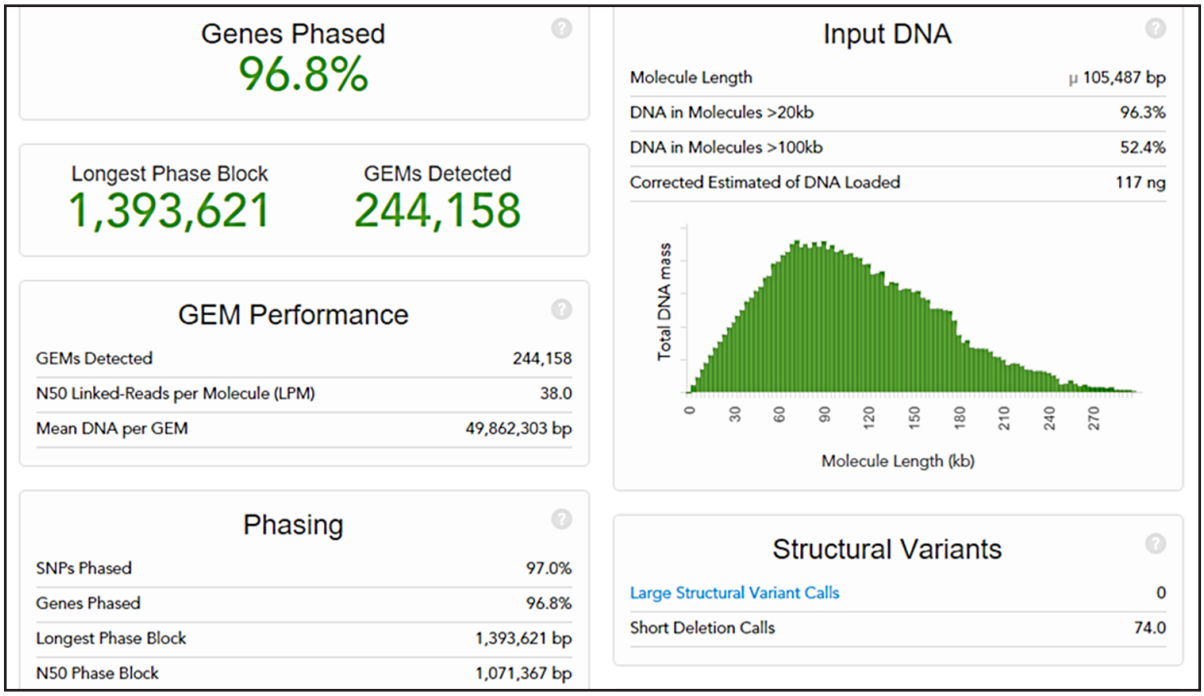
Coverage Plot

MHC Phase Block: 11717 SNPs

3 Haplotype Switch Errors (0.026%)

50 Variant Errors (0.43%)

IGV view of the MHC Phase



Long Ranger Report