January 11, 2021
CAP-0001

CUSTOM PANEL OVERVIEW	
Number of Targets:	1567 (1518 SNPs, 49 InDels)
Number of Probes:	2809
Probe Footprint (bps):	253,731 bps
Partial Targets NOT Covered:	3 InDels
Whole Targets NOT Covered:	184 (175 SNPs, 9 InDels)

PROJECT RATIONAL AND GOALS

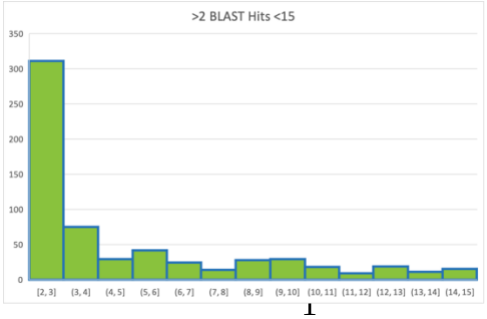
NIST seeks to validate putative low-frequency variants identified in NIST's GIAB HG002 reference material, and to further discern de novo mutations that occur overtime during cell line propagation. CHOP will sequence the capture libraries.

1) Design a hybrid capture panel comprised of SNPs and InDels targets ☑
2) Send Kit to NIST for DNA sample processing and subsequent sequencing at CHOP
3) Provide access to DNANexus to generate error-corrected BAM and VCF/MUT files to carry out basic mutagenesis analysis and assess CHIP clonal populations

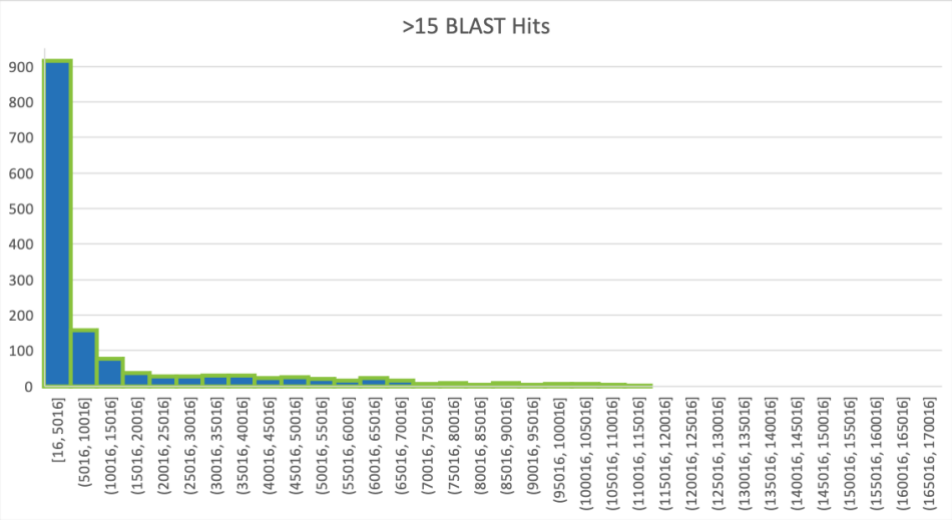
CUSTOM PANEL DESIGN DESCRIPTION

Custom Panel was designed for NIST/CHOP targeting 1,567 individual SNPs and InDels defined by NIST. Target files were provided by NIST in the human hg38 genome build which will facilitate downstream analysis using DNANexus. A total of 2,816 probes were designed to NIST's target regions after a rigorous filtration strategy. Targets were filtered as follows:

1. Probes were designed to the original 2,522 SNPs provided by NIST resulting in 4,828 probes. Probes were then BLAST-ed against the human genome (hg38).
    a. 2,698 probes had a single BLAST hit
    b. 624 had ≥ 2 BLAST hits ≤ 15



c. 1,478 probes with >15 hits in the genome were filtered



d. 1,518 SNPs remained for further analysis
e. Further filtration of SNPs and InDels for off-target effects and predicted secondary structured resulted in limited removal of a few targets, as noted in table and depicted below.

A total of 2816 probes will cover a footprint of ~254Kb of the human genome

Figure 1. No Coverage. SNP 126.

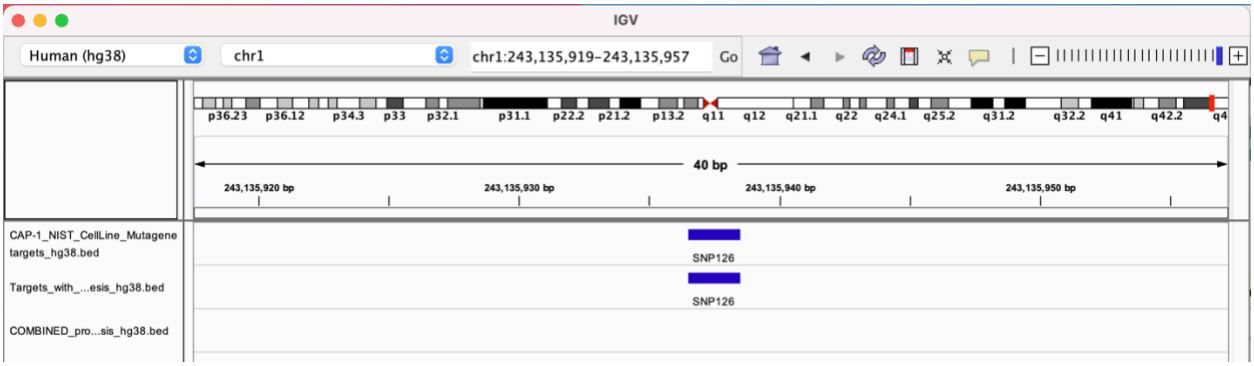


Figure 2A. Partial Coverage. InDel 9.