PANEL INFORMATION

# TwinStrand Duplex Sequencing™ Mutagenesis Panel (Human-50), v1.0

**Tube Label:** PAN — Mutagenesis Panel (Human-50) v1.0

**Part Number:** 06-1005-XX

**Pipeline ID:** human-muta-v1.0

| | |
|---|---|
| **Total Panel Size:** | 48 kb |
| **Number of Target Intervals:** | 20 |
| **Target Type:** | representative genomic regions |
| **Genome Reference:** | hs38DH (Human) |

This human mutagenesis panel is designed to quantify and characterize background and induced mutations in the nuclear genome of human DNA samples. Probes target 20 representative genomic regions not believed to be under significant positive or negative somatic selection. Outputted mutation data are used to estimate genome-wide per-nucleotide mutation frequency, spectrum and trinucleotide signatures for different genotoxic exposures.

## PANEL INFORMATION

| Target Regions in Panel | | |
|---|---|---|
| Chromosome | Start | End |
| chr1 | 84,597,127 | 84,599,527 |
| chr2 | 40,162,767 | 40,165,167 |
| chr4 | 22,386,244 | 22,388,644 |
| chr6 | 155,239,014 | 155,241,414 |
| chr7 | 11,732,774 | 11,735,174 |
| chr8 | 51,513,056 | 51,515,456 |
| chr9 | 23,709,463 | 23,711,863 |
| chr10 | 128,969,037 | 128,971,437 |
| chr11 | 108,510,787 | 108,513,187 |
| chr12 | 114,115,043 | 114,117,443 |
| chr13 | 75,803,912 | 75,806,312 |
| chr14 | 74,661,755 | 74,664,155 |
| chr15 | 46,089,737 | 46,092,137 |
| chr16 | 51,754,103 | 46,092,137 |
| chr17 | 70,672,726 | 70,675,126 |
| chr18 | 5,749,264 | 5,751,664 |
| chr19 | 31,831,021 | 31,833,421 |
| chr20 | 24,153,684 | 24,156,084 |
| chr21 | 23,665,976 | 23,668,376 |
| chr22 | 48,262,370 | 48,264,770 |

Coordinates in the BED format (0 start, half-open)

## Panel-Specific Protocol Recommendations

| | | | |
|---|---|---|---|
| **Cot-1 DNA Recommended:** | Human: 03-2020-XX | **Number of Final PCR Cycles:** | 14 (PCR-2), 6 (PCR-3) |
| **Number of Captures Required:** | 1 | **Conditioning Reagent:** | Strongly recommended |
| **Typical On-Target Percentage:** | 90%-95% (1 capture), 95%-99% (2 captures) | | |

## Panel-Specific Protocol Notes

**DNA Conditioning:** For mutagenesis panels, which require extreme sensitivity in the $10^{-7}$ mutations per nucleotide range, the DNA conditioning step should always be performed.

**1 vs. 2 Captures:** With probe panels of this size, which have been specifically designed to avoid pseudogenes and other difficult regions, a single round of hybrid capture is generally sufficient to obtain 90% or greater on-target bases. However, we recommend starting with 2 captures the first time through the protocol to generate data that can be compared against. An increase in off-target reads means that raw sequencing must be increased proportionally to generate the same amount of informative data, so a trade-off exists between time saved with only 1 capture vs. additional sequencing cost. The user should consider the pros and cons in the context of their specific situation.

**Final PCR Cycles:** The number of final PCR cycles listed in the table above assumes 500 ng of high quality genomic DNA input at the beginning of library preparation. However, the number of final PCR cycles necessary may vary with genomic DNA input mass, DNA quality, probe panel size and nature, number of captures, thermocycler calibration and other factors. The values in the table should be considered a starting point. If 2 rounds of hybrid capture are performed then PCR-3 is the final PCR, and with 1 round of capture PCR-2 is the final PCR.

**Statistical Power:** Mutagenesis experiments should typically target a specific number of Informative Duplex Bases per sample such that a sufficient number of mutagenic events are observed. In general, a minimum of 10-15 mutations are ideal for simple frequency measurements, and at least 100 mutations are needed to assess trinucleotide spectra. When calculating overall mutation frequency, unique mutation events are the numerator, and Informative Duplex Bases are the denominator. Mutation frequency in genomic DNA from control tissue samples can be as low as $1 \times 10^{-7}$, but that must be determined empirically for each sample type. In the absence of prior data, we recommend targeting between 500 million and one billion Duplex Bases to start.

| DNA Input vs. Projected Duplex Sequencing Data | | | | | | |
|---|---|---|---|---|---|---|
| DNA Input (ng)* | Number of Clusters (Million)** | Number of Paired-End Reads (Million)** | Illumina Sequencing Cost (NovaSeq, S4, List Price)*** | Mean Duplex Depth**** | Max Duplex Depth**** | Informative Duplex Bases (Million)**** |
| 250 | 50 | 100 | $144 | 5,500x | 7,000x | 310 |
| 500 | 100 | 200 | $288 | 11,000x | 14,000x | 620 |
| 1,000 | 200 | 400 | $576 | 22,000x | 28,000x | 1,240 |

* Assumes high molecular weight genomic DNA (quantified pre-fragmentation). DNA requirements to achieve a given Duplex Depth when using cell-free DNA or heavily damaged DNA such as that from FFPE will be lower and higher, respectively.
** Approximate sequencing required to achieve a peak tag family size ~10.
*** S4 list price = $7,200/lane (2.5 billion clusters/lane) as of December 2019.
**** Data estimates assume 150 bp paired-end reads

TwinStrand Duplex Sequencing data yields are dependent on accurate quantification, and the quality of input DNA. We recommend Qiagen silica column-based kits or “salting out” methods for DNA extraction. Phenol-chloroform extraction can damage DNA and reduce data yields. DNA should be quantified using a double-stranded DNA-binding fluorescent dye (e.g. Qubit fluorometer, Thermo Fisher Scientific) rather than UV spectrometry which can overestimate concentration if contaminants are present, which will reduce data yields. If possible, check genomic DNA quality using a 0.5% agarose gel, or a fragment analyzer system (e.g. Agilent TapeStation system). DNA integrity number (DIN) >7 on a fragment analyzer or a single high molecular weight band on a gel generally indicates good quality DNA, although oxidation or other forms of non-fragmentation chemical damage may still be present and not appreciable.