

Sample ID: 108R\_SETX\_F10\_KOLF2.1J

Sample Barcode: 208138750025 R05C01 020824

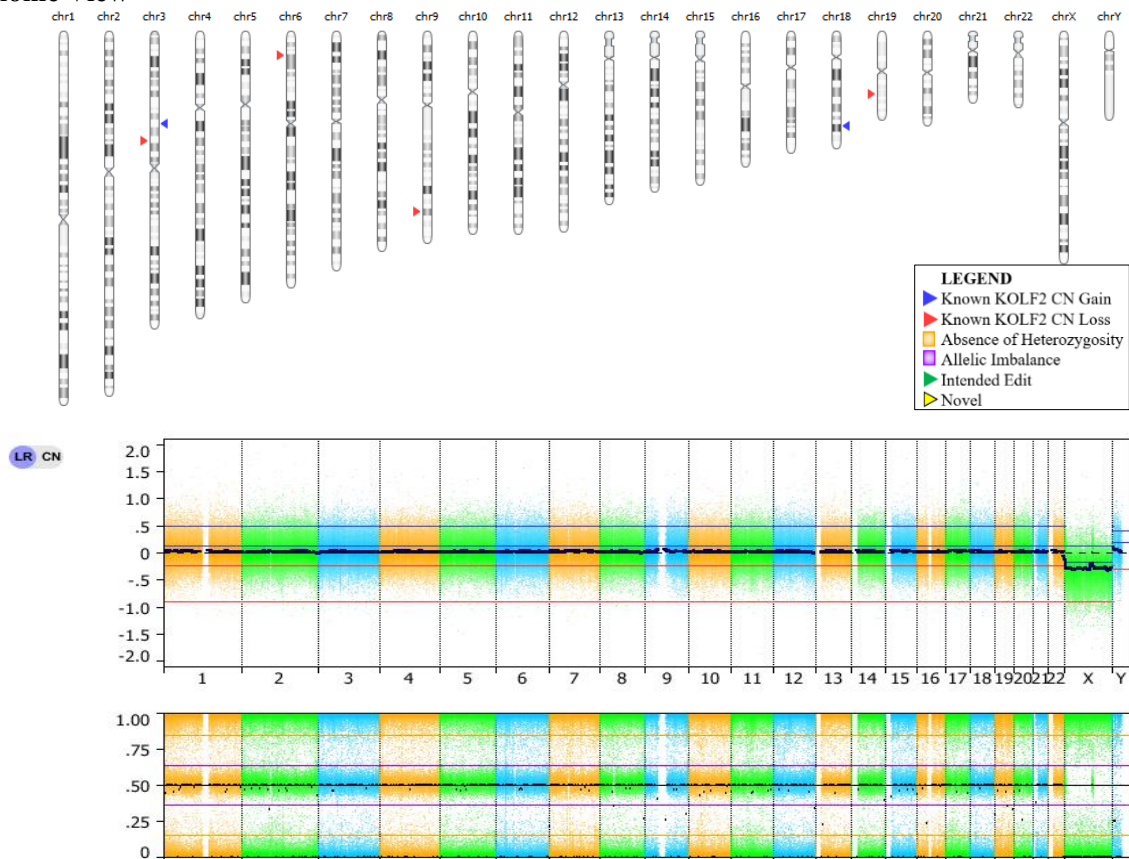
Cell Line: KOLF2.1J, Male

Array Result: PASS

Target Edit: SETX L389S Revertant

at hg19: chr9:135,205,819

### Whole Genome View



Novel Sample-Specific Regions: None detected

### Known KOLF2-C1 CNVs

Below are additional CNVs that have been identified. These regions are known aberrations in the KOLF2-C1 cell line, which is the original iPSC line from which KOLF2.1J was generated. These CNVs are present in all genetically engineered lines derived from KOLF2-C1 and KOLF2.1J.

| Event                | Chr Region - PacBio (hg19) * | Size (bp) | Cytoband | Approx CN | Genes          | Detected** |
|----------------------|------------------------------|-----------|----------|-----------|----------------|------------|
| CN Gain <sup>1</sup> | chr3:61135038-61911274       | 776,236   | 3p14.2   | 3.0       | FHIT, PTPRG    | Automatic  |
| CN Loss <sup>1</sup> | chr3:72289391-72432282       | 142,891   | 3p13     | 1.0       | RYBP           | Automatic  |
| CN Loss <sup>1</sup> | chr6:15490359-15723361       | 233,002   | 6p22.3   | 1.0       | JARID2, DTNBP1 | Automatic  |
| CN Loss <sup>1</sup> | chr9:119241091-119372704     | 131,613   | 9q33.1   | 1.0       |                | Automatic  |
| CN Gain <sup>1</sup> | chr18:62113242-62270679      | 157,437   | 18q22.1  | 3.0       |                | Automatic  |
| CN Loss <sup>2</sup> | chr19:41355704-41387475      | 31,771    | 19q13.2  | 1.0       | CYP2A7         | Automatic  |

\*CNVs were confirmed with PacBio Long Read sequencing data. CN Losses were identified with **pav** (doi:10.1126/science.abf7117, PMID: 33632895). CN Gains were identified with **pbsv** (https://github.com/PacificBiosciences/pbsv).

\*\*Known KOLF2-C1 CN events are automatically detected by VIA™ in >90% of samples (*Automatic*). When a CN event is not automatically detected, the region is manually reviewed (*By Review*) to confirm presence of the known CN event.

1: CN events originally identified by Aziku et al. doi.org/10.1101/2023.06.26.546614 (bioRxiv preprint)

2: Unpublished CN event identified by C. Beck.

**METHODS:**

The Illumina Infinium Global Diversity Array with Cytogenetics-8 (GDA Cyto) contains ~1.8 million probes throughout the human genome. It provides extensive targeted coverage of more than 4800 key genes across the genome and high exonic coverage in regions of disease relevance, providing highly accurate copy number variation calls and an average resolution of 1.5 Mb. Supplemental probes were selected based on the input from a cytogenetics research consortium and strategically localized near exons, in spatial consideration of existing probe placement, to deliver optimal performance in copy number variation analysis. The current build used is the human genome reference build hg19. Analysis was performed using VIA software v7.0 (Bionano Genomics).

The current microarray platform detects chromosomal imbalances (i.e., gains and losses), but is not designed to specifically detect balanced chromosomal rearrangements, inversions, methylation abnormalities, or aberrations involving genomic regions not targeted by this platform.

| Processing Settings   |  |
|---|--|
| Data Type:  | GxA-Cyto GTC                               |
| Processing Type:  | GDA-Cyto_v8                                |
| Systematic Correction Type:                                       | Quadratic Correction                       |
| Systematic Correction File:                                       | Catlg_ILM_GDACyto_20047166_A1_20210504.txt |
| Allow manual centering:   | TRUE                                       |
| Recenter Probes Type:   | Median                                     |
| Auto Gender Detection Enable:                                     | TRUE                                       |
| Type:   | SNP FASST2 Segmentation                    |
| Max Contiguous Probe Spacing (Kbp):                               | 1000                                       |
| Min number of probes per segment:                                 | 50   |
| Significance Threshold:   | 1.00E-13                                   |
| Segment Boundaries:   | Inner probes                               |
| Amplification (4+:2):   | 0.5  |
| Gain (3:2):   | 0.13                                       |
| Loss (1:2):   | -0.24                                      |
| Homozygous Loss (0:2):  | -0.9                                       |
| M vs M, X/Y Loss (0:1):   | -0.3                                       |
| M vs F, X Loss (0:2):   | -0.9                                       |
| M vs M, X/Y Gain (2:1):   | 0.2  |
| M vs F, X Gain (2:2):   | -0.18                                      |
| M vs M, X/Y Amplification, F vs M, X Gain (3:1):                  | 0.41                                       |
| F vs M, X Amplification (4+:1):                                   | 1.7  |
| Homozygous Frequency Threshold:                                   | 0.96                                       |
| Homozygous Value Threshold:                                       | 0.85                                       |
| Heterozygous Imbalance Threshold:                                 | 0.36                                       |
| Minimum LOH Length (KB):  | 3000                                       |
| Minimum SNP Probe Density (Probes/MB):                            | 500  |
| Robust Variance Sample QC Calculation Percent outliers to remove: | 0.5  |
| Mosaic Labeling Type:   | Aberrant Cell Fraction                     |
| Mosaic Labeling:  | TRUE                                       |
| Mosaic Labeling Label anything lower than (%):                    | 65   |
| Pre-Classification Decision Tree:                                 | KOLF_v1                                    |
| Segmental Duplications Track Version                              | 20210331                                   |
| VIA Report Version  | KOLF_V1                                    |