# Utilization of Twist CNV Backbone Spike-in Panel with exome for cytogenetic research

#P16.070.A

Tina Han, Alessandro Davassi, Esteban Toro

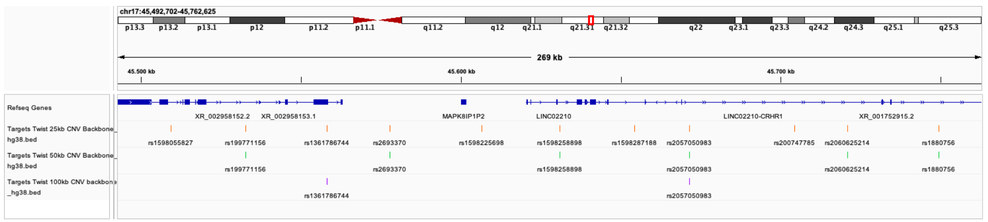
Twist Bioscience, South San Francisco, CA, USA

## Introduction

Chromosomal microarray analysis (CMA), also known as array comparative genomic hybridization (aCGH), has been the gold standard for detection of copy number variation (CNV). Next generation sequencing (NGS) is becoming a standard tool for clinical lab, enabling platform consolidation into an all-in-one NGS assay to replace CMA and aCGH. We describe here a method to leverage Twist Bioscience's double-stranded DNA probes that can be individually tuned to enrich regions of interest with exceptional uniformity, reducing the overall cost of sequencing.

## Material and Methods

With a proprietary algorithm we positioned probes in the intergenic and large intronic regions to create a genome-wide backbone of probes that fill the gaps between exome targets. Twist CNV Backbone Spike-in Panels were designed on common and polymorphic SNPs in multiple populations at 3 different probe densities, ranging from highest resolution to lowest at 25kb, 50kb, and 100kb spacing to accommodate various requirements. These probes target common SNPs polymorphic in multiple populations and are evenly distributed in the intergenic and intronic regions. These panels were designed based on Twist Exome 2.0 plus Comprehensive Exome Spike-in Panel. Here is an example of probe densities at 25kb (top panel in orange), 50kb (middle panel in green), and 100kb (bottom panel in purple):




Poof of concept data includes 2 cohorts: Normal Coriell (n=12) and Coriell Copy Number Variation (CNV) Reference Panel (CNVPANEL01, n=43). CNVPANEL01 was selected based on their relevance for cytogenetic diagnosis, the DNA samples in this panel have each been extracted from a cell line harboring a clinically significant chromosomal aberration. These samples have been identified and are being characterized in a collaborative effort involving the CDC's Genetic Testing Reference Materials Coordination Program (Get-RM), clinical cytogeneticists, microarray suppliers and the NIGMS Human Genetic Cell Repository. Data from the multiple laboratories involved in the project will be made available on the NIGMS Repository catalog web site. To date, Coriell has genotyped all samples on the Affymetrix Genome-Wide Human SNP Array 6.0 platform and performed G-banded karyotyping analysis and, in many cases, fluorescence in situ hybridization (FISH). The ISCN for each sample is listed, describing the results of G-banding and microarray analyses, and, when available, FISH. Additionally, Affymetrix Genome-Wide Human SNP Array 6.0 genotyping data for these samples is available from dbGaP (dbGaP Study Accession: phs000269.v1.p1).

All samples were prepared with Twist Library Preparation Enzymatic Fragmentation Kit 2.0, Twist Universal Adapter System, and Twist Target Enrichment Standard Hybridization v2. Samples were sequenced on Illumina NovaSeq 6000 as paired-end 150 reads (PE150). Raw sequencing data as fastq file were demultiplexed using UDIs with Illumina bcl2fastq (2.20) and 1 mismatch was allowed. Secondary and tertiary analyses were performed on enGenome's eVai platform.

## Results - Variant Calling

Here we present data that illustrates target enrichment efficiency using Twist's complete workflow through diverse samples that harbor a clinically significant chromosomal aberration. We demonstrate the 100% concordance of CNVs called between NGS and aCGH data. Moreover, we show how aCGH data output could be replicated using this NGS solution.

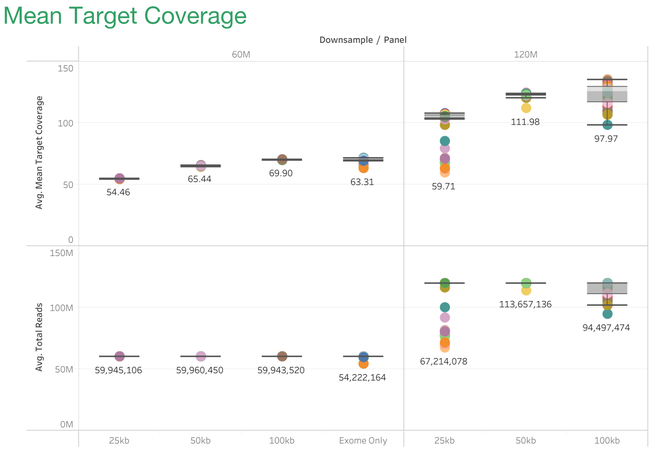
The average number of SNVs, INDELs, and CNVs called at each probe density by combining the CNV Backbone Spike-in Panel and Exome 2.0 plus Comprehensive Exome Spike-in Panel:

| CNV Backbone Spike-in Panel: | 25 kb | 50 kb | 100 kb |
|---|---|---|---|
| Panel Design Size (Mb) | 8.32 | 3.33 | 1.39 |
| SNVs & INDELs Called | 265138 | 207765 | 181956 |
| ALL CNVs Called | 472 | 446 | 411 |
| CNVs less than 50 kb called | 417 | 385 | 361 |
| Mean Target Coverage | 150 | 171 | 161 |

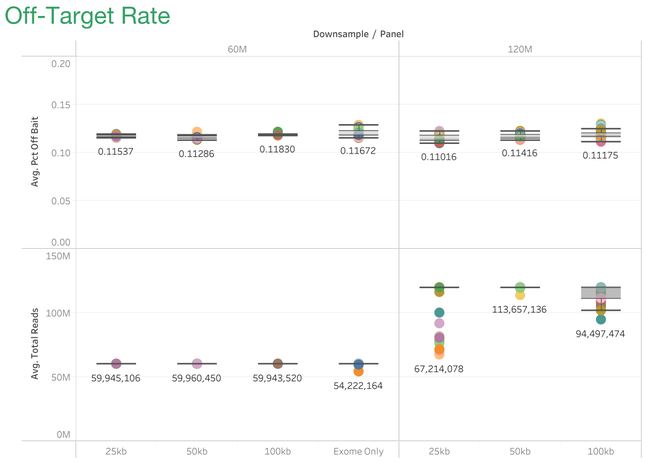
## Results: Capture Performance

Here we present data that illustrates exceptional capture performance using Twist's complete workflow either downsampled as 60M or 120M reads per sample. With 60M reads, 60x mean target coverage, 89% on-target rate, fold-80 base penalty at 1.3, and >96% and >91% bases were covered at least 20x and 30x, respectively. Duplication rate was lower than 10%. Data did not exhibit any obvious AT- or GC-dropout.

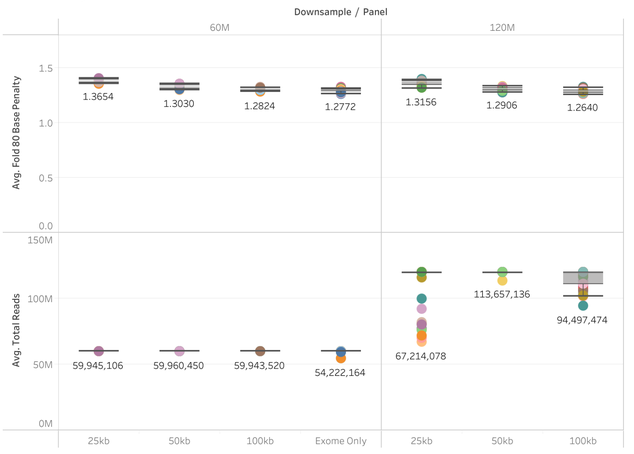
### Mean Target Coverage



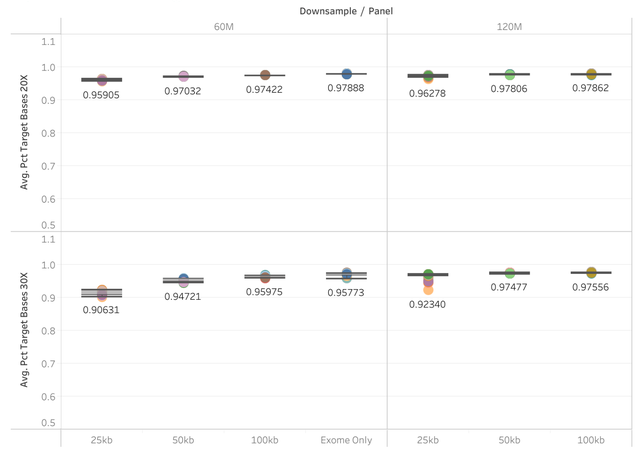

### Off-Target Rate



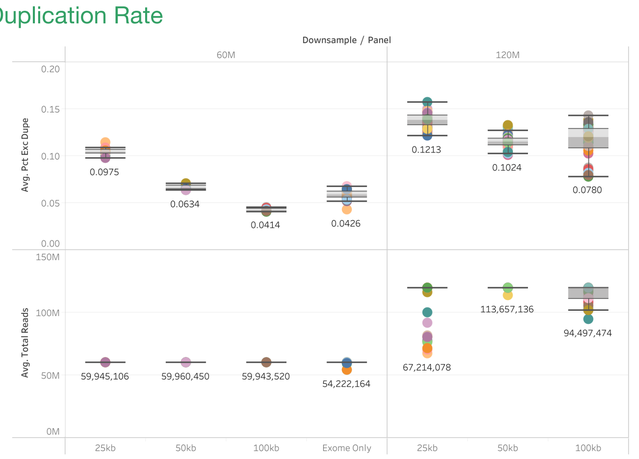

### Uniformity as Fold-80 Base Penalty



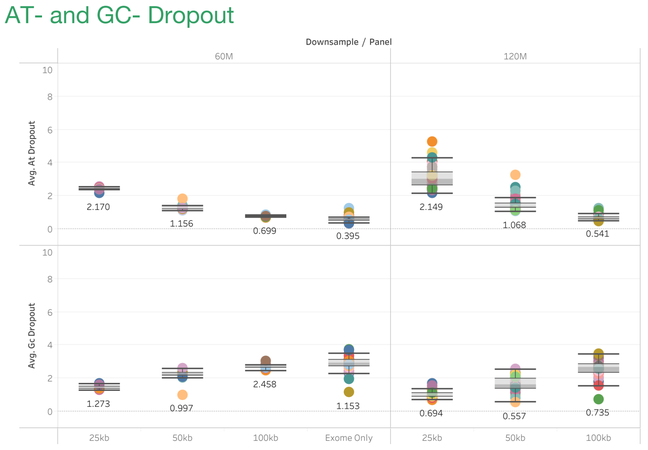

### Depth of Coverage as % Target Bases Covered ≥ 20X and 30x




### Duplication Rate



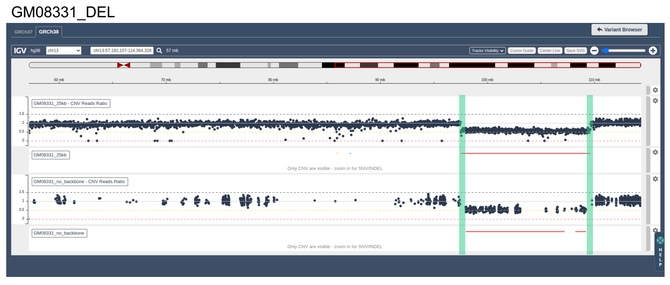

### AT- and GC- Dropout




## Results: CNV Visualization *with* or *without* 25kb CNV Backbone Spike-in Panel

GM09216_DEL

GM08331_DEL



GM06870_DUP

Three examples showing CNV Backbone Spike-in Panel could better determine the breakpoint of deletion or duplication events than Exome by itself. The coverage was expressed as reads-ratio. In each example, the **upper** track is CNV Backbone Spik-in Panel with Twist Exome 2.0 plus Comprehensive Exome Spike-in Panel. The **lower** track is Twist Exome 2.0 plus Comprehensive Exome Spike-in Panel only. The green bars highlighted the breakpoints.

## Conclusions

The data demonstrated how Twist's all-in-one assay lets researchers sequence more deeply and more efficiently into each sample. The data demonstrated the combination of Twist CNV Backbone Spike-in Panels with Twist Exome Panels enables detection of genome-wide detection of CNVs or LOH, on top of SNVs and small INDELs that come with Twist Exome 2.0 plus Comprehensive Exome Spike-in Panel. Further data also demonstrated the ability to detect large CNV event that normally requires CMA or array CGH.

## Conflict of Interest Statement