Introducing new technology for high performance whole exome sequencing with unparalleled hybridization time flexibility and capability for low to high capture-input

6725

Daniel Saatman[1]*, Jinglie Zhou[1], Ekaterina Star[1], Roshani Codipilly[1], Alexandra Wang[1]

[1] Research & Development, Integrated DNA Technologies, Redwood City, CA

* Corresponding author: dsaatman@idtdna.com

Introduction

Whole exome sequencing (WES) via hybridization capture is a targeted sequencing approach to specifically enrich the protein-coding regions of the human genome. The human exome comprises ~1% of the entire genome but contains the majority of relevant variants for researchers. Targeting this region saves on sequencing costs and data storage while also allowing greater sequencing depth of regions of interest. However, current hybridization capture workflows are challenging to execute in the laboratory due to high-temperature sample and reagent handling, pre-heating of multiple buffers, long hybridizations, and narrow library input ranges. In this study we present exome sequencing data using the Integrated DNA Technologies (IDT) novel xGen™ Hybridization and Wash v3 kit (xGen v3), which addresses these common pain-points while generating excellent performance with library input and hyb time flexibility.

Methods

To evaluate targeted sequencing performance, we first generated libraries from 25 ng of 150 bp-sheared Coriell NA12878 gDNA using the IDT xGen cfDNA & FFPE DNA Library Prep Kit. This was followed by hybridization capture with the IDT xGen Exome Hyb Panel v2 (Exome v2) across a wide range of sample inputs and hybridization times. We used low to high total input—1-plex 100 ng to 12-plex 6000 ng—and hybridized for as short as one hour up to overnight using the novel xGen v3 kit. This performance is benchmarked with 1-plex 500 ng captures using the commercially available xGen Hybridization and Wash v2 kit (xGen v2). Post capture libraries were sequenced on the Illumina® NextSeq2000 sequencer.

Versatile and simplified workflow

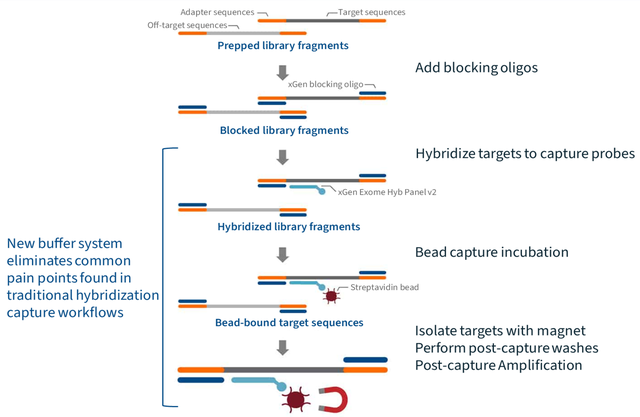


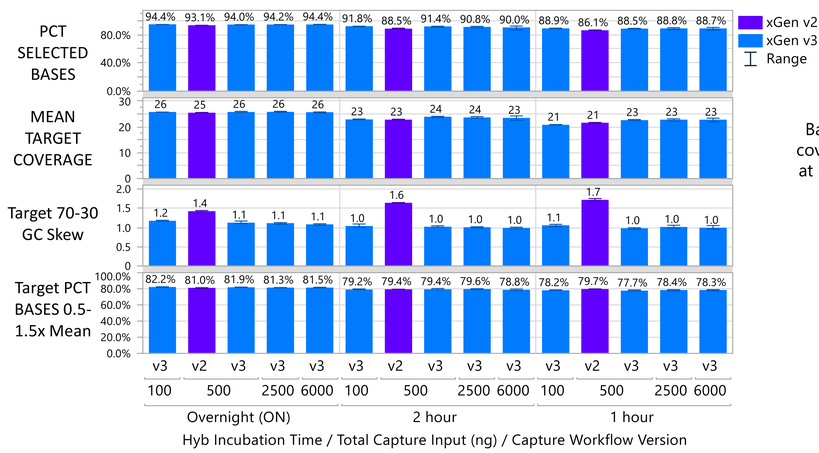
Figure 1. A depiction of the traditional hybridization capture workflow, a targeted sequencing approach to enrich and amplify DNA regions of interest.

	xGen v2	xGen v3
Minimum recommended library input	500 ng	100 ng
Maximum recommended library input	6000 ng	6000 ng
Minimum recommended hybridization time	4 hours	1 hour
Wash buffers to mix/dilute	6	1
Capture and wash hands-on steps	17	9
Hot buffer handling steps	7	0
Hot buffers to pre-heat	3	0
Post-capture sample washes	6	3
Lab & user-dependent: estimation for 32 samples Manual Processing (min)	117	82

Table 1. Comparison of the current and new IDT xGen Hybridization-Capture reagents and workflow. The xGen v3 workflow features a shortened protocol with fewer hands-on steps, room temperature buffers with no pre-heating, and more flexible input and hybridization durations.

Comparison of exome sequencing for multiplexed samples from 1-hour to overnight hybridizations

Good capture performance from low to high capture input, and from short to long hybridizations



Improved coverage at all hybridization times for the xGen v3 workflow

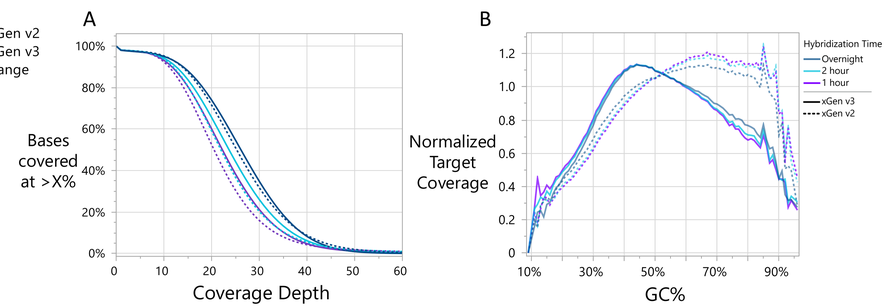


Figure 4. Target coverage depth and GC distribution for xGen v3 captures compared to xGen v2. (**A**) Comparison of both workflows using 500 ng capture input at all hybridization times. The xGen v3 workflow attains higher coverage than the equivalent v2 workflow condition at 20 M total reads per sample. (**B**) Normalized target coverage by GC content shows balanced GC distribution for the xGen v3 workflow compared to xGen v2 by comparing 500 ng singleplex captures at different hybridization times.

Figure 2. Comparison of targeted sequencing metrics across different capture inputs, hybridization times, and capture workflows. Libraries were constructed from 25 ng of Coriell gDNA and enriched using Exome v2 at the following plexity: 1-plex (100 ng and 500 ng input), 5-plex (2500 ng), and 12-plex (6000 ng). For comparison, 500 ng input 1-plex captures using the xGen v2 kit are also shown. Three technical replicates were used for the multiplexed conditions and the 100 ng 1-hour hyb; the rest of the singleplex conditions used four technical replicates. Subsampled to 20 M total reads per sample and analyzed using Picard metrics. Similar on-target rate (PCT SELECTED BASES), mean target coverage, GC skew, and coverage uniformity between all hybridization times and inputs for xGen v3 with better performance compared to xGen v2.

Multiplexed samples have high coverage and unbiased GC distribution

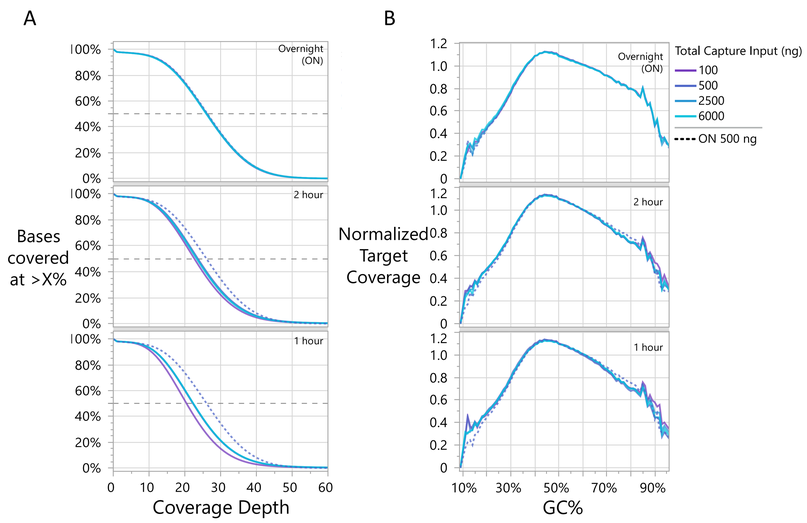
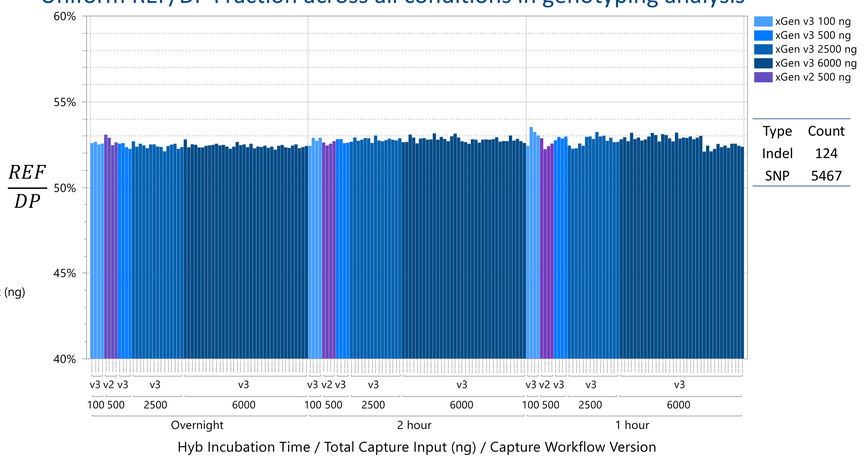


Figure 3. Target coverage depth and GC distribution for xGen v3 captures. (**A**) Equivalent coverage for all inputs / plexities for overnight hybridizations, with a small drop in coverage for low input samples at shorter hybridizations. Subsampled to 20 M total reads per sample. (**B**) Normalized target coverage by GC content shows negligible impact from total capture input or shorter hybridization.

Uniform REF/DP Fraction across all conditions in genotyping analysis



Type	Count
Indel	124
SNP	5467

Figure 5. REF/DP fraction (reference allele depth divided by total depth) for various capture conditions. Libraries were constructed from 25 ng of Coriell gDNA and enriched using Exome v2 at the following plexity: 1-plex (100 ng and 500 ng input), 5-plex (2500 ng), and 12-plex (6000 ng). For comparison, 500 ng input 1-plex captures using the xGen v2 kit are also shown. REF and ALT calls made using GATK HaplotypeCaller with default parameters. All technical replicates are shown side-by-side. No overall difference in reference allelic fraction (REF/DP) across all conditions tested.

Conclusion

- The novel xGen v3 workflow achieved similar targeting, coverage, uniformity, and GC-balance across all samples with minimal impact from low inputs and short hybridizations.
- We calculated the reference allelic fraction and found no considerable differences between capture conditions, indicating the potential of this kit for genotyping applications with low-input samples and/or short hybridizations.
- The flexibility in hybridization duration and library input along with the streamlined workflow allows researchers to meet diverse experimental demands and effectively capture the human exome.

Daniel Saatman
Research Scientist
Integrated DNA Technologies
dsaatman@idtdna.com