Improved Sensitivity and Rapid Confirmation of Variants via Orthogonal Sequencing of Exomes

John F Thompson, Niru Chennagiri, Alex Frieden, Daniel Lieber, Edgardo Lopez, Tristen Ross, Eric White, Timothy Yu Claritas Genomics, Cambridge, MA 02139 USA

CLARITAS GENOMICS

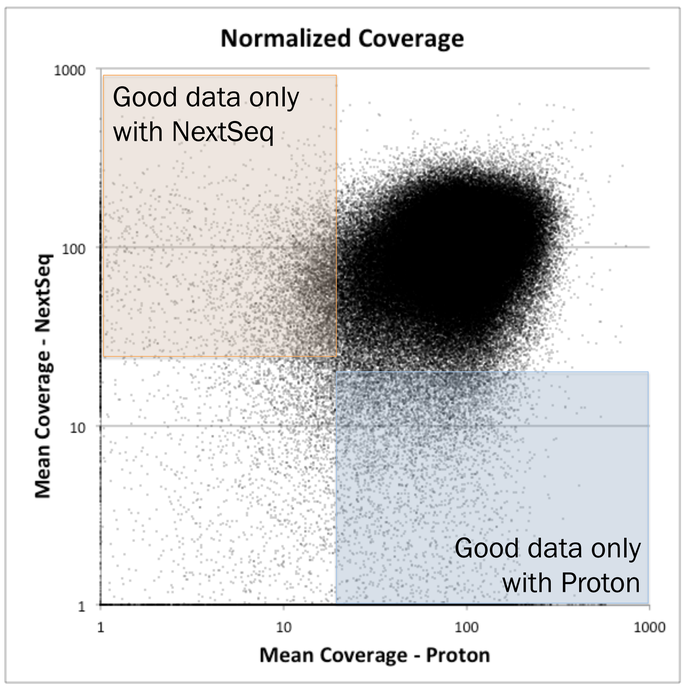
Abstract

Whole exome and genome sequencing studies have revolutionized the diagnosis of genetic disorders by providing information about pathogenic variants. However, each next-generation sequencing (NGS) platform has limitations in terms of coverage and error profiles. In addition, the large number of variants identified in each patient makes confirmation and interpretation of the data very challenging. ACMG practice guidelines recommend that "all disease-focused and/or diagnostic testing include confirmation of the final result using a companion technology", a daunting task given the number of variants involved. Thus, the ability to confirm many variants quickly is critical.

To maximize sensitivity and to minimize the time needed, Claritas Genomics has developed an approach using complementary and orthogonal sequencing technologies to analyze variants from whole exome sequences. Interpretation is restricted to phenotypically-defined, clinically relevant regions of interest (ROI). Each individual's exome is sequenced using two separate DNA capture and sequencing methods. DNA is captured using a hybridization-based approach and sequenced using Illumina NGS. In parallel, DNA is captured using an amplification-based approach and sequenced using Ion Torrent NGS. Because independent and complementary technologies are used, the need for Sanger sequencing for confirmation is nearly eliminated. With the whole exome, ~90% of the variants are detected on both platforms and thus are orthogonally confirmed. When this analysis is carried out on DNA with an extensive truth set (NA12878), 100% of the orthogonally confirmed variants are true positives. In addition, 8% of the true variants are found on only one platform and not called on the other due to insufficient coverage. The use of two capture methods and sequencing platforms extends the region over which variants can be detected so the region examined is more than either platform alone is capable. This improves sensitivity of detection, particularly with insertions and deletions because both platforms miss a significant number.

Because the majority of variants are orthogonally confirmed and interpreted for pathogenicity promptly, there is no delay for Sanger confirmation and a rapid report can be provided to the physician. This provides the potential for physicians to act quickly on information with patients benefitting from timely feedback.

Platform-Specific Coverage



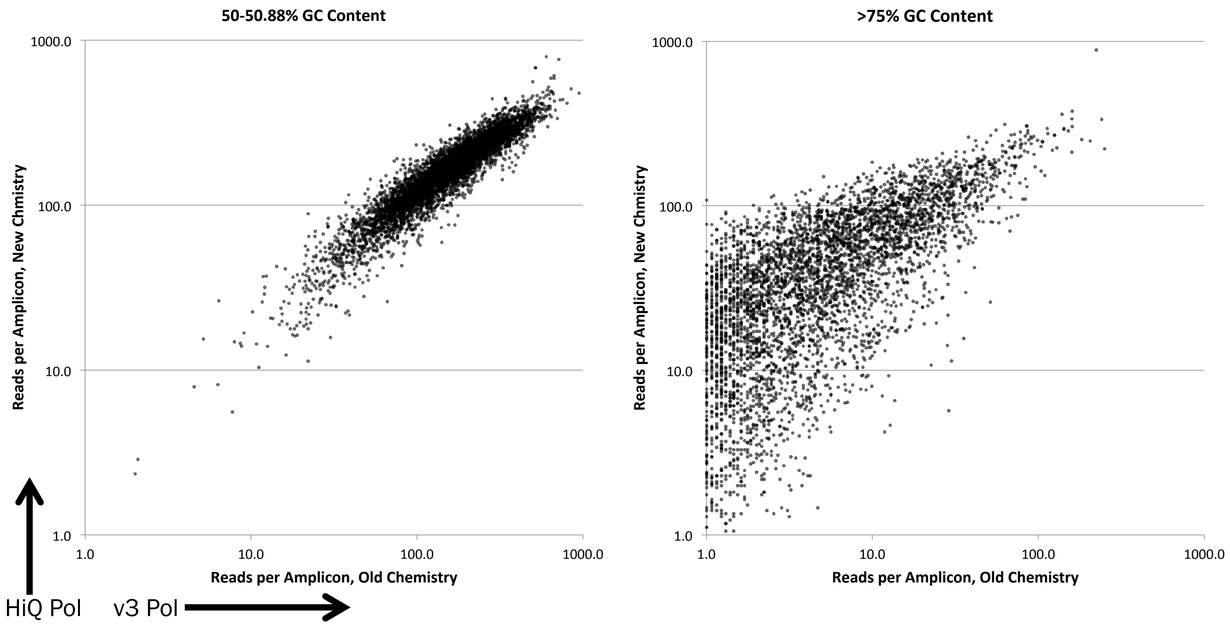
About 90% of the 187,475 exons examined have >20x coverage on both platforms when normalized to 100x coverage on each. These regions have the potential for being orthogonally confirmed by the complementary NGS platforms. About 2.3% of exons are covered at less than 20x on both platforms. Some fraction of these exons have the potential for variants being detected. The remaining exons are covered well on one platform but not the other. This represents an increase in sensitivity of 3-4% when both platforms are used compared to either platform alone.

Complementary Platforms Improve Sensitivity

>97% targeted exome regions are covered >20x via NextSeq or Proton. Apparent sensitivity in the table below is higher because this is for the NIST reference region only which omits many difficult to sequence regions.

Platform	Sensitivity	PPV
Illumina NextSeq	97.6%	99.6%
Ion Proton	94.7%	99.6%
Combined	99.0%	99.5%

New/Old AmpliSeq Exome Coverage: *TS4.0 vs TS4.4/HiQ*



Medium GC: Good coverage that becomes more uniform with HiQ

High GC: Poor coverage that becomes better with HiQ

Amplicon Coverage

The number of reads in each amplicon for both old and new Proton methods are shown. The new methods have >80% fewer amplicons with very high coverage (>500 reads) and >50% fewer amplicons with very low coverage (<20 reads). Thus, many more amplicons have a desirable coverage range (20-500 reads) with the new chemistry (97.3%) relative to the old chemistry (93.4%).

vTSv4.0 TSv4.4>	>500	100-500	20-100	10-20	<10	Total TS4.0
>500	331	2110	0	0	0	2441
100-500	99	156,526	11,342	3	1	167,971
20-100	0	29,991	75,678	525	32	106,226
10-20	0	403	5686	1532	96	7717
<10	0	119	3915	2077	3192	9303
Total TS4.4	430	189,149	96,621	4137	3321	293658

Claritas Clinical Exome Variants Confirmed Orthogonally, PPV=100%

Orthogonal Sequencing using independent enrichment and sequencing methods leads to 100% PPV with NA12878 for variants found on both platforms. These results are for protein coding exons and 10 bp of adjacent sequence that are called in the NIST reference.

Category	Type	Total	FP	TP	PPV	% Of Total
Orthogonally Confirmed (High quality calls on both platforms)	All	15437	0	15437	100.0%	88.6%
	InDel	208	0	208	100.0%	
	SNV	15229	0	15229	100.0%	
Reliable (High quality call on Proton, NoPass on Illumina)	All	431	0	431	100.0%	2.5%
	InDel	16	0	16	100.0%	
	SNV	415	0	415	100.0%	
Likely True Positive (Called on Illumina, no coverage/ref on Proton or called on Proton, no coverage on Illumina)	All	1416	80	1336	94.4%	8.1%
	InDel	211	16	195	92.4%	
	SNV	1205	64	1141	94.7%	
Likely False Positive (NoPass call on Illumina, reference on Proton or call on Proton, reference on Illumina)	All	139	112	27	19.4%	0.8%
	InDel	42	31	11	26.2%	
	SNV	97	81	16	16.5%	

Conclusions

Using Orthogonal Sequencing methods improves sensitivity and allows variants to be rapidly called for improved clinical sequencing performance. The improved Ion Torrent sequencing quality matches well with the independent Illumina methods to provide comprehensive exome coverage.

See other Claritas Genomics posters:
1619, 1933, 2070, 2071, 2085

CLARITAS GENOMICS INC.
99 Erie Street | Cambridge, MA 02139

clientservices@claritasgenomics.com
p: 617.553.5880 p: 855.373.9003 f: 617.553.5842