High Accuracy and Expanded Yield from Next-Generation Testing of Multiple Cancer Risk Genes

Stephen Lincoln[1], Allison Kurian[2], Kristen Shannon[3], Michael Anderson[1], Yuya Kobayashi[1], Geoffrey Nilsen[1], Kevin Jacobs[1], Leif Ellisen[3], James Ford[2], and Federico Monzon[1]

[1]Invitae, San Francisco, CA [2] Stanford University Medical Center, Palo Alto, CA [3]Massachusetts General Hospital, Boston, MA

Background

Gene panels assayed using next-generation sequencing (NGS) technologies are moving from research labs into clinical use, with the potential to provide improved diagnostic yield quickly and at low cost. It is important to understand the performance and clinical implications of these emerging tests by comparison with traditional methods and sequential testing.

Methods

Germ-line DNA samples were collected from multiple sources: (a) Research specimens from consented patients referred for HBOC (Hereditary Breast and Ovarian Cancer) testing at two major cancer centers. These patients were independently tested for BRCA1 and/or BRCA2 using traditional methods as part of their care. (b) Reference samples from public biobanks, chosen both to have broad coverage of genes previously assayed and also enriched for pathogenic variants known to be technically challenging for NGS; (c) A small number of reference samples for which high-quality whole-genome sequence data was publicly available.

Samples were blinded and tested on a custom 220 gene NGS panel including 29 hereditary cancer risk genes. Laboratory protocols and custom bioinformatics methods were implemented to detect a spectrum of DNA variation types from the NGS data alone: These include single nucleotide variants (SNVs), small and large sequence insertions and deletions (indels), and small and large deletion and duplication (del/dup) events. NGS results were compared to the previous testing, and any analytical disagreements were send to a third-party lab to arbitrate, also using traditional methods.

Sequenced to Date	Analyzed and Unblinded In This Report	Description
241	187	MGH HBOC Retrospective Cohort
282	0	MGH HBOC Prospective Cohort
413	413	Stanford HBOC Cohorts
105	105	Public Reference Materials*
7	7	Public Well-Characterized Genomes
1048	**712**	**Total to Date**

* Mutations In hereditary cancer genes: BRCA1, BRCA2, APC, CDKN2A, MEN1, MLH1, MSH2, PTEN, and RET. Other genes: CFTR, ASPA, FANCC, HEXA, GJB1, IKBKAP, HFE, Factor 2, and Factor 5. Large cytogenetic changes associated with Smith-Magenis, Down, Cri-du-chat, Wolf-Hirschhorn, Angelman, and other syndromes are also present.

In 712+105 Clinical and RM Samples	In 7 Whole Genomes (Panel Genes Only)	DNA Variant Type
123	2020	Single nucleotide variants
125	50	Small indels and del-ins (≤10 bp)
4		Large indels (>10 bp and <1 full exon)
20		Small del/dup events (1 exon to 1 gene)
22		Cytogenetic scale copy number changes
23		Variants in/near specific homopolymer tracts*

Acknowledgements: We thank John Major, Curtis Kautzer, Jon Sorenson, Eric Oliveras, Shera Kash, Shan Yang, Reece Hart, Heather Trumbower, Emily Hare, and Michele Cargill for their substantial contributions. We thank Randy Scott and Jill Hagenkord for their inspiration.

Analytic Results

Of the thousands examined, only 5 discordant variants existed between NGS and the previous testing. In all 5 cases the NGS result was shown to be correct by the third-party lab. Thus, 100% analytical sensitivity and 100% analytical specificity was observed between NGS and all of the confirmed traditional results to date.

Net Result		In Clinical and RM Samples	In Genome Sequences
Error type	Number observed	Gene x individual combinations with data from 2 labs to compare	Base-pairs or variants in 220 targeted genes
NGS False Positives	**0**	**1193** gene sequence tests **439** gene del/dup tests	**4,341,624 bp** tested in total
NGS False Negatives	**0**	**314** known positive variants within test's analytical range	**2,070** known positive variants

Some classes of variation posed specific challenges for NGS and led to the development of new bioinformatic solutions:

BRCA2: c.9203del126
Deletion spanning part of but not all of a full exon, into neighborig intron past our pulldown target (or other's typical PCR primer) sites

PLP1: del exon 3-4
Deletion spanning parts, but not all, of 2 neighboring exons and all of the intervening intron

↓

Solution: Split-read algorithm implemented detect these types of event

MSH2: c.942+3A>T
Known pathogenic splice changing mutation next to a 25bp intronic poly-A tract, difficult to sequence on any platform

↓

Solution: PolyMNP algorithm (initially developed to call poly-TG/poly-T site in CFTR) modified to call this site

Clinical Results

Clinical interpretations were also highly concordant, although with a slightly higher rate of VUS (Variants of Unknown Significance) in our BRCA1/2 data (6% vs. 4%). This was expected: We followed ACMG classification guidelines and used only public, and not proprietary reference databases, unlike the traditional BRCA1/2 tests we compared against. We also did not have access to family members in this study to help resolve VUS by segregation analysis.

Exactly One Interpretation Difference for a Pathogenic Variant:

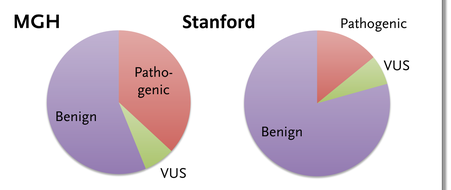
BRCA1: c.4986+3G>C in Stanford case 6396

- Reported in the literature as pathogenic but without evidence meeting ACMG criteria (PubMed:12491499)
- Weak evidence of splicing change using prediction algorithms
- Diagnostically reported as as Pathogenic by the previous lab, but without stated evidence in diagnostic report

More HBOC cases saw a reportable finding with the 29 gene panel vs. BRCA1/2 alone. Some, but not all, of these findings were clinically actionable. By virtue of testing more genes, additional VUS were also uncovered, with about 50% of patients having a VUS in one or more of the 29 genes. This too was expected as some genes in the panel have much less data available in the scientific and medical literature than others.

BRCA1 and BRCA2

	MGH	Stanford
Pathogenic or Likely pathogenic	69	58
Negative	105	326
Uncertain Significance	13	27
Total Cases	187	413



Diagnostic Genes (#observed in 198 cases)		Research Genes
MLH1 (1)	MUTYH (5)	SLX4 (2)
CDKN2A (1)	NBN (2)	BLM (1)
CDH1 (1)	ATM (2)	PRSS1 (1)

First 198 Stanford Cases: Kurian et al. ASCO Breast, 2013

Conclusion

NGS can produce high quality clinical results comparable with traditional testing. Additional useful information can be produced by testing additional genes but with a corresponding increase in VUS. The high NGS specificity observed suggests that validating clinical variants in patients using an orthogonal platform may, over time and in certain circumstances, become less necessary.