Comparing Gene Panel and Augmented Exome Tests Using a Gold-Standard Dataset

#2090T

G. Chandratillake, G. Bartha, A. Patwardhan, D. Church, S. Garcia, R. Chen, J. West.
Personalis, Inc. 1330 O'Brien Drive, Menlo Park, CA 94025.

Contact:
gemma.c@personalis.com

Introduction

Assessing the sensitivity and specificity of next-generation sequencing (NGS) based tests poses a challenge as it requires reference samples with known variants in all genomic regions that are included in the assay.

The Genetic Testing Reference Materials Coordination Program (GeT-RM) is a CDC initiative that aims to improve the availability of appropriate and characterized reference materials for genetic testing. We compared the gene panel test data submitted to the GeT-RM by clinical labs for reference sample NA12878, to the gold-standard data produced for this sample by the NIST hosted Genomes in a Bottle Consortium (GIAB). We also compared the results to those achieved with an augmented exome sequencing assay developed by our lab: the ACE Clinical Exome Test (Patwardhan et al. 2015).

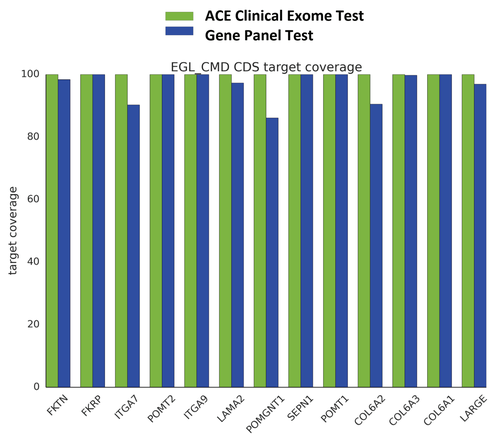
Results

1) Assessment of gene targeting by gene panels tests and ACE Clinical Exome Test

Figure 1.

To assess the potential of the tests to detect variants in the genes included in the panels, we determined the coverage of the coding sequence (CDS) by the gene panel tests and by the ACE Clinical Exome Test (as defined by the targets file), using BEDtools. The union of all CDS sequences in NCBI annotation 105 was used to define the CDS. While targeting of the CDS for the panel genes was often incomplete, targeting of those genes was more comprehensive for the ACE Clinical Exome Test. It is possible that missing CDS for the panels is filled in with Sanger sequencing, however such data is not available so could not be assessed.

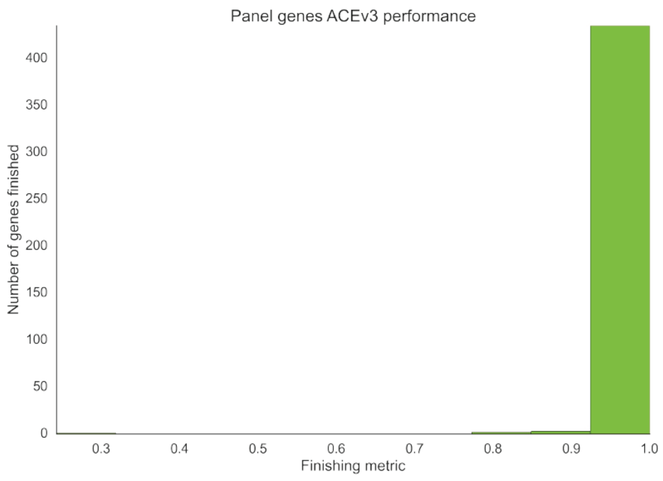
Coverage of a representative panel shown in detail below. Other panels shown on right.

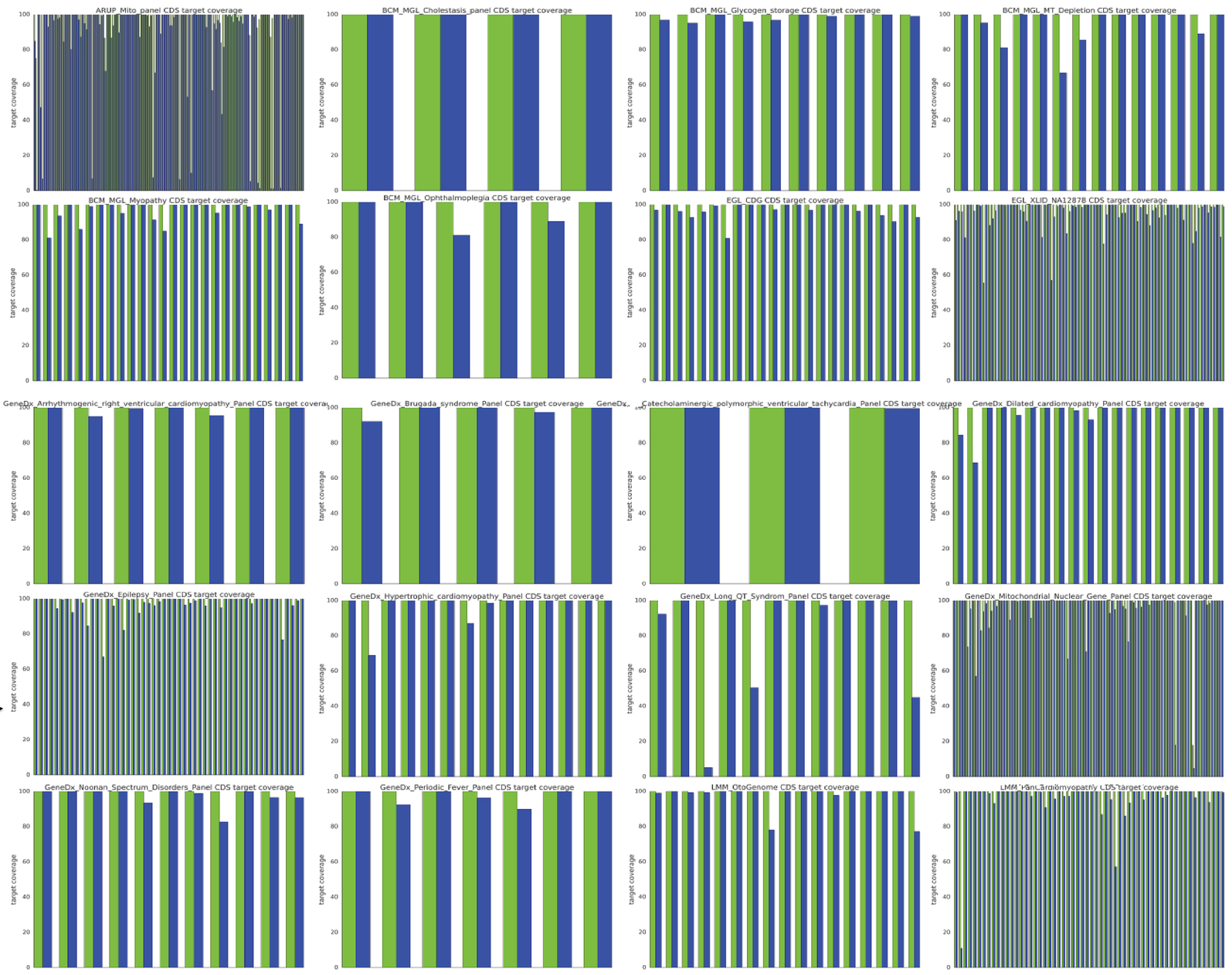


2) Assessment of gene finishing of panel genes by ACE Clinical Exome

Figure 2.

For the >400 genes included in the assessed gene panels, the coverage of those genes by the ACE Clinical Exome Test was assessed using the metric of "gene finishing," a determination of the proportion of coding bases of the gene covered at >20x. Almost all of these clinically-relevant genes were 100% finished using the ACE Clinical Exome Test.





3) Assessment of accuracy of gene panel tests and ACE Clinical Exome Test

Table 1.

The panel test target regions were intersected with the high confidence regions from the GIAB data for NA12878 to create a truth call set. Panel test data deposited in GeT-RM and ACE Clinical Exome data were assessed for measures of accuracy relative to this gold-standard dataset.

1: ARUP_Mito_panel*
2: BCM_MGL_Cholestasis_panel#
3: BCM_MGL_Glycogen_storage
4: BCM_MGL_MT_Depletion
5: BCM_MGL_Myopathy
6: BCM_MGL_Ophthalmoplegia
7: EGL_CDG
8: EGL_CMD
9: EGL_XLID_NA12878
10: GeneDx_Arrhythmogenic_right_ ventricular_ cardiomyopathy_Panel
11: GeneDx_Brugada_syndrome_Panel
12: GeneDx_Catecholaminergic_ polymorphic_ ventricular_tachycardia_Panel
13: GeneDx_Dilated_cardiomyopathy_Panel
14: GeneDx_Epilepsy_Panel
15: GeneDx_Hypertrophic_cardiomyopathy_Panel
16: GeneDx_Long_QT_Syndrom_Panel
17: GeneDx_Mitochondrial_Nuclear_Gene_Panel
18: GeneDx_Noonan_Spectrum_Disorders_Panel
19: GeneDx_Periodic_Fever_Panel
20: LMM_OtoGenome
21: LMM_PanCardiomyopathy

* Panel target includes large amount of non-coding sequence hence low sensitivity of ACE (ACE sensitivity over ACE target is 98.9% for SNVs, 92.7% for indels.)

For this gene panel test data, there appears to have been a transcription error (e.g. c. HGVS alleles rather than genomic alleles) since, in many cases where the gene is on the negative strand, if the ALT bases is complemented, the call is correct.

Panel		TPs		FPs		FNs		Sens. %		FDR %		PPV %	
		Panel	ACE	Panel	ACE	Panel	ACE	Panel	ACE	Panel	ACE	Panel	ACE
1*	SNV	258	59	565	1	287	*486	47.3	*10.8	68.7	1.7	31.3	98.3
	indel	7	57	5	2	63	*13	10.0	*81.4	41.7	3.4	58.3	96.6
2#	SNV	1	10	8	0	9	0	10.0	100.0	88.9	0.0	11.1	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
3	SNV	16	21	4	0	5	0	76.2	100.0	20.0	0.0	80.0	100.0
	indel	0	1	1	0	1	0	0.0	100.0	100.0	0.0	0.0	100.0
4	SNV	14	15	0	0	1	0	93.3	100.0	0.0	0.0	100.0	100.0
	indel	1	1	0	0	0	0	100.0	100.0	0.0	0.0	100.0	100.0
5	SNV	53	59	0	0	6	0	89.8	100.0	0.0	0.0	100.0	100.0
	indel	1	2	1	0	1	0	50.0	100.0	50.0	0.0	50.0	100.0
6	SNV	7	10	3	0	3	0	70.0	100.0	30.0	0.0	70.0	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
7	SNV	25	26	1	0	6	5	80.6	83.9	3.8	0.0	96.2	100.0
	indel	0	1	0	0	2	1	0.0	50.0	N/A	0.0	N/A	100.0
8	SNV	49	54	7	0	5	0	90.7	100.0	12.5	0.0	87.5	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
9	SNV	30	31	76	0	1	0	96.8	100.0	71.7	0.0	28.3	100.0
	indel	0	1	0	0	1	0	0.0	100.0	N/A	0.0	N/A	100.0
10	SNV	21	22	0	0	1	0	95.5	100.0	0.0	0.0	100.0	100.0
	indel	0	1	3	0	1	0	0.0	100.0	100.0	0.0	0.0	100.0
11	SNV	7	7	0	0	0	0	100.0	100.0	0.0	0.0	100.0	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
12	SNV	15	16	0	0	1	0	93.8	100.0	0.0	0.0	100.0	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
13	SNV	17	19	0	0	2	0	89.5	100.0	0.0	0.0	100.0	100.0
	indel	1	2	0	0	1	0	50.0	100.0	0.0	0.0	100.0	100.0
14	SNV	28	34	0	0	6	0	82.4	100.0	0.0	0.0	100.0	100.0
	indel	2	2	0	0	0	0	100.0	100.0	0.0	0.0	100.0	100.0
15	SNV	10	10	0	0	0	0	100.0	100.0	0.0	0.0	100.0	100.0
	indel	1	2	0	0	1	0	50.0	100.0	0.0	0.0	100.0	100.0
16	SNV	10	10	0	0	0	0	100.0	100.0	0.0	0.0	100.0	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
17	SNV	86	91	5	0	5	0	94.5	100.0	5.5	0.0	94.5	100.0
	indel	4	8	1	0	4	0	50.0	100.0	20.0	0.0	80.0	100.0
18	SNV	2	2	0	0	0	0	100.0	100.0	0.0	0.0	100.0	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
19	SNV	12	16	3	0	4	0	75.0	100.0	20.0	0.0	80.0	100.0
	indel	0	0	0	0	0	0	N/A	N/A	N/A	N/A	N/A	N/A
20	SNV	48	49	2	0	1	0	98.0	100.0	4.0	0.0	96.0	100.0
	indel	0	1	1	0	1	0	0.0	100.0	100.0	0.0	0.0	100.0
21	SNV	88	88	22	0	0	0	100.0	100.0	20.0	0.0	80.0	100.0
	indel	0	4	3	0	4	0	0.0	100.0	100.0	0.0	0.0	100.0
Total ex. 1 & 2	SNV	538	580	123	0	47	5	92.0	99.1	18.6	0.0	81.4	100.0
	indel	10	26	10	0	17	1	37.0	96.3	50.0	0.0	50.0	100.0

Methods

Panel selection

The data for all gene panel tests pertaining to Mendelian disease diagnosis deposited in GeT-RM for reference sample NA12878, and for which GIAB data was available, were included in this analysis. Due to lack of GIAB reference data, two tests comprising entirely of the mitochondrial genome (MT) were excluded (GeneDX_NA12878_Mitochondrial_Genome & BCM_MGL_NA12878_Whole_Mitochondrial_Genome) and MT data was excluded from two further tests (GeneDX_NA12878_Dilated_cardiomyopathy_Panel & LMM_NA12878_OtoGenome). Two panels relating to somatic variant detection and drug metabolism were also excluded (GPS_WUSTL_Cancer_panel & GPS_WUSTL_Drug_metabolism_panel).

Conclusion

The ACE Clinical Exome Test Exhibits Equivalent or Superior Performance to many NGS Gene Panel Tests

References

1. Patwardhan et al. Genome Medicine 2015, 7:71
2. GeT-RM web site at NCBI: http://www.ncbi.nlm.nih.gov/variation/tools/get-rm/
3. GIAB reference: http://www.ncbi.nlm.nih.gov/pubmed/24531798

Personalis®
Pioneering Genome-Guided Medicine