Tracking Copy Number Variant Data with Geneticist Assistant[®] NGS Workbench

March 2020

Megan McCluskey, Edward Bouton, Zhaorong Ma, Lidong Luo

Introduction

The detection of various types of sequence variants are important for the study and diagnosis of many diseases, with Next Generation Sequencing technologies providing an excellent tool for the research and clinical detection of variants, including SNVs and small indels (1, 2). New techniques have emerged for analyzing copy number variations as well (3, 4). As these results increasingly become available, tracking detected CNV events is important for establishing and maintaining a knowledge base of these variants.

Geneticist Assistant NGS Workbench provides a database solution for reviewing and tracking variants from targeted panels to whole exome sequencing. After reviewing variants, a pathogenicity can be assigned – Deleterious, Likely Deleterious, Likely Benign, Benign, or Unknown, or using customizable terminology. The variant pathogenicity is saved to the database and shown for future samples containing the variant. Geneticist Assistant also tracks how many times a variant has occurred. In addition to importing SNVs and indels, Geneticist Assistant can also import copy number variants. Cross-platform data is compatible for import using any VCF, GVF, and optionally BAM files.

Sample Variant Data			Copy Number Variation Data		Sample Da	ata Quality Me	tric Report						
Copy Number Variation of "E_Output":													
	ID^	Chrom:Sta	rt - End	Copy Number Varia	tion Type	Classification	Times Observ	ved Per Panel	Frequency Observed Per Panel	Gene	CDS Numbers		
1		1: 237865259 -	237865379	copy_number_loss		Unassigned	3		0.5	RYR2	66		
2		1: 237923057 -	237923172	copy_number_gain		Likely Deleterious	2		0.333	RYR2	83		
3		1: 237957148 -	237957317	copy_number_gain		Deleterious	3		0.5	RYR2	95		
4		4: 114213553 -	114213690	copy_number_gain		Unassigned	2		0.333	ANK2	21		
5		7: 150674907 -	150675021	copy_number_gain		Unassigned	2		0.333	KCNH2	1		
6		19: 35521706 -	35521784	copy_number_loss		Likely Benign	3		0.5	SCN1B	5' UTR - 1		

Figure 1: Copy Number Variation Data displayed in Geneticist Assistant including the copy number call, annotation, statistics and a user-set classification.

Method

Geneticist Assistant can import copy number variant data in GVF format. The copy number variant data can be linked to a sample and/or patient that can also include SNP/indel variants from a VCF file. A BAM file can also be imported for each sample to include coverage data for the panel regions.

🔛 New Run							? ×		
Run Name: Cardiac CNV Sam	ples								
Required Settings:				Optional Settings:					
Chemistry:	Halopley	~	New	Run Date/Time:	2/27/2020 1:14 PM				
chemistry.	Паюрісх		incritti	Sample Collection Da	te 🖲 Vcf File 🔵 Bam Fi	ile O Select 2/27/2	020 1:14 PM 😫		
Instrument:	Illumina	~	New	User Group:	Supervisor		~		
Reference:	Human 37		~	Sample Group:	Subpanel:				
	- Tuniario,			Submit Coverage					
Panel:	Cardiac	~	New	Import from VCF	Pathogenicity and Status Comment Artifact Type				
Panel Group:	Default			Submit Hotspots					
				Accepted VCF filters:	PASS		Edit		
Quality Metric Profile:	NextGENe Phase 1	~	New	Identify Patient ID					
Owner:	Administrator	~	New	Add New Patient					
				from Sample Name					
Multiple Samples In VCF				() from File					
Select Directory	Select Variant and/or Cove	rage and/or Copy Number Va	riant Files	Split Sample Name by Separator (underscore)					
E Output Mutation Rep	ort1:						^		
					Run Date/Time	2/27/2020 1:22 PM			
Sample Name:	E_Output_Mutatio	E_Output_Mutation_Report1			Sample Collection Dat	e 5/24/2019 9:33 AM			
Variant File(*.vcf):	mo_Data/Haloplex/	converted/E_Output/Reports	/E_Output_Muta	ation_Report1.vcf	. Sample Group				
Coverage/Pile Up File(*.b	oam) F:/Data/Demo_Da	ta/Haloplex/converted/E_Out	put/E_Output.b	am	Submit Coverage:	Subpanel			
Copy Number Variant File	(*.gvf) F:/Data/Demo Da	ta/Haloplex/E Output CNVRe	port.gvf		Patients:	E	~		
					Import Pathogenicity from VCF				
Reference:	Human 37 🗸				Submit Hotspots				
Panel:	Cardiac			V Panel Group: Defail	ult Accepted Filters:	PASS	Edit 🗸		
<							>		
						OK	Cancel		

Figure 2: GVF files containing copy number variants can be imported along VCF files and, optionally BAM files for the same sample/patient.

Conclusion

Geneticist Assistant enhances copy number variant analysis results by maintaining a database of detected copy number variants. For each CNV event Geneticist Assistant reports the total number of samples that have been imported with the variant, as well as the frequency at which the variant occurs both relative to the panel and for all samples for which CNV data has been imported. After reviewing a copy number variant, a classification can be assigned and displayed for future occurrences, improving efficiency by reducing the amount of review time required for later samples.

In conjunction with NextGENe software, Geneticist Assistant provides a complete pipeline for the analysis of SNVs, indels, and copy number variants from next generation sequencing data.

References

- 1. Bell, Callum J., et al. Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation Sequencing. Science Translational Medicine 12 Jan 2011: Vol. 3, Issue 65, 65ra4.
- 2. Boycott, Kym M., et al. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nature Reviews Genetics 2013: Vol. 14, 681–691.
- 3. Kerkhof, J., et al. Clinical Validation of Copy Number Variant Detection from Targeted Next-Generation Sequencing Panels. The Journal of Molecular Diagnostics 19, 905–920.
- 4. Plagnol, Vincent, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 28.21 (2012): 2747-2754.

