



APPLICATION NOTE

Technical validation of compatibility between GeneMind SURFSeq 5000 high- throughput sequencing data and VarSome Clinical

CONTENTS

Introduction	3
Methods	4
Results	5
Conclusions	7

Introduction

As new technology makes its way onto the market, it is vital to ensure compatibility between products and robust results across a range of protocols and assays. As such, we evaluated the compatibility of sequencing data generated by [GeneMind](#) instruments with [VarSome Clinical](#), to confirm that VarSome Clinical can accurately process these sequencing files and produce variant calls meeting the quality standards required for diagnostic use.

The technical validation was performed using the [GeneMind SURFSeq 5000](#) high-throughput sequencing platform and four different capture kits: the [Twist Bioscience Human Core Exome kit](#), the [Roche KAPA HyperExome V2 Probes](#), the [4bases Exome capture kit](#), and the GeneMind Biosciences Co., Ltd. GMExome V2 panel.

The results of this validation support the use of GeneMind SURFSeq 5000 sequencing data within the VarSome Clinical environment for clinical germline variant analysis.

GeneMind SURFSeq 5000 high-throughput sequencing platform

The GeneMind SURFSeq 5000 is a bench-top high-throughput sequencing platform. It has a dual flow cell to support five different modes/runs and is compatible with most commonly used applications. It has a unique reversible base termination and sequencing reaction system, combined with a high-sensitivity fluorescence signal detection system to maximize sequencing accuracy.

VarSome Clinical

VarSome Clinical is an IVDR-certified platform for clinical-grade variant interpretation, which calls, annotates, and classifies genetic variants in both germline and somatic workflows, with modules based on ACMG/AMP guidelines. It supports SNVs, indels, CNVs, structural variants, TMB & MSI, and RNA fusion annotation. The platform integrates data from [150+ sources](#) including ClinVar, gnomAD, dbNSFP 5.3, LOVD, DECIPHER, OncoKB™, CKB, and COSMIC.

Methods

A technical validation was performed to assess the compatibility of the VarSome Clinical germline analysis pipeline with sequencing data generated by the GeneMind SURFSeq 5000 high-throughput sequencing platform. Validation was performed on the well-characterized reference sample NA12878 (HG001) from the Genome in a Bottle consortium using four target enrichment workflows (the Twist Bioscience Human Core Exome kit, the Roche KAPA HyperExome V2 Probes, the 4bases Exome capture kit, and the GeneMind Biosciences Co., Ltd. GMExome V2 panel), for which high-confidence benchmark variant datasets are publicly available. Validation using the 4bases Exome capture kit was additionally performed on the NA24385 (HG002) and NA24149 (HG003) reference samples. The products are summarized in Table 1 below.

Table 1: Details of target enrichment workflows used to technically validate compatibility of the VarSome Clinical germline analysis pipeline with data generated by the GeneMind SURFSeq 5000 high-throughput sequencing platform.

Company	Product	Kit / Panel	Assay Target Region Length	
GeneMind Biosciences Co., Ltd.	GeneMind Whole Exome Sequencing Solution (GMExome V2)	GMExome V2 Panel	hg38	53554132 bases
Twist Bioscience	Human Core Exome	Human Core Exome Kit	hg19	33324490 bases
			hg38	33163529 bases
F. Hoffmann-La Roche Ltd	KAPA HyperExome V2 Probes and Kits	KAPA HyperExome V2 Probes	hg19	43114641 bases
			hg38	43181552 bases
4bases SA	4bases Exome Capture Kit	4bases Exome Capture Kit	hg38	37428616 bases

Methods

In all cases, libraries were prepared according to the manufacturer's recommended protocols and sequenced using the GeneMind SURFSeq 5000 high-throughput sequencing platform in PE150 configuration.

Variant calling was performed using the VarSome Clinical germline analysis pipeline aligned against the hg38 genome assembly for all kits, and additionally against the hg19 genome assembly for the Roche and Twist kits. Detected variants, with a minimum coverage depth threshold of 20x, were benchmarked against the GIAB high-confidence truth set (NIST v4.2.1), restricted to the high-confidence callable regions and the capture kit target regions.

VarSome Clinical version 13.15.0.1 was used in all examples.

Validation was performed using VarSome Clinical Thalia validate utility v13.15.2.

Performance statistics were calculated as follows:

Sensitivity = $TP / (TP + FN)$, measures how good the model is at finding all the positive cases.

Precision = $TP / (TP + FP)$, measures how trustworthy the model's positive predictions are.

Specificity = $TN / (TN + FP)$, measures how good the model is at identifying negative cases.

Results

Across all kits, performance of VarSome Clinical was assessed against the presence of variants within the truth set. Performance metrics across variant classes are presented in Table 2.

Table 2: Validation performance statistics for all kits and genome builds across all variants, and SNVs and Indels.

Kit	Genome build	Truth set	Variant Class	Total number of variants	Sensitivity	Precision	Specificity
Twist Bioscience Human Core Exome	hg19	NA12878	All variants	13,327	99.80%	99.60%	100%
			SNVs	12,934	99.80%	99.60%	100%
			Indels	371	98.10%	99.70%	100%
	hg38	NA12878	All variants	19,635	99%	99.80%	100%
			SNVs	19,325	99%	99.80%	100%
			Indels	283	98.20%	98.90%	100%
Roche KAPA HyperExome V2 Probes	hg19	NA12878	All variants	21,491	99.80%	99.60%	100%
			SNVs	20,207	99.90%	99.60%	100%
			Indels	1,053	98.20%	99.60%	100%
	hg38	NA12878	All variants	30,166	99.10%	99.80%	100%
			SNVs	28,573	99.20%	99.80%	100%
			Indels	989	97.00%	99.40%	100%
GeneMind GMExome V2 panel	hg38	NA12878	All variants	54,065	99.30%	99.80%	100%
			SNVs	51,305	99.40%	99.80%	100%
			Indels	1,735	98.90%	99.40%	100%
4Bases Exome capture kit	hg38	NA12878	All variants	19,813	99.40%	99.60%	100%
			SNVs	19,317	99.40%	99.60%	99.90%
			Indels	426	99.10%	98.60%	100%
		NA24385	All variants	21,815	99.30%	99.50%	100%
			SNVs	21,232	99.30%	99.50%	99.90%
			Indels	496	98%	98.40%	100%
		NA24149	All variants	23,199	99.40%	99.60%	100%
			SNVs	22,581	99.40%	99.6	99.90%
			Indels	528	97.90%	98.50%	100%

Conclusions

The results show that VarSome Clinical can successfully process sequencing data generated by the GeneMind SURFSeq 5000 high-throughput sequencing platform, using four different assays; Twist Bioscience Human Core Exome, Roche KAPA HyperExome V2 Probes, the 4bases Exome Capture Kit, and the GeneMind GMExome V2 panel. In all cases, variants present within the truth set were detected with $\geq 99\%$ sensitivity and precision, and $\geq 99.9\%$ specificity. Specifically, indels were detected at a rate of $\geq 97\%$ in all cases, and SNVs were detected at a rate of $\geq 99\%$ in all cases.

Importantly, this level of performance was consistent across the different assay designs and variant classes evaluated, supporting the robustness of VarSome Clinical when analyzing data generated by the GeneMind SURFSeq 5000 platform.

Therefore, these findings provide strong evidence of technical compatibility between VarSome Clinical and GeneMind SURFSeq 5000 sequencing data, with performance that meets quality thresholds expected for clinical-grade NGS variant analysis.



VSC-WP008-V02

Built with Swiss precision by



EPFL Innovation Park - I
1015 Lausanne,
Switzerland