# **GeneticistAssistant** NGS Interpretative Workbench TM

# Features:

# Variant Database

- Historical Database
- Pathogenicity Calling Information
- Pathogenicity Call Supporting Information
- Linkage to External Databases
- Automated Quality Control
- Accessibility
- User Management, Audit Trail, Access Control

# **Cool Tools**

- Custom Report Builder
- Customer Web Portal
- Automated Informatics Pipeline
- Customizable Workflow Builder
- Sample Comparison
- Custom Filtering
- Artifact Flagging
- Process Quality Control
- Positive Control Verification
- Automatic BED file builder with regions of clinical significance

**Developed in collaboration** with Mayo Clinic







# Efficient...Saves Time & Resources, Controls...Real-time Administration & Reporting, For...Disease Panels and Whole Exome Sequencing data, Compatible...with data from all NGS Systems

Developed in collaboration with the Laboratory Medicine, Information Technology and Health Science Research departments of Mayo Clinic, Geneticist Assistant NGS Interpretative Workbench is a unique tool for the management, control, visualization, functional interpretation and historical knowledge base of next generation sequencing Whole Exome data or Disease Panels targeted at specific genes for the purpose of identifying potentially pathogenic variants associated with specific conditions such as hereditary colon cancer and others.

Geneticist Assistant is compatible with data processed from all leading next generation sequencing platforms including Ion Torrent, Illumina and Roche platforms. The program accepts standardized BAM and VCF files, and includes information from the following sources:

#### **Functional Prediction information:**

SIFT, PolyPhen-2, LRT, MutationTaster, FATHMM, CADD & MutationAssessor

Disease association: ClinVar, OMIM, CIViC & COSMIC\*

Conservation scores: phyloP, GERP++, phastCons & SiPhy

**Population frequencies:** 1000 Genomes, Exome Variant Server, and ExAC

Additionally, information from proprietary databases such as **Alamut** and LOVD (Leiden Open Variation Database) are easily accessible through embedded links. Information from other publicly available databases are easily imported into the workbench.

The new administration function provides a real-time tracking of current statuses; historical information; automated email notifications within a completely customizable workflow built to model your actual activities.

Unique tools include **Custom Filtering**, **Patient Comparison**, i.e. **Trio Comparison**, **CAP Validation Assistance**, **automated BED file builder** which automatically highlights areas of clinical significance, **Positive Control Verification**, and in conjunction with NextGENe software can form a completely **automated informatics pipeline**.

\*Requires separate license

# **Historical Database Development**

Geneticist Assistant NGS Interpretative Workbench records variant pathogenicity determination on all found variants, eliminating time consuming duplication of researching the variant, thus speeding diagnosis while reducing costs. As the database is used the number of variants requiring pathogenicity calling is quickly reduced to a few novel variants.

	iants of '272																
ID	Chr : ChrPos		Pathogenicity Likely Deleterious	Gene	Exon Number	Туре	Variant Frequency	Coverage	HGVS Protein	Panel	HGVS Coding c.1458T>C	Te 10	nes Observed Per Panel Time 10	es Observed Per Run 5 11	Samples Per Panel Times Observed Per I 10	Panel Group 11	Samples Per Panel Group
	5:112162854 5:112164561		Likely Deleterious Benign		12	synonymous synonymous		69 60	p.Tyr486= p.Ala545=	DLMP	c.14581>C c.1635G>A	10	10	11	10	11	
9	10:88635779					missense	1	99	p.Pro2Thr	DLMP	c.4C>A	5	5		nts.mitci	11	
14	14:75513883		Benign	MLH3		missense	1	55	p.Asn826Asp	DLMP	c.2476A>G	11	11	Jivana	into anifection	11	
15	17:7579472	rs1042522	Deleterious	TP53	4	missense	1	46	p.Pro72Arg	DLMP	c.215C>G	9	rirPos	Rs	Pathogenicity	11	
16	17:63533768		Deleterious	AXIN2		synonymous	0.5	50	p.Pro462=	DLMP	c.1386C>T	8				11	
17	17:63533789		Likely Benign	AXIN2		synonymous		54	p.Pro455=	DLMP	c.1365A>G	9	12162854	rs2229992	Likely Deleterious	APC 11	
18	17:63554591		Likely Benign	AXIN2	2	missense	1	111	p.Pro50Ser	DLMP	c.148C>T	8	6	0.04.004		100	
38	14:75513828 2:48010488		Unknown	MLH3 MSH6	2	missense	1	55 64	p.Pro844Leu p.Gly39Glu	DLMP	c.2531C>T c.116G>A	2/	5:112164561	<u>rs351771</u>	Benign	APC	1
~	2.40010400	11041011	onknown	mano		maxine	•		ploybold	0 Civil	CarloorA	1	10:88635779	rs11528010	<b>Likely Deleterious</b>	BMPR1/	4 3
													14:75513883	rs175081	Benign	MLH3	2
													17:7579472	rs1042522	Deleterious	TP53	4
		L										7	17:63533768	rs1133683	Deleterious	AXIN2	6
												/	17:63533789	<u>rs9915936</u>	Likely Benign	AXIN2	
													63554591	rs2240308	Likely Benign	AXIN2	
													3828	rs175080	Unknown	1	
														4042821	Unknow		

Historical information on every found variant is recorded and available for instant recall. Additionally prior pathogenicity determination is logged by specific disease panel and globally for all disease panels. The variant review tab provides previously determined variant type, pathogenicity, variant frequency, HGVS Nomenclature, times observed, number of times observed in disease panel and panel group.

Use of the workbench will quickly reduce unnecessary pathogenicity research duplication, speeding diagnoses and reducing costs.

# **Pathogenicity Calling Information**

Geneticist Assistant NGS Interpretative Workbench provides Variant Interpretation, Functional Prediction, Conservation Scores and Disease Associations on each found variant from over 17 sources providing the information in a single view. **Once a call has been** made and confirmed, the research is stored in the database and applied to future recurrences of the variant either in the same disease panel or in any other panel, significantly reducing time and effort on future iterations of the variant in future analyses.

C:/Users/soft/De	esktop/G	A/reference	es/Human	37/ESP65	600SI-	V2-SS/	137.vcf												
Chromosome	17	AA	G					EA	AC	6409,	2191					IGVS_CDNA_VAR		118.1:c.98C>G,N	
Chromosome Positi	on 7579472	AA_AC	1784,2620					EA	AGE								01126113.2:	c.215C>G,NM_0	01126112.2:c.2
ID	rs104253	2 AA AGE						EA	STC	2409.	1591,300						46.5:c.215C		
Ref	G	AA GTC	386, 1012, 804					EXO	ME CHIP	yes						IGVS_PROTEIN_VA			
Alt	C	CA	http://www.no	bi.nim.nih.oo	/sites/w	arvu?gene		FG	-	NM O	01126118, 1:miss	ense.NM 0	01126114.2	timissense.h	M 0011261		(P72R),NM_C	001126113.2:p.(F	72R),NM_001
Oual	0	CDS SIZES	NM 00112611	8.1:1065.NM	0011261	114.2:1026	NM 001126113	3.2:1041.N		13.2:	missense,NM_00	1126112.2:	missense,N	4_000546.5	:missense	IAF		000546.5:p.(P72P	0
Filter	PASS		M_001126112	2:1182,NM_0	00546.5	:1182	-	GL		TP53								5086,36.9963	
HGVS Genomic		CG	1.9					GS		103,1	03, 103, 103, 103					211		aging:0.745,post 745,possibly-dam	
HGVS Coding		CP	0.0					GTC		2795,	2603,1104						damaging:0.	745,possibly-dam 745	aging:0.745,p
HGVS Protein		DBSNP	dbSNP_86					GTS		CC,CC	G,GG					AC	8193,4811	/15	
		DP	92					GWA	S_PUBMED	· .							0100, 011		
/Users/soft/De	esktop/G	A/reference	es/Human	37/dinvar	_00-k	atest.vc	f												
hromosome	17	CAF	[0.3981.0.60]	9]					CLNSRCID	.1201	77 191170.0005	LSD	LSD	SLO	SLO				
Chromosome Positi	on 7579472	CLNACC	RCV00001314	4.1 RCV0000	34639.1	RCV00007	9202.1		COMMON	1		OM	OM	SSR	0				
ID	rs10425	22 CLNALLE	1						65	G5		OTHERK	G OTHERKG	TPA	TPA				
Ref	G	CLNDBN	CODON 72 P	OLYMORPHIS	1\x2c (r:	s1042522)	not providedIA	HighlyPenetrant	GSA	G5A		PH3	PH3	VC	SNV				
Alt	C	CLNDSDB	. I. MedGen						GENEINFO	TP53	3:7157	PM	PM	VLD	VLD				
Qual	0	CLNDSDBID	CN169374						GNO	GNO		PMC	PMC	VP	0x0501	7800000017051f1101	01		
Filter		CLNHGVS	NC 000017.10	):a.7579472G	×				HD	HD		RS	1042522	WGT	0				
HGVS Genomic		CLNORIGIN	1						KGPROD	KGPF	ROD	RSPOS	7579472	dbSNP8u	ildID 86				
HGVS Coding		CLINSIG	2 2 2						KGPhase1	KGPH	nase 1	RV	RV						
HGVS Protein		CLNSRC	. Emory_Unive	rsity OMIM_A	delc_Var	riant			KGPilot12	3 KGPI	lot123	SAO	1						
:/Users/soft/De	esktop/G	A/reference	es/Human	37/dbNSI	P2.4_	variant	.chr17												
chr 17		Jniprot_aapo	72;72;72;72;72	aapos	72		Polyphen2	HVAR_ 0.4198			N N	Reliabili	ty_index	9	phyloP46wa		5iPhy_29way	9.7733	1000Gp1_/
pos(1-coor) 75794	72			aapos_SIF		P00000269			ssor_pr			CADD_r	aw	0.823228	placental		logOdds		R_AC
ref G		interpro_dom				:P72R		HVAR_ B;B;B;B	B FATHM	M_sco	-5.23;-5.74;-5.	CADD n	aw rankso	0.16678	nkscore		5iPhy_29way		1000Gp1_4
alt C		ain		aapos_FA		P00000410			re		45;-5.74;-5.74;				phyloP100v		_log0dds_ran		R_AF
aaref P		ds_strand		MM			P LRT_score	0.3708			-5.45;-4.11;-2. 05			8.316	y_vertebra		kscore		1000Gp1_#
aaalt R			CCC		000	00352610: L;ENSP0000	P LRT_conve nkscore	rted_ra 0.0444	1		0.99319	GERP++		1.87	phyloP100v v vertebra			2.010640	_AC
hg18_pos(1- 75201		SLR_test_sta			026	9305:P72R	LPT need		kscore	n_ran	0.99519	GERP++		1.87	rankscore		UniSNP_ids	rs1042522;rs2 229076:rs3174	1000Gp1_4
coor)			2		ENS		Statute To		EATHM	M pre	D;D;D;D;D;D;D;D	GERP++ score	_RS_rank	0.25490	phastCons4	6 0.002000		747;rs4134781	ESP6500 A
genename TP53		old-			846	:P72R;ENS		acci _ac olo	d	_	;D		6way_pri		way_prima	te		:rs17844988:rs	AE
Uniprot_acc E7EMP 7-2:P0		legenerate	•			00391127: LENSP0000	P	ster_c 0.8072	2 RadialS	VM_so	-0.9287	mate	6way_pri	0.561000	phastCons4			17857747;rs17 882155:rs6038	ESP6500_E
mples Associat					10				ore						way prima			0021020000	AC
D* Name	B	tun Date Time	Add Da	ite Time	Run	Panel	PanelGroup	Reference	Patient E	xternal	ID Status			Va	riant File			Cover	age/Pile Up Fil
800463.variants.	filter 5/14/	2014 11:05:33 A	M 5/14/2014	11:09:54 AM	Demo	DLMP	default	Human 37			New	C:/Use	rs/soft/Der	ktop/GA/N	Aavo data/80	0463.variants.filter.v	cf C:/Users/sc	ft/Desktop/GA	Mayo data/
800418.variants.							default	Human 37			New					0418.variants.filter.v			
800458 variants							default	Human 37			New					0458.variants.filter.v			
cooxid/verrarits.	····ci 3/14/	2014 12.00.00 M		ALLENDO MINI	e cano	PENIP	acrean	mannall 37			-VEW	0,056	14, 2010/0/0	100p, 0001		eroenenentisiniten.v	cr cy users/st	ne ocatop/ 040	mayo_uata/o

Geneticist Assistant Workbench provides a complete overview of information regarding variant pathogenicity in one detailed view. Prior samples which exhibited variant are also detailed.

dbSNP Exome Variant Server

Functional Prediction: SIFT PolyPhen-2 LRT MutationTaster MutationAssessor FATHMM CADD Conservation Scores: phyloP phastCons GERP++ SiPhy

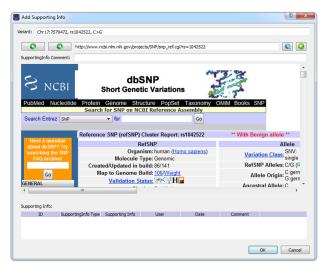
Disease Association: COSMIC\* ClinVar & OMIM CIViC Alamut\* LOVD (Leiden Open Variation Database) And others **Population frequencies:** 1000 Genomes Exome Variant Server

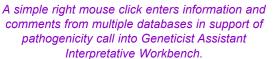
ExAC

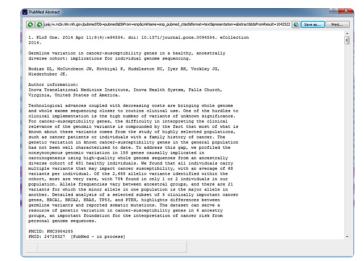
# **Pathogenicity Call Supporting Information**

Supporting information for a pathogenicity call is easily added to the database by a right mouse click in the variant tab. Data from any source such as dbSNP can be added for future recall.

Geneticist Assistant NGS Interpretative Workbench also includes a "mini web browser" which allows a user to search and link scientific information from any web source such as NCBI in support of the pathogenicity call which can be recalled at any time by authorized users. PubMed abstracts can be automatically downloaded into the workbench.







PubMed abstracts can be automatically downloaded into the workbench.

#### Linkage to External Databases

Retrieving further information from external proprietary tools such as Alamut, UCSC Genome Browser, or the LOVD database is a simple click away. (Alamut requires a license)

Varia	ants of '8004	63.igv-sor	ted_Output_M	lutation_	Report1	L_filter	ed': *Filters	; Ар	plied
ID	Chr : ChrPos^	Rs	Pathogenicity	Pathogenic	ity Status	Gene		1	HGVS Coding
28	3:37056045	<u>rs182733777</u>	Unassigned		Variant De	etails		790	+10A>G
29	3:37081751	rs267607840	Unassigned		Show Vari	iants Filte	ered by Panel	.63	3A>G
6	3:37083740	rs9876116	Benign		Edit Varia	nt	•	.66	i8-19A>G
13	14:75505016	<u>rs175075</u>	Benign		External		•	r -	View In Alamut
14	14:75513883	<u>rs175081</u>	Benign		Export		•		View In UCSC
31	14:75514489	rs28756986	Unassigned		Reports				View PubMed Abstract
15	17:7579472	rs1042522	Deleterious		Load Colu	impe		4	View Publyled Abstract
1443	17:7579669	rs17878362	Unassigned		Load Cold				-32_96+47delGGGCTGGGGA
18	17:63554591	rs2240308	Likely Deleterious	Confirmed		AXIN2	NM_004655.3:	c.148	IC>T
19	18:48584856	rs386387676	Likely Benign			SMAD4	NM_005359.5:	c.904	+45_904+46insTT

*Alamut* licensees can quickly retrieve information without error prone and tedious retyping by simply selecting variant of interest and clicking on the drop down menu.

LOVD	Data:									
Symbol	ID	Position mRNA	Position Genomic	Variant DNA	Variant DBID	Times Reported	Chromosome	Allele	Affects Function (Reported)	Affects Function (Concluded)
IVD	16587	NM_002225.3:c.1276_1278	chr15:40710457_40710459	c.(1276_1278del)	IVD_000013	1	N/A	Unknown	Effect unknown	Effect unknown

Retrieving information from the LOVD database is a simple linked operation.

# **Automated Quality Control**

Geneticist Assistant NGS Interpretative Workbench automatically monitors coverage depth, flagging regions to the base level that do not meet your pre-set requirements. The software will track over time the amplicon or regions' performance, providing feedback on the sequence performance, which may alert you to areas that require performance improvement.

	Settings	-						8 ×
	Directories	Quality Control	Pathogenicity Colors	Tab Preferences	Variant Preferences	Alamut Settings	HGMD Settings	License Settings
	Minimum Cove	erage:			50			•
	Average Cov	erage:			100			-
	Percent Cove	red(%):			100.00			-
ľ							ОК	Cancel

Quality control requirements are easily set in the Quality Control tab, the software will then monitor the sequence performance to the base level, indicating regions of non-performance.

Region Name	Chrom:Start - End 🔦 % Cov	ered Average Coverage	Minimum Coverage	Status	Average % Covered	Average Average Coverage	Average Minimum Coverage	Passed	Passed Percent	Failed	Tota
MSH2:NM_000251	2:47630301 - 47630571 100%	493.03	188	Passed -	100%	353	131	10	90.9091%	1	11
MSH2:NM_000251	2: 47635510 - 47635724 100%	722.87	354	Passed -	100%	516	270	10	90.9091%	1	11
MSH2:NM_000251	2: 47637203 - 47637541 100%	674.83	169	Passed -	100%	525	143	10	90.9091%	1	11
MSH2:NM_000251	2: 47639523 - 47639729 100%	777.8	393	Passed -	100%	579	312	10	90.9091%	1	11
MSH2:NM_000251	2: 47641378 - 47641587 100%	636.01	127	Passed	100%	450	104	10	90.9091%	1	11
MSH2:NM_000251	2: 47643405 - 47643598 100%	867.31	485	Passed -	100%	662	367	10	90.9091%	1	11
MSH2:NM_000251	2: 47656851 - 47657110 100%	894.67	431	Passed •	100%	717	347	10	90.9091%	1	11
MSH2:NM_000251	2: 47672657 - 47672826 100%	560.26	343	Passed -	100%	436	284	10	90.9091%	1	11
MSH2:NM_000251	2: 47690140 - 47690323 100%	718.4	458	Passed •	100%	495	313	10	90.9091%	1	11
MSH2:NM_000251	2: 47693767 - 47693977 100%	565.8	13	Failed -	100%	463	81	6	54.5455%	5	11
MSH2:NM 000251	2:47698074 - 47698231 100%	546.21	368	Passed -	100%	391	251	10	90,9091%	1	11

Quality data is presented for both the current sample and a complete history of analysis of all samples for a disease panel. Metrics provided include Minimum Coverage, Average Coverage, % Coverage Across Region and Pass/Fail Status of current run. Historical data includes average coverage of all runs, average percent coverage, absolute Pass/Fail counts, total samples for the region and passed percentage. Sequencing that often fails is easily reviewed, allowing user to determine and correct cause of sequencing failures.

ID^	Name	Run Date Time	Add Date Time	Run	Panel	PanelGroup	Reference	# Regions	# Regions Passed	Patient External ID	Status	Missed Clinical Variants
8	800466.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:18:38 AM	Demo	DLMP	default	Human 37	154	151	XYZ789	Complete 🔻	Yes
9	800402.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:20:03 AM	Demo	DLMP	default	Human 37	154	150	ABC123	New 🔻	Yes
10	800451.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:21:26 AM	Demo	DLMP	default	Human 37	154	152	BC-13-15487	QC Passed 🔻	Yes
11	800474.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:22:46 AM	Demo	DLMP	default	Human 37	154	153	BC-13-20683	Reviewed 💌	Yes
12	272305.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:24:08 AM	Demo	DLMP	default	Human 37	154	4	BC-13-20476	New 🔻	Yes

Importantly, Geneticist Assistant NGS Interpretative Workbench, monitors areas of clinical significance providing a quick review of missed clinical variants as determined by the ClinVar database information.

#### Accessibility

Geneticist Assistant NGS Interpretative Workbench is comprised of a local installed database, either Linux or Windows<sup>®</sup>, and a client Windows program which provides the easy-to-use, graphical user interface. All data is stored locally, accessible only to authorized users. Off-site collaborators or sister facilities can securely (HTTPS security protocol) access the database via the internet.







- 1. Database and Client may reside on single computer
- 2. Geneticist Assistant can be accessed by any computer having client within institution network
- 3. Off-site collaborators or sister facilities can securely (HTTPS security protocol) access the database via the internet.

# User Management, Audit Trail, Access Control

Geneticist Assistant NGS Interpretative database employs a customizable password system (such as an 8 character alpha-numeric password) to protect data integrity. Database records all log-in and log-off and all user-activity by user, which can be recalled by administrative personnel. Access to various information contained in the database can be granted or limited by individuals, and groups. Geneticist Assistant NGS Interpretative Workbench records and tracks all changes and comments for future recall.

Geneticist Assistant Workbench employs a customizable password system (such as an 8 character alpha-numeric password) to gain access to the database.

💮 Geneticist	Assistant - Login
Server:	https://localhost
Username:	Administrator
Password:	
	OK Cancel

lsemane:				
assword:				
Confirm Password:				
irst Name (optional):				
ast Name (optional):				
inal:				
Administrator Privi	leges	C Staff	📝 Act	ive
Permissions:				
Type	View	C Add	🛅 Edit	C Delete
Chemistry	<b>1</b>	12		E
Instrument				<b></b>
Panel				
Panel Group	<b>1</b>	<b></b>		
Pathogenicity	<b>1</b>	<b>1</b>		E2
Patient		<b></b>		
Run		<b>1</b>		
Sample				
Variant	171	171	12	10

Access to various information within Geneticist Assistant can be granted by individual and groups.

Geneticist Assistant Workbench records and tracks all changes and comments made to the database by users for future recall.

ID	14	Coverage	344	Protein		NP_001035197.1	Times Observed Per Panel	11	Ref
Chromosome	14	Pathogenicity	Benign	Coding Ba	se	2476	Times Observed Per Panel Group	11	
<b>Chromosome Position</b>	75513883	Pathogenicity	Status	Codon Po	sition	1	Samples Per Panel	11	
Chr : ChrPos	14:75513883	Variant Freque	ency 1	AA Positi	on	826	Samples Per Panel Group	11	
Rs	rs175081	Zygosity	homo	HGVS Ger	omic	g.75513883T>C	<b>Times Observed Per Patient Per Pane</b>	0	
Ref	т	Read Balance	0	HGVS Cod	ling	c.2476A>G	Samples Per Patient Per Panel	0	
Ref AA	Asn	Gene	MLH3	HGVS Pro	tein	p.Asn826Asp	Patient Variant Frequency	NA	
Alt	с	Gene Strand	-	Variant C	omment		Trans	Ti	
Alt AA	Asp	Exon Number	2	Times Ob	served Per Run	11	GMAF		
Туре	missense	Transcript	NM_00104010	8.1 Panel		DLMP	Alt	С	
Pathogenicity Chan	ges:								
Type	Value	User	Date 🖍	Comment					
Pathogenicity Change	Benign	Administrator 5	5/14/2014 10:58:18 AM	1					
Pathogenicity Status Cha	ange	Administrator 5	5/14/2014 10:58:18 AM	1					

# Cool Tools

#### **Custom Report Builder**

Geneticist Assistant's Report Designer allows users to create highly customizable report templates for the quick and easy creation of standardized reports for each sample/patient. Using the Report Designer users can select the content to be included in the report and define formatting for the report such as report headers, page headers, as well as the inclusion of a lab logo image.

Custom tables can be created to pull data such as variant information and patient details directly from the Geneticist Assistant database. Custom text fields can also be added to include descriptions such as methods, clinical information and/or a disclaimer. Any custom section can be added when creating a report template through the Report Designer. The report template can then be saved for later use in saving reports. Multiple report templates can also be created for different report types. Templates can also be saved for individual sections within a report to allow the quick implementation of the same content when creating a new report template.

	Page heade	5		
[Sample.Patient.External [Sample.Patient.Last Nar [Sample.Patient.First Nar	Boftware PowerTools for G	netic Analysis	Institution Department Street Address City, State Zip Phone #	
			Laboratory Director	
	Data Group He	ader		3]
Patient Information				
ID: [Sample.Patient.External ID]	Last Name: [Sample.Patient.Last Name] DOB: [Sample.Patient.Date of Birth]	Gender:	2: atient.First Name] ?atient.Gender]	
Method				
Method description				
Variants	Master head	er		-
HGVS Coding	Pathoger	nicity	Variant Comment	
	Master ban			t
	[Sample.Variant.Pathog	enicity]	[Sample.Variant.Variant Comme	nt]
[Sample.Variant.HGVS Coding]				(;
[Sample.Variant.HGVS Coding]	Data Group Fo	oter		
[Sample.Variant.HGVS Coding]	Data Group Fo	oter		
	Data Group Fo	oter		

## **Customer Web Portal**

Geneticist Assistant offer access to a customer web portal that can be used for tracking and managing tests ordered from referring institutions. The web portal is directly linked to the Geneticist Assistant database so that information regarding patients and sample submissions can be viewed in Geneticist Assistant and included in reports.

#### The web portal provides:

- Customizable interface
- Production and recognition of sample barcodes
- · Secure encryption of patient information
- Patients and sample submissions linked directly with Geneticist Assistant
- · Printing of packing slips for sample tracking

#### **Patient Tracking**

Patient information, including a patient ID, DOB, gender, relationships, and phenotype, can be imported to the Geneticist Assistant database. Each imported sample can then be assigned to a patient

Packing Slip		23 Jan 2017	10:43:33 AM	
Patient			Ordering Provid	ler
External Id	800426		First Name	Sue
Mother	800418		Last Name	Smith
Father	800402		Phone	1234567890
Gender	Female		Email	sue@hospital.org
			Street	2 Main St
Specimen			City	City
Date Collected	2017-01-23		State	PA
Units	15		Zip	16803
Sample Type	Blood			
			Genetic Counse	elor
Tests Requested			First Name	John
Tests	1		Last Name	Jones
			Phone	3456789012
Reason For Testi	ng		Email	John@counselor.org
Indication	unknown			
Clinical Diagnosis	Breast tumor			ital Or Laboratory
			Organization	Oncology Lab
			Phone	6789012345
			Street	5 Main St
			City	City

Add Patient	Batch Import						
ad Patient	bacch import						
atient File (*.c	sv, ".txt, ".tsv):						
File Properties	8						
Relationship (	Column Starts at: 8	•					
Add New I	Phenotype Terms in File	e Automatically					
Data Delimiter	13						
Comma(",	. <b>D</b>						
Tab("\t")							
Date of Birth	Format:						
MM/dd/yyyyy	<ul> <li>Example: 09/23/</li> </ul>	2015 or 9/23/2015					
File Column Fo	ormat:						
-							
File follow	is standard column ord	er					
	is standard column ord is standard column nan						
File follow	is standard column nan	nes	ifv				
File follow		nes	sfy				
File follow	is standard column nan	nes	sfy				
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec					
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patient_relationships
<ul> <li>File follow</li> <li>Files does</li> <li>Result:</li> </ul>	is standard column nan i not follow above stan	nes dards, I want to spec		patient_gender	patient_race	patient_phenotype	patent_relatorships

#### **Compare Samples**

Create a comparison of multiple samples to view differences in variant calls and/or variant frequencies. Output from **different pipelines** can be compared by importing VCF files from each pipeline and comparisons of **family members**, such as a **trio comparison**, can be created.

୍ମ	Family	y Comparison				? ×				
1	Phenoty	pe:				•				
	Inheritar	nce pattern:	Autosomal Recessive							
			Compound	heterozygous						
1	Show col	lumns:	V AF	🔽 Co	v [	RB RB				
	Filter	variants by panel:	Build37_CCDS	_Exons_Merge	Overlaps	-				
	Sample ID	Sample Name	Patient External ID	Relationship	Phenotype	Zygosity				
	87	UDP3168_Mutation_Report1_Filtered	UDP3168	Father 🔻	Unaffected 🔻	Heterozygous				
	86	UDP3165_Mutation_Report1_Filtered	UDP3165	Mother 🔻	Unaffected 🔻	Heterozygous				
	84	UDP2753_Mutation_Report1_Filtered	UDP2753	Son 💌	Affected 🔻	Homozygous				
	85	UDP2755_Mutation_Report1_Filtered	UDP2755	Daughter 🔻	Affected 🔻	Homozygous				
					ОК	Cancel				

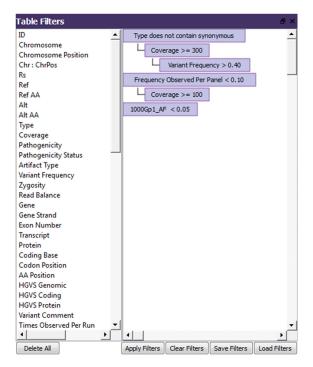
For family comparisons, specify the relationships and phenotypes for each patient to create a comparison based on a selected inheritance pattern.

Family	Compar	son of 4	Samples:													
AF	Cov	O AF	O Cov	AF^	Cov	AF	Cov	ID^ Chr : C	hrPos Rs	Gene	Ref AA	Alt AA	Type	Coverage	Pathogenicity	HGVS Coding
0.494	166	0.437	119	1.000	121	0.984	127	1640 1:11444	3899 <u>rs17464525</u>	AP4B1	Gly	Gly	synonymous	121	Unassigned	NM_001253852.1:c.576C>T
0.527	93	0.448	58	1.000	56	0.984	64	1745 1:10830	7727 <u>rs7528153</u>	VAV3	Thr	Ser	missense	56	Unassigned	NM_006113.4:c.892A>T
0.409	22	0.316	19	1.000	7	1.000	11	3167 2:20691	1228 <u>rs2909111</u>	INO80D	Ala	Val	missense	7	Deleterious	NM_017759.4:c.1073C>T
0.636	11	0.667	24	0.773	22	0.800	10	3306 2:11130	4496 <u>rs71231856</u>	RGPD6	Val	Val	synonymous	22	Benign	NM_001123363.3:c.1560G>A
0.471	121	0.506	85	1.000	97	1.000	79	3382 2:21055	7380 <u>rs6720659</u>	MAP2	His	His	synonymous	97	Unassigned	NM_002374.3:c.486C>T
0.485	130	0.426	122	1.000	135	0.992	118	3390 2:21055	7542 <u>rs741007</u>	MAP2	Thr	Thr	synonymous	135	Unassigned	NM_002374.3:c.648G>A
0.466	103	0.469	113	0.987	76	0.959	74	3396 2:21055	8162 <u>rs741006</u>	MAP2	Arg	Lys	missense	76	Likely Benign	NM_002374.3:c.1268G>A
0.374	187	0.379	145	1.000	169	1.000	130	3400 2:21055	9960 <u>rs2239672</u>	MAP2	Val	Val	synonymous	169	Likely Benign	NM_002374.3:c.3066G>T
0.425	80	0.594	64	0.982	55	1.000	59	3428 2:21145	5637 <u>rs1047883</u>	CPS1	Thr	Ala	missense	55	Benign	NM_001122633.2:c.1048A>G
0.436	78	0.597	62	0.982	57	1.000	59	3432 2:21145	5639 <u>rs2229589</u>	CPS1	Thr	Thr	synonymous	57	Unassigned	NM_001122633.2:c.1050C>T
0.413	184	0.500	182	0.981	157	1.000	126	3440 2:21148	1257 <u>rs2287599</u>	CPS1	Gly	Gly	synonymous	157	Unassigned	NM_001122633.2:c.2697C>G
0.625	72	0.519	79	1.000	86	1.000	82	3594 2:15966	3616 <u>rs10497199</u>	DAPL1	Ala	Thr	missense	86	Likely Deleterious	NM_001017920.2:c.196G>A

Comparison results show variant coverage and allele frequency values for each patient to quickly identify differences and shared variants.

# **Custom Filtering Options**

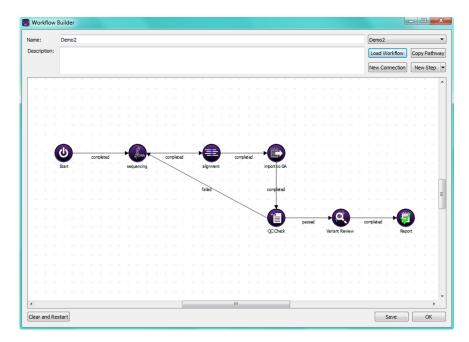
Variants lists, as well as any other data tables in Geneticist Assistant, can be filtered based on a combination of any data fields.



Drag and drop any data field to use for filtering. Multiple filters can be combined and the combined filter can be saved for later use.

#### **Customizable Workflow Builder**

Geneticist Assistant NGS Interpretative Workbench now includes a completely customizable workflow builder that enables you to model your physical NGS workflow. A workflow can then be designated for any cases entered in Geneticist Assistant.



Geneticist Assistant NGS Interpretative Workbench features a completely customizable workflow builder that enables users to model physical workflow in "silico".

# **Process Quality Control**

#### Control Charting for real time and historic evaluation

Track run-to-run variability of control samples. Data is tracked for each individual target region. The data can be used to determine drift in the analytical quality both globally as well as for specific genes and target regions. In addition, the data can be used to easily determine changes between manufacturer reagent lots. The tabular format can easily be exported in csv format to create control charts and graphs.

Geneticist Assistant										
File Panels References Views	Reports Filt	ers Tools Back	up Help							
Sample '800456.variants.filter'										
Control: Number of Samples having the variant	Control: Coverage Min	Control: Coverage Max	Control: Coverage Median	Control: Coverage Mean	Control: Coverage Standard Deviation	Control: Variant Frequency Min	Control: Variant Frequency Max	Control: Variant Frequency Median	Control: Variant Frequency Mean	Control: Variant Frequency Standard Deviati
7	209.0	342.0	254.0	273.857142857	42.6930190728	0.5	0.5	0.5	0.5	0.0
4	210.0	316.0	289.0	276.0	43.7207044774	0.5	0.5	0.5	0.5	0.0
5	299.0	374.0	328.0	325.8	27.4765354439	0.5	0.5	0.5	0.5	0.0
4	171.0	204.0	178.5	183.0	13.5830777072	0.5	0.5	0.5	0.5	0.0
8	522.0	950.0	624.0	658.375	135.858324644	0.5	1.0	1.0	0.8125	0.242061459138
5	344.0	513.0	447.0	430.0	72.420991432	0.5	1.0	0.5	0.7	0.244948974278
5	335.0	454.0	368.0	383.8	44.4675162338	0.5	0.5	0.5	0.5	0.0
5	318.0	449.0	403.0	390.4	52.6026615296	0.5	0.5	0.5	0.5	0.0
8	252.0	368.0	318.0	310.875	49.8308074889	0.5	1.0	0.75	0.75	0.25
11	418.0	614.0	480.0	506.181818182	76.5124091918	0.5	1.0	1.0	0.909090909091	0.19284730396
10	274.0	514.0	367.5	382.3	68.2071110662	0.5	1.0	1.0	0.85	0.229128784748
10	349.0	580.0	456.0	462.4	71.2210642998	0.5	1.0	0.5	0.7	0.244948974278
11	368.0	567.0	422.0	452.636363636	72.0407368688	0.5	1.0	1.0	0.909090909091	0.19284730396
10	259.0	407.0	331.5	329.5	53.2357962277	1.0	1.0	1.0	1.0	0.0
12	301.0	564.0	371.5	385.416666667	78.6675052921	1.0	1.0	1.0	1.0	0.0
12	294.0	501.0	367.0	379.25	61.7833917813	1.0	1.0	1.0	1.0	0.0

Geneticist Assistant records variants in control samples allowing instant review and long term monitoring of process.

erage Regions of '800456.var	ante fiter'																
erage Regions of au0456.var	ants.niter 🔯													-			
Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:		Control:	Control:	Core	ntrol:	-			
mber of Samples having the	region % Covered Mea	in Average Coverage	Min Average Coverage	e Max Average Coverage Me	dian Average Coverage M	ean Average Coverage Standard Devi	ation Minimum Coverage	Min Minimum Coverage M	lax Minimum (		ian Minimum Coverage	Mean Minimum Coverage					
	100.0	51.24	114.28	82.75	85.5754545455	16.1435319801	47.0	87.0	62.0		66.6363636364	11.1866212892					
	100.0	105.78	202.27	148.87	149.944545455	23.9406798581	71.0	142.0	90.0		94.7272727273	18.27114421					
	100.0	116.97	225.87	181.35	177.254545455	39.2522956743	88.0	182.0	143					MSH2			
	100.0	134.24	218.99	174.07	175.94	29.3754671916	81.0	162.0	125								
	100.0	145.13	271.25	208.86	212.593636364	35.9100703204	72.0	121.0	90.0				NIV	1_000251 Ex1			
	100.0	151.21	263.35	214.02	209.098181818	35.1119642438	54.0	104.0	1.08	450							
	100.0	152.74	289.92	232.67	230.353636364	40.9629551652	61.0	151.0	107	450							
	100.0	168.65	286.41	207.58	215.865454545	31.2543871433	88.0	201.0	149								
	100.0	172.75	316.94	209.16	221.563636364	48.7027082078	105.0	186.0	132	400							
	100.0	190.92	359.8	305.66	293.169090909	48.88431603	85.0	198.0	152	400	-					+	
	100.0	191.92	434.59	278.0	291.047272727	64.1764522222	20.0	108.0	69.1								
	100.0	208.87	424.51	324.1	326.473636364	62.6100899903	135.0	306.0	213	350							
	100.0	218.1	328.62	285.62	277.56	41.5869035329	128.0	197.0	156	0.00							
	100.0	225.69	430.14	329.96	343.230909091	63.5996592979	202.0	339.0	270								
	100.0	226.77	424.5	317.92	327.592727273	68.8947959627	189.0	332.0	266	300							
	100.0	228.47	431.23	313.69	318.674545455	61.8472433961	153.0	300.0	197	500							Average
	100.0	243.24	428.52	321.99	327,504545455	55.7850469316	107.0	191.0	151								
	100.0	243.28	422.13	313.55	329.319090909	58.0158958554	140.0	256.0	190 80	250							sta
	100.0	245.44	474.26	356.43	346.044545455	58,4178418206	173.0	407.0	298								
	100.0	246.03	509.92	340.3	341,333636364	82.0740493006	172.0	341.0	243	5							
	100.0	246.29	407.21	322.77	327.93	49.3383144496	137.0	237.0	182	200							
	100.0	246.45	463.52	364.97	357.074545455	63.941956117	125.0	235.0	170	,							
	100.0	251.27	403.32	367.91	352.924545455	56.5526281623	214.0	403.0	280								
	100.0		415.62	341.96	334.183636364	47.0382456137	197.0	350.0	268	150							
		265.68															
	100.0	266.1	467.04	324.99	359.060909091	77.3076385094	155.0	317.0	216								
	100.0	269.68	446.83	339.11	339.839090909	48.3933982076	138.0	239.0	177	100							
	100.0	270.26	478.13	341.18	351.183636364	59.5303369984	169.0	310.0	227		-			_			
	100.0	275.86	406.83	328.18	328.220909091	41.1776659356	155.0	226.0	192								
	100.0	278.07	510.23	368.85	371.664545455	71.0563689633	155.0	308.0	208	50							
	100.0	281.56	494.29	416.74	392.120909091	69.0054948345	91.0	190.0	148								
	100.0	286.15	532.53	395.23	399.988181818	68.3592862373	122.0	228.0	167								
	100.0	288.62	520.45	429.34	411.145454545	71.414780411	180.0	369.0	306	0							
	100.0	289.85	566.58	408.83	428.201818182	83.2768140621	78.0	150.0	97.1		6/5/2014	6/6/2014	6/7/2014	6/8/2014	6/9/2014	6/10/2014	6/11/20
	100.0	309.9	625.52	401.31	419.301818182	90.7533255908	218.0	461.0	277			-, -, **			-,-/.024	-,	77 8 87 8 9
	100.0	309.94	587.55	406.55	423.75	83.8805817815	189.0	399.0	253					Date			
	100.0	312.8	607.05	448.13	468.277272727	101.345068238	251.0	505.0	345.0		381.090909091	83,5109731989		1			
	100.0	318.12	546.1	431.42	425.593636364	73.961455843	194.0	365.0	257.0		269.090909091	49.1018506325					

Control Sample Coverage is automatically captured by Geneticist Assistant on each run providing real time review of process while developing a historical overview to highlight any changes in the process over time.

#### **Positive Control Verification**

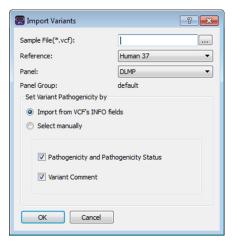
Many users opt to incorporate a positive control, such as NIST Genome in a bottle, with each sequencing run. Geneticist Assistant captures the positive control data, permitting a quick review of the run's efficacy and captures time-based data so that negative trends can be quickly observed and remedied.

C:/Users/soft/Desk	ctop/GA	/referenc	es/Huma	an 37/NISTInteg	gratedCal	ls_14datasets_	131103_
Chromosome	17	DPSum	494	PLILLWG	393,42,0	TrancheSSEmin2	0
<b>Chromosome Position</b>	63533789	HRun	2	PLIIPCRFree	1628,129,0	YesPLtot	10
ID		HapNoVar	0	PLIONEX	170,21,0	allalts	С
Ref	т	NoPLTot	0	PLPlatGen	6514,520,0	datasetcalls	11
Alt	С	PL454WG	369,39,0	PLXIII	897,72,0	geno	3
Qual	15292	PLCG	671,78,0	PLminsum	1295	genoMapGood	10
Filter	PASS	PLHSWEx	67,6,0	PLminsumOverDP	2.62	platformbias	none
HGVS Genomic		PLHSWG	918,93,0	TrancheABQDmin2	0	platformnames	ill,454,ion,cg
HGVS Coding		PLILL250	650,60,0	TrancheAlignmin2	0	platforms	4
HGVS Protein		PLILLCLIA	3015,235,0	TrancheMapmin2	0	varType	SNP

Geneticist Assistant captures positive control data which is very useful in determining efficacy of sequencing run and for determining quality trending.

# Import Existing Knowledge Base

For variants with previously determined pathogenicity, a VCF file can be imported to automatically update the pathogenicity for these variants in the Geneticist Assistant database.



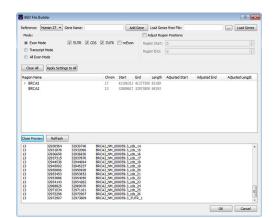
# **Flag Artifacts**

Geneticist Assistant NGS Interpretative Workbench allows users to flag variants that have been identified as artifacts and indicate the type of artifact, for example due to errors caused by chemistry or alignment. The variant can then automatically be flagged as an artifact when found in subsequent samples and can be easily filtered.



### Automatic BED file builder

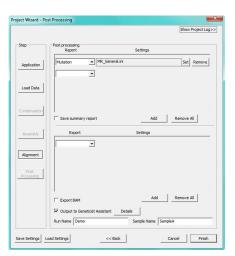
Geneticist Assistant includes the BED File Builder Tool which can be used to create custom BED files for any panel. Simply enter the name of each gene to be included, or load a text file with multiple genes, choose the desired transcript, indicate the type of regions to be included and optionally choose to include a set number of bases at either end of each region.



# **Complete Analysis Pipeline**

In conjunction with NextGENe® software

Geneticist Assistant can be used in conjunction with NextGENe's AutoRun Tool to provide a seamless pipeline for analysis, review and database submission. NextGENe can be configured to automatically access and begin processing data from the sequencing platform, and to then export results to the Geneticist Assistant database. Geneticist Assistant can also be configured to automatically import data from other analysis packages through a simple script.



# **Recommended Hardware Requirements**

Server:
2 cores
4 GB RAM
100 GB hard drive space available (solid state drive recommended)
64bit Linux (Ubuntu 12.04 or higher is recommended) or Windows Vista, 7, 8, 10 or Server 2003 through Server 2012 R2

#### Client:

2 cores 8 GB RAM 250 GB hard drive 64bit Windows Vista, 7, 8, Server 2003 through Server 2012 R2

> For more information or to arrange a free webinar or trial of **Geneticist Assistant NGS Interpretative Workbench** please visit www.softgenetics.com or email: info@softgenetics.com



www.softgenetics.com



SoftGenetics Oakwood Centre 100 Oakwood Avenue Suite 350 State College PA 16803 USA info@softgenetics.com www.softgenetics.com

© 2017 Registered Trademarks are property of their respective owners.

For Clinical Research