

GeneticistAssistant™

NGS Interpretative Workbench

Features:

Variant Database

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

Operational Management

- Customizable Workflow Builder
- Real-time Tracking and Reporting

Cool Tools

- CAP Validation Assistance
- Process Quality Control
- Positive Control Verification
- Automatic BED file builder with regions of clinical significance
- Automated Informatics Pipeline

*Developed in collaboration
with Mayo Clinic*

SOFTGENETICS®
Software PowerTools for Genetic Analysis

GeneticistAssistant™

NGS Interpretative Workbench



*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 & COSMIC

Conservation scores:

phyloP, GERP++, phastCons & SiPhy

Population frequencies:

1000 Genomes and Exome Variant Server

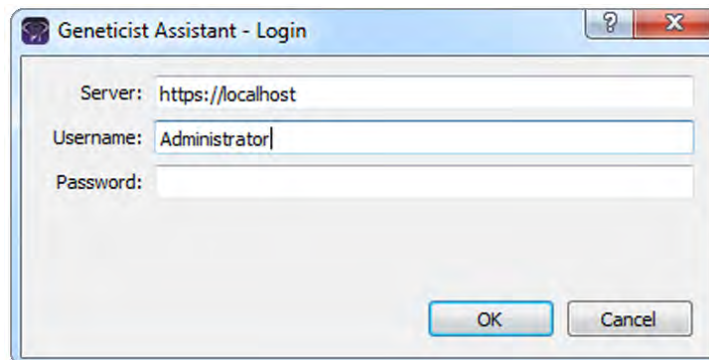
Additionally, information from proprietary databases such as **Alamut** and LOVD (Leiden Open Variation Database) are easily accessible through embedded links. Information from other publicly available databases are easily imported into the workbench.

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 **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**.

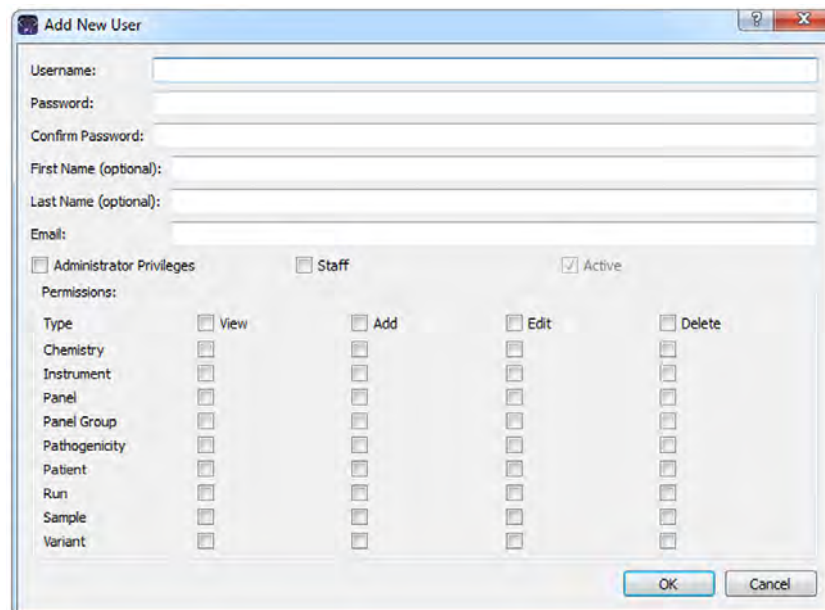
User Management, audit trail, access control

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.



The image shows a login window titled "Geneticist Assistant - Login". It contains three input fields: "Server:" with the value "https://localhost", "Username:" with the value "Administrator", and "Password:" which is empty. At the bottom right, there are two buttons: "OK" and "Cancel".

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



The image shows a "Add New User" dialog box. It has several input fields: "Username:", "Password:", "Confirm Password:", "First Name (optional):", "Last Name (optional):", and "Email:". Below these are three checkboxes: "Administrator Privileges", "Staff", and "Active" (which is checked). Under the "Permissions:" section, there are four columns of checkboxes for "View", "Add", "Edit", and "Delete". The rows are: "Type", "Chemistry", "Instrument", "Panel", "Panel Group", "Pathogenicity", "Patient", "Run", "Sample", and "Variant". At the bottom right, there are "OK" and "Cancel" buttons.

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

Variant:									
ID	14	Coverage	344	Protein	NP_001035197.1	Times Observed Per Panel	11	Ref T	
Chromosome	14	Pathogenicity	Benign	Coding Base	2476	Times Observed Per Panel Group	11		
Chromosome Position	75513883	Pathogenicity Status		Codon Position	1	Samples Per Panel	11		
Chr: ChrPos	14: 75513883	Variant Frequency	1	AA Position	826	Samples Per Panel Group	11		
Rs	rs175081	Zygosity	homo	HGVS Genomic	g.75513883T>C	Times Observed Per Patient Per Panel	0		
Ref	T	Read Balance	0	HGVS Coding	c.2476A>G	Samples Per Patient Per Panel	0		
Ref AA	Asn	Gene	MLH3	HGVS Protein	p.Asn826Asp	Patient Variant Frequency	NA		
Alt	C	Gene Strand	-	Variant Comment		Trans	Ti		
Alt AA	Asp	Exon Number	2	Times Observed Per Run	11	GMAF			
Type	missense	Transcript	NM_001040108.1	Panel	DLMP	Alt	C		
Pathogenicity Changes:									
Type	Value	User	Date	Comment					
Pathogenicity Change	Benign	Administrator	5/14/2014 10:58:18 AM						
Pathogenicity Status Change		Administrator	5/14/2014 10:58:18 AM						

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



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/Desktop/GA/references/Human 37/ESP6500SI-V2-SSA137.vcf											
Chromosome	17	AA	G	EA_AC	6409,2191	HGVS_CDHA_VAR	NM_001126118.1:c.98C>G,NM_001126114.2:c.1126113.2:c.215C>G,NM_001126112.2:c.2146.5:c.215C>G				
Chromosome Position	7579472	AA_AGE	1784,2620	EA_AGE	2409,1591,300	HGVS_PROTEIN_VAR	NM_001126118.1:p.(P33R),NM_001126114.2:p.(P72R),NM_001126113.2:p.(P72R),NM_001126112.2:p.(P72R),NM_000546.5:p.(P72R)				
ID	rs1042522	AA_GTC	.	EA_GTC	yes	MAF	25.4767,40.5086,36.9963				
Ref	G	CA	386,1012,804	EXOME_CHIP	yes	PH	possibly-damaging:0.745,possibly-damaging:0.745,possibly-damaging:0.745,possibly-damaging:0.745,possibly-damaging:0.745				
Alt	C	CA	http://www.ncbi.nlm.nih.gov/sites/arcv/gene	FG	NM_001126118.1:missense,NM_001126114.2:missense,NM_001126113.2:missense,NM_001126112.2:missense,NM_000546.5:missense	TAC	8193,4811				
Qual	0	CDS_SIZES	1.9	GL	TP53						
Filter	PASS	CG	1.9	GS	103,103,103,103						
HGVS Genomic	CG	CP	0.0	GTC	2795,2603,1104						
HGVS Coding	DBSNP	dbSNP_86		GTS	CC,CG,GG						
HGVS Protein	DP	92		GWAS_PUBMED	.						
C:/Users/soft/Desktop/GA/references/Human 37/clinvar_00-latest.vcf											
Chromosome	17	CAF	[0,3981,0,6019]	CLNSRCID	,2077 191170,0005	LSD	LSD	SLO	SLO		
Chromosome Position	7579472	CLNACC	RCV000013144.1 RCV000034639.1 RCV000079202.1	COMMON	1	OM	OM	SSR	SSR		
ID	rs1042522	CLNALLE	1	GS	GS	OTHERKG	OTHERKG	TPA	TPA		
Ref	G	CLNDBN	CODON_72_POLYMORPHISM(yz2c_fr1042522) not_provided AllHighlyPenetrant	GSA	GSA	PH3	PH3	VC	SNV		
Alt	C	CLNDSDB	. . MedGen	GENEINFO	TP53;7157	PH	PM	VLD	VLD		
Qual	0	CLNDSDBID	. . Ch169374	GNO	GNO	PMC	PMC	VP	0x05017800000017051f110101		
Filter	.	CLNHGVS	NC_000017.10:p.7579472G>C	HD	HD	RS	1042522	WGT	0		
HGVS Genomic	CLNHORIGIN	1		KGPROD	KGPROD	RSPOS	7579472	dbSNPBuildID	86		
HGVS Coding	CLNSIG	2 2 2		KGPhase1	KGPhase1	RV	RV				
HGVS Protein	CLNSRC	. Emory_University OMIM_Allelic_Variant		KGPlot123	KGPlot123	SAO	1				
C:/Users/soft/Desktop/GA/references/Human 37/dbNSFP2.4_variant.chr17											
chr	17	Uniprot_aapo	72;72;72;72	aapos	72	Polyphen2_HVAR	0.41988	MutationAssessor	ssor_pred	Reliability_index	9
pos(1-coor)	7579472	s		aapos_SIFT	ENSP0000000269	rankscore		FATHMM_pre		CADD_raw	0.823228
ref	G	Interpro_dom		aapos_FATH	305:P72R	Polyphen2_HVAR	B;B;B;B	re		CADD_raw_ranksc	0.16678
alt	C	ain		HH	ENSP000000410	pred		ore		ore	
aaaref	P	cds_strand	-		739:P72R;ENSP	LRT_score	0.370853			CADD_phred	8.316
aaalt	R	refcodon	CCC		00000352610:P	LRT_converted_ra	0.04441			GERP++_NR	1.87
hg18_pos(1-coor)	7520197	SLR_test_statistic			72R;ENSP00000	nkscore				GERP++_RS	1.87
genename	TP53	codonpos	2		0269305:P72R;	LRT_pred	U			GERP++_RS_rank	0.25490
Uniprot_acc	ENSR66P0463	fold-degenerate	0		ENSP000000398	MutationTaster_sc	0.0			phastCons46	0.002000
	7-2:P04637-3:P				846:P72R;ENSP	ore				way_primate	0.02628
					00000391127:P	MutationTaster_c	0.80722			phastCons46	0.02628
					72R;ENSP00000	ore				way_primate	0.02628
Samples Associated With Variant:											
ID	Name	Run Date/Time	Add Date/Time	Run	Panel	PanelGroup	Reference	Patient External ID	Status	Variant File	Coverage/Pile Up File
2	800463.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:09:54 AM	Demo	DLMP	default	Human 37		New	C:/Users/soft/Desktop/GA/Mayo_data/800463.variants.filter.vcf	C:/Users/soft/Desktop/GA/Mayo_data/800463.variants.filter.pileup
3	800418.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:11:19 AM	Demo	DLMP	default	Human 37		New	C:/Users/soft/Desktop/GA/Mayo_data/800418.variants.filter.vcf	C:/Users/soft/Desktop/GA/Mayo_data/800418.variants.filter.pileup
4	800458.variants.filter	5/14/2014 11:05:33 AM	5/14/2014 11:12:53 AM	Demo	DLMP	default	Human 37		New	C:/Users/soft/Desktop/GA/Mayo_data/800458.variants.filter.vcf	C:/Users/soft/Desktop/GA/Mayo_data/800458.variants.filter.pileup

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

Sources included:

Variant Interpretation:

dbSNP
Exome Variant Server

Functional Prediction:

SIFT
PolyPhen-2
LRT
MutationTaster
MutationAssessor
FATHMM
CADD

Conservation Scores:

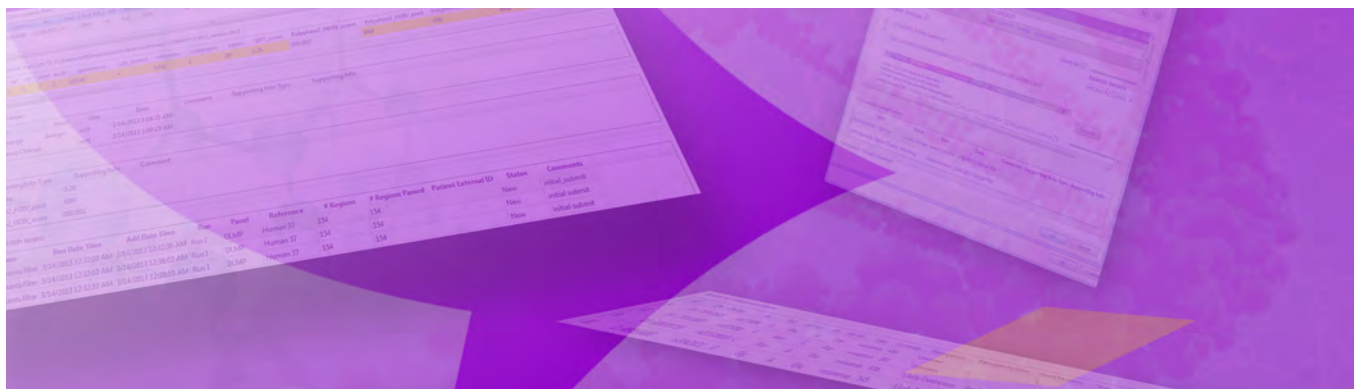
phyloP
phastCons
GERP++
SiPhy

Disease Association:

COSMIC
ClinVar
Alamut (license required)
LOVD (Leiden Open Variation Database)
And others

Population frequencies:

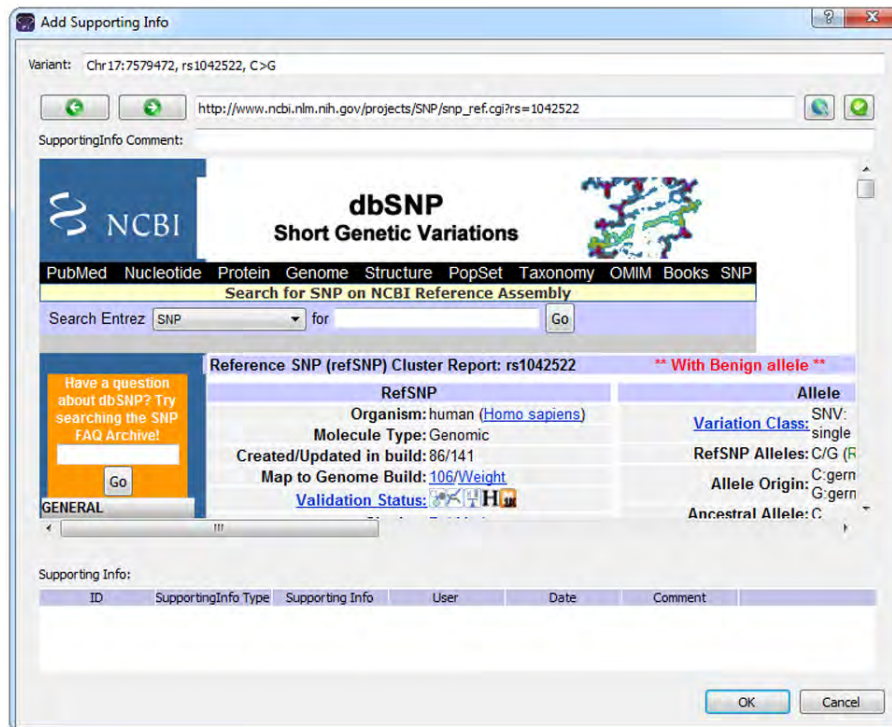
1000 Genomes
Exome Variant Server



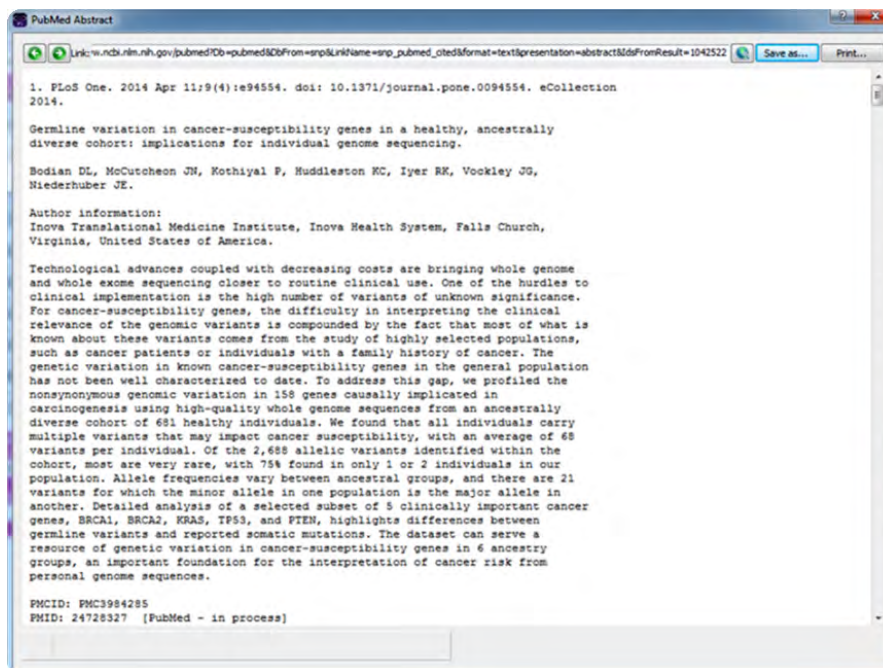
Pathogenicity Call Supporting Information

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.



A simple right mouse click enters information and comments from multiple databases in support of pathogenicity call into Geneticist Assistant Interpretative Workbench.



PubMed abstracts can be automatically downloaded into the workbench.



Linkage to External Databases

Retrieving further information from external proprietary databases such as Alamut or the LOVD database is a simple click away. (Alamut requires a license)

Variants of '272305.variants.filter': *Filters Applied								
ID	Chr : ChrPos	Rs	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	
45	2 : 48010488	rs1042821	Unknown	MSH6	1	<div>Variant Details</div> <div>View In Alamut</div> <div>View PubMed Abstract</div> <div>Delete Variant</div> <div>Add Selected Variants to Action List</div> <div>Add Selected Variants to Custom Report</div> <div>Export Selected to VCF</div> <div>Add All Variants to Custom Report</div> <div>Add Comment</div> <div>Update Pathogenicity</div> <div>Update Pathogenicity Status</div> <div>Select Columns</div>		
38	14 : 75513828	rs175080	Unknown	MLH3	2			
7	5 : 112162854	rs2229992	Likely Deleterious	APC	12			
9	10 : 88635779	rs11528010	Likely Deleterious	BMPRI1A	3			
17	17 : 63533789	rs9915936	Likely Benign	AXIN2	6			
18	17 : 63554591	rs2240308	Likely Benign	AXIN2	2			
15	17 : 7579472	rs1042522	Deleterious	TP53	4			
16	17 : 63533768	rs1133683	Deleterious	AXIN2	6			
8	5 : 112164561	rs351771	Benign	APC	14			
14	14 : 75513883	rs175081	Benign	MLH3	2			
Status Changes of '272305.variants.filter':								
Type	Value	User	Date					
Sample Status Change	New	Administrator	5/14/2014 11:24:08 AM					
Patient of '272305.variants.filter':								

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.

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.

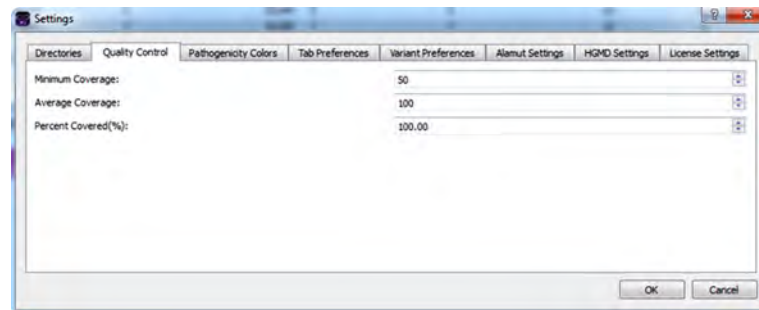
Variants of '272305.variants.filter':													
ID	Chr : ChrPos	Rs	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVs Protein	Panel	HGVs Coding	Times Observed Per Panel	Times Observed Per Run
7	5 : 112162854	rs2229992	Likely Deleterious	APC	12	synonymous 0.5	60	60	p.Tyr485Ile	DLMF	c.1458T>C	10	10
8	5 : 112164561	rs351771	Benign	APC	14	synonymous 0.5	60	60	p.Ala545Ser	DLMF	c.835G>A	10	10
9	10 : 88635779	rs11528010	Likely Deleterious	BMPRI1A	3	missense 1	99	99	p.Pro27Thr	DLMF	c.4C>A	5	5
14	14 : 75513883	rs175081	Benign	MLH3	2	missense 1	55	55	p.Asn26Asp	DLMF	c.2476A>G	11	11
15	17 : 7579472	rs1042522	Deleterious	TP53	4	missense 1	46	46	p.Pro72Arg	DLMF	c.215C>G	9	9
16	17 : 63533768	rs1133683	Deleterious	AXIN2	6	synonymous 0.5	50	50	p.Pro462Leu	DLMF	c.1386C>T	8	8
17	17 : 63533789	rs9915936	Likely Benign	AXIN2	6	synonymous 0.5	54	54	p.Pro435Ser	DLMF	c.1365A>G	9	9
18	17 : 63554591	rs2240308	Likely Benign	AXIN2	2	missense 1	111	111	p.Pro90Ser	DLMF	c.148C>T	8	8
38	14 : 75513828	rs175080	Unknown	MLH3	2	missense 1	55	55	p.Pro844Leu	DLMF	c.2531C>T	6	6
45	2 : 48010488	rs1042821	Unknown	MSH6	1	missense 1	64	64	p.Gly39Glu	DLMF	c.1160A>A	3	3

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.

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.



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	% Covered	Average Coverage	Minimum Coverage	Status	Average % Covered	Average Average Coverage	Average Minimum Coverage	Passed	Passed Percent	Failed	Total
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®, 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.

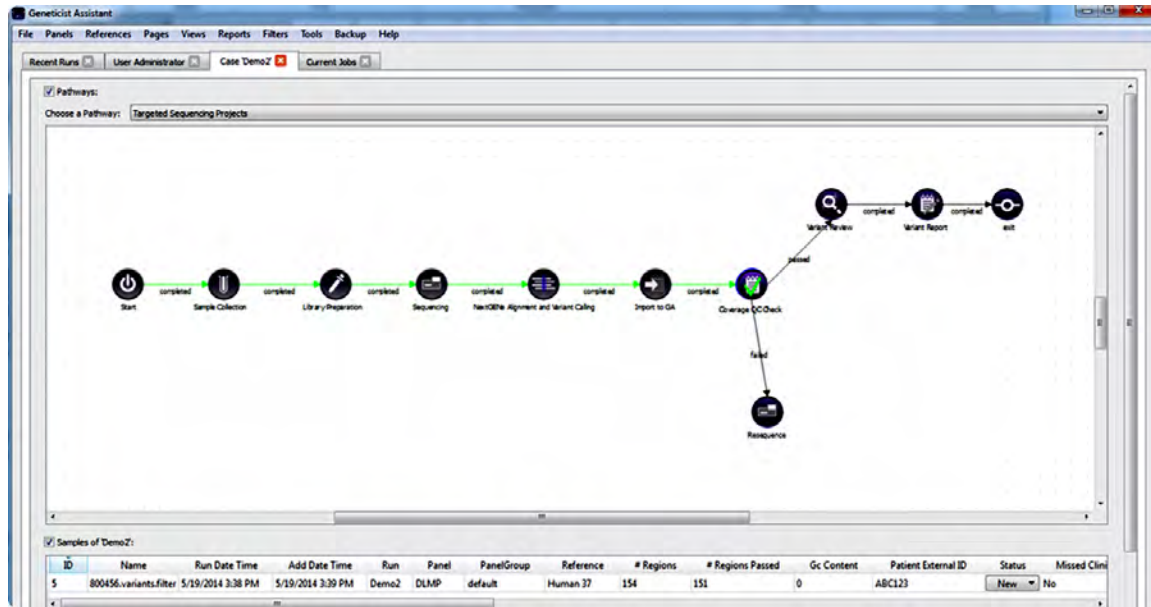


1. Database and Client may reside on single computer
2. Geneticist Assistant can be accessed by any computer having client within institution network
3. Off-site collaborators or sister facilities can securely (HTTPS security protocol) access the database via the internet.

Operational Management

Customizable Workflow Builder

Geneticist Assistant NGS Interpretative Workbench now includes a completely customizable workflow builder that enables you to model your physical NGS workflow. The program automatically **tracks, in real-time**, patient samples from receipt through final reports, providing a complete overview of department by department workload inventory, production statistics and trending. The administration module features **automatic email notification**, identifies regions of low performance and **provides a sound basis for ongoing process improvement**.



Geneticist Assistant NGS Interpretative Workbench features a completely customizable workflow builder that enables users to model physical workflow in “silico”. The software monitors and provides real time, department by department reporting of current sample load & status, captures historical production data for each department, enabling managers to quickly review status of samples and production trending.

Real-time Tracking and Reporting

User: Administrator

Responsible Cases: 3

Watching Cases: 0

Open Cases: 2

Old Open Cases: 0

Closed Cases:

30 Days

Previous 30 Days

2014:

1

0

2013:

0

0

Cases at each step :

Cases

Average Time

Overall:

Total Cases: 3

Open Cases: 2

Old Open Cases: 0

Closed Cases:

30 Days

Previous 30 Days

2014:

1

0

2013:

0

0

Cases at each step :

Cases

Average Time

Targeted Sequencing Projects

Coverage QC Check: 1

Test Pathway

Lab: 1

Administrator: Open Cases:

Case	Due Date	Priority	Start Date	Time Since Start	Owner
Demo		Normal	5/19/2014 11:40 AM	3 hours 20 minutes	Administrator
Demo2	5/22/2014	Normal	5/19/2014 1:16 PM	1 hours 44 minutes	Administrator

Geneticist Assistant provides real time statistics by department and for the complete system. Data is available to all authorized personnel, eliminating the need to manually research the status of each sample. Varying priorities can be set on a per sample basis, historical information by department or system wide is available by time period.



Email alert of a sample status change is easily accomplished with Geneticist Assistant.

Cool Tools

CAP Validation Assistance

File

Panel

References

Views

Reports

Filters

Tools

Backup

Help

Current Job

Recent Runs

Historical Performance

Start: 4/12/2013

End: 3/1/2014

Period: 1

Months

Filter:

Template Usage

	3/2014	2/2014	1/2014	12/2013	11/2013	10/2013	9/2013	8/2013	7/2013	6/2013	5/2013	4/2013
Chr Template	7	1	0	13	12	26	880548755	5	3	2	0	13
Codon Cancer Template	0	0	1	10	7	7602384	880548935	0	0	0	1777719784	1877225
Median Template	8	9	1	1	5	7289961	880548835	5	10	2	2	8
Paired-End Template	3	2	0	11	8	7602384	880548863	5	8	7	1	3
Default Template	10	12	2	9	11	8553714	880548835	5	0	4	11	5

Historical Performance

	3/2014	2/2014	1/2014	12/2013	11/2013	10/2013	9/2013	8/2013	7/2013	6/2013	5/2013	4/2013
Cases	16	17	13	13	13	13	13	14	14	16	16	17
Average Variants	12	15	18	17	16	16	16	13	13	16	17	18
Average Coverage	114	110	115	117	115	119	113	113	119	111	111	118
Average Alignment Time	1411	1143	117	2338	1032	627	151	742	1436	1711	137	619
Average Perfectly Aligned	45329	45379	45617	45739	45471	45415	45601	45268	45208	45185	45102	45103
Average Aligned Reads	75805	75809	758176	758163	758084	758622	758129	758129	758129	758129	758084	758081
Average Unaligned Reads	14880	13080	13112	13118	13147	13115	13088	13162	13111	14819	13464	13419
Average Reads	76241	76181	76121	76129	76129	76141	76129	76129	76129	76129	76129	76129
Average Filtered Reads	689	689	679	679	689	672	678	678	689	678	679	689
Average Paired Reads	76723	77163	76525	76575	76586	77177	76516	77163	76516	76586	76586	76586
Average Deletions	16213	16887	16270	16555	16500	16388	16729	15963	16619	15946	16561	15914
Average Insertions	278	282	276	284	281	271	253	262	246	282	268	252
Average Mutations	8874	8878	8872	8869	8867	8858	8862	8868	8864	8864	8877	8857
Average Covered Sites	7611	7617	7615	7618	7619	7611	7611	7614	7619	7611	7619	7618

In a single report, Geneticist Assistant NGS Interpretative Workbench gathers quality statistics over time, often required during a CAP validation. Frequent review of the quality statistics often reveals negative trends permitting immediate corrective actions to be taken.

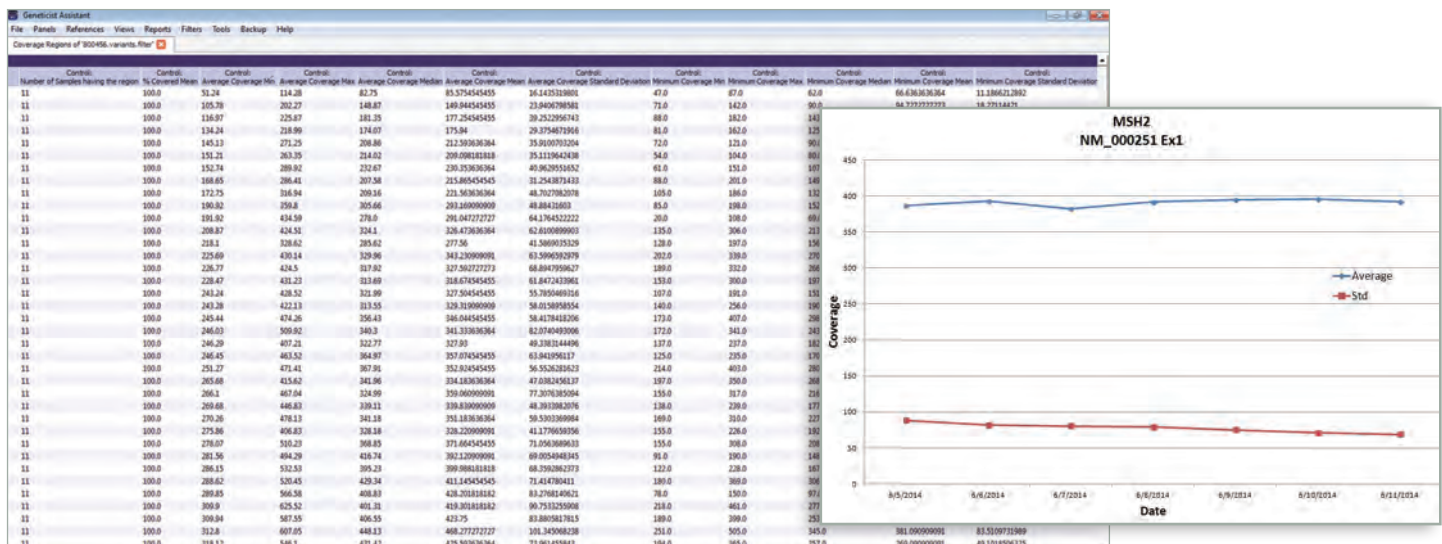
Process Quality Control

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 Filters Tools Backup Help												
Sample: 100-456-variants.filter												
	Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:	Control:
	Number of Samples Having the Variant	Coverage Min	Coverage Max	Coverage Median	Coverage Mean	Coverage Standard Deviation	Variant Frequency Min	Variant Frequency Max	Variant Frequency Median	Variant Frequency Mean	Variant Frequency Standard Deviation	Variant Frequency Standard Deviation
1	209.0	242.0	254.0	273.87142357	42.6320290728	0.5	0.5	0.5	0.5	0.0	0.0	0.0
4	219.0	316.0	389.0	276.0	43.707044774	0.5	0.5	0.5	0.5	0.0	0.0	0.0
5	299.0	374.0	328.0	325.8	27.4765354439	0.5	0.5	0.5	0.5	0.0	0.0	0.0
8	171.0	204.0	178.5	183.0	13.5830777072	0.5	0.5	0.5	0.5	0.0	0.0	0.0
9	522.0	958.0	624.0	658.375	135.85026644	0.5	1.0	1.0	1.0	0.8125	0.242061459138	0.242061459138
5	344.0	513.0	447.0	430.0	72.420081432	0.5	1.0	0.5	0.5	0.7	0.24498974278	0.24498974278
9	335.0	454.0	368.0	383.8	44.875162338	0.5	0.5	0.5	0.5	0.0	0.0	0.0
5	318.0	449.0	403.0	390.4	52.6026615296	0.5	0.5	0.5	0.5	0.0	0.0	0.0
8	252.0	368.0	318.0	310.875	49.8208074889	0.5	1.0	0.75	0.75	0.25	0.25	0.25
11	418.0	614.0	480.0	566.181818182	76.5124091918	0.5	1.0	1.0	1.0	0.890909090909	0.13047330396	0.13047330396
10	274.0	514.0	387.5	382.2	68.2071150462	0.5	1.0	1.0	1.0	0.85	0.229126784748	0.229126784748
10	349.0	580.0	456.0	462.4	71.2229642098	0.5	1.0	0.5	0.5	0.7	0.24498974278	0.24498974278
11	368.0	567.0	422.0	452.636363636	72.0407388688	0.5	1.0	1.0	1.0	0.909090909091	0.13047330396	0.13047330396
10	259.0	407.0	331.5	329.5	53.2257962277	1.0	1.0	1.0	1.0	1.0	1.0	1.0
12	351.0	564.0	371.5	385.416666667	78.6679552911	1.0	1.0	1.0	1.0	1.0	1.0	1.0
12	294.0	501.0	367.0	379.25	61.7833817813	1.0	1.0	1.0	1.0	1.0	1.0	1.0

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



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

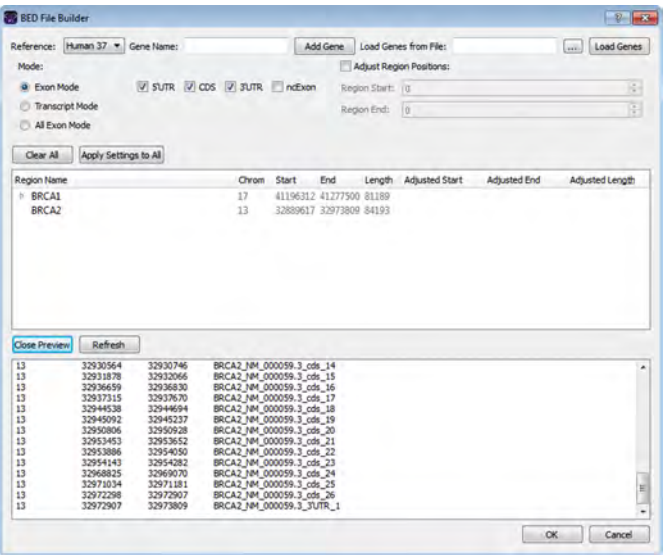
Many users opt to incorporate a positive control, such as NIST Genome in a bottle, with each sequencing run. Geneticist Assistant captures the positive control data, permitting a quick review of the run's efficacy and captures time-based data so that negative trends can be quickly observed and remedied.

C:/Users/soft/Desktop/GA/references/Human 37/NISTIntegratedCalls_14datasets_131103

Chromosome	17	DPSum	494	PLILLWG	393,42,0	TrancheSSEmin2	0
Chromosome Position	63533789	HRun	2	PLIILPCRFree	1628,129,0	YesPLtot	10
ID	.	HapNoVar	0	PLIonEx	170,21,0	allalts	C
Ref	T	NoPLTot	0	PLPlatGen	6514,520,0	datasetcalls	11
Alt	C	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.

Automatic BED file builder

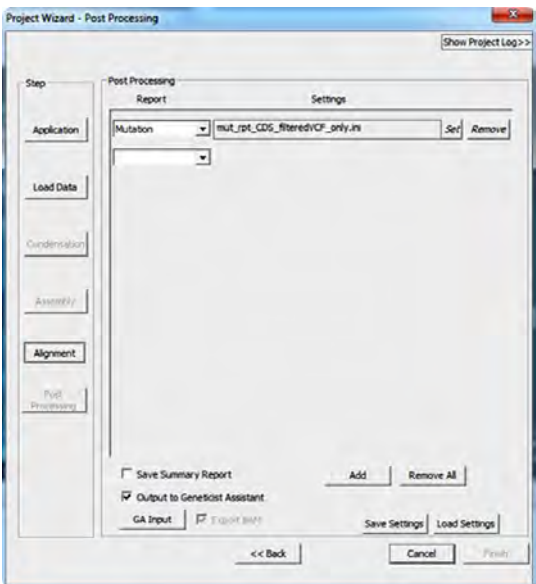


Geneticist Assistant includes the BED File Builder Tool which can be used to create custom BED files for any panel. Simply enter the name of each gene to be included, or load a text file with multiple genes, choose the desired transcript, indicate the type of regions to be included and optionally choose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



Recommend Hardware Requirements

Server:

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

Client:

2 cores
2 GB RAM
100 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

SOFTGENETICS®
Software PowerTools for Genetic Analysis

www.softgenetics.com

SOFTGENETICS®

Software PowerTools for Genetic Analysis

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

© 2014 Registered Trademarks are property of their respective owners.

For Research Use Only