

Geneticist Assistant™

NGS Interpretative Workbench

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*

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, 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**.

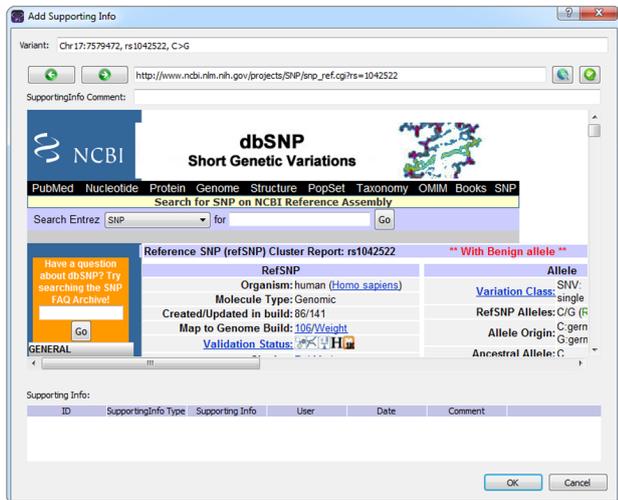
**Requires separate license*



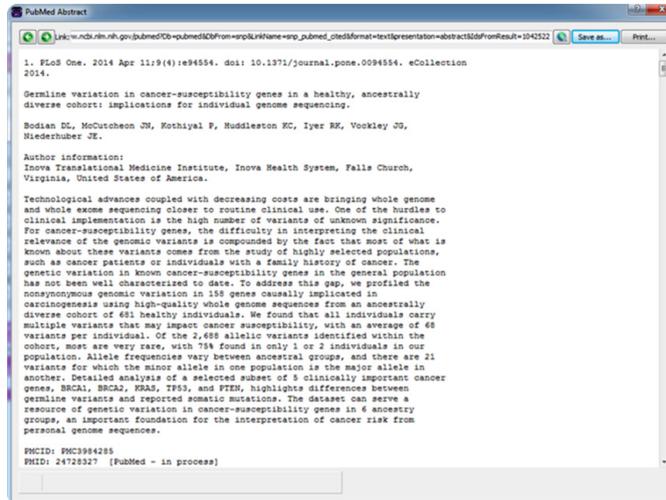
Pathogenicity Call Supporting Information

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 tools such as Alamut, UCSC Genome Browser, or the LOVD database is a simple click away. (Alamut requires a license)

Variants of '800463.igv-sorted_Output_Mutation_Report1_filtered': *Filters Applied						
ID	Chr : ChrPos	Rs	Pathogenicity	Pathogenicity Status	Gene	HGVs Coding
28	3:37056045	rs182733777	Unassigned			90+10A>G
29	3:37081751	rs267607840	Unassigned			633A>G
6	3:37083740	rs9876116	Benign			668-19A>G
13	14:75505016	rs175075	Benign			
14	14:75513883	rs175081	Benign			
31	14:75514489	rs28756986	Unassigned			
15	17:7579472	rs1042522	Deleterious			
1443	17:7579669	rs17878362	Unassigned			6+32_96+47delGGCTGGGGA...
18	17:63554591	rs2240308	Likely Deleterious	Confirmed	AXIN2	NM_004655.3:c.148C>T
19	18:48584856	rs386387676	Likely Benign		SMAD4	NM_005359.5:c.904+45_904+46insTT

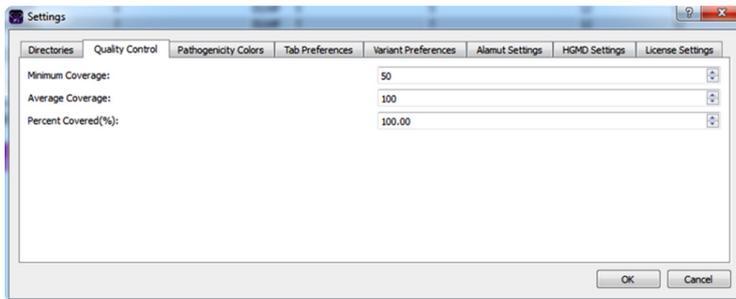
Alamut licensees can quickly retrieve information without error prone and tedious retyping by simply selecting variant of interest and clicking on the drop down menu.

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.



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

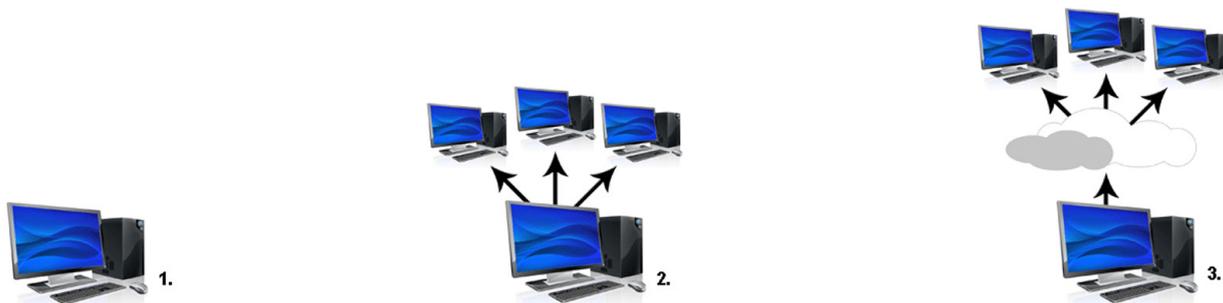
Quality data is presented for both the current sample and a complete history of analysis of all samples for a disease panel. Metrics provided include Minimum Coverage, Average Coverage, % Coverage Across Region and Pass/Fail Status of current run. Historical data includes average coverage of all runs, 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.

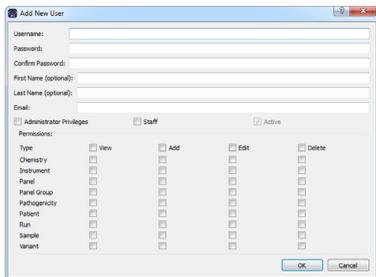
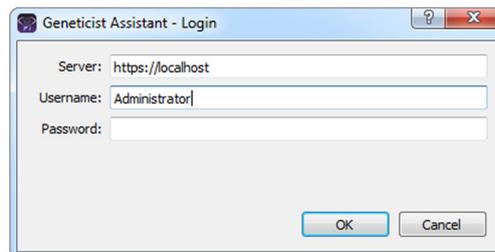


1. Database and Client may reside on single computer
2. Geneticist Assistant can be accessed by any computer having client within institution network
3. Off-site collaborators or sister facilities can securely (HTTPS security protocol) access the database via the internet.

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.



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.

Variant:							
ID	14	Coverage	344	Protein	NP_001035197.1	Times Observed Per Panel	11
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	rs126381	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	DUMP	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		

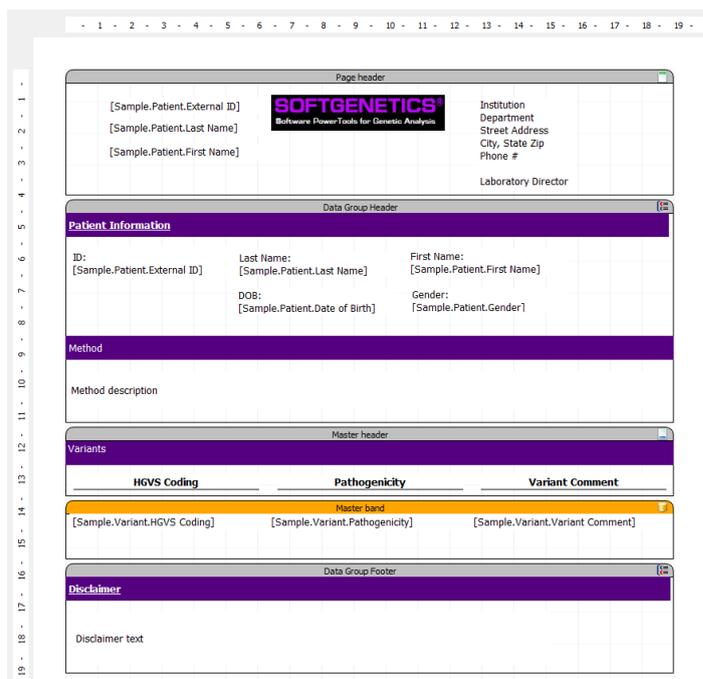


Cool Tools

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.



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

Packing Slip 23 Jan 2017 10:43:33 AM

Patient External Id 800426 Mother 800418 Father 800402 Gender Female	Ordering Provider First Name Sue Last Name Smith Phone 1234567890 Email sue@hospital.org Street 2 Main St City PA State PA Zip 16803
Specimen Date Collected 2017-01-23 Units 15 Sample Type Blood	Genetic Counselor First Name John Last Name Jones Phone 3456789012 Email John@counselor.org
Tests Requested Tests 1	Referring Hospital Or Laboratory Organization Oncology Lab Phone 6789012345 Street 5 Main St City PA State PA
Reason For Testing Indication unknown Clinical Diagnosis Breast tumor	

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

Add New Patient

Relationship Column Starts at: 8

Add New Phenotype Terms in File Automatically

Data Delimiter: Comma(,) Tab("\t")

Date of Birth Format: MM/dd/yyyy Example: 09/23/2015 or 9/23/2015

File Column Format: File follows standard column order File follows standard column names Files does not follow above standards, I want to specify...

Result:

patient_id	patient_last_name	patient_first_name	patient_dob	patient_gender	patient_race	patient_phenotype	patient_relationships
------------	-------------------	--------------------	-------------	----------------	--------------	-------------------	-----------------------

OK Cancel

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.

Family Comparison

Phenotype:

Inheritance pattern: Autosomal Recessive

Compound heterozygous

Show columns: AF Cov RB

Filter variants by panel: Build37_CCDS_Exons_MergeOverlaps

Sample ID	Sample Name	Patient External ID	Relationship	Phenotype	Zygoty
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

OK Cancel

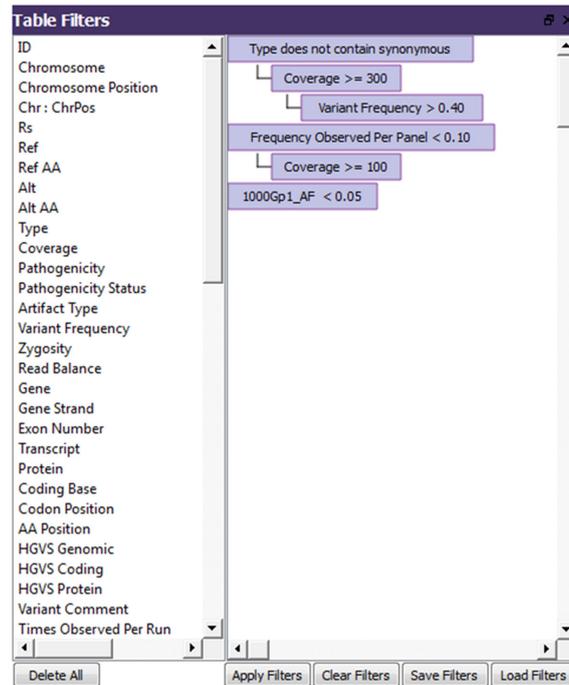
For family comparisons, specify the relationships and phenotypes for each patient to create a comparison based on a selected inheritance pattern.

Family Comparison of 4 Samples:												ID#	Chr	ChrPos	Rs	Gene	RefAA	Alt AA	Type	Coverage	Pathogenicity	HGVSc Coding
AF	Cov	AF	Cov	AF	Cov	AF	Cov	AF	Cov	AF	Cov											
0.494	166	0.437	119	1.000	121	0.984	127	1640	1:114443899	rs17464525	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	1:108307727	rs7528153	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	2:206911228	rs2909111	IN080D	Ala	Val	missense	7	Deleterious	NM_017759.4:c.1073C>T					
0.636	11	0.667	24	0.773	22	0.800	10	3306	2:111304496	rs71231856	RGP06	Val	Val	synonymous	22	Benign	NM_001123363.3:c.1560G>A					
0.471	121	0.506	85	1.000	97	1.000	79	3382	2:210557380	rs6720659	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	2:210557542	rs741007	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	2:210558162	rs741006	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	2:210559960	rs2239672	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	2:211456637	rs1047883	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	2:211456639	rs2229589	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	2:211481257	rs2287599	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	2:159663616	rs10497199	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

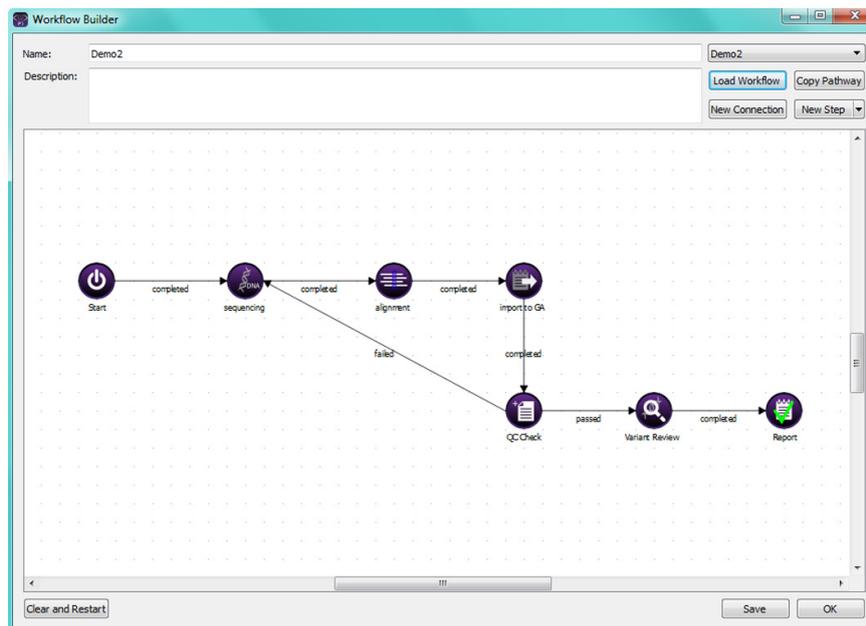
Variants lists, as well as any other data tables in Geneticist Assistant, can be filtered based on a combination of any data fields.



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".

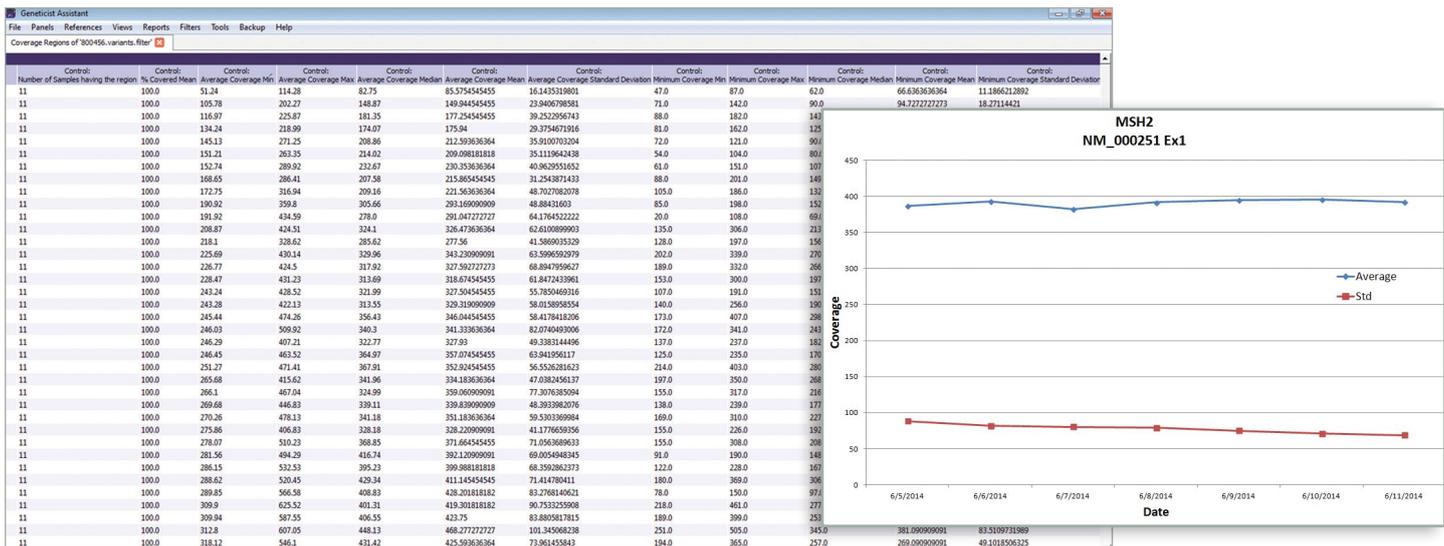
Process Quality Control

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.

Control:	Control:	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			
7	209.0	342.0	254.0	273.857142857	42.6930150728	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	320.0	325.8	27.4765334429	0.5	0.5	0.5	0.5	0.0			
4	171.0	204.0	178.5	183.0	13.583077072	0.5	0.5	0.5	0.5	0.0			
8	522.0	950.0	624.0	658.375	135.85824644	0.5	1.0	1.0	0.8125	0.242061459138			
5	344.0	513.0	447.0	430.0	72.4209991432	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	49.8300714889	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	72.0407368898	0.5	1.0	1.0	0.909090909091	0.19284730396			
10	259.0	316.0	328.5	328.5	53.237962277	1.0	1.0	1.0	1.0	0.0			
12	301.0	564.0	371.5	385.416666667	78.6675052021	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.



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

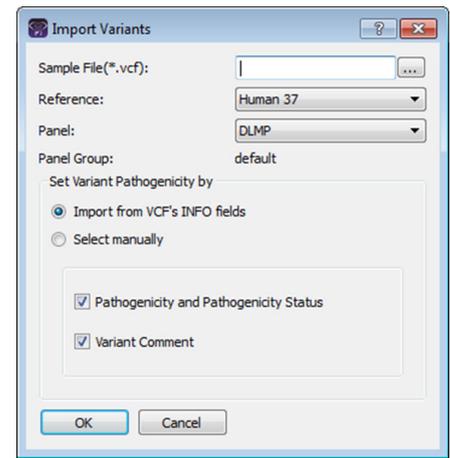
Many users opt to incorporate a positive control, such as NIST Genome in a bottle, with each sequencing run. Geneticist Assistant captures the positive control data, permitting a quick review of the run's efficacy and captures time-based data so that negative trends can be quickly observed and remedied.

C:/Users/soft/Desktop/GA/references/Human 37/NISTIntegratedCalls_14datasets_131103					
Chromosome	17	DPSum	494	PLILLWG	393,42,0
Chromosome Position	63533789	HRun	2	PLIIPCRFree	1628,129,0
ID	.	HapNoVar	0	PLIonEx	170,21,0
Ref	T	NoPLTot	0	PLPatGen	6514,520,0
Alt	C	PL454WG	369,39,0	PLXIII	897,72,0
Qual	15292	PLCG	671,78,0	PLminsum	1295
Filter	PASS	PLHSWEx	67,6,0	PLminsumOverDP	2.62
HGVS Genomic		PLHSWG	918,93,0	TrancheABQDmin2	0
HGVS Coding		PLILL250	650,60,0	TrancheAlignmin2	0
HGVS Protein		PLILLCLIA	3015,235,0	TrancheMapmin2	0
				TrancheSSEmin2	0
				YesPLtot	10
				allalts	C
				datasetcalls	11
				geno	3
				genoMapGood	10
				platformbias	none
				platformnames	ill,454,ion,cg
				platformnames	4
				varType	SNP

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

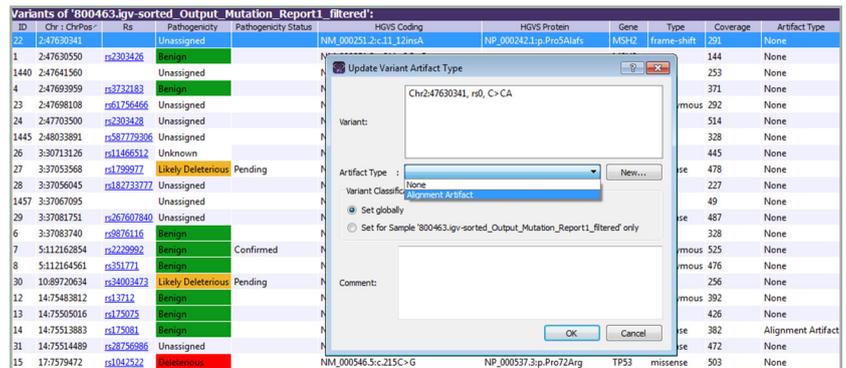
Import Existing Knowledge Base

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.



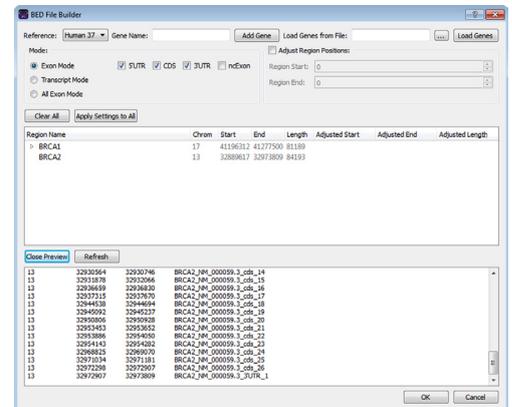
Flag Artifacts

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

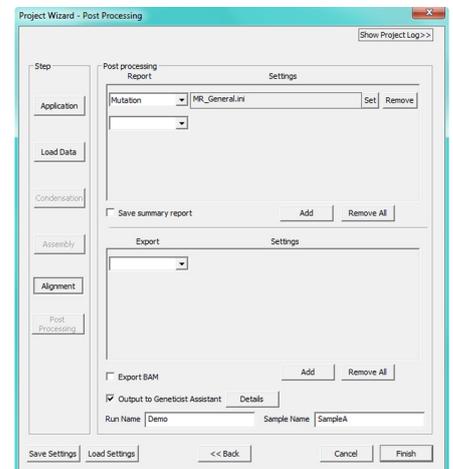
Geneticist Assistant includes the BED File Builder Tool which can be used to create custom BED files for any panel. Simply enter the name of each gene to be included, or load a text file with multiple genes, choose the desired transcript, indicate the type of regions to be included and optionally choose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

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

© 2017 Registered Trademarks are property of their respective owners.

For Clinical Research