SOFTGENETICS®

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

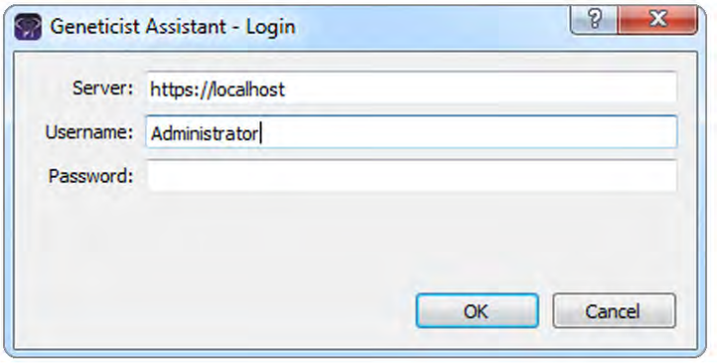
Additionally, information from proprietary databases such as **Alamut** and LOVD (Leiden Open Variation Database) are easily accessible through embedded links. Information from other publicly available databases are easily imported into the workbench.

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.

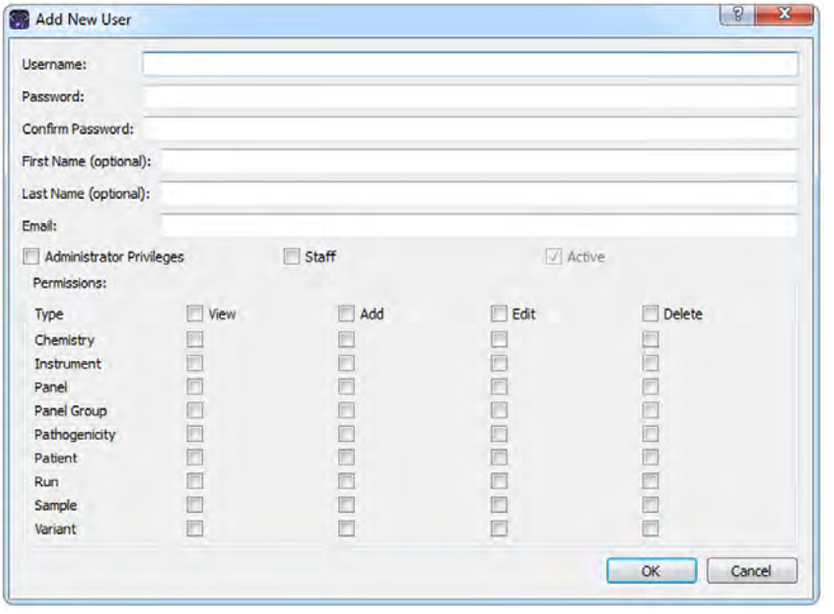
Unique tools include **CAP Validation Assistance, automated BED file builder** which automatically highlights areas of clinical significance, **Positive Control Verification**, and in conjunction with NextGENe software can form a completely **automated informatics pipeline**.

User Management, audit trail, access control

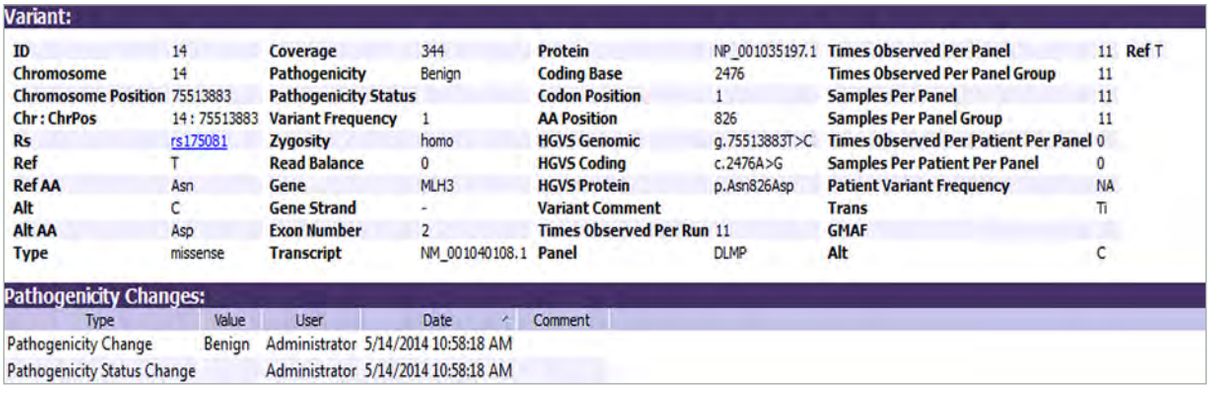
Geneticist Assistant NGS Interpretative database employs a customizable password system (such as an 8 character alpha-numeric password) to protect data integrity. Database records all log-in and log-off and all user-activity by user, which can be recalled by administrative personnel. Access to various 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.



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



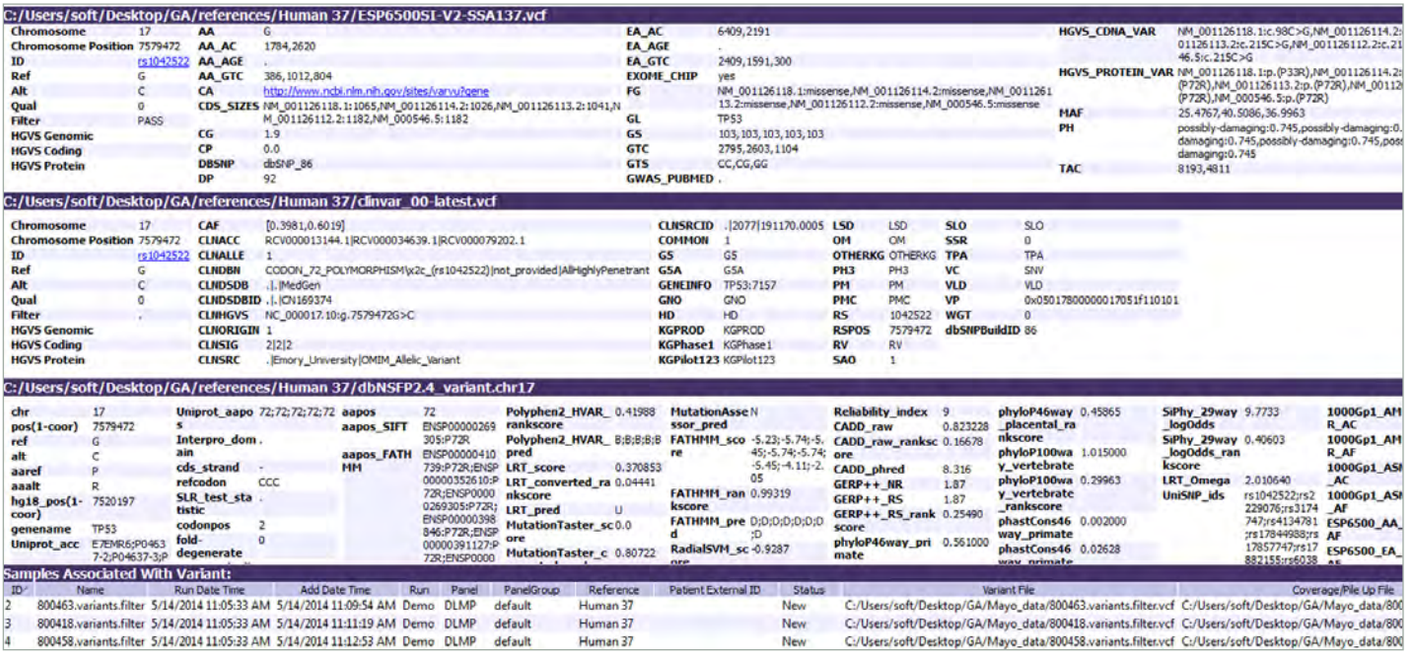
Access to various information within Geneticist Assistant can be granted by individual and groups.



Geneticist Assistant Workbench records and tracks all changes and comments made to the database by 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.



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

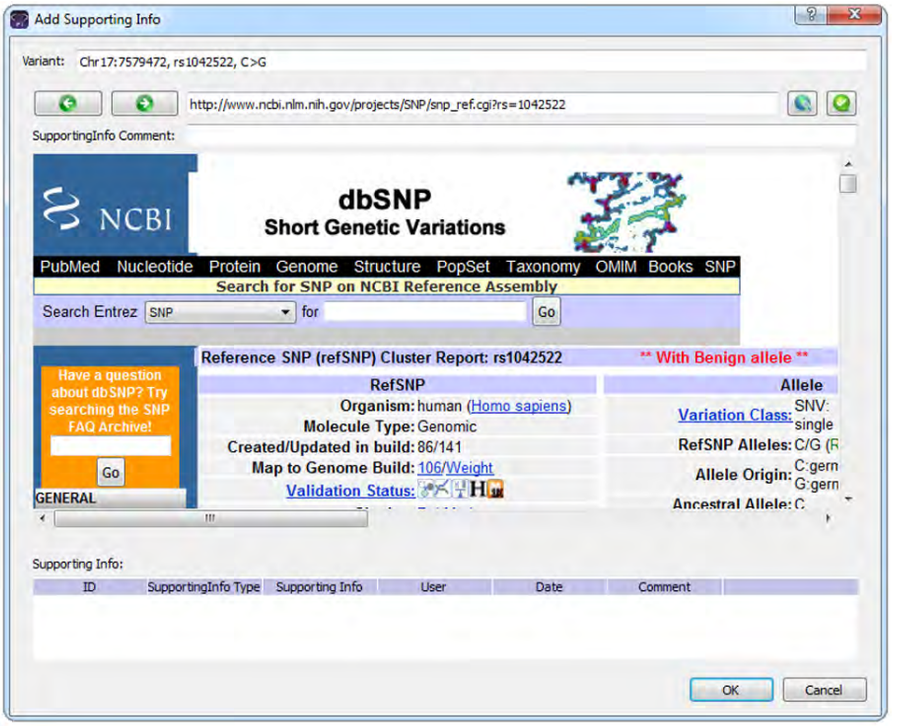
Disease Association:
COSMIC
ClinVar
Alamut (license required)
LOVD (Leiden Open Variation Database)
And others

Population frequencies:
1000 Genomes
Exome Variant Server

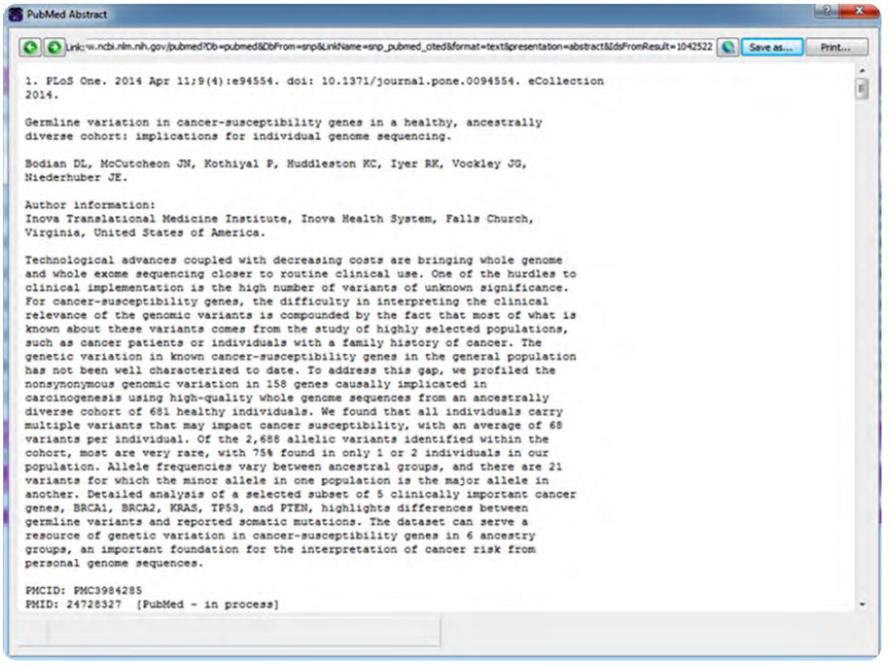
Pathogenicity Call Supporting Information

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.



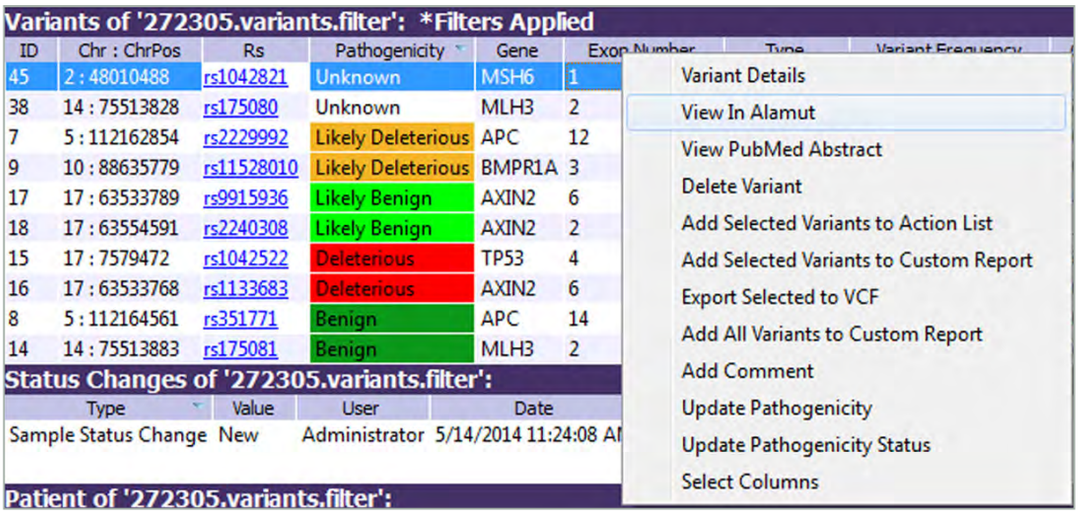
A simple right mouse click enters information and comments from multiple databases in support of pathogenicity call into Geneticist Assistant Interpretative Workbench.



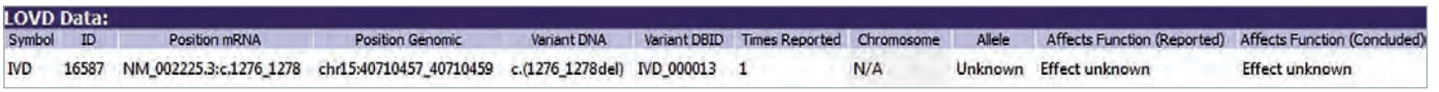
PubMed abstracts can be automatically downloaded into the workbench.

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)



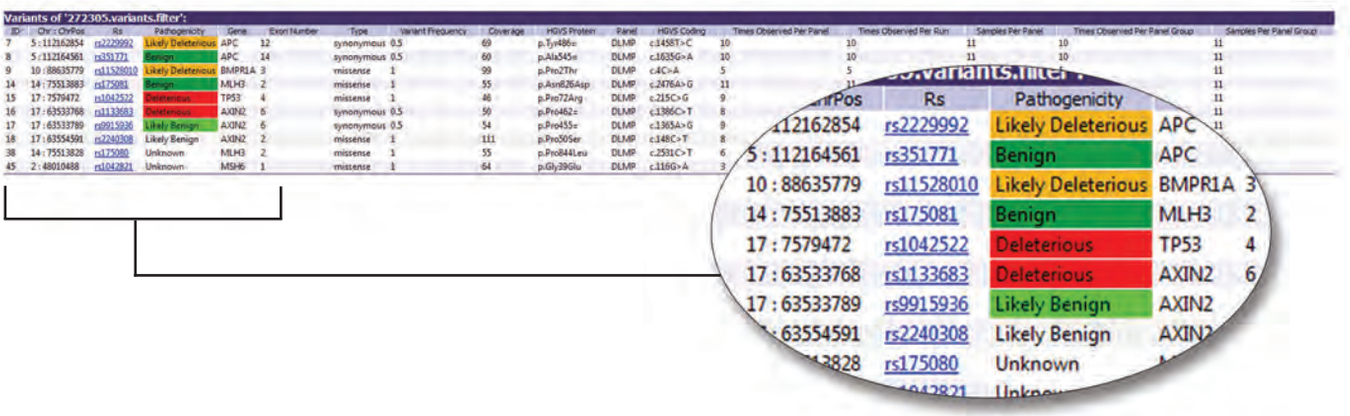
__Alamut__ licensees can quickly retrieve information without error prone and tedious retyping by simply selecting variant of interest and clicking on the drop down menu.



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.

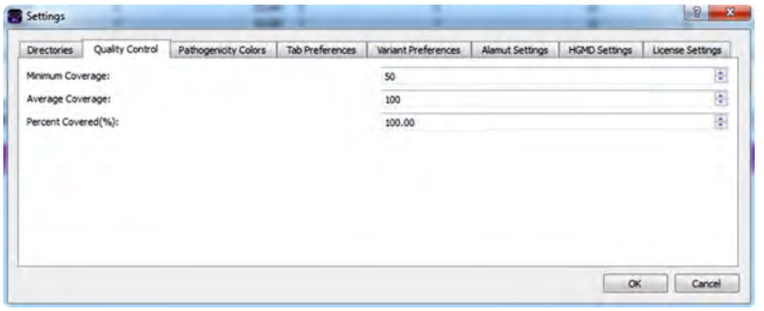


Historical information on every found variant is recorded and available for instant recall. Additionally prior pathogenicity determination is logged by specific disease panel and globally for all disease panels. The variant review tab provides previously determined variant type, pathogenicity, variant 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

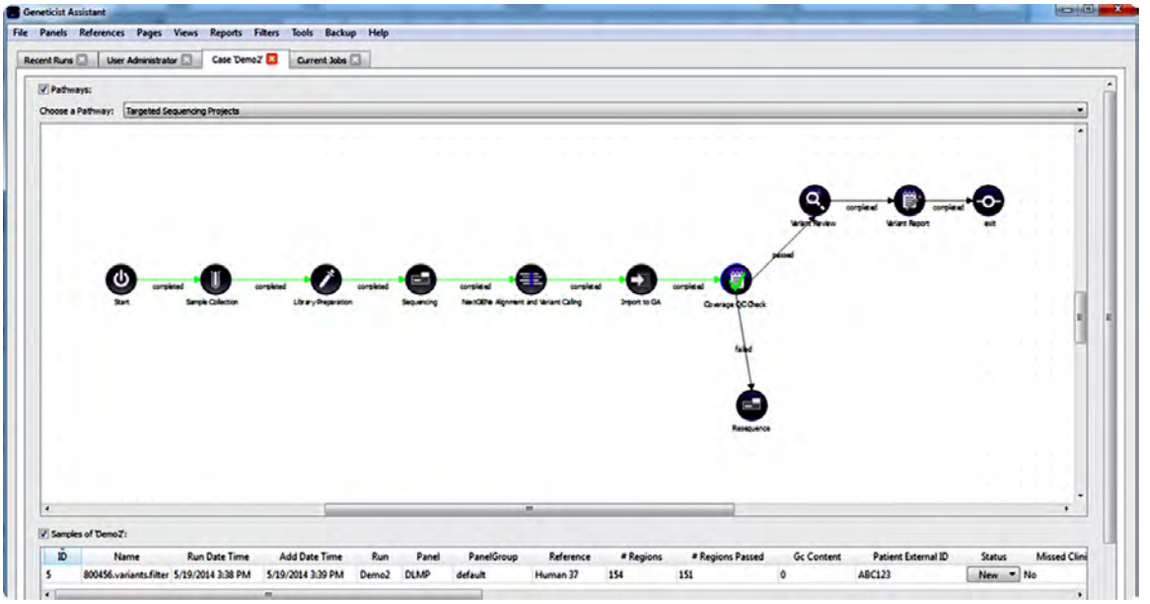
Geneticist Assistant NGS Interpretative Workbench is comprised of a local installed database, either Linux or Windows®, and a client Windows program which provides the easy-to-use, graphical user interface. All data is stored locally, accessible only to authorized users. Off-site collaborators or sister 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

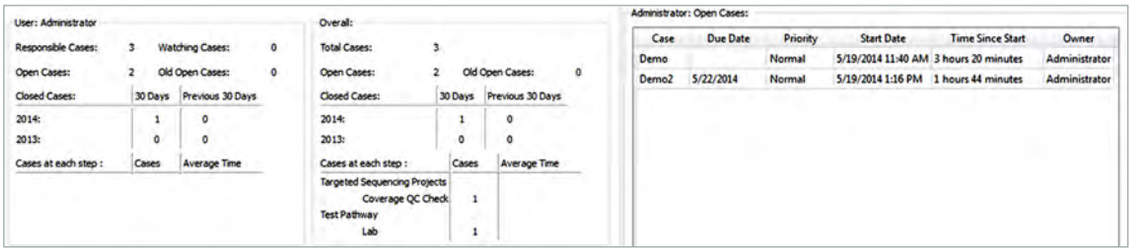
Customizable Workflow Builder

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**.

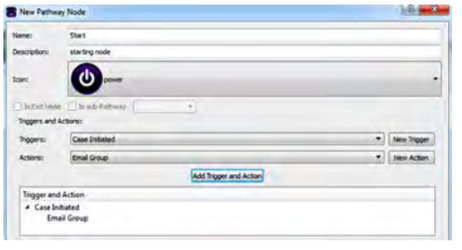


Geneticist Assistant NGS Interpretative Workbench features a completely customizable workflow builder that enables users to model physical workflow in "silico". The software monitors and provides real time, department by department reporting of current sample load & status, captures historical production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

Real-time Tracking and Reporting



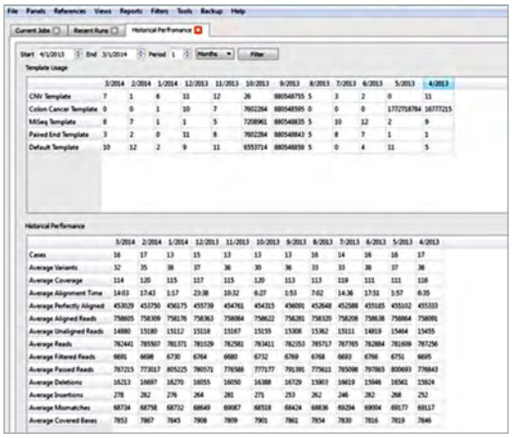
Geneticist Assistant provides real time statistics by department and for the complete system. Data is available to all authorized personnel, eliminating the need to manually research the status of each sample. Varying priorities can be set on a per sample basis, historical information by department or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

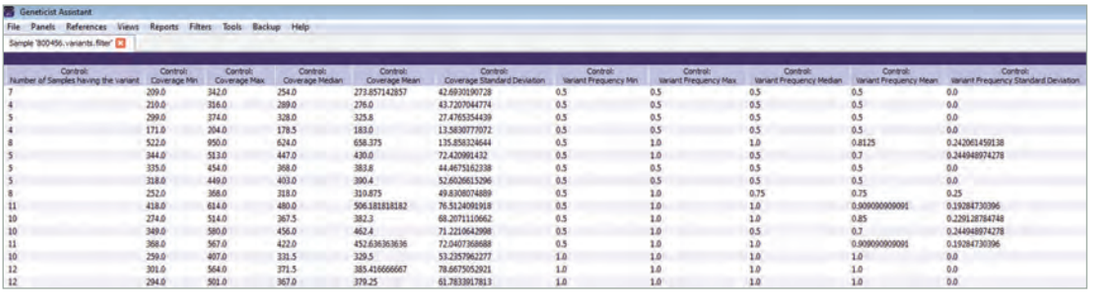


In a single report, Geneticist Assistant NGS Interpretative Workbench gathers quality statistics over time, often required during a CAP validation. Frequent review of the quality statistics often reveals negative trends permitting immediate corrective actions to be taken.

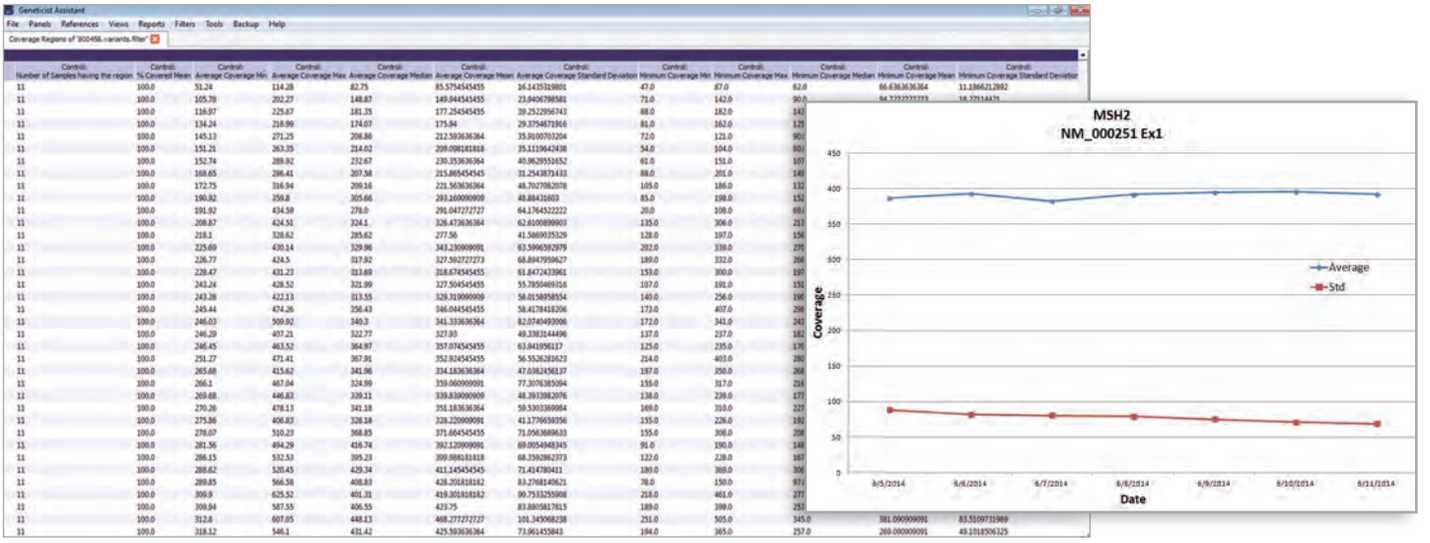
Process Quality Control

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 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

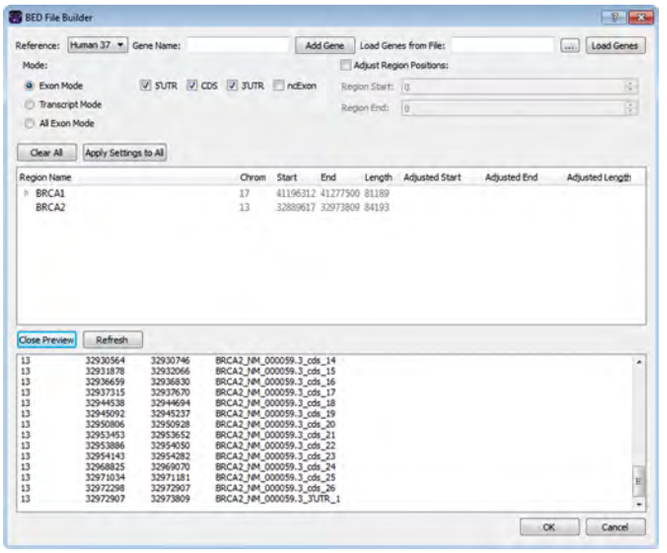
Many users opt to incorporate a positive control, such as NIST Genome in a bottle, with each sequencing run. Geneticist Assistant captures the positive control data, permitting a quick review of the run's efficacy and captures time-based data so that negative trends can be quickly observed and remedied.

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	**PLIllPCRFree**	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	**PLXIll**	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

For Research Use Only