March 2017

# Enlis Genome Research - Release Notes

Software version:  1.11

Annotation version 10

## Version 1.11 Changes:

- Major Annotation updates:  Enlis Annotation version updated to version 10

- • Updated clinical annotation
- • Updated gene categories
- • Updated to dbSNP 146

- Main Application:  Added Regulatory element impact to small variations, based on Encode data.  View on variation detail page.



- Variation Filter tool:  Added Regulatory Element filter



- Phenotype Explorer: Fixed bug where a variant insertion would be set as NoCall if the reference base next to the insertion was NoCall.  This insertion would be counted correctly in

the summary section, but not be listed in the table.  This also applies to the generated report PDF - shown in the Variation Summary section, but not in the Variation Details section.

- Phenotype Explorer:  Fixed bug where the tool would stall if there were only insertion variants listed in the table.

- Phenotype summary:  Added stop button to Phenotype Summary for faster switching to another genome

## Version 1.9 Changes:

- Main Application:  Added **Phenotype Summary Tool**.  A quick way to view and sort all of the known variant-phenotype associations for each genome.

- Sort the phenotype variants by phenotype name, confidence level, clinical significance type, allele frequency, chromosome position, and gene symbol.
- Counts of all the clinical significance types - e.g. Pathogenic, Risk factor, Drug Response, or Protective
- Pie graph shows the sequencing coverage of known phenotype variations - i.e. Out of 42,043 variations known to affect a phenotype, this genome has successfully sequenced 98.1%.
- Each phenotype variant lists annotation data like allele frequency, gene impact, dbSNP ID, phenotype descriptions and modes of inheritance.
- A button to generate a phenotype PDF within the Phenotype Explorer tool

- Phenotype Explorer tool: Added the ability to automatically generate a **Phenotype Variant Report PDF** based on the selected phenotypes.  The PDF includes a summary of the analyzed variants, description of the phenotype, detail on the variants that were found within the selected genome, and a complete list of the variants that were analyzed.

- Phenotype Explorer tool:  Improved search string matching by ignoring common words

- **Major Annotation updates**:  Enlis Annotation version updated to version 8

- **New Gene models**.    Previous gene models based on NCBI mappings have been completely replaced by a better set of gene models based on Gencode.  The new models have significantly fewer mapping errors (which can lead to incorrect gene impact annotation).  The new models are also more comprehensive for non-coding genes.
- Added 22,827 new genes, including 993 more protein coding genes and 21,834 non-coding genes.
- Added 10,145 additional variant to phenotype classifications for Clinical Significance
- Added 615 new gene categories and updated all existing


- Genome Import:  **Added support for importing CNVs in VCF files**

- Genome Import: Fixed bug in handling of male genome zygosity on the X chromosome



- Chromosome Map:  Non-protein coding genes are now colored yellow.  Protein-coding genes remain green.

- Chromosome Map:  Added direction indicators within the gene arrows for easier clarity of gene strand when zoomed in.



- Gene page:  Added gene type to header.  Previously the gene page had only indicated protein-coding genes.  Now in addition to protein-coding genes the gene page will list the type of non-coding gene as well - i.e.

> rRNA -- lincRNA -- misc_RNA -- antisense -- pseudogene -- processed_transcript -- miRNA -- sense_overlapping -- sense_intronic -- Incomplete protein coding -- 3prime_overlapping_ncrna -- IG_V_pseudogene -- snoRNA -- polymorphic_pseudogene -- IG_C_gene -- IG_D_gene -- IG_C_pseudogene -- IG_J_gene -- IG_J_pseudogene -- IG_V_gene -- Mt_rRNA -- Mt_tRNA -- snRNA -- TR_C_gene -- TR_J_gene -- TR_J_pseudogene -- TR_V_gene -- TR_V_pseudogene -- TR_D_gene --

- Gene page:  Added a star indication of the principal transcript(s) to the tables of protein-coding variation

- Gene page:  Added "Jump to section" dropdown menu at the top of the page

- Gene page :  Added genomic location text and alternate locations if present next to chromosome map

- Variation Filter tool:  Fixed bug in X-linked recessive trio filter that excluded some results

- Main Application:  Added default font for clarity and consistency between platforms

- Main Application:  New framework adds additional speed improvements for the entire application

- Main Application:  Adjusted 4 star confidence measure to include a single 1/2 star - indicating that the variation phenotype connection was classified by single submitter without evidence, or that there are multiple conflicting interpretations.

- Apple OSX version - Both the Genome Research and Genome Import applications are now signed to comply with OSX security requirements.  There may be a delay the first time the applications are opened as the signatures are verified.

## Version 1.8 Changes:

- Main Application:  Added **Phenotype Explorer** tool.  Search for keywords that match over 6000 phenotypes (diseases and traits).  Then, immediately assess the currently loaded genomes at the specific positions and variations that are associated with that phenotype.  For example: Looking at the included sample genome NA12877 for "Cystic Fibrosis" - Out of 211 known pathogenic variations, this individual matches 1 pathogenic variation, 2 of the positions are not called in the sequence data, and the individual matches reference at 208 positions.

- **Major Annotation updates**:  Enlis Annotation version updated to version 7

- Allele Frequency data is now based on a diverse population of over 60,000 exomes and around 3000 whole genomes
- Added Allele Frequency data for the mitochondria from ~30000 Genbank sequences
- Added 20,000 additional variant to phenotype classifications for Clinical Significance
- Added 951 new gene categories and updated all existing
- Updated to dbSNP 142

- Main Application:  Added **genome wide predicted deleterious scores** from the **DANN** algorithm.  This algorithm uses a "deep neural network" and a wide variety of biological training data to score every possible single nucleotide variation.  Variations that are predicted deleterious are annotated at 3 different score levels.

- Level 1 - Deleterious score is in the top 2.15% of all possible variations
- Level 2 - Deleterious score is in the top 0.62% of all possible variations
- Level 3 - Deleterious score is in the top 0.31% of all possible variations

( see http://bioinformatics.oxfordjournals.org/content/31/5/761 )

- Main Application:  Added a **4 star confidence measure** for variations with a **clinical significance**:
- One star:  Variation is classified by single submitter
- Two stars: Variation is classified by multiple submitters
- Three stars: Variation is reviewed by expert panel
- Four stars: Variation is reviewed by professional society

- **Genome Import - VCF filtering** during import.  Three options:

- Skip variants with a read depth less than ___.
- Skip variants with a quality score less than ___.
- Only keep variants with a filter field equal to:  [comma delimited list of valid]

These options are not used by default, but can be set in the options window.

- Variation Filter tool:  Added filter for DANN predicted deleterious variations, with the option to select different score levels

- Variation Filter tool:  **Added "Ignore No-call positions"** to "Not in single genome" and "Not in genome set" filters.  This allows for improved comparisons between genomes where one only wants to view positions that were called in both genomes

- Variation Filter tool:  Fixed bug in family trio analysis where valid positions were incorrectly removed from the results when using "Ignore positions where any genome is not called"

- Variation Filter tool:  Significant speed improvements to family trio analysis when using "Ignore positions where any genome is not called"


- Variation Position page:  Added Genotype summary and graph

- Variation Position page: Added an Export button to export the table of variations

- Variation Position page:  Significant speed improvements to finding no call regions


- Gene page:  Added tissue expression data for 9 additional tissues, 53 total

- Gene page:  Added tissue expression data for 112 additional genes

- Gene page:  Clicking on the summary headers will now scroll to see the details further on down the page.  For instance, click on "Protein variations:" header in the sequence summary to jump to the details of the protein-coding variation.

- Main Application:  Added Genome Report summary to view and export all the genome report information for all loaded genomes.  Access from the Tools menu.


## Version 1.7 Changes:

- Main Application:  Added **ClinVar annotation** to identify known pathogenic or benign variations .  This release includes over 120,000 variant to phenotype classifications.  Variants are indicated as:  pathogenic, likely pathogenic, risk factor, association, benign, likely benign, drug response, histocompatibility, protective, or uncertain significance.  ClinVar annotations are integrated into filtering and position page display.

- Main Application:  Added **Homozygous Regions Detector** tool.  Find regions of the genome with "runs" of consecutive homozygous variants.  For rare disease analysis, these regions may indicate a consanguineous union, and provide a starting point for finding recessive disease.  In tumor samples, these regions may indicate loss of heterozygosity.  Multiple options are available to fine tune the analysis, and presets for whole genome data or exome data are provided.  Export your results as a table or PNG image.

- Main Application:  Added **Publication citations to individual variations**.  Publications that support a variant to phenotype classification are listed on the position pages.  Link to Pubmed, or if the associated PDF is freely available, link directly to the PDF.

- Main Application:  Added **PDF display tabs**.  Publication PDFs can be view directly within the application.

- Main Application:  Added **Web tabs**.  Links to internet resources will now open in a web tab within the application.  This makes it easy to quickly check an internet resource, like dbSNP, without switching to a new application.

- Main Application:  Added **List Comparison** tool.  A simple tool to compare two lists of items and find items that are in common, or exclusive to one list or the other.  Especially useful for comparing lists of genes or dbSNP identifiers.

- Variation Filter tool:  Added columns for Allele Frequency, Mammalian Conservation, and Clinical significance/ traits.  Removed dbSNP column.

- Variation Filter tool:  Table export now includes additional columns.

- Variation Filter tool:  Added new filter - <u>Clinical significance</u> (pathogenic, likely pathogenic, risk factor, association, benign, likely benign, drug response, protective, or uncertain significance)

- Variation Filter tool:  Added new filter - Has <u>ClinVar ID</u>

- Variation Filter tool:  Added new filter - Variant has a <u>Citation</u>

- Variation Filter tool:  Added new filter - <u>Near Pathogenic</u> - find variations that within a specified number of basepairs to a known pathogenic, likely pathogenic, or risk factor variation.


- Compound Heterozygous tool:  Added quality score setting, to allow one to ignore variations that have a low quality score.

- Compound Heterozygous tool:  Adjusted the Enlis scoring system - Genes will score higher if one or more of the variants is a known pathogenic, likely pathogenic, or risk factor variation.


- Variation Categories tool:  Updated Clinvar variation category for pathogenic variation

- Gene Categories tool:  Updated default custom gene categories for Omim Genes, MicroRNA, CGD disease genes, and Cancer Genes


- Main application:  Changed tab closing rules.  When closing a tab, the new active tab will be the parent opener tab, if it is still available.


- **Genome Import:  Significant speed improvements**.  Import of VCF, Complete Genomics data, and other variation files is 30% - 600% faster depending on import size.  (BAM file import speed is unchanged)

- Genome Import:  Added support for importing structural variation from VCF

- Genome Import:  Added support for Complete Genomics VCF

- Genome Import:  Importer tool upgraded to 64 bit (32bit still available)


- Enlis **Annotation version updated** to version 6 --

- Annotation version 6:  Added ClinVar, Clinical significance, Phenotype (disease or trait), Publication citations

- Annotation version 6:  Gene annotation updated

- Annotation version 6:  Added 475 new gene categories

- Annotation version 6:  Updated tissue expression data, added 8 new tissues.

## Version 1.6 Changes:

- Main Application:  Added **BAM file connection** feature to genome file.  A BAM file accessible in the filesystem can be connected to a .genome file.  BAM connections will automatically be created in the Import Tool when importing with a BAM file.  Edit or add BAM files connections in the Genome Report Page.

- Main Application:  **Includes IGV** (Integrative Genomics Viewer)  for visualization of BAM data

- Variation Position page:  Added link to view the position in a connected BAM file

- Structural Variation page:   Added link to view the position in a connected BAM file

- Copy Number Variation page:   Added link to view the position in a connected BAM file

- Main Application:  Added **Genomic position locator tool**.  Open it from the tools menu. Use a gene symbol, or accession number and a nucleotide or amino acid number to find a genomic position.

- Main Application:  Now a **64 bit application**, which allows analysis of larger datasets

- Main Application:  Improved font scaling (em units)

- Variation Filter tool:  Added the ability to **save and load filter sets**.  Frequently used filter combinations and be saved and recalled with the menu in the upper right corner.

- Variation Filter tool:  Added new filter - Tissue Expression

- Variation Filter tool:  Added new filter - Omim Genes

- Variation Filter tool:  Added new filter - Clinvar Pathogenic Positions


- Gene page:  **Added tissue expression data** on 44 different tissues.  Added a graph of the expression data.


- Gene Categories tool:  Added a tissue expression section to view genes expressed in 44 different tissues.  Each tissue type has a separate category for three types of expression data:  High Expression, Medium/High Expression, and Non-ubiquitously Expressed.  In the High and Medium/High Expression categories, genes are sorted highest expression first.  In the Non-ubiquitously Expressed categories, genes are sorted by how unique they are to that tissue, most unique first.

- Gene Categories tool:  Custom gene categories will not be sorted alphabetically by default.  There is a new checkbox to sort alphabetically if desired.  This allow encoding information in the order of the genes listed, for instance, listing genes expressed in a tissue by highest expression first.


- Genome Import:  Improvements to Importer tool to prevent crashing when importing many genomes

- Genome Import:  Improved COSMIC variant detection

- Genome Import:  Fixed a bug in the import of VCF positions with multiple alternate alleles


- Variation Filter tool:  Fixed bug in Mac version where variation filter would not search more than one chromosome. (affected some OSX versions)

- Main Application:  More space to view matching gene categories in the autocomplete gene category lists within the gene category and variation filter tools.

- Variation Categories tool:  Updated Clinvar variation category for pathogenic variation

- Gene Categories tool:  Updated default custom gene categories for Omim Genes, MicroRNA, CGD disease genes, and Cancer Genes

- Enlis **Annotation version updated** to version 5 --

- Annotation version 5:  +167 genes

- Annotation version 5:  Updated to dbSNP version 138

- Annotation version 5:  Updated Gene categories - now contains > 20000 categories

# Version 1.5 Changes:

- Main Application:  Added **Compound Heterozygous gene search tool**.  Select a genome to search and the associated parent genomes.  The tool will find genes with protein-coding variations from each parent and rank the genes based on severity of impact.  Options are available to decide which type of variations to analyze and category of genes to search.

- Main Application:  Improved startup time

- Main Application:  Upgraded software framework results speed improvements in some tools.

- Main Application:  Fixed problem in the loading and analysis of > 1500 genomes.  Variation filter using genome sets is now tested to > 10000 genome files.

- Genome Import - Added optional **Bam file import** for calculating call coverage.  Most small variation files, like a VCF, do not include call coverage data.  By selecting a related .bam file, the import tool will identify the regions of the genome that were successfully sequenced and those regions that were not.

- Genome Import - Fix handling of a rare case where an insertion or deletion into a stop codon still creates a stop codon - Now characterized as SYNONYMOUS instead of FRAMESHIFT

- Genome Import - Fix handling of a rare case where an insertion occurs at the first base of an exon

- Genome Import - Improved handling of VCF files

- Genome Import - Fixed bug where the last line of a VCF file was not imported correctly

- Genome Import - Improved accuracy of gene based annotation


- Enlis **Annotation version updated** to version 4 --

- Annotation version 4:  Greatly expanded the number of genes:

> Total number of genes:  +11,144
> Number of protein-coding genes:  +107
> Genes with at least 1 transcript: +348
> MicroRNA genes:  +199
> Pseudogenes:  +7,314
> RNA, transfer: +22
> RNA, ribosomal: +19
> RNA, misc:  +671
> Immunoglobulin gene/pseudogene: +354


- Annotation version 4:  Added Mitochondria genes

- Annotation version 4:  Updated Gene categories - now contains > 19000 categories


- Variation Position page:  **Added HGVS notation**.  Variations are annotated with HGVS notation for chromosome position, cDNA impact, and protein impact.  Click on an annotation to copy it to the clipboard.


- Variation Filter tool:  Added a "clicked row" indicator so that after opening a new tab with a link, one can see that last clicked row upon returning

- Variation Filter tool:  Added ad hoc "Currently loaded genomes" to the genome sets list for the "IN Genome Set" and "NOT IN Genome Set" filters

- Variation Filter tool:  Added ability to specify a custom percentage for the genome sets for the "IN Genome Set" and "NOT IN Genome Set" filters

- Variation Filter tool:  Only use the read depth and quality score filters for the IN genomes, so that unexpected results are not found when comparing genomes

- Gene page:  Added ability to export the protein-coding variation table

- Gene page:  New external link to Uniprot

- Gene page:  Fixed bug in display of protein variation table when a gene has both coding and non-coding transcripts


- Gene Categories tool:  Custom category add - works with space or newline delimiters in addition to commas

- Gene Categories tool:   Custom category add - will automatically convert alternate gene symbols to known official HGNC gene symbols, or warn if a gene symbol was not found

- Gene Categories tool:  Added "Disease Genes - CGD" default custom category.  A category based on genes found in the Clinical Genomic Database (http://research.nhgri.nih.gov/CGD/) Useful for searching in the Variation Filter tool, or the new Compound Het tool.

- Gene Categories tool:  Added "MicroRNA Genes" default custom category.

- Gene Categories tool:  Updated "Cancer Genes" default custom category.


- Variation Categories tool:  Updated built-in ClinVar variation category based on release 2013-05-06 and edited list so that  it now it only contains positions where variants are identified as "probable-pathogenic" "pathogenic" or "drug-response" (~13000 variants)


- Main window search - Will offer symbol suggestions if an alternate gene symbol is not found (i.e. "HER2" >> "ERBB2")

- Main window search - Added ability to enter a chromosome position or range "chr4:50000" Ranges 10nt or less will take you to the position page.  Longer than 10nt will take you to the chromosome maps


- Organize Genomes window - Genome names are colored red if they are not found in the filesystem

## Version 1.4 Changes:

- Enlis Annotation version updated to version 3 --

- Annotation version 3: Reference GRCh37 - New Annotation - Global allele population frequency for small variants.  Frequency is based on data from the 1000 genomes project and the NHLBI Exome Sequencing Project (ESP) Exome Variant Server

  http://www.1000genomes.org/

  http://evs.gs.washington.edu/EVS/

- Annotation version 3: Reference GRCh37 - New Annotation - Mammalian conservation.  Small variants are annotated if the position exhibits high mammalian conservation based on GERP score (Genomic Evolutionary Rate Profiling).  For these purposes, positions with a GERP score >= 3.0 are considered highly conserved.

  http://mendel.stanford.edu/SidowLab/downloads/gerp/index.html

- Annotation version 3: Reference GRCh37 - New Annotation - Catalogue of somatic mutations in cancer (Cosmic).  Small variants are annotated with a Cosmic identifier if that position exists within the Cosmic database.  Cosmic database version 62.

  http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/

- Annotation version 3: Reference GRCh37 - New Annotation - Protein prediction of deleterious animo acid changes.  SNPs are annotated as predicted deleterious if the position and nucleotide change exhibits a SIFT score <= 0.05.

  http://sift.jcvi.org/

- Annotation version 3: Reference GRCh37 - New Annotation -  Variant read depth.  Read depth data on small variants will be imported, if it is available.

 - Annotation version 3: Reference GRCh37 - New Annotation -  Variant quality score.  Quality score data on small variants will be imported, if it is available.

- Annotation version 3: Reference GRCh37 - New Annotation -  Haplink - Not used for now.

- Annotation version 3: Reference GRCh37 - Updated to dbSNP version 137.


- Main Application:  Upgraded software framework results in significant speed improvements in nearly every tool.

- Main Application:  Save custom notes on any variation position, gene, structural variant or copy number variant.  New tool - View All Notes.  See all the notes that have been saved, organized by the most recent.

- Main Application:  Fixed drag and drop of genome file now works when there are multiple periods "." in the filename

- Main Application:  Added the "Tools" menu to the main menu bar, allowing quick access for opening a tool in a new tab.


- Chromosome Map:  Added display of reference sequence when zooming in closer than 200bp.

- Chromosome Map:  Added display of small variation zygosity

- Chromosome Map:  Added more detailed information on small variations when zooming in closer than 200bp, including variant sequence and length indication on variations that span multiple nucleotides.

- Chromosome Map:  Fixed rare bug where an exon would be drawn too large.

- Chromosome Map:  Fixed rare bug where a no call block would be drawn too large or disappear at some zoom levels.

- Chromosome Map:  Fixed rare bug where a gene arrow would be drawn too large or disappear at some zoom levels.

- Chromosome Map:  Fixed rare bug where a non-coding gene with multiple transcripts would show some exons as dark-blue (coding)


- Variation Position page:  Added display columns for Read depth, Quality score, Allele frequency, Mammalian conservation, Protein deleterious prediction, and Cosmic identifiers.

- Variation Position page:  Added external links on variation page.  Directly view the position with the UCSC genome browser.   Search pubmed or google for a dbSNP identifier.

- Genome Import - Improved calculation of full genome call percent for all import types

- Genome Import -  Positions not listed in the source data (Often intervening reference sequence) will now be now considered "No Call" when importing small variants from import types: VCF, GVF, and DNANexus.

- Genome Import -  Added support for importing Illumina's "genome VCF" (.genome.vcf or gVCF) format.  This format incorporates data on positions in the genome that were successfully sequenced but match the reference genome.

- Genome Import -  During importation of a VCF file, variants are no longer excluded based on the Filter column.  Previously, variants would not be imported unless the Filter column contained "PASS" or ".".  Low quality variants can now be filtered within the main application - or alternatively, low quality variants can be removed before importation.


- Variation Filter tool:  New filter added - Global Allele Frequency

- Variation Filter tool:  New filter added - Mammalian conservation

- Variation Filter tool:  New filter added - COSMIC cancer database

- Variation Filter tool:  New filter added - Protein predicted deleterious (SIFT score <= 0.05)

- Variation Filter tool:  New filter added - Quality score

- Variation Filter tool:  New filter added - Read depth


- Gene page:  Summary shows a column for the number of variations that are rare (<1% Global allele frequency), rather than dbSNP novel

- Gene page:  Added display columns for Allele frequency, Mammalian conservation, and Protein deleterious prediction


- Gene categories tool:  Shows a column for the number of variations that are rare (<1% Global allele frequency), rather than dbSNP novel.

- Variation categories tool:  Added an example category for the preview release of Clinvar. ClinVar aggregates information about sequence variation and its relationship to human health. This is a preliminary release from 2012-11-05.  http://preview.ncbi.nlm.nih.gov/clinvar/

## Version 1.3 Changes:

- Annotation version 2: Reference GRCh37 - Updated to dbSNP version 135.

- Annotation version 2: Reference GRCh37 - Updated gene models to the latest Refseq data. New genes are added and significantly improved the identification and positioning of closely related genes. Updated gene function descriptions, gene-disease annotations, protein-protein interaction lists.  Improved the identification and handling of genes in the pseudo-autosomal regions of chrX and chrY.

- Annotation version 2: No longer packaging NCBI36 annotation database in the default install package.  These annotation databases are available upon request.

- Annotation version 2: No longer annotating variations at an arbitrary 0-2000 bp 5' of genes as "TSS-UPSTEAM". Instead there is a new and significantly more flexible "Near gene Start/End" filter in the Variation Filter tool.

- Main Application: Added new Variation Categories tool

> Quickly assess all of the loaded genomes at specified variation positions

> Includes support for viewing small variation (SNP, SUB, INS, DEL), overlapping copy number variation (CNV), overlapping structural variation (SV), and no calls

> Create custom lists of variation positions - users are encouraged to create their own categories for positions that are relevant to their research

> Pre-defined example categories include: Cystic Fibrosis, Eye color, Warfarin Metabolism, and dbSNP Pathogenic

> Export to an Excel compatible tab-delimited file

Integrated with the variation filter tool

- Main Application: Major revamp of the Gene Categories tool

Now contains an extensive list of nested pre-defined gene categories that are derived from gene ontology. These pre-defined categories are divided into several main sections: "Biological Process", "Molecular Function", "Cellular Component", and "Disease/Trait"

New columns for Call Coverage - Protein-coding, exon, and full gene - genes with coverage less than 95% are highlighted

Export to an Excel compatible tab-delimited file

- Main Application: Genetic Disease associated genes are now shown in the Gene Categories tool, rather than in a separate tool page

- Main Application: Fixed bug on windows where two genomes from the same file location could be added to the program simultaneously if their path differed in case when added

- Main Application: Added warning to the home tab if loaded genomes contain different annotation versions

- Main Application: Fix bug that prevented drag-and-drop of .genome files upon first install

- Chromosome Map: Added support for bookmarking a specific map position

- Chromosome Map: Fixed bug that caused gene arrows to disappear when zooming in close to an extremely large gene

- Gene page: Added a list of all the gene categories to which the gene belongs

- Variation Position page: Will now show any overlapping copy number changes or structural variation, in addition to small variation and no calls.

- Variation Position page:  Will show the reference nucleotide, even if none of the loaded genomes contain a variation.


- Variation Filter tool:  New filter added - Near gene Start/End

- Variation Filter tool:  New filter added - BED file

- Variation Filter tool:  New filter added - Variation Category

- Variation Filter tool:  Fixed rare bug in the identification of a variation that overlaps the start of a gene when a gene category/list filter is selected


- Genome Import: Added support for updating an Enlis .genome file to the latest annotations.

- Genome Import:  Added support for importing from Affymetrix Human SNP Array 6.0 exported .txt format

- Genome Import: Added support for selecting multiple files for import or update.  This works for updating Enlis .genome files, or importing single file data - VCF, GVF, DNANexus, or Affy.

- Genome Import: No longer asking for a genome name when importing single file data - VCF, GVF, DNANexus, or Affy.  The .genome filename will be set to the same name as the source file. GenomeID and Display Name can be edited on the Genome Report page.  For Complete Genomics import, the ASM-ID will be automatically determined and used as the filename of the .genome.  CASAVA import will still ask for a genome name when starting import.

- Genome Import:  Fixed bug that prevented canceling an import if the cancel action occurred during the extraction of a gzipped file during CASAVA import

- Genome Import:  CASAVA import, added support for chromosome folder names of the format "X.fa" in addition to "chrX" and "chrX.fa"

- Genome Import: Improved Complete Genomics import to automatically extract the ASM-ID. In addition, the sequencing date and DNA sample ID are automatically extracted and placed in the Genome Details section of the Genome Report page

- Genome Import: Command-line Genome Import tool has been created and is available upon request

- Genome Import: Command-line version includes the ability to load annotation databases into memory to speed import.  Experimental

- Genome Import: Fixed a rare bug where the count of UTR variations for a gene was inaccurate

- Genome Import: Fixed a bug related to identifying a "SPAN" gene component of a variation that spans the start or end of a gene


## Version 1.2 Changes:


- Variation Filter:  Variation filter completely redesigned to allow searching multiple genomes simultaneously.

>    - Rather than showing all available filters, they are added to the current query with "Add filter" buttons

>    - Created new filters for "In Single Genome", "Not in Single Genome", "In Genome Set", "Not in Genome Set", "Zygosity"

>    - Added family trio analysis with predefined filters for finding variations which match "All Mendelian Variations", "All Non-Mendelian Variations", "Autosomal Recessive Model", "Autosomal De-Novo Dominant Model", or "X-linked Recessive Model"


- Variation position page:  Now lists zygosity on one line "(x1)" or "(x2)" instead of listing two separate lines for a zygosity of two

- Gene page:  Protein-coding variations now list zygosity in one line "(x1)" or "(x2)" instead of two  separate lines for a zygosity of two

- Gene page:  The chromosome position link within the protein-coding variations table now opens the variation position page for the listed variation, rather than the chromosome map at that location.  The chromosome map of that position can be reached from the variation position page by clicking the "Goto map" link

- Main search box:  Added support for looking up remote dbSNP rsID location for genome reference HomoSapiens_NCBI36

- Main application:  Added ability to delete genome sets from the menu using Control+click or Command+click (Mac) or middle-button click

- Main application:  Added ability to delete bookmarks from the menu using Control+click or Command+click (Mac) or middle-button click

- Main application:  Added ability to bookmark chromosome maps at a specific position.  Use "Bookmark this page" from the Bookmarks main menu.

- Main application:  Added Organize genomes and Clear Genomes list buttons to the home tab

- Main application:  Added Organize genomes window to reorder the list of loaded genomes and to create and edit genome sets

- Main application:  Genome sets are now stored in a separate genomeSets.json file

- Main application:  Added "Add Genome Set ->" and "Replace with Genome Set ->" items to the main Genome menu.  Allows quick appending of an existing genome set into the currently loaded list, or replacing the currently loaded genome list with an existing genome set.

- Main application:  Removed the dependence on the mRNAtoGeneSymbol database.  It is no longer needed.

- Main application:  Custom gene categories are now stored in the user's profile folder, and do not require operating system administrator privileges to create.

- Structural variation detail page:  Fixed high memory usage when viewing a structural variation page with many genomes loaded


- Genome Import:  Added support for Complete Genomics file format 2.0

- Genome Import:  Added support for GVF and DNANexus variation download

- Genome Import:  Added "Sequence Type" selector when importing CASAVA results.  "Whole Genome" or "Exome / Targeted capture" will allow the copy number variation procedure to be run only on whole genome data as required by the algorithm.

- Genome Import:  For Complete Genomics import, Structural and Copy Number variation are now annotated for gene impact with the Enlis gene model, rather than using the existing Complete Genomics gene impact data.  This allows for consistent data throughout the program.

- Genome Import:  For CASAVA import, if they exist, multiple sites.txt files in each chromosome folder are now recognized.  Multiple sites.txt files may be present when the assembly was done with sections of each chromosome, rather than the entire chromosome at once.

## Version 1.1 Changes:

- Added NCBI 36 / hg18 support

    Added annotation databases for genome import of NCBI36 / hg18 data

    Added support for displaying NCBI36 / hg18 data in the main Genome Research program

    Added warning if genomes based on different references are loaded at the same time.

- Added export of table results from the Variation Filter and Genome Difference tools

    New "Export result" button above the table of results

    Export of table results in custom TSV or custom CSV formats (Microsoft Excel compatible)

    Export of table results in GFF3 or BED formats

- Changes to Variation Filter tool

    Added the variation's zygosity (x1 or x2) to the Var Type field  in the results

    Homozygous variations were previously shown on two separate lines, but now these are consolidated into one line - displaying a zygosity of (x2)

    Added the Result display for counts of both "number of variations" and "number of positions".  For number of variations - a position with a variation of zygosity 2 is counted twice.

## Version 1.0:

Initial release