Applied Maths proudly presents **BioNumerics version 7.6**. Although considered a "minor" upgrade in our upgrade policy, it features important whole genome SNP functionality, completing the whole genome sequence analysis possibilities within the software. This enables a simple analysis workflow, starting from sequence reads, over reference mapping to cluster analysis based on the SNPs detected between the genomes in the comparison. The calculation engine, which became available from BioNumerics 7.5, was updated by adding reference mapping functionality to it. Please have a look in the list below for an elaborate overview of all innovations for this upgrade.

NEW FEATURES IN BIONUMERICS v7.6

GENERAL FUNCTIONALITY

- Example databases can be downloaded directly from the Startup screen.
- SQLite is now the default database engine option when creating a new database.
- Data exchange functionality has been expanded:
  - Export of sequence read sets to FASTQ or FASTA formatted files.
  - Exchange of batch sequence assembly settings with the sequence experiment types.
  - Export and import of information field states and colors.

SEQUENCES

IMPORT

- Import routines "Download sequences from internet" and "Import sequence read sets as links" can use NCBI/SRA identifiers stored in an entry information field of choice.
- Sequences from EMBL/GenBank files can optionally be concatenated per file and imported as multiple contigs in the same sequence experiment.

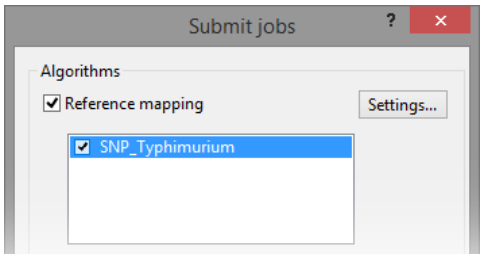
SEQUENCE VIEWER

- A new panel displays information about the contig(s) present in the sequence.

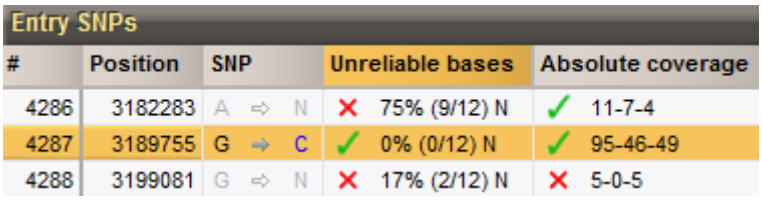
SEQUENCE READ SETS

- Power Assembly analysis templates request the k-mer value at runtime, allowing to adjust the k-mer value before launching the de novo Power Assembly project.
- Power Assembly analyses are adjusted to enable the import of multiple (paired-end) read files per entry, typically generated as Illumina NextSeq data.
- Custom analysis templates can be starting from the sequence read sets of the selected entries in the main window. All custom-defined Power Assembly projects as well as predefined mapping, resequencing and de novo assembly templates can be used.

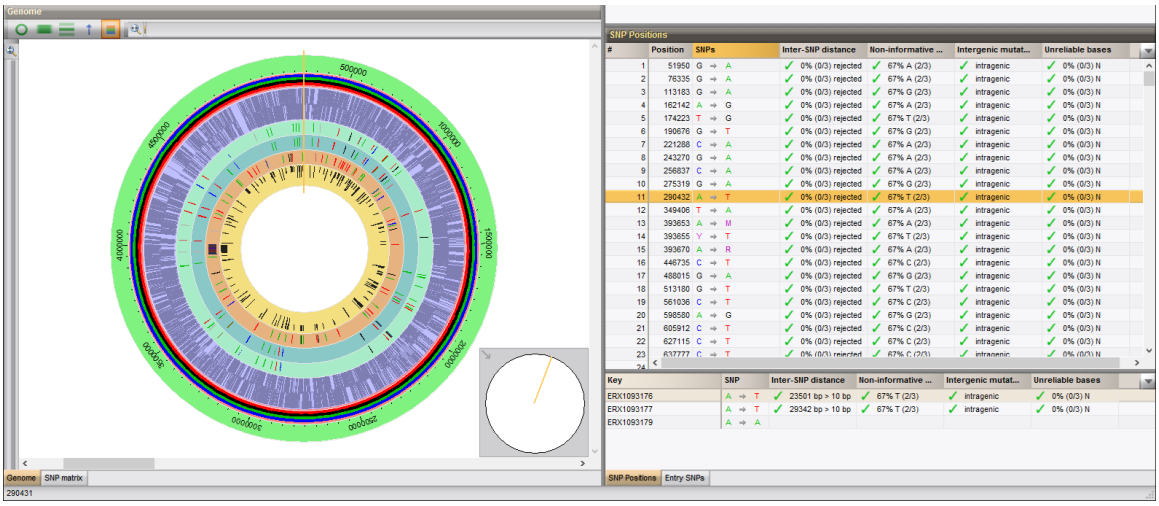
WHOLE GENOME SNP ANALYSIS



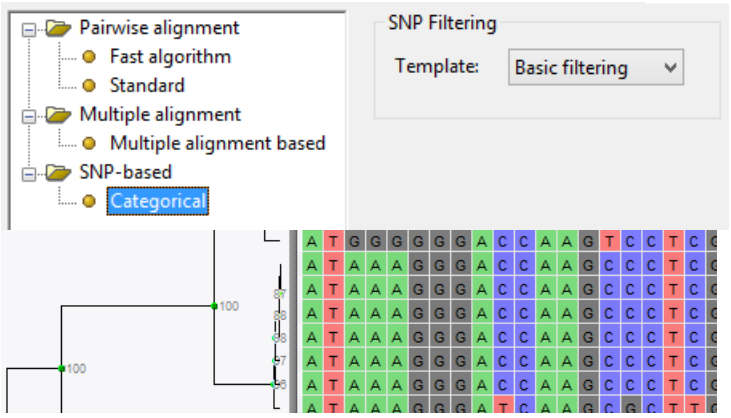
- Reference mappings against a reference of choice can be performed, either locally or on the calculation engine. Two different mappers are available on the calculation engine, i.e. the Bowtie2 algorithm and the BWA algorithm implemented in the Applied Maths mapper algorithm.
- From the entry selection in the main window and from within the comparison, wgSNPs can be identified based on reference mapped sequences. A plethora of SNP filters can be defined and saved as SNP templates, which can be saved and shared within the database and exchanged between coworkers. Filtering options include: filtering for ambiguity, reliability, abundance, base quality, inter-SNP distance, SNP type, coverage, etc.



- SNPs can be assessed in the SNP filtering window and evaluated graphically on the circular genome viewer or in the SNP matrix. Detailed information on the SNP positions for all entries and individually for the selected entry only can be viewed from within the same window.



- Routine wgSNP analysis can be performed directly from the database entry selecting by selecting one of the SNP templates which results in the SNP set being presented in the dedicated Comparison window.



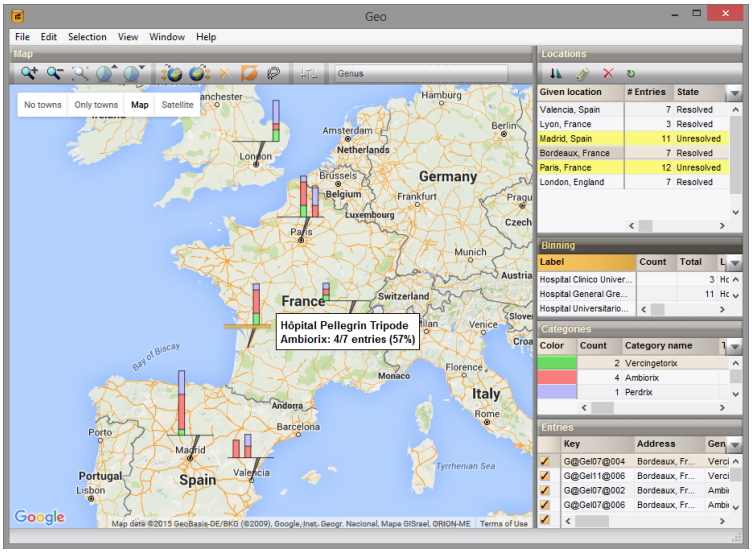
- Follow up analysis, including clustering and further exploration of the SNP set is possible from the available comparison tools within the Comparison window.

PLUGINS

WHOLE GENOME SEQUENCE TOOLS PLUGIN

- The whole genome MLST client plugin has been merged with the wgSNP functionality in the Whole Genome Sequence (WGS) tools plugin.
- Enables direct upload of sequence read sets to the Calculation Engine, i.e. no need to make the data public or to use third-party tools for uploading fastq.gz files to cloud storage. The actual upload is performed by a separate executable, meaning that you can continue working with BioNumerics while the files are being transferred.
- Additional quality measure for sequence read set quality, de novo assembly, assembly-free and assembly-based wgMLST became available in the quality character experiment and the Quality Assessment window.
- FIPS-compliant communication is available for the Nomenclature server and Calculation Engine.
- Two new experiments are created upon installation of the whole genome tools plugin: one character experiment for the storage of trimmed SRS statistics, if calculated, and one character experiment for the storage of the allele call type per locus (AF-only, AB-only, multiple calls, etc.).

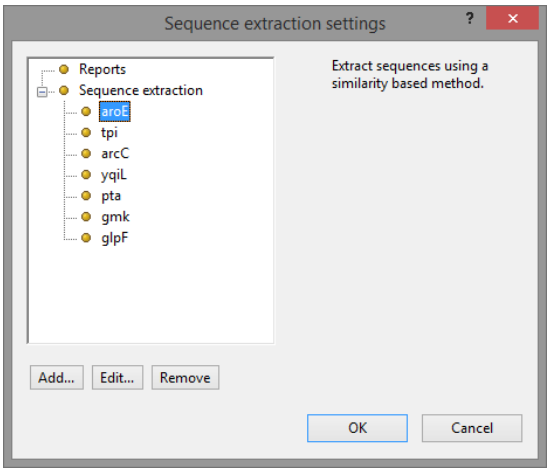
GEOGRAPHICAL MAPPING PLUGIN



- Graphs can now additionally be created based on character data, and starting from the Comparison window.
- The new stacked bar graph option allows additional binning based on a text, number or date field.
- Dendrograms can be plotted on the geographical maps.
- Pie charts and bar graphs scale proportionally to the number of entries per location.
- New selection tools became available, including a lasso tool and functionality to invert the selection or to clear all.
- Resolve locations now shows the locations on the map for better assessment.

SEQUENCE EXTRACTION PLUGIN

- This new plugin offers functionality to extract a particular subsequence of interest from any genome sequence, including unannotated draft genomes.
- The extraction protocol is based on a BLAST similarity search, optionally combined with start/stop codon detection or forward/reverse primer detection.
- Sequence extractions are ran in batch for the selected entries and for a set of user-defined genes or partial gene sequences (e.g. in case of MLST).

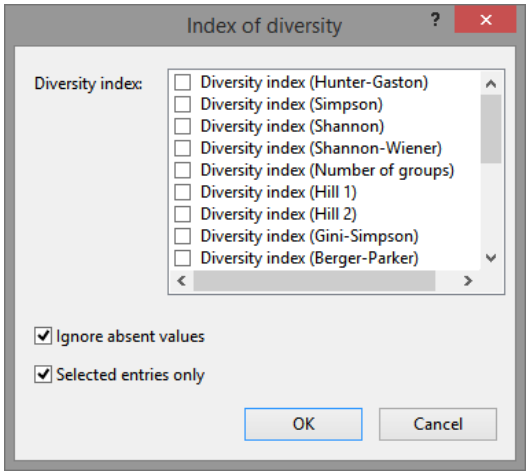


CLUSTERING AND COMPARISON

COMPARISON WINDOW

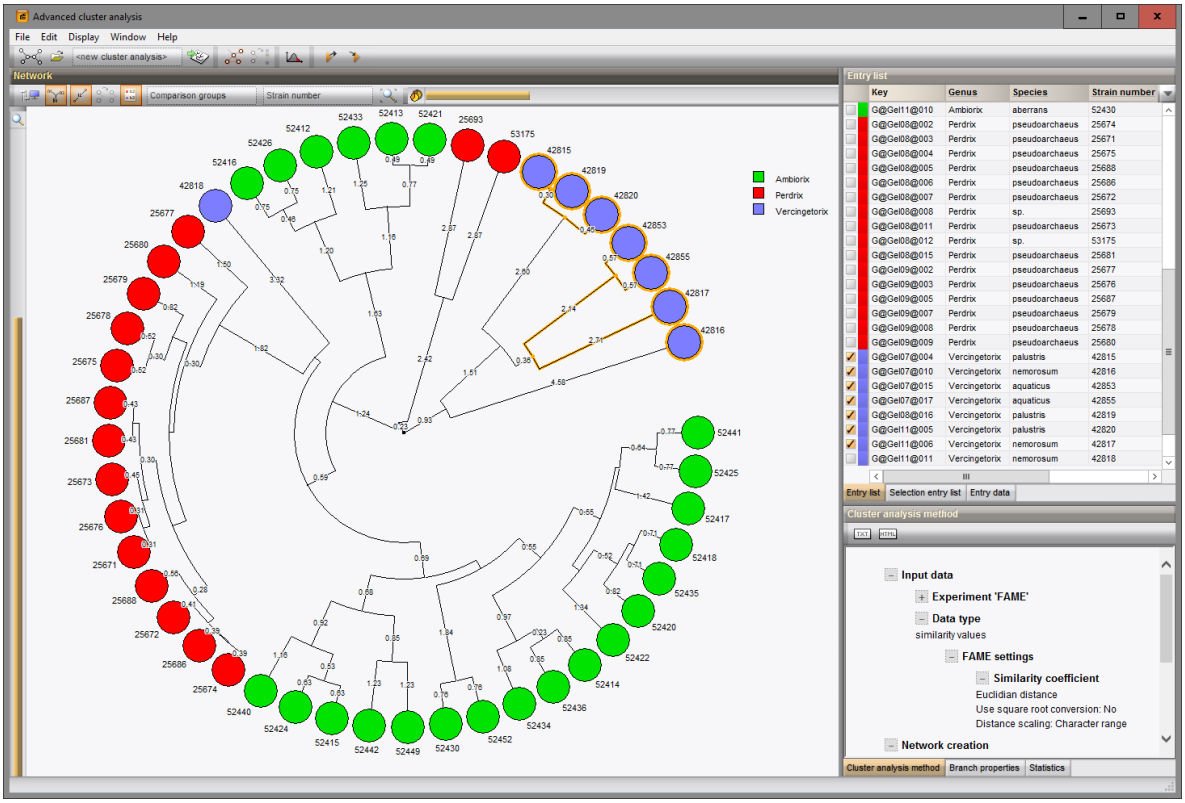
Character data:

- Various character selection and filtering tools became available from the Comparison window.
- A weighted categorical clustering can be performed based on the character weights defined in the character experiment type.
- Character diversity indices can be calculated from the Comparison.

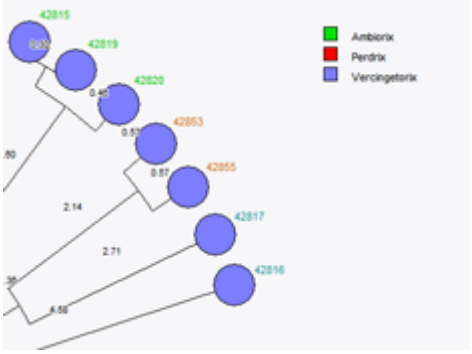


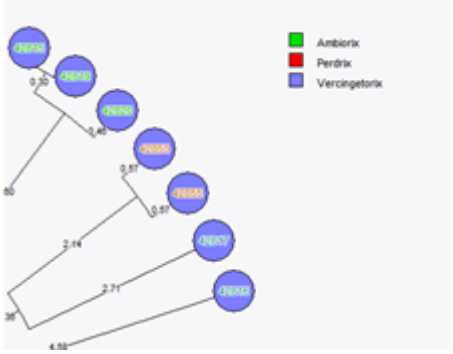
ADVANCED CLUSTERING WINDOW

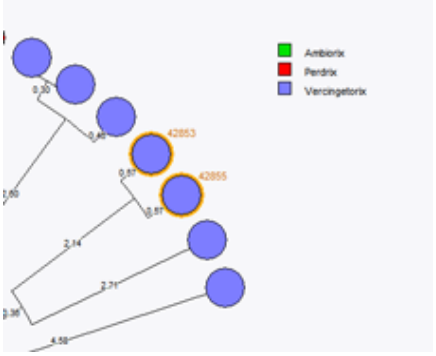
- A new dendrogram layout option is now available in the Advanced cluster analysis window that allows to create a circular dendrogram of your data.



- Square root option for scaling the branch lengths is available as new display setting.
- New label display options are available:
  - Show labels on top of the nodes.
  - Show labels for a selection of nodes.
- New coloring options are available :
  - Color node labels.
  - Color branches.







MATRIX MINING WINDOW

- A categorical similarity coefficient has been added to the Matrix Mining window, allowing analysis for e.g. wgMLST or wgSNP data

ABOUT THE BIONUMERICS PROGRAM

The BioNumerics platform is a modular software environment for the integrated analysis of all your biological data. BioNumerics can be configured exactly to match your research needs!

Contact us:

Europe and worldwide

Applied Maths NV
Keistraat 120
9830 Sint-Martens-Latem, Belgium
Phone: (+32) 9 2222 100
Fax: (+32) 9 2222 102
E-mail: info@applied-maths.com

North America

Applied Maths, Inc.
11940 Jollyville Rd., Suite 115N
Austin, TX 78759, U.S.A.
Phone: (+1) 512-482-9700
Fax: (+1) 512-482-9708
E-mail: info-US@applied-maths.com

