NGS at geneXplain GmbH

Introducing GenomeTraveler

geneXplain GmbH

Products

- Bioinformatics and systems biology: geneXplain platform
- Bioinformatics and NGS: IMC and GenomeTraveler
- Cheminformatics: PASS, PharmaExpert, GUSAR

Systems Biology and Cheminformatics

- geneXplain platform
  - Collection, storage and analysis of experimental data
  - Network clustering and search for master regulators
  - Graphical programming of workflows
  - Possibility to write new scripts and add-ons
- PASS and PharmaExpert
  - Qualitative biological activity prediction for libraries of chemical compounds
  - Drug-drug interactions and identification of most promising candidates
- GUSAR
  - Quantitative biological activity prediction for libraries of chemical compounds
  - (Q)SAR model building

geneXplain GmbH

Products

- Bioinformatics and systems biology: geneXplain platform
- **Bioinformatics and NGS:** IMC and **GenomeTraveler**
- Cheminformatics: PASS, PharmaExpert, GUSAR

Contents

Primary Analysis

- Getting started with NGS data

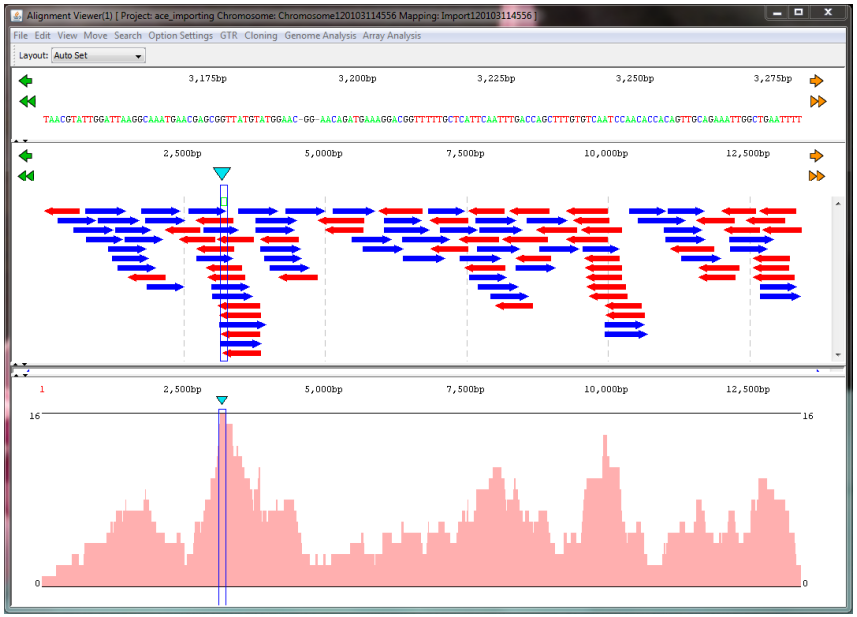
General Information

- GT supports NGS data from the following sources:
  - ABI SOLiD (CSFastA)
  - Illumina GA (FastQ)
  - Roche/454 GS FLX/Junior

- It is possible to map NGS sequences against sequence data in any of these formats:
  - ABI/SCF
  - GenBank
  - FastA

de novo Assembly

- *de novo* assembly can be done with 200 bp or larger reads, saved in the following formats:
  - ABI/SCF/SFF/PHD
  - FastA/FastQ/CSFastA
  - Solexa Raw
  - GenBank/EMBL
  - TXT
- GT also supports the import of assembly files from:
  - Velvet (AFG)
  - Phrap (ACE)

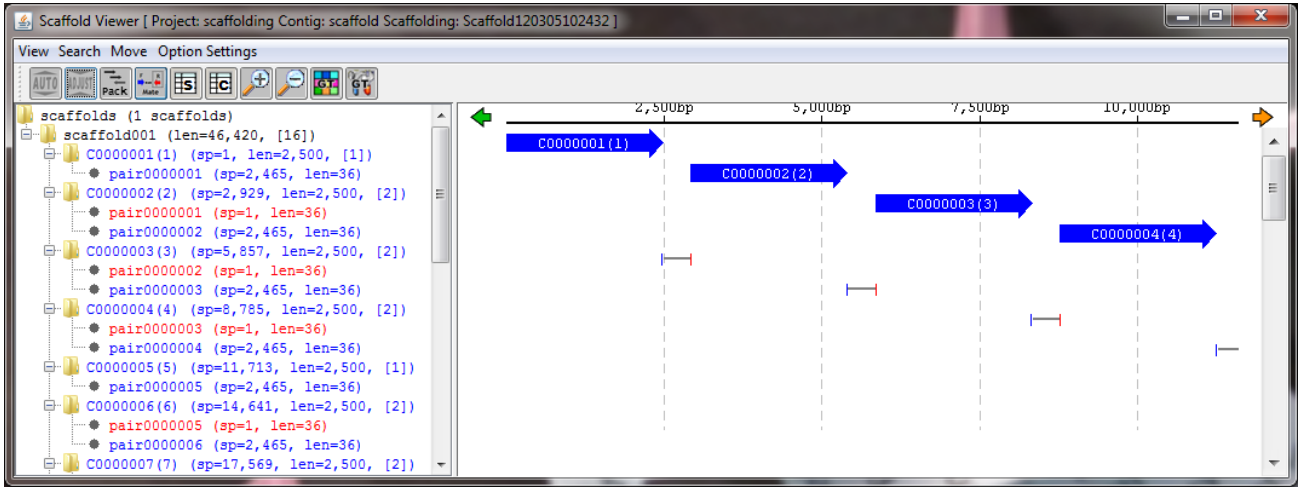


Example: assembly results imported from Phrap.

Scaffolding

The *de novo* assembly of large genomes results in a large number of contigs. These can be combined using paired-end short reads.

Contig positions in a scaffold can be viewed in GT's "Scaffold Viewer".



Mapping Functions

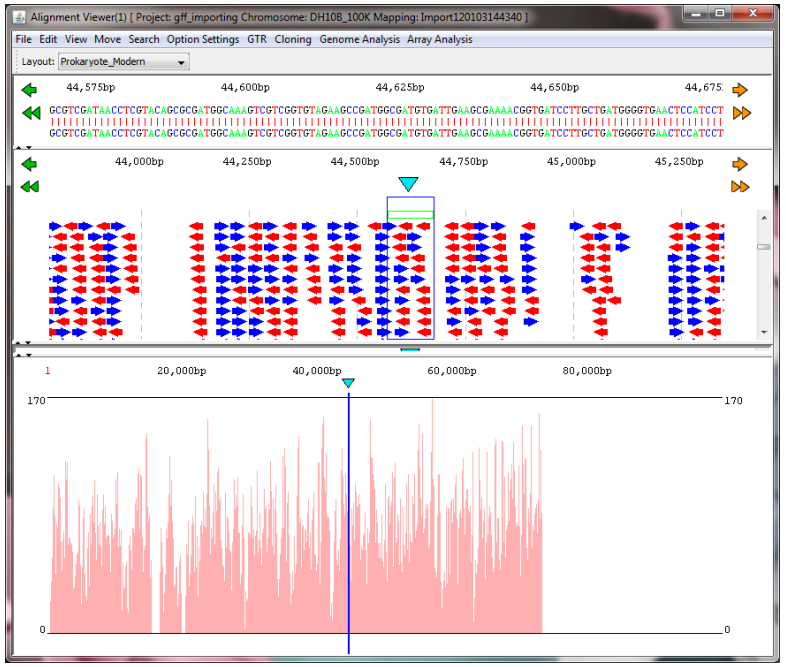
The mapping algorithm is based on NCBI Blast results for reference sequences compared with NGS short reads.

- fragment short reads
- paired-end short reads
- multiple sites
- multiple chromosomes or contigs

Import of Mapping Results

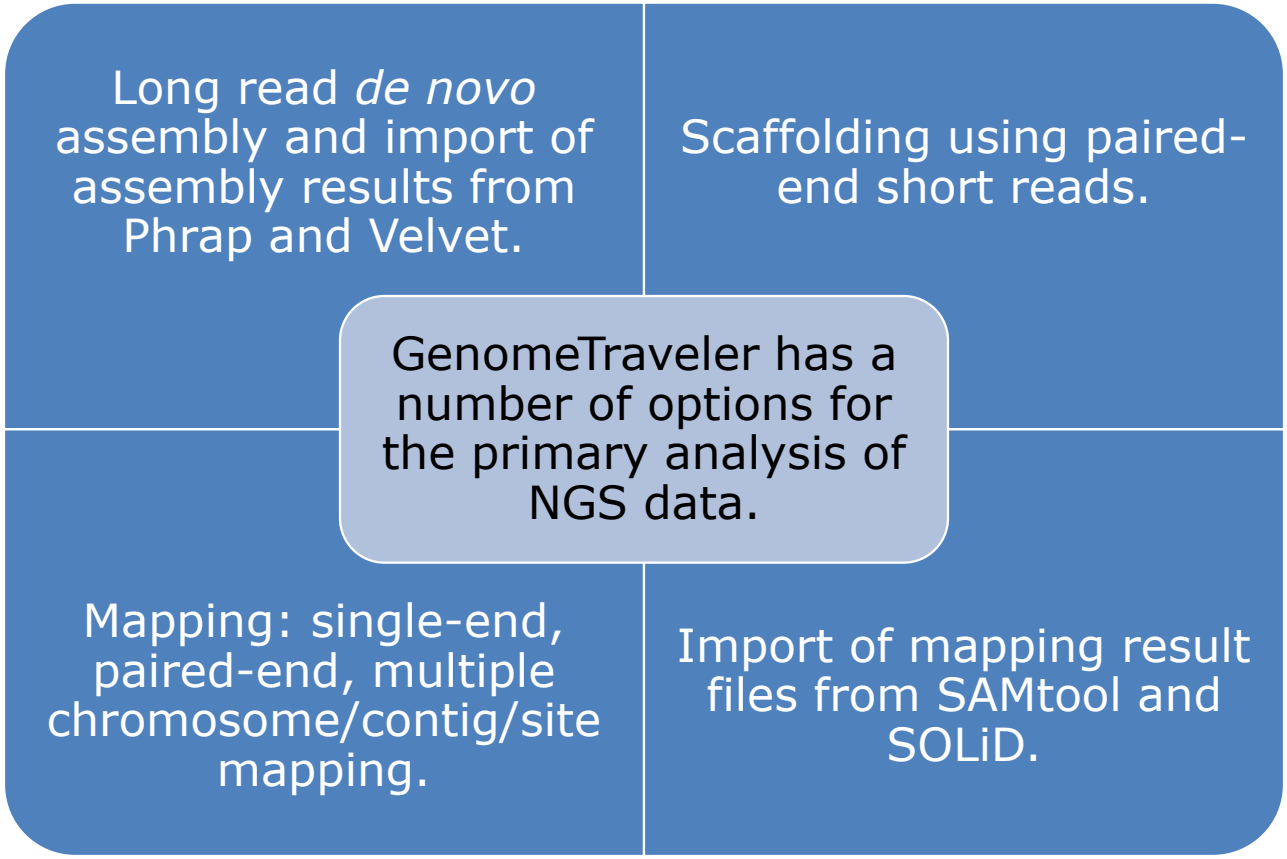
- GenomeTraveler can also import mapping results from other software, namely:
  - SAMtool (SAM/BAM)
  - SOLiD (GFF)

- Mapping and import results can be viewed in the "Profile and Alignment Viewer" of GT.



Example: imported GFF file.

Primary Analysis Summary



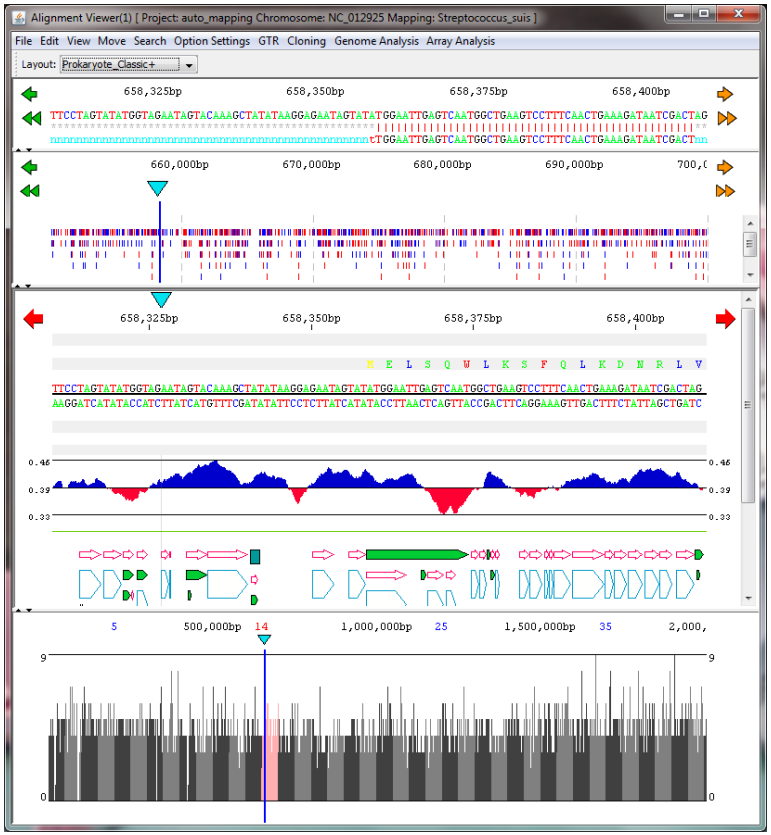
Secondary Analysis

- Analysis of mapped data

General Information

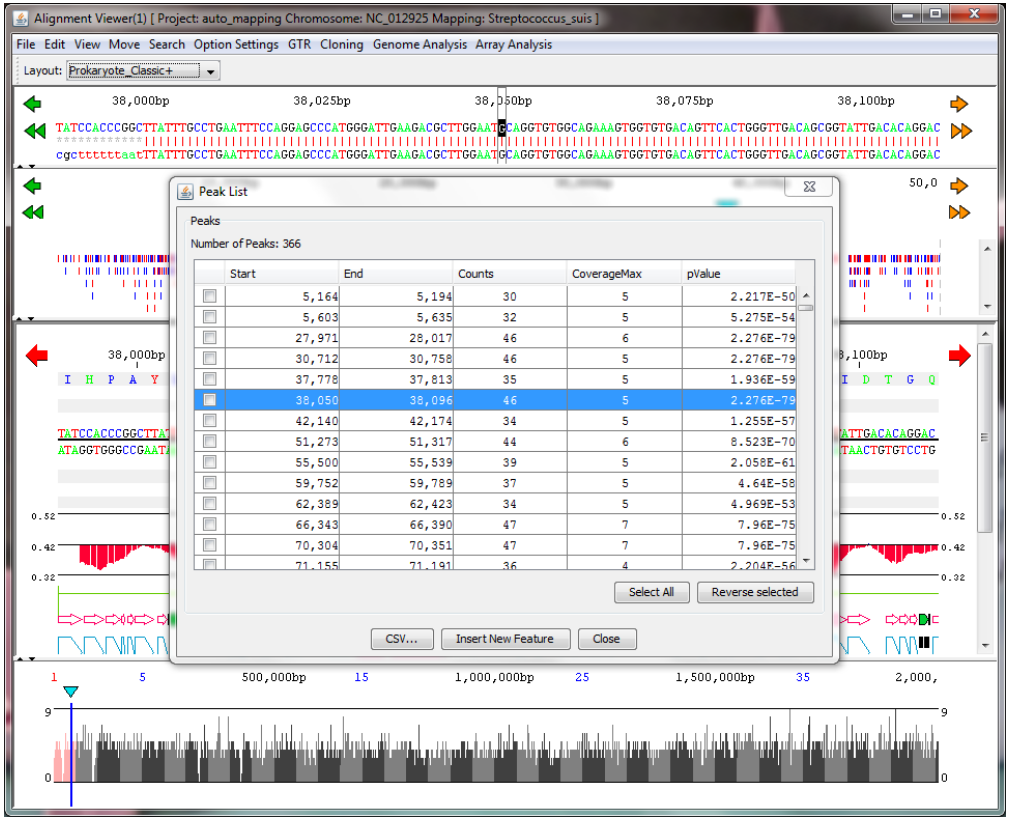
- Once NGS data is mapped, it can be analyzed in the "Profile and Alignment Viewer".

- The viewer consists of four panels:
  - Whole genome navigation panel
  - Profile panel
  - Genome (feature) map
  - Sequence alignment



Profile Analysis

- Peak detection can be done for one, a few or all contigs in one alignment.
- Peaks can be saved as CSV file or inserted as new features into the genome map.

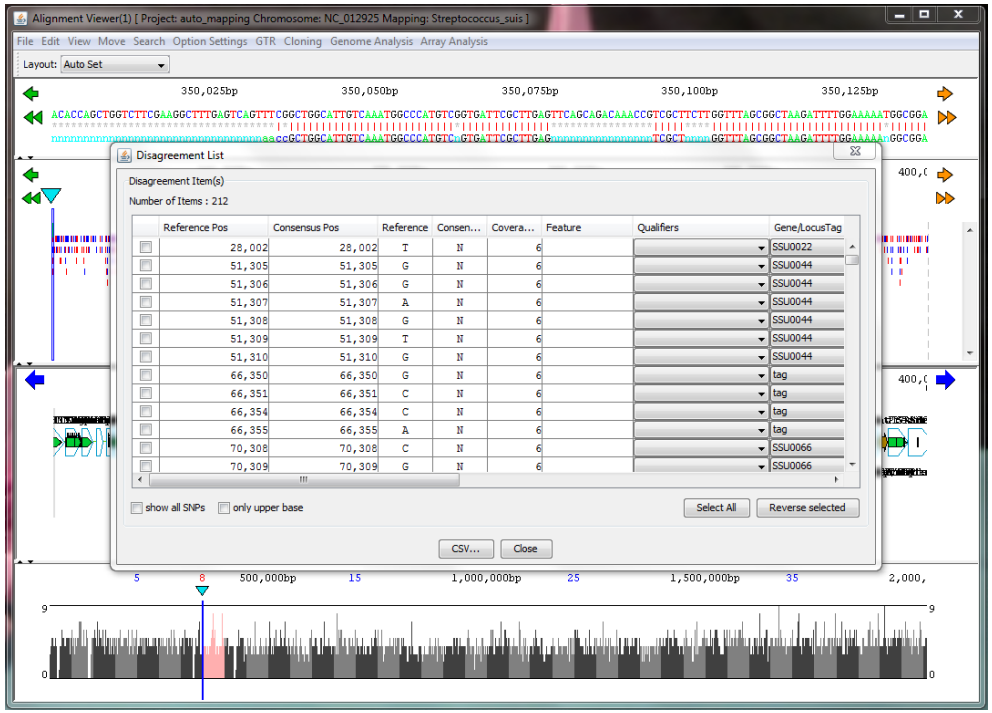


Mutation Analysis

- The "Disagreement List" is a list of bases which differ between the reference and the consensus of mapped reads (example).

- SNP detection
  - Based on disagreement list
  - Feature mapping
  - Link generation with JSNP

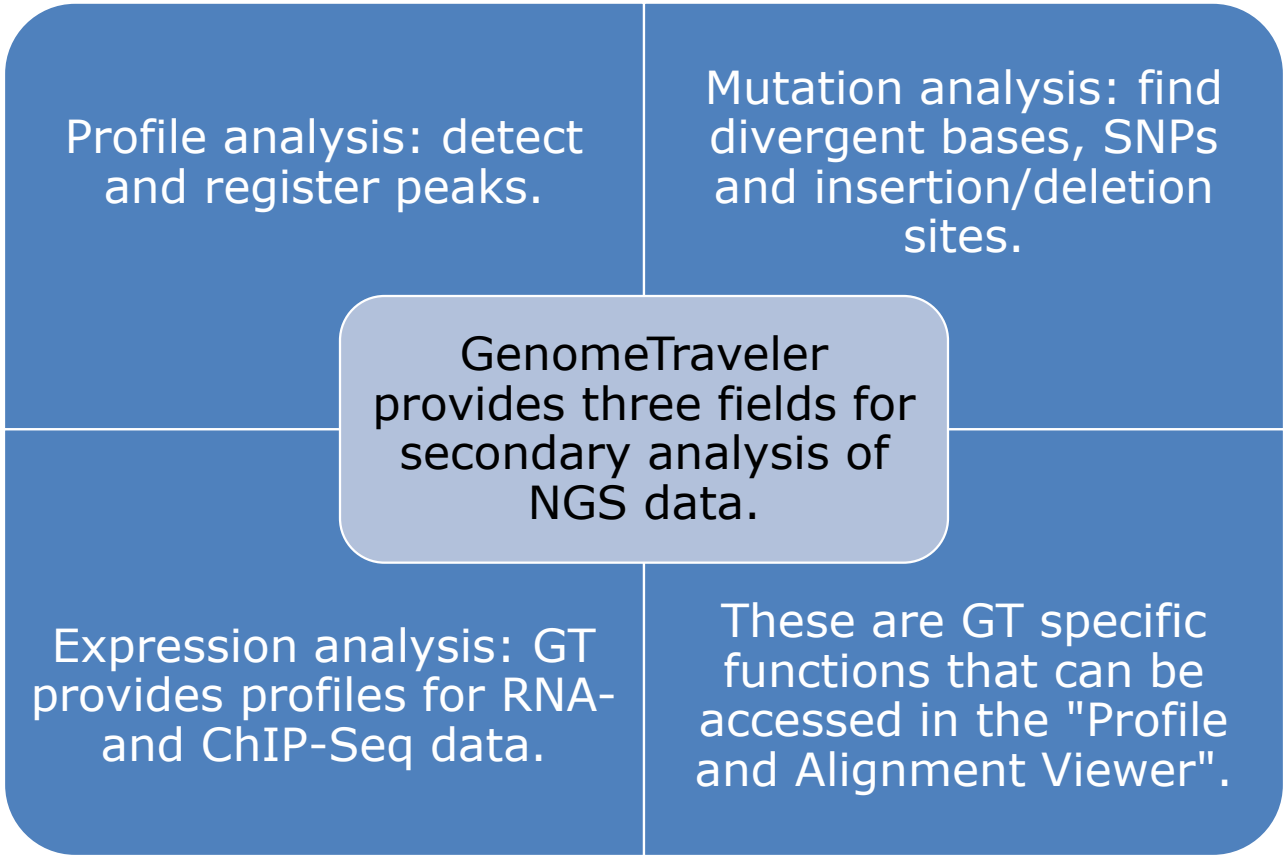
- Indel detection



Expression Analysis

- GT provides profiles and a comparison with expression profiles of tiling arrays for:
  - RNA-Seq
  - ChIP-Seq

Secondary Analysis Summary



Comparative Genomics

- Functions inherited from IMC

General Information

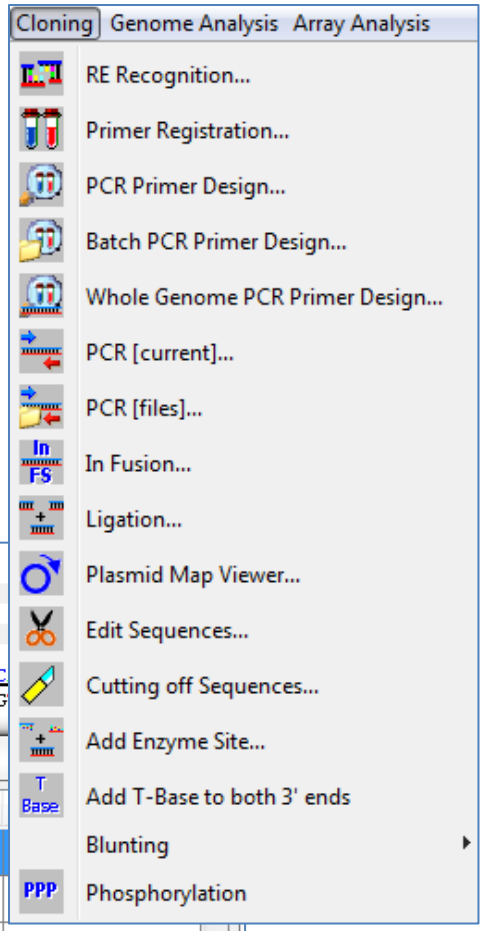
GenomeTraveler is based on IMC (*in silico* Molecular Cloning), which offers three major *in silico* experimentation areas.

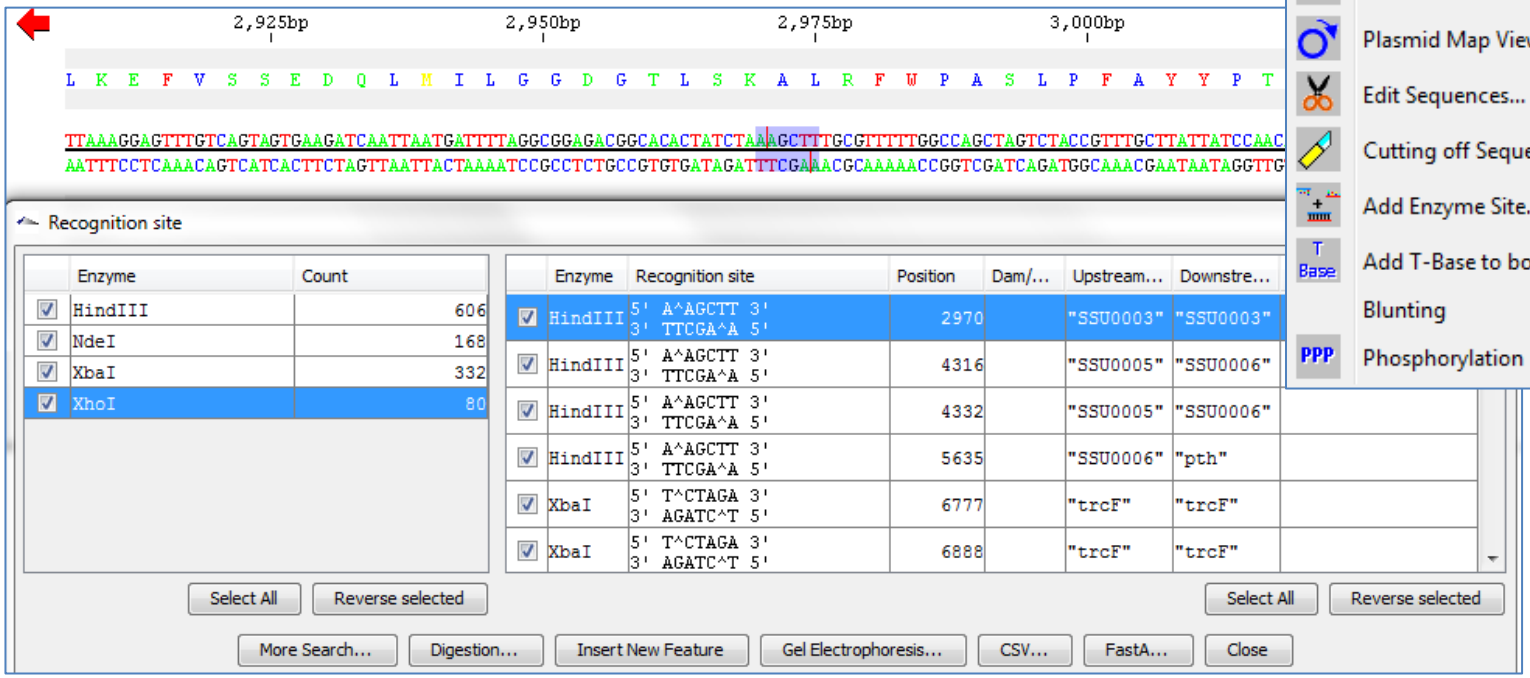
These are also featured in GT and can be accessed in the "Profile and Alignment Viewer".

in silico Cloning

The "Cloning" menu lists a number of possible actions for *in silico* molecular cloning experiments. Included are all necessary functions for:

- PCR (including gel electrophoresis of results)
- Ligation and sequence editing
- Restriction enzymes and digestion (example below)

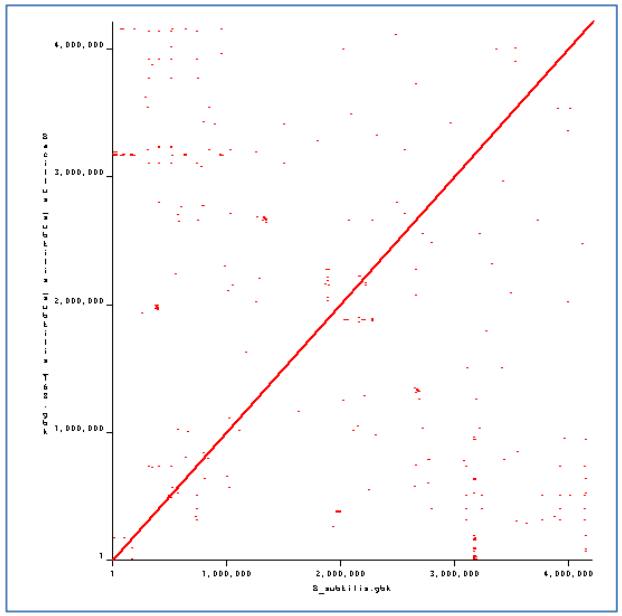


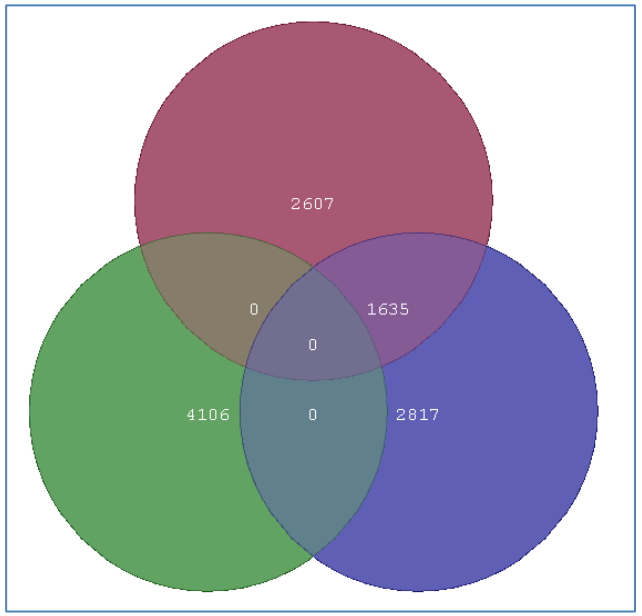


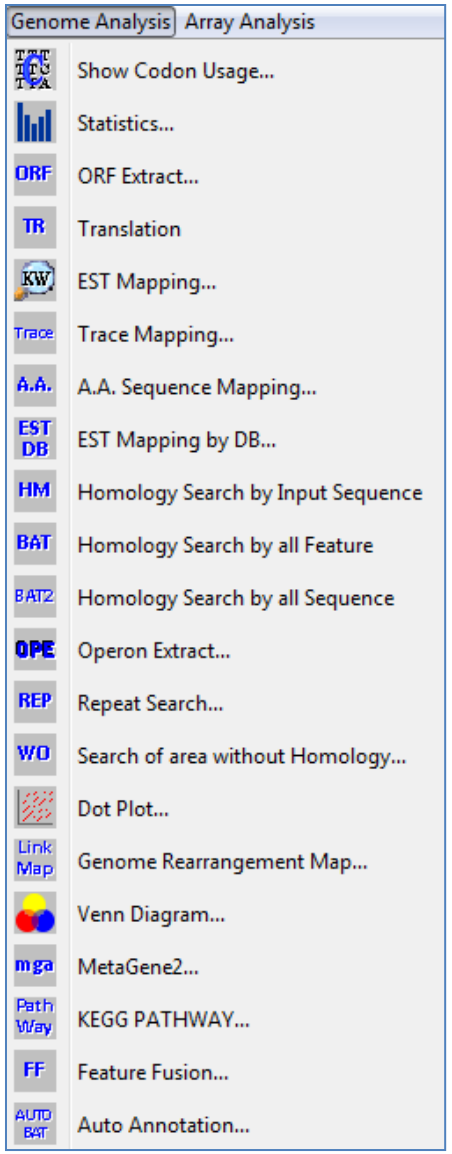
Genome Analysis

The "Genome Analysis" menu is filled with functions from statistics to visualization to working with sequences:

- Translation
- Annotation
- EST and trace mapping
- Homology searches
- Dot plot, Venn diagram (examples below), KEGG pathways







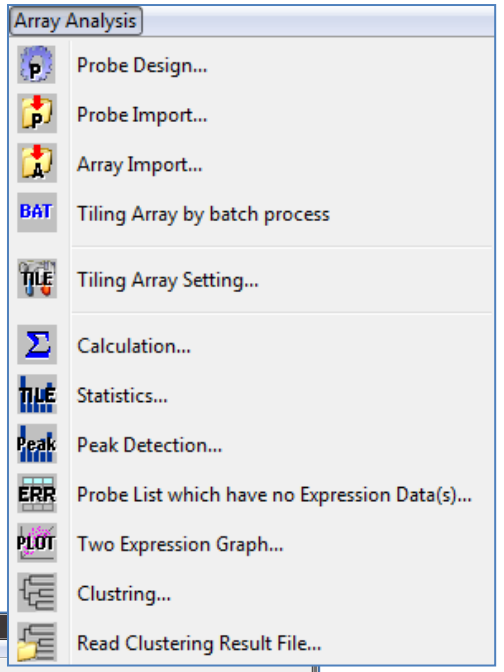
Genome Analysis

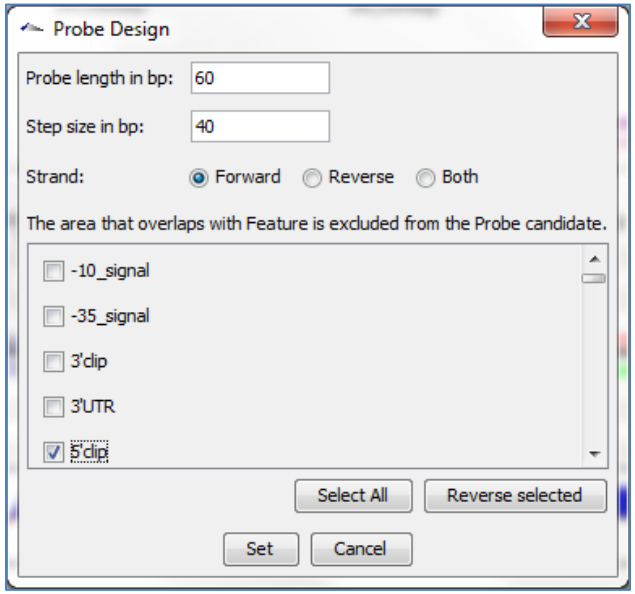
The genome rearrangement map, as part of the genome analysis functionality, compares the distribution of features on your mapped sequence compared to any reference sequence you set (in GenBank/EMBL/GPFF formats), as seen in the example below.

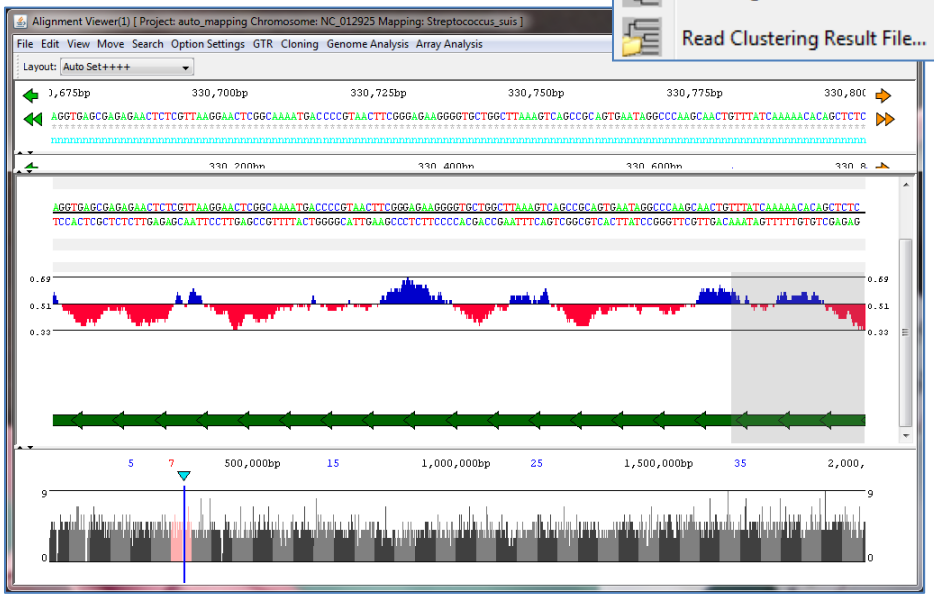
Array Analysis

The "Array Analysis" menu contains all functions necessary for the creation and analysis of array data:

- Probe design (example below)
- Import of probe and array data
- Statistics and peak detection
- Two expression graphs and clustering



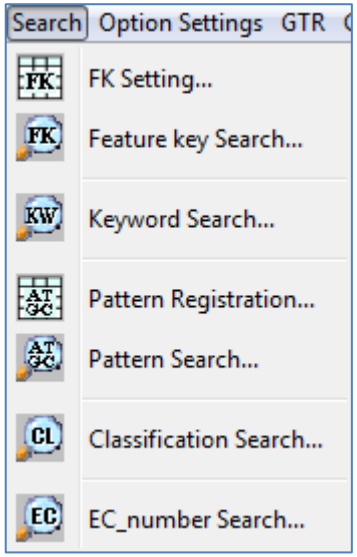


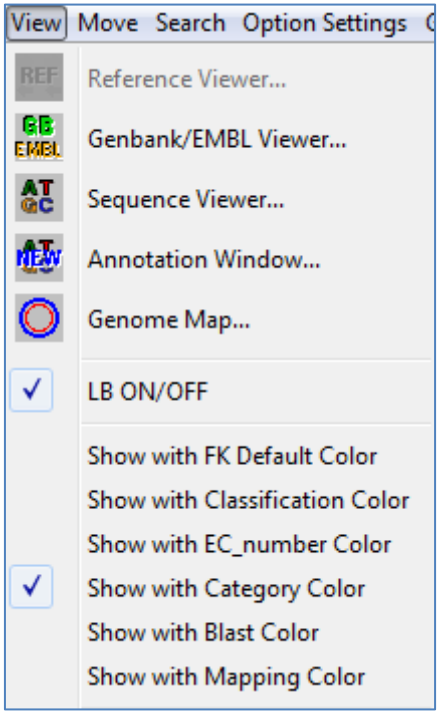


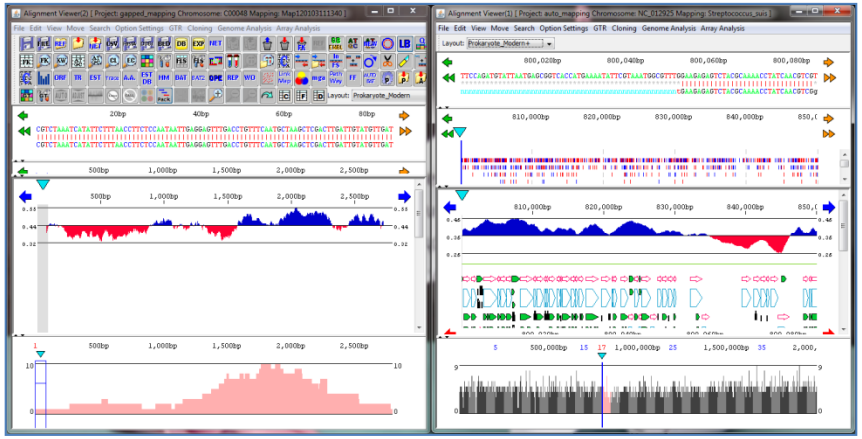
Additional Functions

The "Profile and Alignment Viewer" has more to offer, though:

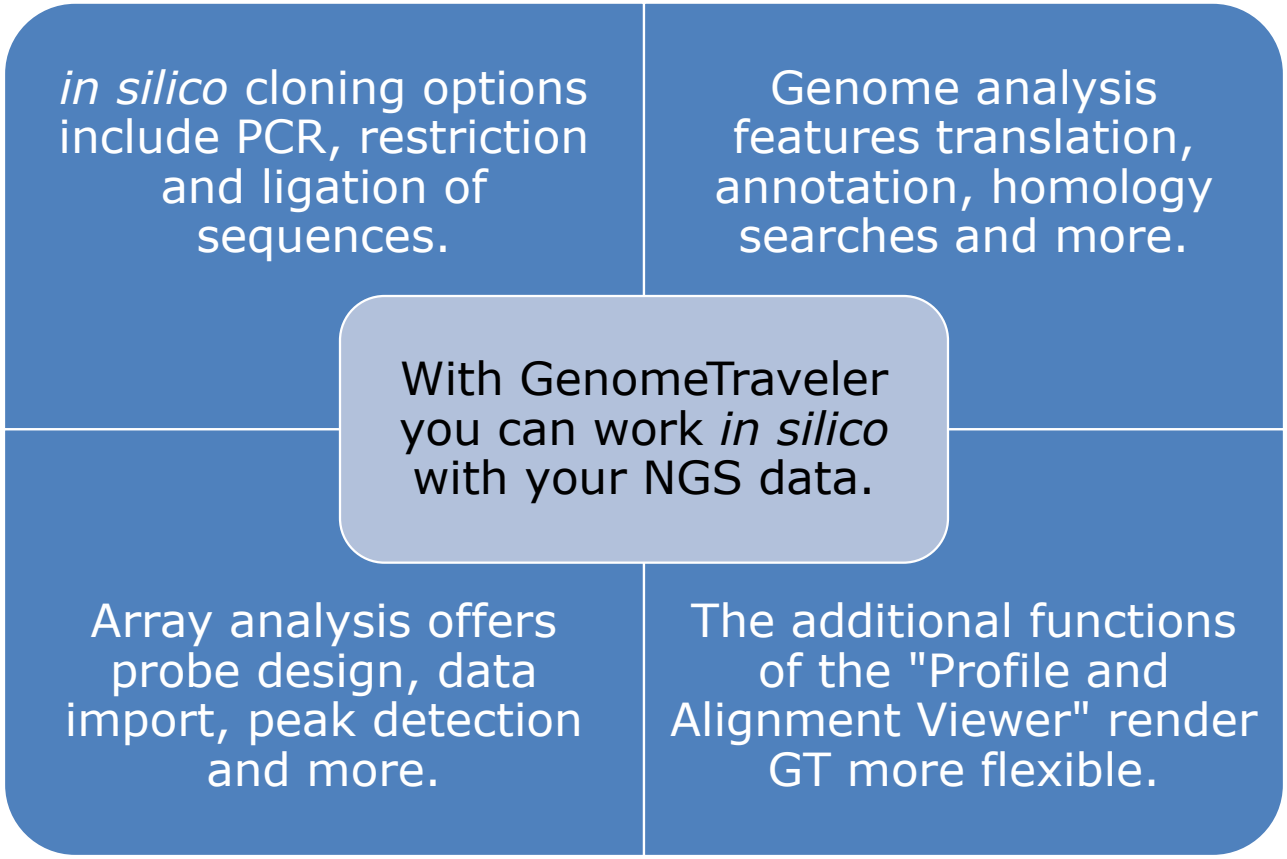
- Search for pattern matches, keywords, and more.
- View your sequence in different ways.
- Customize the genome feature map and hide the program's quick access buttons.







Comparative Genomics Summary



Why You Should Try GenomeTraveler

- Advantages of the Software

Key Notes

- GenomeTraveler has a number of options for the primary analysis of NGS data.
- GenomeTraveler provides three fields for secondary analysis of NGS data.
- With GenomeTraveler you can work *in silico* with your NGS data.

- You can download a free two-week trial version of GenomeTraveler at http://www.genexplain.com/genome-traveler