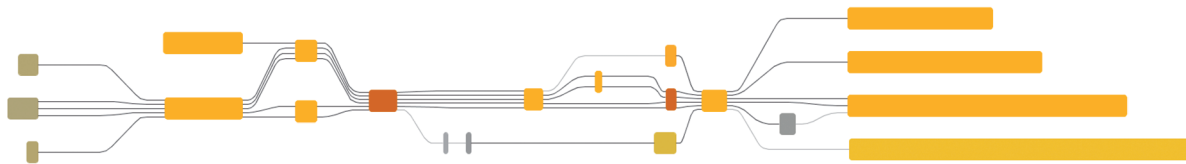


Understanding transcript expression and finding novel variants



Data content

The interpretation of RNA-Seq data requires accurate, comprehensive annotation. This is now more true than ever, as sequencing delivers unbiased coverage with ever increasing read quality and depth. To take full advantage of this, Genomatix provides a comprehensive promoter and transcript annotation database. In addition, a substantial literature annotation database (expert and automated curation) is used to generate regulatory networks and interpret novel findings in light of previous publications.

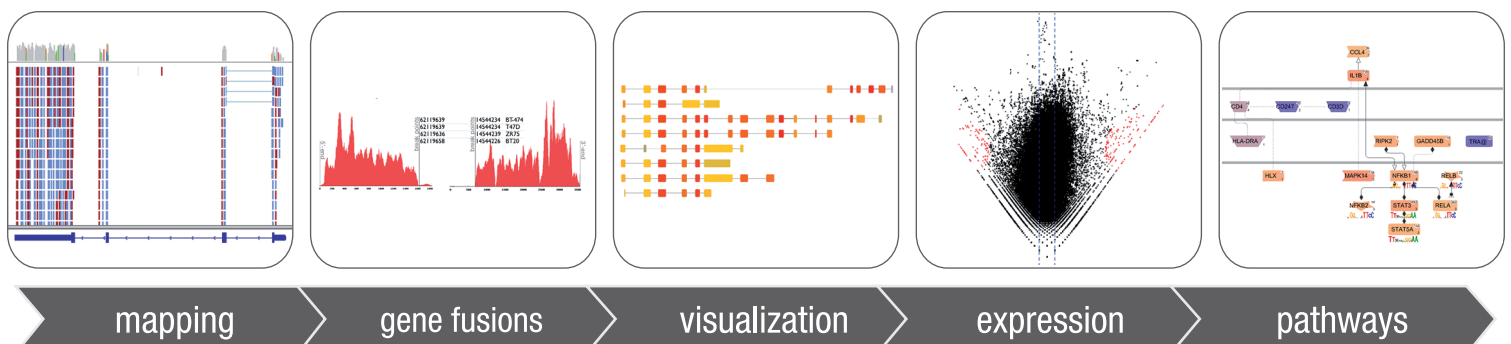
Genomatix mapping & libraries

For mapping RNA-Seq reads, Genomatix supports multiple libraries (mapping references) each tailored for a specific analysis. To obtain gene/transcript coverages and normalized expression values, the transcriptome library is used. To get information on novel splice-junctions, reads are mapped to the splice-junction library. To find novel exons and transcripts the genome library can be used. When mapping to the genome, we support multi-spliced alignments with additional splice-signal detection.

Biological insight

It has always been a strength of Genomatix to provide powerful tools for the downstream analysis of high throughput data - to get to the biology behind it. For RNA-Seq the following tools are of particular interest: gene fusion detection, alternative splicing analysis, gene regulatory networks, signaling pathways and networks (GePS), gene enrichments, differential expression and a large body of literature annotations supporting each finding.

workflow

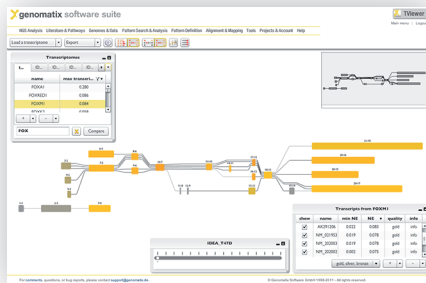


Genomatix visualization and user interface



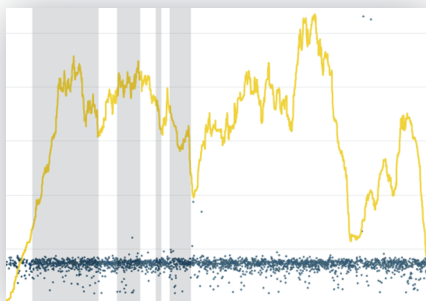
Mapping and quality control

The mapping can be performed via the intuitive graphical user interface of the Genomatix Mining Station and allows the user to select between libraries and mapping types. Multiple levels of quality control ensure read integrity and biological plausibility of the results. To the left a classification result indicating the exon enrichment of an RNA-Seq mapping is shown.



Genomatix Transcriptome Viewer

Our dedicated RNA-Seq viewer has two basic views: the standard view with each transcript drawn separately, and the splicing graph view that merges identical exons (see left). Exons and splice-junctions are drawn according to their read coverage. Transcript coverage information is shown in the paired-end transcript coverage plot.



Paired-end transcript coverage plot

The paired-end viewer shows the coverage of sequenced fragments and the distance of the respective paired-end reads along the transcript. Fragments are only considered if both mates uniquely map. This shows if the transcript is completely covered with reads and if they are within the expected distance.

mapping statistics paired end RNA-Seq			
library	all	unique	spliced
genome	94%	82%	5%
transcriptome	88%	8%	-
splice junction	18%	8%	-
assembled reads (velvet)			
genome	97%	93%	19%

Mapping statistics

Mapping statistics of eight breast cancer cell lines. The reads were mapped to all Genomatix libraries. The low number of unique hits on the transcriptome and splice-junction library is due to exon sharing between transcripts. In addition, the RNA-Seq reads were assembled with velvet and mapped against the genome using our multi spliced-alignment.