

Finding genomic variants

and understanding their downstream effects



Mapping

The detection of SNPs and structural variants depends on good read quality and a reliable mapping approach. The Genomatix mapper aligns reads with a multi-seeding approach, leading to higher sensitivity and specificity. For each mapping we report all reads that map uniquely to the genome but also those that map multiple times or even ambiguous. To support the detection of structural variants, spliced alignments are used, enabling us to map individual reads to different chromosomes.

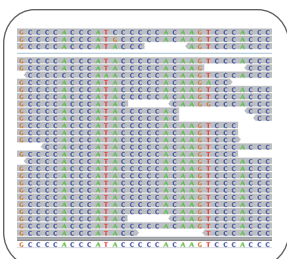
Genomic variants

After the mapping different genomic variants can be detected, including homozygous and heterozygous SNPs, structural variants and copy number variations (CNVs). SNPs are categorized into: initiating, missense, nonsense, read-through, (non)-synonymous or "changes in TF binding sites". CNV regions can be detected in our copy variant visualizer, where different conditions can be compared. Structural variant detection based on paired-end sequencing is another feature of the software.

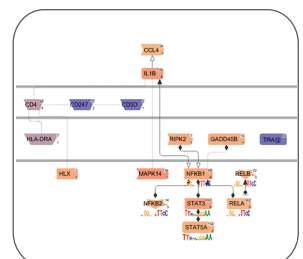
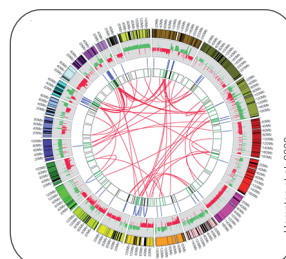
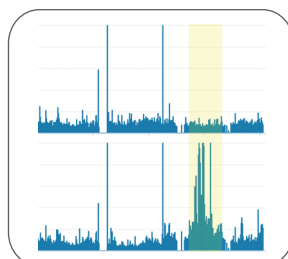
Biological insight

SNPs, CNVs and structural variants can all affect gene regulatory pathways and signaling networks. Thus - to understand the biological implications - we offer a whole suite of downstream analysis tools for gene regulatory networks, gene enrichment, comparative genomics, pathways and networks, literature mining as well as extensive annotation that can all easily be combined for a multiple-evidence based interpretation of your data..

workflow



```
ATCGCTGAGCATAG
ATCGCTCAGTATAG
ATCGCTCAGTATAG
ATCGCTCAGCATAG
ATCGCTCAGCATAG
```



mapping

SNPs

CNVs

variants

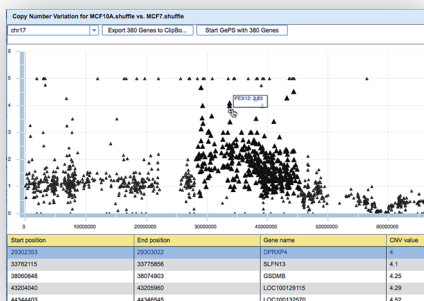
biology

User interface and downstream analysis



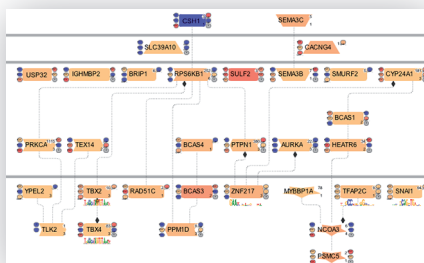
Mapping

Mapping is performed using our Genomatix mapping algorithm which allows to specify flexible settings for each sequencing experiment. In comparison to several public domain solutions our technology produces higher quality results in most cases. For more information please contact us to get a copy of our mapping benchmark.



Copy number variations (CNV)

Compare the normalized coverage of chromosomal regions from different experiments to quickly identify genes with different levels of coverage. To build networks from genes, simply select them and start the Genomatix Pathway System (GePS) shown below.



Biological interpretation

After detecting genomic variations, one usually wants to determine their biological effect. To the left you see a network that was generated in GePS by selecting a genomic region with high copy numbers in MCF-7 breast cancer cell lines. The association with disease terms revealed that the enriched genes are related to breast neoplasms.

mapping statistics DNA-Seq			
	unique	all	spliced
BT-474	85%	94%	0.4%
MCF7	85%	94%	0.5%
MCF 10A	82%	90%	0.4%
T47D	85%	94%	0.4%

Mapping statistics

Mapping statistics of four breast cancer cell lines from the Illumina iDEA challenge 2011. On average, 84% of the reads mapped uniquely to the genome, in total 92% (including multiple hits) could be mapped. The number of spliced-alignments - which can be used to support findings of chromosomal rearrangements - was below 1% in all cases.