strandngs

Streamlining NGS Data Management & Analysis

Reads to Discovery

RNA-Seq | DNA-Seq | ChIP-Seq | Small RNA-Seq | Methyl-Seq | MeDIP-Seq

www.strand-ngs.com

Analyze | Visualize | Annotate | Discover

strand
New Generation Healthcare

strandngs
Streamlining NGS Data Management & Analysis

Data Import

Vendor Platforms:
- Illumina
- Ion Torrent
- Roche 454
- ABI SOLiD
- Pacific Biosciences

File Formats:
- FASTA/FASTQ
- Unaligned BAM
- SAM/BAM
- BED, Counts data
- MiSeq run folder
- VCF/VAL

Library Layouts:
- Single end
- Paired end
- Mate paired
- Directional single/paired end

Alignment

- BWT based algorithm
- SSE/GPU based fast implementation
- Short and long reads alignment
- Single and paired end alignment
- Split read alignment
- Arbitrary gaps and mismatches
- Multiple matches
- Quality/adapter based trimming

Data QC

Pre-alignment QC Plots:
- Base level quality distribution
- Read level quality distribution
- Base composition / quality plots

Post-alignment QC Plots:
- Alignment score distribution
- Mapping quality distribution
- Illumina lane/tile QC plots
- Read length distribution plots
- Read quality distribution plots

Filter Low Quality Reads:
- Filter by read metrics
- Filter duplicate reads
- Filter multiple mapping reads
- Filter by tile quality
- Filter by region lists

RNA-Seq

- Gene, exon and transcript level quantification
- Differential gene expression
- Differential gene splicing
- SNPs, MNPs and InDels
- Novel genes
- Novel splice junctions
- Gene Fusions

DNA-Seq

- WGS, WES, Targeted panels
- SNPs, InDels and SVs
- Annotate with dbSNP, COSMIC
- Effect on transcripts
- SIFT, Polyphen2 predictions
- Multi-sample SNP analysis
- Copy number analysis

ChIP-Seq

- Peak detection using PICS and MACS
- TF regulation binding sites
- Identify affected genes
- Histone modification sites
- ChIP sample vs control
- Motif detection
- Scan motifs in the genome

Small RNA-Seq

- Quantification of miRNA, tRNA, snRNA, snoRNA and scRNA
- Novel small RNA prediction
- Differential gene expression
- Target mRNA prediction using TargetScan, PicTar, microRNA.org, PITA

Methyl-Seq

- Detect hyper- and hypo-methylation
- Detect DMCs and DMRs
- Perform intra-sample analysis
- Perform methylation effect analysis

MeDIP-Seq

- Normalize using a Calibration curve
- Detect hypo- and hyper-methylation
- Annotate methylation call

Strand NGS - Server Edition

- Collaborative analysis
- Centralized storage
- Scalable compute
- Web-based interface for system administration
- Easy and flexible deployment

Biological Interpretations

- GO enrichment, GSEA, GSA
- Single experiment OR Multi-omics analysis
- Identify significant pathways
- Curated / literature-derived pathway rendering
- Intuitive data overlay
- Create custom pathways

Discovery

Rich visualizations:
- Gene view
- Variant support view
- Elastic genome browser

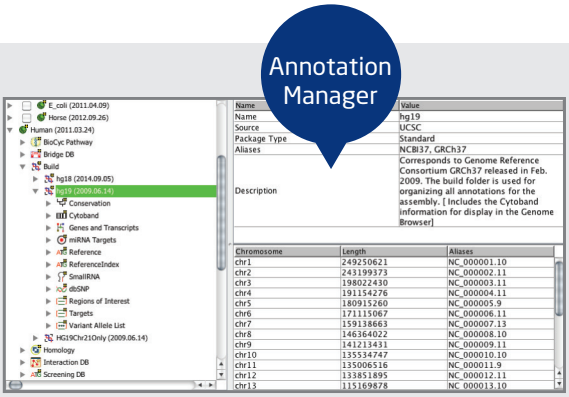
Managed Annotations:
- Pre-packaged RefSeq, UCSC, NSEMBL, and dbSNP annotations
- Custom annotations
- BioCyc pathways

Customization:
- Jython scripts
- R-scripts
- Configurable pipeline

strandngs
Streamlining NGS Data Management & Analysis

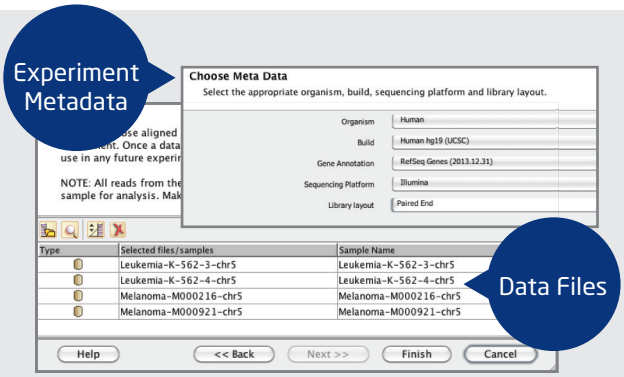
Comprehensive Annotations

- Pre-packaged gene and transcript annotations from UCSC, RefSeq and ENSEMBL for all model organisms.
- SNP annotations from dbSNP and COSMIC, SIFT/Polyphen scores from dbNSFP, small RNA annotations, miRNA target prediction databases, screening databases and more.
- Ability to create annotations for other organisms from gbk/gtf files or FASTA files.



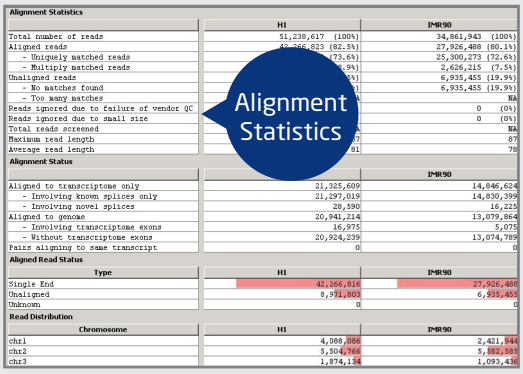
Import Data

- File Formats – FASTA/FASTQ, Unaligned BAM, SAM/BAM, Counts data.
- Library Layouts – Single End, Paired End, Mate Paired as well as Directional layouts.
- Vendor Platforms – Illumina, Ion Torrent, 454 Roche, SOLiD, PacBio.



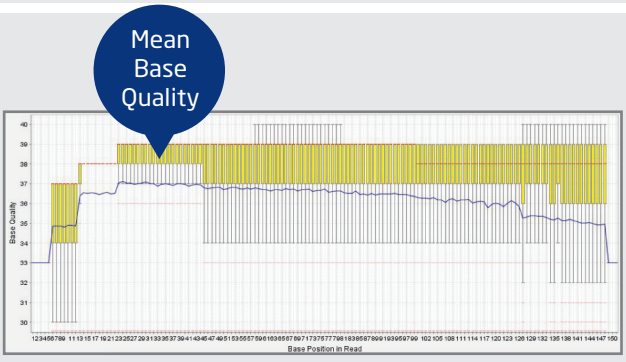
Alignment with Strand NGS aligner

- Alignment for small RNA, DNA-Seq, ChIP-Seq and RNA-Seq data.
- Targeted region alignment for resequencing applications. Handles variable length reads, arbitrary number of gaps and mismatches and paired reads as well.
- Split read alignment for detecting long InDels and translocations.
- Options for trimming adaptors, low quality bases and screening reads against standard databases.
- Aligns ~8 million DNA reads against hg19 per hour per core on a 64GB RAM machine.



Quality Control Plots

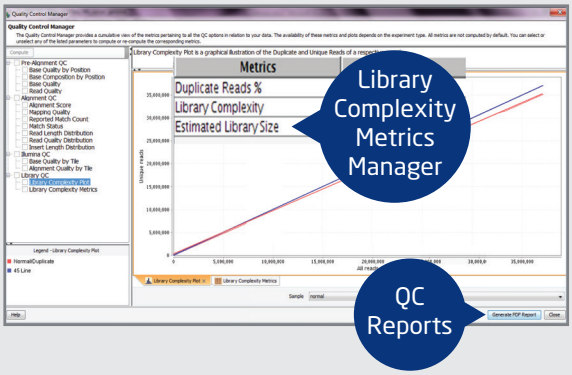
- Base and read quality distributions.
- Base composition and quality distributions by position in read.
- Read length and insert length distributions.
- Alignment score and mapping quality distributions.
- Mate status QC plot.
- Vendor-specific QC plots.
- Targeted region QC.



strandngs
Streamlining NGS Data Management & Analysis

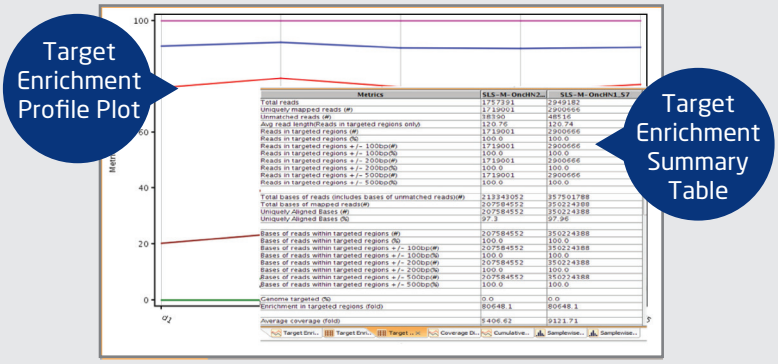
Quality Control Manager: Library QC

- Multiple quality inspection options including pre-alignment, post-alignment, vendor-specific QC and Library QC.
- Automatically generates a QC report for every sample.
- Export QC report as a pdf document.



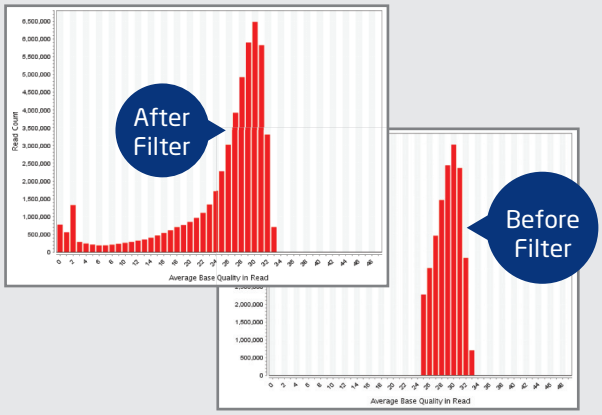
Quality Control Manager: Target Enrichment QC

- Efficiently analyze data generated from targeted resequencing experiments.
- Target region-based quality control for your customized panels.
- Assess coverage distribution across target regions including target enrichment profile plot, metrics, summary table, and sample-wise base frequency coverage distribution.



Filter Low Quality Reads

- Filter by read quality metrics.
- Filter by targeted regions of interest.
- Filter duplicate reads and multiple mapping reads.
- Filter by samples.

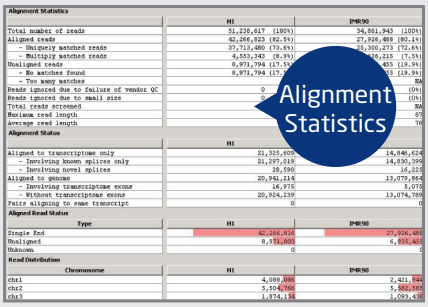


RNA-Seq Analysis

strandngs
Streamlining NGS Data Management & Analysis

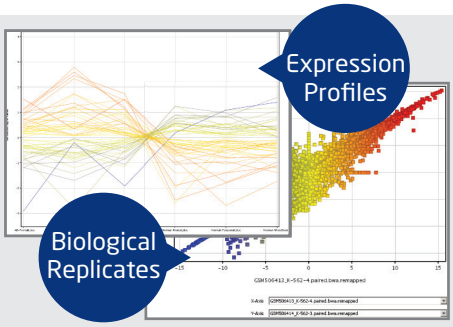
Alignment and Data QC

- Alignment against the transcriptome as well as the whole genome to enable detection of novel splice forms and genes.
- Handles variable length reads and paired reads as well.
- Allows arbitrary number of gaps and mismatches.
- Options for trimming adaptors, low quality bases and screening reads against standard screening databases.
- Perform RNA specific QC for genic regions and gene types.



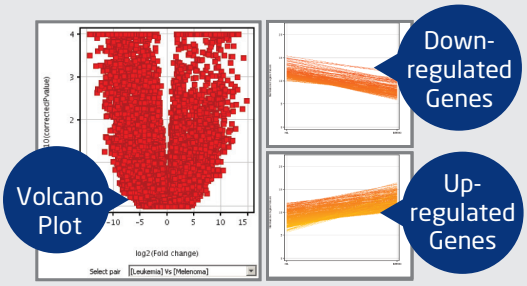
Quantification & Normalization

- Expression values at gene, exon and transcript levels.
- Appropriate handling of partially overlapping reads and multiple mapping reads.
- DESeq, RPM, TMM and Quantile methods for normalization.
- Experiment grouping supports large scale project handling.



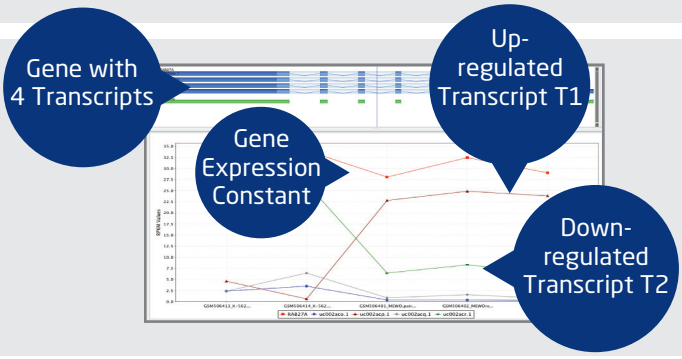
Differential Expression

- t-Test, Mann-Whitney, N-way ANOVA for identifying differentially expressed genes under different experimental conditions.
- Multiple Testing Correction using Benjamini Hochberg, Storey, Bonferroni, etc.
- Fold Change computation and visualization.



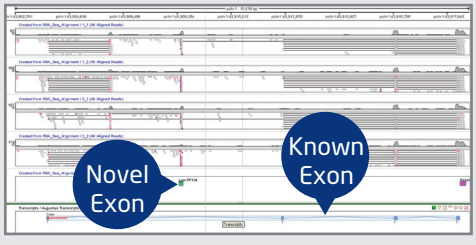
Differential Splicing

- EM algorithm to de-convolute gene counts to transcript counts.
- Identification of differentially spliced genes.
- Visualization in the gene view.



Novel Discovery

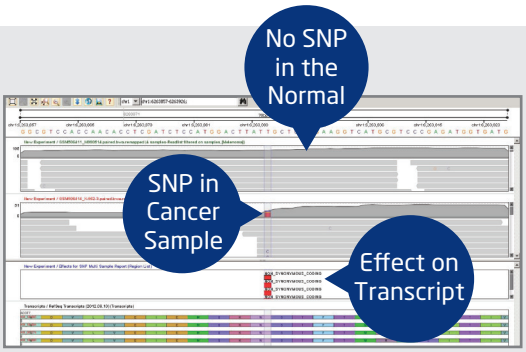
- Detection of novel genes, novel splice junctions, and novel exons in known genes.
- Differential expression analysis of novel genes and contribution of novel exons to differential splicing.
- Prioritization of novel regions using conservation score.



strandngs
Streamlining NGS Data Management & Analysis

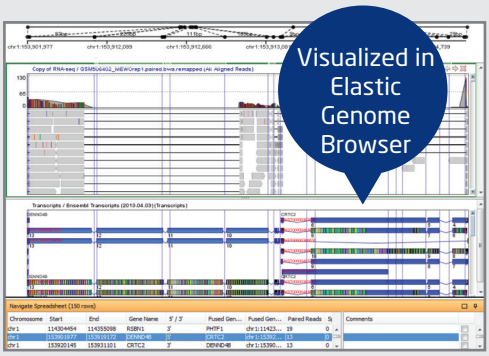
Transcriptome Variant Detection

- SNP, MNP, and InDel detection.
- Annotation with dbSNP to identify known/novel mutations.
- Prediction of effects such as non-synonymous coding, frameshift, splice-site, etc. on transcripts.
- Identify significant SNPs using an intuitive filtering framework.



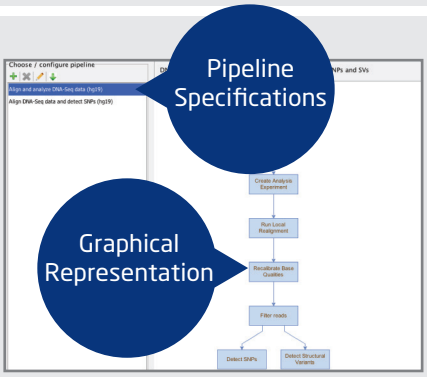
Gene Fusions

- Gene fusions detected from spliced and paired reads.
- Detection of read through transcripts.
- Gene fusions involving paralogs and pseudogenes marked as such to help filter out false positives.



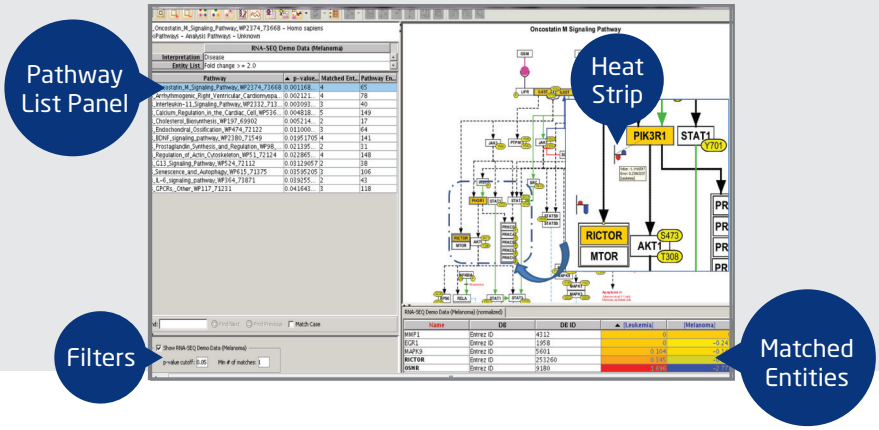
Pipeline Execution

- Pipelines that run in the background.
- Analysis pipeline includes filters and quantification.
- Pipelines support alignment of raw reads and direct import of aligned samples.
- Customization of pipelines for individual runs.



Pathway Analysis

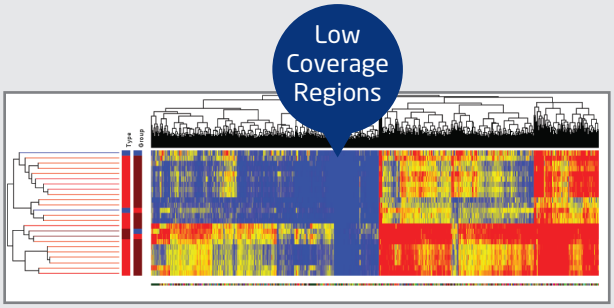
- Single Experiment Analysis (SEA) identifies enriched pathways for the genes from a single experiment type.
- Multi-Omic Pathway Analysis (MOA) from multiple genomics and transcriptomics experiments.
- Overlay differentially expressed entities on curated pathways.
- Choose from curated pathways like WikiPathways, BioCyC pathways, BioPAX pathways or literature-derived networks like NLP and MeSH.
- Find significant pathways for differentially expressed genes.



strandngs
Streamlining NGS Data Management & Analysis

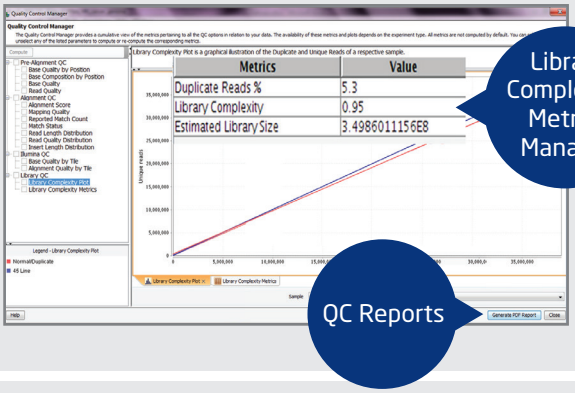
Targeted Re-sequencing Analysis

- Import BED file of target regions.
- Filter reads outside target regions.
- Evaluate efficacy of targeted re-sequencing.
- Identify target regions with low coverage across samples.
- Detect SNPs and other variants on targeted regions.



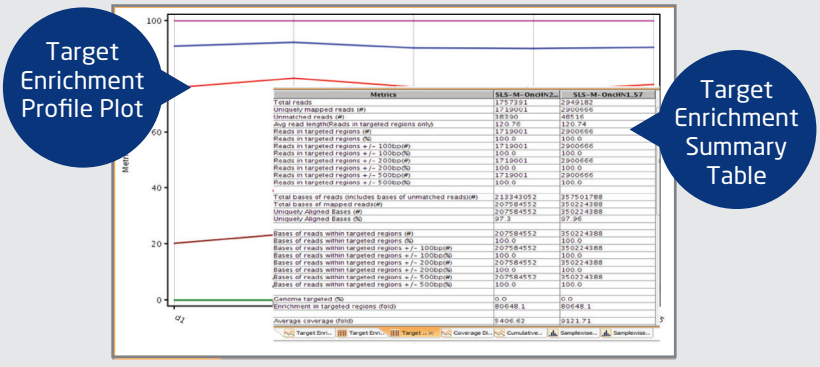
Quality Control Manager: Library QC

- Multiple quality inspection options including pre-alignment, post-alignment, vendor-specific QC and Library QC.
- Automatically generates a QC report for every sample.
- Export QC report as a pdf document.



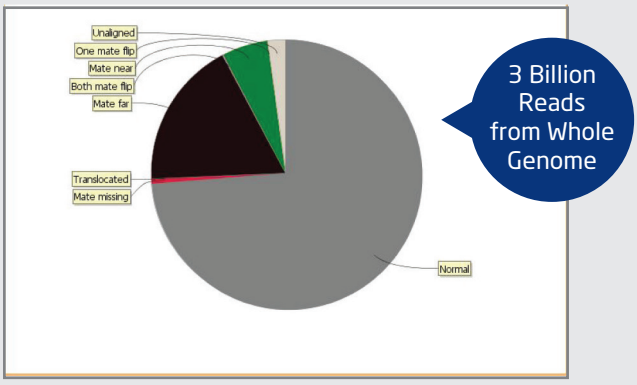
Quality Control Manager: Target Enrichment QC

- Efficiently analyze data generated from target enrichment sequencing experiments.
- Target region-based quality control for your customized panels.
- Assess coverage distribution across target regions, including target enrichment profile plot, metrics, summary table, and sample-wise base frequency coverage distribution.



Whole Genome Analysis

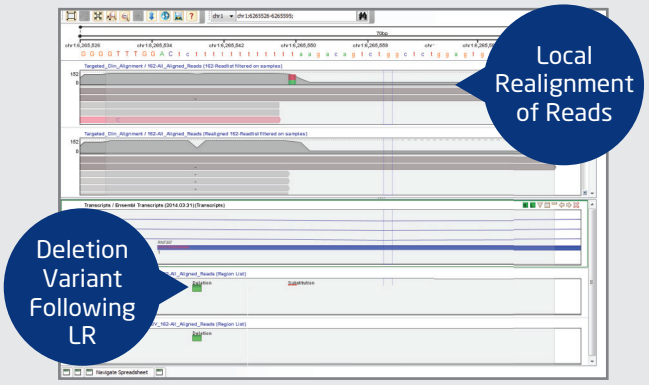
- Perform Whole Genome analysis on human or other organisms on your desktop.
- Regular desktop machine with 4GB RAM, 4 cores, and 2TB hard disk.



strandngs
Streamlining NGS Data Management & Analysis

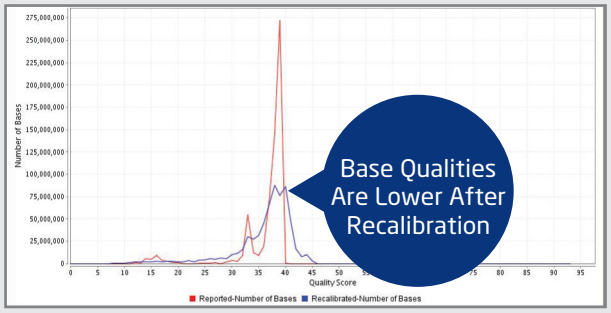
Local Realignment (LR)

- Reads with alignment artifacts around InDels can be realigned using information from multiple reads.
- Helps in getting rid of spurious substitutions and reduces false-positive variant calls.



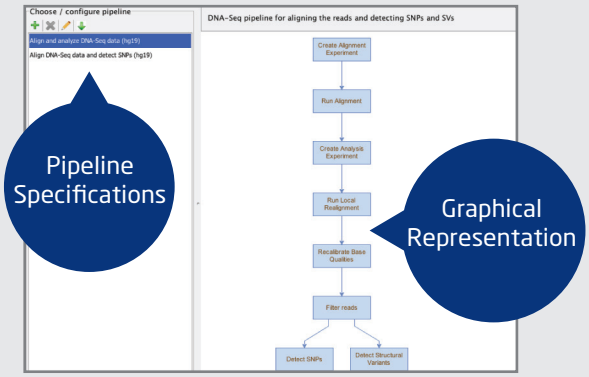
Base Quality Score Recalibration (BQSR)

- Recalibrates base quality scores to reduce errors and systemic biases.
- Uses contexts like machine cycle and dinucleotide context to recalibrate the reported base quality scores.
- Helps reduce false-positive variant calls.



Pipeline Execution

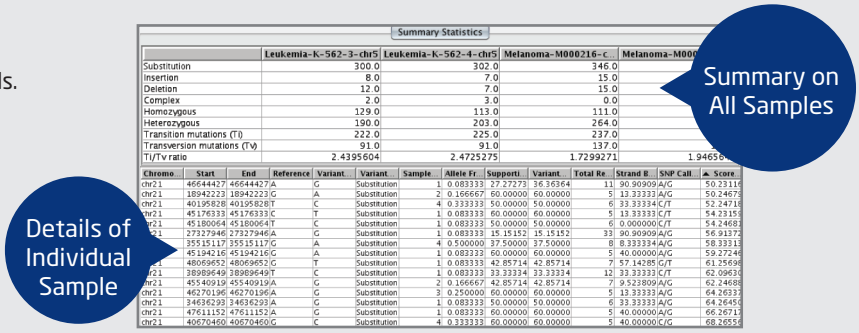
- Analysis pipelines that include filters, local realignment, base quality recalibration, and SNP detection.
- Additional pipelines that start from alignment of raw reads or direct import of aligned samples.
- Customization of pipelines for individual runs.



strandngs
Streamlining NGS Data Management & Analysis

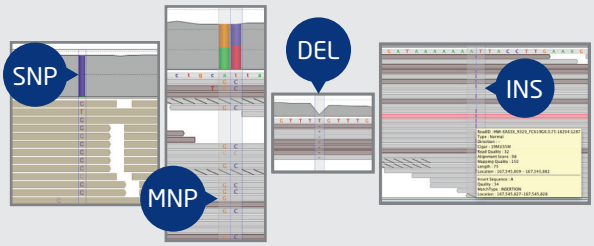
Variant Detection

- SNP Detection algorithm to detect SNPs, MNPs, and small InDels.
- View summary statistics of variants across samples.
- Visualize details of variants in each sample, along with dbSNP annotations.
- Support for VCF and VAL import.



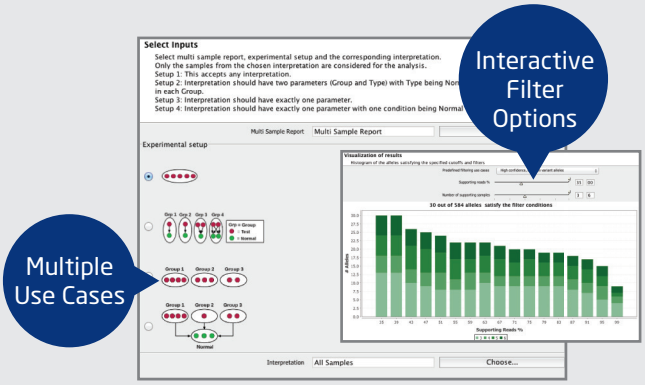
Visualize Variants

- Drag and drop SNP results into the genome browser.
- Visualize SNPs, MNPs, and InDels along with read coverage, as well as other annotations.



Find Significant SNPs

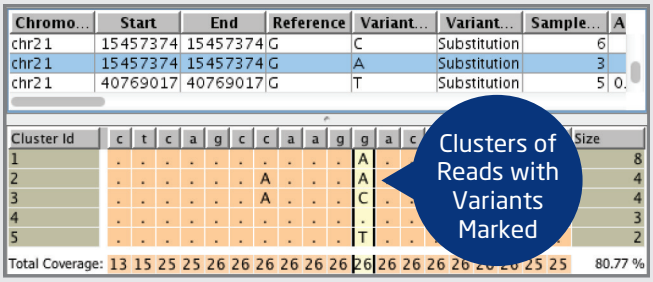
- Multiple use cases including normal-tumor, multi-group comparisons, low-frequency mutations, rare variant analysis, and somatic mutations.
- Cluster significant SNPs and samples to detect patterns visually.
- Identify significant SNPs using an intuitive filtering framework.



strandngs
Streamlining NGS Data Management & Analysis

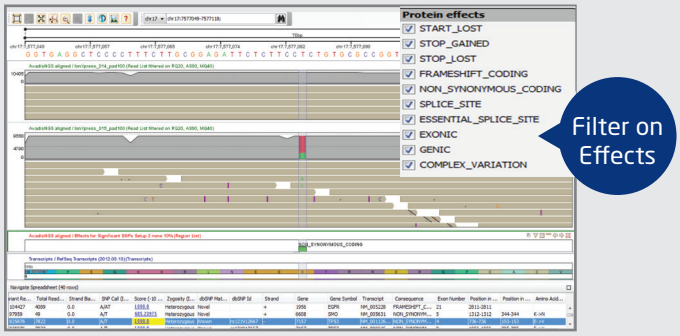
Variant Support View

- Intuitive visualization to verify individual SNPs.
- Color by base quality or mapping quality.
- Cluster reads to make variant locations stand out.
- Annotate clusters with strand information.



SNP Effect Analysis

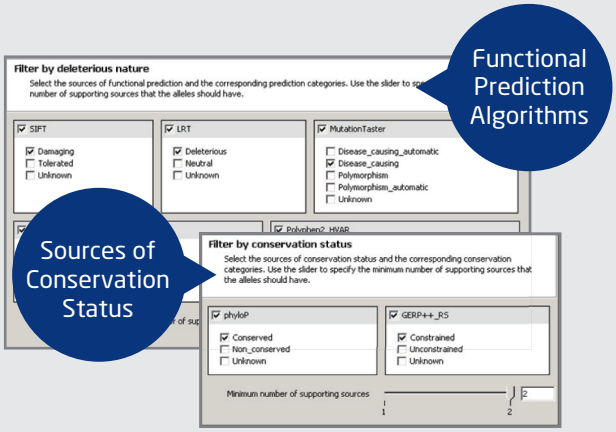
- Identify effect of detected SNPs on transcripts.
- Variety of effects including non-synonymous coding, splice site, stop gained, etc.
- Visualize along with the amino acid sequences for transcripts.
- Ability to filter on interesting effects.
- Annotate with COSMIC and other external databases.



strandngs
Streamlining NGS Data Management & Analysis

Find Damaging Variants

- Filter variants based on functional annotations for non-synonymous variants from dbNSFP.
- Identify damaging variants based on prediction scores from SIFT, Polyphen, LRT, Mutation Taster, and Mutation Assessor.
- Filter based on conservation scores from phyloP and GERP++_RS.
- Filter based on allele frequencies from 1000 genomes.



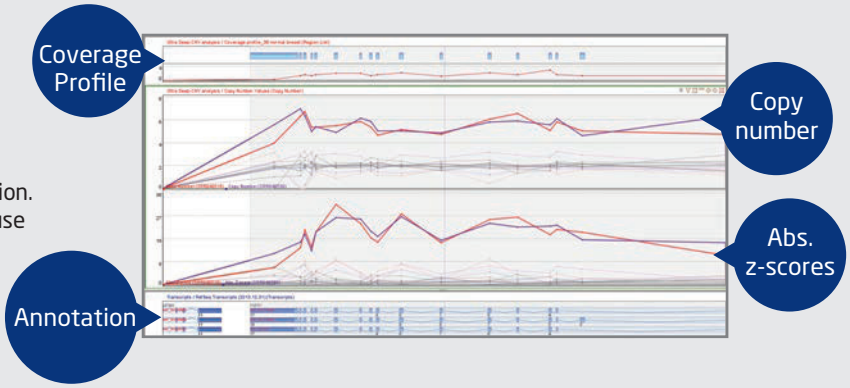
Structural Variant Analysis

- Structural variant detection on paired end and split read data.
- Identify large structural variants including large insertions, deletions, inversions, and translocations.
- Verify using Elastic Genome Browser.

strandngs
Streamlining NGS Data Management & Analysis

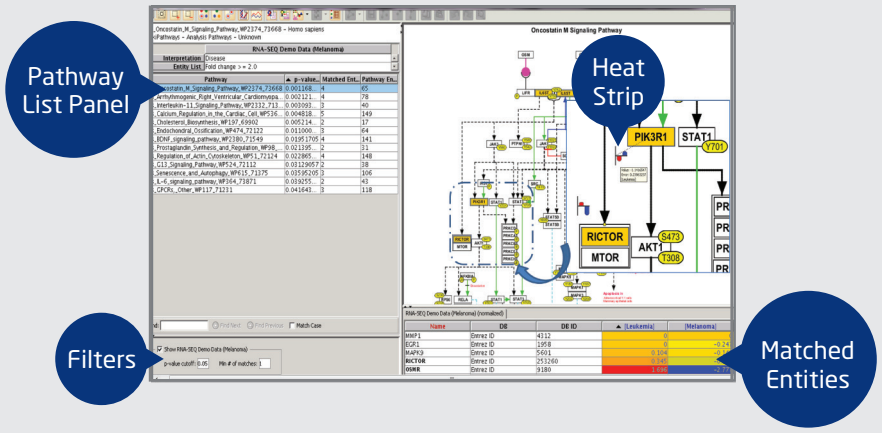
Copy Number Variant (CNV) Analysis

- Detects CNV regions in group of samples.
- Includes GC bias correction as a preprocessing step.
- Includes estimation of sample ploidy and normal cell contamination.
- Ability to create coverage profiles from a group of samples and use it for subsequent CNV analysis across experiments.
- Verify CNV's using the Genome Browser and specially designed web based visualization



Pathway Analysis

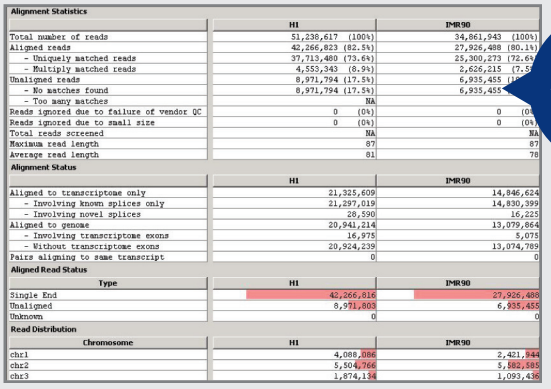
- Single Experiment Analysis (SEA) identifies enriched pathways for the genes from a single experiment type.
- Multi-Omic Pathway Analysis (MOA) from multiple Genomics and transcriptomics experiments.
- Overlay differentially expressed entities on curated pathways.
- Choose from curated pathways like WikiPathways, BioCyC pathways, BioPAX pathways or literature-derived networks like NLP and MeSH.
- Find significant pathways for differentially expressed genes.



strandngs
Streamlining NGS Data Management & Analysis

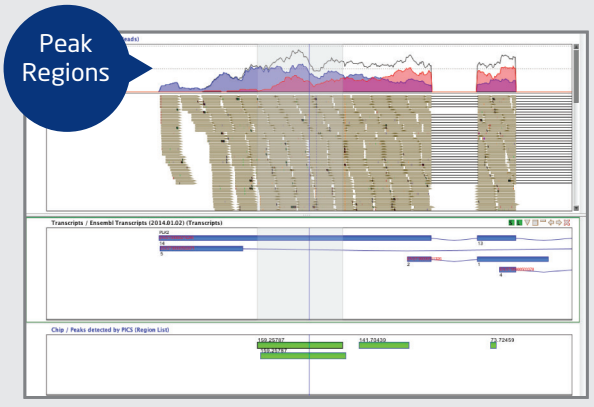
Alignment with Strand NGS aligner

- Handles variable length reads and paired reads as well.
- Allows arbitrary number of gaps and mismatches.
- Options for trimming adaptors, low quality bases, and screening reads against standard screening databases.



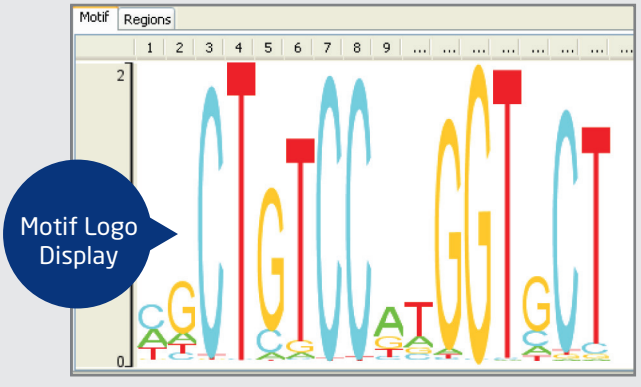
Peak Detection

- Detect enriched regions in ChIP sample as compared to the control.
- Detect peaks of transcription factor regulatory sites using PICS/MACS.
- Identify genes regulated by TF binding sites.
- Detect histone modification sites using the enriched region detection method.



Motif Detection

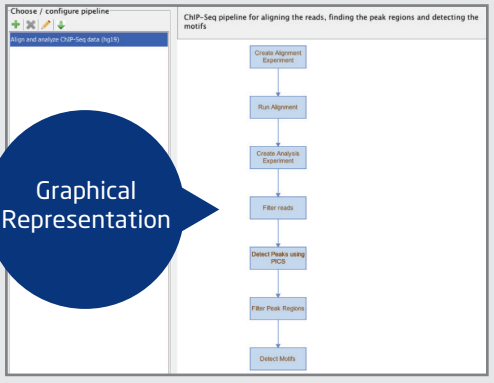
- Identify motifs in detected peak regions using GADEM.
- Import peak regions to detect motifs.
- Import motifs in JASPAR format.
- Scan for motifs in the entire genome or in regions of interest.



strandngs
Streamlining NGS Data Management & Analysis

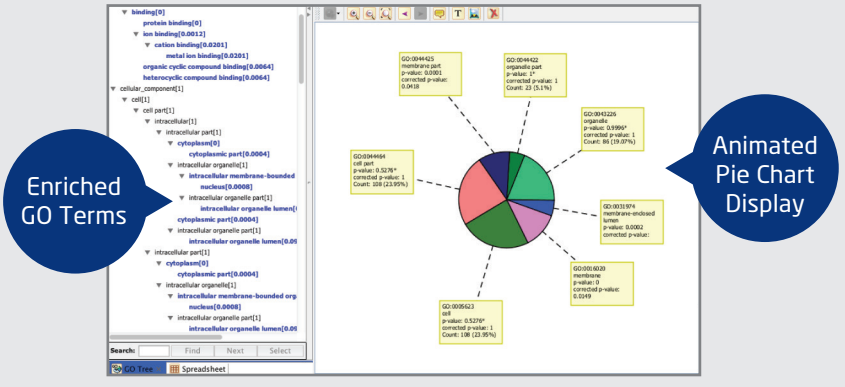
Pipeline Execution

- Pipelines that run in the background.
- Analysis pipeline that includes filters and peak detection.
- Additional pipelines that start from alignment of raw reads or direct import of aligned samples.
- Customization of pipelines for individual runs.



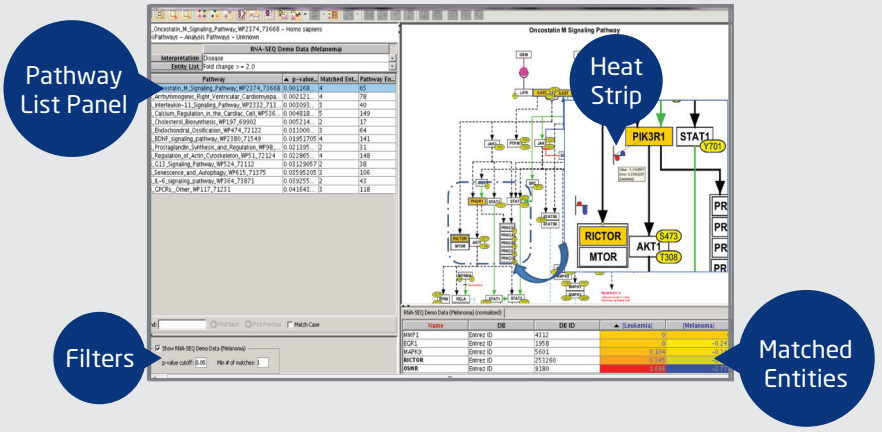
GO Enrichment

- Annotate regions with genic information to discover those affected by peak regions.
- GO Enrichment analysis to detect enriched Gene Ontology terms.



Pathway Analysis

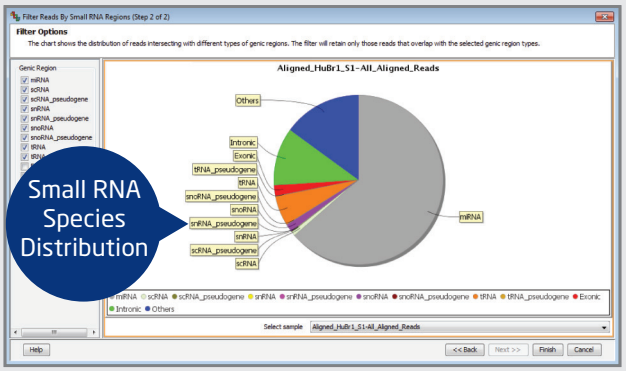
- Single Experiment Analysis (SEA) identifies enriched pathways for the genes from a single experiment type.
- Multi-Omic Pathway Analysis (MOA) from multiple genomics and transcriptomics experiments.
- Overlay differentially expressed entities on curated pathways.
- Choose from curated pathways like WikiPathways, BioCyC pathways, BioPAX pathways or literature-derived networks like NLP and MeSH.
- Find significant pathways for differentially expressed genes.



strandngs
Streamlining NGS Data Management & Analysis

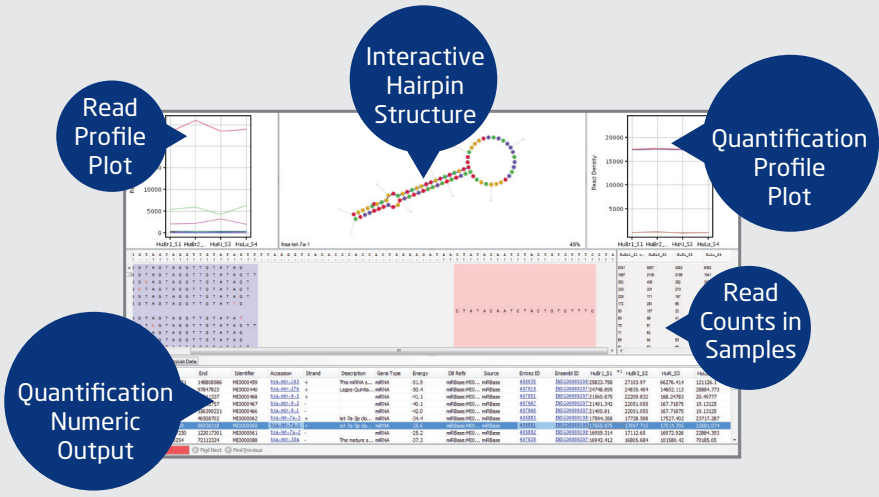
Alignment, QC and Filtering

- Read alignment with Strand NGS aligner.
- Option for trimming adapters, low quality bases and screening reads against standard screening databases
- Read distribution across genic regions and small RNA species.



Quantification & Normalization

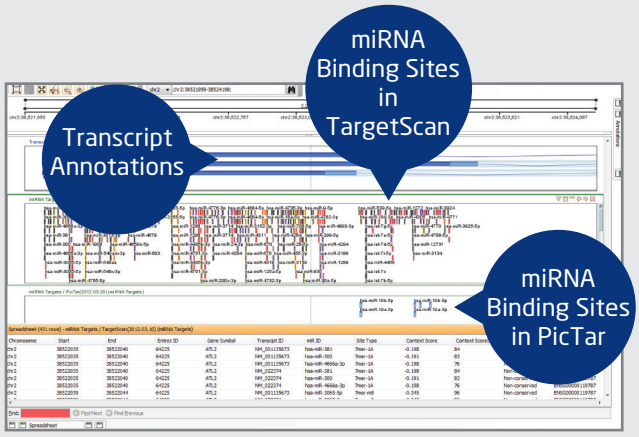
- Expression values for known genes, novel genes, mature miRNAs.
- Ability to pick reads aligning exactly with the 5′ end of mature miRNAs.
- Take into account padding and multiply mapping reads.
- DESeq, TMM, Quantile, and Sample Count based methods for normalization.
- Small RNA Gene View to visualize quantification results.



strandngs
Streamlining NGS Data Management & Analysis

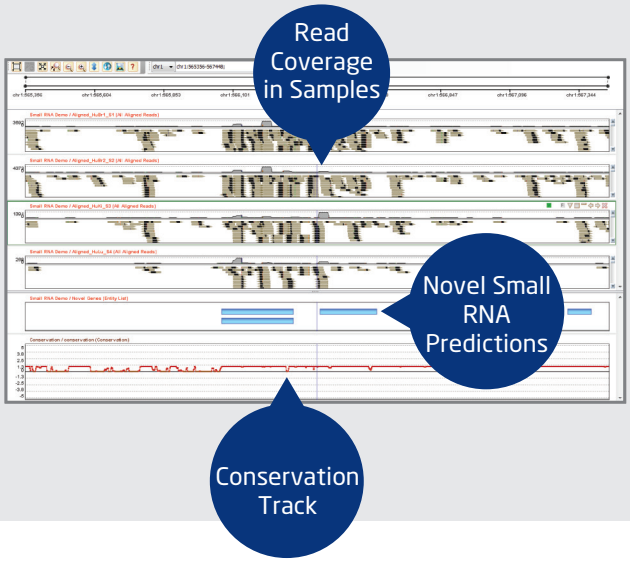
miRNA Target Analysis

- TargetScan, PicTar, TarBase, microRNA.org, and PITA databases for target prediction analysis.
- Identify targets common to multiple databases.
- Perform downstream analyses (GO, GSA, GSEA, Pathway Analysis) on target set of mRNA genes.



Novel small RNA Discovery

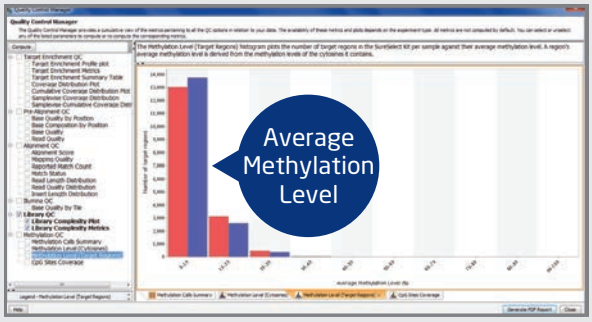
- Predict type of novel gene as one of miRNA, snoRNA, scRNA or tRNA.
- Identify high-confidence predictions with Confidence values and Conservation scores.
- Find annotation discrepancies of known genes from the read coverage patterns.



strandngs
Streamlining NGS Data Management & Analysis

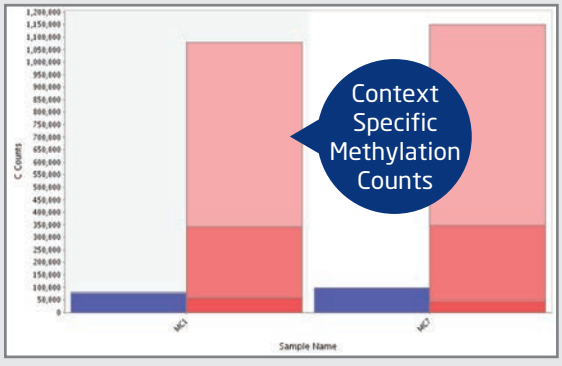
Data Import and QC

- Allows import of Bismark aligned bisulfite treated sequencing data.
- Quality metrics for pre- and post-alignment, target enrichment and library complexity.
- Compute methylation-specific QC to get information on methylation levels and CpG sites coverage for your samples.



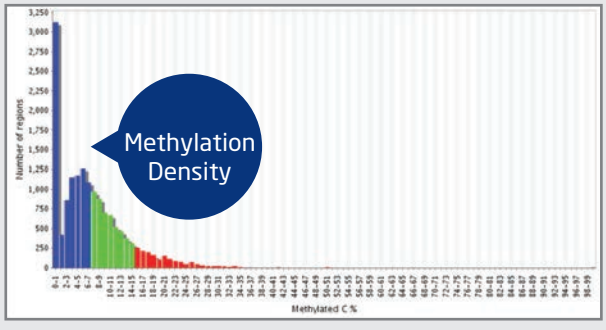
Methylation Detection

- Identify methylated cytosines for specific loci on the genome for individual samples.
- Algorithm considers bisulfite conversion error rate, sequencing error, read coverage, base quality and methylation fraction.



Intra Sample Analysis

- Identify regions with low, moderate, and high methylation density within an individual sample for target regions of interest.



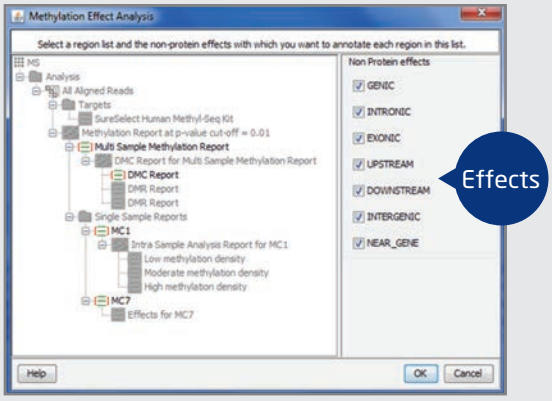
Differential Methylation

- Identify differentially methylated cytosines (DMCs) across experimental conditions/samples.
- Discover differentially methylated regions (DMRs) across experimental conditions/samples.

strandngs
Streamlining NGS Data Management & Analysis

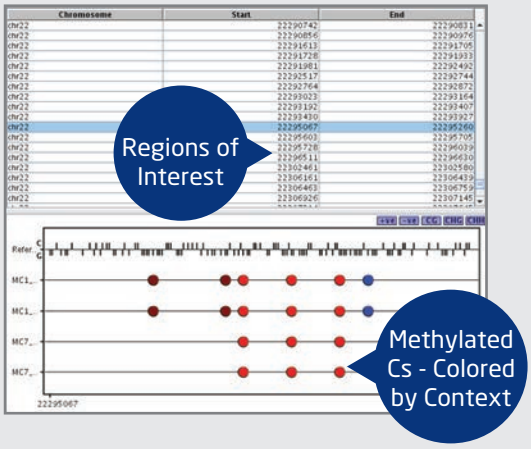
Methylation Effect Analysis

- Determine the Genomic context of the methylated or differentially methylated cytosines resulting from methylation detection and DMC analyses.
- Annotate these cytosines with the selected non-protein effects.



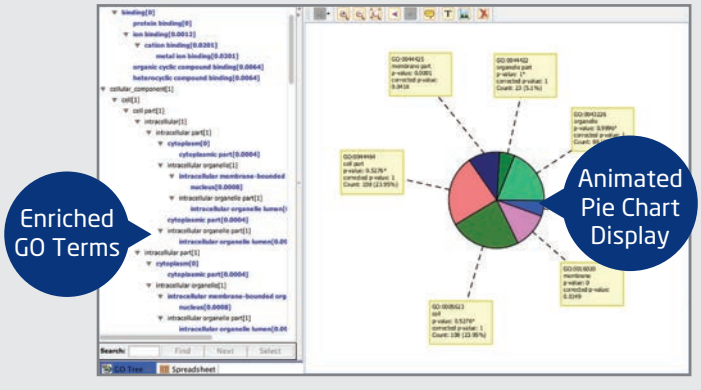
Methylation-Specific Views

- Lollipop Plot - Visualize methylated or differentially methylated cytosines by regions and samples.
- Interactive genome browser - Customized for displaying bisulfite converted reads.
- Within the GB, the methylation level histogram helps visualize the proportion of methylated cytosines compared to unmethylated cytosines in the read coverage.



GO Enrichment

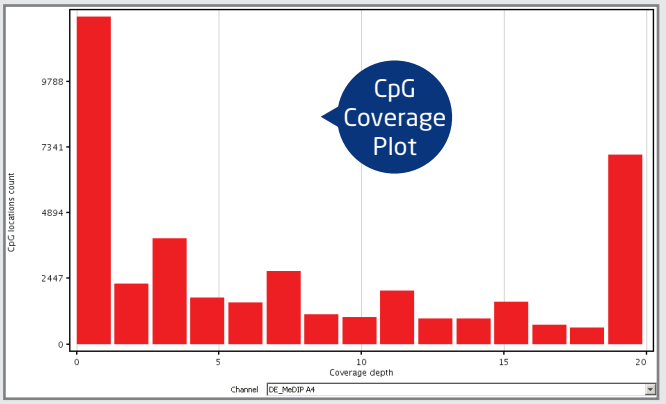
- Genes discovered to be affected by SNPs, SVs, peak regions, or any imported set of genes.
- GO Enrichment analysis on gene lists to detect enriched Gene Ontology terms.



strandngs
Streamlining NGS Data Management & Analysis

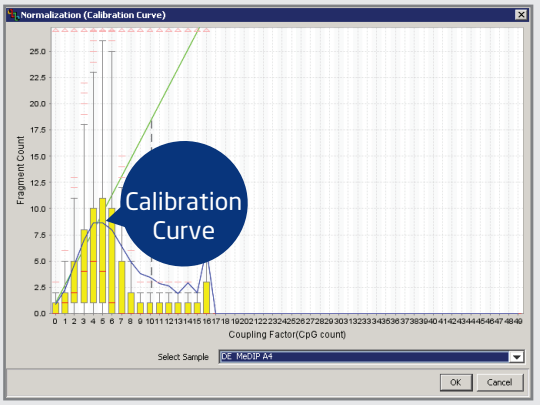
QC and CpG coverage analysis

- Quality metrics for pre- and post-alignment, target enrichment and library complexity.
- Compute CpG sites coverage for your samples.



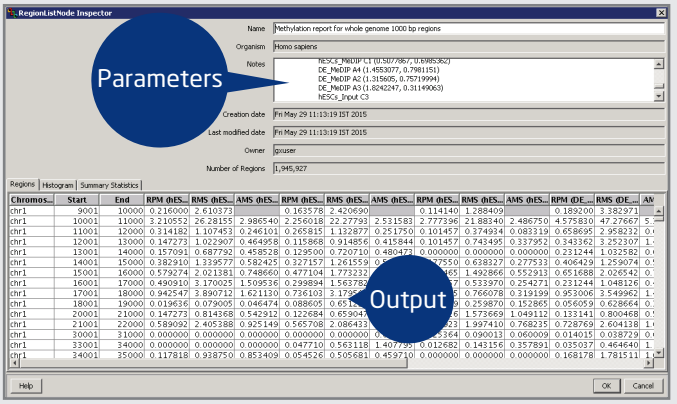
Normalization (Calibration curve)

- Estimate the linear dependency of methylation signals and CpG counts by calculating the slope and the intercept for every sample.



Methylation Detection

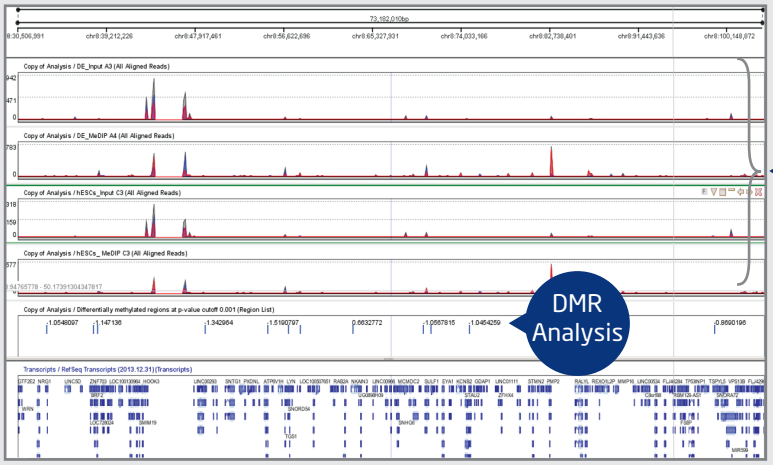
- Identifies methylated regions across the genome or within regions of interest.
- Computes for every region, the reads per million (RPM), relative methylation score (RMS) and absolute methylation score (AMS).



strandngs
Streamlining NGS Data Management & Analysis

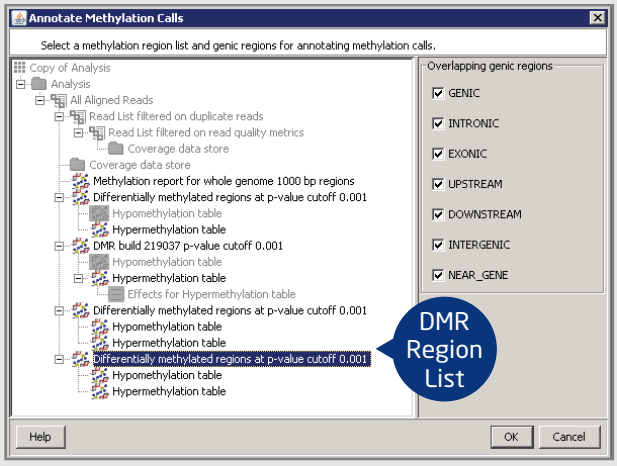
Differential methylation analysis (DMR)

- Identifies DMRs based on selected RMS ratio and p-value thresholds.
- Merges consecutive hypo and hyper-methylated regions if the new region still satisfies the specified threshold.
- Filters DMRs based on RPM, RMS or AMS and p-value thresholds.
- Creates separate lists for hypermethylated and hypomethylated regions from the identified DMRs.
- Creates associated gene lists based on an upstream / downstream padding (bp) selected for the genes.



Annotate methylation calls

- Determines the genomic context of the methylated or differentially methylated regions.
- Annotates the methylation calls for the selected genic region.



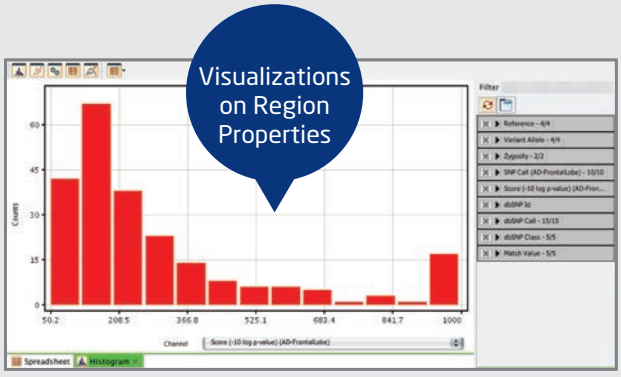
strandngs
Streamlining NGS Data Management & Analysis

Compare Region Lists

- Choose two region lists to compare.
- Find overlapping and non-overlapping region with appropriate distance criteria.

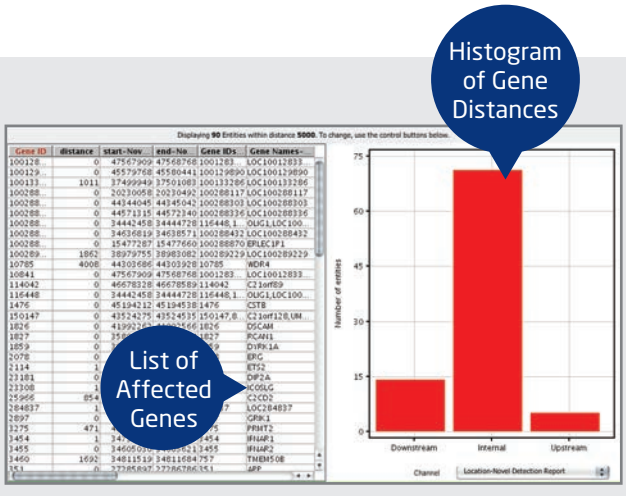
Region List Operations

- Visualize plots on various columns of a region list.
- Create new columns from existing columns using formulas.
- Perform filter operations on the columns to retain relevant regions.



Translate from regions to genes

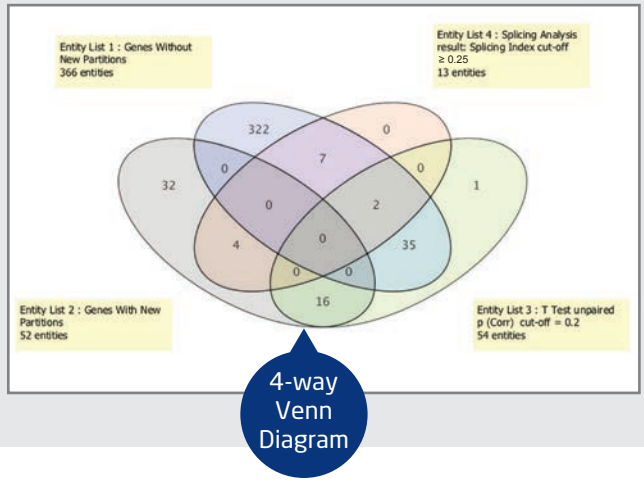
- For any region list find adjacent or overlapping genes.
- Find affected genes for the detected regions (eg, SNPs, SVs, peaks).



strandngs
Streamlining NGS Data Management & Analysis

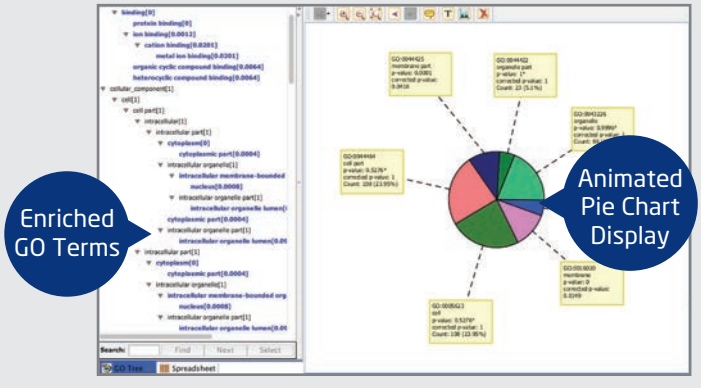
Compare Gene Lists

- Venn Diagrams to compare gene lists.
- Compare gene lists from different experiments and organisms.
- Ability to compare imported gene lists.
- Ability to save individual regions as new gene lists.



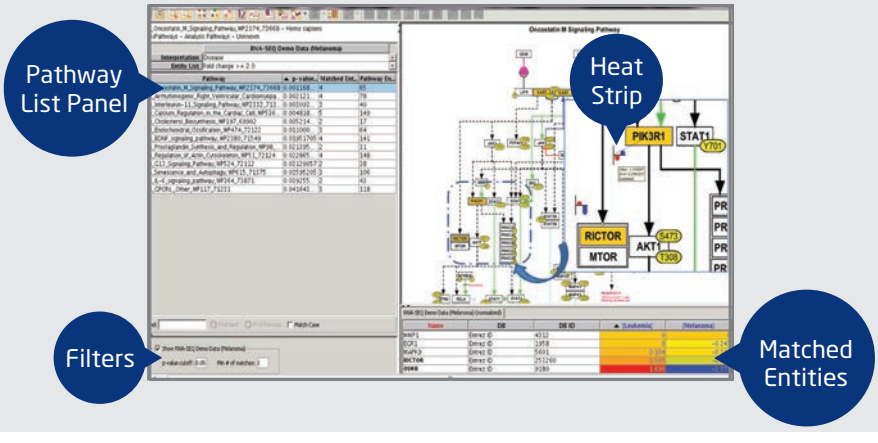
GO Enrichment

- Genes discovered to be affected by SNPs, SVs, peak regions, or any imported set of genes.
- GO Enrichment analysis on set of genes to detect enriched Gene Ontology terms.

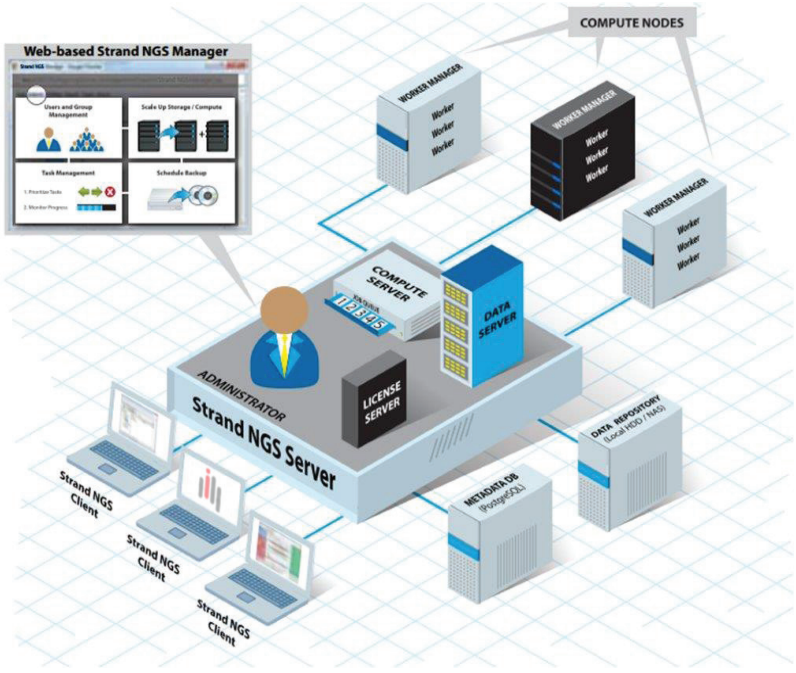


Pathway Analysis

- Single Experiment Analysis (SEA) identifies enriched pathways for the genes from a single experiment type.
- Multi-Omic Pathway Analysis (MOA) from multiple genomics and transcriptomics experiments.
- Overlay differentially expressed entities on curated pathways.
- Choose from curated pathways like WikiPathways, BioCyC pathways, BioPAX pathways or literature-derived networks like NLP and MeSH.
- Find significant pathways for differentially expressed genes.

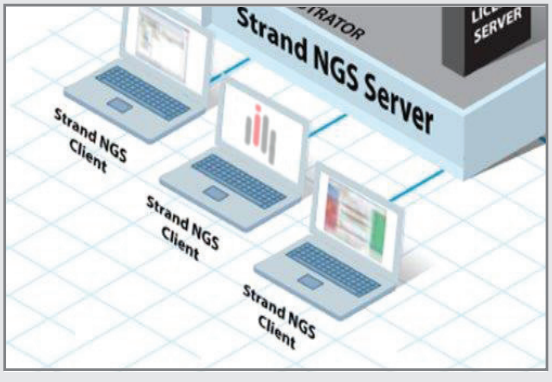


strandngs
Streamlining NGS Data Management & Analysis



Collaborative Analysis

- Allows different people working on the same project to easily collaborate with each other.
- Samples, analysis, and workflow pipelines can be shared.
- Objects can be shared with an individual user or a group.
- Sharing can be done with either Read-only or Read-Write permission.



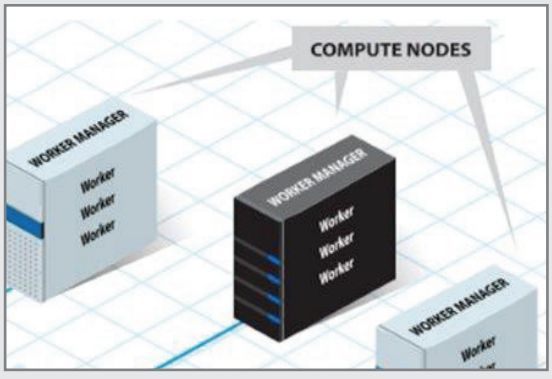
Centralized Storage

- Data and analysis results from multiple users stored in one place in a safe, backed-up manner.
- Data can be backed up regularly; possible to repeat backup weekly or monthly.
- Option to do the backup incrementally.
- Allows storing data in a new location, in case the current location is full, making incremental expansion of storage possible.

strandngs
Streamlining NGS Data Management & Analysis

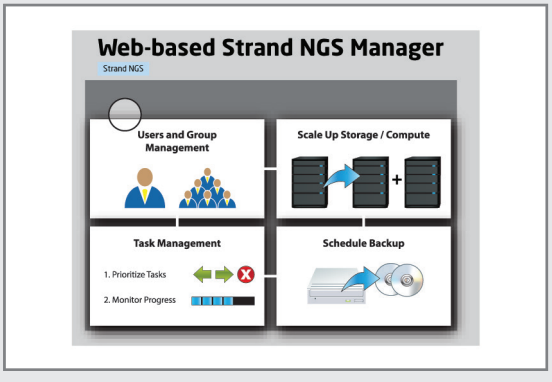
Scalable Compute

- Multiple compute nodes managed by the server to handle heavy computing.
- Compute can be scaled up by simply adding new machines to the network and configuring them as compute nodes.
- Higher reliability due to the distributed nature of the compute.
- Execution of tasks on a first-come, first-serve basis with a possibility for the administrator to reorder them if necessary.
- Status dashboard to show the task status and logs.



Web-based Interface for System Administration

- Allows creating and managing users and groups.
- Facilitates managing the compute resources.
- Allows the administrator to suspend, reschedule, reorder, or delete tasks.
- Group administrator sees only the tasks from his group members.
- System logs and usage can be monitored.



Easy and Flexible Deployment

- Major components of the Server Edition have separate installers and can be installed independently.
- Possible to have the Server, Compute, and the Storage on one big machine or on different machines depending on the available resources and the needs of the enterprise.
- Can have either node-locked licenses or shared floating licenses depending on the need.

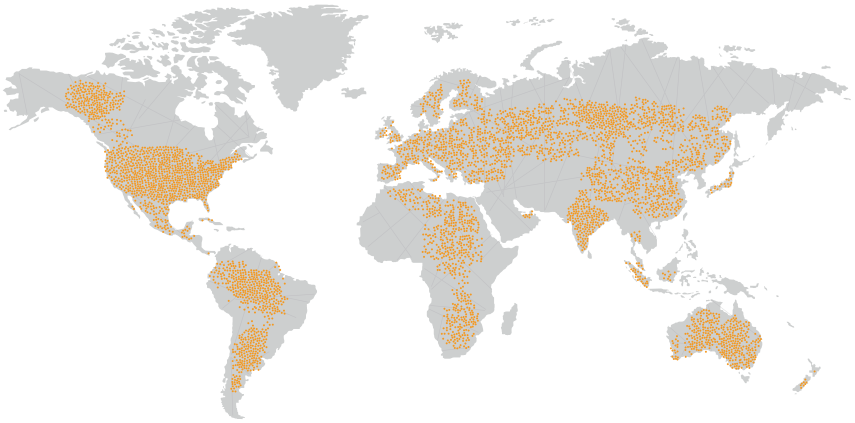
New Generation Healthcare

About Strand

A History of Innovative Genomic Research

Strand Life Sciences is a global genomic profiling company and leader in precision medicine diagnostics, aimed at empowering cancer care and genetic testing for inherited diseases. Strand works with physicians and hospitals to enable faster clinical decision support for accurate molecular diagnosis, prognosis, therapy recommendations, and clinical trials. The Strand Center for Genomics & Personalized Medicine is India's 1st and only CAP & NABL accredited NGS laboratory.

www.strandls.com



A Trusted Partner to Companies Worldwide

For 15 years, our genomics products and solutions have facilitated the work of leading researchers and medical geneticists in over 2,000 laboratories and 100 hospitals around the world.

Strand Life Sciences Pvt. Ltd
5th Floor, Kirloskar Business Park, Bellary Road, Hebbal, Bangalore 560024
Phone:+91-80-40 (787263) Fax: +91-80-4078-7299

Contact us
sales@strandngs.com
USA: 1-800-752-9122 ROW: +1-650-353-5060

support@strandngs.com
USA: 1-800-516-5181 ROW: +1-650-288-4559