



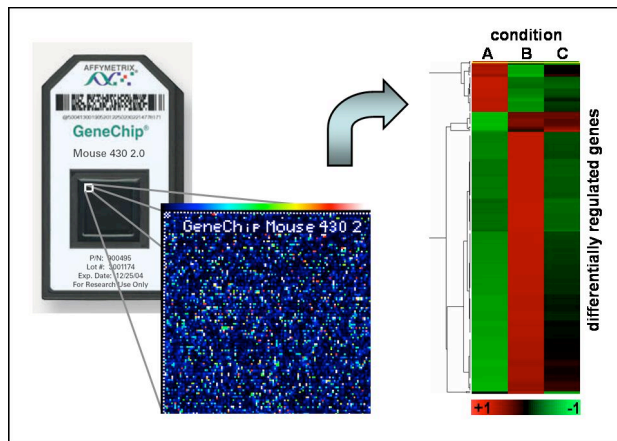
Performing Genomics Analyses on Your Own with AltAnalyze

Nathan Salomonis

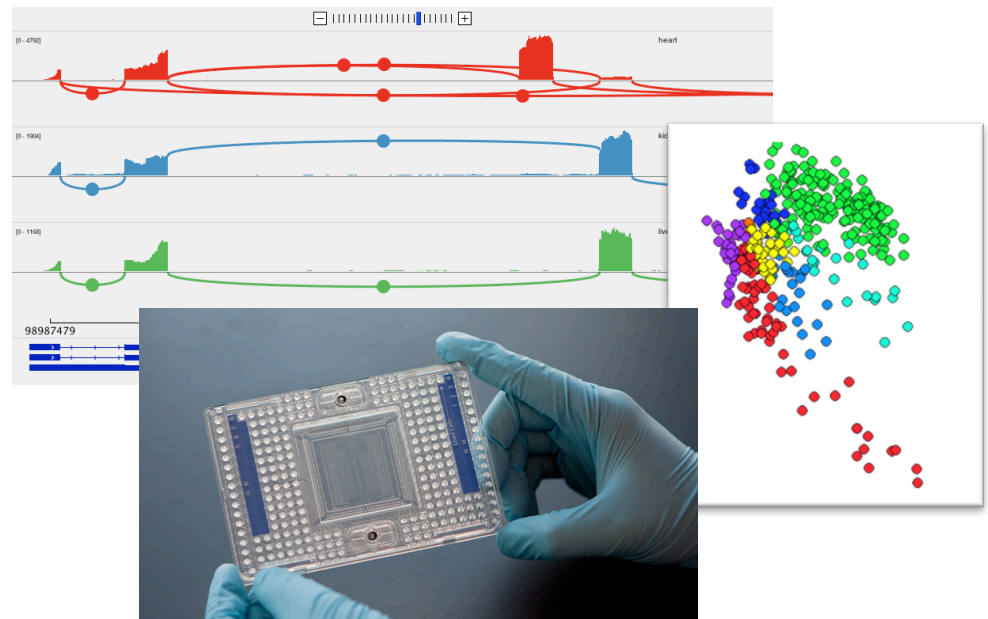
Department of Biomedical Informatics,
CCHMC

Frequently Used Genomics at CCHMC

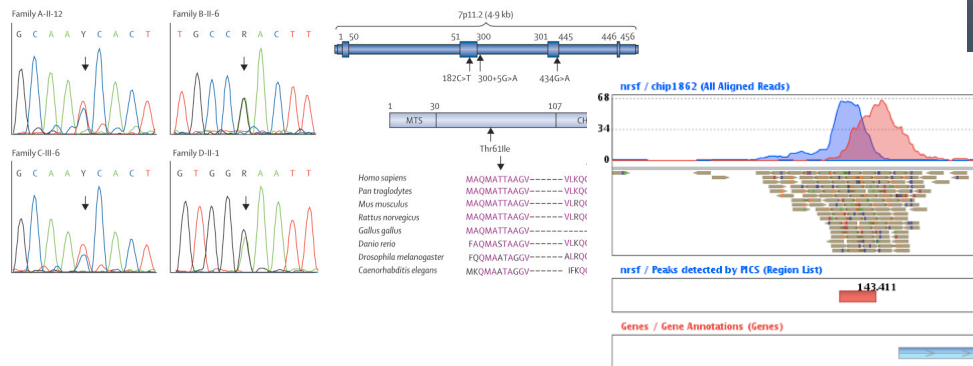
Conventional and Splicing Arrays



Bulk and Single-Cell RNA-Seq



DNA-Seq and ChIP-Seq



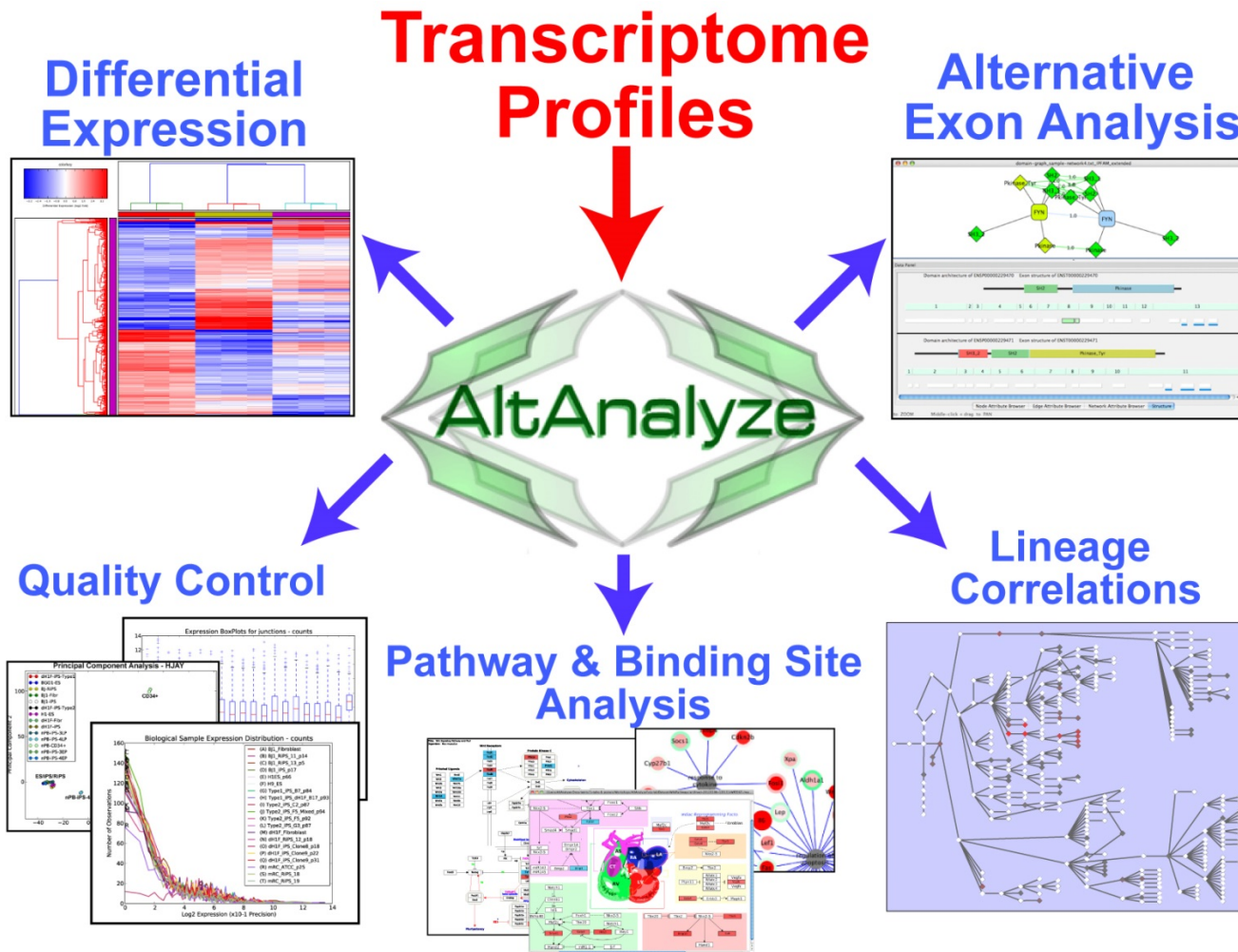
Available Easy-To-Use Tools Developed at CCHMC and UC

- AltAnalyze
 - RNA-Seq and microarray analysis
- BioWardrobe
 - ChIP-Seq and RNA-Seq co-analysis
- CASSI (CCHMC Cluster)
 - Workflow for variant analysis and annotation.
- CisBP/CisBP-RNA (in development)
 - Predict impact of variants on DNA or RNA binding
- NetWalker
 - Network and gene expression analysis
- Sincera
 - Single-cell RNA-Seq analysis

AltAnalyze Supported Input Data

- ***RNA-Seq***
 - FASTQ files (single or paired-end – Kallisto Algorithm)
 - BAM files (recommended – RPKM algorithm)
 - TCGA junction files
 - SOLID junction/exon expression files (TAB)
 - Pre-computed expression values
- ***Microarrays***
 - Affymetrix 3', splicing, microRNA (CEL files – RMA)
 - Agilent arrays (FE)
 - Pre-computed expression values
- ***Tabular expression values***
- **Not currently:**
 - Methylation array, ERRBS, CHIP-Seq, miR-Seq, mass spec, DNA-seq, SNP arrays.

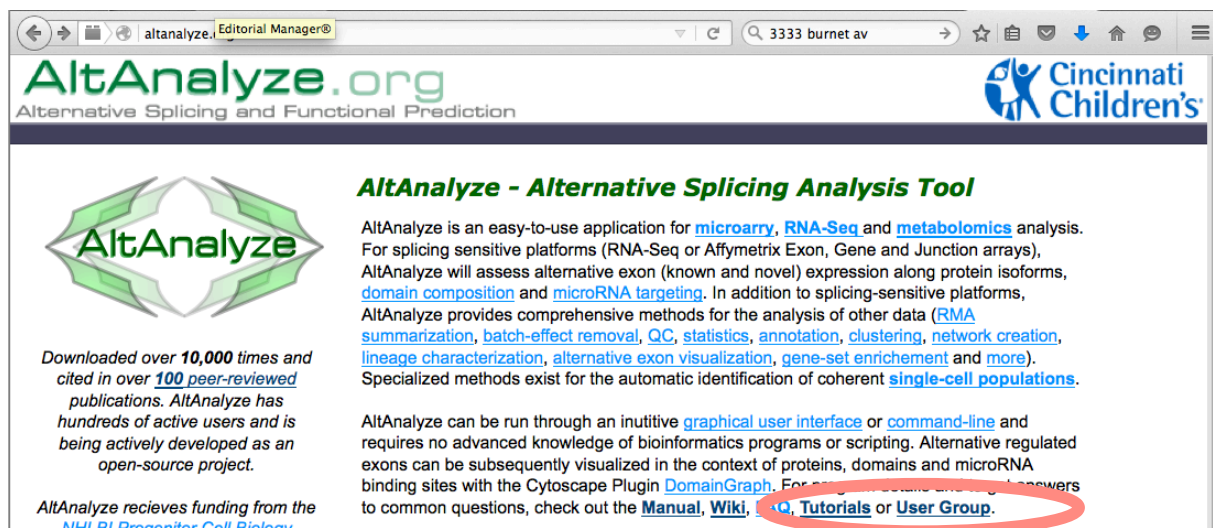
Automating Genomics Analysis in AltAnalyze



Automated and à la carte Options

- Gene Expression Analysis
 - Workflow
- Single-Cell Profiling
 - Workflow and à la carte
- Enrichment analyses
 - Workflow and stand-alone modules
- Network Analysis
 - Workflow and stand-alone modules
- Cell and Tissue-type Deconvolution
 - Workflow and stand-alone modules
- Alternative Exon Analysis
 - Workflow

Documentation



AltAnalyze.org
Alternative Splicing and Functional Prediction

AltAnalyze - Alternative Splicing Analysis Tool

AltAnalyze is an easy-to-use application for [microarray](#), [RNA-Seq](#) and [metabolomics](#) analysis. For splicing sensitive platforms (RNA-Seq or Affymetrix Exon, Gene and Junction arrays), AltAnalyze will assess alternative exon (known and novel) expression along protein isoforms, [domain composition](#) and [microRNA targeting](#). In addition to splicing-sensitive platforms, AltAnalyze provides comprehensive methods for the analysis of other data ([RMA summarization](#), [batch-effect removal](#), [QC](#), [statistics](#), [annotation](#), [clustering](#), [network creation](#), [lineage characterization](#), [alternative exon visualization](#), [gene-set enrichment](#) and [more](#)). Specialized methods exist for the automatic identification of coherent [single-cell populations](#).

AltAnalyze can be run through an intuitive [graphical user interface](#) or [command-line](#) and requires no advanced knowledge of bioinformatics programs or scripting. Alternative regulated exons can be subsequently visualized in the context of proteins, domains and microRNA binding sites with the Cytoscape Plugin [DomainGraph](#). For [common questions](#), check out the [Manual](#), [Wiki](#), [FAQ](#), [Tutorials](#) or [User Group](#).

Downloaded over **10,000** times and cited in over **100** peer-reviewed publications. AltAnalyze has hundreds of active users and is being actively developed as an open-source project.

AltAnalyze receives funding from the [NHHLBI Progenitor Cell Biology Consortium](#) (U01HL099997).

Downloads

[Development version 2.0.9 - 9/28/15](#)

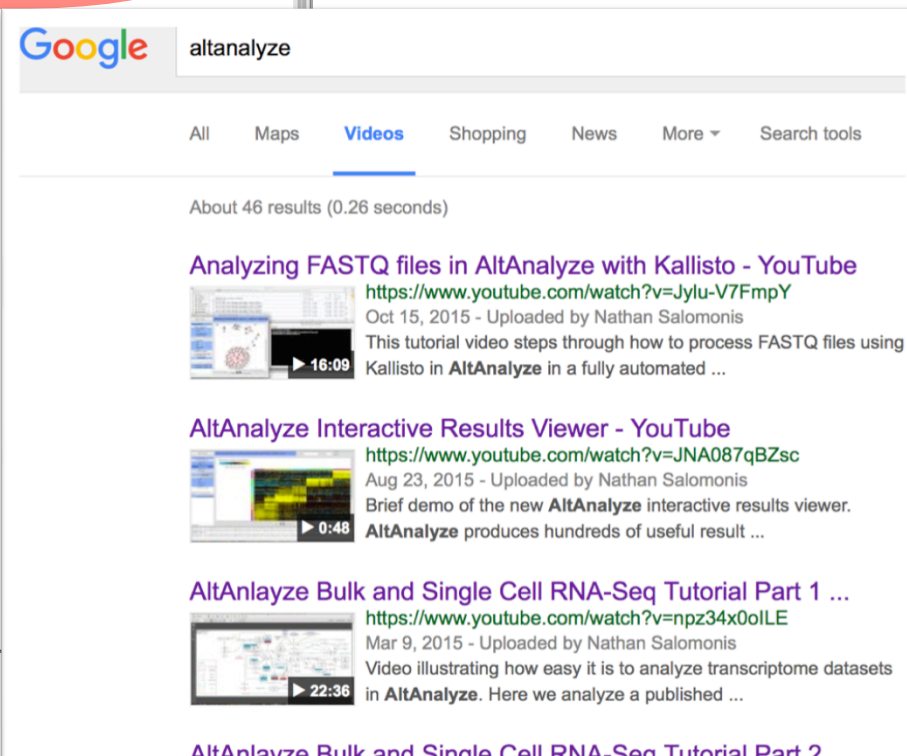
[Release version 2.0.9.3 - 9/28/15](#)

- [Mac OSX Installer \(.dmg\)](#)
- [Windows Archive \(.zip\) 32bit 64bit](#)
- [Ubuntu Archives \(.zip\) 32bit 64bit](#)
- [Cross Platform \(.zip\)](#)

[archived versions](#)

Documentation

- [Manual and Installation Guide](#)
- [Tutorials](#)
- [News and Updates](#)
- [How to Cite](#)
- [Contact Us](#)
- [AboutUs](#)
- [Blog](#)
- [FAQ](#)



Google altanalyze

All Maps **Videos** Shopping News More Search tools

About 46 results (0.26 seconds)

Analyzing FASTQ files in AltAnalyze with Kallisto - YouTube
<https://www.youtube.com/watch?v=Jylu-V7FmpY>
Oct 15, 2015 - Uploaded by Nathan Salomonis
This tutorial video steps through how to process FASTQ files using Kallisto in AltAnalyze in a fully automated ...

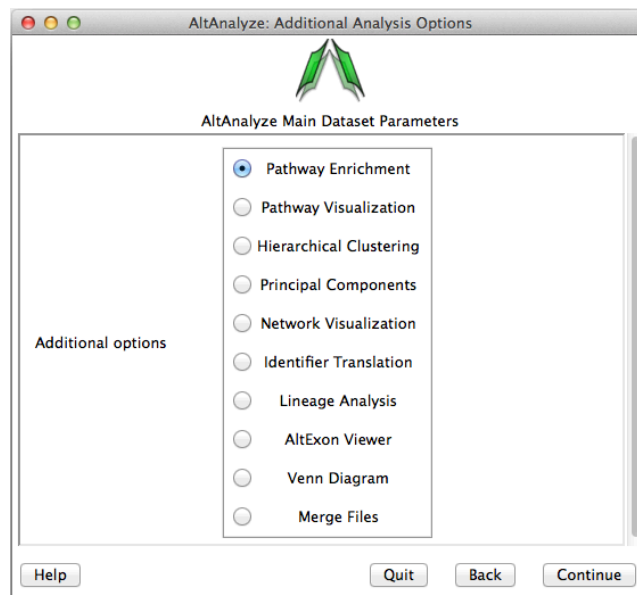
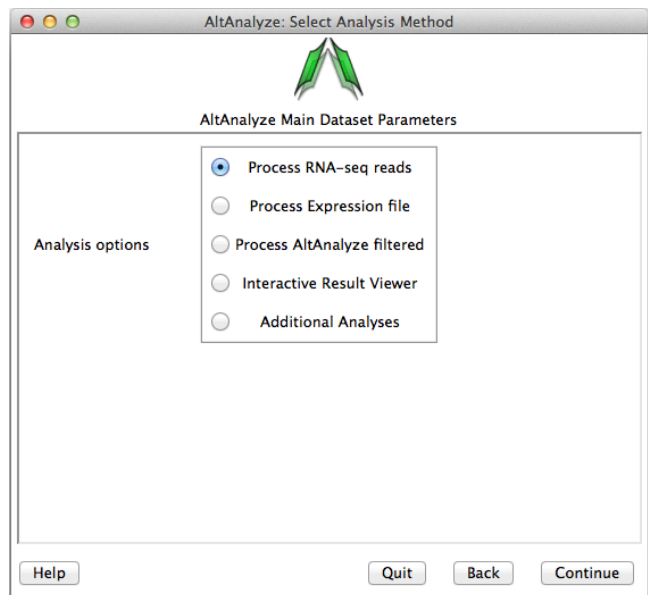
AltAnalyze Interactive Results Viewer - YouTube
<https://www.youtube.com/watch?v=JNA087qBZsc>
Aug 23, 2015 - Uploaded by Nathan Salomonis
Brief demo of the new AltAnalyze interactive results viewer. AltAnalyze produces hundreds of useful result ...

AltAnalyze Bulk and Single Cell RNA-Seq Tutorial Part 1 ...
<https://www.youtube.com/watch?v=npz34x0o1LE>
Mar 9, 2015 - Uploaded by Nathan Salomonis
Video illustrating how easy it is to analyze transcriptome datasets in AltAnalyze. Here we analyze a published ...

AltAnalyze Bulk and Single Cell RNA-Seq Tutorial Part 2

Three Ways to Run AltAnalyze

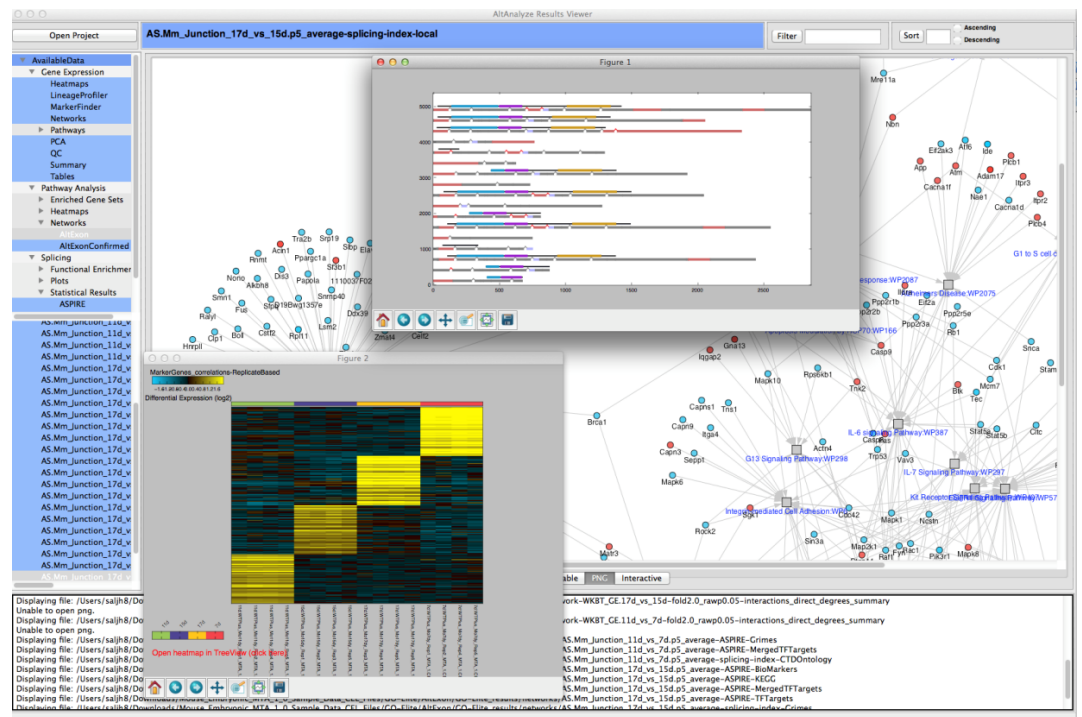
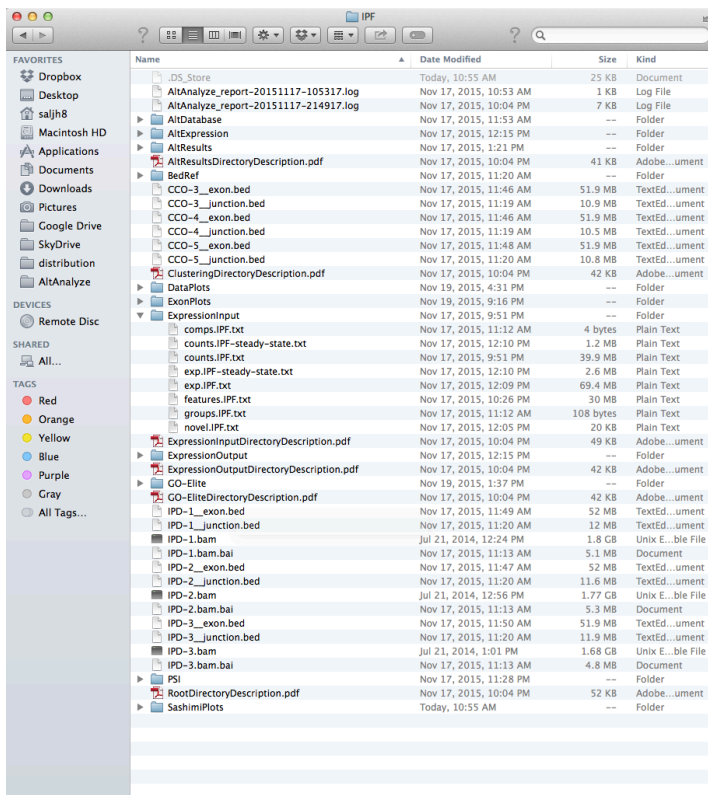
- Easy-to-use workflows (GUI)
- À la carte option (GUI)
- Command-line (all of the above)



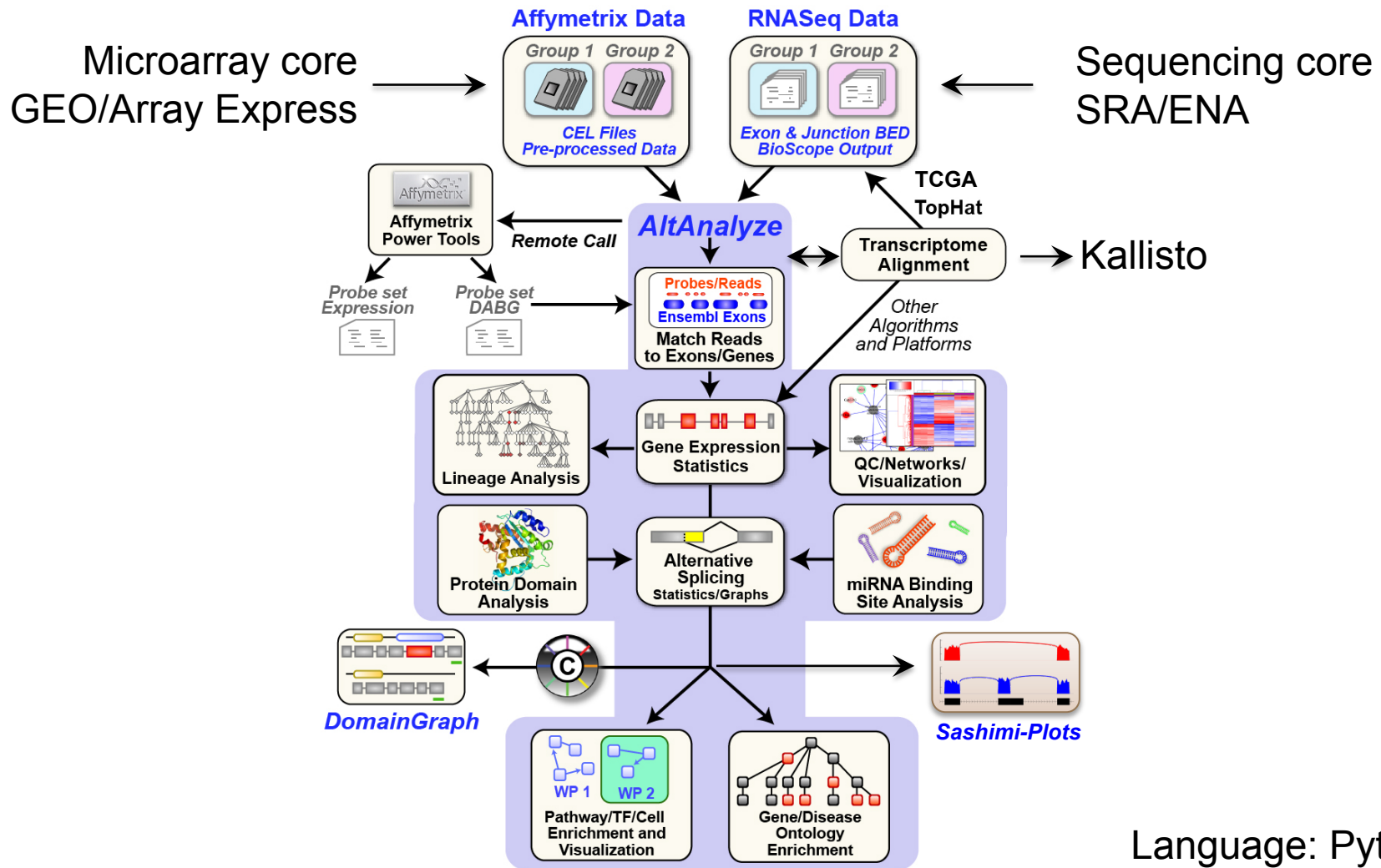
```
saljh8 ~ — bash — 39x24
P13-8438:~ saljh8$ python AltAnalyze.py
--runMarkerFinder --platform gene --ex
pdir /projects/ExpressionInput/exp.meta
-steady-state.txt --species Hs
```


Ways to View the Results

- Navigate directories on your computer
- AltAnalyze Results Viewer

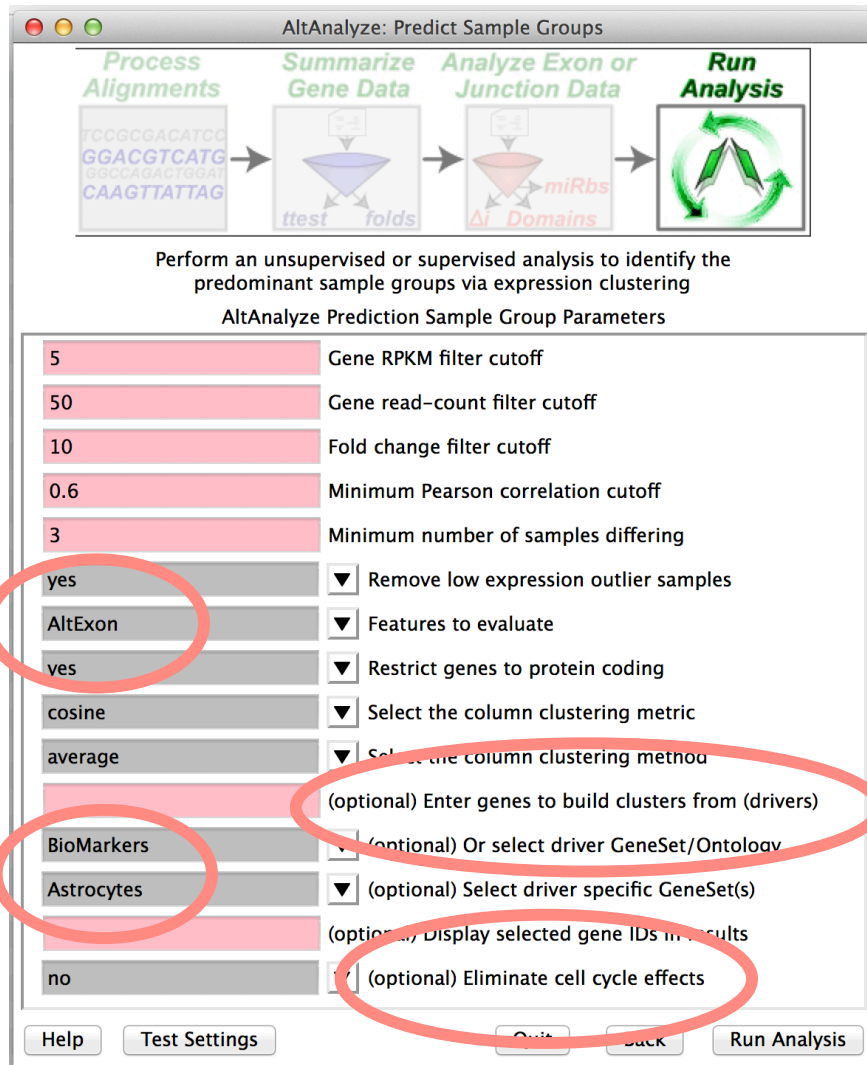


Attempting to Automate Genomics Analysis



Language: Python
Optional calls to R

Single-Cell RNA-Seq and Unsupervised Expression Analysis



AltAnalyze: Predict Sample Groups

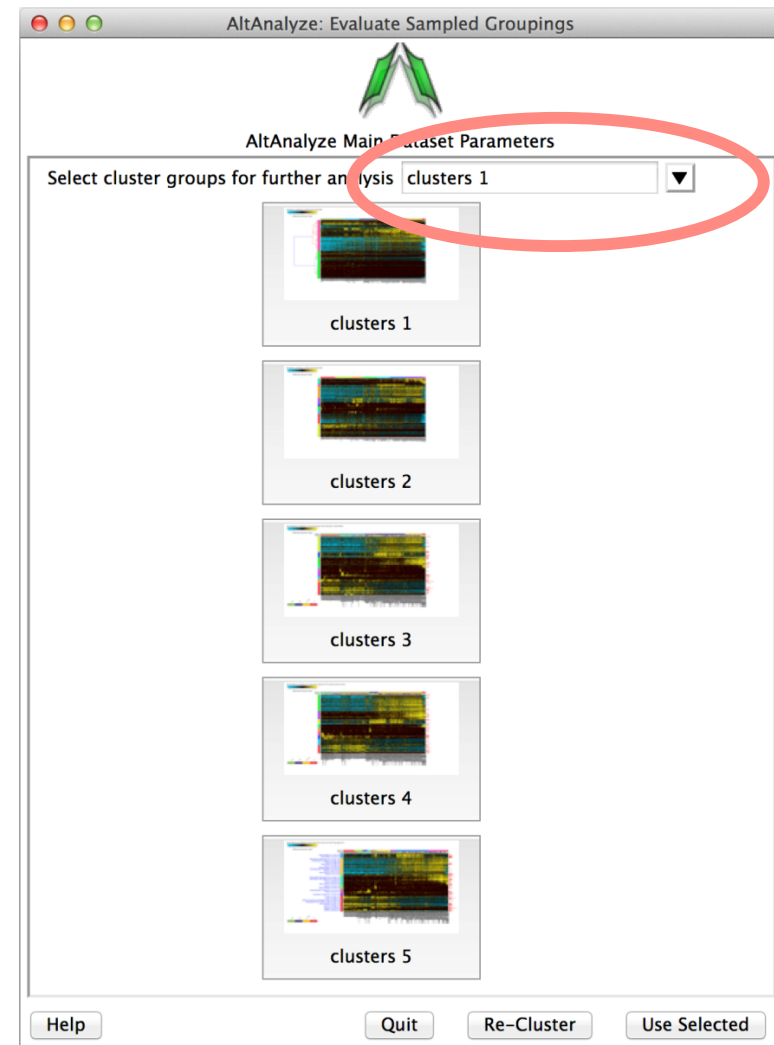
Process Alignments → Summarize Gene Data → Analyze Exon or Junction Data → Run Analysis

Perform an unsupervised or supervised analysis to identify the predominant sample groups via expression clustering

AltAnalyze Prediction Sample Group Parameters

| | |
|------------|---|
| 5 | Gene RPKM filter cutoff |
| 50 | Gene read-count filter cutoff |
| 10 | Fold change filter cutoff |
| 0.6 | Minimum Pearson correlation cutoff |
| 3 | Minimum number of samples differing |
| yes | <input type="checkbox"/> Remove low expression outlier samples |
| AltExon | <input type="checkbox"/> Features to evaluate |
| yes | <input type="checkbox"/> Restrict genes to protein coding |
| cosine | <input type="checkbox"/> Select the column clustering metric |
| average | <input type="checkbox"/> Select the column clustering method |
| | (optional) Enter genes to build clusters from (drivers) |
| BioMarkers | <input type="checkbox"/> (optional) Or select driver GeneSet/Ontology |
| Astrocytes | <input type="checkbox"/> (optional) Select driver specific GeneSet(s) |
| | (optional) Display selected gene IDs in results |
| no | <input type="checkbox"/> (optional) Eliminate cell cycle effects |

Buttons: Help, Test Settings, Quit, Back, Run Analysis



AltAnalyze: Evaluate Sampled Groupings

AltAnalyze Main Dataset Parameters

Select cluster groups for further analysis: clusters 1

clusters 1

clusters 2

clusters 3

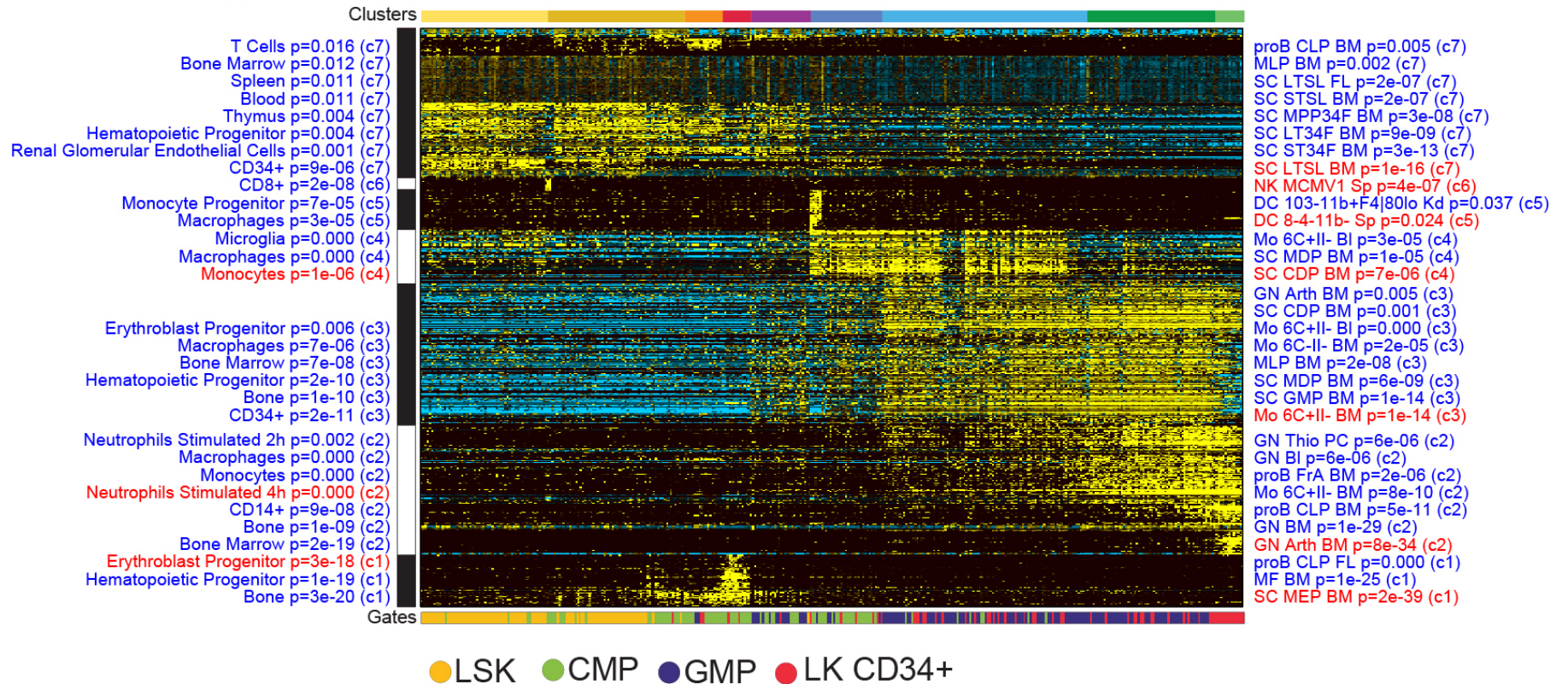
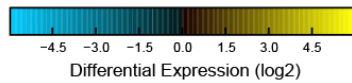
clusters 4

clusters 5

Buttons: Help, Quit, Re-Cluster, Use Selected

<http://altanalyze.org>

Cell-Type Prediction Analysis



Automated Downstream Expression and Splicing Analyses

Python File Edit View Search Table Help Alt 9 100% Sun 9:17 PM Salomonis, Nathan

AltAnalyze Results Viewer

Open Project Clustering-MarkerGenes_correlations-ReplicateBased-hierarchical_cosine_cosine Filter Sort Ascending Descending

- AvailableData
 - Gene Expression
 - Heatmaps
 - ICGS
 - LineageProfiler
 - MarkerFinder
 - Networks
 - Pathways
 - Colored WikiPathway regulated
 - Enrichment Scores
 - UseDir
 - Clustering-MarkerGene
 - Clustering-MarkerGene

Mo 6C+II- BM p=3e-10
proB CLP BM p=5e-13
GN BM p=1e-23
GN Arth BM p=3e-28

Mo 6C+II- BM p=6e-06
GN Arth BM p=5e-06
SC GMP BM p=5e-09

SC GMP BM p=1e-06
Mo 6C+II- BM p=1e-06
SC MDP BM p=3e-09

proB FrA BM p=0.043
DC 4+ Sp p=0.002
SC CDP BM p=0.001

SC STSL BM p=0.002
SC LTSL BM p=0.002
preT DN2A Th p=0.001
MF BM p=8e-05

SC CMP BM p=0.014
proB CLP FL p=0.001
MF BM p=3e-26

SC MEP BM p=7e-42
SC STSL BM p=0.002
SC LT34F BM p=0.000
SC LTSL FL p=4e-09
SC LTSL BM p=1e-14

SC STSL BM p=1e-06
SC MPP34F BM p=4e-07
SC ST34F BM p=7e-11

proB FrA BM p=4e-07
SC MPP34F BM p=6e-08
proB CLP BM p=6e-08

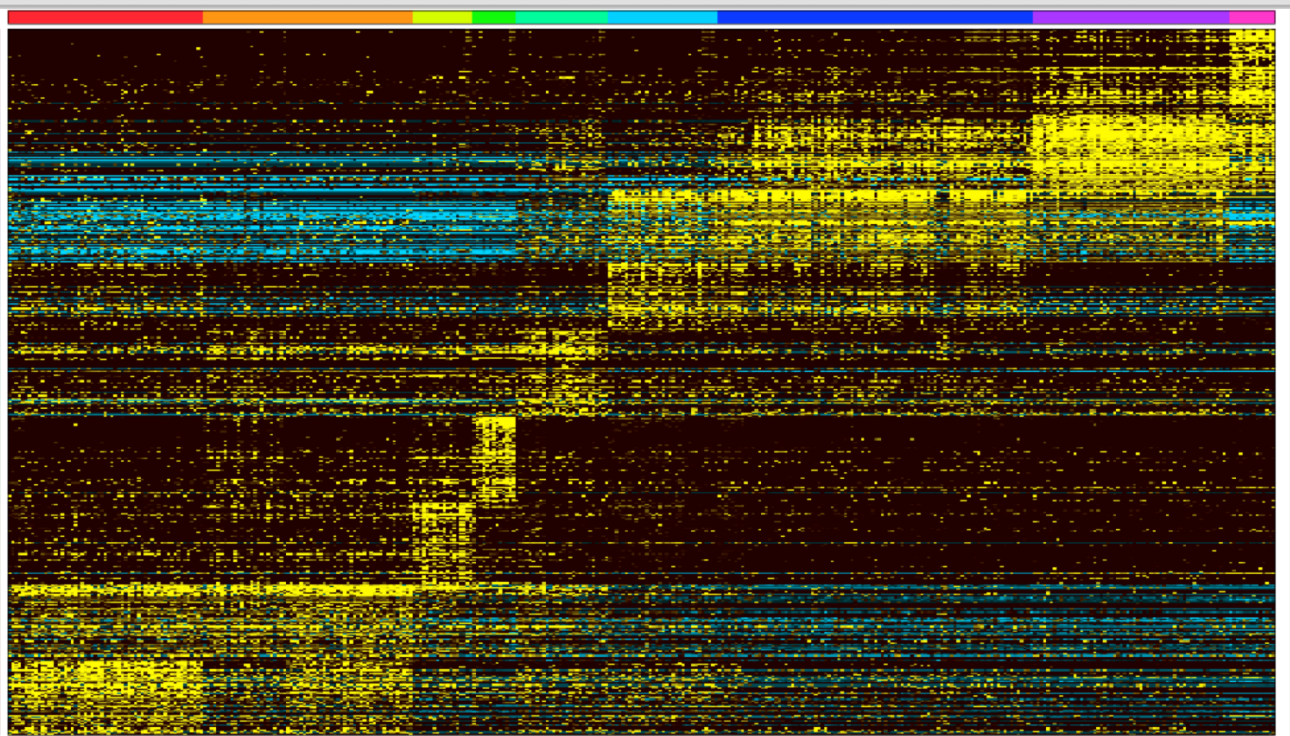


Table PNG Interactive

Unable to open png.
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/networks/AltAnalyze-network-WKBT_GE.9_vs_1-fold2.0_adjp0.05-interactions_direct_degrees_summary
Unable to open png.
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/networks/AltAnalyze-network-WKBT_GE.8_vs_1-fold2.0_adjp0.05-interactions_direct
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/DataPlots/MarkerFinder/Clustering-MarkerGenes_correlations-ReplicateBased-hierarchical_cosine_cosine.txt



Automated Downstream Expression and Splicing Analyses

Python File Edit View Search Table Help Alt 9 100% Sun 9:17 PM Salomonis, Nathan

AltAnalyze Results Viewer

Open Project WP571_FAS pathway and Stress induction of HSP regulation-GE.4_vs_1-fold2.0_adj0.05 Filter Sort Ascending Descending

AvailableData

- Gene Expression
 - Heatmaps
 - ICGS
 - LineageProfiler
 - MarkerFinder
 - Networks
- Pathways
 - Colored WikiPathway regulated
 - Enrichment Scores
 - PCA
 - QC
 - Summary Tables
 - Pathway Analysis
 - WP413_G1 to S cell c
 - WP434_TCA Cycle-G
 - WP435_One Carbon M
 - WP456_GPCRs, Class
 - WP456_GPCRs, Class
 - WP493_MAPK signalin
 - WP519_Proteasome C
 - WP519_Proteasome C
 - WP519_Proteasome C
 - WP544_Exercise-indu
 - WP567_Eukaryotic Tr
 - WP567_Eukaryotic Tr
 - WP571_FAS pathway
 - WP63_Pentose Phosp
 - WP63_Pentose Phosp

Title: FAS pathway and Stress inductio
Availability: CC BY 2.0
Last modified: 10/17/2013
Organism: Mus musculus

Table PNG Interactive

Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/WikiPathways/WP730_Glutathione and one carbon metabolism-GE.7_vs_1-fold2.0_adj0.05
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/WikiPathways/WP85_Focal Adhesion-GE.3_vs_1-fold2.0_adj0.05
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/WikiPathways/WP87_Nucleotide Metabolism-GE.6_vs_1-fold2.0_adj0.05
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/WikiPathways/WP87_Nucleotide Metabolism-GE.7_vs_1-fold2.0_adj0.05
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/WikiPathways/WP571_FAS pathway and Stress induction of HSP regulation-GE.4_vs_1-fold2.0_adj0.05



Automated and à la carte Options

- Gene Expression Analysis
 - Workflow
- Single-Cell Profiling
 - Workflow and à la carte
- **Enrichment analyses**
 - Workflow and stand-alone modules
- **Network Analysis**
 - Workflow and stand-alone modules
- **Cell and Tissue-type Deconvolution**
 - Workflow and stand-alone modules

Gene Set Enrichment Analysis GO-Elite

AltAnalyze Results Viewer

pruned-results_z-score_elite

Filter: Sort: Ascending Descending

| | C | D | E | F | G | H | I | J | K | L | M | |
|------|----------------------------------|----------------|-----------------|--------------------|-----------------|-----------------|-------------------|-------------------|-------------------|----------------------|-------------------|------------------------------|
| 3010 | | | | | | | | | | | | |
| 3011 | Ontology Name(Ontology-ID) | Number Changed | Number Measured | Number in Ontology | Percent Changed | Percent Present | Z Score | FisherExactP | AdjustedP | redundant with terms | inverse redundant | gene symb |
| 3012 | immune system phenotype(MP:129 | 2426 | 2451 | 5.31739488871 | 98.9800081599 | 6.21810756282 | 1.81346825678e-09 | 5.44040477034e-08 | | | | Abcg1 Ahn |
| 3013 | hematopoietic system phenoty | 107 | 2020 | 5.29702970297 | 98.8258317025 | 5.37208765622 | 2.17541733495e-07 | 3.26312600243e-06 | | | | Abcg1 Ahn |
| 3014 | tumorigenesis(MP:0002006) | 46 | 744 | 6.18279569892 | 99.4652406417 | 4.300933688 | 7.06965435356e-05 | 0.000706965435356 | | | | Atp2c1 Aui |
| 3015 | cellular phenotype(MP:000538 | 126 | 2824 | 4.46175637394 | 98.9488437281 | 3.79036813943 | 0.000212331642082 | 0.00159248731561 | | | | Abcg1 Ahn |
| 3016 | liver/biliary system phenotype | 43 | 920 | 4.67391304348 | 99.3520518359 | 2.16274068467 | 0.0411045387446 | 0.205522693723 | | | | Abcg1 Ahp |
| 3017 | | | | | | | | | | | | |
| 3018 | Gene-Set Name | Number Changed | Number Measured | Number in Gene-Set | Percent Changed | Percent Present | Z Score | FisherExactP | AdjustedP | redundant with terms | inverse redundant | gene symb |
| 3019 | Sfp1 | 6 | 23 | 30 | 26.0869565217 | 76.6666666667 | 5.95374162798 | 9.80994131878e-05 | 0.0258001456684 | | | Cd72 Csf1r |
| 3020 | Sp1 | 7 | 61 | 100 | 11.4754098361 | 61.0 | 3.44601513941 | 0.00477631378341 | 0.438503803448 | | | Atp1a3 Ba |
| 3021 | Fos | 14 | 180 | 186 | 7.77777777778 | 96.7741935484 | 3.21916637056 | 0.005293954834 | 0.438503803448 | | | Atp5b Ccl2 |
| 3022 | Nanog | 16 | 245 | 251 | 6.5306122449 | 97.609561753 | 2.68553166384 | 0.0120269195651 | 0.59886947468 | | | B2m Calcr |
| 3023 | Stat1 | 17 | 287 | 2564 | 5.92334494774 | 11.1934477379 | 2.34038167315 | 0.0295815078249 | 0.59886947468 | | | Arhgap19l |
| 3024 | | | | | | | | | | | | |
| 3025 | Gene-Set Name | Number Changed | Number Measured | Number in Gene-Set | Percent Changed | Percent Present | Z Score | FisherExactP | AdjustedP | redundant with terms | inverse redundant | gene symb |
| 3026 | MLP BM | 12 | 32 | 33 | 37.5 | 96.9696969697 | 18.1029241051 | 1.44241289906e-15 | 3.2165807649e-13 | | | MF BM |
| 3027 | SC CMP BM | 14 | 58 | 60 | 24.1379310345 | 96.6666666667 | 15.4177692325 | 5.61316106222e-15 | 6.25867458438e-13 | | | Alas1 Arl2 |
| 3028 | SC MDP BM | 9 | 43 | 48 | 20.9302325581 | 89.5833333333 | 11.3917688134 | 2.37855609191e-09 | 1.76806002832e-07 | | | Cts Elane |
| 3029 | SC MPP34F BM | 8 | 43 | 46 | 18.6046511628 | 93.4782608696 | 10.0418621115 | 5.32035064057e-08 | 2.96609548212e-06 | | | 9030619PC |
| 3030 | Mo 6C+II- BM | 8 | 52 | 60 | 15.3846153846 | 86.6666666667 | 8.99618848843 | 2.51956621247e-07 | 1.12372653076e-05 | | | Cts Elane |
| 3031 | proB CLP BM | 6 | 42 | 47 | 14.2857142857 | 89.3617021277 | 7.44595838401 | 1.39882198116e-05 | 0.000445624716855 | | | proB FrA BM |
| 3032 | Mo 6C+II- BM | 7 | 57 | 60 | 12.2807107544 | 95.0 | 7.34636366596 | 7.19232176419e-06 | 0.000267314625569 | | | 9030619PC |
| 3033 | proB FrA BM | 4 | 25 | 27 | 16.0 | 92.5925925926 | 6.4914378938 | 0.000270412491113 | 0.00603019855182 | | | Cts Elane |
| 3034 | SC ST34F BM | 6 | 56 | 60 | 10.7142857143 | 93.3333333333 | 6.24176744717 | 7.51291318923e-05 | 0.00186153293466 | | | Dntt Gms1 |
| 3035 | MLP FL | 6 | 56 | 60 | 10.7142857143 | 93.3333333333 | 6.24176744717 | 7.51291318923e-05 | 0.00186153293466 | | | SC LT34F BM SC LTSL BM SC ST |
| 3036 | SC CDP BM | 4 | 34 | 37 | 11.7647058824 | 91.8918918919 | 5.39365513776 | 0.000908076898402 | 0.0168750956953 | | | Angpt1 Elt |
| 3037 | Mo 6C+II- BI | 5 | 51 | 60 | 9.80392156863 | 85.0 | 5.37739349472 | 0.000474728937525 | 0.00962405027892 | | | Cts Ddx4 |
| 3038 | SC STSL FL | 4 | 54 | 60 | 7.40740740741 | 90.0 | 3.97434597625 | 0.0050993859825 | 0.0874740826229 | | | Cts F13a1 |
| 3039 | MF BM | 4 | 57 | 60 | 7.01754385965 | 95.0 | 3.82363560199 | 0.00618706756737 | 0.0985511476803 | | | F13a1 Hpf1 |
| 3040 | preT ETP Th | 3 | 35 | 41 | 8.57142857143 | 85.3658536585 | 3.8021520768 | 0.0102833271622 | 0.152878797145 | | | Ddx4 Hif1 |
| 3041 | proB CLP FL | 3 | 46 | 53 | 6.52173913043 | 86.7924528302 | 3.13308576364 | 0.0215585280506 | 0.300471984705 | | | Car1 Cts |
| 3042 | proB FrA FL | 3 | 52 | 60 | 5.76923076923 | 86.6666666667 | 2.8525077032 | 0.0297109840957 | 0.364426505317 | | | Hmga2 Mn |
| 3043 | SC LTSL FL | 3 | 54 | 60 | 5.55555555556 | 90.0 | 2.76831797164 | 0.0327433637422 | 0.364426505317 | | | Hmga2 Hif1 |
| 3044 | SC LTSL BM | 3 | 54 | 60 | 5.55555555556 | 90.0 | 2.76831797164 | 0.0327433637422 | 0.364426505317 | | | Hif Hmga2 |
| 3045 | SC LT34F BM | 3 | 54 | 60 | 5.55555555556 | 90.0 | 2.76831797164 | 0.0327433637422 | 0.364426505317 | | | SC ST34F BM SC STSL BM |
| 3046 | SC STSL BM | 3 | 55 | 60 | 5.45454545455 | 91.6666666667 | 2.7277384676 | 0.0343181910837 | 0.364426505317 | | | SC ST34F BM |
| 3047 | | | | | | | | | | | | SC LTSL BM SC ST34F BM |
| 3048 | Gene-Set Name | Number Changed | Number Measured | Number in Gene-Set | Percent Changed | Percent Present | Z Score | FisherExactP | AdjustedP | redundant with terms | inverse redundant | gene symb |
| 3049 | Signal regulatory protein (SIRP) | 3 | 5 | 6 | 60.0 | 83.3333333333 | 6.14031433317 | 0.000735628965195 | 0.341769542692 | | | Hif Meis1 |
| 3050 | Aurora B signaling | 8 | 38 | 38 | 21.0526315789 | 100.0 | 5.10972078828 | 0.000163101233451 | 0.128197569492 | | | Hif Hmga2 |
| 3051 | N-glycan trimming in the ER ar | 3 | 8 | 8 | 37.5 | 100.0 | 4.6309043459 | 0.00374129679196 | 0.53466532336 | | | Aurkb Birc1 |
| 3052 | Immune System | 30 | 341 | 345 | 8.79765395894 | 98.8405797101 | 4.25676295785 | 0.00012732980017 | 0.128197569492 | | | Calr Mogs |
| 3053 | DNA strand elongation | 6 | 30 | 30 | 20.0 | 100.0 | 4.25092488794 | 0.00146634512698 | 0.384182423269 | | | Adaptive Immune System Antig |
| 3054 | Asparagine N-linked glycosylat | 8 | 48 | 48 | 16.6666666667 | 100.0 | 4.24388973735 | 0.00086964260227 | 0.341769542692 | | | B2m Calr C |

Table PNG Interactive

Displaying file: /Users/saljh8/Desktop/dataAnalysis/Tab/IPF/DataPlots/WP2062_1155useRateMap-Control
 Working...
 Opening project at: /Users/saljh8/Desktop/demo/BoneMarrow
 Resetting grid...
 Currently displaying: SUMMARY
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulatd/GO-Elite_results/overlapping-results_z-score_elite.txt
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulatd/GO-Elite_results/pruned-results_z-score_elite
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulatd/GO-Elite_results/overlapping-results_z-score_elite.txt
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulatd/GO-Elite_results/pruned-results_z-score_elite
 Table Length: 134
 Table Length: 153
 Table Length: 5100
 Table Length: 5100

Gene Set Enrichment Analysis GO-Elite

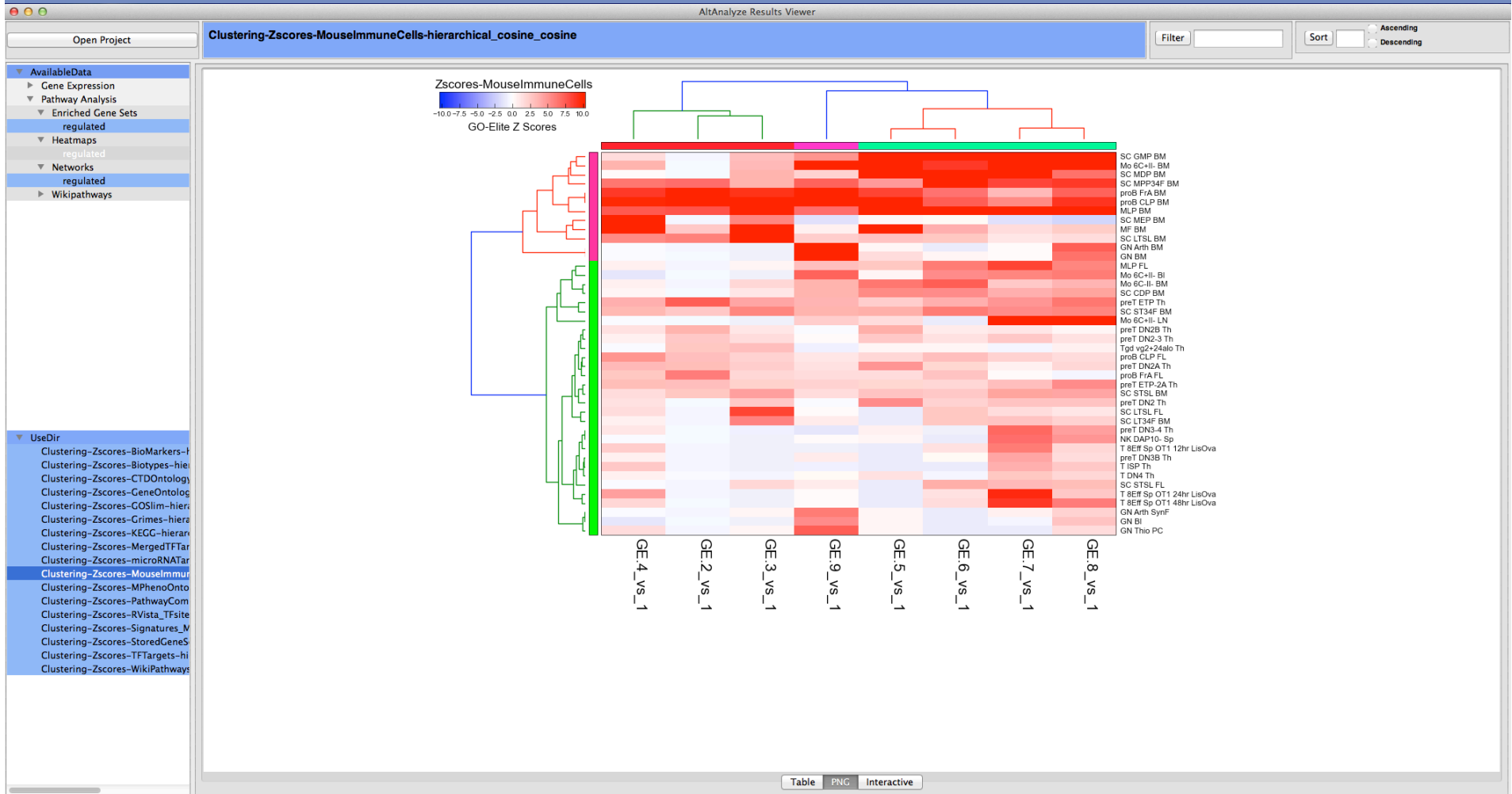


Table Length: 5100

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-BioMarkers-hierarchical_cosine_cosine.txt

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-BioMarkers-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-Biotypes-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-GeneOntology-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-COSim-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-Grimes-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-KEGG-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-Grimes-hierarchical_cosine_cosine

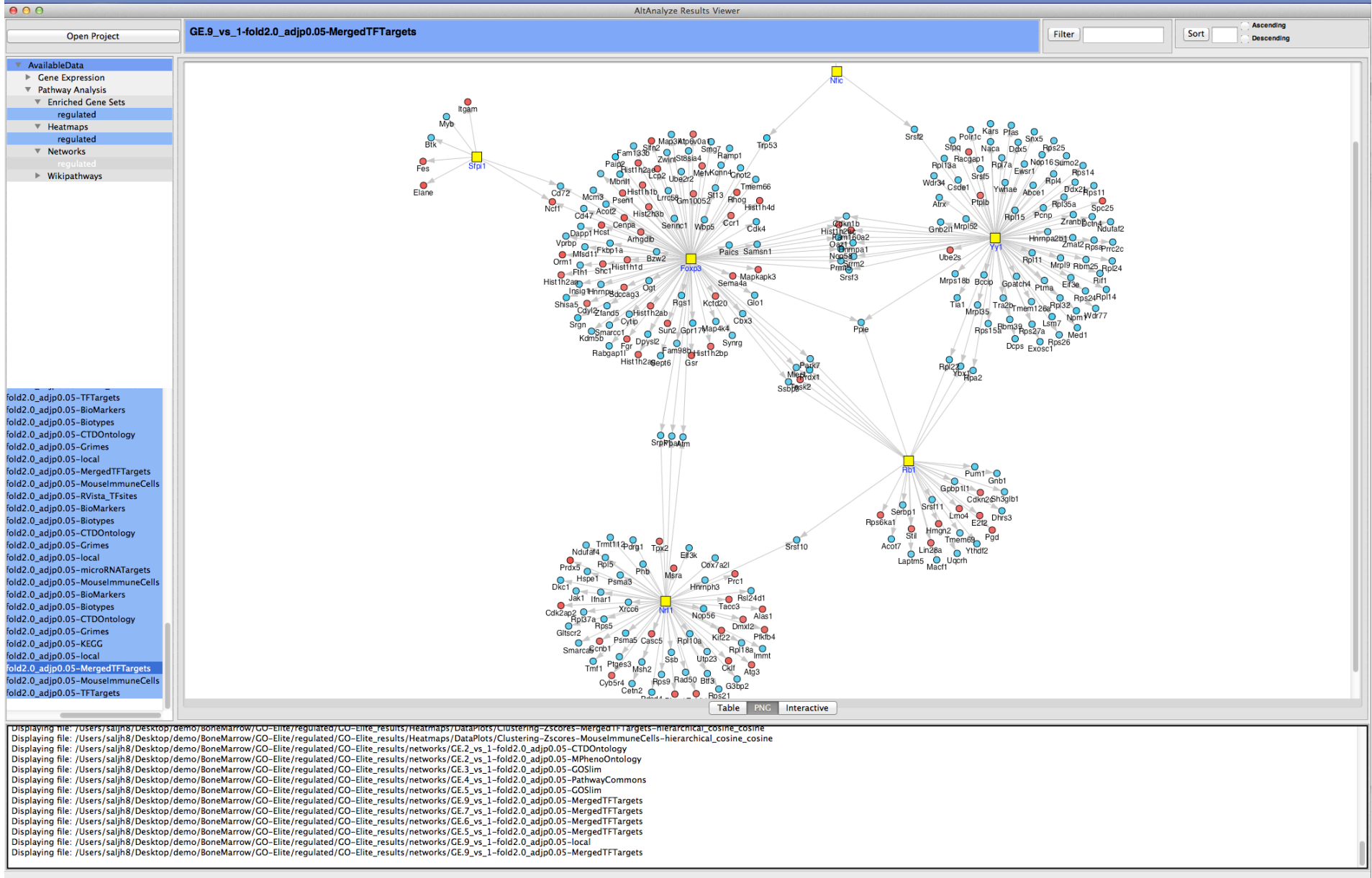
Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-KEGG-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-microRNATargets-hierarchical_cosine_cosine

Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-MergedTFTargets-hierarchical_cosine_cosine

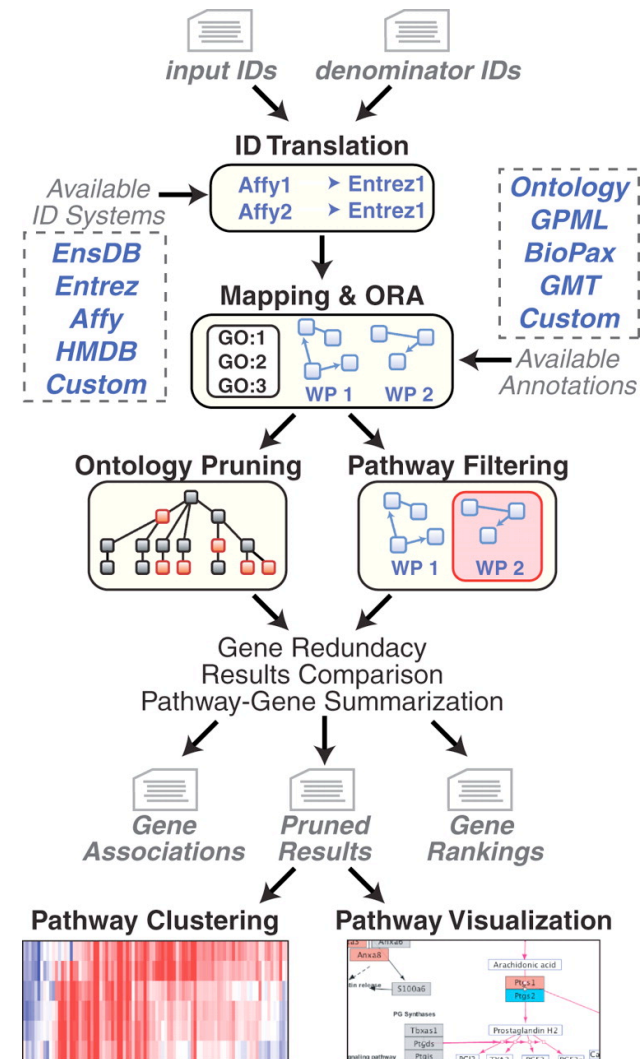
Displaying file: /Users/sajjh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/GO-Elite_results/Heatmaps/DataPlots/Clustering-Zscores-MouseImmuneCells-hierarchical_cosine_cosine

Gene Set Enrichment Analysis GO-Elite



GO-Elite Algorithm

- **Ontology Pruning**
 - Eliminates related significant scoring terms to optimize the output using the hierarchical tree structure.
 - Recommends non-ontology redundant terms based on gene composition overlap.
- **Outputs**
 - Tables
 - Heatmaps
 - Networks
 - **Colored WikiPathways**



Gene Network Analysis

Python File Edit View Search Table Help Alt 9 100% Sun 9:17 PM Salomonis, Nathan

AltAnalyze Results Viewer

Open Project AltAnalyze-network-WKBT_GE.8_vs_1-fold2.0_adjp0.05-interactions_direct Filter Sort Ascending Descending

AvailableData
Gene Expression
Heatmaps
ICGS
LineageProfiler
MarkerFinder
Networks
Pathways
Colored WikiPathway regulated
Enrichment Scores
PCA
QC
Summary Tables
Pathway Analysis

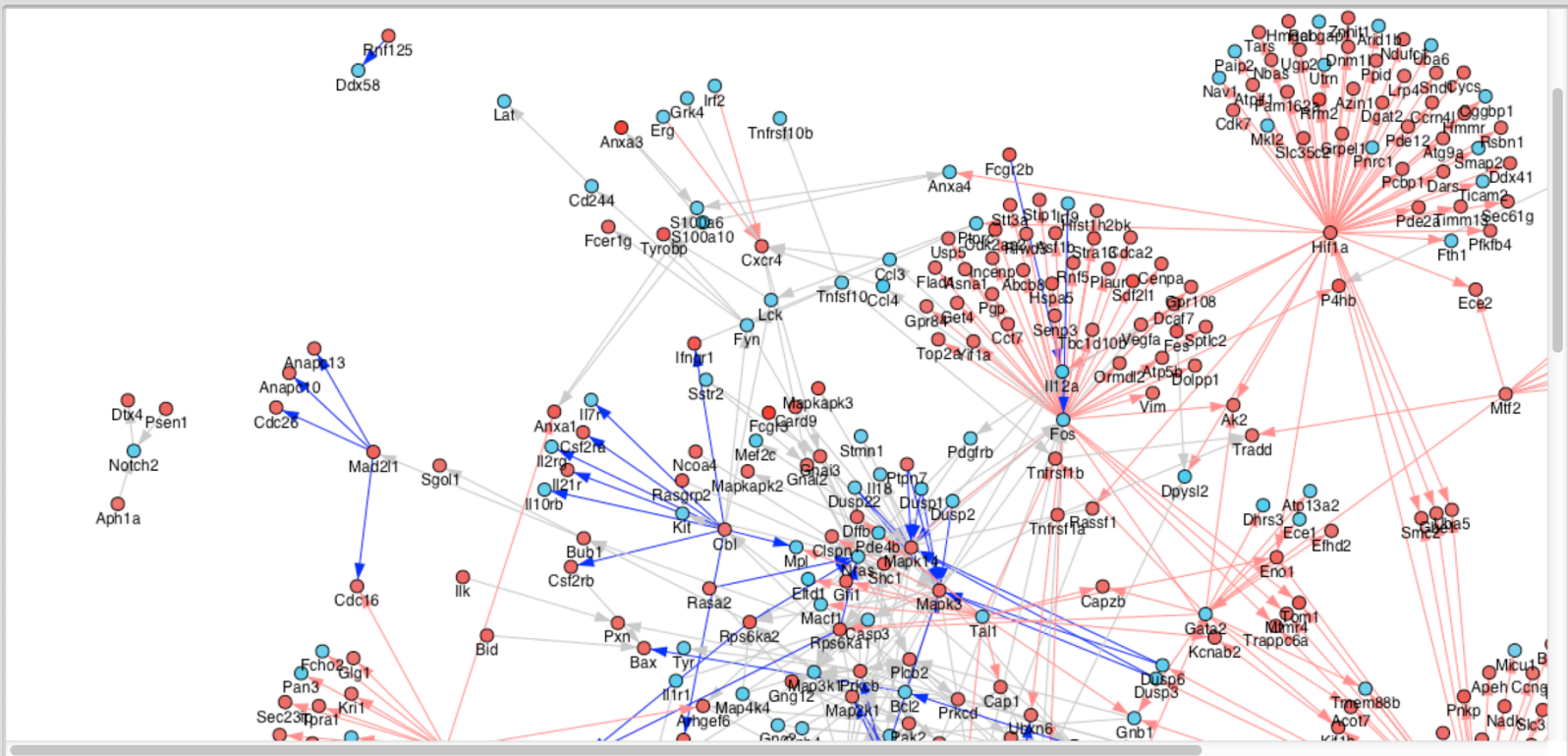
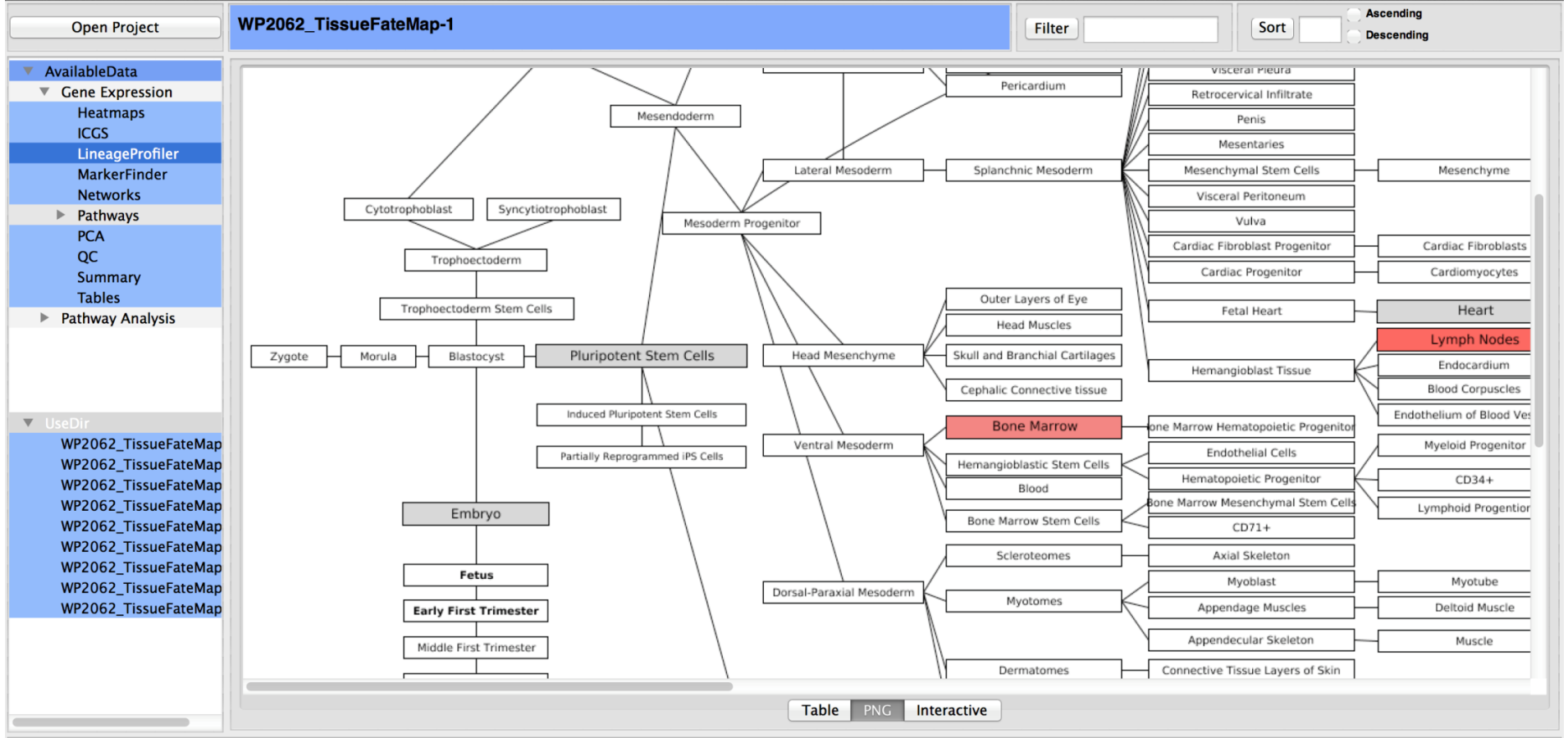


Table PNG Interactive

Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/networks/AltAnalyze-network-WKBT_GE.8_vs_1-fold2.0_adjp0.05-interactions_direct_degrees_summary
Unable to open png.
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/networks/AltAnalyze-network-WKBT_GE.9_vs_1-fold2.0_adjp0.05-interactions_direct_degrees_summary
Unable to open png.
Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/GO-Elite/regulated/networks/AltAnalyze-network-WKBT_GE.8_vs_1-fold2.0_adjp0.05-interactions_direct



LineageProfiler: Cell-type and Tissue Prediction



Currently displaying: SUMMARY
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/DataPlots/Clustering-LineageCorrelations-BoneMarrow-scrNASeq-protein_coding-zscores-groups_filtered-hierarchical_cosine_cosine.txt
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/ICGS/Clustering-BoneMarrow-scrNASeq-CORRELATED-FEATURES-hierarchical_cosine_correlation.txt
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/ICGS/Clustering-exp.BoneMarrow-scrNASeq-G3 Sell Ctl2a Gfi1b Smim5 Gm14434 Psap Ngp Ptrf G-hierarchical_cosine_correlation
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/DataPlots/Clustering-LineageCorrelations-BoneMarrow-scrNASeq-protein_coding-zscores-groups_filtered-hierarchical_cosine_cosine.txt

Automated Downstream Expression and Splicing Analyses

Python File Edit View Search Table Help Alt 9 [Icons] Sun 9:10 PM Salomonis, Nathan

AltAnalyze Results Viewer

Open Project DATASET-BoneMarrow-scrNASeq Filter Sort Ascending Descending

| | A | B | C | D | E | F | G | H |
|----|---------------|---------------|------------------------------------|---------------------|-----------|------------|---------------------------------------|--------------------------------|
| | Probesets | Symbol | Definition | Ensembl_id | Entrez_id | Unigene_id | GO-Process | GO-Function |
| 1 | Elane | Elane | elastase, neutrophil expressed | ENSMUSG00000020125 | | | acute inflammatory response to | cytokine binding // endopep |
| 2 | Ly6c2 | Ly6c2 | lymphocyte antigen 6 complex, | ENSMUSG00000022584 | | | | |
| 3 | Itgb2l | Itgb2l | integrin beta 2-like [Source:MC | ENSMUSG00000000157 | | | cell-matrix adhesion // inflam | receptor activity |
| 4 | 1100001G20Rik | 1100001G20Rik | RIKEN cDNA 1100001G20 gene | ENSMUSG000000051748 | | | response to lipopolysaccharide | metalloenzyme activator acti |
| 5 | Ctsg | Ctsg | cathepsin G [Source:MGI Symbc | ENSMUSG00000040314 | | | defense response to Gram-posit | heparin binding // peptidase |
| 6 | Igsf6 | Igsf6 | immunoglobulin superfamily, n | ENSMUSG000000035004 | | | biological_process | molecular_function |
| 7 | S100a8 | S100a8 | S100 calcium binding protein A | ENSMUSG000000056054 | | | activation of caspase activity // | antioxidant activity // calci |
| 8 | Hp | Hp | haptoglobin [Source:MGI Symbc | ENSMUSG000000031722 | | | acute inflammatory response // | antioxidant activity // cataly |
| 9 | Ermap | Ermap | erythroblast membrane-associ | ENSMUSG000000028644 | | | biological_process | molecular_function |
| 10 | Ms4a3 | Ms4a3 | membrane-spanning 4-domair | ENSMUSG000000024681 | | | biological_process | protein binding |
| 11 | Ly6c1 | Ly6c1 | lymphocyte antigen 6 complex, | ENSMUSG000000079018 | | | | |
| 12 | Fcgr3 | Fcgr3 | Fc receptor, IgG, low affinity III | ENSMUSG000000059498 | | | antibody-dependent cellular cy | IgG binding // IgG receptor |
| 13 | Car1 | Car1 | carbonic anhydrase 1 [Source:M | ENSMUSG000000027556 | | | one-carbon metabolic process | carbonate dehydratase activi |
| 14 | Trem3 | Trem3 | triggering receptor expressed c | ENSMUSG000000041754 | | | biological_process | bacterial cell surface binding |
| 15 | Clec12a | Clec12a | C-type lectin domain family 12 | ENSMUSG000000053063 | | | regulation of immune response | carbohydrate binding // pro |
| 16 | Anxa3 | Anxa3 | annexin A3 [Source:MGI Symbc | ENSMUSG000000029484 | | | defense response to bacterium | calcium ion binding // calci |
| 17 | Fcnb | Fcnb | ficolin B [Source:MGI Symbol;Ac | ENSMUSG000000026835 | | | complement activation, lectin pathway | |
| 18 | Mpo | Mpo | myeloperoxidase [Source:MGI S | ENSMUSG000000009350 | | | defense response to fungus // | heme binding // heparin bin |
| 19 | Ltf | Ltf | lactotransferrin [Source:MGI Sy | ENSMUSG000000032496 | | | cellular iron ion homeostasis // | ferric iron binding // hepari |
| 20 | Ngp | Ngp | neutrophilic granule protein [Sc | ENSMUSG000000032484 | | | biological_process // defense r | cysteine-type endopeptidase |
| 21 | Klf1 | Klf1 | Kruppel-like factor 1 (erythroid) | ENSMUSG000000054191 | | | chromatin remodeling // embr | DNA binding // core promot |
| 22 | Atp1b2 | Atp1b2 | ATPase, Na+/K+ transporting, | ENSMUSG000000041329 | | | cell adhesion // potassium ion | protein binding |
| 23 | S100a9 | S100a9 | S100 calcium binding protein A | ENSMUSG000000056071 | | | actin cytoskeleton reorganizati | antioxidant activity // calci |
| 24 | Camp | Camp | cathelicidin antimicrobial pepti | ENSMUSG000000038357 | | | defense response to Gram-neg | protein binding |
| 25 | Clec4e2 | Clec4e2 | C-type lectin domain family 4 | ENSMUSG000000020148 | | | innate immune response | carbohydrate binding |

Table PNG Interactive

Currently displaying: SUMMARY
 Displaying file: /Users/saljh8/Desktop/demo/BoneMarrow/ExpressionOutput/DATASET-BoneMarrow-scrNASeq.txt
 Plotting values from: Elane
 Plotting values from: Mpo
 Plotting values from: Klf1



Automated Downstream Expression and Splicing Analyses

Python File Edit View Search Table Help Alt 9 100% Sun 9:11 PM Salomonis, Nathan

AltAnalyze Results Viewer

Open Project DATASET-BoneMarrow-scrRNASeq Filter Sort Ascending Descending

AvailableData

- Gene Expression
 - Heatmaps
 - ICGS
 - LineageProfiler
 - MarkerFinder
 - Networks
 - Pathways
 - PCA
 - QC
 - Summary
 - Tables
 - Pathway Analysis
- UseDir
 - DATASET-BoneMarrow-
 - GenMAPP-BoneMarrow-
 - LineageCorrelations-Bo
 - LineageCorrelations-Bo
 - LineageCorrelations-Bo
 - LineageCorrelations-Bo
 - SUMMARY-BoneMarrow

| | A | B | C | D | E | F | G | H |
|----|--------------|---|---|---|---|---|---------------------------------------|--------------------------------|
| 1 | Probesets | | | | | | O-Process | GO-Function |
| 2 | Elane | | | | | | acute inflammatory response to | cytokine binding // endopep |
| 3 | Ly6c2 | | | | | | ell-matrix adhesion // inflamr | receptor activity |
| 4 | Itgb2l | | | | | | response to lipopolysaccharide | metalloenzyme activator acti |
| 5 | 1100001G20Ri | | | | | | defense response to Gram-posit | heparin binding // peptidase |
| 6 | Ctsg | | | | | | ological_process | molecular_function |
| 7 | Igsf6 | | | | | | activation of caspase activity // | antioxidant activity // calciu |
| 8 | S100a8 | | | | | | acute inflammatory response // | antioxidant activity // cataly |
| 9 | Hp | | | | | | ological_process | molecular_function |
| 10 | Ermap | | | | | | ological_process | protein binding |
| 11 | Ms4a3 | | | | | | antibody-dependent cellular cy | IgG binding // Ig receptor |
| 12 | Ly6c1 | | | | | | ne-carbon metabolic process | carbonate dehydratase activi |
| 13 | Fcgr3 | | | | | | ological_process | bacterial cell surface binding |
| 14 | Car1 | | | | | | regulation of immune response | carbohydrate binding // pro |
| 15 | Trem3 | | | | | | defense response to bacterium | calcium ion binding // calciu |
| 16 | Clec12a | | | | | | complement activation, lectin pathway | |
| 17 | Anxa3 | | | | | | defense response to fungus // | heme binding // heparin bin |
| 18 | Fcnb | | | | | | ellular iron ion homeostasis // | ferric iron binding // hepari |
| 19 | Mpo | | | | | | ological_process // defense r | cysteine-type endopeptidase |
| 20 | Ltf | | | | | | romatin remodeling // embr | DNA binding // core promot |
| 21 | Ngp | | | | | | ell adhesion // potassium ion | protein binding |
| 22 | Klf1 | | | | | | ctin cytoskeleton reorganizati | antioxidant activity // calciu |
| 23 | Atp1b2 | | | | | | defense response to Gram-neg | protein binding |
| 24 | S100a9 | | | | | | acute inflammatory response | carbohydrate binding |
| 25 | Camp | | | | | | | |
| 26 | Class3 | | | | | | | |

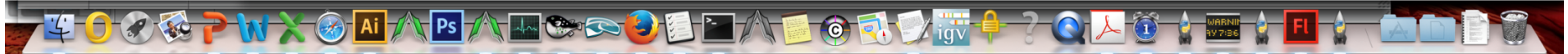
Figure 1

Klf1

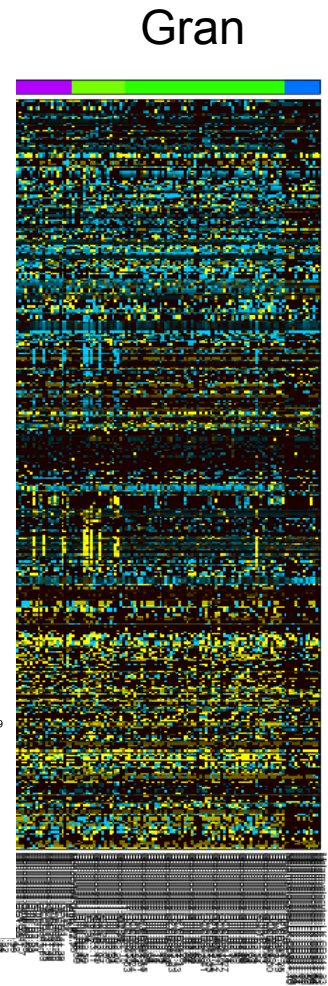
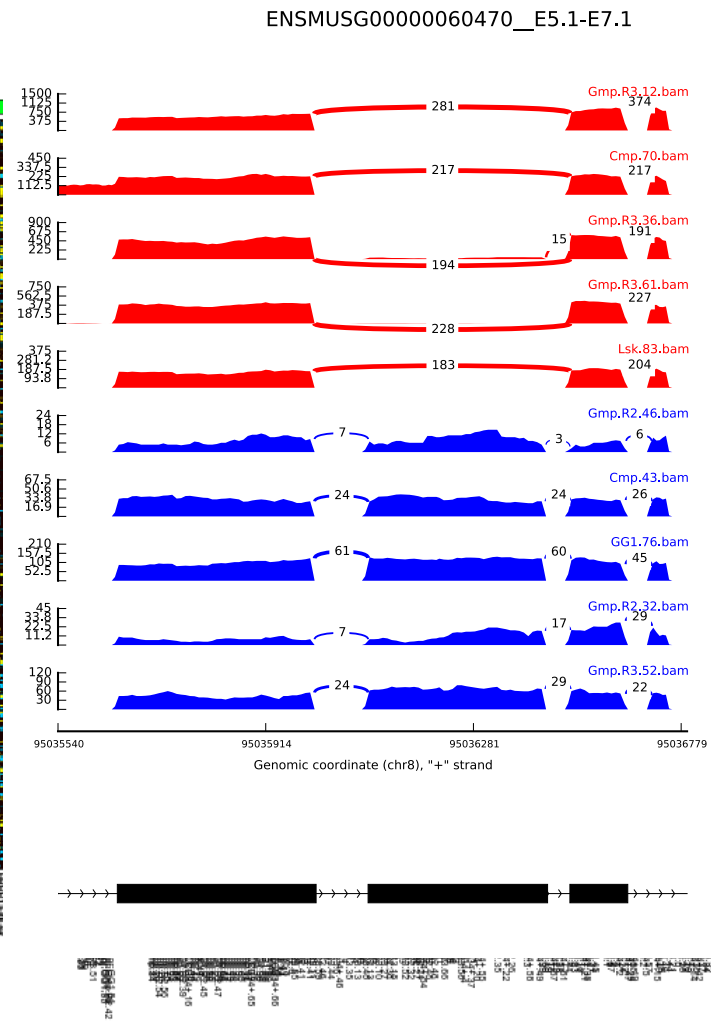
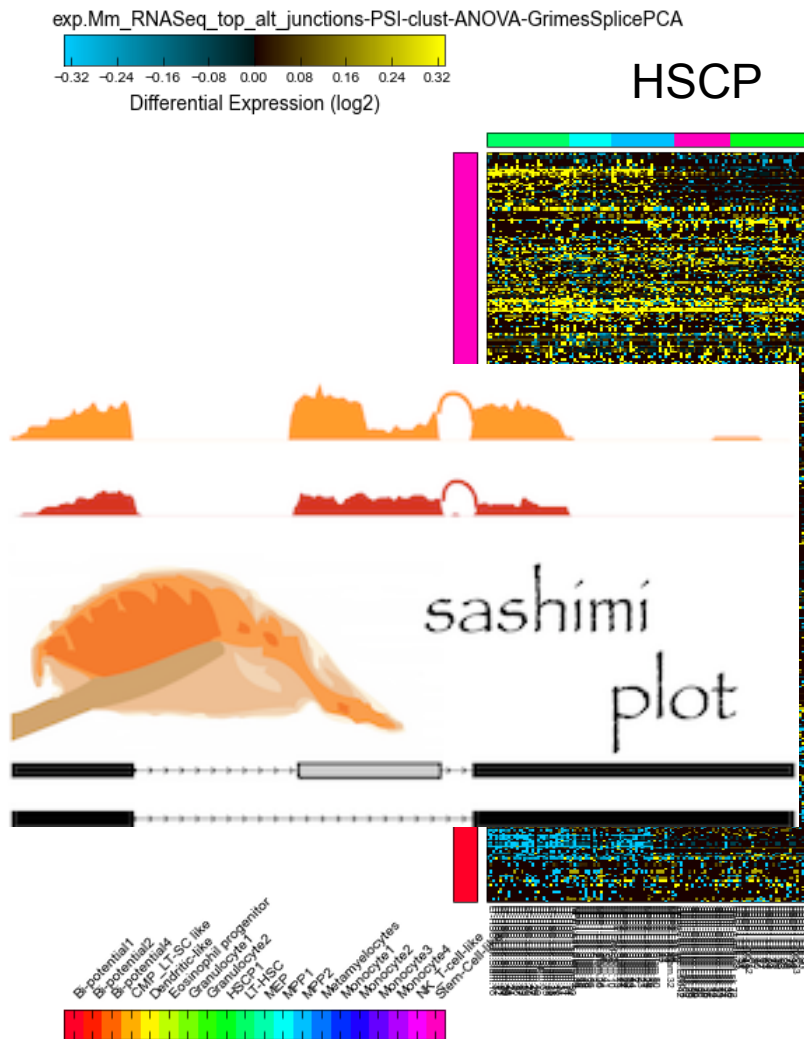
| Sample | Expression (approx.) |
|--------|----------------------|
| avg-1 | 0.1 |
| avg-2 | 0.3 |
| avg-3 | 0.1 |
| avg-4 | 6.5 |
| avg-5 | 0.5 |
| avg-6 | 0.1 |
| avg-7 | 0.1 |
| avg-8 | 0.1 |
| avg-9 | 0.1 |

Table PNG Interactive

Plotting values from: Elane
 Plotting values from: Mpo
 Plotting values from: Klf1
 Plotting values from: Ltf
 Plotting values from: Klf1



Splicing Analyses from the Single-Cell Workflow

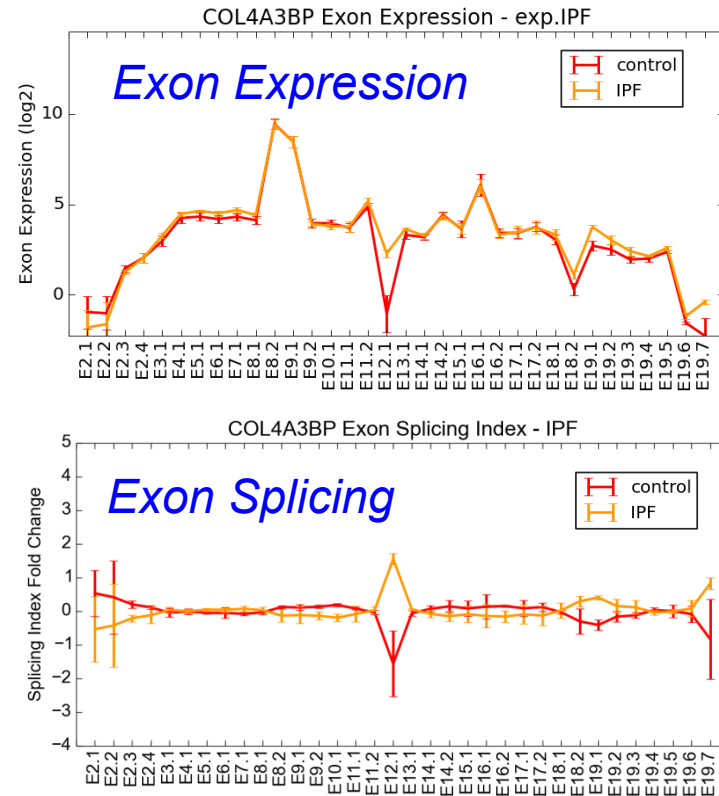


Automated and à la carte Options

- Gene Expression Analysis
 - Workflow
- Single-Cell Profiling
 - Workflow and à la carte
- Enrichment analyses
 - Workflow and stand-alone modules
- Network Analysis
 - Workflow and stand-alone modules
- Cell and Tissue-type Deconvolution
 - Workflow and stand-alone modules
- **Alternative Exon Analysis**
 - Workflow

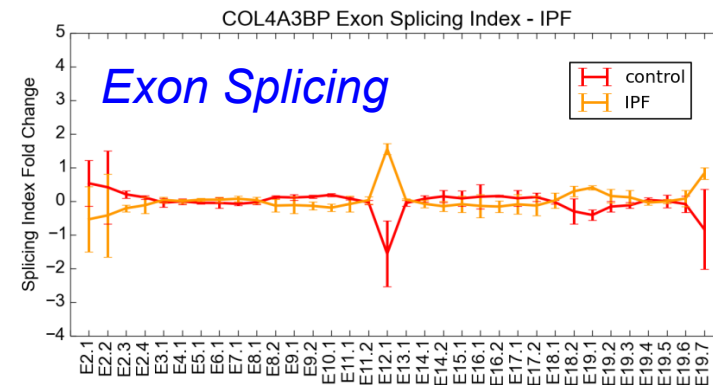
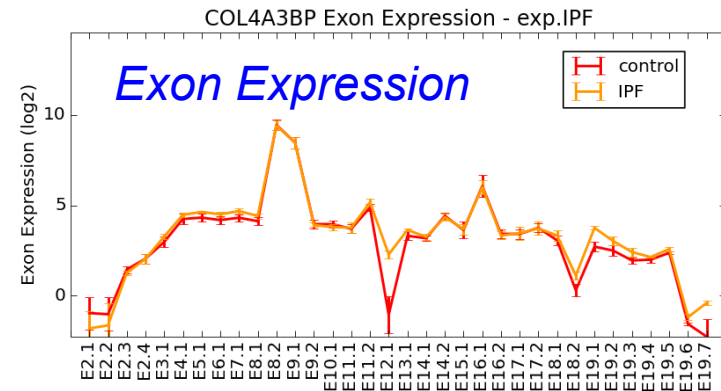
Visualization

- Exon Expression Line Graph
 - Typically RPKM of exons and standard deviation (SD)
- Exon Splicing Line Graph
 - Gene expression normalized exon intensities with SD

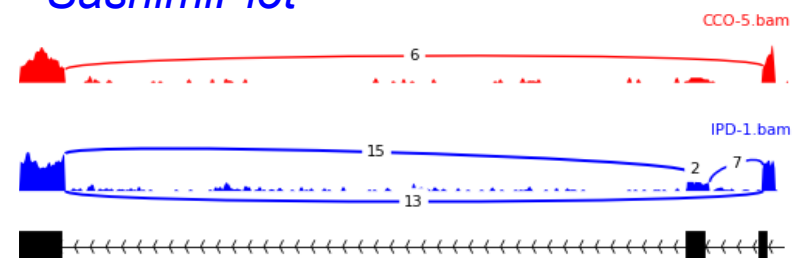


Visualization

- Exon Expression Line Graph
 - Typically RPKM of exons and standard deviation (SD)
- Exon Splicing Line Graph
 - Gene expression normalized exon intensities with SD
- SashimiPlot
 - Read-depth of genomic and junction aligning graph

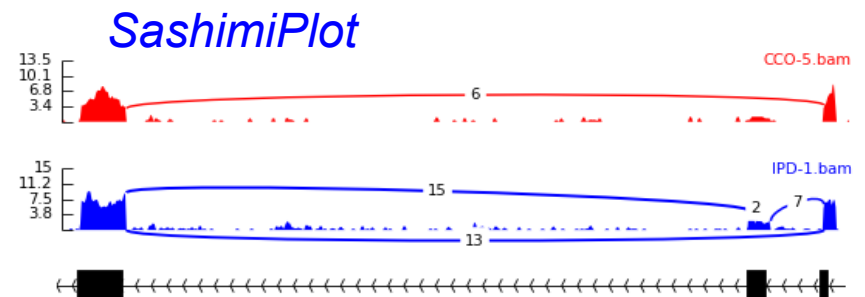
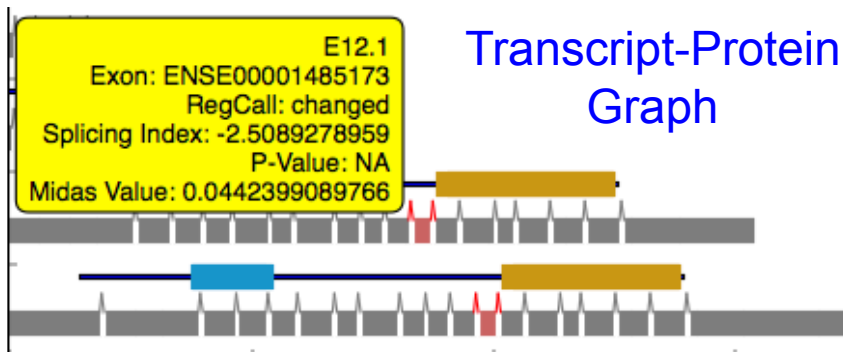
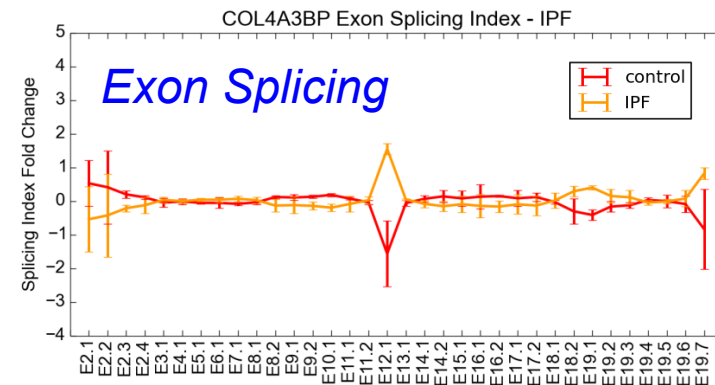
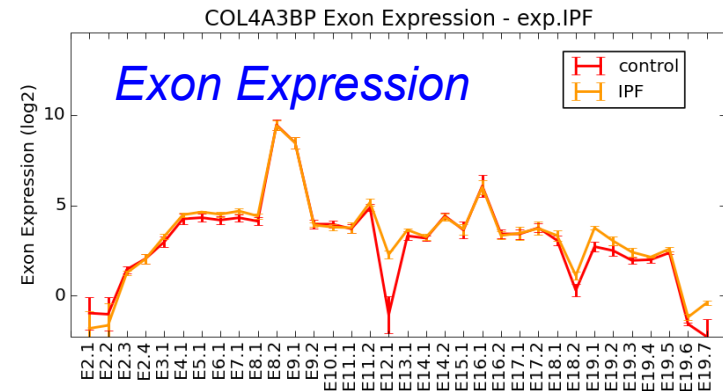


SashimiPlot

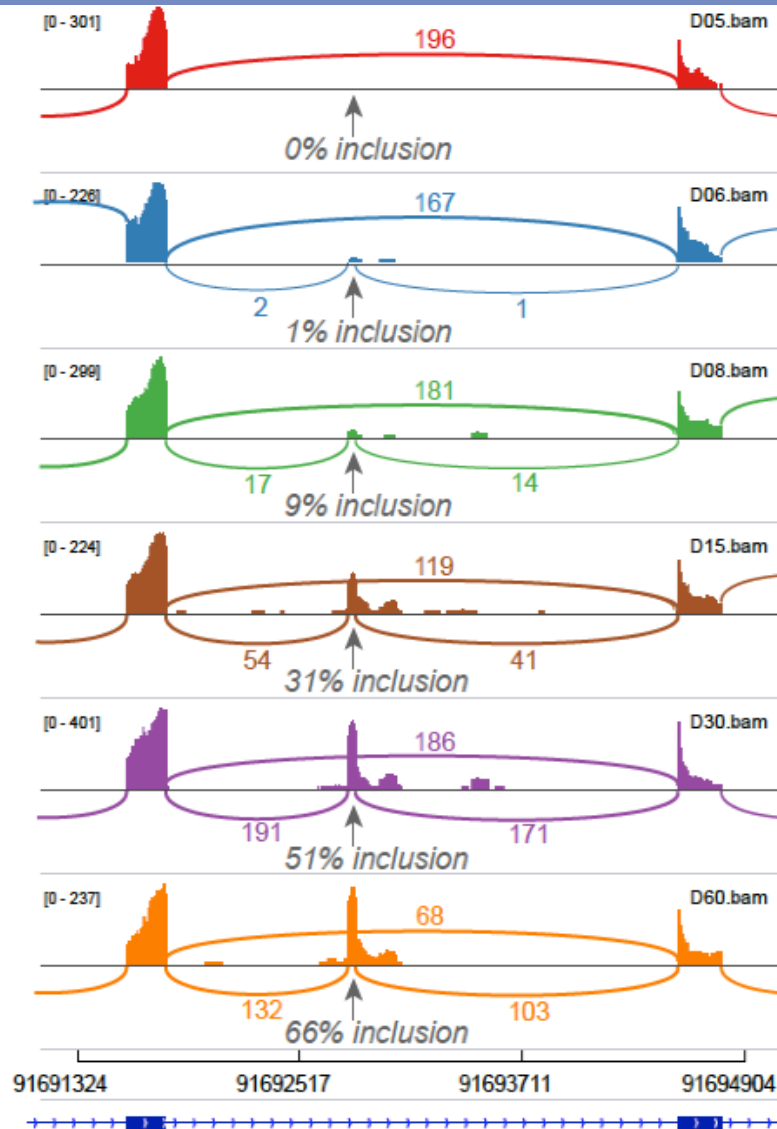


Visualization

- Exon Expression Line Graph
 - Typically RPKM of exons and standard deviation (SD)
- Exon Splicing Line Graph
 - Gene expression normalized exon intensities with SD
- SashimiPlot
 - Read-depth of genomic and junction aligning graph
- Transcript-Protein Structure Graph
 - RNA isoform exon structure with protein coding and domain potential.



Alternative Splicing Detection (continued)



Percent Spliced In Method

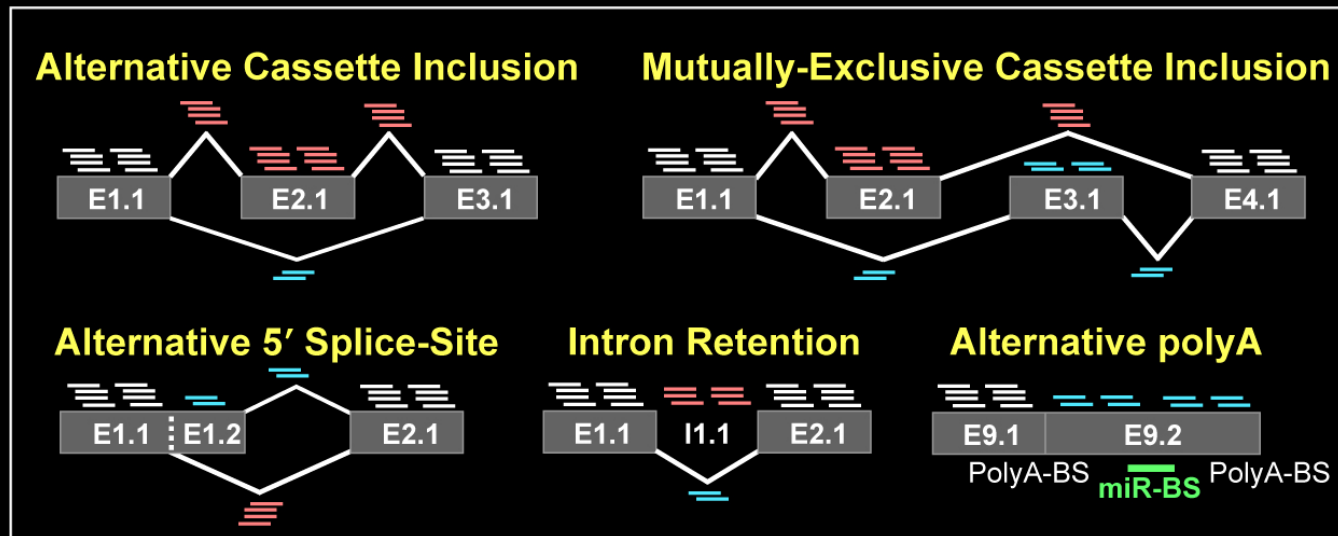
Measure the degree to which an exon is spliced in or out.

0.0 is no inclusion

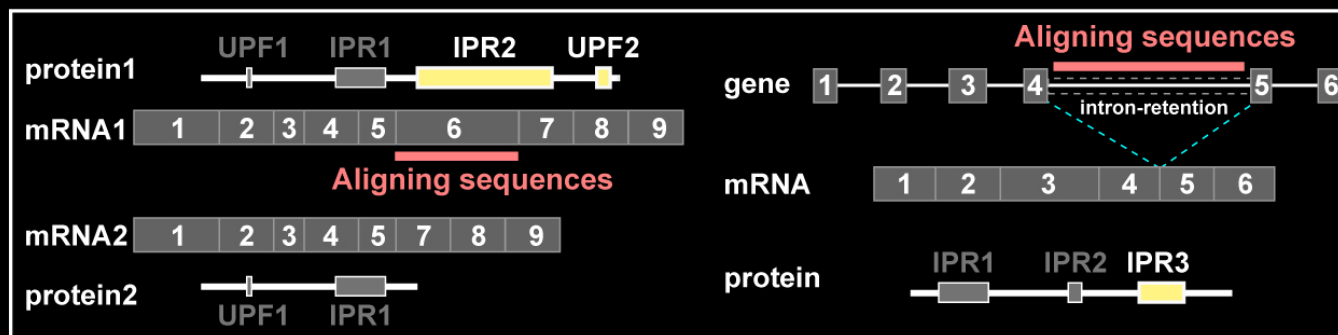
1.0 is complete inclusion

Inclusion is looked at for one junction compared to all that overlap with it in genomic space.

Analysis of Alternative Exon-Events and their Impact (AltAnalyze)



Domain Level Predictions



Acknowledgements

Salomonis Lab

Meenakshi Venkatasubramanian
Stuart Hay
Kashish Chetal



Grimes Lab (CCHMC)

H. Leighton Grimes
Andre Olsson



Singh Lab (CCHMC)

Harinder Singh
Virendra Chaudhri



**CCHMC Single-Cell and DNA
Sequencing Core Facilities**

Aronow Lab (CCHMC)

Bruce Aronow



Funding

NHBLI Progenitor Cell Biology
Consortium:

-Bioinformatics Core