

iPathwayGuide

Expression Analysis with
iPathwayGuide

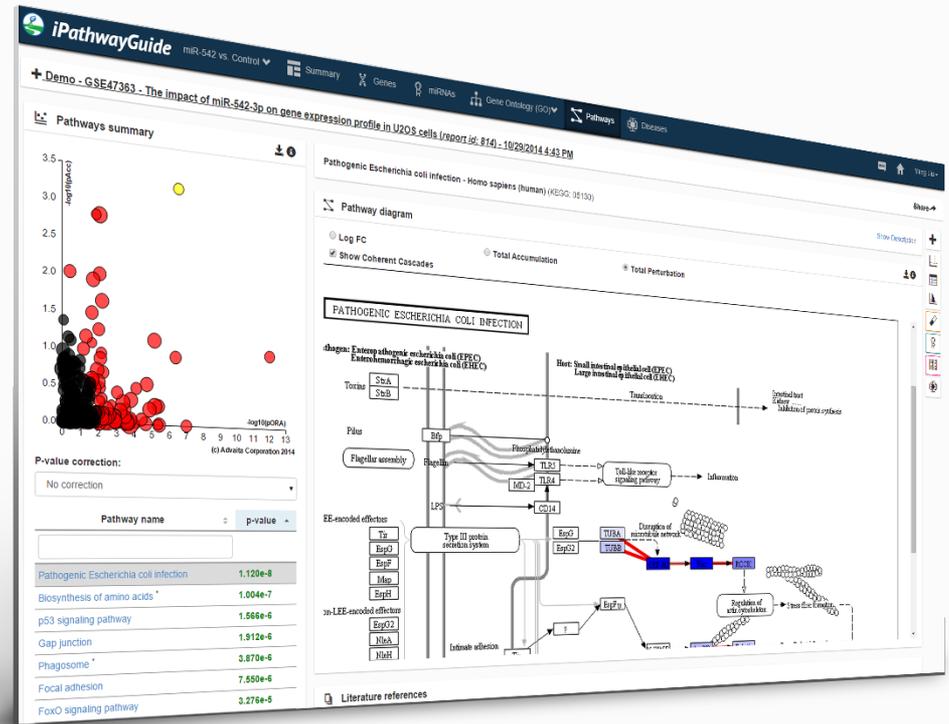


Agenda: 04. Expression Analysis after siRNA

- What is expression analysis
 - Insights from miRNA inference
 - Insights from GO terms– how we make it useful
 - Insights from Diseases
- Methodology
 - miRNA inference
 - GO Analysis: ontologies, elim & weight, additional details
 - Diseases: ontology
 - How to generate a Meta-Analysis
- TRY IT! Demonstrate how iPathwayGuide gives better insights
 - Dataset bkg
 - Step-by-step navigation: Enter from email, Summary, miRNA, GO Terms, Diseases, Printable Report
 - How to generate meta-analysis: purchase report, select contrasts, name conditions, submit
 - Q & A

iPathwayGuide Core Functions

- ✓ DE Genes
- ☐ Predicted miRNAs
- ☐ GO Analysis
 - Biological processes
 - Molecular functions
 - Cellular components
- ✓ Pathway Analysis (Drugs, miRNAs, SNPs)
- ☐ Diseases
- ☐ Meta analysis



EXPRESSION ANALYSIS IN iPATHWAYGUIDE

MIRNA INFERENCE

- Which miRNAs might be active in the experiment?

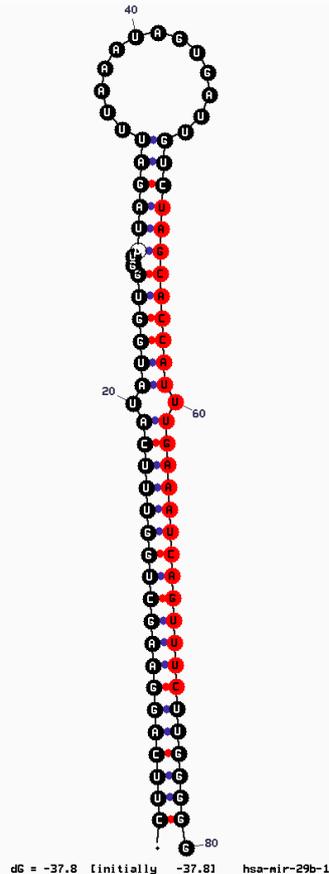
GENE ONTOLOGY ANALYSIS

- Which biological processes, cellular components, or molecular functions are impacted in the experiment?
- What related processes are likely affected?

RELATED DISEASES

- What human diseases are associated with similar changes in expression?

MICRO-RNA INFERENCE ANALYSIS



miRNAs bind specific nucleotide sequences to inhibit transcription

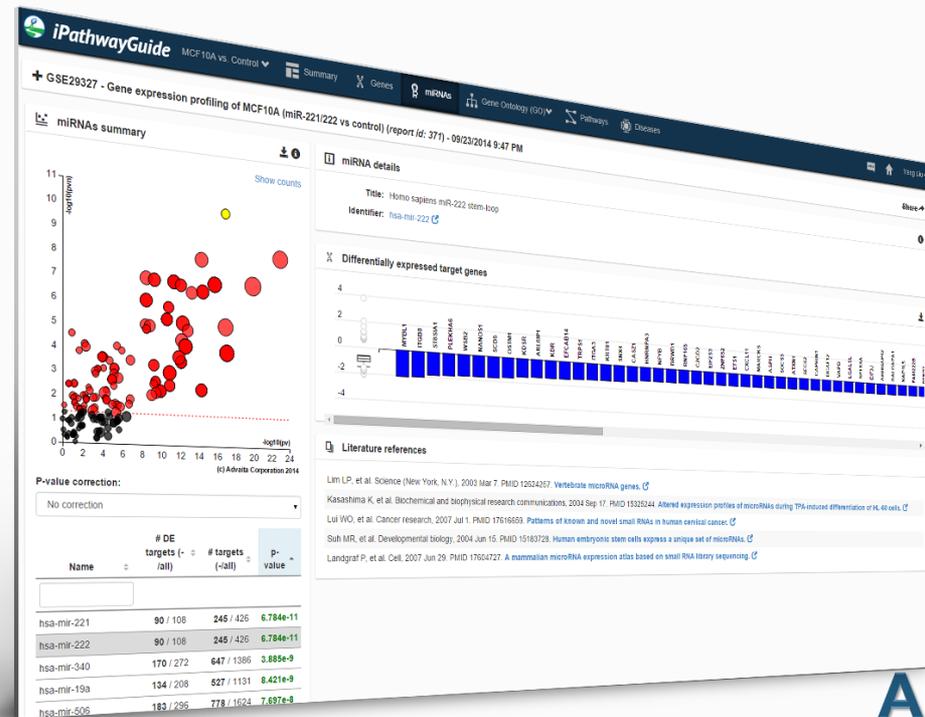
- miRNA usually have many genes (miRBase, TargetScan)
- miRNA activity can be inferred by observing how many target genes are downregulated
- Must correct for multiple comparisons

miRNA Analysis

RESEARCH QUESTION: Which miRNAs are most likely to be active, according to the DE downregulated genes in this experiment?

METHODOLOGY:

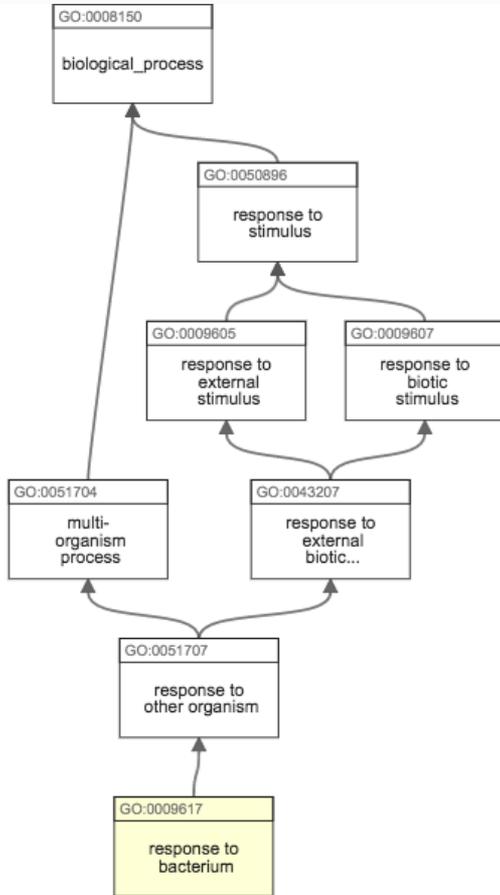
1. Identify miRNAs associated with DE genes
2. Count number of DE down vs total DE targets
3. Calculate enrichment score (pORA) for each miRNA



GENE ONTOLOGY (GO) ANALYSIS

AN ONTOLOGY IS A HIERARCHY OF TERMS

- Each GO TERM describes a
 - BIOLOGICAL PROCESS,
 - MOLECULAR FUNCTION, or
 - CELLULAR COMPONENT
- Associated genes are annotated to each term... and its parent(s)
- Correction factors need to account for nested terms:
 - ELIM and WEIGHT METHODS

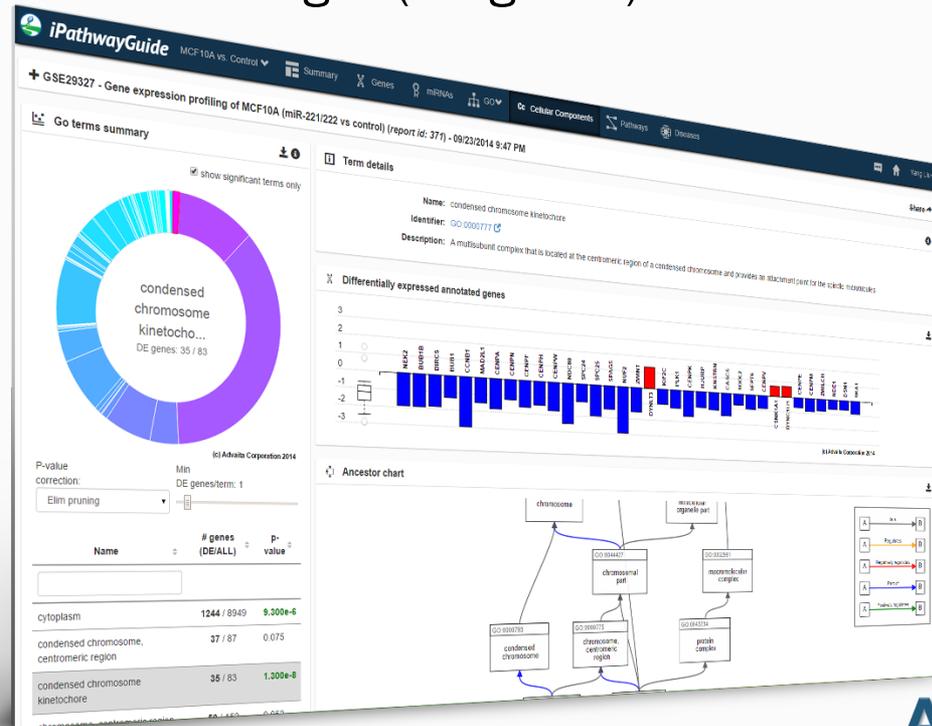


GO Analysis

RESEARCH QUESTION: What processes/ functions/ components are most impacted by the measured expression changes (DE genes)?

METHODOLOGY:

1. Identify GO terms associated with DE genes
2. Calculate enrichment score (pORA) for each term
3. Use *Elim* or *Weight* Methods to find informative terms



iPG: GO Analysis Method

- 3 Domains:

- Biological Processes
- Molecular Functions
- Cellular Components

- Hypergeometric distribution analysis

- 5 Correction Factors

- Uncorrected
- FDR
- Bonferroni (FWER)
- Elim*
- Weight*

* Takes into account repeated genes in the hierarchy (Alexa et al 2006)

The screenshot displays the iPathwayGuide interface for a GO analysis. The main title is "Demo - GSE47363 - The impact of miR-542-3p on gene expression profile in U2OS cells (report id: 2803) - 03/28/2015 03:34 PM".

Go terms summary: A circular bar chart shows the distribution of GO terms. The selected term is "negative regulation of cell proliferation" with 79 DE genes out of 586 total genes.

Term details:

- Name: negative regulation of cell proliferation
- Identifier: GO:0008285
- Description: Any process that stops, prevents or reduces the rate or extent of cell proliferation.

Differentially expressed annotated genes: A bar chart shows the log2 fold change for various genes. Genes with positive values (upregulated) include ILK, PRNP, HES6, ITIH3, TOR1B, HES5, CDKN1A, HES4, MUC2, SIRT1, CLDN3, CLDN4, LINC, LOC100289191, PAK1, PAK2, PAK3, PAK4, PAK5, PAK6, PAK7, PAK8, PAK9, PAK10, PAK11, PAK12, PAK13, PAK14, PAK15, PAK16, PAK17, PAK18, PAK19, PAK20, PAK21, PAK22, PAK23, PAK24, PAK25, PAK26, PAK27, PAK28, PAK29, PAK30, PAK31, PAK32, PAK33, PAK34, PAK35, PAK36, PAK37, PAK38, PAK39, PAK40, PAK41, PAK42, PAK43, PAK44, PAK45, PAK46, PAK47, PAK48, PAK49, PAK50, PAK51, PAK52, PAK53, PAK54, PAK55, PAK56, PAK57, PAK58, PAK59, PAK60, PAK61, PAK62, PAK63, PAK64, PAK65, PAK66, PAK67, PAK68, PAK69, PAK70, PAK71, PAK72, PAK73, PAK74, PAK75, PAK76, PAK77, PAK78, PAK79, PAK80, PAK81, PAK82, PAK83, PAK84, PAK85, PAK86, PAK87, PAK88, PAK89, PAK90, PAK91, PAK92, PAK93, PAK94, PAK95, PAK96, PAK97, PAK98, PAK99, PAK100. Genes with negative values (downregulated) include PAK101, PAK102, PAK103, PAK104, PAK105, PAK106, PAK107, PAK108, PAK109, PAK110, PAK111, PAK112, PAK113, PAK114, PAK115, PAK116, PAK117, PAK118, PAK119, PAK120, PAK121, PAK122, PAK123, PAK124, PAK125, PAK126, PAK127, PAK128, PAK129, PAK130, PAK131, PAK132, PAK133, PAK134, PAK135, PAK136, PAK137, PAK138, PAK139, PAK140, PAK141, PAK142, PAK143, PAK144, PAK145, PAK146, PAK147, PAK148, PAK149, PAK150, PAK151, PAK152, PAK153, PAK154, PAK155, PAK156, PAK157, PAK158, PAK159, PAK160, PAK161, PAK162, PAK163, PAK164, PAK165, PAK166, PAK167, PAK168, PAK169, PAK170, PAK171, PAK172, PAK173, PAK174, PAK175, PAK176, PAK177, PAK178, PAK179, PAK180, PAK181, PAK182, PAK183, PAK184, PAK185, PAK186, PAK187, PAK188, PAK189, PAK190, PAK191, PAK192, PAK193, PAK194, PAK195, PAK196, PAK197, PAK198, PAK199, PAK200.

Ancestor chart: A hierarchical diagram showing the relationship between GO terms. The root is "biological_process" (GO:0008150). Other terms include "biological regulation" (GO:0060007), "regulation of biological process" (GO:0000789), "cellular process" (GO:0009867), "single-organism process" (GO:0044889), "negative regulation of biological..." (GO:0044875), "regulation of cellular process" (GO:0000794), "cell proliferation" (GO:0000283), "negative regulation of cellular..." (GO:0044873), "regulation of cell proliferation" (GO:0044127), and "negative regulation of cell..." (GO:0008285).

Literature references:

- Dai Y, et al. Carcinogenesis, 2008 Sep. PMID 18487222. Prohibitin and the SWI/SNF ATPase subunit BRG1 are required for effective androgen antagonist-mediated transcriptional repression of androgen receptor-regulated genes.
- Guo W, et al. Molecular and cellular biochemistry, 2007 Jan. PMID 17043753. Prohibitin suppresses renal interstitial fibroblasts proliferation and phenotypic change induced by transforming growth factor-beta1.
- Hatcher GJ, et al. Developmental biology, 2001 Feb 15. PMID 11161571. TBX5 transcription factor regulates cell proliferation during cardiogenesis.
- Narita M, et al. Cell, 2006 Aug 11. PMID 16901784. A novel role for high-mobility group A proteins in cellular senescence and heterochromatin formation.
- Milne TA, et al. Proceedings of the National Academy of Sciences of the United States of America, 2005 Jan 18. PMID 15640349. Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors.

DISEASE ANALYSIS

INTERNATIONAL CLASSIFICATION OF DISEASES (ICD-10)

- A hierarchy of human diseases from the WHO, October 2015 (in US, from CMS and NCHS)
- Associated genes are annotated to each disease
 - Not necessarily nested
 - Use standard correction for multiple comparisons

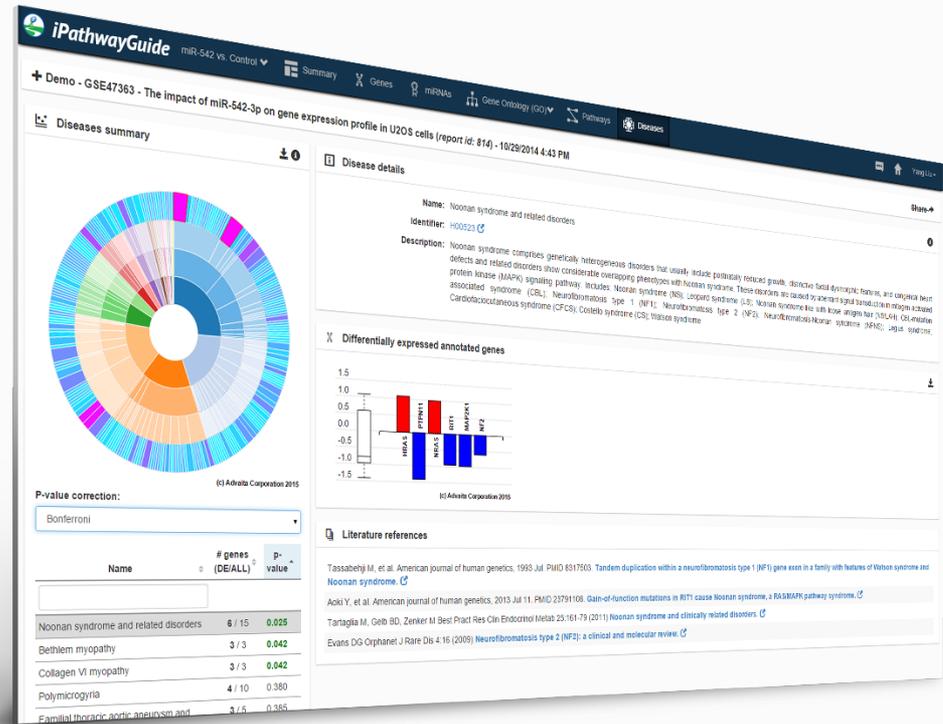


Disease Analysis

RESEARCH QUESTION: What diseases are most impacted by the measured DE Genes?

METHODOLOGY:

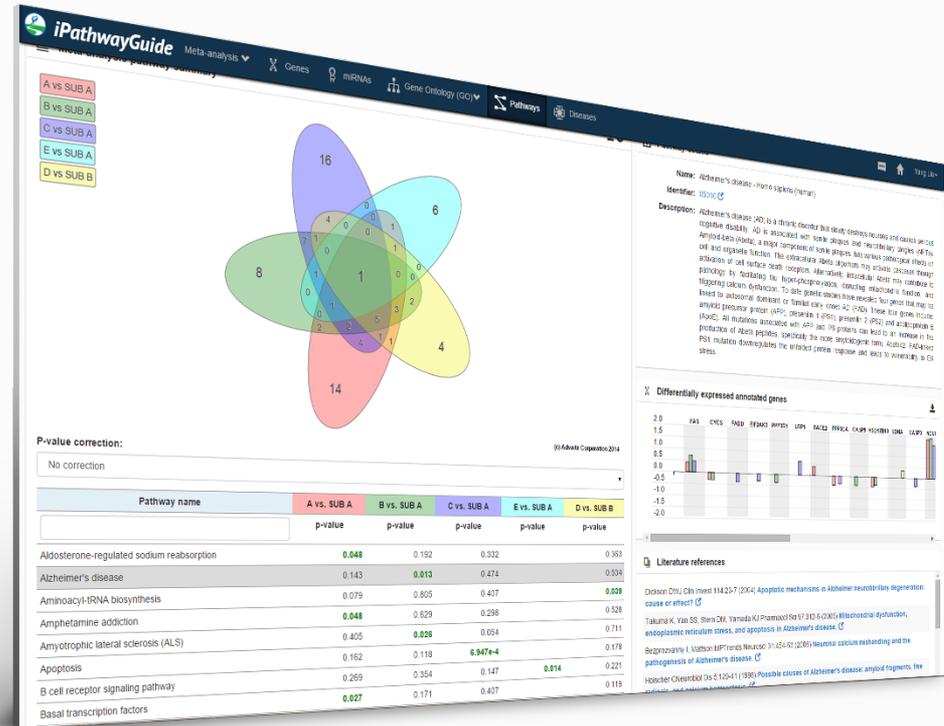
1. Identify ICD terms associated with DE genes
2. Calculate enrichment score (pORA) for each term
3. Use correction factors to eliminate false positives
4. Navigate hierarchy to explore results



Meta Analysis

RESEARCH QUESTION: How do differential expression profiles vary across experimental conditions?

- Compare up to 5
 - time points
 - doses
 - disease subtypes, etc.
- Combine proteomic and transcriptomic analyses
- Quickly identify common or unique patterns
 - Genes
 - miRNAs
 - GO terms
 - Pathways
 - Diseases



DATASET BACKGROUND

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN **AML1/ETO accelerates cell migration and impairs cell-to-cell adhesion and homing of hematopoietic stem/progenitor cells**

Received: 22 March 2016
Accepted: 20 September 2016
Published: 07 October 2016

Marco Saia^{1*}, Alberto Termanini², Nicoletta Rizzi², Massimiliano Mazza³, Elisa Barbieri^{2,4}, Debora Valli², Paolo Ciana³, Alicja M. Gruszka^{2,*} & Myriam Alcalay^{1,3,*}

The AML1/ETO fusion protein found in acute myeloid leukemias functions as a transcriptional regulator by recruiting co-repressor complexes to its DNA binding site. In order to extend the understanding of its role in preleukemia, we expressed AML1/ETO in a murine immortalized pluripotent hematopoietic stem/progenitor cell line, EML C1, and found that genes involved in functions such as cell-to-cell adhesion and cell motility were among the most significantly regulated as determined by RNA sequencing. In functional assays, AML1/ETO-expressing cells showed a decrease in adhesion to stromal cells, an increase of cell migration rate *in vitro*, and displayed an impairment in homing and engraftment *in vivo* upon transplantation into recipient mice. Our results suggest that AML1/ETO expression determines a more mobile and less adherent phenotype in preleukemic cells, therefore altering the interaction with the hematopoietic niche, potentially leading to the migration across the bone marrow barrier and to disease progression.

Approximately 12–15% of cases with adult acute myeloid leukemia (AML) carry the (8:21) translocation, which fuses the *AML1* (also known as *RUNX1*) and *ETO* (otherwise *RUNX1T1* or *MTG8*) genes and results in expression of the AML1/ETO chimeric protein¹. The published expression data demonstrate that AML1/ETO expression induces a distinct gene expression profile that involves the regulation of haematopoietic transcription and

- AML1/ETO fusion is common in AML; necessary but not sufficient to cause AML in mice
- Expressed AML1/ETO in HSC, used cytokines to induce myeloid differentiation
 - EV (empty vector)
 - A22 (transfected cells)
 - D0 (day 0)
 - D3 (day 3 differentiation)

LET'S TRY IT!

- STEPS:
 - ACCEPT SHARE
 - GENERATE META-REPORT
 - GO TERMS
 - DISEASES
 - PRINTABLE REPORT
 - META-ANALYSIS

HOW TO GENERATE A NEW META-ANALYSIS

- Dashboard > +Create new meta-report

+ Create new meta-report

+ Analyze a new experiment

- Find reports to include
 - Missing? Make sure it's purchased
 - Use green arrow to include a contrast (names can be edited later)
 - Repeat for up to 5 contrasts

Project Description	Creation Time
+ GSE48018 - Trivalent Influenza Vaccine - day 3	08/17/2016 03:42 PM
+ GSE77185 - AML1-ELO	12/19/2016 02:20 AM
+ GSE77185 - AML1-ELO: D0 vs A3	12/19/2016 01:49 AM
+ GSE77185 - AML1-ELO: EV vs AE22	12/19/2016 01:43 AM
✗ GSE77185 - AML1/ELO	12/09/2016 04:54 PM

RNA_EV.D0 vs. RNA_EV.A3	
RNA_EV.D0 vs. RNA_AE22.D0	
RNA_EV.D0 vs. RNA_AE22.A3	➤
RNA_EV.A3 vs. RNA_AE22.A3	
RNA_AE22.D0 vs. RNA_AE22.A3	

Selected contrasts:	
RNA_EV.D0 vs. RNA_EV.A3	⏪
RNA_EV.D0 vs. RNA_AE22.D0	⏪
RNA_EV.A3 vs. RNA_AE22.A3	⏪
RNA_AE22.D0 vs. RNA_AE22.A3	⏪

- Add Title & Description, then Create Report

iPG: Predicted miRNAs

- Bar plot displays count of downregulated targets vs upregulated targets
- Gene's plot lists target DE genes from max -FC to max +FC
- Clicking on any gene will navigate to genes pages

The screenshot displays the iPathwayGuide interface for the miRNA mmiu-miR-27b-3p. The top navigation bar includes 'iPathwayGuide', 'RNA_EV00 vs. RNA_EV03', 'Summary', 'Genes', 'miRNAs', 'Gene Ontology (GO)', 'Pathways', 'Diseases', and the user 'Cordelia Ziraldo'.

The main content is divided into two panels:

- miRNAs summary:** A bar chart showing the number of differentially expressed (DE) targets for various miRNAs. The y-axis lists miRNAs: mmiu-miR-27b-3p, mmiu-miR-27a-3p, mmiu-miR-323-3p, mmiu-miR-154-5p, mmiu-miR-493-3p, mmiu-miR-128-3p, mmiu-miR-136-5p, mmiu-miR-488-3p, mmiu-miR-362-3p, and mmiu-miR-329-3p. The x-axis represents the number of DE targets, ranging from 0 to 40. A legend indicates that blue bars represent downregulated targets and red bars represent upregulated targets. A 'Show p-values' link is also present.
- miRNA details:**
 - Identifier:** mmiu-miR-27b-3p
 - Previous Ids:** mmiu-miR-27b
 - Sequence:** UUCACAGUGGCUAGUUCUGC
 - Stem-loop:** mmiu-miR-27b
 - Family:** miR-27
 - Description:** Mus musculus miR-27b stem-loop
 - Sequence:** AGGUCAGAGCUAGCUGAUUGUGAAGCUGAUUGUUCGCCUUGUUCACAGUGGCUAGUUCGACCU
- Differentially expressed target genes:** A bar chart showing the log2 fold change (log2 FC) for various target genes. The y-axis ranges from -8 to 8. The x-axis lists genes including Gata3, Egr1, Cxcl1, Cxcl2, Cxcl3, Cxcl4, Cxcl5, Cxcl6, Cxcl7, Cxcl8, Cxcl9, Cxcl10, Cxcl11, Cxcl12, Cxcl13, Cxcl14, Cxcl15, Cxcl16, Cxcl17, Cxcl18, Cxcl19, Cxcl20, Cxcl21, Cxcl22, Cxcl23, Cxcl24, Cxcl25, Cxcl26, Cxcl27, Cxcl28, Cxcl29, Cxcl30, Cxcl31, Cxcl32, Cxcl33, Cxcl34, Cxcl35, Cxcl36, Cxcl37, Cxcl38, Cxcl39, Cxcl40, Cxcl41, Cxcl42, Cxcl43, Cxcl44, Cxcl45, Cxcl46, Cxcl47, Cxcl48, Cxcl49, Cxcl50, Cxcl51, Cxcl52, Cxcl53, Cxcl54, Cxcl55, Cxcl56, Cxcl57, Cxcl58, Cxcl59, Cxcl60, Cxcl61, Cxcl62, Cxcl63, Cxcl64, Cxcl65, Cxcl66, Cxcl67, Cxcl68, Cxcl69, Cxcl70, Cxcl71, Cxcl72, Cxcl73, Cxcl74, Cxcl75, Cxcl76, Cxcl77, Cxcl78, Cxcl79, Cxcl80, Cxcl81, Cxcl82, Cxcl83, Cxcl84, Cxcl85, Cxcl86, Cxcl87, Cxcl88, Cxcl89, Cxcl90, Cxcl91, Cxcl92, Cxcl93, Cxcl94, Cxcl95, Cxcl96, Cxcl97, Cxcl98, Cxcl99, Cxcl100.
- Literature references:** A list of references related to the miRNA and its targets, including:
 - Ahn HW, et al., Mol Hum Reprod. 16:463-471(2010), PMID:20215419 MicroRNA transcriptome in the newborn mouse ovaries determined by massive parallel sequencing.
 - Chiang HR, et al., Genes Dev. 24:992-1009(2010), PMID:20413612 Mammalian microRNAs: experimental evaluation of novel and previously annotated genes.
 - Dostie J, et al., RNA. 9:180-186(2003), PMID:12554860 Numerous microRNAs in neuronal cells containing novel microRNAs.
 - Houbaviy HB, et al., Dev Cell. 5:331-339(2003), PMID:12919686 Embryonic stem cell-specific microRNAs.
 - Lago-Quintana M, et al., Curr Biol. 12:735-739(2002), PMID:12007417 Identification of tissue-specific microRNAs from mouse.
 - Landgraf P, et al., Cell. 129:1401-1414(2007), PMID:17604727 A mammalian microRNA expression atlas based on small RNA library sequencing.
 - Poy MN, et al., Nature. 432:226-230(2004), PMID:15538371 A pancreatic islet-specific microRNA regulates insulin secretion.

iPG: GO Analysis

- Graph shows relative significance
 - Cyan = less significant
 - Magenta – most significant
- Can filter by # of DE Genes per term
- Genes plot shows directly annotated genes to GO Term
- GO tree shows hierarchy (interactive)

GO terms summary

show significant terms only

G-protein coupled receptor sig...
DE genes: 52 / 175

P-value correction: Min
Elim pruning: 5
DE genes/term: 5

Name	# genes (DE/ALL)	p-value
G-protein coupled receptor signaling pathway	52 / 175	3.900e-10
inflammatory response	71 / 247	3.900e-9
response to lipopolysaccharide	34 / 126	1.200e-7
defense response to bacterium	27 / 80	2.900e-6
leukocyte migration involved in inflammatory response	7 / 10	5.200e-6
negative regulation of angiogenesis	12 / 31	9.400e-6
positive regulation of angiogenesis	16 / 55	2.200e-5
antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-dependent	5 / 6	3.900e-5
immune response	122 / 542	3.900e-5
cell surface receptor signaling pathway	160 / 833	1.200e-4

Differentially expressed annotated genes

Ancestor chart

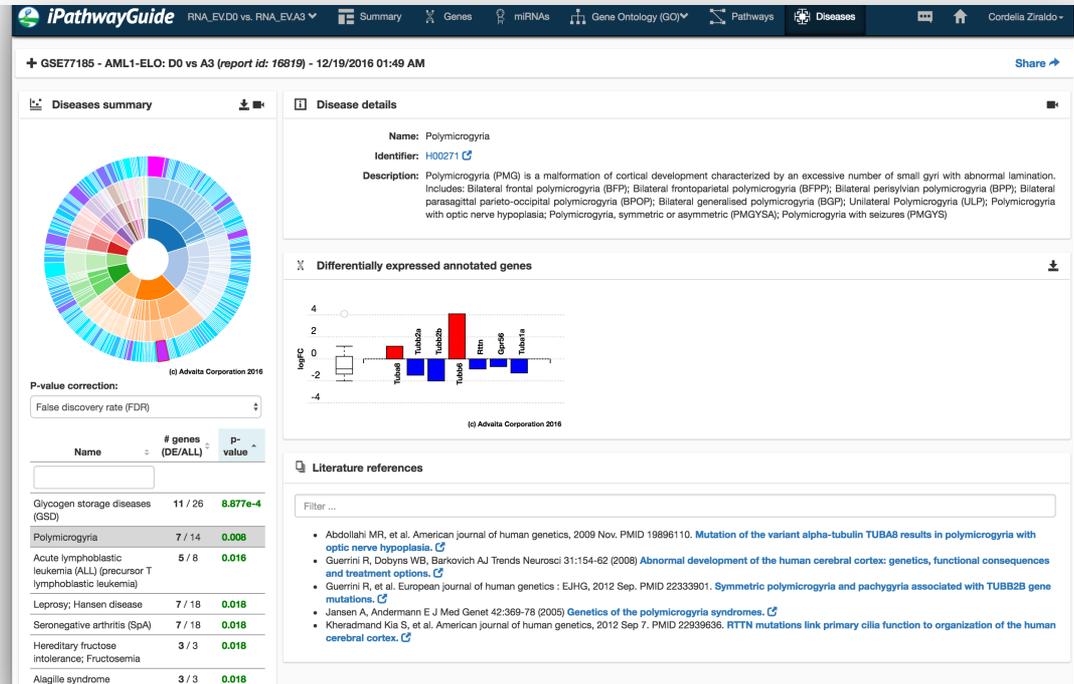
Literature references

Page 1 of 31

- Abe H, et al. Gene. 1999; PMID:9931443 Molecular cloning, chromosome mapping and characterization of the mouse CRT3D gene, a putative member of the leukocyte chemoattractant receptor family. [View article](#)
- Ames RS, et al. FEBS letters. 1999 Oct 21; PMID:9895825 In Xenopus oocytes the human C1a and C1a receptors elicit a promitogenic response to the anaphylatoxin C3. [View article](#)
- An S, et al. The Journal of biological chemistry. 2002 Jan 7; PMID:11817212 Src tyrosine kinase-induced cell proliferation, survival, and related signaling events mediated by G protein-coupled receptor. [View article](#)

iPG: Diseases

- Graph shows relative significance
 - Cyan = less significant
 - Magenta – most significant
- Can filter diseases by classification
- Genes plot shows directly annotated genes to disease



iPG: Diseases

- Graph shows relative significance
 - Cyan = less significant
 - Magenta – most significant
- Can filter diseases by classification
- Genes plot shows directly annotated genes to disease

iPathwayGuide RNA_EV.D0 vs. RNA_EV.A3 Summary Genes miRNAs Gene Ontology (GO) Pathways Diseases Cordelia Ziraldo

+ GSE77185 - AML1-ELO: D0 vs A3 (report id: 16819) - 12/19/2016 01:49 AM Share

Diseases summary

Disease details

Name: Polymicrogyria
 Identifier: H00271

Description: Polymicrogyria (PMG) is a malformation of cortical development characterized by an excessive number of small gyri with abnormal lamination. Includes: Bilateral frontal polymicrogyria (BFP); Bilateral frontoparietal polymicrogyria (BFPP); Bilateral perisylvian polymicrogyria (BPP); Bilateral parasagittal parieto-occipital polymicrogyria (BPOP); Bilateral generalised polymicrogyria (BGP); Unilateral Polymicrogyria (ULP); Polymicrogyria with optic nerve hypoplasia; Polymicrogyria, symmetric or asymmetric (PMG/SA); Polymicrogyria with seizures (PMG/S)

Differentially expressed annotated genes

logFC

TUBA8
 TUBA8
 TUBA8
 TUBA8
 TUBA8
 TUBA8
 TUBA8
 TUBA8

Literature references

Filter ...

- Abdollahi M, et al. American journal of human genetics, 2009 Nov. PMID 19896110. [Mutation of the variant alpha-tubulin TUBA8 results in polymicrogyria with optic nerve hypoplasia.](#)
- Guerrini R, Dobyns WB, Barkovich AJ Trends Neurosci 31:154-62 (2008) [Abnormal development of the human cerebral cortex: genetics, functional consequences and treatment options.](#)
- Guerrini R, et al. European journal of human genetics : EJHG, 2012 Sep. PMID 22333901. [Symmetric polymicrogyria and pachygyria associated with TUBB2B gene mutations.](#)
- Jansen A, Andermann E J Med Genet 42:369-78 (2005) [Genetics of the polymicrogyria syndromes.](#)
- Kheradmand Kia S, et al. American journal of human genetics, 2012 Sep 7. PMID 22939636. [RTTN mutations link primary cilia function to organization of the human cerebral cortex.](#)

Diseases summary table:

Name	# genes (DE/ALL)	p-value
Glycogen storage diseases (GSD)	11 / 26	8.877e-4
Polymicrogyria	7 / 14	0.008
Acute lymphoblastic leukemia (ALL) (precursor T lymphoblastic leukemia)	5 / 8	0.016
Leprosy; Hansen disease	7 / 18	0.018
Seronegative arthritis (SpA)	7 / 18	0.018
Hereditary fructose intolerance; Fructosemia	3 / 3	0.018
Atagile syndrome	3 / 3	0.018

iPG: Printable Report

- Generate Report from Summary Page
- Click Print to Save or Print
- Top 5 Pathways, top GO Terms, miRNAs, Diseases, all included with accompanying diagrams
- Complete Methods & References

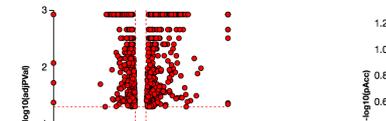
Printer friendly summary Print 



Title:	GSE77185 - AML1-ELO: D0 vs A3
Description:	EV.D0 vs EVA3, AE22.D0 vs AE22.A3
Organism:	Mus musculus (10090)
Contrast:	RNA_EV.D0 vs. RNA_EVA3
Creation time:	12/19/2016 01:49 AM

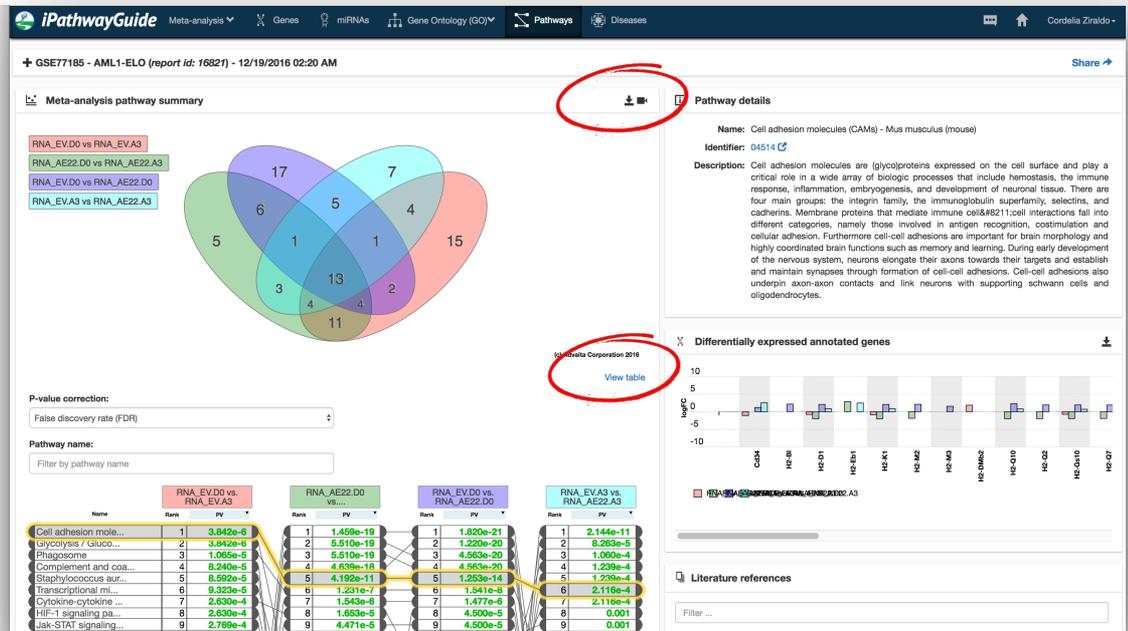
1. Introduction

In this experiment, 941 differentially expressed genes were identified out of a total of 11065 genes with measured expression. These were obtained using a threshold of 0.05 for statistical significance (p-value) and a log fold change of expression with absolute value of at least 0.6. These data were analyzed in the context of pathways obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Release 78.0+/06-02, Jun 16) (Kanehisa et al., 2000; Kanehisa et al., 2002), gene ontologies from the Gene Ontology Consortium database (2016-Mar14) (Ashburner et al., 2000; Gene Ontology Consortium, 2001), miRNAs from the miRBase (Release 21) and TARGETSCAN (TargetsCan version: Mouse:7.1, Human:7.0) databases (Agarwal et al., 2015; Nam et al., 2014; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2014; Friedman et al., 2009; Grimson et al., 2007), and diseases from the KEGG database (Release 78.0+/06-02, Jun 16) (Kanehisa et al., 2000; Kanehisa et al., 2002). In summary, 92 pathways were found to be significantly impacted. In addition, 2288 Gene Ontology (GO) terms, 3 miRNAs, and 34 diseases were found to be significantly enriched based on uncorrected p-values.

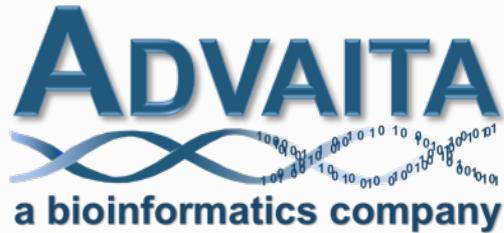


iPG: Meta Analysis

- View comparisons across any module
- Use Gene Bar Plot to identify potential biomarkers
- Download all figures and tables
- View Significant entities as a table or rank diagram



Stop Point Questions



Advaita Corporation
www.AdvaitaBio.com
46099 Five Mile Road, Plymouth, MI 48170
info@advaitabio.com
+1 (734) 922 - 0110