EMBL-EBI

Data standards, curation and ontologies

Last updated: 15 May 2018

Laura Huerta - lauhuema@ebi.ac.uk

Experiment annotation

In this activity you are going to work for **15 minutes** as a biocurator. Have a look at the following data corresponding to an experiment from Gene Expression Omnibus (GEO) database:

Series GSE44596	Query DataSets for GSE44596
Status	Public on Feb 20, 2014
Title	The effect of Sparstolonin B (SsnB) on gene expression in HCAECs
Organism	Homo sapiens
Experiment type	Expression profiling by array
Summary	Sparstolonin B is a novel bioactive compound isolated from Sparganium stoloniferum, an herb historically used in Traditional Chinese Medicine as an anti-tumor agent. SsnB has previously demonstrated anti-angiogenic properties. In functional assays, SsnB inhibited endothelial cell tube formation (Matrigel method) and cell migration (Transwell method) in a dose-dependent manner. We used microarrays to examine how SsnB affected the gene expression of human coronary artery endothelial cells (HCAECs), focusing in particular on pathways related to angiogenesis.
Overall design	Three plates of HCAECs were exposed to 100 micromolar SsnB and three plates of HCAECs were exposed to Vehicle Control (1:1000 dilution of DMSO). After 24 hours, RNA was extracted for microarrays and gene expression was analyzed.
Contributor(s)	Bateman H, Barth J, Lessner S
Citation(s)	Bateman HR, Liang Q, Fan D, Rodriguez V et al. Sparstolonin B inhibits pro-angiogenic functions and blocks cell cycle progression in endothelial cells. *PLoS One* 2013;8(8):e70500. PMID: 23940584
Submission date	Feb 23, 2013
Last update date	May 05, 2017
Contact name	Henry Rhodes Bateman
E-mail	hbateman@uscmed.sc.edu
Organization name	University of South Carolina
Department	Cell Biology & Anatomy
Lab	Bldg 1, RM C-36
Street address	6439 Garners Ferry Road
City	Columbia
State/province	SC
ZIP/Postal code	29209
Country	USA
Platforms (1)	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
Samples (6) Less...	GSM1087598 HCAEC_SsnB_rep1 GSM1087599 HCAEC_SsnB_rep2 GSM1087600 HCAEC_SsnB_rep3 GSM1087601 HCAEC_DMSO_rep1 GSM1087602 HCAEC_DMSO_rep2 GSM1087603 HCAEC_DMSO_rep3

As it is a real example, you can go to GEO database https://www.ncbi.nlm.nih.gov/geo/ and search for GSE44596 in case you need more information:

Introductory questions:

- What is the experiment about?
- Is it a microarray or a RNA-sequencing experiment?
- In which organism is the experiment done?
- What is the biological material used in the experiment?
- How many samples are authors testing?
- What are the two groups of samples compared?
- How many biological replicates are there in each group?
- Are there any technical replicates?
- What is the main variable that authors are studying?

Sample annotation

Now that we have a better understanding of the experiment we are going to work on annotating the samples with all the relevant information that we can find. To look for more details about each sample just click on each one.

The idea is to get all the information we need to fill in the following table:

	Sample characteristics		Experimental variables		
Assay Name	Characteristics [organism]	Characteristics [cell type]	FactorValue [compound]	FactorValue [dose]	Unit [concentration unit]
GSM1087598					
GSM1087599					
GSM1087600					
GSM1087601					
GSM1087602					
GSM1087603					