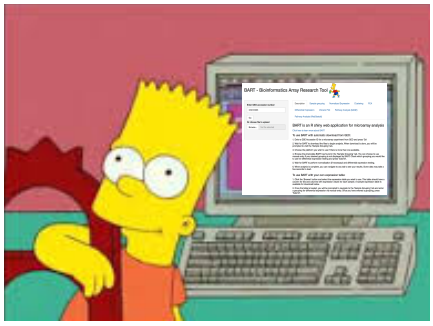
BART Tutorial:



Here we will demonstrate how to use BART using example GEO series GSE20986. First, we type the GSE accession number into the side panel and hit 'Go' to begin the download process. Once the download and parsing of the data is done (this can take a few minutes), a blue pop-up will appear prompting us to navigate to the sample grouping tab. We then click on the sample grouping tab and analyze the characteristics/phenotypical data that BART finds for each sample. We can decide one column to use, multiple columns to use, or manually insert a grouping of our choice into the 'Manual_Entry' column for grouping of samples. BART will use the characteristics defined by the column(s) selected to group samples for differential expression testing. However, there must be at least two groups to compare and at least two samples/replicates per group for differential expression testing, therefore BART will throw an error and request a new grouping if the entered grouping does not meet these qualifications. You must also select whether you would like to use the raw CEL files or the expression table found on GEO as a starting point for BART's analysis. If CEL files are chosen for analysis but are

BART - Bioinformatics Array Research Tool

Enter GEO accession number

GSE20986

Go

Or choose file to upload

Browse... No file selected

Platform:

GPL570

Group based on:

- source
- tissue
- cell type
- Manual_Entry
- Batch_Effect

Select which starting material to use for analysis:

- CEL files (if available)
- Expression table

Submit

Description | Sample grouping | Normalized Expression | Clustering | PCA | Differential Expression | Volcano Plot | Pathway Analysis (GAGE) | Pathway Analysis (WebGestalt)

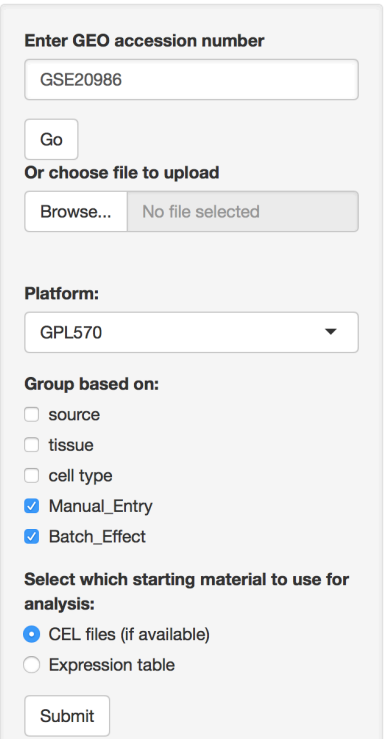
Please enter a grouping for differential expression testing and click 'Submit'

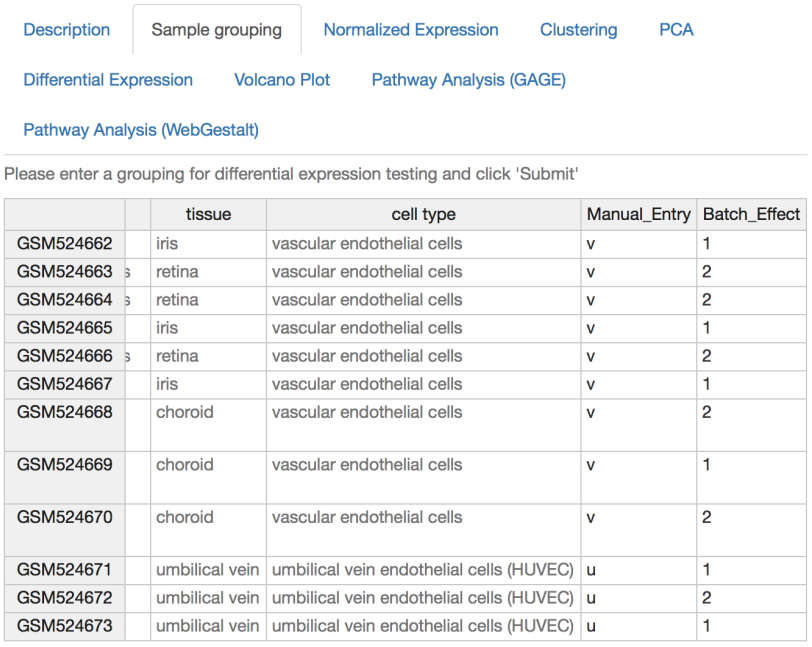
	source	tissue	cell type	Manual_Entry	Batch_E
GSM524662	vascular endothelial iris cells	iris	vascular endothelial cells	edit here	edit here
GSM524663	vascular endothelial retina cells	retina	vascular endothelial cells	edit here	edit here
GSM524664	vascular endothelial retina cells	retina	vascular endothelial cells	edit here	edit here
GSM524665	vascular endothelial iris cells	iris	vascular endothelial cells	edit here	edit here
GSM524666	vascular endothelial retina cells	retina	vascular endothelial cells	edit here	edit here
GSM524667	vascular endothelial iris cells	iris	vascular endothelial cells	edit here	edit here
GSM524668	vascular endothelial choroid cells	choroid	vascular endothelial cells	edit here	edit here
GSM524669	vascular endothelial choroid cells	choroid	vascular endothelial cells	edit here	edit here
GSM524670	vascular endothelial choroid cells	choroid	vascular endothelial cells	edit here	edit here
GSM524671	umbilical vein endothelial cells	umbilical vein	umbilical vein endothelial cells (HUVEC)	edit here	edit here
GSM524672	umbilical vein endothelial cells	umbilical vein	umbilical vein endothelial cells (HUVEC)	edit here	edit here
GSM524673	umbilical vein endothelial cells	umbilical vein	umbilical vein endothelial cells (HUVEC)	edit here	edit here

not available or not in the correct format, BART gives notification and proceeds with analysis using the expression table on GEO.

Batch Effect Correction

If you wish to apply a batch effect correction, you should have previous knowledge of the data and be aware of which samples belong to which batch. We recommend running through BART analysis without batch effect correction first if you are unsure that it is necessary. Then, after observing BART's PCA plot and heatmap you may wish to redo the analysis with a batch effect correction. You can apply a batch effect correction by typing in the batch grouping into the 'Batch_Effect' column and must also indicate a grouping for samples as described above. Keep in mind that the batches should not be the same as the grouping. Only the column(s) you select for grouping will be utilized by BART, therefore, if you do not wish to use the batch effect column, then you can leave it blank and not check the box in the side panel. If you choose to utilize both batch effect correction and manual entry of grouping, then your input may look like this:





		tissue	cell type	Manual_Entry	Batch_Effect
GSM524662		iris	vascular endothelial cells	v	1
GSM524663	s	retina	vascular endothelial cells	v	2
GSM524664	s	retina	vascular endothelial cells	v	2
GSM524665		iris	vascular endothelial cells	v	1
GSM524666	s	retina	vascular endothelial cells	v	2
GSM524667		iris	vascular endothelial cells	v	1
GSM524668		choroid	vascular endothelial cells	v	2
GSM524669		choroid	vascular endothelial cells	v	1
GSM524670		choroid	vascular endothelial cells	v	2
GSM524671		umbilical vein	umbilical vein endothelial cells (HUVEC)	u	1
GSM524672		umbilical vein	umbilical vein endothelial cells (HUVEC)	u	2
GSM524673		umbilical vein	umbilical vein endothelial cells (HUVEC)	u	1

Viewing Your Results:

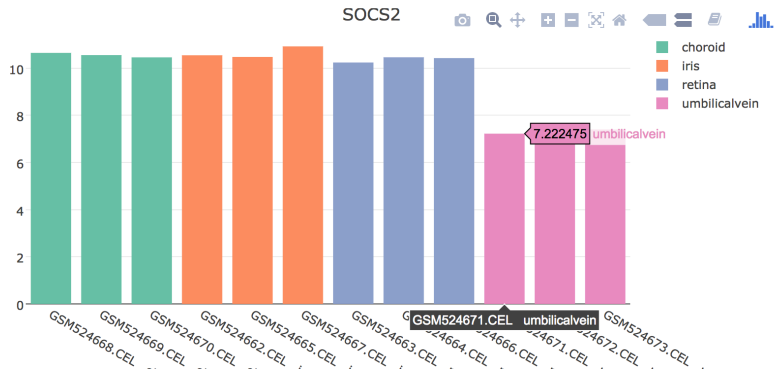
BART will show a gray notification when the analysis is complete and you may navigate to any tab to see your results. In the normalized expression tab, you can search for a gene or probe ID of interest and click on a row in the table in order to display an interactive bar graph of the normalized gene expression in each sample below the table.

Show 10 entries Search: SOCS2

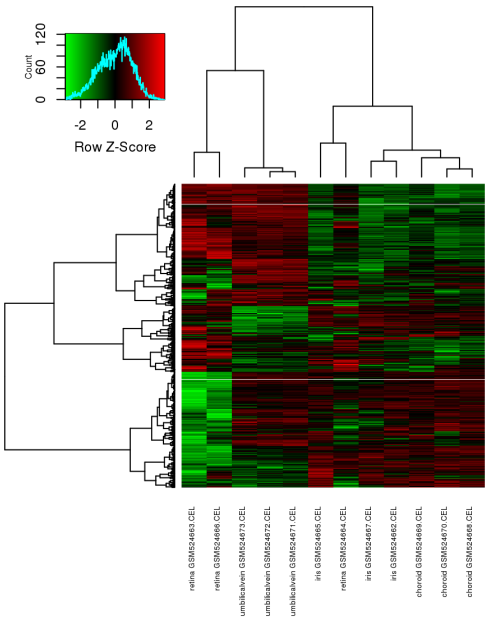
	Symbol	GSM524668.CEL choroid	GSM524669.CEL choroid	GSM524670.CEL choroid	GSM524662.CEL iris	GSM524665.CEL iris	GSM524667.CEL iris
1559315_s_at	SOCS2-AS1	4.84170393659247	4.98010377964211	4.96446873141651	4.72781278451038	5.00448776108882	5.37980930807377
203372_s_at	SOCS2	9.21028653152388	9.20272794982732	9.25167350501548	9.09201011377022	8.89006278877527	9.72797900513616
203373_at	SOCS2	10.6488220924779	10.555230946344	10.4587220301986	10.5481007379352	10.4768606360196	10.9247731764643
232539_at	SOCS2	3.72258585335614	3.74537897094541	3.33098004874651	4.1747409924903	3.89658117137144	3.57225866705772

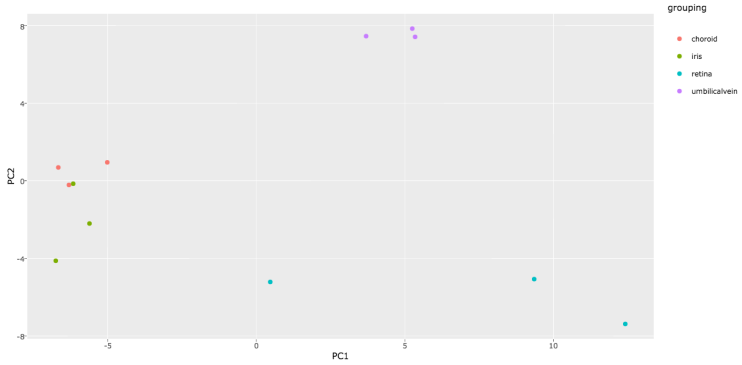
Showing 1 to 4 of 4 entries (filtered from 54,675 total entries) Previous 1 Next

Used RMA normalized CEL files



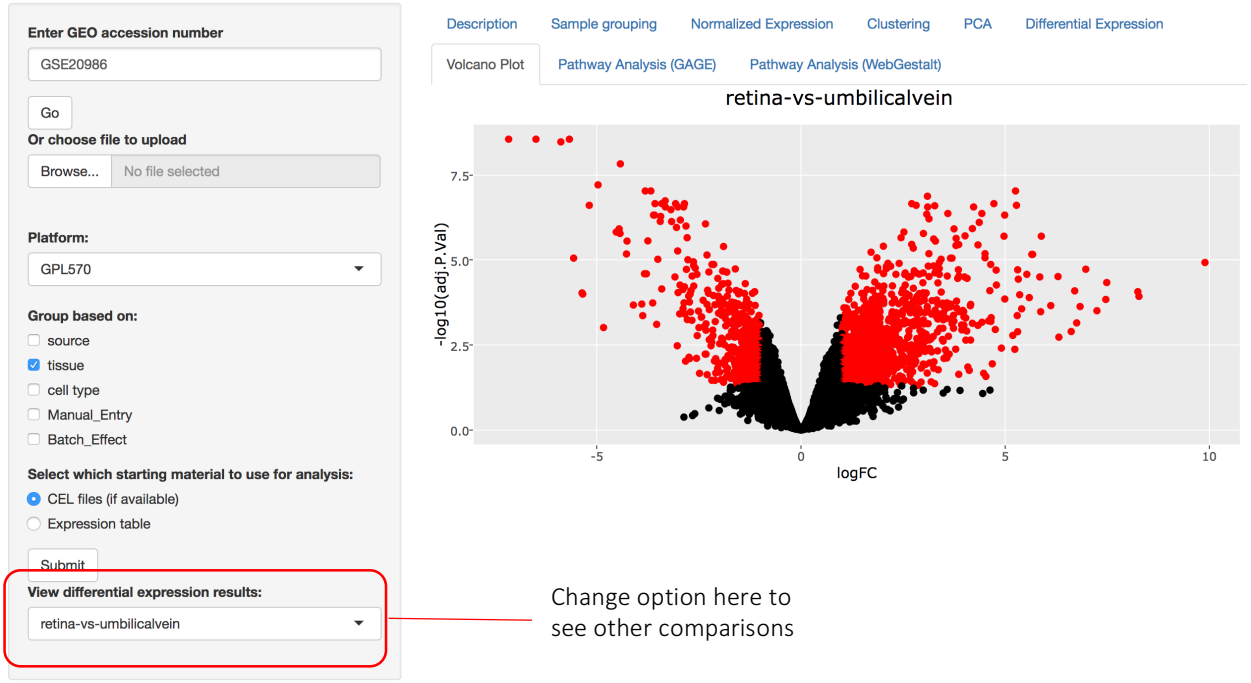
The hierarchical clustering and PCA tabs will show how well the replicates group together.





At this point, BART will have performed differential expression analysis on all combinations of groups. Click on the differential expression tab to see the results of the first comparison. It may take a few seconds to load the differential expression table. You can change which comparison to view on the drop-down menu available on the side panel. All of the tabs including and to the

right of the differential expression tab are sensitive to a change in this option, since they all perform analysis on the selected individual differentially expressed gene list. In this case, there is only one significantly differentially expressed gene between iris and choroid cells. We can change the option to see the differentially expressed genes between retina and umbilical vein tissue. We see that there are over 3,000 differentially expressed entries so we visualize the results in the 'Volcano Plot' tab.



We can also perform GSA with GAGE and ORA with WebGestalt in the corresponding tabs. The GAGE pathway analysis tab will show the enriched KEGG pathways based on the log fold changes for all of the genes while the WebGestalt pathway analysis tab will perform over-representation analysis with KEGG using the significantly differentially expressed gene list. The WebGestalt tab will also show a GOSlim summary of the differentially expressed genes. All plots and tables are available for download.

Using BART with a personal gene expression table:

Simply upload a table with the first column corresponding to gene names or identifiers, the first row corresponding to sample names, and the rest of the cells corresponding to an expression value (normalized or un-normalized) for each gene in each sample. If you are unsure about the formatting of this table, download the example expression table found on the BART description page. If your upload file size is too large, try compressing the file with gzip first. You must use manual entry to enter groups and batches (if necessary). The resulting analysis will be the same. Currently, we do not support pathway analysis for user uploaded tables.