Expression Atlas: Quick tour

Maria Keays [1]
Gene Expression
Beginner
0.5 hour


An undergraduate-level understanding of biology is an advantage.

This course was last updated in September 2016.

Learning objectives:

- Basic understanding of what the Expression Atlas and how it can be used
- Know where to find more information about the Expression Atlas

What is Expression Atlas?

The Expression Atlas [3] ('Atlas') provides information on gene expression [4] patterns under different biological conditions. Gene expression data is re-analysed in-house to detect genes showing interesting baseline and differential expression patterns. The results of these analyses can be searched using gene annotations and/or sample attributes.

What can I do with it?

You can search using gene annotations to find out about the expression of your favourite gene(s) in normal, untreated conditions, and also in various experimental conditions. You can also search using sample attributes, to find out which genes are expressed in a particular tissue, e.g. human kidney, or which genes show interesting expression patterns under a certain experimental condition, e.g. after a gene knock-out, or in a particular disease.

Why do we need it?

Atlas exists to provide easy and free access to the results of our in-house analyses of a wide range of gene expression data. Its main aim is to allow biologists to find out more about their genes of interest, and/or to discover gene expression patterns underlying the biological phenomenon that they study. Atlas could act as a starting point for researchers gathering information prior to conducting their own experiments, or as a means of corroborating existing research results.
Where does the data come from?

All data in Expression Atlas are drawn from the ArrayExpress [5] archive of functional genomics [6] data (Figure 1). ArrayExpress stores and accesses data from all kinds of functional genomics experiments performed in laboratories around the world. The data files are stored along with metadata [7] describing the biological samples and protocols used.

Figure 1 The data in Expression Atlas are derived from ArrayExpress. ArrayExpress is a database of functional genomics experiments that are either submitted by researchers or imported from GEO. The experiments are manually curated and re-analysed for inclusion in Expression Atlas.

Expression Atlas curators select eligible datasets and manually curate them to a high standard, annotating the metadata using the Experimental Factor Ontology [8] (EFO). Currently, only gene expression microarray and RNA-seq datasets are selected for the Atlas. The data are then analysed using the Expression Atlas in-house automated pipeline.

Navigating and accessing the homepage

From the Expression Atlas homepage [3] (Figure 2), you can search across Atlas using gene annotations (e.g. identifiers, A) and/or sample properties (e.g. experimental conditions, B). You can also browse through the data sets analysed in Atlas (C), and check out the help pages [9] (D) if you get stuck.
Each experiment in Expression Atlas is classed as being either "baseline" or "differential". We will look at what this means on the next few pages.

**Baseline experiments**

Baseline experiments show gene expression [4] levels in "normal", untreated conditions. They are intended to answer questions such as "Which genes are expressed in a normal human lung?" [10]. The majority of baseline experiments are RNA-seq [11] based, and gene expression levels are represented by FPKM [12] values.

**The baseline experiment view page**

You can browse baseline experiments by clicking the link on the homepage (Figure 3).
Figure 3 Browsing baseline experiments in Expression Atlas

Clicking on a baseline experiment name takes you to the baseline experiment overview page (Figure 4). If you come from a gene page, only that gene will be shown. Click "Reset" (A) to show all the genes.

Gene expression levels are shown in the heatmap [13] table, with highest expressed genes at the top. Each row represents a gene, and each column represents an experimental condition, e.g. a tissue or a cell type. You can filter the heatmap to show certain genes by typing gene annotations in the "Gene query" box (B), and clicking "apply". You can also change the minimum expression level shown using the slider (C).
Differential experiments

Many gene expression experiments are conducted with a view to comparing some "abnormal" condition against an appropriate "normal" condition, for example, comparing maize plants grown in drought conditions against those grown in normal watering conditions [14]. Experiments of this type are termed "differential", as they are looking at the difference in the expression level of each gene in the two conditions. The expression levels are measured, and then a statistical test is performed for each gene, to determine whether its expression level is significantly different between the two conditions. Expression Atlas has both microarray and RNA-seq differential experiments.

The differential experiment view page
Like baseline experiments, differential experiments can be browsed from the homepage. The differential experiment overview page (Figure 5) is very similar to the baseline one. The main difference is that for differential expression experiments, the heatmap [13] summarises the difference in expression level between two conditions for each gene, rather than each gene's expression level. Again, each row in the heatmap represents a gene.

The genes with the largest difference in expression between the two conditions in the comparison, measured as log2 fold-change, are shown at the top. Each column in the heatmap represents a comparison between two conditions in the experiment. Some experiments only have one comparison (hence one column), while others may have more than this.

Figure 5 The differential expression experiment page.

You can filter the heatmap to only show genes that you are interested in (A). You can also change the thresholds for the maximum adjusted p-value [15] (B) and the minimum log2 fold-change shown (C), to increase or decrease the number of results shown. You can also hover over the comparisons for more information about the samples or over the heatmap to see a summary of the statistics for a given comparison and gene.

Please read our help pages [9] for more information.

Searching Expression Atlas

You can search Expression Atlas in a number of ways:

- Searching by gene [16];
- Searching by condition [17];
- Combining searches [18].
Searching by gene

The Atlas homepage - searching for a gene

To find out about the expression of your gene(s) of interest, type gene names, identifiers, and/or annotations into the "Gene query" box (Figure 6A). Suggestions will appear as you type to help you, though you are free to ignore these if you wish. You can also limit your search to within one species, by selecting one from the "Organism" dropdown menu (Figure 6B).

![Search for a gene in Expression Atlas.](Illustration)

Click "Search" to see your results.

The gene view page

You can see all information about a gene in Atlas on its gene page (Figure 7). This page consists of three tabs for baseline expression, differential expression and search term information. The name of the search term information tab matches the bioentity that you searched for (BRCA1 information in the example below).
Figure 7 Navigate the gene view page using the tabs at the top of the page.

The **baseline expression tab** (Figure 8) summarises all baseline expression data in Atlas for the gene. To zoom in on the heatmap, click and drag left/right, or tap with two fingers and pinch. Clicking on an experiment name will take you to the experiment overview page in Atlas.

Figure 8 The baseline expression tab on the geneview page.

The **differential expression tab** (Figure 9) summarises all differential expression data in Atlas for the gene. Clicking on a comparison name will take you to the experiment overview page in Atlas.
The **search term information tab** (Figure 10) contains information about the gene that is pulled in from multiple sources, including Ensembl [19], the Gene Ontology [20], InterPro [21] and UniProt [22].

**Figure 9** The differential expression tab on the gene view page

**Figure 10** The search term information tab on the gene view page

**Searching by condition**

**Searching by condition**
From the Expression Atlas homepage [23] you can also search by experimental conditions and/or sample properties (Figure 11). Type these in the "Sample properties" search box (A) and click "Search". You can also optionally select a species from the "Organism" dropdown menu (B) to limit your search to experiments about that species only.

**Figure 11** Search for experimental conditions in Expression Atlas.

**Multi-experiment search results page**

If your search returns only one experiment you will be taken directly to the experiment overview page.

If your search returns results about more than one experiment, you will be taken to a page with two tabs (Figure 12). This is similar to the gene overview page, but without the tab containing search term information. Here, the baseline expression tab (Figure 12) contains information about baseline expression experiments matching your query.
**Figure 12** Searching by condition and the baseline expression results tab.

The differential expression tab (Figure 13) contains information about differential expression experiments that were found.
Figure 13 Searching by condition and the baseline expression results tab.

Similarly to the gene overview page, clicking on the name of a baseline experiment or comparison will take you to the experiment overview page.

Combining searches

You can also search for a combination of gene(s) and condition(s), e.g. searching for expression of the H19 [24] gene in human liver (Figure 14).
Figure 14 Combining a gene and condition search.

When you click "Search", you will be taken to a two-tab search results page summarising the baseline and differential expression results matching your query (Figure 15).

Figure 15 Results page for a search combining gene and condition.
Summary

- Expression Atlas contains baseline and differential gene expression data from a variety of organisms across many different experimental conditions.
- You can search Atlas to find out more about the normal expression of your gene(s) of interest.
- You can find out about the behaviour of your gene(s) of interest in different experimental conditions.
- Atlas can tell you which genes are normally expressed in a particular tissue or cell type.
- It can also help you to find out which genes are affected by certain experimental conditions.

Your feedback

Please tell us what you thought about this course. Your feedback is invaluable and helps us to improve our courses and thus enhance your learning experience.

Learn more

Find out more

- Expression Atlas is developed and maintained by the Functional Genomics team at EMBL-EBI, led by Alvis Brazma [26].

Recommended courses

- All data in Expression Atlas is taken from ArrayExpress [5]. To learn more about it, try the ArrayExpress Quick Tour [27].
- The EBI runs courses about gene expression throughout the year. Please check the listings [28] for more information.

Recommended reading


Get help and support on Expression Atlas

- Please check out our help pages [9] for more information.
- Still stuck? Send us an email [29] and we'll do our best to help you out!
Contributors

Maria Keays [1]
EMBL-EBI
Bioinformatician - Brazma team: Functional Genomics

Maria studied Biology at the University of Bath, going on to an MSc in Bioinformatics at the University of Leicester followed by a PhD in evolutionary biology at the University of St. Andrews. She joined the Functional Genomics group at EMBL-EBI in 2011, and works on various projects for ArrayExpress and Expression Atlas.

Melissa Burke [30]
EMBL-EBI
Scientific Training Officer (e-learning): Training Team

Melissa is the Scientific Training Officer (e-learning) for the Training Team at the EMBL-EBI. She joined the Training Team in July 2016 after having worked as a Scientific Curator for ArrayExpress/Expression Atlas at the EMBL-EBI. She has a PhD in Molecular Parasitology and has worked internationally as a postdoctoral researcher specialising in the functional genomics of infectious diseases.

All course materials in Train online are free cultural works licensed under a Creative Commons Attribution-ShareAlike 4.0 International license
Source URL: http://www.ebi.ac.uk/training/online/course/expression-atlas-quick-tour-0

Links
[1] http://www.ebi.ac.uk/training/online/trainers/mkeays
[2] https://www.ebi.ac.uk/gxa/home
[7] http://www.ebi.ac.uk/training/online/glossary/metadata
[14] http://www.ebi.ac.uk/gxa/experiments/E-GEOD-16567?accessKey=&serializedFilterFactors=&queryFactorType=&rootContext=&heatmapMatrixSize=50&displayLevels=false&displayGeneDistribution=false&arrayDesignAccession=&geneQuery=&exactMatch=true&exactMatch=on&regulation=UP_DOWN&queryFactorValues=g2_g7&queryFactorValues=1&specific=on&cutoff=0.05&foldChangeCutOff=1.0
[15] http://www.ebi.ac.uk/training/online/glossary/p-value
[21] https://www.ebi.ac.uk/interpro/
[23] http://www.ebi.ac.uk/gxa
[26] https://www.ebi.ac.uk/about/people/alvis-brazma
[27] http://www.ebi.ac.uk/training/online/course/arrayexpress-quick-tour-1
[28] https://www.ebi.ac.uk/training/
[29] mailto:arrayexpress-atlas@ebi.ac.uk