

Supplementary Data

Title: Defining developmental potency and cell lineage trajectories by expression profiling of differentiating mouse embryonic stem cells

Authors: Kazuhiro Aiba, Timur Nedorezov, Yulan Piao, Akira Nishiyama, Ryo Matoba, Lioudmila V. Sharova, Alexei A. Sharov, Shinya Yamanaka, Hitoshi Niwa, and Minoru S. H. Ko

Supplementary Data:

Examples of gene expression patterns (Pages 3 – 10 in this document).

Additional Supplementary Data Legends:

Supplementary Table S1: A list of genes and their expression patterns for primitive endoderm lineage. The explanation of legends for each column can be found in the help file of the NIA Array Analysis software (<http://lgsun.grc.nia.nih.gov/ANOVA/help.html>). In brief, “logChange” is the degree of gene expression change, which is determined as a slope of regression of log-transformed gene expression level for a gene onto the cell coordinates on each cell lineage trajectory; and “logExpression (logExpr)” is simply an average log-transformed gene expression level for a gene in cells. In the Excel file, genes are ordered by the product of logChange and logExpr (i.e., logChange x logExpr), because the importance of genes associated with each lineage trajectory depends both on the magnitude of change of their expression levels in the course of differentiation and on the average expression (i.e., highly-expressed genes are more important than lowly-expressed genes).

Supplementary Table S2: A list of genes and their expression patterns for trophoblast lineage. See the legend for Supplemental Information 3.

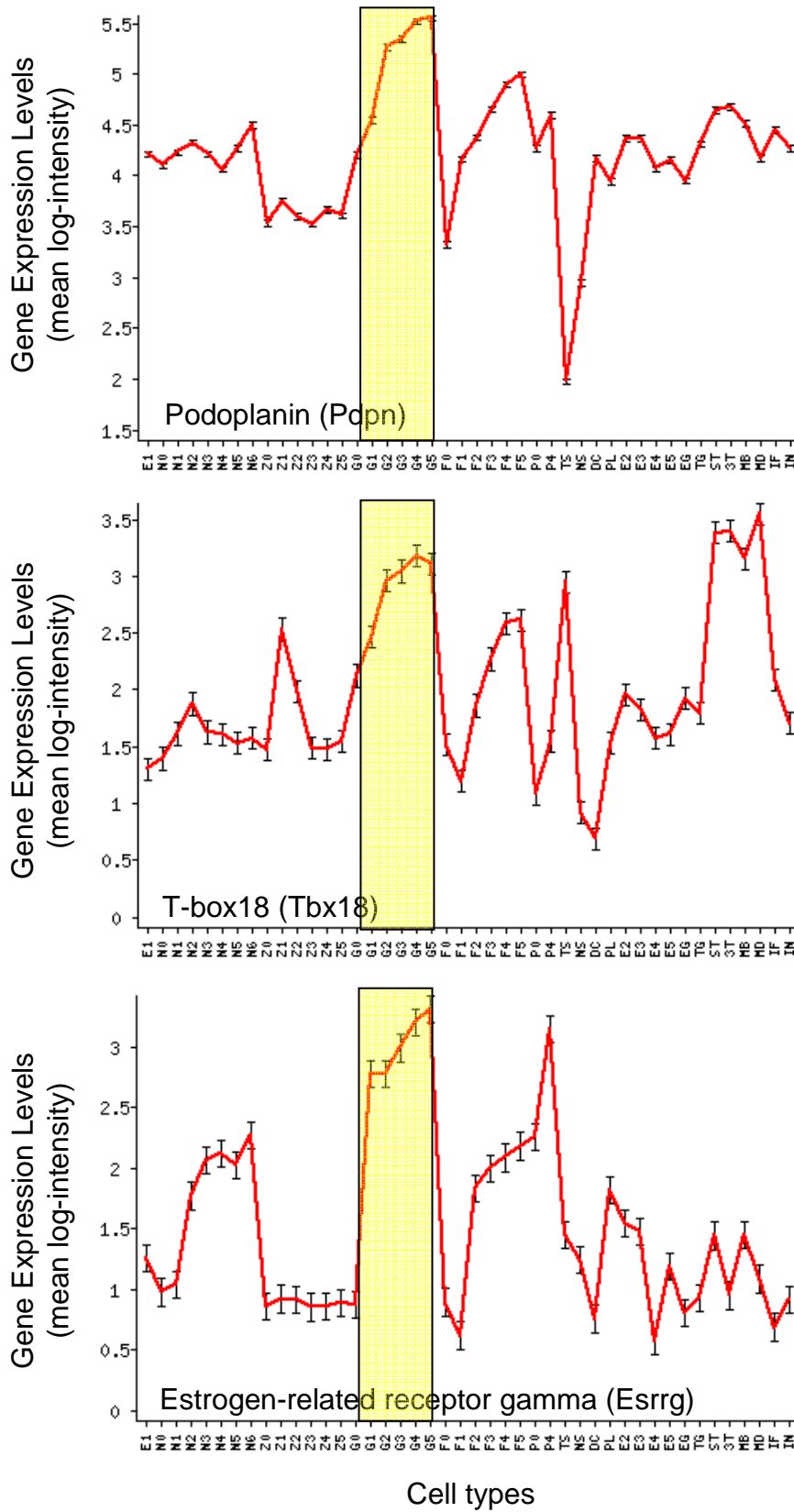
Supplementary Table S3: A list of genes and their expression levels for primitive ectoderm/neural lineage. See the legend for Supplemental Information 3.

Supplementary Video S1: A movie file (.avi format) is available at <http://lgsun.grc.nia.nih.gov/data/Publications-Supplemental-download.html>. This provides a view of 3D-PCA figure from different angles by rotating a Figure 1b. (Tested for Quicktime or Windows Media Player on PC or Macintosh.). A virtual reality modeling language (vrmf) file is also available at <http://lgsun.grc.nia.nih.gov/data/Publications-Supplemental-download.html>. This provides yet another view of 3D PCA figure (Fig. 1b). You can rotate and zoom-in 3D PCA figure. To open this file, you need to download a free viewer/client for vrmf file: either Cortona by Parallel Graphics (<http://www.parallelgraphics.com/products/cortona/>) or Cosmo player by Karmanaut (<http://www.karmanaut.com/cosmo/player/>). Once one of the clients is installed on your computer, you open a web browser (e.g., Internet Explorer) and open the vrmf file from the web browser.

Supplementary Video S2: A movie file (.avi format) is available at <http://lgsun.grc.nia.nih.gov/data/Publications-Supplemental-download.html>. This provides a view of 3D-PCA figure from different angles by rotating a Figure 3b. (Tested for Quicktime or Windows Media Player on PC or Macintosh.). A virtual reality modeling language (vrmf) file is also available at <http://lgsun.grc.nia.nih.gov/data/Publications-Supplemental-download.html>. This provides yet another view of 3D PCA figure (Fig. 3b). See the legend for Supplementary Video S1.

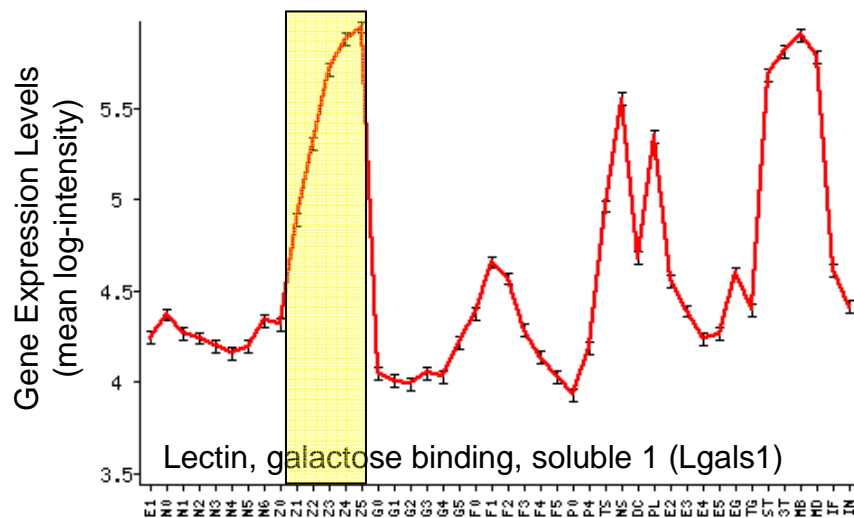
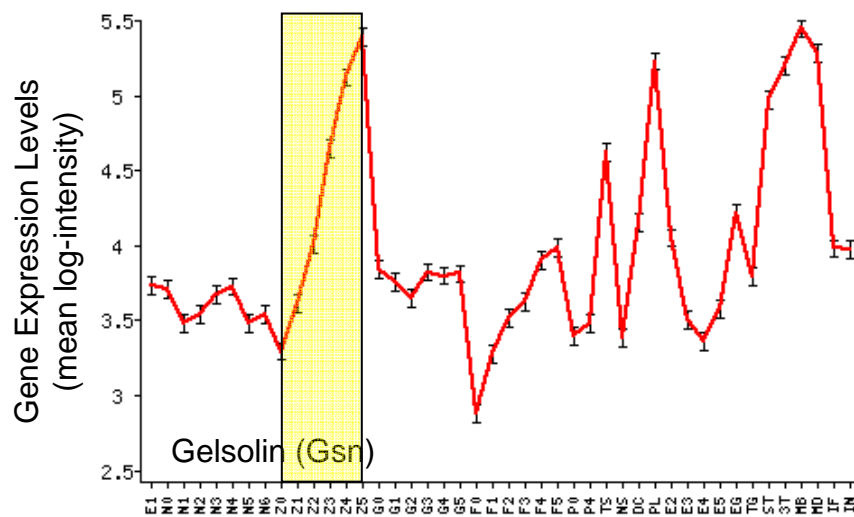
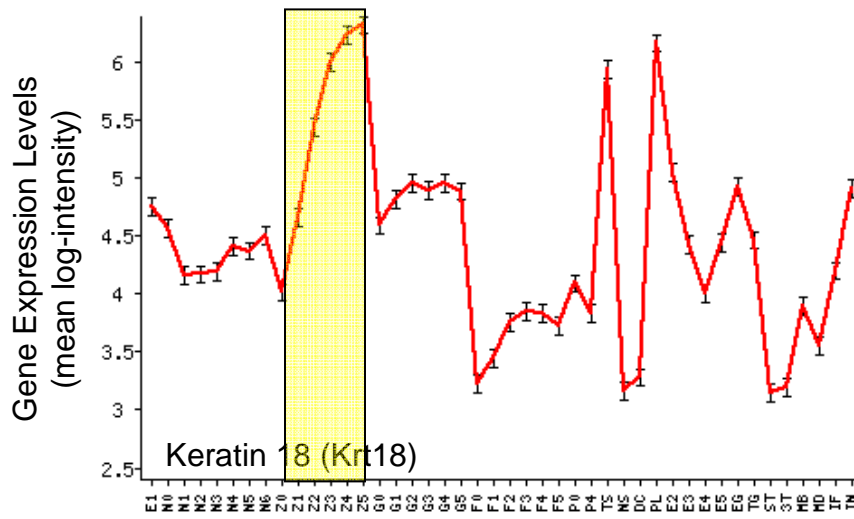
Supplementary Data: Examples of gene expression patterns

(Page 3) Examples of genes whose expression changes positively correlate with the primitive endoderm trajectory



Supplementary Data: Examples of gene expression patterns

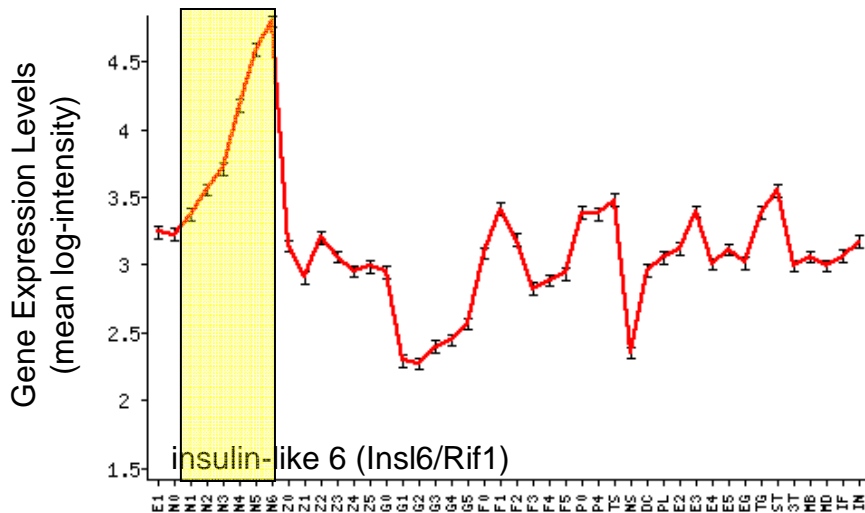
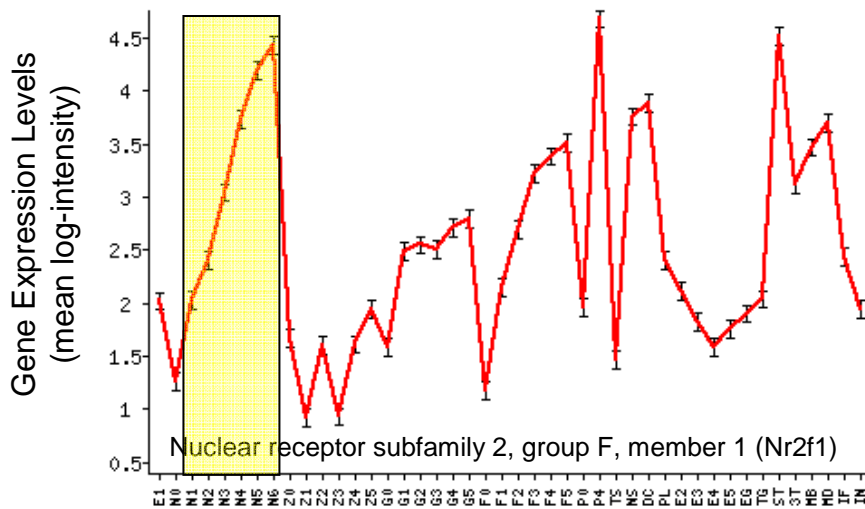
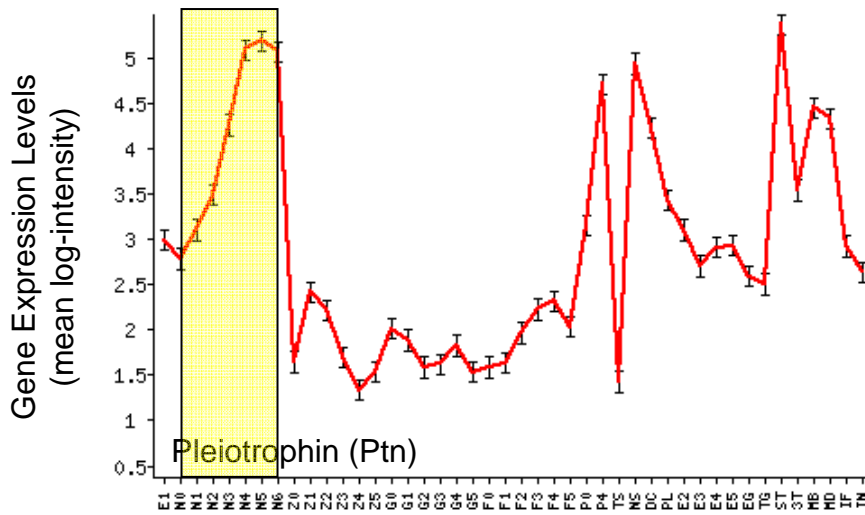
(Page 4) Examples of genes whose expression changes positively correlate with the trophoblast trajectory



Cell types

Supplementary Data: Examples of gene expression patterns

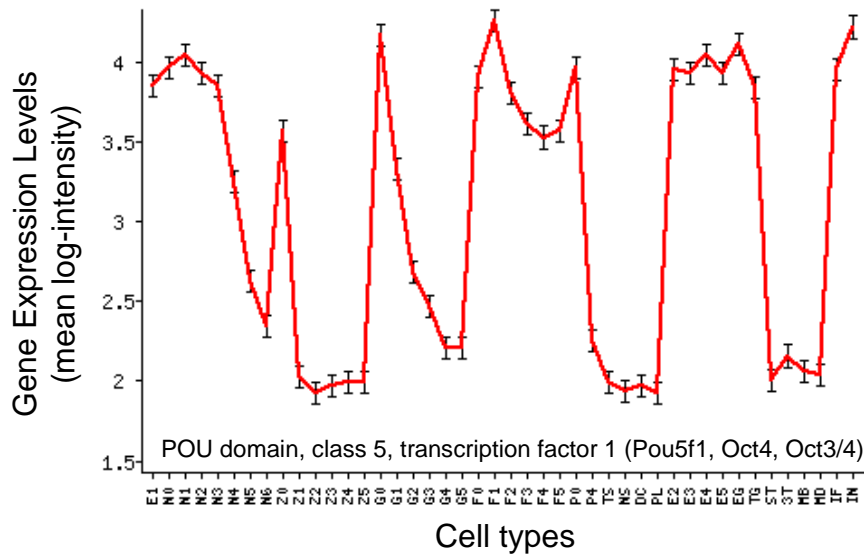
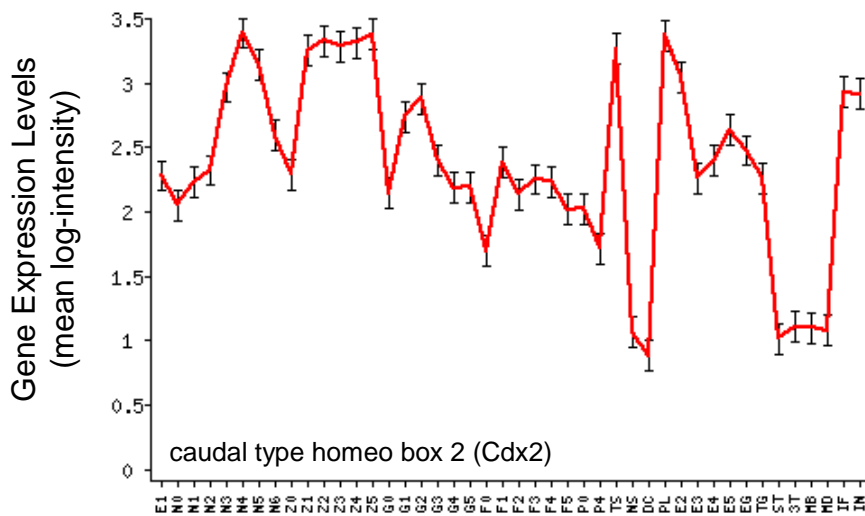
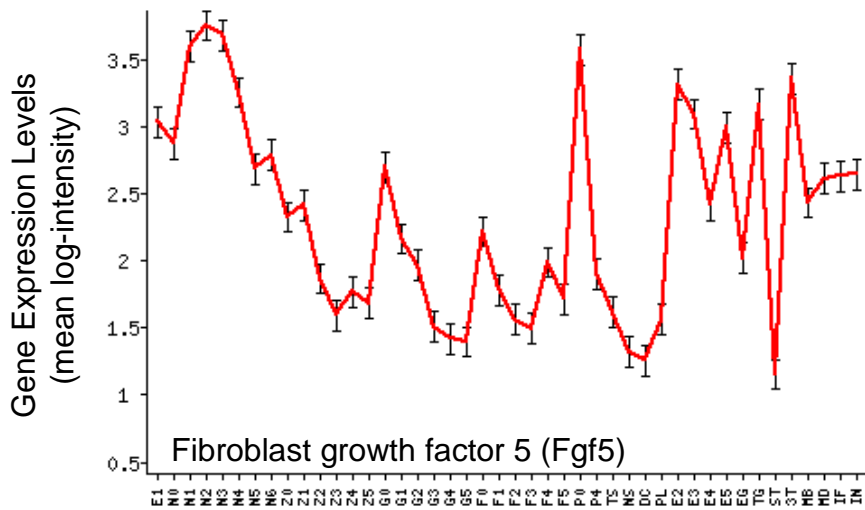
(Page 5) Examples of genes whose expression changes positively correlate with the primitive ectoderm/neural ectoderm trajectory



Cell types

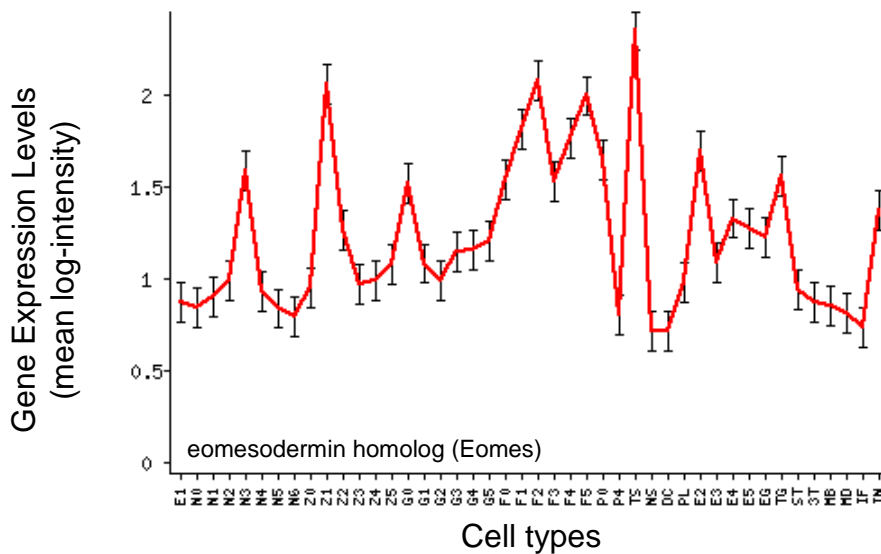
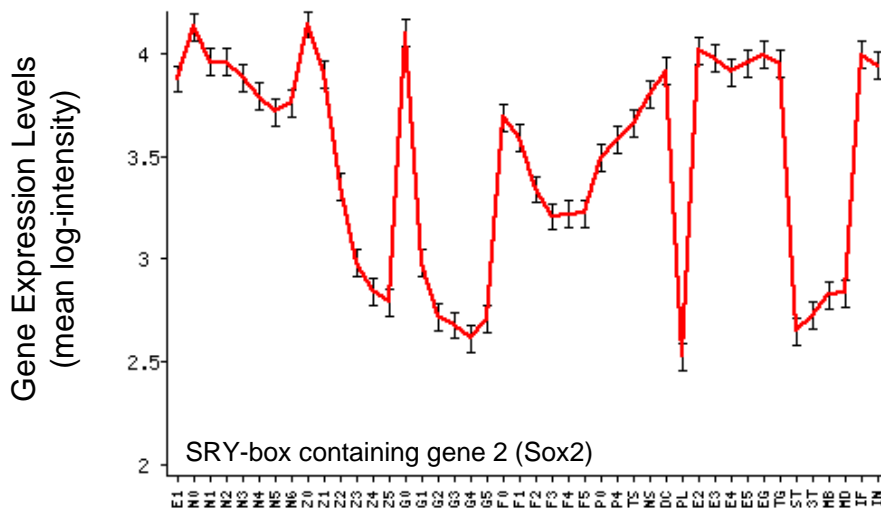
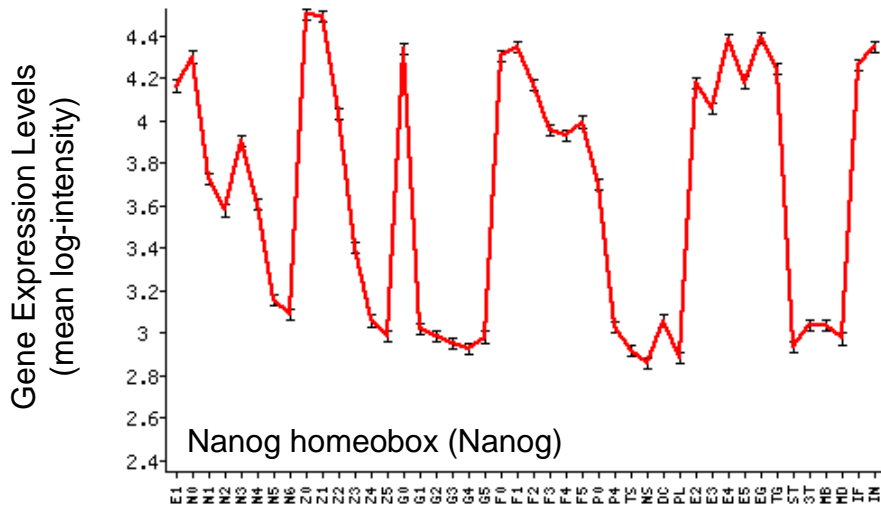
Supplementary Data: Examples of gene expression patterns

(Page 6) Examples of other notable genes



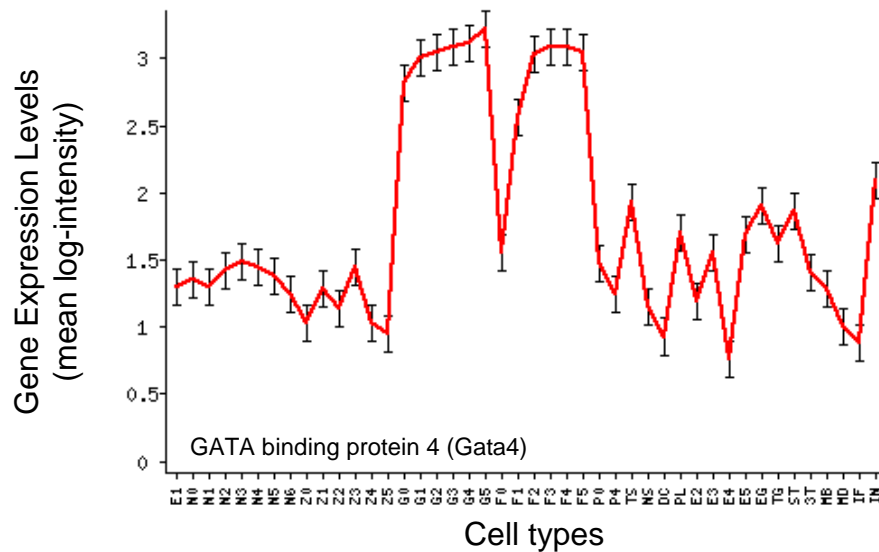
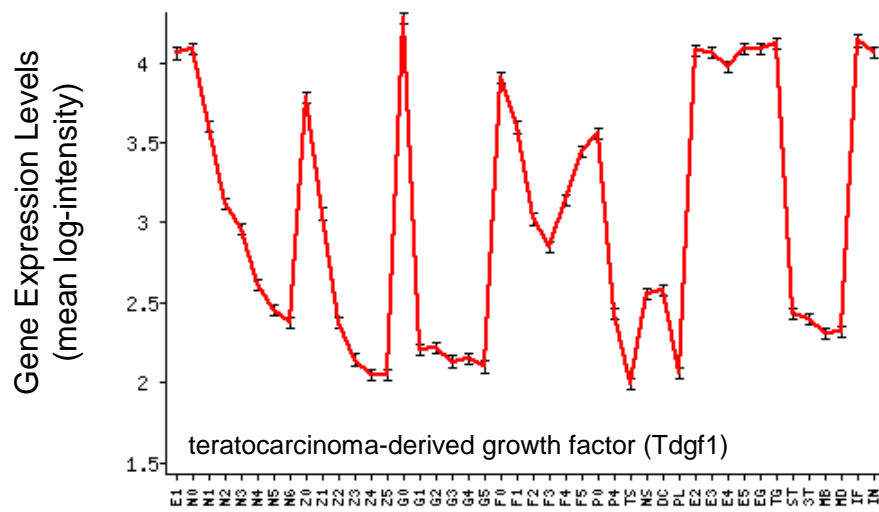
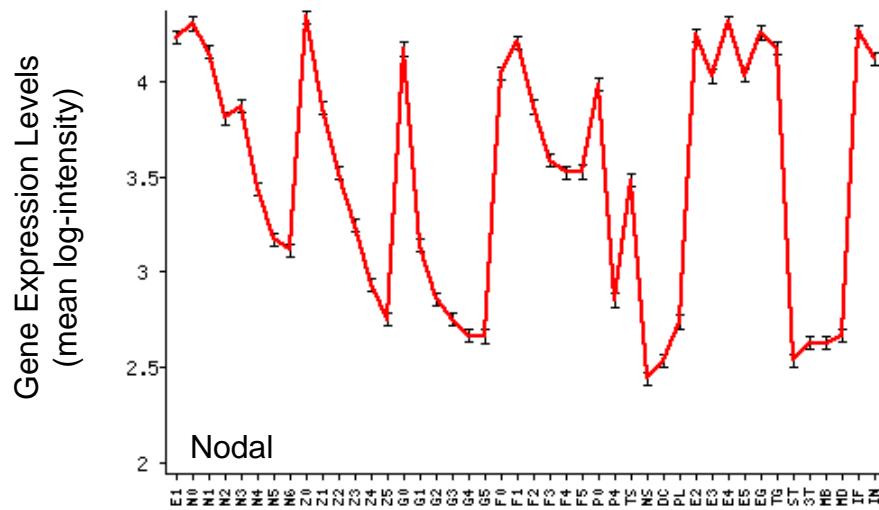
Supplementary Data: Examples of gene expression patterns

(Page 7) Examples of other notable genes



Supplementary Data: Examples of gene expression patterns

(Page 8) Examples of other notable genes



Supplementary Data: Examples of gene expression patterns

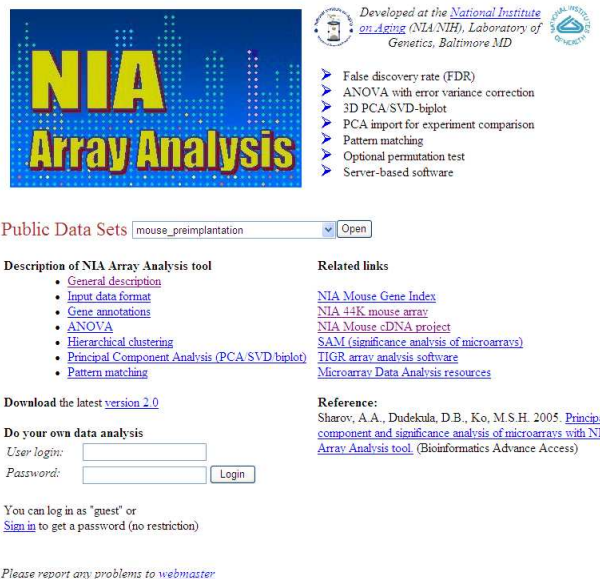
(Page 9) How to search for the expression patterns of a specific gene.

Method 1: Download the Supplemental Information 7, “All DNA microarray data” from the journal web site. Open the file by Microsoft Excel. Use the search function of Excel to search a specific gene name, e.g., Pou5f1. Select a row containing all the gene expression levels in different cell types and use the plot making function of Excel to visualize the data.

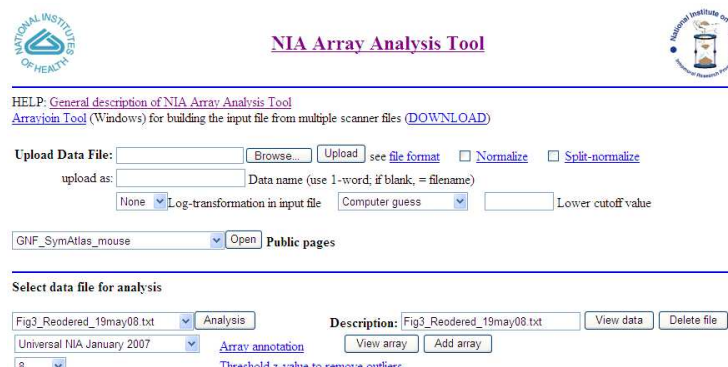
Method 2: Use the public website of NIA Array Analysis tool.

Step 1: Open <http://lgsun.grc.nia.nih.gov/ANOVA/> by an web browser.

Step 2: Login name: Trajectory Password: Trajectory1
Note the first letter should be uppercase.



Step 3: Choose the file, “Fig3_Reodered_19may08.txt.” Click “Analysis.”



Supplementary Data: Examples of gene expression patterns

(Page 10) How to search for the expression patterns of a specific gene.

Step 4: Type a gene symbol, e.g., “Oct4” or “Pou5f1.” Gene aliases should work, but official gene symbols (<http://www.informatics.jax.org/> or <http://www.ncbi.nlm.nih.gov/>) will work better.

The screenshot shows the NIA Array Analysis Tool interface. At the top, there are logos for the National Institutes of Health and the National Institute on Aging. Below the logos, there is a button labeled "Do another analysis". The main heading is "Results: Fig3_Reordered_19may08.txt". Underneath, there are links for "Results" and "ANOVA". There are three "Download" buttons corresponding to "Input file", "Means (significant genes only)", and "TIGR-MEV file (Stanford format)". A search bar is visible with a red arrow pointing to it. At the bottom, there is a link for "Principal Component Analysis (PCA)".

Step 5: Click “Hist” of a correct gene symbol.

The screenshot shows the NIA Array Analysis Tool interface with search results for "Oct4". The heading is "Search Results for 'Oct4'". Below the heading is a table with columns: Plot, Probe id, Symbol, Average log-intensity, Error Variance, F, P, FDR, Annotation, GeneIndex 'U' Cluster, RefSeq Accession, GenBank Accession, MGI, OtherGeneSymbol (s), and Comments. A red arrow points to the "Hist" button in the first row of the table.

Plot	Probe id	Symbol	Average log-intensity	Error Variance	F	P	FDR	Annotation	GeneIndex 'U' Cluster	RefSeq Accession	GenBank Accession	MGI	OtherGeneSymbol (s)	Comments
Hist	Z00074951-1	Pou5f1	3.0741	0.00958	168.445	0.00000	0.00000	POU domain, class 5, transcription factor 1	U017906	NM_013633.1	NM_013633.1	MGI:101893	11	

Step 6: You will see the expression profiles in the next window.

The screenshot shows the NIA Array Analysis Tool interface with the expression profile for Z00074951-1. The heading is "Expression of Z00074951-1". Below the heading is a table with columns: Probe id, GeneSymbol, Annotation, GeneIndex 'U' Cluster, RefSeq Accession, GenBank Accession, MGI, OtherGeneSymbol (s), and Comments. Below the table is a line graph titled "Mean log-intensity" showing the expression profile across various tissues. A red arrow points to the graph. To the right of the graph is a table with columns: Tissue, Log-mean, and SD.

Probe id	GeneSymbol	Annotation	GeneIndex 'U' Cluster	RefSeq Accession	GenBank Accession	MGI	OtherGeneSymbol (s)	Comments
Z00074951-1	Pou5f1	POU domain, class 5, transcription factor 1	U017906	NM_013633.1	NM_013633.1	MGI:101893	11	

Tissue	Log-mean	SD
E1	3.8459	0.0692
N0	3.9593	0.0692
N1	4.0395	0.0692
N2	3.9265	0.0692
N3	3.8453	0.0692
N4	3.2463	0.0692
N5	2.6214	0.0692
N6	2.339	0.0692
Z0	3.5627	0.0692
Z1	2.0297	0.0692
Z2	1.9202	0.0692
Z3	1.9709	0.0692
Z4	1.8035	0.0692