



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Datasets of genes coexpressed with FBN1 in mouse adipose tissue and during human adipogenesis



Margaret R. Davis^{a,1}, Erik Arner^b, Cairnan R.E. Duffy^{c,2},
Paul A. De Sousa^c, Ingrid Dahlman^d, Peter Arner^d,
Kim M. Summers^{a,*}

^a The Roslin Institute, University of Edinburgh, Easter Bush, EH25 9RG, UK

^b RIKEN Center for Life Science Technologies (Division of Genomic Technologies) (CLST (DGT)),
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

^c Centre for Clinical Brain Sciences, University of Edinburgh, Chancellors Building, 49 Little France Crescent,
Edinburgh, EH16 4SB, UK

^d Department of Medicine, Huddinge (Med H), Karolinska Universitetssjukhuset Huddinge,
141 86 Stockholm, Sweden

ARTICLE INFO

Article history:

Received 21 June 2016

Received in revised form

28 June 2016

Accepted 29 June 2016

Available online 5 July 2016

ABSTRACT

This article contains data related to the research article entitled “Expression of *FBN1* during adipogenesis: relevance to the lipodystrophy phenotype in Marfan syndrome and related conditions” [1]. The article concerns the expression of *FBN1*, the gene encoding the extracellular matrix protein fibrillin-1, during adipogenesis *in vitro* and in relation to adipose tissue *in vivo*. The encoded protein has recently been shown to produce a short glucogenic peptide hormone, (Romere et al., 2016) [2], and this gene is therefore a key gene for regulating blood glucose levels. *FBN1* and coexpressed genes were examined in mouse strains and in human

DOI of original article: <http://dx.doi.org/10.1016/j.ymgme.2016.06.009>

* Corresponding author.

E-mail address: kim.summers@roslin.ed.ac.uk (K.M. Summers).

¹ Present address: Department of Pathology and Cell Biology, Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, QC, Canada H3C 3J7

² Present address: School of Chemistry, EaStCHEM, University of Edinburgh, Joseph Black Building, West Mains Road, Edinburgh, EH9 3JJ, UK

<http://dx.doi.org/10.1016/j.dib.2016.06.055>

2352-3409/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

cells undergoing adipogenesis. The data show the genes that were coexpressed with *FBN1*, including genes coding for other connective tissue proteins and the proteases that modify them and for the transcription factors that control their expression. Data analysed were derived from datasets available in the public domain and the analysis highlights the utility of such datasets for ongoing analysis and hence reduction in the use of experimental animals.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

Specifications Table

Subject area	<i>Biology</i>
More specific subject area	<i>Formation of adipose tissue from mesenchymal stem cells and adipocyte precursors</i>
Type of data	<i>Figures, Excel file</i>
How data was acquired	<ol style="list-style-type: none"> 1. A list of mouse genes that were correlated with <i>Fbn1</i> was derived from BioGPS (http://biogps.org, accessed 18 June 2015) using the “Correlation” tab, with a minimum correlation coefficient for inclusion of 0.70 2. Human data for a time course of adipogenesis in the FANTOM5 project (http://fantom.gsc.riken.jp/zenbu/) were downloaded on 8 May 2015 and clustered using BioLayout Express^{3D} Version 3.2 (available for download at http://www.biobioinformatics.org). Clusters of coexpressed genes are presented in order of cluster size 3. Human data for expression of 1380 transcription factor genes during the adipogenesis time course were downloaded from the FANTOM5 project (http://fantom.gsc.riken.jp/zenbu/) on 28 May 2015 and clustered using BioLayout Express^{3D} (http://www.biobioinformatics.org). Clusters of coexpressed genes are presented in order of size
Data format	<i>Analysed</i>
Experimental factors	<p>The mouse analysis used data from a previous study [3] in which RNA was extracted from epididymal adipose tissue from 28 inbred mouse strains. Mice were 25 weeks of age. RNA from two to four mice per strain was pooled</p> <p>The human analysis used data from the FANTOM5 project [4]. Human mesenchymal stem cells were cultured in the presence of adipogenic medium for up to 14 days [5–7]. Expression results for time points at induction of differentiation and 3 h, 1, 2, 4, 8, 12 and 14 days after induction of differentiation were downloaded for this analysis. Further details are available at http://fantom.gsc.riken.jp/5/sstar/MSc_to_adipocyte_(human).</p>
Experimental features	<ol style="list-style-type: none"> 1. Mouse data. The record for the mouse <i>Fbn1</i> gene was accessed at http://biogps.org [8]. The dataset chosen was eQTL – Fat (GNF1M). The probeset was <i>gnf1m00711_a_at</i>. Under the “Correlation” tab, a correlation cutoff of 0.75 was chosen. The list of genes was downloaded and is presented in the spreadsheet

2. Human data. Gene-based expression levels using promoter analysis by cap analysis of gene expression (CAGE) [9] were obtained from <http://fantom.gsc.riken.jp/zenbu/> for the mesenchymal stem cell to adipocyte differentiation experiment. 7633 genes showing non-zero expression in at least one time point were analysed using BioLayout Express3D with a correlation coefficient cutoff of 0.85 and an MCL inflation value of 2.2 [10,11]
3. Human transcription factor data. Gene-based expression levels based on promoter analysis by CAGE were obtained from <http://fantom.gsc.riken.jp/zenbu/> for the mesenchymal stem cell to adipocyte differentiation experiment. Expression patterns for 1380 genes were clustered with a correlation coefficient cutoff of 0.62 and an MCL inflation value of 2.2

Data source location	n/a
Data accessibility	The data are included with this article

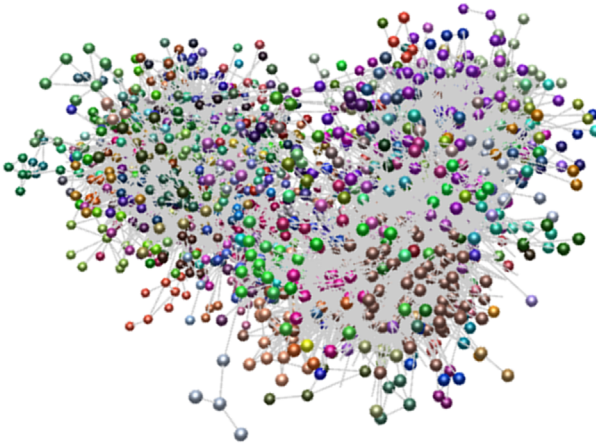
Value of the data

- Adipogenesis involves the downregulation of generic mesenchymal extracellular matrix (ECM) genes and upregulation of those genes encoding proteins specific to adipose tissue. The data provide a comprehensive list of the genes whose expression is altered during adipogenesis, which will assist in finding gene networks that are up- and downregulated during this process.
- The mouse data focus on fibrillin-1, the precursor of a newly described glucogenic hormone. In the mouse, *Fbn1* mRNA varies between strains and the data present a list of genes whose expression is correlated with that of *Fbn1* *in vivo* (in mice). These data provide insight into the function of fibrillin-1 in adipogenesis and can be used also to determine the appropriate mouse strains for further experiments on obesity.
- The data also show how transcription factors are altered during the process of adipogenesis in human cells, which allows researchers to determine the regulatory molecules that are likely to control expression of key genes.
- These data will be useful in determining key genes downregulated during adipogenesis and can be compared with other experiments and *in vivo* data from humans and other animals.

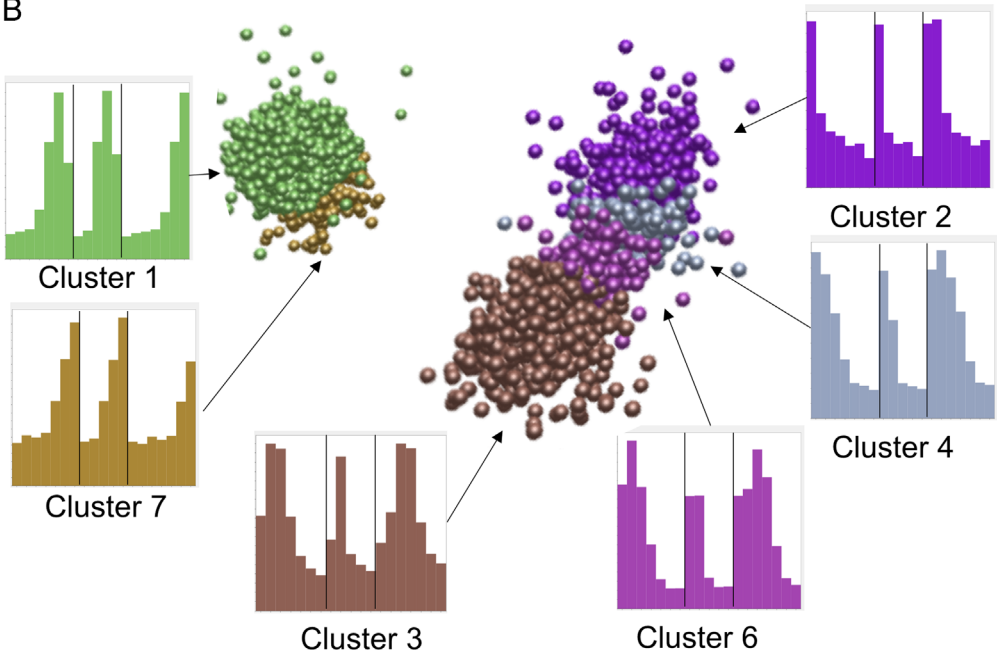
1. Data

The main data consist of gene lists. The first list contains genes whose expression was correlated with that of *Fbn1* in different mouse strains. The second list shows all the clusters derived by BioLayout Express^{3D} for the differentiation of mesenchymal stem cells to adipocytes. Each cluster represents a different pattern of expression, which is described in the table. The patterns show a transition from generalised mesenchymal extracellular matrix to a specialised adipose extracellular matrix, with high expression of mitochondrial and lipid processing genes by the end of the time course. The third list shows the clusters of transcription factors derived from the same experiment. A similar transition in transcription factor expression was observed as adipogenesis progressed.

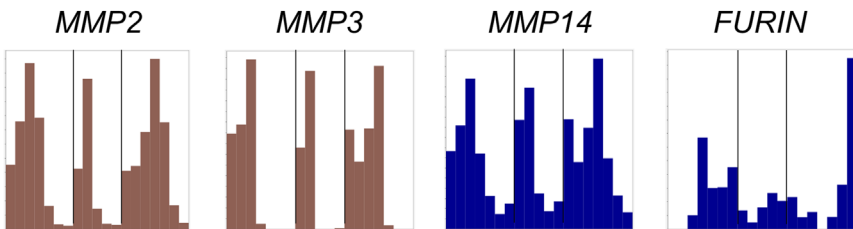
A



B



C

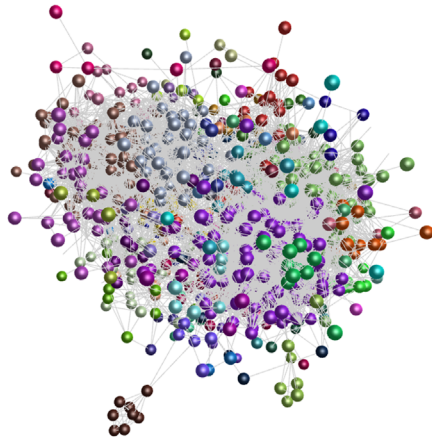


2. Experimental design, materials and methods

1. Mouse gene expression data were obtained as described in [1]. The list is provided in the Excel workbook, sheet “Mouse genes”, which contains the gene ID, gene symbol, microarray reporter ID and the Pearson correlation coefficient for the correlation of expression pattern across mouse strains with *Fbn1*. A high correlation coefficient indicates that the gene had a very similar expression pattern, that is it was high in epididymal fat of mice where expression of *Fbn1* was also high.
2. Human gene expression during the transition from mesenchymal stem cells to adipocytes in culture was obtained from a publicly available time course of adipogenesis available from the FANTOM5 website, as described in [1]. Gene expression patterns were clustered using BioLayout Express^{3D} [1]. Fig. 1A shows the generated network graph of the main element. Fig. 1B shows the up- and downregulated clusters within the same element. The node representing *FBN1* is located in the centre of Cluster 06 (pink). Histograms show expression patterns of these nodes. Fig. 1C shows expression patterns of some key processing enzymes for fibrillin-1. The list of genes in each cluster is provided in the Excel workbook, sheet “Human cluster genes”, which contains the gene symbol, the cluster number (clusters of four or more genes) and a description of the average expression pattern in the cluster (for clusters with at least 7 nodes; the remaining clusters were too small for meaningful interpretation).
3. A separate file containing gene based transcription factor expression levels was downloaded from the FANTOM5 website for clustering of transcription factor expression patterns. A Pearson correlation coefficient of 0.62 with an MCL inflation value of 2.2 was used to cluster the transcription factors. This ensured that all 1380 available transcription factors were included in the analysis. There were 73 clusters with four or more nodes and 209 genes whose expression pattern did not cluster with at least 3 other genes. Fig. 2A shows the generated network graph of the main element. Fig. 2B shows the up- and downregulated clusters within the same element. Histograms show expression patterns of these nodes. The list is provided in the Excel workbook, sheet “Human TF genes”, which contains the gene symbol, the cluster number and a description of the average expression pattern in the cluster (for clusters with at least 10 nodes).

Fig. 1. Network visualisation and clustering of gene expression patterns based on transcription initiation during human adipogenesis *in vitro*. Expression levels were derived from transcription start site data [4] available at the FANTOM5 website. Clustering was performed using BioLayout Express^{3D} with a correlation coefficient threshold of 0.85 and an MCL inflation value of 2.2. Spheres represent genes and edges correlations in expression patterns between them with a correlation coefficient ≥ 0.85 . (A) Network graph of the main element, showing genes (spheres) and correlations between them of ≥ 0.85 (grey lines). Nodes in clusters of genes with similar expression patterns as determined by the MCL clustering algorithm, with an inflation value of 2.2, are shown in the same colour. Each replicate was entered separately. (B) Network graph of the main up- and downregulated clusters, with edges removed for ease of visualisation. The average expression pattern of each of the clusters is shown in the histograms. Colours in the histograms are the same as the nodes they represent. Results for individual replicates are shown, separated by lines, and the order of samples across the X axis is: Replicate 1: 3 h, 1 day, 2 days, 4 days, 8 days, 12 days, 14 days post induction of differentiation; Replicate 2: 0 h, 2 days, 8 days, 12 days, 14 days post induction; Replicate 3: 0 h, 3 h, 1 day, 2 days, 4 days, 8 days, 12 days post induction. (C) Expression pattern of proteases that cleave fibrillins. *MMP2* (maximum 6000 tpm) and *MMP3* (maximum 160 tpm) were in cluster 3; *MMP14* (maximum 850 tpm) and *FURIN* (maximum 13 tpm) did not cluster.

A



B

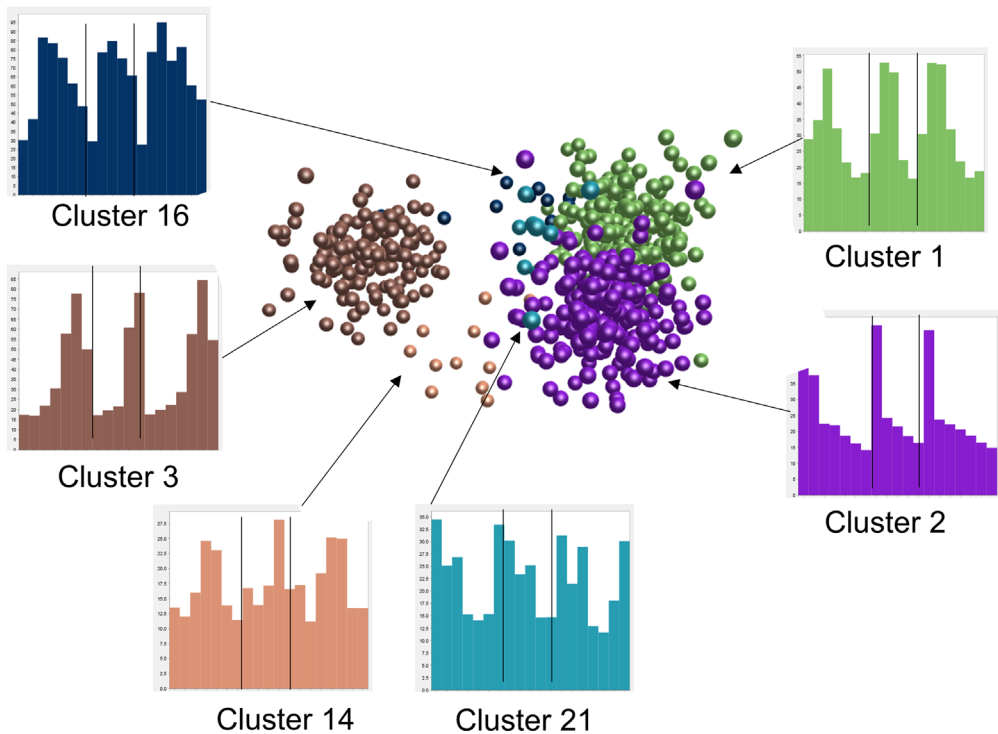


Fig. 2. Network visualisation and clustering of gene expression patterns for 1380 transcription factor genes, based on transcription initiation during human adipogenesis *in vitro*. Expression levels were derived from transcription start site data [4] available at the FANTOM5 website. Clustering was performed using BioLayout Express^{3D} with a correlation coefficient threshold of 0.62 and an MCL inflation value of 2.2. Spheres represent genes and edges correlations in expression patterns between them with a correlation coefficient ≥ 0.62 . (A) Network graph of the main element, showing genes (spheres) and correlations between them of ≥ 0.62 (grey lines). Nodes in clusters of genes with similar expression patterns as determined by the MCL clustering algorithm, with an inflation value of 2.2 are shown in the same colour. Each replicate was entered separately. (B) Network graph of the main up- and down-regulated clusters, with edges removed for ease of visualisation. The average expression pattern of each of the clusters is shown in the histograms. Colours in the histograms are the same as the nodes they represent. Results for individual replicates are shown, separated by lines, and the order of samples across the X axis is: Replicate 1: 3 h, 1 day, 2 days, 4 days, 8 days, 12 days, 14 days post induction of differentiation; Replicate 2: 0 h, 2 days, 8 days, 12 days, 14 days post induction; Replicate 3: 0 h, 3 h, 1 day, 2 days, 4 days, 8 days, 12 days post induction.

Acknowledgements

The Roslin Institute (MRD and KMS) is supported by Institute Strategic Programme Grants from the Biotechnology and Biological Resources Council (BBSRC) of the United Kingdom, grant numbers BB/J004235/1 and BB/J004316/1. EA was supported by a Research Grant from MEXT to the RIKEN Center for Life Science Technologies. PA and ID receive support from the Swedish Research Council.

Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.06.055>.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.06.055>.

References

- [1] M.R. Davis, E. Arner, C.R.E. Duffy, P.A. De Sousa, I. Dahlman, P. Arner, K.M. Summers, Expression of FBN1 during adipogenesis: relevance to the lipodystrophy phenotype in Marfan syndrome and related conditions, *Mol. Genet. Metab.* (2016), in press; 20 June 2016.
- [2] C. Romere, C. Duerrschmid, J. Bournat, P. Constable, M. Jain, F. Xia, P.K. Saha, M. Del Solar, B. Zhu, B. York, P. Sarkar, D.A. Rendon, M.W. Gaber, S.A. LeMaire, J.S. Coselli, D.M. Milewicz, V.R. Sutton, N.F. Butte, D.D. Moore, A.R. Chopra, Asprosin, a fasting-induced glucogenic protein hormone, *Cell* 165 (2016) 566–579.
- [3] C. Wu, D.L. Delano, N. Mitro, S.V. Su, J. Janes, P. McClurg, S. Batalov, G.L. Welch, J. Zhang, A.P. Orth, J.R. Walker, R.J. Glynne, M.P. Cooke, J.S. Takahashi, K. Shimomura, A. Kohsaka, J. Bass, E. Saez, T. Wiltshire, A.I. Su, Gene set enrichment in eQTL data identifies novel annotations and pathway regulators, *PLoS Genet* 4 (2008) e1000070.
- [4] E. Arner, C.O. Daub, K. Vitting-Seerup, R. Andersson, B. Lilje, F. Drablos, A. Lennartsson, M. Ronnerblad, O. Hrydziusko, M. Vitezic, T.C. Freeman, A.M. Alhendi, P. Arner, R. Axton, J.K. Baillie, A. Beckhause, B. Bodega, J. Briggs, F. Brombacher, M. Davis, M. Detmar, A. Ehrlund, M. Endoh, A. Eslami, M. Fagiolini, L. Fairbairn, G.J. Faulkner, C. Ferrai, M.E. Fisher, L. Forrester, D. Goldowitz, R. Guler, T. Ha, M. Hara, M. Herlyn, T. Ikawa, C. Kai, H. Kawamoto, L.M. Khachigian, S.P. Klinken, S. Kojima, H. Koseki, S. Klein, N. Mejhert, K. Miyaguchi, Y. Mizuno, M. Morimoto, K.J. Morris, C. Mummery, Y. Nakachi, S. Ogishima, M. Okada-Hatakeyama, Y. Okazaki, V. Orlando, D. Ovchinnikov, R. Passier, M. Patrikakis, A. Pombo, X.Y. Qin, S. Roy, H. Sato, S. Savvi, A. Saxena, A. Schwegmann, D. Sugiyama, R. Swoboda, H. Tanaka, A. Tomoiu, L.N. Winteringham, E. Wolvetang, C. Yanagi-Mizuochi, M. Yoneda, S. Zabierowski, P. Zhang, I. Abugessaisa, N. Bertin, A.D. Diehl, S. Fukuda, M. Furuno, J. Harshbarger, A. Hasegawa, F. Hori, S. Ishikawa-Kato, Y. Ishizu, M. Itoh, T. Kawashima, M. Kojima, N. Kondo, M. Lizio, T.F. Meehan, C.J. Mungall, M. Murata, H. Nishiyori-Sueki, S. Sahin, S. Nagao-Sato, J. Severin, M.J. de Hoon, J. Kawai, T. Kasukawa, T. Lassmann, H. Suzuki, H. Kawaji, K.M. Summers, C. Wells, F. Consortium, D.A. Hume, A.R. Forrester, A. Sandelin, P. Carninci, Y. Hayashizaki, Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells, *Science* 347 (2015) 1010–1014.
- [5] A.M. Pettersson, B.M. Stenson, S. Lorente-Cebrian, D.P. Andersson, N. Mejhert, J. Kratzel, G. Astrom, I. Dahlman, A.V. Chibalin, P. Arner, J. Laurencikienė, LXR is a negative regulator of glucose uptake in human adipocytes, *Diabetologia* 56 (2013) 2044–2054.
- [6] A.M. Rodriguez, D. Pisani, C.A. Dechesne, C. Turc-Carel, J.Y. Kurzenne, B. Wdziekonski, A. Villageois, C. Bagnis, J.P. Breittmayer, H. Groux, G. Ailhaud, C. Dani, Transplantation of a multipotent cell population from human adipose tissue induces dystrophin expression in the immunocompetent mdx mouse, *J. Exp. Med.* 201 (2005) 1397–1405.
- [7] L.E. Zaragosi, G. Ailhaud, C. Dani, Autocrine fibroblast growth factor 2 signaling is critical for self-renewal of human multipotent adipose-derived stem cells, *Stem Cells* 24 (2006) 2412–2419.
- [8] C. Wu, C. Orozco, J. Boyer, M. Leglise, J. Goodale, S. Batalov, C.L. Hodge, J. Haase, J. Janes, J.W. Huss 3rd, A.I. Su, BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources, *Genome Biol.* 10 (2009) R130.
- [9] A.R.R. Forrester, H. Kawaji, M. Rehli, J.K. Baillie, M.J.L. de Hoon, H.V. T. Lassmann, K.I. V. M. Lizio, M. Itoh, R. Andersson, C.J. Mungall, T.F. Mehan, A promoter level mammalian expression atlas, *Nature* 507 (2014) 462–470.
- [10] T.C. Freeman, L. Goldovsky, M. Brosch, S. van Dongen, P. Maziere, R.J. Crocock, S. Freilich, J. Thornton, A.J. Enright, Construction, visualisation, and clustering of transcription networks from microarray expression data, *PLoS Comput. Biol.* 3 (2007) 2032–2042.
- [11] A. Theodoridis, S. van Dongen, A.J. Enright, T.C. Freeman, Network visualization and analysis of gene expression data using BioLayout Express^{3D}, *Nat. Protoc.* 4 (2009) 1535–1550.