A Topographical Map of Spatiotemporal Patterns of Gene Expression

Eileen E.M. Furlong[1,*]
[1]European Molecular Biology Laboratory, D-69117 Heidelberg, Germany
*Correspondence: furlong@embl.de
DOI 10.1016/j.devcel.2008.04.007

A recent study by Folkes et al. in *Cell* generated a 3D atlas of gene expression for the *Drosophila* blastoderm embryo using a new approach for image registration. This virtual embryo allows in silico multiplexing of in situ hybridizations and lays the groundwork for new insights into gene regulatory networks.

The past decade has seen an exponential growth in the volume of data describing developmental gene expression patterns. Expression profiling using microarrays or SAGE has provided quantitative information on mRNA expression during defined temporal windows of development (Arbeitman et al., 2002; Tomancak et al., 2007). However the degree of spatial information is very limited. In situ hybridization is therefore an essential complementary approach. Although traditionally limited to small numbers of genes, robotic liquid handling has allowed in situ hybridizations to be conducted at a medium- to high-throughput scale (Pollet et al., 2005; Satou et al., 2001; Sprague et al., 2008; Tomancak et al., 2007). While this provides an invaluable resource to evaluate the potential function of these genes, the nonquantitative nature of these data make them very difficult to integrate into larger scale models of regulatory networks. High-resolution fluorescent in situ hybridizations (FISH), performed using linear detection methods, can provide quantitative expression data, but it is difficult to expand beyond the simultaneous detection of five or six transcripts (Kosman et al., 2004). Therefore, despite the wealth of information, there is a large gap between the genomic scale of quantitative approaches and the spatial detail of fluorescent in situ hybridization.

To elucidate transcriptional networks, a current challenge is therefore to capture quantitative spatiotemporal information of all genes expressed in a developing embryo in a unified topographical map. This will reveal cohorts of genes with tightly regulated expression in discrete populations of cells. Such a 3D expression atlas would also be invaluable for network analysis to formulate relational models between the location of a transcription factor's expression compared to that of its target genes. Although an obvious essential step to decipher gene regulatory networks, this type of information remains largely unknown for the vast majority of developmental regulators. A number of efforts have begun to bridge this gap by constructing a virtual embryo where in situ hybridization patterns of hundreds or thousands of genes are overlaid and directly compared in silico by virtual multiplexing of in situs. The crux of this approach lies in the ability to align the images from thousands of embryos to bring the data into register at a pixel level. For some model organisms, image registration was achieved by taking advantage of very stable anatomical features; e.g., the mouse brain (Lein et al., 2007). Alternatively, in animals where the cell number is relatively small and/or the cell size is large, the cell-to-cell registration can be achieved manually; e.g., Ciona (Tassy et al., 2006). However, these approaches are not appropriate for organisms whose embryos have more varied anatomical appearance and/or contain a large number of cells.

A study by Fowlkes et al. in a recent issue of *Cell* addressed this by developing a sophisticated two-step registration process to compile hundreds of independent in situ measurements into an average expression pattern, mapped to a virtual embryo (Fowlkes et al., 2008). *Drosophila* blastoderm embryos were stained by double fluorescent in situ hybridizations using probes directed against a reference gene (either *ftz* or *eve*, two key regulators during this stage of development) and gene X or Y. A general DNA marker was used to locate each nucleus. To minimize inter and intra in situ batch variation, the fluorescence measurements were normalized to put them on a common scale. This is a challenging problem in itself. The authors used the center of mass coordinates of all nuclei and the average fluorescence levels of the two genes in the nucleus and surrounding cytoplasm, referred to collectively as a point cloud, to map the nuclear location onto a 3D lattice. For spatial registration, the point clouds from hundreds of tightly stage-matched embryos were used to make a reference virtual embryo with average shape and nuclear distribution. Successive iterations refined the mapping of the reference gene's expression to each nucleus within the virtual embryo, providing an equivalence map for all cells across all embryos.

Although there are no cell divisions during these stages of development, the nuclei move and thus have dynamic locations rather than static coordinates. Therefore, the boundaries of *ftz* or *eve* expression could not be used to identify corresponding nuclei over time. To address this, the authors implemented a second step of temporal registration. The point cloud data provides an average embryo shape and nuclear density pattern at a given blastoderm stage. By comparing these data across neighboring time points, a numerical model was generated to account for the distance each nucleus had to travel to explain the measured changes in nuclear density. It will be interesting to see if this type of registration approach can be extended to developmental stages undergoing cell division and more complex embryos with multiple cell layers.

Registering the reference gene expression from hundreds of embryos permitted the registration of the second gene's expression, producing an average 3D expression map for gene X or Y. Although the in situ hybridizations of gene X and Y were conducted independently, their expression patterns can be overlaid and compared as their average expression patterns are mapped to a common model (the virtual embryo). This allows spatial multiplexing in silico at an almost infinite scale.

The authors performed this expression mapping for 95 genes during a 50 min interval prior to gastrulation—an impressive body of work. The results are very exciting: 3D virtual images that look almost identical to simultaneous in situs performed in the same embryo. It is very difficult to quantitate how similar "almost identical" is. The authors assessed this by comparing the variability in expression of two genes (*giant* and *eve*) measured in the same embryo (a standard double in situ hybridization) to that inferred by their registration method for the two genes imaged in different embryos. The variability in expression was measured by grouping nuclei that have similar expression levels of *giant* into bins and then looking at the average expression levels of *eve*. The maximum standard deviation of *eve* expression in all bins was 0.21 for the "real" double in situs and 0.3 in the virtual embryo, while the maximum expression levels of *eve* and *giant* deviate by <7% between the two approaches. By these two criteria, the virtual embryo provides an accurate average quantitation of gene expression in 3D.

Many studies have tried to reconstruct gene regulatory networks based on expression data under the assumption that a transcription factor and a target gene will be temporally coexpressed for some period of time or have a small temporal shift in their expression. These studies are primarily based on microarray profiling data, as this has been the only global quantitative data available. The extraction of relative gene expression data using virtual embryos opens a new avenue for inferring regulatory connections at a cell-based level. Given the much higher spatial resolution, these data, in combination with other data such as transcription factor binding site occupancy, should vastly improve the predictive power of regulatory models. There is a clear need to move toward a quantitative protein atlas of transcription factor expression, integrated with quantitative measures of their target genes' mRNA expression. The computational methods developed by Fowlkes et al. (2008) are an important first step in this direction.

REFERENCES

Arbeitman, M.N., Furlong, E.E., Imam, F., Johnson, E., Null, B.H., Baker, B.S., Krasnow, M.A., Scott, M.P., Davis, R.W., and White, K.P. (2002). Science *297*, 2270–2275.

Fowlkes, C.C., Luengo Hendriks, C.L., Keränen, S.V., Weber, G.H., Rübel, O., Huang, M.-Y., Chatoor, S., DePace, A.H., Simirenko, L., Henriquez, C., et al. (2008). Cell *133*, 364–374.

Kosman, D., Mizutani, C.M., Lemons, D., Cox, W.G., McGinnis, W., and Bier, E. (2004). Science *305*, 846.

Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Nature *445*, 168–176.

Pollet, N., Muncke, N., Verbeek, B., Li, Y., Fenger, U., Delius, H., and Niehrs, C. (2005). Mech. Dev. *122*, 365–439.

Satou, Y., Takatori, N., Yamada, L., Mochizuki, Y., Hamaguchi, M., Ishikawa, H., Chiba, S., Imai, K., Kano, S., Murakami, S.D., et al. (2001). Development *128*, 2893–2904.

Sprague, J., Bayraktaroglu, L., Bradford, Y., Conlin, T., Dunn, N., Fashena, D., Frazer, K., Haendel, M., Howe, D.G., Knight, J., et al. (2008). Nucleic Acids Res. *36*, D768–D772.

Tassy, O., Daian, F., Hudson, C., Bertrand, V., and Lemaire, P. (2006). Curr. Biol. *16*, 345–358.

Tomancak, P., Berman, B.P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S.E., and Rubin, G.M. (2007). Genome Biol. *8*, R145.

Together We Stand: Genes Cluster to Coordinate Regulation

Grigoris Amoutzias[1] and Yves Van de Peer[1,*]
[1]Department of Plant Systems Biology, VIB and Department of Molecular Genetics, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium
*Correspondence: yves.vandepeer@psb.ugent.be
DOI 10.1016/j.devcel.2008.04.006

Although most eukaryotic genomes lack operons, occasionally clusters of genes are discovered that are related in function. Now, a metabolic operon-like gene cluster has been described in *Arabidopsis thaliana* that is needed for triterpene synthesis.

Traditionally, genes were considered to be randomly distributed in eukaryotic genomes, in stark contrast with prokaryotes, where unrelated genes with related functions are clustered in so-called operons. However, in the last few years, genomic data reveal that also in eukaryotic genomes, genes can be physically clustered. The best known example is probably *Caenorhabditis elegans*, where almost 15% of the genes are organized in polycistronic operons, but also in yeast,