BIOINFORMATICS APPLICATIONS NOTE Vol. 30 no. 9 2014, pages 1319–1321
doi:10.1093/bioinformatics/btu007

Gene expression

Advance Access publication January 9, 2014

# myFX: a turn-key software for laboratory desktops to analyze spatial patterns of gene expression in *Drosophila* embryos

Ivan Montiel[1], Charlotte Konikoff[1], Bremen Braun[1], Mary Packard[2], Sian L. Gramates[2], Qian Sun[1,3], Jieping Ye[1,3] and Sudhir Kumar[1,4,5,*]

[1]Center for Evolutionary Medicine and Informatics, Biodesign Institute, Arizona State University (ASU), Tempe, AZ 85287, USA, [2]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA, [3]School of Computing, Informatics, and Decision Systems Engineering, ASU, Tempe, AZ 85287, USA, [4]School of Life Sciences, ASU, Tempe, AZ 85287, USA and [5]Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia

Associate Editor: Janet Kelso

## ABSTRACT

**Summary:** Spatial patterns of gene expression are of key importance in understanding developmental networks. Using *in situ* hybridization, many laboratories are generating images to describe these spatial patterns and to test biological hypotheses. To facilitate such analyses, we have developed biologist-centric software (*myFX*) that contains computational methods to automatically process and analyze images depicting embryonic gene expression in the fruit fly *Drosophila melanogaster*. It facilitates creating digital descriptions of spatial patterns in images and enables measurements of pattern similarity and visualization of expression across genes and developmental stages. *myFX* interacts directly with the online *FlyExpress* database, which allows users to search thousands of existing patterns to find co-expressed genes by image comparison.

**Availability and implementation:** *myFX* is freely available at http://www.flyexpress.net.

**Contact:** s.kumar@asu.edu

Received on July 17, 2013; revised on October 16, 2013; accepted on January 5, 2014

Analyzing embryonic gene expression in developmental space and time has greatly contributed to our understanding of developmental regulatory networks. Genes co-expressed in the developing embryo are of special interest because they may share common regulatory elements and participate in similar developmental processes (Konikoff *et al.,* 2012; Kumar *et al.,* 2011; Lécuyer *et al.,* 2007; Tomancak *et al.,* 2002, 2007). *In situ* hybridization data from *Drosophila melanogaster* are a model for such image-based analyses because of the vast number of experimental studies revealing the functions of many genes and an abundant collection of high-throughput *in situ* images capturing the expression time-course of these genes throughout embryogenesis (Kumar *et al.,* 2002; Lécuyer *et al.,* 2007; Tomancak *et al*. 2002, 2007). *Drosophila* embryos have been efficiently digitally standardized, allowing direct image-to-image comparisons of spatial expression patterns (Frise *et al*. 2010; Kumar *et al*. 2011). Overall, *Drosophila* presents a unique opportunity to assess both developmental spatial patterns and functional significance. Therefore, many laboratories use the *Drosophila* model system to investigate spatial patterns and test biological hypotheses.

We have developed a new software package (*myFX*) to facilitate investigations of these spatial patterns of gene expression. This software is intended to be a user-friendly platform to store, annotate, visualize and analyze images depicting expression patterns of developmentally relevant genes in *Drosophila* embryogenesis. It is closely linked with our *FlyExpress* resource (Kumar *et al*. 2011) to enhance function discovery by searching thousands of existing images of gene expression patterns.

For ease of use, *myFX* provides fast installation on Mac OSX and Microsoft Windows, which produces a local image repository and SQL database without requiring users to have any knowledge of databases or other advanced computer techniques. The *myFX* interface works in a similar manner on both operating systems and accepts image data in TIFF, JPEG and BMP image formats, among others. Images can be imported into *myFX* by browsing file folders or by dragging-and-dropping. Here, we describe key functionalities of *myFX* through two common use cases.

In one, a user has a novel *in situ* image and would like to standardize it and search for co-expressed genes. Once a user uploads images into *myFX* (Fig. 1A), automatic processing for size standardization and image alignment necessary for developmental comparisons is initiated (Konikoff *et al*. 2012; Kumar *et al*. 2011) (Fig. 1B). *myFX* contains facilities for the user to refine standardization and alignment using a built-in image editor, where the images can be rotated, flipped (horizontally and vertically) and scrubbed (Fig. 1C). This means that the user can work interactively with *myFX* to manually enhance quality of image standardization.

Next, *myFX* creates three spatial profiles for each processed embryo, where black pixels denote expression presence and white pixels indicate expression absence. The middle profile represents gene expression in the image, whereas the left and right profiles are under- and over-representation of staining, respectively. These spatial profiles are necessary for expression comparisons and other downstream analyses, including searching the online *FlyExpress* database (or a local database, see later) for genes with similar patterns of expression (Fig. 1D). Such analyses also require that the developmental stage and anatomical orientation of

*To whom correspondence should be addressed.


**Fig. 1.** (**A**) An image containing *Dichaete* expression, which is being dragged-and-dropped into the *myFX* visual interface for processing. (**B**) A screenshot of the *myFX* image grid that displays thumbnails of original and processed images, along with their standardized expression patterns. (**C**) The *myFX* interface provides an image editor with capability to correct automatically standardized expression patterns. (**D**) Users can launch searches for co-expressed genes by accessing the menu associated with expression profiles. (**E**) Each processed image is associated with three unique expression profiles, which can be used for searching. Pixels from the selected pattern are automatically mapped onto a GEM (Genomewide Expression Map), which represents global expression for genes of a given stage range and anatomical orientation in the *FlyExpress* digital library. (**F**) A user may drag-and-align the pixels onto the GEM to enhance the accuracy of image alignments. They may also select the sampling of pixels to be used. (**G**) A results table showing a ranked list of genes, sorted according to overlap between the query and expression of genes captured in the GEM

the embryo (e.g. lateral, dorsal or ventral) be specified, which is easily input into *myFX*.

To search for co-expressed genes, clicking on a spatial profile produces a search dialog for launching an online search (Fig. 1E). Selecting an expression profile and data source results in a set of search pixels overlaid on a *FlyExpress* Genomewide Expression Map (GEM) below (Fig. 1F). Search pixels represent key points of gene expression in a spatial profile and can be aligned on the GEM by manually dragging the pixel collection. The GEMs represent global expression among genes for a given stage and anatomical orientation (Konikoff *et al.* 2012). Using our novel GEM-based searching system, co-expressed genes are found by taking the collection of pixels in a spatial profile, mapping them onto the appropriate GEM and asking for genes that share the highest numbers of co-expressed pixels with the query.

This system replaces the current image-to-image similarity search based on pairwise comparison of spatial profiles (Kumar *et al.* 2011) and obviates the need to pre-compute an enormous pairwise image similarity matrix that currently contains hundreds of millions of values and requires a large-scale database indexing system. To enhance the speed and quality of GEM-based searching, we also developed new multipoint searching capabilities between *myFX* and *FlyExpress*. A user can find genes with similar expression patterns in *FlyExpress* from an *ad hoc* user defined multipoint search from GEMs created in *myFX*. This capability is available thanks to new database updates to *FlyExpress*.

Launching the search produces a new window, where results are displayed in a list format (Fig. 1G). For example, 2237 genes showing varying extents of co-expression with the *Dichaete* query are produced. The number to the left of each gene indicates the

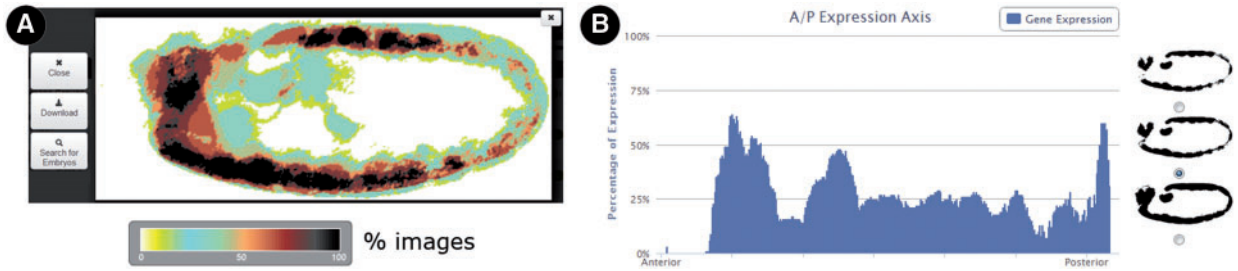


**Fig. 2.** (A) Expression profiles facilitate creation of custom GEMs, which visually represent expression in a set of images. A GEM of the genes in Figure 1 is shown. (B) Expression levels can also be visualized along anterior-posterior and dorsal-ventral axes. In the histogram, percentage of *Dichaete* expression is determined by the number of black pixels to total pixels for the expression profile selected

number of pixels exhibiting co-expression with the query. In the example, 84 genes show co-expression with *Dichaete* at >90% of selected pixels. Additional genes involved in neuroblast development (such as *SoxN* and *pdm2*, Buescher *et al*. 2002; Tsuji *et al*. 2008) are among the top-ranked results. In addition to these biologically relevant examples, some uncharacterized genes exhibiting high expression overlap are also among the results. These genes are now candidates for experimental testing to determine potential involvement, if any, in neuroblast or related developmental networks.

In the second usage scenario, a *myFX* user has many more images for processing and analysis. *myFX* will automatically process all images, and users can refine and annotate each image and its expression. Any spatial expression pattern may be used to launch a search for co-expressed genes within the local database, or online via *FlyExpress*. Also, users can organize images into groups in *myFX*, which can be used to generate custom GEMs (Fig. 2A). In GEMs, expression levels are rendered as heatmaps, with darker colors indicating greater amounts of co-expression in a given area (Kumar *et al*. 2011). Spatial profiles can also be used to visualize expression levels along an embryo's anterior–posterior or dorsal–ventral axis (Fig. 2B). These histograms can be used to graphically compare the time-course of expression for a specific gene in the context of embryonic axes or to compare expression for a group of genes expressed at the same stage in the developing embryo.

*myFX* offers advantages over existing tools for managing and analyzing *in situ* experimental data, such as FlyGUI (Crombach *et al*. 2012) and SPEX$^2$ (Puniyani *et al*. 2010). For example, *myFX* is a turn-key application that can be easily installed by any user on Windows and Mac OS, which are popular operating systems in *Drosophila* developmental biology laboratories. Also, *myFX* can process whole-mount *in situ* images from a wider variety of file formats. It is also user-friendly, as drag-and-drop uploading and sorting will permit biologists to easily organize their images in *myFX*. Moreover, *myFX* provides automated image standardization with tools for manual refinement, allowing expression profiles to be edited if desired. In addition, *myFX* is unique in its ability to search for genes with overlapping patterns locally within the user's own images and on the web at *FlyExpress*. For sets of expression patterns, *myFX* can also uniquely generate expression summaries, called GEMs, which summarize expression in a collection of images belonging to a specific stage (or stage range) and anatomical orientation.

Overall, our *myFX* software represents our effort to give biologists image storage and additional gene expression analysis tools for use on their own desktops. We expect this new resource will open examinations of patterns of gene expression to a wider audience of biologists, allowing them to quickly identify new candidates for hypothesis-driven experimentation, and thereby lead to enhanced discoveries on biological mechanisms that direct important biological processes.

## ACKNOWLEDGEMENT

The authors thank Michael McCutchan, Adam Orr, Stuart Newfeld, Bill Gelbart and Thom Kaufman for their valuable advice and help during the development of this software and feedback on initial versions of this article.

*Funding*: US National Institutes of Health (HG002516-09 to S.K. and LM010730-03 to J.Y.). The source code of *myFX* is available upon request.

*Conflict of Interest*: none declared.

## REFERENCES

Buescher,M. *et al*. (2002) Formation of neuroblasts in the embryonic central nervous system of *Drosophila melanogaster* is controlled by *SoxNeuro*. *Development*, **129**, 4193–4203.

Crombach,A. *et al*. (2012) Medium-throughput processing of whole mount *in situ* hybridisation experiments into gene expression domains. *PLoS One*, **7**, e44658.

Frise,E. *et al*. (2010) Systematic image-driven analysis of the spatial *Drosophila* embryonic expression landscape. *Mol. Syst. Biol*., **6**, 345.

Konikoff,C. *et al*. (2012) Comparison of embryonic expression within multigene families employing the FlyExpress discovery platform reveals significantly more spatial than temporal divergence. *Dev. Dyn*., **241**, 150–160.

Kumar,S. *et al*. (2002) BEST: a novel computational approach for comparing gene expression patterns from early stages of *Drosophila melanogaster* development. *Genetics*, **162**, 2037–2047.

Kumar,S. *et al*. (2011) FlyExpress: visual mining of spatiotemporal patterns for genes and publications in *Drosophila* embryogenesis. *Bioinformatics*, **27**, 3319–3320.

Lécuyer,E. *et al*. (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell*, **131**, 174–187.

Puniyani,K. *et al*. (2010) SPEX2: automated concise extraction of spatial gene expression patterns from Fly embryo ISH images. *Bioinformatics*, **26**, 47–56.

Tomancak,P. *et al*. (2002) Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol*., **3**, RESEARCH0088.1–88.14.

Tomancak,P. *et al*. (2007) Global analysis of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol*., **8**, R145.

Tsuji,T. *et al*. (2008) Neuroblast entry into quiescence is regulated intrinsically by the combined action of spatial Hox proteins and temporal identity factors. *Development*, **135**, 3859–3869.