Gene Expression Patterns 27 (2018) 36–45

Contents lists available at ScienceDirect

Gene Expression Patterns

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

High-throughput method for extracting and visualizing the spatial gene expressions from *in situ* hybridization images: A case study of the early development of the sea anemone *Nematostella vectensis*

A.M. Abdol, Andrew Bedard, Imke Lánský, J.A. Kaandorp*

University of Amsterdam, Science Park 904, 1098XH Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
In situ hybridization
Spatial gene expression profile
Image processing
Gastrulation

ABSTRACT

Studying the spatial gene expression profiles from *in situ* hybridization images of the embryo is one of the first steps toward the comprehensive understanding of gene interactions in an organism. In the case of *N. vectensis*, extracting and collecting these data is a challenging task due to the difficulty of detecting the cell layer through the transparent body plan and changing morphology during the blastula and gastrula stages. Here, first, we introduce a method to algorithmically identify and track the cell layer in *N. vectensis* embryo from the late blastula to the late gastrula stage. With this, we will be able to extract spatial expression profiles of genes alongside the cell layer and consequently reconstructing the 1D representation of gene expression profiles. Furthermore, we use the morphological configurations of the embryo extracted from confocal images, to model the dynamics of embryos morphology during the gastrulation process in 2D. Ultimately, we provide a visualization tool for studying and comparing the extracted spatial gene expression profiles over the simulated embryo. We anticipate that our method of extraction and visualization to be a starting point for quantifying and collecting more *in situ* images from various sources, which can potentially accelerate our understanding of gene interactions in the early development of *N. vectensis*. The method allows researchers to visualize and compare the different gene expressions from different *in situ* images or different experiments. As an example, we were able to show the complementary expression of *NvFoxA-NvSnailA* and *NvBra-NvErg* in the *central domain* and *central/external rings* during the development which suggests the possible repression effects between each pair; as it has been discovered by functional analysis.

1. Introduction

In situ hybridization and *protein immunolocalization*, are among commonly used methods for detecting the expression territory of a gene in an organism or tissue. They are the primary tools for studying the expression regions of specific genes and reviewing the results of functional gene analysis e.g. knockdown experiments. The standard approach to examine the results of *in situ* experiments is to inspect the images of stained embryos and reason about the outcome of the study. However, as performing the *in situ* hybridization experiments are becoming more affordable and imaging techniques are becoming more accurate, biologists are producing a significant number of *in situ* images every day. As an excellent example of this, in the case of *Drosophila melanogaster*, carefully controlled fluorescent *in situ* hybridization through the course of the embryonic development provided invaluable snapshots of gene expression profiles, *spatiotemporal gene expressions* (Pisarev et al., 2009; Poustelnikova et al., 2004).

The traditional methods of studying the presence/absence of a gene or analyzing the expression territory of a gene by inspecting the *in situ* images are proven to be immensely instructive (Rivera-Pomar and Jackle, 1996; Ingham, 1988; Akam, 1987; Nusslein-Volhard et al., 1987; Hemavathy et al., 2004). In few cases where the *in situ* hybridization images were used to extract the spatiotemporal expression profiles of genes into numerical data, the data has been used to develop and test new mathematical models of Gene Regulatory Network (GRN), *Connectionist Model of Development* (Mjolsness et al., 1991). Availability of spatiotemporal gene expressions data of *D. melanogaster*, combined with the new model and improvement in optimization algorithms helped with the reconstruction of several GRNs in different developmental stages of *D. melanogaster* (Reinitz and Sharp, 1995; Jaeger et al., 2004; Perkins, 2007; Surkova et al., 2008). However, despite the growing number of *in situ* images from different organisms,

* Corresponding author.
E-mail addresses: a.m.abdol@uva.nl (A.M. Abdol), andrewjeanbedard@gmail.com (A. Bedard), imke.lansky@gmail.com (I. Lánský), j.a.kaandorp@uva.nl (J.A. Kaandorp).
URL: http://amirmasoudabdol.name (A.M. Abdol).

https://doi.org/10.1016/j.gep.2017.10.005
Received 19 June 2017; Received in revised form 12 October 2017; Accepted 17 October 2017
Available online 06 November 2017
1567-133X/

spatiotemporal models of GRN are often restricted to a few model organisms, e.g. *D. melanogaster*, this is mainly due to the laborious procedure of obtaining the data as well as the challenging task of extracting and quantifying the gene expression profiles from *in situ* images of the embryo.

Here we aim to introduce an automated algorithm for extracting gene expression profiles of an another popular model organism, *Nematostella vectensis*. This sea anemone is a cnidarian model organism that is intensely used to understand the evolution of biological novelties, development, ecology and since recently also regeneration (Layden et al., 2016a; Darling et al., 2005). The genome of *N. vectensis* has been sequenced and has revealed strong similarities regarding gene content and genome organization with the ones from vertebrates (Putnam et al., 2007). Thus, understanding the evolution of its GRNs has become a major task in deciphering the evolution of developmental processes. However, due to the existence of different germ layers and the changes in the embryo morphology from the blastula to gastrula stage, obtaining the spatiotemporal gene expression profiles poses an even bigger challenge compared to extraction of gene expression profiles from *D. melanogaster*'s blastoderm.

Previously, from a growing number of *in situ* hybridization images, Botman and colleagues introduced a labour-intensive method for extracting gene expressions from *N. vectensis* despite the changing morphology during the development. Thereafter, they have investigated the possibility of reverse-engineering a GRN involving in the gastrulation process (Botman et al., 2014, 2016; Kraus and Technau, 2006; Magie et al., 2007). In another study, Abdol et al. demonstrated that given the spatial and temporal gene expressions, they are able to group genes that are active in the same place and at the same time. Furthermore, by solely analyzing the data, they predicted the gene interactions where have been found through functional analysis of genes, namely the repression of *NvBra* by *NvErg* (Röttinger et al., 2012; Abdol et al., Kaandorp).

The availability of spatial genes profiles will allow researchers to tackle a new set of problems and use new data processing methods to study different aspects of development and gene regulations with more systematic and focused methods than the manual interpretation of *in situ* images, especially in a case of less known organism e.g. *N. vectensis*.

Here, in contrast to Botman's method of extraction which involved manual cell layer selection, expression extraction and post-processing, we are introducing a new algorithm to automatically detect the cell layer, identify boundaries of embryos, predict the embryo's current developmental stage, discretize the embryo into pseudo-cells and finally constructing the expression profile of genes. We think our method will accelerate the extraction of spatial gene expression profiles from the *in situ* images of *N. vectensis*'s embryo during blastula and gastrula stages of the development. Moreover, we will show that our method is able to process images from different labs with various experimental setups.

2. Material and methods

N. vectensis belongs to the sister group of bilaterian animals, the cnidarian (sea anemone, coral, jellyfish, *Hydra*). It follows a relatively simple embryonic development in which it preserves its radial symmetry in early developmental stages. In sexual reproduction, development starts with a fertilized egg. After a series of cleavages, the embryo forms a hollow sphere of cells known as the blastula. The development then continues to the gastrulation process in which morphological movements establish the *endodermal* and *ectodermal* germ layers (Magie et al., 2007; Kraus et al., 2007). As shown in Fig. 1, the *N. vectensis* transparent body wall makes it a suitable subject for various microscopy techniques. However as most imaging techniques capture a cross section of the embryo at a certain depth, the shadow of cells in different layers is always being captured as well, we call this the *look-through* effect. This effect combined with slight irregularities of the embryo in different development stages impose a crucial problem for detecting the cell layer and consequently extracting gene expression profile during developmental stages in our study the blastula and gastrula.

Fig. 2 shows examples of *in situ* hybridization images in which detecting the inner and outer boundaries of the cell layer is crucial for extracting correct expression profile of the gene. In Fig. 2A, inaccurate boundary detection leads to gene expression measurements where cells are not expected whereas in Fig. 2B, incorrect detection of the cell layer (and consequently the mouth's morphology) results in wrong representation of the embryo. Although the inner hollow region of the embryo provides an excellent guide where the cell layer is located during the blastula and through the gastrulation process, algorithmic detection of this region is a laborious task considering variations in the cell layer's width, embryo's stages. The process of extracting gene expression patterns from the fruit fly's embryo focuses on the blastoderm stage where the embryo's morphology does not change and the embryo can be described as an embryo sac filled with nuclei. This allows for a simpler procedure in which only the outer boundary of the embryo needs to be detected and thereafter the gene expression can be measured inside the embryo (Crombach et al., 2012).

In the rest of this section, we will introduce an algorithm to improve the extraction of gene expression profiles from *in situ* images of the *N. vectensis* embryo's. We start by detecting the inner region of the embryo using feature extraction algorithms that allows us to detect the outer and inner boundaries of the cell layer. Finally, we will reconstruct the gene expression profile by measuring the intensity of colors in the region where we already marked as the cell layer.

2.1. Detecting cell layer

In order to detect the cell layer, first, we need to distinguish the cell layer from the rest of the image. In other words, we need to detect the boundaries of the cell layer both from outside and inside of the embryo. The outside boundary distinguishes the cell layer from the background, and the inside boundary distinguishes the cell layer from the hollow region of the blastula. In order to diminish the mentioned *look-through* effect on the edge detection procedure, we have used an edge detection algorithm called *Phase Stretch Transform* (PST) (Asghari and Jalali, 2015). The image transformation of PST emulates the propagation of light through a diffractive medium by applying a phase function to the image in the frequency domain (Broughton and Bryan, 2008). The PST can detect edges in our images because its underlying phase function is frequency dependent, meaning it applies more phase to higher frequency features of the images. Through three parameters, the PST allows us to tune the algorithm for tolerating more noise in edge detection at the cost of *lower spatial resolution* as well as emphasizing on the detection of sharper edges at the cost of increased *edge noise*. Fig. 3B shows the output of PST on a selected *in situ* hybridization image.

Although the features extracted using PST, Fig. 3B, resemble the embryo, they cannot directly be used to detect the cell layer boundaries in *N. vectensis*. The amount of unnecessarily extracted features in the cell layer and inside the hollow region makes it is hard to detect the boundaries of the cell layer. In order to resolve this issue, we used an image processing filter known as the Frangi filter which was originally conceived as a vessel enhancement filter for Digital Subtraction Angiography (DSA) images. The Frangi filter uses the local structure of an image's Hessian to determine the vesselness of its features, namely, determining adjacent areas for which the Hessian describes a direction for which the intensity values changes by the least amount (Frangi et al., 1998). As shown in Fig. 3C, although the Frangi filter reduces the overall noise detected in the embryo, especially in the interior, the end output of Frangi filter still cannot be interpreted as edges. Therefore, we have used PST over the image, already processed by the Frangi filter to improve the edge detection, Fig. 3D. Comparing Fig. 3B and D, we see a better representation of the cell layer and improved interior detection.