Marine Genomics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Marine Genomics

Gene regulation in amphioxus: An insight from transgenic studies in amphioxus and vertebrates

Iryna Kozmikova, Zbynek Kozmik *

Institute of Molecular Genetics of the Academy of Sciences, Videnska 1083, 142 20 Prague 4, Czech Republic

ARTICLE INFO

Article history:
Received 28 February 2015
Received in revised form 10 June 2015
Accepted 10 June 2015
Available online xxxx

Keywords:
Transgenic animal
Vertebrates
Chordates
Gene regulation
Evolution

ABSTRACT

Cephalochordates, commonly known as amphioxus or lancelets, are the most basal subphylum of chordates. Cephalochordates are thus key to understanding the origin of vertebrates and molecular mechanisms underlying vertebrate evolution. The evolution of developmental control mechanisms during invertebrate-to-vertebrate transition involved not only gene duplication events, but also specific changes in spatial and temporal expression of many genes. To get insight into the spatiotemporal regulation of gene expression during invertebrate-to-vertebrate transition, functional studies of amphioxus gene regulatory elements are highly warranted. Here, we review transgenic studies performed in amphioxus and vertebrates using promoters and enhancers derived from the genome of *Branchiostoma floridae*. We describe the current methods of transgenesis in amphioxus, provide evidence of Tol2 transposon-generated transgenic embryos of *Branchiostoma lanceolatum* and discuss possible future directions. We envision that comparative transgenic analysis of gene regulatory sequences in the context of amphioxus and vertebrate embryos will likely provide an important mechanistic insight into the evolution of vertebrate body plan.

1. Introduction

The ultimate goal if we were to fully understand the animal evolution lies in the discovery of the regulatory codes in the genomes. It has become apparent that considerable differences in morphology and overall complexity of body plans among animals are not mirrored at the level of gene number. In fact, a great part of the gene set is shared among Bilateria and their sister group, the Cnidaria, consisting of only two germ layers and a limited number of cell types, suggesting that the common ancestor of eumetazoans already had a highly complex gene repertoire (Kortschak et al., 2003). A recent study found that enhancers in cnidarian *Nematostella vectensis* are characterized by the same combinations of histone modifications as in bilaterians, and that these enhancers preferentially link to developmental control genes (Schwaiger et al., 2014). These results suggest that at least some complex features of gene regulation were present in the common ancestor of eumetazoa.

It is well established that the precise spatial, temporal, and quantitative regulation of gene expression is essential for proper animal development. Numerous studies have identified cis-regulatory mutations with functional consequences for morphology, physiology, and behavior (Wray, 2007). Changes in gene regulation are thus one of the major potential driving forces of species evolution. Indeed, the evolution of new body plans is often driven by changes in the regulation of gene expression (Carroll, 2008). The regulatory machinery controlling body plan formation is comprised of an intricate array of transcription factors (TF) that interact with cis-acting regulatory DNA (cis-regulatory elements, CREs), such as promoters and enhancers. Identifying the divergence and conservation among functional gene regulatory elements is an important goal of the comparative evo-devo approach. This is most often done by DNA sequence comparisons of distant or closely related species *in silico*. Recent progress in sequencing of whole genomes of multiple metazoa has provided a rich resource for such an analysis and large numbers of evolutionarily conserved non-coding elements (CNEs) were identified (Hufton et al., 2009). Despite advances in the design of computational algorithms to identify CREs in animal genomes, experimental cis-regulatory analysis remains the most important task although it is time-consuming and laborious. The traditional way to discover CRE experimentally is based on the approach where a sequence suspected to contain gene regulatory activity is placed in the context of a basal promoter driving a reporter gene such as *lacZ* or *EGFP*. In case of developmental control genes, most *in vivo* studies of their cis-regulation have relied on transgenesis as a means to assess the activity of potential promoters or enhancers in the context of developing embryo. To address the extent of cis-regulatory changes and their impact on gene regulatory networks (GRN) among the species of interest, various transgenic experiments are likely going to provide an important mechanistic insight. First of all, a homologous CRE should be tested in each of the model systems individually, in which case transgenesis is

* Corresponding author at: Institute of Molecular Genetics of the AS CR, v.v.i., Videnska 1083, 142 20 Prague 4, Czech Republic. Tel.: +420 241062110; fax: +420 241063125.
E-mail address: kozmik@img.cas.cz (Z. Kozmik).

http://dx.doi.org/10.1016/j.margen.2015.06.003
1874-7787/

performed in the same species from which CRE is isolated. In addition, in some cases the reciprocal transgenic tests may be very informative. In such trans-species transgenic experiments, CREs derived from one organism are tested in another (likely related) organism and *vice versa*. For example, putative CRE identified in the genome of invertebrate chordate amphioxus can be tested in a model vertebrate such as fish, chicken or mouse. Likewise, the well-characterized vertebrate CREs may be interrogated in amphioxus to reveal the presence of ancestral regulatory information. Further resolution of GRN requires identification of epistatic interactions among the network players (members). This can be accomplished by systematic identification and characterization of cis-regulatory elements that control expression of specific genes within a particular GRN.

Cephalochordates (also called amphioxus or lancelets) form one of the three chordate subphyla, along with urochordates and vertebrates. Recent reinterpretation of amphioxus phylogenetic position placing it at the base of chordates and as the sister taxon to vertebrates plus tunicates (Bourlat et al., 2006; Delsuc et al., 2006) highlights the importance of amphioxus in understanding chordate- and vertebrate-specific features at the macroevolutionary scale. Amphioxus genomic, morphological, and developmental characteristics are probably highly similar to those of the chordates (Bertrand and Escriva, 2011). In fact, the adult anatomy of amphioxus is vertebrate-like but much simpler. Amphioxus possesses typical chordate characteristics, such as a dorsal hollow neural tube and notochord, a ventral gut and a pharynx with gill slits, segmented axial muscles, gonads, a post-anal tail, a pronephric kidney, and presumed homologs of the thyroid gland (the endostyle) and adenohypophysis (the so-called pre-oral pit). Although lacking some vertebrate-specific structures, amphioxus has been instrumentally informative in studies of vertebrate innovations such as neural crest, vertebrate head or paired lateral eyes (Bertrand et al., 2011; Vopalensky et al., 2012; Yu et al., 2008). The anatomical simplicity has been mirrored by the simplicity of the amphioxus genome. It is generally well accepted that the two rounds of whole-genome duplication occurred specifically in the vertebrate lineage and that the genome of amphioxus provides a useful glimpse at the 'pre-duplicated' version of the ancestral chordate genome (Holland et al., 2008).

Here, we focus on animal transgenesis as one way of cis-regulatory analysis of the amphioxus genome. We review the existing transgenic studies performed in amphioxus and vertebrates using CREs (promoters and enhancers) derived from genomes of *Branchiostoma floridae*, *Branchiostoma belcheri* and *Branchiostoma lanceolatum*. We describe the current limitations of amphioxus transgenesis, provide the first example of successful transposon-mediated transgenesis and propose future directions. Finally, we discuss the potential of using amphioxus transgenesis in comparative studies aimed at understanding the cis-regulation in the chordate lineage.

2. Transgenic studies in amphioxus

To date, the genomes of the two species of *Branchiostoma* have been sequenced. First of all, a complete sequence of 520-megabase genome of the Florida lancelet *B. floridae* was determined (Holland et al., 2008; Putnam et al., 2008) and confirmed that cephalochordates had not undergone the two rounds of whole-genome duplication that occurred in vertebrates. This has opened up the possibility to quickly identify and locate genomic regions of interest. Recently, the genome sequence of *B. belcheri* was published (Huang et al., 2014), further expanding the genomic resources and allowing cross-comparative analysis of non-coding regions of the cephalochordate genomes. The most straightforward way to interrogate amphioxus cis-regulatory elements in the context of amphioxus embryo is by transgenic studies. Foreign DNA can be introduced into amphioxus embryos by microinjection of unfertilized eggs (Holland and Yu, 2004; Liu et al., 2013; Yu et al., 2004). However, only a handful of CREs have been tested in amphioxus so far (see Table 1).

Table 1
List of CREs experimentally verified by transgenesis in amphioxus.

Amphioxus gene	Type of CRE	Functional	Reference
FoxD	Promoter	Yes	Yu et al. (2004)
Engrailed	Promoter, enhancer	Yes	Beaster-Jones et al. (2007)
Znf504/703	CNE (enhancer)	Yes	Holland et al. (2008)
Actin	Promoter	Yes	Feng et al. (2014)
Znf504/703	CNE (enhancer)	Yes	Feng et al. (2014)
Chordin	Promoter	Yes	This study

In their pioneering study, Yu et al. (2004) have shown that expression of a lacZ reporter construct including 6.3 kb of the amphioxus *FoxD* upstream regulatory region recapitulates expression of the endogenous gene in the nerve cord, somites, and notochord. Further analysis identified a 1.6 kb region necessary for the nerve cord and somite expression, while the remaining 4.7 kb of the upstream regulatory region was sufficient for notochord expression. The shortest tested fragment encompassing only 0.7 kb of the proximal promoter did not show any activity and may represent a suitable minimal promoter for future enhancer tests (see Discussion). Cis-regulatory analysis of amphioxus *FoxD* in vertebrates (Yu et al., 2008) provided additional insight into the evolution of neural crest and paralogous *FoxD* genes after duplication in the vertebrate lineage. Amphioxus has a single copy of the *FoxD* gene, whereas vertebrate genomes carry multiple paralogous *FoxD* genes. Of these paralogs, only FoxD3 has been co-opted into the neural crest gene regulatory network by vertebrate-specific acquisition of a cis-regulatory element directing *FoxD3* expression to the neural crest (Van Otterloo et al., 2013). Such element is likely not present in the 6.3 kb of amphioxus *FoxD* upstream regulatory region capable of recapitulating the endogenous amphioxus *FoxD* expression since the corresponding reporter gene failed to direct expression to chick neural crest (Yu et al., 2008).

A study of cis-regulation of the amphioxus *engrailed* gene provided an insight into the evolution of muscle-specific enhancer (Beaster-Jones et al., 2007). The upstream regulatory region (7.8 kb) of amphioxus *engrailed* directs expression coincident with the areas of expression of the endogenous gene. Within this region, a 1.2 kb muscle-specific enhancer was identified that shows sequence similarity to the mouse *En2* muscle enhancer. Interestingly, the amphioxus enhancer directs expression not only to somites in amphioxus, but also to larval muscles in *Ciona intestinalis*, despite the fact that endogenous *engrailed* gene is not expressed in muscle tissue of *Ciona*. This result illustrates the fact that the transcription factors and gene regulatory networks are generally highly conserved. Constraints imposed on gene regulatory networks directing expression to specific tissues may allow the loss or gain of some components, but overall the gene regulatory networks remain largely intact (Hinman et al., 2003). The lack of native *engrailed* expression in *C. intestinalis* muscle suggests that this gene has lost the muscle-specific enhancer that is conserved in amphioxus and mouse, leading to the loss of *engrailed* from muscle-specific GRN in *Ciona*.

The cis-regulatory activity of CNE located near the amphioxus *ZNF503/703* gene was tested by transgenesis in amphioxus and mice and proved positive as an enhancer in both animal models (Holland et al., 2008). The amphioxus *ZNF503/703* reporter gene construct was highly active in amphioxus notochord and somites and at a lower level in the ectoderm and central nervous system. This reporter activity coincides with known expression of the endogenous amphioxus *ZNF503/703* gene in the central nervous system, somites, notochord, and pharyngeal endoderm (Holland et al., 2008). It is interesting to note that the two corresponding CNEs derived from human *ZNF503* and *ZNF703* genes were also investigated by amphioxus and mouse transgenesis (Holland et al., 2008). The expression driven by the three CNEs (amphioxus *ZNF503/703*, human *ZNF503*, human *ZNF703*) was not identical in the two species. Although human *ZNF503* and *ZNF703* CNEs directed tissue-specific expression in both amphioxus and mouse, the pattern was distinct from the one produced by the

Please cite this article as: Kozmikova, I., Kozmik, Z., Gene regulation in amphioxus: An insight from transgenic studies in amphioxus and vertebrates, Mar. Genomics (2015), http://dx.doi.org/10.1016/j.margen.2015.06.003