Evo-Devo gene toolkit update: at least seven Pax transcription factor subfamilies in the last common ancestor of bilaterian animals

Markus Friedrich[a,b,*]

[a] Department of Biological Sciences, Wayne State University, 5047 Gullen Mall, Detroit, MI 48202, USA
[b] Department of Anatomy and Cell Biology, Wayne State University, School of Medicine, 540 East Canfield Avenue, Detroit, MI 48201, USA

*Author for correspondence (e-mail: mf@biology.biosci.wayne.edu)

INTRODUCTION

The transcription factor gene family defined by the Paired box (Pax) DNA-binding protein domain represents an important component of the developmental genetic tool kit that organizes animal body plan diversity (Knoll and Carroll, 1999; Chi and Epstein, 2002; Carroll et al. 2013). Reconstructing the diversification of *Pax* gene subfamilies in relation to animal body plan evolution has thus been of considerable interest. The progress made over the past two decades has been relatively incremental, however, due to the evolutionary and functional complexity of this exceptionally modular transcription factor gene family, as most recently summarized by Underhill (2012). Like in the case of the much larger super family of homeodomain transcription factors, gene tree reconstruction approaches have been challenged by the small size of the highly constrained Pax domain (128 amino acids). This handicap is only mildly mitigated by the presence of additional conserved domains in select *Pax* gene subfamilies, like, most dramatically, a second DNA binding region in the form of a homeodomain (Fig. 1). The final challenge to the reconstruction of *Pax* gene trees is the dynamic of homolog loss and duplication events to the effect that the scattered conservation of some of the bilaterian subfamilies obscured their ancientness prior to the advent of broader taxonomic sampling studies (Paixão-Côrtes et al. 2013; Feiner et al. 2014) (Fig. 1). Given that the last common ancestor (LCA) of protostomes and deuterostomes represents a particularly crucial node in the tree of life, it is important to note that studies of *Pax* gene family evolution continue to differ in the recognition of subfamilies that map to this critical juncture (Hill et al. 2010; Wang et al. 2010; Underhill, 2012; Feiner et al. 2014; Franke et al. 2015).

The earliest studies of *Pax* gene diversity recognized four subfamilies in the bilaterian LCA, named after the gene names of their vertebrate constituents: *Pax1/9*, *Pax2/5/8*, *Pax3/7*, and *Pax4/6* (Fig. 1) (Breitling and Gerber, 2000; Sun et al. 2002). Subsequent investigations of transcription factor diversity in non-chordate deuterostomes increased this number to five,

based on the discovery of orthologs of the *Drosophila Pox neuro* factor in the echinoderm *Strongylocentrotus purpuratus* (sea urchin) and the hemichordate *Saccoglossus kowalevski* (acorn worm) (Howard-Ashby et al. 2006; Lowe et al. 2006; Matus et al. 2007; Freeman et al. 2008). The studies in sea urchin also produced evidence for yet another ancient branch in the *Pax* gene tree, which had first been discovered in *Drosophila*: that of the *eyegone* (*eyg*) subfamily (Fig. 1) (Howard-Ashby et al. 2006; Matus et al. 2007). The *eyg* gene was initially misinterpreted as a *Pax4/6* subfamily member due to its dramatic eye depletion phenotype in *Drosophila* and structural similarities to the truncated Pax domain in the vertebrate *Pax6 (5a)* isoform (Jun et al. 1998; Pichaud and Desplan, 2002). Although this notion persists (Tanaka-Matakatsu et al. 2015), recent studies have revealed that *eyg* antagonizes rather than promotes eye development in non-*Drosophila* insects (Zarinkamar et al. 2011) and is characterized by a canonical Pax domain in the acorn worm (Friedrich and Caravas, 2011). The conservation of *eyg* in the acorn worm in addition to sea urchin also further cemented the bilaterian status of the *Pax eyg* gene subfamily, which, however, continues to go unnoticed in recent *Pax* gene family surveys (Hill et al. 2010; Wang et al. 2010; Underhill, 2012; Franke et al. 2015). This oversight, of course, is explained by the biasing effect of the absence of *eyg* in vertebrates due to evolutionary gene loss during early chordate evolution.

The most recent addition to the bilaterian *Pax* gene subfamily repertoire has been emphasized in a study of *Pax* gene conservation and expression in velvet worms, a protostome phylum closely allied with arthropods (Fig. 1) (Franke et al. 2015). Deep sequencing of an embryonic cDNA library of the velvet worm model species *Euperipatoides rowelli* uncovered homologs of most of the above-mentioned *Pax* gene subfamilies except for *eyg*, which was not specifically explored. Interestingly, the *E. rowelli* transcriptome includes an additional *Pax* gene that solidly groups with members of a previously suspected subfamily (Hill et al. 2010), now designated *Pax alpha* (Franke et al. 2015). Similar to the *Pox neuro* and *eyg*
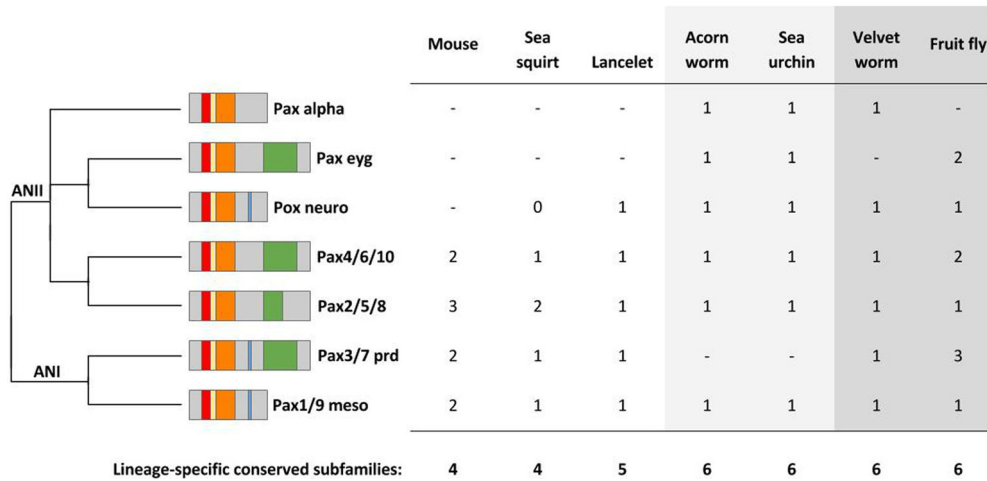
	Mouse	Sea squirt	Lancelet	Acorn worm	Sea urchin	Velvet worm	Fruit fly
Pax alpha	-	-	-	1	1	1	-
Pax eyg	-	-	-	1	1	-	2
Pox neuro	-	0	1	1	1	1	1
Pax4/6/10	2	1	1	1	1	1	2
Pax2/5/8	3	2	1	1	1	1	1
Pax3/7 prd	2	1	1	-	-	1	3
Pax1/9 meso	2	1	1	1	1	1	1
Lineage-specific conserved subfamilies:	4	4	5	6	6	6	6

Fig. 1. Bilaterian Pax transcription factor gene subfamilies. Color code of protein domains: red, PAI domain; yellow, linker domain; orange, RED domain; blue, octapeptide motif; green, homeodomain. Numbers in tables indicate singleton ortholog conservation or subfamily expansions that are lineage specific. Shading code in table: white, chordate species; light gray, non-chordate deuterostomes; dark gray; protostome species. Definition of ancestral superfamilies I and II, i.e., ANI and ANII, respectively, based on Sun et al. (1997, 2002). Other aspects of the tentative gene tree adapted from Hill et al. (2010), Friedrich and Caravas (2011), Underhill (2012), and Franke et al. (2015). Homolog conservation compiled from Howard-Ashby et al. (2006), Bassham et al. (2008), Freeman et al. (2008), Wang et al. (2010) and Paps et al. (2012).

subfamilies, the latter is absent from vertebrates but is otherwise conserved in a broad range of bilaterian lineages including non-chordate deuterostomes (i.e., sea urchin and acorn worm) (Fig. 1). Thus, combined, the total count of Pax transcription factors in Bilateria now amounts to seven.

Taken together, the available data paint a more complex picture of *Pax* gene organization in the bilaterian LCA than still commonly considered (Underhill, 2012; Feiner et al. 2014; Franke et al. 2015). It further becomes apparent that invertebrates are generally richer in ancient *Pax* gene subfamilies than vertebrates, being only one subfamily short of the complete set of bilaterian *Pax* genes (Fig. 1). Of note, the *Pax* gene count of velvet worms might increase further, pending the outcome of specific searches for *eyg* orthologs (Franke et al. 2015).

On a final note: As *Pax* gene families have expanded with the increase of studies tracking their origins, so have approaches to naming, classifying, and archiving members of this gene family. As a result, it has become challenging and time-consuming to track the documentation of specific homologs. Paps et al. (2012), for instance, list three *Pax4/6* homologs and one *eyg* homolog in *Drosophila*, while there are two *Pax4/6* and two *eyg* homologs in this species (Bao and Friedrich, 2009). The same study lists six *Pax* genes for the sea urchin while the current sea urchin genome database contains seven distinct *Pax* gene models (Cameron et al. 2009). It thus seems timely for the community to build a consolidated *Pax* homolog database similar to that for homeodomain transcription factors in order to aid future studies of *Pax* gene function and diversity (Zhong and Holland, 2011).

REFERENCES

Bao, R., and Friedrich, M. 2009. Molecular evolution of the Drosophila retinome: exceptional gene gain in the higher Diptera. *Mol. Biol. Evol.* 26: 1273–1287.

Bassham, S., Cañestro, C., and Postlethwait, J. H. 2008. Evolution of developmental roles of Pax2/5/8 paralogs after independent duplication in urochordate and vertebrate lineages. *BMC Biol.* 6: 35.

Breitling, R., and Gerber, J. K. 2000. Origin of the paired domain. *Dev. Genes Evol.* 210: 644–650.

Cameron, R. A., Samanta, M., Yuan, A., He, D., and Davidson, E. 2009. SpBase: the sea urchin genome database and web site. *Nucleic Acids Res.* 37: D750–D754.

Carroll, S. B., Grenier, J. K., and Weatherbee, S. D. 2013. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design.* John Wiley & Sons, Oxford.

Chi, N., and Epstein, J. A. 2002. Getting your Pax straight: Pax proteins in development and disease. *Trends Genet.* 18: 41–47.

Feiner, N., Meyer, A., and Kuraku, S. 2014. Evolution of the vertebrate Pax4/6 class of genes with focus on its novel member, the Pax10 gene. *Genome Biol. Evol.* 6: 1635–1651.

Franke, F. A., Schumann, I., Hering, L., and Mayer, G. 2015. Phylogenetic analysis and expression patterns of Pax genes in the onychophoran Euperipatoides rowelli reveal a novel bilaterian Pax subfamily. *Evol. Dev.* 17: 3–20.

Freeman, R. M., Jr, et al. 2008. cDNA sequences for transcription factors and signaling proteins of the hemichordate Saccoglossus kowalevskii: efficacy of the expressed sequence tag (EST) approach for evolutionary and developmental studies of a new organism. *Biol. Bull.* 214: 284–302.

Friedrich, M., and Caravas, J. 2011. New insights from hemichordate genomes: prebilaterian origin and parallel modifications in the paired domain of the Pax gene eyegone. *J. Exp. Zool. B Mol. Dev. Evol.* 316: 387–392.

Hill, A., et al. 2010. Origin of Pax and Six gene families in sponges: single PaxB and Six1/2 orthologs in Chalinula loosanoffi. *Dev. Biol.* 343: 106–123.

Howard-Ashby, M., Materna, S. C., Brown, C. T., Chen, L., Cameron, R. A., and Davidson, E. H. 2006. Identification and characterization

of homeobox transcription factor genes in Strongylocentrotus purpuratus, and their expression in embryonic development. *Dev. Biol.* 300: 74–89.

Jun, S., Wallen, R. V., Goriely, A., Kalionis, B., and Desplan, C. 1998. Lune/eye gone, a Pax-like protein, uses a partial paired domain and a homeodomain for DNA recognition. *Proc. Natl. Acad. Sci. USA* 95: 13720–13725.

Knoll, A. H., and Carroll, S. B. 1999. Early animal evolution: emerging views from comparative biology and geology. *Science* 284: 2129–2137.

Lowe, C. J., et al. 2006. Dorsoventral patterning in hemichordates: insights into early chordate evolution. *PLoS Biol.* 4: e291.

Matus, D. Q., Pang, K., Daly, M., and Martindale, M. Q. 2007. Expression of Pax gene family members in the anthozoan cnidarian, Nematostella vectensis. *Evol. Dev.* 9: 25–38.

Paixão-Côrtes, V. R., Salzano, F. M., and Bortolini, M. C. 2013. Evolutionary history of chordate PAX genes: dynamics of change in a complex gene family. *PLoS ONE* 8: e73560.

Paps, J., Holland, P. W. H., and Shimeld, S. M. 2012. A genome-wide view of transcription factor gene diversity in chordate evolution: less gene loss in amphioxus? *Brief. Funct. Genomics* 11: 177–186.

Pichaud, F., and Desplan, C. 2002. Pax genes and eye organogenesis. *Curr. Opin. Genet. Dev.* 12: 430–434.

Sun, H., et al. 2002. Identification of essential amino acid changes in paired domain evolution using a novel combination of evolutionary analysis and in vitro and in vivo studies. *Mol. Biol. Evol.* 19: 1490–1500.

Sun, H., et al. 1997. Evolution of paired domains: isolation and sequencing of jellyfish and hydra Pax genes related to Pax-5 and Pax-6. *Proc. Natl. Acad. Sci.* 94: 5156–5161.

Tanaka-Matakatsu, M., Miller, J., and Du, W. 2015. The homeodomain of Eyeless regulates cell growth and antagonizes the paired domain-dependent retinal differentiation function. *Protein Cell* 6: 68–78.

Underhill, D. A. 2012. PAX proteins and fables of their reconstruction. *Crit. Rev. Eukaryot. Gene Expr.* 22: 161–177.

Wang, W., Zhong, J., and Wang, Y.-Q. 2010. Comparative genomic analysis reveals the evolutionary conservation of Pax gene family. *Genes Genet. Syst.* 85: 193–206.

Zarinkamar, N., Yang, X., Bao, R., Friedrich, F., Beutel, R., and Friedrich, M. 2011. The Pax gene eyegone facilitates repression of eye development in Tribolium. *Evodevo* 2: 8.

Zhong, Y.-F., and Holland, P. W. H. 2011. HomeoD B2: functional expansion of a comparative homeobox gene database for evolutionary developmental biology. *Evol. Dev.* 13: 567–568.