Mechanisms of Development 144 (2017) 163–170

Contents lists available at ScienceDirect

Mechanisms of Development

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

Evolution of the Rax family of developmental transcription factors in vertebrates

Daniela P. Orquera [a], Flávio S.J. de Souza [a,b,*]

[a] Instituto de Investigaciones en Ingeniería Genética y Biología Molecular, Consejo Nacional de Investigaciones Científicas y Técnicas, 1428 Buenos Aires, Argentina
[b] Departamento de Fisiología, Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 27 February 2016
Received in revised form 1 November 2016
Accepted 7 November 2016
Available online 10 November 2016

Keywords:
Eye development
Hypothalamus
Gene duplication
rx1
rx2
rx3

ABSTRACT

Rax proteins comprise a small family of paired-type, homeodomain-containing transcription factors with essential functions in eye and forebrain development. While invertebrates possess only one *Rax* gene, vertebrates can have several *Rax* paralogue genes, but the evolutionary history of the members of the family has not been studied in detail. Here, we present a thorough analysis of the evolutionary relationships between vertebrate *Rax* genes and proteins available in diverse genomic databases. Phylogenetic and synteny analyses indicate that *Rax* genes went through a duplication in an ancestor of all jawed vertebrates (Gnathostomata), giving rise to the ancestral vertebrate *Rax1* and *Rax2* genes. This duplication event is likely related to the proposed polyploidisations that occurred during early vertebrate evolution. Subsequent genome-wide duplications in the lineage of ray-finned fish (Actinopterygii) originated new *Rax2* paralogues in the genomes of teleosts. In the lobe-finned fish lineage (Sarcopterygii), the N-terminal octapeptide domain of Rax2 was lost in a common ancestor of tetrapods, giving rise to a shorter version of Rax2 in this lineage. Within placental mammals, the *Rax2* gene was lost altogether in an ancestor of rodents and lagomorphs (Glires). Finally, we discuss the scientific literature in the light of *Rax* gene evolution and propose new avenues of research on the function of this important family of transcriptional regulators.

1. Introduction

The *Rax* - also called *Rx* - gene was first described in three independent reports that appeared in 1997. Furukawa et al. (1997) identified Rax as a paired-type homeodomain protein that was expressed in the anterior neural plate and later in the eyes and hypothalamus of mouse embryos, while Casarosa et al. (1997) described a *Rax* homologue, *Xrx1*, from the frog *Xenopus laevis*. Mathers et al. (1997) described the expression of *Rax* homologues in the mouse, zebrafish and *Xenopus laevis*, showing that expression pattern of the gene was conserved in vertebrates. They also found that engineered mice that lacked *Rax* expression were eyeless and had forebrain abnormalities, establishing *Rax* as a gene with crucial functions in early mammalian development (Mathers et al. 1997). Indeed, a hypomorph *Rax* allele is found in the *eyeless* (*ey1*) mouse, a mutant strain that lacks eyes and optic cups (Tucker et al. 2001). Early surveys identified two or three *Rax* genes in the genomes of some vertebrate model organims like teleost fishes, *X. laevis* and chicken (Mathers et al. 1997; Ohuchi et al. 1999), and a second *Rax*-like gene (called *RAX2* or *QRX*) was later identified in humans (Wang et al. 2004). Interestingly, this second gene is absent from the mouse genome, which harbours only one *Rax* gene (Wang et al. 2004; Zhong and Holland 2011). Both human *Rax* paralogues are important for eye development, since mutations in *RAX* (Voronina et al. 2004; Gonzalez-Rodriguez et al., 2010; Lequeux et al. 2008; Abouzeid et al. 2012) and *RAX2* (Wang et al. 2004; Yang et al. 2015) are associated with absence of eyes or eye defects in human patients.

Apart from studies in mouse mutants and human patients, the function of *Rax* genes has been studied in several vertebrate models, including teleosts like medaka and zebrafish (Winkler et al. 2000; Loosli et al. 2003; Rojas-Muñoz et al. 2005), the frogs *X. laevis* and *X. tropicalis* (Andreazzoli et al. 1999, 2003; Pan et al. 2006; Wu et al. 2009; Giudetti et al. 2014; Fish et al. 2014) and the chicken (Ohuchi et al. 1999; Chen and Cepko 2002). *Rax* genes have also been studied in the context of eye evolution, as in African cychlids (Schulte et al. 2014) and the cavefish, *Astyanax mexicanus* (McGaugh et al. 2014). The interpretation and comparison between all these studies is made difficult by the inconsistent gene names and the lack of a thorough evolutionary analysis and classification of vertebrate *Rax* genes, even when the phylogeny of a few members of the Rax family has been described (Wang et al. 2004; Wu et al. 2009; Schulte et al. 2014). Here, we analyse *Rax* genes and proteins from several model organisms and sequenced genomes and present a thorough picture of the evolutionary history of this important group of transcription factors.

* Corresponding author at: Instituto de Investigaciones en Ingeniería Genética y Biología Molecular, Consejo Nacional de Investigaciones Científicas y Técnicas, 1428 Buenos Aires, Argentina.
E-mail address: fsouza@dna.uba.ar (F.S.J. de Souza).

http://dx.doi.org/10.1016/j.mod.2016.11.002
0925-4773/

2. Results

2.1. Analysis of Rax protein sequences

Invertebrates, including cnidarians, protostomes and chordates like amphioxus and sea squirts, possess only one *Rax* gene (Mazza et al. 2010), but vertebrates can have more than one *Rax* paralogue. To obtain a clearer picture of the evolution of the *Rax* gene family, we scoured sequenced vertebrate genomes for *Rax* homologues, paying special attention to species in key phylogenetic positions. In particular, we obtained predicted Rax peptide sequences from species like the Japanese lamprey (*Lethenteron japonicum*), a jawless vertebrate; the elephant shark (*Callorhinchus milii*), a cartilaginous fish; the spotted gar (*Lepisosteus oculatus*), a ray-finned fish representing a sister group to teleost fishes; and the West Indian Ocean coelacanth (*Latimeria chalumnae*) a lobe-finned fish that represents a sister lineage to tetrapods (amphibians, reptiles, birds and mammals). In addition, we also retrieved Rax sequences from teleost fishes and several tetrapod genomes, in particular from the mammalian phylogenetic tree. Fig. 1 shows a schematic tree of vertebrate phylogeny with the main events of *Rax* evolution that are discussed below.

An alignment of vertebrate Rax proteins distinguishes two groups of these proteins, namely a Rax1 and a Rax2 subgroup (Supplementary Fig. S1), which are also evidenced in phylogenetic and genomic analyses (see below). As observed previously (Furukawa et al. 1997; Wu et al. 2009; Mazza et al. 2010), Rax proteins possess three conserved regions of special interest: i) a N-terminal octapeptide motif that is also found in transcription factors of the Pax family, among others; ii) the central homeodomain (HD) region that binds to DNA, and iii) the C-terminal OAR domain, so called due to its presence in the transcription factors Otp, Aristaless and Rax, among others (Furukawa et al. 1997). All Rax proteins in our alignment possess the HD and the OAR domains, but Rax2 proteins of all tetrapods are shorter and lack the octapeptide motif. Since both Rax1 and Rax2 of non-tetrapods have the octapeptide (Suppl. Fig. 1), this indicates that the loss of this segment happened in a lobe-finned fish (Fig. 1).

2.2. Rax phylogeny

To understand the evolutionary history of vertebrate *Rax* genes, we estimated the phylogeny of Rax proteins using the maximum likelihood method as implemented in PhyML 3.0 (Guindon et al. 2010). Statistical support for the branches was estimated using the bootstrap (Felsenstein 1985). Since many Rax proteins lack the octapeptide domain, the alignment used for phylogeny estimation was mostly restricted to the homeodomain and C-terminal regions (including the OAR motif) of the proteins.

The resulting phylogenetic tree (Fig. 2) shows that Rax proteins can be divided into two large groups, which we named the Rax1 and Rax2 subgroups. The Japanese lamprey Rax protein falls within the Rax2 subgroup. No other *Rax* gene could be retrieved from the Japanese lamprey draft genome, and no *Rax* gene was found in the draft genome of the sea lamprey, *Petromyzon marinus* (Smith et al. 2013). Due to the incomplete nature of the genome assemblies of agnathans, however, it cannot be discarded that another *Rax* gene might be identified in the future in this clade. The sequence of the Rax protein of the Japanese lamprey is quite divergent in relation to the other Rax proteins, as indicated by the long branch length of the lamprey Rax in the tree of Fig. 2. It is important to note that the Rax1 and Rax2 subgroups are retrieved even when the lamprey sequence is excluded from the phylogenetic analysis (data not shown), indicating that this divergent sequence is not altering the phylogenetic reconstruction and the resulting tree topology.

As for jawed vertebrates, all species analysed possess Rax1 and Rax2 representatives, except for some mammals (see below). The most basal phylogenetic group is represented by the elephant shark, a cartilaginous fish. The elephant shark possesses two *Rax* genes, one in each subgroup. The basal ray-finned fish, the spotted gar, as well as the basal lobe-finned fish, the coelacanth, also possess one *Rax1* and one *Rax2* representative each. This indicates that the duplication of the *Rax* genes occurred very early in vertebrate evolution, and that the genomes of the ancestors of both teleosts and tetrapods carried two *Rax* genes (Fig. 1). However, teleosts have more than two *Rax* genes; two that belong to the Rax2 subgroup, namely *rx1* and *rx2*, and one member of the Rax1 subgroup, *rx3*. This is true for the zebrafish (*Danio rerio*), medaka (*Oryzias latipes*) and pufferfish (*Tetraodon nigroviridis*), which are shown in the tree in Fig. 2, as well as for other species like the Mexican blind cavefish (*Astyanax mexicanus*), the Nile tilapia (*Oreochromis niloticus*) and other teleosts (data not shown). This observation, together with synteny analyses (see below) suggests that the three *Rax* genes present today in teleosts originated in the whole-genome duplication (WGD) that happened in a teleost ancestor after the divergence of the spotted gar lineage (Fig. 1; Glasauer and Neuhauss 2014). *rx1* and *rx2* were retained after WGD, while one *rx3* paralogue was subsequently lost from the teleost ancestor.

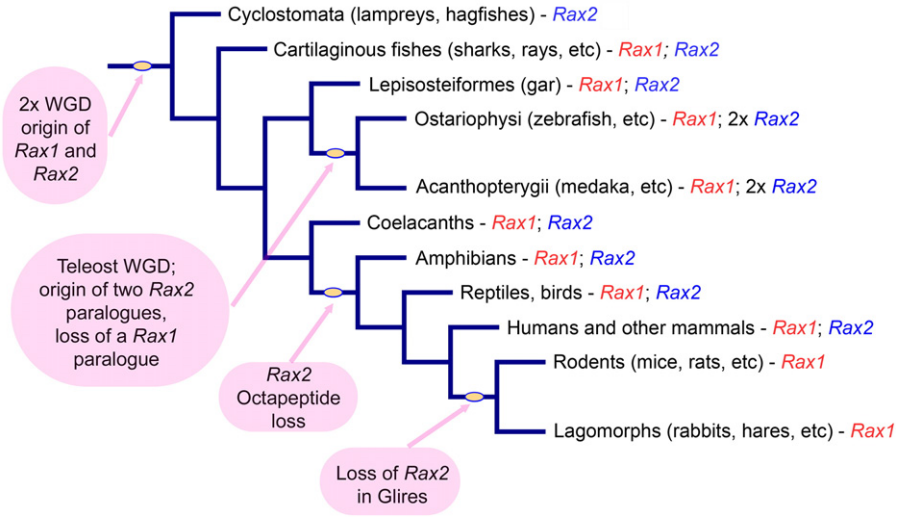

Fig. 1. Rax evolution in vertebrates. The tree schematically represents the phylogenetic relationships of vertebrate clades relevant for the present study. The relative timing of important events in the evolution of *Rax* genes are indicated. Note that the timing of the two rounds of whole genome duplications (2× WGD) in a vertebrate ancestor is still controversial; it is represented here as happening in a common ancestor of cyclostomes and jawed vertebrates (see Smith et al. 2013; Mehta et al. 2013; Smith and Keinath 2015). The presence and number of *Rax1* (red) and *Rax2* (blue) paralogues in each clade is indicated.