Biochimica et Biophysica Acta 1859 (2016) 770–779

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Evolution and functions of Oct4 homologs in non-mammalian vertebrates☆

Daria Onichtchouk *

Developmental Biology Unit, Institute of Biology I, Faculty of Biology, and Center for Biological Signaling Studies (BIOSS), Albert-Ludwigs-University, Freiburg, Germany

ARTICLE INFO

Article history:
Received 6 February 2016
Received in revised form 22 March 2016
Accepted 23 March 2016
Available online 4 April 2016

Keywords:
Pct4
Pou5f1
Pou5f3
Pluripotency
Zebrafish
Xenopus

ABSTRACT

PouV class transcription factor Oct4/Pou5f1 is a central regulator of indefinite pluripotency in mammalian embryonic stem cells (ESCs) but also participates in cell lineage specification in mouse embryos and in differentiating cell cultures. The molecular basis for this versatility, which is shared between Oct4 and its non-mammalian homologs Pou5f1 and Pou5f3, is not yet completely understood. Here, I review the current understanding of the evolution of PouV class transcription factors and discuss equivalent and diverse roles of Oct4 homologs in pluripotency, differentiation, and cell behavior in different vertebrate embryos. This article is part of a Special Issue entitled: The Oct Transcription Factor Family, edited by Dr. Dean Tantin.

1. Introduction

The POU domain is a bipartite DNA-binding domain which consists of the C-terminal POU-type homeodomain (POU_H) and an N-terminal POU-specific domain (POU_S). The POU_S domain is unique to POU proteins; the POU_H domain is related to the classic homeodomain, first identified in *Drosophila* homeotic gene products [46]. Both subdomains are indispensable for sequence-specific DNA binding. In spite of that, POU_H and POU_S segments are structurally independent domains connected via a flexible linker: this organization is thought to contribute to the flexibility of protein–DNA and protein–protein interactions of POU transcription factors [128,143]. The POU-domain gene family is divided into 6 classes of transcription factors, PouI-PouVI, involved in cellular decisions in multiple cell lineages [106,143]. Oct proteins are a subclass of the POU transcription factor family including PouII, PouIII, and PouV classes and the founding members of the family were mammalian Oct1 and Oct2 proteins [127]. Oct proteins recognize an 8 bp consensus sequence [ATGC(A/T)AAT] termed the "octamer motif" ([55,64], reviewed in [135]). The 5′ ATGC motif associates with the POU-specific domain, while the 3′ half site (A/T)AAT associates with the POU-type homeodomain. The Oct4 protein encoded by the founding PouV class member, *Pou5f1*, was first purified biochemically [118] and then cloned independently in a screen for transcription factors, which are expressed only in pluripotent cells such as embryonic stem (ES) cells and inner cell mass [113], and a screen for factors, which bind to "octamer motif" sequences, previously shown to be enriched in mouse embryonic enhancers [103,119]. Originally, Oct4 expression was found to be restricted to female germ line, totipotent, and pluripotent stem cells in pre-implantation embryos ([103,113,117,119]); however, later studies demonstrated that the Oct4 protein is still present after implantation during gastrulation and afterwards [27,29]. Functionally, Oct4 appears to be a gatekeeper for cell pluripotency [126] and activation of endogenous Oct4 is necessary for inducing pluripotent cells from somatic cells in reprogramming experiments ([131]; recently reviewed by [108]). Due to the immense clinical potential of reprogramming approaches and application of embryonic stem cells, Oct4 received high attention from the scientific community and became one of the most frequently published transcription factors today. A search of the PubMed database (National Center for Biotechnology Information, NIH, Bethesda) for the terms "Oct4" or "Pou5f1" shows that the number of citations per year has been increasing apace from 2000, and exceeded 500 citations per year in 2013 (Fig. 1). The majority of Oct4/Pou5f1 research is performed in pluripotent cell cultures. Studies of vertebrate Oct4 homologs outside of mammalian clades received less attention: the number of publications on non-mammalian PouV class genes does not exceed a total of 70 since the first publication ([47], Fig. 1).

Paradoxically, on top of its well-known central role in establishing and maintaining pluripotency in cell cultures, Oct4 was recently implicated in differentiation of all embryonic lineages from pluripotent cells [2,107]. Early research on mouse Oct4-deficient embryos [95] suggested

☆ This article is part of a Special Issue entitled: The Oct Transcription Factor Family, edited by Dr. Dean Tantin.

* Institute of Biology 1, Albert-Ludwigs-University Freiburg, Hauptstrasse 1, D-79104 Freiburg, Germany.
E-mail address: daria.onichtchouk@biologie.uni-freiburg.de.

http://dx.doi.org/10.1016/j.bbagrm.2016.03.013
1874-9399/

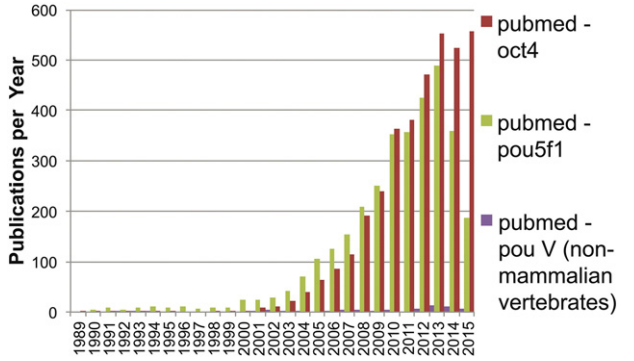

Fig. 1. The time dynamics of publications either using the terms "Oct4" and "Pou5f1" or the names of non-mammalian PouV class homologs* (*including all cases, when a non-mammalian gene was named Pou5f1 or Oct4).

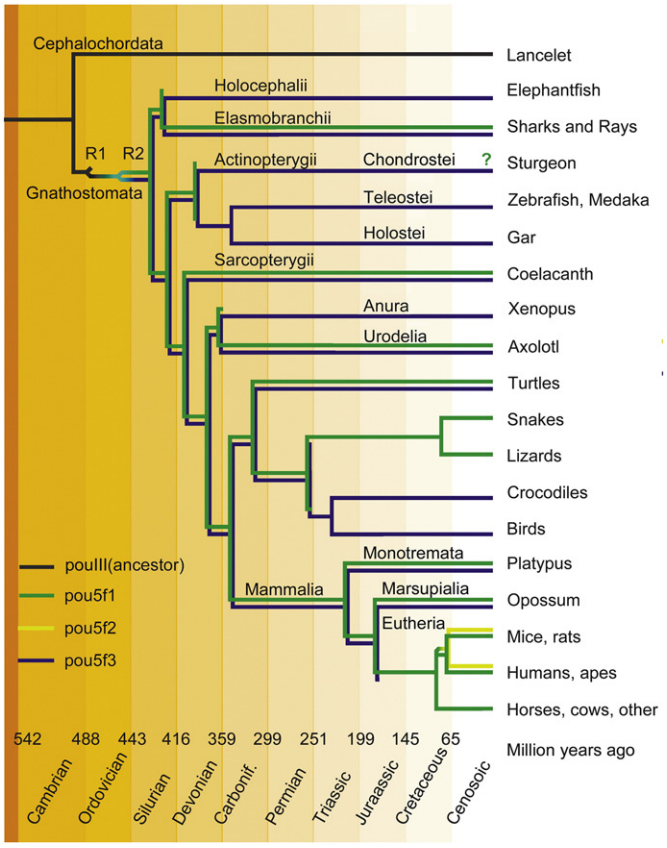

Fig. 2. Evolution of PouV class genes. Ancient PouV gene duplication and extinction of paralogs *pou5f1* and *pou5f3* according to [36] was combined with time scale of the hypothetical splits between main taxa [91,129] and the Tree of Life drawn by Olaf Olander, which is available at http://www.tellapallet.com/tree_of_life.htm under Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. R1 and R2 designate whole genome duplications [26].

that Oct4 is critical for the first differentiation event in development: separation of inner cell mass and trophectoderm. This is in line with experiments in mouse ES cells, where the reciprocal inhibition between transcription factors Oct4 and Cdx2 is necessary for trophectoderm specification [98,100]. In the context of pluripotent cell cultures, different amounts of Oct4 elicit opposite effects: low levels of Oct4 lock the cells in a naïve pluripotent state, while higher Oct4 levels promote differentiation [62,97,107]. In mouse embryos, maternal Oct4 is not crucial for the establishment of pluripotency, as demonstrated by recent experiments using conditional genetic ablation [27,38,80,148]. In contrast to its role in ES cells, Oct4 is involved in the early steps of extra-embryonic endoderm specification in mouse blastocysts [38,80,148] and promotes cell viability at gastrulation [27]. It is difficult to understand how one transcription factor can regulate such apparently opposite processes as pluripotency maintenance and lineage-specific differentiation, unless we assume Oct4 operating at successive developmental stages within different gene networks. While the pluripotency gene network maintaining the stem cell character and self-renewal in ES cells is well characterized (reviewed in [154]), Oct4-dependent regulatory mechanisms of cell fate specification are still unclear [27,38].

Well-established model organisms of developmental biology, zebrafish, medaka, *Xenopus*, and chick, become ever more important to elucidate the Oct4-dependent molecular developmental mechanisms, due to their experimental advantages such as extrauterine development, large size of their embryos, and amenability to *in-vivo* experimental manipulation. The question of how comparable the functions of PouV class factors across vertebrates are has gained focused interest within the scientific community and is now a topic of intense debate [17,34–37,99]. In this review, I will summarize the current knowledge about evolution and biological roles of vertebrate PouV class genes and discuss it with an emphasis on their equivalent and diverse functions, related to mammalian Oct4.

2. Vertebrate Oct4 homologs, Pou5f1, and Pou5f3

Gnathostomes (jawed vertebrates) possess one or both closely related, maternally expressed PouV class proteins: Pou5f1 (the ortholog of mouse Oct4) and Pou5f3 (a paralog of mouse Oct4). Pou5f1 and Pou5f3 will be collectively referred below as "Oct4 homologs." The presence of two PouV class genes in the vertebrate lineage is supported by i) different arrangement of genes around *pou5f3* and *pou5f1* loci across different taxa (syntenic evidence) and by ii) two sequence-based signatures: Oct4/Pou5f1, but not Pou5f3, possesses the conserved sequence MAGH at its N-terminus and the deletion of a single arginine residue within the POU_S domain [35]. A current view on Oct4 homolog evolution is illustrated in the Fig. 2. The history of the discovery of Oct4 homologs and changing views as to their evolution, which are reflected by changes in their nomenclature, are outlined in the following paragraph.

Three Oct4/*Pou5f1*-related *Xenopus* genes were isolated by virtue of sequence homology to the POU DNA-binding domain. The first gene, termed *Oct-60*, is primarily expressed as maternal transcript in oocytes and is present in early embryos until the gastrula stage of development. Transcripts from a second and third POU-domain gene, termed *Oct-25* and *Oct-91*, reach the highest levels during early and late gastrulation, respectively, and fade afterwards. Sequence comparisons with other members of the POU-domain family showed that Oct-25, Oct-60, and Oct-91 are most closely related to Oct4 and to each other, and they were assigned to the POU-V subclass of the POU-domain family [47]. A PouV class zebrafish gene, designated somewhat misleadingly *pou2* was cloned by similarity to Oct4, and is maternally expressed with its transcripts present up to the gastrula stage [134] and later at the mid-hindbrain boundary [43]. More *Pou5f1*-related genes expressed in oocytes and early developmental stages were cloned from Axolotl (*AxOct-4*: [5]), medaka (*Ol-pou5f1*: [139]), and chicken (*cPouV*: [78]). Orthology relationships between mammalian *Pou5f1* and non-mammalian maternal PouV class genes were at that time unclear, which was reflected in a variety of gene names given to PouV class genes, which added confusion to the field. Only a single *pou2*-like gene was isolated from medaka and zebrafish genomes, both of which are expressed early in development, which suggested functional similarities of its gene product with mammalian Oct4. Therefore, some investigators considered fish *pou2* and mouse *Pou5f1* as orthologs (i.e. [15,110,115]), and some described them more cautiously as homologs [6,89]. The finding that genomes of basal mammals, exemplified by platypus and opossum, contains two PouV class genes, *Pou5f1*-like and *pou2*-like, led Niwa et al. [99]) to hypothesize that mammalian *Pou5f1* arose by duplication of an ancestral *pou2*-like gene in early mammalian