The POU-er of gene nomenclature

Stephen R. Frankenberg1,*, Dale Frank2, Richard Harland3, Andrew D. Johnson4, Jennifer Nichols5, Hitoshi Niwa6, Hans R. Schöler7, Elly Tanaka8, Chris Wylie9 and Joshua M. Brickman10,*

ABSTRACT

The pluripotency factor POU5F1 (OCT4) is well known as a key regulator of stem cell fate. Homologues of POU5F1 exist throughout vertebrates, but the evolutionary and functional relationships between the various family members have been unclear. The level to which function has been conserved within this family provides insight into the evolution of early embryonic potency. Here, we seek to clarify the relationship between POU5F1 homologues in the vertebrate lineage, both phylogenetically and functionally. We resolve the confusion over the identity of the zebrafish gene, which was originally named pou2, then changed to pou5f1 and again, more recently, to pou5f3. We argue that the use of correct nomenclature is crucial when discussing the degree to which the networks regulating early embryonic differentiation are conserved.

Class V POU (POUV) transcription factors are important regulators of potency, differentiation and early development in vertebrates. They constitute one of five classes of POU domain-containing proteins defined by similarity within both the homeodomain and the POU-specific domain (Rosenfeld, 1991). Mouse POU5F1 (also called OCT3 or OCT4) was the first class V member identified (Lenardo et al., 1989; Okamoto et al., 1990; Rosner et al., 1990; Schöler et al., 1989, 1990). It is a central regulator of embryonic stem cell (ESC) pluripotency (Nichols et al., 1998; Niwa et al., 2000; Yuan et al., 1995) and the most essential of the four factors originally identified as being able to induce reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006; Yu et al., 2007). Since the initial characterisation of POU5F1, homologues have been identified in many vertebrates, including frog, chicken, axolotl, teleost fishes and sturgeon.

As more POUV genes were identified, it became apparent that the POUV family has a complex evolutionary history. Some eutherian mammals possess a second, single-exon POUV gene, POU5F2 (previously called SPRM-1), which has a role in spermatogenesis (Andersen et al., 1993; Pearse et al., 1997) and presumably arose in an ancestral eutherian by retroviral insertion of a copy of POU5F1. Marsupial and monotreme genomes contain two POUV genes: POU5F1 and another homologue now called POU5F3. Both the marsupial POUV genes (POU5F1 and POU5F3) are expressed in early development in domains similar to those described for Pou5f1 in the mouse (Frankenberg et al., 2010, 2013). Notably, some vertebrate lineages have orthologues of both POU5F1 and POU5F3, whereas others (including eutherian mammals) have only POU5F1 and others only POU5F3 (Fig. 1). The basis for this pattern of evolution is unclear, but is possibly due to distinct roles for each parologue that became redundant in a taxon-specific manner.

In zebrafish, a POUV gene was identified and originally, albeit confusingly, named pou2 (Takeda et al., 1994). With the availability of the complete sequence, it became clear that pou2 encodes a class V POU protein (Burgess et al., 2002) and, based on an assumption of orthology with mammalian POU5F1 (Lunde et al., 2004), zebrafish pou2 was subsequently renamed pou5f1. However, the zebrafish gene is not a true orthologue of mammalian POU5F1 but is instead more closely related to mammalian POU5F3, as indicated by conservation of both synteny and sequence (Frankenberg et al., 2010; Frankenberg and Renfree, 2013; Niwa et al., 2008; Tapia et al., 2012). Moreover, a recent study (Frankenberg and Renfree, 2013) has now demonstrated unequivocally that the gene duplication event giving rise to POU5F1 and POU5F3 occurred before the divergence of extant cartilaginous and bony fishes, showing conclusively that the zebrafish gene is a true POU5F3 orthologue. On the weight of this evidence, the Zebrafish Nomenclature Committee re-named the zebrafish gene pou5f3, reflecting its proper place in the POUV family. We now support the application of this nomenclature to all vertebrate orthologues of POU5F3 (see Table 1).

The degree to which the function of the POUV proteins is conserved in evolution is variable. This has been tested by the capacity of different POUV proteins to rescue the loss of endogenous POU5F1 activity in ESC self-renewal (Hammachi et al., 2012; Morrison and Brickman, 2006) and in the generation of iPSCs (together with three other mammalian factors) (Tapia et al., 2012). Based on these assays, orthologues of POU5F1 and POU5F3 show varying degrees of functional conservation in inducing pluripotency and supporting self-renewal. In particular, in Xenopus, which has three POU5F3 genes (pou5f3.1, pou5f3.2 and pou5f3.3) that presumably arose by tandem duplication, the expression pattern and activity of these genes have diversified such that only one of them – pou5f3.1 – is expressed in primordial germ cells and has ‘OCT4-like’ activity in both reprogramming and ESC self-renewal (Livigni et al., 2013; Venkatarama et al., 2010; Tapia et al., 2012). Other POU5F3 genes, including the two other Xenopus genes, have varying degrees of OCT4-like activity in such assays, but it is notable that zebrafish pou5f3 has little activity in either reprogramming or the support of Pou5f1 mutant ESCs (Laval et al., 2007; Morrison and Brickman, 2006; Niwa...
et al., 2008; Tapia et al., 2012). Zebrafish pou5f3 also participates in zygotic genome activation (Lee et al., 2013; Leichsenring et al., 2013) and, although this has not been explored in mammals, the observation that mouse embryos develop to the blastocyst stage in the absence of both maternal and zygotic Pou5f1 (Frum et al., 2013; Le Bin et al., 2014; Nichols et al., 1998; Wu et al., 2013) suggests that this function might not be conserved. Although there is a degree of redundancy in function between Pou5f1 and Pou5f3, as is particularly evident from species that have lost one of these genes, this is not always the case. Understanding the relationship between the evolution of this family and the degree to which the network regulating pluripotency is conserved requires an unambiguous representation of the phylogenetic relationships through nomenclature.

Despite clear evidence that the zebrafish PouV gene is a Pou5f3 orthologue, and the recent renaming of the gene to reflect this, the name pou5f1 is still persistently used in the literature (Bensch et al., 2013; Kotkamp et al., 2014; Lee et al., 2013; Leichsenring et al., 2013; Lippok et al., 2014; Rodriguez-Mari et al., 2013). Onichtchouk (2012) had argued that, because the duplication event giving rise to the two mammalian paralogues was not proven to pre-date the divergence of tetrapod and teleost lineages, pou5f1 is a valid name for the zebrafish gene despite it being more similar to mammalian Pou5f3 than to mammalian Pou5f1. However, a subsequent study (Frankenberg and Renfree, 2013) demonstrated unequivocally that the duplication event occurred before the divergence of extant cartilaginous and bony fishes. Clearly, the use of the correct name does not alter the underlying biology or the degree to which pluripotency is or is not conserved in teleosts. However, the use of common names for orthologous genes is entrenched in the language of biology as a way of conveying that these genes are also likely to have orthologous roles and functions. Disruption of this convention – as has occurred with the zebrafish pou5f3 gene – can be misleading and cause confusion among readers who might be unfamiliar with the genes being discussed. We therefore urge researchers in the field to adopt the naming agreed by the Zebrafish Nomenclature Committee and proposed here for other vertebrate homologues.

The use of divergent model organisms is a powerful approach for understanding basic developmental mechanisms that are fundamental within broad taxonomic groups, and the 20th century literature tended to emphasize similarities across evolution. To some extent, the focus has now shifted to understanding differences in the deployment of genes so as to understand the underlying basis for species differences. Many of us in our own research are interested in the evolution of PouV genes and are asking questions such as: What are the distinctions between Pou5f1 and Pou5f3 that explain why one or other gene has become extinct in so many vertebrate lineages? How do these distinctions relate to differences among species in how they regulate potency and early development or specify their germ cells? Addressing such questions requires a clear understanding of the evolutionary relationships between homologues and the appropriate use of nomenclature to discuss them, and we therefore hope that this article helps to resolve any confusion regarding gene naming conventions for the PouV family, as well as highlighting potential pitfalls in gene nomenclature more generally.

Acknowledgements
We thank Ira Blitz (UC Irvine) for critical discussions.

Competing interests
The authors declare no competing financial interests.

References


required for lineage priming in the developing inner cell mass of the mouse blastocyst. Development 141, 1001-1010.


