© 2019. Published by The Company of Biologists Ltd | Development (2019) 146, dev168633. doi:10.1242/dev.168633

STEM CELLS AND REGENERATION

RESEARCH ARTICLE

Co-option of the PRDM14-CBFA2T complex from motor neurons to pluripotent cells during vertebrate evolution

Masanori Kawaguchi[1,*], Kota Sugiyama[1,*], Kazumi Matsubara[1,‡], Che-Yi Lin[2,‡], Shigehiro Kuraku[3,‡], Shota Hashimoto[1], Yoshiaki Suwa[1], Luok Wen Yong[2], Koji Takino[1], Shota Higashida[1], Daisuke Kawamura[1], Jr-Kai Yu[2] and Yoshiyuki Seki[1,§]

ABSTRACT

Gene regulatory networks underlying cellular pluripotency are controlled by a core circuitry of transcription factors in mammals, including POU5F1. However, the evolutionary origin and transformation of pluripotency-related transcriptional networks have not been elucidated in deuterostomes. PR domain-containing protein 14 (PRDM14) is specifically expressed in pluripotent cells and germ cells, and is required for establishing embryonic stem cells (ESCs) and primordial germ cells in mice. Here, we compared the functions and expression patterns of PRDM14 orthologues within deuterostomes. Amphioxus PRDM14 and zebrafish PRDM14, but not sea urchin PRDM14, compensated for mouse PRDM14 function in maintaining mouse ESC pluripotency. Interestingly, sea urchin PRDM14 together with sea urchin CBFA2T, an essential partner of PRDM14 in mouse ESCs, complemented the self-renewal defect in mouse *Prdm14* KO ESCs. Contrary to the *Prdm14* expression pattern in mouse embryos, *Prdm14* was expressed in motor neurons of amphioxus embryos, as observed in zebrafish embryos. Thus, *Prdm14* expression in motor neurons was conserved in non-tetrapod deuterostomes and the co-option of the PRDM14-CBFA2T complex from motor neurons into pluripotent cells may have maintained the transcriptional network for pluripotency during vertebrate evolution.

This article has an associated 'The people behind the papers' interview.

KEY WORDS: Co-option, Pluripotent cells, Motor neuron, PRDM14, CBFA2T

INTRODUCTION

Germ cell specification in multicellular organisms is subdivided into two modes: preformation and epigenesis (Johnson et al., 2003; Extavour and Akam, 2003). In different species, germ cell specification can occur via either of two modes. During preformation, cytoplasmic germ cell determinants, known as the germ plasm, are asymmetrically localised in oocytes and predetermine the cellular competency for germ cell formation, e.g. in flies, worms, teleost fishes and frogs. In contrast, germ cells are segregated from pluripotent cells, which can differentiate into both germ cells and somatic cells, by receiving induction signals from surrounding tissues in epigenesis, e.g. in salamanders, mice and humans (Chatfield et al., 2014; Ohinata et al., 2009; Sasaki et al., 2015). The phylogenetic distribution of preformation and epigenesis in metazoans implies that epigenesis is an ancestral mode of germ cell formation and that preformation evolved repeatedly and independently (Johnson et al., 2003). Because the establishment of pluripotent cells in early embryos is essential for germ cell development during epigenesis, understanding the emergence and diversity of gene regulatory networks (GRNs) needed for pluripotency is important for determining the ancestral mechanisms of germ cell formation.

Pluripotent mouse stem cells, otherwise known as mouse embryonic stem cells (mESCs), were first established by culturing the inner cell mass of a blastocyst at embryonic day (E) 3.5 (Evans and Kaufman, 1981), and have served as a useful model for characterising the pluripotency of GRNs. The self-renewal of mESCs is regulated by a core circuitry of transcription factors: POU5F1 (also known as OCT3 or OCT4), NANOG, SOX2 and the KLF family of proteins (Niwa, 2014). *Sox2* and the *Klf* family of genes are essential for maintaining ESC pluripotency; however, these factors are also expressed in somatic lineages in mice (Avilion et al., 2003; Jiang et al., 2008; Kuo et al., 1997). In contrast, *Pou5f1* has been shown to act as a master regulator of pluripotency, and its expression is only associated with pluripotent cells and germline cells in mice (Yeom et al., 1996; Yamaguchi et al., 2005). *Pou5f1* and its paralogue *Pou5f3* might have emerged from a common ancestor of gnathostomes, while *Pou5f1* was eliminated from teleosts, frogs and birds (Frankenberg et al., 2014). In contrast, the genome of the axolotl encodes both *Pou5f1* and *Pou5f3*, and these genes are associated with embryonic pluripotent cells and germ cells (Tapia et al., 2012). Conditional knockout (KO) of *Pou5f1* leads to apoptosis of primordial germ cells (PGCs) in mice (Kehler et al., 2004), whereas knock down of *Pou5f3.3* expands the early number of PGCs in *Xenopus* embryos (Butler et al., 2018). These studies suggest that *Pou5f1* is required for germ cell formation in species exploiting epigenesis, whereas *Pou5f1* can be abrogated in species exploiting preformation (Johnson and Alberio, 2015).

Previously, we have shown that PR domain-containing 14 (PRDM14) is expressed in pluripotent cells and germ cells in mice, and is required for the establishment of primordial germ cells (Yamaji et al., 2008). *Prdm14*-deficient mESCs expand only in the presence of inhibitors of the extracellular signal-regulated kinase (ERK) and glycogen synthase kinase 3 β (GSK3β) (together known as '2i') plus leukaemia inhibitory factor (LIF), but not in serum plus LIF (Yamaji et al., 2013). ESCs maintained in 2i plus LIF are transcriptionally similar to peri-implantation epiblast cells at around E4.5 and are called naïve ground state pluripotent stem cells

[1]Department of Biomedical Chemistry, School of Science and Technology, Kwansei Gakuin University, Sanda, Hyogo 6691337, Japan. [2]Institute of Cellular and Organismic Biology, Academia Sinica, Taipei 11529, Taiwan. [3]Laboratory for Phyloinformatics, RIKEN Center for Biosystems Dynamics Research, 2-2-3 Minatojima-minami, Kobe 650-0047, Japan.
*These authors contributed equally to this work
‡These authors contributed equally to this work

§Author for correspondence (yseki@kwansei.ac.jp)

J.-K.Y., 0000-0001-8591-0529; Y.S., 0000-0001-5745-5549

Received 12 June 2018; Accepted 20 December 2018

Development (2019) 146, dev168633. doi:10.1242/dev.168633

(Nichols and Smith, 2009). The autocrine of fibroblast growth factor 4 (FGF4) activates ERK signalling and destabilises the transcriptional networks required for pluripotency (Kunath et al., 2007). PRDM14 represses *Fgfr1/2* expression to maintain ESC pluripotency under the serum-plus-LIF condition (Yamaji et al., 2013). *Prdm14* expression is closely linked to naïve pluripotent cells, the inner cell mass (ICM) and ESCs in mice, and its expression is shut off during the early transition period from naïve to primed pluripotency (Yamaji et al., 2008). Epiblast-like cells (EpiLCs), which are derived from ESCs, are transcriptionally similar to post-implantation epiblast cells at around E5.5, and are commonly called primed pluripotent cells (Hayashi et al., 2011). Prolonged *Prdm14* expression disturbs the spontaneous differentiation of mouse and human ESCs induced by embryoid body formation (Okashita et al., 2015; Tsuneyoshi et al., 2008). Furthermore, *Prdm14* overexpression in epiblast-like cells with primed pluripotency induces their conversion into ESCs with naïve pluripotency mediated by the re-establishment of transcriptional networks for naïve pluripotency (Okashita et al., 2016). Contrary to the PRDM14 functions in pluripotent cells in mammals, *Prdm14* is expressed in primary motor neurons but not in pluripotent cells, and its expression is required for primary motor neuron maturation in zebrafish embryos (Liu et al., 2012). Here, we have compared the functions and expression patterns of PRDM14 orthologues in order to understand their evolutionary conservation and the functional diversity among different deuterostome lineages.

RESULTS

Phylogenetic tree and synteny of *Prdm14* orthologues

The phylogenetic distribution of Prdm genes has been well documented in a previous study (Vervoort et al., 2016). To construct a phylogenetic tree of *Prdm14* orthologues among metazoa, we first searched for *Prdm14* orthologues in the genomes of diverse animals using the web tool aLeaves (Kuraku et al., 2013). Unexpectedly, we identified *Prdm14* orthologues in the genome of the sea anemone *Nematostella vectensis*, which belongs to the non-bilaterian phylum Cnidaria. In the superphylum Lophotrochozoa, we identified *Prdm14* orthologues in the genomes of *Capitella teleta*, *Helobdella robusta*, *Lottia gigantea* and *Schistosoma mansoni*, but not in the planarian *Schmidtea mediterranea*. In contrast, we did not identify *Prdm14* orthologues in *Caenorhabditis elegans* and *Drosophila melanogaster* that belonged to the superphylum Ecdysozoa. Recent phylogenetic analysis has suggested that the Priapulid worm (Priapulida phylum) is the earliest branching ecdysozoan that, together with Kinorhyncha and Loricifera phyla, forms a sister group to nematodes (*C. elegans*) and arthropods (*Drosophila*) (Martín-Durán and Hejnol, 2015). Interestingly, the *Priapulus caudatus* genome was found to encode a putative *Prdm14* orthologue, implying that *Prdm14* might have been lost during evolution within the ecdysozoans. Among the deuterostomes, sea urchins in Echinodermata and acorn worms in Hemichordata have *Prdm14* orthologues in their genomes, but *Prdm14* orthologues were not identified in currently available *Ciona intestinalis* genome assemblies. Next, we generated a phylogenetic tree of *Prdm14* orthologues (Fig. 1A and Fig. S1) and analysed the synteny of genomic regions containing the identified *Prdm14* orthologues (Fig. 1B). The genomic localisation of *Prdm14* between *Ncoa2* and *Slco5a1* is conserved among vertebrates, except for zebrafish – a teleost fish. Interestingly, in the genomes of medaka and spotted gar (which belong to a different subgroup of Teleostei), *Prdm14* is located between *Ncoa2* and *Slco5a1*, as described above for non-teleost vertebrates, suggesting that the gene order in the genomic region downstream of *Prdm14* might have been rearranged recently in the otocephalan lineage leading to zebrafish. In the amphioxus genome, *Prdm14* and *Slco5a1* are localised in different scaffolds of the genome.

Establishment of PRDM14 orthologue-expressing mouse *Prdm14*-knockout (KO) ESCs

Mouse *Prdm14* (m*Prdm14*) is indispensable for the self-renewal of ESCs cultured with serum plus LIF (Yamaji et al., 2013). Two inhibitors of MAPK/ERK kinase and GSK3β (referred to as '2i') in combination with LIF overcome the loss of self-renewal activity in m*Prdm14* KO ESCs. These findings indicate that mPRDM14 is required for the suppression of ESC differentiation in the presence of serum plus LIF. To assess the capacity of PRDM14 orthologues to sustain the self-renewal of ESCs cultured with serum plus LIF, we established four cell lines: m*Prdm14* KO ESCs harbouring an empty vector, sea urchin PRDM14 (sPRDM14), zebrafish PRDM14 (zPRDM14) or mPRDM14 under the 2i plus LIF condition (Fig. 2A,B). We first compared the expression levels of the pluripotency-associated genes *Dnmt3a*, *Dnmt3b* and *Dnmt3l*, and the differentiation-associated genes using quantitative RT-PCR (qRT-PCR) analysis in wild-type ESCs and in m*Prdm14* KO ESCs expressing s*Prdm14*, *Prdm14* or m*Prdm14*, or in m*Prdm14* KO ESCs with the empty vector (Fig. 2C). Among the pluripotency-associated genes, only *Tcl1*, which has been identified as a direct target of PRDM14 (Yamaji et al., 2013), was consistently downregulated in m*Prdm14* KO ESCs expressing the empty vector or sPRDM14 compared with wild-type ESCs. *Tcl1* downregulation in m*Prdm14* KO ESCs was reversed by the exogenous expression of zPRDM14 or mPRDM14. It has been shown that *Dnmt3a*, *Dnmt3b* and *Dnmt3l* repression by PRDM14 is essential for sustaining global hypomethylation in mouse ESCs (Yamaji et al., 2013). In m*Prdm14* KO ESCs expressing the empty vector or sPRDM14, *Dnmt3a*, *Dnmt3b* and *Dnmt3l* expression levels were consistently upregulated versus those in wild-type ESCs, whereas the exogenous expression of zPRDM14 or mPRDM14 completely reversed *Dnmt3a*, *Dnmt3b* and *Dnmt3l* repression in m*Prdm14* KO ESCs. Comparing the transcriptional activities of PRDM14 orthologues suggested that zPRDM14, but not sPRDM14, has the similar transcriptional activity to mPRDM14. Aligning the primary amino acid sequences of sPRDM14, zPRDM14 and mPRDM14 demonstrated that the PR domain and zinc-finger domain are relatively conserved, but the N-terminal region of PRDM14 shows low conservation among these orthologues (Fig. S1B). Because the zinc-finger domain of PRDM14 binds specific DNA sequences (Chia et al., 2010), we analysed the binding capacity of sPRDM14, zPRDM14 and mPRDM14 at target genes in mESCs using chromatin immunoprecipitation (ChIP)-qPCR. Although we did not detect the transcriptional activity of sPRDM14, ChIP-qPCR analysis clearly demonstrated that sPRDM14 bound to target genes similar to zPRDM14 and mPRDM14 in mESCs, except for *Fgfr1* (Fig. 2D). The reduction of PRDM14 binding at *Fgfr1* might be one of the causes for the defect of ESC self-renewal by sPRDM14 in serum plus LIF condition. On the other hand, these findings suggest the possibility that sPRDM14 cannot recruit complexes involving transcriptional regulation at target genes in mESCs.

zPRDM14, but not sPRDM14, rescued the self-renewal of mouse *Prdm14*-deficient ESCs

We next explored the growth rate and self-renewal capacity of m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14 in growth medium containing LIF and either 2i or serum. Growth rates were comparable among these

Development (2019) 146, dev168633. doi:10.1242/dev.168633

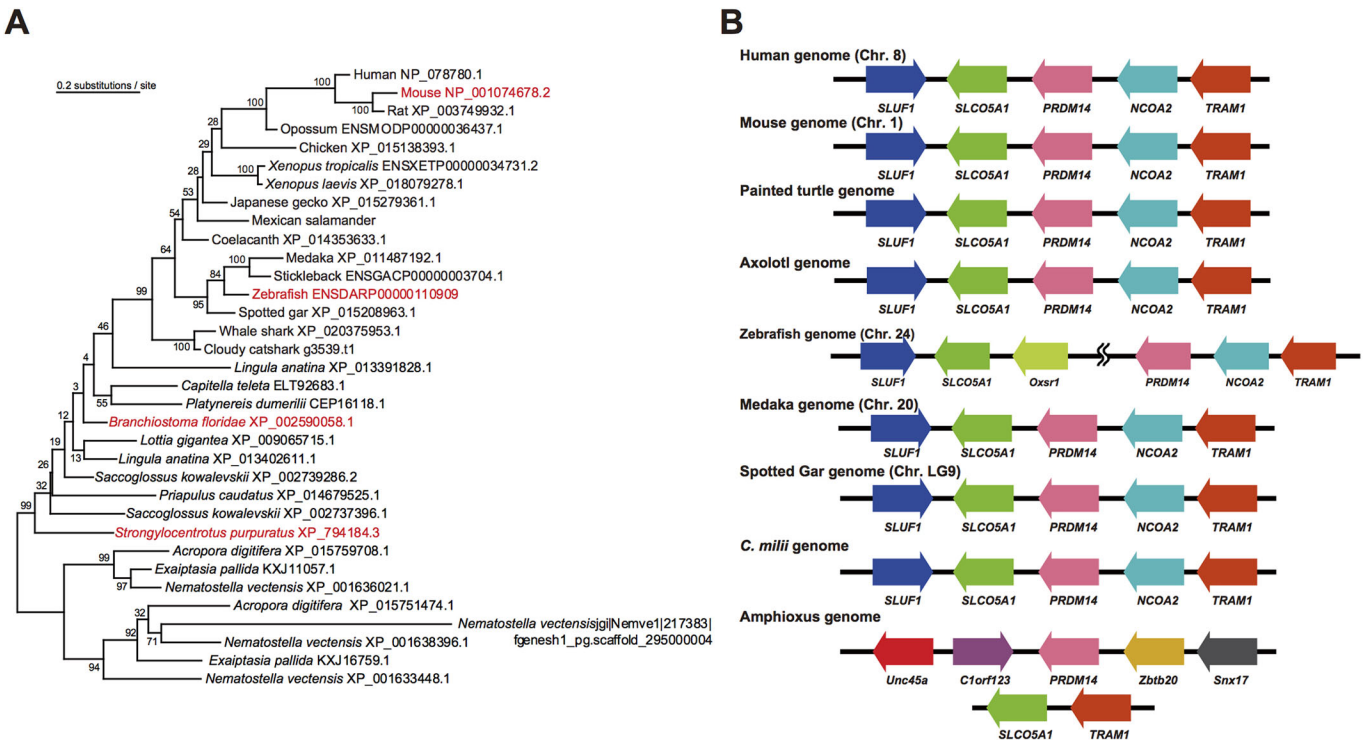

Fig. 1. Molecular phylogeny and synteny conservation for the metazoan *Prdm14* gene. (A) A maximum-likelihood tree was constructed using amino acid sequences of 34 metazoan *Prdm14* orthologues. Orthologues in red were used in this study. The grouping of these sequences as *Prdm14* orthologues was confirmed with our preliminary phylogenetic tree inference, including *Prdm4*, *Prdm6* and *Prdm12* (Fig. S1A). (B) Synteny conservation of genomic regions containing *Prdm14* genes among deuterostomes.

ESCs in medium containing 2i plus LIF (Fig. 3A). However, when the cells were grown in medium containing serum plus LIF, although m*Prdm14* KO ESCs expressing zPRDM14 or mPRDM14 continued to proliferate beyond 8 days, m*Prdm14* KO ESCs expressing the empty vector or sPRDM14 did not (Fig. 3A). We compared the expression of pluripotency markers in ESCs grown in medium containing serum plus LIF or 2i plus LIF. Pluripotency markers were maintained by exogenous zPRDM14 or mPRDM14 expression, but not by sPRDM14 or the empty vector (Fig. 3B). To determine the self-renewal activity of m*Prdm14* KO ESCs expressing each PRDM14 orthologue, we performed alkaline phosphatase (AP) staining of these cells cultured with serum plus LIF for 9 days (4 passages) or 14 days (7 passages) (Fig. 3C,D). Although we observed a small number of m*Prdm14* KO ESCs expressing sPRDM14 with AP activity in the presence of serum plus LIF up through 9 days, the cells did not expand after four passages. In contrast, we could expand m*Prdm14* KO ESCs expressing zPRDM14 or mPRDM14 with AP activity for at least 2 weeks (7 passages). These cells retained pluripotency markers, which were expressed at slightly higher levels than in wild-type cells (Fig. 3E).

zPRDM14 and mPRDM14, but not sPRDM14, induce DNA demethylation at pluripotency markers and germline-specific genes

Our group and another group have previously shown that mPRDM14 is involved in DNA hypomethylation through the repression of *Dnmt3a*, *Dnmt3b* and *Dnmt3l*, and the recruitment of ten-eleven translocation (TET) proteins, which oxidise 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) in ESCs (Yamaji et al., 2013; Okashita et al., 2014). First, we compared the protein expression levels of DNMT3A, DNMT3B and DNMT3L in wild-type ESCs and m*Prdm14* KO ESCs harbouring the empty vector or expressing sPRDM14, zPRDM14 or mPRDM14 (Fig. 4A). The expression levels of DNMT3A, DNMT3B and DNMT3L were consistently reduced in m*Prdm14* KO ESCs expressing zPRDM14 or mPRDM14 (but not sPRDM14) compared with those in m*Prdm14* KO ESCs expressing empty vector, in agreement with the observed mRNA expression levels (Fig. 2C). Next, to assess the transcriptional repression activity of PRDM14 orthologues at the *Dnmt3b* promoter, we performed luciferase assays using an 8.6 kb, upstream region of *Dnmt3b*. Consistent with the western blot results described above, zPRDM14 and mPRDM14, but not sPRDM14, repressed *Dnmt3b* promoter activity (Fig. 4B). Because mPRDM14 can activate the pluripotency marker *Tcl1* and the germline-specific genes *Dazl*, *Sycp3* and *Asz1* by inducing DNA demethylation (Okashita et al., 2014), we analysed the expression and DNA methylation of these genes in wild-type ESCs and in m*Prdm14* KO ESCs expressing the empty vector or the sPRDM14-, zPRDM14-, or mPRDM14-expressing vector (Fig. 4C-E). Expression of these genes was consistently reduced in m*Prdm14* KO ESCs harbouring the empty vector compared with wild-type ESCs, with elevated DNA methylation around the transcription start sites of these genes. Consistent with the compensation assay data for ESC self-renewal and for DNMT3A, DNMT3B and DNMT3L repression (shown above), zPRDM14 and mPRDM14, but not sPRDM14, drove upregulation and demethylation of these genes.

sPRDM14 can interact with mouse TET, but not with mouse CBFA2T2

mPRDM14 can interact with TET1/2 and CBFA2T2 to induce DNA demethylation and transcriptional regulation, respectively

Development (2019) 146, dev168633. doi:10.1242/dev.168633

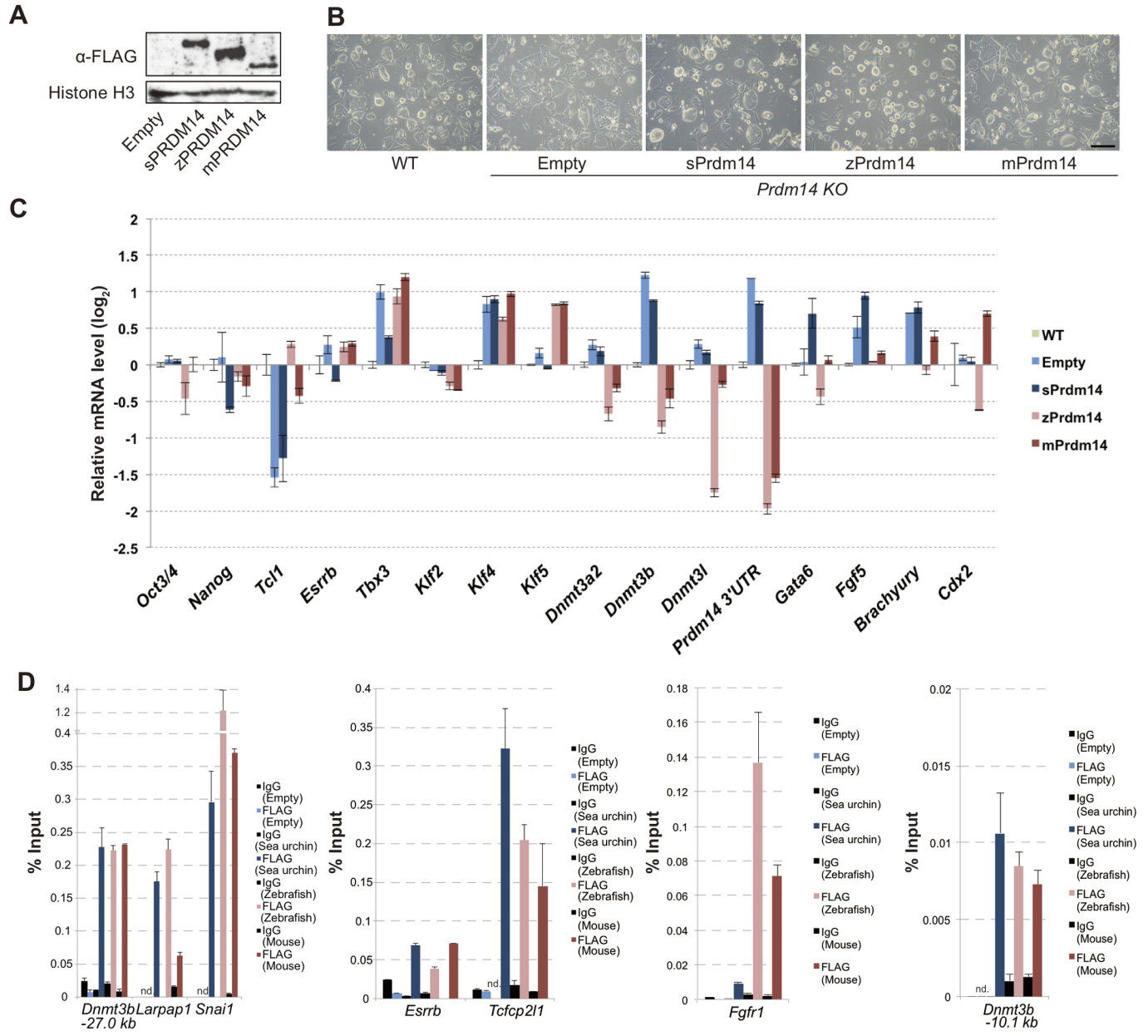

Fig. 2. zPRDM14, but not sPRDM14, compensate for transcriptional regulation in m*Prdm14* KO ESCs. (A) Western blot analysis of PRDM14 orthologues. (B) Morphology of wild-type ESCs and m*Prdm14* KO ESCs expressing the empty vector, s*Prdm14*, z*Prdm14* or m*Prdm14*. Scale bar: 50 µm. (C) Relative expression levels of the indicated genes in wild-type ESCs and m*Prdm14* KO ESCs expressing the empty vector, s*Prdm14*, z*Prdm14* or m*Prdm14*. Each Ct value was subtracted from the Ct values for wild-type ESCs and is shown as the $\log_2$ with the s.e.m. (two technical replicates). (D) ChIP-qPCR analysis for the enrichment of the indicated PRDM14 orthologues in m*Prdm14* KO ESCs. The enrichment of each orthologue is indicated by the percent input with the SEM (two technical replicates). nd, not detected.

(Okashita et al., 2014; Nady et al., 2015). Therefore, we analysed the interaction of sPRDM14, zPRDM14 and mPRDM14 with mouse TET2 (mTET2) and mouse CBFA2T2 (mCBFA2T2) by co-immunoprecipitation (co-IP) analysis. zPRDM14 and mPRDM14 interacted with endogenous mouse TET2 and CBFA2T2, whereas sPRDM14 interacted only with TET2 (Fig. 5A). We have previously shown that PRDM14 actively demethylates DNA at pluripotency-marker genes (including *Esrrb* and *Tfcp2l1*) through the TET-BER pathway to enhance POU5F1 recruitment during the transition from primed cells to naive pluripotency (Okashita et al., 2016). To compare the capacities of sPRDM14, zPRDM14 and mPRDM14 to recruit TET proteins to *Esrrb* and *Tfcp2l1*, we performed ChIP with antibodies against FLAG (FLAG-PRDM14), TET1 and TET2 at *Esrrb* and *Tfcp2l1* in m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14. Consistent with the co-IP experiments, we observed that sPRDM14, zPRDM14 and mPRDM14 enhanced TET1 and TET2 binding at *Esrrb* and *Tfcp2l1* (Fig. 5B). PRDM14 can repress the transcription of target genes, including *Fgfr1* and *Dnmt3*, by recruiting polycomb repressive complex 2 (PRC2) (Yamaji et al., 2013). Therefore, we investigated the recruitment of SUZ12, which is an essential component of PRC2, and the enrichment of H3K27me3 at *Fgfr1*

Development (2019) 146, dev168633. doi:10.1242/dev.168633

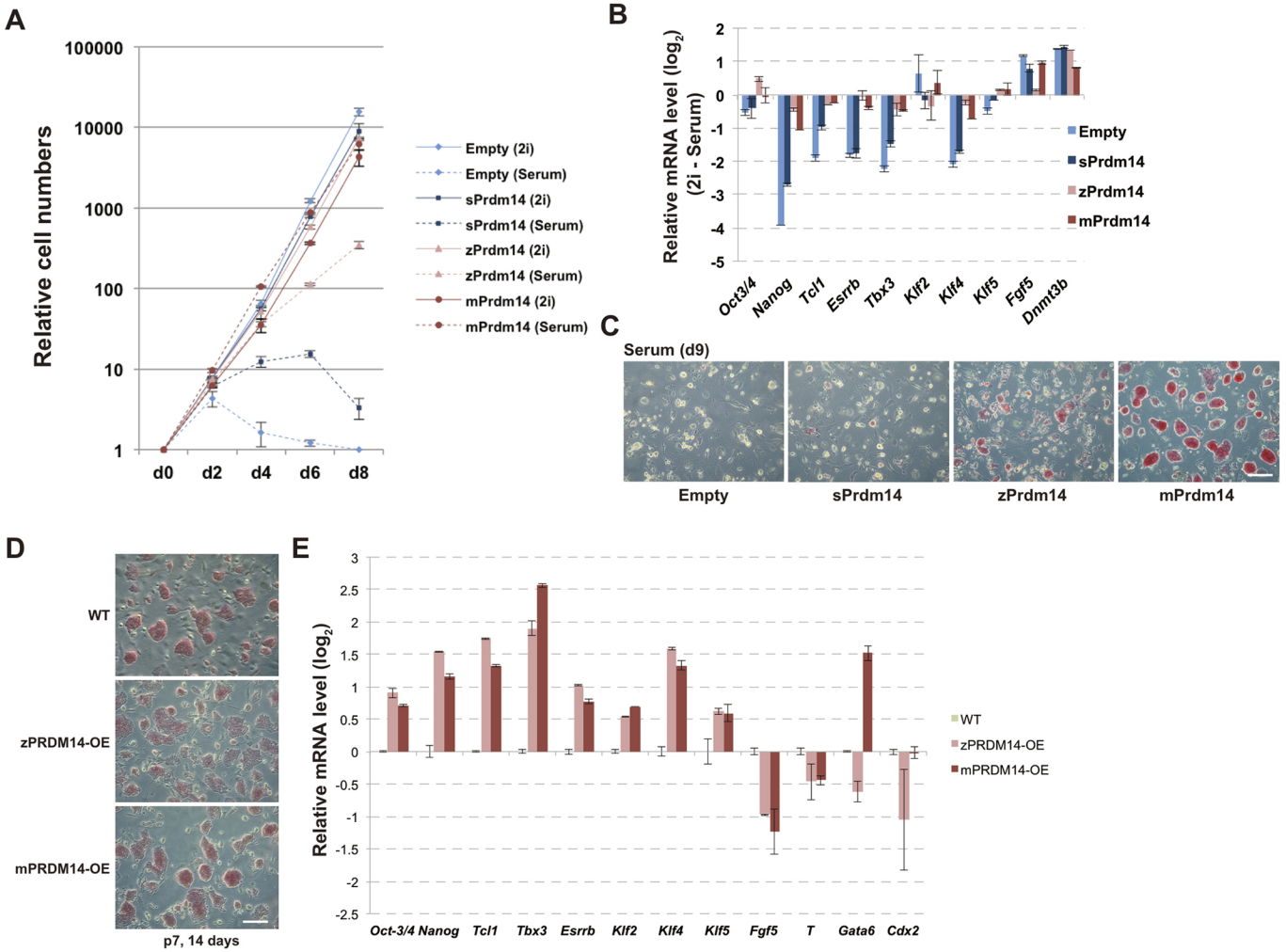

Fig. 3. zPRDM14, but not sPRDM14, maintains the self-renewal of m*Prdm14* KO ESCs under the serum plus LIF condition. (A) The growth rate of m*Prdm14* KO ESCs expressing the empty vector, s*Prdm14*, z*Prdm14* or m*Prdm14* under the serum plus or 2i plus LIF condition. The mean values are shown with the s.e.m. (three biological replicates). (B) Relative expression levels of the indicated genes. Each Ct value measured under the 2i plus LIF condition was subtracted from that measured under the serum plus LIF and is shown with the s.e.m. (two technical replicates). (C,D) AP activity of m*Prdm14* KO ESCs expressing the empty vector, s*Prdm14*, z*Prdm14* or m*Prdm14* cultured in medium containing serum plus LIF for 9 days (C) or 2 weeks at seven passages (D). Scale bars: 50 µm. (E) Relative expression levels of the indicated genes in wild-type ESCs and m*Prdm14* KO ESCs expressing z*Prdm14* or m*Prdm14* and cultured for 2 weeks. Each Ct value was subtracted from that in wild-type ESCs and is shown with the s.e.m. (two technical replicates).

and *Dnmt3b* in m*Prdm14* KO ESCs expressing each PRMD14 orthologue (Fig. 5C). SUZ12 recruitment and H3K27me3 at these genes were enriched in m*Prdm14* KO ESCs expressing mPRDM14 or zPRDM14 compared with m*Prdm14* KO ESCs-expressing the empty vector or sPRDM14. These findings suggest that PRDM14 acquired the capacity to bind CBFA2T2, which might be essential for the recruitment of the PRC2 complex, during the evolutionary route to Chordata after the splitting from the common ancestor of Chordata and Echinodermata.

The amphioxus PRDM14 orthologue can compensate for mPRDM14 functions in maintaining pluripotency and transcriptional regulation

As shown above, PRDM14 derived from the teleost zebrafish completely compensated for mPRDM14 functions in DNA demethylation and pluripotency maintenance. Next, we examined the functional conservation of the *Prdm14* gene isolated from the Cephalochordate amphioxus (amPRDM14) (Fig. 6A). Surprisingly, amPRDM14-overexpressing m*Prdm14* KO ESCs retained AP activity for 9 days when grown in medium containing serum plus LIF (Fig. 6B). To examine the conservation of its function in transcriptional regulation, we compared global gene expression changes between control (empty), sPRDM14-, amPRDM14- and mPRDM14-overexpressing m*Prdm14* KO ESCs by microarray analysis. We plotted the relative expression levels of genes that were upregulated (n=273) and downregulated (n=288) by mPRDM14 in sPRDM14- and amPRDM14-overexpressing m*Prdm14* KO ESCs (Fig. 6C,D). Genes upregulated by mPRDM14 tended to be upregulated by sPRDM14 and amPRDM14, which is consistent with the data showing sPRDM14 and amPRDM14 can interact with TET2 (Fig. 5A and data not shown). Interestingly, a scatter plot comparing genes downregulated by mPRDM14 showed a weak negative correlation in sPRDM14-overexpressing m*Prdm14* KO ESCs (Fig. 6C). sPRDM14 could bind to target genes, but not to mCBFA2T2 (Figs 2E and 5A), suggesting that sPRDM14 overexpression exerted a dominant-negative effect on transcriptional repression by endogenous CBFA2T2 in m*Prdm14* KO ESCs. In

Development (2019) 146, dev168633. doi:10.1242/dev.168633

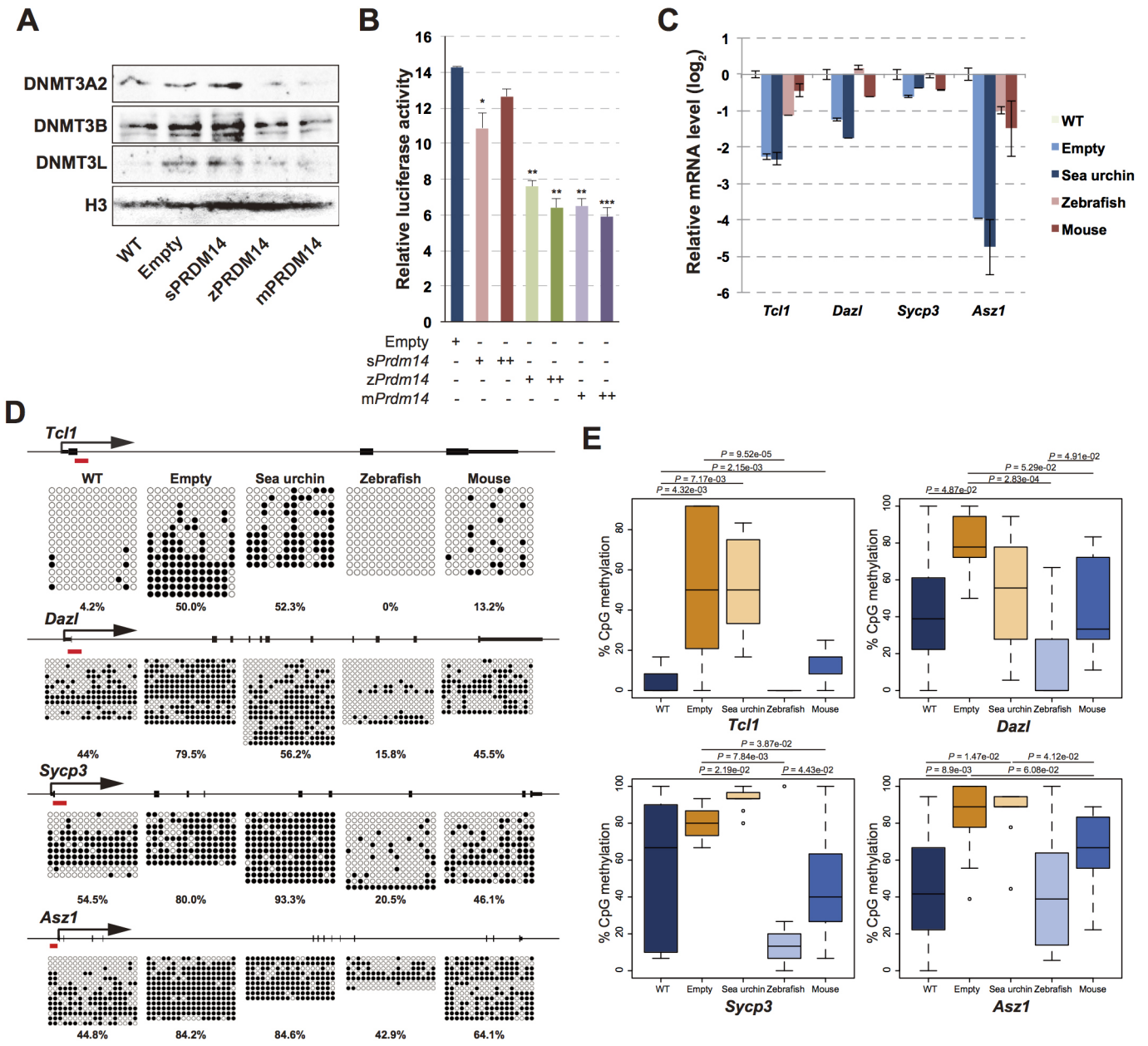

Fig. 4. zPRDM14, but not sPRDM14, induces hypomethylation at pluripotency-associated genes and germline-specific genes. (A) Western blotting of DNMT3A, DNMT3B and DNMT3L in wild-type ESCs and in m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14. (B) Luciferase activity driven by *Dnmt3b* against each PRDM14 orthologue. Relative luciferase activities of the pGL3-basic vector are shown with the s.e.m. (biological triplicates). *P* values were calculated using Student's *t*-test. **P*<0.05, ***P*<0.01, ****P*<0.001. (C) qRT-PCR analysis of pluripotency-associated genes and germline specific genes in wild-type ESCs and m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14. Each Ct value was subtracted from that in wild-type ESCs and is shown with the s.e.m. (two technical replicates). (D) DNA-methylation status of pluripotency-associated genes and germline-specific genes in wild-type ESCs and m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14 were measured by bisulphite sequencing. The red bars indicate the primer positions. The white circles indicate unmethylated cytosine or 5-formylcytosine or 5-carboxylcytosine, and black circles indicate 5-methylcytosine or 5-hydroxymethylcytosine. (E) Boxplot based on the bisulphite sequencing data. The central bars indicate the medians and lower and upper limits of the boxes, marking the 25th and 75th percentiles. *P* values were calculated using the Mann–Whitney *U*-test.

contrast, supporting the self-renewal assay results of amPRDM14-expressing ESCs, genes downregulated by mPRDM14 tended to also be downregulated by amPRDM14 (Fig. 6D). qRT-PCR analysis confirmed the consistent upregulation or downregulation of mPRDM14 target genes by axPRDM14 (Fig. 6E) and co-IP analysis clearly showed that amPRDM14 interacted with CBFA2T2 (Fig. 6F). These findings suggest that the acquisition of transcriptional regulation activity of PRDM14 orthologues tightly correlates with the CBFA2T2-binding capacity.

Pre-PR domain is responsible for the incompatibility of sPRDM14 with mCBFA2T2

A previous report has clearly revealed that mPRDM14 interacts with mCBFA2T2 through the pre-PR domain and PR domain (Nady et al., 2015). To identify the responsible amino acids for the incompatibility of sPRDM14 with mCBFA2T2, we reconstructed a model structure of sea urchin and mouse PRDM14 orthologues that form a complex with mouse CBFA2T2 using SWISS-MODEL, based on the previous crystal structure data for mPRDM14 (pre-PR

Development (2019) 146, dev168633. doi:10.1242/dev.168633

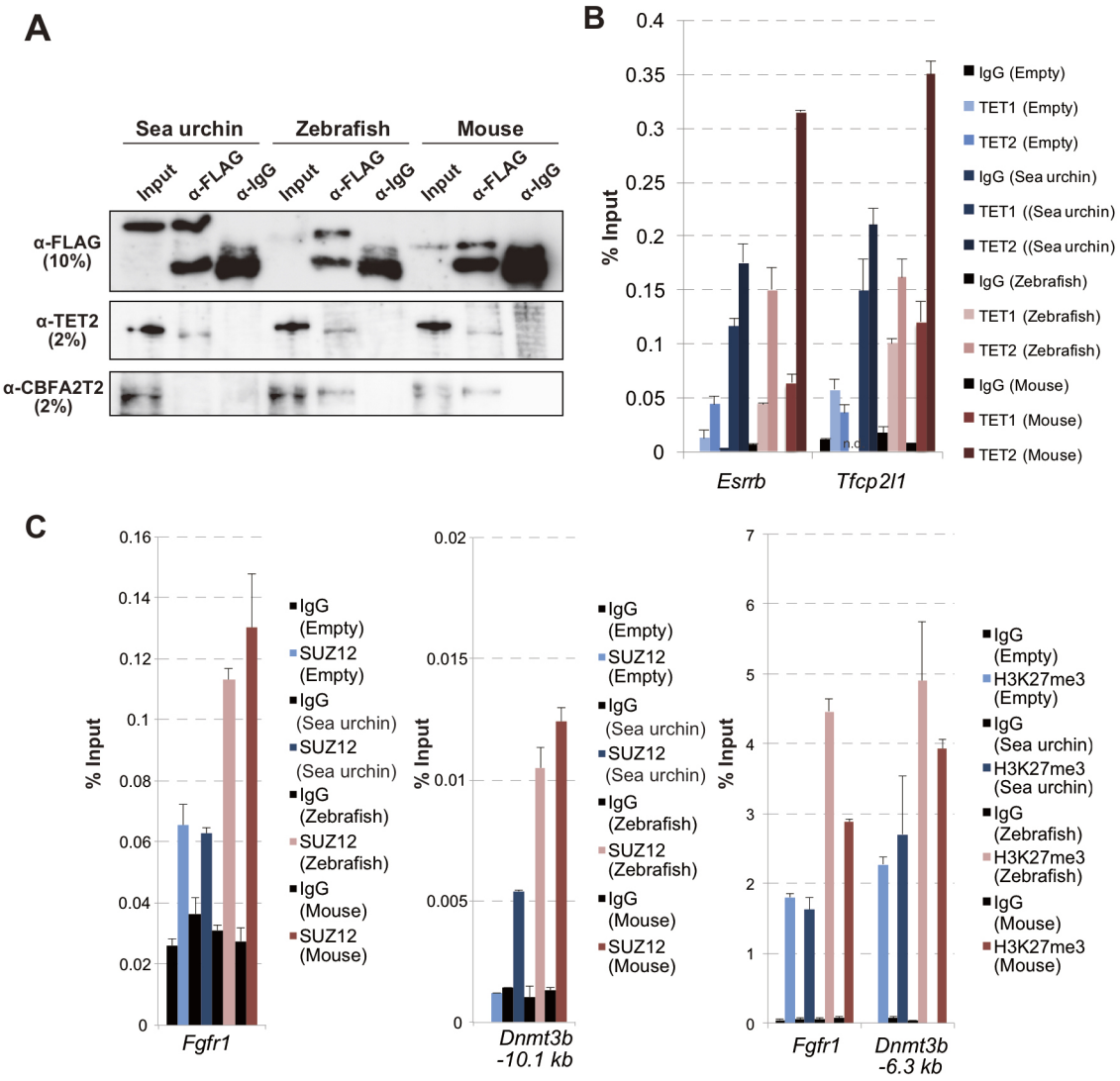

Fig. 5. zPRDM14, but not sPRDM14, interacts with mCBFA2T2. (A) Co-IP analysis of each PRDM14 orthologue with TET2 and CBFA2T2. (B) ChIP-qPCR analysis of TET1 and TET2 binding at pluripotency-associated genes in wild-type ESCs or m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14. The enrichment of TET1 and TET2 are indicated by the percent input with the s.e.m. (two technical replicates). (C) ChIP-qPCR analysis of SUZ12 and H3K27me3 at *Fgfr1* and *Dnmt3b*. The enrichment of SUZ12 and H3K27me3 are shown by the percent input with the s.e.m. (two technical replicates).

and PR domain) and the mCBFA2T2 complex (Nady et al., 2015) (Fig. 7A,B). Interestingly, the density and strength of the negative charge on the interface of the PRDM14 pocket interacting with mCBFA2T2 were significantly weaker in sPRDM14 compared with that in mPRDM14. Previous mutation analysis of mPRDM14 and mCBFA2T2 have clearly shown that the basic amino acids, Arg105 and Lys109 (located at the N-terminal region of mCBFA2T2) bind to a mPRDM14 pocket formed by the pre-PR and PR domain through electrostatic interaction (Nady et al., 2015). These findings suggest that the weak negative charge on the surface of the sPRDM14 pocket might be responsible for the incompatibility of sPRDM14 with mCBFA2T2. To identify the contribution of the pre-PR domain and PR domain in the binding capacity of PRDM14 with mCBFA2T2 and the maintenance of ESC pluripotency, we performed compensation assays for mouse PRDM14 using two chimeric proteins consisting of the N-terminal region of sPRDM14 (including the pre-PR domain) and C-terminal region of mPRDM14 (including the PR domain), and vice versa (Fig. 7C,E). We generated a model structure of the chimeric protein consisting of the two PRDM14 orthologues (Fig. 7D). Interestingly, the negative charge on the surface of the interacting pocket with CBFA2T2 depended on the pre-PR domain rather than the PR domain (Fig. 7C). To examine whether the model structure of the chimeric protein supports the function in the maintenance of ESC pluripotency, we established chimeric PRDM14-overexpressing m*Prdm14* KO ESCs (Fig. 7D). Consistent with the model structure of the chimeric protein (Fig. 7C), both self-renewal activity and interaction capacity of PRDM14 with mCBFA2T2 strongly depended on the pre-PR domain (Fig. 7E,F). Furthermore, s/mPRDM14-overexpressing m*Prdm14* KO ESCs partially retained AP activity, suggesting that the PR domain is also involved in the self-renewal activity of mPRDM14.

Ectopic expression of sPRDM14 with sea urchin CBFA2T (sCBFA2T) compensates for mPRDM14 and mCBFA2T2 functions in mESCs

sPRDM14 could not interact with endogenous mCBFA2T2 and compensate for mPRDM14 functions in maintaining mESC

Development (2019) 146, dev168633. doi:10.1242/dev.168633

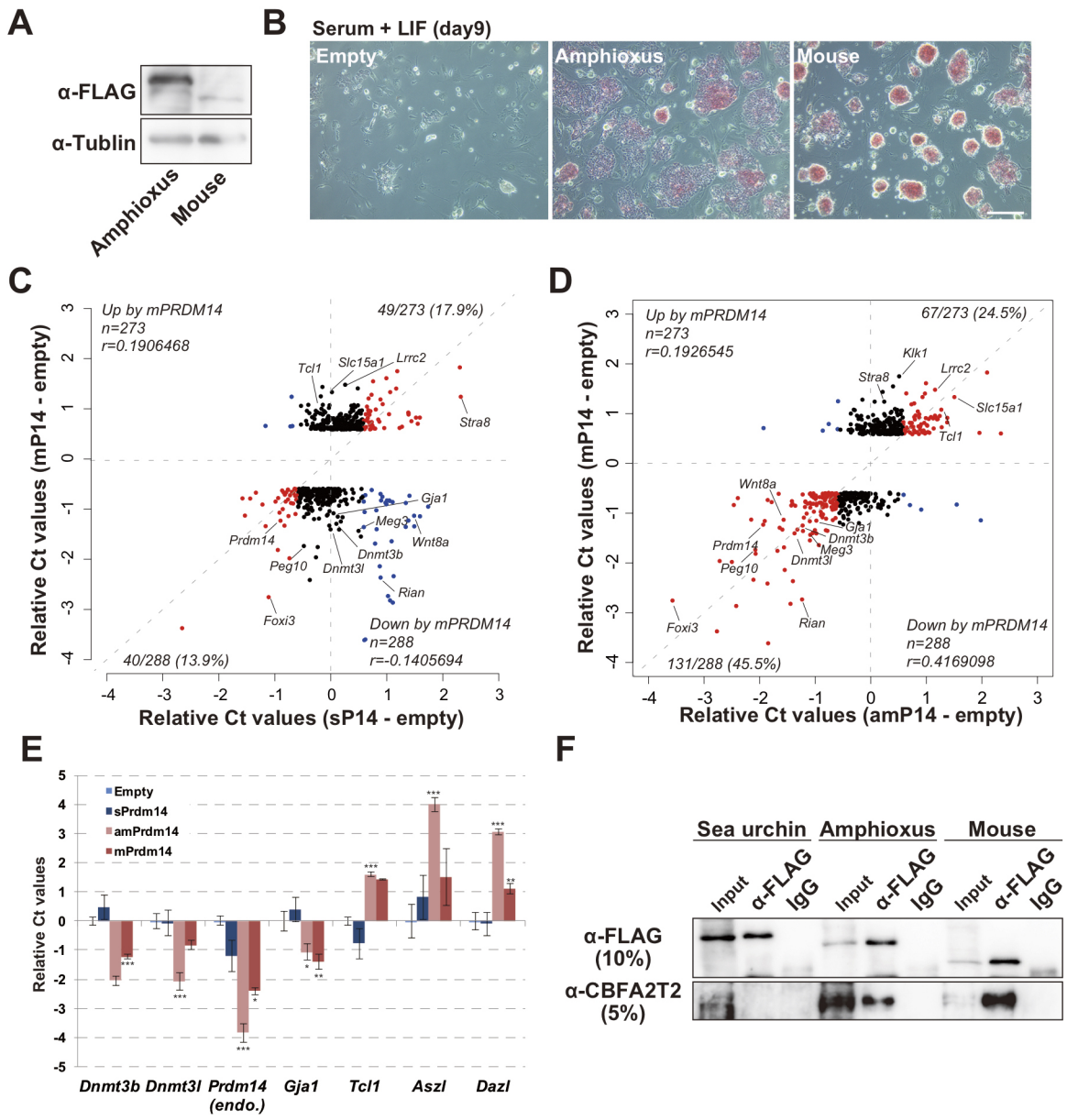

Fig. 6. amPRDM14 can compensate for lost mPRDM14 function in ESCs. (A) Western blot analysis of amPRDM14 and mPRDM14 expression. (B) Cellular morphology and AP activity in ESCs cultured in medium containing serum plus LIF for 9 days. Scale bar: 50 µm. (C,D) Scatter plot of microarray data showing the relative intensities of genes upregulated or downregulated by mouse PRDM14. The horizontal axis shows the relative intensities of microarray probes from sPRDM14-expressing (C) or amPRDM14-expressing (D) m*Prdm14* KO ESCs relative to m*Prdm14* KO ESCs expressing the empty vector. The red dots indicate genes that were co-upregulated or co-downregulated by expression of mPRDM14 and either sPRDM14 or amPRDM14. The blue dots indicate genes whose expression levels were negatively correlated after exposure to mPRDM14 and either sPRDM14 or amPRDM14. (E) Relative expression levels of the indicated genes in m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, zPRDM14 or mPRDM14. Each Ct value is subtracted from that found in m*Prdm14* KO ESCs expressing the empty vector. The error bars indicate the s.e.m. of biological triplicates. *P* values were calculated by Tukey's multiple-comparison test. **P*<0.05, ***P*<0.01, ****P*<0.001. (F) Co-IP analysis of sPRDM14, amPRDM14 and mPRDM14 with CBFA2T2.

pluripotency (Figs 3C and 5A). To investigate the possibility of the incompatibility between sPRDM14 and mCBFA2T2, we tried to establish m*Prdm14* KO ESCs co-expressing sPRDM14 and sCBFA2T. First, we obtained the sequence of *Cbfa2t2* orthologues and paralogues in the genome among metazoans. Cnidaria (sea anemone), protostomes and non-vertebrate deuterostomes possess only one *Cbfa2t* homologue in their genomes, whereas vertebrate genomes contain *Cbfa2t1*, *Cbfa2t2* and *Cbfa2t3*. Three Cbfa2t gene families might have emerged from a single *Cbfa2t* gene by two rounds of whole-genome duplication in the vertebrate lineage after the split of amphioxus from their common ancestor (Dehal and Boore, 2005) (Fig. 8A,B). The nervy homology region 1 (NHR1) domain of mCBFA2T2 is necessary and sufficient for the mPRDM14-mCBFA2T2 interaction (Nady et al., 2015). The alignment of mCBFA2T1/2/3 and sCBFA2T showed that the NHR1 domain is highly conserved, including amino acids involved in hydrogen bonding with PRRDM14 (Figs 7B and 8B). Next, we established *Prdm14* KO ESCs co-expressing sPRDM14 and sCBFA2T to examine the self-renewal activity of ESCs (Fig. 8C). Switching from the 2i plus LIF condition to the serum plus LIF condition promoted differentiation of m*Prdm14* KO ESCs harbouring the empty vector or expressing

Development (2019) 146, dev168633. doi:10.1242/dev.168633

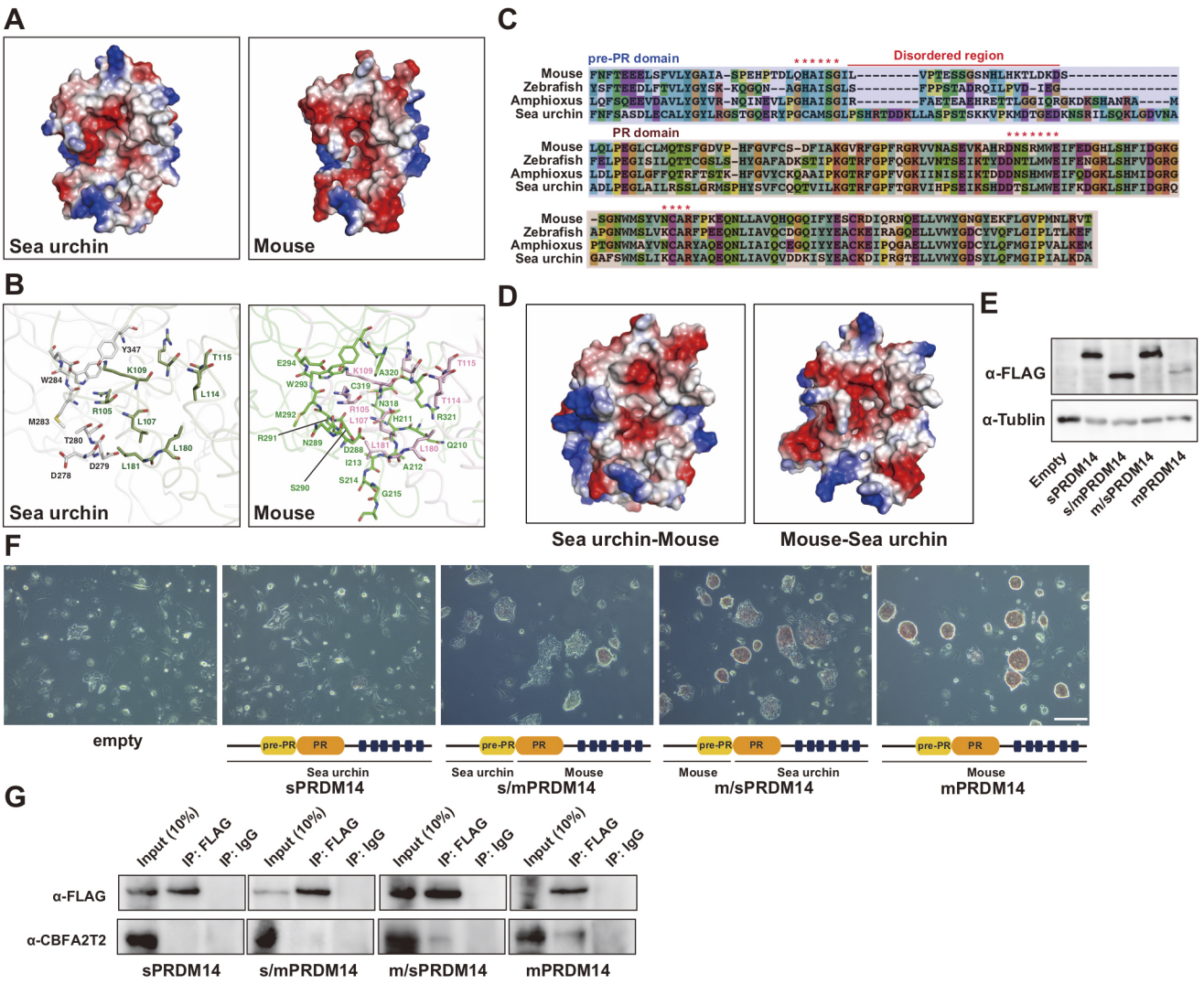

Fig. 7. The pre-PR domain is responsible for the incompatibility of sPRDM14 with mCBFA2T2. (A) Electrostatic charge distribution of the binding pocket of mPRDM14 and sPRDM14 with mCBFA2T2. (B) Residues involved in the negative charge of the binding surface of sPRDM14 and mPRDM14 with mCBFA2T2 are shown as sticks. (C) Alignment of the amino acid sequences of the pre-PR and PR domain derived from mPRDM14, zPRDM14, amPDM14 and sPRDM14. Red asterisks indicate the amino acids involved in hydrogen bonding with mCBFA2T2. (D) Electrostatic charge distribution of the binding pocket of chimeric protein constructed by sPRDM14 and mPRDM14. (E) Western blot analysis of sPRDM14, s/mPRDM14, m/sPRDM14 and mPRDM14. (F) Colonies stained for AP activity in *Prdm14* KO ESCs expressing the empty vector, sPRDM14, s/mPRDM14, m/sPRDM14 or mPRDM14, after being transferred to medium containing serum plus LIF for 9 days. Scale bar: 50 µm. (G) Co-IP of sPRDM14, s/mPRDM14, m/sPRDM14 or mPRDM14 with mCBFA2T2.

sPRDM14 (Fig. 8D). Interestingly, co-expressing sCBFA2T and sPRDM14 in m*Prdm14* KO ESCs maintained pluripotency under the serum plus LIF condition, similar to m*Prdm14* KO ESCs expressing mPRDM14. As mentioned above, the interaction of PRMD14 orthologues with CBFA2T2 closely correlated with the maintenance of ESCs. Co-IP analysis showed that sPRDM14 interacted with sCBFA2T in ESCs (Fig. 8E). sPRDM14 could not interact with endogenous mCBFA2T2 (Fig. 5A); however, the interaction of sPRDM14 with endogenous mCBFA2T2 was recovered by exogenous expression of sCBFA2T (Fig. 8F). CBFA2T has been shown to oligomerise through its NHR2 domain, which stabilises PRDM14 binding to chromatin in mESCs (Tu et al., 2016), suggesting that sPRDM14 interacts with mCBFA2T2 through sCBFA2T-mCBFA2T2 oligomerisation (Fig. 8G). Indeed, sCBFA2T interacted with both sPRDM14 and endogenous mCBFA2T2 (Fig. 8F). Finally, to investigate whether sCBFA2T can compensate for mCBFA2T2 function, we deleted exon 3 of m*Cbfa2t2* in sPRDM14 and sCBFA2T-expressing m*Prdm14* KO ESCs (Fig. S2). The sPRDM14-sCBFA2T complex compensated for the mPRDM14-mCBFA2T2 complex in the maintenance of ESC pluripotency (Fig. 8H). Supporting the compensation of sPRDM14 with sCBFA2T for mPRDM14 function, key target genes regulated by mPRDM14 were consistently regulated by the combination of sPRDM14 with sCBFA2T in m*Prdm14* KO ESCs (Fig. 8I). To compare the global changes in gene expression, we performed microarray analysis. Genes downregulated by mPRDM14 showed no correlation in sPRDM14-expressing m*Prdm14* KO ESCs (Figs 6C and 8J). Interestingly, a significant positive correlation of downregulated genes by mPRDM14 was observed in m*Prdm14* KO ESCs co-expressing sPRDM14 and sCBFA2T (Fig. 8K), which strongly supports the self-renewal activity of m*Prdm14* KO ESCs co-expressing sPRDM14 and sCBFA2T (Fig. 8D). These findings clearly indicate that sPRDM14 can cooperate with sCBFA2T, but not mCBFA2T2, in transcriptional regulation required for the maintenance of ESC pluripotency.

Development (2019) 146, dev168633. doi:10.1242/dev.168633

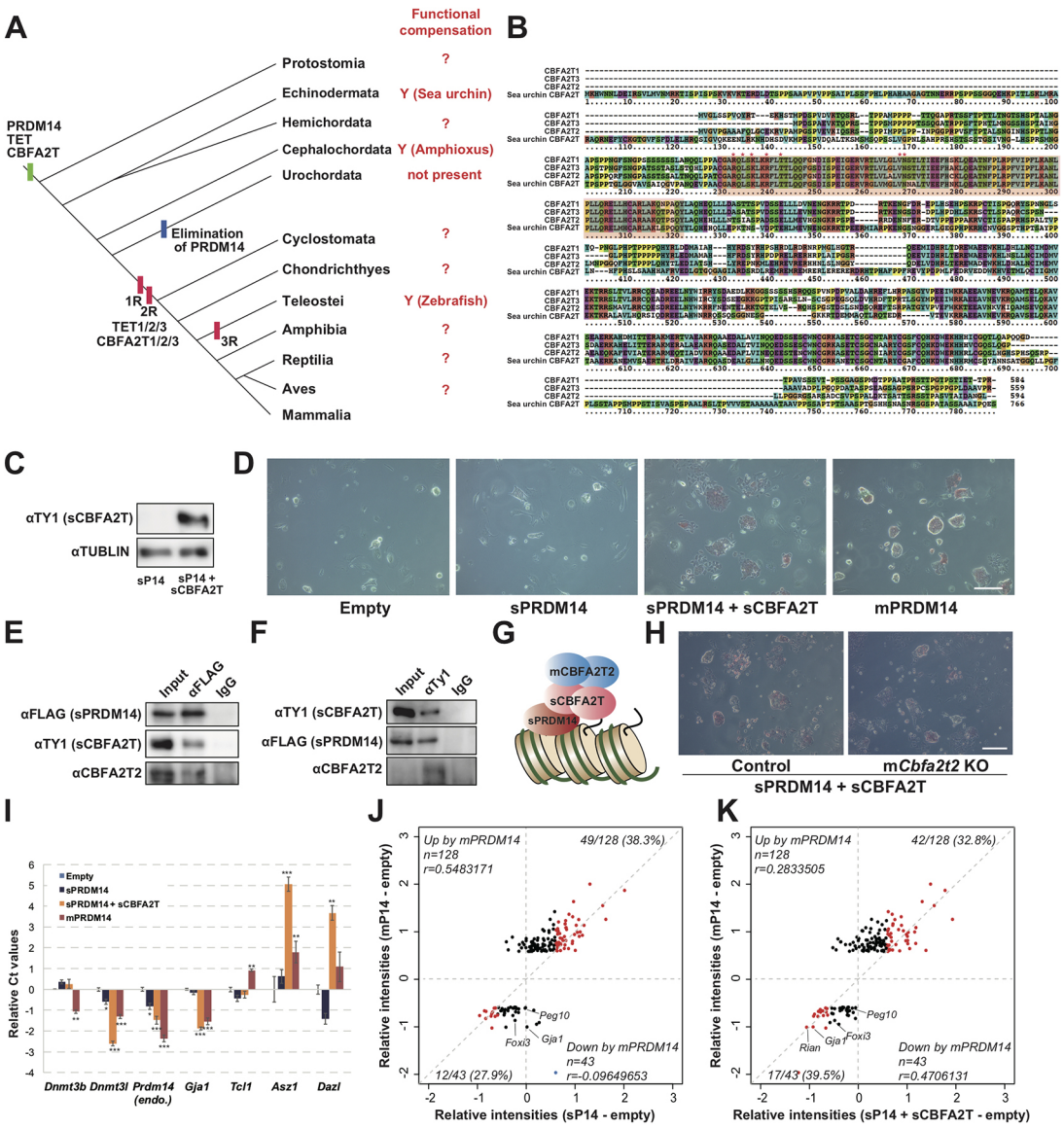

Fig. 8. Combined expression of sPRDM14 and sCBFA2T compensate for lost mPRMD14 function. (A) Distribution of Prdm14, Tet and Cbfa2t family genes in metazoans and functional compensation results (?, currently not examined; Y, functional compensation). (B) Alignment of the amino acid sequences of mCBFA2T1, mCBFA2T2, mCBFA2T3 and sCBFA2T. The red shading underneath the sequences indicates the NHR1 domain. Red asterisks show the amino acids involved in the hydrogen bonding between mCBFA2T2 and mPRDM14. (C) Western blot against sCBFA2T in *Prdm14* KO ESCs expressing sPRDM14. (D) Colonies stained for AP activity in *Prdm14* KO ESCs expressing the empty vector, sPRDM14, mPRDM14 or sPRDM14 with sCBFA2T, after being transferred to medium containing serum plus LIF for 9 days. Scale bar: 50 µm. (E) Co-IP of sPRDM14 with sCBFA2T and mCBFA2T2. (F) Co-IP of sCBFA2T with sPRDM14 and mCBFA2T2. (G) A model of the transcriptional regulation by sPRDM14 with sCBFA2T in m*Prdm14* KO ESCs. (H) Colonies stained for AP activity in m*Prdm14* KO ESCs expressing sPRDM14 and sCBFA2T2, and m*Prdm14* and m*Cbfa2t2* KO expressing sPRDM14 and sCBFA2T2, after being transferred to medium containing serum plus LIF for 9 days. Scale bar: 50 µm. (I) Relative expression levels of the indicated genes in m*Prdm14* KO ESCs expressing the empty vector, sPRDM14, sPRDM14 with sCBFA2T or mPRDM14. Each Ct value is subtracted from that found in m*Prdm14* KO ESCs expressing the empty vector. The error bars indicate the s.e.m. of biological triplicates. *P* values were calculated by Tukey's multiple-comparison test. **P*<0.05, ***P*<0.01, ****P*<0.001. (J) Scatter plot of microarray data showing the relative intensities of genes upregulated or downregulated by mouse PRDM14. The horizontal axis shows the relative intensities of microarray probes from sPRDM14-expressing (I) or sPRDM14 and sCBFA2T co-expressing (K) mouse *Prdm14* KO ESCs relative to *Prdm14* KO ESCs expressing the empty vector. The red dots indicate genes that were co-upregulated or co-downregulated by mPRDM14 and sPRDM14 or sPRDM14 with sCBFA2T. The blue dot in J indicates a gene whose expression levels were negatively correlated after exposure to mPRDM14 and sPRDM14.

Prdm14 is expressed in motor neurons of early amphioxus embryos

Prdm14 expression is associated with germline and pluripotent cells in mice or motor neurons in zebrafish (Yamaji et al., 2008; Liu et al., 2012). However, the expression pattern of *Prdm14* has not been examined thoroughly in other species, especially the invertebrate chordates. To explore the conservation of *Prdm14* expression in chordates, we studied the developmental expression of *Prdm14* in amphioxus. Using whole-mount *in situ* hybridisation, am*Prdm14* was not detectable at the cleavage and blastula stages, but started to become expressed in the ectoderm at the gastrula stage in amphioxus embryos (Fig. 9A-C). At the neurula stage, we detected am*Prdm14* expression in pairs of cells localising in the neural tube (Fig. 9D-F). This expression pattern was reminiscent of that for *Mnx/Hb9*, which is expressed in developing motor neurons (Fig. 9G) (Ferrier et al., 2001). To confirm the expression of

Development (2019) 146, dev168633. doi:10.1242/dev.168633

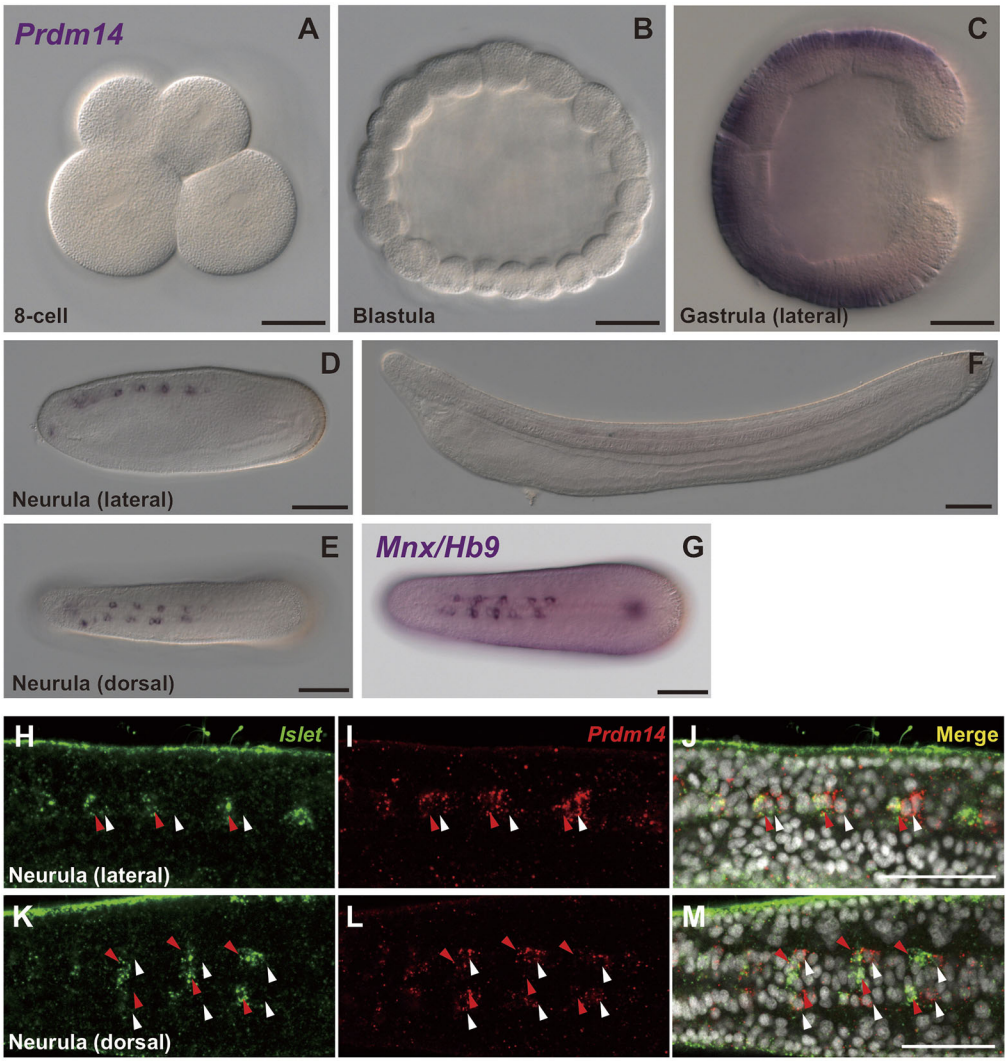

Fig. 9. am*Prdm14* was expressed in motor neurons in amphioxus embryos. (A-G) Distribution of *Prdm14* and *Mnxb* RNA in amphioxus embryos. (H-M) Fluorescent double labelling of *Islet* RNA (green) and *Prdm14* RNA (red) in amphioxus embryos at the neurula stage. White arrowheads indicate *Prdm14*-positive/*Islet*-negative cells. Red arrowheads indicate cells positive for both *Prdm14* and *Islet*. Scale bars: 50 µm.

Prdm14 in motor neurons, we performed double labelling of *Prdm14* and the primary motor neurons (PMNs) marker *Islet*, which is expressed in all PMNs (Jackman et al., 2000). Islet was co-expressed with *Prdm14* in the anterior cell of the two *Prdm14*-positive cells (Fig. 9H-M), indicating that *Prdm14* is expressed in the motor neurons of amphioxus embryos. In amphioxus, PGCs are marked by *Vasa* and *Nanos* expression, and are accumulated at the posterior end of the embryo at the neurula stage (Wu et al., 2011). However, we did not detect *Prdm14* expression at the posterior end of the embryo (Fig. 9D,E), suggesting that amphioxus *Prdm14* is produced in motor neurons, but not in PGCs. These findings provide evidence that the specific expression of *Prdm14* in motor neurons is conserved between amphioxus and zebrafish.

DISCUSSION

The phylogenetic distribution of *Prdm14*, *Tet* and *Cbfa2t2* orthologues indicates that these factors arose at least in the common ancestor of bilaterians and that *Prdm14* might have been lost during the evolution within the Ecdysozoa (Figs 1A and 8A). Comparative analysis of *Prdm14* expression among deuterostomes suggests that PRDM14-TET and PRDM14-CBFA2T2 complexes were co-opted from motor neuron into pluripotent cells around the time when tetrapods emerged. Our data show that all PRDM14 orthologues that we tested can interact with TET2 and we have previously shown that mouse PRDM14 cooperate with TET protein to maintain and acquire pluripotency in mouse pluripotent cells (Okashita et al., 2016). Furthermore, the contribution of TET proteins in early mouse embryo has been intensively investigated by several groups (Dai et al., 2016; Khoueiry et al., 2017). Contrary to the case of mouse, although *Prdm14* is also expressed in pluripotent cells of early chick embryo, the level of 5-hydroxymethylcytosine, which is produced from 5-methylcytosine by TET proteins, is quite low in pre-streak chick embryo (Almeida et al., 2012). Therefore, the co-option of PRDM14-TET complex into pluripotent cells might depend on the expression of TET proteins in pluripotent cells during the emergence of mammals. The data presented here show that amPRDM14 and zPRDM14 interacted with endogenous mCBFA2T2 and compensated for mPRDM14 function in maintaining ESC pluripotency. Furthermore, the combination of sPRDM14 and sCBFA2T expression compensated for mPRDM14 and mCBFA2T2 functions, suggesting that PRDM14 cooperated with CBFA2T in transcriptional regulation in the common ancestor of deuterostomes. Sea stars exploit an inductive mechanism for germ cell formation; hence, part of the embryonic gut may maintain multipotency to give rise to PGCs (Wessel et al., 2014). Interestingly, *Prdm14* is expressed in the embryonic gut, but not in PGCs in sea star embryos (Fresques et al., 2014). Our study provides clear evidence that sPRDM14 cooperates with sCBFA2T

Development (2019) 146, dev168633. doi:10.1242/dev.168633

to maintain mESC pluripotency, suggesting that the PRDM14-CBFA2T complex serves essential functions in maintaining embryonic gut multipotency in echinoderm embryos.

Intriguingly, *Prdm14* is expressed in the neuronal system in amphioxus, zebrafish and *Xenopus* embryos (Eguchi et al., 2015), and is required for motor neuron maturation in zebrafish embryos (Liu et al., 2012). The CBFA2T family is also expressed in the neural system in *Xenopus* and chicken embryos, and its expression is involved in neurogenesis, including motor neurons (Cao et al., 2002; Aaker et al., 2009). Our data provide evidence that the dependence of PRDM14 function on CBFA2T2 is conserved within deuterostomes, suggesting that PRDM14 might also cooperate with CBFA2T to control neurogenesis in non-mammalian deuterostomes. Contrary to the expression pattern of *Prdm14* in neuronal systems in non-mammalian vertebrates, *Prdm14* expression is lost in the nervous system, and is restricted in pluripotent cells and primordial germ cells in mammals (Yamaji et al., 2008; Tang et al., 2015). These findings suggest that *Prdm14* expression shifted from the neuronal system to pluripotent and germ cells en route from a vertebrate common ancestor to mammals during evolution (Fig. 8A). *Prdm14* is expressed in primary motor neurons (PMNs) but not in secondary motor neurons (SMNs) in zebrafish embryos (Liu et al., 2012). In contrast, PMNs have not been described in amniotes and it is thought that SMNs in zebrafish embryo closely resemble motor neurons in amniotes, including mouse (Eisen, 1994). These findings suggest that the loss of *Prdm14* expression in motor neurons might be linked to the loss of PMNs during the emergence of amniotes.

Previous ChIP-Seq analysis showed that POU5F1 is localised at a distal enhancer of the *Prdm14* locus both in mESCs and human ESCs (hESCs) (Okashita et al., 2016; Tsankov et al., 2015). Determining whether POU5F1 positively regulates *Prdm14* expression in mouse and human pluripotent stem cells is important for understanding how *Prdm14* integrated into the pluripotent network during evolution, which warrants further investigation. We propose two major possibilities to explain the change in the *Prdm14* expression pattern during evolution. The first possibility is that the expression pattern of upstream factors of *Prdm14* shifted from motor neurons to pluripotent cells and germ cells. *Pou5f1* belongs to the PouV class of genes, which is one of several classes in the POU-domain gene family. Based on currently available genomic information, the amphioxus genome contains PouIII genes, but not PouV genes, and PouV-class genes might have evolved from PouIII genes in the common ancestor of gnathostomes (Frankenberg and Renfree, 2013; Onichtchouk, 2016). Interestingly, the amphioxus PouIII genes, Brn1, Brn2 and Brn4 are expressed in the neuroectoderm at the gastrula stage and then in motor neurons at the neurula stage, suggesting that Brn1, Brn2 and Brn4 may control *Prdm14* expression in amphioxus motor neurons (Candiani et al., 2002). Therefore, the emergence of POU5F1 from PouIII genes in the vertebrate genome and their restricted expression in pluripotent cells and germ cells might have enabled restricted *Prdm14* expression in pluripotent cells and germ cells. Another possibility is that *Prdm14* expression changes arising during evolution were caused by the evolution of cis-regulatory elements neighbouring the *Prdm14* locus. Although the mouse genome contains POU5F1-recognition sequences in the distal enhancer of *Prdm14* (Galonska et al., 2015), the evolutionary origin and conservation of this element have not been investigated in non-mammalian vertebrates. We consider amphibians and reptiles to be key organisms for understanding how *Prdm14* was co-opted into the network for pluripotency-associated transcription factors during evolution.

In the early development of amniotes, including reptile and mammals, the pluripotent state of epiblast cells should be maintained for the long term compared with non-amniote vertebrates to ensure the maturation of extra-embryonic tissues. Integration timing of the PRDM14-CBFA2T2 axis from neurogenesis to pluripotent cells in the evolution of deuterostomes is closely associated with the emergence of amniotes. *Prdm14* disruption destabilises pluripotency networks (Chia et al., 2010; Yamaji et al., 2013), whereas PRDM14 overexpression stabilises pluripotency networks in mESCs and hESCs (Okashita et al., 2015; Tsuneyoshi et al., 2008), suggesting that co-opting PRDM14 into transcriptional networks for pluripotency in pluripotent cells might have served some roles in the emergence of amniotes during evolution. Therefore, we consider it to be an essential challenge to uncover the *de novo* mechanisms stabilising pluripotency transcriptional networks during the emergence of mammals. In summary, our work comparing the expression pattern and function of PRDM14 offers a framework for addressing how pluripotency transcriptional networks evolve in deuterostomes.

MATERIALS AND METHODS

Phylogenetic tree inference

Amino acid sequences of putative homologues were retrieved from aLeaves (Kuraku et al., 2013), incorporating Ensembl release 84. Multiple-sequence alignment was performed with MAFFT software, version 7.299b (Katoh and Standley, 2013) using the '-linsi' option. The aligned sequence sets were processed using Trimal software, version 1.4, rev15 (Capella-Gutierrez et al., 2009), with the '-automated1' and '-nogaps' options. Molecular phylogenetic trees were inferred with the maximum-likelihood method using the RAxML program, version 8.2.8 (Stamatakis, 2014) with the '-m PROTCATWAG -f a -# 1000' option. Bootstrap probabilities were obtained with 1000 re-samplings.

Establishment of PRDM14 orthologue-expressing cells

m*Prdm14* KO ESCs (Yamaji et al., 2013) were maintained in N2B27 basal medium with 3 µM CHIR99021 (Sigma) and 0.4 µM PD0325901 (Wako) on mitomycin-treated STO feeder cells. DNA encoding FLAG-tagged s*Prdm14*, am*Prdm14* (Eurofins Genomics) and s*Cbfa2t* (GENEWIZ) were chemically synthesised based on their known sequences and were cloned into the pCAGGS-IRES-Puro vector or the pCAGGS-IRES-Blast vector. *Prdm14* KO ESCs were transfected with each expression vector and selected in medium containing 0.5 µg/ml puromycin or 5 µg/ml blasticidin.

Gene expression analysis

Total RNA was extracted using TRIzol (Thermo Fisher Scientific) according to the manufacturer's instructions. cDNA was synthesised using ReverTra Ace qPCR RT Master Mix (TOYOBO), according to the manufacturer's instructions. Subsequently, cDNA was amplified with FastStart SYBR Green Master (Roche) using gene-specific primers (Table S1), and fluorescence was measured using the Light Cycler 96 system (Roche).

Total RNA was purified using the PureLink RNA Mini Kit (Thermo Fisher Scientific) and labelled using the GeneChiP WT Terminal Labeling and Control Kit (Thermo Fisher Scientific). The labelled single-stranded cDNA was hybridised to the GeneChiP Mouse Gene 2.1 ST Array Strip (Thermo Fisher Scientific). The signal intensities of microarray probes were scanned and quantitated using the GeneAtlas Personal Microarray System (Thermo Fisher Scientific). CELL files of Gene ChIP data were normalised using the Robust Multi-array Average program with the default settings in the Affymetrix Expression Console. Microarray data have been deposited in GEO under accession number GSE112904.

ChIP experiments

ChIP was performed as described previously (Tachibana et al., 2008). Briefly, DNA–protein complexes were crosslinked in phosphate-buffered saline containing 1% formaldehyde for 10 min at room temperature.

Development (2019) 146, dev168633. doi:10.1242/dev.168633

Fixation was terminated by the addition of 125 mM glycine. The fixed cells were lysed in sodium dodecyl sulphate (SDS) lysis buffer and sonicated with a BioRuptor. The ChIP samples were mixed with antibodies against the FLAG epitope (Sigma), TET1 (Millipore), TET2 (Abcam), SUZ12 (Abcam) and H3K27me3 (Millipore), as well as SureBeads Protein A or G Magnetic Beads (BioRad) overnight at 4°C. DNA-protein-SureBead Protein A or G complexes were collected using a magnetic rack (BioRad). De-fixed DNA was purified by phenol-chloroform extraction, followed by ethanol precipitation. Precipitated DNA was used as a template for qPCR analysis with gene-specific primers (Table S1).

AP staining

m*Prdm14* KO ESCs expressing the empty vector or a PRDM14 orthologue were maintained in N2B27 basal medium with 2i plus LIF. To evaluate the self-renewal activity of ESCs, 5×10^4 cells were plated in 12-well plates in medium containing serum plus LIF and passaged every 3 days. Nine days after cultivation in medium containing serum plus LIF, the cells were subjected to AP staining, as described previously (Cox and Singer, 1999).

Western blot and IP experiments

Cells were lysed by boiling in SDS sample buffer (Wako), and proteins were separated on SDS-polyacrylamide gels. Next, the proteins were electroblotted to polyvinylidene difluoride membranes, and the membranes were probed using the following primary antibodies: anti-FLAG (Sigma, F1804; 1/500), anti-DNMT3A (Abcam, ab13888; 1/500), anti-DNMT3B (Abcam, ab176166; 1/500), anti-DNMT3L (Abgent, AP1062a; 1/500), anti-histone (Abcam, ab1791; 1/5000) and anti-TY1 (Sigma, SAB4800032; 1/2000). Following binding of the primary antibody, the membrane was incubated with horseradish peroxidase-coupled secondary antibody (Santa Cruz, sc-2005; 1/2500). Detection was achieved using the ImmunoStar Zeta luminescent reagent (Wako).

For IP analysis, the cells were lysed in lysis buffer [50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% NP40, and 0.1% Triton-X]. The lysates were incubated with an anti-FLAG antibody (Sigma; F1804) and SureBeads Protein A or Protein B (BioRad) for 1 h at room temperature. The FLAG-tagged proteins bound to SureBeads Protein A or Protein B were captured using a magnetic rack and the precipitated proteins were eluted with 20 mM glycine.

DNA methylation analysis

Genomic DNA was purified from ESCs using the Wizard SV Genomic DNA Purification System (Promega) according to the manufacturer's instructions. Bisulphite treatment was performed using the EpiTect Fast DNA Bisulfite Kit (Qiagen) according to the manufacturer's instructions. Bisulphite-treated DNA was amplified by TAKARA EpiTaq HS (TAKARA) with gene-specific primers (Table S1). Glucosylation of genomic DNA followed by methylation-sensitive qPCR analysis was performed as described previously (Okashita et al., 2016). The sequences of the primers used in this study are shown in Table S1.

Model structure building and refinement

The model structure for the sea urchin PRDM14 complex with CBFA2T2 was built by the homology modelling method using the previously reported mouse PRDM14-CBFA2T2 complex structure (Nady et al., 2015) (Protein Data Bank code; 5ECJ) as a modelling template. Homology modelling was performed with SWISS-MODEL (Biasini et al., 2014). The initial model structures were refined by rigid-body refinement, energy minimisation and simulated annealing with no experimental energy terms to remove model bias. The refinement was performed by CNS version 1.3 (Brünger et al., 1998). The figures containing molecular structures and electrostatic molecular surface were prepared with PyMOL. The region corresponding to the disorder region in mouse PRDM14 (amino acid number 217-235) in sea urchin PRDM14 model structure was omitted from the molecular structure.

Whole-mount *in situ* hybridisation with amphioxus embryos and image acquisition

To synthesise riboprobes, amphioxus cDNA fragments were amplified as templates as previously described (Wu et al., 2011). DIG-labelled *Prdm14* and *Mnx/Hb9* antisense riboprobes and fluorescein-labelled *Islet* antisense riboprobes were synthesised using T7 RNA polymerase (Roche). Whole-mount *in situ* hybridisation was performed on the 4% paraformaldehyde-fixed amphioxus embryos as described previously (Wu et al., 2011). The AP substrates NBT and BCIP (Roche) were used for colour detection. For double fluorescent *in situ* hybridisation, anti-DIG-POD and anti-fluorescein-POD antibodies (Roche) were used to detect the riboprobes, followed by the amplification of fluorescence signal using the TSA Plus Cyanine 3 and Fluorescein Evaluation Kit (PerkinElmer). 4′,6-diamidino-2-phenylindole (DAPI) (Invitrogen, 1 µg/ml in PBS containing 0.1% Tween) was used for nuclear staining. The results were photographed using either a Zeiss AxioCam MRC camera mounted on a Zeiss Imager A2 microscope or a Zeiss confocal system (Zeiss LSM 880 with Airyscan).

Acknowledgements

We thank Andrew Johnson for providing genomic information for axolotl *Prdm14* and for stimulating discussions. We thank Kentaro Kajiwara for the constructing of plasmids. We thank Mitinori Saitou for providing the m*Prdm14* KO ESCs.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: Y. Seki; Methodology: Y. Seki; Validation: Y. Seki; Investigation: Y. Seki, M.K., K.S., C.-Y.L., S.K., Y. Suwa, S. Hashimoto, L.W.Y., K.T., S. Higashida, D.K.; Writing - original draft: Y. Seki; Writing - review & editing: Y. Seki, S.K., K.M., J.-K.Y.; Supervision: Y. Seki; Project administration: Y. Seki; Funding acquisition: Y. Seki, J.-K.Y.

Funding

This study was supported by a Grant-in-Aid for Young Scientists (A) (Japan Society for the Promotion of Science KAKENHI grant number 24681040), by a Grant-in-Aid for Scientific Research (B) (Japan Society for the Promotion of Science KAKENHI grant number 18H02422), by the Promotion of Joint International Research (Japan Society for the Promotion of Science KAKENHI grant number 15KK0262), by Scientific Research on Innovative Areas, 'Epigenome dynamics and regulation in germ cells' (Ministry of Education, Culture, Sports, Science, and Technology KAKENHI grant numbers 26112514 and 16H01223), by Scientific Research on Innovative Areas, 'Mechanisms regulating gamete formation in animals' (Ministry of Education, Culture, Sports, Science, and Technology KAKENHI grant number 16H01258), by support from Academia Sinica, Taiwan (to J.-K.Y.) and by support from the Ministry of Science and Technology, Taiwan (grant number MOST 102-2311-B-001-011-MY3 to J.-K.Y.).

Data availability

Microarray data have been deposited in GEO under accession number GSE112904.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.168633.supplemental

References

Aaker, J. D., Patineau, A. L., Yang, H.-J., Ewart, D. T., Gong, W., Li, T., Nakagawa, Y., Mcloon, S. C. and Koyano-Nakagawa, N. (2009). Feedback regulation of NEUROG2 activity by MTGR1 is required for progression of neurogenesis. *Mol. Cell. Neurosci.* **42**, 267-277.

Almeida, R. D., Loose, M., Sottile, V., Matsa, E., Denning, C., Young, L., Johnson, A. D., Gering, M. and Ruzov, A. (2012). 5-hydroxymethyl-cytosine enrichment of non-committed cells is not a universal feature of vertebrate development. *Epigenetics* **7**, 383-389.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126-140.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L. et al. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252-W258.

Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921.

Butler, A. M., Owens, D. A., Wang, L. and King, M. L. (2018). A novel role for sox7 in Xenopus early primordial germ cell development: mining the PGC transcriptome. *Development* **145**, dev155978.

Candiani, S., Castagnola, P., Oliveri, D. and Pestarino, M. (2002). Cloning and developmental expression of AmphiBrn1/2/4, a POU III gene in amphioxus. *Mech. Dev.* **116**, 231-234.
Cao, Y., Zhao, H. and Grunz, H. (2002). XETOR regulates the size of the proneural domain during primary neurogenesis in Xenopus laevis. *Mech. Dev.* **119**, 35-44.
Capella-Gutierrez, S., Silla-Martinez, J. M. and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973.
Chatfield, J., O'reilly, M.-A., Bachvarova, R. F., Ferjentsik, Z., Redwood, C., Walmsley, M., Patient, R., Loose, M. and Johnson, A. D. (2014). Stochastic specification of primordial germ cells from mesoderm precursors in axolotl embryos. *Development* **141**, 2429-2440.
Chia, N.-Y., Chan, Y.-S., Feng, B., Lu, X., Orlov, Y. L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.-S. et al. (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* **468**, 316-320.
Cox, W. G. and Singer, V. L. (1999). A high-resolution, fluorescence-based method for localization of endogenous alkaline phosphatase activity. *J. Histochem. Cytochem.* **47**, 1443-1456.
Dai, H.-Q., Wang, B.-A., Yang, L., Chen, J.-J., Zhu, G.-C., Sun, M.-L., Ge, H., Wang, R., Chapman, D. L., Tang, F. et al. (2016). TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. *Nature* **538**, 528-532.
Dehal, P. and Boore, J. L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* **3**, e314.
Eguchi, R., Yoshigai, E., Koga, T., Kuhara, S. and Tashiro, K. (2015). Spatiotemporal expression of Prdm genes during Xenopus development. *Cytotechnology* **67**, 711-719.
Eisen, J. S. (1994). Development of motoneuronal phenotype. *Annu. Rev. Neurosci.* **17**, 1-30.
Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
Extavour, C. G. and Akam, M. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* **130**, 5869-5884.
Ferrier, D. E. K., Brooke, N. M., Panopoulou, G. and Holland, P. W. H. (2001). The Mnx homeobox gene class defined by HB9, MNR2 and amphioxus AmphiMnx. *Dev. Genes Evol.* **211**, 103-107.
Frankenberg, S. and Renfree, M. B. (2013). On the origin of POU5F1. *BMC Biol.* **11**, 56.
Frankenberg, S. R., Frank, D., Harland, R., Johnson, A. D., Nichols, J., Niwa, H., Scholer, H. R., Tanaka, E., Wylie, C. and Brickman, J. M. (2014). The POU-er of gene nomenclature. *Development* **141**, 2921-2923.
Fresques, T., Zazueta-Novoa, V., Reich, A. and Wessel, G. M. (2014). Selective accumulation of germ-line associated gene products in early development of the sea star and distinct differences from germ-line development in the sea urchin. *Dev. Dyn.* **243**, 568-587.
Galonska, C., Ziller, M. J., Karnik, R. and Meissner, A. (2015). Ground state conditions induce rapid reorganization of core pluripotency factor binding before global epigenetic reprogramming. *Cell Stem Cell* **17**, 462-470.
Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532.
Jackman, W. R., Langeland, J. A. and Kimmel, C. B. (2000). islet reveals segmentation in the Amphioxus hindbrain homolog. *Dev. Biol.* **220**, 16-26.
Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., Robson, P., Zhong, S. and Ng, H.-H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat. Cell Biol.* **10**, 353-360.
Johnson, A. D. and Alberio, R. (2015). Primordial germ cells: the first cell lineage or the last cells standing? *Development* **142**, 2730-2739.
Johnson, A. D., Crother, B., White, M. E., Patient, R., Bachvarova, R. F., Drum, M. and Masi, T. (2003). Regulative germ cell specification in axolotl embryos: a primitive trait conserved in the mammalian lineage. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**, 1371-1379.
Katoh, K. and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780.
Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Lomelí, H., Nagy, A., Mclaughlin, K. J., Schöler, H. R. et al. (2004). Oct4 is required for primordial germ cell survival. *EMBO Rep.* **5**, 1078-1083.
Khoueiry, R., Sohni, A., Thienpont, B., Luo, X., Velde, J. V., Bartoccetti, M., Boeckx, B., Zwijsen, A., Rao, A., Lambrechts, D. et al. (2017). Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat. Genet.* **49**, 1061-1072.
Kunath, T., Saba-el-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S. and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**, 2895-2902.
Kuo, C. T., Veselits, M. L., Barton, K. P., Lu, M. M., Clendenin, C. and Leiden, J. M. (1997). The LKLF transcription factor is required for normal tunica media formation and blood vessel stabilization during murine embryogenesis. *Genes Dev.* **11**, 2996-3006.
Kuraku, S., Zmasek, C. M., Nishimura, O. and Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **41**, W22-W28.
Liu, C., Ma, W., Su, W. and Zhang, J. (2012). Prdm14 acts upstream of islet2 transcription to regulate axon growth of primary motoneurons in zebrafish. *Development* **139**, 4591-4600.
Martín-Durán, J. M. and Hejnol, A. (2015). The study of Priapulus caudatus reveals conserved molecular patterning underlying different gut morphogenesis in the Ecdysozoa. *BMC Biol.* **13**, 29.
Nady, N., Gupta, A., Ma, Z., Swigut, T., Koide, A., Koide, S. and Wysocka, J. (2015). ETO family protein Mtgr1 mediates Prdm14 functions in stem cell maintenance and primordial germ cell formation. *eLife* **4**, e10150.
Nichols, J. and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-492.
Niwa, H. (2014). The pluripotency transcription factor network at work in reprogramming. *Curr. Opin. Genet. Dev.* **28**, 25-31.
Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T. and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. *Cell* **137**, 571-584.
Okashita, N., Kumaki, Y., Ebi, K., Nishi, M., Okamoto, Y., Nakayama, M., Hashimoto, S., Nakamura, T., Sugasawa, K., Kojima, N. et al. (2014). PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development* **141**, 269-280.
Okashita, N., Sakashita, N., Ito, K., Mitsuya, A., Suwa, Y. and Seki, Y. (2015). PRDM14 maintains pluripotency of embryonic stem cells through TET-mediated active DNA demethylation. *Biochem. Biophys. Res. Commun.* **466**, 138-145.
Okashita, N., Suwa, Y., Nishimura, O., Sakashita, N., Kadota, M., Nagamatsu, G., Kawaguchi, M., Kashida, H., Nakajima, A., Tachibana, M. et al. (2016). PRDM14 drives OCT3/4 recruitment via active demethylation in the transition from primed to naive pluripotency. *Stem Cell Rep.* **7**, 1072-1086.
Onichtchouk, D. (2016). Evolution and functions of Oct4 homologs in non-mammalian vertebrates. *Biochim. Biophys. Acta* **1859**, 770-779.
Sasaki, K., Yokobayashi, S., Nakamura, T., Okamoto, I., Yabuta, Y., Kurimoto, K., Ohta, H., Moritoki, Y., Iwatani, C., Tsuchiya, H. et al. (2015). Robust in vitro induction of human germ cell fate from pluripotent stem cells. *Cell Stem Cell* **17**, 178-194.
Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.
Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H. and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J.* **27**, 2681-2690.
Tang, W. W. C., Dietmann, S., Irie, N., Leitch, H. G., Floros, V. I., Bradshaw, C. R., Hackett, J. A., Chinnery, P. F. and Surani, M. A. (2015). A unique gene regulatory network resets the human germline epigenome for development. *Cell* **161**, 1453-1467.
Tapia, N., Reinhardt, P., Duemmler, A., Wu, G., Araúzo-Bravo, M. J., Esch, D., Greber, B., Cojocaru, V., Rascon, C. A., Tazaki, A. et al. (2012). Reprogramming to pluripotency is an ancient trait of vertebrate Oct4 and Pou2 proteins. *Nat. Commun.* **3**, 1279.
Tsankov, A. M., Gu, H., Akopian, V., Ziller, M. J., Donaghey, J., Amit, I., Gnirke, A. and Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518**, 344-349.
Tsuneyoshi, N., Sumi, T., Onda, H., Nojima, H., Nakatsuji, N. and Suemori, H. (2008). PRDM14 suppresses expression of differentiation marker genes in human embryonic stem cells. *Biochem. Biophys. Res. Commun.* **367**, 899-905.
Tu, S., Narendra, V., Yamaji, M., Vidal, S. E., Rojas, L. A., Wang, X., Kim, S. Y., Garcia, B. A., Tuschl, T., Stadtfeld, M. et al. (2016). Co-repressor CBFA2T2 regulates pluripotency and germline development. *Nature* **534**, 387-390.
Vervoort, M., Meulemeester, D., Béhague, J. and Kerner, P. (2016). Evolution of Prdm genes in animals: insights from comparative genomics. *Mol. Biol. Evol.* **33**, 679-696.
Wessel, G. M., Fresques, T., Kiyomoto, M., Yajima, M. and Zazueta, V. (2014). Origin and development of the germ line in sea stars. *Genesis* **52**, 367-377.
Wu, H.-R., Chen, Y.-T., Su, Y.-H., Luo, Y.-J., Holland, L. Z. and Yu, J.-K. (2011). Asymmetric localization of germline markers Vasa and Nanos during early development in the amphioxus Branchiostoma floridae. *Dev. Biol.* **353**, 147-159.
Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N. and Tada, T. (2005). Nanog expression in mouse germ cell development. *Gene Expr. Patterns* **5**, 639-646.
Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y. and Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat. Genet.* **40**, 1016-1022.
Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K. and Saitou, M. (2013). PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368-382.
Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hubner, K. and Scholer, H. R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881-894.

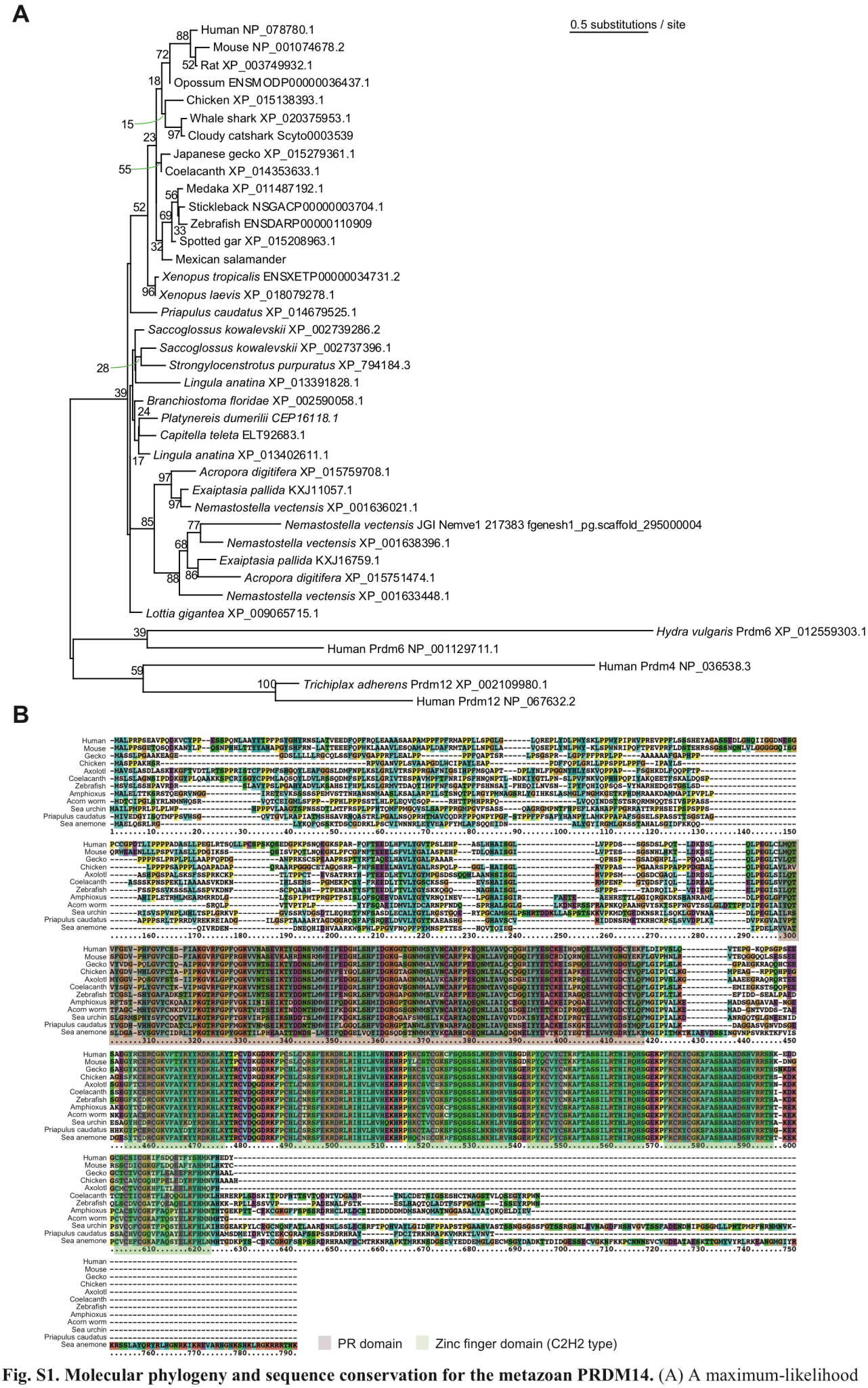

Fig. S1. Molecular phylogeny and sequence conservation for the metazoan PRDM14. (A) A maximum-likelihood tree was constructed using amino acid sequences of Prdm14 orthologues and human Prdm4/6/12, hydra Prdm6 and Trichoplax adhaerens Prdm12 as and outgroup. (B) Sequence alignment of PRDM14 orthologue proteins. The pink restriction box indicates the position of the PR domain and the green restriction box indicates the position of the zinc finger domain.

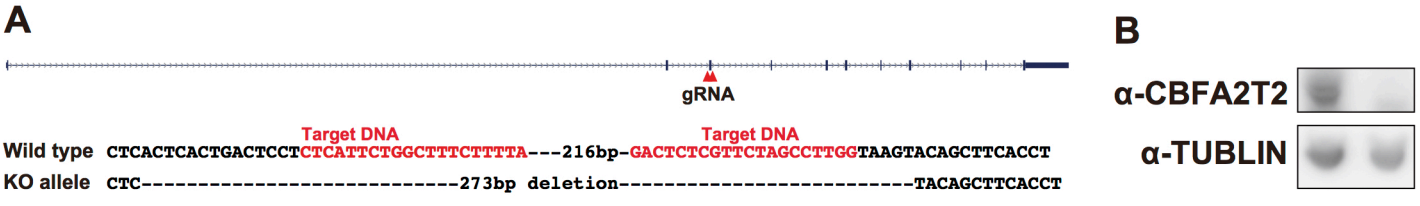

Fig. S2. Establishment of m*Cbfa2t2* KO cell line using CRISPR/Cas9 sysetm. (A) CRISPR design and the deletion sequence of m*Cbfa2t2* KO cell line. (B) Western blot analysis of CBFA2T2.

Table S1. Primer lists for qRT-PCR, bisulfite sequencing and ChIP-qPCR.

qRT-PCR

Gene	Strand	Sequence(5'-3')
Oct3/4	Forward	CTCTCCCATGCATTCAAACTG
	Reverse	CCCCTGTTGTGCTTTTAATCC
Nanog	Forward	TTTGGAGGTGAATTTGGAAGC
	Reverse	TCACCTGGTGGAGTCACAGAG
Esrrb	Forward	CGTGTGACAAGGAGACAGGAG
	Reverse	TCCAGCCACAACGTCATTATC
Tcl1	Forward	TGGCCTCACTAGAACAAGAGG
	Reverse	CTCGGTCAAGGATGGAAGC
Klf2	Forward	CCCCAGGAAAGAAGACAGGAG
	Reverse	AGGCATTTCTCACAAGGCATC
Tcl1	Forward	TGGCCTCACTAGAACAAGAGG
	Reverse	CTCGGTCAAGGATGGAAGC
Tbx3	Forward	TGATGTTTTAAGAGCCGATGC
	Reverse	AGGATAATGGGACTTCCGTTG
Klf4	Forward	GACCAGGATTCCCTTGAATTG
	Reverse	ACCAAGCACCATCATTTAGGC
Klf5	Forward	TGGAAGTCCCGATAGACAAGC
	Reverse	GTGGCAGGTAAATTTGGGTTG
Dnmt3a2	Forward	CAGACGGGCAGCTATTTACAG
	Reverse	TGGTTCTCTTCCACAGCATTC
Dnmt3b	Forward	CTCGCAAGGTGTGGGCTTTTGTAAC
	Reverse	CTGGGCATCTGTCATCTTTGCACC
Dnmt3l	Forward	CTGGTGAAGAACTGCCTTCTC
	Reverse	AAACTGTGGAGGGAAGAGACC
Prdm14 CDS	Forward	TGTGGTACGGAAATGGCTATG
	Reverse	AAACACCTTTCCACAGCGTTC
Prdm14 3'UTR	Forward	GGAATCCATTCAGACCAGGAG
	Reverse	GCACATAGTCGCTGGCTACAG
Gata6	Forward	TTGCCTCCAAATCATGTGCTTC
	Reverse	GCCTCCAGGATAGACCAAATG
Fgf5	Forward	ATGAGTGCATCTGCTCTGCTC
	Reverse	CGTCTGTGGTTTCTGTTGAGG
T	Forward	AAGGACAGAGAGACGGCTGTG
	Reverse	AAAGTAGGACAGGGGGTGGAC
Cdx2	Forward	GTAAATGCCAGAGCCAACCTG
	Reverse	GGCTTGTTTGGCTCGTTACAC
Dazl	Forward	GATGGACATGAGATCATTGGAC
	Reverse	ATACCAGGGAGCAATCCTGAC
Sycp3	Forward	CGAGCAGTTCATAAAGAGTTTG
	Reverse	TCTTGCTGCTGAGTTTCCATC

Asz1	Forward	GAGTGGGCTTCTCCCAGAAA
	Reverse	GGTCATTTTCCCGCTCATTC
Gja1	Forward	GTGCAAGTGTGTAAGCGTGTG
	Reverse	CACAAAGATCCATGAGGAAGG

Bisulfite sequenceing

Gene	**Strand**	**Sequence(5'-3')**
Tcl1	Forward	GAAATAGGAGGGTTAGGGAGATTTTAGATG
	Reverse	TTTCTTTTAAACACCAACATTAAAACCCAT
Dazl	Forward	CCACTTCTCTTTTCTACACC
	Reverse	CCACTCTTATCCTCCAAACC
Sycp3	Forward	AAGGGTTAGGTTTTTTTAGA
	Reverse	AACTTCCTACCTAAATACCCAA
Asz1	Forward	TTGGAGAGAAAAAGATTTTT
	Reverse	TACCATAACTCCAAACTATTCT

ChIP-qPCR

Gene	**Strand**	**Sequence(5'-3')**
Esrrb	Forward	GGTTGCTTTCTTTTGCTGGTG
	Reverse	TGATCCTTTGGAGTGGAGGAC
Tfcp2l1	Forward	TCATCCTTATCCTCCCAGCAG
	Reverse	GGGGAAGAGGAAAGTGGATTC
Fgfr1	Forward	GCTTGGCCTTGGATGAATTGTTGGC
	Reverse	AGCCAGGTTGGCCTTTTGTCCT
Dnmt3b (-10.1 kb)	Forward	CAGGAAATGCGTGCCTAGAGG
	Reverse	AGGCTTTTCACTTGAGGGCTG
Dnmt3b (-6.3 kb)	Forward	CTGCCACTACCACCAACAAAC
	Reverse	TTTCCACAGGAACACTCATGC