Essential role of the TFIID subunit TAF4 in murine embryogenesis and embryonic stem cell differentiation

Diana Langer[1], Igor Martianov[1,*], Daniel Alpern[1,2,*], Muriel Rhinn[3,*], Céline Keime[1], Pascal Dollé[3], Gabrielle Mengus[1] & Irwin Davidson[1]

TAF4 (TATA-binding protein-associated factor 4) and its paralogue TAF4b are components of the TFIID core module. We inactivated the murine *Taf4a* gene to address Taf4 function during embryogenesis. Here we show that *Taf4a*$^{-/-}$ embryos survive until E9.5 where primary germ layers and many embryonic structures are identified showing Taf4 is dispensable for their specification. In contrast, Taf4 is required for correct patterning of the trunk and anterior structures, ventral morphogenesis and proper heart positioning. Overlapping expression of *Taf4a* and *Taf4b* during embryogenesis suggests their redundancy at early stages. In agreement with this, *Taf4a*$^{-/-}$ embryonic stem cells (ESCs) are viable and comprise Taf4b-containing TFIID. Nevertheless, *Taf4a*$^{-/-}$ ESCs do not complete differentiation into glutamatergic neurons and cardiomyocytes *in vitro* due to impaired preinitiation complex formation at the promoters of critical differentiation genes. We define an essential role of a core TFIID TAF in differentiation events during mammalian embryogenesis.

[1] Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, 1 Rue Laurent Fries, 67404 Illkirch, France. [2] L'École polytechnique fédérale de Lausanne, Route Cantonale, 1015 Lausanne, Switzerland. [3] Department of Development and Stem Cells, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, 1 Rue Laurent Fries, 67404 Illkirch, France. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to I.D. (email: irwin@igbmc.fr).

Formation of the RNA polymerase II (Pol II) preinitiation complex (PIC) requires a set of general transcription factors among which is TFIID, a multiprotein complex composed of the TATA-binding protein (TBP) and a set of TBP-associated factors (TAFs)[1]. TFIID assembles from a 'core complex' comprising TAF4, TAF5, TAF6, TAF9 and TAF12, followed by TAF8-TAF10, and then association with a second module comprising TBP, TAF1, TAF2 and TAF7 (refs 2,3). In mammals, core TAFs are almost ubiquitously expressed, although cell-specifically expressed paralogues of TBP and a subset of TAFs have been described and their functions defined by mouse knockouts. Taf7l plays a critical role in male germ cell development and in adipocytes[4–6], Taf4b is essential for male and female fertility[7,8] and Taf9b regulates neuronal gene expression[9]. Trf2 and Trf3, the two TBP paralogues, play essential roles in the male and female germ lines, respectively[10,11].

TAF4 forms a histone fold heterodimer with TAF12 (refs 12–14) that associates with the TAF6–TAF9 heterodimer and TAF5 to form the TFIID core module. TAF4 is crucial for the structural integrity of the core module and TFIID[1,15]. Its paralogue TAF4B also heterodimerizes with TAF12 and integrates into TFIID, thus maintaining TFIID integrity in the absence of TAF4 (ref. 16). While we have used somatic inactivation to address the function of murine Taf4 in mouse embryonic fibroblasts (MEFs), the adult murine epidermis or neonatal liver[16–19], its role in embryogenesis is unknown.

Here we characterise $Taf4a^{-/-}$ (gene encoding Taf4) embryos showing they survive until E9.5, but exhibit severe growth retardation and specific defects in anterior and ventral patterning and morphogenesis. We also show that $Taf4a^{-/-}$ embryonic stem cells (ESCs) are viable, but are unable to properly differentiate into glutamatergic neurons and cardiomyocytes *in vitro* due to impaired PIC formation at the promoters of critical differentiation genes. Thus, while Taf4b can compensate for the absence of Taf4 during the early stages of embryogenesis and in ESCs, Taf4 plays specific roles in differentiation *in vivo* and *in vitro*.

Results

Mid-gestation embryonic lethality of $Taf4a^{-/-}$ embryos. Mice with floxed and null $Taf4a$ alleles[16] ($Taf4a^{lox/-}$ and $Taf4a^{+/-}$) were viable, fertile and did not exhibit any apparent abnormalities. Breeding of $Taf4a^{+/-}$ mice did not yield any $Taf4a^{-/-}$ pups, indicating death *in utero*[16]. We, therefore, analysed embryogenesis at various developmental stages after crosses of $Taf4a^{+/-}$ animals.

$Taf4a^{-/-}$ embryos could be distinguished from $Taf4a^{+/+}$ and $Taf4a^{+/-}$ embryos by their smaller overall size at E6.5 (Fig. 1a–c). By E7.5 mutant embryos displayed a more compact appearance (Fig. 1d–g), and by E8.5, a shortened body axis and under-developed head and trunk regions (Fig. 1h–k). Between E8.5 and E9.5, mutant embryos showed virtually no increase in overall size, but developmental processes continued with for example formation of a primitive heart. Nevertheless, the overall morphology of $Taf4a^{-/-}$ embryos showed that their development did not proceed much beyond that of E8.5 wild-type (WT) embryos (Fig. 1k–m and see below). At E9.5, mutant embryos were severely growth retarded and subsequently resorbed (Fig. 1l–o). Genotyping indicated that $Taf4a^{-/-}$ embryos were present in approximately Mendelian ratios until E9.5 after which time they died and were longer detected (Supplementary Table 1).

Histology analysis showed that WT E7.5 embryos had completed division of the proamniotic cavity into the exocoelomic and the amniotic cavities, separated by the amniotic membrane (Fig. 1r), whereas in $Taf4a^{-/-}$ embryos this compartmentalization had not
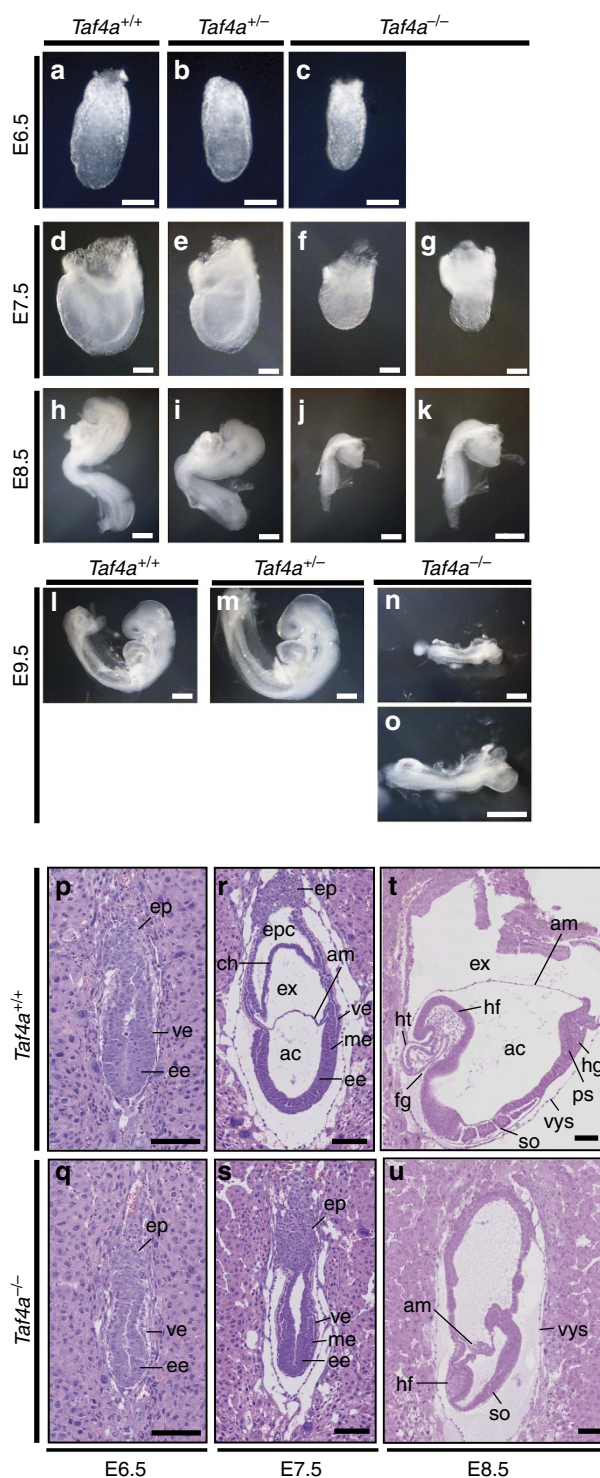
Figure 1 | Taf4 is required during mid-gestation. (**a–c**) E6.5 stage embryos of the indicated genotypes. (**d–g**) E7.5-stage embryos of the indicated genotypes, **g** shows a blow up of **f**. Scale bar, 100 μm. (**h–k**). E8.5-stage embryos of the indicated genotypes, **k** shows a blow up of **j**. (**l–o**) E9.5-stage embryos of the indicated genotypes, **o** shows a blow up of **n**. Scale bar, 200 μm. (**p–u**) Haematoxylin and eosin-stained sections through WT or mutant embryos at the indicated stages. Scale bar, 100 μm. ep, ectoplacental cone; ve, visceral endoderm; ee, embryonic ectoderm; me, mesoderm;, epc, ectoplacental cavity; ch, chorion; ex, exocoelomic cavity; am, amnion; ac, amniotic cavity; hf, headfold; ht, heart; fg, foregut; hg, hindgut; so, somites; vys, visceral yolk sac; ps, primitive streak.

Table 1 | Quantification of observed phenotypes in Taf4a$^{-/-}$ embryos.

Gestational age	Cardiac defects				Other defects	
	No Heart	Heart in amniotic cavity	Heart in exocoelmic cavity	Heart outside visceral yolk sac	Allantois not fused	Embryo not turned
E8.5	22/24 (92%)	ND	ND	ND	NA	NA
E9.5	18/66 (27%)	1/25 (4%)	21/25 (84%)	3/25 (12%)	19/21 (90%)	66/66 (100%)

ND, not determined; NA, not applicable

occurred resulting in an open proamniotic cavity (Fig. 1s). Taf4a$^{-/-}$ embryos showed reduced size and overall resembled WT embryos at the egg cylinder stage at E6.5 (Fig. 1p,q). WT E8.5 embryos underwent major developmental events, including foregut invagination, concurrent headfold and heart tube development, and formation of the first somite pairs (Fig. 1t). In contrast, Taf4a$^{-/-}$ embryos displayed severely under-developed anterior structures with no apparent heart structure formation or foregut invagination. Headfold-like mesenchymal tissues were compact, showed poor morphogenesis and reduced size. A large section of the trunk region was missing and somites were small, compact and tightly spaced (Fig. 1u and see below Fig. 3). At E8.5 the developing amniotic structure appeared agglutinated and failed to elongate properly to surround the embryo.

Taf4a$^{-/-}$ embryos showed growth defects at E6.5 and most severe by E9.5. To examine this, we measured 5-bromo-2'-deoxyuridine (BrdU) incorporation, but no major change in proliferation rates was seen (Supplementary Fig. 1A). In contrast, we observed increased apoptosis in mutant embryos assessed by terminal deoxynucleotidyl nick-end labelling on whole-mount embryos and paraffin sections (Supplementary Fig. 1B,C). Increased apoptosis occurring as early as E7.5, rather than reduced cell proliferation, therefore contributes to the smaller size of the Taf4a$^{-/-}$ embryos.

Overlapping embryonic expression of Taf4a and Taf4b. Whole-mount in situ hybridization indicated ubiquitous expression of Taf4a (Supplementary Fig. 2A). A specific signal was observed with the anti-sense probe throughout the embryo and the extraembryonic region at E6.5 and E7.5. By E8.5, Taf4a mRNA was detected in the entire embryo and in the yolk sac. Consistent with the in situ hybridization, Taf4 protein was detected throughout the E7.5 WT embryo, but was absent from the mutant embryos (Supplementary Fig. 2B).

As Taf4a$^{-/-}$ embryos die much later than germ line knockouts for Tbp or other core Tafs, it is possible that Taf4b may at least partially compensate for Taf4 loss. In situ hybridization showed ubiquitous Taf4b expression at E6.5 that was more pronounced in the extraembryonic region (Supplementary Fig. 2A). At E7.5 Taf4b was expressed through-out the epiblast and extraembryonic regions. Notably, however, expression in the visceral endoderm and the definitive endoder-mal layer was weaker than Taf4a. Distinct enrichment in Taf4b expression in regions corresponding to the ectoplacental cavity and in a ring of extraembryonic ectoderm at the mid proximo-distal domain was observed. At E8.5 Taf4b expression closely resembled that of Taf4a, with the exception of the yolk sac where staining was weaker than Taf4a.

The overlapping expression suggests Taf4b may partially compensate for lack of Taf4a and thus account for the later death of Taf4a mutants. Indeed, Taf4b expression was detected throughout the Taf4a mutant embryos at E8.5 (Supplementary Fig. 2C). Quantitative reverse transcription–PCR (RTq–PCR)

analysis confirmed that its expression was virtually unchanged in the mutant embryos (Supplementary Fig. 2D). All other TFIID Tafs were comparably expressed in WT and mutant embryos. Only Taf5, Taf6 and Taf13 showed >2-fold increased expression in the mutant embryos (Supplementary Fig. 2D).

Defective heart formation in Taf4a$^{-/-}$ embryos. The yolk sac comprises a bilayer structure of mesoderm- and endoderm-derived cells crucial for embryo survival. The first blood cells and blood islands develop in the mesodermal compartment of the yolk sac[20–22]. Before E8.5 mouse embryos are in a lordotic position such that the embryo initially faces the outside of the egg cylinder (ref. 23 and Fig. 2a). Around E8.5 the embryo rotates around its own axis placing the heart in a ventral position (Fig. 2a,b) and the amnion enlarges to surround the whole embryo. The exocoelome lining consisting of visceral endoderm and mesoderm also expands and eventually surrounds the whole embryo as the visceral yolk sac.

Taf4a$^{-/-}$ embryos failed to turn by E9.5 developing partially outside the yolk sac that was reduced compared with WT (Fig. 2c–f). Mutant embryos lacked the blood-filled vascular plexus and the developing blood vessels seen in WT yolk sacks, suggesting defective erythroid cell generation (Fig. 2c–f). In mutant yolk sacks, endothelial cells were present, but did not form organized vascular structures (Supplementary Fig. 3A). Immunolabeling with PECAM-1, an endothelial cell-specific maker, revealed formation of a primitive vascular system only on the dorsal surface of the Taf4a$^{-/-}$ embryos adjacent to the neural tube that failed to pervade the interior of the embryo (Supplementary Fig. 3A). Defective allantois fusion to the chorionic plate was also observed. At E9.5, fusion is complete in 100% of the WT embryos, compared with only 10% of mutant embryos (Fig. 2k,l and Table 1).

E9.5 WT embryos displayed a well-organized heart tube with erythrocytes visible in the inflow region (Fig. 2d), whereas mutant embryos showed primitive cardiac structures bulging from the headfold base often developing outside the amnion in the exocoelomic cavity (Fig. 2g–i and Table 1). In rare cases, heart structures developed outside the visceral yolk sac (Fig. 2j–l). While 100% of WT embryos developed a primitive heart tube by E8.5, only 8% of mutant embryos showed a small bulge indicating a severely retarded initiation of heart tube formation (Fig. 2g–l and Table 1). At E9.5 all WT hearts showed contractions on dissection, whereas this was never observed in mutant embryos.

In situ hybridization showed specfic expression of cardiac transcription factor Nkx2–5 in the prospective myocardium of the atria, ventricles, and outflow tract in E8.5 WT embryos (Fig. 2m,o). In contrast, mutant embryos exhibited a crescent-like expression domain with no evident heart chamber formation indicating that although cardiac lineage specification occured, heart tube morphogenesis was defective (Fig. 2n,p). Expression of a second cardiac transcription factor Tbx5 was similar to Nkx2–5 although this factor was expressed at higher levels in inflow tract
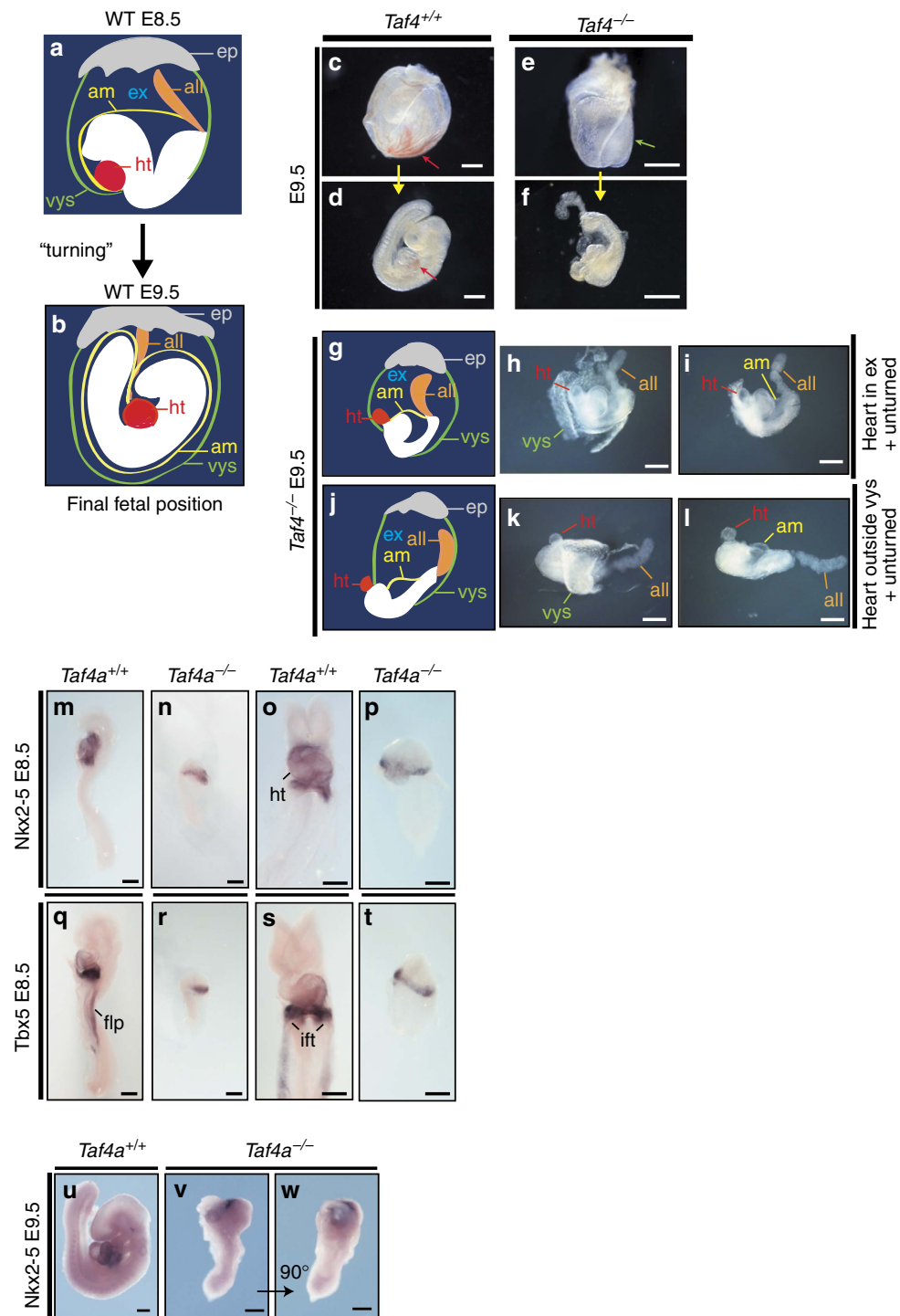
Figure 2 | Taf4 is required for correct positioning of head and heart. (**a,b**). Schematic representations illustrate normal development at E8.5 and E9.5 highlighting positions of head and heart during embryo turning. (**c–f**) WT and mutant embryos at E9.5 before and after dissection from the yolk sac highlighting the blood-filled yolk sac vascular plexus in control embryos (red arrow in **c**) and the embryo protruding partially out of the yolk sac (green arrow in **e**). WT embryos show a strong accumulation of blood in the atrium of the heart (red arrow in **d**), which was absent in $Taf4a^{-/-}$ embryos (**f**). (**g–i**) Schematic illustration of observed mutant phenotypes at E9.5. All mutant embryos fail to undergo turning at E9.5 and fusion of the allantois to the chorion. In addition to this, mutant embryos develop heart structures either in the exocoelomic cavity (**h–i**) or outside of the visceral yolk sac (**k–l**). (**m–w**) *In situ* hybridization of wild-type and $Taf4a^{-/-}$ embryos for the cardiac transcription factors *Nkx2-5* at E8.5 (**m–p**) or *Tbx5* at E8.5 (**q–t**) and E9.5 (**u–w**). ep, ectoplacental cone; all, allantois; ht, heart; flp, forelimb precursors; ift, inflow tract. Scale bar, 500 µm (**c–f**), 300 µm (**h–i**), 200 µm (**m–w**).

structures and showed additional staining in the prospective forelimb field (Fig. 2q,s). In mutant embryos, Tbx5 showed a crescent-like expression domain, with no expression in the prospective forelimb (Fig. 2r,t). These crescent-like domains in

the E8.5 mutant embryos closely resembled the expression pattern of heart-specific markers in the cardiac crescent of E7.5 WT embryos. Importantly however, histology analyses and *in situ* hybridization indicated mis-localization of the primitive heart
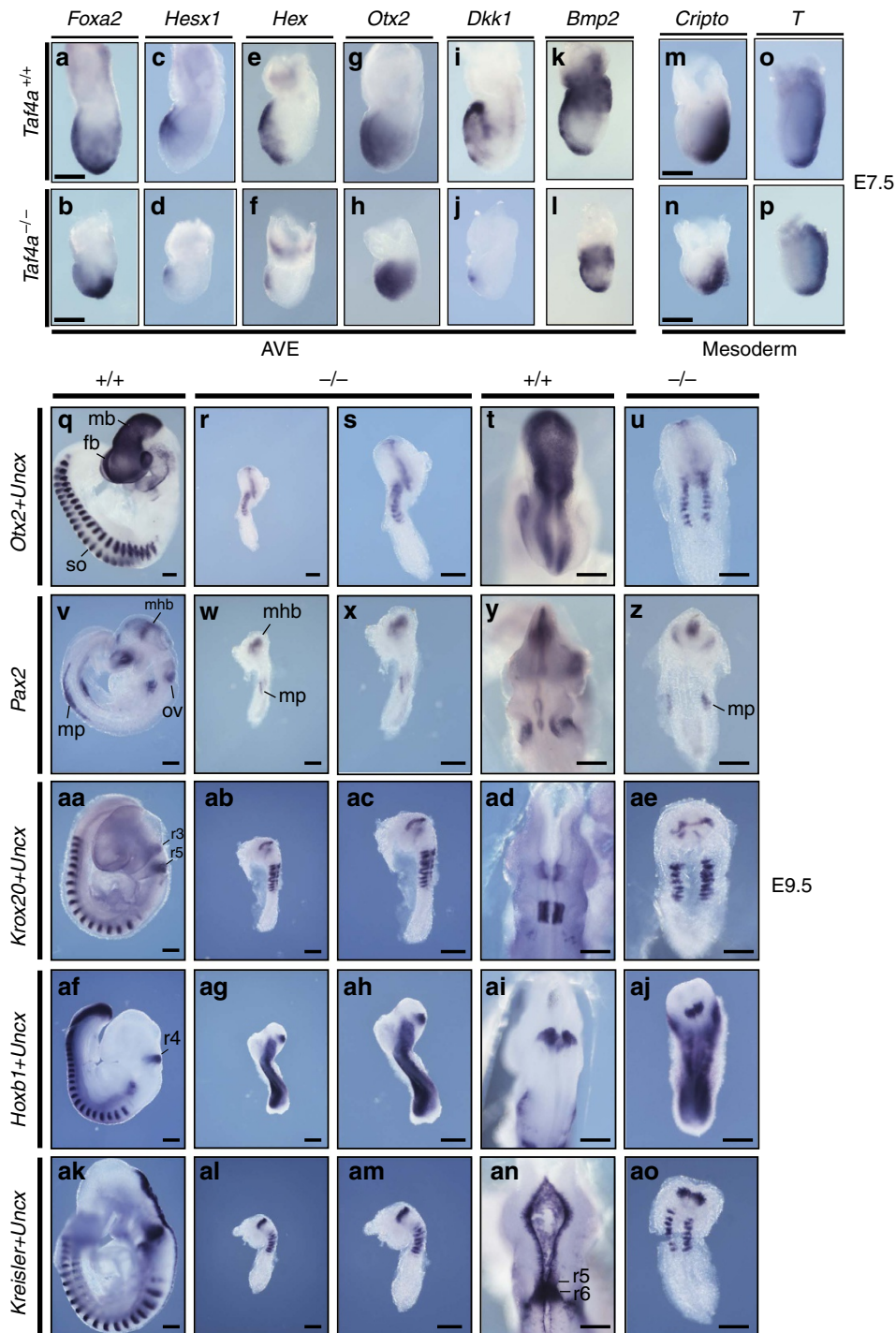
Figure 3 | Abnormal patterning of *Taf4a*$^{-/-}$ embryos. (**a–o**) (upper row). *In situ* hybridization with the indicated probes for anterior visceral endoderm (AVE) and mesoderm markers on E7.5 WT embryos. (**b–p**) (lower row). *In situ* hybridization with the indicated probes on E7.5 *Taf4a*$^{-/-}$ embryos. (**q–ao**). *In situ* hybridization with neuronal markers in E9.5 WT or *Taf4a*$^{-/-}$ embryos (lateral views **q–ak**; ventral view R dorsal views T-AN) and (lateral views RP-AL, higher magnifications of same embryos **s–am**; ventral view U; dorsal views **z–ao**). Scale bar, 200 μm in all panels.

structures anterior to the headfolds rather than in the ventral position in keeping with the absence of turning of the mutant embryos. This was highlighted by *in situ* hybridization with Nkx2–5 at E9.5 labelling the developing heart structure at the anterior pole of the embryo next to the headfolds rather than in a ventral position as in WT (Fig. 2u–w). *Taf4a*$^{-/-}$ embryos therefore display defective cardiac development that, together with lack of chorio-allantoic fusion and defective yolk sac and

intra-embryonic vascularization, likely accounts for embryonic death at mid-gestation.

Abnormal anterior patterning of *Taf4a*$^{-/-}$ embryos. We next used specific markers to investigate defective anterior development of *Taf4a*$^{-/-}$ embryos, The anterior visceral endoderm (AVE) regulates initial steps of head, heart and foregut

positioning. Defective expression of AVE markers leads to severe anterior truncation phenotypes[24]. Expression domains of several AVE markers in E7.5 $Taf4a^{-/-}$ embryos were either reduced (*Hesx1*, *Hex* and *Dkk*1) or showed retarded formation (*Foxa2*, *Otx2*, Fig. 3a–l). In contrast, expression domains of mesodermal markers *Brachyury* (*T*) and *Cripto* were comparable in WT and mutant embryos (Fig. 3m–p).

To investigate differentiation of specific anterior structures, we studied expression of forebrain, midbrain and hindbrain markers. *Otx2* is required for induction and maintenance of the anterior neural plate[25,26] and specification of forebrain and

midbrain regions. The *Otx2* expression domain was diminished in E9.5 $Taf4a^{-/-}$ mutants, particularly the forebrain was strongly reduced (Fig. 3q–u). *Pax2* was expressed in the developing midbrain–hindbrain region, the otic vesicle, and the mesonephros in WT embryos, whereas in mutant embryos midbrain–hindbrain staining was reduced, and the presence of nephrogenic precursor cells was observed. In contrast, no staining of presumptive otic vesicle structures was observed (Fig. 3v–z).

In WT E9.5 embryos, *Krox20* marked rhombomeres 3 and 5. In mutant embryos, labelled structures were thinner and
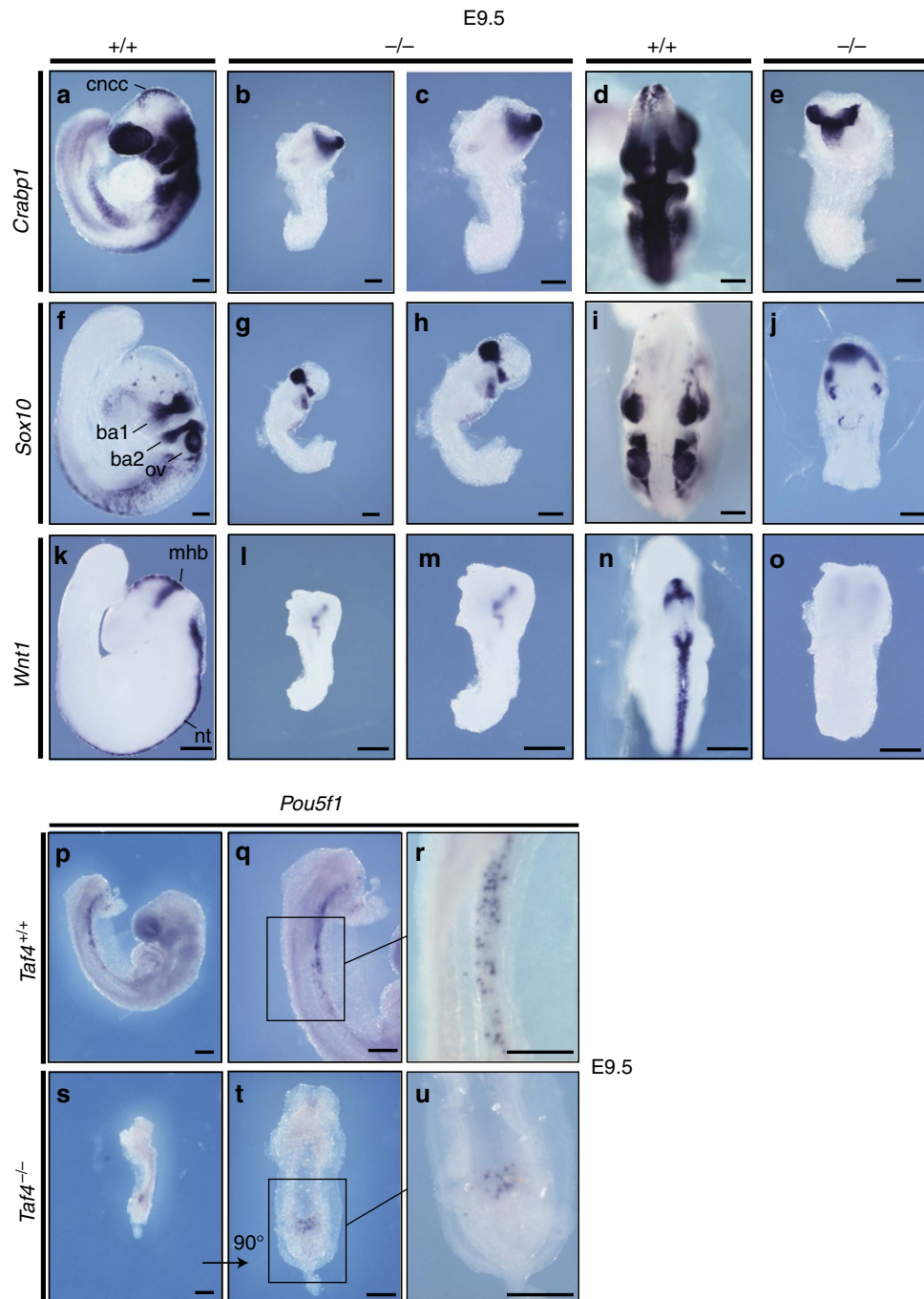
Figure 4 | Neural crest cell markers and primordial germ cells in *Taf4* mutant embryos. (**a**-**o**) *In situ* hybridizations for neural crest markers *Sox10*, *Crabp1* and *Wnt1* in E9.5 control (lateral views **a**,**f**,**k** dorsal views **d**,**i**,**n**) and $Taf4a^{-/-}$ embryos (lateral views **b**,**g**,**l** higher magnifications of the same embryos **c**,**h**,**m** dorsal views E, J, O). (**p**-**u**). *In situ* hybridizations at E9.5 with *Pou5f1* (*Oct4*) as a marker for primordial germ cells (PGCs). ba1, first branchial arch; ba2, second branchial arch; ov, otic vesicle; cncc, cranial neural crest cells; mhb, midbrain–hindbrain boundary; nt, neural tube. Scale bar, 200 μm in all panels.

disorganized with rhombomere 5 more affected than rhombomere 3 (Fig. 3aa–ae). *Hoxb1*, a marker for rhombomere 4, was clearly detected in mutant embryos (Fig. 3af–aj), whereas *Mafb* (*Kreisler*), expressed in rhombomeres 5 and 6 in WT embryos (Fig. 3ak,an), stained only a single stripe in *Taf4a*$^{-/-}$ mutants (Fig. 3al,am,ao), which in some instances, was thinner than the *Hoxb1*-positive stripe (compare Fig. 3ag,al). Thus, although the anterior region of mutant embryos was severely atrophied, this did not reflect a complete lack of forebrain/midbrain/hindbrain specification, all of which were present albeit both reduced in size and disorganized. *Uncx* staining showed fewer (typically 8–11) and irregularly shaped somites in E9.5 *Taf4a*$^{-/-}$ mutants.

The above analyses revealed an immediate proximity of the hindbrain and somite markers (*Uncx*) and a reduction or absence of the hindbrain posterior to rhombomere 5 along with a significant portion of the trunk. We analysed the anterior–posterior axis and gut development in mutant embryos. *Foxa2* and *Shh* labelled the notochord and developing hindgut in E9.5 WT embryos, but only one structure in mutant embryos (Supplementary Fig. 3B). Brachyury (*T*) staining indicated the labelled structure was the notochord showing interruption at one or several places, perhaps due to local apoptosis or tissue distortion (Supplementary Fig. 3C). The notochord showed stronger anterior staining by *Brachyury* and *Foxa2* indicating defective elongation along the anterior–posterior axis. Nevertheless, mutant embryos maintained strong *Brachyury* staining in their posterior region, which was proportionally larger than in WT embryos at E8.5 and E9.5. This could indicate a defect in specification and differentiation of axial mesendodermal structures. *Taf4a*$^{-/-}$ embryos therefore fail to develop a detectable gut tube consistent with histology analysis that showed lack of a foregut pocket E8.5 (Fig. 1u) and their shortened notochord is consistent with the absence of an elongated trunk region.

Reduced neural crest domain in *Taf4a*$^{-/-}$ embryos. Having defined defective anterior and trunk development, we next investigated neural crest cells (NCCs) normally induced at the midbrain level and as a wave extending towards the tail region. Defects in NCCs may contribute to the defects observed in the cephalic and trunk region of the *Taf4a*$^{-/-}$ embryos.

In E9.5 WT embryos, *Crabp1* labelled the neural tube and NCCs originating in the posterior hindbrain and anterior trunk regions that give rise to cranial NCCs migrating into the branchial arches and posterior regions (Fig. 4a,d). A severely reduced *Crabp1* expression domain was seen in mutant embryos, indicating they were completely devoid of migrating NCCs in both the cranial and posterior regions (Fig. 4b,c,e). The residual y-shaped expression domain may correspond to part of the neural tube and primitive branchial arch 1. Similarly, *Sox10*, a marker with more restricted expression domains in the otic vesicle and the NCCs migrating into the branchial arches (Fig. 4f,i), also showed severely reduced expression in the mutant embryo and no migrating NCCs were observed (Fig. 4g,h,j). Furthermore, *Wnt1* expression was absent from the spinal cord and NCCs (Fig. 4k–o). *Taf4a*$^{-/-}$ embryos therefore lack a large portion of the NCCs migrating towards craniofacial structures and more posterior trunk regions.

Taf4 is dispensable for primordial germ cell generation. Given the role of Taf4 in reprogramming (ref. 27 and see below), we specifically investigated the presence of primordial germ cells (PGCs) in *Taf4a*$^{-/-}$ embryos. PGCs colonize the developing gonads by migration giving rise to oocytes and spermatocytes in adults[28]. PGCs can be identified at around E7.25 (refs 29,30) and begin to migrate through the primitive streak to the developing hindgut, and by ≈ E9.5 towards the genital ridges. Mutant embryos lack a properly developed hindgut mesentery at E9.5. We investigated whether mutants were able to initiate PGC development. At E7.5 *Pou5f1* (*Oct4*) is exclusively expressed by the newly formed PGCs, and expression persists as they migrate and proliferate. We readily detected *Pou5f1*-expressing PGCs in the WT hindgut mesentery at E9.5 (Fig. 4p–r). In mutant embryos, *Pou5f1*-expressing PGCs could be observed, but were around 10-fold less numerous than in WT (Fig. 4s–u). These PGCs accumulated at the presumptive entry site into the hindgut lacking in mutant embryos. Thus, PCG specification takes place in absence of Taf4, but they cannot migrate and pursue their proliferation due to the absence of the hindgut.

***Taf4a*$^{-/-}$ ESCs are viable and comprise TFIID containing Taf4b**. The lethality, reduced size and pleiotropic defects of the *Taf4a*$^{-/-}$ embryos preclude biochemical and functional analysis of specific differentiation events. Nevertheless, as death occured at mid-gestation, it should be possible to derive *Taf4a*$^{-/-}$ ESCs and use them to address the role of Taf4 in specific differentiation programmes. We isolated ESCs from blastocysts of *Taf4a*$^{+/-}$ heterozygous crosses and established two homozygous null lines and two WT lines from these crosses. The absence of Taf4 protein in *Taf4a*$^{-/-}$ ESCs was confirmed by immunostaining, while pluripotency markers Pou5f1, Nanog and Sox2 were normally expressed (Supplementary Fig. 4A–C). WT and *Taf4a*$^{-/-}$ ESCs showed comparable alkaline phosphatase staining, proliferation rates and cell cycle profiles (Supplementary Fig 4D–F). Comparative RNA sequencing (RNA-seq) revealed only minor differences in gene expression. Around 40 genes were upregulated although only 12 showed more than a threefold change, whereas 166 genes were downregulated (Supplementary Fig. 4G) including Noggin and several genes associated with Wnt signalling. However, their altered expression did not elicit detectable changes in cell physiology and no change in pluripotency gene expression was observed. Together these data show that *Taf4a*$^{-/-}$ ESCs are viable and comparable in most essential parameters to WT ESCs.

To characterise ESC TFIID and to compare it with human ESCs and MEFs, we performed western blot analysis. WT mouse ESCs express Tbp and all tested Tafs, including Taf4b. For equivalent amounts of protein extract, expression of many Tafs was higher in ESCs than in MEFs (Fig. 5a), consistent with previous reports[27]. In two human ESC lines, no expression of TAF1, TAF4 and TAF5 was seen, although TAF4b was detected. These observations confirm previous reports of non-canonical TFIID composition in human ESCs[31], and highlight fundamental differences in the biochemistry of mouse and human ESCs.

Comparison of WT and *Taf4a*$^{-/-}$ ESCs indicated comparable expression of most TAFs. Tbp was immunoprecipitated (IP) from both lines and equivalent amounts of Tbp loaded on immunoblots (Supplementary Fig. 5). Taf4b was co-precipitated with Tbp in both the presence and absence of Taf4 indicating that Taf4b maintained TFIID integrity in *Taf4a*$^{-/-}$ ESCs. Nevertheless, reduced levels of several TAFs, with respect to Tbp, were observed in the Tbp IP from the *Taf4a*$^{-/-}$ ESCs suggesting that Taf4b expression was not sufficient to maintain an overall TFIID level comparable to WT ESCs.

We then performed gel filtration and analysed the presence of Tbp and TAFs in the high-molecular mass fractions. Tbp and TAFs coeluted in a high molecular mass complex of around 2 mDa in WT extracts (Fig. 5b). Taf4b was also present in these fractions, although a subfraction eluted in a lower molecular mass complex. In *Taf4a*$^{-/-}$ extracts, Taf4b was present along with
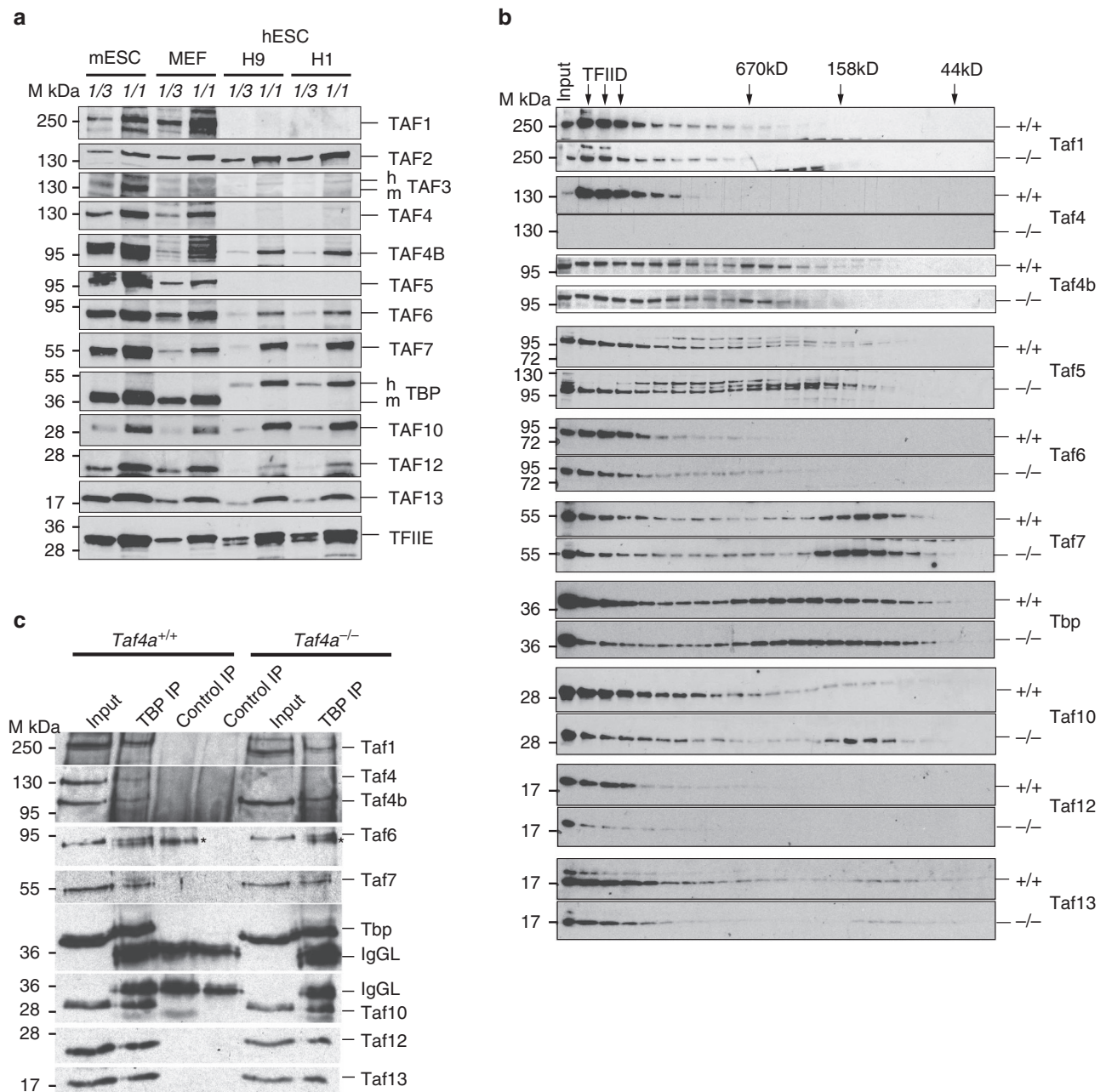
Figure 5 | Biochemical analysis of TFIID in *Taf4a* ^{−/−} ES cells. (a). Immunoblots showing expression of TFIID subunits in ES cells, MEFs and two human ES cell lines, H1 and H9. (**b**). Immunoblots on gel filtration fractions of nuclear extracts from WT and mutant ES cells. (**c**). Immunoblots of Tbp IPs from the pooled TFIID fractions indicated by arrows in **b**. * indicates a non-specific band migrating just below Taf6.

Tbp and other tested TAFs in high-molecular mass fractions, although Tbp and many TAFs were present in reduced amounts in particular Taf7 and Taf10 that were increased in lower molecular weight fractions.

To investigate whether the high-molecular mass fractions correspond to TFIID, the indicated fractions were pooled and subjected to Tbp IP. In fractions from WT and mutant ESCs, Tafs 1, 4b, 6, 7, 10, 12 and 13 were co-precipitated with Tbp showing formation of a TFIID complex in both cell types (Fig. 5c). In fractions from WT cells, both Taf4 and Taf4b were precipitated, while in the fractions from mutant cells only Taf4b was present. Taf4b therefore integrates in TFIID in *Taf4a* ^{−/−} ESCs, but these cells contain lower overall levels of TFIID as the presence of Taf4b cannot quantitatively compensate Taf4 loss.

To further verify potential redundancy between Taf4 and Taf4b, we transfected WT and *Taf4a* ^{−/−} ESCs with control small interfering RNA (siRNA) and siRNAs directed against Taf4b. After control siRNA and siTaf4b silencing in WT ESCs and replating the transfected cells, comparable numbers of alkaline phosphatase staining colonies were observed expressing comparable levels of Nanog, showing that Taf4b knockdown did not compromise cell viability (Supplementary Fig. 6A–C). In contrast, Taf4b silencing in *Taf4a* ^{−/−} ESCs led to a strong reduction in the number of colonies and surviving cells showed normal Taf4b and Nanog expression. Hence, Taf4b silencing in *Taf4a* ^{−/−} ESCs strongly compromised cell viability and only cells that had escaped transfection formed colonies. These results indicate that ESCs lacking either Taf4 of Taf4b were viable. In

contrast, loss of both proteins through genetic deletion of Taf4 and subsequent Taf4b silencing resulted in a loss of viability. Taf4 and Taf4b are therefore redundant in terms of assuring cell ESC viability.

Defective neuronal differentiation of $Taf4a^{-/-}$ ES cells. $Taf4a^{-/-}$ ESCs were grown as embryoid bodies (EBs) to determine their spontaneous differentiation capacity over a period of 14 days. $Taf4a^{-/-}$ EBs initially appeared indistinguishable from WT EBs, but at later stages, gave rise to fewer cystic-like structures (Supplementary Fig. 7A) that resemble the developing yolk sac *in vivo*[32]. Hence, *in vitro* differentiated $Taf4a^{-/-}$ ES cells recapitulated to some extent the phenotype of the $Taf4a^{-/-}$ embryos. Taf4 was expressed throughout WT EB differentiation, while Taf4b expression was strongly decreased, but Taf4b was maintained in the mutant cells (Supplementary Fig. 7B).

We investigated whether $Taf4a^{-/-}$ ESCs could give rise to the three primary germ layers. Expression of the pluripotency markers Nanog and Dppa3 was reduced, but incompletely repressed compared with WT (Supplementary Fig. 7B). The early endoderm and mesoderm markers Gata4 and Gata6 were induced, whereas late endodermal lineage markers like Foxa2, and Afp were strongly downregulated (Supplementary Fig. 7B). Mesendodermal markers Brachyury (T) and Goosecoid (Gsc) showed a strong upregulation in day 6 EBs before their expression returned to normal levels, indicating that Taf4 may regulate the fine-tuning of their expression. Expression of the early cardiac marker Nkx2–5 and its downstream target gene Anf was also reduced. $Taf4a^{-/-}$ ESCs can therefore initiate differentiation into early mesoendodermal-type cells, whereas later stage differentiation markers were more strongly affected.

Expression of the early ectodermal marker Fgf5 was induced, but expression of later neuronal markers such as Pax6 and Mtap2 and Dcx was strongly reduced (Supplementary Fig. 7B). To investigate the role of Taf4 in neuronal differentiation, we differentiated WT and $Taf4a^{-/-}$ ESCs into glutamatergic neurons by treating EBs with all-*trans* retinoic acid (RA) (Supplementary Fig. 8A, refs 33,34). Expression of the vesicular glutamate transporter Vglut2 as well as Taf4 was observed in such neurons cultured for 12 days (Supplementary Fig. 8B).

Dissociated RA-treated WT EBs differentiated into neuronal precursors shortly after plating and adopted neuronal morphology with abundant axons and prominent Tubb3 staining after 2 days (Fig. 6a,b). In contrast, mutant EBs failed to generate neuronal precursors with an axonal network and most cells died within a day. After 1–2 days, the few remaining cells appeared flattened and enlarged. Western blot showed only weak Tubb3 expression in mutant cells, but also revealed defective repression of Pou5f1 and Nanog whose expression was strongly repressed after differentiation of WT EBs (Fig. 6c). Taf4 is therefore essential for RA-induced neuronal differentiation and for pluripotency gene repression.

Immunoblot analysis showed a mild overall reduction in Tbp and Taf expression by day 8 of WT EB differentiation (Fig. 6d). On differentiation of mutant EBs, the overall reduction in Taf levels was less marked with no decreased Taf4b expression, whereas Taf3 and Taf5 were reduced.

To demonstrate that defective differentiation was due to Taf4 loss, we generated $Taf4a^{-/-}$ ESCs stably re-expressing exogenous full-length Taf4 (Fig. 6e). WT ESCs isolated in parallel after transfection with control vector showed normal neuronal differentiation with an abundant Tubb3-stained axonal network, whereas the $Taf4a^{-/-}$ EBs transfected with control vector showed no neuronal differentiation (Fig. 6f). Importantly, the rescue $Taf4a^{-/-}$ line showed comparable differentiation to the

WT lines. Defective differentiation was therefore directly attributable to Taf4 loss and restored by its re-expression. We also attempted to rescue differentiation by the expression of Taf4 (372–1,083) and Taf4 (805–1,083); however, no stable clones could be isolated suggesting that expression of truncated Taf4 was toxic in these cells. The integrity of Taf4 is therefore required to rescue the differentiation defects.

We performed RNA-seq on two independent WT and mutant ESC lines before (EB4) and 24 (EB5) and 96 h (EB8) after RA treatment. At EB4 585 deregulated genes were observed (354 down and 231 up, Fig. 7a). Similarly, at EB5, 255 upregulated and 287 downregulated genes were detected, but the highest number of deregulated genes was observed at EB8 (1,212 up and 1,008 down). Thus, Taf4b can replace Taf4 to maintain cellular housekeeping functions and pluripotency and to a limited extent the initial stages of differentiation, but Taf4 is required for full differentiation with progressively more severe alterations in gene expression as differentiation progressed (Fig. 7a).

At EB8, genes involved in neurogenesis are strongly downregulated, whereas many pluripotency genes are prominent in the upregulated class (Fig. 7b,c). Downregulated genes are highly associated with the brain, while upregulated genes are associated with ESCs and spermatogenesis reflecting the presence of stem-like genes expressed in PGCs and spermatogonia (Supplementary Fig. 8C). Clustering identified genes with distinct kinetics of expression in WT cells and altered expression in $Taf4a^{-/-}$ cells. Three clusters show the most prominent changes. Genes in cluster 2 were upregulated when WT ES cells were grown as EBs, but were most strongly induced after RA treatment at EB5 and were highly enriched in brain and neurogenesis functions (Fig. 7d). In mutant ESCs, their activation was strongly reduced whereby expression at EB8 was comparable to that of WT EB4. Genes in cluster 3 were also strongly enriched in brain and neurogenesis, but were mainly activated between EB5 and EB8 and were strongly reduced in mutant cells. In contrast, genes in cluster 4 are ES-expressed and pluripotency genes downregulated in WT EBs, then switched off during their differentiation, whereas in $Taf4a^{-/-}$ EBs, their expression persisted during differentiation.

Similar to $Taf4a^{-/-}$ cells *in vitro*, expression of several neuronal genes was downregulated in $Taf4a^{-/-}$ embryos showing that Taf4 is also required for their normal expression *in vivo* (Supplementary Fig. 8D). Expression of Pou5f1, and Nanog was, however, repressed in $Taf4a^{-/-}$ embryos, but that of other pluripotency associated genes like Zfp42 and Dppa5a remained elevated at E9.5.

The above data show that WT ESCs differentiated into neuronal precursors that switched off pluripotency gene expression, whereas $Taf4a^{-/-}$ ESCs adopt a unique state with partially induced expression of differentiation genes, and partially repressed expression of pluripotency genes.

Defective PIC formation at neuronal genes. The above classification based on kinetics of gene expression also reflects different regulatory mechanisms as revealed by ChIP-seq. TFIIB, Pol II, H3K4me3 and H3K27ac were present at promoters of genes in cluster 1 in WT ESCs. Consistent with increased expression of these genes during differentiation, TFIIB, Pol II and H3K27ac were increased at EB8 (Fig. 7e). In contrast, lower occupancy was seen in $Taf4a^{-/-}$ ESCs and little change was observed at EB8 in accordance with their reduced activation. TFIIB, Pol II, H3K4me3 and low levels of H3K27ac were observed at cluster 2 promoters in WT and mutant ESCs and activation of these genes was associated with increased occupancy in WT EB8, but not in $Taf4a^{-/-}$ EB8. Increased elongating Pol II was also observed in
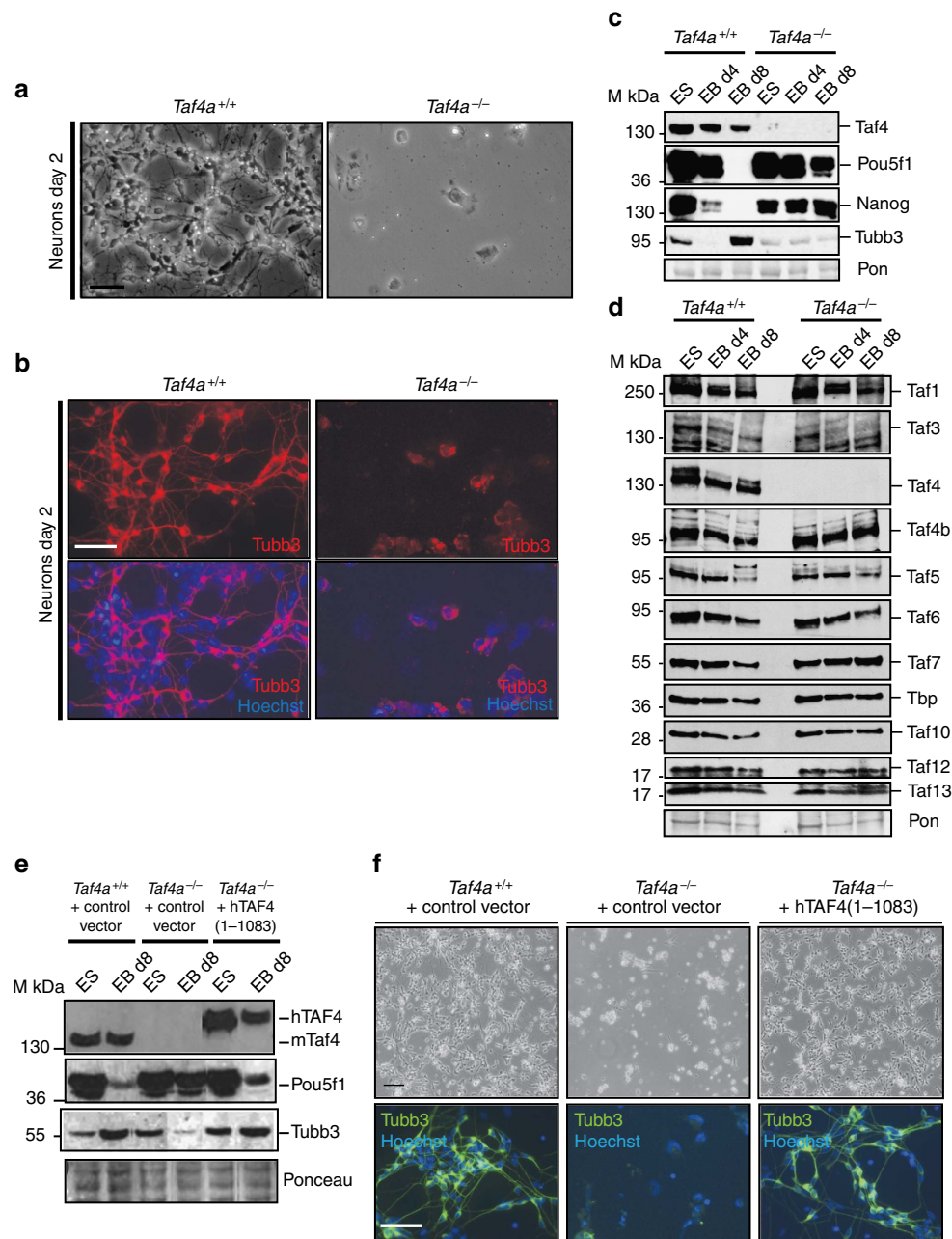
Figure 6 | Taf4 is required for RA-driven neuronal ES cell differentiation. (**a**) Phase-contrast microscopy of RA-treated EBs differentiated into neurons 2 days after plating. (**b**) Tubb3 labelling of RA-treated EBs differentiated into neurons 2 days after plating. (**c**,**d**). Immunoblots on extracts from differing ES cells and EBs with the indicated antibodies. The membranes were stained with Ponceau as loading controls. (**e**). Immunoblots on extracts from ES cell lines engineered to re-express recombinant Taf4. (**f**). Phase-contrast microscopy and Tubb3 labelling of RA-treated EBs differentiated into neurons 2 days after plating. All images were taken at ×20 magnification. Scale bars 100 μm.

WT, but not mutant EB8. The reduced activation of these genes therefore results from diminished PIC formation and elongating Pol II.

In contrast to clusters 1 and 2, where basal levels of PIC formation and Pol II occupancy were observed in ESCs, later stage activation of cluster 3 genes was associated with high levels of *de novo* TFIIB and Pol II recruitment and a potent increase in H3K27ac in WT EB8, but not mutant EB8 (Fig. 7e). Nevertheless, H3K4me3 was observed at these promoters in both WT and mutant ESCs suggesting that even in the absence of PIC formation they were epigenetically marked for future activation. Downregulation of genes in cluster 4 during differentiation was

associated with strongly diminished TFIIB, Pol II, H3K27ac and H3K4me3 in WT EB8 showing diminshed PIC formation that was less marked in mutant cells. Consequently, Tbp was redistributed during differentiation of WT ESCs with around 14,000 sites showing preferential or unique occupancy at EB8, while occupancy of many ES sites was strongly reduced (Supplementary Fig. 9A). Tbp was enriched at the TSS of 2,086 genes mainly present in clusters 2 and 3 and strongly enriched in ontology terms associated with brain and the nervous system consistent with redistribution of Tbp from ES-expressed/pluripotency genes to those involved in neuronal differentiation (Supplementary Fig. 9B).
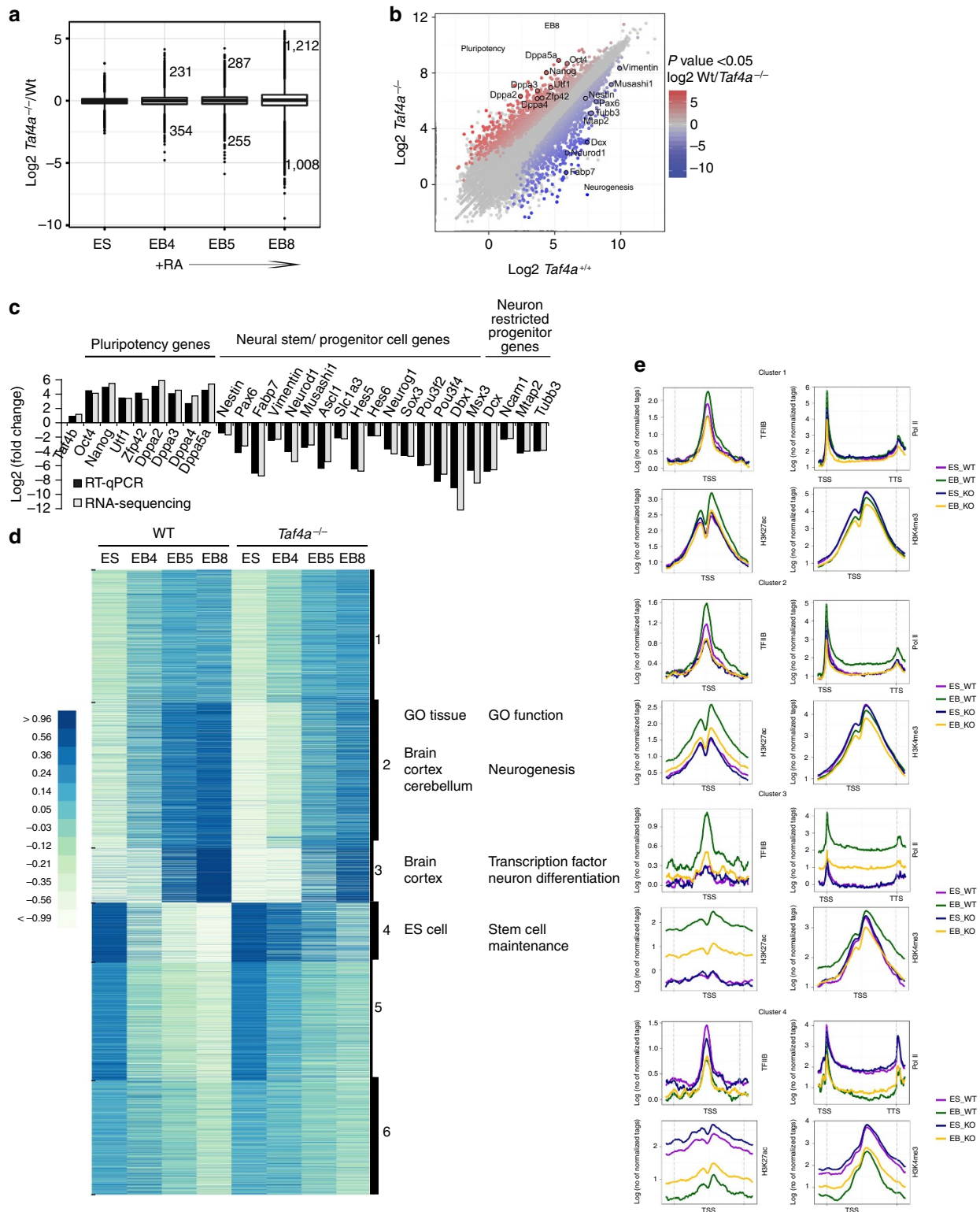
Figure 7 | Taf4 is required for activation of neurogenic genes in differentiating ES cells. (**a**). Gene expression changes at different stages of differentiation in *Taf4a*[−/−] compared with WT ES cells. The numbers of genes with altered expression and their fold changes at each stage are indicated. (**b**). Global comparison of RNA-seq data in *Taf4a*[−/−] compared to WT at the EB8 stage. The expression of several neurogenic and pluripotency genes are highlighted. (**c**). Comparison of the changes in expression of selected marker genes measured by RNA-seq and by RT–qPCR. (**d**). Kinetics of gene expression and their alterations during differentiation are represented as a clustered heatmap. The ontology terms associated with the genes in clusters 2–4 that are most affected by Taf4 loss are indicated. (**e**). Metaprofiles for the indicated ChIP-seq data sets corresponding to genes in clusters 2–4 are shown. For TFIIB, H3K4me3 and H3K27ac the profiles are centred on the TSS. For Pol II the profiles are show over the gene from the transcription start site (TSS) to the transcription termination site (TTS).

In contrast, in $Taf4a^{-/-}$ EB8, little occupancy of the neuronal associated WT EB8 sites was observed (Supplementary Fig. 9A). Instead, Tbp was enriched at EB8 at around 3,000 genes associated with embryos and a variety of tissues from the different germ layers (Supplementary Fig. 9C). A smaller group of around 3,000 sites appeared more strongly occupied in mutant EB8 that were associated with 400 genes with only poorly enriched ontology scores.

Together these observations are consistent with the idea that Tbp was redistributed from ES-expressed/pluripotency to neurogenic genes during differentiation of WT ESCs, whereas in $Taf4a^{-/-}$ ESCs, Tbp redistribution was diminished, and it remained associated with ES-expressed genes. Taf4 is therefore required for de novo PIC formation necessary for activation of neurogenic genes and for efficient dissociation of the PIC at pluripotency genes.

Taf4 is essential for contractile cardiomyocyte differentiation. In the spontaneous differentiation experiments, expression of the early cardiac markers Nkx2–5 and Anf was reduced and in accordance with this, we observed a lack of spontaneously contractile cardiomyocytes from the $Taf4a^{-/-}$ EBs, whereas contractile cardiomyocytes were frequently seen from WT EBs. To specifically examine cardiomyocyte differentiation, EBs were cultured for 5 days in medium containing 20% FCS, before plating on gelatin-coated plates for a further 9 days.

Spontaneously contracting cardiomyocytes were readily identified around 3–5 days after plating of WT EBs, whereas no contractile activity of mutant cells was observed during the 9-day period. In agreement with this, Tnnt2 immunostaining revealed formation of well-defined sarcomeric structures in the differentiated WT cells, whereas such structures were severely reduced, disorganized and patchy in the $Taf4a^{-/-}$ cells (Fig. 8a). Importantly, re-expression of full-length Taf4 restored proper Tnnt2 expression and formation of organized and functional sarcomeric structures restoring observable contractile areas (Fig. 8b, Supplementary Fig. 10A). Taf4 is therefore essential for differentiation of contractile cardiomyocytes in vitro.

RNA-seq from day 9 differentiated WT and $Taf4a^{-/-}$ ESCs identified 636 upregulated genes and 1,550 downregulated genes (Fig. 8c,d) associated with muscle contraction, mesoderm development, cell adhesion and signal transduction (Supplementary Fig. 10B). Expression of many early and late cardiomyocyte markers, including those involved in electrophysiology of contraction, were strongly reduced in mutant cells (Fig. 8d and Supplementary Fig. 10B). Nevertheless, as seen above, pluripotency gene expression was not properly repressed during cardiomyocyte differentiation. Taf4 is therefore required for normal activation of the gene expression programme required for differentiation of functional cardiomyocytes.

Discussion
We define an essential role of Taf4 in mouse embryogenesis. Previous inactivation of Tbp or ubiquitous core Tafs led to pre-implantation or early post-implantation embryonic death[35–38]. Due to this early lethality, TAF function at later stages of mammalian embryogenesis where complex lineage specification and differentiation processes occur is so far unknown. In contrast, $Taf4a^{-/-}$ embryos survive until E9.5 and display many developmental defects. The longer survival of these embryos may be explained by redundancy with Taf4b that maintains TFIID integrity and all functions required for early development and gastrulation, after which the essential role of Taf4 in specific developmental processes can be observed. This idea is supported by the overlapping expression of Taf4a and Taf4b at early stages

of development, the expression of Taf4b in the $Taf4a^{-/-}$ embryos and more directly by our observation that Taf4 and Taf4b are redundant for ESC viability. Moreover, $Taf4b^{-/-}$ mice are viable and show defects only at the adult stage[7,39]. Thus, while Taf4 can fully replace Taf4b during embryogenesis, Taf4b can compensate for Taf4 loss in ESCs and at early, but not later stages of embryogenesis. The observed redundancy may reflect the integration of Taf4b in TFIID in $Taf4a^{-/-}$ ESCs to maintain its integrity. However, we cannot exclude other interpretations such as the presence of Taf4b in an alternative TFIID-like complex functioning during early development, and a requirement for Taf4-containing TFIID only at later stages to regulate expression of a distinct subset of genes that govern further development.

$Taf4a^{-/-}$ embryos are characterized by several defects. Embryos displayed an overall atrophy resulting in part from increased apoptosis, the absence of a large trunk region and a strongly reduced forebrain domain. However, a striking feature is the absence of foregut invagination associated with lack of gut tube formation and aberrant ventral morphogenesis. In WT embryos, ventral folding morphogenesis achieves gut internalization, proper ventral localization of the heart and enclosure of the fetus in extraembryonic membranes[40]. $Taf4a^{-/-}$ embryos are characterized by incomplete enclosure in the yolk sac, misplacement of the heart that develops anterior to the head and a lack of embryo turning. This aspect of the phenotype strongly resembles embryos lacking Bmp2 that arrest before turning and show many similar features in particular misplacement of the heart anterior to the head[40]. Analysis of $Taf4a^{-/-}$ embryos thus confirms the importance of correct ventral morphogenesis in positioning of head and heart. Nevertheless, Taf4 is not required for the critical visceral endoderm expression of Bmp2 that persists in mutant embryos, suggesting that it may rather be required for the downstream transcriptional response of the Bmp2 pathway.

While the above phenotypes highlight processes defective in $Taf4a^{-/-}$ embryos, others occur almost normally. Taf4 is dispensable for many patterning events, activation of early neural and mesodermal makers and the posterior development of the embryo that, in contrast to the anterior, is almost normal. Importantly, Pou5f1-expressing PGCs are formed in $Taf4a^{-/-}$ embryos although they cannot migrate properly due to the absence of a developed hindgut. Generation of PGCs involves complex transcriptional and epigenetic reprogramming events including reactivation of pluripotency genes as well as changes in histone and DNA methylation[41–43]. Taf4 is therefore not required for this naturally induced reprogramming.

This observation contrasts with a reported role of Taf4 in maintaining ESC pluripotency and promoting induced pluripotent stem cell reprogramming in vitro. Pijnappel et al.[27], reported that Taf4 overexpression enhanced the ability of Pou5f1, Nanog, Sox2 and Myc to reprogram human and mouse fibroblasts to induced pluripotent stem cells in vitro. Furthermore, they reported that Taf4 or Taf4b short hairpin RNA knockdown resulted in spontaneous differentiation of ESCs implying that both are essential for maintaining stemness. They further proposed Taf4 as an integral component of the pluripotency network as Pou5f1 and Nanog bind a critical enhancer driving high Taf4a expression in ESCs. In contrast, Bahat et al.[44] used siRNA knockdown and reported that Taf4b plays a critical role in maintaining ES cell stemness, whereas Taf4 knockdown had no effect. Thus, there are conflicting data on the role of Taf4 and Taf4b in maintaining ESC pluripotency, perhaps related to the techniques used or the efficiencies of knockdown achieved.

Our use of genetically modified mice and ESCs clearly shows that Taf4 is dispensable for pre-implantation development in vivo and that $Taf4a^{-/-}$ ESCs are almost indistinguishable from WT
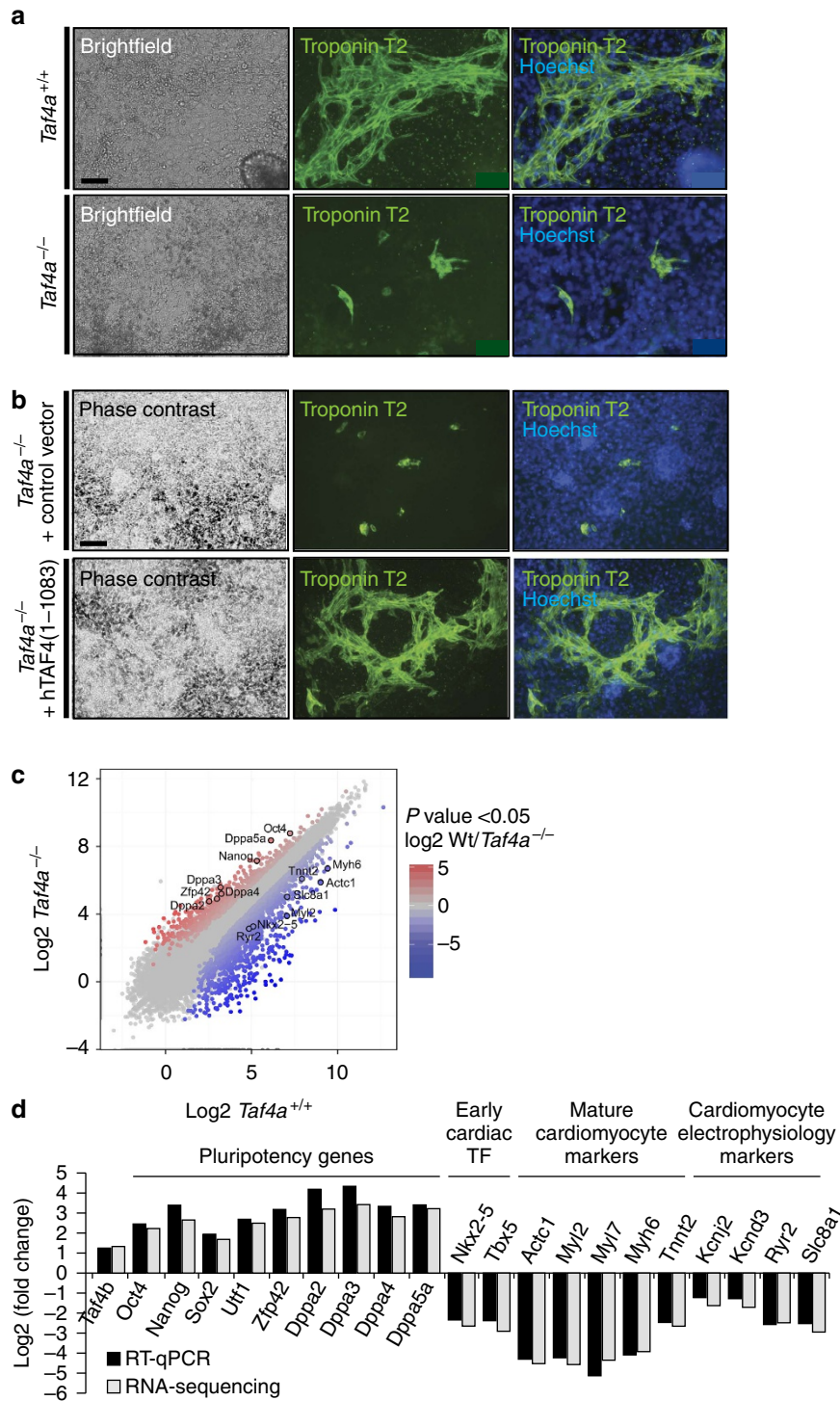
Figure 8 | Taf4 is required for cardiomyocyte differentiation. (**a**) Phase-contrast microscopy and Tnnt2 labelling of EBs differentiated into cardiomyocytes 9 days after plating. (**b**) Phase-contrast microscopy and Tnnt2 labelling of ES cells engineered to re-express exogenous Taf4 differentiated into cardiomyocytes 9 days after plating. All images were taken at ×20 magnification. (**c**) Global comparison of RNA-seq data in $Taf4a^{-/-}$ compared with WT at the EB8 stage. The expression of several cardiomyocyte markers and pluripotency genes are highlighted. (**d**) Comparison of the changes in expression of selected marker genes measured by RNA-seq and by RR–qPCR. Scale bars 50 μm.

ESCs *in vitro* when grown under pluripotent conditions. As $Taf4b^{-/-}$ mice are viable, genetics indicates that neither Taf4 nor Taf4b are essential for pluripotency, but rather are redundant for this function. Moreover the appearance of PGCs in the $Taf4a^{-/-}$ embryos indicates that Taf4b can also replace Taf4 in this reprogramming event.

Rather than playing a role in ESC pluripotency, our data show that Taf4 acts during differentiation. Similar to what is seen *in vivo*, $Taf4a^{-/-}$ EBs initiate expression of early markers of all three germ layers, but Taf4 is essential for the activation of later stage differentiation genes. Analysis of RA-driven neurogenic differentiation of $Taf4a^{-/-}$ ESCs confirmed that expression of

early-induced genes is affected less than later markers. The activation kinetics reflect somewhat different regulatory mechanisms as promoters of early-induced genes show appreciable PIC formation in the undifferentiated state that is increased during differentiation, whereas late-induced genes show little PIC formation in ESCs, their activation being almost completely dependent on *de novo* PIC formation. In mutant ESCs reduced PIC formation is seen at early genes, but is not increased during differentiation, and *de novo* PIC formation at later genes is almost completely abolished. Thus while Taf4b-containing TFIID can assemble the PIC in ESCs, Taf4 is critical for *de novo* PIC formation on later stage neurogenic genes.

The defective neuronal and cardiomyocyte differentiation of *Taf4a*$^{-/-}$ ESCs is reminiscent of the role of Taf3 in endodermal differentiation of ESCs[45]. Short hairpin RNA silencing showed that Taf3 is not required for ESC pluripotency, but plays a specific role in endoderm differentiation. Here we show that Taf4 is not required for ESC pluripotency, but plays a more general role in differentiation of all three germ layers. It is possible that the overall reduced amounts of TFIID in *Taf4a*$^{-/-}$ ESCs may impact their differentiation. Pijnappel *et al.*[27] (and Fig. 5a in this study) showed higher levels of Tbp and TAFs in ESCs compared with fibroblasts and suggested that high TFIID levels are required for pluripotency. Reduced TFIID would therefore be expected to affect pluripotency rather than differentiation. For this reason, we postulate that the defective differentiation of *Taf4a*$^{-/-}$ ESCs and the defects seen in Taf4-null embryos reflect loss of a specific function of Taf4 rather than an overall decrease in TFIID.

Bahat *et al.*[44] reported that Taf4 knockdown rendered ESCs refractory to the effects of RA, whereas we show here incomplete differentiation of RA-treated *Taf4a*$^{-/-}$ ESCs. However, Bahat *et al.*[44] treated ESCs directly with RA without prior growth as EBs, an essential feature of the neurogenic differentiation protocol. Nevertheless, an interesting feature of differentiating *Taf4a*$^{-/-}$ neurons and cardiomyocytes is the persistent expression of pluripotency genes. Bahat *et al.*[44] showed that Taf4b, but not Taf4, interacts with Pou5f1. The interactions of Pou5f1 with the Taf4b-containing TFIID in the *Taf4a*$^{-/-}$ ESCs may promote expression of the pluripotency network to the detriment of redistribution to differentiation genes thereby explaining the persistent expression of pluripotency genes in absence of Taf4.

Methods

Recombinant mice. C57Bl6J mice carrying *Taf4a* null alleles where exons 11 and 12 have been deleted were previously described[16]. *Taf4a*$^{-/-}$ embryos were derived from crosses of 6 to 7-week-old *Taf4a*$^{+/-}$ animals. Mice were kept in accordance with institutional guidelines regarding the care and use of laboratory animals and in accordance with National Animal Care Guidelines (European Commission directive 86/609/CEE; French decree no.87–848). All procedures were approved by the French national ethics committee.

Histology and immunohistochemistry. Deciduae were fixed in Bouin's fluid for 14 h before being dehydrated and embedded in paraffin. The embedded decidua were cut in 5-µm sections and stained with hematoxylin and eosin. Other sections were used for terminal deoxynucleotidyl nick-end labelling assay using the ApopTag Peroxidase *In Situ*.
Apoptosis Detection Kit (Millipore) as described by the manufacturer. To analyse proliferation, pregnant females were peritoneal injected with BrdU (50 mg kg^{-1} final) and sacrificed after 90 min. Paraffin sections on deciduae obtained from these litters were boiled in citrate buffer for antigen retrieval and staining was performed with primary anti-BrdU (555627, BD Biosciences). Pecam staining on whole-mount embryos and yolk sacs was performed as described[46].

Whole-mount *in situ* hybridization. Mouse embryos were dissected in ice-cold PBS and fixed O/N in 4% paraformaldehyde (PFA)/PBS. After several washes in PBS-0.1% Tween-20 (PBT) to eliminate traces of formaldehyde, embryos were bleached for 1 h in 3% H$_2$O$_2$/PBT and washed 3 × 5 min in PBT before being

digested with Proteinase K (10 µg ml^{-1}) for 30 s, 1 min 30 s or 2 min 30 s (E7.5, E8.5 and E9.5). Digestion was stopped by 5 min incubation in 2 mg ml^{-1} glycine/PBT. Embryos were washed again 3 × 5 min in PBT before post-fixing for 20 min in 0.2% glutaraldehyde/4% PFA/PBS. After further washes they were incubated in prewarmed hybridization buffer and prehybridized for 2 h at 65 °C. The buffer was then replaced with fresh prewarmed hybridization buffer containing the digoxigenin labelled RNA probes and incubated O/N at 65 °C. The next day, embryos were washed twice in the hybridization buffer at 65 °C before treating them with RNaseA to reduce background. The embryos were then blocked for 2 h in 2% blocking solution (Roche)/TBS-0.1% Tween-20 (TBST) and incubated O/N in the same solution containing 1:2,500 anti-digoxigenin antibody (Roche). The next day the embryos were washed in TBST, before washing them in NTMT buffer (2 × 10 min) and developing the signal with BM Purple (Roche). Most *in situ* hybridisations were performed on three independent WT and mutant embryos.

Generation, culture and differentiation of *Taf4a*$^{-/-}$ ES cells. *Taf4a*$^{+/-}$ mice were crossed and E3.5 blastocysts were flushed out the uterus and cultured on mitomycin treated MEF feeders. The ICM outgrowths were trypsinized and passaged on 24-well plates. ES cells were further cultured on inactivated MEF feeders under standard conditions (DMEM 4.5 g l^{-1} glucose/Glutamax supplemented with 15% heat inactivated FCS, LIF, nonessential amino acids and β-mercaptoethanol). *Taf4a*$^{-/-}$ ES cells were identified by PCR as previously described[16]. For spontaneous differentiation, ES cells were feeder depleted and allowed to form EBs in DMEM 4.5 g l^{-1} glucose/Glutamax supplemented with 10% heat inactivated FCS, nonessential amino acids and β-mercaptoethanol in non-adherent bacterial plates. EBs were cultured for a total of 14 days with medium change every other day. Cells were differentiated by RA into neurons as described[33,34]. For differentiation into cardiomyocytes, 4 × 10^6 cells were grown as EBs in EB medium containing 20% heat inactivated FCS and cultured for 5 days. EBs were then transferred to gelatin-coated plates and cultured for 9 more days. For rescue experiments, a linearized control plasmid or a plasmid containing recombinant human or mouse TAF4 complementary DNA (cDNA) was electroporated into *Taf4a*$^{-/-}$ ES cells and selected with puromycin. Resistant clones were further amplified and differentiated as described above.

Immunostainings and alkaline phosphatase staining. For immunostaining, cells were fixed for 30 min in 4% paraformaldehyde at room temperature, followed by incubation in blocking solution (PBS, 0.1% Triton X-100, 5% normal goat serum) for at least 2 h. Primary antibodies were incubated O/N in blocking solution and then washed 3 times with PBS, 0.1% Triton X-100 for 10 min per step to reduce unspecific binding before incubation with secondary antibodies. The Alkaline Phosphatase Detection Kit (SCR004; Millipore) was used following the manufacturers instructions.

Cell growth and cell cycle analysis. To determine growth rates, 5 × 10^4 cells from either WT or mutant ES cells, were plated gelatin-coated six-well plates in triplicate and counted for 4 consecutive days. For cell cycle analysis, cells were fixed in 70% ice-cold ethanol O/N, washed with PBS and then stained with propidium iodide for 4 h. Fluorescence was detected using the BD FACSCalibur flow cytometer and the obtained data were analysed with the CellQuest software.

Gene expression analyses. Total RNA from ES cells or EBs was isolated using the GenElute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich). cDNA was generated using Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Scientific), following the manufacture's instructions. PCR used the SYBR Green PCR Master Mix from Qiagen and the LightCycler 480 Real-Time PCR-System (Roche). Relative expression levels were normalized to GAPDH. RNA-seq analyses were performed essentially as described[47,48]. All RT-qPCR experiments were performed in triplicate. Messenger RNA-seq was performed essentially as previously described[48]. Sequence reads mapped to reference genome mm9/NCBI37 using Tophat[49]. Data normalization and quantification of gene expression was performed using the DESeq 2 Bioconductor package[50]. Two biological replicates from two independent WT and *Taf4a*$^{-/-}$ ESC lines were analysed. Gene ontology analyses were performed using the functional annotation clustering function of DAVID (http://david.abcc.ncifcrf.gov/).

Preparation of ES cell nuclear extracts, IP and immunoblot. ES cells were resuspended in swelling buffer (25 mM Tris pH 7,9, 10 mM KCl, 1.5 mM MgCl$_2$, 1 mM DTT), incubated for 10 min on ice, lysed using a dounce homogenizer and centrifuged. Pelleted nuclei were resuspended in extraction buffer (50 mM Tris pH 7,5, 400 mM NaCl, 0.5% NP-40, 5% Glycerol and 1 × Roche complete protease inhibitor) and proteins were extracted for 2 h at 4 °C with rotation. For IP, 500 µg nuclear extract were pre-cleared for 2 h with 50 µl of protein G sepharose beads, before addition of with 1 µg of anti-TBP (3G3) antibody O/N at 4 °C. The next day, bound complexes were recovered by incubating the extract with 25 µl protein G sepharose beads for 2 h at 4 °C. Beads containing bound proteins were washed three times with extraction buffer for 15 min per wash and once with the buffer containing 150 mM NaCl before being boiled in in SDS loading buffer for 5 min at

95 °C. Bound proteins were analysed by Western blot using the indicated anti-bodies. A control IP for each extract was performed in parallel using a non-specific antibody against glutathione-S transferase (GST). Uncropped western blots are provided for all immunoblots in Supplementary Figure 11.

Gel filtration. 500 µl of nuclear extracts prepared as described above containing 6 mg of protein were injected onto a equilibrated Superose 6 (10/300) column and run at 0.4 ml min^{-1}. 500 µl fractions were collected and analysed by western blot.

Antibodies. The following antibodies were used: anti-OCT4 (sc-9081; Santa Cruz), anti-NANOG (RCAB001P; Reprocell), SOX2 (L1D6A2; Cell Signaling), previously described or recently developed in house antibodies against TBP (3G3), TAF1, TAF3 (ref. 15), TAF4 (32TA and also sc-136093; Santa Cruz), TAF4B (ref. 16), TAF5 (1TA), TAF6 (25TA), TAF7 (19TA), TAF12 (22TA), and TAF13 (16TA), TAF10 (sc-102125; Santa Cruz) and in house 6TA2B11 (ref. 51), anti-Pol II (sc-9001x; Santa Cruz), anti-Tnnt2 (MS-295-P1; Thermo scientific). We also used a polyclonal TAF4B antibody (2057) raised against a peptide corresponding to amino acids 692–710 of human TAF4B that is 100% conserved in mouse Taf4b. This antibody also recognizes Taf4 as this region comprises only two amino acid differences with Taf4b. The following antibody dilutions were used for immunostaining. Pou5f1 and Nanog, 1:100; Sox2, and Taf4, 1:50; Tubb3 and Tnnt2, 1:1,000. For immunoblots all antibodies were diluted 1:1,000 except polyclonal serum 2057 that was diluted 1:2,000.

Chromatin IP sequencing. Chromatin IP experiments were performed on 0,4% PFA-fixed chromatin according to standard protocols as previously described[15]. Chromatin IP sequencing libraries were prepared and sequenced on an Illumina Hi-seq2500 as single-end 50-base reads. After sequencing, peak detection was performed using the MACS software (ref. 52 http://liulab.dfci. harvard.edu/MACS/). Global clustering, meta analyses and quantitative comparisons were performed using seqMINER[53] and R (http://www.r- project.org/).

References

1. Wright, K. J., Marr, 2nd M. T. & Tjian, R. TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc. Natl Acad. Sci. USA* **103,** 12347–12352 (2006).
2. Cler, E., Papai, G., Schultz, P. & Davidson, I. Recent advances in understanding the structure and function of general transcription factor TFIID. *Cell Mol. Life Sci.* **66,** 2123–2134 (2009).
3. Bieniossek, C. et al. The architecture of human general transcription factor TFIID core complex. *Nature* **493,** 699–702 (2013).
4. Zhou, H. et al. Dual functions of TAF7L in adipocyte differentiation. *eLife* **2,** e00170 (2013).
5. Cheng, Y. et al. Abnormal sperm in mice lacking the taf7l gene. *Mol. Cell. Biol.* **27,** 2582–2589 (2007).
6. Zhou, H., Wan, B., Grubisic, I., Kaplan, T. & Tjian, R. TAF7L modulates brown adipose tissue formation. *Elife* **3,** e02811 (2014).
7. Falender, A. E. et al. Maintenance of spermatogenesis requires TAF4b, a gonad-specific subunit of TFIID. *Genes Dev.* **19,** 794–803 (2005).
8. Voronina, E. et al. Ovarian granulosa cell survival and proliferation requires the gonad-selective TFIID subunit TAF4b. *Dev Biol.* **303,** 715–726 (2007).
9. Herrera, F. J., Yamaguchi, T., Roelink, H. & Tjian, R. Core promoter factor TAF9B regulates neuronal gene expression. *Elife* **3,** e02559 (2014).
10. Martianov, I. et al. Late arrest of spermiogenesis and germ cell apoptosis in mice lacking the TBP-like TLF/TRF2 gene. *Mol. Cell.* **7,** 509–515 (2001).
11. Gazdag, E. et al. TBP2 is essential for germ cell development by regulating transcription and chromatin condensation in the oocyte. *Genes Dev.* **23,** 2210–2223 (2009).
12. Gangloff, Y., Romier, C., Thuault, S., Werten, S. & Davidson, I. The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem. Sci.* **26,** 250–257 (2001).
13. Werten, S. et al. Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20). *J. Biol. Chem.* **277,** 45502–45509 (2002).
14. Alpern, D. et al. TAF4, a subunit of transcription factor II D, directs promoter occupancy of nuclear receptor HNF4A during post-natal hepatocyte differentiation. *Elife* **3,** e03613 (2014).
15. Thuault, S. et al. Functional analysis of the TFIID-specific yeast TAF4 (yTAF(II)48) reveals an unexpected organization of its histone-fold domain. *J. Biol. Chem.* **277,** 45510–45517 (2002).
16. Mengus, G. et al. TAF4 inactivation in embryonic fibroblasts activates TGFbeta signalling and autocrine growth. *EMBO J.* **24,** 2753–2767 (2005).
17. Fadloun, A. et al. The TFIID subunit TAF4 regulates keratinocyte proliferation and has cell-autonomous and non-cell-autonomous tumour suppressor activity in mouse epidermis. *Development* **134,** 2947–2958 (2007).
18. Fadloun, A. et al. Retinoic acid induces TGFbeta-dependent autocrine fibroblast growth. *Oncogene* **27,** 477–489 (2008).
19. Martianov, I. et al. TAF4 inactivation reveals the 3 dimensional growth promoting activities of collagen 6A3. *PLoS ONE* **9,** e87365 (2014).
20. Wilt, F. H. Erythropoiesis in the chick embryo: the role of endoderm. *Science* **147,** 1588–1590 (1965).
21. Miura, Y. & Wilt, F. H. Tissue interaction and the formation of the first erythroblasts of the chick embryo. *Dev. Biol.* **19,** 201–211 (1969).
22. Palis, J., McGrath, K. E. & Kingsley, P. D. Initiation of hematopoiesis and vasculogenesis in murine yolk sac explants. *Blood* **86,** 156–163 (1995).
23. Kaufman. *The Atlas of Mouse Development* (Academic Press, 1998).
24. Perea-Gomez, A., Rhinn, M. & Ang, S. L. Role of the anterior visceral endoderm in restricting posterior signals in the mouse embryo. *Int. J. Dev. Biol.* **45,** 311–320 (2001).
25. Rhinn, M. et al. Sequential roles for Otx2 in visceral endoderm and neuroectoderm for forebrain and midbrain induction and specification. *Development* **125,** 845–856 (1998).
26. Ang, S. L. et al. A targeted mouse Otx2 mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122,** 243–252 (1996).
27. Pijnappel, W. W. et al. A central role for TFIID in the pluripotent transcription circuitry. *Nature* **495,** 516–519 (2013).
28. Richardson, B. E. & Lehmann, R. Mechanisms guiding primordial germ cell migration: strategies from different organisms. *Nat. Rev. Mol. Cell Biol.* **11,** 37–49 (2010).
29. Tam, P. P. & Snow, M. H. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J. Embryol. Exp. Morphol.* **64,** 133–147 (1981).
30. Ginsburg, M., Snow, M. H. & McLaren, A. Primordial germ cells in the mouse embryo during gastrulation. *Development* **110,** 521–528 (1990).
31. Maston, G. A. et al. Non-canonical TAF complexes regulate active promoters in human embryonic stem cells. *Elife* **1,** e00068 (2012).
32. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol.* **87,** 27–45 (1985).
33. Bibel, M. et al. Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nat. Neurosci.* **7,** 1003–1009 (2004).
34. Bibel, M., Richter, J., Lacroix, E. & Barde, Y. A. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat. Protoc.* **2,** 1034–1043 (2007).
35. Martianov, I., Viville, S. & Davidson, I. RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science* **298,** 1036–1039 (2002).
36. Mohan, I. W., Scheer, E., Wendling, O., Metzger, D. & Tora, L. TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol. Cell. Biol.* **23,** 4307–4318 (2003).
37. Gegonne, A. et al. The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells. *Mol. Cell. Biol.* **32,** 1984–1997 (2012).
38. Voss, A. K. et al. Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. *Development* **127,** 5449–5461 (2000).
39. Falender, A. E., Shimada, M., Lo, Y. K. & Richards, J. S. TAF4b, a TBP associated factor, is required for oocyte development and function. *Dev. Biol.* **288,** 405–419 (2005).
40. Madabhushi, M. & Lacy, E. Anterior visceral endoderm directs ventral morphogenesis and placement of head and heart via BMP2 expression. *Dev. Cell* **21,** 907–919 (2011).
41. Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28,** 812–828 (2014).
42. Saitou, M. & Yamaji, M. Primordial germ cells in mice. *Cold Spring Harb. Perspect. Biol.* **4** (2012).
43. Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139,** 15–31 (2012).
44. Bahat, A. et al. TAF4b and TAF4 differentially regulate mouse embryonic stem cells maintenance and proliferation. *Genes Cells* **18,** 225–237 (2013).
45. Liu, Z., Scannell, D. R., Eisen, M. B. & Tjian, R. Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* **146,** 720–731 (2011).
46. Schlaeger, T. M., Qin, Y., Fujiwara, Y., Magram, J. & Sato, T. N. Vascular endothelial cell lineage-specific promoter in transgenic mice. *Development* **121,** 1089–1098 (1995).
47. Laurette, P. et al. Transcription factor MITF and remodeller BRG1 define chromatin organisation at regulatory elements in melanoma cells. *eLife* **4,** doi: 10.7554/eLife.06857 (2015).
48. Herquel, B. et al. Trim24-repressed VL30 retrotransposons regulate gene expression by producing noncoding RNA. *Nat Struct. Mol. Biol.* **20,** 339–346 (2013).

49. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111 (2009).

50. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11,** R106 (2010).

51. Indra, A. K. *et al.* TAF10 is required for the establishment of skin barrier function in foetal, but not in adult mouse epidermis. *Dev. Biol.* **285,** 28–37 (2005).

52. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).

53. Ye, T. *et al.* seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* **39,** e35 (2011).

Acknowledgements

Author contributions

D.L. performed all embryo analyses in collaboration with M.R and P.D. D.L also performed analyses and differentiation of ES cells and analysed the RNA-seq with C.K., D.A. and I.M. performed and analysed the ChIP-seq data. G.M. generated and managed the recombinant mice, isolated the ES cell lines and performed embryo dissections I.D., D.L., M.R., P.D. and G.M. conceived experiments, analysed data and wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.