Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

XIST restraint by AFF3 enables human neural and musculoskeletal development

Ryohichi Sugimura

rios@hku.hk

Centre for Translational Stem Cell Biology/ The University of Hong Kong. https://orcid.org/0000-0001-5701-3628

Yang Xiang

The University of Hong Kong

Guanchen Wu

The University of Hong Kong https://orcid.org/0009-0008-9868-3677

Hongji Li

The University of Hong Kong

Alex To

The University of Hong Kong https://orcid.org/0000-0001-7872-228X

Qiyan Wang

The University of Hong Kong

Amanda NM Looi

The University of Hong Kong

Yan Xue

The University of Hong Kong

Chloe Kan

The University of Hong Kong

Yiming Chao

The University of Hong Kong

Xu Jiang

Southeast University

Ke Fang

Southeast University

Xiaohong Cen

The University of Hong Kong

Yiyi Xuan

InnoHK Centre for Translational Stem Cell Biology

Degong Ruan

The University of Hong Kong

Chao Liang
The University of Hong Kong
Ziqi Huang
The University of Hong Kong
Siyuan Chen
The University of Hong Kong
Feiming Li
The University of Hong Kong
Yamei Jiang
Chinese University of Hong Kong – Shenzhen
Sang Jin Lee
The University of Hong Kong https://orcid.org/0000-0002-8007-8785
Pentao Liu
University of Hong Kong https://orcid.org/0000-0001-5774-9678
Jin Liu
The Chinese University of Hong Kong, Shenzhen https://orcid.org/0000-0002-5707-2078
Zhuojuan Luo
Southeast University https://orcid.org/0000-0002-0715-0648
Chengqi Lin
Southeast University https://orcid.org/0000-0002-2121-584X

Article

Keywords: AFF3, XIST, stem cell-based embryo models, neural development, musculoskeletal development

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11233698/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

XIST restraint by AFF3 enables human neural and musculoskeletal development

Yang Xiang [1], Guanchen Wu [1], Hongji Troy Li [1], Alex To [1], Qiyan Wang [1], Amanda NM Looi [1], Yan Xue [1], Chloe Kan [1], Yiming Chao [1], Xu Jiang [3], Ke Fang [3], Xiaohong Cen [1], Yiyi Xuan [4], Degong Ruan [4, 5], Chao Liang [2], Ziqi Huang [2], Siyuan Chen [2], Feiming Li [2], Yamei Jiang [6], Sang Jin Lee [2], Pengtao Liu [1, 4], Jin Liu [6], Zhuojuan Luo [3#], Chengqi Lin [3#], Ryohichi Sugimura [1, 4#]

1 School of Biomedical Sciences, The University of Hong Kong, Hong Kong, China;

2 Faculty of Dentistry, The University of Hong Kong, Hong Kong, China;

3 Southeast University, China;

4 InnoHK Centre for Translational Stem Cell Biology, Hong Kong, China;

5 The University of Hong Kong-Shenzhen Hospital Translational Medicine Research Centre, China;

6 Chinese University of Hong Kong – Shenzhen, China.

Corresponding zjluo@seu.edu.cn; cqlin@seu.edu.cn; rios@hku.hk

Highlights

- AFF3 loss impairs neural and musculoskeletal development in female gastruloids;
- AFF3 restrains XIST expression during gastrulation;
- Premature and excessive upregulation of XIST represses autosomal neural and musculoskeletal development programs;
- Genetic and pharmacological attenuation of XIST restores both neural and musculoskeletal development in gastruloids and neurogenesis in vivo.

Keywords

AFF3, XIST, stem cell-based embryo models, neural development, musculoskeletal development

Abstract

The underlying developmental mechanisms of defects in human organogenesis remain poorly defined. Pathogenic variants in AFF3, the super elongation complex subunit, cause a congenital disorder of neural and limb development, yet their developmental basis is unclear. Stem cell-based embryo model gastruloids show that AFF3, enriched in neural progenitors and neuromesodermal progenitors, is required for female neural and musculoskeletal development. AFF3 loss depletes neural and musculoskeletal lineages. Mechanistically, AFF3 restrains premature and excessive activation of the X-inactivation mediator XIST. Abnormal XIST upregulation represses autosomal genes required for neural and musculoskeletal development. Genetic deletion or pharmacological attenuation of XIST restores gastruloid morphogenesis, rescues neural and musculoskeletal development via autosomal genes, and improves cerebral organoid formation and neurogenesis after transplantation. These findings identify AFF3-mediated XIST restraint as an essential safeguard for early human organogenesis, and its failure can be traced back to gastrulation.

Introduction

Congenital disorders encompass a wide spectrum of structural and functional abnormalities, including those of the central nervous and musculoskeletal systems. Neural and musculoskeletal lineage commitment is a particularly vulnerable developmental window (*1*–*3*), yet its underlying mechanisms in humans remain poorly understood. Progress has been limited by restricted access to early human embryonic material, motivating the use of stem cell-based embryo models. Among these models, gastruloids recapitulate key features of human gastrulation and the formation of the three germ layers, providing a powerful platform for dissecting defects in both neural (ectodermal) and musculoskeletal (mesodermal) lineages (*4*).

Inherited diseases offer insight into normal development, including that of early neural and musculoskeletal lineages. The autosomal gene AFF3 encodes a scaffold subunit of the super elongation complex (SEC), a regulator of RNA polymerase II elongation that controls imprinting and cell-fate specification (*5*). Other SEC components have established roles in neurodevelopment. For example, AFF4 acts through interactions with MECP2, mutations in which cause Rett syndrome (*6*). Pathogenic variants in AFF3 show neural and musculoskeletal defects in patients. Dominant-negative variants in a degron motif prevent proteolysis of AFF3, causing KINSSHIP syndrome, with severe intellectual disability, epilepsy, shortening of limbs, and malformation of kidneys. Loss-of-function variants, by contrast, cause milder intellectual disability and shortening of limbs (*7*–*9*). This divergence indicates that dominant-negative and loss-of-function variants act through distinct mechanisms. AFF3 loss-of-function mouse models show variable phenotypes between strains, from no phenotype to moderate

neuroanatomical defects (*9*). This highlights the limits of murine models of AFF3 and the importance of studying its pathogenic variants in a human context.

AFF3 has been reported to repress XIST in differentiated cells (*10*). XIST mediates X chromosome inactivation (XCI) and equalizes X-linked gene dosage between females and males (*11*, *12*). This equalization in the embryo is established dynamically: XIST is biallelically expressed in pre-implantation human embryos (*13*), and its expression declines during gastrulation (*14*). A recent study showed that XIST can suppress autosomal developmental genes in pluripotent stem cells (*15*), suggesting that excessive XIST activation could disrupt development. Consistent with this, pentasomy X (XXXXX) patients express XIST in excess. Their presentation overlaps with that of AFF3 patients, including congenital intellectual disability and limb malformation (*11*, *16*). These findings suggest that XIST dosage could influence neural and musculoskeletal development in embryos. Yet whether AFF3 regulates XIST during human development, and whether failure of this regulation causes neural and musculoskeletal defects, have not been addressed.

Here, using stem cell-based gastruloids, we find that AFF3 restrains XIST to enable neural and musculoskeletal lineage commitment. Mechanistically, AFF3 deletion permits premature and excessive XIST upregulation, which represses autosomal gene expression required for neural and musculoskeletal development. Pharmacogenetic attenuation of XIST restores both lineages in gastruloids. It also rescues cerebral organoid formation, as well as neurogenesis after transplantation in vivo. Together, these findings indicate that AFF3-mediated XIST restraint is required for human neural and musculoskeletal development.

Results

AFF3 is enriched in NPC and NMP in human embryos

To examine the role of AFF3 in early embryogenesis, we generated gastruloids from human expanded potential stem cells (*17*). These gastruloids recapitulated key features of post-implantation development, including the formation of amnion- and yolk sac (YS)-like structures. SOX17+ YS endoderm was formed on the ventral side, while a T+ primitive streak-like domain emerged on the posterior side. PODXL+ amnion surrounded the outer layer of the neuroectoderm, whereas FOXA2+ T+ notochord-like cells (*18*) were located in the middle of the gastruloids. Notably, SOX2+ neuroectoderm formed a characteristic neural progenitor rosette-like structure and robustly expressed AFF3. The NKX6-1+ neuroectodermal cells, which later differentiate into the ventral hindbrain, expressed AFF3 and occupied the ventral region of the gastruloids (Fig. 1A).

We then resolved AFF3 expression across cell populations in gastruloids by single-cell RNA-seq (scRNA-seq) (Fig. 1B). We identified neural progenitor cells (NPC) and neuromesodermal

progenitors (NMP), defined by SOX2 and TBXT, which also expressed NKX1-2, NFEM, and CDX2. NMPs contribute to trunk neural crest, spinal cord, and paraxial mesoderm (PXM) lineages, contributing to musculoskeletal tissues (*19*). We detected PXM markers TBX6, MESP1, and MESP2 (*20*, *21*) around the NMP area (Fig. 1C). AFF3 was enriched in both NPC and NMP as well as PXM (Fig. 1B-D).

We next confirmed that AFF3 enrichment in gastruloids reflected its expression in human development. NPC, NMP, and PXM populations were identified in gastrulation-stage human embryos at gestational weeks (GW) 3 and 4, and aligned with those in gastruloids (*14*) (Fig. 1E). AFF3 was likewise enriched in both NPC and NMP as well as PXM populations, consistently between human embryos and gastruloids (Fig. 1F). Moreover, the expression pattern of NPC, NMP, and PXM markers was highly concordant between gastruloids and gastrulating human embryos at GW 3, indicating that gastruloids faithfully capture early neuroectoderm patterning (Fig. 1G; Supp. Fig. 1B-D). AFF3 expression was also detected in the neuroectoderm and mesoderm of mouse gastrulation-stage embryos at E7.5 (*14*) (Supp. Fig. 1A), suggesting evolutionary conservation between species. Together, these data support gastruloids as a model for studying AFF3 in early human development.

AFF3 loss impairs neural and musculoskeletal development in female gastruloids

To investigate the role of AFF3 during gastruloid development, we generated AFF3-/- (KO) gastruloids and compared their phenotypic and transcriptional profiles with those of wild-type controls (Supp. Fig. 2A, 2B, 2E, 2F). Loss of AFF3 led to clear morphological abnormalities in female gastruloids (Fig. 2A-B). TemporalVAE transcriptomic staging (*22*) mapped female AFF3 KO gastruloids to Carnegie Stage (CS) 6, compared with CS7/8 for wild-type, suggesting delayed developmental progression (Supp. Fig. 3A-C). Consistent with these differences, RNA-seq analysis identified extensive gene expression changes in female AFF3 KO gastruloids (Fig. 2C). Heatmap analysis further revealed that AFF3 KO gastruloids reduced neural and musculoskeletal genes for both myogenic (PAX3, PAX7, RBM24, ACTN3, TNNT1, DMD) and skeletal (GAS1, BOC, CDON) development (*23*–*25*) (Fig. 2D). We further validated the reduction of neural markers PAX6 and SOX2 with immunofluorescence (Fig. 2E). Among cell types inferred from RNA-seq, AFF3 KO uniquely reduced neural and musculoskeletal populations (Fig. 2F). In parallel, AFF3 KO gastruloids showed increased expression of non-neural and non-musculoskeletal markers associated with all three germ layers, as well as hematopoietic and cardiac lineages (Supp. Fig. 8A, 8B). Together, these findings indicate that AFF3 is required for neural and musculoskeletal development in females.

To further define the neural and musculoskeletal defects, we performed scRNA-seq in combination with LARRY molecular barcoding (Fig. 2G). Because scRNA-seq captures the LARRY barcode, individual stem cells can be traced and their lineage trajectories reconstructed at single-cell resolution during gastruloid development (*26*). Consistent with

the morphological and bulk RNA-seq phenotypes, single-cell analysis revealed a significant reduction in NPC and neuronal populations in female AFF3 KO gastruloids compared with wild-type controls, alongside enrichment of cardiac mesoderm (CM), primordial germ cells (PGC), vascular endothelial cells (VEC), blood progenitors (BP) and yolk sac endoderm (YSE) (Fig. 2H). This suggests biased differentiation of NMP to lateral plate mesoderm (VEC, BP, CM) rather than their normal path to neural and musculoskeletal lineages. Lineage tracing with LARRY showed that AFF3 KO epiblast cells were preferentially diverted toward BP and CM, as well as toward extraembryonic amnion and trophectoderm (TE). By contrast, wild-type epiblast cells are predominantly committed to NPC and neural fates (Fig. 2I). Together, these findings show that loss of AFF3 impairs neural and musculoskeletal development and redirects female gastruloid cells toward alternative developmental trajectories.

Premature and excessive upregulation of XIST upon AFF3 loss

To define the mechanism underlying the neural and musculoskeletal development defects caused by AFF3 loss, we examined candidate AFF3-target genes. Because AFF3 has been implicated in restraining XIST expression (*27*), we first asked whether dysregulated XIST contributed to the female-specific phenotype observed in AFF3 KO gastruloids. XIST expression is developmentally dynamic: preimplantation human embryos express XIST from both X chromosomes (*13*). Expression is then resolved to a single allele and remains at a low level before rising when X-inactivation is established (*28*). Reflecting this temporal regulation, XIST was expressed at lower levels in the GW3 human embryos than at GW4 (Fig. 3A-B). Likewise, human expanded pluripotent stem cells (hEPSCs) undergo partial X chromosome reactivation (XCR) with loss of XIST expression (*29*) and XIST was correspondingly low in our hEPSCs (Fig. 3C). AFF3 deletion disrupted this low expression pattern in a sex-specific manner: XIST was upregulated in female hEPSCs and gastruloids across two independent lines (Supp. Fig. 4A), whereas male hEPSCs and gastruloids showed no comparable increase (Fig. 3C). This indicates that AFF3 restrains XIST expression in female cells, but male cells may depend on different mechanisms such as XIST promoter DNA methylation (*30*). In line with elevated XIST, multiple X-linked genes were suppressed in AFF3 KO female gastruloids, whereas male gastruloids again showed no significant change (Supp. Fig. 4B). These data show that XIST is restrained during early development and that AFF3 maintains this repression in female cells. scRNA-seq revealed aberrant XIST upregulation in AFF3 KO female gastruloids, but minimal expression in WT gastruloids (Fig. 3D). RNA FISH confirmed robust XIST upregulation in day 8 AFF3 KO female gastruloids, whereas signals were largely absent from WT controls (Fig. 3E-F). It also showed that XIST formed a broad nuclear cloud, suggestive of its dispersal beyond the X chromosome. This phenomenon has been described in the context of XIST-mediated X chromosome dampening, where XIST represses autosomal targets via interchromosomal interactions (*15*). We further examined XACT expression patterns in WT and AFF3 KO female gastruloids. XACT, a long non-coding RNA that coats the active X (*31*), was present either

biallelically or monoallelically in WT, the former indicating XaXa status in early embryonic development (Fig. 3E). All KO cells expressed XACT monoallelically, indicating XaXi status. XIST was present on the XACT+ chromosome, suggesting abnormal XIST upregulation (Fig. 3E-F). Upon AFF3 loss, XIST induction may begin from the XACT+ active and open X, then spread to the other X chromosome. To test whether AFF3 directly regulated the XIST locus, we combined immunofluorescence for AFF3 with DNA FISH for the X chromosome, which showed colocalization with the X chromosome in individual cells (Fig. 3G). AFF3 binds to both X chromosomes and represses XIST expression in early embryonic development; AFF3 loss leads to abnormal XIST expression, as shown by monoallelic expression in the XACT+ active X chromosome, as well as biallelic expression (Fig. 3H). To prevent inactivation of both X chromosomes, one copy of the X chromosome may be protected by an XIST antagonist such as TSIX (*32*), which we examined below. These findings indicate that AFF3 prevents premature and excessive XIST activation during gastrulation in female embryos.

Premature and excessive XIST represses autosomal neural and musculoskeletal development

Next, we asked how AFF3 represses XIST expression. We performed CUT&Tag using an AFF3 antibody. We identified AFF3-binding peaks around the XIST locus, including one site near the XIST transcription start site and another in a regulatory element, L1PA, which belongs to the LINE1 family (*33*), indicating direct occupancy by AFF3 (Fig. 4A). H3K9me3 was present in L1PA and the XIST gene, consistent with its high-level occupancy on the female X chromosome (*14*). L1PA variant is associated with Alzheimer's disease in females, suggesting a potential role of XIST regulation in neurodegenerative disorders in late life (Supp. Fig. 9A-D). Upon AFF3 loss, RNA Pol II occupancy at L1PA increased, suggesting that AFF3 restrains XIST expression (Fig. 4A-B).

We next asked how abnormal XIST activation impairs neural and musculoskeletal development. To directly test the functional contribution of XIST, we generated AFF3/XIST double-knockout (DKO) gastruloids (Supp. Fig. 2C, 2D) and AFF3 KO gastruloids treated with the XIST inhibitor X1 (*34*) and examined both at day 8 (Fig. 4C-D). Strikingly, XIST genetic deletion or pharmacological attenuation rescued the morphological defects caused by AFF3 loss (Fig. 4D-E). Among genes downregulated in AFF3 KO gastruloids, we identified the neural crest & musculoskeletal regulator EPHB1 (*35*)(*36*) and the neuronal guidance factor NAV1 (*37*) (Fig. 4F-G). Notably, both genes are autosomal targets of XIST (*15*). RNA-seq further showed that nine autosomal XIST-target genes (*15*) were restored in both DKO and X1-treated KO gastruloids relative to KO, including skeletal development genes CDON, BOC, and GAS1 (*38*) (Fig. 4H, Supp. Fig. 8D). These rescued genes were significantly enriched for neural and musculoskeletal development (Fig. 4I). In contrast, GPM6B, ZNF185, MID1, and MECP2, the X-linked neural genes (*39*–*43*) examined, showed no comparable restoration (Supp. Fig. 5A).

These findings support that abnormal XIST upregulation in AFF3 KO female gastruloids represses autosomal programs for neural and musculoskeletal development (Fig. 4J).

Dominant-negative AFF3 results in dysregulated neurogenesis

AFF3 loss-of-function variants cause mild to moderate developmental delay and shortened limbs. In contrast, AFF3 dominant-negative variants cause severe developmental delay, intellectual disability, seizures, epilepsy, shortened limbs, and kidney malformation (*9*). The phenotypic spectra partially overlap but are largely distinct; particularly severe developmental delay, intellectual disability, seizures, and epilepsy are specific to dominant-negative patients (Fig. 5A-B). This suggests potentially different mechanisms underlying loss-of-function and dominant-negative effects. To test this, we generated AFF3 dominant-negative gastruloids. As expected, XIST inhibitor X1 did not alter the transcriptional signature (Fig. 5C). XIST expression did not change in dominant-negative female gastruloids (Fig. 5D). Instead of the neural loss seen in AFF3 KO gastruloids, dominant-negative gastruloids showed increased expression of excitatory and reduced expression of inhibitory neural genes (Fig. 5E-F). This is consistent with the seizure and epilepsy symptoms seen in dominant-negative patients (*7*). Thus, XIST inhibition rescues loss-of-function but not dominant-negative phenotypes, consistent with a variant-specific mechanism.

XIST targeting rescues cerebral organoid formation

To test whether the findings from AFF3 loss-of-function gastruloids extend to a model of cortical development, we generated cerebral organoids from WT, AFF3 KO, and AFF3/XIST DKO hEPSCs, allowing us to assess both the AFF3 loss phenotype and its rescue by XIST deletion. AFF3 KO organoids were markedly smaller, with a cystic, sac-like morphology, whereas WT organoids formed dense, cortex-like structures (Fig. 6A-B). As in gastruloids, XIST deletion rescued these defects, with DKO organoids resembling WT and showing no obvious developmental phenotype. Histological analysis further revealed cortical organization in WT and DKO organoids, but not in KO organoids, which instead showed a mesenchymal-like pattern (Fig. 6B), echoing the non-neural diversion seen in gastruloids. Consistent with this loss of neural identity, the neural progenitor marker PAX6 (*44*) was detected within the cortex-like regions of WT and DKO organoids but was largely absent from KO organoids (Fig. 6B).

We next performed single-cell transcriptomic profiling of the organoids to define the cellular basis of these phenotypes (Fig. 6C). Consistent with the morphological data, AFF3 KO organoids showed a significant reduction in neuronal lineage populations, whereas deletion of XIST restored these populations toward WT levels (Fig. 6C). In contrast, AFF3 KO organoids showed a significant increase in non-neuronal populations, including blood progenitor (BP),

endothelial cell (EC), kidney (Kid), meso-endoderm (ME), smooth muscle cell (SMC), and yolk sac endoderm (YSE) (Supp. Fig. 7B). SFRP2, a gene required for neural development (*45*), and an autosomal target of XIST, was re-expressed upon XIST deletion (Fig. 6D). To identify candidate genes underlying this rescue, we intersected the 1,599 genes restored in DKO organoids with previously defined XIST-target autosomal genes (*15*) . This analysis identified 44 overlapping genes (Fig. 6E), recapitulating the autosomal XIST-target rescue seen in gastruloids and suggesting that a subset of these targets is functionally relevant to the neural defect. Functional annotation of the overlapping genes revealed significant enrichment for cell fate specification, neuronal development, and synaptic membrane terms (Fig. 6F).

Because XIST mediates X-linked silencing, we asked whether the neural defect arises from altered X-linked gene dosage rather than autosomal repression. However, X-linked neurogenesis genes, including ZNF185, MID1, MECP2, GPM6B, and CA5B (*46*, *47*), remained comparably expressed across WT, KO, and DKO organoids despite XIST upregulation (Supp. Fig. 5). This is consistent with only one X chromosome being inactivated upon AFF3 loss (Fig. 3E). The concurrent upregulation of XIST antisense transcript, TSIX, may counteract inactivation of the second X (*32*) (Supp. Fig. 6), as anticipated above. The AFF3-regulated DLK1-DIO3 imprinted locus (*48*), including the neurogenesis-associated gene DLK1 (*49*, *50*), was likewise unchanged (Supp. Fig. 7A). Together, these results argue against canonical X-linked dosage or imprinting as the driver. We demonstrate that premature XIST activation disrupts neural development through repression of autosomal programs, which can be reversed by XIST removal.

XIST targeting rescues neurogenesis in vivo

We next examined whether XIST inhibition could restore the neural program in vivo. We transplanted WT, WT+X1, KO, and KO+X1 organoids onto the cortex of NSG mice (Fig. 7A). Two weeks after transplantation, organoids had integrated with the cortex. We then harvested the organoids with adjacent murine brain tissue and assessed morphology and neural lineage marker expression by H&E and immunofluorescence staining (Fig. 7B-D). Consistent with previous data (Fig. 6), AFF3 KO organoids were markedly smaller prior to transplantation (Fig. 7B). After transplantation, KO organoids exhibited morphological abnormalities and ossification indicating mesenchymal lineage shift, while X1-treated KO organoids were morphologically similar to WT organoids (Fig. 7C). X1 treatment significantly increased SOX2 (neural progenitor cells), EOMES (intermediate neural progenitors), and TUJ1 (mature neurons) in KO organoids compared with untreated organoids (Fig. 7D-E). Together, these data demonstrate that pharmacological inhibition by X1 rescues neurogenesis in transplanted cerebral organoids, indicating that XIST restraint by AFF3 is crucial for proper neural development within a complex host environment.

Discussion

Our study introduces several conceptual advances. We show that early-life XIST restraint protects neural and musculoskeletal development: AFF3 loss leads to premature and excessive expression of XIST, with consequent defects in both lineages. Pharmacogenetic targeting verified that XIST is causative, in culture and after transplantation in vivo. Excessive XIST represses autosomal genes required for neural and musculoskeletal development. XIST evolved to equalize X chromosome dosage between females and males; our findings indicate that this same activity carries a risk of developmental defects when it is deployed too early. A key role of AFF3 in early embryonic development is to restrain XIST during the vulnerable window of gastrulation. The loss of NPC and NMP upon AFF3 deletion, and their recovery when XIST is removed, indicate that these populations depend on XIST being held low while they are specified. Because NMPs feed both the neural tube and the somitic mesoderm, a single failure of XIST restraint in this population can account for defects in two lineages at once. These findings place the origin of organogenesis failure at gastrulation, well before the affected organs form, and so reinforce the value of studying gastrulation to understand organogenesis, extending the utility of stem cell-based embryo models.

Restraint of XIST by AFF3 accounts for the loss-of-function class of variants, but not for the dominant-negative class, and our findings reveal that the two act through different mechanisms. Patients with the two classes present with partially overlapping symptoms such as intellectual disability and shortened limbs (*7*, *9*). However, the greater severity of the dominant-negative syndrome, and the epilepsy specific to it, already pointed to a distinct underlying mechanism. This study showed that loss-of-function leads to abnormal XIST upregulation resulting in perturbed neural and musculoskeletal development. In contrast, our data reveal how dominant-negative variants lead to dysregulated neurogenesis biased toward excitatory neurons, potentially explaining epilepsy in patients, without abnormal XIST upregulation. XIST inhibition rescued the loss-of-function phenotype but did not alter the transcriptional signature of dominant-negative gastruloids, accounting for the female bias reported in loss-of-function but not in dominant-negative patients. What drives the dominant-negative phenotype remains to be determined.

More broadly, our findings show that XIST can disrupt early development. By leveraging rare diseases, we can ask how general this mechanism might be. Indeed, pentasomy X (XXXXX) patients exhibit excessive XIST expression as well as a similar spectrum of symptoms to those caused by AFF3 mutations (*16*), including intellectual disability, limb malformation, and craniofacial skeletal abnormalities of neural crest origin. The risk we identify sits against a background of advantage. Two copies of the X chromosome serve as a genetic backup benefiting females. Females are less susceptible to Fragile X intellectual disability because of the backup copy of FMR1 (*51*). Rett syndrome is lethal to most male embryos, whereas females can survive because of a second MECP2 allele, albeit with neurodevelopmental defects (*52*). Some X-linked genes escape inactivation, offering females protection against

metabolic disease, as well as enhanced longevity and greater regenerative capacity compared with males (*53*). To balance X chromosome dose, XIST is uniquely expressed in female cells, with a rare exception in male cancers or XXY Klinefelter Syndrome (*54*). Our study discovered a potential risk posed by XIST during neural and musculoskeletal development, which may be a trade-off of the benefits of having two copies of the X chromosome. Beyond the early embryo, we further found that an XIST enhancer variant is associated with Alzheimer's disease in females. This suggests that XIST regulation could underlie both neurodevelopment and neurodegeneration.

Two main limitations should be considered. First, although gastruloids and cerebral organoids capture key aspects of early human development, they do not fully reproduce the spatial architecture, extraembryonic signaling, maternal environment, vascularization, immune inputs, or long-range tissue interactions of the embryo. Future studies in patient-derived stem cell models, or models carrying patient-specific mutations, will be essential to establish the developmental timing, generality, and disease relevance of these mechanisms. Second, some male patients with AFF3 loss-of-function present with mild intellectual disability. These cases are rare; only nine male patients with intellectual disability are reported (*7*, *9*). Our data showed that AFF3 loss in male cells does not upregulate XIST, so the mechanism we describe cannot account for these patients. The underlying mechanisms in these patients will require follow-up studies.

References

1. J. Briscoe, A. Pierani, T. M. Jessell, J. Ericson, A Homeodomain Protein Code Specifies Progenitor Cell Identity and Neuronal Fate in the Ventral Neural Tube. *Cell* **101**, 435–445 (2000).

2. V. Graham, J. Khudyakov, P. Ellis, L. Pevny, SOX2 Functions to Maintain Neural Progenitor Identity. *Neuron* **39**, 749–765 (2003).

3. L. H. Pevny, S. Sockanathan, M. Placzek, R. Lovell-Badge, A role for SOX1 in neural determination. *Development* **125**, 1967–1978 (1998).

4. A. M. Arias, Y. Marikawa, N. Moris, Gastruloids: Pluripotent stem cell models of mammalian gastrulation and embryo engineering. *Developmental Biology* **488**, 35–46 (2022).

5. Z. Luo, C. Lin, A. Shilatifard, The super elongation complex (SEC) family in transcriptional control. *Nat Rev Mol Cell Biol* **13**, 543–547 (2012).

6. J. Y. Sonn, W. Kim, M. Iwanaszko, Y. Aoi, Y. Li, G. Qi, L. Parkitny, J. L. Brissette, L. Weiner, J. Botas, I. Al-Ramahi, A. Shilatifard, H. Y. Zoghbi, MeCP2 interacts with the super elongation complex to regulate transcription. *Science Advances* **11**, eadt5937 (2025).

7. S. Bassani, J. Chrast, G. Ambrosini, N. Voisin, F. Schütz, A. Brusco, F. Sirchia, L. Turban, S. Schubert, R. Abou Jamra, J.-U. Schlump, D. DeMille, P. Bayrak-Toydemir, G. R. Nelson, K. N. Wong, L. Duncan, M. Mosera, C. Gilissen, L. E. L. M. Vissers, R. Pfundt, R. Kersseboom, H. Yttervik, G. Å. M. Hansen, M. F. Smeland, K. M. Butler, M. J. Lyons, C. M. B. Carvalho, C. Zhang, J. R. Lupski, L. Potocki, L. Flores-Gallegos, R. Morales-Toquero, F. Petit, B. Yalcin, A. Tuttle, H. Z. Elloumi, L. McCormick, M. Kukolich, O. Klaas, J. Horvath, M. Scala, M. Iacomino, F. Operto, F. Zara, K. Writzl, A. Maver, M. K. Haanpää, P. Pohjola, H. Arikka, A. J. A. Kievit, C. Calandrini, C. Iseli, N. Guex, A. Reymond, Variant-specific pathophysiological mechanisms of AFF3 differently influence transcriptome profiles. *Genome Med* **16**, 72 (2024).

8. Y. Inoue, N. Tsuchida, N. Okamoto, S. Shuichi, K. Ohashi, S. Saitoh, A. Ogawa, K. Hamada, M. Sakamoto, N. Miyake, K. Hamanaka, A. Fujita, E. Koshimizu, S. Miyatake, T. Mizuguchi, K. Ogata, Y. Uchiyama, N. Matsumoto, Three KINSSHIP syndrome patients with mosaic and germline AFF3 variants. *Clinical Genetics* **103**, 590–595 (2023).

9. N. Voisin, R. E. Schnur, S. Douzgou, S. M. Hiatt, C. F. Rustad, N. J. Brown, D. L. Earl, B. Keren, O. Levchenko, S. Geuer, S. Verheyen, D. Johnson, Y. A. Zarate, M. Hančárová, D. J. Amor, E. M. Bebin, J. Blatterer, A. Brusco, G. Cappuccio, J. Charrow, N. Chatron, G. M. Cooper, T. Courtin, E. Dadali, J. Delafontaine, E. Del Giudice, M. Doco, G. Douglas, A. Eisenkölbl, T. Funari, G. Giannuzzi, U. Gruber-Sedlmayr, N. Guex, D. Heron, Ø. L. Holla, A. C. E. Hurst, J. Juusola, D. Kronn, A. Lavrov, C. Lee, S. Lorrain, E. Merckoll, A. Mikhaleva, J. Norman, S. Pradervand, D. Prchalová, L. Rhodes, V. R. Sanders, Z. Sedláček, H. A. Seebacher, E. A. Sellars, F. Sirchia, T. Takenouchi, A. J. Tanaka, H. Taska-Tench, E. Tønne, K. Tveten, G. Vitiello, M. Vlčková, T. Uehara, C. Nava, B. Yalcin, K. Kosaki, D. Donnai, S. Mundlos, N. Brunetti-Pierri, W. K. Chung, A. Reymond, Variants in the degron of AFF3 are associated with intellectual disability, mesomelic dysplasia, horseshoe kidney, and epileptic encephalopathy. *The American Journal of Human Genetics* **108**, 857–873 (2021).

10. Y. Zhang, C. Wang, X. Liu, Q. Yang, H. Ji, M. Yang, M. Xu, Y. Zhou, W. Xie, Z. Luo, C. Lin, AFF3-DNA methylation interplay in maintaining the mono-allelic expression pattern of XIST in terminally differentiated cells. *J Mol Cell Biol* **11**, 761–769 (2019).

11. C. J. Brown, A. Ballabio, J. L. Rupert, R. G. Lafreniere, M. Grompe, R. Tonlorenzi, H. F. Willard, A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44 (1991).

12. A. Loda, E. Heard, Xist RNA in action: Past, present, and future. *PLOS Genetics* **15**, e1008333 (2019).

13. S. Petropoulos, D. Edsgärd, B. Reinius, Q. Deng, S. P. Panula, S. Codeluppi, A. Plaza Reyes, S. Linnarsson, R. Sandberg, F. Lanner, Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165**, 1012–1026 (2016).

14. B. Zeng, Z. Liu, Y. Lu, S. Zhong, S. Qin, L. Huang, Y. Zeng, Z. Li, H. Dong, Y. Shi, J. Yang, Y. Dai, Q. Ma, L. Sun, L. Bian, D. Han, Y. Chen, X. Qiu, W. Wang, O. Marín, Q.

Wu, Y. Wang, X. Wang, The single-cell and spatial transcriptional landscape of human gastrulation and early brain development. *Cell Stem Cell* **30**, 851-866.e7 (2023).

15. I. Dror, T. Chitiashvili, S. Y. X. Tan, C. T. Cano, A. Sahakyan, Y. Markaki, C. Chronis, A. J. Collier, W. Deng, G. Liang, Y. Sun, A. Afasizheva, J. Miller, W. Xiao, D. L. Black, F. Ding, K. Plath, XIST directly regulates X-linked and autosomal genes in naive human pluripotent cells. *Cell* **187**, 110-129.e31 (2024).

16. L. M. Moraes, L. C. Cardoso, V. L. Moura, M. A. Moreira, A. N. Menezes, J. C. Llerena, H. N. Seuánez, Detailed analysis of X chromosome inactivation in a 49,XXXXX pentasomy. *Mol Cytogenet* **2**, 20 (2009).

17. X. Gao, M. Nowak-Imialek, X. Chen, D. Chen, D. Herrmann, D. Ruan, A. C. H. Chen, M. A. Eckersley-Maslin, S. Ahmad, Y. L. Lee, T. Kobayashi, D. Ryan, J. Zhong, J. Zhu, J. Wu, G. Lan, S. Petkov, J. Yang, L. Antunes, L. S. Campos, B. Fu, S. Wang, Y. Yong, X. Wang, S.-G. Xue, L. Ge, Z. Liu, Y. Huang, T. Nie, P. Li, D. Wu, D. Pei, Y. Zhang, L. Lu, F. Yang, S. J. Kimber, W. Reik, X. Zou, Z. Shang, L. Lai, A. Surani, P. P. L. Tam, A. Ahmed, W. S. B. Yeung, S. A. Teichmann, H. Niemann, P. Liu, Establishment of porcine and human expanded potential stem cells. *Nat Cell Biol* **21**, 687–699 (2019).

18. Y. J. Passamaneck, L. Katikala, L. Perrone, M. P. Dunn, I. Oda-Ishii, A. Di Gregorio, Direct activation of a notochord cis-regulatory module by Brachyury and FoxA in the ascidian Ciona intestinalis. *Development* **136**, 3679–3689 (2009).

19. T. Rito, A. R. G. Libby, M. Demuth, M.-C. Domart, J. Cornwall-Scoones, J. Briscoe, Timely TGFβ signalling inhibition induces notochord. *Nature* **637**, 673–682 (2025).

20. D. L. Chapman, I. Agulnik, S. Hancock, L. M. Silver, V. E. Papaioannou, Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Dev Biol* **180**, 534–542 (1996).

21. Y. Takahashi, S. Hiraoka, S. Kitajima, T. Inoue, J. Kanno, Y. Saga, Differential contributions of Mesp1 and Mesp2 to the epithelialization and rostro-caudal patterning of somites. *Development* **132**, 787–796 (2005).

22. Y. Liu, F. Cai, M. Barile, Y. Chang, D. Cao, Y. Huang, TemporalVAE: atlas-assisted temporal mapping of time-series single-cell transcriptomes during embryogenesis. *Nat Cell Biol* **27**, 1982–1992 (2025).

23. M. L. Echevarría-Andino, N. E. Franks, H. E. Schrader, M. Hong, R. S. Krauss, B. L. Allen, CDON contributes to Hedgehog-dependent patterning and growth of the developing limb. *Dev Biol* **493**, 1–11 (2023).

24. J. Yang, L.-H. Hung, T. Licht, S. Kostin, M. Looso, E. Khrameeva, A. Bindereif, A. Schneider, T. Braun, RBM24 is a major regulator of muscle-specific alternative splicing. *Dev Cell* **31**, 87–99 (2014).

25. M. Buckingham, F. Relaix, The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions. *Annu Rev Cell Dev Biol* **23**, 645–673 (2007).

26. I. Singh, D. Fernandez-Perez, P. S. Sanchez, A. E. Rodriguez-Fraticelli, Pre-existing stem cell heterogeneity dictates clonal responses to the acquisition of leukemic driver mutations. *Cell Stem Cell* **32**, 564-580.e6 (2025).

27. Y. Zhang, C. Wang, X. Liu, Q. Yang, H. Ji, M. Yang, M. Xu, Y. Zhou, W. Xie, Z. Luo, C. Lin, AFF3-DNA methylation interplay in maintaining the mono-allelic expression pattern of XIST in terminally differentiated cells. *J Mol Cell Biol* **11**, 761–769 (2019).

28. S. Yokobayashi, Y. Yabuta, M. Nakagawa, K. Okita, B. Hu, Y. Murase, T. Nakamura, G. Bourque, J. Majewski, T. Yamamoto, M. Saitou, Inherent genomic properties underlie the epigenomic heterogeneity of human induced pluripotent stem cells. *Cell Reports* **37**, 109909 (2021).

29. Y. Wang, N. Yang, W. Sun, C. Zhao, X. Hu, S. Lu, S. Cao, N. Wang, T. Hai, G. Feng, C. An, H. Wang, Characterisation of X chromosome status of human extended pluripotent stem cells. *Cell Proliferation* **56**, e13468 (2023).

30. J. Lobo, S. P. Nunes, A. J. M. Gillis, D. Barros-Silva, V. Miranda-Gonçalves, A. van den Berg, M. Cantante, R. Guimarães, R. Henrique, C. Jerónimo, L. H. J. Looijenga, XIST-Promoter Demethylation as Tissue Biomarker for Testicular Germ Cell Tumors and Spermatogenesis Quality. *Cancers* **11**, 1385 (2019).

31. C. Vallot, C. Patrat, A. J. Collier, C. Huret, M. Casanova, T. M. Liyakat Ali, M. Tosolini, N. Frydman, E. Heard, P. J. Rugg-Gunn, C. Rougeulle, XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell* **20**, 102–111 (2017).

32. J. T. Lee, Regulation of X-Chromosome Counting by Tsix and Xite Sequences. *Science* **309**, 768–771 (2005).

33. X. Li, N. Liu, Advances in understanding LINE-1 regulation and function in the human genome. *Trends Genet* **41**, 577–589 (2025).

34. R. Aguilar, K. B. Spencer, B. Kesner, N. F. Rizvi, M. D. Badmalia, T. Mrozowich, J. D. Mortison, C. Rivera, G. F. Smith, J. Burchard, P. J. Dandliker, T. R. Patel, E. B. Nickbarg, J. T. Lee, Targeting Xist with compounds that disrupt RNA structure and X inactivation. *Nature* **604**, 160–166 (2022).

35. A. Smith, V. Robinson, K. Patel, D. G. Wilkinson, The EphA4 and EphB1 receptor tyrosine kinases and ephrin-B2 ligand regulate targeted migration of branchial neural crest cells. *Current Biology* **7**, 561–570 (1997).

36. A. Compagni, M. Logan, R. Klein, R. H. Adams, Control of Skeletal Patterning by EphrinB1-EphB Interactions. *Developmental Cell* **5**, 217–230 (2003).

37. T. Maes, A. Barceló, C. Buesa, Neuron Navigator: A Human Gene Family with Homology to *unc-53*, a Cell Guidance Gene from *Caenorhabditis elegans*. *Genomics* **80**, 21–30 (2002).

38. M. L. Echevarría-Andino, N. E. Franks, H. E. Schrader, M. Hong, R. S. Krauss, B. L. Allen, CDON contributes to Hedgehog-dependent patterning and growth of the developing limb. *Dev Biol* **493**, 1–11 (2023).

494 39. I. Dror, T. Chitiashvili, S. Y. X. Tan, C. T. Cano, A. Sahakyan, Y. Markaki, C. Chronis,
495 A. J. Collier, W. Deng, G. Liang, Y. Sun, A. Afasizheva, J. Miller, W. Xiao, D. L.
496 Black, F. Ding, K. Plath, XIST directly regulates X-linked and autosomal genes in naive
497 human pluripotent cells. *Cell* **187**, 110-129.e31 (2024).

498 40. H. Michibata, T. Okuno, N. Konishi, K. Kyono, K. Wakimoto, K. Aoki, Y. Kondo, K.
499 Takata, Y. Kitamura, T. Taniguchi, Human GPM6A is associated with differentiation
500 and neuronal migration of neurons derived from human embryonic stem cells. *Stem*
501 *Cells Dev* **18**, 629–639 (2009).

502 41. F. Hazan, C. Yılmaz Uzman, D. A. Emecen, M. M. Yılmazer, T. Öztürk, M. M. Bilen,
503 A. Hüvez, R. İlhan, F. Genel, P. Ballar, A Candidate Gene for Cerebro-Oculo-Nasal
504 Syndrome: A Zinc-Finger Gene ZNF185 Located at Xq28. *Mol Syndromol*, doi:
505 10.1159/000549375 (2025).

506 42. T. Lu, R. Chen, T. C. Cox, R. X. Moldrich, N. Kurniawan, G. Tan, J. K. Perry, A.
507 Ashworth, P. F. Bartlett, L. Xu, J. Zhang, B. Lu, M. Wu, Q. Shen, Y. Liu, L. J.
508 Richards, Z. Xiong, X-linked microtubule-associated protein, Mid1, regulates axon
509 development. *Proc Natl Acad Sci U S A* **110**, 19131–19136 (2013).

510 43. M. L. Gonzales, J. M. LaSalle, The role of MeCP2 in brain development and
511 neurodevelopmental disorders. *Curr Psychiatry Rep* **12**, 127–134 (2010).

512 44. X. Zhang, C. T. Huang, J. Chen, M. T. Pankratz, J. Xi, J. Li, Y. Yang, T. M. LaVaute,
513 X.-J. Li, M. Ayala, G. I. Bondarenko, Z.-W. Du, Y. Jin, T. G. Golos, S.-C. Zhang, Pax6
514 Is a Human Neuroectoderm Cell Fate Determinant. *Cell Stem Cell* **7**, 90–100 (2010).

515 45. J. Kele, E. R. Andersson, J. C. Villaescusa, L. Cajanek, C. L. Parish, S. Bonilla, E. M.
516 Toledo, V. Bryja, J. S. Rubin, A. Shimono, E. Arenas, SFRP1 and SFRP2 Dose-
517 Dependently Regulate Midbrain Dopamine Neuron Development In Vivo and in
518 Embryonic Stem Cells. *Stem Cells* **30**, 865–875 (2012).

519 46. M. Bertin, H. Todorov, S. Frank, S. Käseberg, R. Menon, E. Gabassi, C. Foerster, N.
520 Bobon, F. Furlanetto, A. Soliman, H. M. B. Ibrahim, V. Engelhardt, L. Birschmann, H.
521 Brennenstuhl, B. Lohrer, A. Mas-Sanchez, E. Cesare, J. Winter, J. Krummeich, J.
522 Winkler, B. Winner, E. Weis, S. Diederich, K. Luck, P. Lunt, S. Gerber, P. Baumann,
523 N. Elvassore, B. Berninger, M. F. Basilicata, S. Schweiger, S. Falk, M. Karow,
524 Dynamic allele usage of X-linked genes ameliorates neurodevelopmental disease
525 phenotypes in brain organoids. *Nat Commun* **17**, 599 (2026).

526 47. C. Thapa, Female iPSC X-chromosome inactivation (XCI) erosion and its
527 transcriptomic effects during CRISPR gene editing and neural differentiation.[Preprint]
528 (2026). https://doi.org/10.64898/2026.02.27.708613.

529 48. Z. Luo, C. Lin, A. R. Woodfin, E. T. Bartom, X. Gao, E. R. Smith, A. Shilatifard,
530 Regulation of the imprinted Dlk1-Dio3 locus by allele-specific enhancer activity. *Genes*
531 *Dev.* **30**, 92–101 (2016).

532 49. D. Müller, P. Cherukuri, K. Henningfeld, C. H. Poh, L. Wittler, P. Grote, O. Schlüter, J.
533 Schmidt, J. Laborda, S. R. Bauer, R. M. Brownstone, T. Marquardt, Dlk1 Promotes a

Fast Motor Neuron Biophysical Signature Required for Peak Force Execution. *Science* **343**, 1264–1266 (2014).

50. R. Montalbán-Loro, G. Lassi, A. Lozano-Ureña, A. Perez-Villalba, E. Jiménez-Villalba, M. Charalambous, G. Vallortigara, A. E. Horner, L. M. Saksida, T. J. Bussey, J. L. Trejo, V. Tucci, A. C. Ferguson-Smith, S. R. Ferrón, Dlk1 dosage regulates hippocampal neurogenesis and cognition. *Proceedings of the National Academy of Sciences* **118**, e2015505118 (2021).

51. D. C. Crawford, J. M. Acuña, S. L. Sherman, FMR1 and the fragile X syndrome: Human genome epidemiology review. *Genet Med* **3**, 359–371 (2001).

52. R. E. Amir, I. B. Van den Veyver, M. Wan, C. Q. Tran, U. Francke, H. Y. Zoghbi, Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23**, 185–188 (1999).

53. A. P. Arnold, K. Reue, M. Eghbali, E. Vilain, X. Chen, N. Ghahramani, Y. Itoh, J. Li, J. C. Link, T. Ngun, S. M. Williams-Burris, The importance of having two X chromosomes. *Phil. Trans. R. Soc. B* **371**, 20150113 (2016).

54. A. Sadagopan, I. T. Nasim, J. Li, M. Achom, C.-Z. Zhang, S. R. Viswanathan, Somatic XIST activation and features of X chromosome inactivation in male human cancers. *Cell Systems* **13**, 932-944.e5 (2022).

55. E. C. Jonlin, M. Fujita, R. Isasi, K. Kato, M. Munsie, K. Muto, K. Niakan, K. Saha, J. Sugarman, L. Turner, I. Hyun, What does "appropriate scientific justification" mean for the review of human pluripotent stem cell, embryo, and related research? *Stem Cell Reports* **20** (2025).

56. A. C. H. Chen, Y. L. Lee, H. Ruan, W. Huang, S. W. Fong, S. Tian, K. C. Lee, G. M. Wu, Y. Tan, T. C. H. Wong, J. Wu, W. Zhang, D. Cao, J. F. C. Chow, P. Liu, W. S. B. Yeung, Expanded Potential Stem Cells from Human Embryos Have an Open Chromatin Configuration with Enhanced Trophoblast Differentiation Ability. *Advanced Science* **10**, 2204797 (2023).

57. C.-E. Choi, C. Liang, Y. Shamiya, S. J. Lee, A. Paul, Co-Delivery of Ca-MOF and Mg-MOF Using Nanoengineered Hydrogels to Promote In Situ Mineralization and Bone Defect Repair: In Vitro and In Vivo Analysis. *Advanced Healthcare Materials* **14**, e02630 (2025).

58. C. Liang, Z. Wu, Z. Huang, S. Chen, F. Li, Y.-D. Kwon, D. Lee, S.-H. An, S. Choi, S.-H. Kim, S. J. Lee, Fully and spontaneously degradable carboxymethyl chitosan/oxidized hyaluronic acid-based membrane promoting three-dimensional tissue remodeling *in vitro* and *in vivo*. *International Journal of Biological Macromolecules* **371**, 152640 (2026).

Methods

Culture and γ-irradiation of SNL feeder cells

SNL feeder cells were kindly provided by Pengtao Liu. When cultures reached 80%–90% confluency, cells were dissociated with TrypLE (Thermo Fisher Scientific, 12605036) for 3 min at 37°C. Enzymatic digestion was terminated with high glucose DMEM (Thermo Fisher Scientific, 12100046) supplemented with 1× Penicillin-Streptomycin-Glutamine (Thermo Fisher Scientific, 10378016) and 10% fetal bovine serum (Gibco, A3840001). Cells were collected by centrifugation at 300 × g for 5 min, resuspended in M10 medium (DMEM with 10% FBS), and seeded onto cell culture plates pre-coated with 0.2% gelatin. For feeder preparation, live SNL cells were expanded in 15-cm dishes for 3–5 days, harvested, and resuspended in 15 mL M10 medium. Cell suspensions were then irradiated using an MDS Nordion Gammacell 3000 Elan at a dose of 30 Gy per batch.

Culture of human expanded potential stem cells (hEPSCs)

Human EPSCs were cultured in human EPSC medium (EPSCM) as previously described, with modifications (*17*). Briefly, EPSCM consisted of a 1:1 mixture of DMEM/F12 (Thermo Fisher Scientific, 21331020) and Neurobasal medium (Thermo Fisher Scientific, 21103049), supplemented with 1× Penicillin-Streptomycin-Glutamine (Thermo Fisher Scientific, 10378016), 1× non-essential amino acids (Thermo Fisher Scientific, 11140050), N2 supplement (Thermo Fisher Scientific, 17502-048), B27 supplement (Thermo Fisher Scientific, 17504-044), Insulin-Transferrin-Selenium-Ethanolamine (ITS-X; Thermo Fisher Scientific, 51500056), 1× sodium pyruvate (Thermo Fisher Scientific, 11360070), 1× trace elements B (Corning, MT99175CI) and C (Corning, MT99176CI), 1× chemically defined lipid concentrate (Gibco, 11905031), 0.045% bovine serum albumin fraction V (Thermo Fisher Scientific, 15260037), and 0.1 mM β-mercaptoethanol (Gibco, 21985–023). To generate complete human EPSCM, the medium was further supplemented with 64 μg/mL L-ascorbic acid (Sigma, A8960), 3.3 μg/mL reduced glutathione (Sigma, G4251), 100 ng/mL human LIF (Peprotech, 300–05), 1 μM CHIR99021 (Tocris, 4423), 2.5 μM XAV939 (Sigma, X3004), 2.5 μM endo-IWR-1 (Tocris, 3532), and 0.1 μM A419259 (Tocris, 3914).

To prepare SNL feeder plates, 1 mL of 0.2% autoclaved gelatin (Sigma, G2500) was added to each well of a 12-well plate, and plates were incubated at 37°C overnight (≥12 h). The following day, one vial of SNL feeder cells was thawed in a 37°C water bath and transferred to a 15-mL conical tube containing 1 mL M10 medium (DMEM supplemented with 10% FBS). Cells were centrifuged at 300 × g for 3 min, the supernatant was discarded, and the pellet was resuspended in 1.5 mL M10 medium. After removal of the gelatin solution, 1 mL M10 medium was added to each well, followed by seeding of 100 μL of the resuspended SNL cell suspension per well. Cells were distributed evenly across the plate and cultured at 37°C until use. Under these conditions, one vial of SNL feeder cells was sufficient for seeding 18 wells of a 12-well plate, yielding ~50%–80% confluency after 2 days. Feeder plates were typically prepared 2

days before passaging to allow the feeder cells to acquire a flat, fibroblast-like morphology, and were used between days 2 and 7 after plating.

Passaging hEPSC: Cells were passaged at ~80% confluency. Cultures were first aspirated and washed 3 times with 1 mL PBS per well of a 12-well plate. PBS was removed, and 0.5 mL TrypLE (equilibrated to room temperature) was added to each well. Cells were incubated at 37°C for 2 min until they adopted a rounded morphology. Dissociation was terminated by adding 1 mL M10 medium per well, and cells were gently pipetted to obtain a single-cell suspension. The cell suspension was transferred to a 15-mL conical tube and centrifuged at 300 × g for 3 min. After removal of the supernatant, cells were resuspended in 1 mL EPSCM supplemented with 10 μM Y-27632. Meanwhile, culture medium was aspirated from the feeder plate and replaced with 1 mL fresh EPSCM containing 10 μM Y-27632 per well. Cells were then seeded onto the prepared feeder plates at a split ratio of 1:5 to 1:15, and plates were gently tapped on all sides to ensure even distribution of the cell suspension. Medium was replaced 24 h after passaging with fresh EPSCM lacking Y-27632. On day 2, the medium was changed again, and thereafter cultures were fed daily until the next passage.

We followed the ISSCR ethical regulations for human embryo model research (*55*). IRB approval was obtained from HKU (UW 25-535).

hEPSC pre-differentiation

hEPSCs were maintained in EPSCM, with medium changed daily, as previously reported (*17*). Before embryoid body (EB) formation, cells were pre-differentiated in DFK medium consisting of DMEM/F12 (Thermo Fisher Scientific, 21331020) supplemented with 10% KSR (Thermo Fisher Scientific, 10828028) for 3 days, with daily medium changes. To ensure efficient pre-differentiation, the initial cell density was maintained below 30% confluency.

EB formation

After 3 days of pre-differentiation, DFK medium was aspirated and cells were washed with PBS (Thermo Fisher Scientific, 10010031). Cells were then incubated with 500 μL TrypLE at 37°C for 2 min and examined under a microscope. At this stage, most feeder cells had typically detached. The TrypLE was aspirated and replaced with an additional 500 μL of fresh TrypLE for 1–2 min to further digest the hEPSCs. Dissociation was terminated by adding 2 mL DFK medium, and cells were collected with gentle pipetting to minimize mechanical stress. The cell suspension was centrifuged at 300 × g for 3 min, after which the supernatant was carefully removed. Cells were resuspended in 1 mL DFK medium supplemented with 10 μM Y-27632 (Selleckchem, S1049). For aggregation, 4,000 cells in 150 μL medium were seeded into each well of a low-attachment U-bottom 96-well plate. On the following day, 100 μL fresh DFK medium without Y-27632 was added to each well. Plates were maintained at 37°C for 3 days.

Gastruloid formation

EBs were collected into 15-mL conical tubes and allowed to settle for 3 min or centrifuged at 100 × g for 1 min. The supernatant was carefully removed, and the EBs were resuspended in 1 mL Medium A from the STEMdiff Hematopoietic Kit (STEMCELL Technologies, 05310). EBs were then transferred to low-attachment 24-well plates and cultured in 1 mL Medium A per well (day 0). Cultures were maintained in Medium A for 3 days, with a half-medium change performed on day 2. Thereafter, EBs were cultured in 1 mL Medium B from the same kit for 5 days, with half-medium changes every other day. Throughout differentiation, gastruloids were maintained in low-attachment plates and placed on an orbital shaker at ≤100 rpm from the Medium A stage through the Medium B stage.

Genome editing in human EPSCs by CRISPR-Cas9

Preparation of gRNA. For AFF3 targeting, sgRNAs were manually designed against exons 3, 4, and 7. For XIST, sgRNAs were designed based on a previous study (*15*). Candidate sgRNA sequences were obtained using CHOPCHOP (https://chopchop.cbu.uib.no/) and synthesized by Tsingke Biotechnology. Editing efficiency was initially predicted using the CRISPOR algorithm (https://crispor.gi.ucsc.edu/) and subsequently validated experimentally in hEPSCs by electroporation.

CRISPR-Cas9 genome editing. Female C5 and male EM-4 hEPSC lines (*17*, *56*) were maintained for 4–5 days before electroporation until they reached 70%–80% confluency. Electroporation was performed using the Lonza P3 Primary Cell 4D-Nucleofector Kit (Lonza, V4XP-3032) according to the manufacturer's instructions. Following electroporation, cells were replated, and the medium was replaced after 24 h. Edited hEPSCs were then expanded for several days, passaged, diluted to 5,000 cells/mL, and reseeded into multiwell plates.

Clonal isolation and genotyping. After 10 days, individual colonies were manually picked and subjected to genotyping. Approximately 100 single-cell-derived clones per cell line were screened by gel electrophoresis and Sanger sequencing to confirm the intended deletions. Clones harbouring distinct insertion-deletion (indel) mutations were selected for downstream experiments, yielding 2–3 positive clones per cell line. Sanger sequencing results for all knockout clones are presented in Supp. Fig. 2B and 2D.

Extraction of genomic DNA from human EPSCs and polymerase chain reaction (PCR)

Following gene editing, polyclonal populations were expanded for 1 week and passaged at least twice before single-clone isolation. Cells were then diluted to enable clonal outgrowth, ideally at a density yielding no more than 50 colonies per dish. After 4–5 days of culture, individual round colonies were identified microscopically and manually picked using 10 µL

pipette tips. Colonies were dissociated with TrypLE and transferred to 96-well plates, with half of each colony used for continued culture and the other half reserved for genotyping.

For single-colony genotyping, fresh lysis buffer was prepared by combining 910 µL water, 40 µL 10% NP-40, and 50 µL Proteinase K. A total of 15 µL cell suspension was added to 20 µL lysis buffer in PCR tubes. Cell lysis was carried out in a thermal cycler under the following conditions: 65°C for 30 min, followed by 95°C for 10 min. After lysis, samples were centrifuged to pellet insoluble debris. PCR was then performed according to the manufacturer's instructions for the specific DNA polymerase used.

Bulk RNA extraction and qPCR

Samples containing at least 1 × 10^5 cells were collected in 1.5 mL microcentrifuge tubes and lysed in TRIzol reagent (Thermo Fisher Scientific, 15596026) for 5–10 min. Chloroform was then added at a TRIzol: chloroform ratio of 500 µL:100 µL, followed by vigorous mixing and centrifugation at 12,000 × g at 4°C for 5 min. After phase separation, the aqueous phase containing total RNA was transferred to a fresh tube, and RNA was precipitated with 250 µL of isopropanol. The RNA pellet was washed twice with 75% ethanol, air-dried for 10–20 min, and dissolved in RNase-free water, followed by DNase treatment. Purified RNA was stored at −80°C until further use. RNA concentration and quality were assessed using a NanoDrop 2000c spectrophotometer.

cDNA was synthesized using PrimeScript RT Master Mix (Takara, RR036A) with 1 µg total RNA in a 10 µL reaction volume. The resulting cDNA was diluted and used for quantitative PCR analysis. RT-qPCR was performed using TB Green Premix Ex Taq II (Takara, RR820A) on a CFX96 Real-Time PCR System (Bio-Rad). Technical triplicates were included for each sample. GAPDH was used as the internal reference gene to calculate relative gene expression.

Flow cytometry analysis and cell sorting

AFF3 KO and wild-type hEPSC lines were differentiated using the gastruloid or cerebral organoid protocols described above. Following in vitro differentiation, AFF3 knockout and wild-type gastruloids or organoids were harvested separately, dissociated into single cells, and transferred to FACS buffer. After digestion, cells were washed with PBS, centrifuged, and resuspended in homemade FACS buffer (PBS supplemented with 2% FBS).

For extracellular staining, cells were first incubated with 5 µL Fc-blocking antibody (BioLegend, 422302) for 10 min at room temperature to reduce non-specific binding. Cells were then stained for 30 min at 4°C in the dark. After staining, cells were washed twice with ice-cold FACS buffer, centrifuged at 300 × g for 5 min at 4°C, and resuspended in FACS buffer containing 1 µg/mL DAPI (Abcam, ab228549). Cell sorting was performed using a BD FACSAria Fusion cell sorter, and flow cytometry analysis was carried out on a BD LSRFortessa within 24 h.

Library preparation and bulk RNA sequencing

Total RNA was extracted using TRIzol reagent as described above (Bulk RNA extraction and qPCR). RNA-seq library preparation was performed by the HKU Centre for PanorOmic Sciences (CPOS) core facility. Library quality was assessed using an Agilent 2100 Bioanalyzer, and sequencing was carried out on an Illumina HiSeq platform.

Raw sequencing reads were processed using Trim-Galore to remove low-quality bases and adapter sequences. Filtered reads were aligned to the human reference genome (GRCh38) using HISAT2. Gene-level read counts were generated with featureCounts and normalized using deepTools. Differentially expressed genes were identified using the following criteria: fold change > 1.5 and adjusted p-value < 0.05. Volcano plots and Gene Ontology (GO) analyses were generated based on these filtered genes. Heatmaps of selected genes were visualized using Z-score normalization.

Single-cell RNA-seq and dataset analysis

WT and AFF3 KO hEPSCs were differentiated using the gastruloid and cerebral organoid protocols described above. Following in vitro differentiation, AFF3 knockout and wild-type gastruloids or organoids were harvested separately, dissociated into single cells, and transferred to FACS buffer. After digestion, cells were washed with PBS, centrifuged at 100 × g for 5 min, and resuspended in FACS buffer (PBS supplemented with 2% FBS) containing 1 μg/mL DAPI. DAPI-negative cells were sorted for subsequent single-cell RNA sequencing. Cell quality control, library construction, and sequencing were performed by the HKU Core Facility. A total of 206.2 Gb and 213.9 Gb of sequencing data were generated for the AFF3 knockout and wild-type samples, respectively.

Raw scRNA-seq data were pre-processed using Cell Ranger (v7.0.0) and aligned to the human reference genome (GRCh38). Gene-by-cell count matrices were generated and analysed using Seurat (v4.2.0) in R. Quality control was performed independently for each dataset. Cells with fewer than 500 detected genes, fewer than 2,000 RNA counts, or greater than 15% mitochondrial gene expression were excluded. Putative doublets were identified using DoubletFinder (v2.0.3) and removed.

After quality control, potential batch effects between genotypes were corrected using Seurat's FindIntegrationAnchors workflow. The top 3,000 variable genes were selected for principal component analysis–based dimensionality reduction. Cell clusters were identified and visualized using UMAP. Cell type annotation was performed based on cluster-enriched marker genes identified by Seurat FindAllMarkers together with canonical lineage markers. Differentially expressed genes (DEGs) between genotypes in gastruloids were identified using Seurat FindMarkers with the binomial likelihood ratio test. DEGs were defined as genes with q-value < 0.05 and $|\log_2\text{fold change}| > 0.1$. Gene set enrichment analysis for GO and KEGG

pathways was performed using gseGO and gseKEGG from clusterProfiler (v4.4.4) on all genes ranked by DEG test statistics. Significantly enriched pathways were defined by p-value < 0.05.

LARRY infection and library amplification

The LARRY-v2 lentiviral barcode library was obtained from Addgene (*26*). For virus production in 20 dishes, 91 µg of LARRY library plasmids, 65 µg of psPAX2, and 120 µg of pMD2.G were diluted in 11 mL of Opti-MEM and mixed thoroughly. In a separate tube, 1.5 mL of 1 mg/mL PEI (3:1 PEI: DNA, w/w) (Polysciences, 23966-1) was diluted in 11 mL of Opti-MEM and mixed well. The PEI mixture was added dropwise to the DNA mixture, vortexed for 15 s, and incubated at room temperature for exactly 15 min. The transfection mixture was then added dropwise (1 mL per dish) to HEK293T cells at 70–80% confluence. The dishes were incubated at 37 °C with 5% CO_2. After 24 h, the medium was changed to reduce PEI toxicity. Viral supernatant was collected at 48 h and 72 h post-transfection, pooled, and centrifuged at 3,000 × g for 10 min at 4 °C to remove cell debris. The clarified supernatant was centrifuged at 20,000 × g for 2 h at 4 °C using either a swinging-bucket or a fixed-angle rotor. The supernatant was carefully removed, and the virus pellet was resuspended in PBS or DMEM. Aliquots were stored at –80 °C, and repeated freeze–thaw cycles were avoided.

Viral titre was determined by serial dilution on HEK293T cells. hEPSCs at the pre-differentiation stage were infected at different multiplicities of infection (MOI: 0.1, 1, 10). Because the MOI values were based on HEK293T cells, we selected the optimal MOI that yielded approximately 10% infected hEPSCs. Wild-type (WT) and AFF3 knockout (KO) hEPSCs were subsequently infected with the LARRY virus at the chosen MOI. Following the in vitro differentiation workflow, gastruloids were dissociated into single-cell suspensions using Accumax. DMEM supplemented with 10% FBS was added, and the cells were centrifuged at 800 × g for 8 min at 4 °C. The supernatant was carefully removed without disturbing the pellet. Cells were washed twice with FACS buffer (800 × g, 5 min) and resuspended in FACS buffer (PBS with 0.5% BSA and 2 mM EDTA). DAPI solution was added, and the cells were incubated for 5 min at 4 °C. Immediately after cell sorting, the sorted cells were processed using the Chromium Single Cell 3’ Reagent Kit v3 (10x Genomics) strictly according to the manufacturer’s instructions.

Leftover cDNA from the AFF3 WT and AFF3 KO scRNA-seq samples was used for library preparation. A first PCR was set up with 100 ng of cDNA template, 0.5 µM pre-enrichment forward primer (CTGAGCAAAGACCCCAACGAGAA), and 0.5 µM of the corresponding 10x dual-index TruSeq reverse primer (ACACTCTTTCCCTACACGACGCTCTTCCGATCT) using a high-fidelity polymerase master mix. The cycling program was: 98 °C for 3 min; 8 cycles of 98 °C for 20 s, 58 °C for 15 s, and 72 °C for 20 s; 72 °C for 3 min; hold at 4 °C. The product was used directly as a template for a second PCR with 0.5 µM TruSeq_LARRY forward primer (GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCTAGGAGAGACCATATGGGATC) and the same sample-specific dual-index TruSeq reverse primer, applying the identical cycling

program. Because non-specific products were present, gel extraction was performed, and the DNA was eluted in 10–20 µL of elution buffer. The purified PCR products were indexed using the 10x Dual Index Kit. Library size distribution was confirmed by Bioanalyzer, and concentration was quantified by Qubit or qPCR. Indexed libraries were pooled equimolarly and sequenced on an Illumina NovaSeq or NextSeq platform.

Immunofluorescence from paraffin sectioning

Slides containing paraffin sections of gastruloids and brain organoids were dewaxed in xylene, and rehydrated sequentially in 100%, 95%, 90% and 70% EtOH. Antigen retrieval was performed with antigen retrieval buffer (Dako, S1700) by microwave until boiling at high-power heating, followed by 15 min of low-power heating. After cooling to room temperature, samples were permeabilzed with 0.3% Triton X-100 and 1% BSA in PBS for 30 min, and blocked with 10% FBS in PBS for 30-60 min. Primary antibodies were diluted according to manufacturer's recommendation in PBS with 10% FBS and incubated in wet box at 4°C overnight. Samples were washed three times for 5 min each with 0.2% PBS-T. Secondary antibodies were diluted (1:2000) in PBS with 10% FBS and incubated avoiding light at room temperature for 30 min. Samples were washed three times for 5 min each with 0.2% PBS-T, incubated with DAPI (1:1000) for five minutes avoiding light, and mounted with Fluoromount-G (Southern Biotech, 0100-35), sealed, and imaged.

DNA FISH and IF

Cells were cytospun at 1,000 × g for 3 minutes, and slides were air-dried for at least 1–2 h. The slides were fixed with 4% PFA for 15 minutes, then washed twice with PBS. 1× antigen retrieval buffer was preheated to 80°C by microwave for 20 minutes. Slides were allowed to cool down, rinsed quickly with PBS, and a hydrophobic barrier was drawn around the cells. Cells were then blocked with 5% goat serum in a humidified chamber for 20 minutes at room temperature. 20 µL of diluted primary antibody was applied to cover the cells, and slides were incubated overnight at 4°C in a humidified chamber. The next day, slides were washed with either PBST (3 × 5 minutes) or PBS (1 × 5 minutes). Diluted secondary antibody was added and slides were incubated for 1 h at room temperature. From this step onward, all procedures were performed in the dark. Slides were fixed again with 4% PFA for 15 minutes and washed twice with PBS. The slides were equilibrated in 2× SSC (2 × 5 minutes) and dehydrated sequentially in 70%, 85%, and 100% ethanol (2 minutes each), then air-dried. 10 µL of probe, mixed and briefly centrifuged, was applied; a coverslip was added and sealed with rubber cement. The slides were placed in a PCR cycler: 82°C for 10 minutes (co-denaturation), then 45°C for 2 hours (hybridization). The slides were washed in pre-warmed (68°C) 0.3% NP-40 / 0.4× SSC for 2–6 min, rinsed in ddH_2O (1 min), dehydrated in 100% ethanol (1 min), and

air-dried. DAPI solution (1:1,000 in PBS) was added onto the slides for 5–10 minutes. The slides were washed with PBS and ddH_2O. The slides were mounted with antifade medium and imaged immediately.

CUT&TAG experiments

Gastruloid-derived cells were used for the CUT&TAG Assay with duplicates per sample. The experiments were strictly performed according to the manufacturer's instructions for the Hyperactive Universal CUT&Tag Kit (Vazyme, cat. no. TD904-C2).

Reagents and buffer preparation.

10× Enhancer Buffer and 10× Termination Buffer were used as supplied by the manufacturer. 5× Permeabilization Buffer was diluted to 1× with nuclease-free ddH_2O (200 µL 5× stock + 800 µL water). NE Buffer was prepared fresh as 20 mM HEPES-KOH pH 7.5, 0.5 mM spermidine, 10 mM KCl, 0.1 % (v/v) Triton X-100, 10 % (v/v) glycerol, and 1 mM PMSF added immediately before use. ConA Binding Buffer contained 20 mM HEPES-KOH pH 7.5, 10 mM KCl, 1 mM $CaCl_2$, and 1 mM $MnCl_2$. Wash Buffer was 20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM spermidine, 0.1 % (w/v) BSA, supplemented with 1× protease inhibitor (one cOmplete™ EDTA-free protease inhibitor tablet dissolved in 1 mL ddH_2O as a 50× stock; used 20 µL per 1 mL of buffer). Blocking Buffer was Wash Buffer supplemented with 2 mM EDTA (from 0.5 M EDTA, pH 8.0). Digitonin was prepared as a 5 % (w/v) stock in DMSO; it was contact-toxic and was handled with appropriate personal protective equipment. Antibody Buffer was Wash Buffer containing 0.05 % digitonin (add 0.5 µL of 5 % digitonin per 50 µL Wash Buffer) and was pre-chilled on ice. Dig-wash Buffer was Wash Buffer containing 0.05 % digitonin (add 8 µL of 5 % digitonin per 792 µL Wash Buffer) and was prepared fresh. Dig-300 Buffer contained 20 mM HEPES-KOH pH 7.5, 300 mM NaCl, 0.5 mM spermidine, 0.1 % BSA, 0.05 % digitonin, and 1× protease inhibitor; it was made by combining 100 µL 10× Dig-300 Buffer, 2 µL 5 % digitonin, 20 µL 50× protease inhibitor, and 878 µL ddH_2O. 1× Tagmentation Buffer was Dig-300 Buffer supplemented with 10 mM $MgCl_2$ (add 1/100 volume of 1 M $MgCl_2$). Stop Buffer was assembled per sample immediately before use by mixing 2.25 µL 0.5 M EDTA, 2.75 µL 10 % SDS, and 0.5 µL 20 mg/mL Proteinase K.

Cell preparation and treatment

Gastruloid-derived cells were harvested, counted, centrifuged at 600 × g for 5 min at room temperature, and washed once with PBS. For whole-cell TF enhancer treatment, 1e5 cells were resuspended in 100 µL 1× Enhancer Buffer (10 µL 10× Enhancer Buffer + 90 µL PBS) and incubated for 10 min at room temperature. 10 µL 10× Termination Buffer was added, mixed by inversion, and incubated for another 10 min at room temperature. Cells were pelleted (600 × g, 5 min), the supernatant discarded, and the cells resuspended in 100 µL 1× Permeabilization Buffer. Cell mixtures were then incubated at 62 °C for 10 min, after which 900 µL 1× Permeabilization Buffer was added, mixed by inversion, and incubated at 37 °C for

30 min. Cells were then centrifuged, the supernatant discarded, and the pellet resuspended in 100 µL Wash Buffer. Alternatively, for nuclei extraction, the washed cell pellet was resuspended in 100 µL ice-cold NE Buffer, pipetted gently, and incubated on ice for 10 min. The cells were centrifuged at 600 × g for 5 min, the supernatant discarded, the cells washed once with 500 µL PBS, and the nuclei resuspended in 100 µL Wash Buffer. Both preparations yielded a suspension ready for ConA bead binding and were processed immediately.

Concanavalin A bead activation and cell binding

For each sample, 100 µL ConA Binding Buffer was added to a well of an 8-tube strip. The ConA Beads were thoroughly resuspended and transferred into the binding buffer. The tubes were mixed gently, placed on a magnetic separation rack until the solution cleared (~2 min), and the supernatants were discarded. The tubes were removed from the magnet, and then 100 µL ConA Binding Buffer was added, and the samples were gently resuspended without vortexing. After that, 100 µL cell/nuclei suspension was transferred to the tube containing the activated beads, mixed by inversion, and incubated at room temperature for 10 min. The tubes were briefly centrifuged at <100 × g to collect the liquid, then placed on the magnet until clear, and the supernatant was carefully removed.

Blocking, primary and secondary antibody incubations

Bead–cell complexes were resuspended in 100 µL Blocking Buffer and the samples incubated at room temperature for 5 min. The tubes were then placed on the magnet, the supernatant discarded, and the cell/nuclei suspension resuspended in 50 µL pre-chilled Antibody Buffer. The primary antibody solutions were then added at the recommended dilution (1:100). The samples were mixed by inversion, briefly centrifuged, and incubated overnight at 4 °C with gentle rotation. The next day, the strips were placed on the magnet, the supernatant was removed, and 50 µL of secondary antibody diluted 1:100 in Dig-wash Buffer was added. The samples were mixed by inversion and incubated at room temperature for 30–60 min with rotation. Following incubation, the strips were placed on a magnet, the supernatant was discarded, and the beads were washed three times with 200 µL Dig-wash Buffer. The used primary antibodies were Rabbit polyclonal anti-Histone H3K4Me3 (Abcam, ab8580); H3K9me3 antibodies (Abcam, ab176916); H3K27me3 antibodies (Abcam, ab192985); H3K27ac antibodies (Abcam, ab4729); Rabbit anti-RNA polymerase II (Abcam, ab26721); Rabbit polyclonal anti-AFF3 (Sino Biological, 205373-T02).

pA/G-Tnp transposome binding and optional treatment

The pA/G-Tnp Pro (protein A/G-transposase fusion pre-loaded with adaptors) was diluted to 0.04 µM in Dig-300 Buffer. The supernatants were removed, and then 100 µL of diluted transposome was added; the samples were mixed by inversion and incubated at room temperature with rotation for 1 h. After incubation, the strips were placed on a magnet, the supernatant was discarded, and the beads were washed three times with 200 µL Dig-300 Buffer.

Tagmentation and DNA release

Washed beads were resuspended in 50 µL of 1× Tagmentation Buffer; the samples were mixed thoroughly and incubated at 37 °C for 60 min in a thermal cycler with the heated lid off to prevent excessive evaporation. To stop the reaction and digest proteins, 5.5 µL of freshly prepared Stop Buffer was added directly to the mixture. The samples were mixed by inversion and incubated at 55 °C for 30 min, then at 70 °C for 20 min to inactivate Proteinase K. The strips were briefly centrifuged and placed on a magnetic rack for 2–3 min. The supernatant (~55 µL) was carefully transferred to a clean 1.5 mL tube, avoiding the bead pellet.

DNA purification with SPRI beads

0.9 volumes (approximately 50 µL) of VAHTS DNA Clean Beads (or equivalent SPRI beads) were added to the supernatant. Samples were thoroughly mixed by pipetting or vortexing and incubated at room temperature for 5 min. The tube was placed on a magnetic rack until the solution cleared (~5 min); then the supernatant was carefully removed and discarded without disturbing the beads. The beads were washed twice with 200 µL of freshly prepared 80 % ethanol. The beads were air-dried at room temperature for 3–5 min until no liquid remained and the pellet was apparent. The tube was removed from the magnet. The DNA was then eluted by adding 22 µL nuclease-free ddH_2O, pipetting and mixing 10 times, and the DNA solution was incubated at room temperature for 5 min. 15 µL was used for library amplification, and the remainder was stored at –20 °C.

Library amplification and final cleanup

A 50 µL PCR reaction was prepared containing 15 µL purified DNA, 25 µL 2× CAM (CUT&Tag amplification master mix), and 2.5 µL each of 10 µM i5 and i7 indexing primers (N5XX and N7XX). The following program was used: 72 °C for 3 min (strand displacement), 95 °C for 3 min, then 16 cycles of 98 °C for 10 s, 60 °C for 5 s, and 72 °C for 1 min, with a final hold at 4 °C. The amplified library was purified by adding 100 µL VAHTS DNA Clean Beads (2× volume), mixing, and incubating for 5 min at room temperature. The library was washed twice with 200 µL 80 % ethanol as described above, the beads were air-dried, and the target DNA was eluted in 22 µL ddH_2O. 20 µL was transferred to a fresh tube. The library concentration was quantified with a fluorometric assay (e.g., Qubit), and fragment size distribution was evaluated on a Bioanalyzer or TapeStation.

CUT&TAG data analysis

Raw paired-end sequencing reads were first subjected to quality control and adapter trimming with fastp (v.0.24.0) using default settings; an optional FastQC report was generated to document read quality before and after trimming. The cleaned reads were then aligned to a suitable reference genome (GRCh19/hg19) using Bowtie2 (v.2.5.4) with parameters --very-sensitive -X 2000 --no-mixed --no-discordant --no-unal. These settings enforced concordant paired-end mapping, extended the maximum insert size to 2 kb to accommodate

nucleosomal fragments, and discarded unpaired or discordant alignments. PCR duplicates were removed from the resulting BAM file with Picard MarkDuplicates (v.3.1.0). The final filtered BAM was sorted and indexed with SAMtools.

Normalized coverage tracks were generated from each deduplicated BAM with deepTools bamCoverage (v.3.5.6) using spike-in normalization, a bin size of 10 bp, and the --centerReads option to center the signal on the inferred fragment midpoint. For every experimental group containing duplicates, the individual BigWig files were averaged with bigwigAverage (same bin size) to produce a consensus track. To visualize differential binding, the control consensus signal was subtracted from the treatment consensus signal using bigwigCompare at a 10-bp bin size; the resulting differential BigWig was viewed directly in IGV.

Peak calling was performed independently on each replicate using MACS2 (v.2.2.9.1). The parameters were -q 0.01, --keep-dup auto, --call-summits, and -g hs. A set of high-confidence consensus peaks was then derived by keeping only those regions identified in duplicates through sequential bedtools intersect operations. This conservative strategy minimized false positives and yielded a robust peak set for further annotation. Genomic annotation of the consensus peaks was performed with the R package ChIPseeker (v.1.42.1) together with the corresponding TxDb annotation package (TxDb.Hsapiens.UCSC.hg38.knownGene for human). Promoters were defined as the interval ±3 kb around the transcription start site.

Motif enrichment analysis was carried out on the consensus peaks using HOMER (v.5.1). A 400-bp window spanning –300 to +100 bp relative to the peak summit was extracted, and the findMotifsGenome.pl script was run with the -len 8,10,12 option to scan for motifs of 8, 10 and 12 nucleotides.

RNA fluorescence in situ hybridization (RNA FISH)

Cell preparation

Day 8 gastruloids were collected from 6-well plates into 15-mL tubes and washed once with 1× PBS. Samples were dissociated in 1 mL Accumax per tube by pipetting 10 times, followed by incubation for 5 min at 37 °C. Cells were pipetted 10 additional times and incubated for a further 5 min at 37 °C. When no visible clusters remained, dissociation was quenched with M10 medium. Cells were pelleted at 500 × g for 5 min at room temperature (RT), the supernatant was removed, and cells were resuspended in 1 mL sorting buffer (2% FBS in PBS). Cells were stained with DAPI and live single cells were isolated by FACS.

Cytospin

Sorted cells were deposited onto glass slides by cytocentrifugation (10,000 cells per slide; 800 × g, 3 min). Slides were air-dried in a fume hood.

Slide pretreatment

A hydrophobic barrier was drawn around each cell spot using a wax pen. Cells were fixed with 150 µL 4% paraformaldehyde (Beyotime, P0099) for 5 min, then washed twice with 1× DEPC-treated PBS (2 min each). Endogenous peroxidase activity was quenched with 150 µL 0.3% H_2O_2 in 1× DEPC PBS for 10 min at RT, protected from light. Slides were rinsed in a staining box containing 15 mL 1× DEPC PBS for 3 min, and this rinse was repeated once. Cells were treated with 150 µL 0.2 M HCl for 5 min. Slides were washed once in 1× DEPC PBS (2 min), then incubated with 150 µL of 1:8,000 diluted sample pre-treatment solution pre-warmed to 37 °C, for 30–60 s. Slides were washed twice with 1× DEPC PBS (2 min each).

Probe hybridization and amplification

An RNA FISH kit (KunYu Biotech, RFT003-30) from KunYu Company was used. Hybridization buffer I was pre-warmed to 40 °C (141 µL per sample). Hybridization probe I (9 µL per sample) was denatured at 85 °C for 2 min, placed on ice, then combined with hybridization buffer I and pre-warmed to 40 °C. Residual DEPC PBS was removed from sample slides and dried with Kimwipe. A FixoGum barrier was applied, and 150 µL hybridization mix was added per sample. Slides were incubated overnight at 40 °C in a water bath. Secondary and tertiary hybridizations were performed sequentially, with pre-warmed wash buffer (45 °C; three washes, 5 min each) between steps, using hybridization buffers II–III with denatured probes II–III (cat) (85 °C, 2 min) followed by incubation at 40 °C for 1 h each.

Quaternary hybridization, fluorogenic development, and imaging

Three quaternary hybridization rounds were performed with buffers IV–VI (40 °C; 150 µL per sample) and probes IV#1–IV#3 (KunYu Biotech, RFP00247T3, RFP00287T2, RFP00288T4), with incubations of 1 h (IV#1) or 50 min (IV#2–IV#3), each separated by three 45 °C wash-buffer washes. After quaternary rounds #1–#2, fluorogenic buffer was applied (150 µL; 15 min at RT in the dark; dilution 1:1,000 for #1 and 1:3,000 for #2), followed by 3 times of wash-buffer washes and blocking buffer (cat) incubation (150 µL; 15 min at 40 °C, dark). After quaternary round #3, fluorogenic buffer #3 (1:3,000) was applied (150 µL; 15 min, RT, dark). Slides were counterstained with DAPI (150 µL; 5 min, RT, dark), rinsed once with 1× DEPC PBS and 1× DEPC H_2O, mounted with Fluoromount-G (Southern Biotech, 0100-35), sealed, and imaged.

Cerebral organoid formation

hEPSCs were maintained for 2–3 days and then dissociated into single cells. Cells were centrifuged at 300 × g for 3 min, and after removal of the supernatant, the pellet was resuspended in Organoid Formation Medium supplemented with 10 µM Y-27632. The cell concentration was adjusted to approximately 9 × 10^4 cells/mL, and 100 µL of the suspension was seeded into each well of a U-bottom ultra-low-attachment 96-well plate, corresponding to 9,000 cells per well. Plates were incubated at 37°C for 5 days, with 100 µL of fresh Organoid Formation Medium added every other day.

On day 5, organoid morphology was assessed by microscopy, after which organoids were transferred to a low-attachment 24-well plate for induction. One to two organoids were placed into each well containing 1 mL Induction Medium and cultured at 37°C for 2 days. On day 7, Matrigel was thawed on ice for 1 h, and no more than 16 organoids were embedded at a time. Following polymerization of the Matrigel droplets at 37°C for 45 min, embedded organoids were transferred to a 6-well ultra-low-attachment plate containing Expansion Medium and cultured for 3 days.

On day 10, Expansion Medium was completely removed and replaced with 3 mL Maturation Medium per well. Plates were then placed on an orbital shaker in a 37°C incubator. Maturation Medium was changed every 2 days, depending on organoid morphology and growth rate. Organoids were collected for analysis by day 40 and processed for flow cytometry or immunolabelling.

Animal procedures for cerebral organoid implantation on brain

All animal experiments complied with ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines. Studies were also carried out in accordance with the Guidance on the operation of the Animals (Scientific Procedures) Act 1986 and associated guidelines and the NIH (National Research Council) Guide for the Care and Use of Laboratory Animals. All animal studies were conducted under the animal research protocol No. 25-057, approved by the Committee on the Use of Live Animals in Teaching and Research (CULATR) at the University of Hong Kong. The studies adhered to the Animals (Control of Experiments) Ordinance (Hong Kong) and guidelines from the Centre for Comparative Medical Research (CCMR), Li Ka Shing Faculty of Medicine, The University of Hong Kong.

NSG (CCMR) mice (6–8 weeks old, male) were obtained from the Experimental Animal Centre of the University, had access to food and water ad libitum, and were maintained under pathogen-free conditions. All procedures were performed under anesthesia using intraperitoneal injections of a ketamine/xylazine mixture (100 mg/kg ketamine, AlfaMedic Ltd., 013004; 10 mg/kg xylazine, AlfaMedic Ltd., 013006). Anesthesia was confirmed by checking body reflexes, and ophthalmic ointment (Terramycin®, Pfizer) was applied to prevent corneal drying. The calvarial implantation sites on the mouse head were decontaminated with alternating applications of 70% alcohol and povidone-iodine

(Betadine®) using sterile cotton swabs. The animals received buprenorphine (0.05 mg/kg, ATCvet Code, QN02AE01,) 30-60 min before surgery and twice daily for three days thereafter. Meloxicam was provided in the drinking water (0.25mL in 200 mL drinking water; Troy Laboratories, Australia) for 3 days post-surgery, followed by 4 days post-surgery.

The detailed animal procedures are as follows (*57*). Hair in the surgical area was removed using a shaver and hair removal cream (Reckitt Benckiser, 276428). A midline incision was made over the calvarium, and a full-thickness flap was elevated. A round, critical-size calvarial defect 4 mm in diameter was created with a trepan burr under sterile saline irrigation with suction. The mice were randomly assigned to four experimental groups: (1) WT; (2) KO; (3) WT + X1; and (4) KO + X1. Organoids from each group were then implanted into the defect and covered with a biodegradable, in-house carboxymethyl chitosan/oxidized hyaluronic acid membrane (*58*). The surgical site was sutured with 5-0 nylon monofilament (Ethicon, Inc., 668G). After surgery, mice were placed in an intensive care unit (ICU) until they fully recovered, after which they were returned to their original housing location. 14 days after implantation, mice were euthanized by intraperitoneal overdose of pentobarbital (250 mg/kg; Dorminal®, AlfaMedic Ltd., 013003) for histological processing.

Harvested head tissues were preserved in 10% neutral buffered formalin (NBF) followed by 70% ethanol at 4 °C for paraffin sectioning. Tissues were first decalcified in a 12.5% EDTA solution under shaking for 3 days at room temperature. Following decalcification, all samples were dehydrated, embedded, sectioned, and stained with Harris Haematoxylin (Cancer Diagnostic Inc., SH3777) and eosin (Dako, CS701) (H&E). Immunofluorescence was performed using a sectioned series of slides, as described above.

Quantification and statistical analysis

To quantify the fluorescence intensity of SOX2, EOMES, and TUJ1 within the transplanted organoids, the whole brain sections containing the organoids were immunolabelled, counterstained with DAPI, and then imaged. The DAPI+ cells were detected using QuPath, and then the specific immunohistochemical markers were measured. Mean fluorescence intensity values in regions of interest (ROIs) were exported and used in the violin plot analysis for different groups. One-way ANOVA was used for the grouped analysis, and asterisks were added automatically by Prism9 with * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$.

Data and code availability

The processed data and files generated in this study can be found on the download page of GitHub: https://github.com/michaelwu12337/AFF3-XIST.

Acknowledgement

We appreciate Andy Chun Hang Chen and William Shu Biu Yeung, who kindly provided EM3 and EM4 stem cells for gastruloid generation. We appreciate CPOS HKU for sequencing, flow cytometry, and imaging services. We appreciate Hin Kwok's expert input on library preparation and barcoding. We appreciate Tiago Rito for critical reading of the manuscripts, and Asaf Ronel for manuscript editing. We appreciate colleagues from the Centre for Translational Stem Cell Biology for constructive feedback and technical support.

Funding

We appreciate funding from RGC GRF 17115325, RGC/NSFC JRS, and InnoHK.

Author contribution

YX, GW, HTL, AT, QW, and ANML conducted gastruloid and cerebral organoid experiments. YX and GW generated AFF3 KO and AFF3/XIST DKO human EPSCs. YX, XC, DR, and PL assisted with gastruloid experiment design. CK, CL, ZH, SC, FL, and SJL assisted with cerebral organoid experiments. YX, YC, XJ, YJ, and JL assisted with computational analysis of gastruloids, cerebral organoids, and the AFF3 patient record. KF assisted with CUT&Tag experiments. ZL, CL, and RS supervised the project.

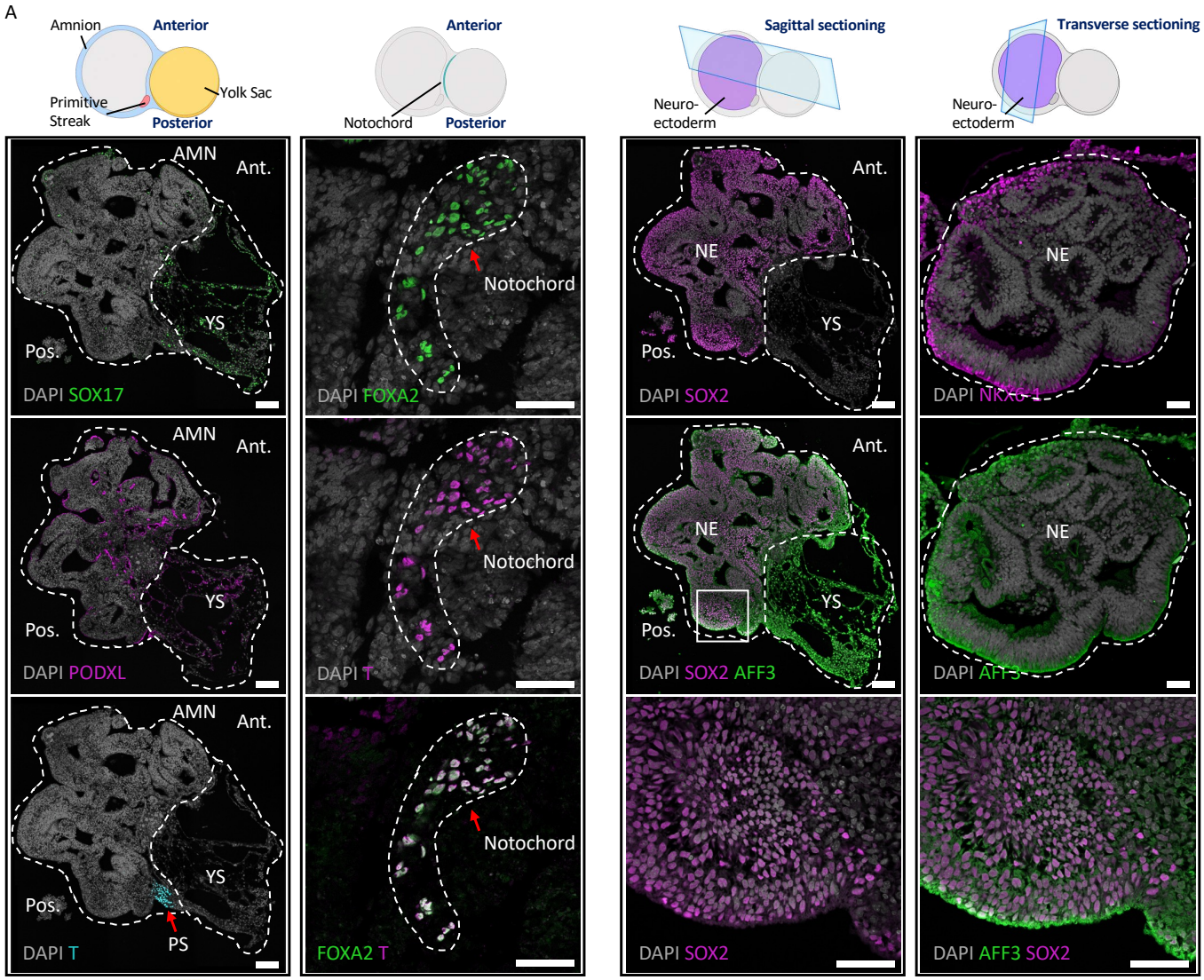
Ethics Declaration

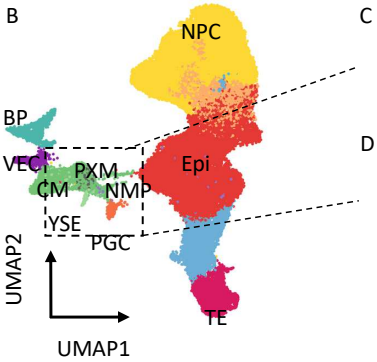
The study involving human EPSC cell lines was approved by the Institutional Review Boards (IRB) of HKU (UW 25-535). Consent to Participate declaration: not applicable.

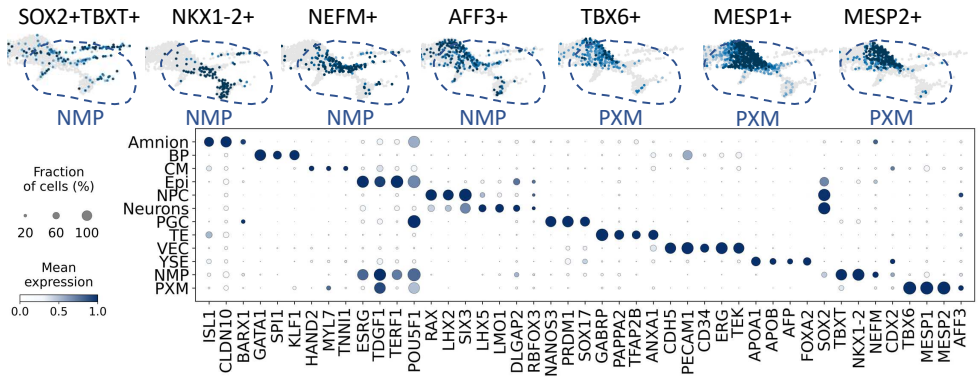
Competing interests

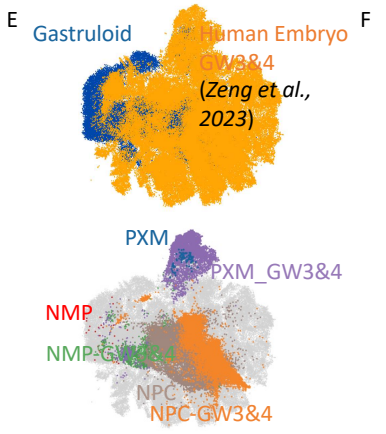
The authors declare no competing interests.

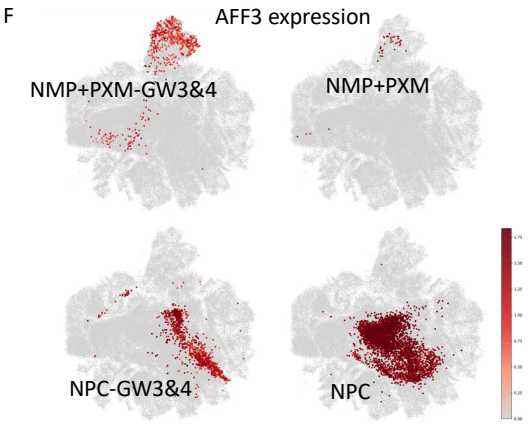
Figure 1


A
Amnion
Anterior
Primitive Streak
Yolk Sac
Posterior
Notochord
Sagittal sectioning
Neuro-ectoderm
Transverse sectioning
AMN
Ant.
Pos.
YS
PS
NE
Notochord
DAPI SOX17
DAPI FOXA2
DAPI SOX2
DAPI NKX6-1
DAPI PODXL
DAPI T
DAPI SOX2 AFF3
DAPI AFF3
FOXA2 T
DAPI AFF3 SOX2


B
NPC
BP
VEC
PXM
CM
NMP
Epi
YSE
PGC
TE
UMAP2
UMAP1


C
SOX2+TBXT+
NKX1-2+
NEFM+
AFF3+
TBX6+
MESP1+
MESP2+
NMP
PXM
D
Fraction of cells (%)
20 60 100
Mean expression
0.0 0.5 1.0
Amnion
BP
CM
Epi
NPC
Neurons
PGC
TE
VEC
YSE
NMP
PXM
ISL1 CLDN10 BARX1 GATA1 SPI1 KLF1 HAND2 MYL7 TNNI1 ESRG TDGF1 TERF1 POU5F1 RAX LHX2 SIX3 LHX5 LMO1 DLGAP2 RBFOX3 NANOS3 PRDM1 SOX17 GABRP PAPPA2 TFAP2B ANXA1 CDH5 PECAM1 CD34 ERG TEK APOA1 APOB AFP FOXA2 SOX2 TBXT NKX1-2 NEFM CDX2 TBX6 MESP1 MESP2 AFF3


E
Gastruloid
Human Embryo GW3&4
(Zeng et al., 2023)
PXM
PXM_GW3&4
NMP
NMP-GW3&4
NPC
NPC-GW3&4


F
AFF3 expression
NMP+PXM-GW3&4
NMP+PXM
NPC-GW3&4
NPC

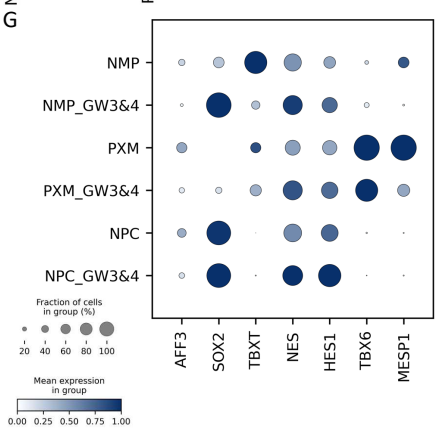

G
NMP
NMP_GW3&4
PXM
PXM_GW3&4
NPC
NPC_GW3&4
Fraction of cells in group (%)
20 40 60 80 100
Mean expression in group
0.00 0.25 0.50 0.75 1.00
AFF3 SOX2 TBXT NES HES1 TBX6 MESP1

Fig. 1. AFF3 is enriched in NPC and NMP in human embryos.

(**A**) Cartoon of Day 8 gastruloids and Immunofluorescence analysis of gastruloids showing SOX17+ yolk sac (YS) endoderm, PODXL+ amnion-like layer (AMN) covering the neuroectoderm, T+ primitive streak (PS) cells, and FOXA2+ T+ notochord-like structure. Both sagittal and transverse sections show AFF3 expression in SOX2+ and NKX6-1+ neuroectoderm (NE). Scale bar = 100 um. (**B**) UMAP showing cell populations in gastruloids. (**C**) Enlarged view showing neuromesodermal progenitor (NMP) with SOX2, TBXT, NKX1-2, NFEM, CDX2, and AFF3 expression. NMP derivative paraxial mesoderm (PXM) markers TBX6, MESP1, and MESP2 are detected around the NMP area. (**D**) Dot plot showing marker genes of each population. AFF3 is enriched in both neural progenitor cell (NPC) and NMP, as well as PXM. (**E**) Integrated UMAP visualisation of single-cell transcriptomes from gastruloids and human embryos at Gestation Week (GW) 3&4 (*14*). NPC, NMP, and PXM populations across gastruloids and human embryos at GW 3&4. (**F**) AFF3 expression in NPC, NMP, and PXM from panel E. (**G**) Gene dot plot showing marker genes to annotate NPC, NMP, and PXM in gastruloids and human embryos. BP, blood progenitor; CM, cardiac mesoderm; Epi, epiblast; NMP, neuromesodermal progenitor; NPC, neural progenitor cell; PGC, primordial germ cell; PXM, paraxial mesoderm; TE, trophectoderm; VEC, vascular endothelial cell; YSE, yolk sac endoderm.

Figure 2

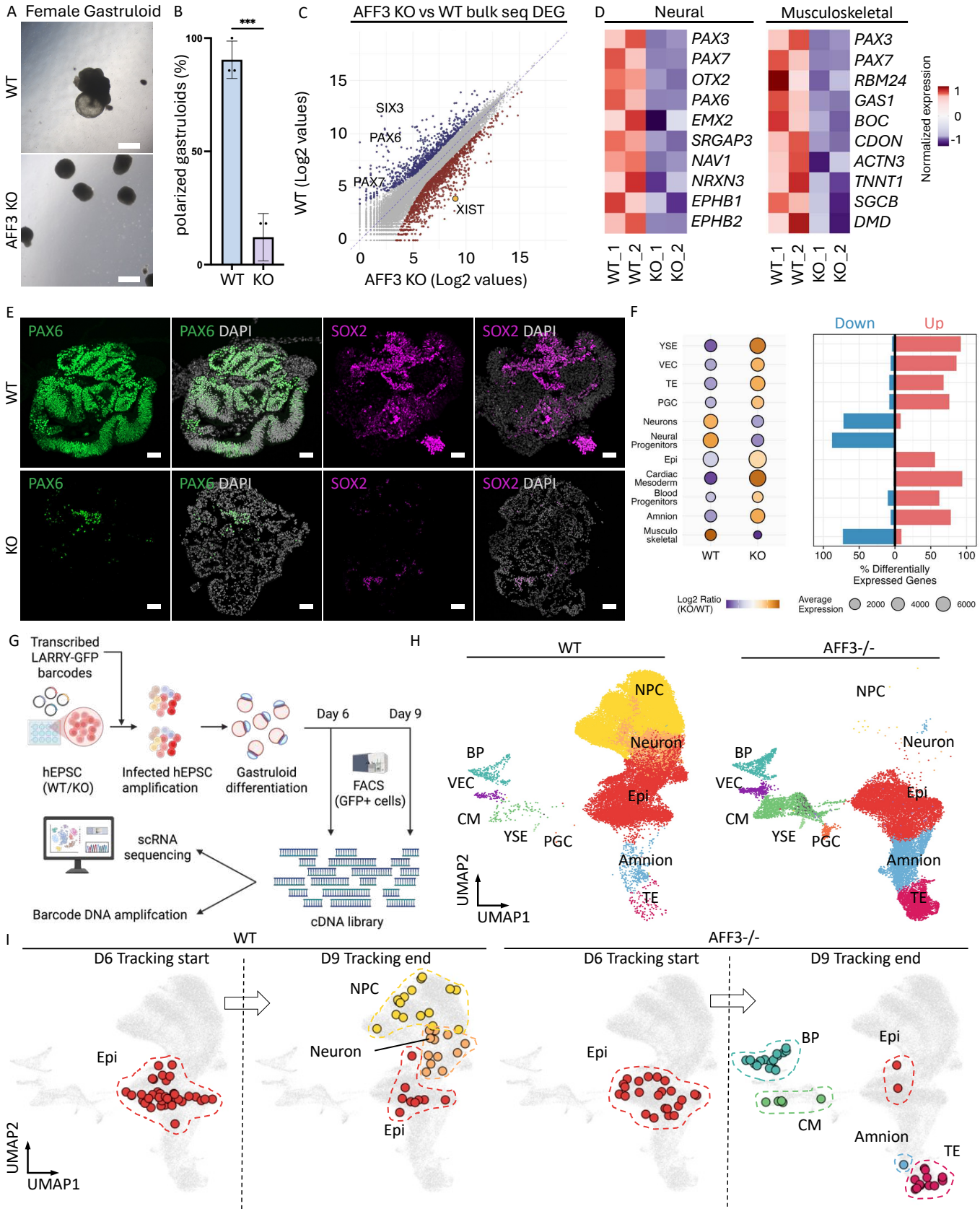

A Female Gastruloid
WT
AFF3 KO
B
***
polarized gastruloids (%)
100
50
0
WT KO
C
AFF3 KO vs WT bulk seq DEG
WT (Log2 values)
AFF3 KO (Log2 values)
SIX3
PAX6
PAX7
XIST
D
Neural
Musculoskeletal
PAX3
PAX7
OTX2
PAX6
EMX2
SRGAP3
NAV1
NRXN3
EPHB1
EPHB2
RBM24
GAS1
BOC
CDON
ACTN3
TNNT1
SGCB
DMD
WT_1 WT_2 KO_1 KO_2
Normalized expression
1
0
-1
E
PAX6
PAX6 DAPI
SOX2
SOX2 DAPI
WT
KO
F
YSE
VEC
TE
PGC
Neurons
Neural Progenitors
Epi
Cardiac Mesoderm
Blood Progenitors
Amnion
Musculo skeletal
WT KO
Down Up
% Differentially Expressed Genes
Log2 Ratio (KO/WT)
Average Expression
2000 4000 6000
G
Transcribed LARRY-GFP barcodes
hEPSC (WT/KO)
Infected hEPSC amplification
Gastruloid differentiation
Day 6 Day 9
FACS (GFP+ cells)
scRNA sequencing
Barcode DNA amplifcation
cDNA library
H
WT
AFF3-/-
NPC
Neuron
Epi
BP
VEC
CM
YSE
PGC
Amnion
TE
UMAP2
UMAP1
I
WT
AFF3-/-
D6 Tracking start
D9 Tracking end
Epi
NPC
Neuron
BP
CM
Amnion
TE
UMAP2
UMAP1

Fig. 2. AFF3 loss impairs neural and musculoskeletal development in female gastruloids.

(**A**) Representative brightfield image of female AFF3 WT/KO gastruloids. Scale bar: 500 um. (**B**) Polarized gastruloids between female AFF3 WT/KO. (**C**) Scatterplot of RNA-seq data showing DEGs between AFF3 WT and AFF3 KO in female gastruloids. Representative genes are highlighted. (**D**) Heatmaps of bulk RNA-seq datasets from female AFF3 WT/KO show that neural and musculoskeletal genes were reduced. (**E**) Immunofluorescence of neural progenitor markers PAX6 (green) and SOX2 (magenta) is reduced in AFF3 KO. (**F**) Proportion of up-regulated or down-regulated cell types specific genes from bulk RNA-seq. Dot plot showing the absolute abundance (dot size) and relative expression (dot color) of cluster-specific genes. Bar plots depicting the percentage of up- and down-regulated cluster-specific genes. Bar length corresponds to the proportion of genes. Neurons, neural progenitors, and musculoskeletal populations are reduced in female AFF3 KO. (**G**) Schematic of single-cell RNA sequencing and LARRY lineage-barcoding analysis. Female hEPSCs were infected with LARRY lentivirus, cultured for 48 hours to allow clonal expansion, split into two replicates, and used for gastruloid formation. Gastruloids were harvested at day 6 and day 9 for single-cell transcriptomic and lineage-tracing analyses. GFP marks LARRY-infected cells in gastruloids. GFP+ cells were sorted for downstream analysis. (**H**) UMAP showing NPC and neuron defects in AFF3 KO, while CM, PGC, VEC, and YSE are enriched. (**I**) LARRY lineage analysis showing epiblast-derived lineage trajectories. In wild-type gastruloids, epiblast cells contribute to neural progenitor and neuronal lineages, whereas in AFF3 KO gastruloids, epiblast-derived clones preferentially contribute to blood progenitor (BP), cardiac mesoderm (CM), and trophectoderm (TE) lineages. BP, blood progenitor; CM, cardiac mesoderm; Epi, epiblast; NPC, neural progenitor cell; PGC, primordial germ cell; TE, trophectoderm; VEC, vascular endothelial cell; YSE, yolk sac endoderm.

Figure 3

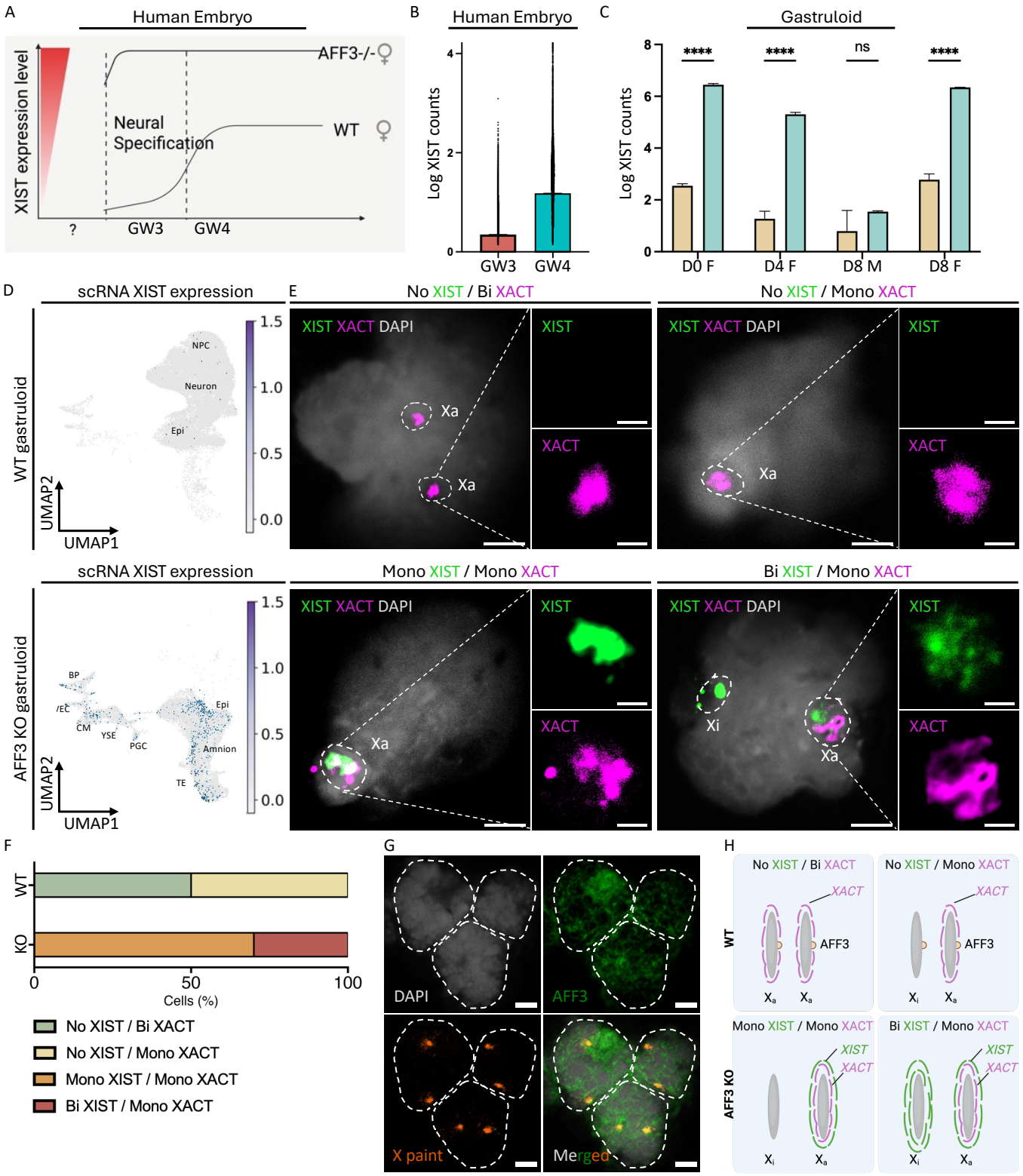

A
Human Embryo
XIST expression level
AFF3-/-
WT
Neural Specification
?
GW3
GW4
B
Human Embryo
Log XIST counts
GW3
GW4
C
Gastruloid
Log XIST counts
****
****
ns
****
D0 F
D4 F
D8 M
D8 F
D
scRNA XIST expression
WT gastruloid
NPC
Neuron
Epi
UMAP2
UMAP1
scRNA XIST expression
AFF3 KO gastruloid
BP
/EC
CM
YSE
PGC
Epi
Amnion
TE
E
No XIST / Bi XACT
No XIST / Mono XACT
Mono XIST / Mono XACT
Bi XIST / Mono XACT
XIST XACT DAPI
XIST
XACT
Xa
Xi
F
WT
KO
Cells (%)
No XIST / Bi XACT
No XIST / Mono XACT
Mono XIST / Mono XACT
Bi XIST / Mono XACT
G
DAPI
AFF3
X paint
Merged
H
WT
AFF3 KO
XACT
XIST
AFF3

Fig. 3. Premature and excessive upregulation of XIST upon AFF3 loss.

(**A**) Model of how AFF3 loss dysregulates XIST expression. XIST expression is repressed during neural specification at week (W) 3 embryo, then upregulated from W4 in WT. KO induced premature and excessive XIST expression. W3 corresponds to Carnegie stages 7 to 9, and W4 corresponds to Carnegie stages 10 to 13. (**B**) Violin plots of XIST in the W3 and W4 human embryo datasets. XIST is upregulated post-gastrulation. XIST expression during human embryogenesis based on single-cell RNA-sequencing data from human embryos (*14*). (**C**) XIST expression levels at different developmental time points in both female WT/KO gastruloids from bulk RNA-seq datasets. XIST is prematurely and excessively upregulated in female AFF3 KO gastruloids. Male WT and AFF3 KO gastruloids at Day (D) 8 are shown as references. (**D**) XIST expression in female WT & KO gastruloid scRNA-seq. XIST expression is upregulated in KO. (**E**) RNA fluorescence in situ hybridisation for XIST in day 8 wild-type and AFF3 KO gastruloid cells. Scale bars indicate 5 µm (left one within the group) and 2 µm (right ones within the group). The number of XIST and XACT spots per nucleus is indicated above the image. (**F**) Percentage of cells from (**E**) with No XIST, Mono XIST, and Bi XIST cells in wild-type and AFF3 KO gastruloids. XACT spot number per nucleus (Mono, Bi) is indicated. n>100 cells per group. The data are representative of 3 independent experiments. (**G**) Immunofluorescence for AFF3 protein and DNA FISH for chrX-paint in AFF3 WT gastruloid-derived cells. Scale bars indicate 5 µm. (**H**) Model of AFF3 regulation of XIST expression. AFF3 binds to both X chromosomes and represses XIST. XaXa status in WT recapitulates the early embryonic development stage. One X chromosome loses XACT, but XIST is not expressed yet. KO nucleus expressed XIST in either monoallelic or biallelic manner. BP, blood progenitor; CM, cardiac mesoderm; Epi, epiblast; NPC, neural progenitor cell; PGC, primordial germ cell; TE, trophectoderm; VEC, vascular endothelial cell; YSE, yolk sac endoderm.

Figure 4

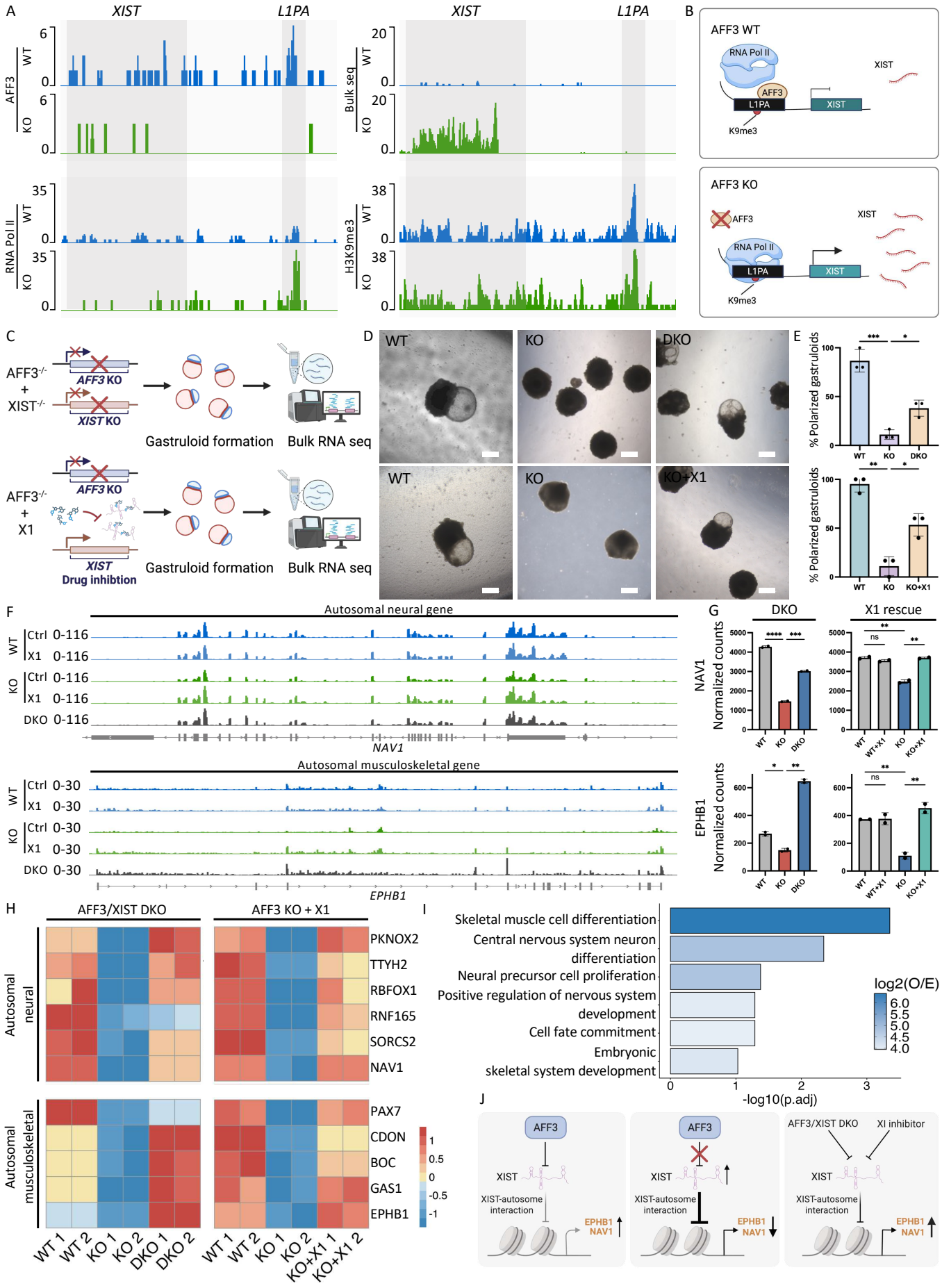

A
XIST
L1PA
AFF3
WT
KO
Bulk seq
RNA Pol II
H3K9me3
B
AFF3 WT
RNA Pol II
AFF3
L1PA
XIST
K9me3
AFF3 KO
C
AFF3-/- + XIST-/-
AFF3 KO
XIST KO
Gastruloid formation
Bulk RNA seq
AFF3-/- + X1
XIST Drug inhibtion
D
WT
KO
DKO
KO+X1
E
% Polarized gastruloids
F
Autosomal neural gene
Ctrl 0-116
X1 0-116
DKO 0-116
NAV1
Autosomal musculoskeletal gene
Ctrl 0-30
X1 0-30
DKO 0-30
EPHB1
G
DKO
X1 rescue
NAV1
EPHB1
Normalized counts
H
AFF3/XIST DKO
AFF3 KO + X1
Autosomal neural
Autosomal musculoskeletal
PKNOX2
TTYH2
RBFOX1
RNF165
SORCS2
NAV1
PAX7
CDON
BOC
GAS1
EPHB1
WT 1
WT 2
KO 1
KO 2
DKO 1
DKO 2
KO+X1 1
KO+X1 2
I
Skeletal muscle cell differentiation
Central nervous system neuron differentiation
Neural precursor cell proliferation
Positive regulation of nervous system development
Cell fate commitment
Embryonic skeletal system development
log2(O/E)
-log10(p.adj)
J
AFF3
XIST
XIST-autosome interaction
EPHB1
NAV1
AFF3/XIST DKO
XI inhibitor

Fig.4. Premature and excessive XIST represses autosomal neural and musculoskeletal development.

(**A**) IGV tracks showing AFF3 protein genomic binding, RNA Pol II genomic binding, XIST RNA, and H3K9Me3 signal in XIST locus from AFF3 WT/KO gastruloid-derived cells. AFF3 protein binds L1PA regulatory element. KO increased RNA Pol II binding at L1PA and XIST expression. H3K9me3 is enriched in both WT and KO. (**B**) Model of how AFF3 regulates XIST expression by binding to L1PA. AFF3 binds to L1PA and inhibits XIST expression. H3K9me3 is present in L1PA. When AFF3 is KO, RNA Pol II facilitates XIST expression. (**C**) Schematic of AFF3 KO / XIST KO (DKO) and AFF3 +X1 drug treatment workflow. (**D**) Representative brightfield image of AFF3 WT/KO/DKO and AFF3 WT/KO/KO+X1 gastruloids. DKO and X1 rescued morphology of KO gastruloids. Scale bar: 500um. (**E**) Quantification of gastruloids with polarization in WT, KO, DKO, and KO+X1. (**F**) IGV tracks showing NAV1 and EPHB1 expression in AFF3 WT, WT+X1, KO, KO+X1, and DKO gastruloid-derived cells. (**G**) Quantification of NAV1 and EPHB1 expression levels in WT, KO, DKO, WT+X1, and KO+X1 from bulk RNA-seq. DKO and X1 rescued the expression of NAV1 and EPHB1. (**H**) Heatmaps showing the expression patterns of rescued autosomal genes in DKO/KO+X1 compared with KO. (**I**) Gene Ontology analysis of rescued autosomal genes between KO and DKO gastruloids from panel H. Positive values indicate pathways upregulated in DKO and KO+X1 gastruloids. Neural and musculoskeletal genes are rescued. (**J**) Model of AFF3 KO, DKO, and KO+X1 treatment rescue autosomal XIST targets EPHB1 and NAV1 expression.

Figure 5

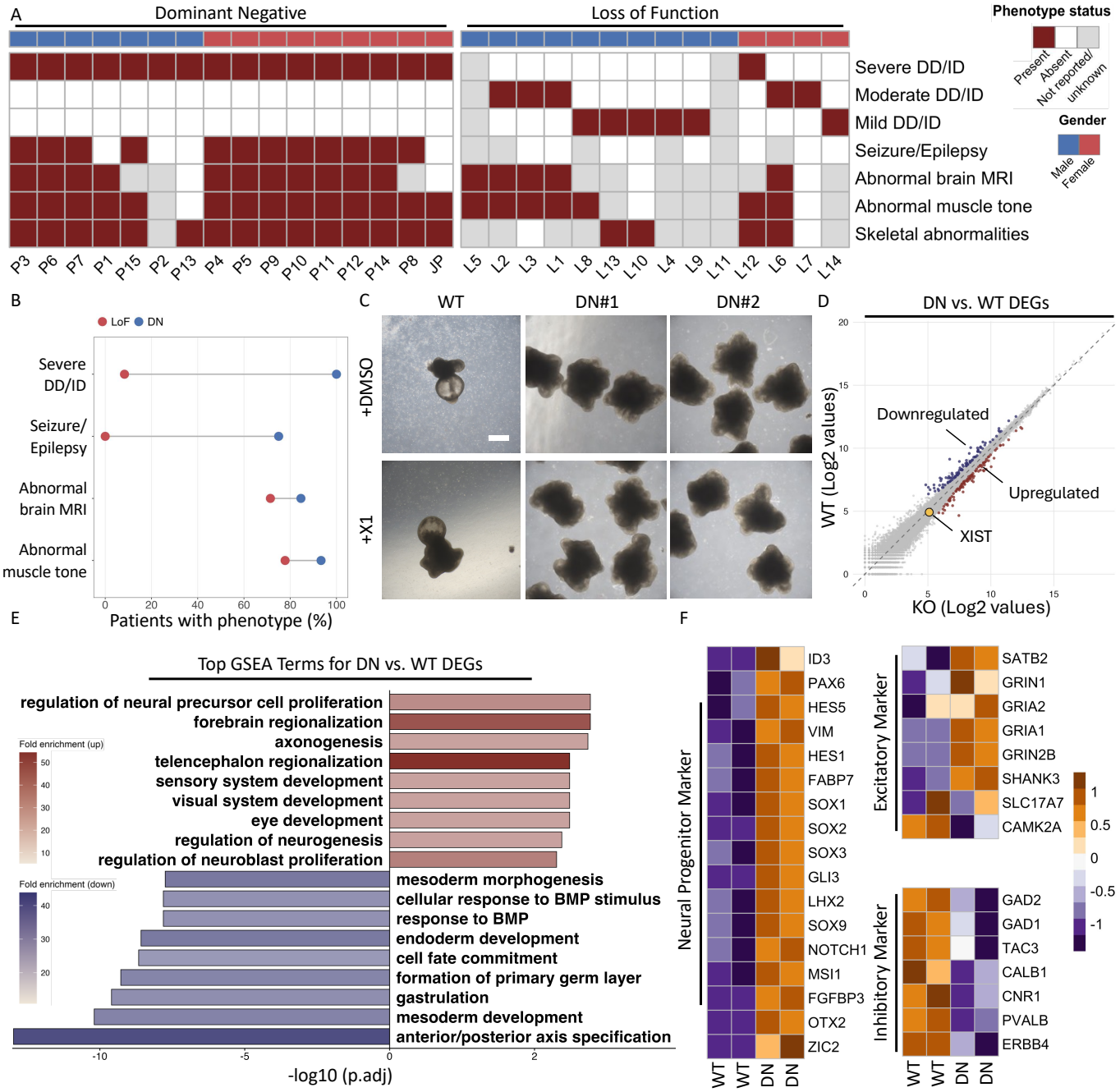

A
Dominant Negative
Loss of Function
Phenotype status
Present
Absent
Not reported/ unknown
Gender
Male
Female
Severe DD/ID
Moderate DD/ID
Mild DD/ID
Seizure/Epilepsy
Abnormal brain MRI
Abnormal muscle tone
Skeletal abnormalities
P3 P6 P7 P1 P15 P2 P13 P4 P5 P9 P10 P11 P12 P14 P8 JP
L5 L2 L3 L1 L8 L13 L10 L4 L9 L11 L12 L6 L7 L14
B
LoF
DN
Severe DD/ID
Seizure/ Epilepsy
Abnormal brain MRI
Abnormal muscle tone
Patients with phenotype (%)
C
WT
DN#1
DN#2
+DMSO
+X1
D
DN vs. WT DEGs
Downregulated
Upregulated
XIST
WT (Log2 values)
KO (Log2 values)
E
Top GSEA Terms for DN vs. WT DEGs
regulation of neural precursor cell proliferation
forebrain regionalization
axonogenesis
telencephalon regionalization
sensory system development
visual system development
eye development
regulation of neurogenesis
regulation of neuroblast proliferation
mesoderm morphogenesis
cellular response to BMP stimulus
response to BMP
endoderm development
cell fate commitment
formation of primary germ layer
gastrulation
mesoderm development
anterior/posterior axis specification
Fold enrichment (up)
Fold enrichment (down)
-log10 (p.adj)
F
Neural Progenitor Marker
ID3
PAX6
HES5
VIM
HES1
FABP7
SOX1
SOX2
SOX3
GLI3
LHX2
SOX9
NOTCH1
MSI1
FGFBP3
OTX2
ZIC2
Excitatory Marker
SATB2
GRIN1
GRIA2
GRIA1
GRIN2B
SHANK3
SLC17A7
CAMK2A
Inhibitory Marker
GAD2
GAD1
TAC3
CALB1
CNR1
PVALB
ERBB4
WT WT DN DN

Fig.5. Dominant-negative AFF3 results in dysregulated neurogenesis.

(**A**) Heatmaps showing clinical patients with heterozygous/homozygous/compound AFF3 dominant-negative (DN) mutant patients and loss-of-function (LOF) mutant patients. DD (developmental delay), ID (intellectual disability). Patient data were extracted from PMID: 38811945 (*7*). (**B**) The dumbbell plot shows that AFF3 dominant-negative (DN) mutant patients preferentially have severe DD/ID and Seizure/Epilepsy compare with LoF mutant patients. (**C**) Representative brightfield image of female AFF3 WT, DN#1, DN#2, and X1-treated gastruloids. Scale bar: 500 um. (**D**) Scatterplot of RNA-seq data showing DEGs between AFF3 WT and AFF3 DN in female gastruloids. Representative genes are highlighted. (**E**) GO analysis showing adjusted p-values for up-regulated/down-regulated DEGs related GO terms in AFF3 DN gastruloids compared with WT gastruloids. (**F**) Heatmaps showing the expression patterns of neural lineages including neural progenitors, general mature neurons, excitatory neurons, and inhibitory neurons in AFF3 DN gastruloids compared with WT gastruloids.

Figure 6

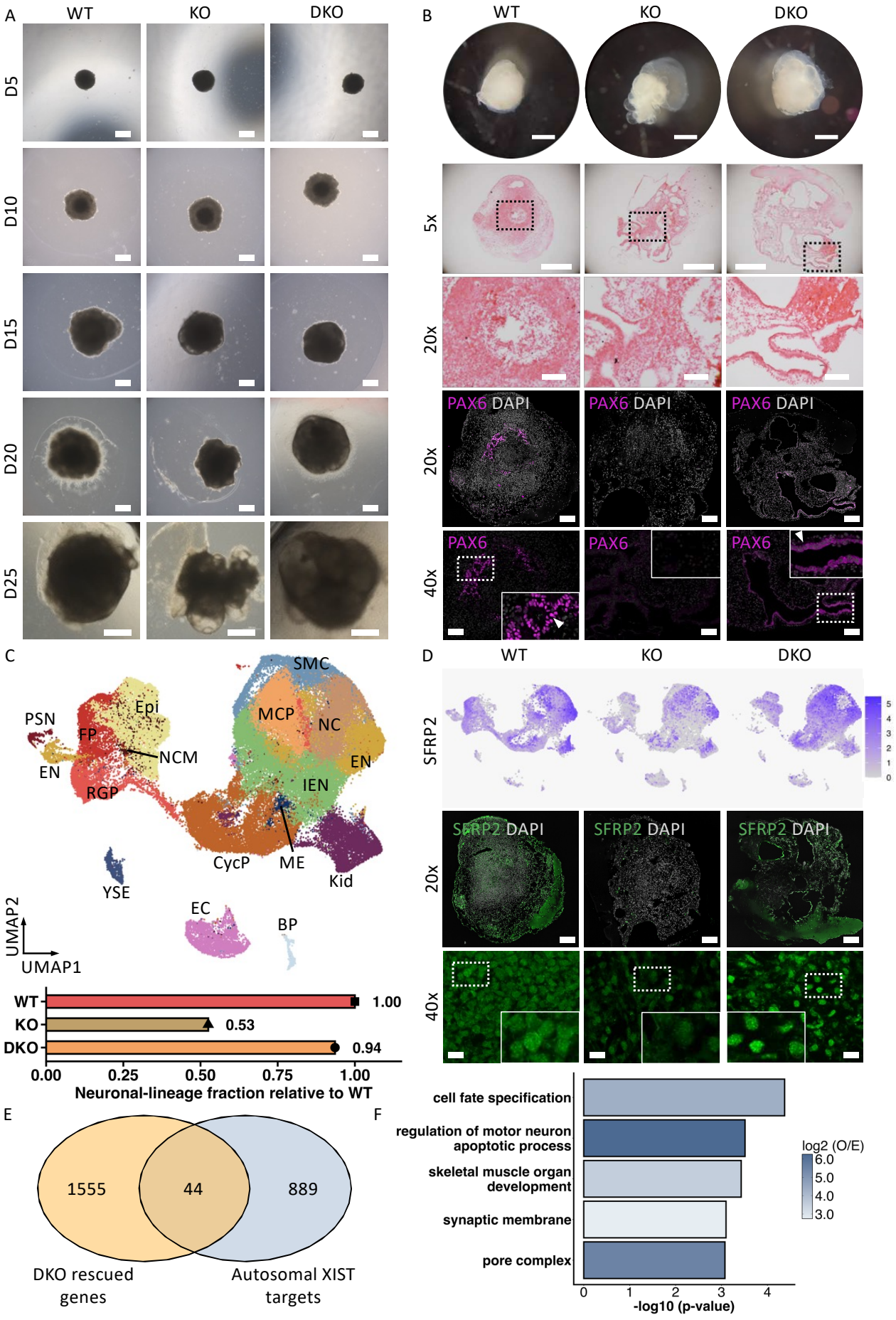

A
WT
KO
DKO
D5
D10
D15
D20
D25
B
WT
KO
DKO
5x
20x
PAX6 DAPI
PAX6
40x
C
SMC
Epi
PSN
FP
MCP
NC
NCM
EN
EN
RGP
IEN
CycP
ME
Kid
YSE
EC
BP
UMAP2
UMAP1
WT
KO
DKO
1.00
0.53
0.94
0.00
0.25
0.50
0.75
1.00
Neuronal-lineage fraction relative to WT
D
WT
KO
DKO
SFRP2
SFRP2 DAPI
20x
40x
E
1555
44
889
DKO rescued genes
Autosomal XIST targets
F
cell fate specification
regulation of motor neuron apoptotic process
skeletal muscle organ development
synaptic membrane
pore complex
log2 (O/E)
6.0
5.0
4.0
3.0
0
1
2
3
4
-log10 (p-value)

Fig. 6. XIST targeting rescues cerebral organoid formation.

(**A**) Representative morphology of wild-type, KO, and DKO cerebral organoids over 25 days of differentiation. KO organoids show impaired morphogenesis, which is rescued by DKO. (**B)** Stereoscopic images of day 25 cerebral organoids, with hematoxylin and eosin staining and immunofluorescence of organoid sections. KO organoids form sac-like structures. Expression of the neural progenitor marker PAX6 is reduced in KO organoids and restored in DKO organoids. (**C**) UMAP visualization of cell populations in day 25 wild-type, KO, and DKO cerebral organoids (top). Quantification of neuronal lineage cells across wild-type, KO, and DKO organoids (bottom). (**D**) UMAP plots and IF staining showing the expression patterns of the representatively rescued gene SFRP2. SFRP2 expression is reduced in KO, and is restored in DKO. Scale bar: 100 um. (**E**) Venn diagram showing overlap between genes rescued by DKO and previously reported autosomal XIST target genes (*15*). The 44 overlapping genes were analysed in (F). (**F**) Gene Ontology analysis of the 44 overlapping autosomal XIST target genes enriching neural and musculoskeletal development. BP: Blood progenitor. CycP: Cycling progenitor. EC: Endothelial cell. EN: Excitatory neuron. Epi: Epiblast. FP: Forebrain progenitor early neuron. Kid: Kidney. IEN: Immature deep layer excitatory neuron. MCP: Meningeal choroid plexus-like cells. ME: Meso-endoderm. NC: Neural crest. NCM: Neural crest melanocyte lineage. PSN: Peripheral sensory neuron. RGP: Radial glia apical progenitor. SMC: Smooth muscle cell. YSE: Yolk sac endoderm.

Figure 7

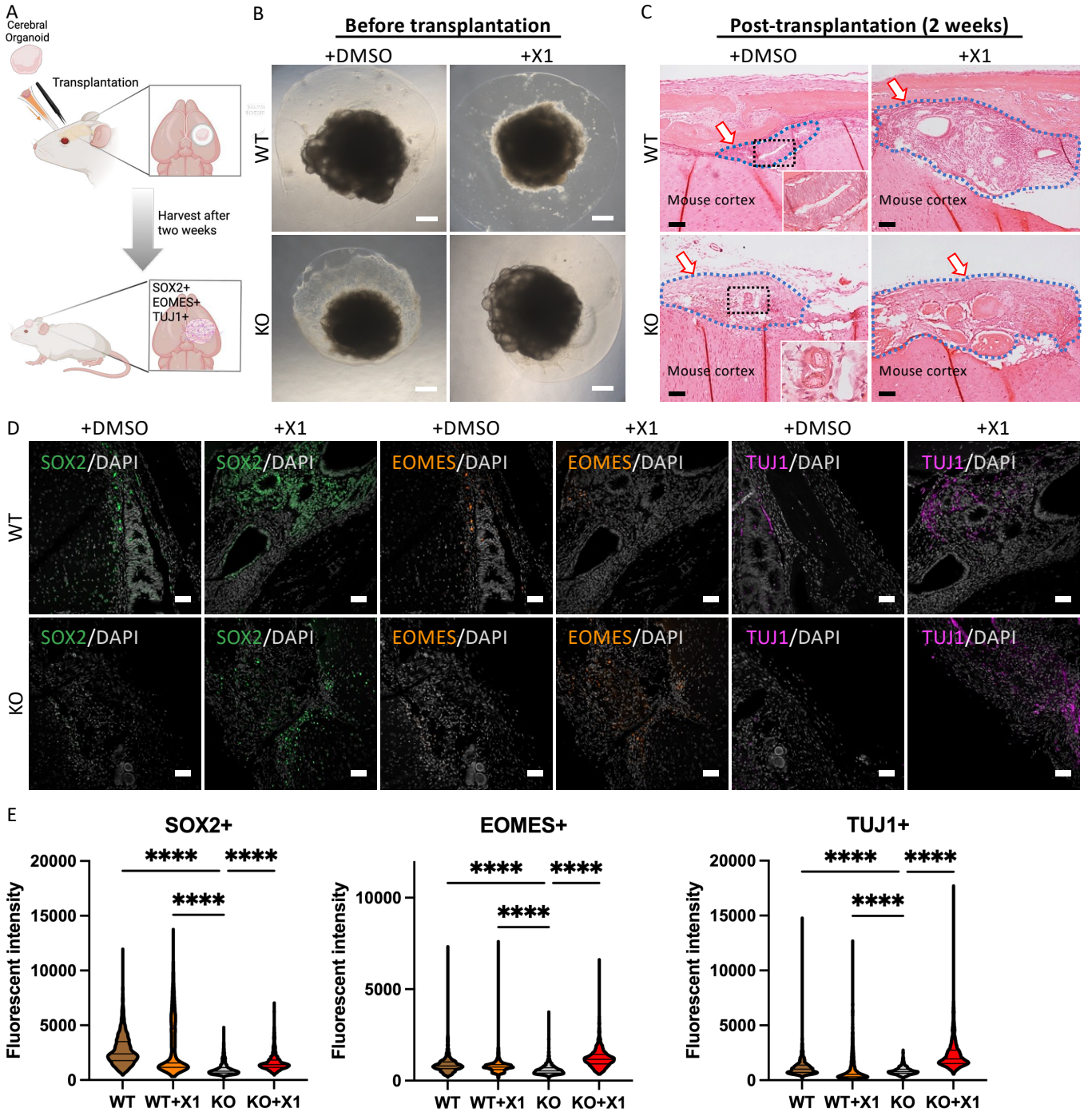
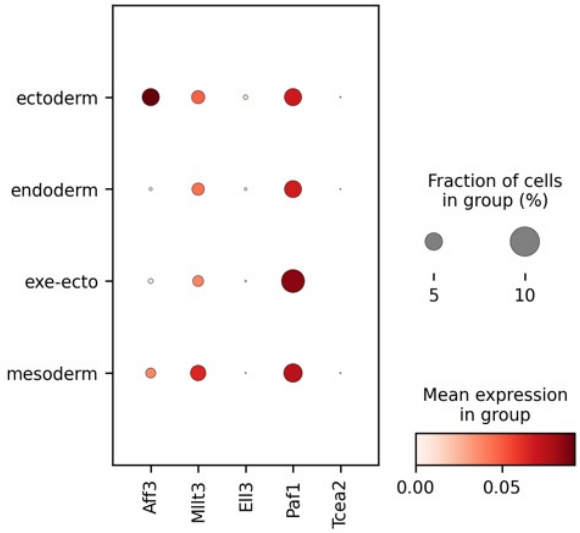

A
Cerebral Organoid
Transplantation
Harvest after two weeks
SOX2+
EOMES+
TUJ1+
B
Before transplantation
+DMSO
+X1
WT
KO
C
Post-transplantation (2 weeks)
+DMSO
+X1
WT
KO
Mouse cortex
D
+DMSO
+X1
+DMSO
+X1
+DMSO
+X1
WT
KO
SOX2/DAPI
EOMES/DAPI
TUJ1/DAPI
E
SOX2+
EOMES+
TUJ1+
Fluorescent intensity
20000
15000
10000
5000
0
****
WT
WT+X1
KO
KO+X1

Fig.7. XIST targeting rescues neurogenesis in vivo.

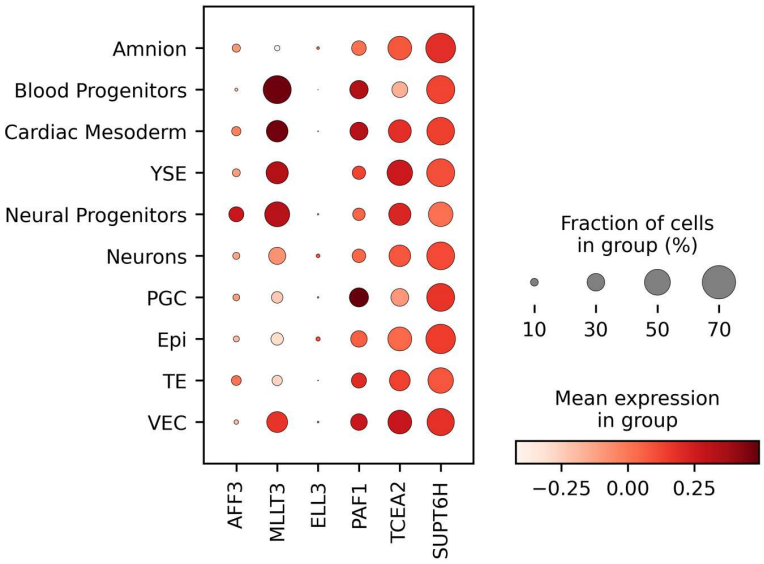
(**A**) Schematic of female AFF3 WT, WT+X1, KO and KO+X1 cerebral organoids in vivo transplantation workflow. (**B**) Representative brightfield images of female AFF3 WT/WT+X1/KO/KO+X1 cerebral organoids at different time points. KO organoid shows reduced size, which is rescued by X1. Scale bar: 500 um. (**C**) Representative H&E images of female AFF3 WT, WT+X1, KO and KO+X1 cerebral organoids after two weeks of transplantation. Scale bar: 100 um. Inset panels show cortex formation in WT and ossicle formation in KO. X1-treated KO organoid grows larger with cortex formation. (**D**) Representative immunofluorescence images of cerebral organoid sections. Grey, DAPI; Green, SOX2; Orange, EOMES; Magenta, TUJ1. KO reduced SOX2+ neural progenitor cells, EOMES+ intermediate neural progenitors, and TUJ1+ mature neurons, which are restored by X1. Scale bar: 50 um. (**E**) The violin plot shows the mean fluorescent intensity of each specific marker in all groups. One-way ANOVA was used for the grouped analysis. N=600~2400. Data are shown as the mean ± SEM. **P < 0.01, ***P < 0.001, and ****P < 0.0001.

Supplementary Materials-Sup Fig1

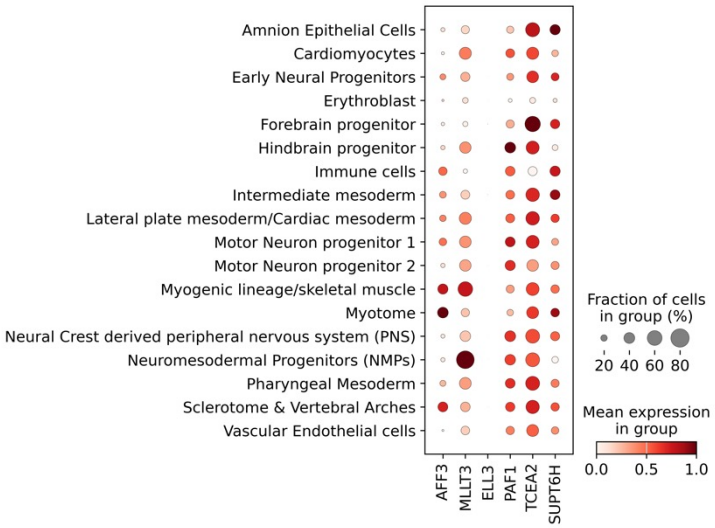
A.



B.



C.



D.

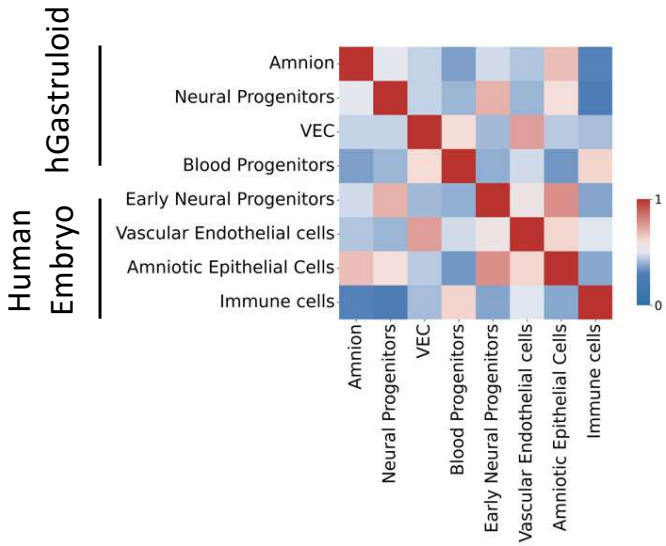
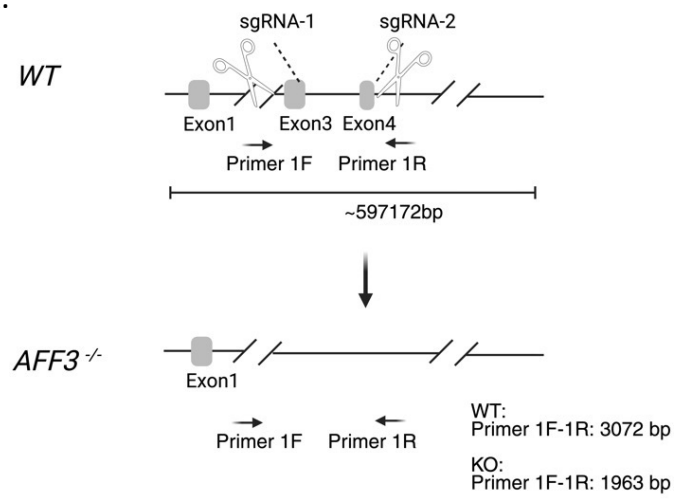


Figure S1. Super elongation factors are expressed in human, mouse embryos, and gastruloids.

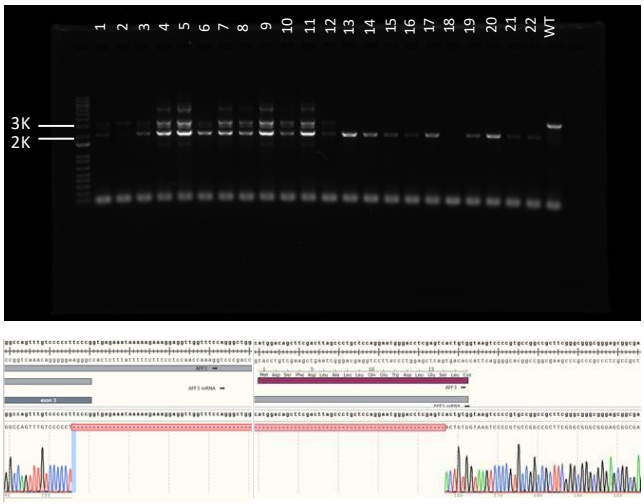
(A) The Dotplot analysis of super-elongation factors in mouse embryo E7.5 datasets. (B) The Dotplot analysis of super-elongation factors in human gastruloid datasets. (C) The Dotplot analysis of super-elongation factors in human embryo scRNA-seq datasets. (D) The heatmap showing the similarity of human embryo datasets to human gastruloids by Pearson correlation analysis.

Supplementary Materials-Sup Fig2

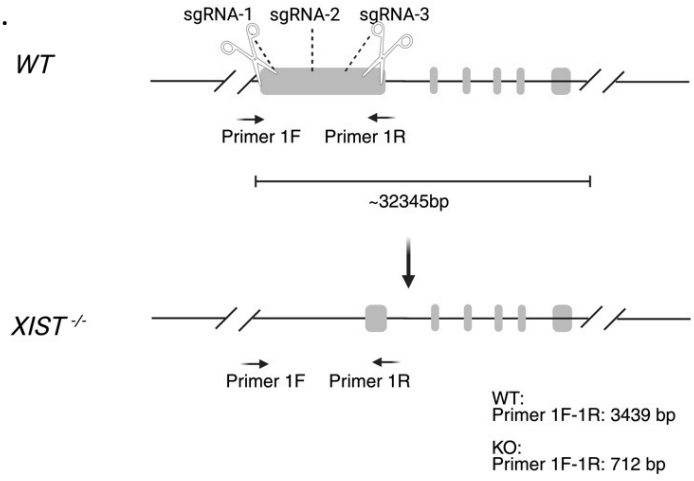
A.



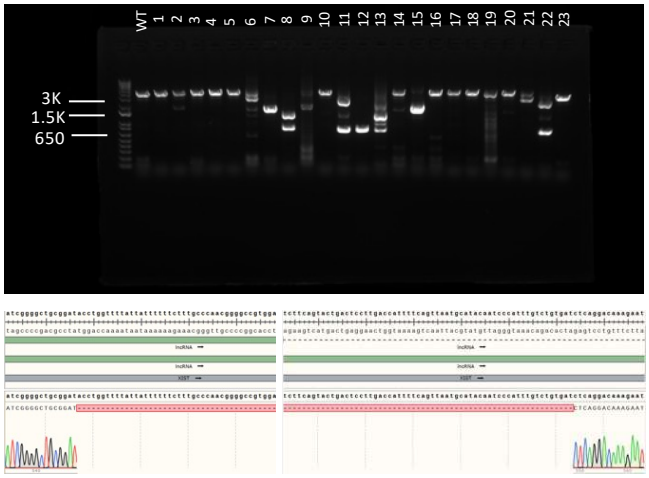
B.



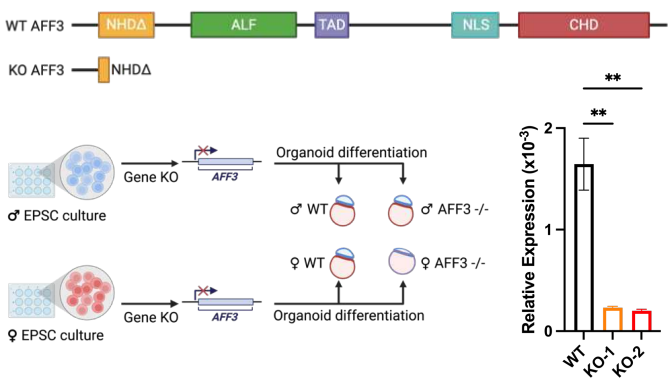
C.



D.



E.



F.

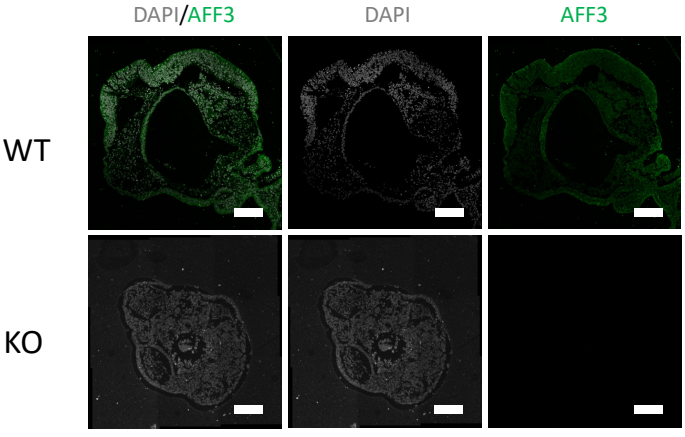
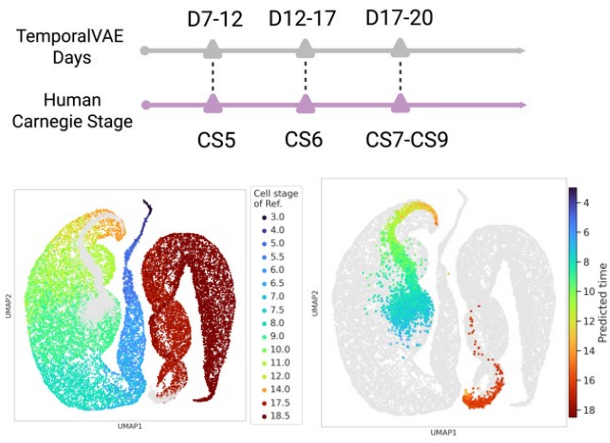


Figure S2. Characterisation of AFF3 KO, XIST KO hEPSCs.

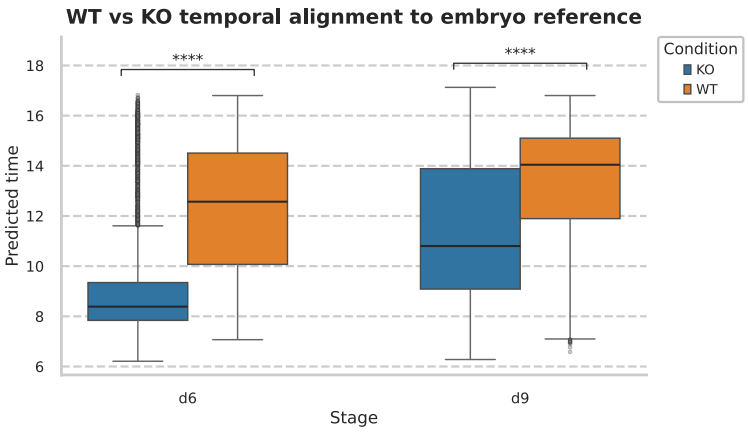
(A) Schematic representation of the AFF3 knockout strategy. Exons 3 and 4 are targeted. (B) Genomic PCR analyses using the primer sets flanking the cleavage sites verifying the genomic DNA deletion of AFF3 (Top) and Sanger sequencing of the two cut PCR fragments validating the genomic DNA deletion of AFF3 (Bottom). (C) Schematic representation of the XIST knockout strategy. Exon 1 is targeted. (D) Genomic PCR analyses using the primer sets flanking the cleavage sites verifying the genomic DNA deletion of XIST (Top) and Sanger sequencing of the two cut PCR fragments validating the genomic DNA deletion of XIST (Bottom). (E) Schematic representation of AFF3 WT/KO organoid generation and RT-qPCR analysis of AFF3 expression in AFF3 WT and two AFF3 KOs in Day6 gastruloids. (F) Immunostaining confirming that the AFF3 KO organoids lack AFF3 expression. Scale bar: 100 um.

Supplementary Materials-Sup Fig3

A.



B.



C.

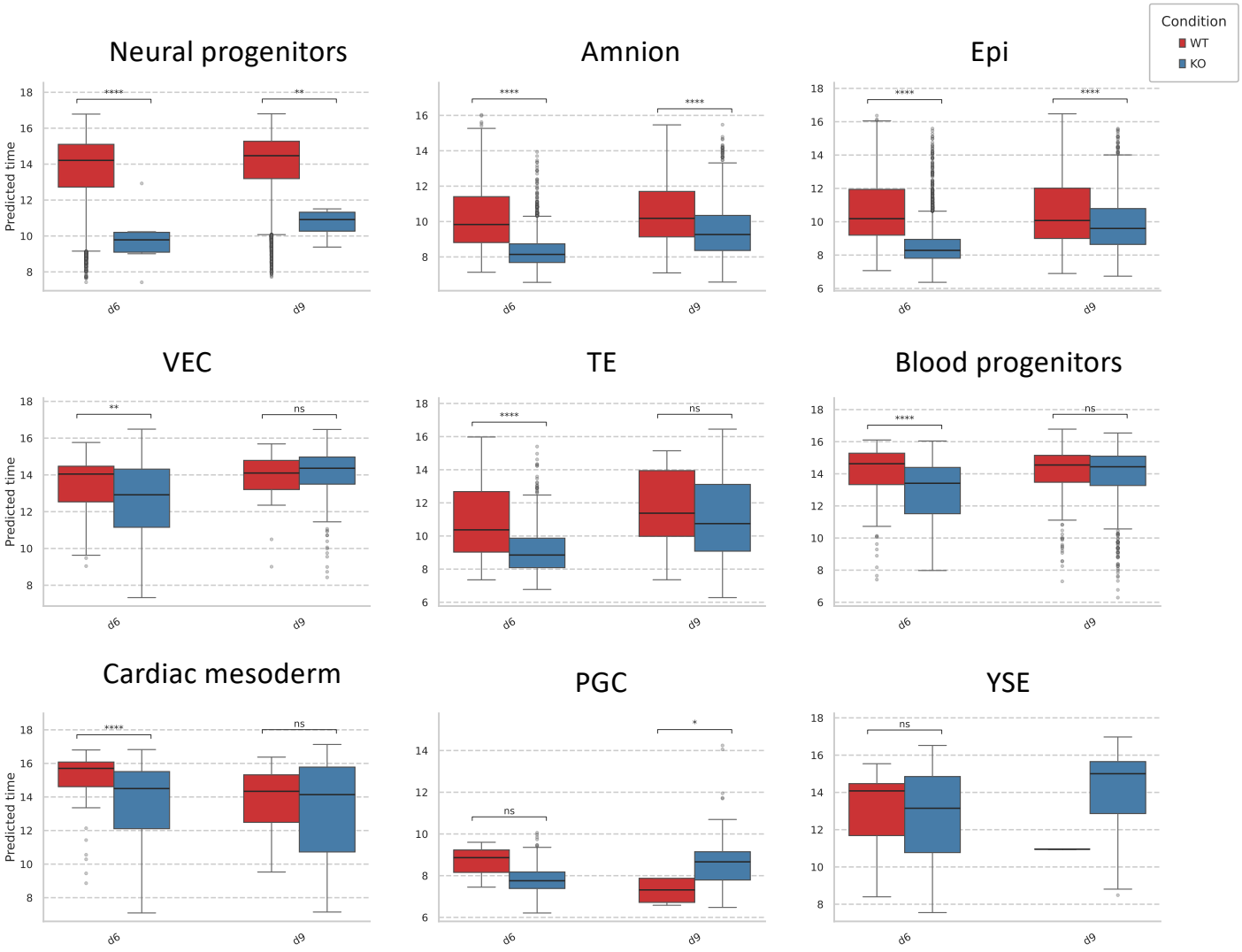
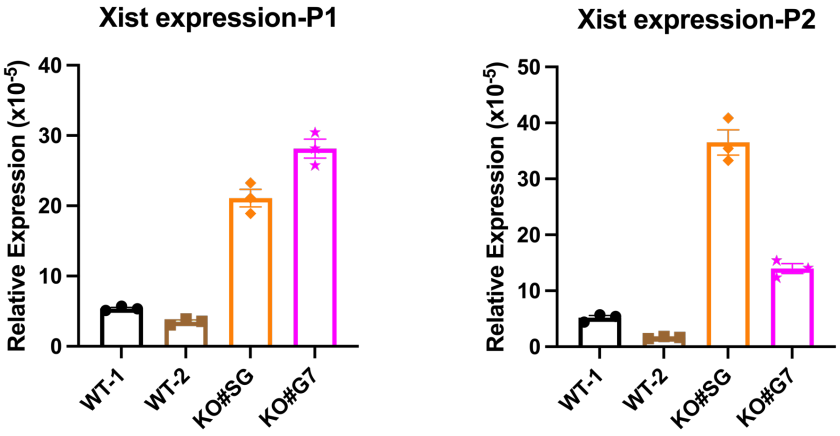


Figure S3. The temporalVAE analysis of human gastruloids.

(A) The conversion of TemporalVAE days to Human Carnegie Stage (top) and TemporalVAE analysis of human gastruloid dataset compared with previously published human early embryo development scRNA-seq datasets (bottom). Human gastruloid scRNA-seq data can be clustered with the embryo reference dataset. (B) The alignment of human WT/KO organoids to reference embryo datasets. WT samples were much more mature than KO on day 6 or 9. (C) The prediction time of specific cell clusters in human gastruloids. PSC: pluripotent stem cells; VEC: vascular endothelial cells; TE: trophectoderm; PGC: primordial germ cell; YSE: yolk sac endoderm. Comparison of celltype level, neural progenitors were mature in the WT (embryo day 14-19).

Supplementary Materials-Sup Fig4

A.



X-linked genes in gastruloids

B.

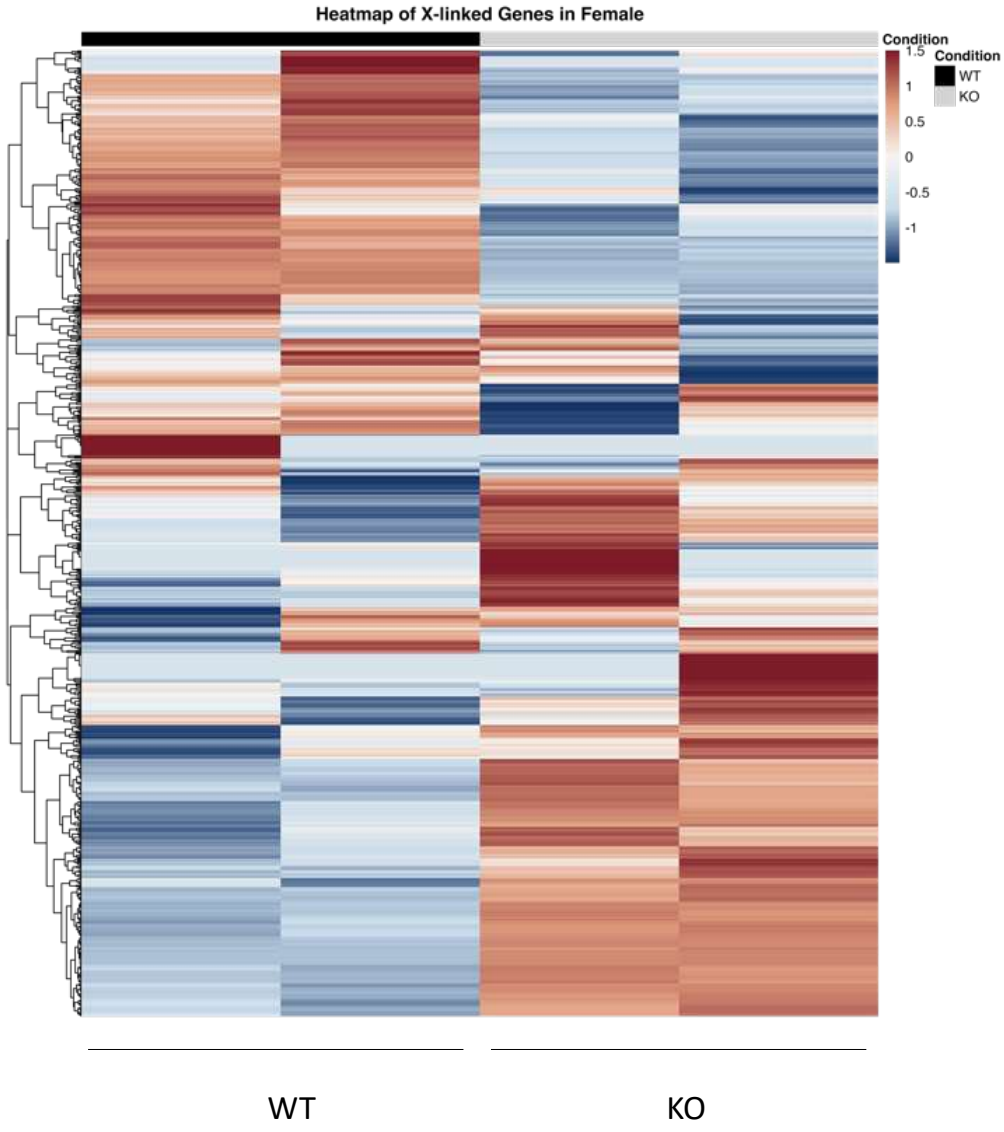
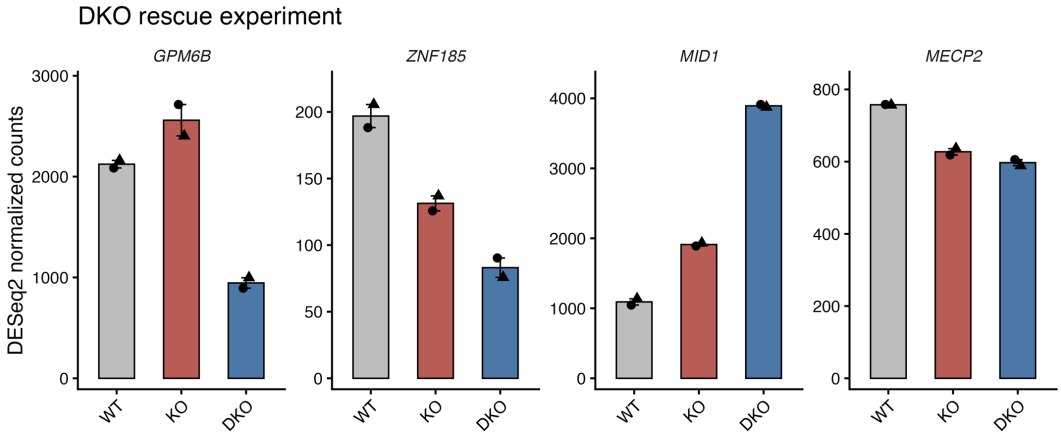


Figure S4. AFF3 KO upregulates XIST expression and dysregulates X-linked gene expression.

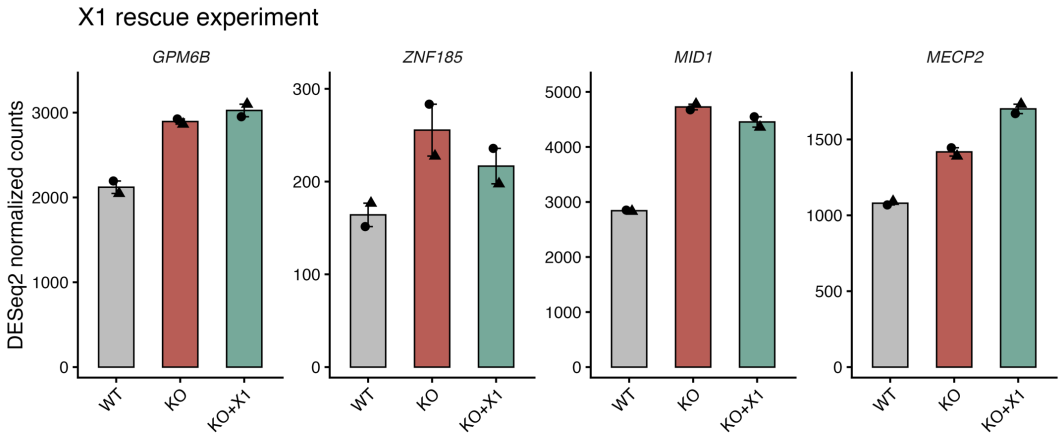
(A) The qPCR analysis of XIST expression in AFF3 WT and several AFF3 KOs. P1 and P2 stand for two sets of primers. (B) The heatmap showing the X-linked gene expression in WT/KO gastruloids.

Supplementary Materials-Sup Fig5

A.



B.



C.

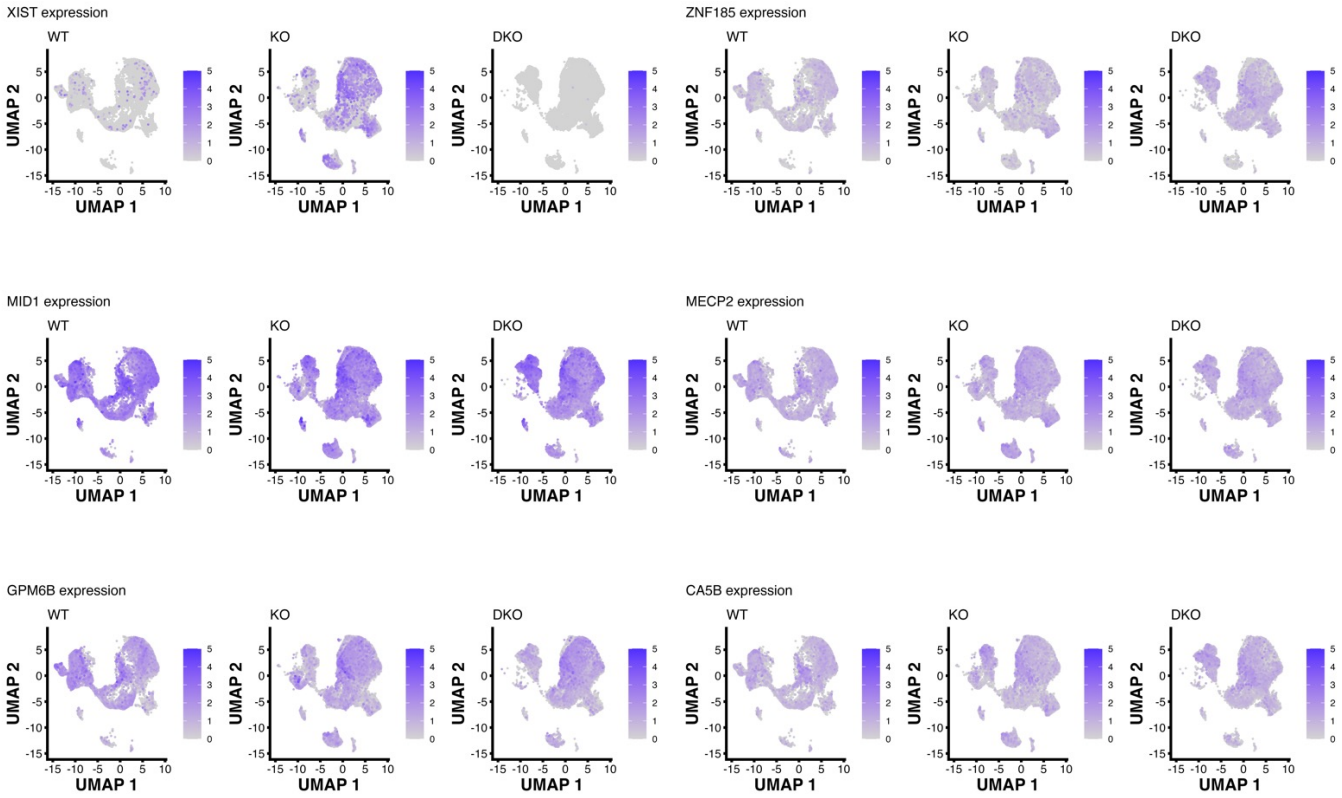


Figure S5. The X-linked neural genes ZNF185, MID1, MECP2, GPM6B, CA5B expression in gastruloids and cerebral organoids.

(A) The bar graph showing the GPM6B, ZNF185, MID1, and MECP2 expression in AFF3 WT/KO/DKO gastruloids. (B) The bar graph showing the GPM6B, ZNF185, MID1, MECP2 expression in AFF3 WT/KO/KO+X1 gastruloids. (C) the UMAP showing the XIST, ZNF185, MID1, MECP2, GPM6B, CA5B expression in AFF3 WT/KO/DKO cerebral organoids.

Supplementary Materials-Sup Fig6

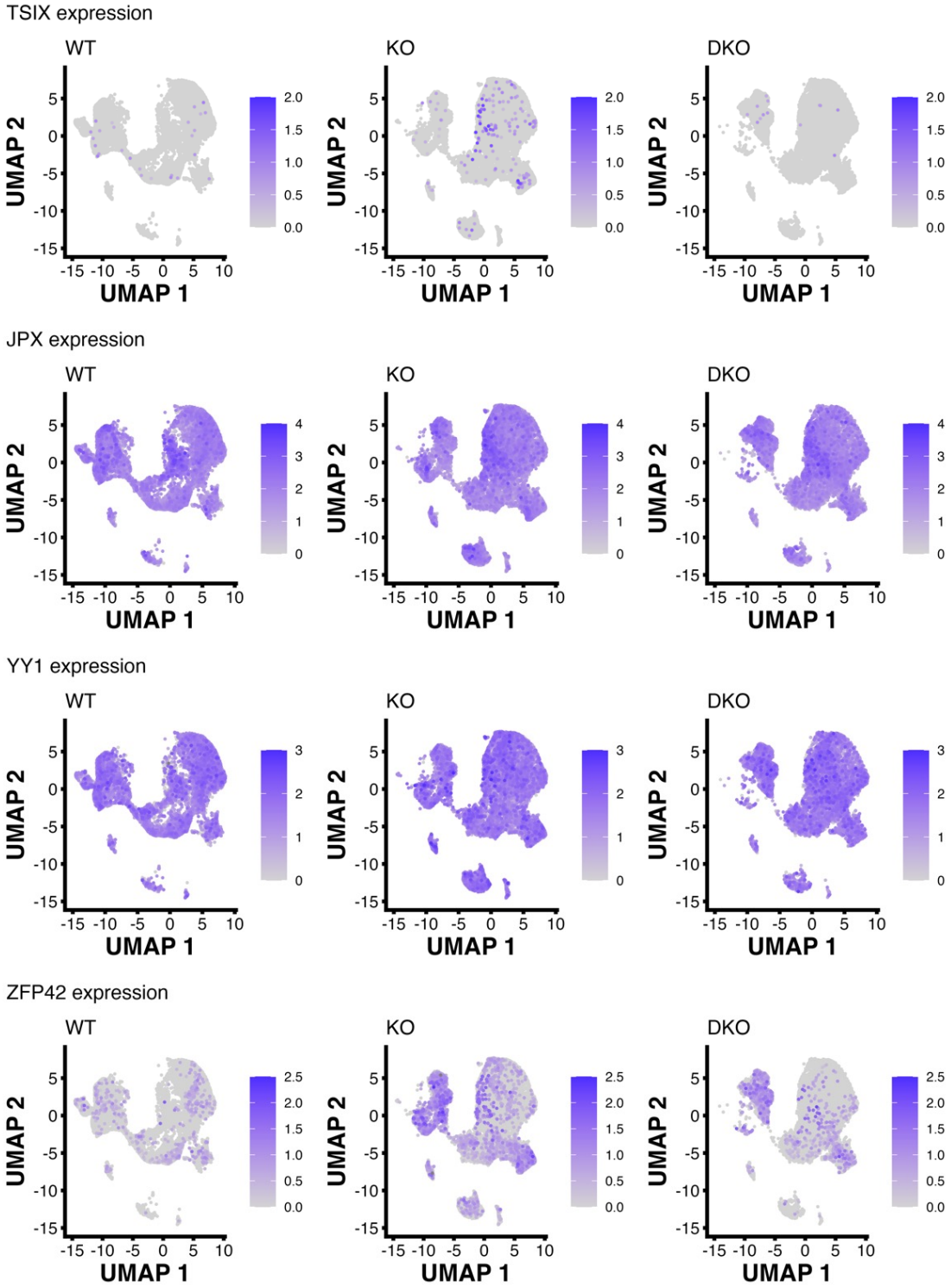
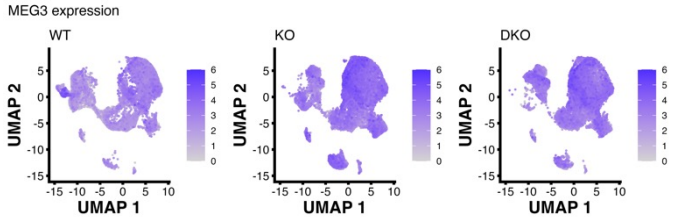


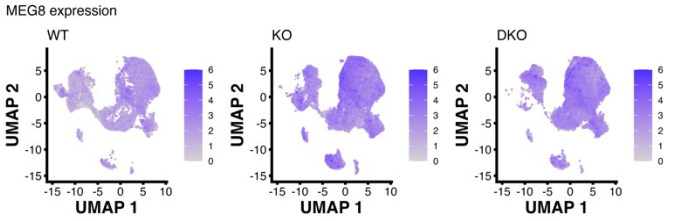
Figure S6. UMAP showing the TSIX, JPX, YY1, ZFP42 (REX1) expression in AFF3 WT/KO/DKO cerebral organoids.

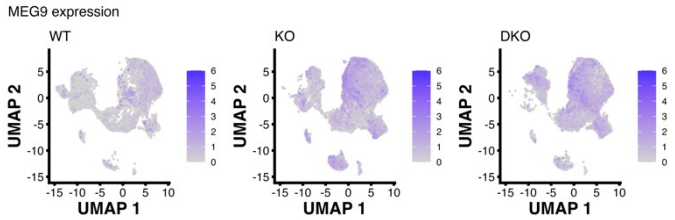
Supplementary Materials-Sup Fig7

A.

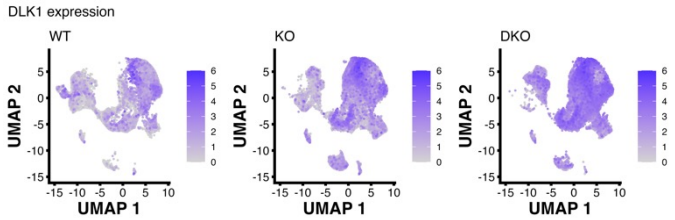
Maternal imprinting genes

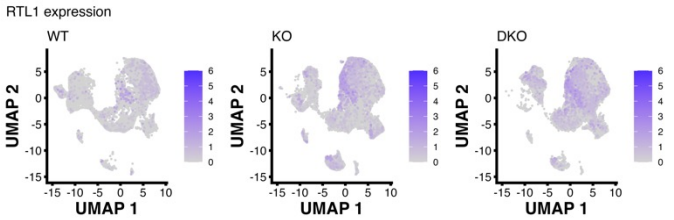


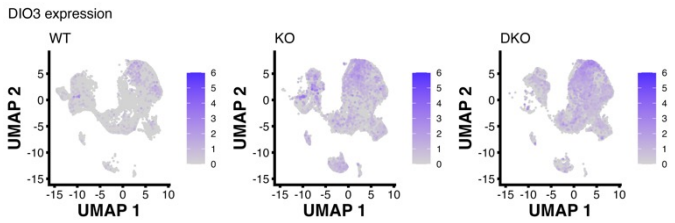




Paternal imprinting genes







B.

Non-neuronal lineage cell fraction relative to WT

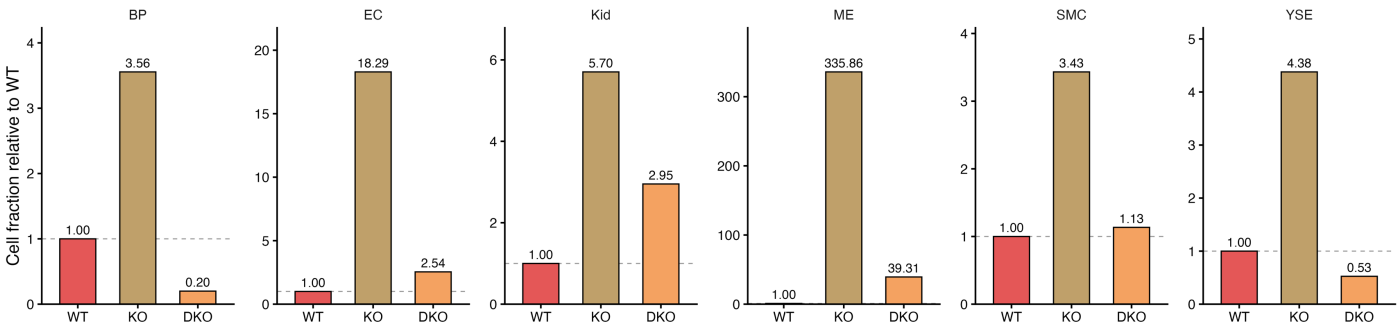
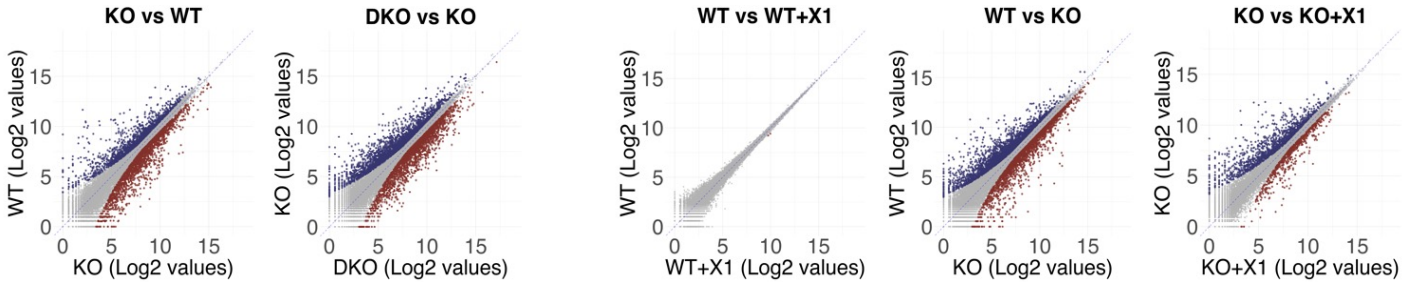


Figure S7. Imprinted gene expression and non-neuronal lineage proportions in AFF3 WT/KO/DKO cerebral organoids.

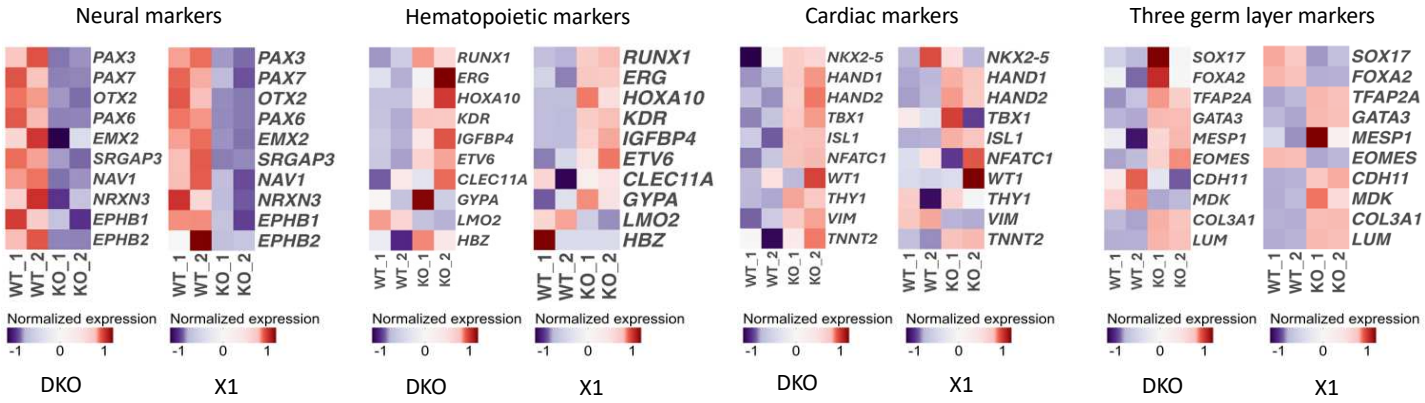
(A) UMAP showing maternal gene expression, including MEG3, MEG8, MEG9, and paternal gene expression, including DLK1, RTL1 expression in AFF3 WT/KO/DKO cerebral organoids. (B) Bar plot showing fraction of non-neural lineages including blood progenitors (BP), endothelial cells (EC), kidney-lineage cells (Kid), meso-endoderm (ME), smooth muscle cells (SMC), and yolk sac endoderm (YSE) in AFF3 WT/KO/DKO cerebral organoids, normalized to WT.

Supplementary Materials-Sup Fig8

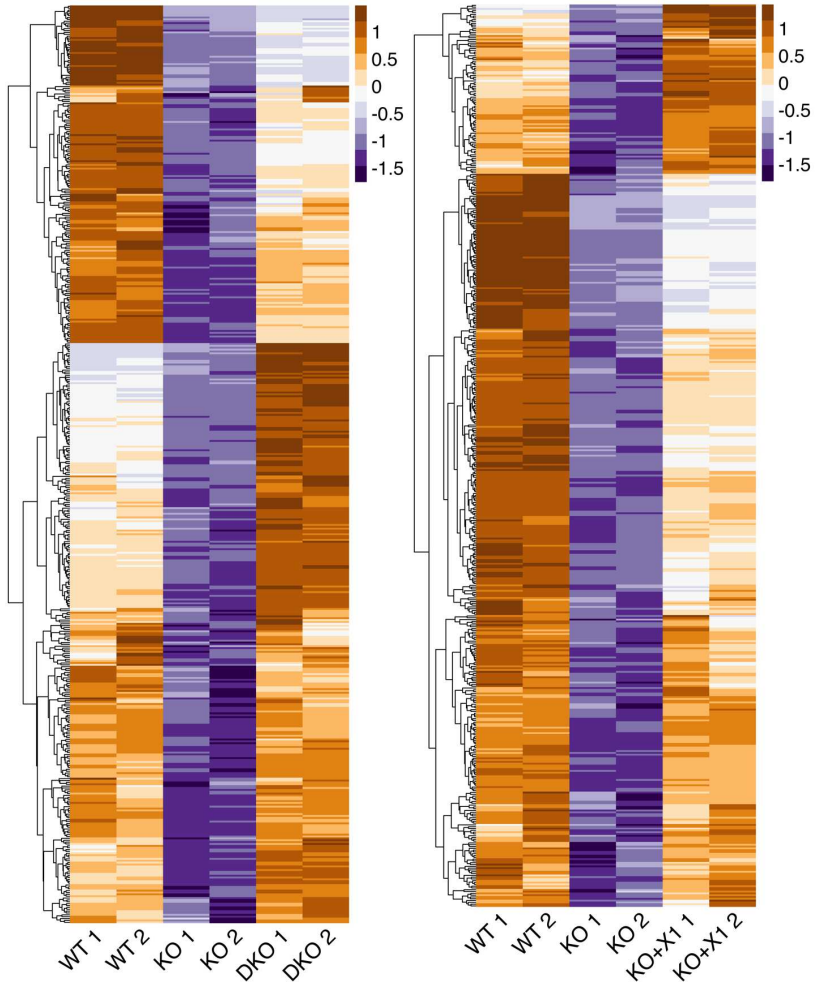
A.



B.



C.



D.

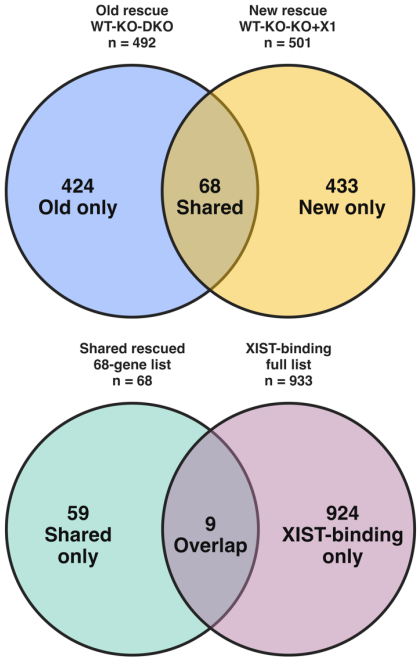


Figure S8. Pharmacological inhibition and genetic knockout of XIST can upregulate neural marker gene expression.

(A) Scatterplot of RNA-seq data showing DEGs in KO vs WT, DKO vs KO, WT vs WT+X1, WT vs KO, KO vs KO+X1 in female gastruloids. (B) Heatmaps showing the expression patterns of neural lineages, hematopoietic lineages, cardiac lineages, and three germ layers in AFF3 KO gastruloids compared with WT gastruloids. (C) Heatmaps showing the expression patterns of DEGs in DKO (left)/KO+X1(right) compared with KO. (D) Venn Diagram showing the overlapping 68 genes between DKO rescued genes and KO+X1 rescued genes (top), and then the final overlapping genes between the 68 genes and XIST targeting genes.

Supplementary Materials-Sup Fig9

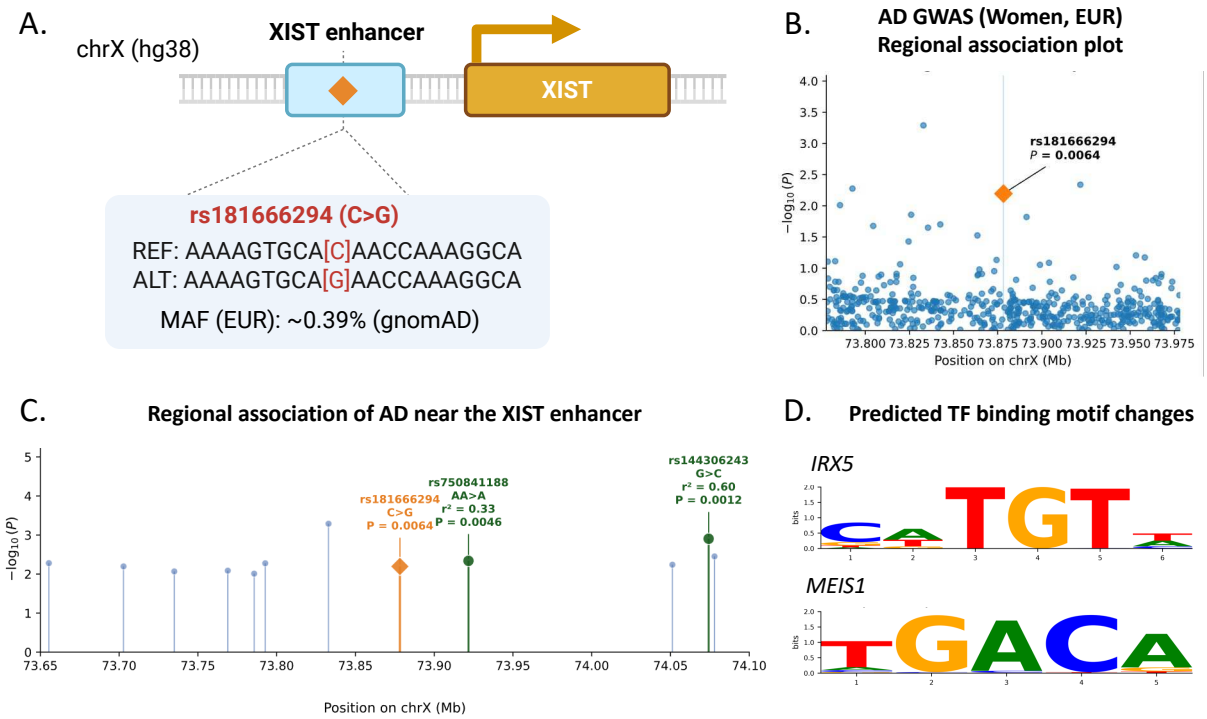


Figure S9. Genetic and functional characterization of a low-frequency variant within the XIST enhancer.

(A) Schematic representation of the XIST enhancer and the candidate variant rs181666294 on chromosome X (GRCh38). rs181666294 is located within the experimentally identified XIST enhancer at chrX:73,877,041–73,878,379 and represents a C>G substitution in the GRCh38 genomic orientation. The surrounding reference (REF) and alternative (ALT) sequences are shown. The alternative allele is present at low frequency in European populations (gnomAD non-Finnish European allele frequency approximately 0.39%; ~0.5% in 1000 Genomes EUR). (B) Regional association plot from a large X chromosome-wide association study of Alzheimer's disease in women of European ancestry. rs181666294 showed a nominal association with Alzheimer's disease in the female-stratified analysis ($P = 0.0064$; $\beta = -0.232$, effect allele frequency = 0.0042). (C) Regional analysis of nominal Alzheimer's disease association signals surrounding the XIST enhancer. rs181666294 is highlighted in orange. Two nearby low-frequency variants showing similar negative Alzheimer's disease association signals in women are highlighted in dark green: rs750841188, a 1-bp deletion located approximately 43 kb from the enhancer ($P = 0.0046$), and rs144306243 (G>C), located approximately 196 kb from the enhancer ($P = 0.0012$). In the 1000 Genomes Phase 3 European reference population, these variants showed moderate linkage disequilibrium with rs181666294 (rs750841188, $r^2 = 0.33$; rs144306243, $r^2 = 0.60$). (D) Predicted effects of rs181666294 on transcription factor binding motifs using JASPAR position frequency matrices. Sequence-based motif analysis predicted that the C>G substitution reduces the relative match score for IRX5 (MA2680.1; REF = 0.868, ALT = 0.626, $\Delta = -0.242$) and MEIS1 (MA0498.3; REF = 0.864, ALT = 0.629, $\Delta = -0.235$). Sequence logos represent the corresponding JASPAR motifs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CellReagent20261001.xlsx
- SupplementaryGELrawData.pdf