Published online 2 March 2023

Nucleic Acids Research, 2023, Vol. 51, No. 10 ***4745–4759***
https://doi.org/10.1093/nar/gkad109

Endogenous retrovirus-derived enhancers confer the transcriptional regulation of human trophoblast syncytialization

Miao Yu[1,†], Xiaoqian Hu[2,†], Zihang Pan[1], Cui Du[3], Jing Jiang[3], Wanshan Zheng[1], Han Cai[1], Yinan Wang[1], Wenbo Deng[1], Haibin Wang[1], Jinhua Lu[1], Ming-an Sun[3,4,*] and Bin Cao[1,*]

[1]Fujian Provincial Key Laboratory of Reproductive Health Research, Department of Obstetrics and Gynecology, Women and Children's Hospital, School of Medicine, Xiamen University, Xiamen, Fujian 361002, China, [2]State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, School of Public Health, Xiamen University, Xiamen, Fujian 361002, China, [3]Institute of Comparative Medicine, College of Veterinary Medicine, Yangzhou University, Yangzhou, Jiangsu 225009, China and [4]Jiangsu Co-innovation Center for Prevention and Control of Important Animal Infectious Diseases and Zoonosis, Joint International Research Laboratory of Agriculture and Agri-Product Safety, the Ministry of Education of China, Yangzhou University, Yangzhou, Jiangsu, China

Received August 23, 2022; Revised February 01, 2023; Editorial Decision February 03, 2023; Accepted February 07, 2023

ABSTRACT

Endogenous retroviruses (ERVs) have been proposed as a driving force for the evolution of the mammalian placenta, however, the contribution of ERVs to placental development and the underlying regulatory mechanism remain largely elusive. A key process of placental development is the formation of multinucleated syncytiotrophoblasts (STBs) in direct contact with maternal blood, through which constitutes the maternal-fetal interface critical for nutrient allocation, hormone production and immunological modulation during pregnancy. We delineate that ERVs profoundly rewire the transcriptional program of trophoblast syncytialization. Here, we first determined the dynamic landscape of bivalent ERV-derived enhancers with dual occupancy of H3K27ac and H3K9me3 in human trophoblast stem cells (hTSCs). We further demonstrated that enhancers overlapping several ERV families tend to exhibit increased H3K27ac and reduced H3K9me3 occupancy in STBs relative to hTSCs. Particularly, bivalent enhancers derived from the Simiiformes-specific MER50 transposons were linked to a cluster of genes important for STB formation. Importantly, deletions of MER50 elements adjacent to several STB genes, including *MFSD2A* and *TNFAIP2*, significantly attenuated their expression concomitant to compromised syncytium formation. Together, we propose that ERV-derived enhancers, MER50 specifically, fine-tune the transcriptional networks accounting for human trophoblast syncytialization, which sheds light on a novel ERV-mediated regulatory mechanism underlying placental development.

INTRODUCTION

The placenta is a critical organ that simultaneously maintains successful pregnancy and supports fetal development by providing substantial resources of oxygen and nutrients, facilitating hormone production and immunological functions (1). The chorionic villus is the basic architectural and functional unit of the human placenta, which is primarily comprised of the interior layer of mononuclear cytotrophoblast (CTB) and the outermost layer of multinucleated syncytiotrophoblast (STB) that directly immersed in maternal blood. Proliferative CTBs give rise to STBs through cell-cell fusion, termed syncytialization, which enables proper placental functions and fetal development, whereas aberrant placental syncytium formation is associated with pregnancy complications, including preeclampsia, recurrent pregnancy loss, and fetal growth restriction (2–4). Although accumulating evidence has shown that hormones, cytokines and transcription factors are indispensable for the process of trophoblast cell fusion (5,6), the underlying genetic and epigenetic mechanisms and concomitant regulatory network governing trophoblast syncytialization remain elusive.

Endogenous retroviruses (ERVs), which belong to transposable elements (TEs), have been proposed as a driving

*To whom correspondence should be addressed. Tel: +86 05922880503; Fax: +86 05922880503; Email: caobin19@xmu.edu.cn
Correspondence may also be addressed to Ming-an Sun. Tel: +86 18362820286; Fax: +86 051487979275; Email: mingansun@yzu.edu.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

force for the evolution of the mammalian placenta (7–10). One of the most striking examples is the co-option of Syncytin genes in the placenta. Syncytins are derived from endogenous retroviral elements, which code envelope proteins in ancient retrovirus to aid virion–cell membrane fusion and evolve to serve as bona fide fusogenic proteins to promote cell-cell fusion in trophoblasts (11). For example, Syncytin-1 in humans binds to specific receptors to facilitate the formation of pseudopod on the plasma membrane and the change of phospholipid composition of the membrane, thereby promoting the fusion between two neighboring trophoblast cells (12). Remarkably, *Syna* and *Synb* knockout mice exhibit the defective formation of the ST-I and ST-II layer respectively (13), which highlights the importance of Syncytins for preserving the structural and functional integrity of the placental barrier by modulating trophoblast syncytialization. Moreover, the emergence of Syncytins in different mammalian species seems to affect the placental morphology, invasiveness of the trophoblast, and cellular composition of placentas, thus accounting for the evolutionary diversification of the placentas (14).

Aside from co-option for direct cellular functions, such as cell-cell fusion and immune modulation, ERVs provide cell- and context-specific transcriptional modulatory machineries, such as promoters and enhancers, splicing and termination sites, and regulatory non-coding RNAs (15). By virtue of their repetitive and interspersed nature, the existence of ERVs profoundly drives genome reconstruction and shapes host gene regulatory networks during mammalian evolution (16,17). Notably, ERVs are susceptible to substantial epigenetic modulation, attracting various epigenetic modifiers such as H3K27ac and H3K4me3, which therefore affects the expression of neighboring genes (18). Cumulative evidence documents the *cis*-regulatory activity of ERVs in early embryogenesis (19,20), differentiated cells and tissues (21), and germline formation (22), among which recent progress spotlights the physiological significance of ERVs in extraembryonic tissues. For example, an anthropoid primate-specific ERV-derived enhancer THE1B controls the placental expression of CRH and alters gestation length and birth timing (23). Sun et al. recently performed an epigenomic survey to identify a group of putative lineage-specific placental enhancers derived from ERVs (24), yet their contributions to cell-specific transcriptional regulation in the placenta are under-appreciated. In addition, how ERV elements rewire the transcriptional networks critical for trophoblast cell lineage development remains undetermined.

In the present study, we profile epigenomic and transcriptomic features of the human trophoblast stem cell (hTSC) and hTSC-derived STB to investigate the potential regulatory roles of ERVs in the process of trophoblast syncytialization. We reveal the prevalence of ERVs preferably enriched for H3K27ac/H3K9me3 bivalent enhancers in hTSCs, which transitions to an active state upon STB differentiation. Remarkably, MER50 ERVs serve as functional enhancers driving the expression of adjacent STB genes. We demonstrate that the downregulation of certain STB genes by CRISPR deletion of selected MER50 elements could hinder syncytium formation. Moreover, our study identifies novel regulators of trophoblast syncytialization, such as TNFAIP2 and RAI14, whose expression and function are substantially influenced by the MER50 enhancers. Our study indicates ERV enhancers fine-tune the transcriptional program accounting for trophoblast fusion apart from the well-established function of coding fusogenic genes.

MATERIALS AND METHODS

hTSC and organoid culture

Human trophoblast stem cells from the first-trimester placenta were isolated and cultured as described previously (25). Briefly, the first-trimester placenta villi were digested using 0.25% trypsin and 10 U/ml DNase I for 0.5 h and 15 min, respectively. The isolated cells were filtered and resuspended in the hTSC medium. The Ethics Committees of Xiamen University approved the study protocol for the collection of human placenta tissues. Written consent was obtained from all sample donators. For the organoid derived from hTSCs, about 5×10^3 cells were embedded into 30 μl per drop Matrigel (Corning, Cat#356231) and carefully placed upside down at 37°C for 20 min to solidify, followed by continuous culture in the trophoblast organoid medium (TOM) (26). Additional details of the hTSC and placental organoid culture are provided in Supplementary methods.

Generation of CRISPR knock-out cell lines

To generate the ΔMER50-M and ΔMER50-T cells, two sgRNA sequences for each sequence were designed to create a specific deletion encompassing the MER50 sequence upstream of *MFSD2A* and *TNFAIP2*, respectively. sgRNAs were introduced into hTSCs by the lentivirus system [lentiCRISPR v2 (Addgene, #52961), psPAX2 (Addgene, #12260), pMD2.G (Addgene, #12259)]. PCR and Sanger sequencing was utilized to validate the efficiency of the corresponding genomic deletion in the individual colonies derived from single cell expansion. Sequences for the sgRNAs and genotyping primers are shown in Supplementary Table S1, S2.

Transcriptome sequencing and analysis

Total RNA was isolated from three biological replicates using TRIzol (Invitrogen, Cat#350508), and RNA integrity and concentration were assessed by the Qubit dsDNA HS Assay (Invitrogen). RNA-Seq libraries were sequenced as 100 bp paired-end reads with the MGISEQ-2000 platform. Raw reads were trimmed with Trim Galore v0.6.4, and then aligned to the reference genome (GRCh38) using STAR v2.7.3 (27). After obtaining gene-level read counts using the *featureCount* function from subread v2.0.0 (28), differentially expressed genes were identified using DESeq2 v1.30.1 (29) with settings: FDR < 0.05 and $|\log2\text{Foldchange}| > 1$.

ChIP-seq

The information on ChIP-Seq data sets as listed in Supplementary Table S3. For this study, ChIP-Seq was performed according to the previous protocol (30). In brief, cell samples (hTSC, STB24h and STB48h) were fixed with 1% formaldehyde solution for 10 min at room temperature,

then 0.125 M glycine was applied for 10 min to stop fixation. Samples were then sequentially incubated in lysis buffer 1 (50 mM HEPES pH 7.5, 1 mM EDTA, 140 mM NaCl, 0.5% NP-40, 10% glycerol, 0.25% Triton X-100) and lysis buffer 2 (10 mM Tris–HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl) on a rotator at 4°C for 10 min. Chromatin fragmentation was performed using Bioruptor Sonicator (Diagenode) in lysis buffer 3 (10 mM Tris–HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl, 0.1% sodium deoxycholate, 0.5% *N*-lauroylsarcosine). Fragmented chromatin was centrifuged to remove debris, and 5% of the volume was used as an input control. The rest was incubated with the anti-H3K27ac (Abcam, Cat#ab4729) or anti-H3K4me3 (Abcam, Cat#ab8580) antibody on a rotator at 4°C overnight and subjected to 20 μl protein A magnetic beads (Invitrogen, Cat#10001D). After washing and elution, the protein-DNA complex was reversed by heating at 65°C overnight. DNA was purified, and Illumina libraries were constructed using the KAPA DNA HyperPrep Kit (Roche, Cat#08963835001) following the manufacturer's instruction and sequenced on an Illumina NovaSeq 6000.

Raw reads were trimmed with Trim Galore v0.6.4, and then aligned to the reference genome (GRCh38) using Bowtie v2.3.5 (31) with default settings. Reads aligned to multiple genomic loci were discarded using samtools v1.13 (32) with the parameter -q 2. PCR duplicates were removed using the *rmdup* function of samtools v1.13 (32). After confirming the data reproducibility, reads from biological replicates were pooled together for further analysis. Peak calling was performed with MACS v2.2.6 (33) with default settings, with H3K9me3 and H3K27me3 called as broad peaks while others as narrow peaks. The called peaks were further cleaned by removing those overlapping ENCODE Blacklist V2 regions (34).

Reference genome and annotation

Reference genome and gene annotation for human (GRCh38) were downloaded from the ENSEMBL database (release 102) (35). Transposable element annotations were downloaded from the RepeatMasker website (http://www.repeatmasker.org/).

Regulatory element annotation

Putative regulatory elements, including promoters and enhancers, were annotated based on genome-wide distribution of histone modifications and distance to TSSs. In brief, we first determined the peaks for H3K27ac, H3K4me3 and H3K9me3 based on corresponding ChIP-Seq data. According to the rationale that H3K4me3 specifically marks active promoters and H3K27ac marks both active promoters and enhancers, we then annotated promoters are defined as H3K4me3 occupied regions, and enhancers as H3K27ac peaks that are >500 bp from TSSs. Annotated promoters and enhancers that overlap H3K9me3 peaks were further annotated as of 'bivalent' state. To compare between regulatory elements shared or specific for hTSC and STB, annotated enhancers were further clarified as distinct groups based on H3K27ac and H3K9me3 marks in hTSC and STB, including: (i) hTSC-specific, for those with H3K27ac only in hTSC; (ii) shared (hTSC-bivalent), for those with H3K27ac in both hTSC and STB, yet also have H3K9me3 in hTSC; (iii) shared (hTSC-active), for those with H3K27ac in both hTSC and STB, and no H3K9me3 in hTSC; (iv) STB-specific, for those with H3K27ac in STB only.

TE enrichment analysis

To determine the overrepresentation of ERV families within distinct groups of regulatory elements, we used the *fisher* function from BEDtools v2.29.2 (36) for Fisher's exact test analysis. To control for Family-Wise Error Rate, the calculated *P*-values were further adjusted with Bonferroni approach.

Dual-luciferase reporter assay

To construct the promoter reporter vector, the MER50 consensus sequence (from Dfam database; https://www.dfam.org/home) was synthesized by Tsingke Biotechnology Co., Ltd and cloned into pGL4.10 (Promega, Cat#E6651). The reporter constructs with no promoter served as a control. Enhancer reporter vectors were generated by inserting the MER50 consensus sequence and downstream mini promoter (miniP: AGAGGGTATATAATGGAAGCTCGACTTCCAG) into the pGL4.10 plasmid. The reporter vector with sole miniP was used as baseline control. HEK293T cells or hTSCs were seeded in 24-well plates one day prior to transfection. Firefly and Renilla (used for normalization) luciferase were co-transfected into cells using Lipofectamine 3000.

Western blot analysis, RNA extraction, reverse-transcription quantitative PCR (RT-qPCR), immunostaining, siRNA transfection, MTS assay and ELISA

Western blot, RNA extraction, RT-qPCR, Immunostaining, siRNA transfection, MTS assay and ELISA were performed using standard procedures. Additional details are available in Supplementary methods and Supplementary Table S4.

Statistical analysis and visualization

All statistical analyses on the transcriptome and epigenome data were performed with R statistical programming language (37). Heatmaps for ChIP-Seq data were generated using DeepTools v3.5.1 (38). RNA-Seq and ChIP-Seq tracks were visualized using IGV v2.11.1 (39). For other analyses, statistical calculations were performed with unpaired Student's *t* test, Mann-Whitney test, and one- or two-way ANOVA using GraphPad Prism 8. Data was documented as the mean ± SEM, based on at least three independent experiments.

RESULTS

ERVs frequently overlap enhancers that transit from bivalent to active state during human STB differentiation

Accumulating evidence demonstrates that ERVs could serve as prominent regulatory elements in diverse developmental

events (17). We previously elucidated that dozens of ERV families are overrepresented in human placental enhancers (24). To further delineate the functions of ERVs during human trophoblast lineage specification, we established an *in vitro* model for hTSC culture and differentiation (Figure 1A) using a protocol modified from the previous study (25). We successfully generated proliferative hTSCs from the first-trimester placenta CTBs and subsequently differentiated hTSCs *in vitro*. Cellular morphology under bright field microscopy and the expression of canonical markers for different trophoblast subtypes (i.e. *ITGA6* and *TEAD4* for hTSC, *CGB* for STB, *HLA-G* for EVT) confirmed that the hTSC we established possesses the potency to differentiate into both STB and EVT lineages (Supplementary Figure S1A–D). Immunofluorescent imaging further verified the morphology of differentiated trophoblast subtypes and the expression of marker genes (hCG for STB and HLA-G for EVT) (Figure 1B, Supplementary Figure S1D). Moreover, hTSCs cultured in three-dimensional (3D) Matrigel droplets underwent self-organization and formed a spheroid-shaped organoid, where the outer layer sustained mononucleated CTB and the interior committed to spontaneous STB differentiation (Supplementary Figure S1E, F). We utilized the hTSC model to study the transcriptomic and epigenomic dynamics during STB differentiation, focusing on the function of ERVs regarding gene regulatory activity and trophoblast cell fusion.

To identify the ERV components with regulatory potentials governing trophoblast differentiation, we assessed the chromatin signature of *cis*-elements in the process of syncytialization. We first annotated the *cis*-elements for hTSC and STB using public epigenomic data (25) based on two active histone marks H3K27ac and H3K4me3 (Supplementary Table S3). We further leveraged published H3K9me3 data and verified its good quality (Supplementary Figure S2). Notably, a higher percentage of overall enhancers in hTSC relative to STB (16.8% versus 3.4%, $P < 2.2e{-}16$) were of bivalency featuring both H3K27ac and the repressive mark H3K9me3 (Supplementary Figure S3A–C). The similar bivalent state on ERVs has also been previously observed in mouse embryonic stem cells (ESCs) (40). A comparison of the shared enhancers in both cell types further confirmed the prevalence of bivalent enhancers in hTSC (Figure 1C). Therefore, we further classified the shared *cis*-elements as hTSC-bivalent and hTSC-active groups (Figure 1D), with the hTSC-bivalent group switching from bivalent to active state during STB differentiation. This is consistent with the overall reduced H3K9me3 level in STBs (Supplementary Figure S3D). Together with those specific for hTSC and STB, four groups of enhancers were defined for further comparison. Overall, ERVs were overrepresented in the enhancers for both hTSC and STB (Figure 1E and Supplementary Figure S3E), particularly the shared enhancers with the bivalent state in hTSC (Figure 1E, F). In contrast, ERVs were rarely enriched in promoter regions (Supplementary Figure S4A, B), which suggests that ERVs *cis*-regulatory elements might be coopted primarily as enhancers in trophoblasts. This histone modification feature in hTSCs fits nicely with the previous findings that rare enrichment of TEs within promoters marked by H3K4me3 in rodent placentas (41). Importantly, while multiple ERV families (e.g. MER21, MER41, MER39, MER50) previously reported as placental-enhancer-enriched (24) were identified as significant hits in our analysis, the top enriched ERV families differed substantially among different groups of enhancers (Supplementary Figure S4C). Analysis of the newly generated epigenomic dataset using our *in vitro* model (Supplementary Table S3) gave rise to highly consistent data (Figure 1G, H, Supplementary Figure S3F and S5). Together, these results suggest that distinct families of ERVs are preferably enriched in different groups of enhancers, with the most significant enrichment for enhancers that transit from bivalent to active state during STB differentiation.

MER50 transposons constitute putative enhancers for the adjacent STB genes

Poised *cis*-elements with bivalent histone modifications profoundly shape the transcriptional landscape of developmental processes by maintaining low expression of cell fate specification genes, while simultaneously keeping them predisposed for rapid activation in response to lineage specification signals (42). After unveiling the epigenetic characteristics of ERVs within different groups of hTSC/STB enhancers, we further examined if certain ERV families are more likely located near genes important for STB differentiation. We first identified the genes up-regulated in STB relative to hTSC using both previously published (25) and new transcriptomic data, which compiled a list of 2455 genes more abundant in STB (Figure 2A). We next profiled the occurrence of ERVs close to these STB genes and identified several ERV families that were significantly enriched (Figure 2B), including MER41A, MER50 and LTR8B, which were also highly enriched within STB enhancers (Supplementary Figures S4 and S5). We further demonstrated that STB enhancers overlapping several ERV families (e.g. MER41A, MER50 and LTR8B) tend to have increased H3K27ac and decreased H3K9me3 occupancy in STB relative to hTSC (Figure 2C), which is consistent with the aforementioned bivalent to active transition of ERV-derived enhancers. Apart from the top MER41A family which has previously been linked to the evolution of human placenta and innate immunity (24,43), the second-ranked MER50 family was of particular interest, since Syncytin-2—a placental gene crucial for STB fusion—seems to be co-opted from one env gene originating from the internal region (MER50-int) of the same ERV family (Supplementary Figure S6A). MER50 and MER50-int transposons are both restricted to the Simiiformes lineage, with the latter frequently flanked by the LTR sequence of MER50. Apart from coding the important fusogenic protein Syncytin-2, this coincidence raises the question of whether MER50 transposons also facilitate the evolution of regulatory elements modulating STB differentiation.

To test this hypothesis, we compared the gene expression pattern regarding their distance to MER50, which reveals that genes with MER50 elements resident within 50 kb from their corresponding TSSs displayed higher expression levels in STB relative to hTSC (Figure 2D). We also examined the association of MER50 with annotated enhancers using both public and our datasets (Figure 2E, F, and Supplemen-

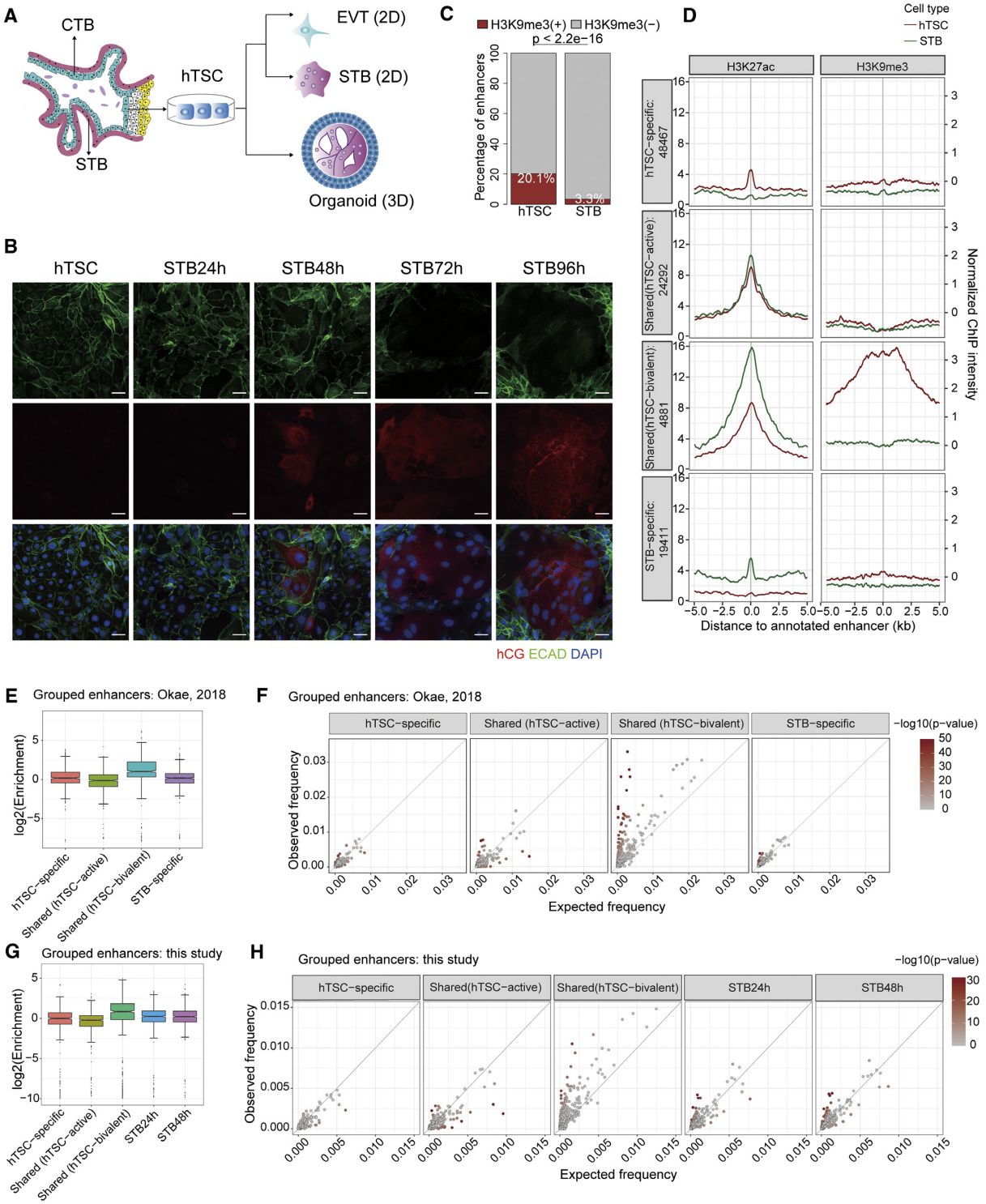

Figure 1. ERV-derived enhancers bear the features of H3K9me3/H3K27ac bivalent histone modifications in human trophoblasts. (**A**) Schematic representation of hTSC derivation and differentiation. hTSC derived from the first-trimester placental villi were maintained in hTSC medium or subject to 2D (STB, EVT) and 3D (organoid) differentiation. (**B**) Representative immunostaining of ECAD (green), hCG (red) and DAPI (blue) in hTSC and STB at the indicated timepoints of syncytialization in 2D culture. Scale bar, 50 μm. (**C**) The relative abundance of annotated enhancers with bivalent marks (H3K27ac and H3K9me3) in hTSC relative to STB. *P*-value is calculated by using Fisher's exact test. (**D**) Averaged ChIP intensity plots showing the alteration of H3K27ac and H3K9me3 marks at different groups of annotated enhancers in hTSC and STB. The numbers for the four groups (i.e. hTSC-Specific, Shared (hTSC-Active), Shared (hTSC-Bivalent) and STB-Specific) of enhancers are 48467, 24292, 4881 and 19411, respectively. (**E**) Boxplots for the enrichment of different ERV families in grouped enhancers for hTSC and STB based on the publicly available epigenomic data. (**F**) Scatter plots for the enrichment of ERV families in different groups of enhancers. The *x*- and *y*-axis represent the expected and observed frequency of different ERV families in enhancers. The color gradient indicates the -log10(p-value) adjusted with Bonferroni correction. (**G**, **H**) Counterpart data to E and F based on newly generated epigenomic data for hTSC and STB at the indicated differentiation timepoint.

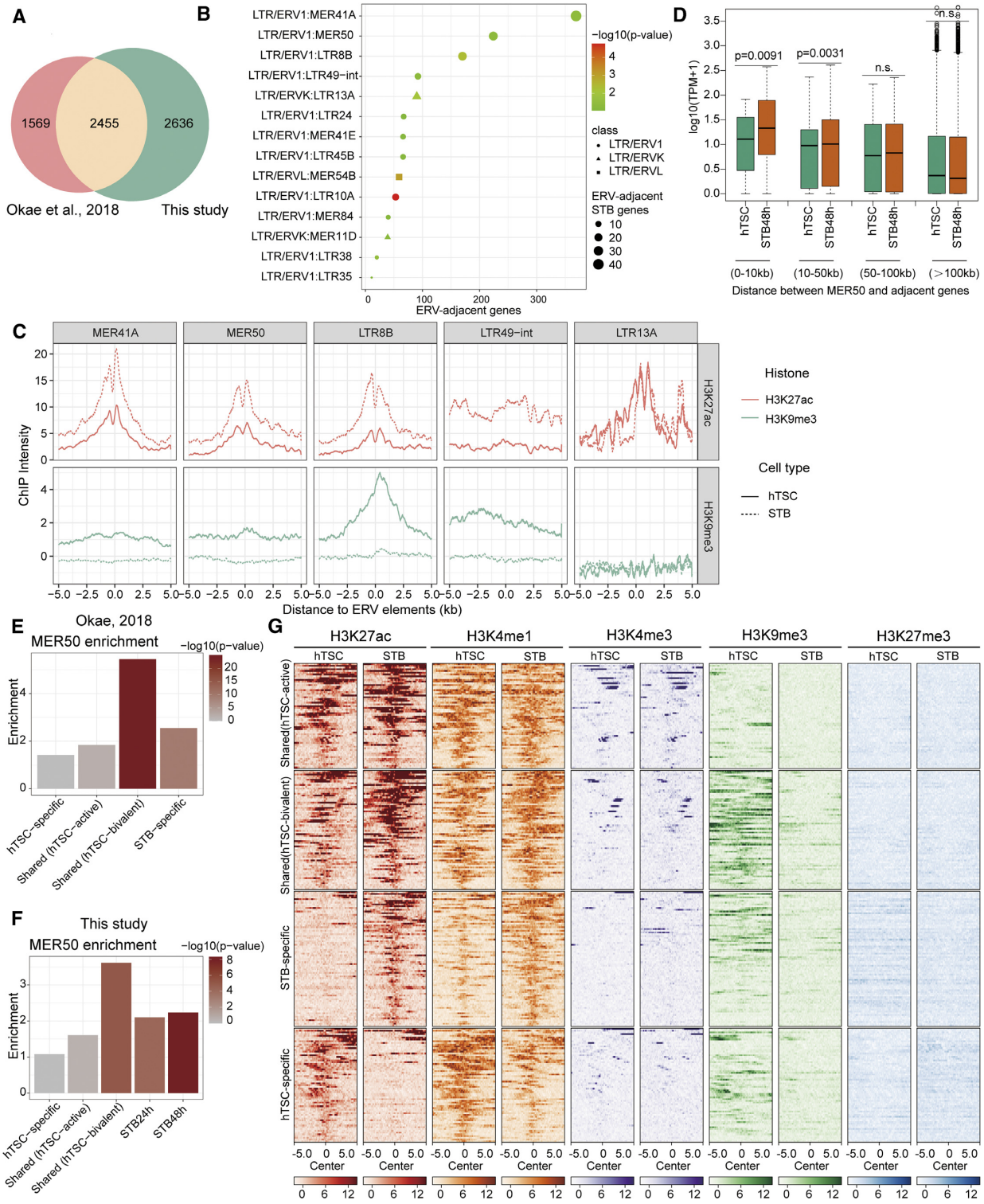

Figure 2. Putative bivalent MER50 enhancers are associated with STB gene expression. (**A**) Venn diagram illustrates 2455 overlapped genes up-regulated in STBs (defined as STB genes) relative to hTSCs based on RNA-Seq data from Okae *et al.* (2018) and the current study. (**B**) The bubble diagram shows the ranking of ERV subtypes close to STB genes (ERVs resident within 10 kb centered on the gene body) by the order of STB gene numbers associated with a given ERV. The *x*-axis represents the numbers for ERV-adjacent genes, and the size of the circle represents the number of ERV-adjacent STB genes. (**C**) Alteration of H3K27ac and H3K9me3 occupancy between hTSC and STB for the STB enhancers overlapping the top 5 ERV families enriched surrounding STB-specific genes. The averaged curves show the normalized ChIP intensity surrounding ERV elements that overlap STB enhancers. (**D**) Expression changes of MER50-associated genes in hTSC and STB. Different groups of genes are defined based on the distance of their TSSs to the nearest MER50 elements. (**E**, **F**) Enrichment of MER50 within different groups of enhancers annotated for hTSC and STB, using public data (E) and new data (F), respectively. The color gradient indicates the $-\log_{10}(P\text{-value})$ adjusted with Bonferroni correction. (**G**) Heatmaps showing the intensity of various histone modifications flanking different groups of MER50-associated enhancers.

tary Figure S6B), which consistently suggests that MER50 is preferably enriched within the enhancers that switch from bivalent/silent to the active state during STB differentiation [i.e. the two groups including Shared (hTSC-bivalent) and STB-specific]. Inspection of the epigenetically-annotated enhancers specific for different human tissues using ENCODE data (Supplementary Table S3) suggests that the enrichment of MER50 is most evident in the placenta (Supplementary Figure S6C), highlighting the cell- or tissue-specific function of MER50 in the placenta. Further visualization of various histone modifications revealed that the shared MER50-associated enhancers (particularly the subset with both H3K27ac and H3K9me3 bivalent modification in hTSC, which makes up around 25% (68 out of 270) of all MER50-derived enhancers) showed increased H3K27ac levels in STB, in parallel with the accumulative STB-specific enhancers (Figure 2G). Even though MER50 also overlap some hTSC-specific enhancers and promoters, their association tends to be random since no significant enrichment was detected (Figure 2E, F, and Supplementary Figure S6D–F). Together, integrative analysis of the transcriptomic and epigenomic data suggests that MER50 elements reside frequently surrounding STB genes and within STB enhancers, indicating that they may be endowed with gene-regulatory activity important for STB differentiation.

MER50 functions to drive the expression of the adjacent STB genes

During the process of biological evolution, ERVs dispersed in genomes with long terminal repeats (LTR) which were once used as viral promoters, acquire novel physiological functions due to their intrinsic regulatory activity as promoters or enhancers to modulate the nearby host gene expression (15). Therefore, we hypothesized that MER50 may serve as enhancers or promoters and drive the synergetic expression of a panel of proximal STB genes in human trophoblasts, which further promotes syncytialization. To this end, we inserted the MER50 consensus sequence upstream of the luciferase reporter gene with or without mini promoter, which could respectively evaluate the enhancer and promoter potentials based on the luciferase signals as a readout (Figure 3A). We first introduced the enhancer or promoter reporter system into 293T cells and measured the activity of MER50 as a *cis*-element. Our results demonstrated that the MER50 could significantly induce downstream luciferase gene expression compared with the controls. In addition, we observed a more robust gene regulatory efficiency of MER50 in human trophoblasts, especially in STBs, which indicates that MER50 enhancers are more active in the setting of trophoblast differentiation (Figure 3B, C). These findings are consistent with the above observations that MER50 was a top-ranking putative ERV element adjacent to STB genes and enriched for the unique H3K27ac/H3K9me3 bivalent histone modification, which together implicates that the MER50 *cis*-element is likely to be involved in STB formation by modulating the expression of STB genes.

To pinpoint MER50-modulated genes critical for trophoblast syncytialization, we profiled the MER50 adjacent STB genes by overlapping 452 adjacent genes of MER50 ($\leq$20 kb of TSSs) with genes upregulated specifically in hTSC-derived STBs. We predicted 40 members of MER50 target genes (Figure 3D), among which a subset of genes has established functions in trophoblast fusion, such as *ERVFRD-1* (44), *MFSD2A* (45), *SPHK1* (46), *PSGs* (47), *IFITMs* (48) and *ERVV-2* (49,50). Furthermore, RNA-Seq data revealed that the expression of MER50 adjacent STB genes was gradually increased as the trophoblast fusion proceed (Figure 3E). Consistently, the representative track view illustrated that active H3K27ac or H3K4me3 histone modification at the MER50 loci correlated with the activation of syncytialization-related genes in human trophoblasts (Figure 3F, G). Taken together, these data corroborate a regulatory link between enhancer-like MER50 ERV and the genes involved in STB formation, which needs further experimental validation to address the corresponding functional significance.

Functional validation of the definitive MER50 enhancer upstream of the *MFSD2A* gene in syncytiotrophoblast formation

To assess if the MER50 elements with enhancer-like epigenomic characteristics are indeed functional enhancers, we focused on the MER50 element adjacent to the *MFSD2A* gene. The syncytiotrophoblast is formed through cytotrophoblast fusion, which is facilitated by fusogenic protein Syncytins. Key fusogens expressed in the human placenta, *ERVW-1* and *ERVFRD-1* (the coding gene for Syncytin-1 and Syncytin-2) (51,52), work in concert with the cognate receptors, SLC1A5 and MFSD2A respectively, to promote cell-cell fusion *in vitro*. Although MFSD2A has been proposed to potentiate trophoblast cell fusion, the molecular mechanisms responsible for the transcriptional regulation of *MFSD2A* in the dynamic process of syncytialization remain unclear. Given that the MER50-MFSD2A element displayed active enhancer features (Figure 3F) and correlated with a robust gene upregulation (Supplementary Figure S7A), we reasoned that the MER50 enhancer confers the transcriptional activation of *MFSD2A*.

To investigate the transcriptional modulation function of the MER50 enhancer, we employed the CRISPR-Cas9 system to delete the MER50 region located 12kb upstream of *MFSD2A* in hTSCs (Figure 4A). The successful deletion of MER50 (termed as ΔMER50-M, ΔM-M for short) was verified by both genomic PCR and Sanger sequencing (Supplementary Figure S7B and C). Single-cell colonies homozygous for ΔMER50-M were generated for further investigation. Notably, cells carrying MER50 excision failed to express MFSD2A, in contrast to wildtype (WT) cells in which MFSD2A transcription was robustly up-regulated in response to cell fusion signal (Figure 4B). We next sought to test whether the deletion of MER50-MFSD2A is sufficient to restrict syncytiotrophoblast formation. Consistent with the previous observation that knockdown of intrinsic MFSD2A expression in BeWo choriocarcinoma cells hinders trophoblast fusion (45), we found the efficiency of syncytialization in ΔMER50-M cells was extremely low, as shown by immunostainings for ECAD, hCG, and SDC1 in both 2D and 3D cultures. (Figure 4C, F and Supplementary Figure S7D). In addition, we compared the dynamic

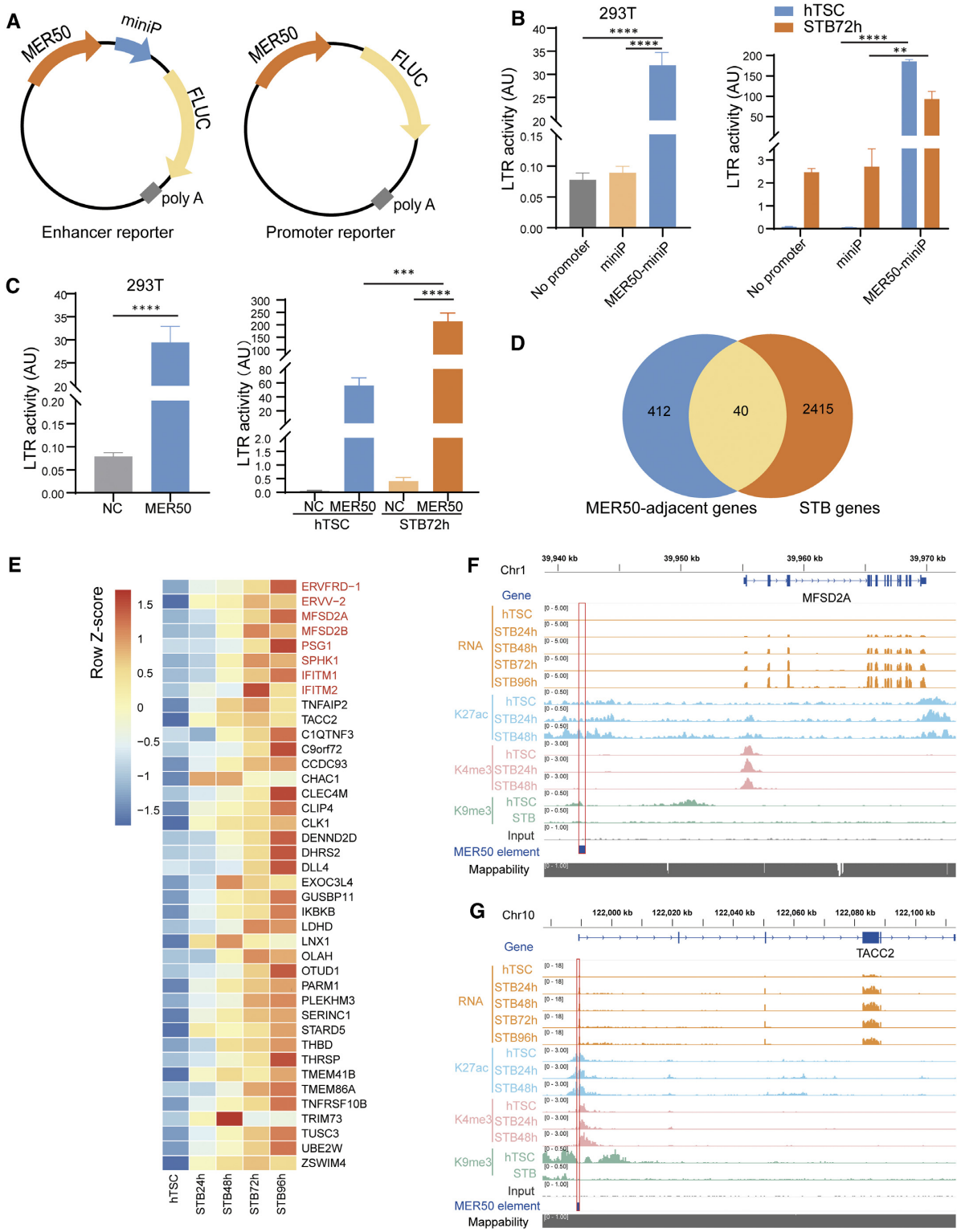

Figure 3. MER50 *cis*-regulatory elements drive the expression of the adjacent STB genes. (**A**) Schematic of the enhancer and promoter reporter system evaluating the transcriptional regulation activity of the MER50 conserved sequence. miniP, minimal promoter; FLUC, Firefly Luciferase. (B, C) Dual-luciferase reporter assays in HEK293T, hTSC, and STB. Cells were transfected with the indicated enhancer (**B**) or promoter (**C**) reporter vectors, and the LTR activity was calculated based on relative fold changes in FLUC activity (FLUC/RLUC). Results represent the mean ± SEM of 3–5 independent experiments. $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$, by Mann–Whitney test and one- or two-way ANOVA. (**D**) Venn diagram depicting the intersection between the coding genes adjacent to MER50 (MER50 resident within 20 kb centered on the gene body) and the STB genes defined in Figure 2A. (**E**) Heatmap representation of dynamic gene expression (*Z*-score-transformed FPKM) of 40 genes overlapped in (D) in the process of trophoblast syncytialization. (**F**, **G**) Genome browser representation for RNA-Seq, H3K27ac, H3K4me3 (data from this study) and H3K9me3 (data from Okae *et al.*) ChIP-Seq on the *MFSD2A* and *TACC2* gene in hTSCs and STBs from continuous differentiation timepoints. MER50 ERV loci are highlighted in the box.

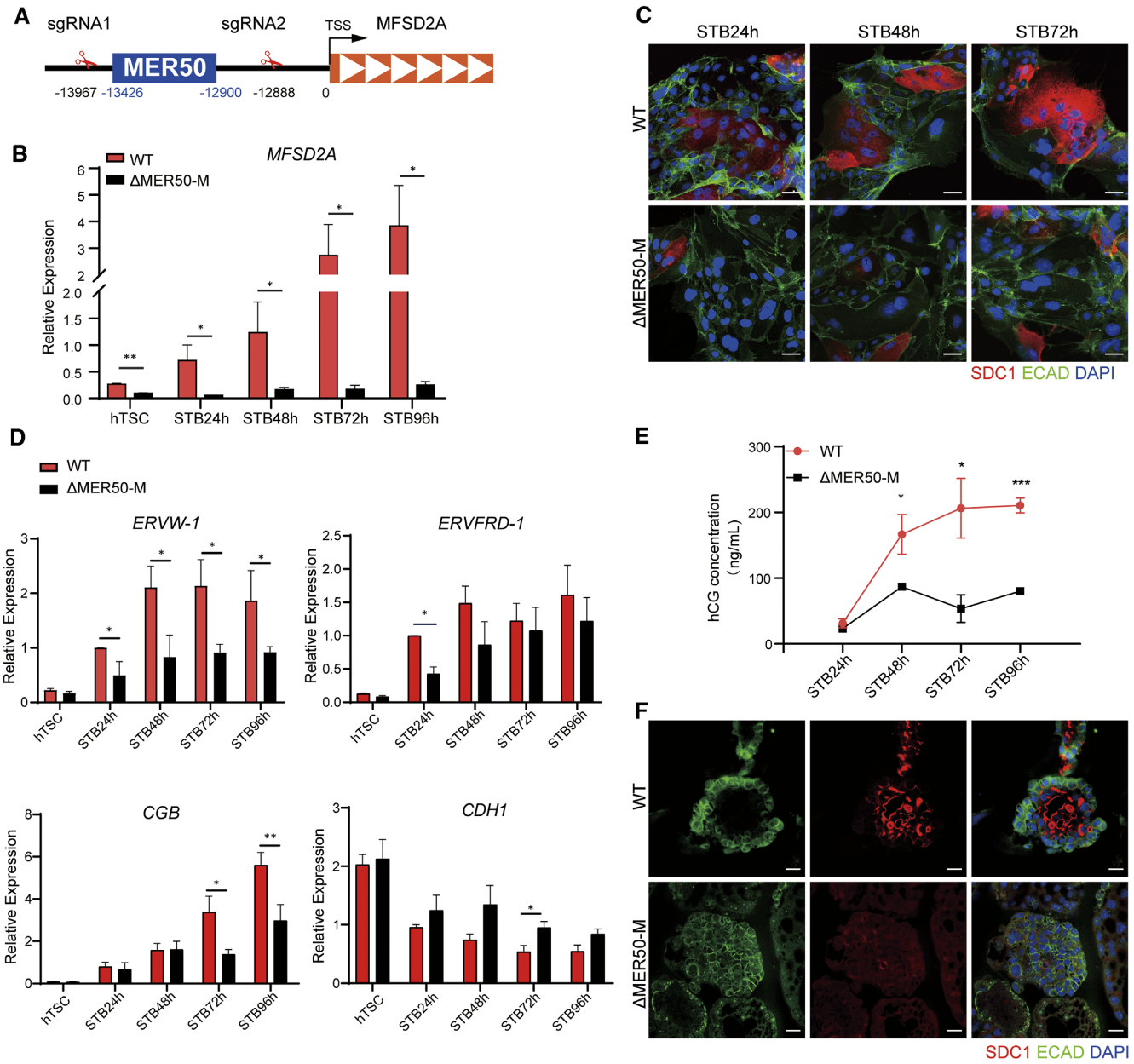

Figure 4. Identification of the MER50 function in regulating MFSD2A expression and STB formation. (**A**) Schematic diagram illustrating CRISPR knockout strategy of the MER50 sequence close to MFSD2A. (**B**) qPCR analysis of *MFSD2A* expression in the WT and ΔMER50-M cells under indicated differentiation status. Results represent the mean ± SEM of three independent experiments. $*P < 0.05$, $**P < 0.01$, Multiple unpaired *t* tests. (**C**) Representative immunostaining of ECAD (green), SDC1 (red) and DAPI (blue) in hTSC and STB derived from the WT/ΔMER50-M hTSC at 24, 48, 72 h. Scale bar, 50 μm. (**D**) RT-qPCR analysis of STB markers (*ERVW-1, ERVFRD-1, CDH1* and *CGB*) in the WT/ΔMER50-M hTSC and STB at 24, 48, 72 and 96 h. Error bars represent mean ± SEM. $*P < 0.05$, $**P < 0.01$, Multiple unpaired *t* tests. At least three biological replicates were examined. (**E**) ELISA detection of hCG concentration in the supernatant from WTs and ΔMER50-M STBs. $*P < 0.05$, $***P < 0.001$; Multiple unpaired *t* tests. $n = 4$ for each group. (**F**) The organoids derived from the WT/ΔMER50-M hTSCs (day 7) were stained for ECAD (green), SDC1 (red) and DAPI (blue). Scale bar, 20 μm.

expression of genes critical for STB formation by qPCR. *ERVW-1* (encode Syncytin-1) was significantly repressed at all timepoints of STB differentiation in ΔMER50-M cells (Figure 4D), while *ERVFRD-1* (encode Syncytin-2) expression showed more significant downregulation at 24h, which supports the function of Syncytin-2 as a prerequisite regulator for ensuring maintenance of post-mitotic syncytia (14,53). Furthermore, defective hCG expression and secretion, a hallmark of functional STBs, was repeatedly observed in ΔMER50-M cells at 48, 72 and 96 h differentiation timepoints (Figure 4E). Collectively, our findings revealed the MER50-enhancer is indispensable for driving the upregulation of MFSD2A expression, which promotes the conversion of cell fate from hTSC to STB.

Inspection on MER50-associated genes uncovered novel regulators of human STB differentiation

Besides some well-known STB-related genes, our data also identified several genes with proximal MER50 enhancers, whose impacts on trophoblast syncytialization were unappreciated. To test whether these genes act as new regulators of trophoblast fusion and their activation is fine-tuned by the corresponding MER50 enhancer, we studied the contribution of two candidate genes, *TNFAIP2* and *RAI14*, regarding STB formation. *TNFAIP2*, a tumor necrosis factor-α (TNFα)-induced gene, was implicated to involve in a variety of physiological processes during cell fate reprogramming and organ development (54–57). To investigate whether it is required for trophoblast differentiation, we therefore used the CRISPR-Cas9 system to generate genomic deletion of the *TNFAIP2* gene in hTSC (TNFAIP2-KO), which was assured by Western blot and Sanger sequencing (Supplementary Figure S8A, B). The TNFAIP2-KO cells grown in the hTSC culture medium displayed indistinguishable stemness from the WT cells, as shown by the immunostaining of TP63 (an hTSC marker) and MTS growth assay (Figure 5A and Supplementary Figure S8C). However, the TNFAIP2-KO cell failed to form hCG-positive STB cells under the cell culture condition of differentiation, in contrast to the control WT cells showing significant downregulation of E-cadherin (ECAD) and TP63, as well as increased hCG concomitant to trophoblast syncytialization (Figure 5A, B). Consistently, Western blot and RT-qPCR of several hTSC and STB markers further supported our finding that knocking out of TNFAIP2 from hTSC significantly decelerated, to a large extent, inhibited syncytiotrophoblast formation (Figure 5C, D), which was further supported by the hCG ELISA data (Figure 5E). In line with the data from 2D cultures, we demonstrated that genomic excision of TNFAIP2 in the placenta organoid, a spontaneous STB differentiation model, resulted in diminished induction of SDC1-positive STB cells interior of the organoid (Supplementary Figure S8D). In addition, we also performed siRNA targeting *TNFAIP2* to further validate our data and found downregulation of endogenous *TNFAIP2* transcripts is sufficient to block trophoblast syncytialization (Supplementary Figure S8E-G). However, in comparison to *MFSD2A* and *TNFAIP2*, hypomorphic mutation of another MER50-associated STB gene, *RAI14*, seemed to accelerate the kinetics of the STB differentiation process, as shown by immunostainings and Western blot of the above-mentioned STB markers (Supplementary Figure S10). Taken together, MER50 target genes may constitute a complex gene regulatory network for syncytium formation.

To evaluate the function of the MER50 element lying close to the newly defined STB gene in terms of transcriptional regulation, we focused on the MER50 element which is located 4.5kb upstream of the *TNFAIP2*. Notably, the potential MER50 enhancer region adjacent to *TNFAIP2* exhibited classic bivalent histone marks of H3K27ac and H3K9me3 modifications, which underwent significant epigenetic changes characterized by increased H3K27ac and decreased H3K9me3 histone modification upon syncytialization (Supplementary Figure S9A). To address the functional significance of the *TNFAIP2*-MER50 enhancer, we designed two sgRNAs to excise the MER50 sequence nearby *TNFAIP2*, named ΔMER50-TNFAIP2 (ΔMER50-T or ΔM-T for short) (Figure 5F, Supplementary Figure S9B, C). Using the single-cell colony expansion with efficient deletion of MER50 (Supplementary Figure S9C), we delineated the function of MER50 regulating *TNFAIP2* expression (Figure 5G, H). RT-qPCR results showed that, to some extent, MER50 deletion indeed led to decreased expression of the nearby *TNFAIP2* at the transcriptional level (Figure 5G). Unexpectedly, the ratio of STB in both 2D and 3D cultures was not dramatically altered in the ΔMER50-T cells shown by immunostainings of STB makers (Supplementary Figure S9D, E), which may be due to the possibility that the baseline level of TNFAIP2 is sufficient to maintain a sufficient level of trophoblast fusion. However, the hCG secretion kinetics was significantly slowed down in ΔMER50-T cells, which implies that the *TNFAIP2*-MER50 enhancer could promote the extent of functional STB formation instead of changing the overall morphology and marker expression of fused trophoblast cell (Figure 5H). Taken together, we discovered TNFAIP2 as a novel player in modulating trophoblast cell fusion, whose expression and function were profoundly influenced by the adjacent MER50-derived enhancer element.

DISCUSSION

In the current study, we reveal a distinct chromatin signature of H3K27ac/H3K9me3 bivalency in ERV-derived enhancers, which displays a transition from bivalent to active state during STB differentiation. We identify MER50 ERV enhancers indispensable for the transcriptional activation of the adjacent STB genes, whereby regulate trophoblast fusion. Importantly, functional evaluation of MER50 enhancers regarding syncytialization and corresponding gene expression pattern emphasizes the coordinated gene regulatory networks extensively driven by the MER50 enhancers in the process of trophoblast syncytialization (Supplementary Figure S11).

The epigenome directs cell fate decisions and modulates differentiation into specialized cell types in both the embryonic and extra embryonic cell lineages (58–60). We identify the unique epigenetic features of bivalent enhancers in hTSCs that are associated with trophoblast lineage specification. Bivalent histone modifications leverage epigenetic regulations of both gene activation and repression at the *cis*-element region, which keeps genes silent but poised for rapid activation when the repressive mark is wiped out (42). Recent studies have revealed that dynamic changes in the bivalent domains play a vital regulatory role in the differentiation of divergent stem cells (61,62). For instance, H3K4me3/ H3K27me3 bivalent promoters in ESCs confer the maintained pluripotency by simultaneously silencing cell fate specification genes while keeping them with the potential of rapid activation upon differentiation triggers (63,64). A previous study suggested that bivalent modification often occurs in transposon-exclusion zones due to global repressive epigenetic modifications of transposons that might interfere with the function of the bivalent domains and thus be eliminated by the selection in ESC (63). However, our research observed that this correlation is quite

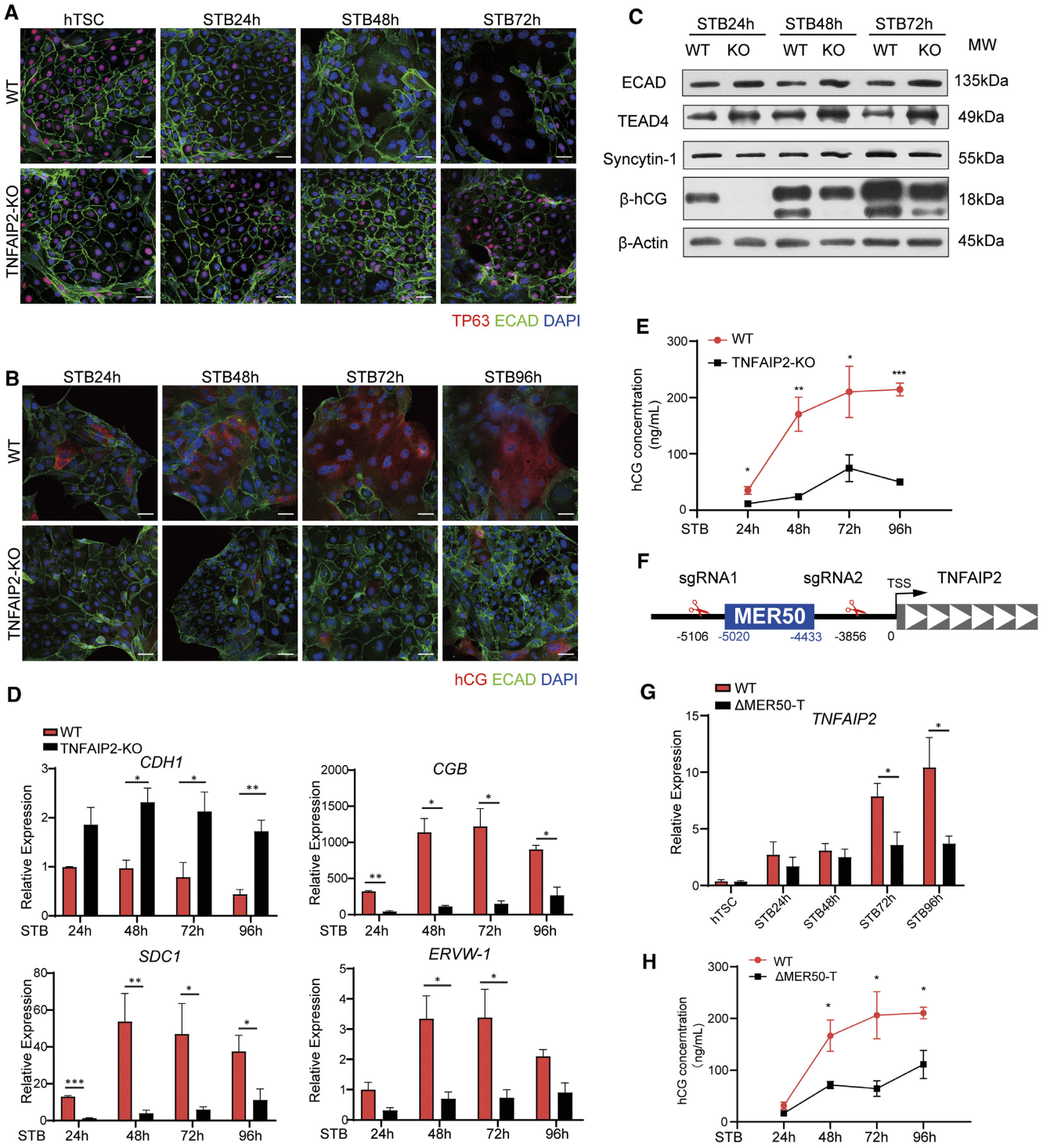

Figure 5. The expression of TNFAIP2 governed by the MER50 enhancer sustains syncytialization extent. (**A**) Representative immunostaining of ECAD (green) and TP63 (red) shows the stemness of the WT/TNFAIP2 KO hTSCs and STBs. Scale bar, 50 μm. (**B**) Immunostaining for ECAD (green) and hCG (red) in the WT/TNFAIP2 KO STB at the indicated timepoints of syncytialization. Nuclei were stained with DAPI. Scale bar, 50 μm. (**C**) Western blots of ECAD, TEAD4, Syncytin-1 and β-hCG in the WT and TNFAIP2 KO cells. β-Actin as loading control. (**D**) mRNA levels of *CDH1, CGB, SDC1* and *ERVW-1* in the WT/TNFAIP2 KO STB at 24, 48, 72, 96 h. Results represent the mean ± SEM of at least three independent experiments. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, multiple unpaired *t* tests. (**E**) ELISA detection of hCG secretion from WT and TNFAIP2 KO STBs at the indicated timepoints. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, multiple unpaired *t* tests. $n = 4$. (**F**) Schematic diagram showing CRISPR deletion of the MER50 enhancer element close to *TNFAIP2* targeted by the two flanking sgRNAs. The WT control in (E) and (H) refers to the same control cell line. (**G**) mRNA levels of *TNFAIP2* in the WT/ΔMER50-T hTSC and STB at the 24, 48, 72, 96 h timpoints of syncytialization. $^{*}P < 0.05$, *t*-test. Results represent the mean ± SEM of at least three independent experiments. (**H**) ELISA detection of hCG in culture medium from the WT and ΔMER50-T STBs at the indicated timepoints. $^{*}P < 0.05$, Multiple unpaired *t* tests. $n = 4$.

different in hTSC, in which ERVs display hotspots of bivalent histone modifications. We illustrated that ERVs are overrepresented in the enhancers for both hTSC and STB, particularly the shared enhancers with bivalency in hTSC, which is coincident with the previous observation that ERVs occupy about 80% of mouse TSC enhancers (41). Of note, although certain promoters are occupied by H3K4me3 and H3K9me3 in the same chromatin region in mouse TSCs (65), our study found rare enrichment of ERV in bivalent promoter regions in human trophoblasts, which might be due to the species-specific epigenetic discrepancy between mouse and human TSCs. Taken together, our study, along with others, characterizes the link between epigenome bivalency within ERV elements and differentiation plasticity. Moreover, the signature of bivalent histone modifications in hTSC suggests that the assessment of the sole monovalent mark may not reflect the true chromatin state and transcriptional activity especially in lineage-commitment cells, highlighting the importance of comprehensive epigenome surveys with diverse histone marks.

Intriguingly, we identified H3K9me3 as the repressive histone mark in the bivalent enhancers instead of the canonical H3K27me3 modification present in bivalent domains in ESCs. Recent studies have shown that PRC2-mediated H3K27me3 restricts human trophoblast induction from naïve ESCs (66,67), in line with the global low level of H3K27me3 detected in mouse extraembryonic stem cells (65), which suggests that H3K9me3 may serve as the primary repressive histone modification both in human and mouse TSCs. As a hallmark of heterochromatin modification, the importance of H3K9me3 for the early stages of trophoblast development is gaining increasing attention, since H3K9me3-marked heterochromatin domains set a barrier for cell fate reprogramming in TSCs (68,69). H3K9me3-mediated heterochromatin might be more significant in TSCs than in other cell types to silence lineage-incompatible genes and repetitive elements to maintain placental development and function (70), considering the overall hypomethylation signature of the TSC genome is less likely to dominate gene repression. In our study, we found the H3K9me3 modification of ERVs in TSCs is critical not simply to prevent retrotransposition, but also to establish the poised ERV enhancers harboring bivalent H3K9me3/H3K27ac occupancy in TSC while remarkably resolved H3K9me3 during differentiation towards STBs. Of note, the biological significance regarding the replacement of H3K27me3 with H3K9me3 in trophoblast lineages remains unclear. Moreover, although lineage-specific DNA methylation and SETDB1 recruitment has been proposed as key factors controlling the transition from H3K27me3 to H3K9me3 in the cell fate specification process (71), the specific H3K9me3 modifiers in trophoblast stem cell and the corresponding contribution to trophoblast differentiation require further investigation.

The transcriptome dynamics concomitant to trophoblast syncytialization raises an intriguing question of how the transcriptional programs are fine-tuned to shape a coordinated gene expression pattern accounting for cell fusion. Recent studies have focused on the co-option of ERV-derived regulatory elements for functional adaptations in the placenta (41,72). One of the transposon element families, MEdium Reiteration frequency interspersed repeats (MERs) are a cluster of DNA transposons and ERVs, which were recently linked to the evolution of mammalian pregnancy. For instance, the previous study suggests that MER41A/B elements associated with lineage-specific human placental enhancers create dozens of serum response factor binding motifs (24). In addition, the MER20 regulates progesterone- and cAMP-dependent gene expression through direct binding of pregnancy-related factors, contributing to a novel gene regulatory network dedicated to establishing pregnancy in mammals (73). The MER50 repeat was identified in the 1990s as a new ERV member in the human genome, which derived from the genome of an ancestor of the Simiiformes and was classified into the ERV1 family (74). Sun *et al.* verified that MER50 is overrepresented in lineage-specific human placental enhancers (24), however, the functional significance of MER50 was unknown. Rigorous functional validations are required to address the importance of ERV enhancers, since the observation that within hundreds of TE with enhancer-like epigenetic signatures only a small proportion can actually display definitive enhancer activity in mouse TSCs (72). Our current study shows solid evidence to prove the definitive function of MER50 in trophoblast differentiation and the corresponding gene regulation. To our knowledge, we identified the first STB formation-related ERV enhancer, that profoundly modulates the transcriptional program underlying syncytium establishment. Given that MER50 is a primate-specific young ERV with recent integrations in the Simiiformes, it will be interesting to define whether the difference in the interspecies divergence of gene expression between primate and rodent placentae is associated with the emergence of this ERV element.

Syncytialization is an important cellular process for placental development, while the underlying genetic and epigenetic regulation mechanisms are still poorly understood (4,75). Our data demonstrate that a panel of MER50-related genes, including *ERVFRD-1*, *MFSD2A*, *TNFAIP2* and *RAI14*, play crucial roles in modulating the syncytialization of hTSC. *TNFAIP2* was previously shown as a canonical inflammatory cytokine TNF-induced gene, whereas recent studies reported the Tnfaip2-dependent induction of lipid biosynthesis essential for mESC differentiation (55) and the regulatory effect of TNFAIP2 on tumorigenesis (76). Our *in vitro* experiments clarified that TNFAIP2 expression driven by the MER50 enhancer is essential for STB formation, whereas how the TNFAIP2 expression directs syncytialization and what are the *TNFAIP2*-MER50 enhancer-related transcriptional factors remain to be investigated. Notably, the contribution of MER50 enhancers to trophoblast cell fusion is bidirectional. For example, mutation of the MER50 target gene *RAI14* accelerates the process of trophoblast fusion, which is in line with the inhibitory effect of other MER50 enhancer-related STB genes, *IFITM2/3*, on trophoblast syncytialization in response to interferon signals (48). Whereas most previous studies showcased a particular ERV regulating a single gene, our results highlight the possibility that certain ERVs, MER50 for example, may exert a more extensive influence on the regulatory networks of cell fate specification by orchestrating a full panel of genes that function in concert.

In summary, the dynamics of the epigenetic landscape of STB differentiation reveals an extensive rewiring of bivalent

ERV enhancer that reinforces syncytialization progression. Our study identifies the first ERV enhancer with bivalent histone modification, MER50, facilitating the transcriptional regulation of adjacent STB genes, among which new syncytialization regulators were determined. Together, our study highlights that ERV-derived enhancers, MER50 specifically, fine-tune the transcriptional program accounting for human trophoblast syncytialization, which sheds light on a novel epigenetic mechanism of placental development.

DATA AVAILABILITY

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE209977.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This study utilized the computational resources of Yangzhou University College of Veterinary Medicine High-Performance Computing cluster. The authors would like to thank Drs Hiroaki Okae and Takahiro Arima for authorizing us access to the hTSC ChIP-Seq data; the members of the Cao and Sun labs for helpful advice and discussions.

FUNDING

National Key Research and Development Program of China [2022YFC2702403 to B.C., 2022YFC2704702 to X.H.]; National Natural Sciences Foundation in China [82130047, 81971414 to B.C., 32270584, 31900422 to M.A.S.]; Natural Sciences Foundation of Fujian Province of China [2020J06003 to B.C.]; 111 Project D18007 and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) (to M.A.S.). Funding for open access charge: National Natural Sciences Foundation in China [82130047, 81971414 to BC].
Conflict of interest statement. None declared.

REFERENCES

1. Turco,M.Y. and Moffett,A. (2019) Development of the human placenta. *Development*, **146**, dev163428.
2. Redman,C.W. and Staff,A.C. (2015) Preeclampsia, biomarkers, syncytiotrophoblast stress, and placental capacity. *Am. J. Obstet. Gynecol.*, **213**, S9–S11.
3. Lea,R.G., Tulppala,M. and Critchley,H.O. (1997) Deficient syncytiotrophoblast tumour necrosis factor-alpha characterizes failing first trimester pregnancies in a subgroup of recurrent miscarriage patients. *Hum. Reprod.*, **12**, 1313–1320.
4. Shao,X., Cao,G., Chen,D., Liu,J., Yu,B., Liu,M., Li,Y.X., Cao,B., Sadovsky,Y. and Wang,Y.L. (2021) Placental trophoblast syncytialization potentiates macropinocytosis via mTOR signaling to adapt to reduced amino acid supply. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2017092118.
5. Papuchova,H. and Latos,P.A. (2022) Transcription factor networks in trophoblast development. *Cell. Mol. Life Sci.*, **79**, 337.
6. Renaud,S.J. and Jeyarajah,M.J. (2022) How trophoblasts fuse: an in-depth look into placental syncytiotrophoblast formation. *Cell. Mol. Life Sci.*, **79**, 433.
7. Emera,D. and Wagner,G.P. (2012) Transposable element recruitments in the mammalian placenta: impacts and mechanisms. *Brief. Funct. Genomics*, **11**, 267–276.
8. Gifford,W.D., Pfaff,S.L. and Macfarlan,T.S. (2013) Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.*, **23**, 218–226.
9. Chuong,E.B. (2013) Retroviruses facilitate the rapid evolution of the mammalian placenta. *Bioessays*, **35**, 853–861.
10. Dunlap,K.A., Palmarini,M., Varela,M., Burghardt,R.C., Hayashi,K., Farmer,J.L. and Spencer,T.E. (2006) Endogenous retroviruses regulate periimplantation placental growth and differentiation. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 14390–14395.
11. Chuong,E.B. (2018) The placenta goes viral: retroviruses control gene expression in pregnancy. *PLoS Biol.*, **16**, e3000028.
12. Dittmar,T., Weiler,J., Luo,T. and Hass,R. (2021) Cell-Cell Fusion Mediated by Viruses and HERV-Derived Fusogens in Cancer Initiation and Progression. *Cancers (Basel)*, **13**, 5363.
13. Dupressoir,A., Vernochet,C., Harper,F., Guegan,J., Dessen,P., Pierron,G. and Heidmann,T. (2011) A pair of co-opted retroviral envelope syncytin genes is required for formation of the two-layered murine placental syncytiotrophoblast. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, E1164–E1173.
14. Lavialle,C., Cornelis,G., Dupressoir,A., Esnault,C., Heidmann,O., Vernochet,C. and Heidmann,T. (2013) Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **368**, 20120507.
15. Branco,M.R. and Chuong,E.B. (2020) Crossroads between transposons and gene regulation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **375**, 20190330.
16. Thompson,P.J., Macfarlan,T.S. and Lorincz,M.C. (2016) Long terminal repeats: from parasitic elements to building blocks of the transcriptional regulatory repertoire. *Mol. Cell*, **62**, 766–776.
17. Chuong,E.B., Elde,N.C. and Feschotte,C. (2017) Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.*, **18**, 71–86.
18. Hsu,P.S., Yu,S.H., Tsai,Y.T., Chang,J.Y., Tsai,L.K., Ye,C.H., Song,N.Y., Yau,L.C. and Lin,S.P. (2021) More than causing (epi)genomic instability: emerging physiological implications of transposable element modulation. *J. Biomed. Sci.*, **28**, 58.
19. Fuentes,D.R., Swigut,T. and Wysocka,J. (2018) Systematic perturbation of retroviral LTRs reveals widespread long-range effects on human gene regulation. *Elife*, **7**, e35989.
20. Macfarlan,T.S., Gifford,W.D., Driscoll,S., Lettieri,K., Rowe,H.M., Bonanomi,D., Firth,A., Singer,O., Trono,D. and Pfaff,S.L. (2012) Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature*, **487**, 57–63.
21. Wilson,K.D., Ameen,M., Guo,H., Abilez,O.J., Tian,L., Mumbach,M.R., Diecke,S., Qin,X., Liu,Y., Yang,H. *et al.* (2020) Endogenous retrovirus-derived lncRNA BANCR promotes cardiomyocyte migration in humans and non-human primates. *Dev. Cell*, **54**, 694–709.
22. Sakashita,A., Maezawa,S., Takahashi,K., Alavattam,K.G., Yukawa,M., Hu,Y.C., Kojima,S., Parrish,N.F., Barski,A., Pavlicev,M. *et al.* (2020) Endogenous retroviruses drive species-specific germline transcriptomes in mammals. *Nat. Struct. Mol. Biol.*, **27**, 967–977.
23. Dunn-Fletcher,C.E., Muglia,L.M., Pavlicev,M., Wolf,G., Sun,M.A., Hu,Y.C., Huffman,E., Tumukuntala,S., Thiele,K., Mukherjee,A. *et al.* (2018) Anthropoid primate-specific retroviral element THE1B controls expression of CRH in placenta and alters gestation length. *PLoS Biol.*, **16**, e2006337.
24. Sun,M.A., Wolf,G., Wang,Y., Senft,A.D., Ralls,S., Jin,J., Dunn-Fletcher,C.E., Muglia,L.J. and Macfarlan,T.S. (2021) Endogenous retroviruses drive lineage-specific regulatory evolution across primate and rodent placentae. *Mol. Biol. Evol.*, **38**, 4992–5004.
25. Okae,H., Toh,H., Sato,T., Hiura,H., Takahashi,S., Shirane,K., Kabayama,Y., Suyama,M., Sasaki,H. and Arima,T. (2018) Derivation of human trophoblast stem cells. *Cell Stem Cell*, **22**, 50–63.
26. Turco,M.Y., Gardner,L., Kay,R.G., Hamilton,R.S., Prater,M., Hollinshead,M.S., McWhinnie,A., Esposito,L., Fernando,R., Skelton,H. *et al.* (2018) Trophoblast organoids as a model for maternal-fetal interactions during human placentation. *Nature*, **564**, 263–267.
27. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

28. Liao,Y., Smyth,G.K. and Shi,W. (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.*, **41**, e108.
29. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
30. Erokhin,M., Gorbenko,F., Lomaev,D., Mazina,M.Y., Mikhailova,A., Garaev,A.K., Parshikov,A., Vorobyeva,N.E., Georgiev,P., Schedl,P. *et al.* (2021) Boundaries potentiate polycomb response element-mediated silencing. *BMC Biol.*, **19**, 113.
31. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
32. 1000 Genome Project Data Processing Subgroup, Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
33. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
34. Amemiya,H.M., Kundaje,A. and Boyle,A.P. (2019) The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.*, **9**, 9354.
35. Yates,A.D., Achuthan,P., Akanni,W., Allen,J., Allen,J., Alvarez-Jarreta,J., Amode,M.R., Armean,I.M., Azov,A.G., Bennett,R. *et al.* (2020) Ensembl 2020. *Nucleic Acids Res.*, **48**, D682–D688.
36. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
37. R.C. Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
38. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
39. Thorvaldsdottir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*, **14**, 178–192.
40. He,J., Fu,X., Zhang,M., He,F., Li,W., Abdul,M.M., Zhou,J., Sun,L., Chang,C., Li,Y. *et al.* (2019) Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells. *Nat. Commun.*, **10**, 34.
41. Chuong,E.B., Rumi,M.A., Soares,M.J. and Baker,J.C. (2013) Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat. Genet.*, **45**, 325–329.
42. Blanco,E., Gonzalez-Ramirez,M., Alcaine-Colet,A., Aranda,S. and Di Croce,L. (2020) The bivalent genome: characterization, structure, and regulation. *Trends Genet.*, **36**, 118–131.
43. Chuong,E.B., Elde,N.C. and Feschotte,C. (2016) Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*, **351**, 1083–1087.
44. Vargas,A., Moreau,J., Landry,S., LeBellego,F., Toufaily,C., Rassart,E., Lafond,J. and Barbeau,B. (2009) Syncytin-2 plays an important role in the fusion of human trophoblast cells. *J. Mol. Biol.*, **392**, 301–318.
45. Toufaily,C., Vargas,A., Lemire,M., Lafond,J., Rassart,E. and Barbeau,B. (2013) MFSD2a, the Syncytin-2 receptor, is important for trophoblast fusion. *Placenta*, **34**, 85–88.
46. Singh,A.T., Dharmarajan,A., Aye,I.L. and Keelan,J.A. (2012) Sphingosine-sphingosine-1-phosphate pathway regulates trophoblast differentiation and syncytialization. *Reprod. Biomed. Online*, **24**, 224–234.
47. Camolotto,S., Racca,A., Rena,V., Nores,R., Patrito,L.C., Genti-Raimondi,S. and Panzetta-Dutari,G.M. (2010) Expression and transcriptional regulation of individual pregnancy-specific glycoprotein genes in differentiating trophoblast cells. *Placenta*, **31**, 312–319.
48. Buchrieser,J., Degrelle,S.A., Couderc,T., Nevers,Q., Disson,O., Manet,C., Donahue,D.A., Porrot,F., Hillion,K.H., Perthame,E. *et al.* (2019) IFITM proteins inhibit placental syncytiotrophoblast formation and promote fetal demise. *Science*, **365**, 176–180.
49. Esnault,C., Cornelis,G., Heidmann,O. and Heidmann,T. (2013) Differential evolutionary fate of an ancestral primate endogenous retrovirus envelope gene, the EnvV syncytin, captured for a function in placentation. *PLoS Genet.*, **9**, e1003400.
50. Roberts,R.M., Ezashi,T., Schulz,L.C., Sugimoto,J., Schust,D.J., Khan,T. and Zhou,J. (2021) Syncytins expressed in human placental trophoblast. *Placenta*, **113**, 8–14.
51. Mi,S., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L., LaVallie,E., Tang,X.Y., Edouard,P., Howes,S. *et al.* (2000) Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*, **403**, 785–789.
52. Blaise,S., de Parseval,N., Benit,L. and Heidmann,T. (2003) Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 13013–13018.
53. Lu,X., Wang,R., Zhu,C., Wang,H., Lin,H.Y., Gu,Y., Cross,J.C. and Wang,H. (2018) Fine-tuned and cell-cycle-restricted expression of fusogenic protein syncytin-2 maintains functional placental syncytia. *Cell Rep.*, **23**, 3979.
54. Yu,Y.L., Chen,M., Zhu,H., Zhuo,M.X., Chen,P., Mao,Y.J., Li,L.Y., Zhao,Q., Wu,M. and Ye,M. (2021) STAT1 epigenetically regulates LCP2 and TNFAIP2 by recruiting EP300 to contribute to the pathogenesis of inflammatory bowel disease. *Clin. Epigenetics*, **13**, 127.
55. Deb,S., Felix,D.A., Koch,P., Deb,M.K., Szafranski,K., Buder,K., Sannai,M., Groth,M., Kirkpatrick,J., Pietsch,S. *et al.* (2021) Tnfaip2/exoc3-driven lipid metabolism is essential for stem cell differentiation and organ homeostasis. *EMBO Rep.*, **22**, e49328.
56. Wolf,F.W., Sarma,V., Seldin,M., Drake,S., Suchard,S.J., Shao,H., O'Shea,K.S. and Dixit,V.M. (1994) B94, a primary response gene inducible by tumor necrosis factor-alpha, is expressed in developing hematopoietic tissues and the sperm acrosome. *J. Biol. Chem.*, **269**, 3633–3640.
57. Jia,L., Shi,Y., Wen,Y., Li,W., Feng,J. and Chen,C. (2018) The roles of TNFAIP2 in cancers and infectious diseases. *J. Cell. Mol. Med.*, **22**, 5188–5195.
58. Hemberger,M., Hanna,C.W. and Dean,W. (2020) Mechanisms of early placental development in mouse and humans. *Nat. Rev. Genet.*, **21**, 27–43.
59. Yan,J. and Huangfu,D. (2022) Epigenome rewiring in human pluripotent stem cells. *Trends Cell Biol.*, **32**, 259–271.
60. Xu,R., Li,C., Liu,X. and Gao,S. (2021) Insights into epigenetic patterns in mammalian early embryos. *Protein Cell*, **12**, 7–28.
61. Sun,H., Wang,Y., Wang,Y., Ji,F., Wang,A., Yang,M., He,X. and Li,L. (2022) Bivalent Regulation and Related Mechanisms of H3K4/27/9me3 in Stem Cells. *Stem Cell Rev. Rep.*, **18**, 165–178.
62. Vastenhouw,N.L. and Schier,A.F. (2012) Bivalent histone modifications in early embryogenesis. *Curr. Opin. Cell Biol.*, **24**, 374–386.
63. Bernstein,B.E., Mikkelsen,T.S., Xie,X., Kamal,M., Huebert,D.J., Cuff,J., Fry,B., Meissner,A., Wernig,M., Plath,K. *et al.* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, **125**, 315–326.
64. Harikumar,A. and Meshorer,E. (2015) Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep.*, **16**, 1609–1619.
65. Rugg-Gunn,P.J., Cox,B.J., Ralston,A. and Rossant,J. (2010) Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 10783–10790.
66. Kumar,B., Navarro,C., Winblad,N., Schell,J.P., Zhao,C., Weltner,J., Baque-Vidal,L., Salazar Mantero,A., Petropoulos,S., Lanner,F. *et al.* (2022) Polycomb repressive complex 2 shields naive human pluripotent cells from trophectoderm differentiation. *Nat. Cell Biol.*, **24**, 845–857.
67. Zijlmans,D.W., Talon,I., Verhelst,S., Bendall,A., Van Nerum,K., Javali,A., Malcolm,A.A., van Knippenberg,S., Biggins,L., To,S.K. *et al.* (2022) Integrated multi-omics reveal polycomb repressive complex 2 restricts human trophoblast induction. *Nat. Cell Biol.*, **24**, 858–871.
68. Nicetto,D. and Zaret,K.S. (2019) Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Curr. Opin. Genet. Dev.*, **55**, 1–10.
69. Hada,M., Miura,H., Tanigawa,A., Matoba,S., Inoue,K., Ogonuki,N., Hirose,M., Watanabe,N., Nakato,R., Fujiki,K. *et al.* (2022) Highly rigid H3.1/H3.2-H3K9me3 domains set a barrier for cell fate reprogramming in trophoblast stem cells. *Genes Dev.*, **36**, 84–102.
70. Wang,C., Liu,X., Gao,Y., Yang,L., Li,C., Liu,W., Chen,C., Kou,X., Zhao,Y., Chen,J. *et al.* (2018) Reprogramming of

H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat. Cell Biol.*, **20**, 620–631.
71. Matsumura,Y., Nakaki,R., Inagaki,T., Yoshida,A., Kano,Y., Kimura,H., Tanaka,T., Tsutsumi,S., Nakao,M., Doi,T. *et al.* (2015) H3K4/H3K9me3 bivalent chromatin domains targeted by lineage-specific DNA methylation pauses adipocyte differentiation. *Mol. Cell*, **60**, 584–596.
72. Todd,C.D., Deniz,O., Taylor,D. and Branco,M.R. (2019) Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells. *Elife*, **8**, e44344.
73. Emera,D. and Wagner,G.P. (2012) Transformation of a transposon into a derived prolactin promoter with function during human pregnancy. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 11246–11251.
74. Jurka,J., Kapitonov,V.V., Klonowski,P., Walichiewicz,J. and Smit,A.F. (1996) Identification of new medium reiteration frequency repeats in the genomes of Primates, Rodentia and Lagomorpha. *Genetica*, **98**, 235–247.
75. Liang,C.Y., Wang,L.J., Chen,C.P., Chen,L.F., Chen,Y.H. and Chen,H. (2010) GCM1 regulation of the expression of syncytin 2 and its cognate receptor MFSD2A in human placenta. *Biol. Reprod.*, **83**, 387–395.
76. Jia,L., Zhou,Z., Liang,H., Wu,J., Shi,P., Li,F., Wang,Z., Wang,C., Chen,W., Zhang,H. *et al.* (2016) KLF5 promotes breast cancer proliferation, migration and invasion in part by upregulating the transcription of TNFAIP2. *Oncogene*, **35**, 2040–2051.