# *Supplemental Material*

## Supplementary Methods

### Generation of targeted alleles

Targeting vectors containing 5' and 3' arms homologous to the *Eomes* locus, a FLP recognition target (FRT) flanked PGK.Neomycin selection cassette and a PGK.DTA (diphtheria toxin A) cassette for negative selection. The ΔVPE targeting vector was generated by recombineering using oligos listed in Table S1, designed to delete 656bp of the VPE. The ΔPSE_b vector includes a 5' 5.8kb SpeI-EcoRV fragment and a 3' 5kb KpnI-EcoRI fragment of the *Eomes* locus, and deletes 2019bp of PSE_b. The PSE vector comprises a 5' 5.6kb AatII-Bsu36I fragment, where the upstream AatII site was introduced by PCR (Table S1), and a 3' 5kb KpnI-EcoRI fragment of the *Eomes* locus, resulting in deletion of 4775bp of the PSE. XhoI (PSE, PSE_b) or ApaLI (VPE) linearized vectors (15ug) were electroporated into CCE ES cells, and $Eomes^{GFP/+}$ cells. Screening of drug resistant ESC clones was carried out by Southern blot analysis with the restriction enzymes and probes summarised in Fig. S2, S3, S4 and S5 using standard protocols (Behringer et al., 2013).

## Supplemental Figures

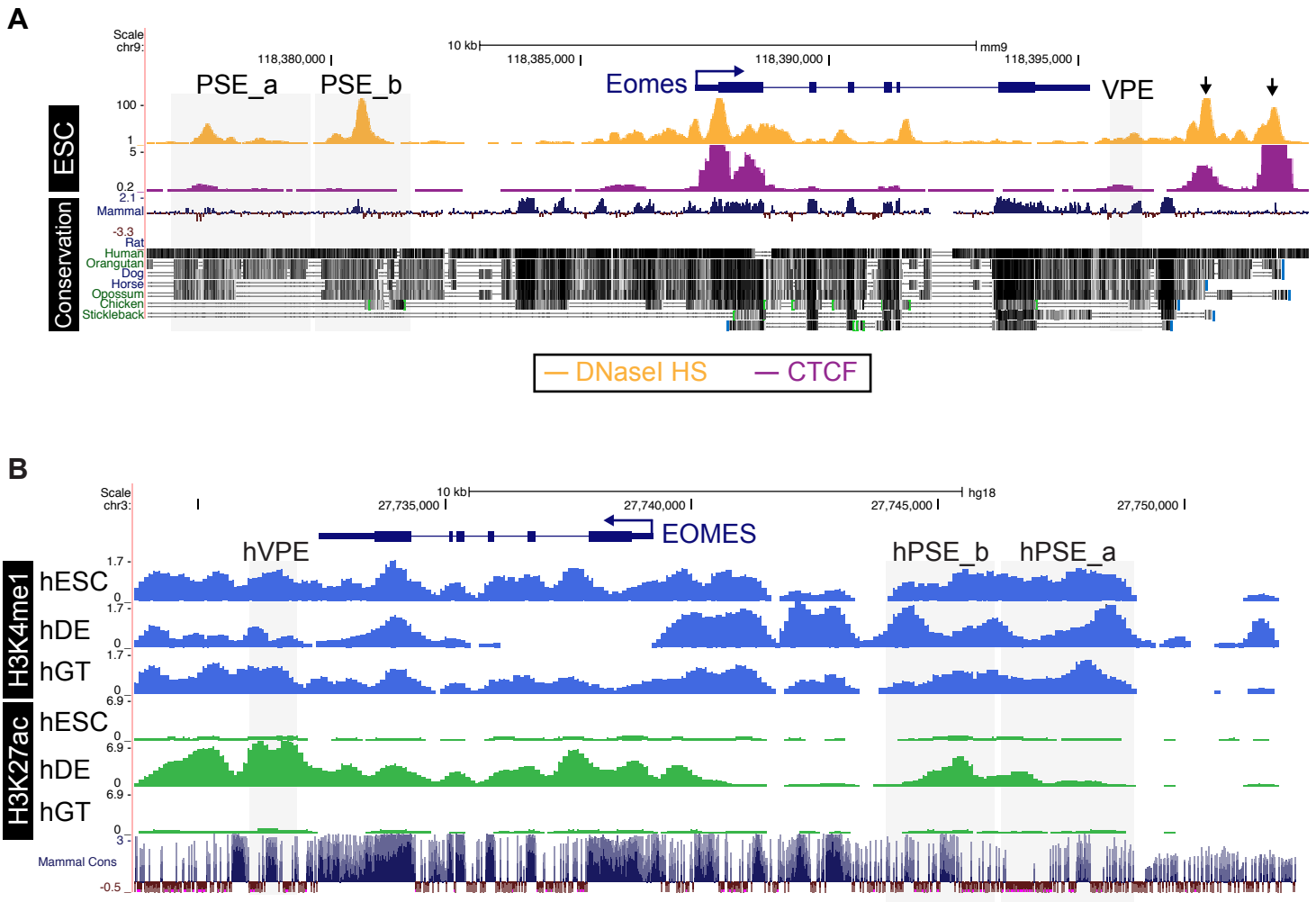



**Figure S1: PSE and VPE enhancers are conserved in human**

(A) DNaseI hypersensitivity (HS) and ChIP-seq of CTCF in ESC (Consortium, 2012). Conservation at the Eomes locus across vertebrates (UCSC browser, mm9). Boxes indicate PSE_a, PSE_b, and VPE enhancer regions, highly conserved amongst mammals. Arrows indicate CTCF bound regions downstream of the VPE. (B) ChIP-seq of H3K27ac and H3K4me1 histone modifications at the Eomes locus in human ESC (hESC), definitive endoderm (hDE) and human gut tube (hGT) (UCSC browser, hg18) (Wang et al., 2015). Homologous regions to the mouse VPE and PSE are associated with these active enhancer marks and are highlighted in grey. Human VPE, PSE_a and PSE_b (hVPE, hPSE_a, hPSE_b)

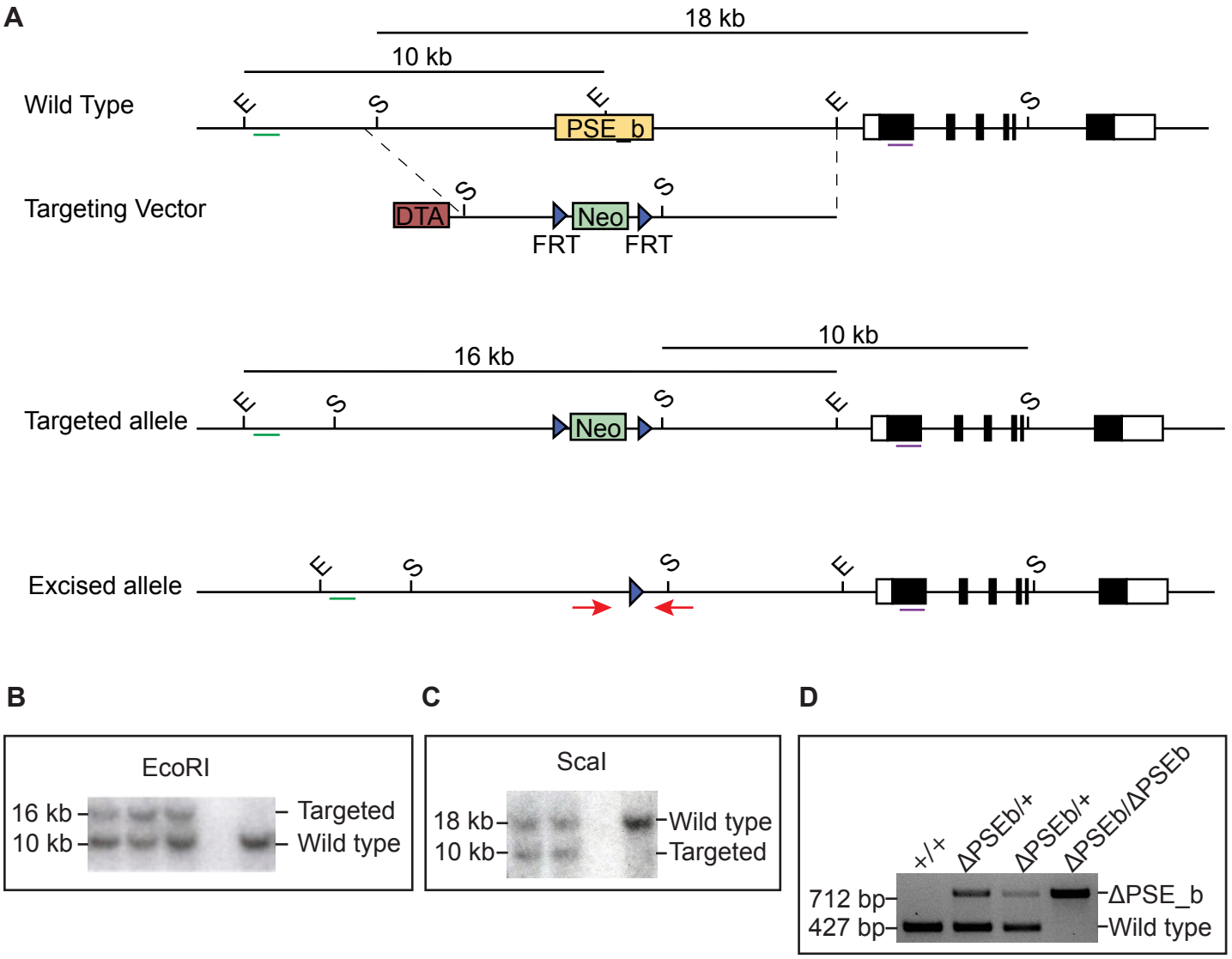



**Figure S2: Targeted deletion of the PSE_b sub-region**

(A) Targeting strategy to delete the 2kb PSE_b region (chr9:118379552-118381570; mm9) by homologous recombination. Southern blot restriction digest used for screening are indicated together with the probes (green and blue bars) and expected fragment sizes for the correctly targeted allele. EcoRI (E), ScaI (S), FLP-recombinase recognition site (FRT) site, Neomycin resistance cassette (Neo), Diphtheria toxin A cassette (DTA). Red arrows indicate primers for verifying FLP excision. (B,C) Southern blot of successfully targeted ESC clones. (D) PCR genotyping of $Eomes^{\Delta PSE\_b}$ mice.

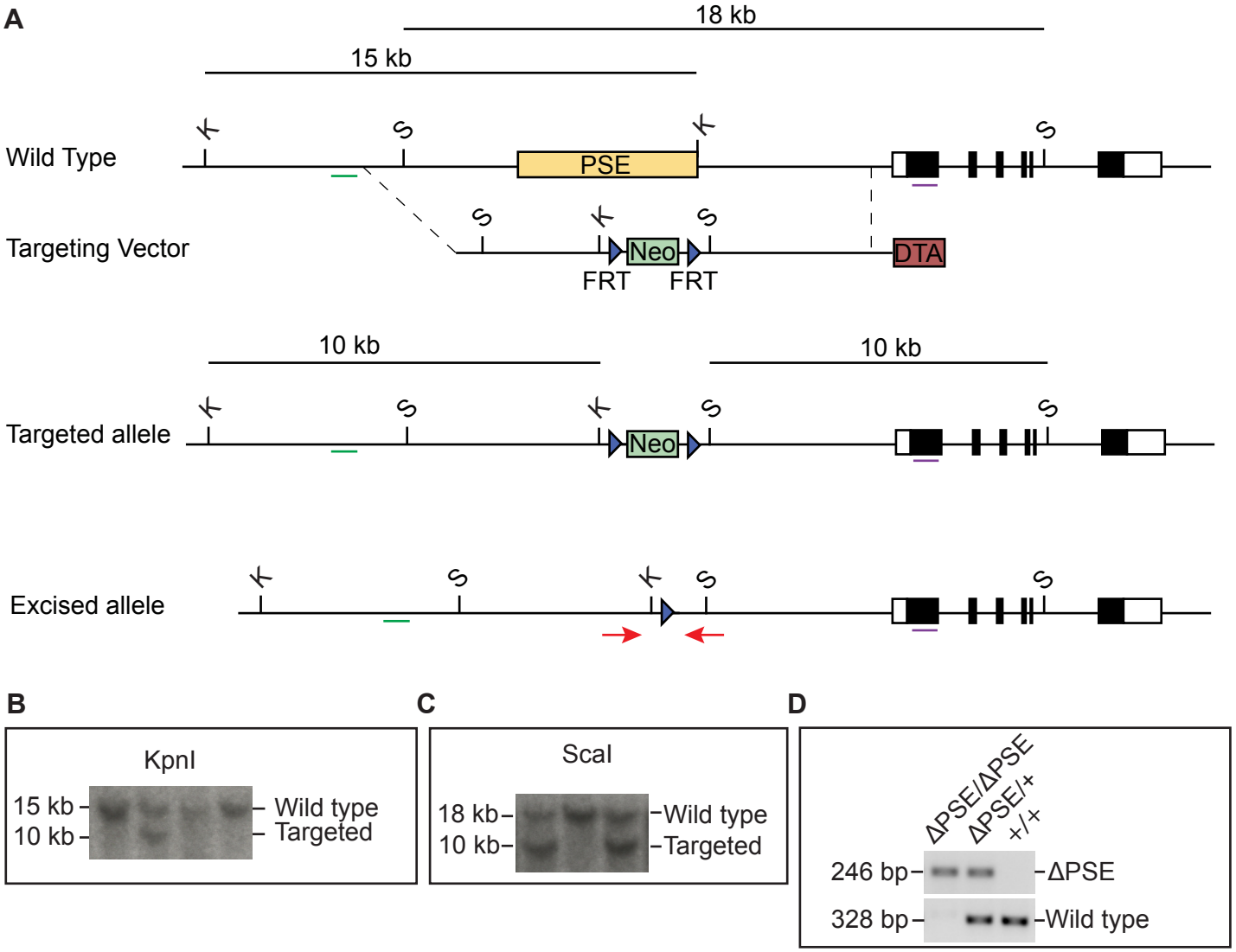
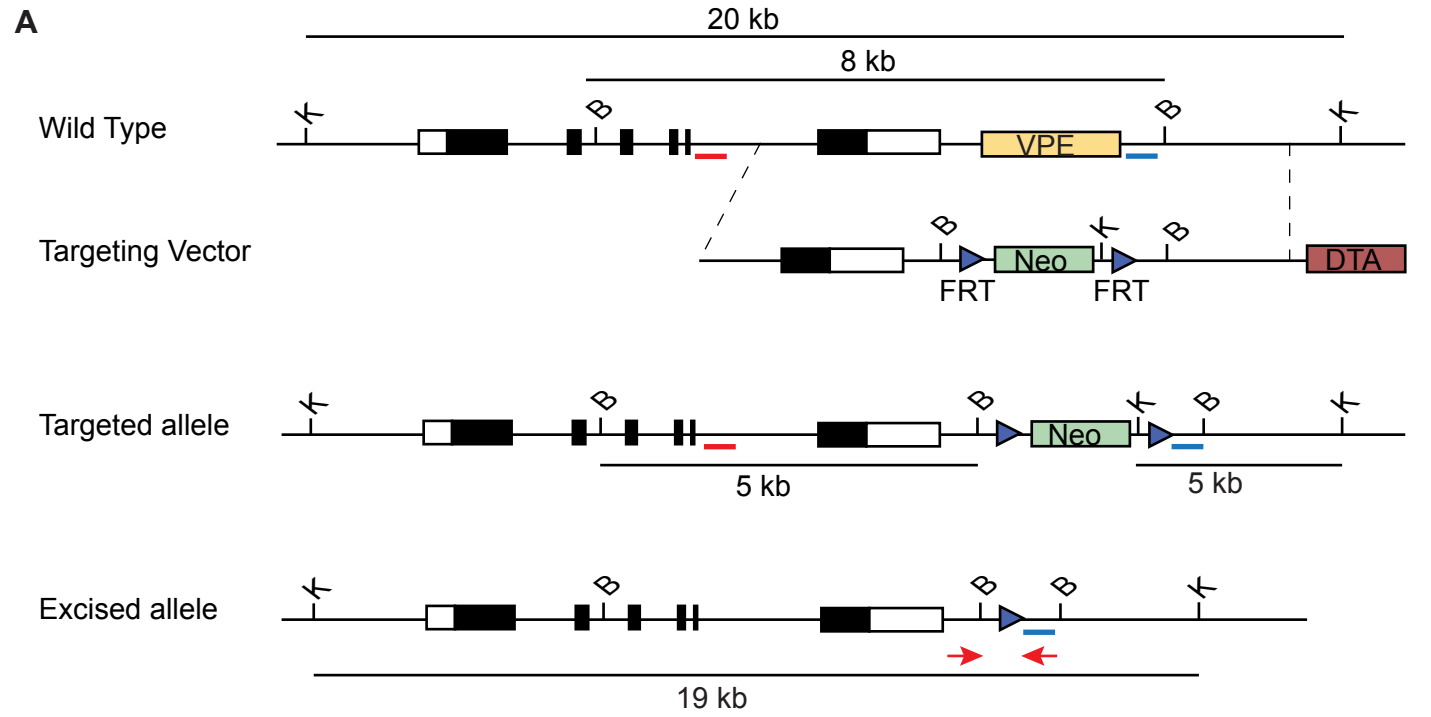
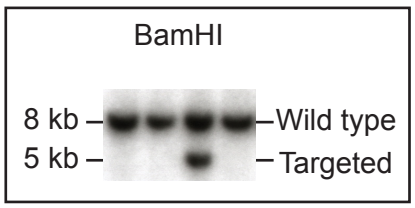
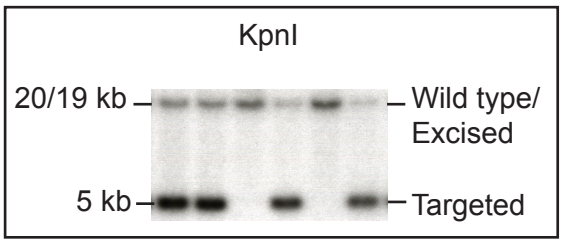


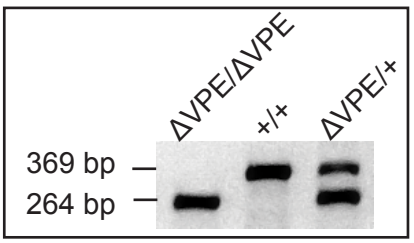

**Figure S3: Targeted deletion of the PSE region**

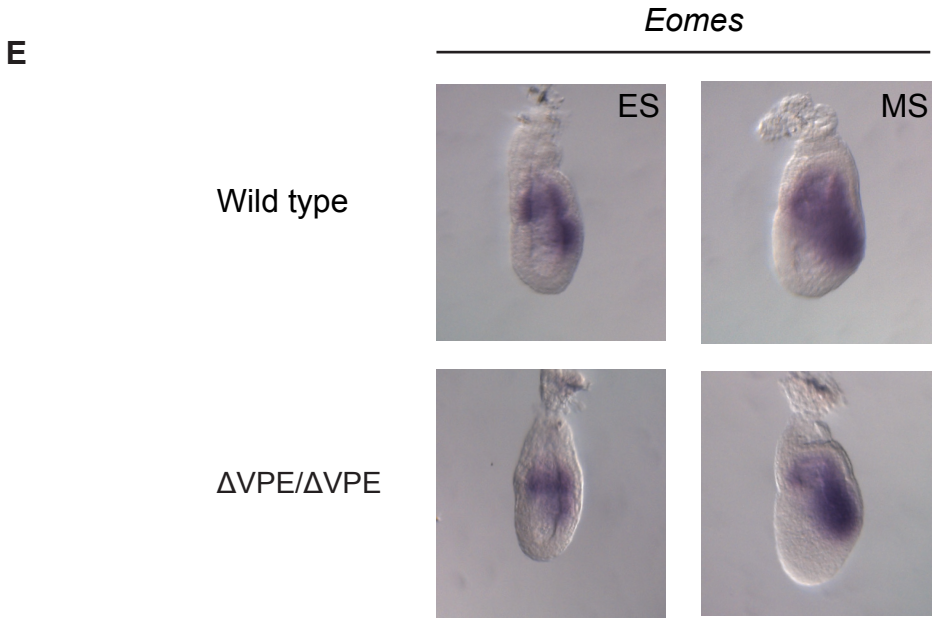
(A) Targeting strategy to delete the 5kb PSE region (chr9:118376796-118381570; mm9) by homologous recombination. Southern blot restriction digest used for screening are indicated together with the probes (green and blue bars) and expected fragment sizes for the correctly targeted allele. KpnI (K), ScaI (S), FLP-recombinase recognition site (FRT) site, Neomycin resistance cassette (Neo), Diphtheria toxin A cassette (DTA). Red arrows indicate primers for verifying FLP excision. (B,C) Southern blot of successfully targeted ESC clones. (D) PCR genotyping of Eomes$^{\Delta PSE}$ mice.



## Figure S4: Targeted deletion of the VPE region

(A) Targeting strategy to delete the 0.7kb VPE region (chr9:118395625-118396280; mm9) by homologous recombination. Southern blot probes (red and blue bars), restriction digests and expected fragment sizes are indicated for the targeted and excised alleles. BamHI (B), KpnI (K), FLP-recombinase recognition site (FRT) site, Neomycin resistance cassette (Neo), Diphtheria toxin A cassette (DTA). Red arrows indicate primers for verifying FLP excision. (B) Southern blot of targeted ESC clones. (C) Southern blot to identify excision of Neo cassette in targeted ESC clones. (D) PCR genotyping ΔVPE allele in mice derived from $Eomes^{\Delta VPE/+}$ intercrosses. (E) Whole-mount *in situ* hybridisation of Eomes transcripts at early mid-streak stages shows Eomes expression domains are unaltered in $Eomes^{\Delta VPE/\Delta VPE}$ compared to wild type embryos.

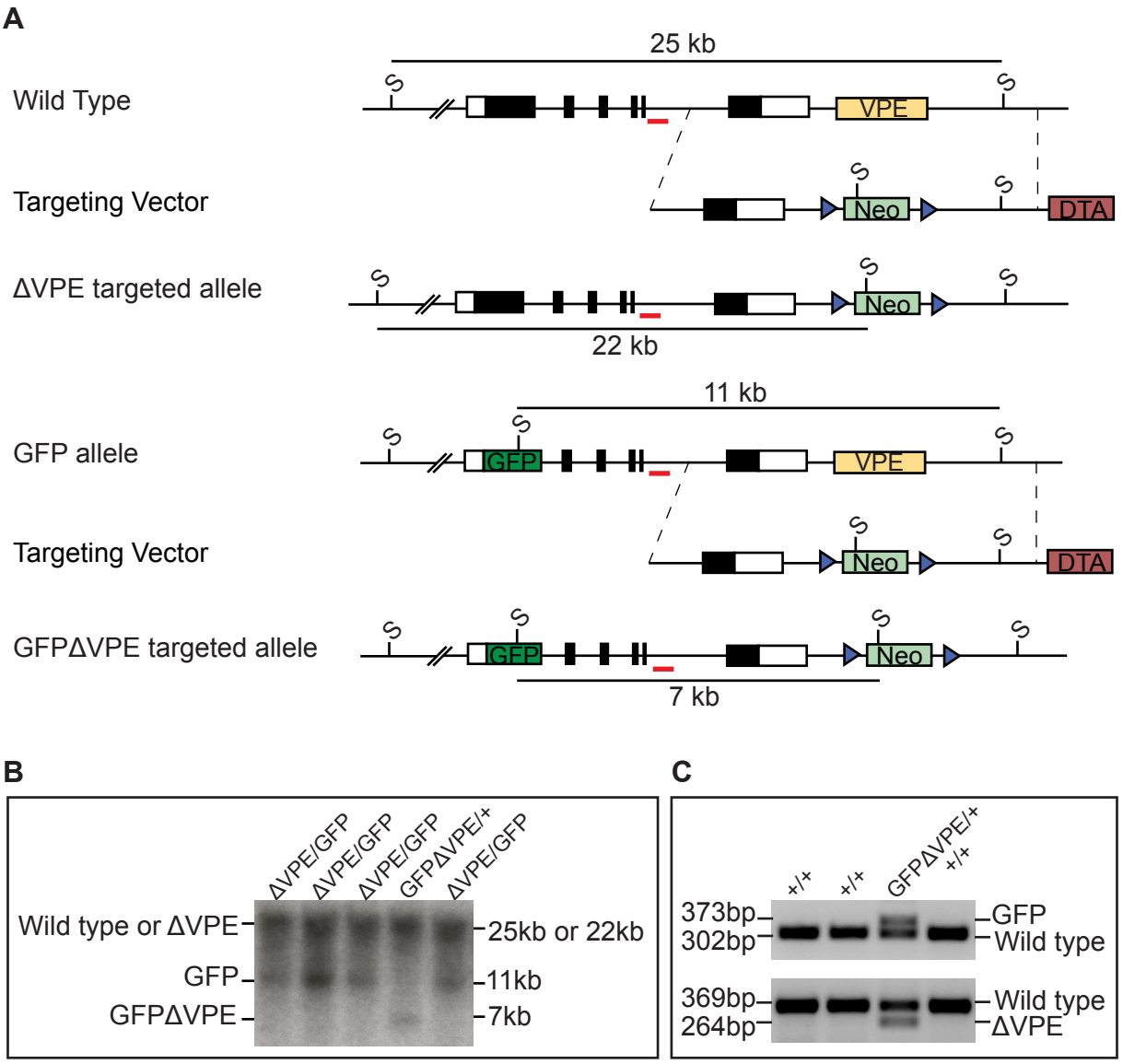



**Figure S5: Generating Eomes$^{GFP}$ allele lacking the VPE region**

(A) Heterozygous *Eomes*$^{GFP/+}$ (Arnold et al., 2009) ESC were re-targeted using the same construct and primary screening strategy as used to delete the VPE. Southern blot strategy used to distinguish targeting the VPE region in either the GFP or wild type alleles, and expected fragment sizes are indicated. SpeI (S). (B) Southern blot showing two different genotypes of successfully targeted clones; *Eomes*$^{GFP\Delta VPE/+}$ and *Eomes*$^{GFP/\Delta VPE}$. (C) PCR genotyping of *Eomes*$^{GFP\Delta VPE/+}$ mice.

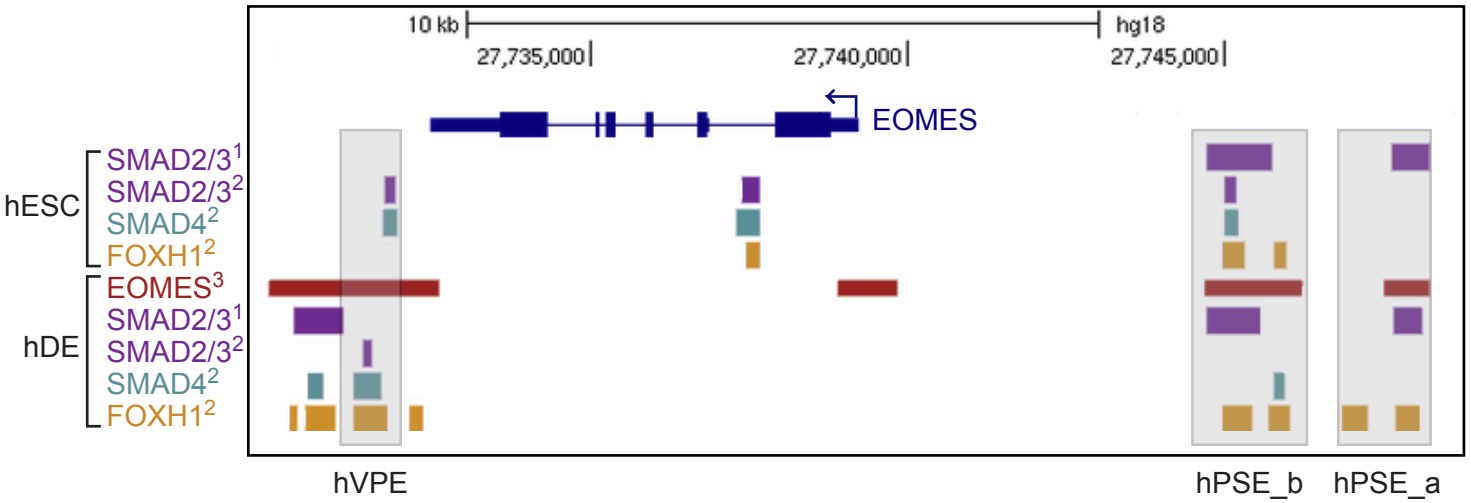



**Figure S6: Regulation of the VPE by Nodal signaling**

(A) Homologous human regions of the mouse VPE and PSE are bound by EOMES and mediators of the Nodal signaling pathway in hESCs and hDE. 1=(Brown et al., 2011) 2=(Kim et al., 2011), 3=(Teo et al., 2011). ChIP-seq data showing regions bound by SMAD2/3 (purple), SMAD4 (green), FOXH1 (orange) and EOMES (red) are represented by coloured bars and were aligned to the EOMES locus on the UCSC Genome browser Human Mar. 2006 (NCBI36/hg18) Assembly (http://genome.ucsc.edu/). Homologous regions to the mouse VPE and PSE are highlighted in grey. Human VPE, PSE_a and PSE_b (hVPE, hPSE_a, hPSE_b). FoxH1 binds the conserved FoxH1 binding site at the VPE in hDE.

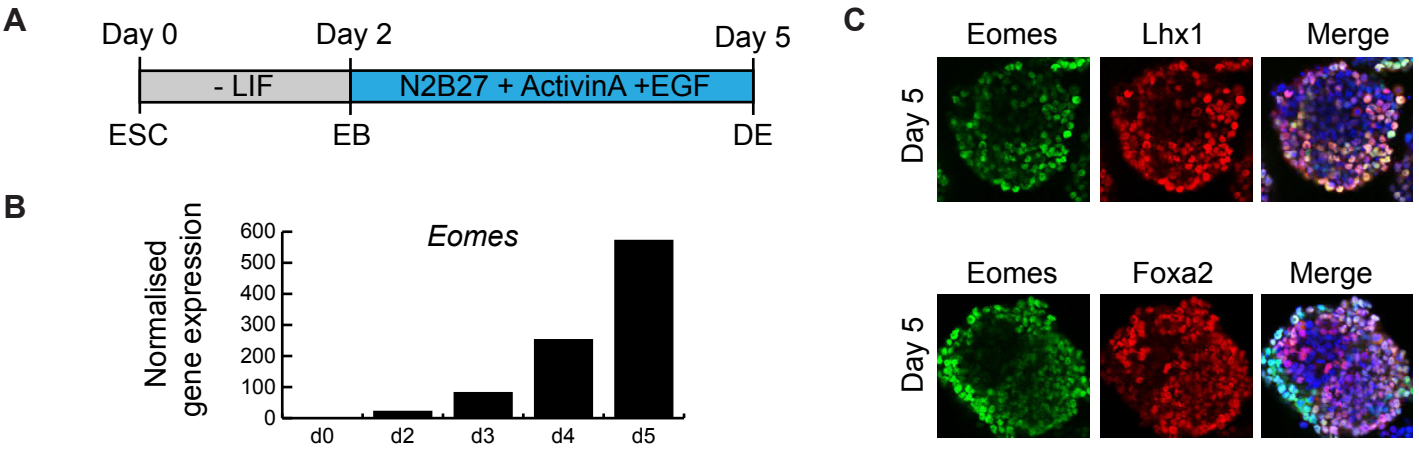



**Figure S7: Definitive endoderm differentiation**

(A) Schematic of protocol to differentiate ESC to definitive endoderm (DE) fate. ESC were grown in the absence of LIF for 2 days to form embryoid bodies (EB) and then differentiated in N2B27 medium, 20ng/ml ActivinA and 20ng/ml EGF for a further 3 days. (B) qPCR of *Eomes* mRNA demonstrates a dramatic increase in expression over the course of the 5 day differentiation regime. Gene expression is normalised to *Gapdh*. (C) 2D confocal images of d5 DE EBs stained with antibodies against definitive endoderm markers Eomes, Lhx1 or Foxa2, and counterstained with DAPI.

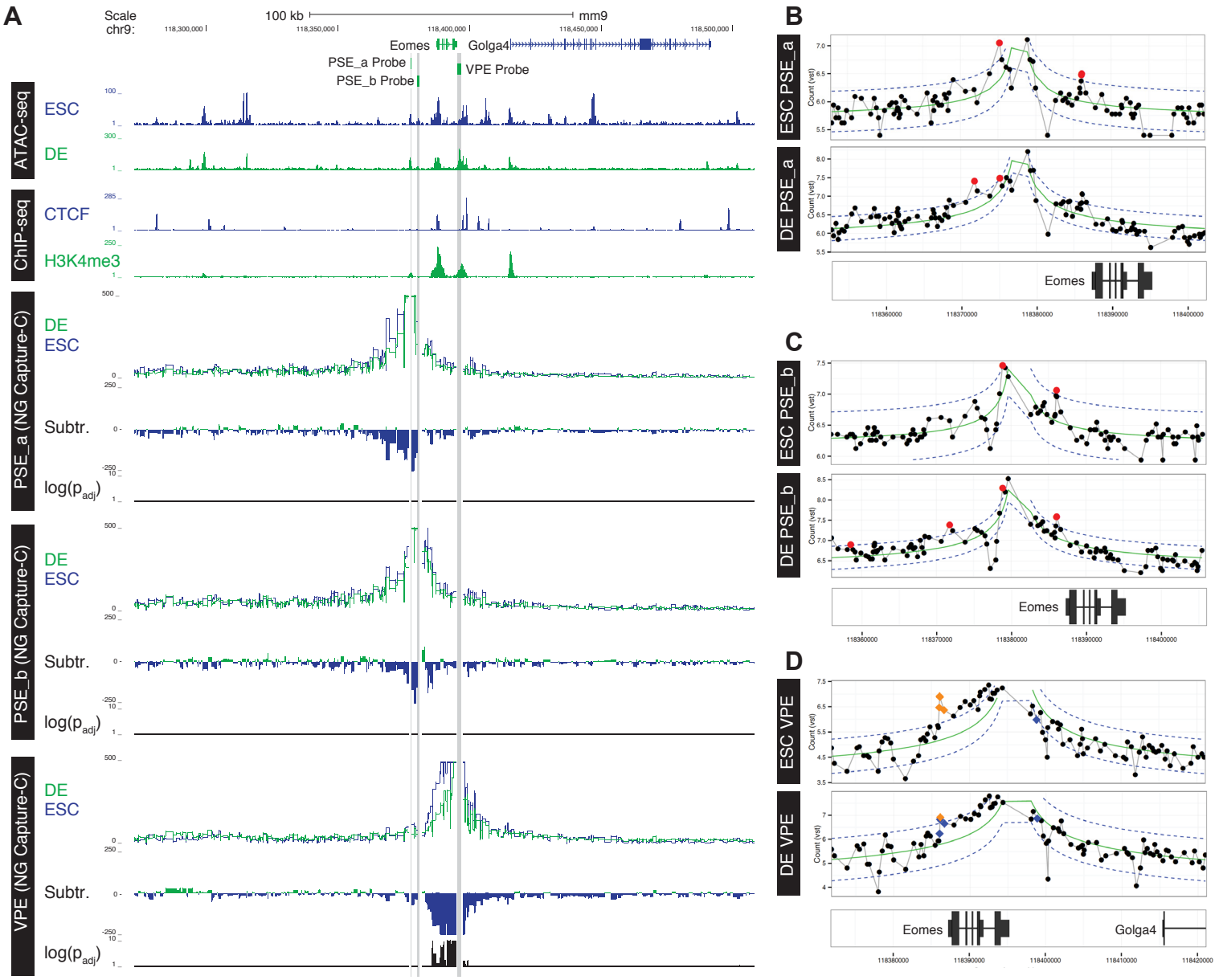



**Figure S8: NG Capture-C from the *Eomes* enhancers.**

(A) NG Capture-C interaction profiles of the PSE_a, PSE_b, and VPE from ESC (blue) and DE (green). Tracks show mean interactions of normalized biological replicates (n=3), subtraction of ESC from DE (Subtr.) and DESeq2 significant differences between DE and ESC (-log($P_{adj}$); p≤0.05). Open chromatin was determined by ATAC-seq in both ESC and DE, ChIP-seq of the boundary element CTCF in ESC is from published data (Handoko et al., 2011) and H3K4me3 ChIP-seq was generated in triplicate from DE. FourCSeq comparison of NG Capture-C between DE and ESC from the PSE_a (B), PSE_b (C), and VPE (D). Red circles mark fragments with more interactions than expected based upon proximity to the promoter (green line), Blue Diamonds show fragments with significantly different interactions between the two conditions (P≤0.05), Orange Diamonds show fragments with enriched reactions that are significantly different between the two conditions.

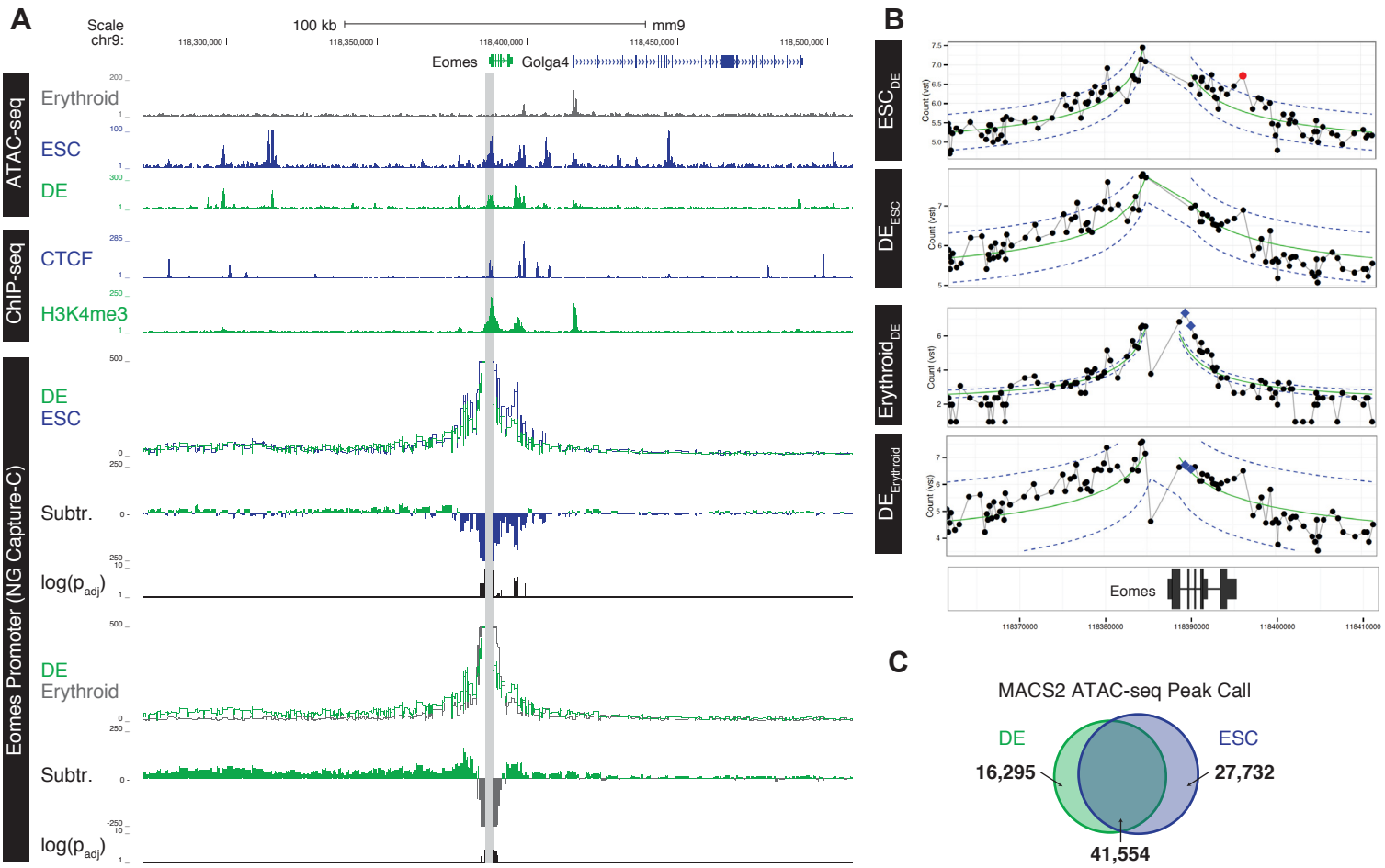



**Figure S9: NG Capture-C from the *Eomes* promoter.**

(A) NG Capture-C interaction profiles of the *Eomes* promoter from terminally differentiated erythrocytes (Ery, grey), ESC (blue) and DE (green). Tracks show mean interactions of normalized biological replicates (n=3), subtraction of ESC and PHS from DE (Subtr.) and DESeq2 significant differences between the cell types (-log($P_{adj}$); p≤0.05). Open chromatin was determined by ATAC-seq in all three cell types (n=3), ChIP-seq of the boundary element CTCF in ESC is from published data (Handoko et al., 2011) and H3K4me3 ChIP-seq was generated in triplicate from DE. (B) FourCSeq comparison of NG Capture-C of the *Eomes* promoter between DE, ESC and Ery. Comparison condition is shown in subscript. Red circles mark fragments with more interactions than expected based upon proximity to the promoter (green line), Blue Diamonds show fragments with significantly different interactions between the two conditions (P≤0.05), Orange Diamonds show fragments with enriched reactions that are significantly different between the two conditions. (C) Comparison of MACS2 peak call for ATAC-seq from DE and ESC.

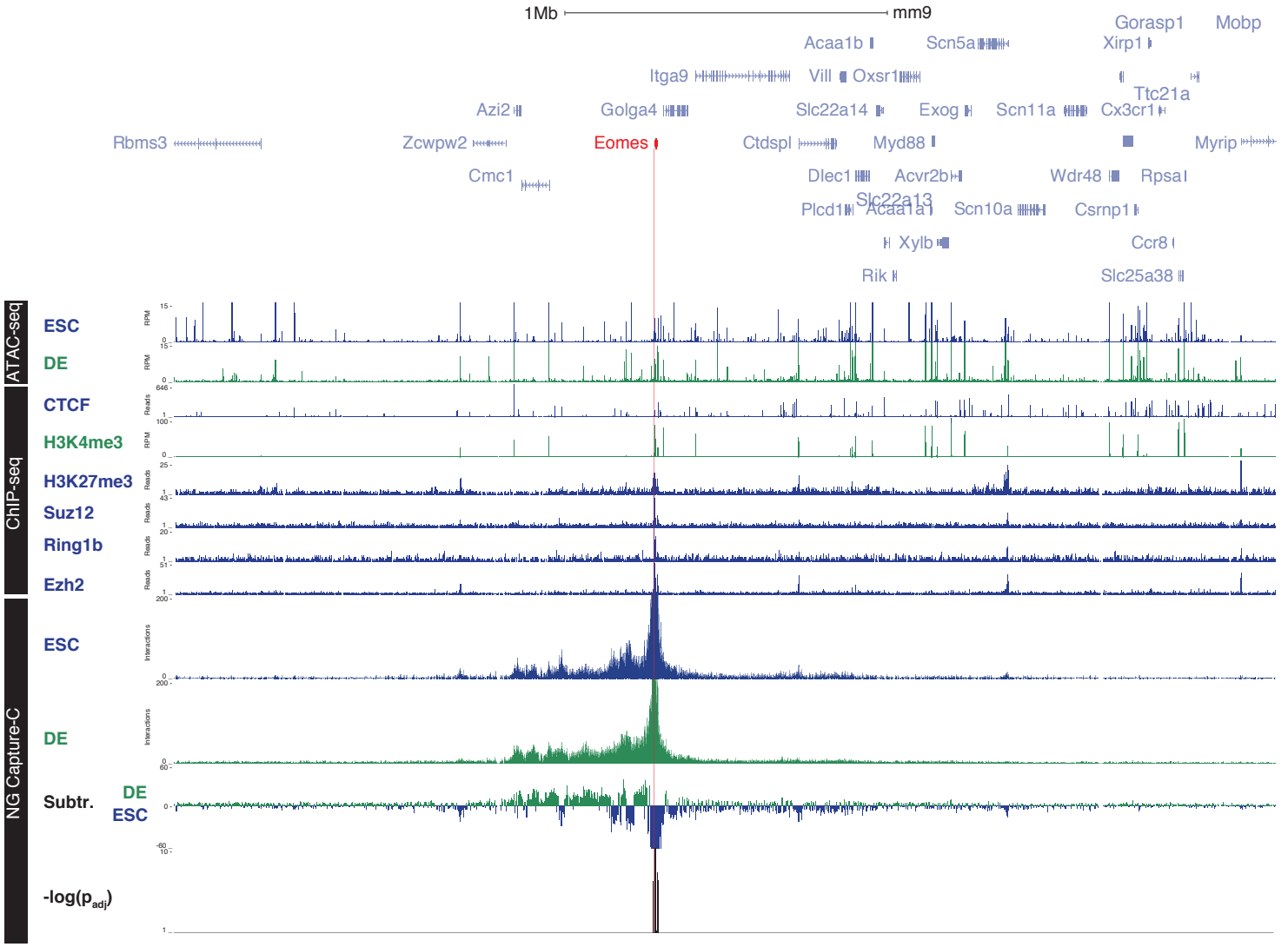



**Figure S10: Long-range NG Capture-C from the *Eomes* promoter.**

NG Capture-C interaction profiles of the *Eomes* promoter (chr9:116890604-120321539) from erythrocytes (grey), ESC (blue) and DE (green). Tracks show mean interactions of normalized biological replicates (n=3), subtraction of ESC from DE (Subtr.) and DESeq2 significant differences between DE and ESC (-log($P_{adj}$); p≤0.05). Location of the Polycomb Repressor Complexes components (Ezh2, Suz12, Ring1b) and associated histone modification (H3K27me3) in ESC are shown (Ku et al., 2008; Mikkelsen et al., 2007). Open chromatin was determined by ATAC-seq in all three cell types (n=3), ChIP-seq of the boundary element CTCF in ESC is from published data (Handoko et al., 2011) and H3K4me3 ChIP-seq was generated in triplicate from DE.

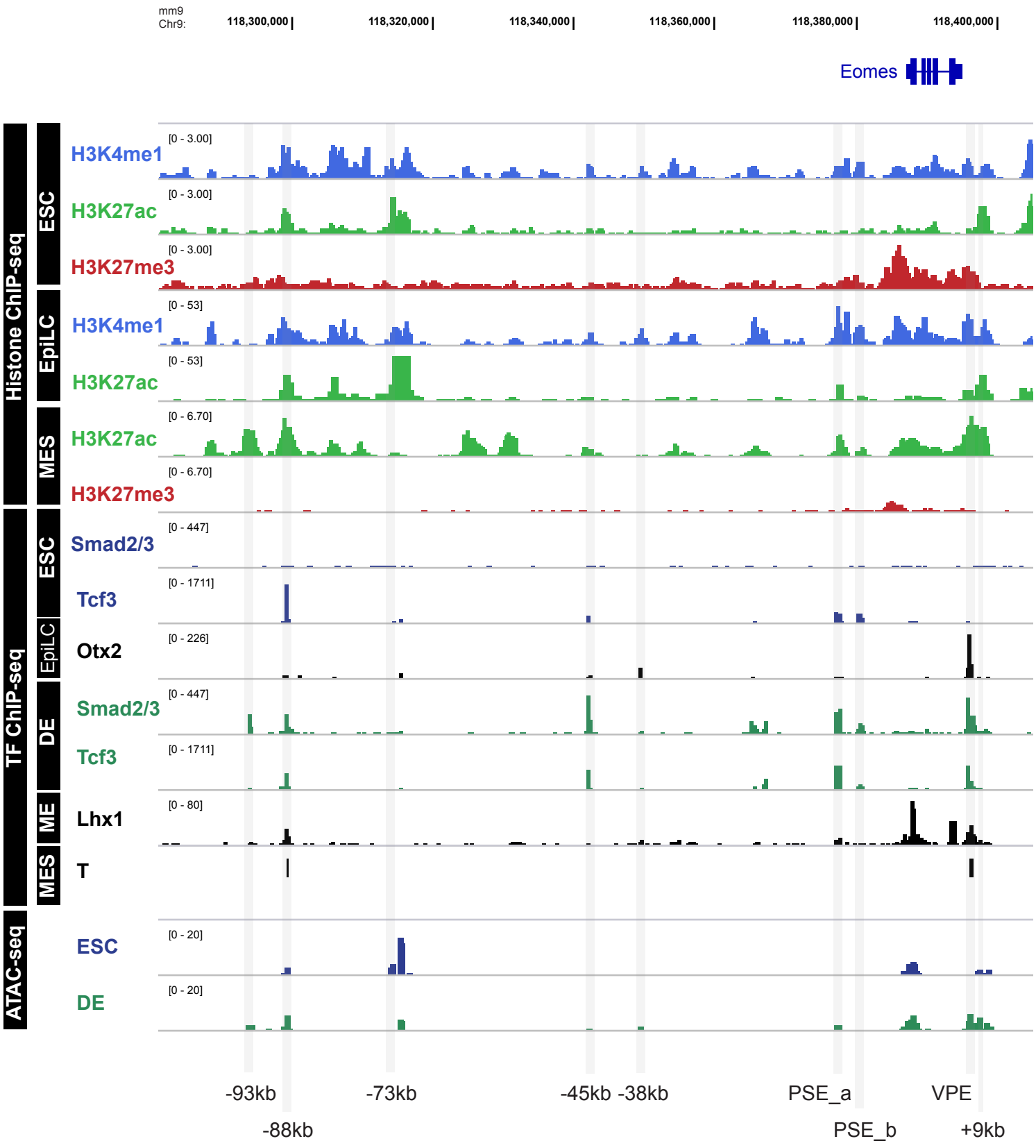


**Figure S11: Mapping enhancers within the *Eomes* compartment.**

ChIP-seq of histone modifications H3K4me1 (light blue), H3K27me3 (red) and H3K27ac (light green) in ESC, epiblast like cells (EpiLC) and mesoderm (MES) (Alexander et al., 2015; Buecker et al., 2014; Consortium, 2012). Open chromatin was generated using ATAC-seq in ESC and DE (n=3). ChIP-seq of TFs involved in endoderm and anterior mesendoderm specification. Smad2/3 and Tcf3 in ESC (blue) and DE (green) (Wang et al., 2017). Otx2 in EpiLC (Buecker et al., 2014), Lhx1 in P19 mesendoderm (ME) (Costello et al., 2015), and Brachyury (T) (Lolas et al., 2014) in MES. Regions of increased chromatin accessibility unique to ESC (-73kb) and those associated with Smad2/3 occupancy uniquely in DE (-93kb, -45kb, -38kb, PSE_a, VPE and +9kb) are highlighted as in Fig. 5B. In addition, a TF binding hotspot accessible in both ESC and DE (-88kb), and the PSE_b, are also highlighted.

## Supplemental References

**Alexander, J. M., Hota, S. K., He, D., Thomas, S., Ho, L., Pennacchio, L. A. and Bruneau, B. G.** (2015). Brg1 modulates enhancer activation in mesoderm lineage commitment. *Development* **142**, 1418-1430.

**Arnold, S. J., Sugnaseelan, J., Groszer, M., Srinivas, S. and Robertson, E. J.** (2009). Generation and analysis of a mouse line harboring GFP in the Eomes/Tbr2 locus. *Genesis* **47**, 775-781.

**Behringer, R., Gertsenstein, M., Nagy, K. and Nagy, A.** (2013). *Manipulating the mouse embryo : a laboratory manual, fourth edition* (Fourth edition. edn): Cold Spring Harbor Laboratory Press.

**Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C. H., Lim, B., Vardy, L., Dunn, N. R., Trotter, M., Pedersen, R., et al.** (2011). Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors. *Stem Cells* **29**, 1176-1185.

**Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T. and Wysocka, J.** (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 838-853.

**Consortium, E. P.** (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74.

**Costello, I., Nowotschin, S., Sun, X., Mould, A. W., Hadjantonakis, A. K., Bikoff, E. K. and Robertson, E. J.** (2015). Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. *Genes Dev* **29**, 2108-2122.

**Handoko, L., Xu, H., Li, G., Ngan, C. Y., Chew, E., Schnapp, M., Lee, C. W., Ye, C., Ping, J. L., Mulawadi, F., et al.** (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* **43**, 630-638.

**Kim, S. W., Yoon, S. J., Chuong, E., Oyolu, C., Wills, A. E., Gupta, R. and Baker, J.** (2011). Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. *Dev Biol* **357**, 492-504.

**Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser, A., Nusbaum, C., Xie, X., Chi, A. S., et al.** (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4**, e1000242.

**Lolas, M., Valenzuela, P. D., Tjian, R. and Liu, Z.** (2014). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. *Proc Natl Acad Sci U S A* **111**, 4478-4483.

**Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., et al.** (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560.

**Teo, A. K., Arnold, S. J., Trotter, M. W., Brown, S., Ang, L. T., Chng, Z., Robertson, E. J., Dunn, N. R. and Vallier, L.** (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev* **25**, 238-250.

**Wang, A., Yue, F., Li, Y., Xie, R., Harper, T., Patel, N. A., Muth, K., Palmer, J., Qiu, Y., Wang, J., et al.** (2015). Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16**, 386-399.

**Wang, Q., Zou, Y., Nowotschin, S., Kim, S. Y., Li, Q. V., Soh, C. L., Su, J., Zhang, C., Shu, W., Xi, Q., et al.** (2017). The p53 Family Coordinates Wnt and Nodal Inputs in Mesendodermal Differentiation of Embryonic Stem Cells. *Cell Stem Cell* **20**, 70-86.

## Table S1: Primers used in this study

| Primer name | Forward sequence | Reverse sequence | Product |
|---|---|---|---|
| *Targeting vectors* | | | |
| VPE Recombineering | GGCTGGGGTTGGG GAAGGAGTGTTTGC CCTGGAGATGCAAG ATTGTGCTCGGATC CAATTAACCCTCAC TAAAGGGC | GGTCCCAGAAGTTTG GAGGACGGGAAAGA CTGTCCACAGCTCAG GTATATCGAAGTTAT AAGCTTGAAGTTCCT ATACTTTC | n/a |
| PSE AatII | TGACGTCTGTGTTC AAAAGCACGAGGG | ACCAGAGACCGTATG TTCCC | 2.7kb |
| *Transgenic reporter* | | | |
| VPE LacZ | GCCCTGGAGATGC AAGATTG | CAGCTCAGGTATATC TTCTGGC | 696bp |
| *Genotyping* | | | |
| VPE WT | TCGTTGAGTGGTGA GCAGGGAG | AGCGAGGACATCCA CGGAAAAC | 369bp |
| VPE Δ | TCGTTGAGTGGTGA GCAGGGAG | TTTGGAGGACGGGA AAGACTG | 264bp |
| PSE WT | AGGGTGGCTCTATA CAGGTG | GCATTGGAGTTGAAG GTGGG | 328bp |
| PSE Δ | AGGGTGGCTCTATA CAGGTG | TCACAAGTCTCTCCT GGCAC | 246bp |
| PSE_b WT | TTGCGTTTGTTGGG TTTTGG | CCATCACTGGGAGA GTAGGC | 427bp |
| PSE_b Δ | GGCTATTGCCTCCA TACAGC | CCATCACTGGGAGA GTAGGC | 712bp |
| LacZ | TTACCAGGCCGAAG CAGCGTTGTTG | GCGGCAGTAAGGCG GTCGGGATAGT | 300bp |
| *RT-PCR* | | | |
| Gapdh | CAATGACCCCTTCA TTGACC | GATCTCGCTCCTGGA AGATG | 145bp |
| Eomes | TGTTTTCGTGGAAG TGGTTCTGGC | AGGTCTGAGTCTTGG AAGGTTCATTC | 323bp |

**Table S2: Antibodies used in this study**

| Name | Catalog number | Company |
|---|---|---|
| Foxa2 | sc-6554 | Santa Cruz |
| Lhx1 | sc-19341 | Santa Cruz |
| TBR2/Eomes | ab23345 | Abcam |
| GFP AlexaFluor 488 | A21311 | Invitrogen |
| Goat IgG AlexaFluor 594 | A11058 | Invitrogen |
| Rabbit IgG AlexaFluor 488 | A21206 | Invitrogen |
| Anti-H3K4me3 | 07-473 | Millipore |

**Table S3.** Long-range *Foxa2* and *Lhx1* promoter interactions identified by NG Capture-C

Click here to Download Table S3

**Table S4.** Probes used for NG Capture-C.

Click here to Download Table S4

**Table S5.** Accession codes used in this study.

Click here to Download Table S5