Stem Cell Reports, Volume *5*

Supplemental Information

Tbx3 Controls Dppa3 Levels and Exit from Pluripotency toward Mesoderm

Avinash Waghray, Néstor Saiz, Anitha D. Jayaprakash, Ana G. Freire, Dmitri Papatsenko, Carlos-Filipe Pereira, Dung-Fang Lee, Ran Brosh, Betty Chang, Henia Darr, Julian Gingold, Kevin Kelley, Christoph Schaniel, Anna-Katerina Hadjantonakis, and Ihor R. Lemischka

Supplemental Information

Tbx3 controls Dppa3 levels and exit from pluripotency towards mesoderm

SUPPLIMENTAL EXPERIMENTAL PROCEDURES

shRNA design and viral infection

The lentiviral shRNA oligos were designed and constructed as described (Lee et al., 2012). Briefly, lentiviral particles were made by expressing the respective shRNA constructs along with pCMVΔ8.9 and pVSVG vectors in 293T cells. The supernatant was collected 72 hours after transfection and filterted through a 0.45µM sterile filters. The filtered supernatant was concentrated using Ultracel 30K ultra centrifuge filters at 3000rpm for 30 minutes as described (Lee et al., 2012). The concentrate along with 6µg/µl of polybrene was used to infect mESCs.

Heterokaryon reprogramming assay

The self-renewal ability of mESCs was assayed by fusing mESCs with human B cells as described (Lee et al., 2012a). Reprogramming of human ESCs was assayed by measuring mRNA expression of key genes, as needed using qRT-PCR (Pereira et al., 2008).

Generation of Tbx3 targeting vector and Tbx3 null mESCs

Tbx3 null mESCs were generated from Tbx3 heterozygote ($Tbx3^{+/N}$) mESCs that have been described previously (Davenport et al., 2003). Tbx3 heterozygous cells ($Tbx3^{+/N}$ cells) were kindly donated by Dr. V. E. Papaioannou from Columbia University Medical

Center. $Tbx3^{N/N}$ mESCs were generated by culturing $Tbx3^{+/N}$ cells in 2000 µg/µl of genetecin in long-term cultures. Colonies were picked and analyzed for a normal karyotype. Karyotype of the $Tbx3^{N/N}$ cells was assayed at the company "Cell line genetics". To generate $Tbx3^{N/P}$ cells the wild type allele in $Tbx3^{+/N}$ cells was re-targeted by homologous recombination using a targeting vector. The targeting vector was generated using the pFlexible-puro vector (a kind gift from Jianlong Wang's lab at Icahn School of Medicine at Mount Sinai). Two homologous arms were cloned into the PmeI/AscI (5') and NotI (3') sites. The arms were PCR amplified from genomic DNA of R1 mESCs using primers listed in table S4. The targeting of drug selected Tbx3 null mESCs ($Tbx3^{N/N}$ and $Tbx3^{N/P}$) was tested by genomic PCR (Table S4) and confirmed by Southern blotting as described (Carvajal-Vergara et al., 2010). T3Blast mESCs were derived from blastocyst stage embryos (E3.75), collected from $Tbx3^{+/-}$ intercrosses, as previously described previously (Czechanski et al., 2014).

Derivation of blastocysts, immunofluorescence and image acquisition

Blastocysts were collected at E3.75 or E4.5 from $Tbx3^{+/-}$ females bred to $Tbx3^{+/-}$ males (C57BL/6xCD1) (Davenport et al., 2003). The day of detection of vaginal plug was considered E0.5. Embryos were flushed from oviducts in FHM medium (Millipore) and fixed subsequently in 4% paraformaldehyde in PBS (Electron Microscopy Sciences). Whenever necessary, zona pellucida was removed by brief incubation in acidic Tyrode's solution (Sigma). Mice were maintained under a 12h light/dark cycle in the designated facilities of Memorial Sloan Kettering Cancer Center (MSKCC). Mouse work was subject

to approval by, and carried out in accordance with guidelines from MSKCC's Institutional Animal Care and Use Committee. Immunofluorescence (IF) was performed as described (Frankenberg et al., 2013). Imaging was carried out on a Zeiss LSM 510 META confocal microscope, using a Plan-Neofluar 40x/1.3 Oil DIC objective. Embryos were mounted in drops of 5µg/ml Hoechst 33342 solution (Life Technologies) and 1µm optical sections were acquired to build whole-embryo z-stacks. Images were segmented for cell counting using the MINS software (Lou et al., 2014) and data was processed as previously described (Schrode et al., 2014). Following imaging, embryos were lysed for genotyping as described previously (Kang et al., 2013). The primers used for both embryo and ES cell genotyping have been previously published (Davenport et al., 2003).

Generation of cell lines and cell culture conditions

All mESCs were cultured in mESC media as described (Ang et al., 2011; Lee et al., 2012). Tbx3R rescue clones were described previously (Ivanova et al., 2006). Dppa3R and genetic complementation system cell lines was made by introducing the pLKO.1-pig based transgenic vector (Lee et al., 2012) along with empty pTRIPz vector (for rtTA3) into $Tbx3^{+/+}$ and $Tbx3^{N/N}$ cells in the presence of Dox. $Tbx3^{N/N}$+Tbx3 Tg cell line was constructed by introducing a Dox inducible Tbx3 vector made by cloning Tbx3 cDNA into the pTRIPz vector (Table S4). $Tbx3^{N/N}$+BirA cell line was generated by introducing the biotin ligase enzyme BirA cDNA into $Tbx3^{N/N}$ cells by electroporation as described (Wang et al., 2006). $Tbx3^{N/N}$+BirA+bfTbx3 cell line was generated by introducing the Flag-Biotin tagged Tbx3 cDNA (Table S4) into $Tbx3^{N/N}$+BirA cells. Tbx3i cell lines were generated

from the KH2 parental cell line (Beard et al., 2006). Tbx3 isoform2 was knocked into the ColA1 locus using the Flp-in expression vector pBS31'-RBGpA-Tbx3Iso2 generated using vectors from the Thermo scientific open biosystems targeting kit. Briefly, ~10^7 cells were electroporated with 15µg each of the pBS31'-RBGpA-Tbx3Iso2 and pCAGGS-FLPe-puro vectors. Hygromycin (140 µg/ml) was added into the media 24 hours after the electroporation. Surviving colonies were picked and positive clones were selected by qRT-PCR with and without Dox. Recombinant Wnt3a was used at the indicated concentrations. All cell culture reagents and chemicals mentioned above were purchased through the Pluripotent Stem Cell shared resource facility (SRF) at Mount Sinai (http://icahn.mssm.edu/research/resources/shared-resource-facilities/human-embryonic-stem-cell). Human ESCs were cultured in human ESC media on matrigel as described (Carvajal-Vergara et al., 2010). Human B cells used for heterokaryon-based reprogramming assay were cultured as described (Pereira et al., 2008).

Quantitative real-time (qRT) PCR analysis

Analysis of mRNA expression by qRT-PCR analysis was done using the Fast SYBR green 2X master mix kit from Applied biosystems. The primers used are mentioned in table S4. The reactions were run on the Roche lightcycler 480 machine.

Western blotting and antibodies

Western blotting was done as described (Lee et al., 2012). List of antibodies used are mentioned in table S4.

Alkaline phosphatase staining, FACS analysis and immunofluorescence

Alkaline phosphatase (AP) staining was done following manufacturer's instructions using the Stemgent or Millipore kits as described (Lee et al., 2012). To test the colony forming ability of mESCs, single SSEA1+ mESCs were sorted on to 96 well plates (Flow Cytometry Core, Icahn School of Medicine at Mount Sinai) and emergence of AP positive colonies was scored. SSEA1 staining was done following manufacturer's instructions using the BD Pharmingen kit. Wild type cells stained with IgM-APC (Table S4) was used as negative control in figure 2K. IF on mESCs was done using the antibodies listed in table S4 as described (Carvajal-Vergara et al., 2010).

ChIP-Seq and mRNA-Seq Sample preparation

ChIP-Seq sample preparation was performed as described (Ang et al., 2011). Briefly, cells were fixed with 11% formaldehyde in mESC media for 10 minutes on a rotating platform. The fixation was stopped by addition of 125nM of Glycine for 5 minutes on a rotating platform. Cells were scrapped, quick-frozen and stored at -80°C. ChIP was done by pulling down the bfTbx3 protein using Dynabeads conjugated to streptavidin (SA) from $Tbx3^{N/N}$+NBirA+bfTbx3 cells. Pull down in $Tbx3^{N/N}$+NBirA cells was used as negative control. DNA library preparation was done using the Illumina ChIP-seq sample prep kit. The libraries were run on HiSeq 2000 machines at the genomics SRF at Icahn school of medicine at Mount Sinai (http://icahn.mssm.edu/departments-and-institutes/genomics/about/resources/genomics-core-facility).

For mRNA-Seq sample preparation, the total RNA was isolated using Trizol reagent following the manufacturer's instructions. Quality and quantity of total RNA was assayed using Nanodrop and Agilent Bioanalyzer. Further, mRNA-seq libraries were prepared using the TrueSeq protocol from Illumina. Briefly, mRNA from 500ng-1 µg of total RNA was isolated using poly T beads. Double stranded cDNA was made using super-script reverse transcriptase and random primers. The cDNA ends were blunted, "A" base added and adapters ligated to each end of the cDNA. Total of 15 cycles of PCR were performed to generate the cDNA libraries. The concentration of the libraries was assayed using Agilent Bioanalyzer and sequenced either using 1*26bp or 1*100bp on an Illumina Genome analyzer II or HiSeq 2000 platform at the genomics SRF at Mount Sinai.

For single cell mRNA-Seq preparation, SSEA1+DAPI- mESCs were sorted and collected. The cDNA synthesis was done following the manufacturers instruction using the C1 Single-Cell Auto Prep System (Fluidigm). Amount of cDNA synthesized was quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies). Sequencing libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina). About 24 libraries were mixed together. The concentration of the mixed libraries was assayed using Agilent Bioanalyzer. The libraries were sequenced using 1*100bp on a HiSeq 2000/2500 platform at the genomics SRF at Mount Sinai and NYU genomics core.

Next-Gen sequencing analysis and other analyses

ChIP-Seq analysis was done on the raw Fastq files received. The fastq files were mapped to the mouse mm9 genome using Bowtie (v1.0) program allowing for 2 base pair

mismatches. The mapped output file was processed through MACS analysis (v1.4) software to determine peaks. Homer software package (http://homer.salk.edu/homer/index.html) was used for annotation of the peaks and further analysis was carried out using the R statistical software. The bigWig files were generated using the Galaxy server (https://usegalaxy.org). The ChIP dataset for β-catenin was annotated from literature (Zhang et al., 2013). β-catenin bound regions are determined as those regions bound in both Flag and Biotin ChIP pull downs in the samples with induction of Wnt signaling. The UCSC genome browser was used to visualize the bigWig files from ChIP-Seq datasets.

For mRNA-Seq analysis the raw fastq files were mapped to the mouse mm9 genome using Tophat (with bowtie2). The output files were processed through Cufflinks and Cuffdiff programs to conduct differential gene expression analysis. The output ".bam" files from Tophat were used to generate the bedGraph files using Homer software package. Briefly, within the homer package ".bam" files were used to generate tagdirectories using makeTagDirectory command. The tagdirectories were then used to generate bedGraph files using the makeUCSCfile command. The UCSC genome browser was used to visualize the bedGraph files from mRNA-Seq datasets. CummeRbund was used as needed for analysis of the raw Fastq files and the output files from Cuffdiff program.

For single cell mRNA-Seq analysis the raw fastq files were mapped to the mouse mm9 genome using Tophat (with bowtie2). The output ".bam" files were processed through

Cuffquant program. The output ".cxb" files from Cuffquant were processed through Cuffnorm program to generate FPKM values. The p value for the anti-correlation of expression betweeb Tbx3 and Dppa3 in cells expressing these TFs was calculated using Student t-distribution test.

R software was used to do the statistics as needed. Quantile analysis was done by calculating the Q-scores to rank the genes as described previously (Creyghton et al., 2010). The distribution of quantile scores was plotted as bar plots and the significance between the different datasets was calculated by Wilcoxon rank sum test with continuity correction using the R statistical software. For both ChIP-Seq and mRNA-Seq analysis the mm9 mouse reference genome used.

Microarray data used to generate the correlation plots in Figure 2A and 2B was from published data (Hailesellasse Sene et al., 2007). Microarray data used to generate the data in Figure S3B and S4E was from published literature (Mikkelsen et al., 2008). Gene set enrichment analysis (GSEA) was performed following standard instructions from the standalone java tool downloaded from www.broadinstitute.org/gsea. FPKM values were used to make the pre ranked lists. The gene sets were annotated from Amigo gene ontology terms and are available on request.

SUPPLIMENTAL REFERENCES

Ang, Y.S., Tsai, S.Y., Lee, D.F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B.*, et al.* (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. Cell *145*, 183-197.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23-28.

Carvajal-Vergara, X., Sevilla, A., D'Souza, S.L., Ang, Y.S., Schaniel, C., Lee, D.F., Yang, L., Kaplan, A.D., Adler, E.D., Rozov, R.*, et al.* (2010). Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. Nature *465*, 808-812.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A.*, et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America *107*, 21931-21936.

Czechanski, A., Byers, C., Greenstein, I., Schrode, N., Donahue, L.R., Hadjantonakis, A.K., and Reinholdt, L.G. (2014). Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nature protocols *9*, 559-574.

Davenport, T.G., Jerome-Majewska, L.A., and Papaioannou, V.E. (2003). Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome. Development (Cambridge, England) *130*, 2263-2273.

Frankenberg, S., Shaw, G., Freyer, C., Pask, A.J., and Renfree, M.B. (2013). Early cell lineage specification in a marsupial: a case for diverse mechanisms among mammals. Development *140*, 965-975.

Hailesellasse Sene, K., Porter, C.J., Palidwor, G., Perez-Iratxeta, C., Muro, E.M., Campbell, P.A., Rudnicki, M.A., and Andrade-Navarro, M.A. (2007). Gene function in early mouse embryonic stem cell differentiation. BMC genomics *8*, 85.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533-538.

Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.K. (2013). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development *140*, 267-279.

Lee, D.F., Su, J., Ang, Y.S., Carvajal-Vergara, X., Mulero-Navarro, S., Pereira, C.F., Gingold, J., Wang, H.L., Zhao, R., Sevilla, A.*, et al.* (2012). Regulation of embryonic and induced pluripotency by aurora kinase-p53 signaling. Cell stem cell *11*, 179-194.

Lou, X., Kang, M., Xenopoulos, P., Munoz-Descalzo, S., and Hadjantonakis, A.K. (2014). A Rapid and Efficient 2D/3D Nuclear Segmentation Method for Analysis of Early Mouse Embryo and Stem Cell Image Data. Stem Cell Reports *2*, 382-397.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49-55.

Pereira, C.F., Terranova, R., Ryan, N.K., Santos, J., Morris, K.J., Cui, W., Merkenschlager, M., and Fisher, A.G. (2008). Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. PLoS genetics *4*, e1000170.

Schrode, N., Saiz, N., Di Talia, S., and Hadjantonakis, A.K. (2014). GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. Developmental cell *29*, 454-467.
Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364-368.
Zhang, X., Peterson, K.A., Liu, X.S., McMahon, A.P., and Ohba, S. (2013). Gene Regulatory Networks Mediating Canonical Wnt Signal Directed Control of Pluripotency and Differentiation in Embryo Stem Cells. Stem cells.

SUPPLEMENTAL TABLES

Table S1. Related to Figures 2 and 5. mRNA-Seq data from $Tbx3^{+/+}$, $Tbx3^{+/N}$, $Tbx3^{N/N}$, shLuc and shTbx3 containing mESCs.

Table S2. Related to Figures 2 and 5. ChIP-Seq targets of Tbx3 in mESCs

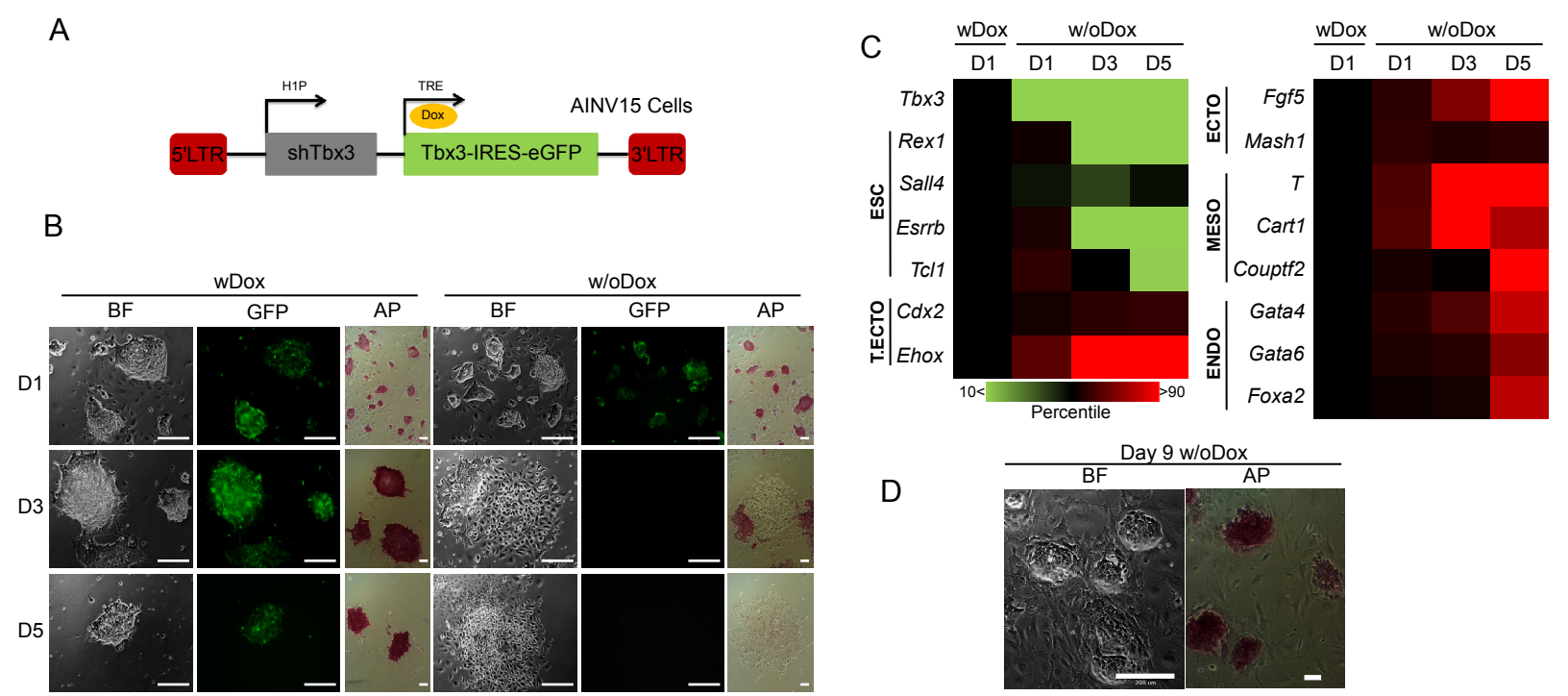
Table S3. Related to Figures 3. Single cell mRNA-Seq data of wild type R1 mESCs

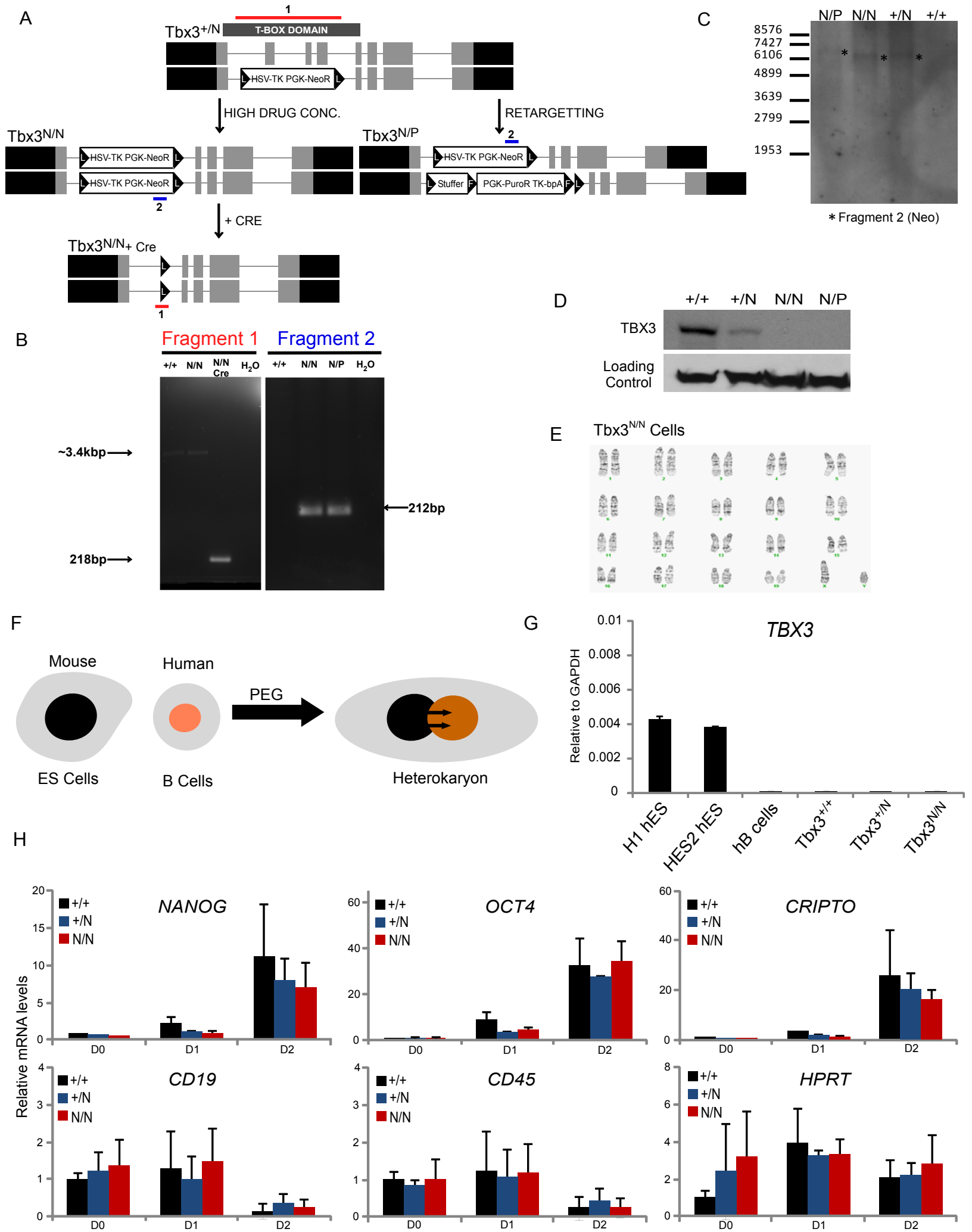
Table S4. Related to all figures. Oligos and antibodies used in the study.

Table S5. Related to Figure 2. T3blast mESC lines and embryo counts

Table S6. Related to Figure 2. Number of E10.5 chimeric embryos

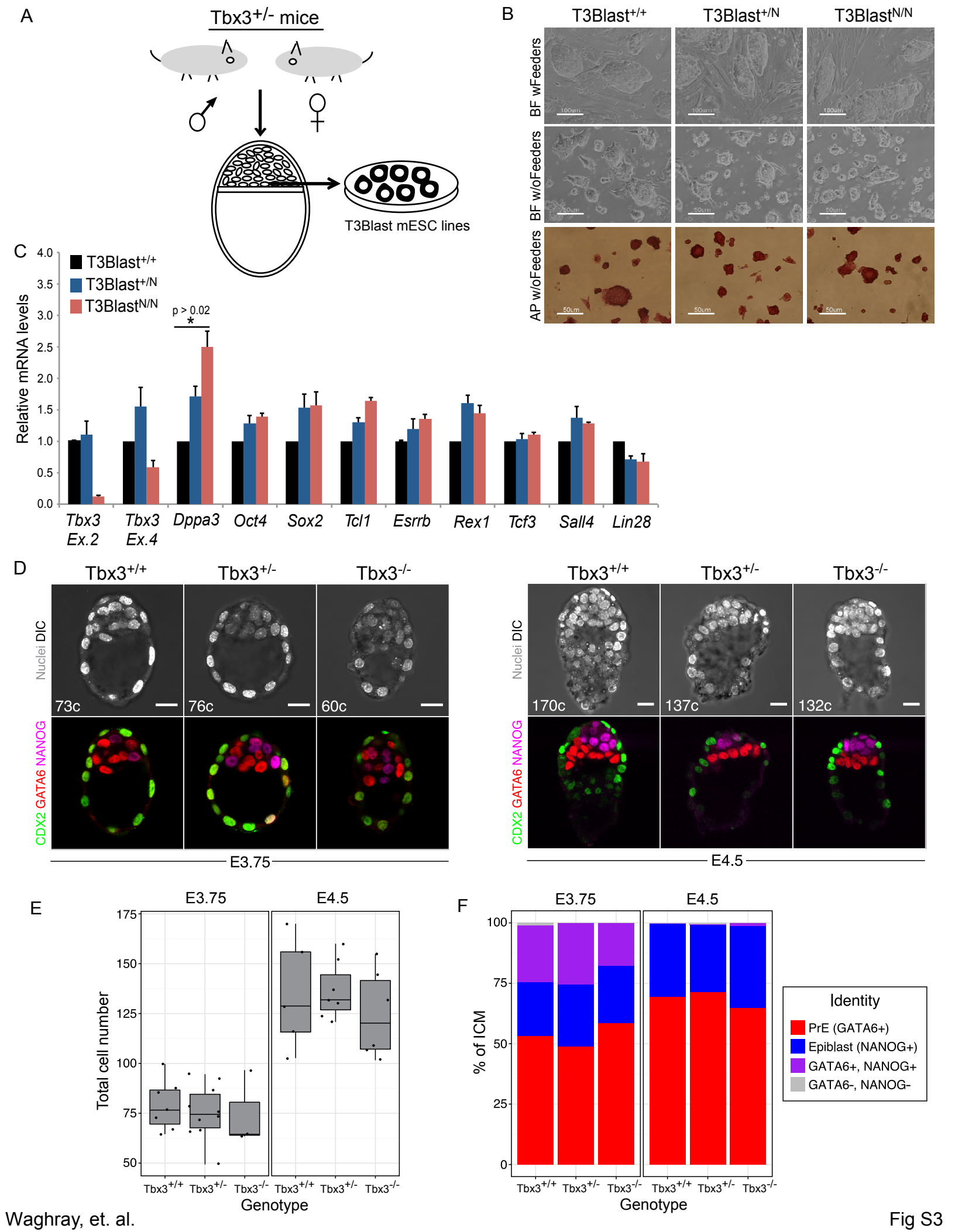
	+/N cells	N/N cells	N/P cells
No. of Blastocyst Injections	1	2	1
# of Chimeric* embryos (Total # of Embryos)	4(8)	11(12)	3(9)
* embryos with eGFP OR tdTomato expression			

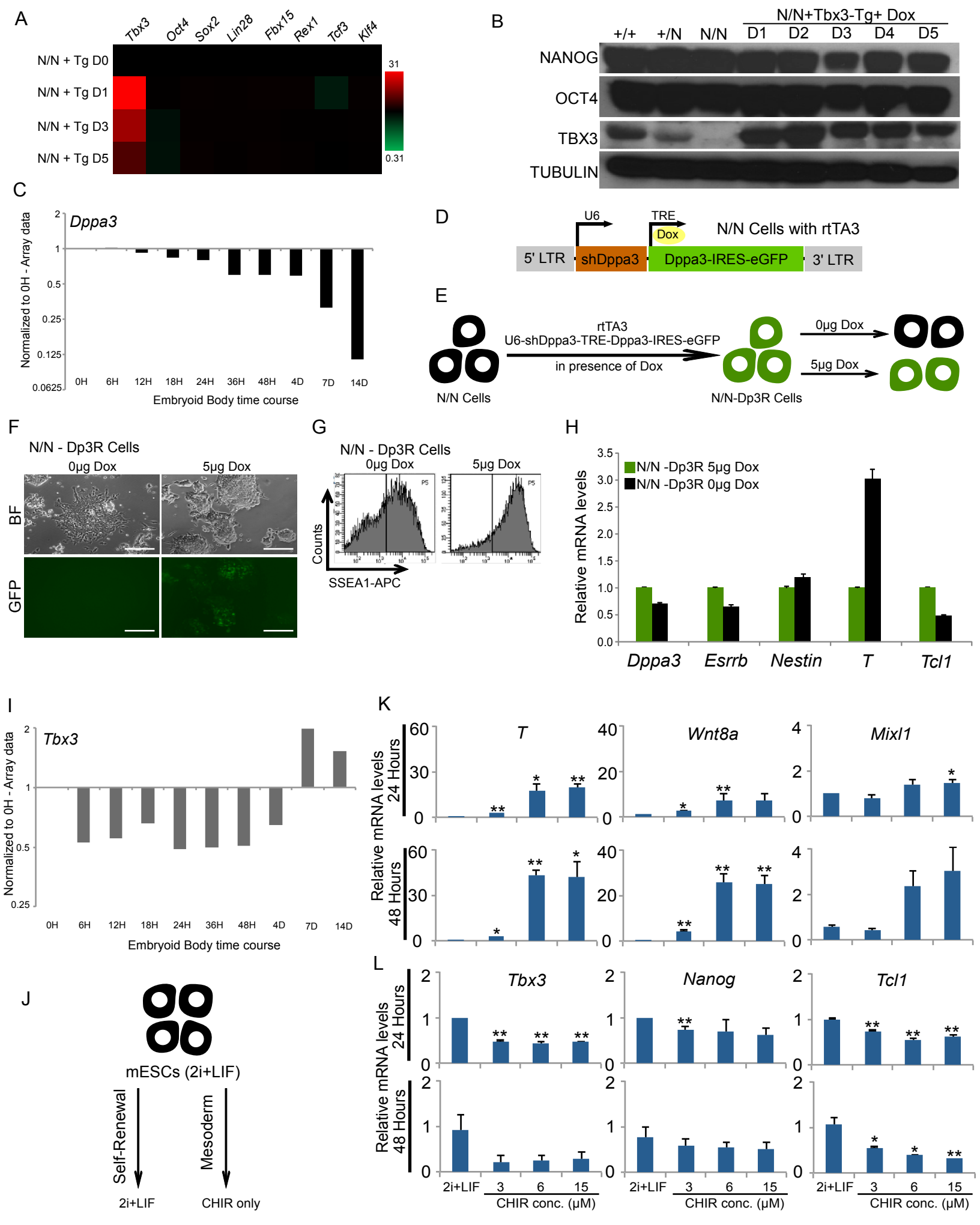

A
H1P
TRE
Dox
AINV15 Cells
5'LTR
shTbx3
Tbx3-IRES-eGFP
3'LTR
B
wDox
w/oDox
BF
GFP
AP
D1
D3
D5
C
wDox
w/oDox
D1
D1
D3
D5
Tbx3
Rex1
Sall4
Esrrb
Tcl1
Cdx2
Ehox
ESC
T.ECTO
10<
>90
Percentile
Fgf5
Mash1
T
Cart1
Couptf2
Gata4
Gata6
Foxa2
ECTO
MESO
ENDO
D
Day 9 w/oDox
BF
AP



Waghray, et. al.

Fig S2


A
Tbx3+/- mice
T3Blast mESC lines
B
T3Blast+/+
T3Blast+/N
T3BlastN/N
BF wFeeders
BF w/oFeeders
AP w/oFeeders
C
T3Blast+/+
T3Blast+/N
T3BlastN/N
p > 0.02
*
Relative mRNA levels
Tbx3 Ex.2
Tbx3 Ex.4
Dppa3
Oct4
Sox2
Tcl1
Esrrb
Rex1
Tcf3
Sall4
Lin28
D
Tbx3+/+
Tbx3+/-
Tbx3-/-
Nuclei DIC
CDX2 GATA6 NANOG
73c
76c
60c
170c
137c
132c
E3.75
E4.5
E
E3.75
E4.5
Total cell number
Genotype
F
E3.75
E4.5
% of ICM
Genotype
Identity
PrE (GATA6+)
Epiblast (NANOG+)
GATA6+, NANOG+
GATA6-, NANOG-


A
Tbx3
Oct4
Sox2
Lin28
Fbx15
Rex1
Tcf3
Klf4
N/N + Tg D0
N/N + Tg D1
N/N + Tg D3
N/N + Tg D5
31
0.31
B
N/N+Tbx3-Tg+ Dox
+/+
+/N
N/N
D1
D2
D3
D4
D5
NANOG
OCT4
TBX3
TUBULIN
C
Dppa3
Normalized to 0H - Array data
2
1
0.5
0.25
0.125
0.0625
0H
6H
12H
18H
24H
36H
48H
4D
7D
14D
Embryoid Body time course
D
U6
TRE
Dox
N/N Cells with rtTA3
5' LTR
shDppa3
Dppa3-IRES-eGFP
3' LTR
E
rtTA3
U6-shDppa3-TRE-Dppa3-IRES-eGFP
in presence of Dox
N/N Cells
N/N-Dp3R Cells
0µg Dox
5µg Dox
F
N/N - Dp3R Cells
0µg Dox
5µg Dox
BF
GFP
G
N/N - Dp3R Cells
0µg Dox
5µg Dox
Counts
SSEA1-APC
H
N/N -Dp3R 5µg Dox
N/N -Dp3R 0µg Dox
Relative mRNA levels
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Dppa3
Esrrb
Nestin
T
Tcl1
I
Tbx3
Normalized to 0H - Array data
2
1
0.5
0.25
0H
6H
12H
18H
24H
36H
48H
4D
7D
14D
Embryoid Body time course
J
mESCs (2i+LIF)
Self-Renewal
2i+LIF
Mesoderm
CHIR only
K
T
Wnt8a
Mixl1
Relative mRNA levels
24 Hours
48 Hours
60
30
0
40
20
4
2
L
Tbx3
Nanog
Tcl1
Relative mRNA levels
24 Hours
48 Hours
2
1
0
2i+LIF
3
6
15
CHIR conc. (µM)

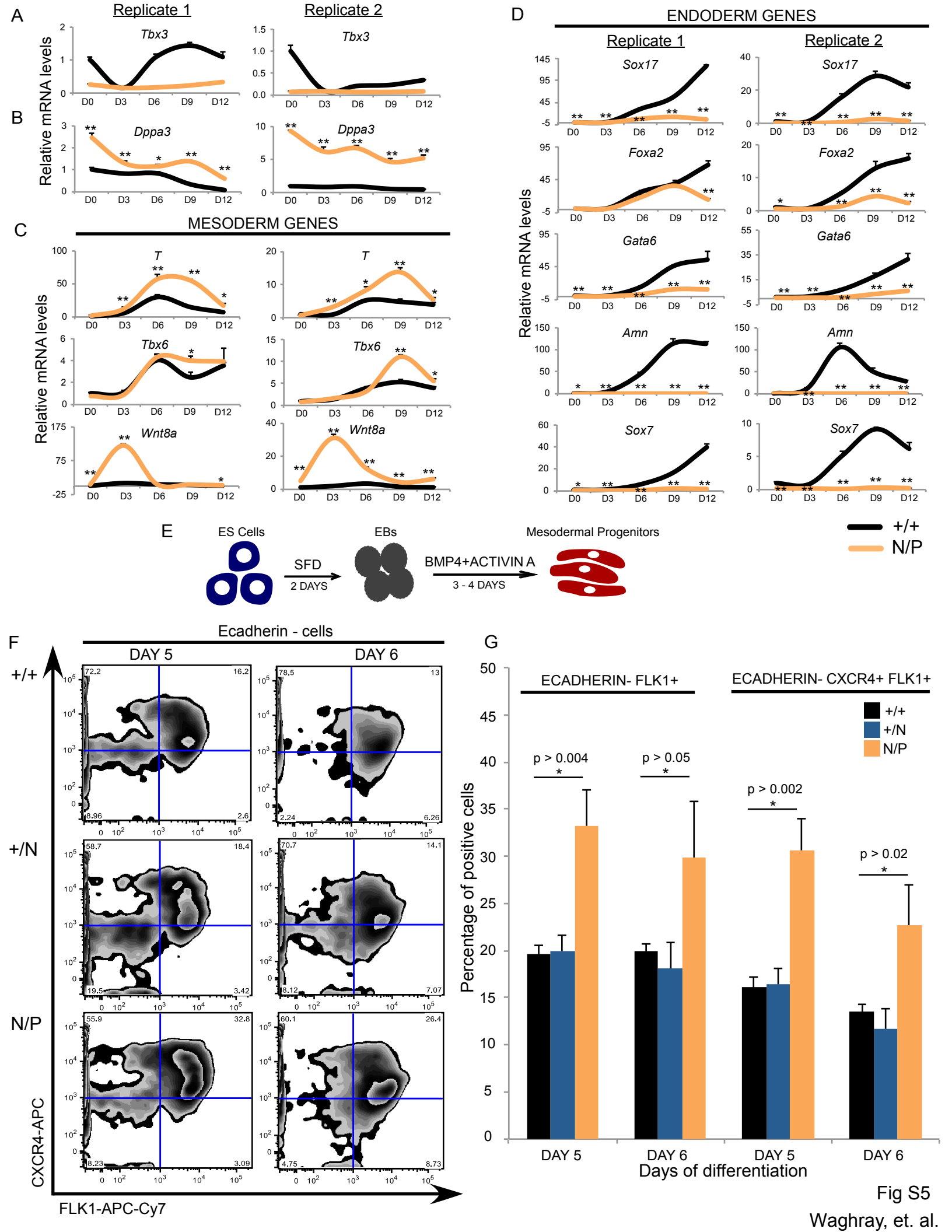
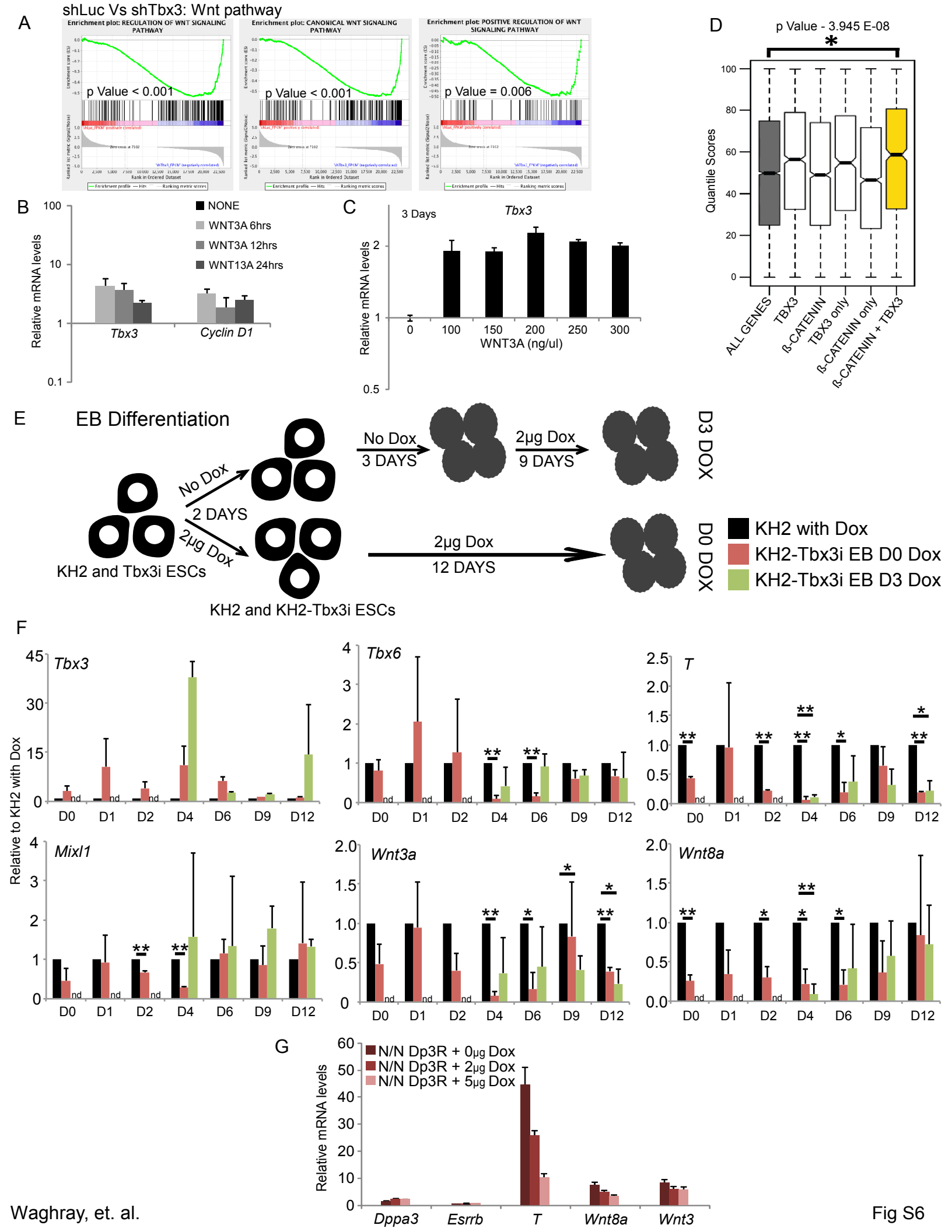


Fig S5

Waghray, et. al.



Waghray, et. al.

Fig S6

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Related to Figure 1. Tbx3 knock down in mESCs. A. The construct used to control Tbx3 levels in Tbx3R cell line. **B.** Short-hairpin (sh)RNA mediated knock down of Tbx3 (w/o Dox) in mESCs causes morphological changes and loss of alkaline phosphatase (AP) staining. Endogenous Tbx3 is knocked down using a constitutively expressed shRNA and is replaced by a Dox-inducible shRNA immune exogenous Tbx3-IRES-eGFP cassette. Tbx3 levels in this system (Tbx3R) are therefore controlled by Dox addition. GFP levels represent expression of inducible Tbx3 transgene. Cells are maintained long term in the presence of Dox (Tbx3) in ESC media containing LIF. Effect of Tbx3 knockdown is assayed 1 day (D1), 3 days (D3) or 5 days (D5) after removal of Dox. **C.** mRNA expression levels of the indicated genes after the knockdown of Tbx3 using the Tbx3R cell line using qRT-PCR. The symbols indicate: ESC- ESC marker, T.ECTO- trophectodermal marker, ECTO- ectodermal marker, MESO- mesodermal marker and ENDO- endodermal markers **D.** Tbx3R cells are grown on gelatin in the absence of Doxycycline (Dox) for 9 days in ESC media containing LIF. Bright field (BF) and AP staining colony pictures are shown.

Figure S2. Related to Figure 2. *In vitro* derived Tbx3 null mESCs. A. Targeting strategy showing the genomic region of Tbx3 that was deleted. Also shown are the locations of the primers used to confirm the targeting. Cre recombinase enzyme was used to delete the neomycin cassette prior to use in blastocyst injection assay in Figure 1G. **B.** Genomic PCR for the indicated fragments to confirm targeting of Tbx3 null mESCs. Fragment 1 spans the deleted region. Fragment 2 amplifies with the neomycin cassette. The 3.4kbp

band from fragment 1 PCR indicates the size of the PCR product prior to deletion of neomycin cassette. **C.** Southern Blotting result using the probe for fragment 2, present in the neomycin cassette. The asterix (*) indicates the position of positive bands. **D.** Level of Tbx3 protein measured by Western blot in indicated cell lines grown in ESC media containing LIF. Non-specific bands are used as loading control. **E.** Karyotype of $Tbx3^{N/N}$ mESCs grown in ESC media containing LIF. **F.** Schematic of the heterokaryon reprogramming assay. **G.** Tbx3 mRNA expression using a human TBX3 specific primer in H1 human (h)ESCs, HES2 hESCs, human B (hB) cells, $Tbx3^{+/+}$ cells, $Tbx3^{+/N}$ cells and $Tbx3^{N/N}$ cells. **H.** mRNA levels of indicated genes over a 2 days after heterokaryon-based reprogramming. NANOG, OCT4 and CRIPTO are used here as ESC markers. CD19 and CD45 are used as B cell markers. HPRT acts as housekeeping gene. Data represents average of three replicates.

Figure S3. Related to Figure 2. Blastocyst derived Tbx3 null mESCs. A. Schematic of $Tbx3^{+/-}$ mice and derivation of T3Blast mESCs. **B.** BF pictures of blastocyst derived T3Blast mESCs grown with (upper row) and without (middle row) feeders. AP staining colony pictures blastocyst derived T3Blast mESCs grown without feeders (lower row). **C.** mRNA **l**evels of indicated ESC markers in mESCs cultured without feeders in the presence of LIF by qRT-PCR. Data is average of two replicates. The p values were generated using the t-test. **D.** Immunofluorescence for representative $Tbx3^{+/+,}$ $Tbx3^{+/-}$ and $Tbx3^{-/-}$ embryos at mid blastocyst (E3.75, left panels) and peri-implantation (E4.5, right panels) stages. CDX2 marks the trophectoderm lineage, GATA6 marks the primitive endoderm (PrE) and NANOG marks the epiblast. Total cell number for each embryo is shown on bright field

images. Scale bars = 20μm. **E.** Box plots showing total cell numbers for all embryos analyzed for each genotype and stage. Each spot represents one single embryo. **F.** Average ICM composition for each genotype and stage, shown as percentage of total ICM cells.

Figure S4. Related to Figure 3. Tbx3 and Dppa3 during mESC differentiation. A. mRNA levels of indicated genes in the $Tbx3^{N/N}$ +Tg cells grown in ESC media containing LIF and Dox (2μg/ml) **B.** Level of indicated proteins in $Tbx3^{+/+}$, $Tbx3^{+/N}$, $Tbx3^{N/N}$ and $Tbx3^{N/N}$ +Tg+Dox cells grown in ESC media containing LIF and Dox (2μg/ml). **C.** mRNA expression pattern of Dppa3 during the EB differentiation time course from published microarray data. **D.** Construct used to generate $Tbx3^{N/N}$-Dp3R cell line from $Tbx3^{N/N}$ cells. **E.** Schematic explaining the strategy for generation of $Tbx3^{N/N}$-Dp3R cells from $Tbx3^{N/N}$ cells followed by removal of Dox (0μg/ml) to knock down Dppa3 in $Tbx3^{N/N}$-Dp3R cells. **F.** Morphology and GFP expression in $Tbx3^{N/N}$ ($Tbx3^{N/N}$-Dp3R) mESCs harboring a Dppa3 genetic complementation system (Dp3R). When these cells are treated with Dox, a Dppa3 transgene compensates for Dppa3 knockdown. GFP reports the levels of transgene expression. To deliver rtTA3, cells were co-infected with an empty pTRIPz vector. **G.** Flow cytometry (FACS) analysis checking levels of SSEA1 protein in $Tbx3^{N/N}$-Dp3R mESCs 7 days after removal of Dox. **H** mRNA levels of indicated genes 7 days after removal of Dox in $Tbx3^{N/N}$-Dp3R cells using qRT-PCR. **I.** mRNA expression pattern Tbx3 during the EB differentiation time course from published microarray data. **G.** Schematic showing set up of experiment for Wnt mediated serum free differentiation into mesoderm lineage. **H-I.** mRNA expression levels of genes during the two day time

course of serum free differentiation. Data is average of two replicates. The p values were generated using the t-test. * p value<0.05 and ** p value < 0.01.

Figure S5. Related to Figure 4. *In vitro* differentiation potential of Tbx3 null mESCs. A-D. The mRNA expression pattern Tbx3 (**A**) and Dppa3 (**B**) genes, mesoderm genes (**C**) and endoderm genes (**D**) during the 12-day EB differentiation time course of $Tbx3^{+/+}$ and $Tbx3^{N/P}$ cells. Data from two independent replicates is shown. The p values were generated using the t-test. * p value<0.05 and ** p value < 0.005. **E.** Schematic of serum free differentiation of mESCs into mesodermal progenitors. **F.** FACS plots displaying levels of CXCR4-APC and FLK1-APC-Cy7 in ECADHERIN-FITC negative cells on Day 5 and Day 6 time points during SFD. **G.** Quantification of the data in section F. FACS data is an average of three independent replicates. The p values were generated using the t-test.

Figure S6. Related to Figure 5. Function of Tbx3 and Wnt signaling in maintenance of mESCs. A. Gene set enrichment analysis to detect statistically significant gene sets in the given datasets. The panels show enrichment of the listed Wnt related gene ontology terms in the given datasets. **B.** Tbx3 mRNA expression in mESCs grown in ESC media containing LIF after a few hours of treatment with Wnt3a. **C.** Tbx3 mRNA expression in mESCs grown in ESC media containing LIF after 3 days of WNT3A treatment at indicated concentrations. **D.** mRNA expression profile from normalized mRNA-Seq dataset from wild type R1 mESCs. The expression data is represented as quantile scores, described in supplemental methods. The different datasets indicated here are: ALL GENES –All genes in the mRNA-Seq dataset; Tbx3 –TBX3 bound genes from ChIP-Seq study; β-catenin - β-

CATENIN bound genes from ChIP-Seq study; Tbx3 only – TBX3 bound genes not bound by β-CATENIN from ChIP-Seq studies; β-catenin only - β- CATENIN bound genes not bound by TBX3 from ChIP-Seq studies; β-catenin+Tbx3 – TBX3 and β- CATENIN co-bound genes. P-Value is calculated using the Wilcoxon rank sum test with continuity correction. **E.** Schematic showing differentiation of KH2 and KH2-Tbx3i mESC lines as EBs. Dox (2μg/ml) is used to induce over expression of Tbx3. Tbx3 is over expressed either prior to differentiation (lower part) or after 3 days of differentiation (upper part). The parental KH2 line is used as control. **F.** mRNA Levels of Tbx3 and mesodermal markers during the EB differentiation process measured by qRT-PCR. Data represents average of two replicates. The p values were generated using the t-test. * p value<0.05 and ** p value < 0.005. **G.** mRNA levels of indicated genes in $Tbx3^{N/N}$ cells after removal of Dox (knock down of *Dppa3*).