Investigating the Role of *Brachyury (T)* in the Control of Chromatin State Within Early Mesodermal Cells

Inaugural-Dissertation
to Obtain the Academic Degree
Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry, and Pharmacy
of Freie Universität Berlin

by

Arica Beisaw
from Tokyo, Japan

September 2014

The present work was carried out at the Max Planck Institute for Molecular Genetics, Department of Developmental Genetics, in Berlin, Germany, between May 2010 and September 2014, under the supervision of Dr. Phillip Grote and Prof. Dr. Bernhard G. Herrmann.

1st Reviewer: **Prof. Dr. Bernhard G. Herrmann**
Max Planck Institute for Molecular Genetics
Ihnestraße 63-73
14195 Berlin

2nd Reviewer: **Prof. Dr. Stephan Sigrist**
Institute of Biology – Freie Universität Berlin
Takustraße 6
14195 Berlin

Date of Defense: January 22, 2015

Summary

The gene *Brachyury, T,* encodes for the founding member of the T-box family of transcription factors and was previously shown to be crucial for the proper specification and development of the mesodermal germ layer during mouse embryogenesis. In addition to their ability to directly activate and repress transcription by binding to T-box binding sites within target genes, T-box factors were also shown to modulate gene expression by their ability to modify the chromatin environment at target genes through a physical interaction with the MLL/Trithorax family of H3K4 methyltransferases, H3K27 demethylases, and BRG1, a component of the SWI/SNF chromatin remodeling complex of enzymes. This interaction is important for the function of most T-box family members, as multiple mutations in T-box factors which are predicted to disrupt this interaction are found in human patients with congenital developmental disorders. I generated a point mutation within the endogenous *T* locus of mouse embryonic stem cells, creating an amino acid substitution (Y88A) in the T protein, to determine whether the interaction of T with histone-modifying enzymes is necessary for proper T function during mesodermal development. Mouse embryos expressing the mutant T^{Y88A} allele resemble *T* knockout embryos and lose the ability to maintain *T* expression, suggesting that the Y88 residue is necessary for proper T function and for maintenance of *T* expression itself. Gene expression profiling of *T*-positive mesodermal cells from embryos expressing the T^{Y88A} allele revealed that 60% of dysregulated genes were unique to the T^{Y88A} mutant when compared to *T* knockout embryos, suggesting that the amino acid substitution disrupts a critical function of T separate from the maintenance of *T* expression. Furthermore, gene expression profiling corroborated previous data that *T* may play a role in hemangioblast commitment, and uncovered a role for T in mammalian erythrocyte specification and development. Genome-wide histone modification profiling in *T*-positive cells differentiated into mesoderm *in vitro* revealed that functional interaction with H3K4 methyltransferases and H3K27 demethylases was not disrupted in the T^{Y88A} mutant, however, H3K27 acetylation at T binding sites was decreased. Overall, the data in this thesis suggest that the Y88 residue in T is necessary for proper T function during mesodermal specification, and is necessary in the control of H3K27 acetylation at target loci in early mesodermal cells.

Zusammenfassung

Brachyury, T, wurde als erstes Gen, das für einen T-Box Transkriptionsfaktor kodiert identifiziert und begründet daher diese Genfamilie. *Brachyury* ist essentiell für die Spezifizierung und Entwicklung des mesodermalen Keimblattes während der Embryonalentwicklung. T-Box Faktoren binden sogenannte T-Box Zielsequenzen in der genomischen DNA und können dadurch in der Nähe gelegene Gene aktivieren oder reprimieren. Darüberhinaus konnte gezeigt werden, dass T-Box Faktoren durch die Bindung von Proteinkomplexen der MLL/Trithorax Familie der H3K4 Histonmethyltransferasen, H3K27 Histondemethylasen und dem Chromatinremodellierungs-Komplex BRG1, einem Mitglied der SWI/SNF Familie, die Genaktivität beeinflussen können. Diese Interaktion ist für die physiologische Funktion der T-Box Faktoren sehr wichtig, so sind mehrere Mutationen im humanen Genom bekannt, die Aminosäureaustausche in T-Box Faktoren zur Folge haben, die diese Interaktionen verhindern und damit bei betroffenen Patienten zu schweren Symptomen führen.

Um festzustellen wie essentiell die Bindung von chromatinmodifizierenden Proteinkomplexen durch T-Box Faktoren ist, wurde das Genom von embryonale Stammzellen der Maus so verändert, dass es für ein T Protein mit einem Aminosäureaustausch (Y88A) kodiert, welches diese Interaktion nicht mehr zulassen sollte. Die sich aus diesen veränderten Stammzellen entwickelnden Mausembryonen weisen morphologisch dieselben Phänotypen wie *T* Null Mutanten auf. Des weiteren wird Brachyury-Expression in den T^{Y88A} Mutanten initiiert, aber nicht aufrecht erhalten, was zeigt, dass Y88 wichtig für die normale Funktion von T ist, und T wichtig für die Aufrechterhaltung der Expression seiner selbst. Der Vergleich von Genexpressionsprofilen zwischen *T* positiven, mesodermalen Zellen aus dem frühen Embryo von T^{Y88A} und *T* Null Mutanten zeigt, dass 60% der deregulierten Gene spezifisch nur in T^{Y88A} Mutanten dereguliert sind. Den meisten dieser Gene kann eine Funktion in der Entwicklung des Hämangioblasten, einer Population multipotenter Stammzellen aus der Endothelzellen und Blutzellen hervorgehen, zugeordnet werden. Die genomweite Analyse und Vergleich von in vitro gewonnenen mesodermalen *T* positiven Zellen der

T^{Y88A} Mutante zeigt ausserdem, dass H3K4- und H3K27-Methylierung nicht aberrant ist. Jedoch ist die H3K27- Acetylierung an Loci an denen T direkt bindet deutlich reduziert.

Zusammengenommen zeigen die hier gewonnen Daten, dass die über die Aminosäure Y88 vermittelte Interaktion von BRACHYURY mit anderen Komplexen essentiell ist für bestimmte Funktionen von BRACHYURY während der Mesodermentwicklung und dass diese Interaktion wichtig ist für die Acetylierung von Chromatin an von BRACHYURY gebundenen genomischen Regionen.

Acknowledgements

Contributions to Experimental Work

The work contained in this thesis was performed with the help from the following people: Karol Macura performed all tetraploid complementation assays to generate mouse embryos from genetically modified embryonic stem cells. Dr. Lars Wittler helped perform microdissections of caudal ends from mutant and control embryos for gene expression profiling. Dr. Phillip Grote helped with sorting of *T-mCherry*-positive cells by FACS. Dr. Frederic Koch generated libraries for RNA-sequencing using RNA isolated from *T-mCherry*-positive cells from $T^{2J/+}$;*Tg*($T^{mCherry}$)*1Bgh* and $T^{2J/Y88A}$;*Tg*($T^{mCherry}$)*1Bgh* embryos. In addition, he provided the gene expression profiling dataset for *T-mCherry*-positive cells from $T^{2J/2J}$ knockout embryos and the genome-wide T binding site dataset from mesodermal cells differentiated *in vitro*. He mapped RNA-Seq reads from *T-mCherry*-positive cells from $T^{2J/+}$ and $T^{2J/Y88A}$ embryos and ChIP-Seq reads from histone-modification profiling of *in vitro* differentiated *T-mCherry*-positive cells to the mouse genome, and provided files for data analysis. Matt Huska provided the Q-Q plots comparing enrichment in dimerization in dysregulated genes versus non-dysregulated genes. The rest of the work contained within this thesis was performed by me.

Abbreviations

°C	degrees Celsius
3'	3 prime end
^{32}P	radioactive phosphorus
3D	three dimensional
5'	5 prime end
ACTH	adrenocorticotropic hormone
AP	alkaline phosphatase
BAC	bacterial artificial chromosome
BFU-E	burst forming unit erythroid
BirA	biotin ligase enzyme
bp	base pair
BSA	bovine serum albumin
CAGGS	CMV immediate-early enhancer, CBA promoter, and CBA intron 1/exon 1 hybrid promoter
cDNA	complementary DNA
CFU-E	colony forming unit erythroid
ChIP	chromatin immunoprecipitation
ChIP-Seq	chromatin immunoprecipitation coupled to high-throughput sequencing
CHX	cycloheximide
Ci	Curie
cm	centimeter
cm^2	square centimeters
CMV	cytomegalovirus promoter
CoIP	co-immunoprecipitation
DNA	deoxyribonucleic acid
DAPI	4',6-diamidino-2-phenylindole
DAVID	Database for Annotation, Visualization, and Integration Discovery
dCTP	deoxycytosine triphosphate
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethylsulfoxide
Dox	doxycycline
DPBS	deionized phosphate buffered saline
E. coli	*Escherichia coli*
E	embryonic day
EB	embryoid body
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EMT	epithelial-to-mesenchymal transition
EpiLC	epiblast-like cell
ESC	embryonic stem cell
FACS	fluorescence activated cell sorting
Flp/FRT	Flp recombinase/flippase recognition target
FlpE	Flp recombinase E

FPKM	fragments per kilobase of transcript
G418	Gentecin 418
GO	gene ontology
H3K27ac	histone 3 lysine 27 acetylation
H3K27me2/3	histone 3 lysine 27 di/trimethylation
H3K27me3	histone 3 lysine 27 trimethylation
H3K4me1	histone 3 lysine 4 monomethylation
H3K4me2	histone 3 lysine 4 dimethylation
H3K4me3	histone 3 lysine 4 trimethylation
HEK293T	human embryonic kidney cell line 293 (SV40 Large T-antigen transformed)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HS	high-sensitivity
IRES	internal ribosomal entry site
ISH	*in situ* hybridization
KDM	lysine (K) demethylase
K-S	Kolmogorov–Smirnov
kb	kilobase
KD4-T	ubiquitous T knockdown ESC line
LIF	leukemia inhibitory factor
LNA	locked nucleic acid
log_2FC	log_2 of the fold change
M	molar
MABT	maleic acid buffer with 0.1% Tween-20
MAMEP	Molecular Atlas of the Mouse Embryo Project
mg	milligram
mL	milliliter
MLL	Mixed Lineage Leukemia
mM	millimolar
mm	millimeter
mm10	*Mus musculus* genome annotation 10
mmol	millimole
mRNA	messenger RNA
mV	millivolts
NaCl	sodium chloride
NaOH	sodium hydroxide
neo	neomycin
ng	nanogram
nmol	nanomole
NMP	neuromesodermal progenitor
OptiMEM	Optimized Modified Eagle Medium
pA	poly A
pBS	pBluescript
PBS	phosphate buffered saline
pBS_DTA	pBluescript vector with diphtheria toxin fragment A
PCR	polymerase chain reaction
PGK	phosphoglycerate kinase promoter

PolII	RNA Polymerase II
PSM	presomitic mesoderm
Q-Q	quantile-quantile
qPCR	quantitative PCR
qRT-PCR	quantitative reverse-transcriptase PCR
RefSeq	Reference Sequence database
RIN	RNA integrity number
RNA	ribonucleic acid
RNA-Seq	high-throughput sequencing of expressed RNA
rpm	revolutions per minute
rRNA	ribosomal RNA
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
shRNA	short hairpin RNA
siRNA	small interfering RNA
SSC	saline-sodium citrate buffer
SWI/SNF	SWItch/Sucrose NonFermentable nucleosome remodeling complex
T	*Brachyury* gene
T^{2J}	*Brachyury* null allele
T^{c}	*Brachyury* curtailed antimorphic allele
T^{Wis}	*Brachyury* Wisconsin antimorphic allele
TAE	Tris-acetate-EDTA buffer
TBS	T binding site
TBST	Tris Buffered Saline with 0.1% Tween-20
$TBX6^{WT}$	wild-type TBX6 protein
TE	Tris-EDTA buffer
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride
tRNA	transfer RNA
TS	Theiler Stage
TSS	transcription start site
T^{WT}	wild-type T protein
T^{Y88A}	Y88A T protein
T^{Y88A}	T allele encoding for the Y88A amino acid substitution
UCSC	University of California, Santa Cruz
UTR	untranslated region
UV	ultraviolet
V	volts
WISH	whole-mount *in situ* hybridization
WT	wild-type
Y88A	tyrosine to alanine amino acid substitution at position 88 of T
α-DIG-AP	anti-digoxigenin antibody coupled to alkaline phosphatase
ΔΔCT	delta delta Ct
µF	microFaraday
µg	microgram
µJ	microjoule
µL	microliter
µM	micromolar

Table of Contents

Summary i

Zusammenfassung ii

Acknowledgements iv

Contributions to Experimental Work v

Abbreviations vi

List of Figures and Tables xi

Introduction 1

1.1 Early Embryonic Development in the Mouse 2

1.2 Mesoderm Development and Axial Elongation 2

1.3 The Role of Brachyury in Mesoderm Development 4

1.4 The T-box Family of Transcription Factors 7

1.5 T Within the Mesodermal Gene Regulatory Network 10

1.6 T as a Regulator of Chromatin State 13

1.7 Does T Exert Its Function as a Master Regulator through Modification of the Chromatin Environment in Early Mesodermal Cells? 21

Results 23

2.1 Generation of the T^{Y88A} Point Mutation *In Vivo* 24

2.2 Rescue of $T^{2J/Y88A}$ Mutation by Transgene Integration of Wild-type *T* 31

2.3 Confirmation of T^{Y88A} as a Null Allele 35

2.4 The Functional Interaction of T with Histone-Modifying Proteins is Necessary for Maintenance of *T* Expression 38

2.5 T^{Y88A} Protein Levels Decrease in $T^{2J/Y88A}$ Embryos During Development 40

2.6 Does the Sudden Loss of *T* Expression from E8.0 Cause the *T* Knockout Phenotype in $T^{2J/Y88A}$ Mutants? 44

2.7 Expression of Downstream Target Genes in $T^{2J/Y88A}$ Mutant Embryos 48

2.8 Comparison of Dysregulated Gene Expression in $T^{2J/Y88A}$ versus $T^{2J/2J}$ Mutants 53

2.9 Comparison of RNA-seq Data to Known T Binding Sites 59

2.10 Investigating the Physical Interaction Between T and Histone- and Chromatin-Modifying Enzymes 61

2.11 Investigating the Functional Consequences of Loss of T Interaction with Histone-Modifying Proteins 65

2.11.1 H3K27me3 and H3K4me3 in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutant Mesodermal Cells 69

2.11.2 H3K4me1 in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutants 71

2.11.3 H3K27ac in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutants 73

2.11.4 Investigating Enhancer Regions in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutants 79

Discussion 81

3.1 The Phenotypic Consequences of T^{Y88A} Mutation 82

3.2 Is Dimerization of T Effected in the T^{Y88A} Mutant? 85

3.3 Gene Expression Profiles in $T^{2J/Y88A}$ Compared to $T^{2J/2J}$ Knockout Embryos 87

3.3.1 Similarities in Gene Expression Profiles in $T^{2J/Y88A}$ Versus $T^{2J/2J}$ Knockout Embryos 87

3.3.2 A Role for *T* in Hemangioblast Commitment and Erythrocyte Development 89

3.4 The Histone Modification Landscape in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutant Mesodermal Cells 94

3.5 A Model for T Function 97

Materials and Methods 99
4.1 BAC Recombineering and Plasmid Generation 100
4.1.1 Recombineering of the *T* BAC 100
4.1.2 Generation of the T^{Y88A} Homologous Targeting Vector 101
4.1.3 Generation of Plasmid Constructs 101
4.2 Tissue Culture 102
4.2.1 ESC Culture Media 103
4.2.2 ESC Culture 104
4.2.3 Generation of Modified ESCs by Electroporation of Linearized BACs and Targeting Constructs 104
4.2.4 Lipofectamine Transfection of FlpE Recombinase 105
4.2.5 Picking Individual ESC Clones 105
4.2.6 Splitting and Freezing of ESC Clones 106
4.2.7 Genomic DNA Isolation for Southern Blot 106
4.2.8 Restriction Digest and Transfer of Genomic DNA 106
4.2.9 Probe Labeling and Hybridization 107
4.2.10 Genotyping PCR of ESC Clones 108
4.2.11 Quantification of BAC DNA Integration 108
4.2.12 *In Vitro* Differentiation of ESCs into Mesodermal Cells 108
4.2.13 Culture and Transfection of Non-ESCs 109
4.2.14 Generation of P19 Cell Lines Stably Over-expressing Tagged T and Differentiation into Mesoderm 109
4.2.15 Generation of Fully ESC-derived Embryos by Tetraploid Complementation 110
4.3 Molecular Biology Techniques 111
4.3.1 Whole-Mount *In Situ* Hybridization (WISH) 111
4.3.2 Isolation of RNA 111
4.3.3 cDNA Synthesis 112
4.3.4 Gene Expression Analysis (qRT-PCR) 112
4.3.5 Sequencing of *T* cDNA 112
4.3.6 Quantification of Protein Concentration 112
4.3.7 Western Blotting 113
4.3.8 Immunohistochemistry 113
4.3.9 Cycloheximide Chase Assay 114
4.3.10 Co-Immunoprecipitation of Proteins 114
4.4 Gene Expression Profiling 116
4.4.1 Embryo Dissection and Isolation of *T-mCherry*-positive Cells 116
4.4.2 RNA Isolation, Library Preparation, and Sequencing 116
4.4.3 RNA-Seq Data Analysis 117
4.5 Profiling the Histone Modification Landscape 118
4.5.1 Isolation of *T-mCherry*-positive Differentiated ESCs and ChIP using Histone Mark Specific Antibodies 118
4.5.2 Library Prep and Sequencing 118
4.5.3 Analysis of ChIP-Seq Data 119
4.6 Primers 120
4.7 Antibodies 121
Appendix 123
References 131

List of Figures and Tables

Figure 1.1: Early embryogenesis in the mouse 3
Figure 1.2: *T* expression and its role in mesoderm formation 6
Figure 1.3: Phylogenetic tree of the T-box family of transcription factors 8
Figure 1.4: T-Box factor structure and conservation 16
Figure 1.5: Potential T modulation of the chromatin environment through interactions with histone- and chromatin-modifying complexes 22
Figure 2.1: Introduction of the Y88A point mutation into the endogenous *T* locus 25
Figure 2.2: Removal of the neomycin selection cassette from $T^{+/Y88A}$ and $T^{2J/Y88A}$ ESCs 26
Figure 2.3: T^{Y88A} heterozygous and mutant embryo phenotypes 28
Figure 2.4: $T^{+/Y88A}$ embryos at E9.5 morphologically resemble wild-type embryos 30
Figure 2.5: BAC integration to rescue the $T^{2J/Y88A}$ mutant phenotype 32
Figure 2.6: Rescue of the $T^{2J/Y88A}$ mutant phenotype by integration of a wild-type copy of *T* 34
Figure 2.7: Generation of $T^{+/Y88A}$*;Tg(*T^{Y88A}*neo)1Bgh* ESCs 35
Figure 2.8: Phenotype of $T^{+/Y88A}$*;Tg(*T^{Y88A}*neo)1Bgh* embryos 37
Figure 2.9: Expression of *T* in the $T^{2J/Y88A}$ mutant embryos 39
Figure 2.10: Levels of T protein in $T^{2J/Y88A}$ mutant embryos 40
Figure 2.11: Testing the stability of T^{Y88A} protein using a cycloheximide chase assay 42
Figure 2.12: Cellular localization of T^{Y88A} mutant protein 43
Figure 2.13: *T* expression in *KD4-T* knockdown embryos 45
Figure 2.14: Notochord and somite formation in *KD4-T* embryos. 47
Figure 2.15: Integration of the *T-mCherry* reporter BAC into $T^{2J/+}$ and $T^{2J/Y88A}$ ESCs 49
Figure 2.16: Sorting and RNA isolation from *T-mCherry*-positive cells 51
Figure 2.17: Verification of RNA-seq data 52
Figure 2.18: Hierarchical clustering of dysregulated genes in $T^{2J/Y88A}$ compared to $T^{2J/2J}$ mutants 54
Figure 2.19: Dysregulated genes with a $\log_2$FC>1.5 at TS13 55
Figure 2.20: GO term enrichment analysis of dysregulated genes in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants 57
Table 1: Changes in expression of hemangioblast genes in the $T^{2J/Y88A}$ mutant 58
Figure 2.21: Comparing dysregulated genes and dimerization of T 60
Figure 2.22: Co-immunoprecipitation of T with KDMs in HEK293T cells 62
Figure 2.23: Co-immunoprecipitation of T with KDMs in differentiated P19 cells 64
Figure 2.24: Differentiation of ESCs into mesoderm recapitulates the *in vivo* situation 66
Figure 2.25: Sorting and shearing of chromatin from *T-mCherry*-positive cells 67
Figure 2.26: ChIP-Seq set-up and library verification 68
Figure 2.27: H3K27me3 profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control 70
Figure 2.28: H3K4me3 profiles at TSS in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control 71
Figure 2.29: H3K4me1 profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control 72
Figure 2.30: H3K27ac profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control 74
Figure 2.31: Analyzing the regions of lower H3K27ac in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant mesodermal cells 76

Figure 2.32: Regions of lower H3K27ac correspond to T binding in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants 77
Figure 2.33: GO term enrichment analysis of genes associated with lower H3K27ac in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants 79
Figure 2.34: Comparison of active enhancer regions around T binding sites in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants 80
Figure 3.1: Model of T function during mesodermal differentiation 98
Supplementary Figure S1: Confirmation of $T^{2J/+}$ heterozygosity 124
Supplementary Figure S2: mCHERRY protein stability 125
Supplementary Figure S3: *T-mCherry* reporter expression compared to *T* expression in *in vitro* differentiated mesodermal cells 126
Supplementary Figure S4: Histone mark profiles at the TSS of dysregulated genes 127
Supplementary Figure S5: H3K27me3 profiles at regions of lower H3K27ac in $T^{2J/2J}$ and $T^{2J/Y88A}$ mutant mesodermal cells 128
Supplementary Table S1: Total vs. Mapped Reads from RNA-Seq and ChIP-Seq 129

Chapter 1.

Introduction

The establishment of a mature, multi-cellular organism from a single-celled fertilized egg is a remarkable process that requires the orchestration of multiple aspects including cellular proliferation, morphogenesis, lineage specification, and pattern formation. Throughout mammalian embryonic development, multiple mechanisms work together to ensure cell-type specific gene expression, in order to control the identity and function of each cell. It is the carefully coordinated action of growth factor signaling pathways, tissue-specific transcription factors, and factors which modify the chromatin environment that allows for the proper development of a functional, multi-cellular organism.

1.1 Early Embryonic Development in the Mouse

After fertilization, the single-celled mouse embryo undergoes multiple cell divisions to give rise to the blastocyst. The blastocyst consists of pluripotent epiblast cells in the inner cell mass which will develop into the embryo proper, along with the trophectoderm and primitive endoderm which both contribute to extraembryonic structures necessary to support normal embryonic development and provide a source of signals to pattern embryonic tissues prior to gastrulation. After implantation into the uterus of the mother at E4.5, the mouse embryo elongates and forms a hollow cup-shaped structure (Figure 1.1A). The process of gastrulation commences at E6.5 to form and ensure proper positioning of the three primary germ layers of the embryo; ectoderm, endoderm, and mesoderm. It is from these primary germ layers that all fetal tissues will arise; the ectoderm gives rise, for example, to neural tissues, neural crest, and epidermis, endoderm gives rise to the gut and associated organs including liver, pancreas, lungs, and the intestinal tract, and the mesodermal germ layer gives rise to the heart, skeleton, muscle, blood and vasculature, kidneys, and gonads (Lawson and Pedersen, 1992; Lawson et al., 1991).

1.2 Mesoderm Development and Axial Elongation

One hallmark of gastrulation includes the formation of a structure known as the primitive streak at the posterior side of the developing embryo. Epiblast cells ingress through the primitive streak by undergoing an epithelial-to-mesenchymal transition (EMT) and migrate to their respective positions to form the mesoderm and definitive endoderm (Figure 1.1B). Interestingly, it has been shown that distinct mesodermal cell lineages are specified according to the time and position of ingression through the primitive streak (Lawson, 1999). For example, the most posterior mesoderm subpopulations, which are exposed to BMP4 signaling from the extraembryonic ectoderm, differentiate into the mesodermal layer of the chorion, visceral yolk sac mesoderm, and blood islands (Winnier et al., 1995). Mesoderm arising from the intermediate to anterior portion of the streak gives rise to lateral plate, cardiac, and

paraxial mesoderm. Finally, epiblast cells that ingress through the most anterior part of the primitive streak give rise to mesodermal tissues of the prechordal plate, node, and notochord (Kinder et al., 1999).

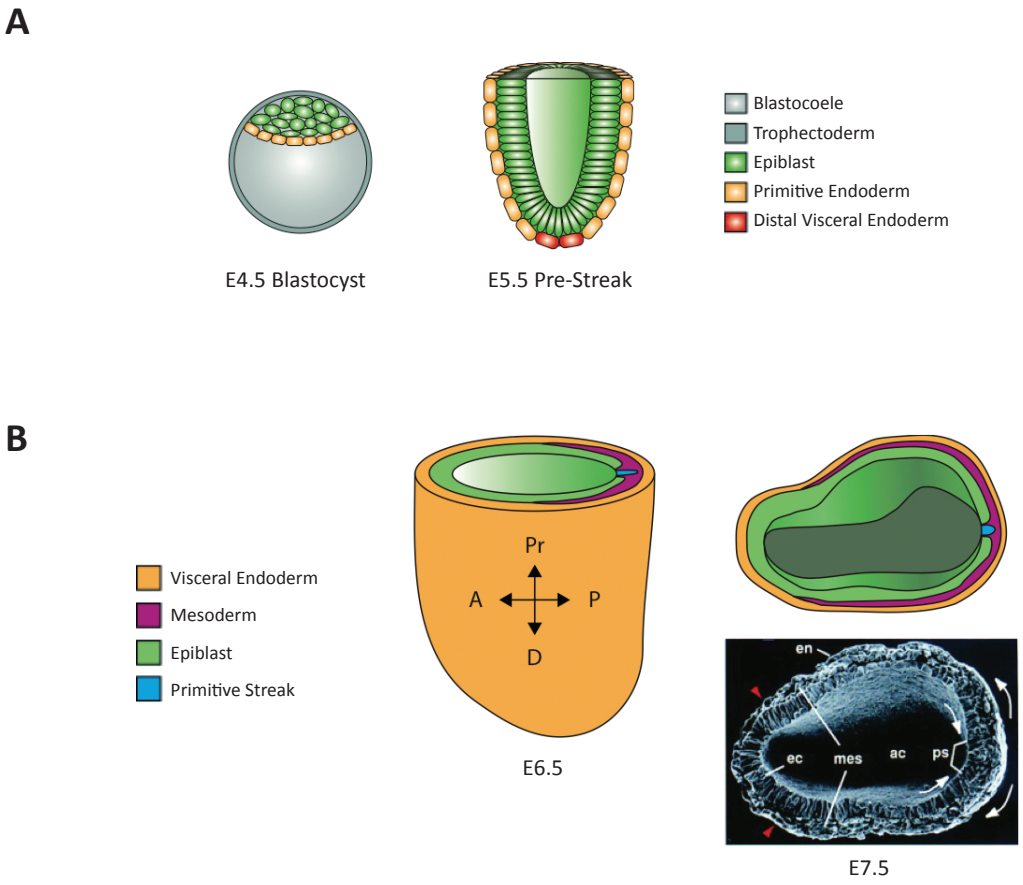


Figure 1.1: Early embryogenesis in the mouse
(A) Schematic illustrating the murine blastocyst (E4.5) and early post-implantation (E5.5) embryo. The distal visceral endoderm acts as a source of signaling molecules to pattern the early embryo. (B) Schematic illustrating a gastrulating murine embryo from E6.5 to E7.5. The micrograph (Sun et al., 1999) represents a transverse section through an E7.5 mouse embryo undergoing gastrulation. Epiblast cells ingress through the primitive streak and undergo EMT to migrate away to respective positions within the embryo (white arrows). Red arrowheads indicate the border to which mesodermal cells migrate. Figure modified from Dr. Eun-Ha Shin (PhD thesis, Max Planck Institute for Molecular Genetics, Berlin). A, anterior; P, posterior; D, distal; Pr, proximal; en, visceral endoderm; ec, epiblast cells; mes, mesoderm; ps, primitive streak; ac, amniotic cavity.

After initial primitive streak formation and establishment of the anterior-posterior axis from E6.5-E7.5, elongation of the body axis and initial differentiation of multiple structures and organs occurs. After turning of the embryo at E8.75, the elongation of the posterior axis is thought to occur as progenitor cells in the caudal end of the embryo differentiate and give rise to structures of the trunk and tail. Interestingly, single-cell labeling from E7.5-E9.0 mouse embryos *in utero,* through the use of a LaacZ intragenic somatic recombination system, revealed the presence of an axial stem cell population which gives rise to both neural tube and paraxial mesodermal tissue in the trunk of the developing mouse embryo (Tzouanacou et al., 2009). This contradicted the previously accepted hypothesis that neural tissues arise from the ectoderm and paraxial mesoderm from the mesodermal germ layer, established during gastrulation. Rather, it appears as if some lineage determination decisions occur later in development in the caudal end of the embryo. *Brachyury*, described below, in conjunction with activated Wnt signaling, has been implicated in the maintenance of this neuromesodermal progenitor (NMP) cell population (Martin and Kimelman, 2012).

1.3 The Role of Brachyury in Mesoderm Development

One of the earliest studied regulators of mesoderm formation, *Brachyury (*gene symbol *T)*, was discovered in 1927 (Dobrovolskaia-Zavadskaia, 1927), when a mouse strain arose from a spontaneous mutation resulting in a 160kb deletion on chromosome 17. Heterozygous *T* mutant mice displayed a shortened tail, some malformations in presacral vertebrae, and fusion of the notochord during embryonic development with the neural tube, cloaca, or tail gut (Grüneberg, 1958), while embryos homozygous for *T* mutation died *in utero*. Early mutant analyses revealed that homozygous *T* mutants displayed severe axis truncation, kinked neural tube formation, loss of notochord tissue, formation of only 7-9 abnormally-formed anterior somites, defects in heart formation, and reduction in allantois size (Figure 1.2B, Chesley, 1935; Gluecksohn-Schoenheimer, 1944; Grüneberg, 1958). This reduction in allantois size results in the failure of the allantois to properly connect to the chorion and leads to embryonic death around E10.0. The gene responsible for the *T* phenotype was cloned in 1990 by Bernhard G. Herrmann

(Herrmann et al., 1990) and was the founding member of the T-box family of transcription factor-coding genes. *T* is first expressed in early gastrulation stage embryos at the posterior end of the primitive streak, and expression extends anteriorly throughout the primitive streak to its region of highest expression in the node (Figure 1.2A). Mesodermal cells lose expression of *T* after ingression through the primitive streak and subsequent migration to lateral parts of the embryo. From E8.0-E8.5, expression of *T* is found in the developing notochord and continues to be expressed in cells ingressing through the primitive streak at the posterior end of the embryo. After the embryo has turned at E9.0, *T* expression is restricted to the notochord and the tail bud, the source of mesodermal progenitor cells which are required for axis elongation. After axis elongation and tail formation have ceased, *T* transcripts are only expressed in the notochord and notochordal derivatives (Herrmann, 1991; Wilkinson et al., 1990). *T* expression has also been found in hemangioblast cells, along with *Flk-1*. These $T^{+}/Flk\text{-}1^{+}$ double-positive cells were found in the posterior primitive streak and proximal parts of the extraembryonic mesoderm, and can give rise to both hematopoietic and endothelial cells of developing blood vessels in the yolk sac (Huber et al., 2004). Immunohistochemical analysis of T protein localization confirmed previous domains of *T* expression and also revealed additional domains of T in the extraembryonic mesoderm, adjacent to the posterior primitive streak and in the midline of the allantois. In addition, T was also found in the heart and chorionic ectoderm (Inman and Downs, 2006a; Kispert and Herrmann, 1994).

Later studies of T function in mesoderm development revealed that T is necessary for mesodermal cells to migrate away from the primitive streak after ingression. Analysis of *T/T* mutant chimeras demonstrated the accumulation of mutant *T* cells in the primitive streak and later in the tail bud, suggesting that *T* mutant cells are defective in cell adhesion or migration (Rashbass et al., 1991; Wilson et al., 1995; Yanagisawa et al., 1981). Further studies also demonstrated that a main role for *T* is in notochord cell identity and maintenance, stemming from a defect in node formation in *T* mutants; *T* mutant embryos display a loss of a functional notochord and pyknotic cells within a notochordal precursor-like structure (Chesley, 1935; Rashbass et al., 1991; Yanagisawa, 1990). It is hypothesized that defects in neural tube and somite formation

seen in *T* mutants is a secondary effect of the loss of the notochord and its ability to signal to neighboring cell types.

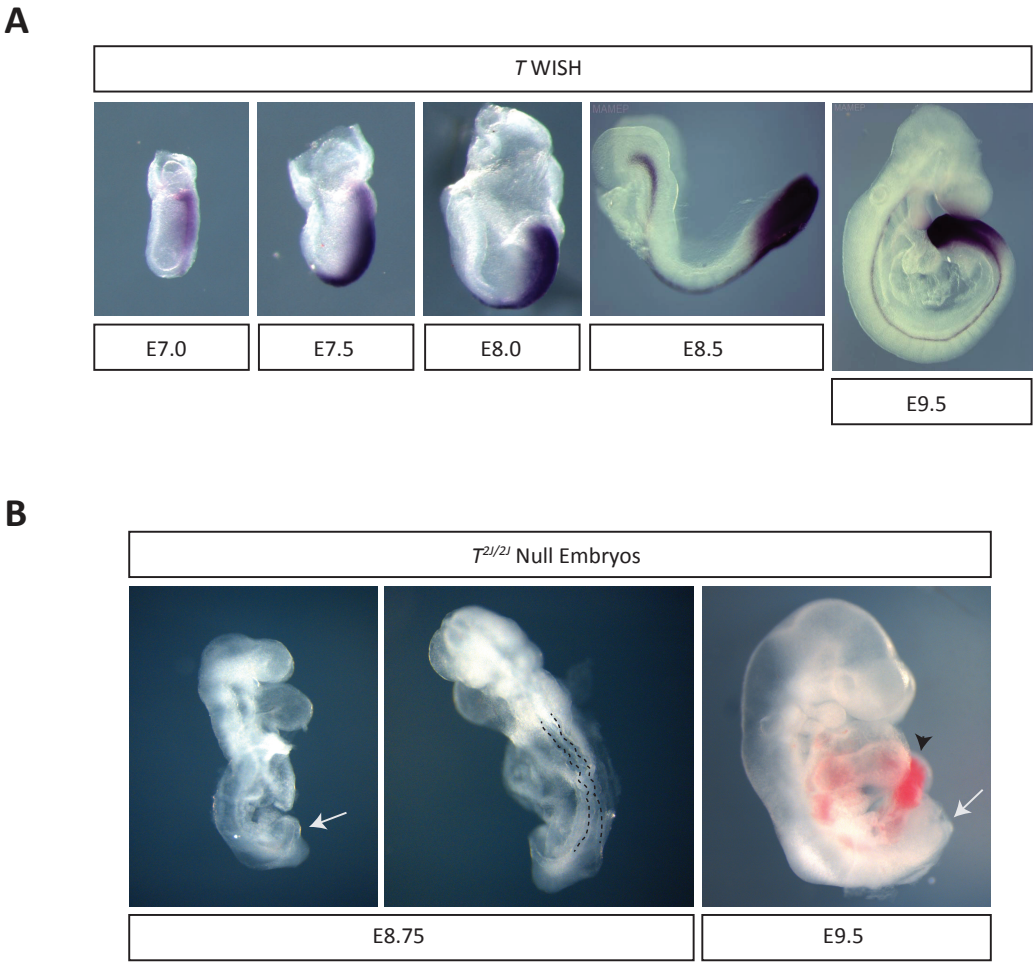


Figure 1.2: *T* expression and its role in mesoderm formation
(A) Whole-mount *in situ* hybridization (WISH) of *T* during early stages of mouse development. *T* is expressed in the primitive streak, notochord, and tail bud. Images from E7.0 to E8.0 are from $T^{2J/+}$ heterozygous embryos, although the expression domain is similar to that seen in wild-type embryos. Images from E8.5 to E9.5 are from wild-type embryos taken from the Molecular Atlas of the Mouse Embryo Project (MAMEP) database (Max Planck Institute for Molecular Genetics, Berlin). (B) $T^{2J/2J}$ null embryos display phenotypes as described in the text. White arrows point to the truncated axis, as a result of the loss of posterior mesoderm. Black dashed lines indicate the kinked neural tube in a dorsal view of the E8.75 embryo. The black arrowhead indicates abnormal allantois formation.

Interestingly, there is also a correlation between the severity of phenotype and dosage of wild-type *T* and it has been postulated that more caudal structures require higher doses of *T* for proper development since development of anterior structures proceeds normally in *T* mutants (Stott et al., 1993; Yanagisawa, 1990). Analysis of various mutant *T* alleles suggested that severity of axis truncation and defects in notochord formation were directly correlated to wild-type *T* dosage (Yanagisawa, 1990). Heterozygous $T^{+/-}$ mice are viable but display a variable shortened-tail phenotype. Embryos heterozygous for the T^{C} or T^{Wis} mutation display an invariable tailless phenotype, due to the antimorphic nature of these C-terminally truncated *T* alleles (Herrmann et al., 1990; Macmurray and Shin, 1988; Shedlovsky et al., 1988). Addition of a single-copy transgene expressing wild-type *T* was shown to rescue both the $T^{C/+}$ tailless phenotype and the $T^{+/-}$ short-tailed phenotype in a dose-dependent manner (Stott et al., 1993), confirming that the dosage of *T* directly correlated to severity of the tail phenotype.

1.4 The T-box Family of Transcription Factors

Biochemical studies demonstrated that the *T* gene encodes a nuclear-localized DNA-binding transcription factor (Kispert and Herrmann, 1993; Kispert et al., 1995). Selection of binding sites from a random pool of oligonucleotides revealed that T binds to DNA containing a palindromic consensus binding site with a core sequence of AGGTGT within each half-site and that this binding is conferred by the N-terminus (T-box) of the protein. It was further demonstrated that T can activate the expression of a reporter construct containing consensus T binding sites *in vitro* and that the C-terminus of the protein contains both transcriptional activation and repression domains. Crystallographic structure analyses of the N-terminal T-box of T with DNA established that T binds to the palindromic DNA site as a dimer, with a C-terminal helix which binds to the minor groove of DNA without causing bending (Müller and Herrmann, 1997).

PCR with degenerate primers and inspection of amino acid sequence conservation within the T-box established that T belonged to a family of related transcription factors that have since been shown to play multiple roles during embryonic

development and morphogenesis (Bollag et al., 1994). To date, 17 T-box family members have been found in mammals (Figure 1.3), many of which have orthologous genes in a variety of multicellular organisms. Many T-box family members have been shown to play crucial roles in multiple aspects of embryonic development. For example, *Eomes* and *T* are necessary for proper specification of the mesendodermal and mesodermal germ layer, respectively (Russ et al., 2000; Ryan et al., 1996). *Tbx6* is essential for proper specification of the presomitic mesoderm into somites (Chapman and Papaioannou, 1998). *Tbx3* has been shown to play a role in mammary gland, limb, and heart development (reviewed in Washkowitz et al., 2012). *Tbx4* and *Tbx5* have been shown to play crucial roles in proper limb development (Rodriguez-Esteban et al., 1999; Takeuchi et al., 1999), while *Tbx5* and *Tbx20* play important roles in cardiac development (Horb and Thomsen, 1999; Szeto et al., 2002). The importance of this family of transcription factors is also exemplified by the multitude of human congenital diseases and a subset of cancers which arise from heterozygous mutations in members of the T-box family (reviewed in Packham and Brook, 2003).

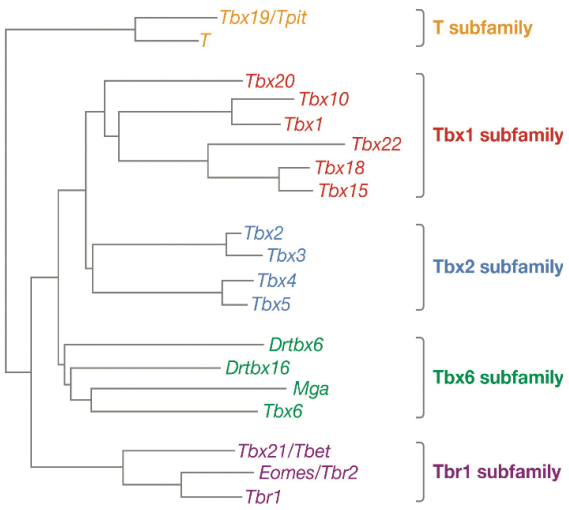


Figure 1.3: Phylogenetic tree of the T-box family of transcription factors
The phylogenetic relationship of T-box factor genes in 5 subfamilies, indicated on the right. All genes are present in mammals with the exception of *Drtbx6* and *Drtbx16,* which are found in zebrafish. Figure modified from (Naiche et al., 2005).

An interesting question arose pertaining to T-box factor specificity and mechanisms of action from studies of *Tbx4* and *Tbx5* in chick limb development. *Tbx4* is expressed in the developing hindlimb while *Tbx5* expression is found in the forelimb (Isaac et al., 1998; Logan et al., 1998). Misexpression of *Tbx4* in the forelimb of the chick resulted in a leg-like structure, with *Tbx5* expression unaffected. Conversely, misexpression of *Tbx5* in the hindlimb repressed *Tbx4* expression and gave rise to a wing-like structure (Rodriguez-Esteban et al., 1999; Takeuchi et al., 1999). Despite the fact that these proteins share 94% amino acid similarity within the T-box, their ability to regulate gene expression and cell fate is vastly different. This, combined with the observation that many T-box factors tested to date can bind to the original *Brachyury* target sequence (Carreira et al., 1998; Hsueh et al., 2000; Tada et al., 1998), suggests that DNA binding at the T-box is not the sole determinant of specificity. Indeed, some specificity arises from the orientation and spacing of T-box binding sites (Conlon et al., 2001). Furthermore, the number of T-box binding sites and cooperative binding of T-box factors may also play a role in activating target gene expression (Kusch et al., 2002). It has also been hypothesized that interacting protein partners may play a role in T-box specificity. For example, in early heart development, TBX5 has been shown to interact with multiple protein partners to control gene expression including NKX2.5, GATA4, TBX20, and MEF2C (Brown et al., 2005; Garg et al., 2003; Ghosh et al., 2009; Hiroi et al., 2001). T has been shown to interact with SMAD1 through its N-terminus (Marcellini et al., 2003) and with MIXL1 to antagonize its activity (Pereira et al., 2011). Furthermore, ChIP-Seq data of T binding reveals colocalization of binding of multiple T-box factors and Smad2/3 with T in developing Xenopus embryos (Gentsch et al., 2013) and multiple transcription factors motifs, including SOX17, MYF, and E2F1, near T binding sites in an *in vitro* model of primitive streak cells derived from mouse embryonic stem cells (ESCs; Lolas et al., 2014).

T has also been shown to homodimerize, both by examining its crystal structure (Müller and Herrmann, 1997) and in yeast-two-hybrid studies (Kusch et al., 2002). The importance of this homodimerization is exemplified by the dominant negative phenotype apparent in $T^{c/+}$ and $T^{Wis/+}$ heterozygous animals, both of which contain C-terminal truncations of the *T* gene. It has also been hypothesized that T may form heterodimers with other T-box transcription factors, as the stretch of amino acids

important for dimerization is conserved within other members of the T-box family. However, the actual interaction of heterodimers and the importance of this interaction has not yet been investigated.

1.5 T Within the Mesodermal Gene Regulatory Network

T lies centrally within a mesodermal gene regulatory network which controls the molecular differentiation of epiblast cells into early mesodermal cells. Experiments performed in *Xenopus* animal cap explant assays demonstrated that *Xbra* (the ortholog of *T* in *Xenopus*) is activated in the absence of protein synthesis by both activin A and basic FGF (Smith et al., 1991). Further studies showed that activation of both TGFβ and FGF signaling were required for initiation of *Xbra* expression in the *Xenopus* embryo (Amaya et al., 1991; Hemmati-Brivanlou and Melton, 1992). The FGF signaling pathway was also shown to be important for maintenance of *Xbra* expression during subsequent mesoderm development, and these molecules function in an autoregulatory loop to maintain expression of each other. Thus, ectopic expression of *Xbra* can induce expression of *eFGF* and vice-versa, while expression of dominant negative forms of FGF receptor or Xbra results in loss of *Xbra* and *eFGF* expression, respectively. Mesodermal explants from *Xenopus* embryos retain the ability to maintain *Xbra* expression, however, disaggregation of mesodermal explants, and subsequent dilution of secreted growth factors such as FGF, resulted in loss of *Xbra* expression. This could be rescued by addition of exogenous FGF to the disaggregated mesodermal explants, demonstrating the importance of FGF signaling in maintaining *Xbra* expression (Amaya et al., 1993; Conlon and Smith, 1999; Conlon et al., 1996; Isaacs et al., 1994; Schulte-Merker and Smith, 1995). In line with these observations, *Xbra* was also shown to bind to a T-box half-site in the *eFGF* promoter to activate transcription and that this regulatory region was found to be conserved in the promoters of mouse and human *eFGF* (Casey et al., 1998). Consequently, the role of the FGF signaling pathway in mesodermal formation and *T* activation/maintenance is conserved in mice. Complete or conditional loss of *Fgfr1* or of both *Fgf4* and *Fgf8* from the primitive streak resulted in defects in axial

extension and downregulation of *T* expression (Boulet and Capecchi, 2012; Deng et al., 1994; Naiche et al., 2011; Wahl et al., 2007; Yamaguchi et al., 1994).

Studies of the *T* promoter in mouse and frog also revealed the presence of TCF/LEF binding sites, required for the activation of canonical Wnt signaling, suggesting that Wnt signaling also plays a role in maintenance of *T* expression (Galceran et al., 2001). *Wnt3a* null embryos and embryos lacking *Tcf1/Lef1,* display truncation of the axis posterior to the first 7-9 somites and ectopic neural structures in place of paraxial mesoderm. Furthermore, WNT3A signaling directly activates *T* expression in the developing mouse embryo via TCF/LEF binding sites in the *T* promoter (Galceran et al., 1999; Takada et al., 1994; Yamaguchi et al., 1999; Yoshikawa et al., 1997). Experiments performed in zebrafish by Martin and Kimelman demonstrated that *ntl* and *bra* (the orthologs of *T* in zebrafish), in combination with canonical Wnt signaling, function together in a positive autoregulatory loop to maintain mesodermal cells in an undifferentiated state during axial elongation. This is thought to occur through the ability of *ntl* and Wnt signaling to cross-regulate one another, and through the ability of *ntl* to activate *cyp26a1*, which degrades differentiation-inducing retinoic acid (Martin and Kimelman, 2008, 2010).

Both canonical Wnt signaling and *T* have been implicated in directing differentiation of NMP cells into mesodermal cell types. Through single-cell, cell-autonomous manipulation of Wnt signaling, Martin and Kimelman demonstrated that inhibition of Wnt signaling resulted in a significant decrease in single-cell derived clones contributing to somitic tissue and an increase in contribution to spinal cord neurons (Martin and Kimelman, 2012). Continuous activation of Wnt signaling resulted in a significant decrease of NMP cells giving rise to a neural fate, demonstrating that Wnt signaling is both necessary and sufficient to direct NMP cells towards a mesodermal cell fate. Furthermore, they showed that Wnt signaling contributes to a second lineage decision between paraxial mesodermal and endothelial vascular progenitor cells in the zebrafish embryo. *T* has also been implicated in differentiation of mesodermal cell types from NMP cells through studies of genome-wide T binding in *Xenopus* embryos and differentiated mouse ESCs. Genome-wide profiling of Xbra binding in *Xenopus* embryos revealed that Xbra binds a multitude of genes implicated in mesodermal differentiation including *tbx6, msgn1, mespa, mespb, myoD, myf5,* and *xbra* itself. Furthermore,

expression of these genes is down-regulated upon knockdown of *xbra* and *xbra3*. Xbra was also shown to bind near genes implicated in neural development, including *sox2, zic2, pax3,* and *ngn3*, all of which were up-regulated upon *xbra/xbra3* knockdown (Gentsch et al., 2013). Surprisingly, the overlap between Xbra binding and dysregulated gene expression upon *xbra/xbra3* knockdown was small, with only 10% of Xbra target genes dysregulated upon knockdown. This may be due to compensation by other T-box factors in the early *Xenopus* embryo, including Eomes and VegT, which were also shown to bind to the same target regions as Xbra. However, this does not rule out the possibility that Xbra (or T) binding to the genome may serve functions other than transcriptional activation. ChIP-seq experiments of T binding in an *in vitro* differentiated model of mouse primitive streak cells revealed similar patterns. T was shown to bind regions of multiple genes involved in mesodermal cell development including *Fgf8, Msgn1, Mesp2, Tbx6,* and *Hoxd8*. Furthermore, T was shown to bind to neural developmental genes including *Pax2, Zic2,* and *Id3*. Gene expression profiling by RNA-seq of mouse embryos homozygous for the T^{c} mutation revealed that multiple mesodermal genes are both bound and regulated by T, including *Fgf8, Tbx6, Msgn1,* and *Hoxd8*. Furthermore, T binds and regulates genes involved in EMT, such as *Sox4* and *Bves*, and *Foxa2*, a gene implicated in node and notochord development (Lolas et al., 2014). Taken together, from the experiments described above, T has been shown to play a crucial role in the mesodermal gene regulatory network, both in maintaining mesodermal cells in an undifferentiated state, and further promoting mesodermal cell differentiation and repressing neural fate in NMP cells.

1.6 T as a Regulator of Chromatin State

Another line of evidence that *T* functions as a master regulator of mesoderm formation is its known functional interaction with chromatin- and histone-modifying proteins. The past decade of research has clearly shown the importance of post-translational modification of histone proteins and their role in organizing chromatin structure and accessibility. The compaction and accessibility of chromatin is known to play a role in regulating gene activity, and opening of chromatin is essential for transcriptional activation of gene expression. It has become clear that there are multiple histone modifications present at any number of loci in the genome, and that it is their combinatorial code which marks specific states of transcriptional activity. Methylation of lysine residues in histone tails is a modification that has received much attention, and it has been appreciated through computational analysis of genome-wide datasets that the methylated residue, the degree of methylation, and their position within the genome can distinguish active versus inactive gene bodies, promoter regions, enhancer regions, and exonic versus intronic boundaries. For example, methylation of lysine (K) 27 of histone 3 (H3K27me3), deposited in promoters and gene bodies by the Polycomb complex of proteins, has been associated with transcriptionally silent regions of the genome. H3K4me3 is found at the proximal promoters of both active and inactive genes, and within the first few exons of actively expressed genes. Furthermore, it has been suggested that H3K4me3 may play a functional role in the initiation of transcription, as TAF3 can bind H3K4me3 and possibly facilitate recruitment of RNA polymerase II (Vermeulen et al., 2007). H3K4me1 is found within introns of gene bodies and has been shown to demarcate enhancer regions in combination with H3K27 acetylation (Hon et al., 2009; Rada-Iglesias et al., 2011). Proteins of the MLL/Trithorax complex have been shown to mediate H3K4 mono-, di-, and trimethylation *in vivo*, including MLL1-4, and SETD1A and SETD1B proteins (Cho et al., 2007; Hughes et al., 2004; Lee and Skalnik, 2008; Lee et al., 2007a; Wu et al., 2008) . SETD7 and SMYD1-4 proteins also have the ability to methylate H3K4 histone tails through their SET enzymatic domains, however their limited ability to methylate nucleosomal substrates, and the fact that they have been shown to have multiple non-histone substrates, argues

that they may not play major roles in methylating H3K4 *in vivo* (Hamamoto et al., 2004; Herz et al., 2013; Nishioka et al., 2002; Wang et al., 2001).

Methylation of histones is a dynamic process, and enzymes with the ability to remove histone methylation have been discovered within the past 10 years. These histone demethylase (KDM; lysine (K) demethylase) proteins have the ability to remove methylation deposited by methyltransferases, and thus contribute to the regulation of the chromatin landscape. Of particular importance for this thesis are the H3K27 demethylases, KDM6A (UTX) and KDM6B (JMJD3), which have the ability to remove H3K27me2/3 marks. These proteins are thought to contribute to transcriptional activation of genes by removing repressive H3K27me3 marks from promoters and gene bodies throughout the genome. These, and other histone-modifying proteins, are not recruited to target loci through sequence-specific recognition of DNA elements alone. There is evidence that interactions with RNA polymerase II machinery, tissue-specific transcription factors, and non-coding RNAs may all play a role in recruiting these histone-modifying enzymes to specific loci (Bertani et al., 2011; Kim et al., 2009; Lee et al., 2009; Rinn et al., 2007; Wang et al., 2011; Zhao et al., 2008; Zhu et al., 2005).

In 2007, Lewis, Miller, and colleagues demonstrated that the T-box family member T-BET, which is critical for T-helper 1 cell differentiation, has the ability to recruit permissive H3K4me2 histone methyltransferase activity to modify the chromatin environment at key lineage determinant genes such as *Ifn-γ* and *Cxcr3* and that this activity is independent from its ability to activate transcription (Lewis et al., 2007). Furthermore, in 2008, the same group demonstrated that T-BET can physically interact with H3K4 methyltransferase complex component RBBP5 (Miller et al., 2008). They also demonstrated the ability of T-BET to interact with H3K27 demethylase activity and localized this activity to two conserved pockets of residues within the T-box domain (termed T-box pocket 1 and 2 in Figure 1.4A). It was shown that both binding to H3K4 methyltransferase complex components and binding to H3K27 demethylase activity were conserved among different T-box family members, including T, TBX6, and EOMES. The importance of the interaction of T-box factors with these histone-modifying proteins was revealed by the presence of multiple mutations within the T-box pocket 1 that were associated with congenital human disease. For example, mutation of tyrosine residue 100 in TBX5 was found in patients with congenital heart disease while mutation of the

analogous residue in TBX3 was found in patients with Ulnar mammary syndrome (Figure 1.4B, Bamshad et al., 1997, 1999; Reamon-Buettner and Borlak, 2004). Introduction of these tyrosine mutations into T-BET (Y182A) and T (Y88A) resulted in loss of interaction with both H3K4 methyltransferase component RBBP5 and H3K27 demethylase activity. From these experiments, it was suggested that in addition to their ability to activate transcription directly, T-box factors are also important for establishing the chromatin environment at their target loci by recruitment of enzymes which remove repressive H3K27me2/3 and enzymes which deposit permissive H3K4me2. Moreover, this function of T-box factors is essential for their proper activity, as mutations within domains which recruit histone-modifying enzymes in multiple T-box factors results in human congenital disease.

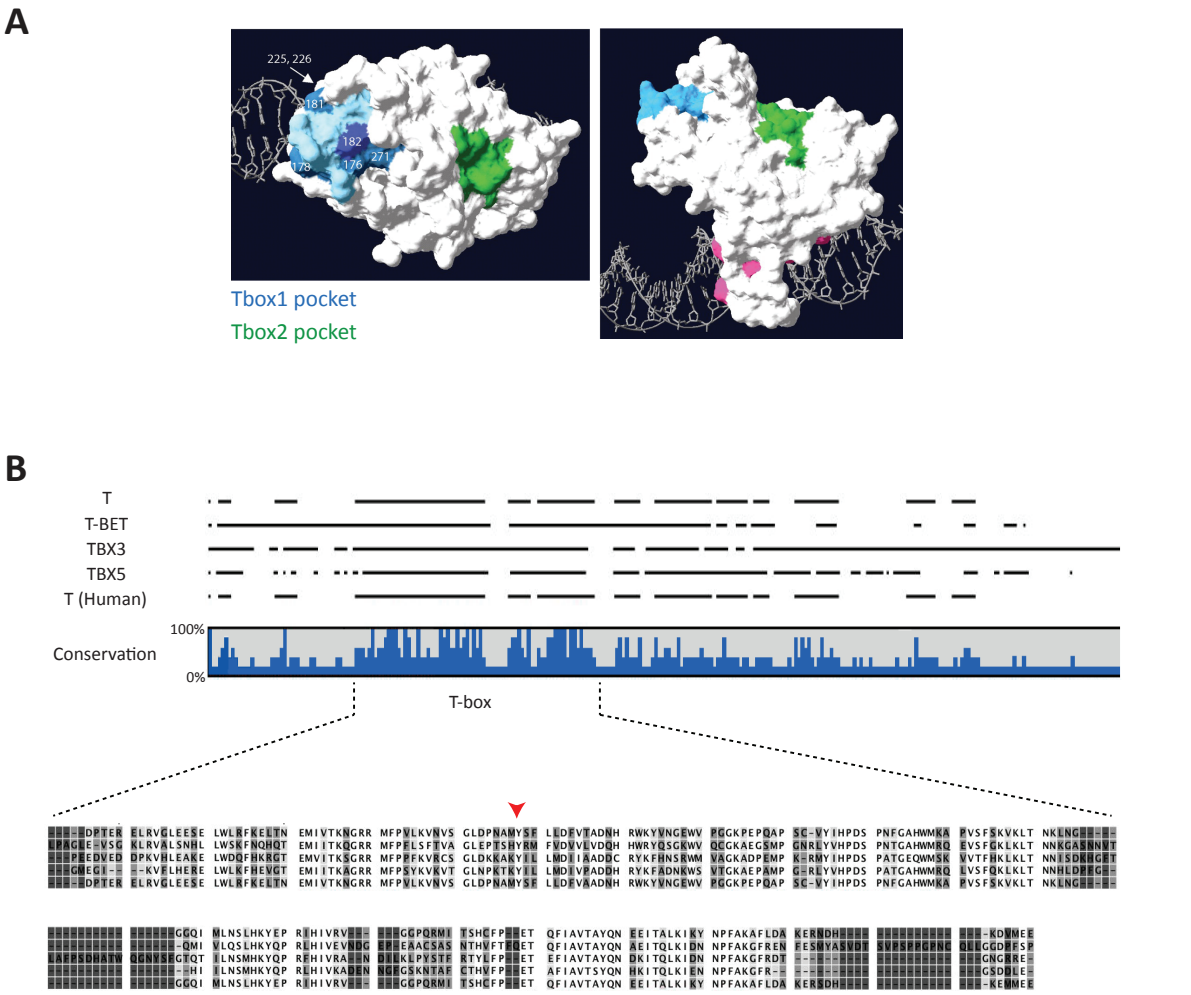


Figure 1.4: T-Box factor structure and conservation
(A) T-BET superimposed onto the crystal structure of TBX3 bound to DNA. Analyses by Miller and colleagues revealed that conserved pockets of residues (here termed Tbox1 and Tbox2 pockets) exist when comparing various members of the T-box family of transcription factors, or when comparing the same T-box factor across multiple species. These conserved pockets lie outside of regions which contact DNA. Figure adapted from (Miller et al., 2008). (B) Conservation analysis revealed that a conserved tyrosine residue is found in the Tbox1 pocket (red arrowhead) which is mutated in multiple human congenital diseases.

Miller and colleagues went on to show that T-box factors also contribute to modification of the chromatin environment at target genes through their interaction with SWI/SNF ATPase-dependent chromatin remodeling complex component BRG1 (Miller et al., 2010). Through a restriction enzyme accessibility assay, they were able to show that interaction of the T-box factors with H3K27 demethylases KDM6A and KDM6B was necessary for binding to BRG1 and accessibility of chromatin at target gene

promoters. From the experiments described above, the T-box family of transcription factors appears to direct lineage specification and cell-type specific gene expression programs through multiple mechanisms, including direct activation of target genes and modification of the surrounding chromatin environment at target genes. This led to our hypothesis that these functions are also utilized by T during mesodermal lineage specification and differentiation. It is tempting to consider that a master regulator of mesoderm formation not only functions through its direct activation of target gene expression, but that it also has the ability to prime chromatin at regions in mesodermal genes that may be activated by other transcription factors later during mesodermal differentiation and development. Furthermore, the interaction with H3K4 methyltransferases, H3K27 demethylases and SWI/SNF chromatin remodeling complexes may be necessary for the proper function of T during embryonic development.

In line with our hypothesis that T functions together with H3K27 demethylase proteins during development, both KDM6A and KDM6B have been shown to have roles in embryonic development and development of structures arising from the mesodermal germ layer. For example, KDM6A (UTX) was shown to be important for regulation of *Hox* gene expression in human embryonic teratocarcinoma cells (Agger et al., 2007). Knockdown of the *Kdm6a (Utx)* ortholog in zebrafish results in axial elongation defects and defects in proper somite formation, which were attributed to the improper regulation of posterior *Hox* genes (Lan et al., 2007). Knockout of *Kdm6a* in mice results in embryonic lethality at E10.5, most likely due to defects in cardiac morphogenesis, including heart looping and myocardial trabeculation, and loss of proper cardiac-specific gene expression. Furthermore, KDM6A was shown to interact with cardiac-specific transcription factors including SRF, GATA4, NKX2.5, and TBX5 and to synergize with these transcription factors to activate expression of cardiac genes (Lee et al., 2012; Shpargel et al., 2012). *Kdm6a* knockout mice also displayed defects in posterior trunk development and in hematopoiesis, including a marked reduction in red blood cells in the Aorta-Gonad-Mesonephros and in the yolk sac vasculature (Shpargel et al., 2012; Wang et al., 2012). In line with this, *Kdm6a* knockdown in murine primary bone marrow cells results in the decreased ability of these cells to form erythroid colonies in CFU-E and BFU-E assays and granulocyte-macrophage colonies (Liu et al., 2012). Furthermore,

knockdown of *Kdm6a* in adult mice resulted in lowered red blood cell counts and hemoglobin levels, along with disrupted maturation in erythroid, megakaryocytic, and granulocyte lineages (Thieme et al., 2013).

Multiple *in vitro* studies have also suggested the importance of *Kdm6a* in the differentiation of cells from the mesodermal lineage. For example, *Kdm6a* was shown to be important for the demethylation of H3K27me3 at muscle-specific genes during *in vitro* differentiation of C2C12 muscle progenitor cells (Seenundun et al., 2010). During the formation of embryoid bodies (EBs), loss of *Kdm6a* results in defective mesodermal lineage differentiation as evidenced by a loss of *T* expression. Moreover, KDM6A can bind to the *T* promoter and promote the recruitment of RNA Polymerase II (PolII) and histone marks associated with active transcription, including H3K27 acetylation and H3K4me3 (Morales Torres et al., 2013; Wang et al., 2012).

It became clear from the *Kdm6a* knockout studies that the demethylation of H3K27me3 is not the sole function of KDM6A during mesodermal development. *Kdm6a* is found on the X chromosome and escapes X inactivation. *Uty*, a paralog of *Kdm6a,* is found on the Y chromosome and has 88% sequence homology to the KDM6A protein. However, unlike KDM6A, UTY lacks detectable histone demethylase activity *in vitro* (Hong et al., 2007; Lan et al., 2007). When *Kdm6a* knockout mice were generated, many groups were able to show that male hemizygous $Utx^{\Delta/Y}$ embryos escaped embryonic lethality and were able to live to adulthood, albeit with a lower frequency than expected and reduced life expectancy (Shpargel et al., 2012; Wang et al., 2012; Welstead et al., 2012). This led to the hypothesis that *Kdm6a* performs demethylase-independent functions during embryonic development that can be rescued by the catalytically inactive UTY. This was further proven in *in vitro* EB formation, where the introduction of a catalytically dead form of *Kdm6a* is able to rescue mesoderm formation and expression of mesodermal genes such as *T* and *Flk1*. Furthermore, the recruitment of PolII and active histone modifications at the *T* locus were also rescued in the presence of a catalytically dead form of *Kdm6a* (Morales Torres et al., 2013; Wang et al., 2012). This also holds true *in vivo*, where introduction of a catalytically dead version of the *C. elegans utx-1* into the *utx-1* knockout background rescues defects in embryogenesis, tail, and gonadal development (Vandamme et al., 2012). Additionally, *utx-1* functions in the same pathway as the *C. elegans* MLL3/4 H3K4me3 methyltransferase ortholog (SET16)

and *set-16* mutants phenocopy *utx-1* mutants. However, concurrent knockdown of *utx-1* and *set-16* did not lead to exacerbated phenotypes, suggesting that these two genes work together in the same pathway. The authors were also able to verify previous data which shows that KDM6A protein is found in a complex with H3K4me3 methyltransferase complex components MLL3/4, WDR5, ASH2L, RBBP5, and PTIP (Issaeva et al., 2007; Lee et al., 2007b).

The coupling of removal of the repressive H3K27me3 mark and the deposition of activating H3K4 methylation by having KDM6A and MLL3/4 in the same complex appears to be an elegant mechanism used by the cell to activate transcription of target genes during differentiation. In addition, the KDM6A/MLL3/4 complex may also play a role in activating enhancer regions during embryogenesis. It was previously shown that the H3K4me1 histone modification is able to mark cell-type specific enhancer regions (Heintzman and Ren, 2009; Visel et al., 2009; Xi et al., 2007). In 2010, Creyghton and colleagues (Creyghton et al., 2010) were able to show that H3K4me1 is associated with poised enhancers and that regions co-occupied by H3K4me1 and H3K27Ac, or with H3K27Ac alone, are active enhancer regions that correlate with higher gene expression in a cell-type specific manner. Studies in *Drosophila melanogaster* proved that the ortholog of *Mll3/4*, *trr*, is necessary for H3K4 monomethylation *in vivo* and that knockdown of *trr* in imaginal wing discs results in a decrease in both H3K4me1 and H3K27Ac (Herz et al., 2012). The same group went on to show that this was conserved in human cells since knockout of both *Mll3* and *Mll4* resulted in a global decrease in H3K4me1, which was more pronounced in predicted enhancer regions (Hu et al., 2013). They also demonstrated that MLL4 was bound to enhancer regions in the genome, and displayed co-occupancy with p300, H3K4me1 and H3K27Ac, which is a known signature for active enhancers. KDM6A may also play a role in activating enhancers within the MLL3/4 complex, as *Drosophila* UTX has been shown to bind directly to CBP, the histone acetyltransferase responsible for acetylating H3K27 (Tie et al., 2012). Taken together, the MLL3/4 complex, in cooperation with KDM6A and p300/CBP, may function to activate enhancers throughout the genome wherein MLL3/4 is able to monomethylate H3K4, KDM6A can remove H3K27me3, and its interaction with CBP can facilitate acetylation of H3K27.

Unlike KDM6A, the role of KDM6B (JMJD3) in the development of mesodermal-derived tissues is less clear. Two studies have shown an involvement of KDM6B in mesoderm development. First, Dahle and colleagues (Dahle et al., 2010) demonstrated that *Nodal* signaling recruits SMAD2/3 and KDM6B to the *Brachyury* locus in ESCs to remove repressive H3K27me3 and activate transcription in response to activated Wnt signaling. Second, Ohtani and colleagues (Ohtani et al., 2013) demonstrated that *Kdm6b*$^{-/-}$ ESCs are impaired both in their ability to differentiate into *Brachyury*$^{+}$ mesodermal cells, and to give rise to cardiac cells *in vitro*. Furthermore, they showed that KDM6B and β-CATENIN proteins physically interact and that β-CATENIN-mediated activation of *Brachyury* is impaired in *Kdm6b*$^{-/-}$ ESCs.

Insights into the mechanism of KDM6B action came from studies performed in differentiating macrophages. De Santa and colleagues (De Santa et al., 2009) demonstrated that the majority of KDM6B binding in the genome occurs at transcription start sites. Furthermore, immunoprecipitation of KDM6B followed by mass spectrometry revealed that KDM6B binds to factors previously shown to be involved in transcriptional elongation, including SPT6, SPT16, CHD7, ISWI, and KIAA1718 (Chen et al., 2012). Knockdown of *Kdm6b* using shRNAs resulted in an increase of H3K27me3 at target genes with a decrease in elongating PolII across the body of target genes. Interestingly, like KDM6A, KDM6B was also shown to be part of the Trithorax/MLL complex through its interaction with RBBP5, WDR5, and ASH2L (De Santa et al., 2007). However, since these components are found in all MLL complexes, it is not clear in which specific MLL complexes KDM6B belongs. Furthermore, it is unclear if KDM6A and KDM6B are interchangeable within these complexes or for which functions they may compensate for one another.

1.7 Does T Exert Its Function as a Master Regulator through Modification of the Chromatin Environment in Early Mesodermal Cells?

Previous data from *Xenopus* and mouse *in vitro* differentiated primitive streak cells revealed that T binds to and directly regulates the expression of multiple genes involved in both mesodermal and neural cell differentiation and development, most likely through proximal promoter binding and recruitment of transcriptional machinery. However, a very small portion of T-bound genomic regions are associated with genes that are dysregulated upon *T* knockout or knockdown, suggesting that T may perform other functions in addition to its role as a transcriptional activator. Experiments from Miller and colleagues revealed that T-box factors have the ability to bind to histone H3K27 demethylase activity, H3K4 methyltransferase complex members, and components of the SWI/SNF chromatin remodeling complex. Our hypothesis is that T functions to modulate the chromatin environment of early mesodermal cells and prime these cells for further mesodermal differentiation and development through interactions with chromatin- and histone-modifying protein complexes. Perhaps T is able to create a permissive (H3K4me high, H3K27me low) histone environment at promoter and distal regulatory regions through its interactions with KDM6A and the MLL3/4 complex (Figure 1.5A), or promote chromatin remodeling at target genes through interaction with KDM6A and BRG1 (Figure 1.5B). Furthermore, since the *Drosophila* ortholog of KDM6A has been shown to bind to p300/CBP, perhaps T can also regulate histone acetylation at target regions through its interaction with KDM6A (Figure 1.5C). The aim of this thesis was to abrogate the interaction between T and these chromatin- and histone-modifying complexes *in vivo,* through introduction of a point mutation (Y88A) at a residue shown to be important for T-box interaction with KDMs and histone methyltransferase complexes. Analysis of T^{Y88A} mutant embryos and *in vitro* differentiated mesodermal cells was performed to provide insight into the function of T interaction with these chromatin-modifying complexes and to determine which functions of T were carried out by modulation of the chromatin environment.

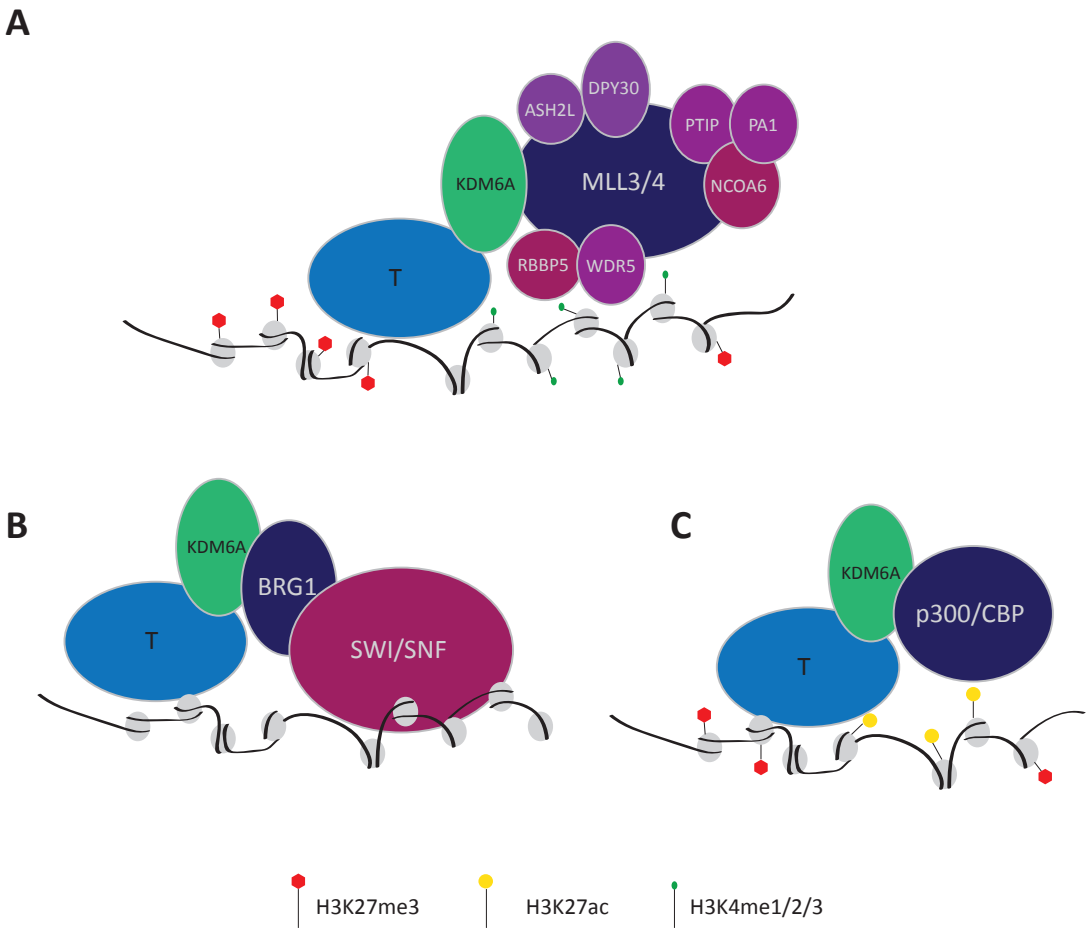


Figure 1.5: Potential T modulation of the chromatin environment through interactions with histone- and chromatin-modifying complexes
(A) T binding to KDM6A and the MLL3/4 complex may promote a permissive (H3K4me high, H3K27me low) environment at proximal and distal regulatory elements throughout the genome. (B) T binding to the SWI/SNF complex, dependent on its interaction with KDM6A, may recruit chromatin remodeling complexes to T target regions throughout the genome. (C) Potential T binding to p300/CBP, mediated by interaction with KDM6A, may modify H3K27 acetylation at T target regions in the genome.

Chapter 2.

Results

2.1 Generation of the T^{Y88A} Point Mutation *In Vivo*

In order to generate a Y88A amino acid substitution in T, site-directed mutagenesis was used to generate a homologous recombination construct mutating the TAC codon coding for tyrosine (Y) to a GCC codon coding for alanine (A). This construct was recombined into the endogenous *T* locus of mouse ESCs, to allow for investigation of this mutation *in vivo*. The homologous recombination targeting construct (Figure 2.1A) contained the mutation within the second exon of the *T* gene, surrounded by 1.3kb and 5.7kb of homologous sequence upstream and downstream of the mutation, respectively, and contained a neomycin selection cassette to allow for selection of recombined clones. A heterozygous $T^{+/Y88A}$ ESC line was generated by targeting one of the endogenous *T* alleles in a wild-type F_1(G4) hybrid ESC line (C57Bl6/129S6, George et al., 2007). To generate a T^{Y88A} mutant ESC line, a hybrid F_1(G4) ESC line of almost identical genetic background and heterozygous for the T^{2J} null mutation ($T^{2J/+}$) was used (gift from Dr. Lars Wittler). There was only one remaining wild-type *T* allele available for targeting in the $T^{2J/+}$ ESCs, and once this allele was targeted, the $T^{2J/Y88A}$ ESC could be used directly to generate hemizygous mutant embryos via tetraploid complementation (Gertsenstein, 2011). This circumvented the need to generate mouse lines and breed heterozygous mice to obtain homozygous $T^{Y88A/Y88A}$ mutant embryos. Detecting the copy number of the *T* allele in the recipient $T^{2J/+}$ cell line confirmed that only one allele was present and tetraploid complementation of clone #10 resulted in embryos with a truncated tail at E12.5, indicative of *T* heterozygosity (Figure S1).

Successful recombination of the targeting construct into ESCs was verified by Southern blot (Figure 2.1B, 2.1C). After verification, removal of the neomycin selection cassette was necessary to avoid any secondary effects due to the insertion of the 2kb selection cassette within the intron of *T*, as it is known that some introns can harbor regulatory elements controlling expression of neighboring genes. To this end, Flp/*FRT* site-directed recombination was used as the neomycin cassette was flanked by *FRT* sites. Removal of the neomycin cassette was verified by loss of G418 resistance and Southern blotting (Figure 2.2).

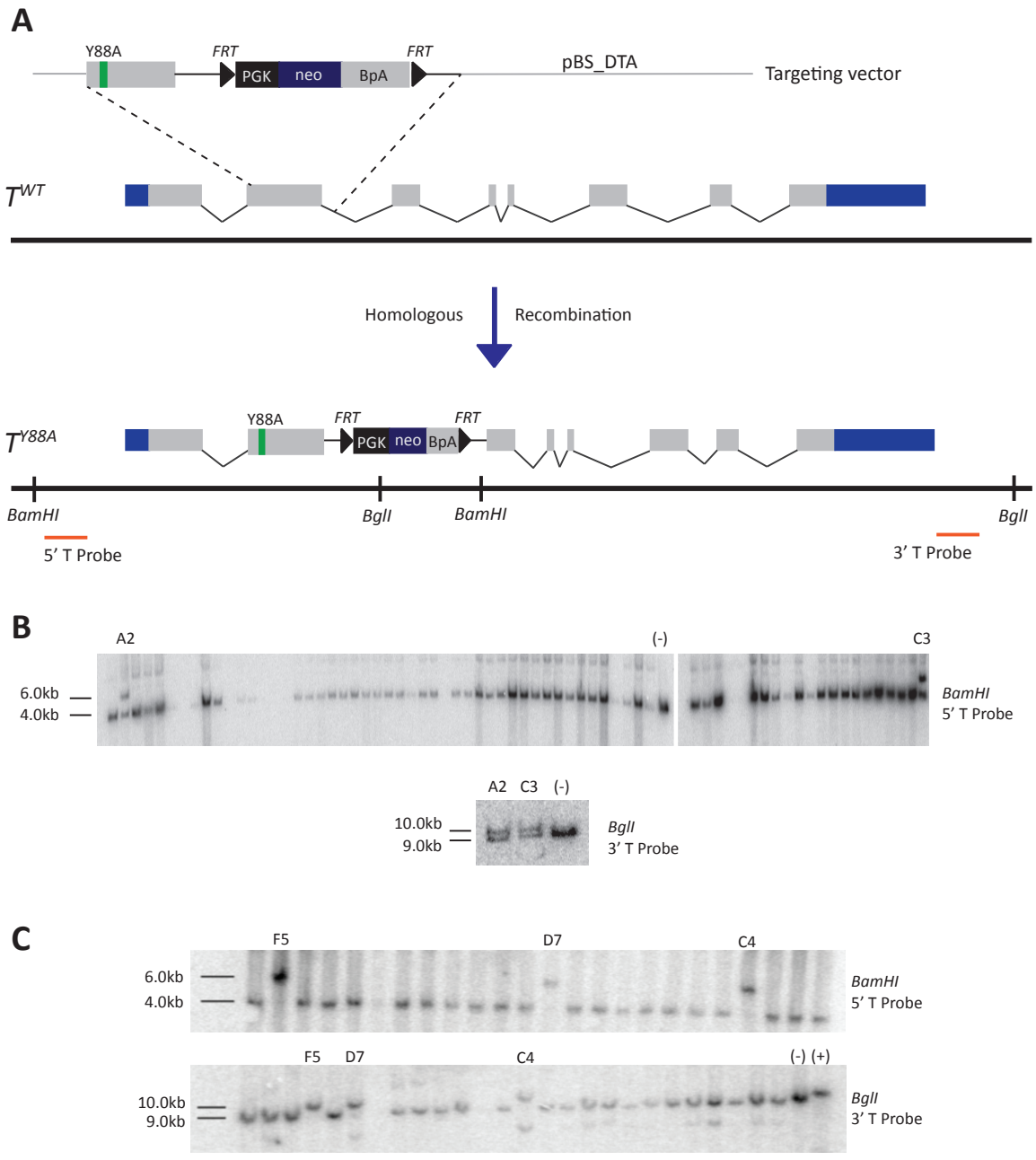
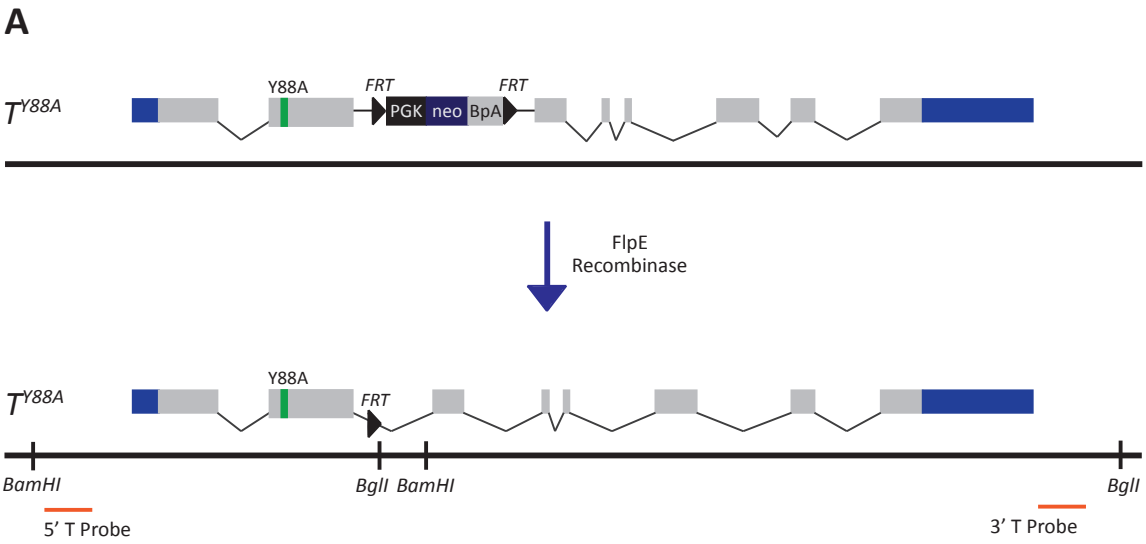
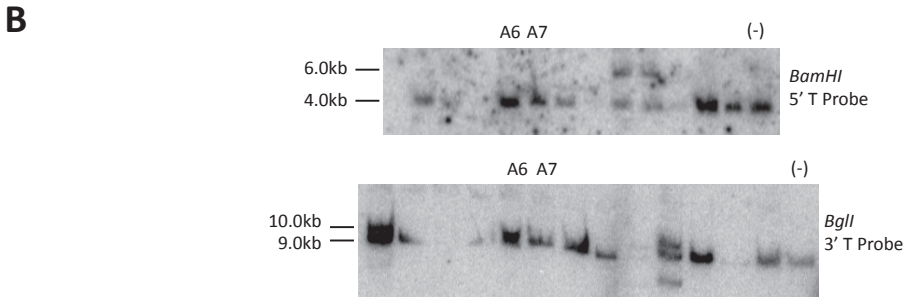


Figure 2.1: Introduction of the Y88A point mutation into the endogenous *T* locus
(A) Schematic illustration of the homologous recombination strategy. *T* exons are shown in gray, the UTR regions in blue, and Y88A point mutation is indicated in green. Red lines illustrate probes used for Southern blot analysis. (B) Southern blot of $T^{+/Y88A}$ ESCs using the indicated 5' or 3' T probes. Genomic DNA from ESC clones was digested with *BamHI* and *BglI* restriction enzymes to probe with the 5' and 3' T probes, respectively. Correct targeting of the point mutation cassette resulted in a band at 6.0kb, in addition to the wild-type band at 4.0kb, using the 5' T probe. Correct targeting resulted in a 10.0kb band, in addition to the 9.0kb wild-type band, using the 3' T probe. Clones A2 and C3 were heterozygous for the T^{Y88A} mutation. (C) Southern blot of $T^{2J/Y88A}$ ESCs, using the same probes as in (B). Correct targeting resulted in one single band at 6.0kb using the 5' T probe and one single band at 10.0kb using the 3' T probe. Clone F5 contained the correctly targeted allele ($T^{2J/Y88A}$). Genomic DNA from wild-type F1 ESCs was used as a negative control (-). DNA from the homologously recombined T^{Y88A} BAC was used as a positive control (+).





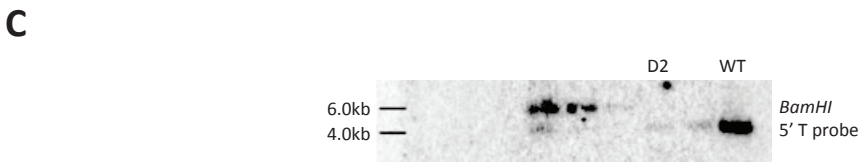


Figure 2.2: Removal of the neomycin selection cassette from $T^{+/Y88A}$ and $T^{2J/Y88A}$ ESCs
(A) Schematic illustrating the mutant T^{Y88A} locus after removal of the neomycin cassette by Flp-mediated recombination. (B) Southern blot of $T^{+/Y88A}$ ESCs after FlpE-mediated recombination. Removal of the neomycin selection cassette resulted in a 4.0kb band, using the 5' T probe after restriction digest of genomic DNA with *BamHI*. In the lower panel, removal of the selection cassette resulted in a 9.0kb band, using the 3' T probe after restriction digest of genomic DNA with *BglI*. Genomic DNA from wild-type F_1 cells was used as a negative control (-). The cassette was successfully removed from clones A6 and A7. (C) Southern blot of $T^{2J/Y88A}$ ESCs after FlpE-mediated recombination. Removal of the selection cassette resulted in a 4.0kb band, using the 5' T probe after restriction digest of genomic DNA with *BamHI*. Genomic DNA from wild-type F_1 cells was used as a negative control (WT). The cassette was successfully removed from clone D2.

The experiments described above resulted in the generation of two ESC lines with a mutation in the endogenous *T* locus creating a Y88A amino acid substitution, $T^{+/Y88A}$ heterozygous and $T^{2J/Y88A}$ mutant ESCs. These ESCs were used to generate embryos via tetraploid complementation and investigate the functional consequences of loss of T interaction with histone-modifying proteins *in vivo*. Embryos harboring the $T^{2J/Y88A}$ mutation resembled *T* knockout embryos, displaying a complete truncation of the posterior axis, loss of somite formation, a kinked neural tube, and loss of the notochord (Figure 2.3A). $T^{2J/Y88A}$ mutant embryos ultimately died between E9.5-10.0 due to defects in allantois formation, an embryonic structure which forms a bridge between the embryo and the maternal placenta and is the precursor to the umbilical cord and associated blood vessels (Inman and Downs, 2007). This result suggested that the functional interaction between T and histone-modifying enzymes is indeed necessary for its normal function during mesoderm development. Furthermore, heterozygous $T^{+/Y88A}$ embryos resembled $T^{+/-}$ heterozygotes; at E9.5, embryos appeared phenotypically normal with no gross morphological defects visible, while at E12.5, the tail was shortened and kinked (Figure 2.3B).

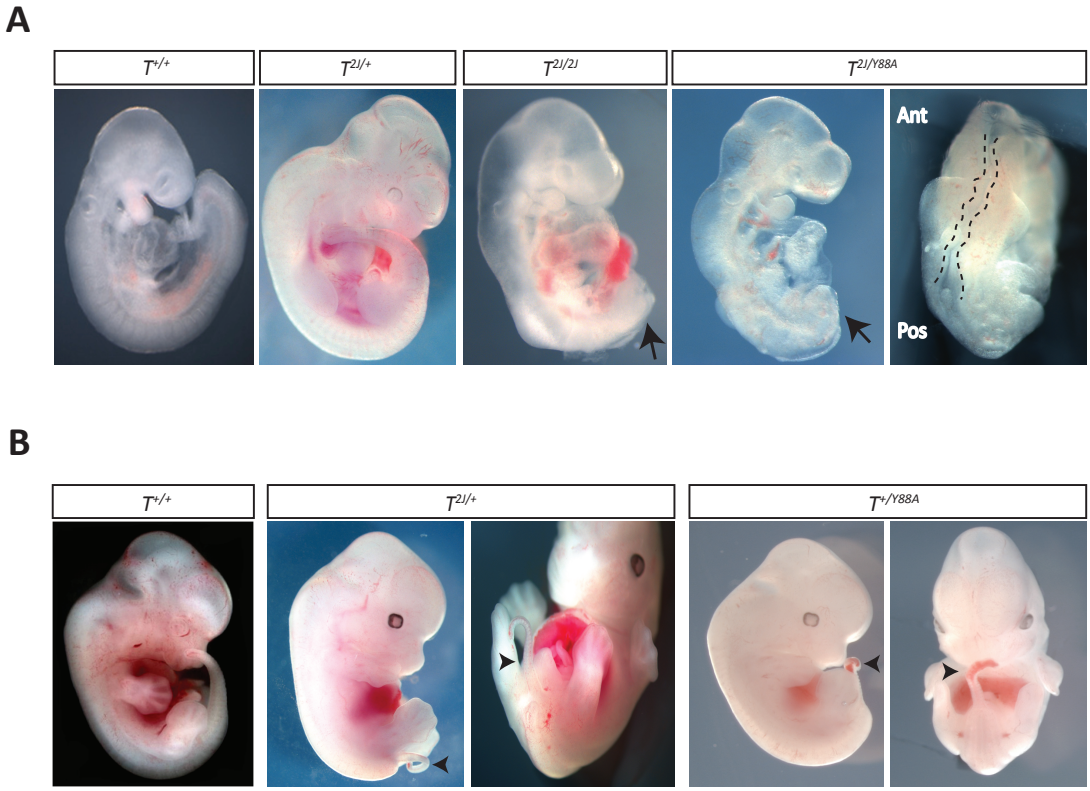


Figure 2.3: T^{Y88A} heterozygous and mutant embryo phenotypes
(A) $T^{2J/Y88A}$ mutant embryos at E9.5 displayed a complete truncation of posterior mesoderm, as indicated by the black arrow, and resembled the $T^{2J/2J}$ knockout embryo. A dorsal view of the $T^{2J/Y88A}$ mutant embryo (farthest panel to the right) shows the kinked neural tube, outlined with a dashed black line. The left panels depict a wild-type $T^{+/+}$ embryo at E9.5 and a heterozygous $T^{2J/+}$ embryo at E10.5, displaying normal axis elongation and posterior mesoderm formation. (B) $T^{+/Y88A}$ embryos at E12.5 are morphologically normal with the exception of a shortened, kinked tail (black arrowhead), as in the $T^{2J/+}$ heterozygous knockout embryos. Ant, anterior; Pos, posterior.

In order to determine whether there were more subtle defects in the $T^{+/Y88A}$ heterozygotes that were not detectable by examining gross morphology alone, WISH was performed to visualize structures whose proper formation is dependent on *T*. For example, *Uncx* is expressed in the caudal end of newly formed somites and is a marker to visualize somite formation (Mansouri et al., 1997). Expression of *Uncx* is normal in $T^{+/Y88A}$ embryos at E9.5 and formation of the somites appeared unaffected when compared to control embryos at this stage (Figure 2.4A). *T* has also previously been shown to be essential for proper notochord formation. Notochord formation was visualized by WISH using a probe directed against *Shh*, a gene expressed in the developing notochord at E9.5 (Echelard et al., 1993). *Shh* is also expressed normally in $T^{+/Y88A}$ at E9.5, suggesting that notochord formation occurs properly in heterozygous mutants (Figure 2.4A). Lastly, expression analysis of *T* itself by WISH demonstrated that *T* was expressed normally at E9.5 in the caudal end and notochord of heterozygous $T^{+/Y88A}$ embryos. Sequencing of *T* cDNA from heterozygous embryos confirmed that both wild-type *T* and T^{Y88A} were expressed, as evidenced by the presence of double peaks in the sequencing chromatogram (Figure 2.4B). Altogether, the results described above suggest that the T^{Y88A} mutant allele acts as a null. When it is expressed alone in the $T^{2J/Y88A}$ mutant, embryos phenocopy *T* knockout embryos, while heterozygous $T^{+/Y88A}$ embryos resemble $T^{+/-}$ heterozygotes. Therefore, it appears as if the functional interaction of T with histone-modifying enzymes is essential for T function; when this interaction is abrogated, T can no longer regulate mesoderm specification and development.

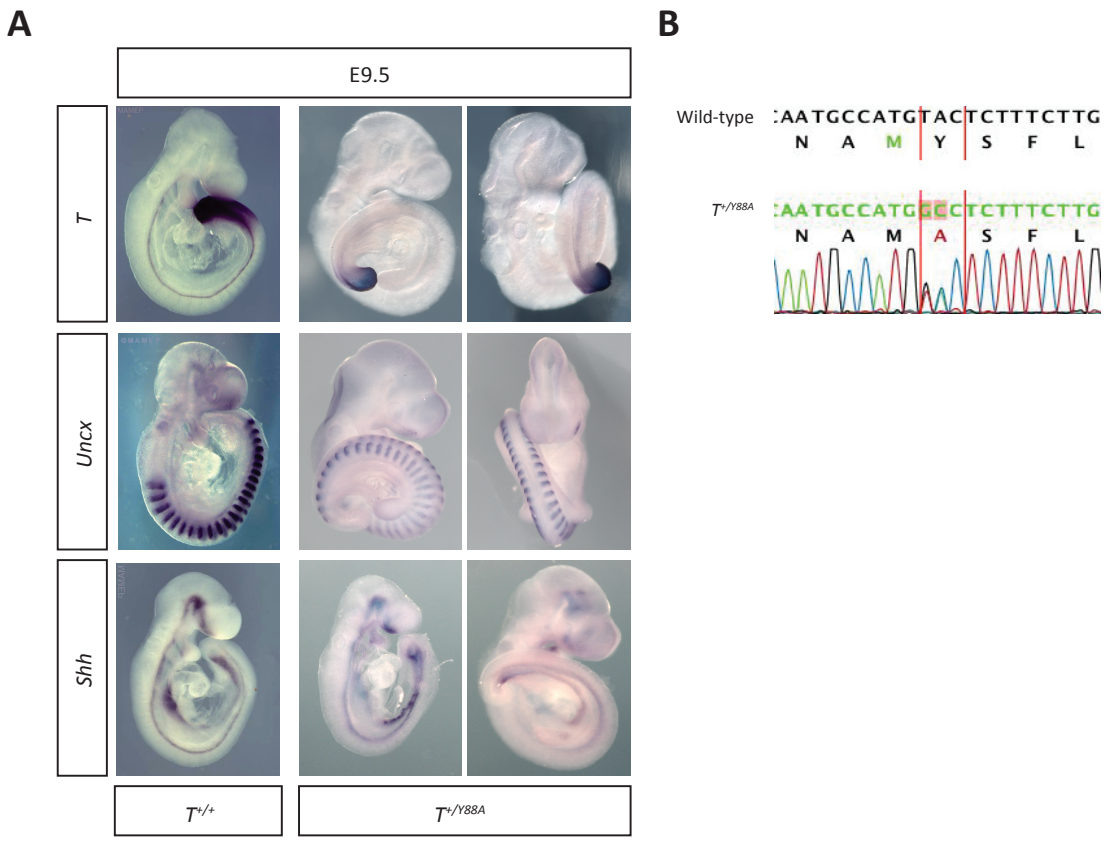


Figure 2.4: $T^{+/Y88A}$ embryos at E9.5 morphologically resemble wild-type embryos
(A) Whole-mount *in situ* hybridization with a *T* probe demonstrated that *T* expression in the tail bud and notochord was normal in the heterozygous $T^{+/Y88A}$ when compared to wild-type $T^{+/+}$. *Uncx* staining of developing somites illustrated that somite formation was normal in the heterozygous $T^{+/Y88A}$. Furthermore, *Shh* staining of the notochord demonstrated that notochord formation in the heterozygous $T^{+/Y88A}$ resembled that of wild-type. Wild-type WISH images were taken from the MAMEP database (Max Planck Institute for Molecular Genetics; Berlin, Germany; http://mamep.molgen.mpg.de). (B) Sequencing of exon 2 of *T* from embryonic cDNA generated from $T^{+/Y88A}$ heterozygotes revealed the expression from both wild-type *T* and T^{Y88A} alleles.

2.2 Rescue of $T^{2J/Y88A}$ Mutation by Transgene Integration of Wild-type *T*

In order to confirm that the Y88A point mutation in T was the cause of the knockout phenotype seen in embryos, a genetic rescue experiment was performed by adding a wild-type copy of *T* back into the genome of $T^{2J/Y88A}$ mutant embryos. In theory, if the mutant phenotype is caused by the Y88A point mutation alone, and not by other secondary effects, adding a wild-type copy of *T* should rescue the phenotype and embryos should appear phenotypically normal. To this end, a bacterial artificial chromosome (BAC) containing the *T* gene and approximately 200kb of total sequence upstream and downstream of *T* was used (RP24-530D23). A neomycin selection cassette was previously integrated into the backbone of the BAC in order to allow for selection of ESC clones (Figure 2.5A, a gift from Dr. Tracie Pennimpede). The *T* BAC was randomly integrated into the genome of $T^{2J/Y88A}$ ESC and positive clones were selected by resistance to G418. Southern blotting using a probe directed against the *neo* resistance gene was used to confirm integration of the wild-type *T* BAC (Figure 2.5B). Clones were further screened by qPCR of the *T* promoter region using ESC genomic DNA to determine the number of BACs that were randomly integrated into the genome. Two primer pairs were chosen, recognizing 5' and 3' of the T gene. Clone E1 contained one integration of the wild-type *T* BAC, as evidenced by the presence of two *T* alleles when compared to heterozygous and wild-type controls; endogenous T^{Y88A} and the wild-type *T* allele from the BAC (Figure 2.5C).

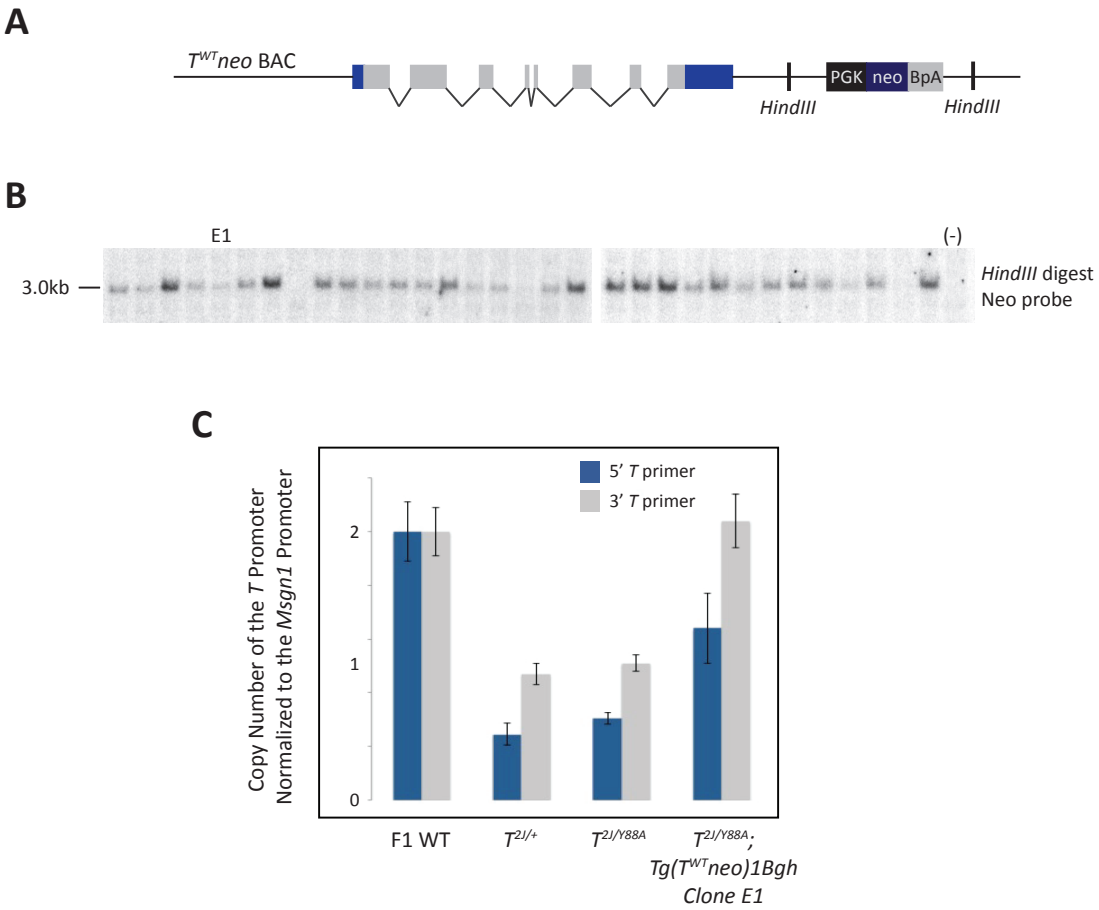


Figure 2.5: BAC integration to rescue the $T^{2J/Y88A}$ mutant phenotype
(A) Schematic illustration of the wild-type *T* BAC integrated into mutant $T^{2J/Y88A}$ ESCs. (B) Southern blot analysis of genomic ESC DNA using a probe recognizing the neomycin resistance gene to verify integration of the wild-type *T* BAC. ESC clones that have integrated the BAC displayed a 3.0kb band when genomic DNA is digested with *HindIII*. Genomic DNA from wild-type F_1 cells was used as a negative control (-). (C) Quantitative real-time PCR using genomic DNA from indicated ESC lines. Primers recognizing regions 5' and 3' of the *T* gene were used. Error bars represent standard error of the mean of technical replicates.

ESC clone E1 was used to generate embryos by tetraploid complementation in order to assess whether adding a wild-type copy of the *T* gene is able to rescue the mutant $T^{2J/Y88A}$ phenotype. $T^{2J/Y88A}$*;Tg(*T^{WT}*neo)1Bgh* embryos appeared phenotypically wild-type, with rescue of the posterior truncation phenotype, somite, and notochord formation (Figure 2.6A). Furthermore, sequencing of cDNA in these embryos demonstrated that both T^{Y88A} and wild-type *T* were expressed (Figure 2.6D). qRT-PCR analysis of *T* expression in embryos revealed that *T* expression was even higher than in $T^{2J/+}$ control heterozygotes (Figure 2.6B), although this can be explained by random integration of the transgene into the ESC genome. Perhaps the transgene was integrated into a more transcriptionally active region of chromatin, leading to higher expression of the transgene compared to the endogenous allele. Lastly, qRT-PCR analysis demonstrates that expression of the downstream target gene *Tbx6*, which was down-regulated in the $T^{2J/Y88A}$ mutant, is restored in transgenic rescue embryos, illustrating that rescue occurred both on the morphological and molecular level (Figure 2.6C).

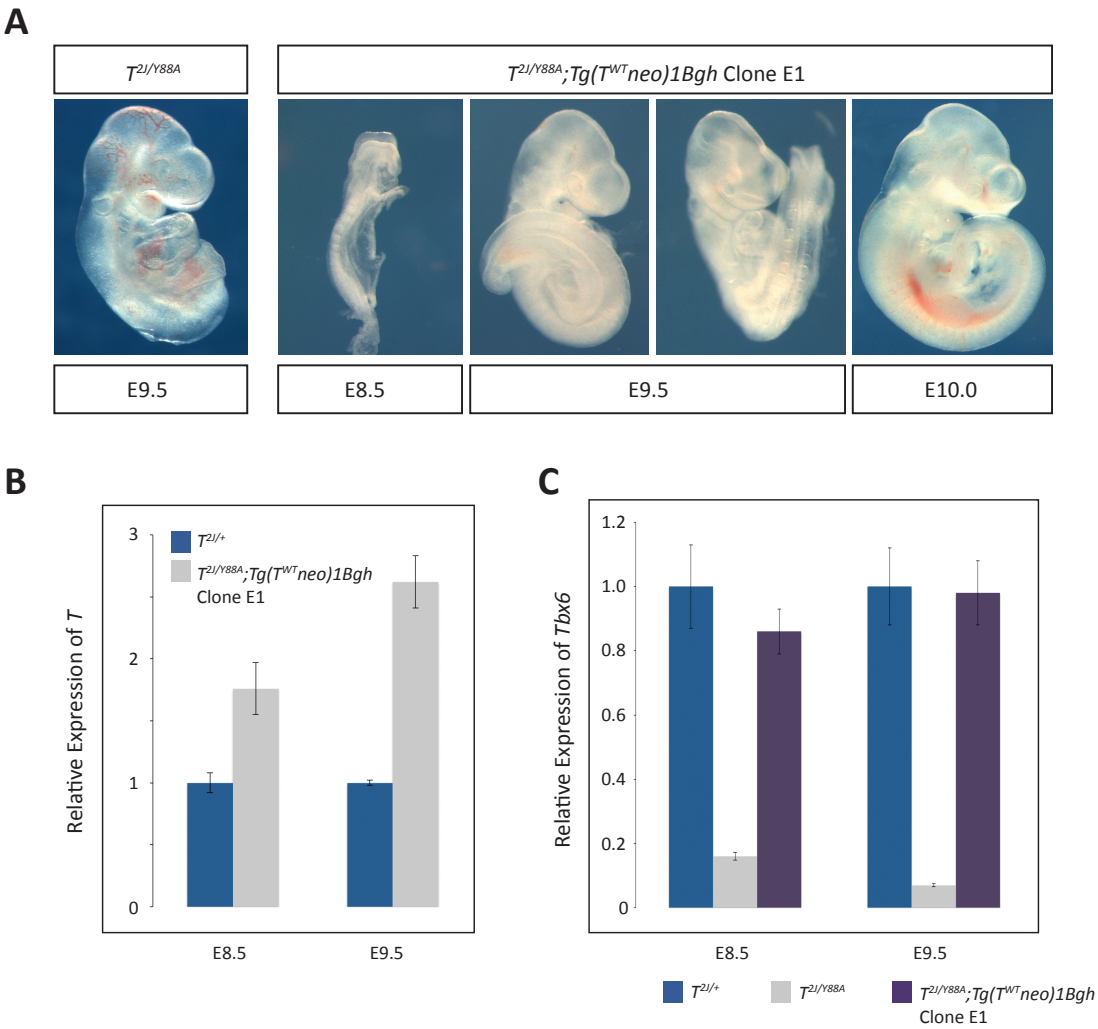


Figure 2.6: Rescue of the $T^{2J/Y88A}$ mutant phenotype by integration of a wild-type copy of *T*
(A) Embryos generated by tetraploid complementation of $T^{2J/Y88A}$;*Tg*(T^{WT}*neo)1Bgh* ESCs displayed complete rescue of the mutant $T^{2J/Y88A}$ phenotype. (B) qRT-PCR was used to measure expression levels of *T* in the $T^{2J/Y88A}$;*Tg*(T^{WT}*neo)1Bgh* embryos. Expression of *T* at E8.5-E9.5 was almost double that of the heterozygous $T^{2J/+}$ control. Error bars represent standard error of the mean in technical replicates. (C) Expression of *Tbx6*, a known *T* target gene which is downregulated in $T^{2J/Y88A}$ mutant embryos, was also restored in the rescue $T^{2J/Y88A}$;*Tg*(T^{WT}*neo)1Bgh* embryos as assayed by qRT-PCR. Error bars represent standard error of the mean in technical replicates. (D) $T^{2J/Y88A}$;*Tg*(T^{WT}*neo)1Bgh* rescue embryos express both wild-type and mutant Y88A *T* as evidenced by a double peak in the sequencing chromatogram of exon 2 of *T* from embryonic cDNA.

2.3 Confirmation of T^{Y88A} as a Null Allele

Since the Y88A amino acid substitution is directly next to a residue shown to be important for the dimerization of T (M87), I wanted to determine if there were dominant-negative effects of the T^{Y88A} allele that may only be visible when multiple copies of the mutant allele are present. If the T^{Y88A} allele was simply a null allele, then the addition of further T^{Y88A} alleles should not affect mesodermal formation and embryos should resemble $T^{+/Y88A}$ heterozygous embryos, as described in the previous section. In order to gain insight into dimerization defects from the T^{Y88A} allele, multiple copies of the T^{Y88A} allele were integrated into the genome of $T^{+/Y88A}$ ESCs using a BAC containing the mutant T^{Y88A} gene and 200kb of upstream and downstream genomic DNA sequence (Figure 2.7A). Positive clones were selected for integration of the T^{Y88A} BAC by resistance to G418, as the BAC contained a neomycin resistance cassette within its backbone. PCR genotyping using BAC specific primers was performed in order to verify that the BAC was integrated into the genome of $T^{+/Y88A}$ ESCs (Figure 2.7B).

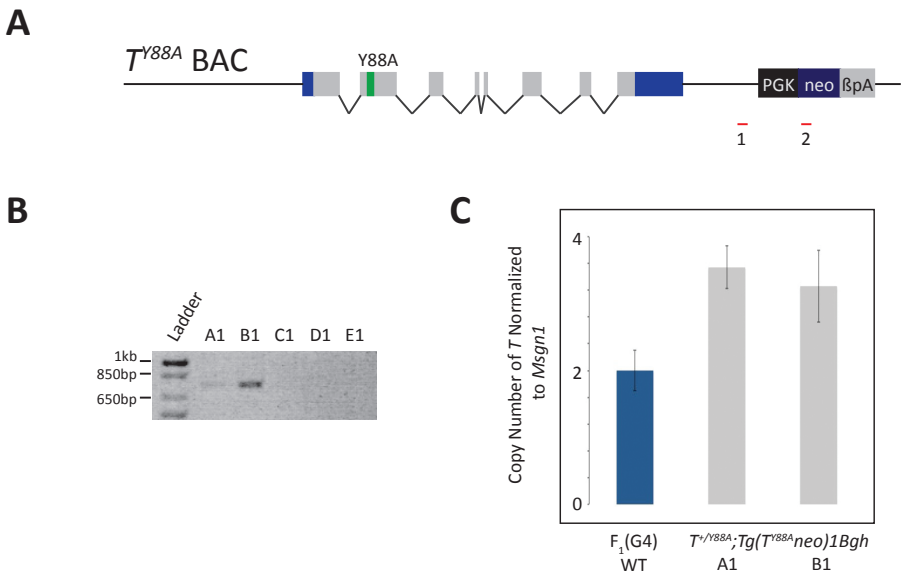


Figure 2.7: Generation of $T^{+/Y88A}$;$Tg(T^{Y88A}neo)1Bgh$ ESCs
(A) Schematic of the T^{Y88A} BAC randomly integrated into the genome of $T^{+/Y88A}$ ESCs. Red lines indicate primers used for PCR genotyping. (B) PCR genotyping of ESC genomic DNA. Clones A1 and B1 displayed an 800bp band indicating successful integration of the T^{Y88A} BAC. (C) qPCR of genomic DNA from ESC clones A1 and B1 demonstrates that both clones contain 2 integrations of the T^{Y88A} BAC when normalized to the *Msgn1* control. Error bars represent standard error of the mean from technical replicates. PGK, phosphoglycerate kinase promoter; neo, neomycin resistance gene; βpA, beta-globin poly-A; WT, wild-type.

Quantitative PCR, with primers recognizing the 3' downstream region of the *T* gene, was used to quantify the number of integrations of the T^{Y88A} BAC into the genome of $T^{+/Y88A}$ ESCs (Figure 2.7C). Clone A1 contained 2 integrations of the T^{Y88A} BAC, for a total of 3 copies of the T^{Y88A} allele and one wild-type *T* allele. $T^{+/Y88A}$*;Tg(*T^{Y88A}*neo)1Bgh* ESCs were used to generate embryos via tetraploid complementation. Embryos were analyzed at E9.5 and E10.5 for *T* expression and for somite and notochord formation. Expression of *T* by WISH was normal in $T^{+/Y88A}$*;Tg(*T^{Y88A}*neo)1Bgh* embryos at E9.5 and E10.5 (Figure 2.8A). However, qRT-PCR demonstrated that the level of *T* expression in E9.5 embryos was 17% of wild-type embryos, compared to 50% of wild-type in $T^{+/Y88A}$ heterozygous embryos (Figure 2.8C). Despite this significant reduction in *T* expression, somite formation and notochord formation appeared normal by WISH of marker genes such as *Uncx* and *Shh* (Figure 2.8A). Interestingly, although embryos at E9.5 and E10.5 appeared normal, at E12.5 64% (16/25) of embryos were necrotic with yolk sacs filled with blood, suggesting that the reduction of *T* seen at E9.5 was embryonic lethal for a large proportion of $T^{+/Y88A}$*;Tg(*T^{Y88A}*neo)1Bgh* embryos (Figure 2.8B), compared to 0% lethality in $T^{+/Y88A}$ heterozygous embryos. The remaining 9 embryos were normal in appearance with the exception of a shortened, kinked tail, much like the $T^{+/Y88A}$ embryos at E12.5. From the data above, it appears as if the T^{Y88A} allele has some semi-dominant-negative effects when added in multiple copies; *T* expression is reduced to 17% of wild-type levels and embryonic lethality occurs in 64% of embryos by E12.5. However, there is most likely not a defect in dimerization, or, less plausibly, dimerization is affected but is not crucial for *T* function, as the phenotypes in $T^{+/Y88A}$*;Tg(*T^{Y88A}*neo)1Bgh* embryos are less severe than would be expected from other dominant-negative *T* alleles.

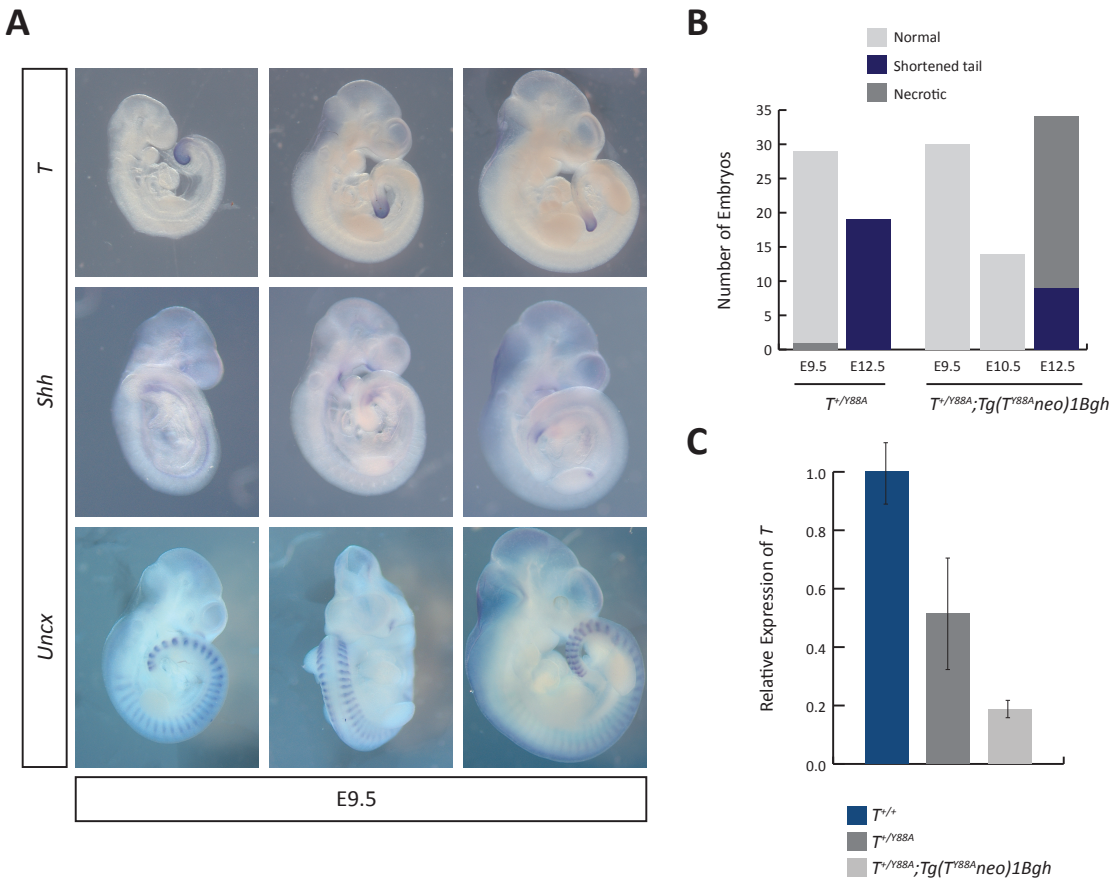


Figure 2.8: Phenotype of $T^{+/Y88A}$;Tg(T^{Y88A}neo)1Bgh embryos
(A) WISH for *T, Shh,* and *Uncx* in $T^{+/Y88A}$;Tg(T^{Y88A}neo)1Bgh embryos from E9.5 illustrates that notochord and somite formation were grossly unaffected. (B) Phenotypes of $T^{+/Y88A}$;Tg(T^{Y88A}neo)1Bgh embryos compared to $T^{+/Y88A}$ heterozygous embryos from E9.5-E12.5. (C) qRT-PCR of *T* expression in E9.5 embryos revealed a significant reduction in *T* expression in $T^{+/Y88A}$;Tg(T^{Y88A}neo)1Bgh embryos, compared to wild-type and $T^{+/Y88A}$ heterozygous embryos. Error bars depict standard error of the mean from technical replicates.

2.4 The Functional Interaction of T with Histone-Modifying Proteins is Necessary for Maintenance of *T* Expression

As described above, the functional interaction of T with histone-modifying proteins is necessary for proper T function during mesodermal specification and development. We chose to investigate $T^{2J/Y88A}$ mutant embryos further in order to shed light onto the mechanisms why these mutant embryos phenocopy *T* knockout embryos. In order to verify that T^{Y88A} mRNA was expressed in mutant embryos, WISH using probes directed against *T* was performed. *T* expression appeared normal in mutant $T^{2J/Y88A}$ embryos when compared to heterozygous $T^{2J/+}$ controls at earlier stages of embryonic development (E7.0-E7.5) before the mutant phenotype becomes apparent (Figure 2.9A). Sequencing of *T* cDNA, derived from mRNA in mutant $T^{2J/Y88A}$ embryos, confirmed the expression of full-length *T* with no mutations other than the Y88A point mutation introduced (Figure 2.9C and data not shown). Interestingly, the WISH of *T* expression at later stages of embryonic development (E8.0) revealed a decrease in *T* expression in mutant $T^{2J/Y88A}$ embryos when compared to controls. This was further validated using qRT-PCR with *T*-specific primers and cDNA from embryos at specific stages during development from E7.5-E9.5. *T* expression decreases between E7.5-E8.0 of development in mutant $T^{2J/Y88A}$ embryos and by E9.5, is almost completely absent (Figure 2.9B). The results described above suggest that the functional interaction of T with histone-modifying proteins appears to be necessary for maintenance of *T* expression itself.

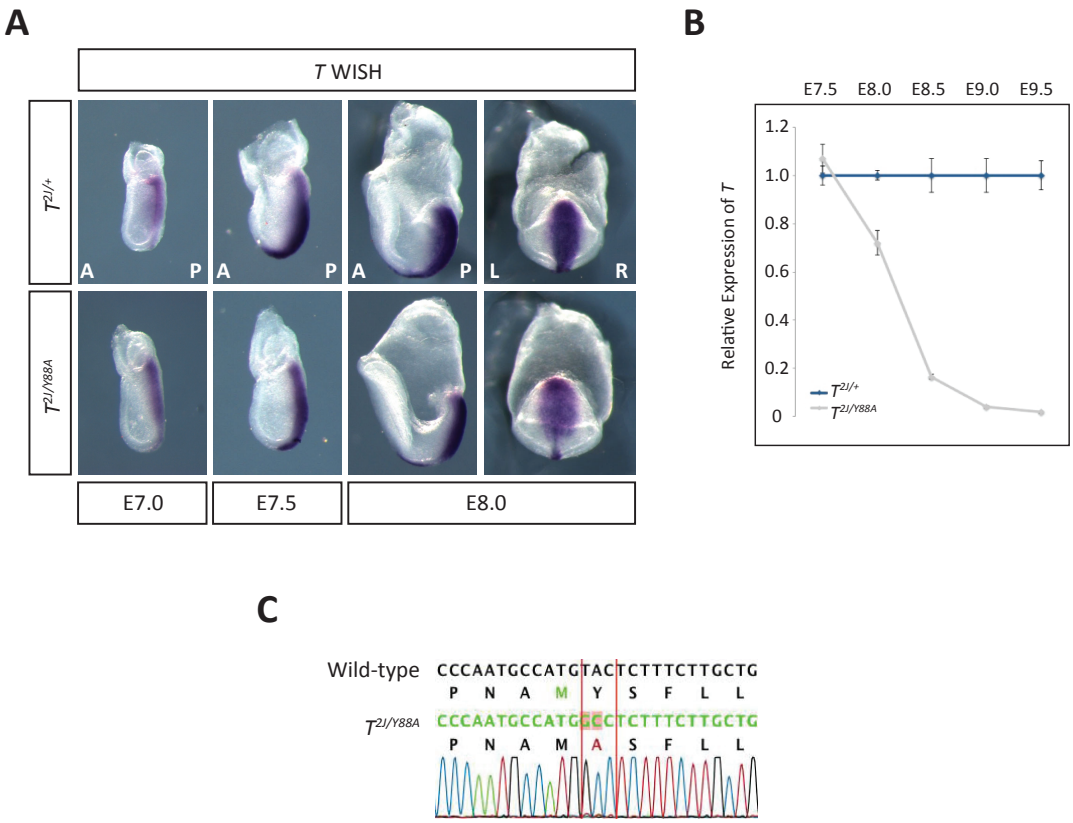


Figure 2.9: Expression of *T* in the $T^{2J/Y88A}$ mutant embryos
(A) WISH in $T^{2J/+}$ compared to $T^{2J/Y88A}$ mutant embryos using a probe detecting *T* transcript. The panels furthest to the right depict *T* expression in embryos viewed dorsally at E8.0. (B) qRT-PCR was used to quantify *T* expression in $T^{2J/+}$ (blue line) compared to $T^{2J/Y88A}$ (gray line) mutant embryos from E7.5-E9.5. Error bars depict standard error of the mean from technical replicates. (C) Sequencing of *T* cDNA from $T^{2J/Y88A}$ mutant embryos indicated that only mutant T^{Y88A} is expressed. Ant, anterior; Pos, posterior; L, left; R, right.

2.5 T^{Y88A} Protein Levels Decrease in $T^{2J/Y88A}$ Embryos During Development

T protein was investigated in $T^{2J/Y88A}$ mutant embryos to determine whether, like mRNA expression, T protein levels were also lower compared to $T^{2J/+}$ control embryos. Embryos were dissected at different embryonic stages from E7.5-E8.5, whole embryo protein lysates were run on an SDS-PAGE gel, and levels of T protein were measured by Western blot. At E7.75, levels of T protein were similar when comparing $T^{2J/Y88A}$ and $T^{2J/+}$ embryos. However, starting at E8.0, levels of T protein decreased in the $T^{2J/Y88A}$ mutant compared to $T^{2J/+}$ control, and by E8.5, T protein was almost completely absent in the mutant compared to control (Figure 2.10).

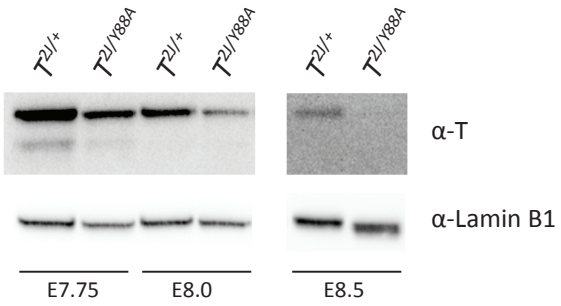


Figure 2.10: Levels of T protein in $T^{2J/Y88A}$ mutant embryos
Levels of T protein were measured by Western blot of whole embryo lysates at indicated time points using an α-T antibody. α-Lamin B1 was used as a loading control.

In order to rule out that T^{Y88A} protein levels are reduced due to inherent instability of the mutant T protein caused by the Y88A point mutation, stability of the wild-type and T^{Y88A} proteins were measured using a cycloheximide (CHX) chase assay. Cycloheximide is a small-molecule compound isolated from *Streptomyces griseus* which inhibits *de novo* eukaryotic protein synthesis by binding to the E site within the ribosome and blocking elongation and translocation of tRNA molecules (Schneider-Poetsch et al., 2010). Wild-type and T^{Y88A} myc-tagged over-expression constructs were generated and transfected into NIH3T3 cells to check for correct tagging of the T protein. Both T^{WT} and T^{Y88A} were properly tagged with the myc tag and could be further used in the CHX chase assay (Figure 2.11A). Wild-type and mutant T-myc constructs were transfected overnight into NIH3T3 cells and the next day CHX was added and whole protein lysates were taken from the cells at time points between 0 and 12 hours. Since no new protein is synthesized in the presence of CHX, monitoring levels of T after the addition of CHX allows one to determine the stability of the wild-type versus T^{Y88A} mutant proteins. In the absence of CHX, levels of both T^{WT} and T^{Y88A} protein were relatively stable, with no decrease in the levels of T protein. However, in the presence of CHX, a decrease in levels of T protein was observed and there was no substantial difference in levels of T^{WT} compared to T^{Y88A} protein (Figure 2.11B, 2.11C). From these experiments, it is concluded that the loss of T^{Y88A} protein observed in mutant embryos is not due to an inherent instability of the mutant T protein.

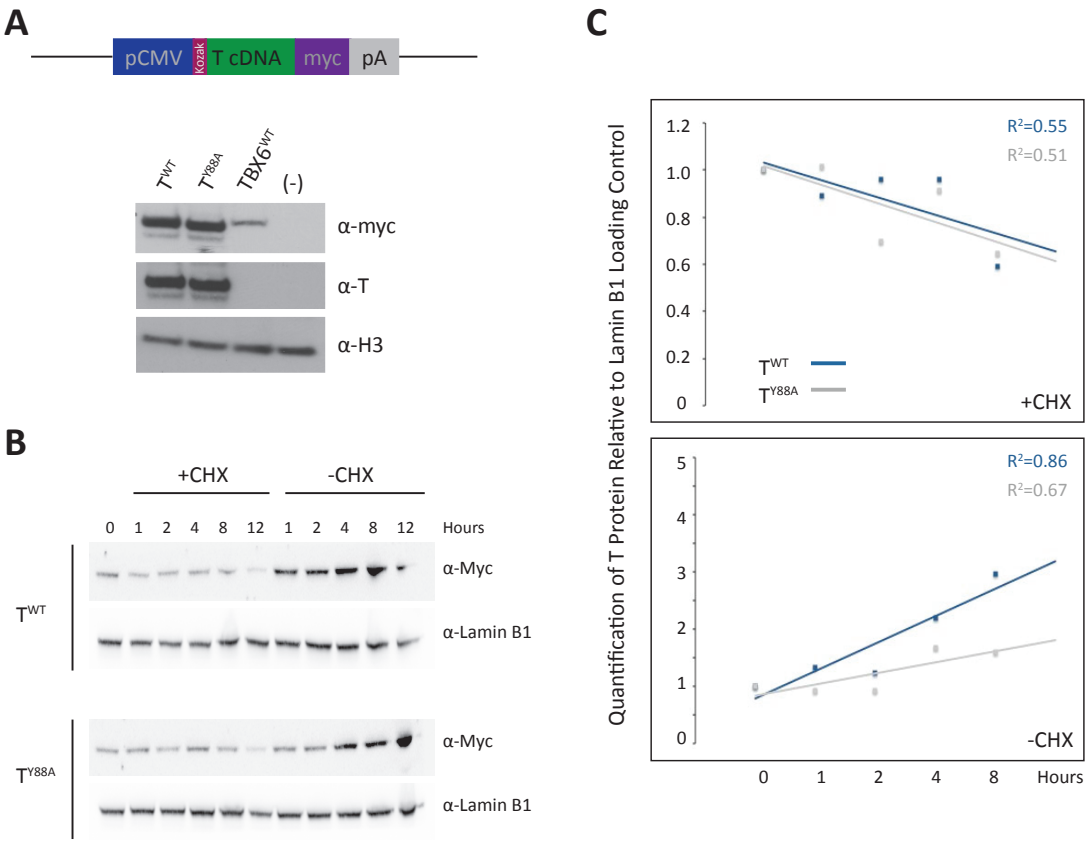


Figure 2.11: Testing the stability of T^{Y88A} protein using a cycloheximide chase assay
(A) Schematic of the plasmid construct transiently transfected into NIH3T3 cells. Western blotting of whole cell lysates with α-myc and α-T antibodies demonstrated that T^{WT} and T^{Y88A} proteins were properly tagged with myc. A construct encoding for wild-type myc-tagged TBX6 was used as a negative control for the α-T antibody. Untransfected NIH3T3 cells were used as a negative control (-) and α-H3 was used as a loading control. (B) NIH3T3 whole cell lysates, transiently transfected with myc-tagged T^{WT} or T^{Y88A} expression constructs and incubated with or without cycloheximide, were subjected to Western blotting with an α-myc antibody at indicated time points. α-Lamin B1 was used as a loading control. (C) Densitometry was used to quantify the amount of T protein relative to the Lamin B1 loading control in the presence and absence of cycloheximide. pCMV, cytomegalovirus promoter; pA, poly-A signal; WT, wild-type; CHX, cycloheximide.

The myc-tagged T constructs were also used to determine localization of T^{WT} and T^{Y88A} protein in cells *in vitro,* as the mislocalization of T^{Y88A} protein could contribute to its loss of proper function in $T^{2J/Y88A}$ mutant embryos. Since T is a transcription factor that regulates gene expression, the expected localization is within the nucleus (Kispert and Herrmann, 1994; Schulte-Merker et al., 1992). To investigate localization of the T^{Y88A} protein compared to wild-type, the C2C12 mouse myoblast cell line was used as these cells are of mesodermal origin, and their nuclei are relatively large and easy to visualize

by microscopy. Cells were transfected with T^{WT} and T^{Y88A} over-expression constructs and T protein was visualized using indirect immunofluorescence. Wild-type T protein was found in the nuclei of C2C12 cells but also localized to the cytoplasm in a large fraction of cells (Figure 2.12). Whether this was an artifact of over-expression of the T protein or a biologically relevant function of T in the cytoplasm of cells remains to be determined. Previous reports have shown that multiple T-box factors contain nuclear export signals and there is some evidence that endogenous T and other T-box factors may localize to the cytoplasm *in vivo* (our unpublished observations, Hong and Hsueh, 2007; Inman and Downs, 2006a; Kulisz and Simon, 2008). Localization of T^{Y88A} mutant protein was also investigated and T^{Y88A} protein was localized both in the nucleus and cytoplasm of C2C12 cells, comparable to wild-type T protein (Figure 2.12). From the results above, it is clear that T protein is reduced in the mutant $T^{2J/Y88A}$ as compared to control, and that this reduction is not due to an inherent instability of the mutant protein. Furthermore, loss of T^{Y88A} protein function, as evidenced by phenocopy of the $T^{2J/2J}$ knockout phenotype, cannot be explained by a mislocalization of the mutant T^{Y88A} protein.

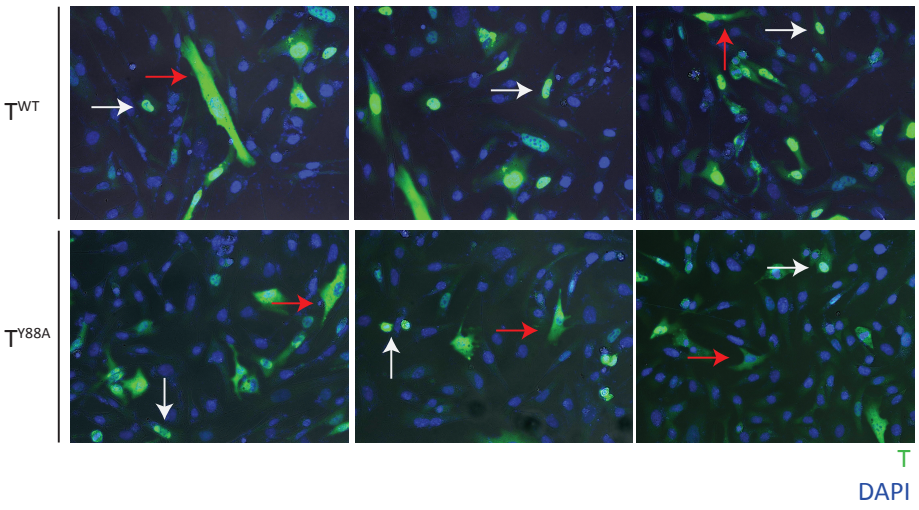


Figure 2.12: Cellular localization of T^{Y88A} mutant protein
C2C12 cells transfected with myc-tagged T^{WT} and T^{Y88A} expression constructs were visualized using an α-myc primary antibody and Alexa-488-coupled secondary antibody (green) and visualized by fluorescence microscopy. Cells were co-stained with DAPI (blue) to visualize nuclei. White arrows indicate cells with nuclear localization of T protein and red arrows indicate cells with cytoplasmic localization of T.

2.6 Does the Sudden Loss of *T* Expression from E8.0 Cause the *T* Knockout Phenotype in $T^{2J/Y88A}$ Mutants?

The knockout *T* phenotype seen in the $T^{2J/Y88A}$ mutant embryos could be due to the point mutation itself disrupting an important function of T, or simply due to the loss of *T* expression from E8.0 onwards. To address this, a system previously developed in our lab was utilized which allows for ubiquitous knockdown of *T* in a temporally-specific manner (Vidigal et al., 2010), in order to compare the phenotype in $T^{2J/Y88A}$ embryos to those that have lost *T* expression from E8.0 onwards. In short, ESCs were previously modified which harbor shRNA hairpins directed against *T* in the ubiquitously expressed *Rosa26* locus under control of a doxycycline-inducible promoter (Figure 2.13A and Vidigal et al., 2010). These *KD4-T* ESCs were used to generate embryos via tetraploid complementation and ubiquitous knockdown of *T* was induced in embryos from either E6.5 or E7.5 onwards by administration of doxycycline (Dox) in the drinking water of the pregnant mother. *T* expression was quantified in embryos at E8.5 and E9.5 to measure *T* knockdown *in vivo*. *T* expression was 40% of that in embryos that were uninduced (-Dox) at E8.5, and at E9.5, *T* expression was completely absent (Figure 2.13B), regardless of whether Dox was administered from E6.5 or E7.5. Expression of *T* by WISH revealed that *T* expression was normal in uninduced embryos and absent in $T^{2J/Y88A}$ mutant embryos (Figure 2.13C). At earlier stages (E8.25-E8.5), *T* was still expressed in embryos with *T* knockdown from E6.5 and E7.5, albeit at lower levels. At E9.0-E9.5, levels of *T* in Dox-induced embryos was either completely absent or severely down-regulated when compared to uninduced embryos (Figure 2.13C), illustrating that ubiquitous knockdown of *T* was achieved successfully.

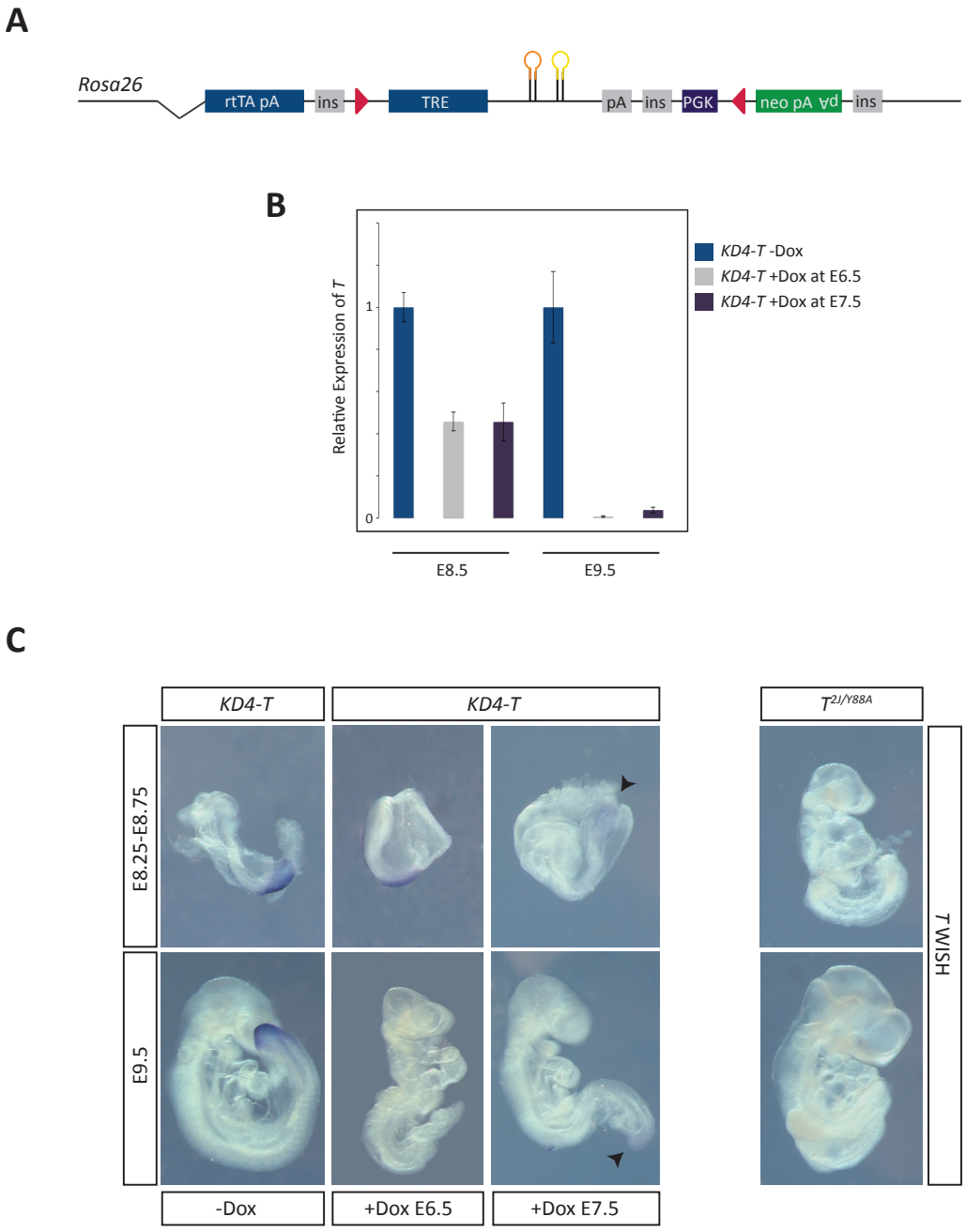


Figure 2.13: *T* expression in *KD4-T* knockdown embryos
(A) Schematic illustrating doxycycline-inducible shRNA hairpins against *T* in the *Rosa26* locus of *KD4-T* ESCs, modified from (Vidigal et al., 2010). (B) qRT-PCR of *T* expression in E8.5 and E9.5 embryos where *T* knockdown was induced at E6.5 (gray bars) or E7.5 (purple bars). Expression is compared to uninduced control *KD4-T* embryos and error bars represent standard error of the mean from technical replicates. (C) WISH for *T* expression in uninduced and induced *KD4-T* embryos compared to $T^{2J/Y88A}$ mutant embryos. Black arrowheads point to lower expression of *T* in *KD4-T* embryos induced at E7.5.

The phenotype of *KD4-T* knockdown embryos at E9.5 was extremely variable. WISH with probes recognizing *Shh* (notochord) and *Uncx* (somites) revealed that in uninduced embryos, notochord and somite formation was normal, as expected (Figure 2.14). In $T^{2J/Y88A}$ mutant embryos, notochord formation as assessed by *Shh* staining was completely disrupted, with *Shh* staining of the presumptive notochord present only at the very anterior end of the developing embryo (arrowheads in Figure 2.14). However, in embryos where knockdown of *T* was induced at E6.5 or E7.5, *Shh* staining of the presumptive notochord ceased at different points throughout development, most likely corresponding to the time when *T* is no longer present (arrowheads in Figure 2.14, *KD4-T* induced panels). Somite formation was also disrupted in *KD4-T* knockdown embryos induced at E6.5 and E7.5. Very disorganized somite-like structures could be visualized by *Uncx* staining, although the phenotype does not appear to be as severe as in $T^{2J/Y88A}$ mutants (Figure 2.14).

The variability of phenotypes seen in *KD4-T* induced embryos was most likely caused by variations in embryo age in the tetraploid complementation assay. The technique, which involves aggregation of ESC clumps with tetraploid host embryos, can result in some variation of embryo age during *in utero* development. Therefore, *KD4-T* embryos with more severe phenotypes were most likely induced by doxycycline at earlier stages as compared to the *T* downregulation in $T^{2J/Y88A}$ mutant embryos. Due to this variability, conclusive evidence that the downregulation of *T* from E8.0-E8.5 in $T^{2J/Y88A}$ mutant embryos is the cause of the mutant phenotype cannot be drawn using the *KD4-T* system as described here.

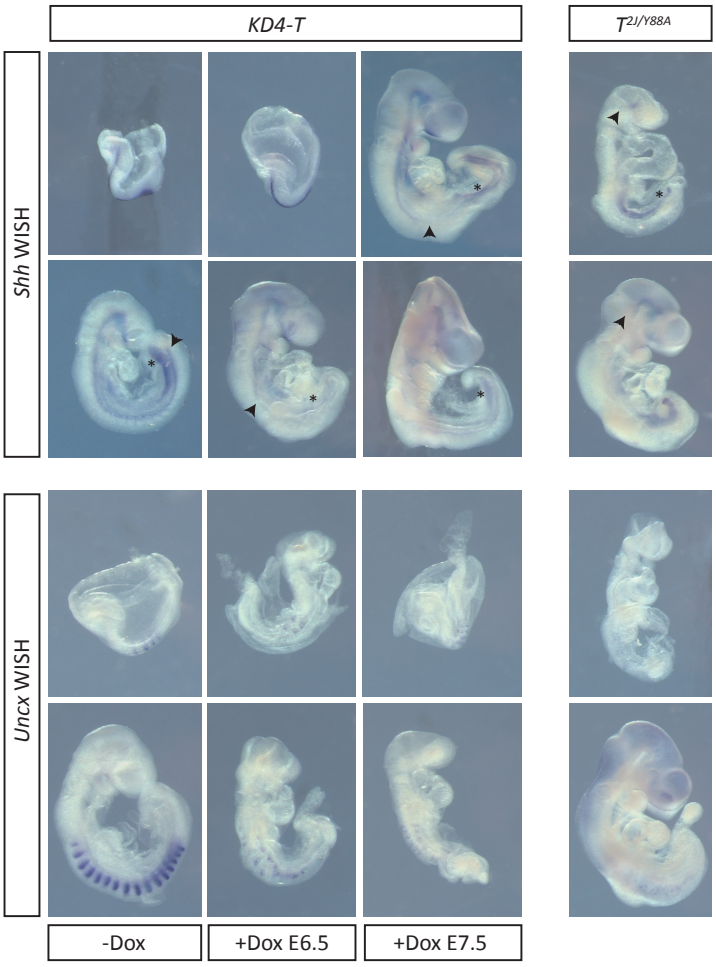


Figure 2.14: Notochord and somite formation in *KD4-T* embryos.
Shh WISH to visualize notochord formation in uninduced *KD4-T* embryos induced at E6.5 and E7.5, and $T^{2J/Y88A}$ mutant embryos. Black arrowheads depict the caudal-most domain of midline *Shh* expression, while asterisks indicate *Shh* staining in the endoderm. Lower panels illustrate *Uncx* WISH to visualize somite formation.

2.7 Expression of Downstream Target Genes in $T^{2J/Y88A}$ Mutant Embryos

$T^{2J/Y88A}$ mutant embryos display phenotypes whose gross morphology resembles that of *T* knockout embryos. I sought to investigate further whether the point mutant embryos resemble knockout embryos on the molecular level and whether there was an ability to distinguish between the function of T as a transcription factor and its function as a modifier of chromatin state. To this end, a fluorescent *T* reporter was integrated into mutant $T^{2J/Y88A}$ ESC to allow for isolation of mesodermal tissue from embryos that would be directly affected by loss of proper *T* function. This isolated tissue was then subjected to gene expression profiling to determine if there were molecular differences between the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants. The transgenic reporter was generated by replacing the first exon of *T* with the coding sequence of the fluorescent mCHERRY protein followed by a neomycin selection cassette with polyA signal within the *T* BAC (Figure 2.15A, gift from Dr. Arnold Schröder). Because the BAC contains regulatory sequences upstream and downstream of the normal *T* gene, *mCherry* expression should recapitulate endogenous *T* expression. The *T-mCherry* reporter BAC was then randomly integrated into $T^{2J/Y88A}$ mutant and $T^{2J/+}$ control ESCs, and integration of the reporter construct was confirmed by Southern blot (Figure 2.15B).

In order to validate that *T-mCherry* reporter expression recapitulates endogenous *T* expression, embryos were generated via tetraploid aggregation from both $T^{2J/Y88A}$ mutant and $T^{2J/+}$ control ESC containing the reporter transgene. Control $T^{2J/+}$*;Tg(*$T^{mCherry}$*)1Bgh* embryos at E9.5 appeared phenotypically normal (Figure 2.15C), and reporter expression was detected in tissues which normally express *T*, including the caudal end and notochord. mCHERRY was also detected in the developing somites, which is most likely due to the fact that the mCHERRY protein is more stable than endogenous T protein. In support of this, Supplementary Figure S2 demonstrates that expression of *mCherry* mRNA under control of the *Tbx6* promoter is detectable only in the caudal end and presomitic mesoderm of embryos at E9.5, in the same domains as *Tbx6* itself. mCHERRY protein is not degraded as quickly as endogenous TBX6, and persists in somitic tissue which arises from the presomitic mesoderm of the embryo. Mutant $T^{2J/Y88A}$*;Tg(*$T^{mCherry}$*)1Bgh* embryos were also dissected at E9.5 and imaged under a

fluorescent microscope to detect mCHERRY. Embryos derived from two separate clones (G8 and C8) expressed the reporter in correct domains of *T* expression (Figure 2.15C), however, we chose clone C8 for further RNA-seq analysis as embryos from clone G8 appeared developmentally delayed.

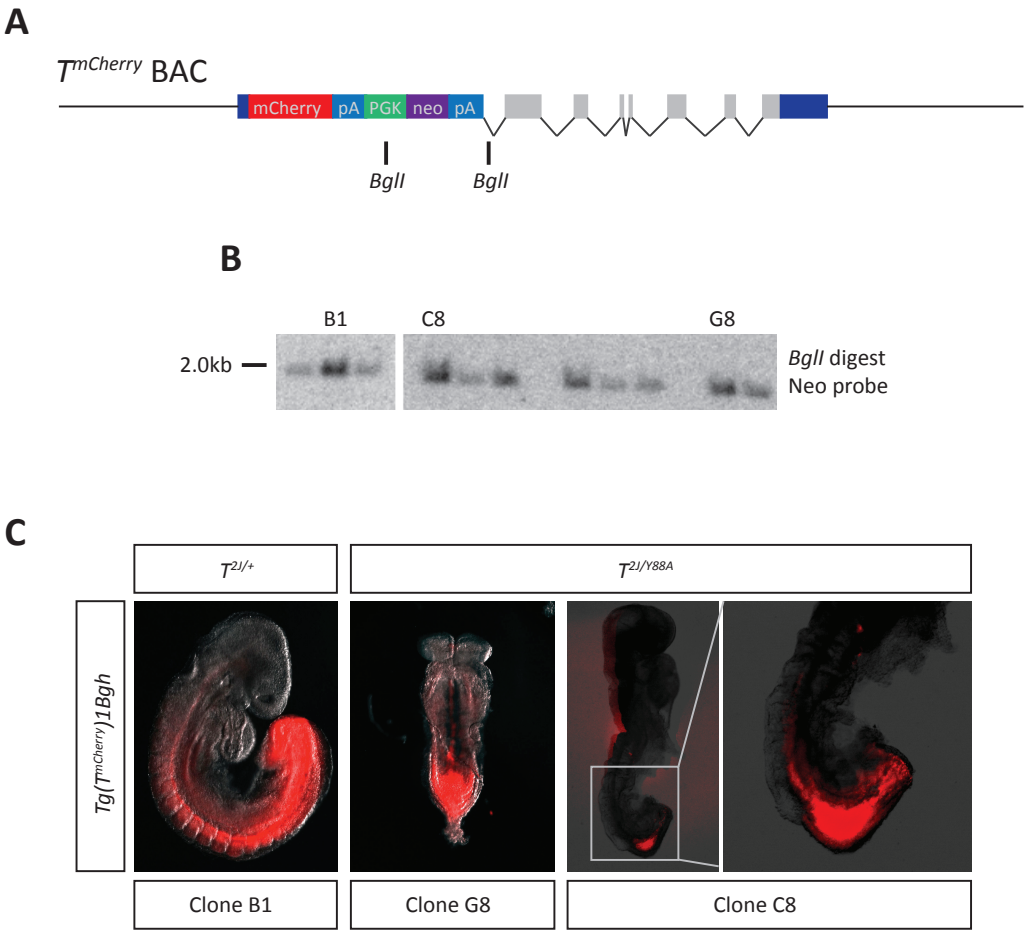
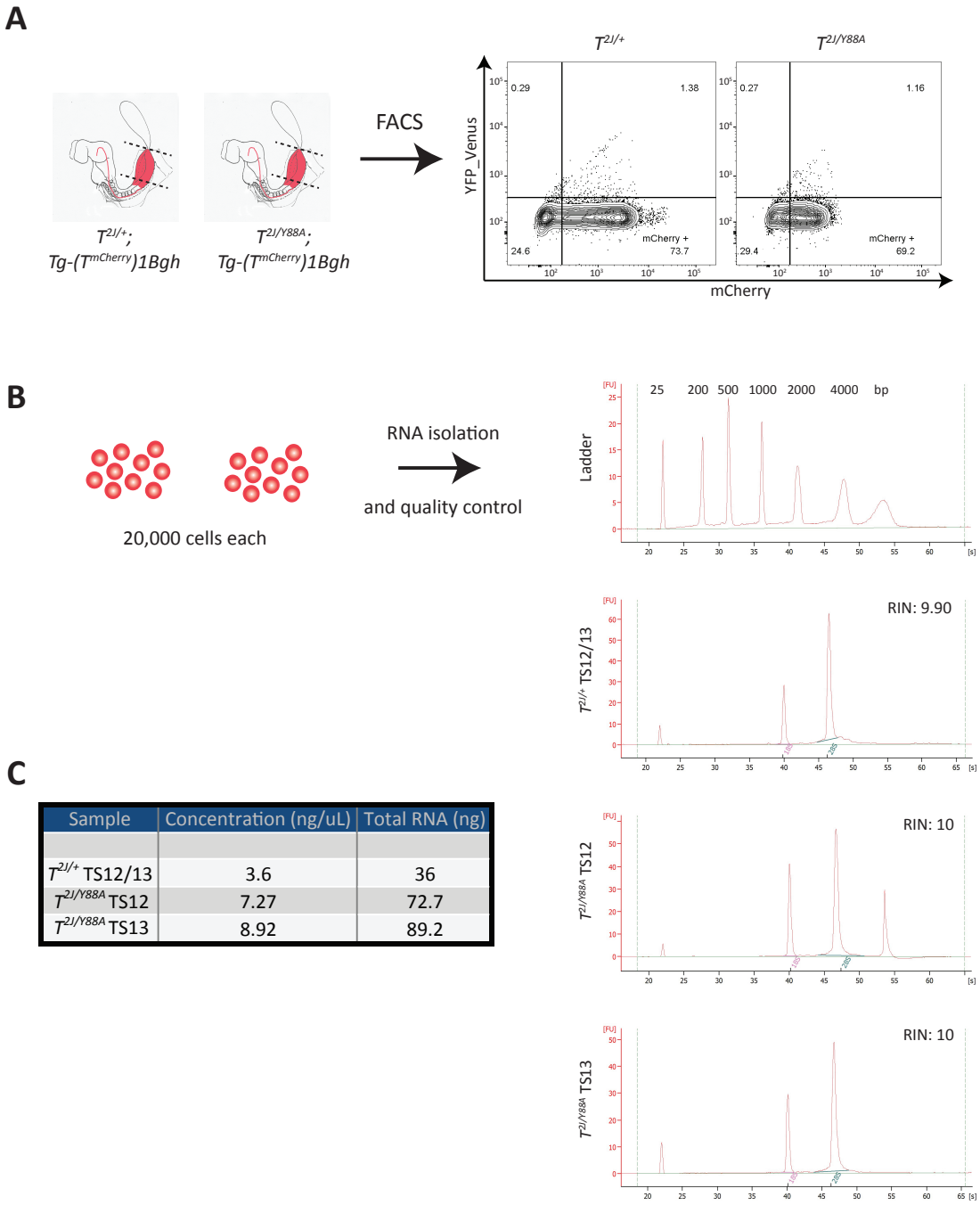


Figure 2.15: Integration of the *T-mCherry* reporter BAC into $T^{2J/+}$ and $T^{2J/Y88A}$ ESCs
(A) Schematic illustration of the $T^{mCherry}$ BAC. The *mCherry* open reading frame and neomycin selection cassette were integrated into the first exon of the *T* gene by BAC recombineering. The remaining *T* exons are displayed in gray and the 3'UTR in blue. (B) Southern blot analysis of ESC clones to check for random integration of the $T^{mCherry}$ BAC. ESC clones that have integrated the BAC displayed a 2.0kb band when digested with *BglI* and probed for neomycin. (C) *T-mCherry* was visualized by fluorescence microscopy in $T^{2J/+}$ and $T^{2J/Y88A}$ embryos. Fluorescent images are overlaid onto bright field images. pA, poly-A; PGK, phosphoglycerate kinase promoter.

Caudal ends from *$T^{2J/Y88A}$;Tg($T^{mCherry}$)1Bgh* and *$T^{2J/+}$;Tg($T^{mCherry}$)1Bgh* embryos were dissected at consecutive developmental stages, Theiler Stages (TS) 12 and 13. *T-mCherry*-positive cells were sorted by FACS (Figure 2.16A) and 20,000 cells per sample were used to isolate RNA. Isolated RNA quantity and quality were verified using a Bioanalyzer (Agilent Technologies) before proceeding to library preparation (Figure 2.16B, 2.16C). rRNA depletion, library preparation, and sequencing were performed as described in the Materials and Methods.

After high-throughput sequencing of the libraries (RNA-Seq), normalized FPKM values were used to determine dysregulated genes in *$T^{2J/Y88A}$* mutant embryos. FPKM values in the mutant were normalized to the *$T^{2J/+}$* heterozygous at TS12 and TS13 and the $\log_2$ of the fold change values ($\log_2$FC) were calculated and dysregulated genes with $\log_2$FC>1.5 were used for further analysis. Verification of the data was performed by examining the expression profile of genes that were previously shown to be *T* target genes; i.e. genes that were dysregulated in *T* knockout mutants and contained a *T* binding site in their promoter regions (Evans et al., 2012; Garnett et al., 2009; Gentsch et al., 2013). Genes involved in neural development, including *Sox1, Sox3,* and *Irx1*, were upregulated in *$T^{2J/Y88A}$* mutant embryos (Figure 2.17A). Furthermore, genes such as *T* itself, *Tbx6, Msgn1, Wnt3a,* and others were downregulated in *$T^{2J/Y88A}$* mutants, in agreement with previously reported data. In concurrence with the WISH and qRT-PCR data of *T* expression in *$T^{2J/Y88A}$* embryos, *T* expression in TS12 (E8.25) embryos is close to wild-type levels and levels of *T* begin to decrease rapidly at TS13 (E8.5, Figure 2.17C). Furthermore, in accordance with the presence of both transactivation and repression domains within the C-terminus of the T protein (Kispert et al., 1995), k-means clustering of dysregulated genes revealed the presence of two groups, genes that are downregulated and genes that are upregulated when compared to controls (Figure 2.17B). The agreement of dysregulated gene expression in the *$T^{2J/Y88A}$* mutant with previously reported data, suggested that the data sets generated were biologically sound and could be used for further analyses.



Sample	Concentration (ng/uL)	Total RNA (ng)
$T^{2J/+}$ TS12/13	3.6	36
$T^{2J/Y88A}$ TS12	7.27	72.7
$T^{2J/Y88A}$ TS13	8.92	89.2

Figure 2.16: Sorting and RNA isolation from *T-mCherry*-positive cells
(A) Caudal ends were dissected from $T^{2J/+}$;*Tg-*($T^{mCherry}$)*1Bgh* and $T^{2J/Y88A}$;*Tg-*($T^{mCherry}$)*1Bgh* embryos at TS12 and TS13. Red color indicates expression domains of the *T-mCherry* reporter. Dotted lines indicate the tissue dissected for sorting by FACS. FACS plots illustrate the population of *T-mCherry*-positive cells sorted for RNA isolation. (B) RNA from 20,000 *T-mCherry*-positive cells was isolated and the quality of RNA was checked using a Bioanalyzer. Plots illustrate quality of RNA samples with a RNA Integrity Number (RIN) assigned by the Bioanalyzer software. (C) Table illustrating quantity of RNA isolated from *T-mCherry*-positive cells. RNA was quantified using the Qubit® High Sensitivity Assay Kit.

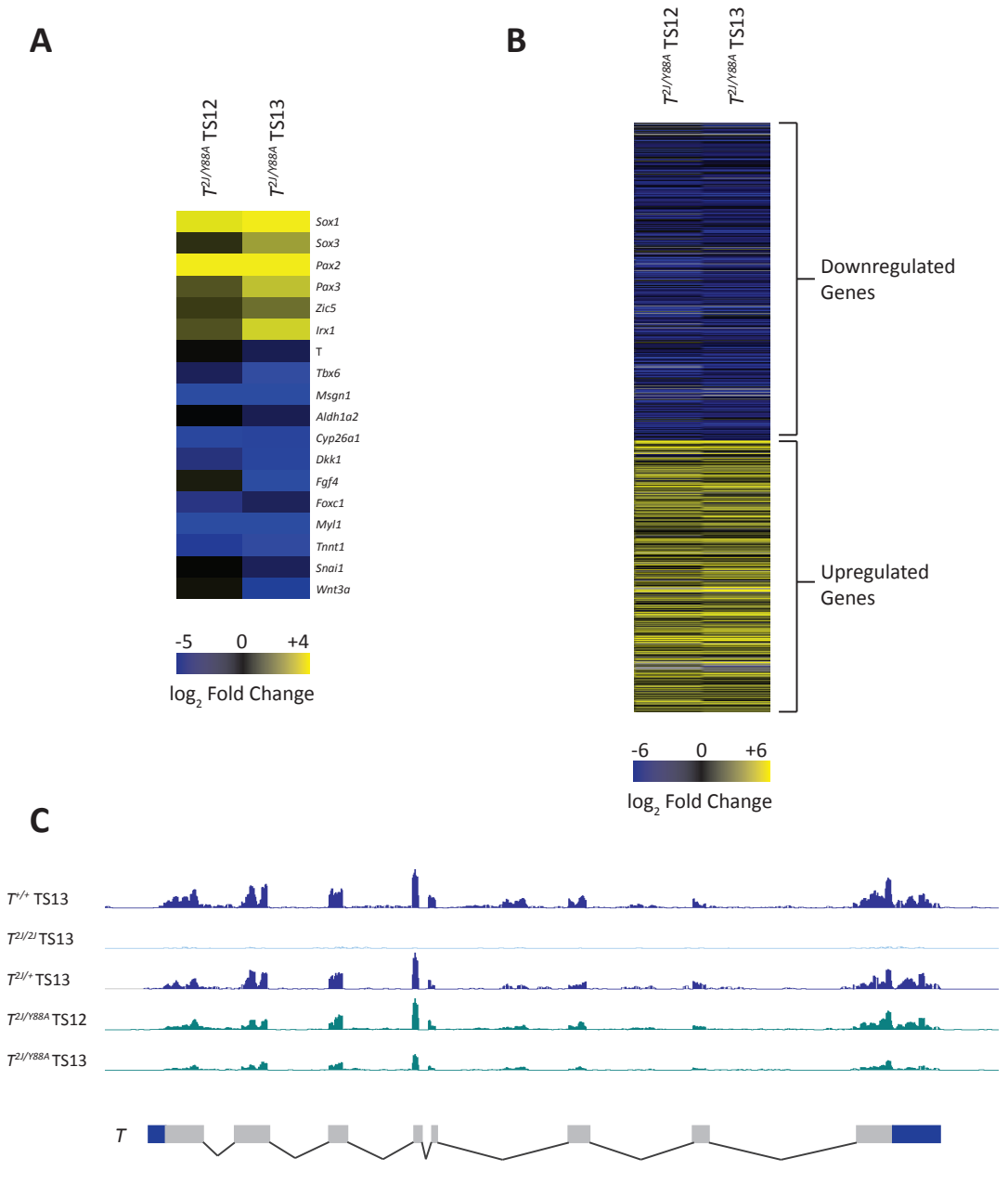


Figure 2.17: Verification of RNA-seq data
(A) Heat map depicting log_2FC values of genes that were previously shown to be *T* target genes. Samples are indicated at the top and gene names on the right. (B) k-means clustering of log_2FC values of all dysregulated genes grouped genes into two clusters, genes that are downregulated and genes that are upregulated. Samples are indicated at the top. (C) RNA-seq tracks of the *T* locus in indicated samples, with control samples in dark blue, $T^{2J/2J}$ knockout sample in light blue and $T^{2J/Y88A}$ samples in green. The y-axis indicates number of mapped reads and all tracks are displayed with the same scale.

2.8 Comparison of Dysregulated Gene Expression in $T^{2J/Y88A}$ versus $T^{2J/2J}$ Mutants

To identify the similarity between $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant gene expression signatures, I took advantage of an available dataset in which 20,000 *T-mCherry*-positive cells were isolated from the caudal ends of $T^{2J/2J}$ knockout and wild-type embryos at the same consecutive developmental stages, TS12 and TS13 (Dr. Frederic Koch). Hierarchical clustering of $\log_2$FC values of dysregulated genes was performed (Figure 2.18A). From this analysis, it became clear that the gene expression signatures from $T^{2J/Y88A}$ mutants at TS12 and TS13 were the most closely related, as they cluster together with a Pearson correlation coefficient of 0.84. The gene expression signatures from the $T^{2J/2J}$ mutants at TS12 and TS13 clustered together, albeit with a lower correlation coefficient than the $T^{2J/Y88A}$ samples, at 0.40. Furthermore, the Pearson correlation coefficient between the mutant $T^{2J/2J}$ and $T^{2J/Y88A}$ samples was 0.35, indicating that despite their similarity in gross morphological phenotypes, the mutants differed significantly based on gene expression profiles. From the hierarchical clustering in Figure 2.18A, it became clear that the $T^{2J/2J}$ TS12 sample seemed to be an outlier, which is most likely due to the wild-type TS12 sample used for normalization, based on principal component analysis of FPKM values (data not shown). For this reason, the rest of the analyses were focused on dysregulated gene expression of $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants at TS13 only.

Hierarchical clustering of dysregulated genes at TS13 was performed (Figure 2.18B). From this analysis, it became clear that there are groups of genes that are dysregulated similarly in the $T^{2J/Y88A}$ versus $T^{2J/2J}$ mutants (Clusters II and VI). However, a large portion of dysregulated genes did not follow the same trend in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants. For example, there were clusters of genes that are upregulated or downregulated in the $T^{2J/2J}$ mutant and not in the $T^{2J/Y88A}$ (Clusters I and VIII) or genes that are upregulated or downregulated in the $T^{2J/Y88A}$ mutant and not in the $T^{2J/2J}$ (Clusters III and V). Interestingly, there were also clusters of genes that were dysregulated in both $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants but in opposite directions (Clusters IV and VII). This suggests that the T^{Y88A} mutant is not simply a null allele, but is also disrupting a critical function of T during early mesoderm development.

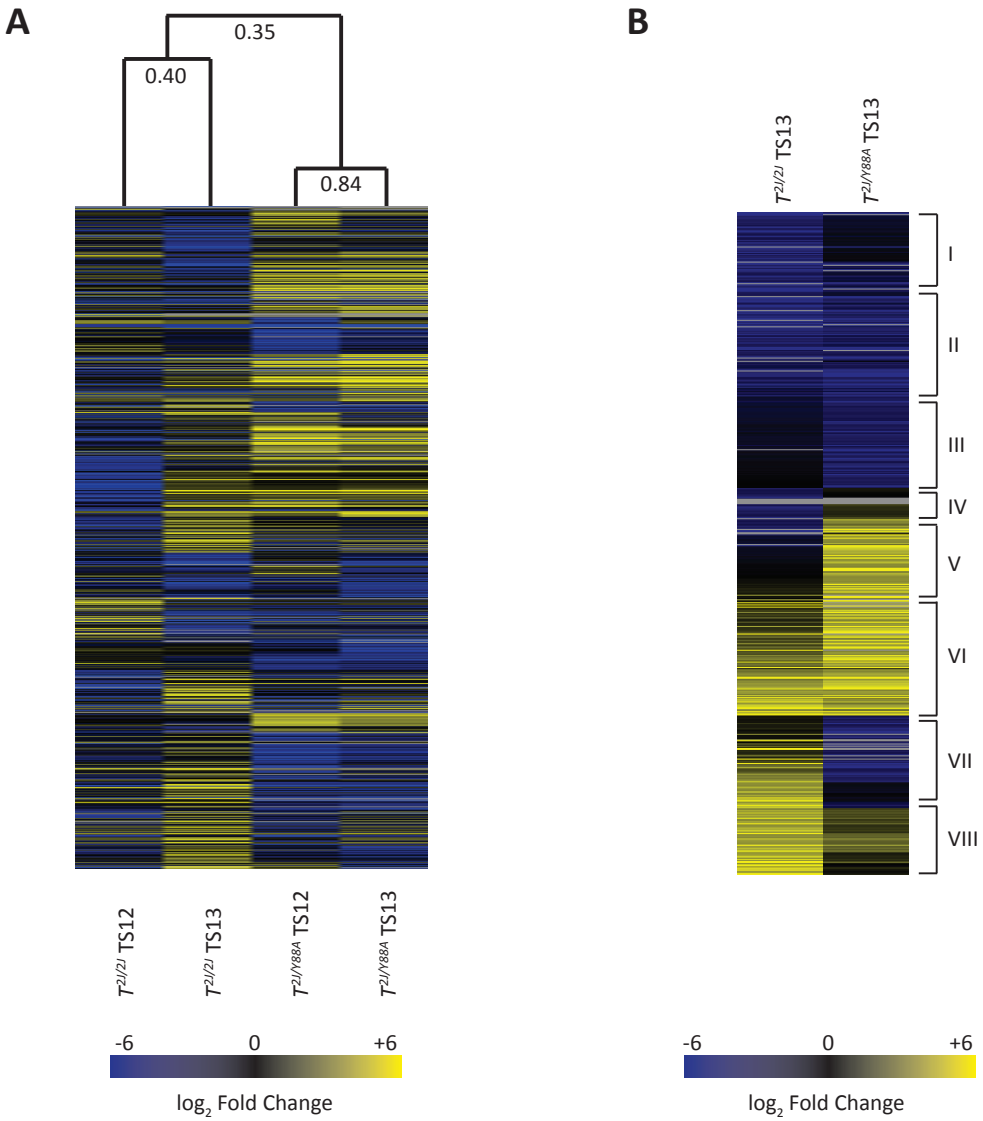


Figure 2.18: Hierarchical clustering of dysregulated genes in $T^{2J/Y88A}$ compared to $T^{2J/2J}$ mutants

(A) Hierarchical clustering of the $\log_2$FC values of dysregulated genes in all samples. The dendrogram illustrates the similarity between samples, with Pearson correlation coefficients displayed. (B) Hierarchical clustering of dysregulated genes in TS13 samples only. Roman numerals represent different clusters of genes as described in the text.

There were a total of 596 dysregulated genes in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants at TS13, and a comparison of these lists revealed that only 34.2% of dysregulated genes in the $T^{2J/Y88A}$ mutant were shared with $T^{2J/2J}$ mutants, while 40% of dysregulated genes in the $T^{2J/2J}$ mutant were shared with the $T^{2J/Y88A}$ mutant (Shared, Figure 2.19). The majority of dysregulated genes were found in only one of the two samples; 202 genes were dysregulated in the $T^{2J/2J}$ mutant only ($T^{2J/2J}$ only, Figure 2.19) and 259 were dysregulated in the $T^{2J/Y88A}$ mutant only ($T^{2J/Y88A}$ only, Figure 2.19), further demonstrating a functional difference between the T^{Y88A} and *T* null alleles.

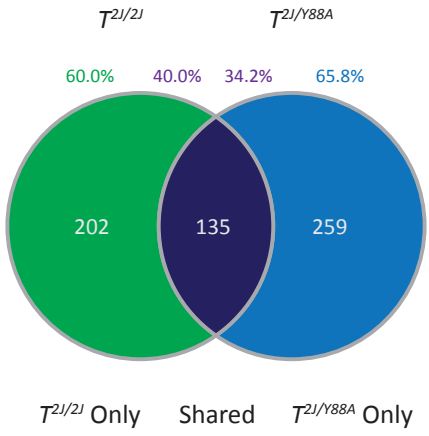


Figure 2.19: Dysregulated genes with a $log_2FC>1.5$ at TS13
Venn diagram illustrating the overlap of dysregulated genes in the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants.

In order to determine if there are biological functions enriched in dysregulated genes that belong to each of the categories described above, gene ontology (GO) term enrichment analysis was performed. Genes that are dysregulated in both the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants are enriched for GO terms including embryonic morphogenesis, cell fate commitment, pattern specification, and segmentation (Figure 2.20A). This included target genes such as *Tbx6, Wnt3a, Cyp26a1, Dll1, Meox1, Pax3, Fgf4,* and *Fgf8*, all of which have previously been shown to have roles in posterior mesoderm formation, patterning, and somitogenesis (Abu-Abed, 2001; Boulet and Capecchi, 2012; Buckingham and Relaix, 2007; Chapman and Papaioannou, 1998; Hrabe de Angelis et al., 1997; Mankoo et al., 2003; Naiche et al., 2011; Takada et al., 1994). Furthermore, dysregulated genes with GO terms such as kidney development, metanephros development, and urogenital system were enriched. These included genes such as *Fgf8, Hnf1b, Osr1, Six1,* and *Pax2*. This corroborates well with previous data published in our lab, demonstrating that *T* is necessary for proper urogenital development (Pennimpede et al., 2012). Another set of GO terms that were enriched in the shared dysregulated genes include skeletal system development and morphogenesis. This included genes such as *Fgf18*, *Six1*, *Col2A1*, and *Tcf15*, consistent with previous studies in our lab showing that late induction of *T* knockdown results in embryos with defects in skeletogenesis (Pennimpede et al., 2012).

A look at some of the other dysregulated genes shared between the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants at TS13 revealed that multiple components of the Wnt signaling pathway were also dysregulated, including *Dkk1, Rspo1, Rspo3, Wnt3a*, and *Frzb*. As the Wnt signaling pathway is known to play a role in axis elongation and somitogenesis, this may help to explain the morphological phenotypes seen in the mutant embryos (Galceran et al., 1999). Moreover, *Tbx4*, a gene shown to play an important role in proper allantois formation and chorio-allantoic fusion, is downregulated in the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants (Naiche and Papaioannou, 2003). This may provide a mechanistic explanation as to why these mutants die at E10.0 due to failure of the allantois to properly attach to the chorion. Lastly, *Shh*, a gene known to play an important role in notochord maintenance (Chiang et al., 1996), is a *T* target that is downregulated in both $T^{2J/2J}$ and $T^{2J/Y88A}$ mutants, which may provide an explanation for the loss of notochord in these mutants.

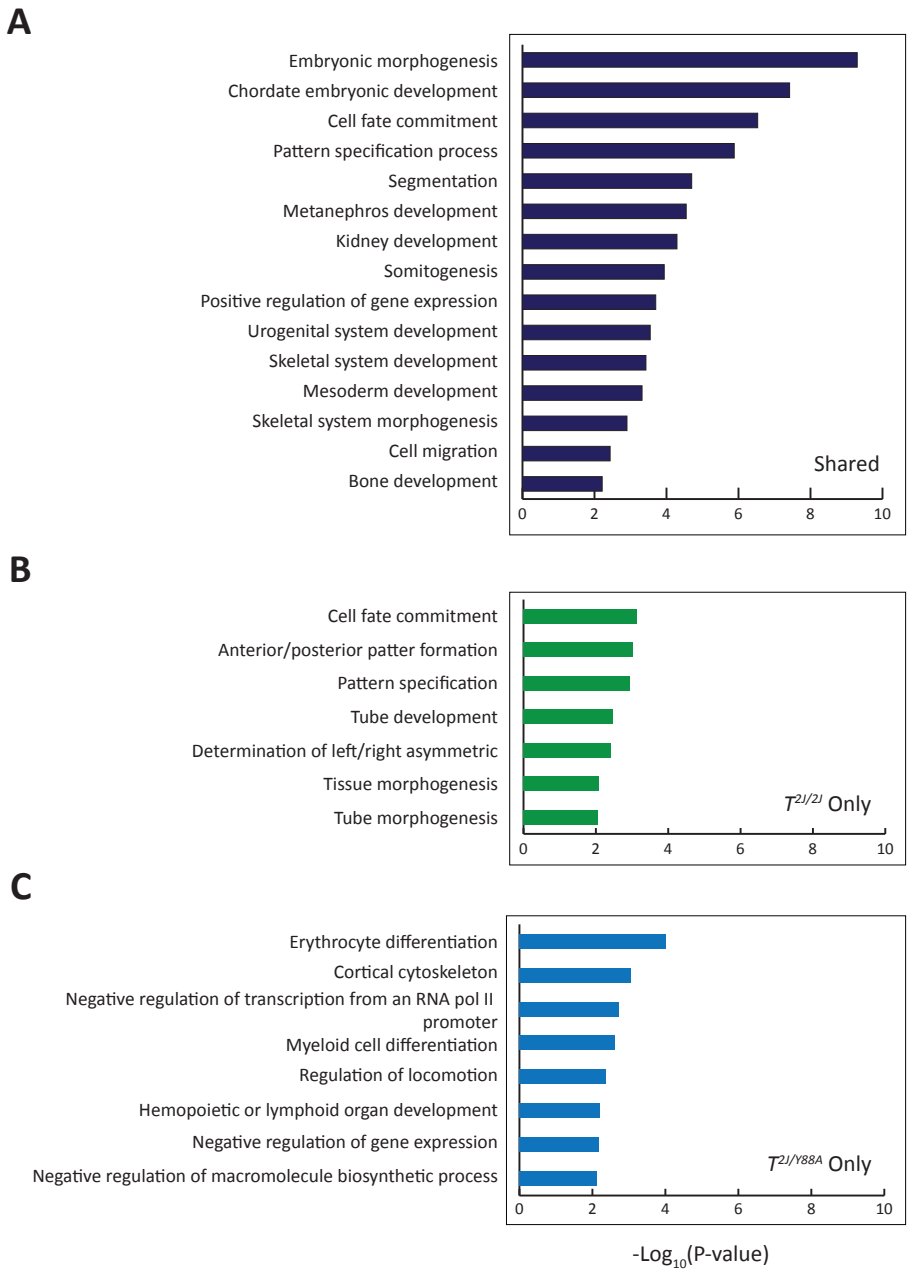


Figure 2.20: GO term enrichment analysis of dysregulated genes in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants

GO terms of genes dysregulated in both $T^{2J/Y88A}$ and $T^{2J/2J}$ (A), $T^{2J/2J}$ only (B) and $T^{2J/Y88A}$ mutants only (C). Graphs depict the $-\text{Log}_{10}$ of the P-value for each GO term, as given by the DAVID bioinformatic tool (Huang et al., 2009).

Perhaps the most interesting group of dysregulated genes are the 259 genes that are dysregulated in the $T^{2J/Y88A}$ mutant but not in the $T^{2J/2J}$. These genes rely on the interaction of T with histone-modifying proteins for normal expression, and are most likely able to be compensated for by other factors in the $T^{2J/2J}$ knockout embryos. The most significantly enriched GO terms associated with these dysregulated genes included erythrocyte differentiation, myeloid differentiation, and hemopoietic or lymphoid organ development. These GO terms included genes such as *Ank1, Alas2, Trim10, Klf1, Dll4, Hbb-b2,* and *Hba-x*, all of which have been previously shown to play important roles in erythrocyte differentiation and development (Drissen et al., 2005; Harada et al., 1999; Hodge et al., 2006; Kauppi et al., 2012; Laranjeiro et al., 2012; Rank et al., 2009; Sadlon et al., 1999). It is thought that hematopoietic and endothelial cells arise from a common precursor cell called the hemangioblast. Furthermore, the hemangioblast is thought to arise from the mesodermal lineage and express *T*. Multiple genes expressed in hemangioblasts are dysregulated in $T^{2J/Y88A}$ mutant embryos (Table 1). Thus, it is interesting to speculate that *T* expression in the hemangioblast serves to prime these cells for differentiation into hematopoietic lineages, and interaction of T with histone-modifying proteins is especially important for hemangioblast commitment and the development of erythrocytes. Furthermore, it appears as if this function of T can be compensated for by other factors in $T^{2J/2J}$ knockout cells, as the expression of hemangioblast and erythrocyte differentiation genes is not significantly dysregulated.

Gene	log_2FC of $T^{2J/Y88A}$ Compared to $T^{2J/+}$ at TS12	log_2FC of $T^{2J/Y88A}$ Compared to $T^{2J/+}$ at TS13
Kdr (Flk1)	-1.175409751	-0.798188719
Tal1 (*Scl)*	-1.842413206	-1.478104652
Lmo2	-1.255019427	-1.835975406
Pecam1	-2.091780715	-1.538799653
Tie1	-2.343490664	-2.048675087
Epor	-2.27978712	-1.23709506

Table 1: Changes in expression of hemangioblast genes in the $T^{2J/Y88A}$ mutant

2.9 Comparison of RNA-seq Data to Known T Binding Sites

I chose to take advantage of another dataset generated by Dr. Frederic Koch, which included a genome-wide analysis of T binding sites in *T-mCherry*-positive mesodermal cells differentiated *in vitro*. The dataset included motif analysis of T binding sites and it was shown that many of the T binding sites contained a palindromic T-box binding site, suggesting that T binds to its target genes as a dimer. A ranked list of T binding sites was generated according to how well the site adhered to the palindromic T-box binding site. As described earlier, since the Y88A mutation in T was neighboring the M87 residue shown to be important for dimerization, I wanted to determine if there was any correlation between the dysregulated genes in $T^{2J/Y88A}$ mutant embryos and dimerization of T; i.e. whether the dysregulated genes in $T^{2J/Y88A}$ mutants were genes where T must bind as a dimer, and the T^{Y88A} mutation is disrupting this dimerization.

To this end, a comparison was made between the dysregulated genes in $T^{2J/2J}$ and $T^{2J/Y88A}$ mutant embryos and the list of T binding sites. Close to half (49%) of the genes dysregulated in both $T^{2J/2J}$ and $T^{2J/Y88A}$ mutants (Shared) were directly bound by T in *in vitro* differentiated mesoderm. Furthermore, 44% of dysregulated genes in the $T^{2J/2J}$ mutant and 37% of dysregulated genes in the $T^{2J/Y88A}$ mutant alone were direct T targets in *in vitro* differentiated mesodermal cells. Interestingly, there was no correlation between strength of T palindromic binding sites and dysregulated gene expression in $T^{2J/Y88A}$ mutant embryos. A plot comparing the distribution of strength of T palindromic sites in differentially expressed versus non-differentially expressed genes illustrates that there is no enrichment of stronger T palindromic sites in dysregulated genes unique to the $T^{2J/Y88A}$ mutant (p=0.996, Figure 2.21B). Furthermore, there was also no enrichment of T palindromic sites in dysregulated genes unique to the $T^{2J/2J}$ mutant (p=0.323, Figure 2.21A). There was a slight enrichment in dysregulated genes shared between the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants (p=0.033, Figure 2.21C), however as the $T^{2J/2J}$ mutant contains no T protein, the dysregulation of these genes cannot be explained by a lack of dimerization of T. From these analyses combined with the genetic analyses in section 2.3, it is concluded that the T^{Y88A} mutation is not likely to disrupt dimerization of T, but instead

another critical function of this protein during mesodermal differentiation and development.

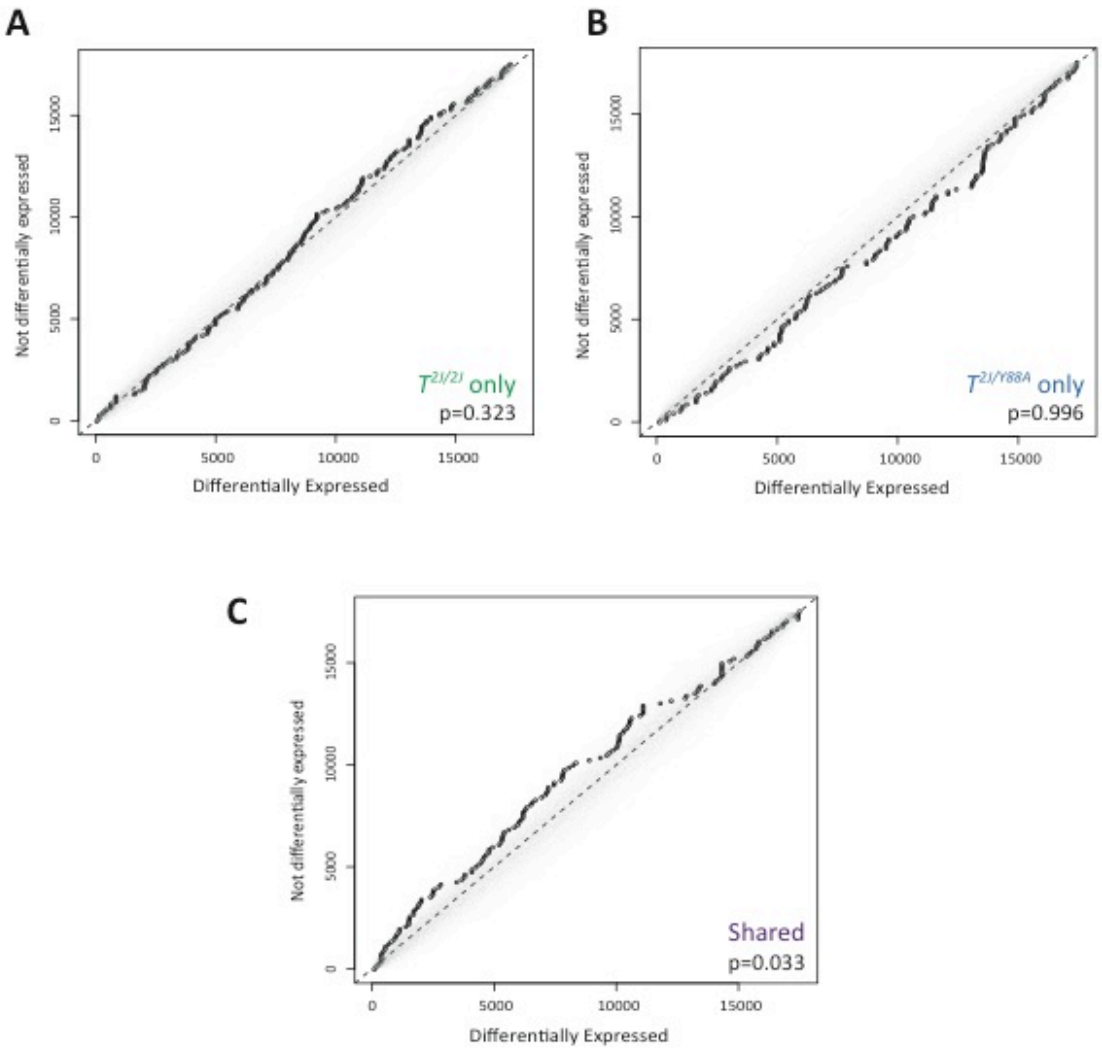


Figure 2.21: Comparing dysregulated genes and dimerization of T
Q-Q plots of distribution of strength of T palindromic binding sites in dysregulated (differentially expressed) genes versus the distribution in non-dysregulated (not differentially expressed) genes in (A) $T^{2J/2J}$ only, (B) $T^{2J/Y88A}$ only, and (C) Shared dysregulated genes. P-values were calculated using the two-sample Kolmogorov-Smirnov (K-S) test. Plots provided by Matt Huska (Max Planck Institute for Molecular Genetics, Berlin, Germany).

2.10 Investigating the Physical Interaction Between T and Histone- and Chromatin-Modifying Enzymes

It was previously shown that T-box factors have the ability to physically interact with H3K27 demethylase activity (Miller et al., 2008), however it was never explicitly shown that the Y→A mutation disrupts the physical interaction between T-box factors and KDM6A/6B. I chose to investigate this interaction further using co-immunoprecipitation assays. I generated constructs over-expressing myc-tagged T^{WT} or T^{Y88A} and V5-Avi-tagged KDM6A or KDM6B (Figure 2.22A) and used them to determine the interaction between T^{WT} and T^{Y88A} with the KDMs. Transient transfection and over-expression in HEK293T cells followed by immunoprecipitation of KDM6A/6B with Streptavidin-coupled beads revealed that T^{Y88A} mutant protein more strongly interacts with KDM6A and KDM6B than the wild-type T protein (Figure 2.22B). The stronger interaction is clear with KDM6B (Figure 2.22B, lane 3 compared to lane 4). Expression of KDM6A was quite weak (Figure 2.22B, input, lanes 1 and 2), however mutant T^{Y88A} was still pulled down more strongly than T^{WT} (Figure 2.22B, IP, lane 1 compared to lane 2). Interestingly, similar results were also obtained with another T-box transcription factor, TBX6. Immunoprecipitation of KDM6A and KDM6B revealed a stronger interaction with TBX6 containing an analogous Y137A amino acid substitution, $TBX6^{Y137A}$, than wild-type TBX6 (data not shown, Maria-Theodora Melissari, Master's thesis, Max Planck Institute for Molecular Genetics, Berlin, Germany).

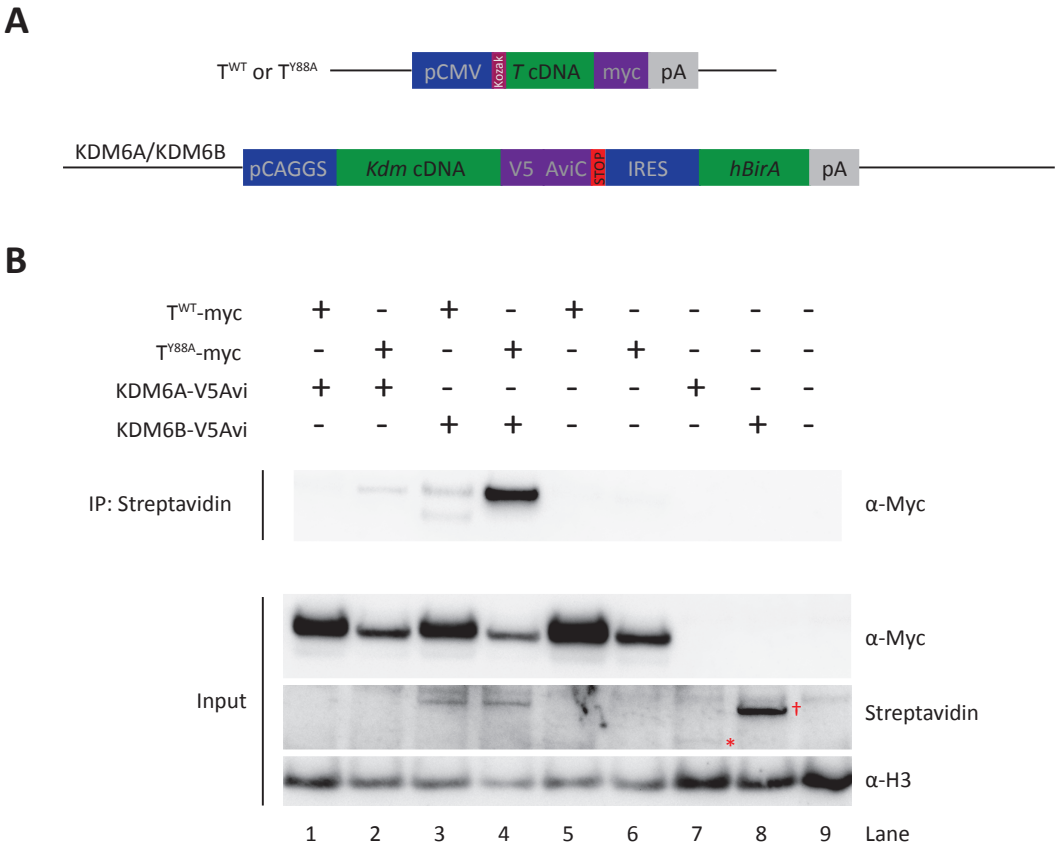


Figure 2.22: Co-immunoprecipitation of T with KDMs in HEK293T cells
(A) Schematic of T and KDM6A/6B over-expression constructs that were co-transfected into HEK293T cells. (B) Co-immunoprecipitation of KDM6A and KDM6B followed by Western blotting revealed a stronger interaction between mutant T^{Y88A} and the KDMs than between T^{WT} and the KDMs. α-Myc was used to detect T and Streptavidin to detect the KDMs. α-H3 was used as a loading control. In the input Streptavidin blot, the asterisk indicates KDM6A while the cross indicates KDM6B.

The result described above was counterintuitive when comparing this to previously published data. However, it does not exclude the possibility that stronger binding to T may disrupt normal H3K27 demethylase function, leading to the loss of demethylase activity in previously published experiments. Next, I sought to determine if this stronger interaction also held true in a mesodermal cell setting. To this end, I generated P19 mouse embryonal carcinoma cell lines stably expressing V5-Avi-tagged wild-type and Y88A T protein. This system was chosen because P19 cells are relatively easy to differentiate into mesodermal cells *in vitro*, by adding 1% DMSO to the cell culture medium and culturing cells in non-adhesive plates to allow for aggregation. Figure 2.23A depicts the over-expression construct that was stably integrated into P19 cells and a Western blot verifying over-expression of tagged T^{WT} and T^{Y88A}. These cells were then used for differentiation into mesodermal cells and subsequent co-immunoprecipitation to look for interaction of T^{WT} and T^{Y88A} with the KDMs and other chromatin-modifying enzymes. Immunoprecipitation of wild-type and mutant Y88A T from mesodermal-like P19 cells revealed no significant changes in binding to KDM6A or BRG1 (Figure 2.23B). Furthermore, an interaction with KDM6B and MLL complex members RBBP5 and WDR5 could not be detected in this system. Interestingly, a slight interaction with H3K27 acetyltransferase p300 could be detected but there was no difference between wild-type and mutant Y88A T. Comparing the coIP results from HEK293T and P19 cells, it appears as if the interaction between T-box factors and KDM proteins may be extremely context-dependent, illustrating the need to investigate this interaction in more relevant mesodermal cell types.

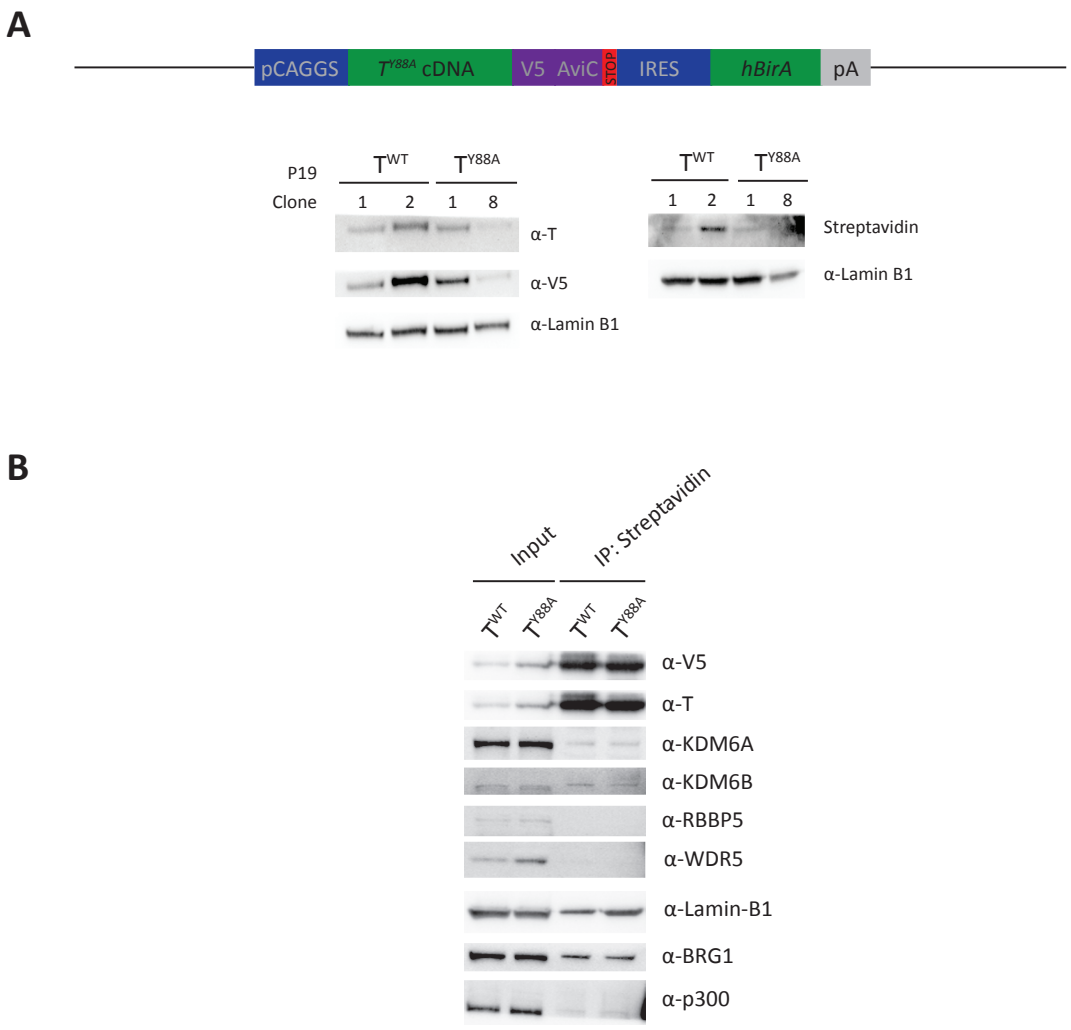


Figure 2.23: Co-immunoprecipitation of T with KDMs in differentiated P19 cells
(A) Schematic illustrating the T over-expression construct stably integrated into P19 cells. Lower panels depict Western blots to verify proper tagging of wild-type and Y88A mutant T in various P19 cell clones. Streptavidin was used to detect the biotin tag, α-V5 was used to detect the V5 tag, and α-Lamin-B1 was used as a loading control. (B) Co-immunoprecipitation of T followed by Western blotting revealed the binding of T to KDM6A, BRG1, and p300, but no differences between wild-type and Y88A mutant T were observed. α-Lamin-B1 was used as a loading control.

2.11 Investigating the Functional Consequences of Loss of T Interaction with Histone-Modifying Proteins

From the data described in previous sections, it appears that the mutant T^{Y88A} protein has the ability to bind KDM6A in mesodermal-like cells, but an interaction of T with members of the MLL complex was not detected. Although T and KDM6A can physically bind, it is not clear whether the mutant protein is disrupting normal KDM6A function. The loss of a functional interaction between T and histone-modifying enzymes in the T^{Y88A} point mutant may cause alterations in the chromatin environment of early mesodermal cells, leading to the disruption of normal mesoderm differentiation and the phenotypic consequences in $T^{2J/Y88A}$ mutants described in the previous sections. To validate this, ChIP-seq using antibodies directed against multiple histone marks (H3K27me3, H3K4me3, H3K27ac and H3K4me1) was performed. For this, I chose to use an *in vitro* differentiation method established in our laboratory to differentiate ESCs into T^{+} mesodermal cells whose gene expression signature recapitulates that of mesoderm within developing embryos. This method allows for the generation of the large amount of mesodermal cells necessary for ChIP-Seq.

Briefly, ESCs were aggregated by the hanging drop method and allowed to form embryoid bodies (EBs) overnight. The next day, EBs were plated onto fibronectin-coated plates and treated with conditioned medium to promote mesodermal differentiation (Schröder, 2010; modified by Dr. Pavel Tsaytler, personal communication). RNA and protein were isolated from EBs undergoing mesoderm differentiation at 24-hour time points throughout the differentiation protocol. In control $T^{2J/+}$ cells, *T* RNA expression is initiated at day 2 of the differentiation procedure and peaks at day 4 of differentiation with levels 20-fold higher than undifferentiated ESCs (Figure 2.24B). In the $T^{2J/Y88A}$ mutant, *T* expression is initiated at day 2 at similar levels when compared to the control. However, expression of *T* is not maintained and by day 5 of differentiation the mRNA in the $T^{2J/Y88A}$ mutant is nearly absent, similar to what was seen in $T^{2J/Y88A}$ mutant embryos *in vivo*. Western blot analysis of T protein levels during the differentiation protocol revealed similar trends; with T protein present at day 2 and peaking at day 4 in the control $T^{2J/+}$ cells and initiation but lack of maintenance of T protein levels in the $T^{2J/Y88A}$

mutant. From these results, it appears as if the ESC differentiation protocol has the ability to recapitulate what was seen previously in control $T^{2J/+}$ and mutant $T^{2J/Y88A}$ embryos *in vivo*, and led to the use of this protocol to generate mesodermal cells for ChIP-Seq experiments.

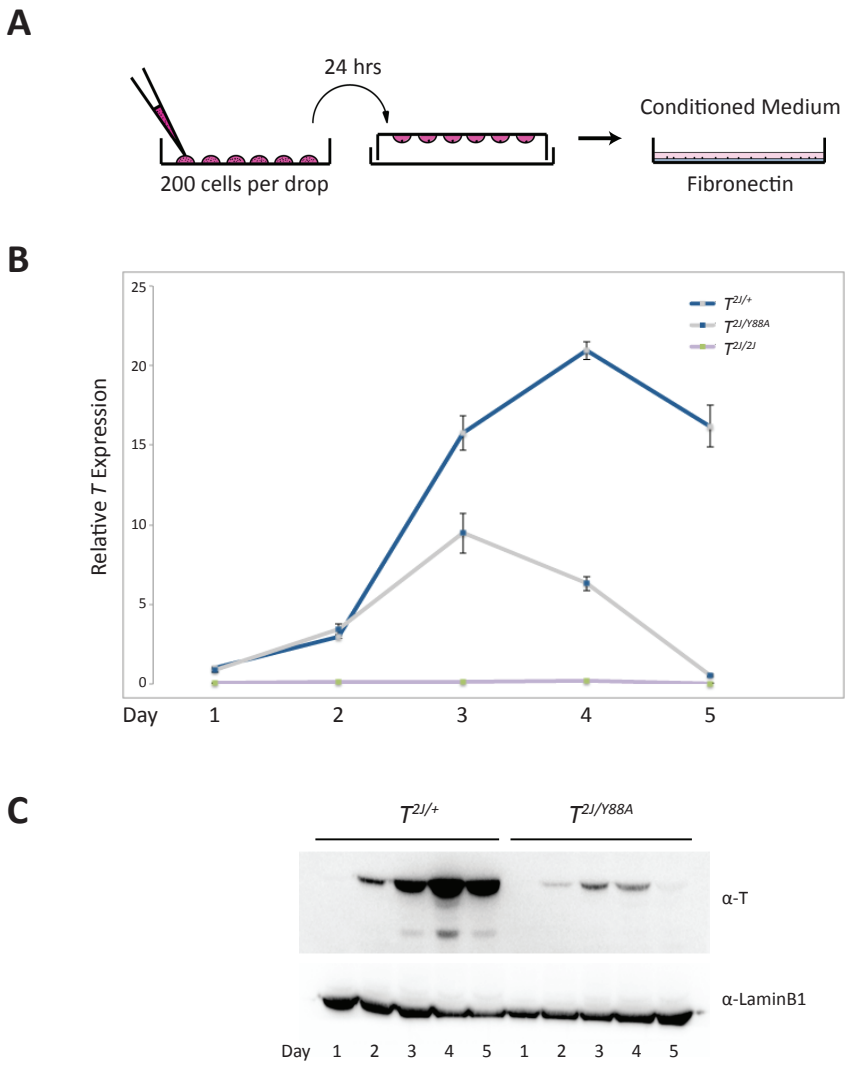


Figure 2.24: Differentiation of ESCs into mesoderm recapitulates the *in vivo* situation
(A) Schematic illustrating the ESC *in vitro* differentiation protocol. Embryoid bodies (EBs) were generated using hanging drops of 200 ESCs per drop. EBs were plated onto fibronectin coated plates and treated with conditioned media for a period of 5 days (B) qRT-PCR of relative *T* expression from differentiated ESCs at the indicated time points. The $T^{2J/2J}$ ESCs were used as a negative control. Error bars indicate standard error of the mean from technical replicates. (C) Western blot analysis of T levels in differentiated ESCs at the indicated time points. α-LaminB1 was used as a loading control.

In order to assay the histone modification landscape in $T^{2J/Y88A}$ and $T^{2J/2J}$ compared to control $T^{2J/+}$, *T-mCherry*-positive mesodermal cells were sorted by FACS and shearing of chromatin to sizes between 200-500bp was verified using a Bioanalyzer (Figure 2.25A and 2.25B). The concentration of sheared chromatin used per ChIP is illustrated in Figure 2.25C. The setup of the ChIP experiment and quality of the library preparations of immunoprecipitated material sent for high-throughput sequencing is shown in Figure 2.26.

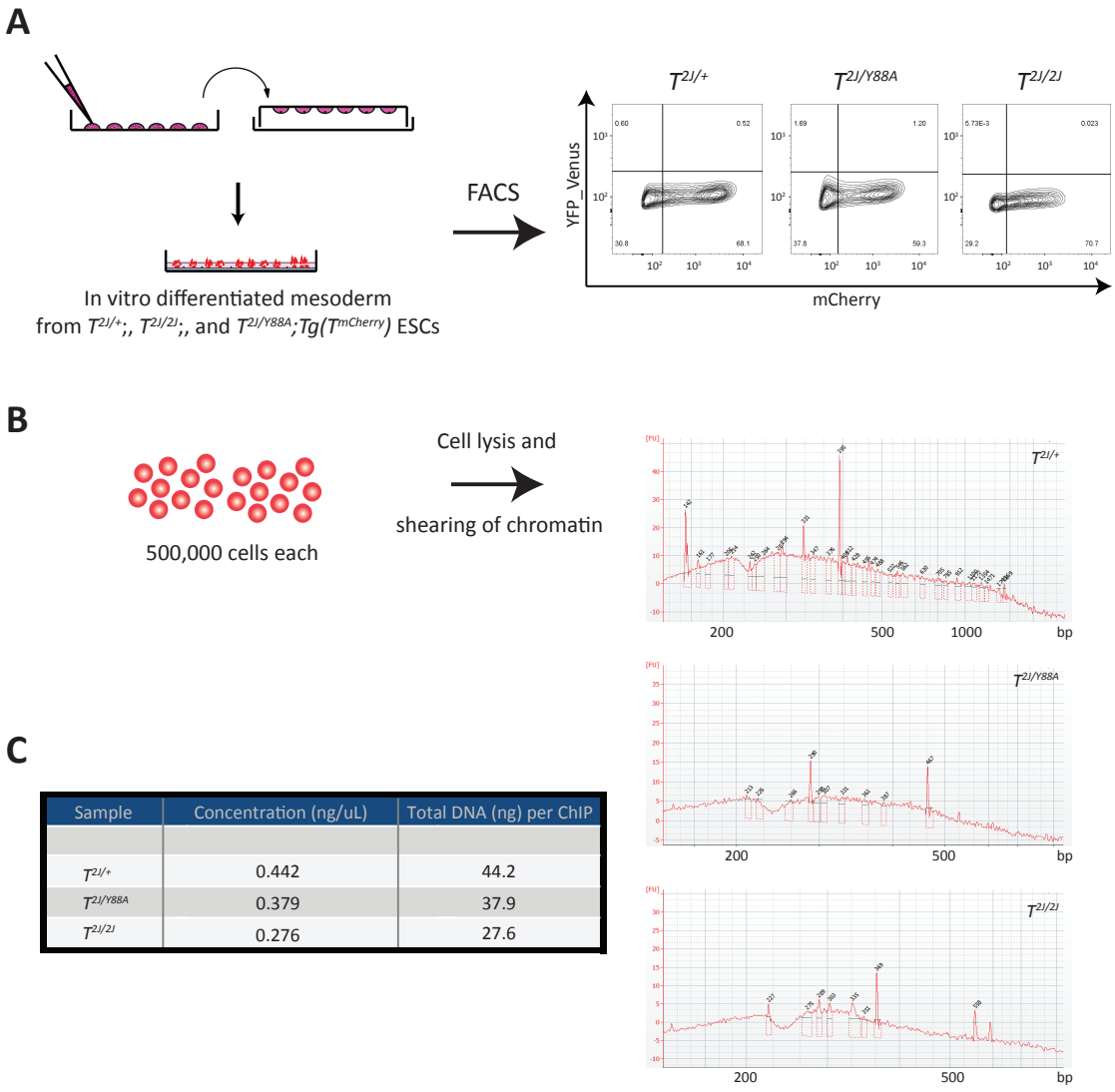


Figure 2.25: Sorting and shearing of chromatin from *T-mCherry*-positive cells
(A) $T^{2J/+}$;Tg-($T^{mCherry}$), $T^{2J/Y88A}$;Tg-($T^{mCherry}$), and $T^{2J/+}$;Tg-($T^{mCherry}$) ESCs were differentiated into mesoderm *in vitro*. *T-mCherry*-positive cells were sorted on day 4 of differentiation. FACS plots illustrate the population of *T-mCherry*-positive cells sorted for ChIP. (B) Sheared chromatin from 500,000 *T-mCherry*-positive cells was isolated and the shearing was checked using a Bioanalyzer and chromatin from 1,000 cells. Plots illustrate quality and size of sheared chromatin, with the majority of fragments falling between 200-500bp. (C) Table illustrating quantity of chromatin isolated from *T-mCherry*-positive cells, using the Qubit® High Sensitivity DNA Assay Kit. The total DNA used per ChIP is depicted in ng.

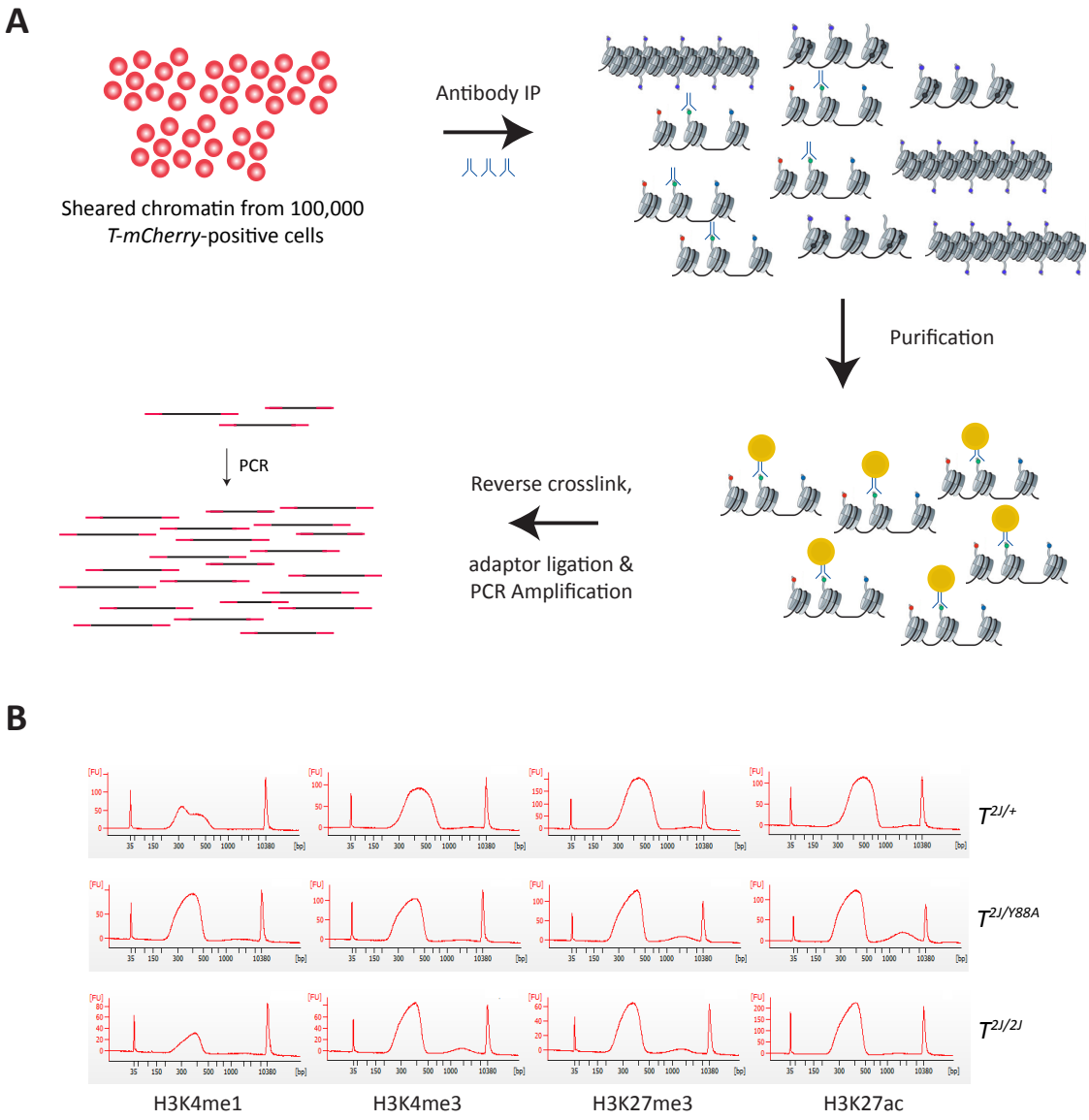


Figure 2.26: ChIP-Seq set-up and library verification
(A) Sheared chromatin from 100,000 *T-mCherry*-positive cells was immunoprecipitated using antibodies recognizing specific histone modifications. After purification and reverse crosslinking, adaptors for sequencing were ligated and the libraries were amplified using PCR. (B) Electropherograms of immunoprecipitated DNA from analysis using the Bioanalyzer revealed that the library preparation for sequencing was successful. The majority of immunoprecipitated DNA was between 300-600bp, corresponding to 200-500bp fragments with 100bp of adaptor sequence.

2.11.1 H3K27me3 and H3K4me3 in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutant Mesodermal Cells

I first examined H3K27me3 profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant mesodermal cells compared to the $T^{2J/+}$ control, as it was expected that differences may arise due to the loss of a functional interaction between T and KDM6A, leading to a loss of histone demethylation. However, heat maps of H3K27me3 at ±4kb from both transcription start sites and T binding sites in the genome revealed that there were only minimal increases in H3K27me3 in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to control (Figure 2.27A, 2.27B). Clusters of genomic regions with higher levels of H3K27me3 were examined more closely and plots of the density of H3K27me3 at ±4kb from both transcription start sites and T binding sites in the genome confirmed that there were no significant changes in H3K27me3 in mesodermal cells from the mutants compared to the control. Furthermore, there were also no differences in levels of H3K4me3, a mark known to demarcate transcription start sites and be deposited by members of the MLL complex of proteins, at ±4kb from transcription start sites in the genome of $T^{2J/Y88A}$ or $T^{2J/2J}$ mutant mesodermal cells when compared to the $T^{2J/+}$ control (Figure 2.28). Examination of both H3K27me3 and H3K4me3 at the transcription start sites of the 596 dysregulated genes from *T-mCherry*-positive cells sorted from $T^{2J/Y88A}$ or $T^{2J/2J}$ mutant embryos also revealed no significant differences. H3K27me3 and H3K4me3 levels were comparable at ±4kb of the TSS of dysregulated genes between mutant $T^{2J/Y88A}$ or $T^{2J/2J}$ and control $T^{2J/+}$ mesodermal cells (Supplementary Figure S4). From these results, it was clear that differences in expression of dysregulated genes in mutant $T^{2J/Y88A}$ or $T^{2J/2J}$ embryos cannot be attributed to changes in H3K27me3 or H3K4me3 at the TSS of these genes.

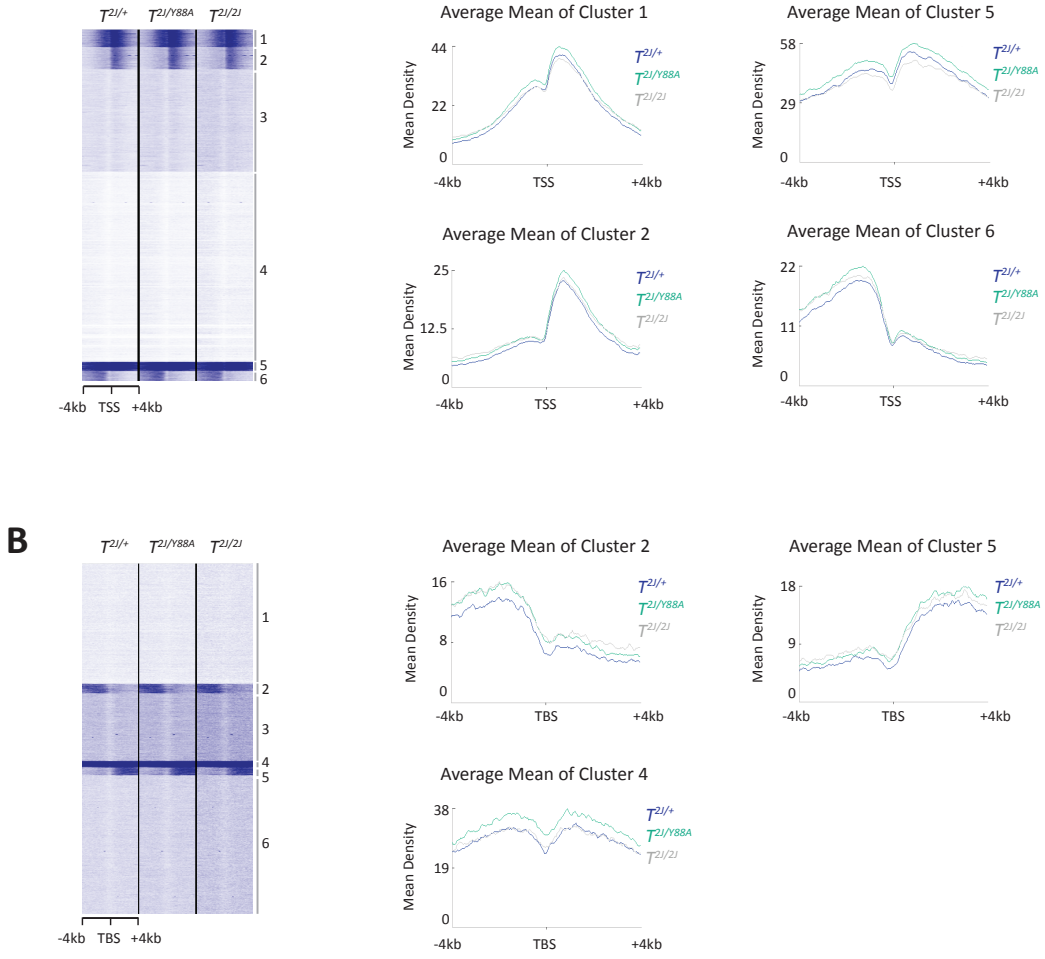


Figure 2.27: H3K27me3 profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control
Heat maps displaying enrichment of H3K27me3 at ±4kb of transcription start sites (A) and T binding sites (B) in mutants compared to control. Numbers to the right of the heat maps depict clusters. Plots depict the average mean density (tag/50bp) of H3K27me3 in clusters with higher overall levels of H3K27me3. TSS, transcription start site; TBS, T binding site.

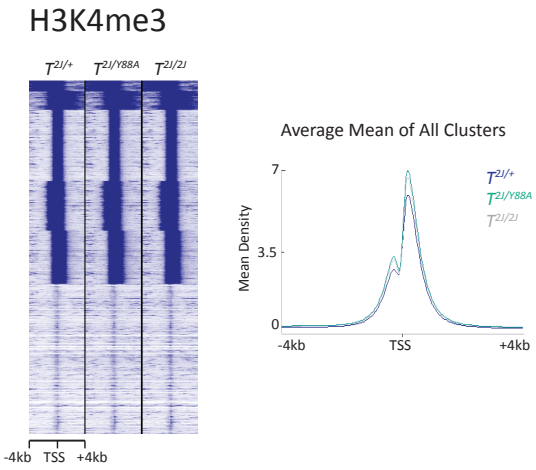


Figure 2.28: H3K4me3 profiles at TSS in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control
Heat map displaying enrichment of H3K4me3 at ±4kb of transcription start sites in mutants compared to control. Plot on the right depicts the average mean density (tag/50bp) of H3K4me3 in all clusters. TSS, transcription start site.

2.11.2 H3K4me1 in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutants

Next, I investigated the profiles of H3K4me1 in mutant $T^{2J/Y88A}$ or $T^{2J/2J}$ mesodermal cells compared to control. H3K4 monomethylation is a histone mark whose role in chromatin organization is not entirely clear. It has previously been shown that H3K4me1 is enriched at distal enhancer elements and marks active enhancers in combination with H3K27ac (Creyghton et al., 2010; Heintzman and Ren, 2009; Visel et al., 2009; Xi et al., 2007). Furthermore, H3K4me1 has been found at transcriptionally active promoters, not directly at transcription start sites, but extending further downstream than H3K4me3 (Barski et al., 2007). Recent studies have shown that H3K4me1 at promoters is associated with transcriptional repression in diverse cell types (Cheng et al., 2014) and can also act as a chromatin boundary by spatially restricting proteins that recognize H3K4me3 and promote transcription. Heat maps of H3K4me1 at ±4kb from both transcription start sites and T binding sites in the genome revealed an overall increase in H3K4me1 in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to control (Figure 2.29A). Plots of the average mean density of H3K4me1 showed that both mutants had higher densities of H3K4me1 across ±4kb of TSS and TBS. To determine if there was an

actual global increase in H3K4me1 in the mutants, or whether it was just an artifact of sequencing, as H3K4me1 was increased at all genomic positions investigated, a Western blot was performed using protein lysates from *T-mCherry*-positive cells. Western blotting with antibodies recognizing H3K4me1 revealed that there was no significant global increase in H3K4me1 in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to control (Figure 2.29B).

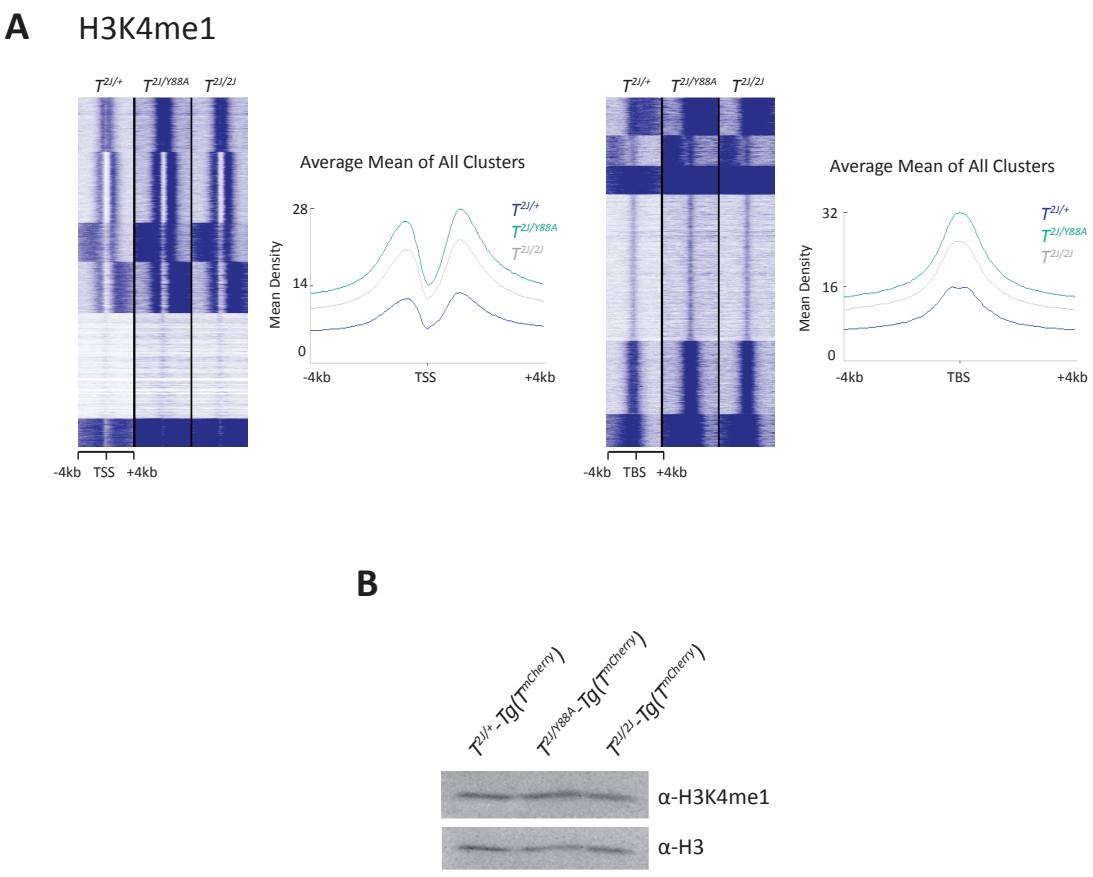


Figure 2.29: H3K4me1 profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control
(A) Heat maps displaying enrichment of H3K4me1 at ±4kb from transcription start sites and T binding sites in mutants compared to control. Plots depict the average mean density (tag/50bp) of H3K4me1 across all TSS and TBS. (B) Western blot from *T-mCherry*-positive sorted cells revealed no global increase in H3K4me1 in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants. α-H3 was used as a loading control. TSS, transcription start site; TBS, T binding site.

2.11.3 H3K27ac in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutants

Next, I examined the profiles of H3K27 acetylation in mutant $T^{2J/Y88A}$ and $T^{2J/2J}$ mesodermal cells compared to control. H3K27ac was previously shown to be enriched at the promoters of transcriptionally active genes and, more recently, in combination with H3K4me1 at active enhancer regions (Creyghton et al., 2010; Wang et al., 2008). There was no difference in H3K27ac at ±4kb from the TSS of any genes in the mutant compared to control mesodermal cells (Figure 2.30A). There was also no difference in H3K27ac levels at the TSS of dysregulated genes (Supplementary Figure S4). However, when examining H3K27ac at ±4kb from T binding sites throughout the genome, there was an apparent decrease in H3K27ac in mutant cells compared to control. Plots of the mean density of H3K27ac of all clusters and of clusters with higher levels of H3K27ac (Clusters 2, 3, 4, and 6 in Figure 2.30B) revealed that the levels of H3K27ac in both $T^{2J/Y88A}$ and $T^{2J/2J}$ cells compared to control were significantly decreased and, more importantly, this phenomenon was present at sites throughout the genome where T is binding. Due to loss of T binding to H3K27 demethylase activity published previously in the T^{Y88A} mutant (Miller et al., 2008), I also examined H3K27me3 at regions within clusters 2, 3, 4, and 6 to determine if there was an increase in H3K27me3 at regions where H3K27ac is lower, as these two marks are mutually exclusive. There was no significant increase in H3K27me3 at regions where lower H3K27ac was found in the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants when compared to control (Supplementary Figure S5), and loss of H3K27ac cannot be attributed to a failure of H3K27me3 to be removed.

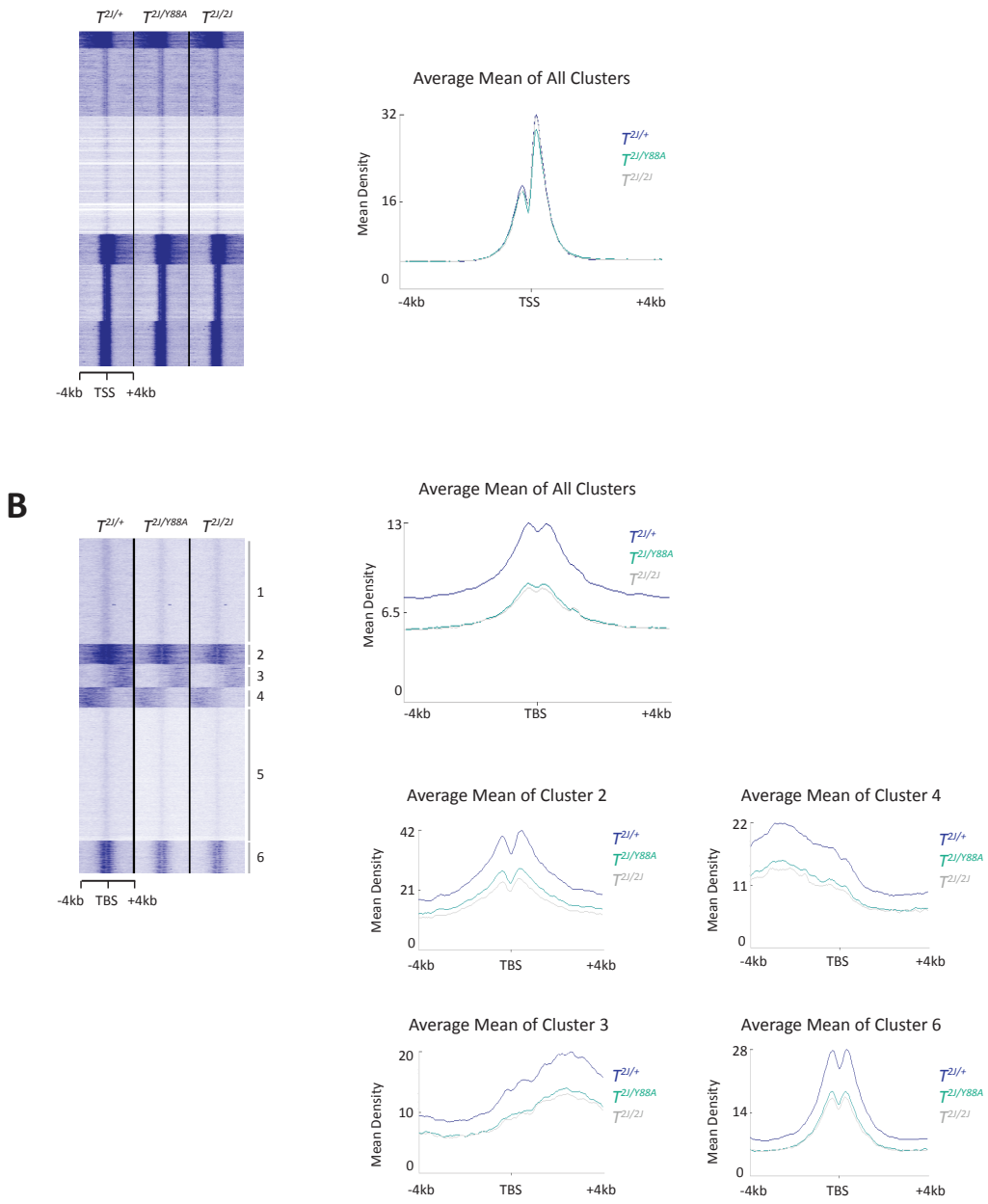


Figure 2.30: H3K27ac profiles in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants compared to $T^{2J/+}$ control
Heat maps displaying enrichment of H327ac at ±4kb from transcription start sites (A) and T binding sites (B) in mutants compared to control. Plots depict the average mean density (tag/50bp) of H3K27ac across all TSS and TBS, and across the TBS of clusters with higher H3K27ac levels (Clusters 2, 3, 4, and 6). TSS, transcription start site; TBS, T binding site.

A closer look was taken at the genomic regions in Cluster 2, 3, 4, and 6 with lower H3K27ac in mutants versus control. The majority of these regions were found in the promoter (-5kb/+2kb of the TSS) or within the bodies of genes (Figure 2.31A). Each region of H3K27ac falling within the promoter or body of a gene was assigned to that specific gene. The remaining 22% of H3K27ac regions were found in intergenic regions of the genome, and these regions were assigned to both the nearest upstream and downstream gene. Examples of H3K27ac tracks in genes with lower H3K27ac are shown in Figure 2.32. H3K27ac is found in the promoter of *Dll3*, and there are lower levels of H3K27ac in the mutant $T^{2J/Y88A}$ and $T^{2J/2J}$ mesodermal cells at the region where T is binding. H3K27ac in genic and intergenic regions of *Cited2* are lower at regions where T is binding, while intergenic regions of *Hand1* demonstrate a slight reduction in H3K27ac at sites of T binding. I also compared the genes with lower H3K27ac in mutant $T^{2J/Y88A}$ and $T^{2J/2J}$ mesodermal cells to dysregulated gene expression from *T-mCherry*-positive cells isolated from mutant embryos. There were a total of 85 dysregulated genes (51 downregulated and 34 upregulated) that were correlated to lower levels of H3K27ac in the mutant mesodermal cells (Figure 2.31B). This included known T targets, such as *Tbx6, Cyp26a1,* and *Snai1* (Evans et al., 2012; Garnett et al., 2009; Gentsch et al., 2013). Interestingly, in all of these examples, lower H3K27ac is found at regions where T is binding. This, coupled with the fact that a fraction of these genes are also dysregulated in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos, suggests that T regulates H3K27ac at a subset of regions where it is binding in the genome of mesodermal cells.

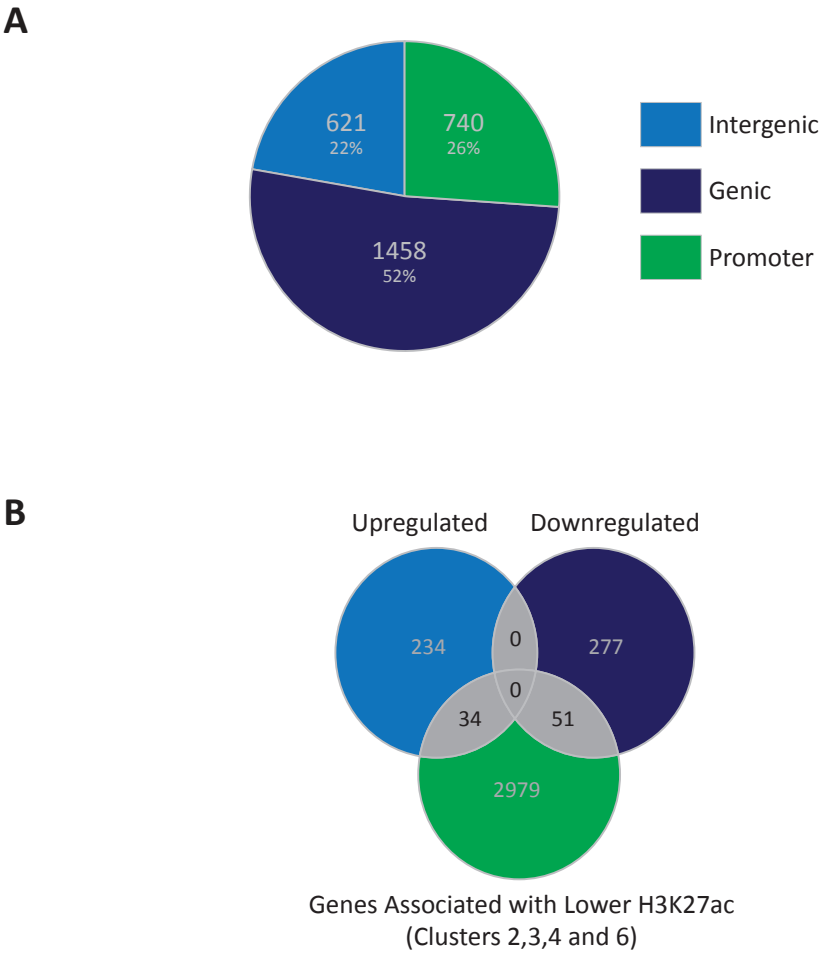


Figure 2.31: Analyzing the regions of lower H3K27ac in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant mesodermal cells

(A) Pie chart illustrating the distribution of regions with lower H3K27ac throughout the genome. (B) Correlation between dysregulated gene expression in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos and regions with lower H3K27ac.

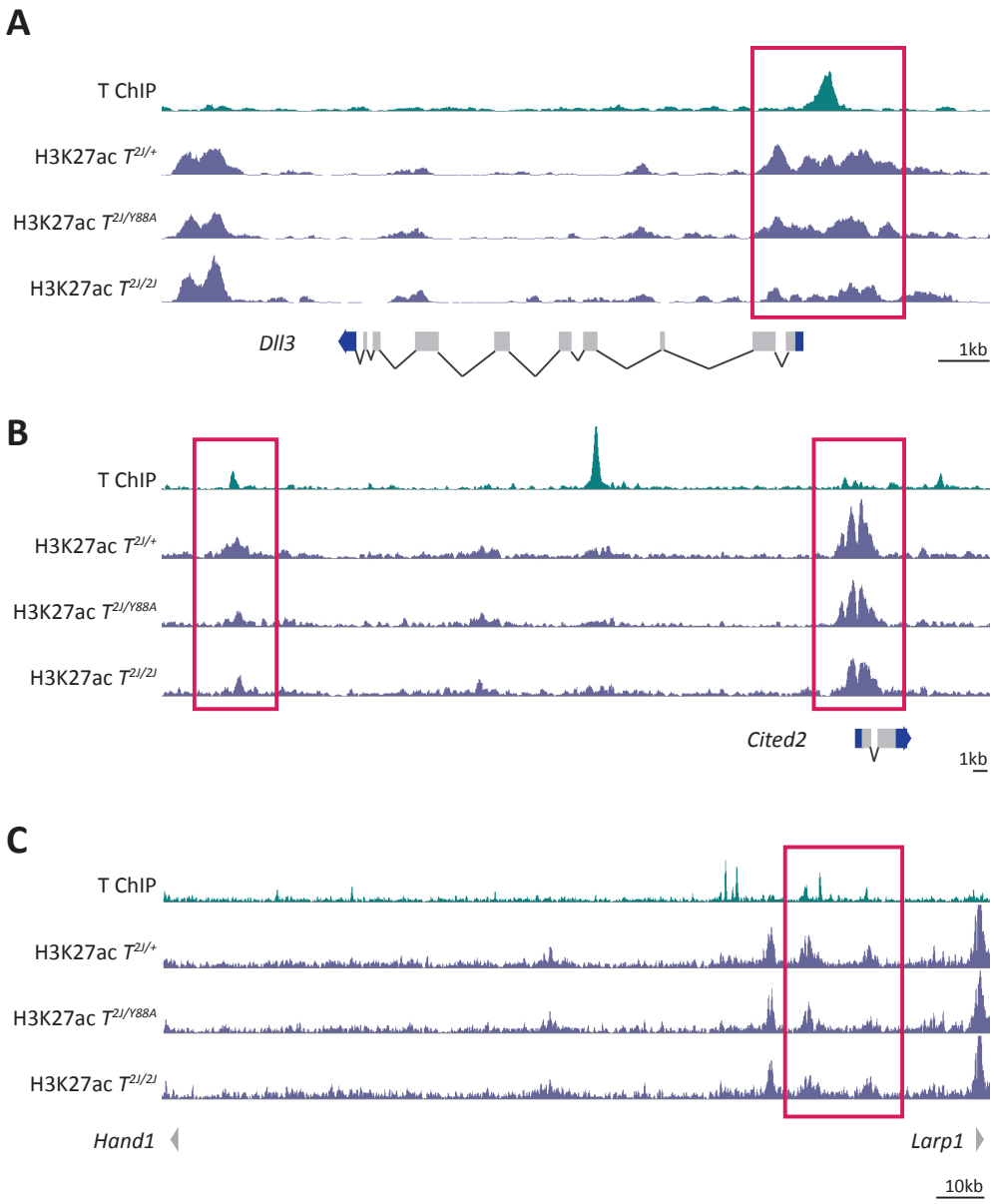


Figure 2.32: Regions of lower H3K27ac correspond to T binding in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants
H3K27ac (purple) and T (green) ChIP-Seq tracks from genes in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant mesodermal cells with lower H3K27ac at T binding sites. Three genes, *Dll3* (A), *Cited2* (B), and *Hand1* (C), are shown. Red boxes indicate the site of T binding and loss of H3K27ac in mutants compared to control. The y-axis indicates number of mapped reads, with all H3K27ac tracks displayed using the same scale.

GO term analysis was performed on the genes associated with regions containing lower H3K27ac in the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants to determine if there were any biological functions enriched within this group of genes. Interestingly, genes associated with lower H3K27ac in the promoter or genic regions in the mutants were enriched for functions pertaining to the cytoskeleton, including actin cytoskeleton organization, cytoskeleton protein binding, actin binding, and cell adhesion (Figure 2.33). Furthermore, there was also enrichment for functions such as embryonic morphogenesis, gastrulation, mesoderm formation, and mesoderm development. This included genes such as *FoxA2*, which has been previously shown to be essential for proper node and notochord formation (Ang and Rossantt, 1994; Weinstein et al., 1994), *Eomes*, an important regulator of mesendodermal cell fate (Arnold et al., 2008; Costello et al., 2011), *Tbx6*, a known regulator of paraxial mesoderm and somite formation (Chapman and Papaioannou, 1998), and several components of the canonical Wnt pathway, including *Wnt3, Lrp6,* and *Lef1*. Genes associated with lower H3K27ac in intergenic regions in the mutant were enriched for functions such as vasculature and blood vessel development. This, together with the dysregulated gene expression in mutant embryos, further strengthens the hypothesis that *T* plays a role in blood cell and blood vessel development. Other functions enriched in the genes associated with lower H3K27ac included regulation of transcription, embryonic morphogenesis, and actin filament organization (Figure 2.33). These analyses suggest that one function of *T* may be to recruit H3K27 acetylases to genes important for cytoskeletal rearrangement, gastrulation, and mesoderm formation.

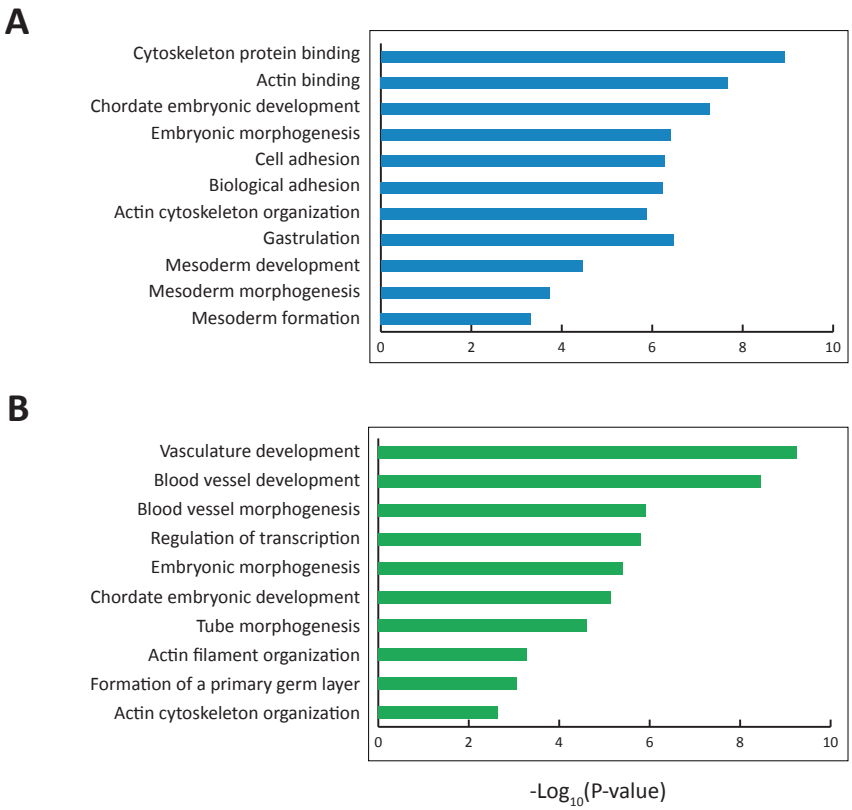


Figure 2.33: GO term enrichment analysis of genes associated with lower H3K27ac in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants

GO term analysis of genes associated with lower H3K27ac in promoter and genic regions (A), and intergenic regions (B) of $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants. Graphs depict the $-\text{Log}_{10}$ of the P-value for each GO term, as given by the DAVID bioinformatic tool.

2.11.4 Investigating Enhancer Regions in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutants

It was previously shown that enhancer elements can be defined by histone modifications, with H3K4me1 in combination with H3K27ac marking active enhancers, and H3K4me1 alone marking poised enhancers. Because of the reduction in H3K27ac at T binding sites throughout the genome of mutant $T^{2J/Y88A}$ and $T^{2J/2J}$, I examined H3K27ac at enhancer elements in mutants versus control mesodermal cells to determine if there were differences in active enhancer elements. Clustering of regions around T binding sites based on colocalization of H3K4me1 and H3K27ac revealed that while the apparent increase in H3K4me1 in mutants was visible, there was no significant decrease in H3K27ac at regions that colocalize with H3K4me1 (Figure 2.34). From these results, it

appears as if T does not regulate H3K27ac at enhancer elements where T is binding, as there is no decrease in H3K4me1/H3K27ac double positive enhancer regions around T binding sites.

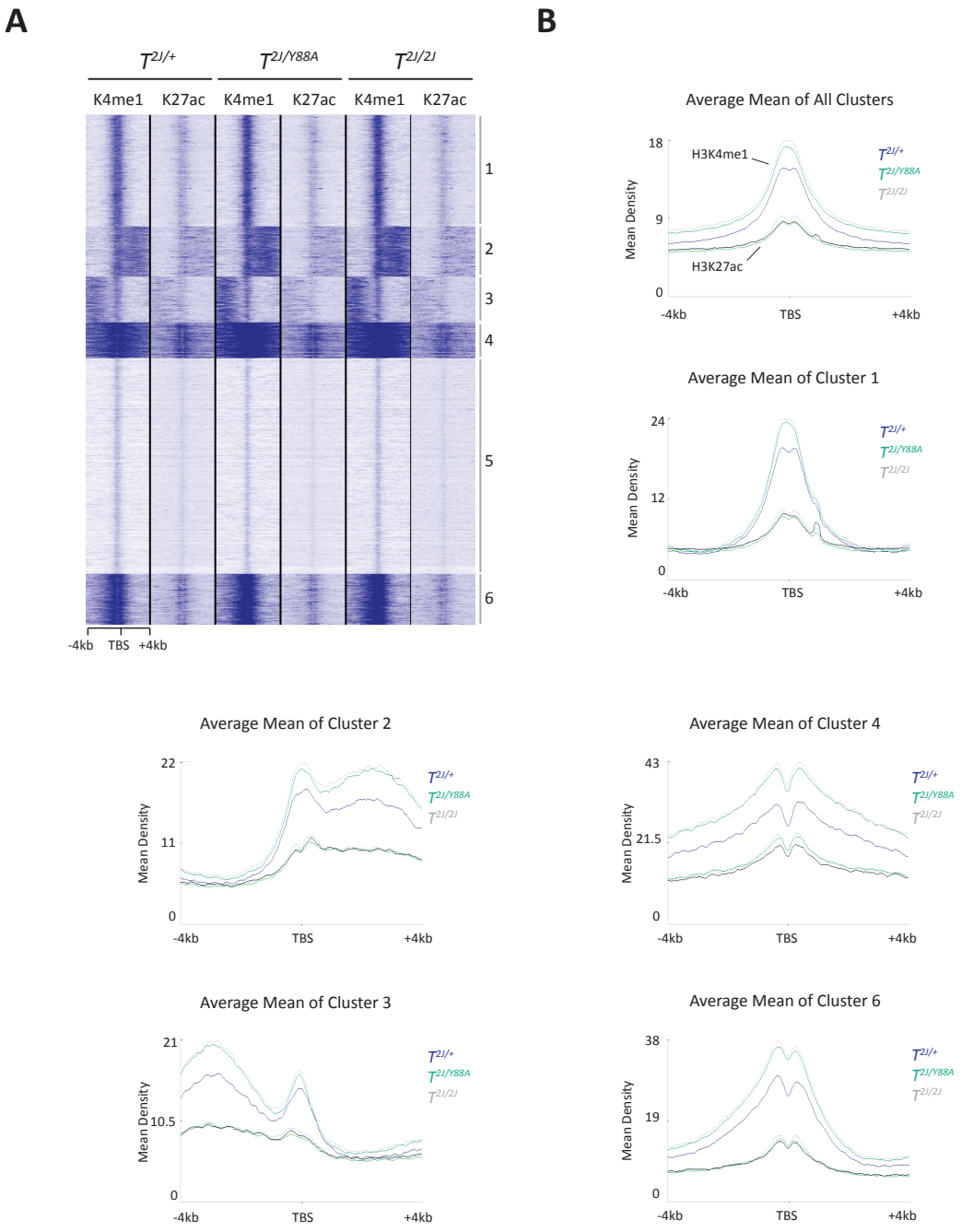


Figure 2.34: Comparison of active enhancer regions around T binding sites in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants

Heat map (A) and average mean density (tag/50bp) plots (B) of H3K4me1/H3K27ac positive enhancer regions. H3K4me1 is shown in the upper curve and H3K27ac in the lower curve. Average mean of cluster 5 was left out due to the low levels of both H3K4me1 and H3K27ac at these regions in the heat map. TBS, T binding site.

Chapter 3.

Discussion

The aim of this thesis was to determine the role of T in the control of chromatin state within early mesodermal cells. To this end, an amino acid substitution (Y88A), previously shown to disrupt the interaction between T and histone-modifying enzymes *in vitro*, was introduced into the endogenous *T* locus to determine the *in vivo* consequences of loss of this functional interaction. Embryos expressing the mutant T^{Y88A} allele resembled *T* knockout embryos, with a complete loss of trunk and tail formation, and were unable to maintain *T* expression. Furthermore, gene expression profiling and survey of the histone modification landscape revealed that the T^{Y88A} mutant allele is not simply a null allele, as it affects H3K27 acetylation at target loci throughout the genome. In addition, genetic evidence pointing to a likely role for T in hemangioblast commitment and erythrocyte development was discovered. Taken together, the results provide evidence that T plays a role in regulating the histone modification landscape in early mesodermal cells and that this function is necessary for proper mesodermal development during embryogenesis.

3.1 The Phenotypic Consequences of T^{Y88A} Mutation

Embryos harboring the T^{Y88A} mutant allele displayed phenotypes almost identical to those in $T^{2J/2J}$ knockout embryos, including a complete truncation of posterior mesoderm, absence of a notochord, kinked neural tube, and loss of somite formation (Figure 2.3A). Furthermore, $T^{2J/Y88A}$ embryos die between E9.5-E10 due to defects in allantois development. From gross morphology of mutant embryos alone, the T^{Y88A} mutant allele appeared to act as a null allele. However, when *T* expression was investigated in $T^{2J/Y88A}$ mutant embryos, it became clear that expression of *T* is normal in early gastrulating embryos from E6.5-E7.5. Only from E8.0 onwards, does *T* expression begin to decrease and by E9.5, expression of *T* is nearly absent from mutant embryos. The first question that arose from this result was whether the phenotypes seen in the $T^{2J/Y88A}$ mutant embryos were solely due to the loss of *T* expression from E8.0, or whether the point mutation was also disrupting a critical function of T. To address this question, I used a system previously developed in our lab to inducibly and ubiquitously knockdown *T* expression with the use of shRNA hairpins *in vivo* (Vidigal et al., 2010). Unfortunately, the results obtained were extremely variable due to the tetraploid complementation technique used to generate embryos from *KD4-T* ESCs (Figure 2.14), and the experiment as performed was not able to demonstrate whether the loss of *T* maintenance alone is the cause of the phenotype seen in T^{2JY88A} embryos. In order to perform this experiment in a more controlled manner, *ex vivo* culture of *KD4-T* embryos (Martin and Cockroft, 2008) to allow for administration of doxycycline at time points corresponding to loss of *T* in $T^{2J/Y88A}$ mutants, can be performed. *Ex vivo* culture has the advantage that approximation of the age of developing embryos is easier, allowing *T* knockdown at time points comparable to $T^{2J/Y88A}$ mutants. Furthermore, it would make it easier to trace *T* expression and understand the dynamics of *T* knockdown in *KD4-T* embryos.

The phenomenon of T-box factors regulating the maintenance of their own expression is not only limited to T. There is evidence that both TBX6 and EOMES regulate the maintenance of their own expression as well. Generation of the analogous Y→A mutation in TBX6 ($Tbx6^{Y137A}$) resulted in embryos displaying the *Tbx6* knockout

phenotype and the loss of *Tbx6* expression in a more spatially recognizable manner (Anna Dorothea Senft and Maria Theodora Melissari, Master's theses, Max Planck Institute for Molecular Genetics, Berlin). *Tbx6* is normally expressed in the caudal end of the embryo at E9.5, and continues to be expressed as these cells leave the caudal end and reside in the presomitic mesoderm (PSM, Chapman et al., 1996). In $Tbx6^{\Delta/Y137A}$ mutant embryos, *Tbx6* expression is limited to the caudal end of E9.5 embryos and expression in the PSM is absent. A likely interpretation is that initiation of *Tbx6* expression in the caudal end of the embryo is intact, and as these cells leave the caudal end to populate the PSM, they are no longer able to maintain *Tbx6* expression (unpublished observations, Maria Theodora Melissari). Interestingly, the fact that an analogous mutation in another T-box factor is able to generate a similar destabilization in expression, as that seen here with T, suggests that the tyrosine (Y) residue, which is conserved across multiple T-box family members, plays a key role in the self-maintenance of T-box factor expression. The question then becomes: how does the Y→A mutation disrupt this mechanism?

Insights into the mechanism of T-box factor self-maintenance came from studies of EOMES function in definitive endoderm development. EOMES has been shown to regulate core definitive endoderm differentiation genes, including itself, through interaction with KDM6B during differentiation of ESCs to definitive endoderm *in vitro* (Kartikasari et al., 2013). *Eomes* expression is first activated through promoter-enhancer loop formation induced by the cooperative action of TBX3 and KDM6B. Once this promoter-enhancer loop has been established, EOMES binds to and activates the expression of core definitive endodermal genes, and itself, through interaction with KDM6B, SMAD2, and transcriptional activation machinery (Kartikasari et al., 2013). Whether the same mechanism is occurring with *T* and *Tbx6* is not as clear. First, expression of *Kdm6b* in *T-mCherry*-positive mesodermal cells isolated from the embryo is rather low, as demonstrated by RNA-Seq data (data not shown), and levels of KDM6B protein are low in *in vitro* differentiated mesodermal P19 cells (Figure 2.23B). Whether KDM6A, which is more highly expressed in mesodermal cells, can compensate and perform similar functions as KDM6B is not known. Furthermore, physical and functional interaction between T and H3K27 demethylase proteins was demonstrated to be intact by co-immunoprecipitation studies in HEK293T and mesodermal-differentiated P19 cells

(Figures 2.22 and 2.23) and by minimal changes in genome-wide H3K27me3 in $T^{2J/Y88A}$ ESCs differentiated *in vitro* into mesoderm (Figure 2.27). Therefore, the loss of *T* maintenance cannot be attributed to loss of T-KDM interaction. Unfortunately, analysis of the histone modification landscape and chromatin architecture at the *T* locus in mesodermal differentiated ESCs was not possible due to the presence of the randomly integrated *T-mCherry* reporter transgene within the ESCs. The reporter contained 200kb upstream and downstream of the *T* gene and it could not be distinguished whether sequencing reads from the histone-modification ChIP-Seq derived from the endogenous *T* locus or the *T-mCherry* reporter, whose integration point in the genome is unknown. However, ChIP using histone mark-specific antibodies coupled to PCR with primers in the *T* regulatory region in cells that do not contain the reporter gene can be performed to determine whether loss of *T* expression can be explained by disruption of the normal chromatin environment in the $T^{2J/Y88A}$ mutant.

I was able to rule out that the loss of *T* expression can be attributed to loss of maintenance of the Wnt/Fgf/T regulatory loop. As described in the introduction, both the Wnt and FGF signaling pathways play a role in the initial activation and maintenance of *T* expression during mesodermal development and that T can positively regulate the expression of *Wnt* and *Fgf* genes. Loss of *T* expression in the $T^{2J/Y88A}$ mutant could occur if this positive regulatory feedback loop is disrupted. However, in the *in vitro* differentiation protocol of ESCs into mesodermal cells, recombinant WNT and FGF proteins are continuously provided in the medium throughout the differentiation protocol. Furthermore, the transcriptional activation of *T* expression is intact, as the percentage of cells that activate the *T-mCherry* transcriptional reporter in $T^{2J/+}$ control versus $T^{2J/Y88A}$ mutant cells is similar when measured by FACS, although the mRNA levels of *T* are not maintained (Supplementary Figure S3). This suggests that transcriptional activation of *T* occurs but post-transcriptional regulation of *T* is disrupted in the $T^{2J/Y88A}$ mutant.

MicroRNAs are small 20-24bp RNA molecules that have been shown to play a role in the post-transcriptional regulation of mRNA molecules, by either inducing degradation or inhibiting translation of target mRNA molecules (reviewed in Vidigal and Ventura, 2012). Interestingly, a number of microRNA promoters were found that were bound by T in *in vitro* differentiated mesodermal cells and were computationally

predicted to target the *T* 3'UTR (miRwalk, miRanda, see Dweep et al., 2011; John et al., 2004). These microRNAs may be targets of repression by T, to allow for maintenance of *T* expression in early mesodermal cells, and whose expression can downregulate *T* in mesodermal cells as they differentiate further. Unfortunately, the RNA-Seq analysis performed here was not suitable for detection of small RNAs and upregulation of these microRNAs in mutant $T^{2J/Y88A}$ mesodermal cells was not measurable. However, one microRNA bound by T with the ability to target the *T* transcript, miR-148a, is interesting as it has been shown to inhibit epithelial-to-mesenchymal transitions (EMT) by targeting *Rock1* and *Mmp7*, mediators of EMT and cell migration, in non-small cell lung cancer cells and gastric cancer tissue, respectively (Li et al., 2013; Sakamoto et al., 2014). Inspection of gene expression profiles in *T-mCherry*-positive cells from $T^{2J/Y88A}$ mutants revealed a decrease in *Rock1* expression (log2FC=-0.48) in $T^{2J/Y88A}$ mutants but not in $T^{2J/2J}$ mutants. Upregulation of *miR-148a* in $T^{2J/Y88A}$ mutant embryos may function to both downregulate *T* and inhibit EMT, a process which occurs as cells ingress through the primitive streak and which is disrupted in $T^{2J/Y88A}$ mutants. To investigate this, treatment of mutant $T^{2J/Y88A}$ ESCs undergoing *in vitro* differentiation into mesoderm with chemically-modified DNA molecules (locked nucleic acid, LNA) which are complementary to and inhibit the function of miR-148a, can determine whether inhibition of *miR-148a* leads to restoration of *T* expression.

3.2 Is Dimerization of T Effected in the T^{Y88A} Mutant?

The Y88A amino acid substitution lies next to M87, a residue which has previously been shown to be important for dimerization of T by assessment of the crystal structure of the T-box of T with DNA (Müller and Herrmann, 1997). Furthermore, mutation of the analogous residue (M86R) in T-PIT, a T-box factor important for pituitary gland development, resulted in congenital isolated ACTH deficiency and was shown to affect T-PIT homodimerization and heterodimerization with PITX1 (Vallette-Kasic et al., 2007). Thus, we considered that the T^{Y88A} mutation may disrupt dimerization of T, and lead to some of the mutant phenotypes seen. To investigate this, I generated mouse ESCs containing one wild-type copy of *T* and three copies of T^{Y88A}, as it was expected that

defects in dimerization of T would cause semi-dominant negative effects that may only be visible when multiple copies of the mutant allele are expressed. Transgenic *T*$^{+/Y88A}$*;Tg(T*^{Y88A}*neo)1Bgh* embryos appeared normal from E9.5-E10.5 with regards to somite, notochord, and posterior mesoderm development (Figure 2.8). However, at E12.5, 64% of embryos were necrotic, suggesting that between E10.5-E12.5, multiple copies of the *T*^{Y88A} allele were not compatible with normal embryonic development, although the cause of death has not yet been thoroughly investigated. Interestingly, expression of *T* at E9.5 was 17% of that in wild-type embryos. Therefore, the death of embryos at E12.5 could be due to a later loss of *T* expression, and not necessarily a defect in dimerization. Perhaps lower levels of *T* are not compatible with maintenance of proper allantoic fusion with the placenta, leading to embryonic death, as it is known that T plays an important role in the development of the allantois. Gene expression profiling also revealed a defect in erythrocyte development in the *T*$^{2J/Y88A}$ mutant embryos, which will be discussed in detail below. It is possible that defects in the ability of embryos to generate a fetal blood supply and blood vessels which support proper circulation can also lead to embryonic death at E12.5 in *T*$^{+/Y88A}$*;Tg(T*^{Y88A}*neo)1Bgh* embryos.

Whether the *T*^{Y88A} mutation is affecting dimerization is still not yet clear. From genome-wide T binding site data, it appears as if most T binding sites in the genome are palindromic, suggesting that T binds to DNA throughout the genome as a dimer (personal communication, Dr. Frederic Koch). This would argue that T dimerization would be quite important for T function. Co-immunoprecipitation experiments with two differently-tagged versions of wild-type and T^{Y88A} to this point have been inconclusive (data not shown). Furthermore, comparison of dysregulated gene expression in *T*$^{2J/Y88A}$ mutant embryos to a ranked list of T palindromic binding sites revealed there was no enrichment in the distribution of strength of T palindromic sites in dysregulated versus non-dysregulated genes (Figure 2.21), implying that dimerization is not effected in the *T*^{Y88A} mutant. However, insights into dimerization of T^{Y88A} still require experimental evidence in order to come to a sound conclusion.

3.3 Gene Expression Profiles in $T^{2J/Y88A}$ Compared to $T^{2J/2J}$ Knockout Embryos

3.3.1 Similarities in Gene Expression Profiles in $T^{2J/Y88A}$ Versus $T^{2J/2J}$ Knockout Embryos

Gene expression profiling of *T-mCherry*-positive cells from $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos revealed that around 40% of genes were commonly dysregulated between the two mutants. Many of these dysregulated genes included known *T* targets such as *Tbx6, Cyp26a1, Fgf4, Snai1,* and *Wnt3a* (Figure 2.17). The downregulation of targets such as *Wnt3a* and *Cyp26a1* most likely contributed to the loss of a mesodermal progenitor cell population, leading to the loss of posterior mesoderm seen in both $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos. Martin and Kimelman suggested that *T* is necessary for the maintenance, but not initial activation, of canonical Wnt signaling via a positive autoregulatory loop (Martin and Kimelman, 2008, 2010). Furthermore, transcriptional activation of *Cyp26a1*, an enzyme involved in the degradation of retinoic acid, is necessary to antagonize mesodermal differentiation, as high levels of retinoic acid induces differentiation of PSM cells into somites through the repression of *Fgf8* (Diez et al., 2003; Moreno and Kintner, 2004; Vermot and Pourquie, 2005). Downregulation of *Snai1* in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos may also contribute to a loss of EMT in cells ingressing through the primitive streak during gastrulation, resulting in an accumulation of mesodermal progenitor cells within the primitive streak. *Snai1* is known to be a potent regulator of EMT by its ability to transcriptionally repress expression of *E-cadherin* (gene symbol *Cdh1*) whose protein product is necessary for formation of adherens junctions in epithelial cells, and whose expression must be downregulated for cells to adopt a mesenchymal, motile fate (Carver et al., 2001). *Tbx6* and *Fgf4* have both previously been shown to be important for PSM formation and differentiation of PSM into somites. *Tbx6* knockout mice form irregular somites in the neck region, and posteriorly, ectopic neural tubes are formed at the expense of somites (Chapman and Papaioannou, 1998). This was shown to occur through the repression of *Sox2* at its N1 enhancer by TBX6 in paraxial mesodermal cells. In *Tbx6* knockout mice, the N1 enhancer is not turned off and *Sox2* expression persists in PSM cells leading to

subsequent neural tube formation. Deletion of the N1 enhancer of *Sox2* in the *Tbx6* knockout background rescued the formation of ectopic neural tubes, and forced expression of *Sox2* in paraxial mesoderm cells was sufficient to form ectopic neural tissue at the expense of somites (Takemoto et al., 2011). Taken together, this provides solid evidence that *Tbx6* is essential for PSM specification and the formation of somites, and its downregulation of $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos results in the failure of somite formation.

Both the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants display a loss of notochord maintenance which is also likely to affect surrounding neural tube and somite patterning in mutant embryos. *Shh* is downregulated in both mutants, and has previously been shown to play an important role in notochord maintenance. *Shh* knockout mice display initial notochord formation in caudal portions of the developing embryo, but the maintenance of the notochord is disrupted in rostral portions of the embryo as differentiation continues (Chiang et al., 1996). This rostral-to-caudal progressive loss of notochord tissue demonstrates that *Shh* is necessary for the maintenance of the notochord during embryonic development. In *T* mutants, formation of a notochordal precursor-like structure is formed, but does not differentiate into a proper notochord and eventually undergoes cell death (Chesley, 1935; Pennimpede et al., 2012; Rashbass et al., 1991; Yanagisawa, 1990). Thus, it is likely that the inability to activate *Shh* expression in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants also results in improper notochord differentiation and failure to develop and maintain a functional notochord. Furthermore, *Shh* knockout embryos also display defects in dorsoventral patterning of the neural tube and patterning of gene expression in the developing somites (Chiang et al., 1996). This may also help to explain the defects in neural tube development of the mutant $T^{2J/Y88A}$ and $T^{2J/2J}$ embryos.

The $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants die at E9.5-E10 due to the failure of the allantois, a precursor of the umbilical cord and vascular portions of the fetal placenta (Inman and Downs, 2007), to fuse and establish a connection with the placenta. Interestingly, the allantois is also one of three embryonic sites known to undergo vasculogenesis, which is the formation of blood vessels without the presence of pre-existing vessels. This occurs through the *de novo* differentiation of endothelial cells from mesenchymal cells, from which hematopoietic stem cells can later arise (Baron and Fraser, 2005). From the gene expression profiling data, *Tbx4* appears to be a target of T in early mesodermal cells as

its expression is downregulated in mutant $T^{2J/Y88A}$ and $T^{2J/2J}$ embryos. *Tbx4* has previously been shown to be an important regulator of allantois formation, as *Tbx4* knockout embryos die at E10.5 as a result of the failure of allantoic fusion to the placenta (Naiche and Papaioannou, 2003), much like the *T* mutant embryos described here. Furthermore, the allantois of *Tbx4* knockout mice is stunted in growth, and although clumps of PECAM-positive endothelial progenitor cells form within the mutant allantois, there is a failure of these cells to elongate into vessels and form a central vessel (Naiche and Papaioannou, 2003). Interestingly, the same lab went on to show that the *Tbx4* vessel phenotype can be rescued by the addition of canonical Wnt agonists, revealing a downstream role of the canonical Wnt pathway in vessel formation within the allantois (Arora et al., 2012). The downregulation of *Tbx4* and multiple components of the canonical Wnt signaling pathway in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants may help to explain the cause of death of these mutant embryos.

3.3.2 A Role for *T* in Hemangioblast Commitment and Erythrocyte Development

While there are many common dysregulated genes in the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos, many of which are likely to explain the shared phenotypes seen in both mutants, the majority of dysregulated genes was not shared between the two mutants. Over 60% of dysregulated genes were unique to each mutant, providing evidence that the T^{Y88A} allele is not simply a null allele, and is disrupting a critical function of *T* in the control of histone acetylation and chromatin state. When examining gene ontology terms associated with dysregulated genes in the $T^{2J/Y88A}$ mutant embryos, one of the most significant biological functions enriched was erythrocyte differentiation and development. During embryonic development, there are two independent waves of erythropoiesis; primitive erythropoiesis, which occurs within the yolk sac of developing embryos at E7.0, and definitive erythropoiesis occurring as hematopoietic progenitors from the yolk sac colonize the fetal liver from E10.0 onwards. During primitive erythropoiesis, the first blood cells in the embryo arise within “blood islands” in the extraembryonic mesoderm of the yolk sac (Haar and Ackerman, 1971), derived from the posterior region of the primitive streak. The emergence of primitive erythroid cells and a vascular network derived from endothelial cells within close proximity in the blood

islands suggested a common progenitor giving rise to both cell types, the hemangioblast (Sabin, 1920). The presence of a hemangioblast precursor cell was validated in 1998, when Choi and colleagues generated blast cell colonies from embryoid bodies and cultured them in medium containing cytokines known to support both hematopoietic and endothelial cell development. They demonstrated that both hematopoietic cells with the potential to give rise to multiple lineages and endothelial cells expressing markers specific for the endothelial cell lineage, *Pecam-1* and *Flk-1,* arose from a common precursor cell (Choi et al., 1998). The same group went on to show that hemangioblast commitment is initiated in the posterior primitive streak of the late-streak/early neural plate stage mouse embryo. Furthermore, they showed that the hemangioblast expresses both *T* and *Flk-1* and that $T^{+}/Flk\text{-}1^{+}$ cells isolated from the embryo have the ability to differentiate into hematopoietic precursor and endothelial cells, arising from a common progenitor (Huber et al., 2004). Lastly, they reported that the hemangioblast cells migrate from the primitive streak region into the presumptive yolk sac, where they undergo relatively fast differentiation into either hematopoietic cells, endothelial cells, or vascular smooth muscle cells, as there was only a very small portion of $T^{+}/Flk\text{-}1^{+}$ cells within the yolk sac.

The molecular evidence supporting the concept of the hemangioblast comes from several studies which confirm the co-expression of hematopoietic and endothelial-specific genes in mouse and human cells. Genes such as *Flk-1 (Kdr), Scl* (*Tal1*), *Lmo-2, Gata2, Epo-R, Tie1, Tie2,* and *Pecam1* were shown to be expressed in both hematopoietic and endothelial cell lineages (Anagnostou et al., 1994; Hashiyama et al., 1996; Kallianpur et al., 1994; Watt et al., 1995). *Flk-1*-null embryos fail to form blood islands at E7.5 and die soon after due to a lack of blood and endothelial cells (Shalaby et al., 1995). Later studies demonstrated that *Flk-1*-null cells still possessed hematopoietic and endothelial potential *in vitro,* and concluded that the loss of blood island formation in *Flk-1*-null embryos results in a failure of hemangioblast cells to migrate to their proper positions within the yolk sac (Schuh et al., 1999). Both *Tal1* and *Lmo2* were also shown to play an essential role in differentiation of the hemangioblast, as knockout embryos displayed a complete lack of hematopoietic cells within the yolk sac (Robb et al., 1995; Shivdasani et al., 1995; Warren et al., 1994). Furthermore, studies of chimeric mutant *Tal1* and *Lmo2* embryos revealed a role for these genes in vascular development

(Visvader et al., 1998; Yamada et al., 2000). Interestingly, the expression of these and other hemangioblast genes are also found to be downregulated in *T-mCherry*-positive cells from the $T^{2J/Y88A}$ mutant embryo (Table 1), providing genetic evidence that *T* is involved in the regulation of hemangioblast commitment and development.

T protein was previously shown to be localized to the primitive streak and to regions of extraembryonic mesoderm closest to the streak. As extraembryonic mesodermal cells become more displaced from the streak, the levels of T protein are downregulated (Inman and Downs, 2006a). From what is known about the hemangioblast from previous studies, it appears as if commitment occurs in the posterior primitive streak and proximal regions of the extraembryonic mesoderm where *T* is expressed. These cells then migrate to more distal regions of the yolk sac to differentiate into hematopoietic and endothelial cells of the blood islands within the yolk sac. From the data above, it appears as if genes which regulate hemangioblast migration and commitment are dysregulated in the $T^{2J/Y88A}$ mutant embryos at TS12 (E8.25) to TS13 (E8.5), suggesting that normal *T* function is necessary for proper hemangioblast development. It will be interesting to investigate the function of *T* in hemangioblast commitment and development *in vivo*. For example, the yolk sacs from $T^{2J/Y88A}$ mutant embryos can be examined for the formation of blood islands with primitive erythrocytes and endothelial cell types. Furthermore, it would be interesting to isolate $T^{+}/Flk1^{+}$ cells from $T^{2J/Y88A}$ mutant embryos and determine their potential to differentiate into the hematopoietic and endothelial cell lineages *in vitro*.

In addition to dysregulated hemangioblast genes, genes involved in downstream erythrocyte differentiation and maturation were also dysregulated in the $T^{2J/Y88A}$ mutant. These included known regulators of erythrocyte differentiation, including *Klf1*. *Klf1* was previously shown to be an important regulator of both primitive and definitive erythropoiesis, as *Klf1* knockout mice display β-thalassemia and die at E14.5 (Nuez et al., 1995; Perkins et al., 1995). *Klf1* was also shown to be important for the expression of several erythroid-specific genes, including adult and embryonic globins, heme biosynthetic enzymes, proteins that stabilize α-hemoglobin, and proteins necessary for proper cytoskeletal structure in erythrocytes (Basu et al., 2007; Hodge et al., 2006; Nilson et al., 2006). Components necessary for the proper architecture of the erythrocyte cytoskeleton and therefore proper blood flow, including *Ank1* (Rank et al.,

2009), were also dysregulated in $T^{2J/Y88A}$ mutants. Furthermore, subunits of the hemoglobin protein, including *Hba-b2* and *Hba-x*, and enzymes involved in the heme biosynthesis pathway, *Alas2*, were also dysregulated in $T^{2J/Y88A}$ mutants. This suggests that proper *T* function is not only necessary for hemangioblast commitment, but also further development and differentiation into the erythrocyte lineage.

Another line of evidence that normal *T* function is necessary for hemangioblast commitment and development comes from data obtained from H3K27ac ChIP-Seq from *T-mCherry*-positive mesodermal cells from $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants. In both mutants, 22% of regions with lower H3K27ac at T binding sites were found within intergenic regions (Figure 2.31A). When GO term analysis was performed on the nearest upstream and downstream genes from these intergenic regions, the most significantly enriched terms were 'vasculature development', and 'blood vessel development and morphogenesis' (Figure 2.33B). These genes were not dysregulated in mutant $T^{2J/Y88A}$ or $T^{2J/2J}$ embryos, but it is plausible that *T* creates the proper chromatin environment at genes necessary for proper blood island development and differentiation within hemangioblast progenitor cells. As development and differentiation of hemangioblast to erythropoietic and endothelial cells within blood islands occurs, *T* is no longer expressed, however its effects on the chromatin environment may still be present, creating a permissive chromatin state for the transcriptional activation of genes necessary for blood vessel development. Some of these genes included *Hand1*, which was previously shown to regulate extraembryonic vasculature development in the yolk sac (Morikawa and Cserjesi, 2004). *Hand1* knockout embryos displayed accumulation of hematopoietic cells between the yolk sac and amnion due to defects in proper vasculogenesis and defects in smooth muscle precursor cell recruitment to endothelial cells before differentiation (Morikawa and Cserjesi, 2004). *Notch1* was also associated with T binding and lower H3K27ac and has previously been shown to play an important role in embryonic angiogenesis. Conditional deletion of *Notch1* from endothelial cells in mice, using a *Tie2-Cre* driver, resulted in embryonic lethality at E10.5 due to vascular defects in the placenta, yolk sac, and embryo proper. *Notch1* conditional deletion from endothelial cells resulted in the formation of an initial primary vascular plexus within the yolk sac, but defects in this structure do not allow it to differentiate into a secondary network of well-organized blood vessels (Limbourg et al., 2005). Interestingly, other

genes such as *Foxo1, Prrx1, Cited2,* and *Sox18*, were associated with lower H3K27ac in intergenic regions and have all been shown to be involved in development of blood vessels (Downes et al., 2009; Ihida-Stansbury et al., 2004; MacDonald et al., 2013; Sengupta et al., 2012). However, their role in vasculogenesis in the embryonic yolk sac has not yet been determined.

Two previous studies demonstrated a functional role for T in hematopoiesis and yolk sac vasculogenesis and blood island formation. First, *in vitro* differentiation of human ESCs into hematopoietic cell types with siRNA-induced downregulation of *Brachyury* resulted in a decrease in $CD45^+$ hematopoietic cells (Cerdan et al., 2012). Second, studies on cultured yolk sacs from T^c/T^c mutant embryos revealed the development of blood island-type structures within the yolk sac, however, they were not packaged inside a continuous *Flk-1*-positive endothelium and presumptive hematopoietic precursors within the *Flk-1*-positive endothelium displayed evidence of pyknosis (Inman and Downs, 2006b). This suggests a role for T in the development of endothelial and hematopoietic progenitor cells within the blood islands of the yolk sac, however, molecular mechanisms as to how T functions to control blood island development have not yet been investigated.

Taken together, the above data points to a role for the involvement of T in hemangioblast commitment and differentiation into erythrocyte and endothelial cell types in the extraembryonic mesoderm of the yolk sac. Gene expression profiling revealed dysregulation of multiple genes normally expressed in hemangioblast cells and in genes important for erythrocyte differentiation and development in the $T^{2J/Y88A}$ mutant. Furthermore, levels of H3K27ac at T binding sites were lower in intergenic regions associated with genes known to play a role in embryonic vasculature development, suggesting that *T* regulates both transcriptional activation and modulation of the chromatin environment at genes involved in blood cell and vessel development. Interestingly, some dysregulation occurs at genes expressed in the hemangioblast in the $T^{2J/2J}$ knockout, but to a much lesser extent, and no dysregulation occurs in genes important for erythrocyte differentiation and development. This suggests that another factor may be able to partially compensate for *T* function in hemangioblast and erythrocyte development. *Eomes*, another member of the T-box family of transcription factors, is also expressed in the primitive streak of developing mouse embryos and is

known to bind a similar motif as T. Perhaps, EOMES is able to compensate for the regulation of hemangioblast commitment and erythrocyte development in $T^{2J/2J}$ knockout embryos, but not in the $T^{2J/Y88A}$, as T^{Y88A} mRNA and protein are still present.

3.4 The Histone Modification Landscape in $T^{2J/Y88A}$ and $T^{2J/2J}$ Mutant Mesodermal Cells

In order to assay the histone modification landscape in $T^{2J/Y88A}$ and $T^{2J/2J}$ mutants, I turned to an *in vitro* differentiation protocol established in our lab to generate large amounts of mesodermal cells from ESCs, marked by *T-mCherry* (Figure 2.24A). I performed ChIP-Seq using antibodies directed against H3K4me3, H3K4me1, H3K27me3, and H3K27ac from *T-mCherry*-positive mesodermal cells. Data from other labs has previously shown that T-box transcription factors, including T, interact directly with components of the Trithorax/MLL complex, namely RBBP5, and functionally interact with H3K27 demethylase proteins, KDM6A and KDM6B *in vitro*. Furthermore, it was also shown that the Y→A mutation within the Tbox1 pocket abrogates the interaction of T-BET with both RBBP5 and with functional H3K27 KDM activity, without disrupting the ability of T-box factors to activate transcription in a non-chromatin context (Miller et al., 2008). From the experiments performed here, I found no significant differences in H3K4me3 or H3K27me3 in mesodermal cells from $T^{2J/Y88A}$ or $T^{2J/2J}$ mutants at genome-wide transcription start sites, T binding sites, or the transcription start sites of dysregulated genes (Figures 2.27, 2.28, and S4). This argues that the Y88 residue in T is not necessary for binding to the MLL complexes or H3K27 KDMs in mesodermal cells, which was also confirmed by co-immunoprecipitation data. CoIP data from overexpressed proteins in HEK cells demonstrated that mutant T^{Y88A} actually has the ability to bind more strongly to KDM6A and KDM6B than wild-type T (Figure 2.22B). However, in a mesodermal setting, the binding of wild-type and mutant Y88A T to KDM6A was similar and an interaction with MLL complex components RBBP5 and WDR5 could not be detected (Figure 2.23B). This was further corroborated by a similar interaction of both wild-type and Y88A mutant T with BRG1. Miller and colleagues previously showed that the interaction between T-box factor, T-BET, and BRG1 was

mediated by KDM6A or KDM6B, since siRNA knockdown of KDM6B resulted in a loss of interaction between T-BET and BRG1 accompanied by a concomitant loss of target gene expression and accessibility of chromatin at target gene loci (Miller et al., 2010). Because there was no loss of BRG1 interaction with mutant T^{Y88A} protein, and no significant changes in H3K27me3 genome-wide in $T^{2J/Y88A}$ mutant mesodermal cells, it was concluded that the Y88A mutation in T does not result in the loss of functional interaction with H3K27 KDM proteins. While this is in contrast to previously published data, experiments carried out by Miller and colleagues were performed by over-expressing *T* in EL4 cells, a mouse lymphoma cell line (Miller et al., 2008). The results described here suggest that this interaction is likely context-dependent.

The only difference found in the histone modification landscape of $T^{2J/Y88A}$ mutant mesodermal cells was a decrease in H3K27ac at T binding sites throughout the genome (Figure 2.30B). This decrease was specific to T binding sites, as H3K27ac at genome-wide transcription start sites and at dysregulated genes was normal (Figure 2.30A and S4). Interestingly, this decrease in H3K27ac was not due to the presence of mutually exclusive H3K27me3 at T binding sites throughout the genome (Supplementary Figure S5). CBP and p300 are the two proteins responsible for H3K27 acetylation *in vivo*, and are both essential for normal embryonic development (Tanaka et al., 2000; Yao et al., 1998). CBP/p300 are thought to act as transcriptional activators, although the mechanism by which these proteins activate gene expression is not entirely clear. Histone acetylation itself is thought to facilitate transcription through relaxation of DNA/histone interaction by neutralization of positive charge of lysine side-chains, recruitment of cofactors via bromodomain-containing proteins that recognize and bind acetylation, and affecting nearby histone tail residues (Bedford and Brindle, 2012). CBP/p300 are thought to facilitate transcriptional activation by interaction between transcription factors with general transcriptional machinery, by acting as a scaffold to promote protein-protein or protein-DNA interactions, or by acetylating non-histone proteins and affecting their function (Black et al., 2006; Chan and La Thangue, 2001). There is also evidence that CBP/p300 bridge transcription factor-bound enhancer regions with general transcriptional machinery at the promoters of genes (Heintzman et al., 2007; Visel et al., 2009). This is substantiated by the observation that CBP/p300 have

the ability to directly or functionally interact with over 400 different proteins (Bedford et al., 2010).

The *Drosophila* ortholog of CBP has been previously shown to bind both KDM6A (UTX) and BRM, the catalytic subunit of the SWI/SNF nucleosome remodeling complex (Tie et al., 2012b), functionally coupling the removal of repressive H3K27me3 with activating H3K27ac and nucleosomal remodeling. Moreover, *Cbp* and *Kdm6a* knockout mice both display defects in primitive and definitive hematopoiesis (Kung et al., 2000; Shpargel et al., 2012; Tanaka et al., 2000; Wang et al., 2012), a function that is also disrupted in $T^{2J/Y88A}$ mutant embryos based on dysregulated gene expression and lower levels of histone acetylation at regulators of blood vessel development. This suggests that KDM6A, CBP, and BRG1 may be recruited as a protein complex to target genes by T, and that this recruitment is necessary for proper transcriptional activation of genes necessary for hemangioblast commitment, and further erythrocyte and blood vessel development. Using T^{Y88A} protein in co-IP assays, binding of KDM6A and BRG1 is normal (Figure 2.23). Furthermore, H3K27me3 in $T^{2J/Y88A}$ mesodermal cells was normal when compared to control $T^{2J/+}$ cells, suggesting that the mutant allele does not disrupt the ability of T to recruit H3K27 KDMs to target loci (Figure 2.27). Although T and KDM6A can physically bind to each other and lead to normal demethylation of H3K27me3 in mesodermal cells, we cannot rule out that KDM6A is performing functions in addition to H3K27 demethylation that are disrupted in the $T^{2J/Y88A}$ mutant. For example, interaction between mutant T^{Y88A} protein and KDM6A may disrupt interaction with CBP, leading to the reduction in H3K27ac observed at T binding sites in $T^{2J/Y88A}$ mutant mesodermal cells. It will be interesting to investigate whether the interaction between T^{Y88A} and CBP is disrupted, either by co-immunoprecipitation studies or by performing ChIP in *T-mCherry*-positive mesodermal cells using antibodies against CBP, to see if there is a loss of CBP binding at genomic regions that correspond to lower H3K27ac in the $T^{2J/Y88A}$ mutant. Perhaps the loss of T-CBP binding contributes to the loss of activation of genes important for hemangioblast and erythrocyte development.

3.5 A Model for T Function

The data from this thesis were used to generate a working model of T function in mesodermal specification and development. Normal T function is necessary for mesodermal progenitor specification, and for paraxial, axial, and blood cell mesoderm development (Figure 3.1A). When T is not present ($T^{2J/2J}$), axial and paraxial mesoderm development are disrupted as evidenced by the loss of the notochord, defects in somite formation, and loss of posterior mesoderm. However, dysregulation of the expression of hemangioblast- and erythrocyte-specific genes does not occur, most likely due to the ability of another factor (perhaps EOMES) to compensate (Figure 3.1C). *Eomes* encodes for a member of the T-box family of transcription factors and binds to a similar T-box motif as T. It is expressed in the posterior primitive streak of early- to mid-streak stage mouse embryos and extends anteriorly as development proceeds. Its expression is restricted to the anterior primitive streak in the mid-late streak stage embryo, and in the late-streak/early head-fold stage embryo is restricted to the chorion and is no longer present in the primitive streak prior to morphological node formation (Figure 3.1B, Ciruna and Rossant, 1999). From this pattern of expression, EOMES may be able to compensate for early hemangioblast commitment in $T^{2J/2J}$ knockout embryos, as it is expressed in the posterior primitive streak at the time commitment occurs. However, EOMES cannot compensate for later axial and paraxial mesoderm development as it is no longer expressed in the proper domains during the commitment of these progenitors. In $T^{2J/Y88A}$ mutant embryos, modified *T* is still expressed, and EOMES is thus unable to compensate at sites of T binding in genes important for hemangioblast commitment. However, the T^{Y88A} mutant allele is not functional, and proper hemangioblast commitment does not occur (Figure 3.1D).

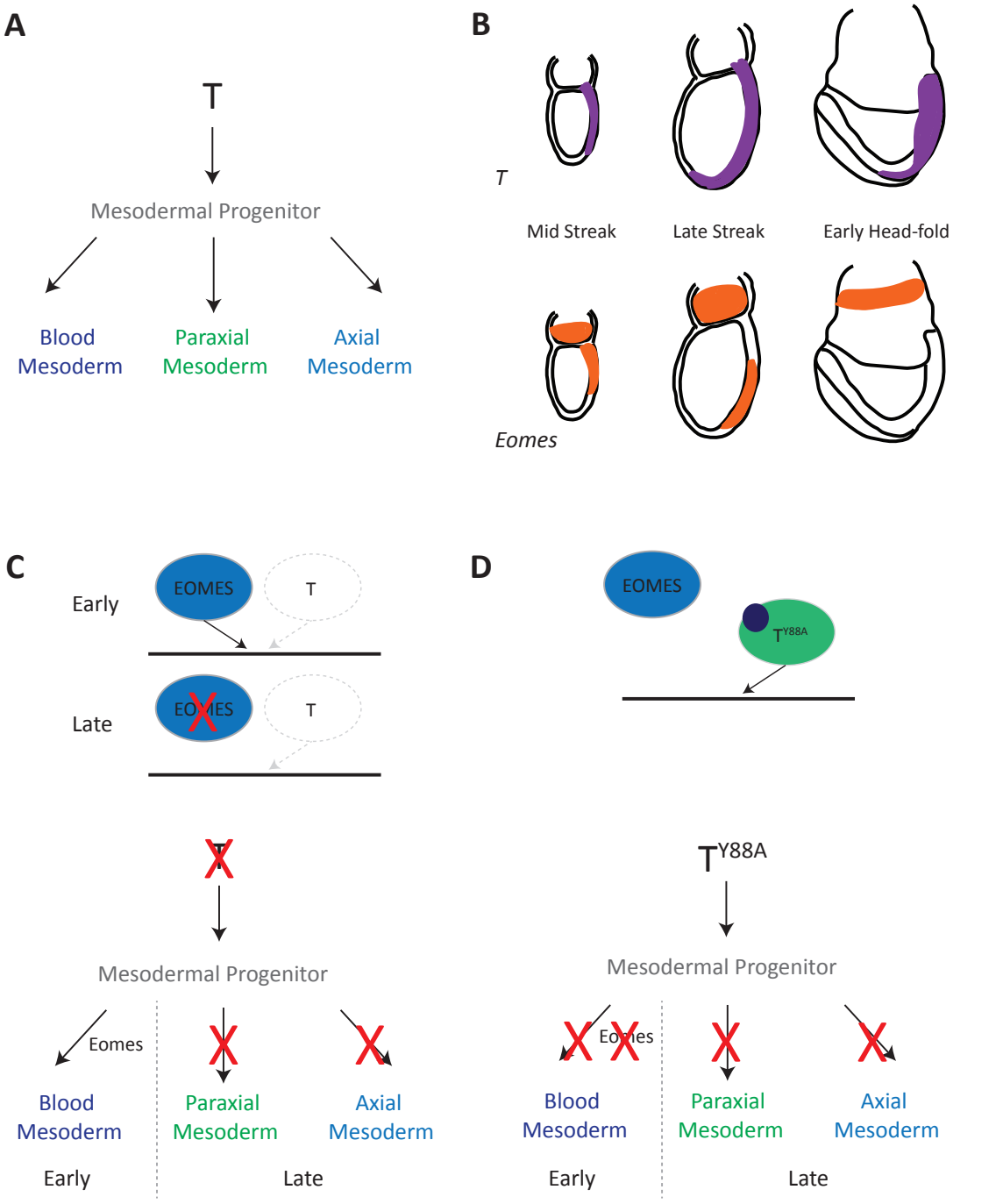


Figure 3.1: Model of T function during mesodermal differentiation
(A) T functions to specify mesodermal progenitor cells and promote further differentiation into blood cell, paraxial, and axial mesoderm derivatives. (B) Schematic illustrating *T* and *Eomes* expression in mouse embryos during mid-streak to early head-fold stages. (C) In $T^{2J/2J}$ knockout embryos, EOMES may be able to compensate for early blood cell mesoderm development, but not for later paraxial and axial mesoderm specification. (D) In $T^{2J/Y88A}$ mutant embryos, EOMES cannot compensate due to the presence of mutant T^{Y88A} protein in early embryos.

Chapter 4.

Materials and Methods

4.1 BAC Recombineering and Plasmid Generation

All standard DNA manipulation techniques were performed according to (Sambrook and Russell, 2001). These included PCR amplification of DNA, restriction enzyme digest of DNA, separation and visualization of DNA by agarose gel electrophoresis with ethidium bromide stain, generation of chemical- and electro-competent *E. coli*, transformation of bacterial strains, and directional cloning.

The following techniques were performed according to manufacturer's instructions: isolation of plasmid DNA from bacterial cultures with Qiaprep Spin Columns (Qiagen), extraction of DNA fragments from agarose gels using Qiaquick Spin Columns (Qiagen), and isolation of BAC DNA from bacterial cultures using Nucleobond BAC 100 columns (Macherey-Nagel). All enzymes used for cloning (restriction enzymes, thermosenstitive alkaline phosphatase, T4 DNA ligase) were purchased from Promega. PCR reactions were performed using Hot Start *TaKaRa Taq* polymerase, according to manufacturer's instructions.

4.1.1 Recombineering of the *T* BAC

The BAC containing the *T* gene (RP24-530D23, Imagenes) was modified to introduce the Y88A point mutation and intronic neomycin selection cassette using Red/ET recombination (Muyrers et al., 1999). First, PCR site-directed mutagenesis and molecular cloning were used to generate a DNA fragment containing part of the *T* gene with the TAC to GCC codon mutation (resulting in a Y88A amino acid substitution), an intronic neomycin selection cassette, and 50bp arms with homology to the *T* BAC. The pSC101-BAD-gbaA(tet) plasmid, containing arabinose-inducible expression of recombinase phage proteins, was transformed into DH10B *E. coli* cells containing the *T* BAC. Recombineering was induced following the protocol provided in the Gene Bridges BAC Modification Kit (Gene Bridges). Positively recombined clones were verified by resistance to chloramphenicol (provided by the BAC backbone) and kanamycin (provided by the recombined selection cassette). Further verification was performed using restriction enzyme digest and sequencing of PCR products from the BAC. Sequencing of exons was performed by MWG Eurofins in Ebersberg, Germany.

4.1.2 Generation of the *T*^{Y88A} Homologous Targeting Vector

Restriction digest of the recombined BAC was used to isolate a fragment of the *T* genomic region containing the Y88A mutation in the 2nd exon of *T*, an intronic neomycin selection cassette, and 1.3kb and 5.7kb of upstream and downstream sequence, respectively. This fragment was cloned into the vector used for homologous recombination, containing a pBluescript backbone and a diphtheria toxin gene for negative selection of ESC clones that randomly integrated the vector sequence into their genome (pBS_DTA). The vector was then linearized with *PacI* for introduction into ESCs.

4.1.3 Generation of Plasmid Constructs

The following plasmids were generated and used for experiments contained in this thesis.

Plasmid	Elements Contained	Purpose
pcDNA3.1_T^{WT}myc	CMV promoter, Kozak consensus site, wild-type *T* coding sequence, 3' in-frame myc tag, polyA signal	CHX chase assay, T localization, CoIP of KDMs and T
pcDNA3.1_T^{Y88A}myc	CMV promoter, Kozak consensus site, *T*^{Y88A} coding sequence, 3' in-frame myc tag, polyA signal	CHX chase assay, T localization, CoIP of KDMs and T
pcDNA3.1_TBX6^{WT}myc	CMV promoter, Kozak consensus site, wild-type *Tbx6* coding sequence, 3' in-frame myc tag, polyA signal	Western blot control
pcDNA6.2_T^{WT}V5AviC	CAGGS promoter, wild-type *T* coding sequence, 3' in-frame V5 and Avi tags, IRES, humanized BirA enzyme coding sequence, polyA signal	Stable integration of tagged T^{WT} in P19 cells
pcDNA6.2_T^{Y88A}V5AviC	CAGGS promoter, *T*^{Y88A} coding sequence, 3' in-frame V5 and Avi tags, IRES, humanized BirA enzyme coding sequence, polyA signal	Stable integration of tagged T^{Y88A} in P19 cells
pcDNA6.2_KDM6A_V5AviC	CAGGS promoter, wild-type *Kdm6a* coding sequence, 3' in-frame V5 and Avi tags, IRES, humanized BirA enzyme coding sequence, polyA signal	CoIP of KDM6A and T
pcDNA6.2_KDM6B_V5AviC	CAGGS promoter, wild-type *Kdm6b* coding sequence, 3' in-frame V5 and Avi tags, IRES, humanized BirA enzyme coding sequence, polyA signal	CoIP of KDM6B and T

4.2 Tissue Culture

All ESC culture was performed based on protocols in (Nagy et al., 2003). All tissue culture was performed in a laminar flow hood and all materials were autoclaved and disinfected with 70% ethanol before work.

4.2.1 ESC Culture Media

Media	Composition
Feeder Medium	Dulbecco's Modified Eagle's Medium (DMEM) containing 4,500mg/mL glucose (Gibco) 10% fetal bovine serum (Gibco) 2mM L-glutamine (Lonza) 50U/mL penicillin; 50µg/mL streptomycin (Lonza)
ESC Medium	Dulbecco's Modified Eagle's Medium (DMEM) containing 4,500mg/mL glucose and without sodium pyruvate (Gibco #41965) 15% fetal bovine serum-ES cell grade (Pan-Biotech GmbH) 2mM L-glutamine (Lonza) 50U/mL penicillin; 50µg/mL streptomycin (Lonza) 1% 100x non-essential amino acids (Gibco) 0.1mM β-mercaptoethanol (Sigma-Aldrich) 1% 100x nucleosides (Specialty media) 1000U/mL murine leukemia inhibitory factor (LIF; Chemicon)
Resuspension Medium	Dulbecco's Modified Eagle's Medium (DMEM), bicarbonate-free (Sigma-Aldrich) 10mM HEPES (Gibco) 15% fetal bovine serum (Gibco) 2mM L-glutamine (Lonza) 50U/mL penicillin; 50µg/mL streptomycin (Lonza) 1% 100x non-essential amino acids (Gibco) 0.1mM β-mercaptoethanol (Sigma-Aldrich) 1% 100x nucleosides (Specialty media)
2x Freezing Medium	Dulbecco's Modified Eagle's Medium (DMEM), bicarbonate-free (Sigma-Aldrich) 20% fetal bovine serum (Gibco) 20% dimethylsulfoxide (DMSO; Sigma-Aldrich)
Serum-free Medium	Dulbecco's Modified Eagle's Medium (DMEM)/F12 (Gibco) Neurobasal Medium (Gibco) 2mM L-glutamine (Lonza) 50U/mL penicillin; 50µg/mL streptomycin (Lonza) 0.5x B-27 supplement, serum-free (Gibco) 0.5x N-2 supplement (Gibco) 0.05% Albumin, bovine Fraction V (7.5%; Sigma-Aldrich) 0.001% 1-thioglycerol (Sigma-Aldrich)

4.2.2 ESC Culture

Mitotically inactive mouse embryonic fibroblast cells (feeder cells) were plated one day prior to plating of ESCs, in feeder medium on a cell culture dish (Corning and TPP) previously coated with 0.1% gelatin (Sigma-Aldrich). Undifferentiated ESCs were cultured on a monolayer of feeder cells ($1x10^6$ feeder cells per 6cm dish) in ESC medium and incubated at 37°C in a humidified 7.5% CO_2 incubator. Media was changed daily.

4.2.3 Generation of Modified ESCs by Electroporation of Linearized BACs and Targeting Constructs

ESCs used for targeting were plated onto a monolayer of feeder cells ($1x10^6$ feeder cells per gelatinized 6cm dish) and grown in ESC medium, as above, until round colonies were visible. Cells were carefully washed twice with cell-culture grade DPBS (Lonza), DPBS was aspirated and 600µL Trypsin/EDTA (Gibco) was added to the cells. Cells were incubated at 37°C for 10 minutes to disrupt cell-cell contacts. Trypsin was inactivated by addition of 3mL ESC medium. Cells were pipetted up and down to create a single cell suspension and ESCs were counted using a hematocytometer (Neubauer; Roth). $2.5\text{-}5x10^6$ ESCS were collected by centrifugation for 5 minutes at 1,000 rpm. Cells were resuspended in 800µL DPBS and transferred to an electroporation cuvette (0.4cm gap; BioRad). 5µg of unpurified, linearized $T^{WT}neo$, $T^{Y88A}neo$, or $T^{mCherry}$ BAC (linearized with *PI-SceI*; New England Biolabs) was used for random integration of BAC DNA. 25µg of *PacI*-linearized pBS_DTA_T^{Y88A} homologous targeting construct was purified by phenol:chloroform extraction and ethanol precipitation and used for homologous recombination of the T^{Y88A} mutation. Linearized DNA was mixed with ESCs in the electroporation cuvette. Cells were electroporated with a GenePulser (BioRad) with the following electroporation conditions: 240V and 500µF, and immediately transferred to a Falcon tube with 10mL ESC medium. Cells were collected by centrifugation, as above, and resuspended in ESC medium before splitting the cells onto a monolayer of neomycin-resistant feeder cells. Cells were incubated overnight at 37°C and after 36 hours, selection for integration of the BAC or homologous targeting construct began by addition of ESC medium supplemented with 250µg/mL G418 (Gentecin, analog of

neomycin sulfate; Gibco). Selection medium was changed daily until neomycin-resistant ESC colonies were visible, approximately 6-8 days after electroporation.

4.2.4 Lipofectamine Transfection of FlpE Recombinase

In order to remove the *FRT*-flanked selection cassette from modified ESCs, transfection of a FlpE mammalian expression vector was used. To this end, 0.5-1x10^6 ESCs were plated onto a monolayer of 1x10^6 feeder cells in a gelatinized 6cm dish. 2.5mL of ESC medium (minus penicillin/streptomycin) was added to the plate. 5.5µg of fresh FlpE expression plasmid was incubated in 375µL OptiMEM (Gibco) and 75µL Lipofectamine 2000 (Invitrogen) was added to 300µL OptiMEM. Both were incubated at room temperature for 5 minutes. The DNA and Lipofectamine 2000 mixtures were mixed together and incubated at room temperature for 15 minutes. The DNA-Lipofectamine mixture was then added to cells with an additional 1.75mL ESC medium (minus penicillin/streptomycin) and incubated at 37°C for 4 hours. After 4 hours, ESCs were counted and split (100, 200, 500 and 1,000 cells) onto a monolayer of feeder cells in a gelatinized 6cm dish. ESC clones were allowed to grow for 6 days before colonies were visible. ESC medium without antibiotic was changed daily.

4.2.5 Picking Individual ESC Clones

Fresh ESC medium was added to cells 4 hours before picking clones. Cells were washed with DPBS before covering the clones with a layer of DPBS for picking. Individual clones were picked using sterile 10µL pipette tips into a single well of a 96-well round-bottom plate with 25µL of trypsin/EDTA. After picking, clones were incubated at 37°C for 10 minutes and trypsin was inactivated by the addition of 100µL of ESC medium. Colonies were disaggregated by pipetting up and down and ESCs were plated onto a monolayer of feeder cells (1x10^6 cells per gelatinized 96-well plate). ESC medium was changed daily during expansion of the clones.

4.2.6 Splitting and Freezing of ESC Clones

Expanded ESC clones were split 1:3 onto 3 separate 96-well plates; two for genomic DNA isolation and one for freezing. ESCs were washed twice in DPBS and 70µL of trypsin/EDTA was added and incubated at 37°C for 10 minutes. Trypsin was inactivated by addition of 140µL of ESC Resuspension medium and cells were pipetted up and down to create a single-cell suspension. 70µL of the 210µL cell suspension was added to a round-bottom plate containing 70µL 2x Freezing medium. Cells were pipetted up and down, the plate was sealed with parafilm and immediately put into a styrofoam box at -80°C to freeze the cells. The remaining 140µL of cell suspension was split onto two 96-well plates containing ESC medium; one coated with 0.1% gelatin and the original plate that was trypsinized. ESCs were expanded by incubation at 37°C before isolation of genomic DNA for verification of target DNA integration by Southern blot.

4.2.7 Genomic DNA Isolation for Southern Blot

In order to test ESC clones for integration of BAC and targeting construct DNA, Southern blotting was performed following the protocol from (Ramirez-Solis et al., 1993). Cells in the 96-well plate were washed twice with DPBS and cells were lysed with 50µL per well of prewarmed lysis buffer (10mM Tris-HCl pH 7.5, 10mM EDTA pH 8.0, 10mM NaCl, 0.5% sarcosyl) supplemented with 1mg/mL Proteinase K (Sigma-Aldrich). Cells were incubated overnight in a humidified chamber at 37°C and 100µL of ice-cold 75mM NaCl in 100% ethanol was added to each well the next day. The cells were left on the bench for 30 minutes without agitation to precipitate DNA. Precipitated DNA was then washed 3x with 70% ethanol and allowed to air-dry. The genomic DNA was then resuspended in 1x TE buffer (10mM Tris-HCl pH 8.0, 1mM EDTA).

4.2.8 Restriction Digest and Transfer of Genomic DNA

A restriction digest master mix was prepared with the following: 1x Restriction Enzyme Buffer (Promega), 1mM spermidine (Promega), 100µg/mL bovine serum albumin (BSA; Promega), 100µg/mL RNaseA (Fermentas) and 40 units of restriction

enzyme (Promega) and added to genomic ESC DNA in the 96-well plate without mixing. The plate was incubated in a humidified chamber at 37°C for 4 hours. The contents of the well were mixed by pipetting up and down and the reaction was incubated overnight in a humidified chamber at 37°C. The following day, genomic DNA fragments were separated on a 0.7% agarose gel in 1x TAE buffer (40mM Tris, 19mM acetic acid, 1.25mM EDTA) for several hours at 70mV. Following documentation, the gel was pretreated by incubation for 30 minutes each in Denaturation buffer (0.5N NaOH, 1.5M NaCl), Neutralization buffer (0.5M Tris-HCl pH 8.0, 1.5M NaCl) and 20x SSC (3M NaCl, 300mM trisodium citrate pH 7.0). DNA was transferred to a blotting membrane (Amersham Hybond-N, GE Healthcare) overnight using a Whatman TurboBlotter according to manufacturer's instructions. The following day, DNA was crosslinked to the membrane using the UV Stratalinker 2400 (500 μJ/cm^2; Stratagene) and the membrane was either air-dried and stored at 4°C or used directly for hybridization.

4.2.9 Probe Labeling and Hybridization

For radioactive labeling of the probe, 25ng of template DNA was diluted in 1x TE buffer and denatured at 95°C for 5 minutes. Labeling was performed using the RediPrime Random Labeling System according to manufacturer's instructions (GE Healthcare) with 5μL of α-^{32}P dCTP (3000Ci/mmol; Redivue, Amersham). Labeled probes were purified using G-25 MicroSpin columns according to manufacturer's instructions (illustra ProbeQuant; GE Healthcare). Membranes were incubated in 15mL ExpressHyb Hybridization solution (Clontech) for 1 hour at 65°C with continuous rotation, before addition of the denatured radioactively-labeled probe and incubation overnight at 65°C with continuous rotation. The next day, membranes were washed twice in 2x SSC at 65°C and twice more with 0.2x SSC at 65°C with continuous rocking. The membrane was sealed in a plastic bag and exposed to a Phosphor screen for 24 hours, before imaging with a PhosphorImager (Amersham).

4.2.10 Genotyping PCR of ESC Clones

Genotyping PCR of ESC clones was performed using genomic DNA isolated as described in Section 4.2.7. Genomic DNA was diluted 1:100 and 1µL was used in a reaction with BAC specific primers to determine integration of BAC DNA. Primers used for genotyping can be found in Section 4.6. PCR was performed using Hot Start *TaKaRa Taq* polymerase according to manufacturer's instructions.

4.2.11 Quantification of BAC DNA Integration

Quantification of the number of BAC DNA integrations was performed using genomic DNA from ESC clones isolated as described in Section 4.2.7. Quantification was achieved using quantitative real-time PCR with SYBR green. Genomic DNA was diluted 1:100 and 1µL was used in a reaction with 5µM *T*-specific primers and GoTaq qPCR Master Mix (Promega). Reactions were performed in a MicroAmp Fast Optical 96-well Plate (Applied Biosystems) and run in triplicates using a StepOne Real-Time PCR System. Data was analyzed using the ΔΔCT method provided in the StepOnePlus Software v2.0.2 (Applied Biosystems). *Msgn1* primers recognizing genomic DNA were used as a normalization control.

4.2.12 *In Vitro* Differentiation of ESCs into Mesodermal Cells

ESCs were grown on feeder cells in a 6cm dish as described in Section 4.2.2. Depletion of feeder cells was achieved by trypsinization, followed by incubation of the ESC/feeder cell mix in a 6-well plate in ESC medium for 25 minutes. Every 25 minutes, the ESC/feeder cell mix was transferred to a new well of a 6-well plate for a total of 3x before plating onto a gelatinized 6cm dish. Since the feeder cells settle and attach to the 6-well plate faster than the ESCs, passaging 3x onto new wells results in feeder cell depletion. Feeder-free ESCs were incubated overnight in ESC medium supplemented with LIF. The next day, feeder-free ESCs were trypsinized and counted, and hanging drops containing 200 cells per drop were plated onto square petri dishes in ESC medium (120mmx120mm; Greiner Bio-one). Plates were inverted and incubated overnight at 37°C, allowing for formation of embryoid bodies (EBs) by aggregation of ESCs. The next

day, EBs were washed twice with Advanced DMEM (Gibco) and collected by centrifugation for 5 minutes at 1000rpm. EBs were plated onto plates coated overnight at 4°C with fibronectin (diluted 1:100 in 1x PBS; Calbiochem #341631), in serum-free medium supplemented with growth factors as per (Schröder, 2010), modified by Dr. Pavel Tsaytler, (personal communication). Medium was changed daily throughout the differentiation protocol (4-5 days total).

4.2.13 Culture and Transfection of Non-ESCs

HEK293T, NIH3T3, and C2C12 cells were cultured in Feeder Medium according to standard protocols. Medium was changed every 1-2 days and cells were passaged regularly. Cells were transiently transfected with over-expression constructs using Lipofectamine 2000 reagent (Invitrogen) according to manufacturer's instructions.

4.2.14 Generation of P19 Cell Lines Stably Over-expressing Tagged T and Differentiation into Mesoderm

Undifferentiated P19 cells were cultured in Feeder medium according to standard protocols. pcDNA6.2_T^{WT}V5AviC and pcDNA6.2_T^{Y88A}V5AviC constructs were linearized by restriction digest with *MluI* and purified by phenol:chloroform extraction and ethanol precipitation. 5µg of linearized, purified DNA was introduced into P19 cells by electroporation, as described in Section 4.2.3 above. Electroporated cells were seeded 1:20, 1:50, 1:100 and 1:200 and cells were selected for stable integration of the over-expression constructs by incubation for 7-8 days in Feeder medium supplemented with 10µg/mL Blasticidin (Invivogen). Single P19 cell clones were then picked, as described in Section 4.2.5 above. Clones were expanded in Feeder medium supplemented with 10µg/mL Blasticidin and tested for T over-expression by Western blot, according to 4.3.7. P19 cells were differentiated by plating of 1.5 million cells per 10cm non-adhesive plate in Feeder medium supplemented with 1% DMSO (Sigma-Aldrich) for 2 days to allow for aggregation of cells.

4.2.15 Generation of Fully ESC-derived Embryos by Tetraploid Complementation

Tetraploid complementation of genetically modified ESCs was performed by Karol Macura according to the protocol described in (Gertsenstein, 2011).

4.3 Molecular Biology Techniques

4.3.1 Whole-Mount *In Situ* Hybridization (WISH)

WISH was performed according to the protocol available online from the Molecular Atlas of the Mouse Embryo Project (MAMEP; Berlin, Germany; http://mamep.molgen.mpg.de/), with minor adjustments described below. All solution compositions can be found in the online protocol. The dissection and fixation of mouse embryos was done according to the MAMEP protocol. The processing of embryos was also performed according to the protocol, but in 2mL screw cap tubes (RNase-free; Sarstedt) instead of in 12-well plates with net wells. RNA ISH probes from the MAMEP database were used and generated according to the online protocol, with cleanup of the probe performed using ProbeQuant G50 Sephadex Columns (Pharmacia). After overnight hybridization with the ISH probe, embryos were washed 2x in Solution I for 30 minutes each at 68°C, followed by 2x washes in Solution 3T for 30 minutes each at 68°C. Embryos were then washed 1x with 50% Solution 3T:50% 1x MABT (100mM Maleic acid, 150mM NaCl, pH 7.5, Tween-20 0.01%) for 20 minutes, followed by 3x washes in 1x MABT with 2mM Levamisole (Sigma-Aldrich) at room temperature. Embryos were then incubated for 1 hour at room temperature in blocking solution (2% Blocking Reagent (Roche), 20% fetal bovine serum, 2mM Levamisole in 1x MABT). Embryos were incubated overnight at 4°C with α-DIG-AP antibody (Roche) in blocking solution. Post-antibody washes were performed as described in the MAMEP online protocol and staining was performed as described, using BM Purple AP Substrate (Roche) as a staining solution. The staining was monitored periodically under a binocular, and staining was stopped as described in the online protocol. Embryos were then imaged using a SteREO Discovery.V12 microscope and AxioCam Color Camera (Zeiss).

4.3.2 Isolation of RNA

RNA from embryo samples and cultured cells was isolated using the RNeasy Mini Plus Kit (Qiagen), according to manufacturer's instructions.

4.3.3 cDNA Synthesis

Reverse transcription of RNA into cDNA was performed using the SuperScript II First Strand Synthesis Kit (Invitrogen) with random hexamer primers (1µL of 50µM random hexamers; Roche), according to manufacturer's instructions. 500ng of RNA was used as a template in each cDNA synthesis reaction.

4.3.4 Gene Expression Analysis (qRT-PCR)

Detection and quantification of individual genes of interest were performed using quantitative reverse-transcriptase PCR with SYBR green. cDNA from Section 4.3.3 was diluted 1:10 and 1µL (~50ng) was used in a reaction with 5µM transcript-specific primers and GoTaq qPCR Master Mix (Promega). Reactions were performed in a MicroAmp Fast Optical 96-well Plate (Applied Biosystems) and run in triplicates using a StepOne Real-Time PCR System. Data was analyzed using the ΔΔCT method provided in the StepOnePlus Software v2.0.2 (Applied Biosystems). Expression of genes of interest were normalized, using *Pmm2* as a reference gene. A list of transcript-specific primers can be found in Section 4.6.

4.3.5 Sequencing of *T* cDNA

PCR amplification of *T* cDNA was performed using 5ng of diluted embryonic cDNA and *T* specific primers listed in Section 4.6. PCR was performed using Hot Start *TaKaRa Taq* polymerase according to manufacturer's instructions. Sequencing of the PCR product was performed by MWG Eurofins in Ebersberg, Germany and sequencing primers can be found in Section 4.6.

4.3.6 Quantification of Protein Concentration

Protein from embryo and cell culture lysates was quantified using the Qubit® Protein Assay Kit and the Qubit® 2.0 Fluorometer, according to manufacturer's instructions.

4.3.7 Western Blotting

Embryo and cell culture samples were lysed in Western lysis buffer (1% SDS, 10mM Tris-HCl pH 7.5, 2mM EDTA, 1x Complete Protease Inhibitor (Roche), 1x PhosStop Phosphatase Inhibitor (Roche)) for 10 minutes on ice. Protein quantification was performed according to Section 4.3.6 and 10-40µg of total protein was loaded onto a precast 4-12% Bis-Tris gel (NuPAGE; Invitrogen). Proteins were separated by electrophoresis and transferred onto a nitrocellulose membrane using the Mini Trans-Blot cell (BioRad) in 1x Tris-Glycine transfer buffer (25mM Tris, 192mM Glycine, 10% methanol). Membranes were then washed once with 1x TBST (50mM Tris-HCl pH 7.5, 150mM NaCl, 0.1% Tween-20) and blocked for 1 hour at room temperature in 5% skim milk/TBST. Membranes were then incubated overnight at 4°C with primary antibody in blocking solution. A list of all antibodies used can be found in Section 4.7. Membranes were then washed 3x for 10 minutes in 1x TBST and incubated with HRP-conjugated secondary antibody in 5% skim milk/TBST for 2 hours at room temperature. Membranes were washed 3x for 10 minutes each in 1x TBST and proteins were detected using the Amersham ECL kit (GE Healthcare) according to manufacturer's instructions. Chemiluminescent signal was visualized using the Fusion SL Advance Chemiluminescence Detection System. Western blot signals were quantified using the Fusion CAPT software.

4.3.8 Immunohistochemistry

Immunostaining of T protein in NIH3T3 cells was performed in 24-well cell culture plates. Cells were transfected with myc-tagged T constructs according to Section 4.2.13. Cells were washed the next day 2x in 1x PBS and fixed for 60 minutes at 4°C with 4% paraformaldehyde in 1x PBS (Sigma-Aldrich). Cells were washed 2x in 1x PBS and blocked using 5% skim milk/PBS for 1 hour at room temperature. Cells were incubated overnight with α-T antibody in 5% skim milk/PBS at 4°C. The next day, cells were washed 3x for 10 minutes each in 1x PBS and incubated with fluorophore-conjugated secondary antibody in 5% skim milk/PBS for 2 hours at room temperature. Cells were

washed again 3x for 10 minutes each with 1x PBS. 1x PBS with 3µM DAPI (Molecular Probes) was added and cells were imaged using a Zeiss Observer.Z1 microscope.

4.3.9 Cycloheximide Chase Assay

NIH3T3 cells were transfected with myc-tagged T^{WT} and T^{Y88A} expression constructs as described in Section 4.2.13. 24 hours after transfection, 50µg/mL cycloheximide in DMSO (Sigma-Aldrich) was added to cells (+CHX samples) while DMSO vehicle was added to the remaining cells (-CHX samples). Lysates from cells were taken at 0 hours (h), 1h, 2h, 4h, 8h and 12h after addition of CHX and a Western blot was run to detect myc-tagged T according to Section 4.3.7. Antibodies used can be found in Section 4.7.

4.3.10 Co-Immunoprecipitation of Proteins

For coIP of over-expressed tagged T and KDMs in HEK293T cells, cells were transfected with over-expression constructs using Lipofectamine 2000 (Invitrogen) according to manufacturer's instructions. 24 hours after transfection, whole cell lysates were collected in Whole Cell Extract Buffer (20mM HEPES, pH 8.0, 350mM NaCl, 0.5% Igepal, 25% glycerol, 0.2mM EDTA, 1.5mM $MgCl_2$, 1mM DTT, 1x Complete Protease Inhibitor (Roche), 1x PhosStop Phosphatase Inhibitor (Roche)), drawn through a 27 gauge needle and incubated on ice for 10 minutes. Lysates were spun at 13,000rpm for 5 minutes at 4°C and protein in the supernatant was quantified according to section 4.3.6. 40µg was taken as an input control for Western blotting. 2mg of protein extract was used per IP and incubated with M270 Streptavidin Dynabeads (Invitrogen) for 2 hours at 4°C for 2 hours with rotation. Beads were then washed 5x for 10 minutes each in stringent wash buffer (20mM HEPES, pH 8.0, 250mM NaCl, 1mM EDTA, 0.1% Igepal, 1x Complete Protease Inhibitor (Roche) and 1x PhosStop Phosphatase Inhibitor (Roche)). Beads were then incubated with 1x NuPAGE® LDS Sample Buffer (Invitrogen) with 50mM DTT and boiled at 95°C for 5 minutes to release bound proteins. Lysates were then loaded onto an SDS-PAGE gel and Western blotting was performed according to Section 4.3.7.

For coIP of tagged T from differentiated P19 cells, differentiated cells were washed 2x in 1x PBS and collected in Whole Cell Extract Buffer as above. 5mg of protein lysate was incubated with M270 Streptavidin Dynabeads (Invitrogen) for 4 hours at 4°C with rotation. Beads were washed 3x for 10 minutes each with low stringency buffer (10mM Tris, pH 7.5, 150mM NaCl, 0.1% Triton X-100, 2mM EDTA, 1x Complete Protease Inhibitor (Roche) and 1x PhosStop Phosphatase Inhibitor (Roche)) with rotation at 4°C. Beads were then incubated with 1x NuPAGE® LDS Sample Buffer (Invitrogen) with 50mM DTT and boiled at 95°C for 5 minutes to release bound proteins. Lysates were then loaded onto an SDS-PAGE gel and Western blotting was performed according to Section 4.3.7.

4.4 Gene Expression Profiling

4.4.1 Embryo Dissection and Isolation of *T-mCherry*-positive Cells

$T^{2J/+}$;Tg($T^{mCherry}$)1Bgh and *$T^{2J/Y88A}$;Tg($T^{mCherry}$)1Bgh* embryos were dissected at TS12 and TS13, corresponding to E8.25 and E8.5, respectively. Caudal ends, excluding somites, notochord, and allantois, were dissected into M2 medium (Sigma-Aldrich) and kept on ice. After dissection, M2 medium was removed and caudal ends were resuspended in 100µL of Trypsin/EDTA + 5µL 10% BSA (Sigma-Aldrich) for 5 minutes on ice with pipetting up and down to create a single-cell suspension. Trypsin was quenched by adding 100µL of 10% BSA solution. 20,000 *T-mCherry*-positive cells were then sorted by FACS using the FACS Aria II SORP flow cytometer.

4.4.2 RNA Isolation, Library Preparation, and Sequencing

RNA was isolated using the RNeasy Micro Plus Kit (Qiagen) according to manufacturer's instructions. Total RNA was quantified using the Qubit® RNA HS Kit (Invitrogen) and RNA quality was verified using a Bioanalyzer 2100 (Agilent), according to manufacturers' instructions. Ribosomal RNA was depleted and libraries were prepared using the ScriptSeq v.2 Complete (Human/Mouse/Rat) Low Input kit (epicentre) according to manufacturer's instructions. Libraries were verified using the High Sensitivity DNA chip and the Bioanalyzer 2100 (Agilent). 10nmol of each library was pooled and sequenced with the Illumina Hi-Seq 2500 platform, with paired end (2x 51 bp) and multiplexed sequencing. Sequences from fastq files were mapped using the RefSeq mm10 annotation from UCSC using TopHat 2.0.8b and only uniquely mapped reads were chosen. Normalization and FPKM values were calculated using Cufflinks 2.1.1 using the RefSeq mm10 annotation from UCSC. Total reads and mapped reads for each sample can be found in Supplementary Table S1.

4.4.3 RNA-Seq Data Analysis

Heat maps, hierarchical clustering, and k-means clustering of gene expression data was performed using Cluster 3.0 software (de Hoon et al., 2004). GO term enrichment analysis was performed using the Database for Annotation, Visualization and Integration Discovery (DAVID) tool (Huang et al., 2009). In order to determine enrichment of palindromic T binding sites in genes with dysregulated expression, a ranked list of T binding sites according to how well they adhere to the palindromic T binding site was generated (Dr. Frederic Koch, Max Planck Institute for Molecular Genetics, Berlin, Germany). This was used to generate Q-Q plots to compare distributions of T palindrome strength in dysregulated versus non-dysregulated genes. To test for significance, P-values were calculated using the two-sample Kolmogorov-Smirnov (K-S) test (Matt Huska, Max Planck Institute for Molecular Genetics, Berlin, Germany).

4.5 Profiling the Histone Modification Landscape

4.5.1 Isolation of *T-mCherry*-positive Differentiated ESCs and ChIP using Histone Mark Specific Antibodies

$T^{2J/+}$;Tg($T^{mCherry}$)1Bgh, *$T^{2J/2J}$;Tg($T^{mCherry}$)1Bgh* and *$T^{2J/Y88A}$;Tg($T^{mCherry}$)1Bgh* ESCs were differentiated according to the protocol in Section 4.2.12. Cells were trypsinized into a single cell suspension and 500,000 *T-mCherry*-positive cells were sorted into 5% BSA using the FACS Aria II SORP flow cytometer. Cells were collected by centrifugation for 5 minutes at 1000rpm and resuspended in 1x PBS with 27µL of 37% formaldehyde (Sigma-Aldrich) to crosslink. Crosslinking was quenched by the addition of 100µL of 2.5M glycine and cells were washed 2x with 1x PBS. The iDeal ChIP-Seq kit (Diagenode) was used to lyse cells and chromatin was sheared using a Standard BioRuptor (Diagenode). The ChIP was performed using the iDeal ChIP-Seq kit according to manufacturer's instructions; with chromatin from 100,000 cells and 1.5µg of antibody recognizing specific histone marks (H3K4me1, H3K4me3, H3K27me3 and H3K27Ac) listed in Section 4.7.

4.5.2 Library Prep and Sequencing

Input and immunoprecipitated DNA size and quality was verified using a Bioanalyzer 2100, according to manufacturer's instructions. Immunoprecipitated DNA was then used to generate libraries for sequencing with the TruSeq ChIP Sample Preparation Kit (Illumina) according to manufacturer's instructions, with minor adjustments. 18x cycles of amplification were used, but with 5 cycles after adaptor ligation and the remaining 13 cycles after purification of ligation products. 10nmol of each sample were pooled and sequencing was performed using the Illumina Hi-Seq 2500 platform. Sequences from fastq files were mapped to the mm10 genome using Bowtie 1.0 and uniquely mapped reads were chosen for further analysis. Total reads and mapped reads for each sample can be found in Supplementary Table S1.

4.5.3 Analysis of ChIP-Seq Data

Uniquely mapped ChIP-Seq data was analyzed using seqMINER v. 1.3.3 and Galaxy open source software (Blankenberg et al., 2010; Ye et al., 2011).

4.6 Primers

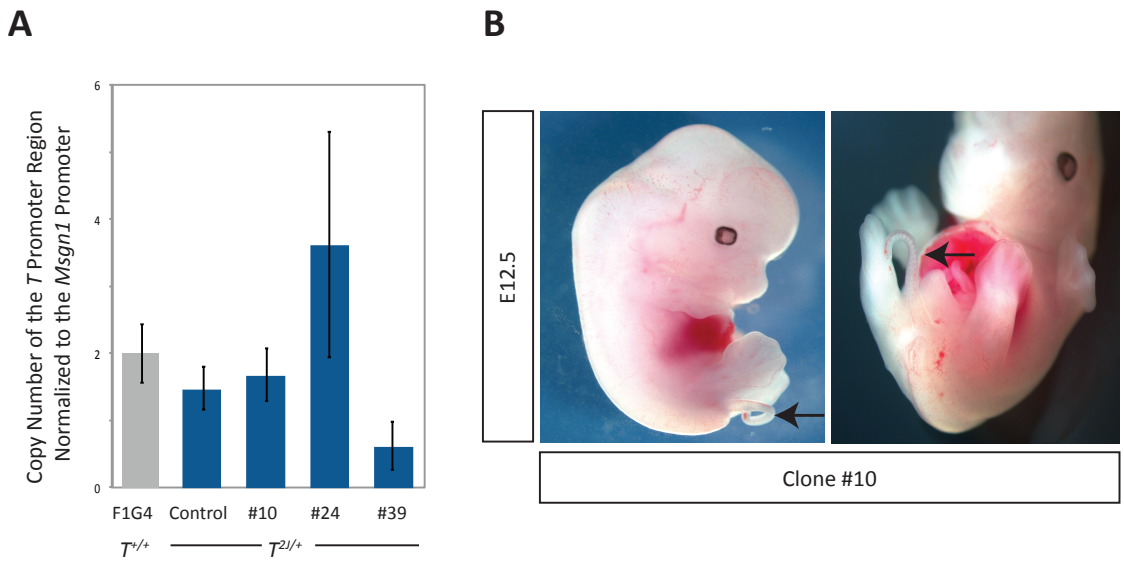
Primer	Sequence	Purpose
TY88A_5'Southern.for	GTCGCTTCTAGCCCGGAAAC	5' *T* Southern Probe Generation
TY88A_5'Southern.rev	GAAGGAAGCCAGAGCTCGCG	
TY88A_3'Southern.for	CACTACACAGTCCAGAGCTCAAGG	3' *T* Southern Probe Generation
TY88A_3'Southern.rev	CACAGACCAGAGACTGGGATACTGG	
T_5'UTR_cDNA	GTCAAGACTCCTGGAAGGTGG	Sequencing of *T* cDNA
T_3'UTR_cDNA	AAGAGCCTGCCACTTTGAGC	
5'His_Screen.for	GCAGGCAAGATAAACGAAGG	Genotyping of $T^{+/Y88A}$;$Tg(T^{Y88A})1Bgh$
Neo5OUT.rev	TGTGCCCAGTCATAGCCG	
Brachy_Lef1.3.F	TTCCTTACAGGAAGCGCGCG	qRT-PCR of 5' genomic *T* region
Brachy_Lef1.3.R	CCCATAAATACAGCCGAGGTGG	
TqPCR.for	CAATGGAGGGGGACAGATCAT	qRT-PCR of *T* expression
TqPCR.rev	CTGGTGATCATGCGTTGCG	
Tbx6(qPCR).for	CTGAAGATCGCAGCCAATC	qRT-PCR of *Tbx6* expression
Tbx6(qPCR).rev	CCCGAAGTTTCCTCTTCACA	
Pmm2(AS251).for	AGGGAAAGGCCTCACGTTCT	qRT-PCR of *Pmm2* expression
Pmm2(AS251).rev	AATACCGCTTATCCCATCCTTCA	

4.7 Antibodies

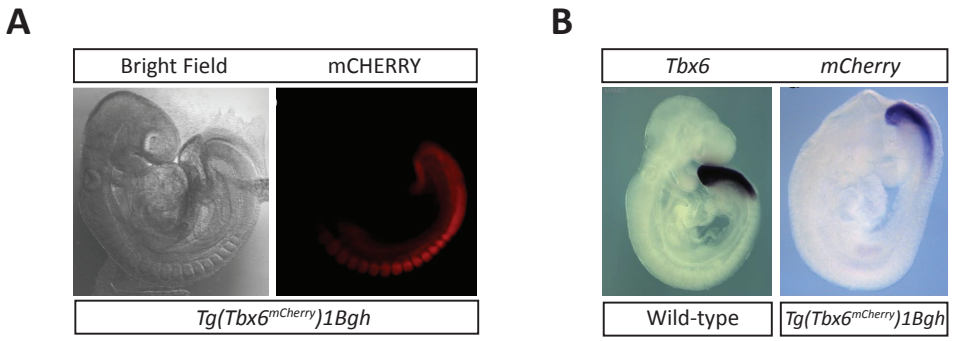
Antibody	Company	Working Concentration	Purpose
α-T (AF2085)	R&D Systems	1:10,000	Western blot
α-Myc (clone 4A6)	Millipore	1:1,000	Western blot, Immunostaining
α-V5	Invitrogen	1:1,000	Western blot
α-Lamin B1 (ab16048)	Abcam	1:4,000	Western blot
α-H3	Abcam	1:5,000	Western blot
Streptavidin-HRP	Generated in the Herrmann Lab	1:5,000	Western blot
α-KDM6A (ab91231)	Abcam	1:500	Western blot
α-KDM6B (ab38113)	Abcam	1:1,000	Western blot
α-RBBP5 (A300-109A)	Bethyl Labs	1:2,000	Western blot
α-WDR5 (A302-429A)	Bethyl Labs	1:2,000	Western blot
α-BRG1 (clone 3G4)	Millipore	1:2,000	Western blot
α-p300 (sc-585)	Santa Cruz	1:500	Western blot
α-Mouse IgG	Cell Signaling	1:5,000	Western blot
α-Rabbit IgG	Cell Signaling	1:5,000	Western blot
α-Goat IgG	Jackson Immuno-research	1:5,000	Western blot
α-H3K4me1 (A1883-0010)	Diagenode	1.5µg 1:500	ChIP Western blot
α-H3K4me3 (ab8580)	Abcam	1.5µg	ChIP
α-H3K27me3 (A1811-001P)	Diagenode	1.5µg	ChIP
α-H3K27Ac (ab4729)	Abcam	1.5µg	ChIP

Chapter 5.

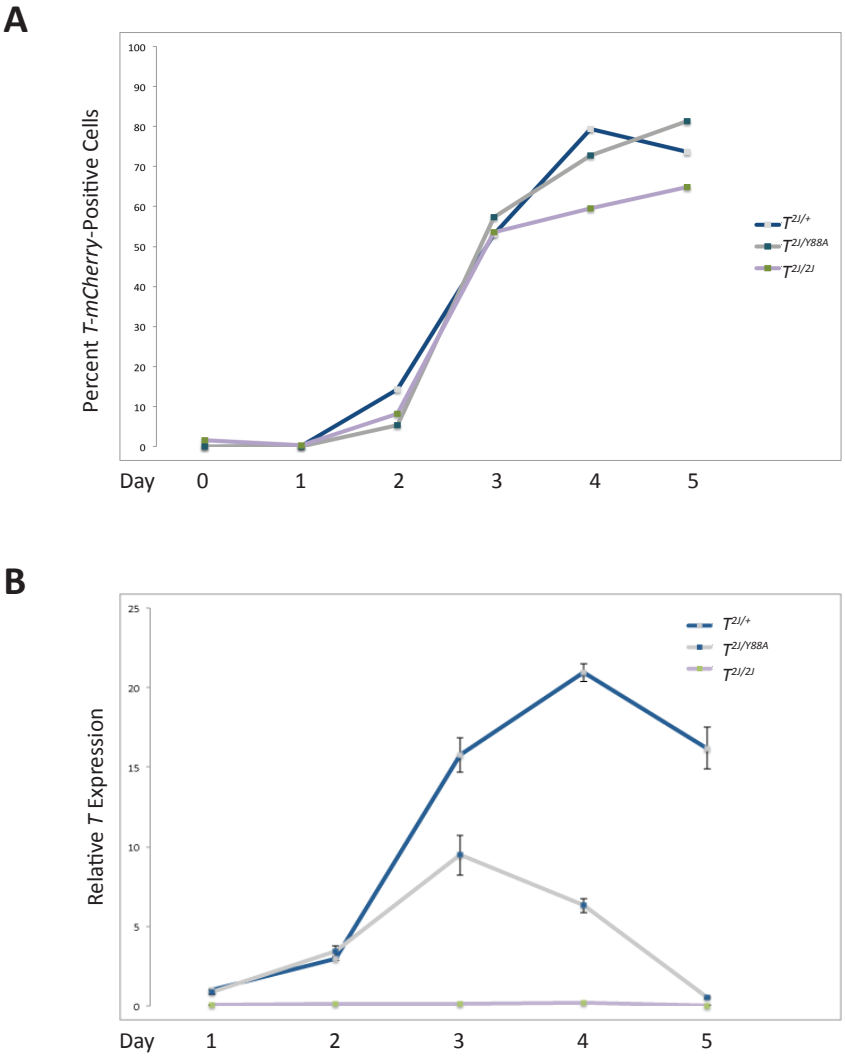
Appendix



Supplementary Figure S1: Confirmation of $T^{2J/+}$ heterozygosity
(A) qPCR using genomic DNA from indicated ESC lines. Primers recognizing regions 5' of the *T* gene were used. Error bars represent standard error of the mean of technical replicates. (B) Tetraploid complementation was used to generate embryos from $T^{2J/+}$ ESC clone #10. Arrows point to the shortened, kinked tail, indicative of *T* heterozygosity.

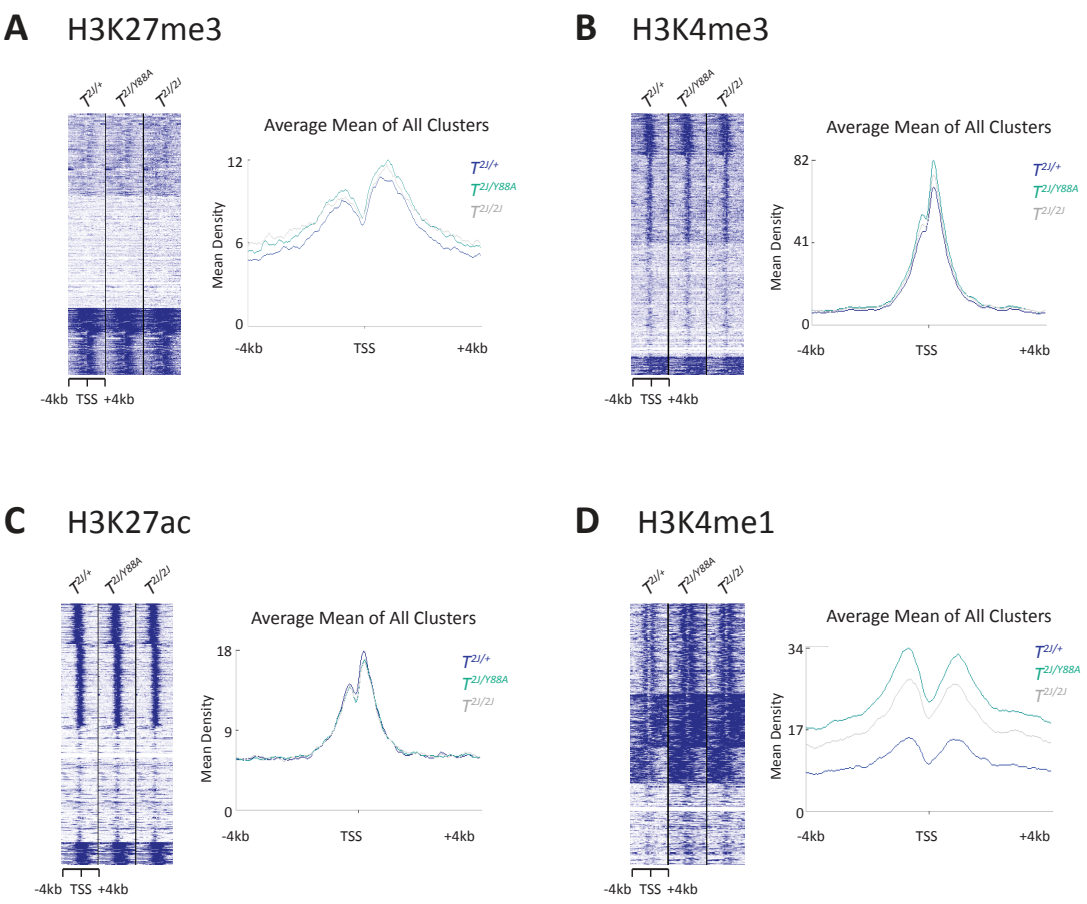


Supplementary Figure S2: mCHERRY protein stability
(A) Bright field and fluorescent images of *Tg(Tbx6^{mCherry})1Bgh* embryos. (B) WISH of *Tbx6* expression in wild-type embryos compared to *mCherry* expression in *Tg(Tbx6^{mCherry})1Bgh* embryos demonstrates that *mCherry* mRNA expression is limited to *Tbx6* expression domains while the mCHERRY protein is more stable and found in the presomitic mesoderm along with developing somites. Images of *Tbx6* expression in the wild-type embryo was taken from the MAMEP database. Images of *mCherry* expression in *Tg(Tbx6^{mCherry})1Bgh* embryo modified from Anna Senft (Master's thesis, Max Planck Institute for Molecular Genetics, Berlin, Germany).

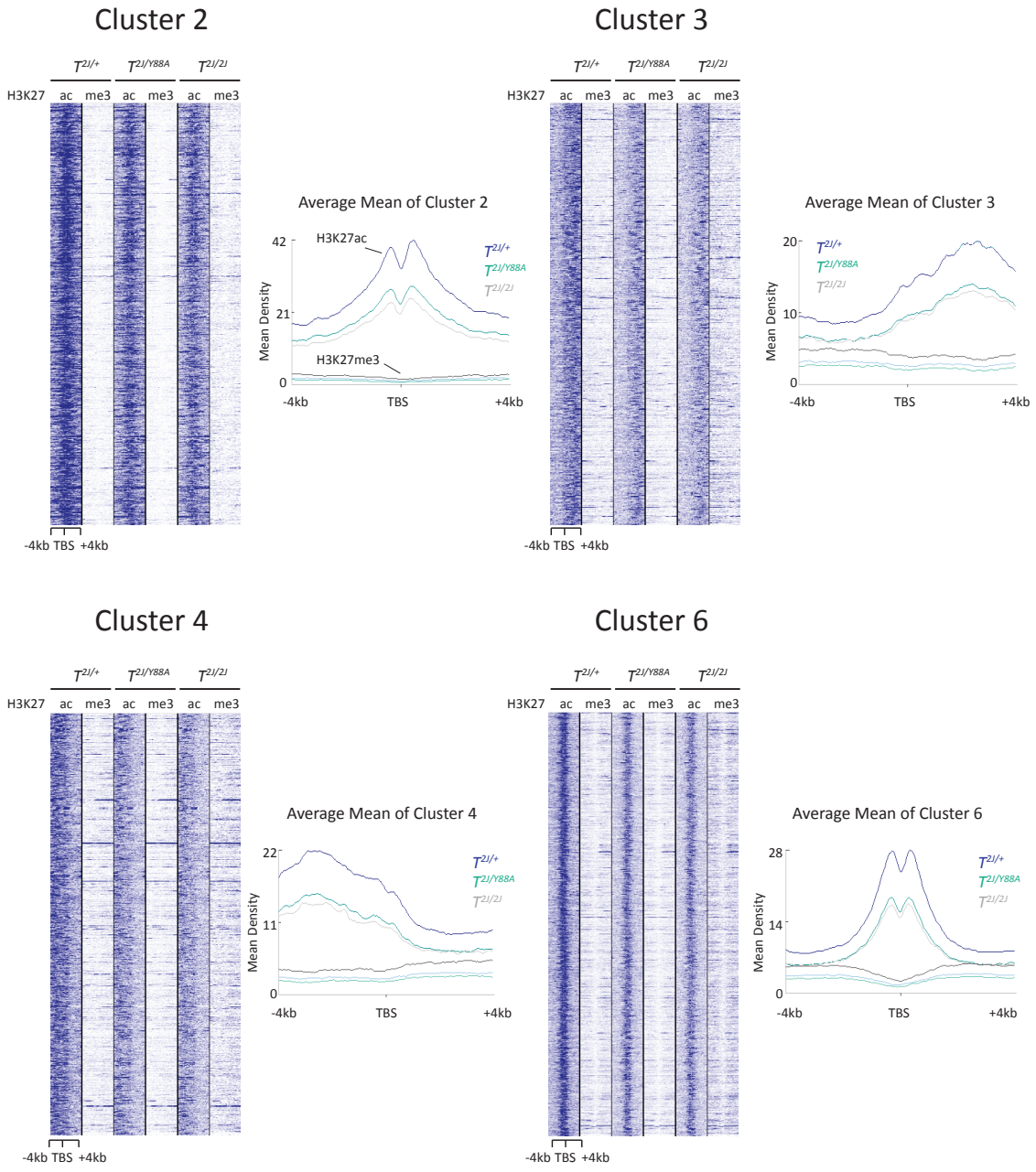


Supplementary Figure S3: *T-mCherry* reporter expression compared to *T* expression in *in vitro* differentiated mesodermal cells

(A) Percent *T-mCherry*-positive cells measured by FACS revealed that transcriptional activation of *T* is intact in $T^{2J/Y88A}$;*Tg*($T^{mCherry}$)*1Bgh* and $T^{2J/2J}$;*Tg*($T^{mCherry}$)*1Bgh* mutant mesodermal cells. (B) qRT-PCR of *T* expression reveals that although transcriptional activation occurs normally, *T* expression is not maintained in $T^{2J/Y88A}$ mutant mesodermal cells.



Supplementary Figure S4: Histone mark profiles at the TSS of dysregulated genes
H3K27me3 (A), H3K4me3 (B), H3K27ac (C), and H3K4me1 (D) profiles at ±4kb from the TSS of genes dysregulated in the $T^{2J/Y88A}$ and $T^{2J/2J}$ mutant embryos. Plots depict the average mean density (tag/50bp) of all clusters at ±4kb from the TSS. TSS, transcription start site.



Supplementary Figure S5: H3K27me3 profiles at regions of lower H3K27ac in $T^{2J/2J}$ and $T^{2J/Y88A}$ mutant mesodermal cells

Heat maps of H3K27ac and H3K27me3 at regions of lower H3K27ac (Figure 2.28B, cluster 2, 3, 4, and 6). Average mean density (tag/50bp) of each cluster shown in plots to the right of heat maps. TBS, T binding site.

Supplementary Table S1: Total vs. Mapped Reads from RNA-Seq and ChIP-Seq

Experiment	Total Reads	Mapped Reads	% Mapped Reads
$T^{2J/+}$ TS12/13 RNA-Seq	153688474	108029549	70.3%
$T^{2J/Y88A}$ TS12 RNA-Seq	116488220	90941490	78.1%
$T^{2J/Y88A}$ TS13 RNA-Seq	150968972	120193323	79.6%
$T^{2J/+}$ H3K4me1	41003799	36394575	88.9%
$T^{2J/+}$ H3K4me3	51778543	32067091	61.9%
$T^{2J/+}$ H3K27me3	41258881	29924508	72.5%
$T^{2J/+}$ H3K27ac	48308179	42205800	87.4%
$T^{2J/Y88A}$ H3K4me1	50896656	45194050	88.8%
$T^{2J/Y88A}$ H3K4me3	57653326	57653326	65.6%
$T^{2J/Y88A}$ H3K27me3	49581312	35089535	70.8%
$T^{2J/Y88A}$ H3K27ac	46652158	40930605	87.7%
$T^{2J/2J}$ H3K4me1	51340905	46220262	90.0%
$T^{2J/2J}$ H3K4me3	56438276	37832462	67.0%
$T^{2J/2J}$ H3K27me3	56903005	41873903	73.6%
$T^{2J/2J}$ H3K27ac	51021769	43899327	86.0%

References

Abu-Abed, S. (2001). The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. Genes Dev. *15*, 226–240.

Agger, K., Cloos, P.A.C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731–734.

Amaya, E., Musci, T.J., and Kirschner, M.W. (1991). Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in Xenopus embryos. Cell *66*, 257–270.

Amaya, E., Stein, P.A., Musci, T.J., and Kirschner, M.W. (1993). FGF signalling in the early specification of mesoderm in Xenopus. Development *118*, 477–487.

Anagnostou, A., Liu, Z., Steiner, M., Chin, K., Lee, E.S., Kessimian, N., and Noguchil, C.T. (1994). Erythropoietin receptor mRNA expression in human endothelial cells. Proc Natl Acad Sci U S A *91*, 3974–3978.

Ang, S., and Rossantt, J. (1994). HNF-3beta Is Essential for Node and Notochord Formation in Mouse Development. Cell *78*, 561–574.

Arnold, S.J., Hofmann, U.K., Bikoff, E.K., and Robertson, E.J. (2008). Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. Development *135*, 501–511.

Arora, R., del Alcazar, C.M., Morrisey, E.E., Naiche, L. a, and Papaioannou, V.E. (2012). Candidate gene approach identifies multiple genes and signaling pathways downstream of Tbx4 in the developing allantois. PLoS One *7*, e43581.

Bamshad, M., Lin, R.C., Law, D.J., Watkins, W.S., Krakowiak, P.A., Moore, M.E., Franceschini, P., Lala, R., Holmes, L.B., Gebuhr, T.C., et al. (1997). Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. Nat. Genet. *16*, 311–315.

Bamshad, M., Le, T., Watkins, W.S., Dixon, M.E., Kramer, B.E., Roeder, A.D., Carey, J.C., Root, S., Schinzel, A., Van Maldergem, L., et al. (1999). The spectrum of mutations in TBX3: Genotype/Phenotype relationship in ulnar-mammary syndrome. Am. J. Hum. Genet. *64*, 1550–1562.

Baron, M.H., and Fraser, S.T. (2005). The specification of early hematopoiesis in the mammal. Curr. Opin. Hematol. *12*, 217–221.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Basu, P., Lung, T.K., Lemsaddek, W., Sargent, T.G., Jr, D.C.W., Basu, M., Redmond, L.C., Lingrel, J.B., Haar, J.L., and Lloyd, J.A. (2007). EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. Blood *110*, 3417–3425.

Bedford, D.C., and Brindle, P.K. (2012). Is histone acetylation the most important physiological function for CBP and p300? Aging (Albany. NY). *4*, 247–255.

Bedford, D.C., Kasper, L.H., Fukuyama, T., and Brindle, P.K. (2010). Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. Epigenetics *5*, 9–15.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol. Cell *43*, 1040–1046.

Black, J.C., Choi, J.E., Lombardo, S.R., and Carey, M. (2006). A mechanism for coordinating chromatin modification and preinitiation complex assembly. Mol. Cell *23*, 809–818.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. In Current Protocols in Molecular Biology, (Wiley Interscience), pp. 19.10.1–21.

Bollag, R.J., Siegfried, Z., Cebra-Thomas, J.A., Garvey, N., Davison, E.M., and Silver, L.M. (1994). An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. Nat. Genet. *7*, 383–389.

Boulet, A.M., and Capecchi, M.R. (2012). Signaling by FGF4 and FGF8 is required for axial elongation of the mouse embryo. Dev. Biol. *371*, 235–245.

Brown, D.D., Martz, S.N., Binder, O., Goetz, S.C., Price, B.M.J., Smith, J.C., and Conlon, F.L. (2005). Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. Development *132*, 553–563.

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

Carreira, S., Dexter, T.J., Yavuzer, U., and Easty, D.J. (1998). Brachyury-Related Transcription Factor Tbx2 and Repression of the Melanocyte-Specific TRP-1 Promoter. Mol. Cell. Biol. *18*, 5099–5108.

Carver, E.A., Jiang, R., Lan, Y., Oram, K.F., and Gridley, T. (2001). The Mouse Snail Gene Encodes a Key Regulator of the Epithelial-Mesenchymal Transition. Mol. Cell. Biol. *21*, 8184–8188.

Casey, E.S., O'Reilly, M. a, Conlon, F.L., and Smith, J.C. (1998). The T-box transcription factor Brachyury regulates expression of eFGF through binding to a non-palindromic response element. Development *125*, 3887–3894.

Cerdan, C., McIntyre, B. a S., Mechael, R., Levadoux-Martin, M., Yang, J., Lee, J.B., and Bhatia, M. (2012). Activin A promotes hematopoietic fated mesoderm development through upregulation of brachyury in human embryonic stem cells. Stem Cells Dev. *21*, 2866–2877.

Chan, H.M., and La Thangue, N.B. (2001). p300/CBP proteins: HATs for transcriptional bridges and scaffolds. J. Cell Sci. *114*, 2363–2373.

Chapman, D.L., and Papaioannou, V.E. (1998). Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. Nature *391*, 695–697.

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

Chen, S., Ma, J., Wu, F., Xiong, L.-J., Ma, H., Xu, W., Lv, R., Li, X., Villen, J., Gygi, S.P., et al. (2012). The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. Genes Dev. *26*, 1364–1375.

Cheng, J., Blum, R., Bowman, C., Hu, D., Shilatifard, A., Shen, S., and Dynlacht, B.D. (2014). A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. Mol. Cell *53*, 979–992.

Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. J. Exp. Zool. *70*, 429–459.

Chiang, C., Litingtung, Y., Lee, E., Young, K.E., Corden, J.L., Westphal, H., and Beachy, P.A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. Nature *383*, 407–413.

Cho, Y.-W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G.R., Copeland, T.D., Kalkum, M., et al. (2007). PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. J. Biol. Chem. *282*, 20395–20406.

Choi, K., Kennedy, M., Kazarov, a, Papadimitriou, J.C., and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. Development *125*, 725–732.

Ciruna, B.G., and Rossant, J. (1999). Expression of the T-box gene Eomesodermin during early mouse development. Mech. Dev. *81*, 199–203.

Conlon, F.L., and Smith, J.C. (1999). Interference with brachyury function inhibits convergent extension, causes apoptosis, and reveals separate requirements in the FGF and activin signalling pathways. Dev. Biol. *213*, 85–100.

Conlon, F.L., Sedgwick, S.G., Weston, K.M., and Smith, J.C. (1996). Inhibition of Xbra transcription activation causes defects in mesodermal patterning and reveals autoregulation of Xbra in dorsal mesoderm. Development *122*, 2427–2435.

Conlon, F.L., Fairclough, L., Price, B.M., Casey, E.S., and Smith, J.C. (2001). Determinants of T box protein specificity. Development *128*, 3749–3758.

Costello, I., Pimeisl, I.-M., Dräger, S., Bikoff, E.K., Robertson, E.J., and Arnold, S.J. (2011). The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. Nat. Cell Biol. *13*, 1084–1091.

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. Proc. Natl. Acad. Sci. U. S. A. *107*, 21931–21936.

Dahle, Ø., Kumar, A., and Kuehn, M.R. (2010). Nodal signaling recruits the histone demethylase Jmjd3 to counteract polycomb-mediated repression at target genes. Sci. Signal. *3*, ra48.

Deng, C.X., Wynshaw-Boris, a, Shen, M.M., Daugherty, C., Ornitz, D.M., and Leder, P. (1994). Murine FGFR-1 is required for early postimplantation growth and axial organization. Genes Dev. *8*, 3045–3057.

Diez, R., Olivera-martinez, I., Goriely, A., Gale, E., Maden, M., and Storey, K. (2003). Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern , Neuronal Differentiation , and Segmentation during Body Axis Extension. Neuron *40*, 65–79.

Dobrovolskaia-Zavadskaia, N. (1927). Sur la mortification spontanee de la queue chez la souris nouveau-nee et sur l'existence d'un caractere (facteur) hereditaire "non-viable."C. R. Seances Soc. Biol. Fil. *97*, 114–116.

Downes, M., François, M., Ferguson, C., Parton, R.G., and Koopman, P. (2009). Vascular defects in a mouse model of hypotrichosis-lymphedema-telangiectasia syndrome indicate a role for SOX18 in blood vessel maturation. Hum. Mol. Genet. *18*, 2839–2850.

Drissen, R., Lindern, M. Von, Kolbus, A., Steinlein, P., Beug, H., Grosveld, F., Driegen, S., and Philipsen, S. (2005). The Erythroid Phenotype of EKLF-Null Mice : Defects in Hemoglobin Metabolism and Membrane Stability. Mol. Cell. Biol. *25*, 5205–5214.

Dweep, H., Sticht, C., Pandey, P., and Gretz, N. (2011). miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J. Biomed. Inform. *44*, 839–847.

Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. a, and McMahon, a P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell *75*, 1417–1430.

Evans, A.L., Faial, T., Gilchrist, M.J., Down, T., Vallier, L., Pedersen, R. a, Wardle, F.C., and Smith, J.C. (2012). Genomic targets of brachyury (T) in differentiating mouse embryonic stem cells. PLoS One *7*, e33346.

Galceran, J., Fariñas, I., Depew, M.J., Farin, I., Clevers, H., and Grosschedl, R. (1999). Wnt3a-/- like phenotype and limb deficiency in Lef1-/-Tcf1-/- mice. Genes Dev. *13*, 709–717.

Galceran, J., Hsu, S.C., and Grosschedl, R. (2001). Rescue of a Wnt mutation by an activated form of LEF-1: regulation of maintenance but not initiation of Brachyury expression. Proc. Natl. Acad. Sci. U. S. A. *98*, 8668–8673.

Garg, V., Kathiriya, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., and Hirayama-yamada, K. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature *21*, 443–447.

Garnett, A.T., Han, T.M., Gilchrist, M.J., Smith, J.C., Eisen, M.B., Wardle, F.C., and Amacher, S.L. (2009). Identification of direct T-box target genes in the developing zebrafish mesoderm. Development *136*, 749–760.

Gentsch, G.E., Owens, N.D.L., Martin, S.R., Piccinelli, P., Faial, T., Trotter, M.W.B., Gilchrist, M.J., and Smith, J.C. (2013). In Vivo T-Box Transcription Factor Profiling Reveals Joint Regulation of Embryonic Neuromesodermal Bipotency. Cell Rep. 1–12.

George, S.H.L., Gertsenstein, M., Vintersten, K., Korets-Smith, E., Murphy, J., Stevens, M.E., Haigh, J.J., and Nagy, A. (2007). Developmental and adult phenotyping directly from mutant embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *104*, 4455–4460.

Gertsenstein, M. (2011). Tetraploid Complementation Assay. In Advanced Protocols for Animal Transgenesis: An ISTT Manual, (Springer Protocols), pp. 357–375.

Ghosh, T.K., Song, F.F., Packham, E. a, Buxton, S., Robinson, T.E., Ronksley, J., Self, T., Bonser, A.J., and Brook, J.D. (2009). Physical interaction between TBX5 and MEF2C is required for early heart development. Mol. Cell. Biol. *29*, 2205–2218.

Gluecksohn-Schoenheimer, S. (1944). The development of normal and homozygous Brachy (T/T) mouse embryos in the extraembryonic coelom of the chick. Proc Natl Acad Sci U S A *30*, 134–140.

Grüneberg, H. (1958). Genetical Studies on the Skeleton of the Mouse. J. Embryol. Exp. Morphol. *6*, 424–443.

Haar, J.L., and Ackerman, G. a (1971). A phase and electron microscopic study of vasculogenesis and erythropoiesis in the yolk sac of the mouse. Anat. Rec. *170*, 199–223.

Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R., and Nakamura, Y. (2004). SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat. Cell Biol. *6*, 731–740.

Harada, H., Harada, Y., Brien, D.P.O., Dennis, S., Naeve, C.W., Downing, J.R., and Rice, D.S. (1999). HERF1 , a Novel Hematopoiesis-Specific RING Finger Protein, Is Required for Terminal Differentiation of Erythroid Cells. Mol. Cell. Biol. *19*, 3808–3815.

Hashiyama, M., Iwama, A., Ohshiro, K., Kurozumi, K., Yasunaga, K., Shimizu, Y., Masuho, Y., Yamaguchi, N., and Suda, T. (1996). Predominant Expression of a Receptor Tyrosine Kinase, TIE, in Hematopoietic Stem Cells and B Cells. Blood *87*, 93–101.

Heintzman, N.D., and Ren, B. (2009). Finding distal regulatory elements in the human genome. Curr. Opin. Genet. Dev. *19*, 541–549.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K. a, et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. *39*, 311–318.

Hemmati-Brivanlou, A., and Melton, D.A. (1992). A truncated activin receptor inhibits mesoderm induction and formation of axial structures in Xenopus embryos. Nature *359*, 609–614.

Herrmann, B.G. (1991). Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. Development *113*, 913–917.

Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. Nature *343*, 617–622.

Herz, H.-M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y.-H., Mickey, K., Voets, O., Verrijzer, C.P., and Shilatifard, A. (2012). Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Genes Dev. *26*, 2604–2620.

Herz, H.-M., Garruss, A., and Shilatifard, A. (2013). SET for life: biochemical activities and biological functions of SET domain-containing proteins. Trends Biochem. Sci. *38*, 621–639.

Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R., and Komuro, I. (2001). Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. Nat. Genet. *28*, 276–280.

Hodge, D., Coghill, E., Keys, J., Maguire, T., Hartmann, B., Mcdowall, A., Weiss, M., Grimmond, S., and Perkins, A. (2006). A global role for EKLF in definitive and primitive erythropoiesis. Blood *107*, 3359–3370.

Hon, G.C., Hawkins, R.D., and Ren, B. (2009). Predictive chromatin signatures in the mammalian genome. Hum. Mol. Genet. *18*, R195–201.

Hong, C.-J., and Hsueh, Y.-P. (2007). Cytoplasmic distribution of T-box transcription factor Tbr-1 in adult rodent brain. J. Chem. Neuroanat. *33*, 124–130.

Hong, S., Cho, Y.-W., Yu, L.-R., Yu, H., Veenstra, T.D., and Ge, K. (2007). Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. Proc. Natl. Acad. Sci. U. S. A. *104*, 18439–18444.

De Hoon, M.J.L., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453–1454.

Horb, M.E., and Thomsen, G.H. (1999). Tbx5 is essential for heart development. Development *126*, 1739–1751.

Hrabe de Angelis, M., McIntyre II, J., and Gossler, A. (1997). Maintenance of somite borders in mice requires the Delta homologue Dll1. Nature *386*, 717–721.

Hsueh, Y., Wang, T., Yang, F., and Sheng, M. (2000). Nuclear translocation and transcription regulation by the membrane-associated guanylate kinase CASK/LIN-2. Nature *404*, 298–302.

Hu, D., Gao, X., Morgan, M. a, Herz, H.-M., Smith, E.R., and Shilatifard, A. (2013). The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol. Cell. Biol. *33*, 4745–4754.

Huang, D.W., Sherman, B.T., and Lempicki, R. a (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Huber, T.L., Kouskoff, V., Fehling, H.J., Palis, J., and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature *432*, 625–630.

Hughes, C.M., Rozenblatt-Rosen, O., Milne, T. a, Copeland, T.D., Levine, S.S., Lee, J.C., Hayes, D.N., Shanmugam, K.S., Bhattacharjee, A., Biondi, C. a, et al. (2004). Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol. Cell *13*, 587–597.

Ihida-Stansbury, K., McKean, D.M., Gebb, S. a, Martin, J.F., Stevens, T., Nemenoff, R., Akeson, A., Vaughn, J., and Jones, P.L. (2004). Paired-related homeobox gene Prx1 is required for pulmonary vascular development. Circ. Res. *94*, 1507–1514.

Inman, K.E., and Downs, K.M. (2006a). Localization of Brachyury (T) in embryonic and extraembryonic tissues during mouse gastrulation. Gene Expr. Patterns *6*, 783–793.

Inman, K.E., and Downs, K.M. (2006b). Brachyury is required for elongation and vasculogenesis in the murine allantois. Development *133*, 2947–2959.

Inman, K.E., and Downs, K.M. (2007). The Murine Allantois : Emerging Paradigms in Development of the Mammalian Umbilical Cord and Its Relation to the Fetus. *258*, 237–258.

Isaac, a, Rodriguez-Esteban, C., Ryan, a, Altabef, M., Tsukui, T., Patel, K., Tickle, C., and Izpisúa-Belmonte, J.C. (1998). Tbx genes and limb identity in chick embryo development. Development *125*, 1867–1875.

Isaacs, H. V, Pownall, M.E., and Slack, J.M. (1994). eFGF regulates Xbra expression during Xenopus gastrulation. EMBO J. *13*, 4469–4481.

Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C.M., Nakamura, T., Mazo, A., Eisenbach, L., and Canaani, E. (2007). Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Mol. Cell. Biol. *27*, 1889–1903.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. PLoS Biol. *2*, e363.

Kallianpur, B.A.R., Jordan, J.E., and Brandt, S.J. (1994). The SCL/Tal-1 Gene is Expressed in Progenitors of Both the Hematopoietic and Vascular Systems During Embryogenesis. Blood *83*, 1200–1208.

Kartikasari, A.E.R., Zhou, J.X., Kanji, M.S., Chan, D.N., Sinha, A., Grapin-Botton, A., Magnuson, M. a, Lowry, W.E., and Bhushan, A. (2013). The histone demethylase Jmjd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation. EMBO J. *32*, 1393–1408.

Kauppi, M., Hilton, A. a, Metcalf, D., Ng, A.P., Hyland, C.D., Collinge, J.E., Kile, B.T., Hilton, D.J., and Alexander, W.S. (2012). Thrombocytopenia and erythrocytosis in mice with a mutation in the gene encoding the hemoglobin β minor chain. Proc. Natl. Acad. Sci. U. S. A. *109*, 576–581.

Kim, J., Guermah, M., McGinty, R.K., Lee, J.-S., Tang, Z., Milne, T. a, Shilatifard, A., Muir, T.W., and Roeder, R.G. (2009). RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell *137*, 459–471.

Kinder, S.J., Tsang, T.E., Quinlan, G. a, Hadjantonakis, a K., Nagy, a, and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. Development *126*, 4691–4701.

Kispert, A., and Herrmann, B.G. (1993). The Brachyury gene encodes a novel DNA binding protein. EMBO J. *12*, 3211–3220.

Kispert, A., and Herrmann, B.G. (1994). Immunohistochemical Analysis of the Brachyury Protein in Wild-type and Mutant Mouse Embryos. Dev. Biol. *161*, 179–193.

Kispert, a, Koschorz, B., and Herrmann, B.G. (1995). The T protein encoded by Brachyury is a tissue-specific transcription factor. EMBO J. *14*, 4763–4772.

Kulisz, A., and Simon, H.-G. (2008). An evolutionarily conserved nuclear export signal facilitates cytoplasmic localization of the Tbx5 transcription factor. Mol. Cell. Biol. *28*, 1553–1564.

Kung, A.L., Rebel, V.I., Bronson, R.T., Ch, L., Sieff, C.A., Livingston, D.M., and Yao, T. (2000). Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. Genes Dev. *14*, 272–277.

Kusch, T., Storck, T., Walldorf, U., and Reuter, R. (2002). Brachyury proteins regulate target genes through modular binding sites in a cooperative fashion. Genes Dev. *16*, 518–529.

Lan, F., Bayliss, P.E., Rinn, J.L., Whetstine, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E., et al. (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. Nature *449*, 689–694.

Laranjeiro, R., Alcobia, I., Neves, H., Gomes, A.C., Saavedra, P., Carvalho, C.C., Duarte, A., Cidadão, A., and Parreira, L. (2012). The notch ligand delta-like 4 regulates multiple stages of early hemato-vascular development. PLoS One *7*, e34553.

Lawson, K. a (1999). Fate mapping the mouse embryo. Int. J. Dev. Biol. *43*, 773–775.

Lawson, K.A., and Pedersen, R.A. (1992). Clonal analysis of cell fate during gastrulation and early neurulation in the mouse. In "Postimplantation Development in the Mouse" Ciba Found. Symp. 165, (Chichester: Wiley), pp. 3–26.

Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891–911.

Lee, J.-H., and Skalnik, D.G. (2008). Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. Mol. Cell. Biol. *28*, 609–618.

Lee, J., Kim, D.-H., Lee, S., Yang, Q.-H., Lee, D.K., Lee, S.-K., Roeder, R.G., and Lee, J.W. (2009). A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. Proc. Natl. Acad. Sci. U. S. A. *106*, 8513–8518.

Lee, J.-H., Tate, C.M., You, J.-S., and Skalnik, D.G. (2007a). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. J. Biol. Chem. *282*, 13419–13428.

Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.-P., Reinberg, D., Di Croce, L., and Shiekhattar, R. (2007b). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science *318*, 447–450.

Lee, S., Lee, J.W., and Lee, S.-K. (2012). UTX, a histone H3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. Dev. Cell *22*, 25–37.

Lewis, M.D., Miller, S. a, Miazgowicz, M.M., Beima, K.M., and Weinmann, A.S. (2007). T-bet's ability to regulate individual target genes requires the conserved T-box domain to recruit histone methyltransferase activity and a separate family member-specific transactivation domain. Mol. Cell. Biol. *27*, 8510–8521.

Li, J., Song, Y., Wang, Y., Luo, J., and Yu, W. (2013). MicroRNA-148a suppresses epithelial-to-mesenchymal transition by targeting ROCK1 in non-small cell lung cancer cells. Mol. Cell. Biochem. *380*, 277–282.

Limbourg, F.P., Takeshita, K., Radtke, F., Bronson, R.T., Chin, M.T., and Liao, J.K. (2005). Essential role of endothelial Notch1 in angiogenesis. Circulation *111*, 1826–1832.

Liu, J., Mercher, T., Scholl, C., Brumme, K., Gilliland, D.G., and Zhu, N. (2012). A functional role for the histone demethylase UTX in normal and malignant hematopoietic cells. Exp. Hematol. *40*, 487–98.e3.

Logan, M., Simon, H.G., and Tabin, C. (1998). Differential regulation of T-box and homeobox transcription factors suggests roles in controlling chick limb-type identity. Development *125*, 2825–2835.

Lolas, M., Valenzuela, P.D.T., 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.

MacDonald, S.T., Bamforth, S.D., Bragança, J., Chen, C.-M., Broadbent, C., Schneider, J.E., Schwartz, R.J., and Bhattacharya, S. (2013). A cell-autonomous role of Cited2 in controlling myocardial and coronary vascular development. Eur. Heart J. *34*, 2557–2565.

Macmurray, A., and Shin, H. (1988). The Antimorphic Nature of the Tc Allele at the Mouse T Locus. Genetics *120*, 545–550.

Mankoo, B.S., Skuntz, S., Harrigan, I., Grigorieva, E., Candia, A., Wright, C.V.E., Arnheiter, H., and Pachnis, V. (2003). The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. Development *130*, 4655–4664.

Mansouri, A., Yokota, Y., Wehr, R., Copeland, N.G., Jenkins, N.A., and Gruss, P. (1997). Paired-related murine homeobox gene expressed in the developing sclerotome, kidney, and nervous system. Dev. Dyn. *210*, 53–65.

Marcellini, S., Technau, U., Smith, J.., and Lemaire, P. (2003). Evolution of Brachyury proteins: identification of a novel regulatory domain conserved within Bilateria. Dev. Biol. *260*, 352–361.

Martin, B.L., and Kimelman, D. (2008). Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. Dev. Cell *15*, 121–133.

Martin, B.L., and Kimelman, D. (2010). Brachyury establishes the embryonic mesodermal progenitor niche. Genes Dev. *24*, 2778–2783.

Martin, B.L., and Kimelman, D. (2012). Canonical Wnt signaling dynamically controls multiple stem cell fate decisions during vertebrate body formation. Dev. Cell *22*, 223–232.

Martin, P., and Cockroft, D.L. (2008). Culture of Postimplantation Mouse Embryos. In Methods in Molecular Biology, Vol. 461: Molecular Embryology: Methods and Protocols, P. Sharpe, and I. Mason, eds. (Totowa, NJ), pp. 7–22.

Miller, S. a, Huang, A.C., Miazgowicz, M.M., Brassil, M.M., and Weinmann, A.S. (2008). Coordinated but physically separable interaction with H3K27-demethylase and H3K4-methyltransferase activities are required for T-box protein-mediated activation of developmental gene expression. Genes Dev. *22*, 2980–2993.

Miller, S. a, Mohn, S.E., and Weinmann, A.S. (2010). Jmjd3 and UTX Play a Demethylase-Independent Role in Chromatin Remodeling to Regulate T-Box Family Member-Dependent Gene Expression. Mol. Cell *40*, 594–605.

Morales Torres, C., Laugesen, A., and Helin, K. (2013). Utx is required for proper induction of ectoderm and mesoderm during differentiation of embryonic stem cells. PLoS One *8*, e60020.

Moreno, T.A., and Kintner, C. (2004). Regulation of Segmental Patterning by Retinoic Acid Signaling during Xenopus Somitogenesis. Dev. Cell *6*, 205–218.

Morikawa, Y., and Cserjesi, P. (2004). Extra-embryonic vasculature development is regulated by the transcription factor HAND1. Development *131*, 2195–2204.

Müller, C.W., and Herrmann, B.G. (1997). Crystallographic structure of the T domain-DNA complex of the Brachyury transcription factor. Nature *389*, 884–888.

Muyrers, J.P., Zhang, Y., Testa, G., and Stewart, a F. (1999). Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Res. *27*, 1555–1557.

Nagy, A., Gertsenstein, M., Vintersten, K., and Behringer, R.R. (2003). Manipulating the Mouse Embryo : A Laboratory Manual (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory).

Naiche, L. a., and Papaioannou, V.E. (2003). Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois. Development *130*, 2681–2693.

Naiche, L. a, Harrelson, Z., Kelly, R.G., and Papaioannou, V.E. (2005). T-box genes in vertebrate development. Annu. Rev. Genet. *39*, 219–239.

Naiche, L. a, Holder, N., and Lewandoski, M. (2011). FGF4 and FGF8 comprise the wavefront activity that controls somitogenesis. Proc. Natl. Acad. Sci. U. S. A. *108*, 4018–4023.

Nilson, D.G., Sabatino, D.E., Bodine, D.M., and Gallagher, P.G. (2006). Major erythrocyte membrane protein genes in EKLF-deficient mice. Exp. Hematol. *34*, 705–712.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P., and Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes Dev. *16*, 479–489.

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the ELKF gene. Nature *375*, 316–318.

Ohtani, K., Zhao, C., Dobreva, G., Manavski, Y., Kluge, B., Braun, T., Rieger, M. a, Zeiher, A.M., and Dimmeler, S. (2013). Jmjd3 controls mesodermal and cardiovascular differentiation of embryonic stem cells. Circ. Res. *113*, 856–862.

Packham, E.A., and Brook, J.D. (2003). T-box genes in human disorders. Hum. Mol. Genet. *12*, 37–44.

Pennimpede, T., Proske, J., König, A., Vidigal, J. a, Morkel, M., Bramsen, J.B., Herrmann, B.G., and Wittler, L. (2012). In vivo knockdown of Brachyury results in skeletal defects

and urorectal malformations resembling caudal regression syndrome. Dev. Biol. *372*, 55–67.

Pereira, L. a, Wong, M.S., Lim, S.M., Sides, A., Stanley, E.G., and Elefanty, A.G. (2011). Brachyury and related Tbx proteins interact with the Mixl1 homeodomain protein and negatively regulate Mixl1 transcriptional activity. PLoS One *6*, e28394.

Perkins, A.C., Sharpe, A.H., and Orkin, S.H. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor ELKF. Nature *375*, 318–322.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. a, Flynn, R. a, and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Ramirez-Solis, R., Davis, A.C., and Bradley, A. (1993). Gene targeting in embryonic stem cells. Methods Enzymol. *225*, 855–878.

Rank, G., Sutton, R., Marshall, V., Lundie, R.J., Caddy, J., Romeo, T., Fernandez, K., Mccormack, M.P., Cooke, B.M., Foote, S.J., et al. (2009). Novel roles for erythroid Ankyrin-1 revealed through an ENU-induced null mouse mutant. Blood *113*, 3352–3362.

Rashbass, P., Cooke, L.A., Herrmann, B.G., and Beddington, R.S. (1991). A cell autonomous function of Brachyury in T/T embryonic stem cell chimaeras. Nature *353*, 348–351.

Reamon-Buettner, S.M., and Borlak, J. (2004). TBX5 mutations in non-Holt-Oram syndrome (HOS) malformed hearts. Hum. Mutat. *24*, 104.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S. a, Goodnough, L.H., Helms, J. a, Farnham, P.J., Segal, E., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311–1323.

Robb, L., Lyons, I., Li, R., Hartley, L., Köntgen, F., Harvey, R.P., Metcalf, D., and Begley, C.G. (1995). Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. Proc. Natl. Acad. Sci. U. S. A. *92*, 7075–7079.

Rodriguez-Esteban, C., Tsukui, T., Yonei, S., Magallon, J., Tamura, K., and Izpisua Belmonte, J.C. (1999). The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. Nature *398*, 814–818.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C., et al. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. Nature *404*, 95–99.

Ryan, K., Garrett, N., Mitchell, a, and Gurdon, J.B. (1996). Eomesodermin, a key early gene in Xenopus mesoderm differentiation. Cell *87*, 989–1000.

Sabin, F.R. (1920). Studies on the origin of blood vessels and red blood corpuscles as seen in the living blastoderm of chicks during the second day of incubation. Contrib. Embryol. *9*, 213–262.

Sadlon, T.J., Dell'Oso, T., Surinya, K.H., and May, B.K. (1999). Regulation of erythroid 5-aminolevulinate synthase expression during erythropoiesis. Int. J. Biochem. Cell Biol. *31*, 1153–1167.

Sakamoto, N., Naito, Y., Oue, N., Sentani, K., Uraoka, N., Zarni Oo, H., Yanagihara, K., Aoyagi, K., Sasaki, H., and Yasui, W. (2014). MicroRNA-148a is downregulated in gastric cancer, targets MMP7, and indicates tumor invasiveness and poor prognosis. Cancer Sci. *105*, 236–243.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning : a laboratory manual (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory).

De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell *130*, 1083–1094.

De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., Burgold, T., Austenaa, L., Bucci, G., Caganova, M., Notarbartolo, S., Casola, S., et al. (2009). Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. EMBO J. *28*, 3341–3352.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O. (2010). Inhibition of Eukaryotic Translation Elongation by Cycloheximide and Lactimidomycin. Nat Chem Biol *6*, 209–217.

Schröder, A.G. (2010). Differenzierung embryonaler Stammzellen der Maus zu paraxialem Mesoderm in vitro. Freie Universität Berlin.

Schuh, A.C., Faloon, P., Hu, Q.L., Bhimani, M., and Choi, K. (1999). In vitro hematopoietic and endothelial potential of flk-1-/- embryonic stem cells and embryos. Proc Natl Acad Sci U S A *96*, 2159–2164.

Schulte-Merker, S., and Smith, J.C. (1995). Mesoderm formation in response to Brachyury requires FGF signalling. Curr. Biol. *5*, 62–67.

Schulte-Merker, S., Ho, R.K., Herrmann, B.G., and Nüsslein-Volhard, C. (1992). The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. Development *116*, 1021–1032.

Seenundun, S., Rampalli, S., Liu, Q.-C., Aziz, A., Palii, C., Hong, S., Blais, A., Brand, M., Ge, K., and Dilworth, F.J. (2010). UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. EMBO J. *29*, 1401–1411.

Sengupta, A., Chakraborty, S., Paik, J., Yutzey, K.E., and Evans-Anderson, H.J. (2012). FoxO1 is required in endothelial but not myocardial cell lineages during cardiovascular development. Dev. Dyn. *241*, 803–813.

Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X., Breitman, M.L., and Schuh, A.C. (1995). Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature *376*, 62–66.

Shedlovsky, A., King, T.R., and Dove, W.F. (1988). Saturation germ line mutagenesis of the murine t region including a lethal allele at the quaking locus. Proc Natl Acad Sci U S A *85*, 180–184.

Shivdasani, R.A., Mayer, E.L., and Orkin, S.H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature *373*, 432–434.

Shpargel, K.B., Sengoku, T., Yokoyama, S., and Magnuson, T. (2012). UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. PLoS Genet. *8*, e1002964.

Smith, J.C., Price, B.M.J., Green, J.B.A., Ridgeway, T., Hill, M., and Aa, L.N.W. (1991). Expression of a Xenopus Homolog of Brachyury (T) Is an Immediate-Early Response to Mesoderm Induction. Cell *67*, 79–87.

Stott, D., Kispert, a, and Herrmann, B.G. (1993). Rescue of the tail defect of Brachyury mice. Genes Dev. *7*, 197–203.

Sun, X., Meyers, E.N., Lewandoski, M., and Martin, G.R. (1999). Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. Genes Dev. *13*, 1834–1846.

Szeto, D.P., Griffin, K.J.P., and Kimelman, D. (2002). HrT is required for cardiovascular development in zebrafish. Development *129*, 5093–5101.

Tada, M., Casey, E.S., Fairclough, L., and Smith, J.C. (1998). Bix1, a direct target of Xenopus T-box genes, causes formation of ventral mesoderm and endoderm. Development *125*, 3997–4006.

Takada, S., Stark, K.L., Shea, M.J., Vassileva, G., McMahon, J. a, and McMahon, a P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev. *8*, 174–189.

Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D.M., Lovell-Badge, R., Papaioannou, V.E., and Kondoh, H. (2011). Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature *470*, 394–398.

Takeuchi, J.K., Koshiba-Takeuchi, K., Matsumoto, K., Vogel-Höpker, a, Naitoh-Matsuo, M., Ogura, K., Takahashi, N., Yasuda, K., and Ogura, T. (1999). Tbx5 and Tbx4 genes determine the wing/leg identity of limb buds. Nature *398*, 810–814.

Tanaka, Y., Naruse, I., Hongo, T., Xu, M., Nakahata, T., Maekawa, T., and Ishii, S. (2000). Extensive brain hemorrhage and embryonic lethality in a mouse null mutant of CREB-binding protein. Mech. Dev. *95*, 133–145.

Thieme, S., Gyárfás, T., Richter, C., Özhan, G., Fu, J., Alexopoulou, D., Muders, M.H., Michalk, I., Jakob, C., Dahl, A., et al. (2013). The histone demethylase UTX regulates stem cell migration and hematopoiesis. Blood *121*, 2462–2473.

Tie, F., Banerjee, R., Conrad, P. a, Scacheri, P.C., and Harte, P.J. (2012). Histone demethylase UTX and chromatin remodeler BRM bind directly to CBP and modulate acetylation of histone H3 lysine 27. Mol. Cell. Biol. *32*, 2323–2334.

Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., and Nicolas, J.-F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev. Cell *17*, 365–376.

Vallette-Kasic, S., Couture, C., Balsalobre, A., Gauthier, Y., Metherell, L., Dattani, M., and Drouin, J. (2007). The TPIT gene mutation M86R associated with isolated adrenocorticotropin deficiency interferes with protein: protein interactions. J. Clin. Endocrinol. Metab. *92*, 3991–3999.

Vandamme, J., Lettier, G., Sidoli, S., Di Schiavi, E., Nørregaard Jensen, O., and Salcini, A.E. (2012). The C. elegans H3K27 demethylase UTX-1 is essential for normal development, independent of its enzymatic activity. PLoS Genet. *8*, e1002647.

Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W.M.P., van Schaik, F.M. a, Varier, R. a, Baltissen, M.P. a, Stunnenberg, H.G., Mann, M., and Timmers, H.T.M. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell *131*, 58–69.

Vermot, J., and Pourquie, O. (2005). Retinoic acid coordinates somitogenesis and left – right patterning in vertebrate embryos. Nature *435*, 215–220.

Vidigal, J.A., and Ventura, A. (2012). Embryonic stem cell miRNAs and their roles in development and disease. Semin. Cancer Biol. *22*, 428–436.

Vidigal, J. a, Morkel, M., Wittler, L., Brouwer-Lehmitz, A., Grote, P., Macura, K., and Herrmann, B.G. (2010). An inducible RNA interference system for the functional dissection of mouse embryogenesis. Nucleic Acids Res. *38*, e122.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J. a, Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854–858.

Visvader, J.E., Fujiwara, Y., and Orkin, S.H. (1998). Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. Genes Dev. *12*, 473–479.

Wahl, M.B., Deng, C., Lewandoski, M., and Pourquié, O. (2007). FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. Development *134*, 4033–4041.

Wang, C., Lee, J.-E., Cho, Y.-W., Xiao, Y., Jin, Q., Liu, C., and Ge, K. (2012). UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. Proc. Natl. Acad. Sci. U. S. A. *109*, 15324–15329.

Wang, H., Cao, R., Xia, L., Erdjument-bromage, H., Borchers, C., Tempst, P., Zhang, Y., Hill, C., and Carolina, N. (2001). Purification and Functional Characterization of a Histone H3-Lysine 4-Specific Methyltransferase. Mol. Cell *8*, 1207–1217.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R. a, Gupta, R. a, et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120–124.

Wang, Z., Zang, C., Rosenfeld, J. a, Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903.

Warren, a J., Colledge, W.H., Carlton, M.B., Evans, M.J., Smith, a J., and Rabbitts, T.H. (1994). The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. Cell *78*, 45–57.

Washkowitz, A.J., Gavrilov, S., Begum, S., and Papaioannou, V.E. (2012). Diverse functional networks of Tbx3 in development and disease. Wiley Interdiscip. Rev. Syst. Biol. Med. *4*, 273–283.

Watt, S.M., Gschmeissner, S.E., and Bates, P.A. (1995). PECAM-1 : Its Expression and Function as a Cell Adhesion Molecule on Hemopoietic and Endothelial Cells. Leuk. Lymphoma *17*, 229–244.

Weinstein, D.C., Ruiz i Altaba, a, Chen, W.S., Hoodless, P., Prezioso, V.R., Jessell, T.M., and Darnell, J.E. (1994). The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. Cell *78*, 575–588.

Welstead, G.G., Creyghton, M.P., Bilodeau, S., Cheng, A.W., and Markoulaki, S. (2012). X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. Proc Natl Acad Sci U S A *109*, 13004–13009.

Wilkinson, D.G., Bhatt, S., and Herrmann, B.G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. Nature *343*, 657–659.

Wilson, V., Manson, L., Skarnes, W.C., and Beddington, R.S. (1995). The T gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. Development *121*, 877–886.

Winnier, G., Blessing, M., Labosky, P. a, and Hogan, B.L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev. *9*, 2105–2116.

Wu, M., Wang, P.F., Lee, J.S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. (2008). Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Mol. Cell. Biol. *28*, 7337–7344.

Xi, H., Shulha, H.P., Lin, J.M., Vales, T.R., Fu, Y., Bodine, D.M., McKay, R.D.G., Chenoweth, J.G., Tesar, P.J., Furey, T.S., et al. (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. PLoS Genet. *3*, e136.

Yamada, Y., Pannell, R., Forster, a, and Rabbitts, T.H. (2000). The oncogenic LIM-only transcription factor Lmo2 regulates angiogenesis but not vasculogenesis in mice. Proc. Natl. Acad. Sci. U. S. A. *97*, 320–324.

Yamaguchi, T.P., Harpal, K., Henkemeyer, M., and Rossant, J. (1994). Fgfr-1 Is Required for Embryonic Growth and Mesodermal Patterning During Mouse Gastrulation. Genes Dev. *8*, 3032–3044.

Yamaguchi, T.P., Takada, S., Yoshikawa, Y., Wu, N., and McMahon, a. P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. Genes Dev. *13*, 3185–3190.

Yanagisawa, K.O. (1990). Does the T gene determine the anteroposterior axis of a mouse embryo? Japanese J. Genet. *65*, 287–297.

Yanagisawa, K.O., Fujimoto, H., and Urushihara, H. (1981). Effects of the brachyury (T) mutation on morphogenetic movement in the mouse embryo. Dev. Biol. *87*, 242–248.

Yao, T.P., Oh, S.P., Fuchs, M., Zhou, N.D., Ch'ng, L.E., Newsome, D., Bronson, R.T., Li, E., Livingston, D.M., and Eckner, R. (1998). Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. Cell *93*, 361–372.

Ye, T., Krebs, A.R., Choukrallah, M.-A., Keime, C., Plewniak, F., Davidson, I., and Tora, L. (2011). seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res. *39*, e35.

Yoshikawa, Y., Fujimori, T., McMahon, a P., and Takada, S. (1997). Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Dev. Biol. *183*, 234–242.

Zhao, J., Sun, B.K., Erwin, J. a, Song, J.-J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science *322*, 750–756.

Zhu, B., Zheng, Y., Pham, A.-D., Mandal, S.S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005). Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Mol. Cell *20*, 601–611.