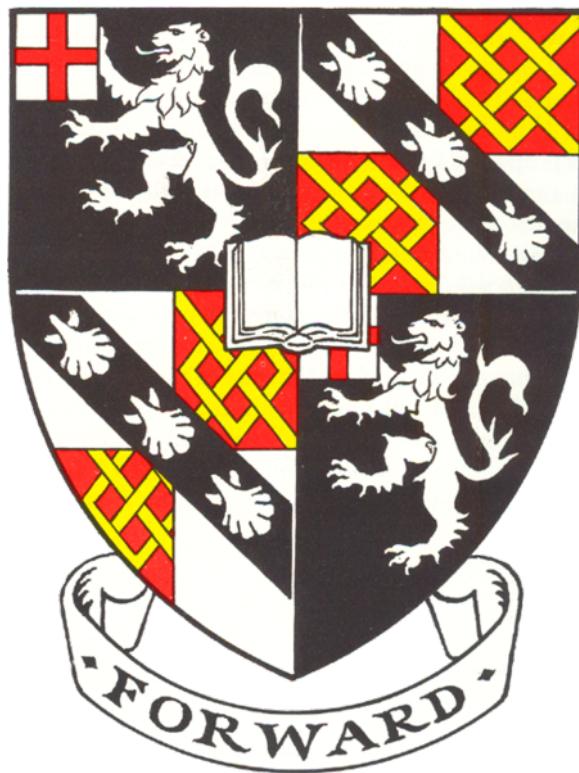


Investigating Human Primordial Germ Cell Specification  
by Manipulation of Regulatory Proteins

**Merrick de Forest Pierson Smela**

Wellcome Trust Cancer Research UK Gurdon Institute  
Department of Physiology, Development, and Neuroscience  
**Churchill College**



University of Cambridge

August 2019

**This dissertation is submitted for the degree of Master of Philosophy.**

## Preface

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. The NANOS3-tdTomato and PRDM14-AID parental cell lines were generated by Dr. Anastasiya Sybirna. The inducible TIR1 vectors were generated in collaboration with Dr. Frederick Wong. Some PCR primers were designed by Drs. Naoko Irie, Anastasiya Sybirna, Walfred Tang, or Frederick Wong. This thesis does not exceed the word limit set by the Degree Committee for the Faculty of Biology.

## Acknowledgements

I would like to thank the Surani group for their helpful support. In particular, Dr. Anastasiya Sybirna introduced me to the AID system and was an excellent mentor to me, and Dr. Frederick Wong also provided important advice on some of my experiments. I would additionally like to acknowledge the Winston Churchill foundation for funding.

## Summary

Although the early human germline is inaccessible for direct study, primordial germ cell-like cells (PGCLCs) can be derived *in vitro* from human pluripotent stem cells. This enables study of human germ cell specification by manipulating the levels of regulatory proteins using genetic tools, including the auxin-inducible degron (AID) and ProteoTuner systems. With these tools, I altered levels of the transcription factor SOX15, which had been proposed to be necessary for germ cell specification. I found that SOX15 was dispensable for establishing germ cell identity but may have a role in maintaining it. I attempted to similarly manipulate the NANOS1 RNA-binding protein, but contrary to previous expectations, it was not expressed in PGCLCs. I furthermore demonstrated a variation on the AID system with decreased leakiness and applied it to the known germline specifier SOX17. Finally, I generated KLF4-AID cell lines which show defects in germ cell specification upon KLF4 depletion and may be useful in future studies of KLF4 function.

# Table of Contents

Title.....	i
Preface, Summary, and Acknowledgements.....	ii
Table of Contents.....	iii
Glossary of Abbreviations.....	iv
Introduction.....	1
Methods.....	5
Results.....	10
SOX15.....	10
NANOS1.....	19
Inducible TIR1 and SOX17-AID.....	20
KLF4-AID.....	23
Discussion.....	26
References.....	31

## Glossary of Abbreviations

AID	Auxin-inducible degron
AP	Alkaline phosphatase
BMP	Bone morphogenetic protein
DAPI	4',6-Diamidino-2-phenylindole
DD	Destabilization domain
DMEM	Dulbecco's modified Eagle's medium
EB	Embryoid body
EGF	Epidermal growth factor
ER	Estrogen receptor
ESC	Embryonic stem cell
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
HMG	High-mobility group
IAA	Indole-3-acetic acid
IRES	Internal ribosomal entry site
KZNF	Krüppel-associated box zinc finger
LIF	Leukemia inhibitory factor
MEF	Mouse embryonic fibroblast
N3tdT	NANOS3-T2A-tdTomato
PBS	Phosphate-buffered saline
PGC	Primordial germ cell
PGCLC	Primordial germ cell-like cell
ROCK	Rho-associated kinase
SCF	Stem cell factor
TE	Transposable element
TGF $\beta$	Transforming growth factor beta
UTR	Untranslated region

## Introduction

The human germline is a continuous lineage of cells tracing back through the generations to the earliest life. Understanding the germline is crucial due to its roles in development, inheritance, and evolution.<sup>1</sup> During embryonic development, the earliest cells to be committed to germ cell fate are known as primordial germ cells (PGCs). These cells are specified shortly after implantation of the blastocyst in the uterus, but just before gastrulation.<sup>2</sup> This stage is inaccessible for direct study in humans, although experiments in mice and other animals have shown that PGCs are specified in response to signaling by bone morphogenetic protein (BMP) paracrine factors. After specification, PGCs migrate to the developing gonad. During this process of migration, the epigenome is reset, leading to global DNA hypomethylation.<sup>3,4</sup> Once in the gonad, PGCs eventually develop into oogonia or spermatogonia, depending on the sex of the fetus. These cells then undergo further differentiation and meiosis to form mature gametes.

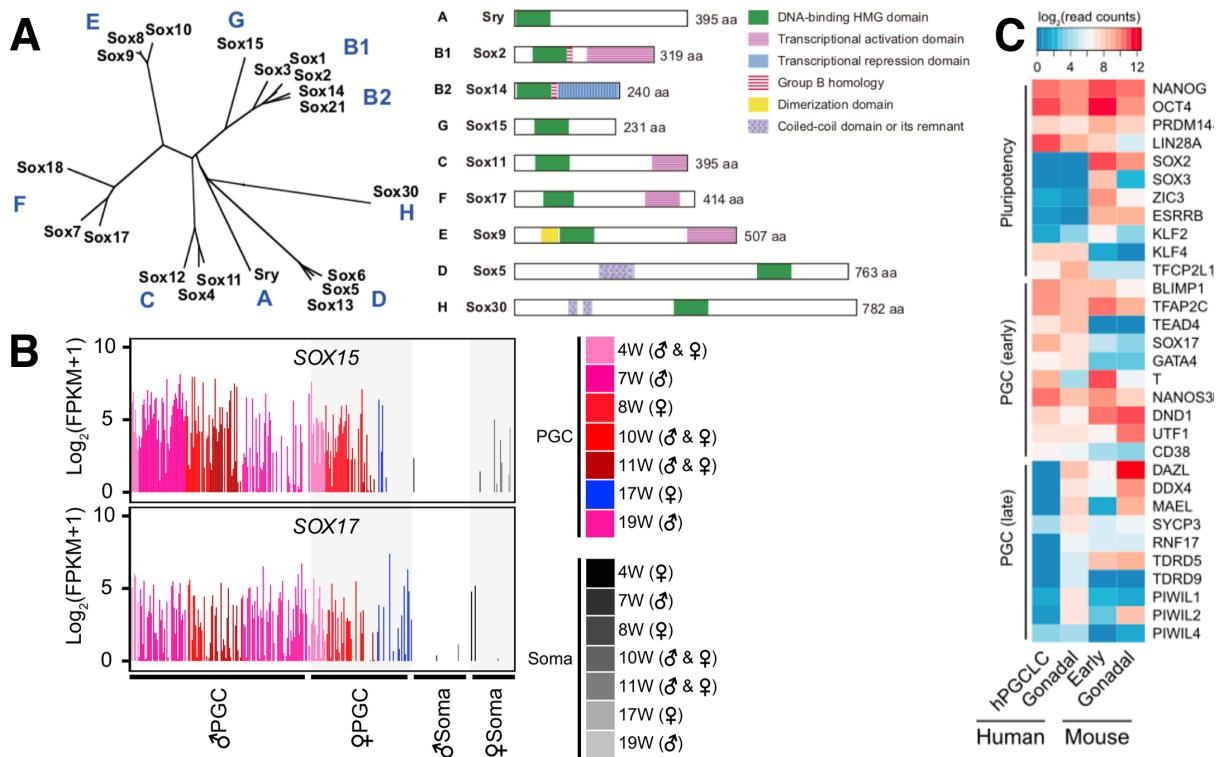
Despite decades of research, the origins of human germ cells are still only partially understood. PGCs are specified during the “black box” of early embryonic development, and model organisms such as mice fail to fully capture many important aspects of human development, including PGC specification. How, then, can human PGCs be studied? The answer is to use a model system in which stem cells differentiate into PGC-like cells (PGCLCs) in response to signaling by BMP and other cytokines. In order to establish germline competency, human embryonic stem cells (hESCs) can be cultured in a specialized growth medium which allows continual maintenance of a germline competent state. Alternatively, stem cells in conventional medium can be initially differentiated towards mesendoderm, and then treated with BMP during a brief window of competency. Both systems produce cells very similar to pre-migratory hPGCs.<sup>5</sup> Mouse PGCLCs can further develop in vitro when co-cultured with E12.5 ovarian somatic cells and treated with a defined set of cytokines and hormones,<sup>6</sup> and a few of the cells can even develop into functional oocytes. Although a similar system has been recently reported to produce human oogonia using coculture with mouse fetal ovarian somatic tissue,<sup>7</sup> the precise factors required for human PGC maturation and epigenetic resetting remain unknown, and the human PGCLC model system does not currently allow large-scale study of cells in the post-migratory state.

Solving this problem requires understanding the differences in gene regulation between human and mouse PGCs. In mice, PGC fate is determined by a core network of three transcription factors: BLIMP1, PRDM14, and AP2γ.<sup>8,9</sup> PRDM14 activates germ cell-specific genes, BLIMP1 represses somatic genes and activates AP2γ, and AP2γ cooperates with the former two factors to enhance their effects.

However, the core PGC gene regulatory network in humans does not completely overlap with its murine counterpart. Although BLIMP1, PRDM14 and AP2γ are still important, they act downstream of SOX17, which is the crucial specifier of germ cell fate.<sup>10</sup> Overexpression of SOX17 and BLIMP1 in hESCs is sufficient to specify germline fate in the absence of cytokines.<sup>11</sup> Differences in gene regulation may reflect the fact that the embryos of mice and other rodents form an egg cylinder after implantation, whereas most other placental mammals develop as a bilaminar disc. The pluripotency network also shows differences between mouse and human.<sup>12</sup> Indeed, pigs, which are not closely related to primates, rely on SOX17 for germline specification, suggesting that the mouse gene regulatory network is likely to be specific to rodents.<sup>11</sup>

SOX17 is not the only regulatory gene that differs between mouse and human PGCs. SOX15, another member of the SOX family, is strongly expressed in human PGCs, but is expressed in mouse gonadal soma only at more advanced stages of development (E11.5 and later).<sup>13</sup> Furthermore, busulfan treatment, which destroys germ cells, had no effect on mouse SOX15 *in situ* hybridization staining in embryonic testis, implying that expression is mostly confined to gonadal somatic cells. In contrast, a recent single-cell transcriptomics study found that SOX15 was more strongly and homogenously expressed than SOX17 among human PGCs before 10 weeks' gestation,<sup>3</sup> and the authors claimed that SOX15 is probably more functionally important for hPGCs *in vivo*.

In both mice and humans, SOX15 is also highly expressed in naïve ESCs, placenta, and muscle satellite cells, with lower expression levels present in many other tissues.<sup>14</sup> The mouse knockout phenotype is relatively mild, with the only known symptom being impaired muscle regeneration after injury.<sup>15</sup> Notably, fertility is normal, ruling out a crucial role for SOX15 in mouse PGC development. SOX15 is the only group G SOX factor in mammals, although its DNA-binding HMG domain is similar to that of group B SOX factors such as SOX2 (Figure 1A). In mouse ESCs, both SOX2 and SOX15 can partner with OCT3/4 to activate



**Figure 1. A:** Phylogenetic analysis of the Sox family of transcription factors, showing subgroups A–H.<sup>16</sup> SOX15 is most similar to subgroup SOXB1 (SOX1, SOX2, and SOX3), especially in its DNA-binding domain. **B:** Single-cell RNA-seq data from Guo *et al.*<sup>3</sup> show that SOX15 is more homogeneously expressed than SOX17 in hPGCs, especially before week 11. **C:** Heat map comparing regulatory gene expression in human and mouse PGCs,<sup>4</sup> and showing differences in SOX2, SOX17, and KLF4.

transcription of target genes.<sup>17</sup> Interestingly, SOX2 is expressed in mouse but not human PGCs (Figure 1B),<sup>4,18</sup> whereas SOX15 shows the opposite pattern.

KLF4 is yet another transcription factor whose expression differs between human and mouse PGCLCs (Figure 1B). This zinc-finger protein is expressed in the human naïve pluripotent state,<sup>19</sup> and is notable for being one of Yamanaka's core reprogramming factors. During mouse PGC specification, KLF4 is inactive due to repression by BLIMP1.<sup>2,20,21</sup> In contrast, KLF4 is highly expressed in human PGCs.<sup>2</sup> The functional role of KLF4 during human germline development is unclear. In naïve ESCs, it is known to bind and activate enhancers within transposable elements (TEs). This directly activates TE expression, but also activates expression of KZNF (Krüppel-associated box zinc finger) repressor proteins which results in indirect TE repression.<sup>19</sup> Conceivably, the role of KLF4 in PGCs may be similar to its role in naïve ESCs.

In this work, I have investigated the functional roles of regulatory genes, including SOX15 and KLF4, in human germline specification. To do this, I manipulated levels of their

proteins using the auxin-inducible degron (AID) and ProteoTuner systems. For studying PGCLC specification, protein-level control is required to achieve the necessary temporal resolution, and also because translational regulation of mRNAs in germline cells<sup>22–24</sup> may confound efforts to manipulate expression at the RNA level. The AID system involves fusing a short degron peptide to the protein of interest, and also expressing a TIR1 E3 ubiquitin ligase.<sup>25</sup> In the presence of auxin (indole-3-acetic acid, IAA), TIR1 will ubiquitylate the degron, leading to destruction of the target protein by the proteasome. This happens rapidly, causing complete depletion within one hour. The ProteoTuner system can give similarly rapid overexpression of a target protein.<sup>26</sup> The protein is expressed as a fusion with a small destabilizing domain (DD) based on an FKBP12 mutant. This domain normally causes protein unfolding and degradation. When the ligand Shield1 is added, the domain is stabilized and the protein quickly accumulates.

Using these tools, I manipulated the levels of SOX15 during PGCLC specification. I found that SOX15 is unnecessary for establishing PGCLC identity, and instead may play a role in maintaining it. Next, I identified some transcriptional effects of SOX15 depletion and overexpression. I also applied the AID system to the NANOS1 RNA-binding protein, which is highly expressed in PGCLCs at the RNA level, but I found that its protein was absent. Additionally, I developed an inducible AID system usable on targets where the normal AID system suffers from leakiness, such as SOX17. In ongoing work, I am using the AID system to investigate the role of KLF4 in PGCLC specification, as well as in naïve pluripotency. Preliminary results indicate that KLF4 is an important regulator of germline identity, although it is not strictly essential in the initial stage of specification. Overall, my research illustrates the utility of genetic tools which rapidly alter protein levels and provides insights into genetic regulation of germline specification.

# Methods

## Cell culture:

hESCs (WIS2 NANOS3-T2A-tdTomato)<sup>11</sup> were cultured in 4i medium,<sup>10,27</sup> containing cytokines TGFβ, bFGF, and LIF, as well as four small-molecule inhibitors for kinases MAPK, MEK, JNK, and GSK3. This medium allows hESCs to be continually maintained in a germline competent state.<sup>10</sup> The hESCs were grown on a layer of irradiated CF1 MEFs (Applied Stem Cell). The MEFs were plated at approximately 15,000 cells/cm<sup>2</sup> on gelatin-coated plates in DMEM supplemented with 10% FBS, 100 U/mL penicillin, and 0.1 mg/mL streptomycin. Medium was changed daily for hESCs. 4i cells were passaged using 0.25% trypsin/EDTA with ROCK inhibitor (10 μM Y-27632, Tocris Bioscience) added to the medium. TCam-2 cells were grown in RPMI + GlutaMAX (Thermo Fisher) supplemented with 10% FBS, 100 U/mL penicillin, and 0.1 mg/mL streptomycin, and passaged with 0.25% trypsin/EDTA. All cells were maintained in an incubator at 37°C and 5% CO<sub>2</sub>. Cell lines used in the experiments tested negative for mycoplasma.

## Generation of mutant cell lines:

For CRISPR/Cas9 experiments, gRNAs were chosen using the online tool at crispr.mit.edu. Oligos were annealed and cloned into eSpCas9(1.1) vector<sup>28</sup> digested with BbsI. Homology arms were amplified by PCR from genomic DNA of the target cell line, with a point mutation introduced to remove the stop codon and CRISPR PAM. The donor plasmids were assembled using InFusion cloning (Clontech). For ProteoTuner overexpression, SOX15 cDNA was cloned into the PB-EF1-DD-IRES-Puro backbone using InFusion. For RNAi experiments, miRNAs were designed using the online tool at rnaidesigner.thermofisher.com. Oligos were annealed and cloned into the pPB-Ins-TRE3G-d2EGFP-miR-pA-EF1a-TET3G-iNeo-Ins plasmid<sup>10</sup> digested with Esp3I. All oligos used for cloning and sequencing are listed in Table 1. Plasmids were delivered using Lipofectamine Stem reagent (Invitrogen) according to manufacturer's instructions. After 48 hours, selection was begun with puromycin (0.5 μg/mL), hygromycin (50 μg/mL), G418 (200 μg/mL for ESCs, 500 μg/mL for TCam-2), and/or FIAU (200 nM) as appropriate, and continued until colonies were picked.

**Table 1:** oligos for cloning and sequencing

Oligo Type	Description	Number	Sequence (5' to 3')
RNAi oligos	SOX15 RNAi A fwd	MPS080	TGCTGAGCTTGTCTTCAACCTGTTGGCCACTGACTGACAACAGGTTAACAGGCTC
	SOX15 RNAi A rev	MPS081	CCTGAGCTTGTCTAACCTGTTGTCAGTCAGTGCCAAAACACAGGTTGAAGCAAAGGCTC
	SOX15 RNAi B fwd	MPS084	TGCTGAACGTAGTAGTCCAACAGGAGAAGTTGGCCACTGACTGACTCTCCTGGACTACAGTT
	SOX15 RNAi B rev	MPS085	CCTGAACGTAGTCTTGGCCAAACTCTCCTGGACTACAGTT
gRNA oligos	AAVS1 gRNA fwd	AS9	caccGGGGCCACTAGGGACAGGAT
	AAVS1 gRNA rev	AS10	aaacATCCTGTCCTAAGTGGCC
	NANOS1 gRNA fwd	MPS011	caccgCTGGAAAGAGCTGCGCTGA
	NANOS1 gRNA rev	MPS012	aaactCACGCCAGCTTCCAGC
	SOX15 gRNA fwd	MPS007	caccgCTGGCATGAGGTTAGAGGT
	SOX15 gRNA rev	MPS008	aaacACCTTAACCCATGGACGc
Primers for cloning and sequencing	AAVS1 fwd	MPS052	ACTAGGGACAGGATTGGTGAC
	AAVS1 rev	MPS051	GCCCCACTGTGGGTGGAG
	AID fwd	AS105	GGGGGAGCAGGAGGCCTCGA
	CIE Fwd (for InFusion)	MPS053	CCTCCACCCCACAGTGGGCAAAACCTTTATGAGGGACAGC
	CIE Rev (for InFusion)	MPS054	CACCAATCTGTCCCTAGTTTATAGATATCAGGGACAGC
	DD 3' InFusion	MPS035	CGAGCGCCGCTCAGTCatccagttctagaqactccac
	DD-EF1 backbone fwd	MPS040	acggggGAGTCAGGTGAAaccatc
	DD-EF1 backbone rev	MPS041	CATGGTGGCGGCTCGAATTCTGTAGGCAGCGGTC
	EF1 backbone fwd	MPS036	CTGAGCGCCGCTCGATAAGC
	EF1 backbone rev	MPS037	CATGGTGGCGGCTCGAGCCGTAG
	Hyg Rev (for InFusion)	MPS050	GTGGGGCGGGCGTCAactattcccttgcctcgac
	IRES Fwd (for InFusion)	MPS049	ATGATAATATGCCACAACCATG
	KLF4 genotyping fwd	MPS066	AAATGCGACCGAGCATTTTC
	KLF4 genotyping rev	MPS067	CCCCCTTGGCATTTGTAAAGTC
	KLF4 seq rev	MPS077	CCACTCACACAGCCCCATCTG
	KLF4-geno_F	AS231	gggcagtctatcgcttgc
	KLF4-geno_R	AS232	tgccttataaccgggtgt
	MC1-DTA fwd	TK267	CAGTGTGGTTTCAAGAGGAAGCA
	Myc Rev (for cloning)	MPS055	CAGATCCTCTCTGAGATGAG
	NANOS1 3' arm 3' end In-Fusion	MPS025	TTGAAAACCACACTGCTGCTTCCCTCC
	NANOS1 3' arm 5' end In-Fusion	MPS024	TATCTAGACCCAGCTAGCCGGCTCCGGCG
	NANOS1 3' DD InFusion	MPS034	ctgcactccCGCGTGCAGCTTCTGCCAGGC
	NANOS1 5' arm 3' end In-Fusion	MPS022	TCCTCCGTAGCGCAGCTTCTGCCAGGC
	NANOS1 5' arm 3' end In-Fusion long	MPS031	CTCCTGCTCTCCCGTAGCGCAGCTTCTGCCAGGC
	NANOS1 5' arm 5' end In-Fusion	MPS021	ccccctcgagggtcGAGGCTTCCCTGGCGCC
	NANOS1 5' myc and InFusion	MPS033	CGAGCCGCCACCATGAAACAAAACATCTCAGAAGAGGATCTGGAGGCTTCCCTGGCGC
	NANOS1 genomic forward	MPS013	GCCCATGGAGGCTTCC
	NANOS1 genomic reverse	MPS014	TTTGTGGACAGGCAACAGC
	NANOS1 genotyping forward	MPS028	ACGGCACACCCATCAACTA
	NANOS1 genotyping fwd short	MPS064	CCCGCCTGGCAAGAGCTG
	NANOS1 genotyping reverse	MPS029	GACGACGTCCCAGTCAG
	NANOS1 insert 5' end In-Fusion	MPS023	CTGGCTACGGAGGAGCAGGAGCGTC
	NANOS1 insert 5' end In-Fusion short	MPS032	ACGGAGGAGCAGGAGGCGTC
	NeoR IF	MPS062	GCCAACGCCACCATGATTGAAACAGATGGATTG
	NeoR IF	MPS063	GAGCTCTAGAGCTCAGAAGAACCTGTCAGAACAGC
	OstIR1 plasmid linearization fwd	MPS047	TGACGCCGCCACGAC
	OstIR1 seq fwd	FS69	CTGAGATCTCTGCGCTGAA
	OstIR1 seq rev	FS82	CAGAATTTCACGAAATTGGGAGC
	OstIR1/IRES plasmid linearization rev	MPS048	GGTTGTGGCCATATTATCATCG
	pBlueScript rev	AS177	TCGAacttcgaggggggggcccgat
	PuroTK seq fwd	NI291	CCGTCCTCATGCACGCTTATC
	Resistance marker switching	MPS060	CATGGTGGCGTGGCTGCAG
	Resistance marker switching	MPS061	TGAGCTCTAGAGCTCGCTGATCAG
	Rox reverse	MPS015	AGCTGGCTAGATATCGCGCG
	Rox seq fwd	NI170	ATTCACTGACGTTAACGATCG
	SOX15 3' arm 3' end In-Fusion	MPS020	TTGAAAACCACACTGATGGCTGGGACCTTTGT
	SOX15 3' arm 5' end In-Fusion	MPS019	TATCTAGACCCAGCTCCCTCATGGACGCACACCTC
	SOX15 3' DD InFusion	MPS039	ctgcactccCGCGTGAAGGGTTAGGGGATCG
	SOX15 5' arm 3' end In-Fusion	MPS017	CCCCGTAGAGGTTGTTAGGGGCATG
	SOX15 5' arm 3' end In-Fusion long	MPS030	GCCTCCTGCTCCCCGTAGAGGTGtGTTAGGGGCATG
	SOX15 5' arm 5' end In-Fusion	MPS016	ccccctcgagggtcGACTACAACTGACCGCCTC
	SOX15 5' myc and InFusion	MPS038	CGAGCCGCCACCATGAAACAAAACATCTCAGAAGAGGATCTGGCGTACCAAGGCTCCAC
	SOX15 genomic forward	MPS009	CGACTACCCGACTACAAAGT
	SOX15 genomic reverse	MPS010	GCTCTGTCTCTGCAACAG
	SOX15 genotyping forward	MPS026	TGGCTCTCCACTCCATACA
	SOX15 genotyping reverse	MPS027	GGCTATCATGGGAGGACTGC
	SOX15 insert 5' end In-Fusion	MPS018	ACACCTCTACGGGGAGCAGGAGGCG
	Venus seq rev	NI020	CTCCTGAAGTCGATGCCCTT

### **PGCLC induction:**

For PGCLC induction, hESCs cultured in 4i medium were dissociated with 0.25% trypsin/EDTA. The cells were suspended in MEF medium to quench the trypsin, and the suspension was filtered through a 50 µm strainer. The cells were pelleted by centrifugation (300 g, 4 minutes) and re-suspended in PGCLC base medium<sup>11</sup> (aRB27 with 10 µM Y-27632, 0.25% w/v poly(vinyl alcohol), and 10 ng/mL hLIF). Cells were counted (Invitrogen Countess) and the suspension was diluted to 40000 live cells/mL in PGCLC medium (PGCLC base plus 500 ng/mL BMP2, 100 ng/mL SCF, and 50 ng/mL EGF). 100 µL of suspension, containing 4000 cells, was added to each well of a 96-well ultra-low-attachment plate (Corning CoStar). Cells were pelleted (300 g, 2 minutes) and the plate was incubated (37°C, 5% CO<sub>2</sub>). For experiments beyond day 6 of culture, a 50% medium change was performed on day 6.

### **AID experiments:**

For AID, indole-3-acetic acid (IAA) was added to the cell culture medium at a final concentration of 100 µM. In PGCLC experiments where the IAA was added after induction, 10 µL of 1.1 mM IAA in PGCLC base medium were added. In these experiments, 10 µL of PGCLC base medium containing no IAA were also added to control wells. For ProteoTuner experiments, Shield1 was used at a concentration of 0.5 µM.

### **Flow cytometry:**

Embryoid bodies (EBs) were collected, washed with PBS, and dissociated by digesting with 0.25% trypsin/EDTA (5 µL per EB) for 10 minutes at 37°C with gentle shaking (600 rpm). For day 6 and older EBs, dissociation was completed by passing the suspension multiple times through a 27-gauge needle. Trypsin was quenched with two volumes of ice-cold sorting medium (3% FBS in PBS) and the cells were pelleted (300 g, 2 minutes). Next, the cells were re-suspended in sorting medium (5 µL per EB) containing AF647 conjugated mouse anti-human CD38 IgG (BD Biosciences 561500) (1 µL per 12 EBs) and incubated at 4°C for 30 minutes. The antibody solution was diluted with two volumes of sorting medium, and the cells were pelleted (300 g, 2 minutes) and resuspended in 500 µL sorting medium plus DAPI (0.1 µg/mL). The suspension was filtered with a 50 µm strainer and analyzed on a flow cytometer (BD LSRFortessa or Sony 800Z). Cells in FACS experiments were sorted directly into 50 µL RNA extraction buffer (Arcturus PicoPure) which was frozen at –80°C for subsequent use. RNA was extracted following the manufacturer's instructions.

### **Statistical analysis of flow cytometry data:**

For flow cytometry data, the hPGCLC fraction was calculated using FlowJo software. For each induced clonal line, the fold change was calculated as the ratio of hPGCLC fraction in treated (with IAA or Shield1, depending on experiment) and untreated samples. This step was performed in order to control for the batch-to-batch variability between different hPGCLC inductions. The fold change values were then compared for experimental cell lines (overexpression or depletion) and control cell lines (either SOX15-AID-Venus with no TIR1, or parental N3tdT). This was done in order to control for any possible nonspecific effects of Shield1 or IAA (a known aryl hydrocarbon receptor agonist<sup>29</sup>). The Wilcoxon rank-sum test was used for comparisons, since by the Kolmogorov-Smirnov test the data were not normally distributed.

### **RNAi in TCam-2 cells:**

For TCam-2 RNAi experiments, cells in a 6-cm dish were treated with doxycycline at 1 µg/mL for 48 hours. Then, cells were washed with PBS and dissociated with 0.25% trypsin/EDTA. Trypsin was quenched with sorting medium, and cells were pelleted and resuspended in 1 mL ice-cold sorting medium. GFP-positive and GFP-negative cells were separated by flow cytometry (Sony 800Z). Cells were pelleted and RNA was isolated with the Qiagen RNeasy Mini kit following the manufacturer's instructions.

### **Quantitative reverse-transcription PCR (qPCR):**

cDNA synthesis was performed using the Quantitect Reverse Transcription kit (Qiagen) following the manufacturer's instructions. qPCR reactions were performed at 10 µL scale in 384-well plate format using the SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich) on a QuantStudio instrument (Applied Biosystems). Primers are listed in Table 2. Two technical replicates were performed for each biological replicate. The  $\Delta\Delta Ct$  method was used for quantification, with *GAPDH* as a reference transcript.

### **Immunofluorescence:**

For immunofluorescence in ESCs, cells were grown on an Ibidi 8-well plate. Cells were fixed with 4% paraformaldehyde at room temperature for 10 minutes. For immunofluorescence in EBs, the EBs were fixed with 4% paraformaldehyde at 4 °C for 1 hour, then washed with PBS and transferred to 10% sucrose in PBS. When EBs had sunk, the process was repeated with 20% sucrose. EBs were then embedded in OCT compound and cryosectioned to 8 µm thickness on SuperFrost Plus slides (VWR). Fixed samples for

immunofluorescence were stained as previously described.<sup>10</sup> Primary antibodies used are listed in Table 3. Imaging was performed with an SP5 confocal laser scanning microscope (Leica) and images were analyzed using Fiji software.<sup>30</sup>

---

**Table 2:** primers used for qPCR

qPCR Primer	Number	Sequence (5' to 3')
AKAP1 fwd	AS315	CCGTTCAAGGGACAGAACGT
AKAP1 rev	AS316	TTGTCACCTGAGCAAGCAGT
AKAP12 fwd	AS325	AACGGTCAAGGAGCCCTAAA
AKAP12 rev	AS326	CATCTTCAGAGTCTCTCTGTC
BEND4 fwd	AS331	TTAGAGAGCATCCCAGTGCCC
BEND4 rev	AS332	TTCAACAGCTGCTTGCTTCC
CBFA2T2 fwd	MPS090	AGTCCAGAGGGAGAGAAAGAGA
CBFA2T2 rev	MPS091	TTCATGAGGGGCACACTGAG
CDX1 fwd	AS335	GACGCCCTACGAGTGGATG
CDX1 rev	AS336	TGTAGACCACGCCGCGTACTTG
CXXC4 fwd	MPS130	CCTGCAAGGGCTCATCAA
CXXC4 rev	MPS131	CCATCGGAATGCTCAGCCG
ELF3 fwd	MPS110	TCAACGAGGGCCTCATGAA
ELF3 rev	MPS111	TCGGAGCGCAGGAACCTG
EMB fwd	AS359	ACCCAATAACAGGTTACCATCA
EMB rev	AS360	CCCTACGTAAAGAGATCAATGGCT
ETV5 fwd	MPS094	GTGTTGTGCTGAGAGACTGGA
ETV5 rev	MPS095	CGACCTGTCAGGCAATGAAGT
FABP7 fwd	MPS120	TGTGCTACCTGGAAGCTGAC
FABP7 rev	MPS121	CTTGAATGCTGAGAGTC
FOXK1 fwd	MPS132	CAGTTACCGCTTGTCAG
FOXK1 rev	MPS133	GAATTCTGCCAGCCTTGTC
GAPDH fwd	NI69	CGCTTCGCCCTCTGCTCCCTG
GAPDH rev	NI70	GGTGACCAAGGCCAAATACGA
GATA6 fwd	Q61	CCCACACACAACCTACAGC
GATA6 rev	Q62	GCGAGACTGACGCTATGTA
HNFB1 fwd	Q89	CAATCCACTCTCAGGAGCG
HNFB1 rev	Q90	ATCGTGGAGAGGCATTGTG
ID2 fwd	MPS124	GCAGCACGTATCGACTACATC
ID2 rev	MPS125	CCACACAGTCCTTGCTGTC
IGFBP5 fwd	MPS118	GCCCCATTGTGACCGCAAAG
IGFBP5 rev	MPS119	GTCAACGACTCCATGCC
IL6R fwd	MPS114	TTGTTTGAGTGGGTCT

qPCR Primer	Number	Sequence (5' to 3')
IL6R rev	MPS115	TGGGACTCCCTGGAAATACTG
KDM3B fwd	MPS096	GCTCGTAATGCTGAGAAGGAGG
KDM3B rev	MPS097	CACATTGCGACAAACCCAGTGG
KIT fwd	NI215	TGCATTCAAGCACAAATGGCACGG
KIT rev	NI216	GTGTGGGGATGGATTGCTCTTGT
LEFTY2 fwd	AS73	CACCTGGACCTCAGGGACTAT
LEFTY2 rev	AS74	CCCAGTCTTGGCCCACTTCAT
MAST2 fwd	MPS100	ATCCGCCAGTTCTCTCTCTGCT
MAST2 rev	MPS101	CCATCTGAATGGTCTCTCTCTC
MST1R fwd	MPS104	TATCCTGCAGGTGGAGCTG
MST1R rev	MPS105	ATGAAATGCCATGCCCTTAG
NANOG fwd	NI99	TGCTGAGATGCCTCACACGGA
NANOG rev	NI100	TGACCGGACCTTGTCTTCTT
OTX2 fwd	MPS073	GACGACGTTACTCGGG
OTX2 rev	MPS074	TCTTAAACCATACCTGCACC
PDCD4 fwd	MPS128	GTGGGACAGTAATGAGCAC
PDCD4 rev	MPS129	CATCTCCAAAGCTCTAGC
POU5F1 fwd	NI19	GCTGGAGCAAAACCCGGAGG
POU5F1 rev	NI20	TGGCCTGTGATATCCCAGGTG
PRDM14 fwd	AS65	CTCGTTCCAGTCACTGGAG
PRDM14 rev	AS66	AGGAAGAGAAATCAGATCCAGAGC
S100A4 fwd	MPS122	GGAGAAAGGCCCTGGATGTGA
S100A4 rev	MPS123	CTCGTTGTCCTGTTGCTG
SOX15 fwd	MPS005	CCAACCGAGCACAGGCTTT
SOX15 rev	MPS006	GTTTCCAGGGAAAGGCCATAG
SOX17 fwd	NI166	GAGCCAAGGGCAGTCCCCTGA
SOX17 rev	NI167	CCTTCCACGACTGCCAGCAT
SOX2 fwd	NI146	TTGCGTGAGTGTGGATGGGATGGTG
SOX2 rev	NI147	GGGAAATGGGAGGGGTGCAAAGAGG
TFAP2C fwd	NI55	CGCTCATGTGACTCTCTGACATCC
TFAP2C rev	NI56	TGGGCCGCCAATAGCATCTTCT
VENTX fwd	AS257	CCATGGCCGGTTGAGTAAG
VENTX rev	AS258	CTCAGGTACTGGTGGTGC

**Table 3:** primary antibodies for immunofluorescence

Target protein	Antibody type	Dilution	Manufacturer
myc tag	Mouse monoclonal	1:4000	Abcam ab32
GFP	Chicken polyclonal	1:1000	Abcam ab13970
SOX17	Goat polyclonal	1:500	R&D AF1924
KLF4	Rabbit polyclonal	1:200	Santa Cruz sc20691
OCT4	Mouse monoclonal	1:200	BD 611203
BLIMP1	Rabbit polyclonal	1:100	CST 9115

## Results

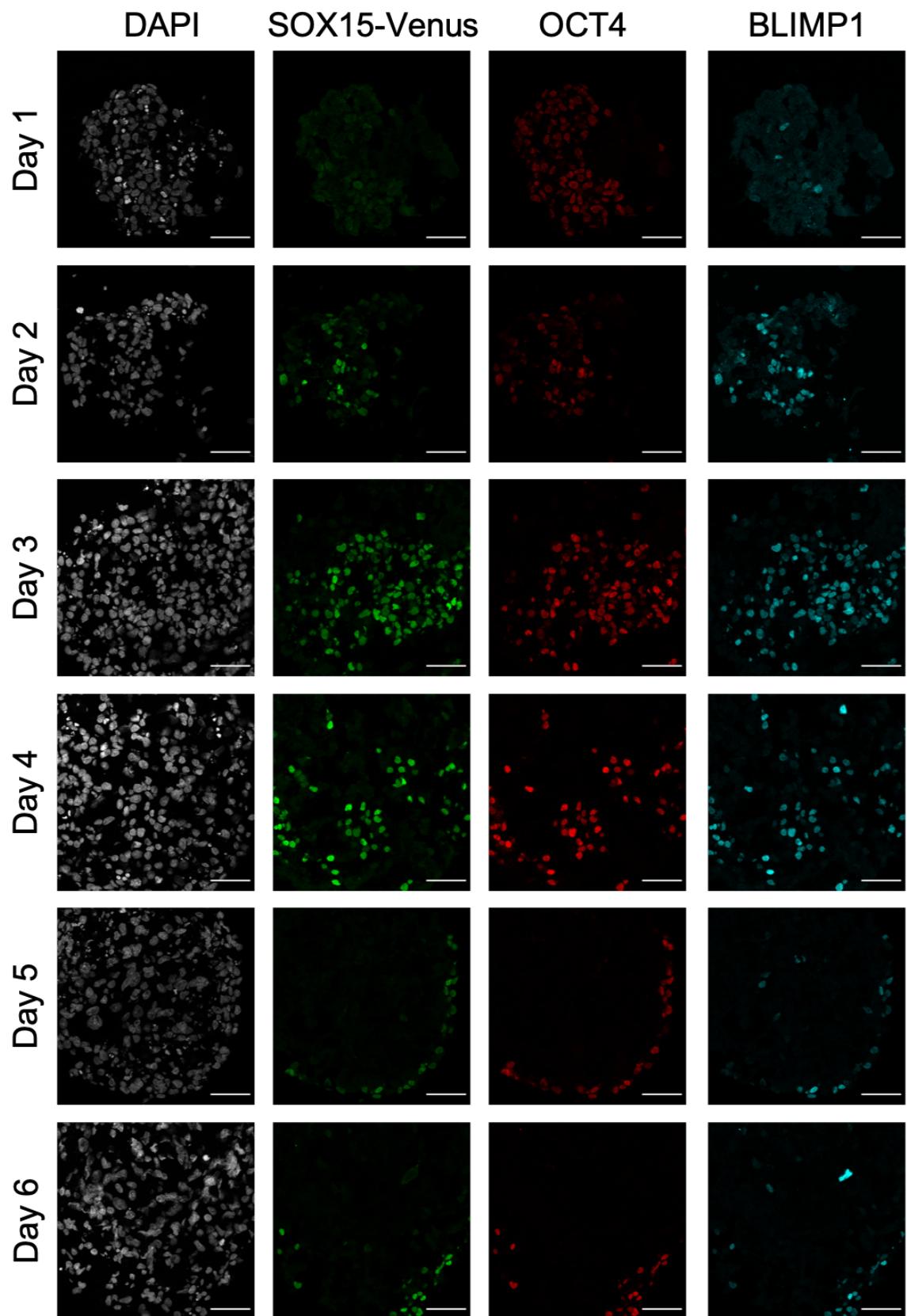
### **SOX15 depletion experiments using AID:**

In order to determine the effects of SOX15 depletion on PGCLC specification, a homozygous knock-in cell line with a C-terminal AID-Venus tag on SOX15 was generated using CRISPR/Cas9 and homologous recombination with a donor plasmid. The parental line had a NANOS3-T2A-tdTomato (N3tdT) reporter, which is expressed specifically in PGCLCs. The SOX15-AID line was subsequently transfected with OsTIR1 using the PiggyBac system, and the selectable marker was excised using transient expression of Dre recombinase.

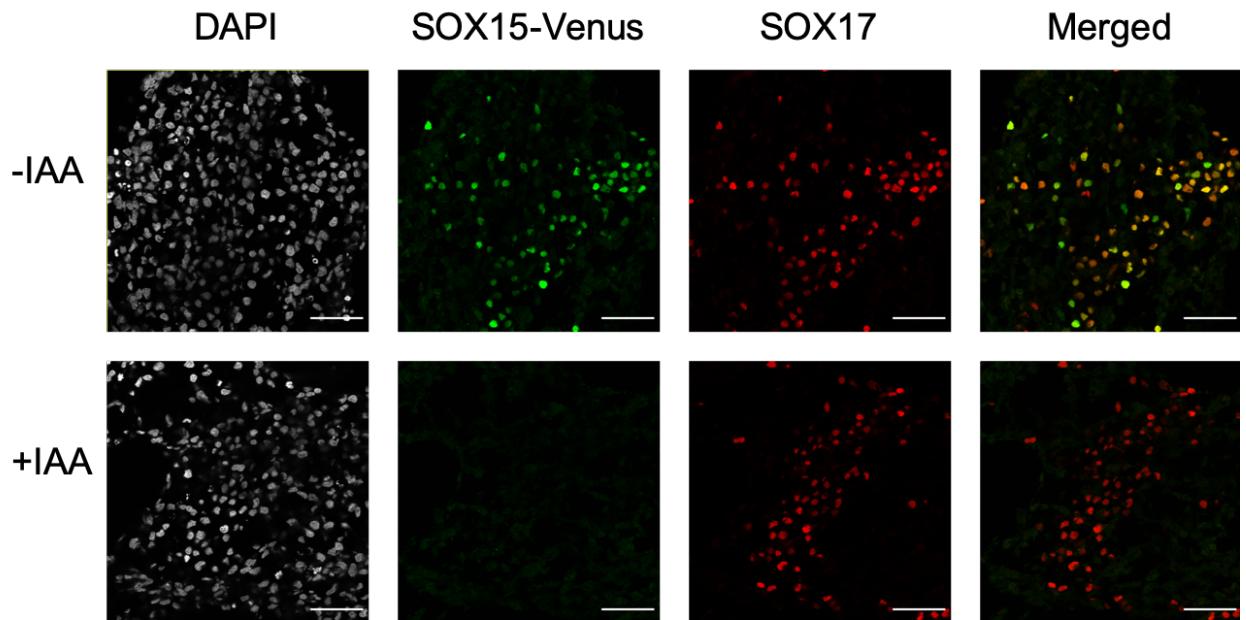
Immunofluorescence experiments confirmed that SOX15 is expressed at the protein level in PGCLCs but not somatic cell lineages (soma) within the EBs. This result is in agreement with previous RNA-seq data.<sup>3,10</sup> Since no suitable antibody for SOX15 was commercially available, an antibody against the Venus tag was used instead. The expression was first observed at a faint level on day 1 after induction, and more strongly on day 2. Expression continued in OCT4/BLIMP1 positive PGCLCs until the end of the time-course experiment (day 6) (Figure 2). Interestingly, beginning on day 5, the PGCLCs were present mainly along the edges of the EBs.

To deplete SOX15, SOX15-AID/OsTIR1 cells were treated with IAA beginning at the same time as PGCLC induction. Immunofluorescence performed on day 4 after induction showed depletion of SOX15-AID-Venus to background levels (Figure 3). SOX15-AID-Venus expression in untreated cells was similar to that observed in the previous experiment. SOX17 was used as a marker for PGCLCs, and this staining showed that PGCLCs were present in both samples. Therefore, SOX15 was evidently not essential for PGCLC identity.

In order to quantify any effects of SOX15 depletion on PGCLC specification efficiency, flow cytometry was performed on cells from dissociated EBs either treated or untreated with IAA. Identification of PGCLCs was performed using a combination of the NANOS3-T2A-tdTomato reporter and antibody staining against the alkaline phosphatase (AP) surface marker. PGCLCs were counted as cells positive for both markers. The results indicated that on day 4 after induction, there was no significant effect of SOX15 depletion on induction efficiency (Figure 4). However, at later time points (days 6 and 8), SOX15 depletion



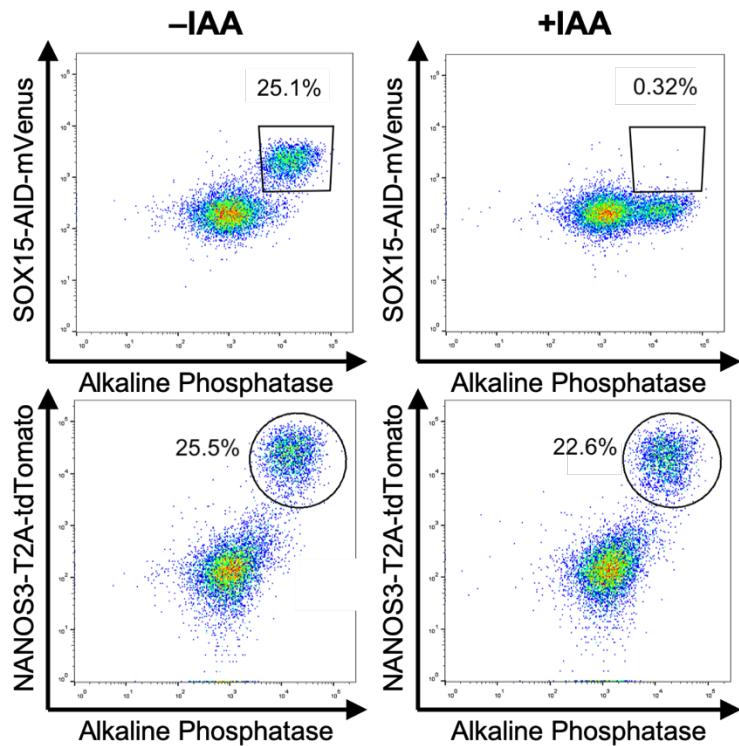
**Figure 2.** Timecourse immunofluorescence of SOX15-AID-Venus EBs stained with DAPI (grey), anti-Venus (green), anti-OCT4 (red), and anti-BLIMP1 (cyan). Scale bar is 50 µm. SOX15 expression is observed faintly on day 1 and robustly on day 2, persisting in OCT4/BLIMP1 positive cells for the remainder of the experiment.



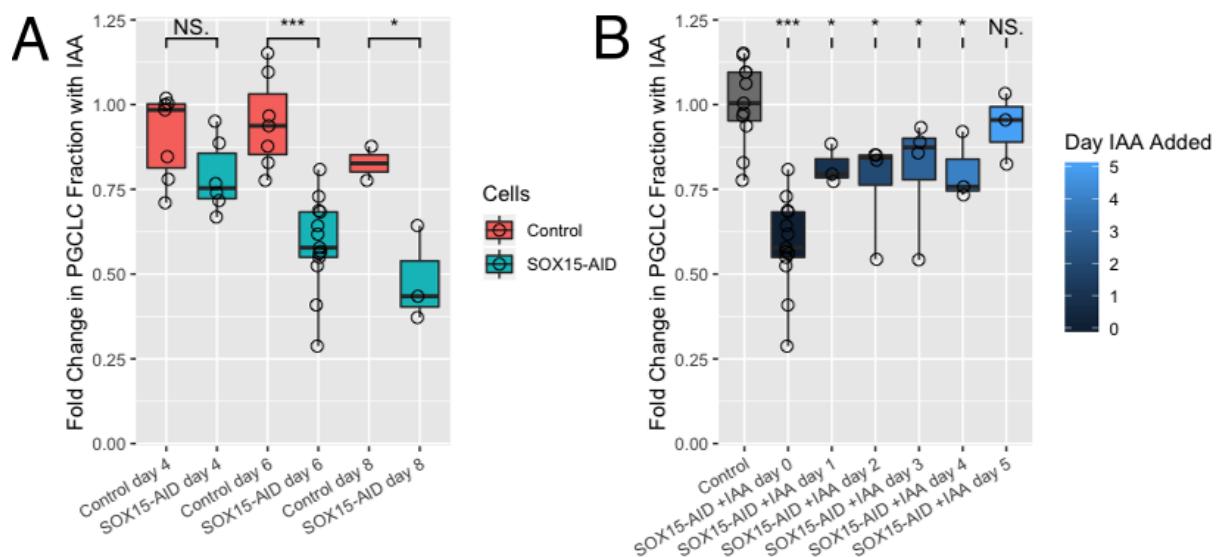
**Figure 3.** SOX15-AID-Venus expression is observed in SOX17-positive PGCLCs in day 4 EBs. A few SOX17-positive Venus-negative definitive endoderm somatic cells are also present. SOX15-AID-Venus, but not SOX17, is depleted to background levels with IAA treatment. Staining for TIR1-myc (not shown) indicates ubiquitous expression. Scale bar is 50  $\mu$ m.

significantly reduced the fraction of PGCLCs, but this still did not completely eliminate all of them (Figure 5A). The progressive decrease in PGCLC fraction upon prolonged SOX15 depletion suggests that SOX15 may have a role in PGCLC maintenance.

The rapid kinetics of SOX15 depletion by the AID system enabled an investigation of the effects of SOX15 depletion starting at various time points after induction. An endpoint at day 6 was chosen because EBs from later time points were not easily disaggregated, leading to low yield of cells. IAA was added on days 0 through 5 and the PGCLC fraction present on day 6 was measured by flow cytometry. As expected, the effect diminished when IAA was added at later time points (Figure 5B). This implies that prolonged depletion of SOX15 is necessary for there to be an effect on PGCLC fraction. It is also notable that there was a significant decrease in PGCLC fraction on day 6 when SOX15 was depleted starting on day 4 or earlier. Interestingly, depletion from day 0 did not produce a significant effect on PGCLC fraction measured on day 4, but depletion from day 4 significantly reduced PGCLC fraction measured on day 6 (Wilcoxon test,  $p = 0.02$ ). Since by day 2 the PGCLC transcriptional network is already largely established,<sup>10</sup> this suggests that SOX15 depletion can interfere with maintenance of PGCLC identity even when specification proceeds normally.



**Figure 4.** Representative flow cytometry analysis of SOX15-AID-Venus protein expression and PGCLC markers (NANOS3 and alkaline phosphatase). IAA treatment results in a near-total reduction of Venus-positive cells on day 4. However, this only causes a slight decrease in AP+/NANOS3+ PGCLCs.

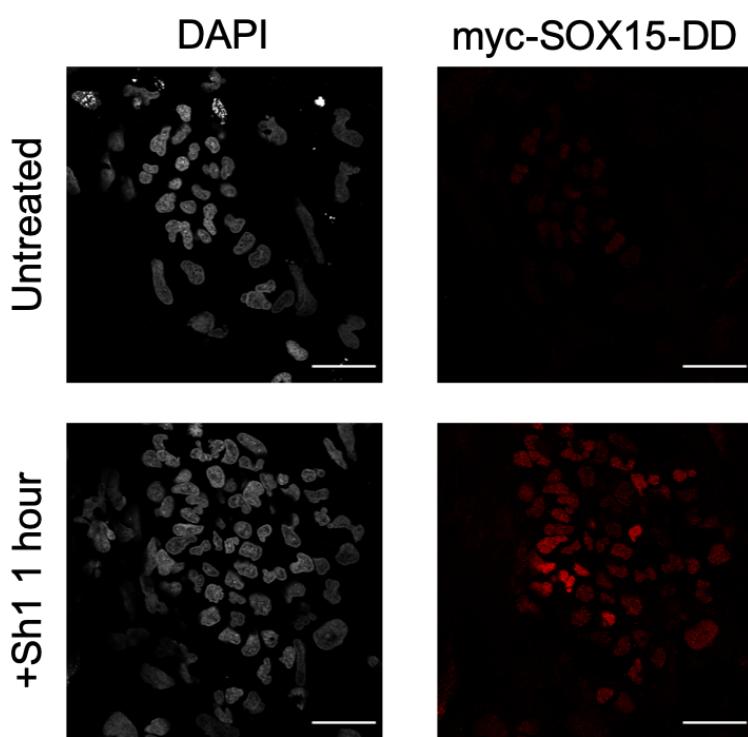


**Figure 5.** SOX15 depletion with IAA treatment decreases NANOS3+/AP+ PGCLC fraction by flow cytometry. Fractions were normalized with respect to untreated samples of the same clones. Statistical comparisons were performed between SOX15-AID/TIR1 clones and control clones without TIR1, which did not deplete SOX15. **A:** IAA treatment from the start of induction, with PGCLC fraction measured on day 4, 6, or 8. **B:** IAA treatment beginning on the day indicated, with PGCLC fraction measured on day 6. Significance values are by Wilcoxon test (\* p < 0.05, \*\* p < 0.01, \*\*\*p < 0.001)

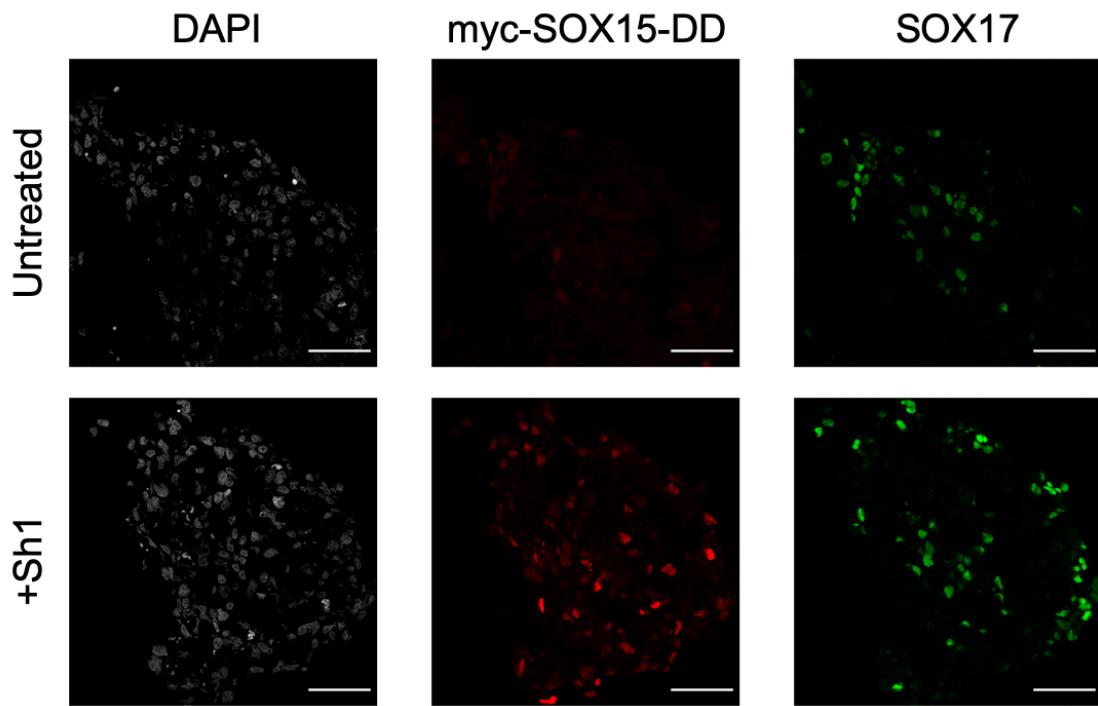
### **SOX15 overexpression experiments using ProteoTuner:**

To further elucidate the functional role of SOX15 in PGCLC specification, overexpression was performed using the ProteoTuner system. This system consists of a destabilizing domain (DD) fused to the protein target, which normally results in protein degradation. Upon addition of a stabilizing ligand (Shield1), protein levels quickly increase. The ProteoTuner system has rapid kinetics<sup>26</sup> similar to those of the AID system.

An expression cassette composed of myc-SOX15-DD-IRES-HygroR under the control of the constitutively active EF1 $\alpha$  promoter was transfected into N3tdT hESCs. PiggyBac transposase was used for delivery. After selection with hygromycin, colonies were picked and clones were tested for Shield-dependent expression by immunofluorescence after 1 hour of treatment (Figure 6). Two suitable clones were identified which homogenously expressed myc-SOX15-DD protein only in the presence of Shield1. Subsequently, these cells were induced to form PGCLCs with Shield1 present or absent during induction. By immunofluorescence on day 4 post-induction, Shield1 treatment resulted in myc-SOX15-DD expression in both PGCLCs and soma (Figure 7).

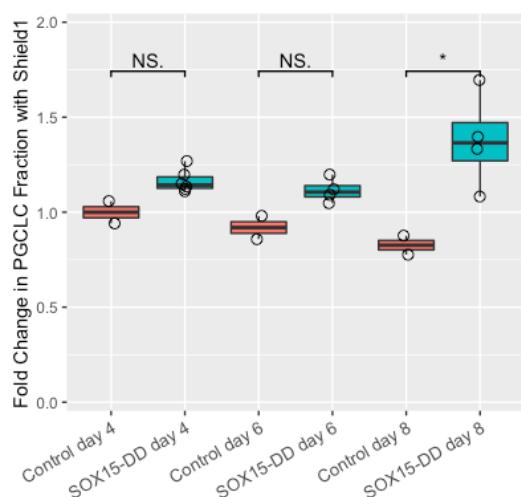


**Figure 6.** hESCs (colonies on a layer of MEFs) express myc-SOX15-DD after one hour of treatment with Shield1. myc-SOX15-DD successfully localizes to the nucleus. Scale bar is 50  $\mu$ m.



**Figure 7.** In a Shield1-dependent manner, EBs overexpress myc-SOX15-DD in both SOX17-positive PGCLCs as well as somatic lineages. Scale bar is 50  $\mu$ m. Samples are from day 4 EBs.

To quantify the effect of SOX15 overexpression on PGCLC induction efficiency, the EBs were dissociated and analyzed by flow cytometry using the same method as in the SOX15 depletion experiments. There was a greater fraction of PGCLCs in EBs overexpressing SOX15 (Figure 8). While this difference was not statistically significant on days 4 and 6, it became significant on day 8. In all conditions, the total number of PGCLCs decreased between days 6 and 8, but this decrease was not as pronounced in the EBs overexpressing SOX15. Taken together with the delayed effects of SOX15 depletion observed in the AID experiments, this further supports a role for SOX15 in PGCLC maintenance.



**Figure 8.** SOX15 overexpression with Shield1 treatment at the beginning of induction increases PGCLC fraction by flow cytometry. NANOS3+/AP+ PGCLC fraction was measured on day 4, 6, or 8, and normalized with respect to untreated samples of the same clones. Statistical comparisons were performed between wild-type controls and clones overexpressing SOX15.

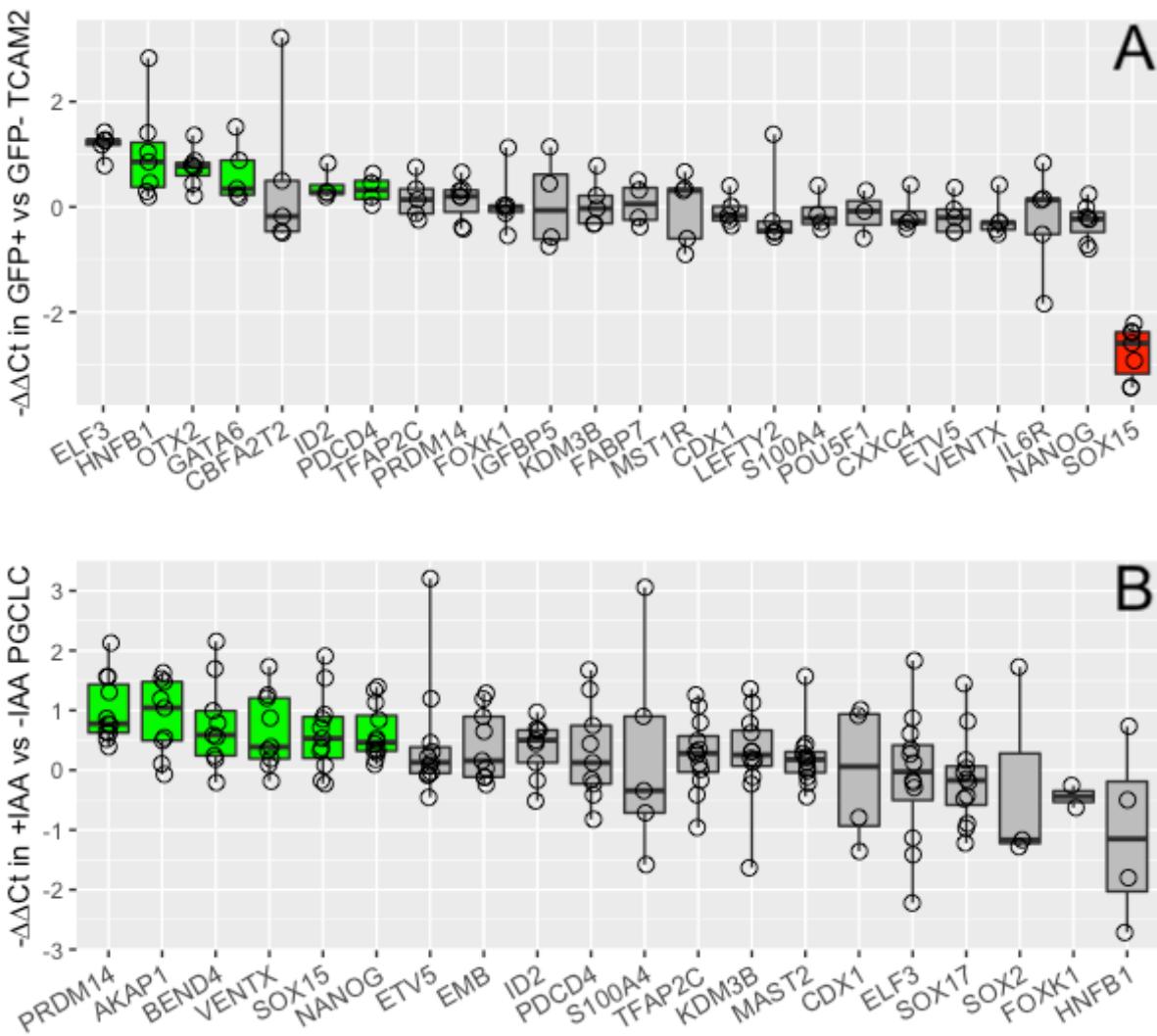
### **Transcriptional effects of SOX15 perturbation:**

In order to minimize the need for time-intensive PGCLC inductions, preliminary experiments were performed using RNAi in TCam-2 cells. This cell line derived from human germ cell seminoma retains germ cell characteristics,<sup>31</sup> including expression of SOX15. The cells were transfected with a vector encoding doxycycline-inducible miRNA and EGFP fluorescent reporter, with an additional neomycin resistance marker. The vector was integrated by PiggyBac transposase. Two separate artificial miRNA constructs with perfect base-pairing to different regions of the SOX15 mRNA 3'-UTR were used, as well as a scrambled control miRNA. After transfection, cells were selected using G418 to remove wild-type cells. Upon treatment with doxycycline, EGFP expression was heterogenous. Therefore, EGFP positive cells were isolated by flow cytometry and further cultured. Although this increased the proportion of EGFP positive cells upon subsequent doxycycline treatment, expression was still heterogenous, suggesting that epigenetic silencing was taking place. Ultimately, experiments were performed with cDNA from EGFP positive cells isolated by flow cytometry.

qPCR experiments on these cell populations showed that one miRNA was highly efficient (79 – 91%) in knocking down SOX15 expression, whereas the other one only gave a moderate decrease in expression (35 – 60%). Scrambled miRNA had no detectable effect. To test transcriptional effects, I assembled a set of candidate genes including both known regulators of germline identity, and previously reported SOX15 targets in other cell types, including human embryonal carcinoma cells,<sup>32</sup> muscle satellite cells,<sup>33</sup> esophageal<sup>34</sup> and pancreatic<sup>35</sup> adenocarcinomas, and mouse ESCs.<sup>17</sup> I designed qPCR primers and validated them by standard curves on cDNA as well as negative control reactions on genomic DNA.

Using these primers, I tested the effect of SOX15 depletion on gene expression levels. With the TCam-2 RNAi system, a subset of the candidate genes (*ELF3*, *HNFB1*, *OTX2*, *GATA6*, *ID2*, and *PDCD4*) had significant increases in expression in cells depleting SOX15 (Figure 9A) (Z test with Holm-Bonferroni correction,  $p < .05$ ). Besides *SOX15* itself, there were no significantly downregulated genes among those tested.

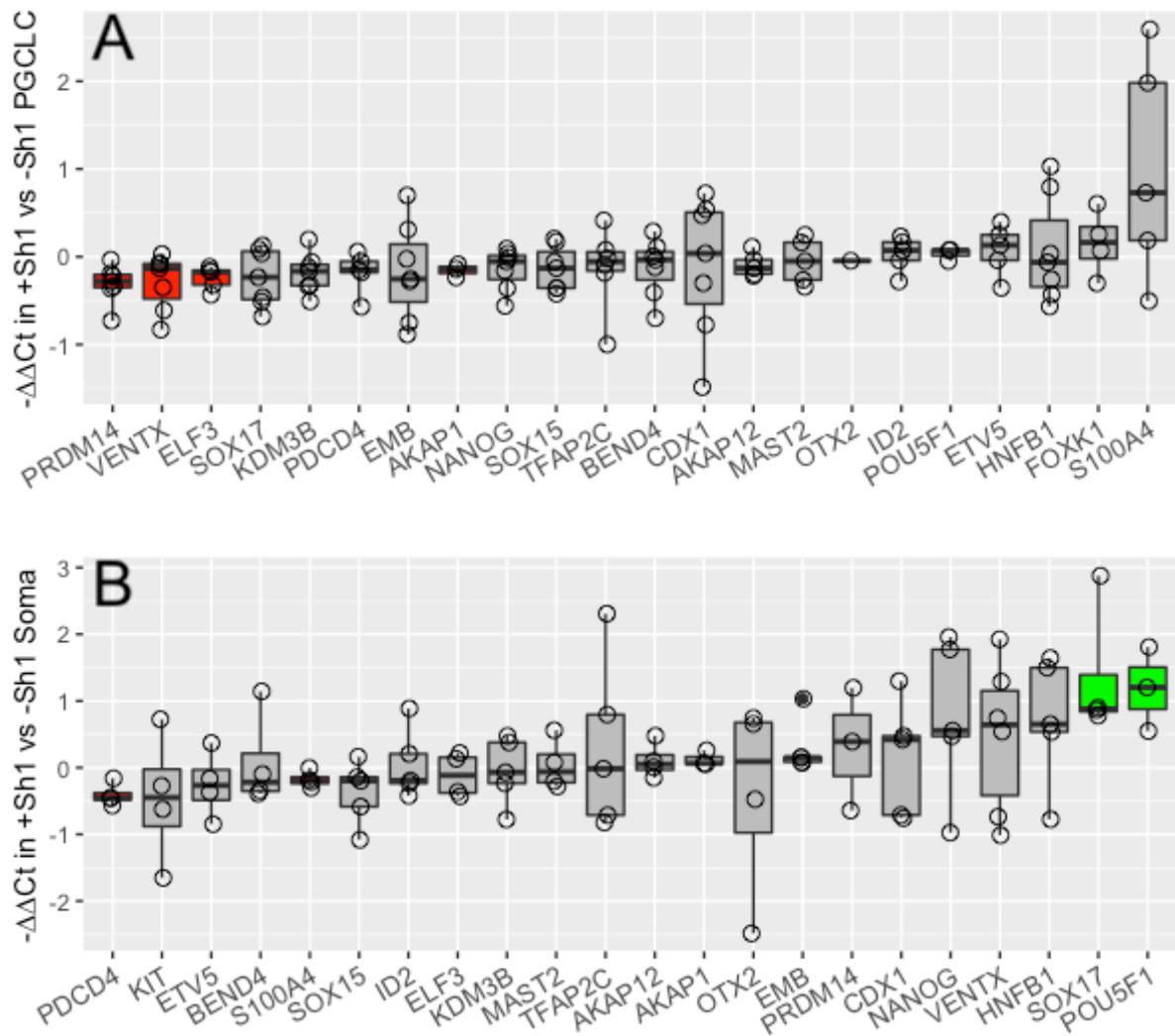
I next investigated transcriptional changes in PGCLCs when SOX15 protein was either depleted by AID or overexpressed by ProteoTuner. I used the day 6 timepoint for experiments, since at day 4 the phenotypic effect was weak and at day 8 the PGCLC yield



**Figure 9.** Transcriptional effects of SOX15 depletion by RNAi in TCam-2 cells (A) or by AID in PGCLCs (B). Each point represents a biological replicate. Green bars represent significantly upregulated genes, and red bars significantly downregulated ones.

was low. By qPCR, depletion of SOX15 from PGCLCs by AID caused significant upregulation of *PRDM14*, *AKAP1*, *BEND4*, *VENTX*, *SOX15*, and *NANOG* (Figure 9B) (Z test with Holm-Bonferroni correction,  $p < .05$ ). Among these, *PRDM14*, *VENTX*, *SOX15*, and *NANOG* are known to be associated with germ cell identity, which suggests a compensatory effect. Notably, the upregulation of *SOX15* implies negative feedback. *AKAP1*, *VENTX*, and *BEND4* are known targets of *PRDM14* in PGCLCs,<sup>36</sup> so their upregulation may be indirect. Surprisingly, the upregulated genes in SOX15-AID PGCLCs were different from those in the TCam-2 RNAi experiment.

In contrast, the SOX15-DD overexpression PGCLCs showed transcriptional changes that were generally the opposite of the SOX15-AID PGCLCs (Figure 10). *PRDM14*, *AKAP1*,



**Figure 10.** Transcriptional effects of SOX15 overexpression by ProteoTuner in EBs: **(A)** PGCLCs or **(B)** soma. Each point represents a biological replicate. Green bars represent significantly upregulated genes, and red bars significantly downregulated ones.

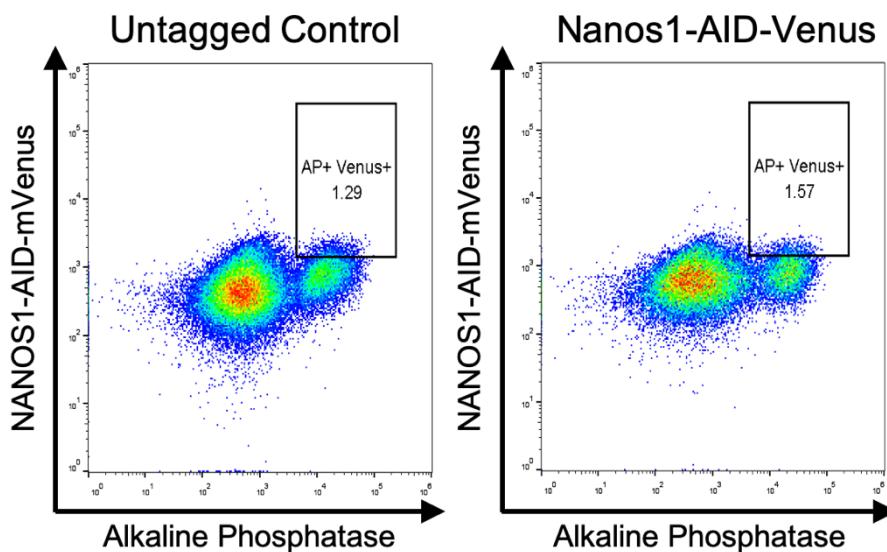
*VENTX*, and *ELF3* were significantly downregulated (Z test with Holm-Bonferroni correction,  $p < .05$ ). The first three of these genes were upregulated in SOX15-AID, and *ELF3* was upregulated in TCam-2 RNAi. Endogenous SOX15 expression was also downregulated on average, although this effect was not statistically significant ( $p = 0.08$ ) due to high variability.

SOX15-DD overexpression also had distinct effects in the somatic cells of the EBs. *SOX17* and *POU5F1* (encoding OCT4) were significantly upregulated, showing an approximately 2-fold increase. These genes are highly expressed in PGCLCs and are crucial for establishing their identity. Although upregulation of these genes was not observed in PGCLCs overexpressing SOX15, this may have been because they were already being transcribed at their maximal rate.

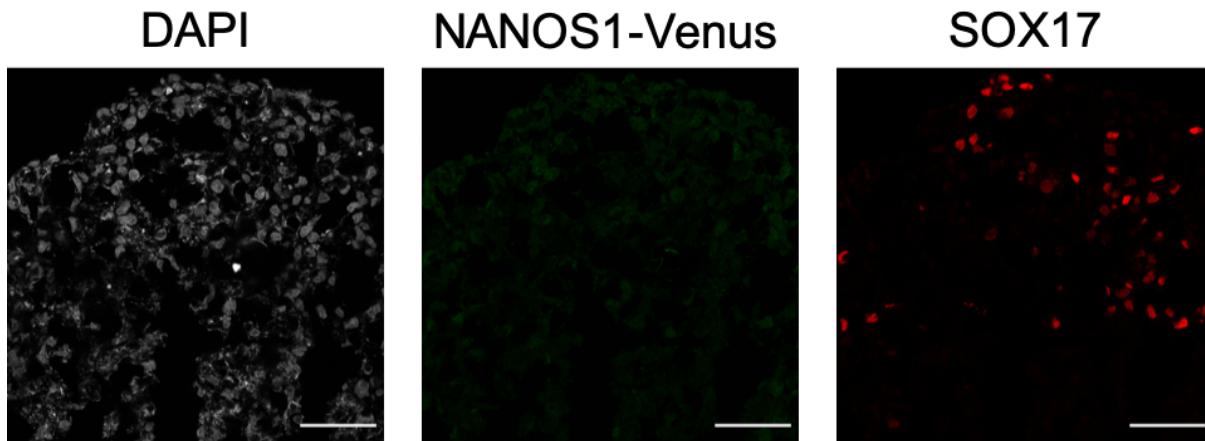
### NANOS1-AID:

Besides transcription factors, RNA-binding proteins also play important roles in germline specification. The NANOS family of proteins, which regulate translation of RNA in germ cells, is highly conserved throughout many different organisms.<sup>22</sup> In humans, there are three NANOS genes, with *NANOS1* and *NANOS3* expression detected by RNA-seq in PGCLCs. *NANOS3* is specifically expressed only in PGCs and PGCLCs. *NANOS1* mRNA is expressed at a higher level than *NANOS3* in PGCLCs, but is also present at moderate levels in other cell types.<sup>10</sup>

To try to determine the functional role of NANOS1 in PGCLCs, I generated a NANOS1-AID-Venus cell line, again from the N3tdT parental line. To my surprise, NANOS1-AID-Venus expression was not present, even in NANOS3-positive PGCLCs. This was confirmed by both flow cytometry (Figure 11) and immunofluorescence (Figure 12). Furthermore, attempted overexpression of myc-NANOS1-DD using the ProteoTuner system was not successful in any of the clones tested by immunofluorescence in ESCs. *NANOS1* mRNA, unlike *NANOS3* mRNA, is very GC-rich with a high degree of secondary structure. It seems that although the mRNA is present in PGCLCs, it is not translated into protein. An alternative explanation is that the AID-Venus tag destabilizes NANOS1 protein, but this tag is known to be well tolerated in the absence of TIR1,<sup>25</sup> so this explanation is unlikely. In any case, I decided to not pursue NANOS1-AID further since there were more promising areas of research.



**Figure 11.** Flow cytometry analysis of NANOS1-AID-Venus and control PGCLCs. No significant Venus signal is observed in either of the samples.

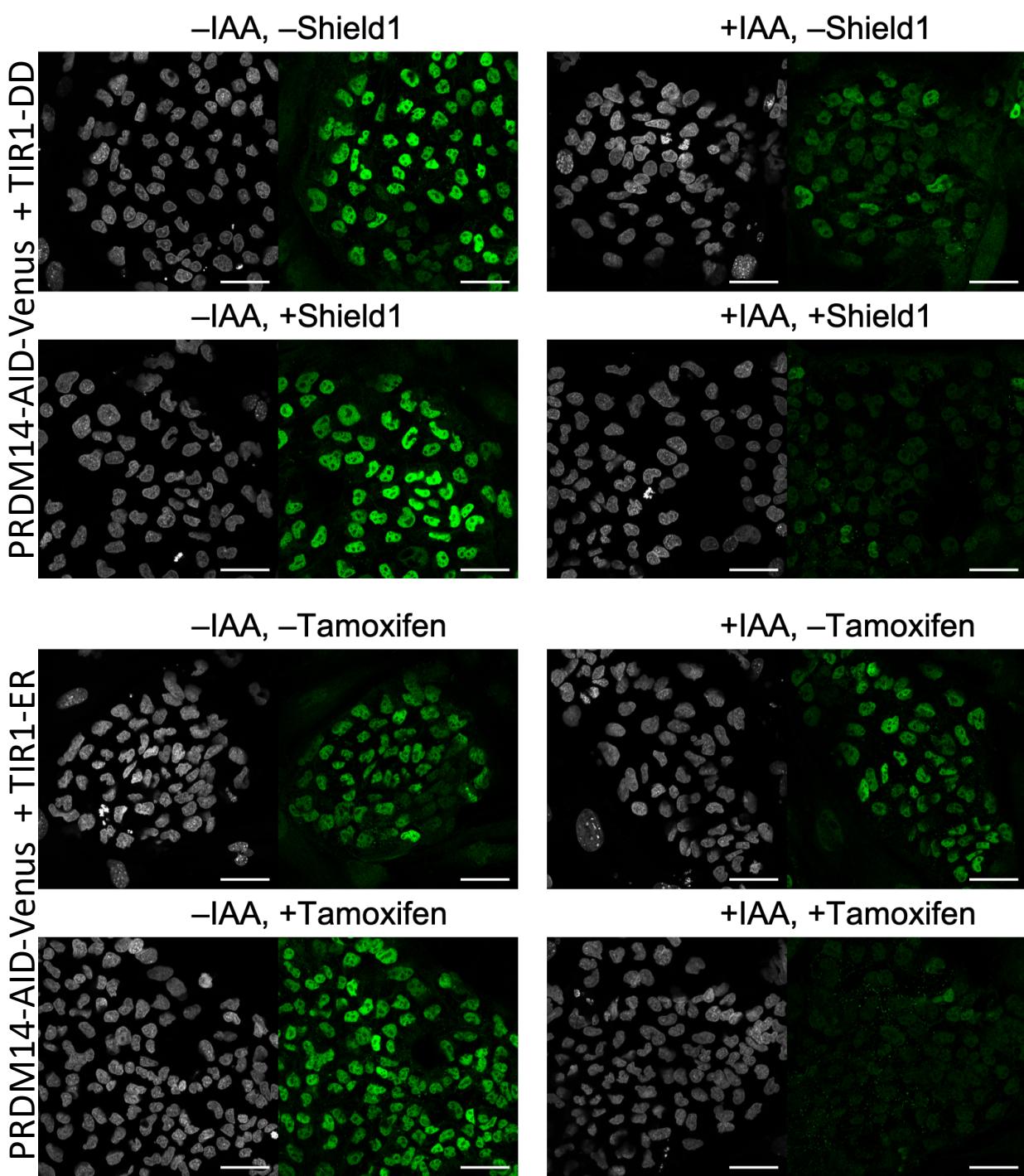


**Figure 12.** NANOS1-AID-Venus expression is not observed by immunofluorescence in a section of a day 4 EB, even in SOX17-positive PGCLCs. Scale bar is 50  $\mu$ m.

#### Inducible TIR1 and SOX17-AID:

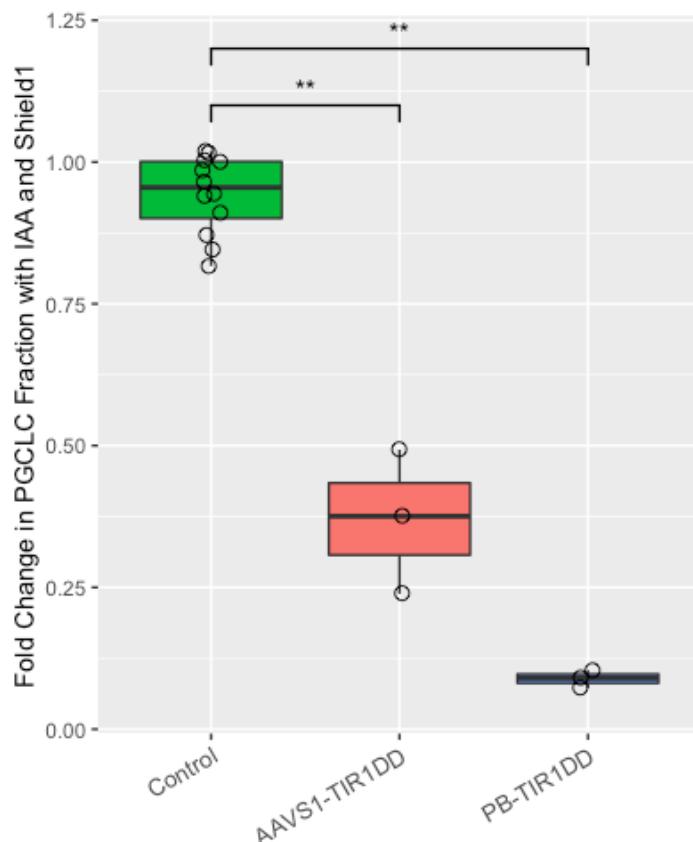
SOX17 is known to play a crucial role in PGC specification and is expressed from an early stage in the process. The AID system was previously applied to SOX17 by other members of the Surani lab. Although the parental SOX17-AID-Venus cell line was able to form PGCLCs, none of the clones transfected with TIR1 was competent for PGCLC specification, even in the absence of auxin. I hypothesized that this was due to leaky degradation of SOX17 by TIR1. Although the AID system generally maintains normal protein levels in the absence of auxin, there are a few known examples of targets where leaky degradation is an issue.<sup>37</sup>

To overcome this leakiness, I created cell lines expressing inducible TIR1, with C-terminal DD or estrogen receptor (ER) fusions. These induction systems were chosen for their rapid kinetics, compatible with AID. As a preliminary test of kinetics, I transfected PRDM14-AID-Venus hESCs with TIR1-DD and TIR1-ER expression vectors, delivered by PiggyBac transposase. Unlike SOX17, SOX15, and KLF4, PRDM14 is expressed in primed hESCs, which means the AID system can be tested directly in those cells. By immunofluorescence, TIR1-DD and TIR1-ER hESC lines depleted PRDM14 after 1 hour of treatment with IAA and Shield1 (for DD) or tamoxifen (for ER) (Figure 13). However, a few cells retained low levels of PRDM14. Since PRDM14-AID does not suffer from the leakiness issue, it was not possible to assess whether the inducible TIR1 still suffered from that issue. Instead, this experiment merely showed that inducible TIR1 could deplete PRDM14 with kinetics similar to constitutive TIR1, albeit with slight heterogeneity.

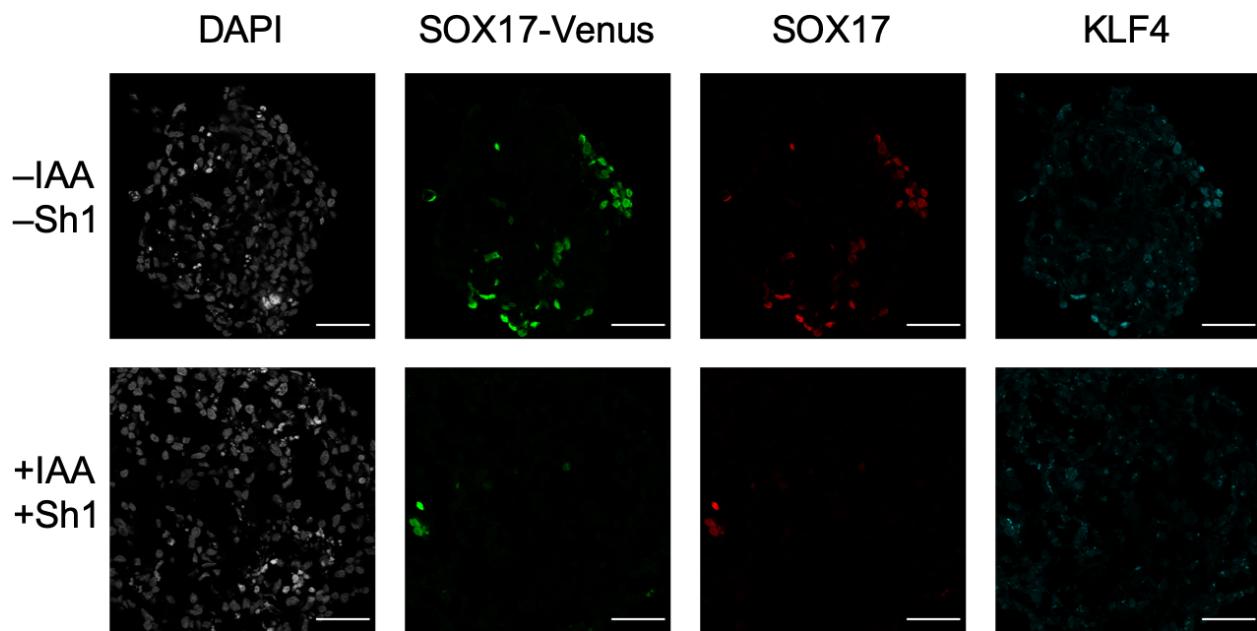


**Figure 13.** Depletion of PRDM14-AID-Venus by inducible TIR1. hESCs were treated with IAA and either Shield1 (for TIR1-DD) or tamoxifen (for TIR1-ER). After one hour of treatment, cells were fixed and stained for immunofluorescence using anti-GFP. Scale bar is 50  $\mu$ m.

Somewhat encouraged by these results with PRDM14-AID, I next applied the inducible TIR1 to SOX17-AID. I first generated a cell line with EF-TIR1-DD knocked in to the AAVS1 locus. This cell line successfully formed PGCLCs with typical efficiency 10–15%. Although this was not quite as efficient as wild-type SOX17 cell lines (typically 30–50%), it still was an improvement over constitutive TIR1 cell lines, which did not form PGCLCs at all. However, the AAVS1-TIR1DD cells showed only a moderate depletion of SOX17 with IAA and Shield1 treatment (Figure 14). Apparently the two copies of TIR1-DD at the AAVS1 locus were insufficient, so I used PiggyBac transposase to deliver additional copies. After screening clones, I identified four that were competent for PGCLC specification but also depleted SOX17 almost completely with IAA and Shield1. As expected, this resulted in drastically reduced specification efficiency (Figure 14) The efficiency in the absence of IAA and Shield1 was similar to the AAVS1-TIR1DD cell line. Notably, the few remaining PGCLCs were all SOX17-positive by flow cytometry and immunofluorescence (Figure 15) This indicates that the presence of these PGCLCs was due to slightly heterogeneous depletion, rather than SOX17 being unnecessary. Although my experiments with SOX17-AID ended here so I could focus on KLF4, hopefully these cells will be a useful resource to other researchers interested in investigating the timing requirements of SOX17 during PGCLC specification.



**Figure 14.** PGCLC induction is blocked by SOX17 depletion. hESCs were induced to form PGCLCs and treated with IAA and Shield1 from the start of the induction. The fraction of NANOS3+/AP+ PGCLCs was measured by flow cytometry after four days, and normalized with respect to untreated samples of the same clones.

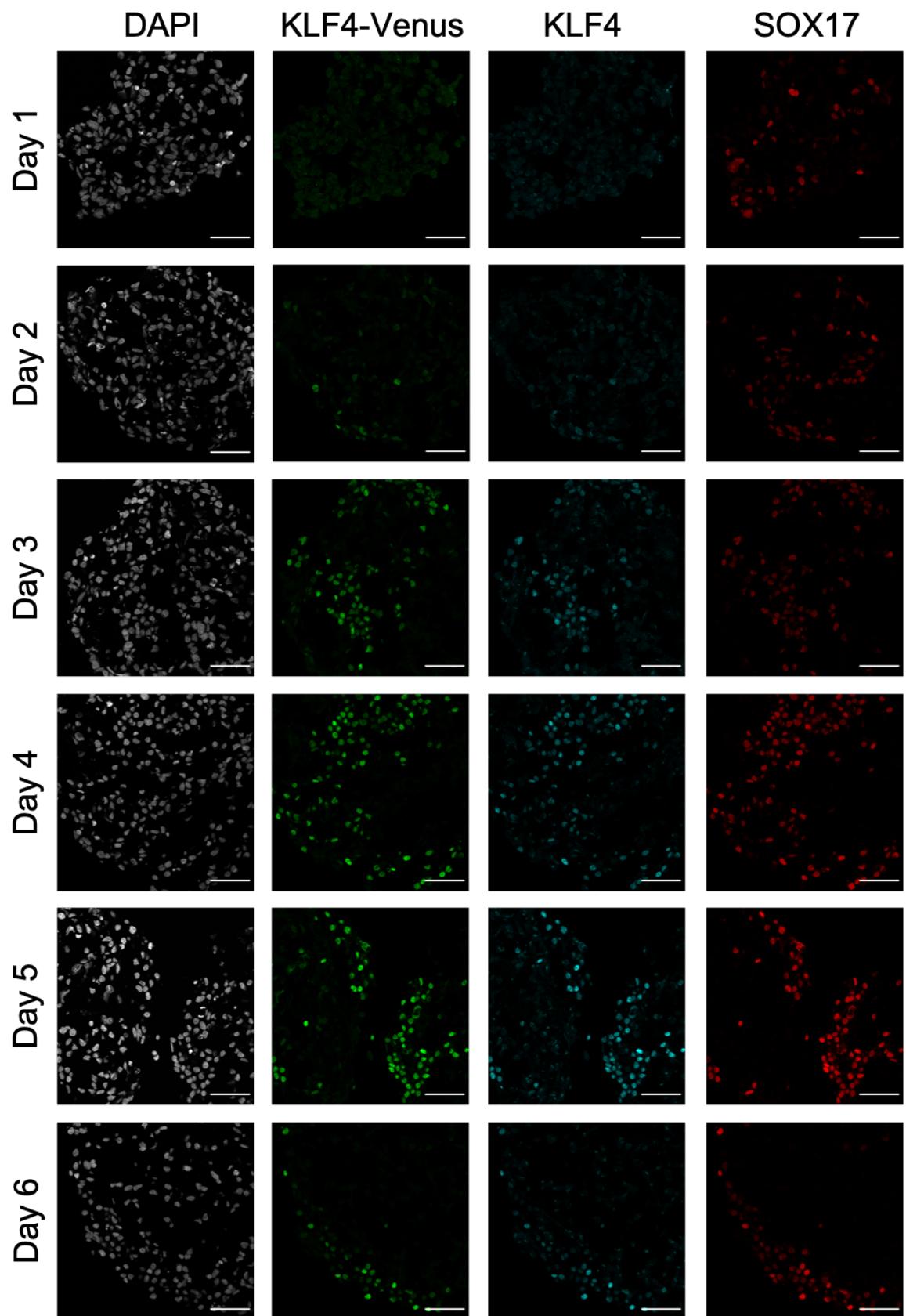


**Figure 15.** SOX17 depletion by AID with TIR1-DD. SOX17-AID-Venus cells with PB-TIR1-DD were induced to form PGCLCs and treated with IAA and Shield1 from the start of the induction. After four days, EBs were fixed, sectioned, and stained using antibodies against GFP, SOX17, and KLF4. The untreated sample contains many SOX17 positive PGCLCs, whereas only a few are visible in the treated sample. Scale bar is 50  $\mu$ m.

#### KLF4-AID:

KLF4, like SOX15 and SOX17, is a transcription factor expressed in the human germline but not the mouse germline.<sup>2</sup> Additionally, it is known to have a role in naïve pluripotency.<sup>19</sup> Thus, it is an attractive target for studying by the AID system. Unfortunately, initial attempts at creating a KLF4-AID cell line yielded only heterozygotes. My next strategy was transfecting a heterozygous KLF4-AID-Venus-Puro cell line with Cas9/gRNA plasmid and KLF4-AID-Venus-Neo donor. For reasons that are still unknown, this was also unsuccessful. The gRNA site was intact in the wild-type allele, so this was not the issue. Eventually, I targeted a heterozygous KLF4-AID-Venus-Neo cell line with KLF4-AID-Venus-Puro. This successfully produced homozygous AID Neo/Puro clones, from which I removed both selectable markers by Dre excision performed in parallel with TIR1 transfection.

As a preliminary test, I induced parental KLF4-AID-Venus cells into PGCLCs and collected samples of EBs on days 1–6. By immunofluorescence, KLF4-Venus expression was present in SOX17-positive PGCLCs, using both anti-GFP and anti-KLF4 antibodies (Figure 16). Whereas SOX17 was expressed from day 1, KLF4 was not visible until day 2, and only reached its full intensity on day 3. Interestingly, on EBs from day 5 and especially day 6, the

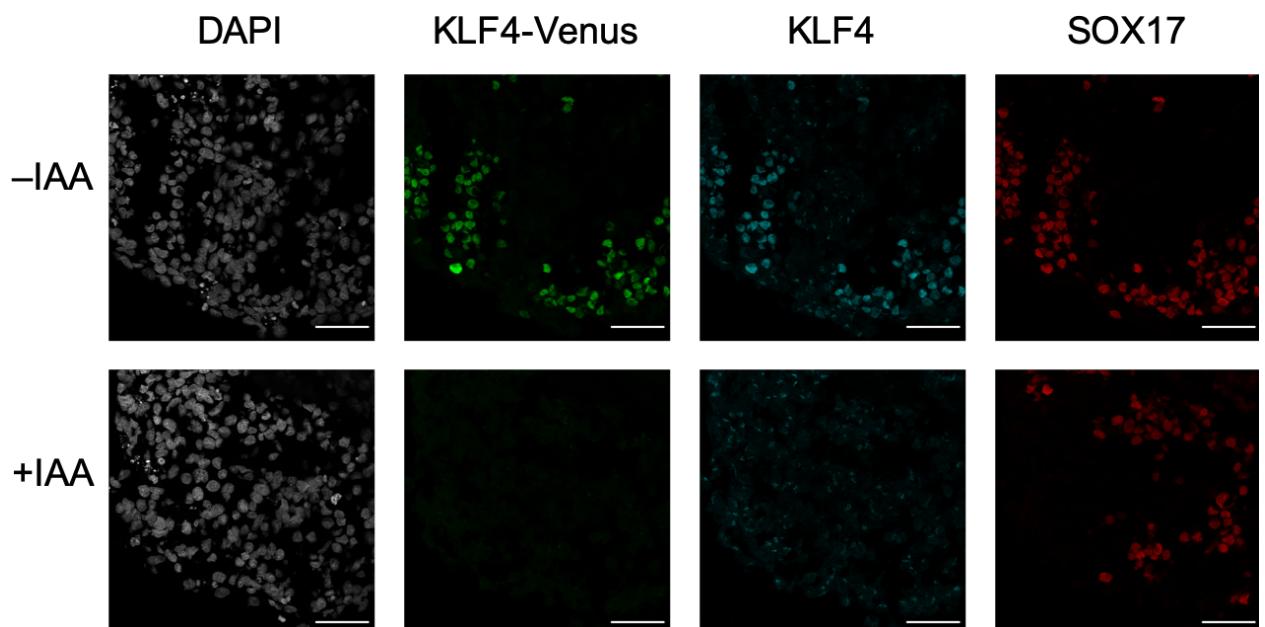


**Figure 16.** Timecourse immunofluorescence of KLF4-AID- Venus EBs stained with DAPI (grey), anti-Venus (green), anti-SOX17 (red), and anti-KLF4 (cyan). Scale bar is 50  $\mu$ m. SOX17 expression is present from day 1, and KLF4 expression is observed to begin in SOX17 positive cells on day 2 and persist for the remainder of the experiment.

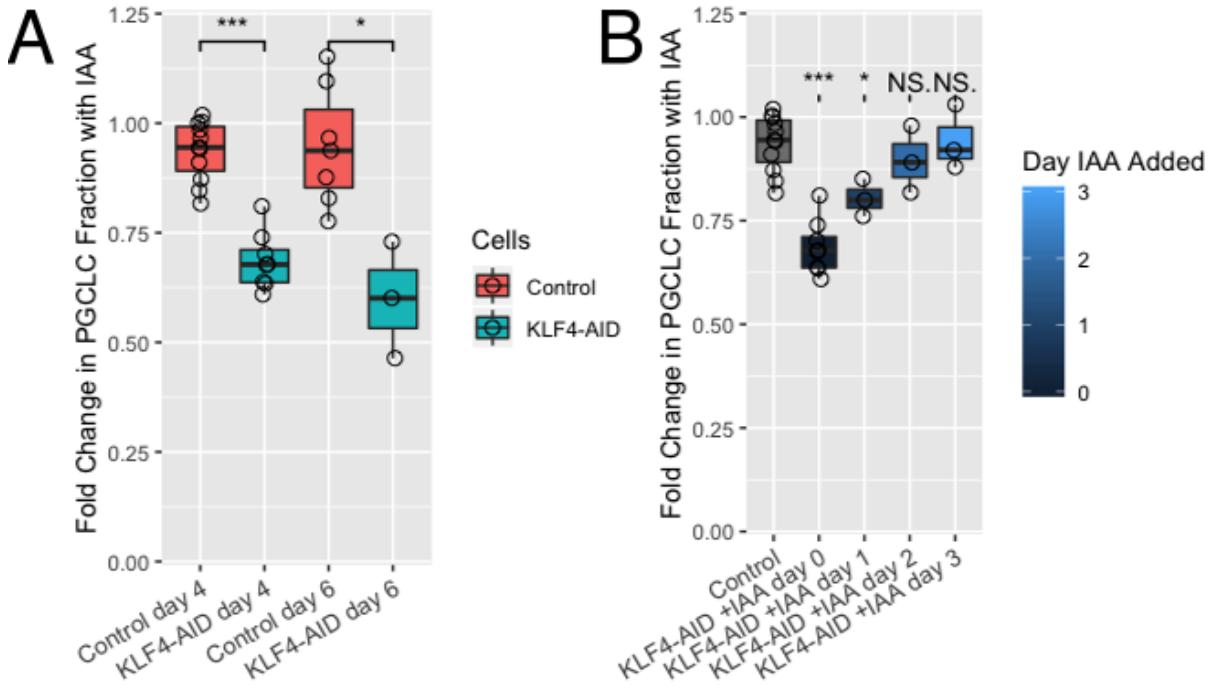
PGCLCs were concentrated around the edges. This matches the pattern observed in the SOX15-AID day 5 and 6 EBs.

Next, I performed PGCLC inductions with KLF4 depletion. I tested a total of seven clones, six of which successfully depleted KLF4 by flow cytometry. I confirmed KLF4 depletion by immunofluorescence using both anti-KLF4 and anti-GFP antibodies (Figure 17). One of the clones had low expression of KLF4 even without IAA, but the remaining four had expression similar to wild-type. I chose the three clones with the highest untreated KLF4 expression to use in subsequent experiments.

By flow cytometry, PGCLC specification efficiency was significantly reduced in PGCLCs depleted of KLF4 (Figure 18A). This effect was observed both on day 4 and day 6 after induction. Unlike with SOX15, the magnitude of the decrease was equivalent at the two timepoints. To assess the temporal requirement for KLF4 during PGCLC specification, I performed another induction, with IAA added on day 0, day 1, day 2, or day 3. Measurement of PGCLC fraction by flow cytometry on day 4 revealed that the effect of KLF4 depletion decreased at later timepoints, and was only statistically significant on day 0 and day 1 (Wilcoxon test,  $p < .05$ ) (Figure 18B).



**Figure 17.** KLF4-AID-Venus expression is observed in SOX17-positive PGCLCs. Panels depict immunofluorescence staining using anti-GFP (green), anti-KLF4 (cyan), and anti-SOX17 (red) on sections of day 4 EBs. KLF4-AID-Venus, but not SOX17, is depleted to background levels with IAA treatment. Scale bar is 50  $\mu$ m.



**Figure 18.** KLF4 depletion with IAA treatment decreases NANOS3+/AP+ PGCLC fraction measured by flow cytometry. Fractions were normalized with respect to untreated samples of the same clones. Statistical comparisons were performed between KLF4-AID/TIR1 clones and control clones without TIR1, which did not deplete KLF4. **A:** IAA treatment from the start of induction, with PGCLC fraction measured on day 4 or 6. **B:** IAA treatment beginning on the day indicated, with PGCLC fraction measured on day 4. Significance values are by two-tailed Wilcoxon test (\* p < 0.05, \*\* p < 0.01, \*\*\*p < 0.001)

## Discussion

### The role of SOX15 in PGCLCs:

Based on previous single-cell RNA-seq data, SOX15 had been suggested to be a critical regulator of human germ cell identity. In both human<sup>3</sup> and porcine<sup>38</sup> PGCs, SOX15 is more homogeneously expressed than SOX17. However, AID depletion of SOX15 during PGCLC specification did not result in significant reduction in specification efficiency, as would be expected if it were essential. Unlike the dramatic effect seen with SOX17 depletion, SOX15 depletion only resulted in a moderate decrease in PGCLC fraction, and this effect was only significant at later timepoints (days 6 and 8). Furthermore, by immunofluorescence, SOX15 is expressed only after day 2 of PGCLC induction. This is later than the expression of SOX17 and BLIMP1, so germ cell identity is already established by this time. These results do not support the claim by Guo *et al.* that SOX15 might be functionally important for germline specification.

However, SOX15 may play a role in maintenance of germ cell identity. AID experiments showed that prolonged SOX15 depletion decreased the PGCLC fraction in EBs, with the effect increasing over time. Furthermore, overexpression of SOX15 increased the PGCLC fraction, again with the effect increasing over time. Limitations in current methods for culturing PGCLCs make it difficult to obtain meaningful results beyond day 8, but it may well be the case that SOX15 is required for long-term maintenance of PGCLCs.

In addition to measuring the changes in PGCLC fraction upon SOX15 depletion or overexpression, I also determined some transcriptional effects by qPCR. In PGCLCs, *PRDM14*, *AKAP1*, *BEND4*, *VENTX*, *SOX15*, and *NANOG* were significantly upregulated when SOX15 was depleted. Furthermore, *PRDM14*, *AKAP1*, *VENTX*, and *ELF3* were significantly downregulated in PGCLCs when SOX15 was overexpressed. It seems that SOX15 protein levels are anticorrelated with PRDM14 expression in PGCLCs. Effects on *AKAP1*, *BEND4*, and *VENTX* are probably best explained as downstream effects, since these genes are known PRDM14 targets. Also, the effects on *SOX15* RNA levels in the AID experiments indicate negative feedback.

Overexpression of SOX15 in somatic lineages of the EBs did not have consistent effects on *PRDM14*, but this is likely because it is repressed by other factors, as its expression was very low. Instead, *PDCD4* was significantly downregulated, and *SOX17* and *POU5F1* upregulated. The expression levels of the latter factors were very low to begin with, but this effect is nonetheless interesting because of their roles in the germline. When SOX15 was depleted in TCam-2 cells by RNAi, *ELF3*, *HNFB1*, *OTX2*, *GATA6*, *ID2*, and *PDCD4* were all significantly upregulated. Among these, *ELF3* was the only gene whose changes were consistent with those in PGCLCs. (*OTX2* expression in PGCLCs was too low to measure, so effects could not be determined.)

Although these transcriptional effects are interesting, they should be interpreted with caution. First, they are quite possibly indirect effects, especially in the case of *AKAP1*, *VENTX*, and *BEND4*. Also, their magnitude is relatively moderate, with the greatest fold change being ~2 (for *PRDM14*). Furthermore, the inconsistencies in the effects between AID depletion in PGCLCs and RNAi depletion in TCam-2, and between SOX15 overexpression in PGCLCs and soma, are cause for concern. More broadly, many of the transcriptional effects reported in other cell types<sup>17,32–35</sup> did not replicate in my experiments. Since the effects evidently differ in different contexts, they may also differ from what would be observed in

*vivo*. Finally, I only tested a total of 32 candidate genes. Although these genes were selected based on their likelihood of being SOX15 targets, they nonetheless represent only a small fraction of the transcriptome, and there are quite likely other SOX15 targets that are more functionally relevant. Overall, I believe that among the effects I have identified, only the effect on *PRDM14* expression is likely to be important for understanding the role of SOX15 in PGCLCs.

The effects of SOX15 perturbation on human PGCLC maintenance and transcriptional activity are best understood in context of its role in other cell types. Although SOX15 has not been investigated nearly as much as other SOX factors, the existing research on SOX15 suggests a role related to preventing improper growth and differentiation. In myogenic progenitors, SOX15 promotes satellite cell maintenance, and thus has an important role in muscle regeneration.<sup>15,33,39,40</sup> In embryonal carcinoma,<sup>32</sup> and esophageal<sup>34</sup> and pancreatic<sup>35</sup> adenocarcinomas, SOX15 acts as a tumor suppressor and lack of SOX15 is associated with aberrant growth. The suppressive action of SOX15 may be mediated through downregulation of Wnt pathway components.<sup>35,41</sup> Notably, Wnt signaling promotes germline competence in mouse, pig, and human pluripotent cells, but after germline specification, excess Wnt signaling is detrimental.<sup>11,36,42</sup> Since *PRDM14* is also known to repress Wnt targets,<sup>36</sup> the anticorrelation of *PRDM14* with respect to SOX15 perturbations may be a compensatory mechanism to maintain Wnt signaling within the range compatible with germline identity.

#### **Lack of NANOS1 expression in PGCLCs:**

NANOS3 protein is strongly expressed in the early germline, so much so that it can be used as a reporter for PGCLCs. Therefore, I expected that the related protein NANOS1 would also be present, since its mRNA levels are even higher than those of *NANOS3* in PGCLCs.<sup>10</sup> Although NANOS family mRNAs are known to be under translational control,<sup>43,44</sup> I assumed that if *NANOS3* mRNA could be translated, then *NANOS1* could be as well. As it turned out, my expectations were incorrect. The tagged NANOS1-AID-Venus protein was absent in PGCLCs, and overexpression of myc-NANOS1-DD was also unsuccessful. The latter result is particularly notable, since only the coding sequence of NANOS1 was used, eliminating the possibility of repressive elements in the UTRs. However, *NANOS1* mRNA also contains some regulatory elements within the coding sequence.<sup>45</sup> In particular, hairpin-like translational control elements near the beginning of the sequence are known to block

translation in *Xenopus*, with the mRNA only being translated in the presence of the Dnd1 helicase.<sup>43</sup> Notably, human DND1 is lowly expressed in PGCLCs. Another possible factor preventing *NANOS1* mRNA translation is SAMD4B, which is strongly expressed in PGCLCs. This is homologous to the mouse Smaug2, which is known to repress *Nanos1* translation.<sup>46</sup> Regardless of the mechanism of translational repression, my main conclusion is that *NANOS1* protein has no role in PGCLCs since it is not present.

#### **Inducible TIR1 and SOX17-AID:**

The AID system is a powerful tool for interrogating protein function, but for some target proteins it suffers from leakiness. Examples of such targets include DHC1,<sup>37</sup> ZNF143,<sup>47</sup> Prp22 (in yeast),<sup>48</sup> and SOX17. Inducible TIR1 systems can help overcome this leakiness, since if TIR1 is only present when needed, it cannot inappropriately degrade the target. An inducible TIR1 system has been reported in the literature;<sup>48</sup> however, this system uses transcriptional regulation for TIR1 induction, which is incompatible with the rapid kinetics of AID. In my experiments, I demonstrated protein-level TIR1 induction using two different fusion proteins: TIR1-DD and TIR1-ER. I then applied TIR1-DD to successfully perform AID targeting SOX17 in PGCLCs, for which constitutive TIR1 had been unacceptably leaky. As expected, depletion of SOX17 resulted in a dramatic loss of PGCLCs. The results of these experiments further support a crucial role for SOX17 in PGCLC specification and provide a method for applying AID to previously problematic targets.

#### **KLF4 in naïve pluripotency and germline identity:**

Although human PGCs and PGCLCs arise from primed pluripotent cells, many of their gene expression patterns resemble a naïve pluripotent state. For example, SOX15, REX1, DPPA3 (Stella), and KLF4 are all expressed both in the germline<sup>3</sup> and in naïve ESCs, but not in primed ESCs.<sup>20,21</sup> Besides SOX15, KLF4 is particularly interesting for its known role in regulating transposable elements (TEs). In naïve ESCs, KLF4 binds LTR and SVA-associated enhancers and opens the surrounding chromatin, permitting transcription. However, it also activates KZNFs which subsequently repress these transposable elements. In early human germline development, DNA is globally demethylated and chromatin is open, yet TEs remain repressed even before piRNAs become active at later stages.<sup>4</sup> Although it is currently unknown how TEs are repressed during this vulnerable window, a reasonable hypothesis is that KLF4 is involved via induction of KZNFs.

During my research, I generated a KLF4-AID-Venus hESC line, which allowed a preliminary study of KLF4 in PGCLCs. By immunofluorescence, KLF4 showed a very faint signal on day 1 after PGCLC induction, which became much stronger on day 2 and persisted in PGCLCs until the end of the experiment (day 6). The timing of KLF4 expression suggests that it is downstream of early germline specifiers such as SOX17 and BLIMP1, which are expressed from day 1 in the course of PGCLC specification. Depletion of KLF4 from PGCLCs by AID resulted in a decrease in PGCLC fraction by approximately 30%. Unlike with SOX15, this effect was of a similar magnitude when measured on days 4 and 6 after induction. These results indicate that KLF4 has an important role in PGCLC specification, which merits further investigation.

Redundancy among Krüppel-like factors may explain why KLF4 depletion only partially reduces PGCLC fraction. In mouse ESCs, KLF4 is doubly redundant with its paralogs KLF2 and KLF5. Deletion of any two of the three factors does not interfere with ESC maintenance.<sup>49</sup> Only when all three are deleted do any issues arise. KLF2 is not expressed in human PGCLCs, but KLF5 is, albeit at moderately lower levels than KLF4.<sup>10</sup> Thus, KLF5 could be partially compensating for KLF4 depletion in PGCLCs. Ongoing research, including RNA-seq experiments, should clarify the role of KLF4 in PGCLCs and the effects of its depletion.

### **Overall conclusions:**

In this research, I used genetic tools to manipulate the levels of regulatory proteins to better understand their roles in human germline specification. In depletion and overexpression experiments on SOX15, I found that it is not required for establishment of germ cell identity but may promote its maintenance. I also identified a few transcriptional effects related to changes in SOX15 protein levels, mainly involving PRDM14 and its targets. I attempted to similarly manipulate the NANOS1 RNA-binding protein, but I found that it was not present in PGCLCs, likely due to translational repression of its mRNA. Additionally, I demonstrated a method to overcome AID leakiness by using inducible TIR1 fusion proteins. I applied this method to deplete SOX17 and observed results consistent with its crucial role as a germline specifier. Finally, I created KLF4-AID ESC lines, which show a moderate reduction in PGCLC specification efficiency in presence of IAA, and which will be useful in future experiments to study the function of this transcription factor.

## References

- (1) Tang, W. W. C.; Kobayashi, T.; Irie, N.; Dietmann, S.; Surani, M. A. Specification and Epigenetic Programming of the Human Germ Line. *Nat. Rev. Genet.* **2016**, *17* (10), 585–600. <https://doi.org/10.1038/nrg.2016.88>.
- (2) Kobayashi, T.; Surani, M. A. On the Origin of the Human Germline. *Development* **2018**, *145* (16), dev150433. <https://doi.org/10.1242/dev.150433>.
- (3) Guo, F.; Yan, L.; Guo, H.; Li, L.; Hu, B.; Zhao, Y.; Yong, J.; Hu, Y.; Wang, X.; Wei, Y.; et al. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. *Cell* **2015**, *161* (6), 1437–1452. <https://doi.org/10.1016/j.cell.2015.05.015>.
- (4) Tang, W. W. C.; Dietmann, S.; Irie, N.; Leitch, H. G.; Floros, V. I.; Bradshaw, C. R.; Hackett, J. A.; Chinnery, P. F.; Surani, M. A. A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. *Cell* **2015**, *161* (6), 1453–1467. <https://doi.org/10.1016/j.cell.2015.04.053>.
- (5) Irie, N.; Surani, M. A. Efficient Induction and Isolation of Human Primordial Germ Cell-Like Cells from Competent Human Pluripotent Stem Cells. In *Germline Stem Cells*; Buszczak, M., Ed.; Springer New York: New York, NY, 2017; Vol. 1463, pp 217–226. [https://doi.org/10.1007/978-1-4939-4017-2\\_16](https://doi.org/10.1007/978-1-4939-4017-2_16).
- (6) Hikabe, O.; Hamazaki, N.; Nagamatsu, G.; Obata, Y.; Hirao, Y.; Hamada, N.; Shimamoto, S.; Imamura, T.; Nakashima, K.; Saitou, M.; et al. Reconstitution in Vitro of the Entire Cycle of the Mouse Female Germ Line. *Nature* **2016**, *539* (7628), 299–303. <https://doi.org/10.1038/nature20104>.
- (7) Yamashiro, C.; Sasaki, K.; Yabuta, Y.; Kojima, Y.; Nakamura, T.; Okamoto, I.; Yokobayashi, S.; Murase, Y.; Ishikura, Y.; Shirane, K.; et al. Generation of Human Oogonia from Induced Pluripotent Stem Cells in Vitro. *Science* **2018**, *362* (6412), 356–360. <https://doi.org/10.1126/science.aat1674>.
- (8) Magnúsdóttir, E.; Dietmann, S.; Murakami, K.; Günesdogan, U.; Tang, F.; Bao, S.; Diamanti, E.; Lao, K.; Gottgens, B.; Azim Surani, M. A Tripartite Transcription Factor Network Regulates Primordial Germ Cell Specification in Mice. *Nat. Cell Biol.* **2013**, *15* (8), 905–915. <https://doi.org/10.1038/ncb2798>.
- (9) Nakaki, F.; Hayashi, K.; Ohta, H.; Kurimoto, K.; Yabuta, Y.; Saitou, M. Induction of Mouse Germ-Cell Fate by Transcription Factors in Vitro. *Nature* **2013**, *501* (7466), 222–226. <https://doi.org/10.1038/nature12417>.
- (10) Irie, N.; Weinberger, L.; Tang, W. W. C.; Kobayashi, T.; Viukov, S.; Manor, Y. S.; Dietmann, S.; Hanna, J. H.; Surani, M. A. SOX17 Is a Critical Specifier of Human Primordial Germ Cell Fate. *Cell* **2015**, *160* (1–2), 253–268. <https://doi.org/10.1016/j.cell.2014.12.013>.
- (11) Kobayashi, T.; Zhang, H.; Tang, W. W. C.; Irie, N.; Withey, S.; Klisch, D.; Sybirna, A.; Dietmann, S.; Contreras, D. A.; Webb, R.; et al. Principles of Early Human Development and Germ Cell Program from Conserved Model Systems. *Nature* **2017**, *546* (7658), 416–420. <https://doi.org/10.1038/nature22812>.
- (12) Hackett, J. A.; Surani, M. A. Regulatory Principles of Pluripotency: From the Ground State Up. *Cell Stem Cell* **2014**, *15* (4), 416–430. <https://doi.org/10.1016/j.stem.2014.09.015>.
- (13) Sarraj, M. A.; Wilmore, H. P.; McClive, P. J.; Sinclair, A. H. Sox15 Is up Regulated in the Embryonic Mouse Testis. *Gene Expr. Patterns* **2003**, *3* (4), 413–417. [https://doi.org/10.1016/S1567-133X\(03\)00085-1](https://doi.org/10.1016/S1567-133X(03)00085-1).

- (14) Yamada, K.; Kanda, H.; Aihara, T.; Takamatsu, N.; Shiba, T.; Ito, M. Mammalian *Sox15* Gene: Promoter Analysis and Implications for Placental Evolution. *Zoolog. Sci.* **2008**, *25* (3), 313–320. <https://doi.org/10.2108/zsj.25.313>.
- (15) Lee, H.-J.; Goring, W.; Ochs, M.; Muhrfeld, C.; Steding, G.; Paprotta, I.; Engel, W.; Adham, I. M. *Sox15* Is Required for Skeletal Muscle Regeneration. *Mol. Cell. Biol.* **2004**, *24* (19), 8428–8436. <https://doi.org/10.1128/MCB.24.19.8428-8436.2004>.
- (16) Kamachi, Y.; Kondoh, H. Sox Proteins: Regulators of Cell Fate Specification and Differentiation. *Development* **2013**, *140* (20), 4129–4144. <https://doi.org/10.1242/dev.091793>.
- (17) Maruyama, M.; Ichisaka, T.; Nakagawa, M.; Yamanaka, S. Differential Roles for *Sox15* and *Sox2* in Transcriptional Control in Mouse Embryonic Stem Cells. *J. Biol. Chem.* **2005**, *280* (26), 24371–24379. <https://doi.org/10.1074/jbc.M501423200>.
- (18) Campolo, F.; Gori, M.; Favaro, R.; Nicolis, S.; Pellegrini, M.; Botti, F.; Rossi, P.; Jannini, E. A.; Dolci, S. Essential Role of *Sox2* for the Establishment and Maintenance of the Germ Cell Line. *STEM CELLS* **2013**, *31* (7), 1408–1421. <https://doi.org/10.1002/stem.1392>.
- (19) Pontis, J.; Planet, E.; Offner, S.; Turelli, P.; Duc, J.; Coudray, A.; Theunissen, T. W.; Jaenisch, R.; Trono, D. Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* **2019**, *24* (5), 724–735.e5. <https://doi.org/10.1016/j.stem.2019.03.012>.
- (20) Leitch, H. G.; Smith, A. The Mammalian Germline as a Pluripotency Cycle. *Development* **2013**, *140* (12), 2495–2501. <https://doi.org/10.1242/dev.091603>.
- (21) Hackett, J. A.; Kobayashi, T.; Dietmann, S.; Surani, M. A. Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. *Stem Cell Rep.* **2017**, *8* (6), 1645–1658. <https://doi.org/10.1016/j.stemcr.2017.05.014>.
- (22) Julaton, V. T. A.; Reijo Pera, R. A. NANOS3 Function in Human Germ Cell Development. *Hum. Mol. Genet.* **2011**, *20* (11), 2238–2250. <https://doi.org/10.1093/hmg/ddr114>.
- (23) Kee, K.; Angeles, V. T.; Flores, M.; Nguyen, H. N.; Reijo Pera, R. A. Human DAZL, DAZ and BOULE Genes Modulate Primordial Germ-Cell and Haploid Gamete Formation. *Nature* **2009**, *462* (7270), 222–225. <https://doi.org/10.1038/nature08562>.
- (24) Jin, Y. N.; Schlueter, P. J.; Jurisch-Yaksi, N.; Lam, P.-Y.; Jin, S.; Hwang, W. Y.; Yeh, J.-R. J.; Yoshigi, M.; Ong, S.-E.; Schenone, M.; et al. Noncanonical Translation via Deadenylated 3' UTRs Maintains Primordial Germ Cells. *Nat. Chem. Biol.* **2018**, *14* (9), 844–852. <https://doi.org/10.1038/s41589-018-0098-0>.
- (25) Natsume, T.; Kiyomitsu, T.; Saga, Y.; Kanemaki, M. T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep.* **2016**, *15* (1), 210–218. <https://doi.org/10.1016/j.celrep.2016.03.001>.
- (26) Maynard-Smith, L. A.; Chen, L.; Banaszynski, L. A.; Ooi, A. G. L.; Wandless, T. J. A Directed Approach for Engineering Conditional Protein Stability Using Biologically Silent Small Molecules. *J. Biol. Chem.* **2007**, *282* (34), 24866–24872. <https://doi.org/10.1074/jbc.M703902200>.
- (27) Gafni, O.; Weinberger, L.; Mansour, A. A.; Manor, Y. S.; Chomsky, E.; Ben-Yosef, D.; Kalma, Y.; Viukov, S.; Maza, I.; Zviran, A.; et al. Derivation of Novel Human Ground State Naive Pluripotent Stem Cells. *Nature* **2013**, *504* (7479), 282–286. <https://doi.org/10.1038/nature12745>.
- (28) Slaymaker, I. M.; Gao, L.; Zetsche, B.; Scott, D. A.; Yan, W. X.; Zhang, F. Rationally Engineered Cas9 Nucleases with Improved Specificity. *Science* **2016**, *351* (6268), 84–88. <https://doi.org/10.1126/science.aad5227>.

- (29) Heath-Pagliuso, S.; Rogers, W. J.; Tullis, K.; Seidel, S. D.; Cenijn, P. H.; Brouwer, A.; Denison, M. S. Activation of the Ah Receptor by Tryptophan and Tryptophan Metabolites <sup>†</sup>. *Biochemistry* **1998**, *37* (33), 11508–11515. <https://doi.org/10.1021/bi980087p>.
- (30) Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* **2012**, *9* (7), 676–682. <https://doi.org/10.1038/nmeth.2019>.
- (31) de Jong, J.; Stoop, H.; Gillis, A. J. M.; Hersmus, R.; van Gurp, R. J. H. L. M.; van de Geijn, G.-J. M.; van Drunen, E.; Beverloo, H. B.; Schneider, D. T.; Sherlock, J. K.; et al. Further Characterization of the First Seminoma Cell Line TCam-2. *Genes. Chromosomes Cancer* **2008**, *47* (3), 185–196. <https://doi.org/10.1002/gcc.20520>.
- (32) Yan, H.-T.; Shinka, T.; Sato, Y.; Yang, X.-J.; Chen, G.; Sakamoto, K.; Kinoshita, K.; Aburatani, H.; Nakahori, Y. Overexpression of SOX15 Inhibits Proliferation of NT2/D1 Cells Derived from a Testicular Embryonal Cell Carcinoma. *Mol. Cells* **24** (3), 323–328.
- (33) Meeson, A. P.; Shi, X.; Alexander, M. S.; Williams, R. S.; Allen, R. E.; Jiang, N.; Adham, I. M.; Goetsch, S. C.; Hammer, R. E.; Garry, D. J. Sox15 and Fhl3 Transcriptionally Coactivate Foxk1 and Regulate Myogenic Progenitor Cells. *EMBO J.* **2007**, *26* (7), 1902–1912. <https://doi.org/10.1038/sj.emboj.7601635>.
- (34) Sulahian, R.; Chen, J.; Arany, Z.; Jadhav, U.; Peng, S.; Rustgi, A. K.; Bass, A. J.; Srivastava, A.; Hornick, J. L.; Shivedasani, R. A. SOX15 Governs Transcription in Human Stratified Epithelia and a Subset of Esophageal Adenocarcinomas. *Cell. Mol. Gastroenterol. Hepatol.* **2015**, *1* (6), 598–609.e6. <https://doi.org/10.1016/j.jcmgh.2015.07.009>.
- (35) Thu, K. L.; Radulovich, N.; Becker-Santos, D. D.; Pikor, L. A.; Pusic, A.; Lockwood, W. W.; Lam, W. L.; Tsao, M.-S. SOX15 Is a Candidate Tumor Suppressor in Pancreatic Cancer with a Potential Role in Wnt/β-Catenin Signaling. *Oncogene* **2014**, *33* (3), 279–288. <https://doi.org/10.1038/onc.2012.595>.
- (36) Sybirna, A. Signalling and Transcriptional Regulation of Human Primordial Germ Cell Specification, University of Cambridge, 2018.
- (37) Yesbolatova, A.; Natsume, T.; Hayashi, K.; Kanemaki, M. T. Generation of Conditional Auxin-Inducible Degron (AID) Cells and Tight Control of Degron-Fused Proteins Using the Degradation Inhibitor Auxinole. *Methods* **2019**. <https://doi.org/10.1016/j.ymeth.2019.04.010>.
- (38) Alberio, R. Personal Communication, 2019.
- (39) Béranger, F.; Méjean, C.; Moniot, B.; Berta, P.; Vandromme, M. Muscle Differentiation Is Antagonized by SOX15, a New Member of the SOX Protein Family. *J. Biol. Chem.* **2000**, *275* (21), 16103–16109. <https://doi.org/10.1074/jbc.275.21.16103>.
- (40) Savage, J.; Conley, A. J.; Blais, A.; Skerjanc, I. S. SOX15 and SOX7 Differentially Regulate the Myogenic Program in P19 Cells. *Stem Cells* **2009**, *27* (6), 1231–1243. <https://doi.org/10.1002/stem.57>.
- (41) Moradi, A.; Ghasemi, F.; Anvari, K.; Hassanian, S. M.; Simab, S. A.; Ebrahimi, S.; Hesari, A.; Forghanifard, M. M.; Boroushaki, M. T.; ShahidSales, S.; et al. The Cross-Regulation between SOX15 and Wnt Signaling Pathway. *J. Cell. Physiol.* **2017**, *232* (12), 3221–3225. <https://doi.org/10.1002/jcp.25802>.
- (42) Hackett, J. A.; Huang, Y.; Günesdogan, U.; Gretarsson, K. A.; Kobayashi, T.; Surani, M. A. Tracing the Transitions from Pluripotency to Germ Cell Fate with CRISPR Screening. *Nat. Commun.* **2018**, *9* (1). <https://doi.org/10.1038/s41467-018-06230-0>.

- (43) Aguero, T.; Jin, Z.; Chorghade, S.; Kalsotra, A.; King, M. L.; Yang, J. Maternal Dead-End 1 Promotes Translation of *Nanos1* by Binding the EIF3 Complex. *Development* **2017**, *144* (20), 3755–3765. <https://doi.org/10.1242/dev.152611>.
- (44) Julaton, V. T. A.; Reijo Pera, R. A. NANOS3 Function in Human Germ Cell Development. *Hum. Mol. Genet.* **2011**, *20* (11), 2238–2250. <https://doi.org/10.1093/hmg/ddr114>.
- (45) Luo, X.; Nerlick, S.; An, W.; King, M. L. Xenopus Germline Nanos1 Is Translationally Repressed by a Novel Structure-Based Mechanism. *Development* **2011**, *138* (3), 589–598. <https://doi.org/10.1242/dev.056705>.
- (46) Amadei, G.; Zander, M. A.; Yang, G.; Dumelie, J. G.; Vessey, J. P.; Lipshitz, H. D.; Smibert, C. A.; Kaplan, D. R.; Miller, F. D. A Smaug2-Based Translational Repression Complex Determines the Balance between Precursor Maintenance versus Differentiation during Mammalian Neurogenesis. *J. Neurosci.* **2015**, *35* (47), 15666–15681. <https://doi.org/10.1523/JNEUROSCI.2172-15.2015>.
- (47) Sathyan, K. M.; McKenna, B. D.; Anderson, W. D.; Duarte, F. M.; Core, L.; Guertin, M. J. An Improved Auxin-Inducible Degron System Preserves Native Protein Levels and Enables Rapid and Specific Protein Depletion. *bioRxiv* **2019**. <https://doi.org/10.1101/585927>.
- (48) Mendoza-Ochoa, G. I.; Barrass, J. D.; Terlouw, B. R.; Maudlin, I. E.; de Lucas, S.; Sani, E.; Aslanzadeh, V.; Reid, J. A. E.; Beggs, J. D. A Fast and Tuneable Auxin-Inducible Degron for Depletion of Target Proteins in Budding Yeast. *Yeast* **2019**, *36* (1), 75–81. <https://doi.org/10.1002/yea.3362>.
- (49) Yamane, M.; Ohtsuka, S.; Matsuura, K.; Nakamura, A.; Niwa, H. Overlapping Functions of Krüppel-like Factor Family Members: Targeting Multiple Transcription Factors to Maintain the Naïve Pluripotency of Mouse Embryonic Stem Cells. *Development* **2018**, *145* (10), dev162404. <https://doi.org/10.1242/dev.162404>.