Rockefeller University

Digital Commons @ RU

Student Theses and Dissertations

2021

A Translational Approach to Modeling Unique Aspects of Germ Cell Development During Self-Organization of the Primate Embryo

Rohan R. Soman

A TRANSLATIONAL APPROACH TO MODELING UNIQUE ASPECTS OF GERM CELL

DEVELOPMENT DURING SELF-ORGANIZATION OF THE PRIMATE EMBRYO

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Rohan R. Soman

June 2021

A TRANSLATIONAL APPROACH TO MODELING UNIQUE ASPECTS OF GERM CELL

DEVELOPMENT DURING SELF-ORGANIZATION OF THE PRIMATE EMBRYO

Rohan R. Soman, Ph.D.

The Rockefeller University 2021

The specification of germ cells during embryonic development is vital not only for the

development of an organism, but quite literally for the survival and propagation of its entire

species. Recent work has demonstrated that several aspects of human primordial germ cell

(hPGC) development are specific to primates, necessitating model systems and *in vivo*

validation that is also species specific. In this work, a synthesis of *in vitro* and *in vivo* techniques

is used to investigate hPGC specification within a human embryonic stem cell model of a

gastrulating embryo known as a gastruloid. An hPGC transcriptomic signature that indicates

migratory potential via canonical and novel mechanisms is indicated, raising several potential

candidates for further investigation into the under-studied migratory phase of germ cell

development. We seek to generate validation assays of hPGC function in an embryonic context

by investigating migratory potential in grafts to the chick embryo, and demonstrate that despite

significant differences between chick and human routes of migration *in vivo*, human *in vitro*-

derived PGCs are in fact able to demonstrate migratory behavior in the chick, following chick

migratory patterns and demonstrating a specific homing towards the chick mesonephros and

gonad. These experiments not only provide a functional validation for *in vitro*-derived hPGCs

that is complementary to molecular and epigenetic analysis, but also hint at the elements of hPGC development that are conserved throughout evolution.

The gastruloid system is then used for further investigation into the hPGC niche. The power of this model system relies upon self-organization due to endogenous signaling in response to an exogenous BMP4 initiating signal, in a manner analogous to that found in the human embryo. We harness this power, using single cell image quantification and genetic tools including CRISPR-Cas9 to probe further into the signaling environment of the hPGC niche in the gastruloid model. We elucidate the role of each leg of the BMP4 – WNT – ACTIVIN/NODAL signaling cascade in development of this niche, which determines gastrulation events, in hPGC specification. These experiments not only demonstrate how BMP4 acts independently in addition to this cascade to directly specify hPGCs, but also how all three pathways work in harmony to generate self-organization of hPGCs within their gastruloid niche. By probing further into the ontogeny of hPGC specification, we find that upregulation of BLIMP1 alone, a canonical regulator of PGCs, is sufficient to induce later markers of hPGC fate, and surprisingly is also sufficient to downregulate SOX2 (a marker of epiblast and early ectodermal fate) and upregulate SOX17, which has been previously placed upstream of BLIMP1.

Finally, we propose the marmoset as a good model of primate reproduction and embryogenesis, in an attempt to describe a non-human primate (NHP) system for validation of hPGC characteristics observed *in vitro*. We systematically investigate marmoset embryonic development *in vivo* using serial, high temporal and spatial resolution ultrasound imaging. We

describe the morphological characteristics of implantation, gastrulation, neurulation, and organogenesis, as well as the curious marmoset phenotype of an elongated peri-gastrulation window, incorporating slowed embryonic growth and rapid extraembryonic development. In doing so, we generate an annotated ultrasound atlas of marmoset embryogenesis, and train models to identify developmental stages and predict fertilization ages from a single frame. In addition, we suggest that the extended peri-gastrulation window in the marmoset will provide a unique opportunity to perform in utero genetic editing, lineage tracing, and even allogenic transplantation to complement *in vitro* studies of hPGC development.

 This body of work provides a synthesis of culture techniques, genetic and molecular tools, and imaging systems that will provide a foundation for the exploration of not only human PGC development, but also hopefully a generalizable roadmap for translation between *in vitro* and *in vivo* studies early primate development.

This thesis is dedicated to my mother, for being the role model and source of courage and

belief, without whom this work would never have been possible.

Acknowledgements

I would like to above all else extend my heartfelt thanks to my thesis advisor, Dr. Ali Brivanlou.
He has been an extraordinary mentor from the moments he first welcomed me into his
laboratory. A role model for the courage required to pave new paths in scientific discovery, he
has taught me to ask the big, audacious questions; always with the end goal of achieving a true,
basic understanding of our own biology; always with the mindset that all obstacles are
surmountable with dedication, the right mindset, and the right colleagues and friends. He has
built an institution with a tradition of following these ideas, building a team with the creativity
and skillset to address the big questions, and building it with style. I am indebted to the friends
and mentors I have found within his large and growing laboratory, both past and present:
Zeeshan Ozair, Gist Croft, Albert Ruzo, Fred Etoc, Tomo Haremaki, Alessia Deglincerti, Melissa
Popowski, Christian Markopoulos, Tati Kanno, Anna Yoney, Iain Martyn, Jakob Metzger, Mijo
Simunovic, Corbyn Nchako, Lauren Gerber, Tiago Rito, Tiago Laundos, Tien Phan, Riccardo De
Santis, Szilvia Galgoczi, Mia Yang, Michael Heke, Jean-Marx Santel, Peter Ingrassia, Adam Souza,
and Danielle Little, in addition to the many other members that have spent time in the
laboratory. In particular amongst all, I would like to thank Tati Kanno, my partner for all
transplantation studies, for her courage and audacity to develop a system based upon the far-
fetched notion that similarities across species are just as important as differences.

TABLE OF CONTENTS

Ultrasound Procedures:

Ultrasound data collection and analysis:

Predictive Modeling

Tissue Histology

Animal Care and Use

3.1 *Micropatterned generation of human gastruloid cultures contain a PGC niche*

3.2 *Single cell transcriptomic analysis of hPGCs indicates a progression towards a migratory molecular signature*

3.3 *Gastruloid-derived hPGCs demonstrate migratory potential in a chick embryonic environment*

4.1 *Two roles of BMP4 in hPGC specification*

4.2 *hPGC fate acquisition requires BMP4 during the first 24 hours of differentiation*

4.3 *hPGC specification is not solely dependent on BMP4 gradients*

4.4 *Secreted WNT and ACTIVIN/NODAL inhibitors pre-pattern an hPGC competent niche*

4.5 *BLIMP1 is sufficient to induce hPGC markers, and itself induces SOX17 in a positive feedback loop*

5.1 *Serial scans reveal early embryonic development with protracted gastrulation*

5.2 *Standard Embryonic Measures as predictors of fertilization age*

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 PGC specification across evolution

The specification of germ cells during embryonic development is vital not only for the

development of an organism, but quite literally for the survival and propagation of its entire

species. From the time of Weismann and Darwin, conflicting mechanisms of germ cell

development across different model organisms have resulted in conflicting theories of how

genetic information is transmitted from one generation to the next. As the ideas of "pan-

genesis" and "blasto-genesis" gave rise to the theories of germ epigenesis and germ evolution

– it has been difficult to reconcile the paramount importance of germ cell function with the

apparent validity of both of these models (*Weismann, 1893*). The incongruity between these

models was as apparent as the divide between proponents of the electric and chemical models

of neural synaptic signaling at the same time (resolved by *Hodgkin and Huxley, 1952*). Over 100

years of study has resolved that, as with the case of neural transmission, both models are true.

The specification of primordial germ cells (PGCs), the progenitors of germ cells, during

embryonic development follows two independent paths of convergent evolution. In most

invertebrates, such as *C. elegans* and *Drosophila*, and some vertebrate species such as *Xenopus*

laevis and *Danio rerio*, PGC specification follows a deterministic path. In this model, a

maternally derived "germ plasm" is segregated and concentrated through several asymmetric

divisions into the eventual PGC population (*see Strome and Updike, 2015*). Though not all

components of germ plasm are conserved, the general function of the germ plasm seems to be

RNA binding capacity, regulating transcription and activating pluripotency (*Oscar and VASA*) while indirectly inhibiting differentiation (*Nanos*) (*Marlow, 2015*). In this fashion, rapidly growing organisms specify germ cells earlier and somatic cells can develop and evolve quickly (*Evans et al., 2014; Johnson and Alberio, 2015*).

In contrast, mammals (including mice, non-human primates, and humans) undergo an inductive model of specification, in which PGCs are segregated from a somatic cell population that is selectively induced by extracellular signaling factors. Homologs of various "germ plasm" components are also found in this model, and are activated in a cascade following morphogen signaling via Bmp4, Wnt3, and Bmp8b – in the mouse (*Lawson et al., 1999; Saitou et al., 2002; Magnusdottir and Surani 2014; Gunesdogan et al., 2014*). Wnt acts as "priming" agent, while Bmp4 has proven to be the initiating event of mouse PGC specification, inducting a subpopulation of Brachyury+ (Bra+) mesoderm at the posterior primitive streak to express Blimp1, Prdm14, and Ap2γ from E6.5-E7.5 (*Ohinata et al., 2005; Yamaji et al., 2008; Saitou et al., 2009*). This triad has proven necessary and sufficient for mouse PGC specification and determine the presence of other markers of pluripotency. At this point, mouse PGCs begin to migrate along the hindgut to the gonadal ridge in a process partially guided by Sdf-1 – Cxcr4 interactions, by which time they express mouse Vasa homolog (Mvh) and Dazl, two highly conserved markers of PGC fate, by the end of migration at E13.5 (*Nikolic et al., 2016, Saitou et al., 2009, Molyneaux et al., 2003*).

PGC specification in birds and reptiles, while inductive in many species, bridges the gap between the preformation model and the induction processes studied in mammals. In the chicken, for example, a germplasm-like structure containing Vasa protein is found in the basal cleavage furrow following the first cleavage, after which Vasa is found in few cells at the center of the blastocyst (*Nakamura et al., 2013*). These cells move anteriorly towards the extra-embryonic tissue to form the germinal crescent at the anterior end of the epiblast during gastrulation, after which they take the curious step of migrating laterally into the developing vasculature and move through the blood stream into the intermediate mesoderm and developing gonadal ridge. Chicken embryonic germ cells can be maintained in culture similarly to mouse cells in the presence of Scf/Lif/bFgf and maintain markers of pluripotency (*Nakano et al., 2011*).

Interestingly, it seems that the ancestral mode of PGC specification across vertebrates is actually the inductive model, with evolutionary genetic analysis indicating that the preformation model (as found in the frog, chick and zebrafish) evolved convergently in species in order to speed the rate of development, reproduction, and speciation (*Evans et al., 2014; Johnson and Alberio, 2015*). This idea may extend beyond the realm of vertebrates as well, as the cricket (*Gryllus bimaculatus*) has also been shown to follow an induction model through Bmp4 signaling in a Bra+ mesendodermal intermediate (*Nakamura et al., 2016*).

In humans and non-human primates (NHPs), PGC induction seems to work most similarly to the mouse. Despite this overarching similarity, significant differences in embryo geometry, organization of extraembryonic tissues, and timing of development necessitate just as rigorous an investigation within primate systems. Human PGCs (hPGCs) were identified in the posterior yolk sac of human embryos of the Carnegie collection nearly 100 years ago (*Witschi et al., 1948*). PGC markers have been found in the posterior epiblast and amnion in peri-gastrulation in non-human primates (*Sasaki et al., 2016; Nakamura et al., 2016*) and tracked during migration along the hindgut endoderm into the intermediate mesoderm and gonadal ridge (*Aeckerle et al., 2014*). Recent studies in human embryos have elucidated migratory hPGCs along the intermediate mesoderm and in the gonadal ridge by immunostaining and at the transcriptome, and epigenetic level, providing a molecular gold standard by which to compare *in vitro*-derived cells (*Mamsen et al., 2012; Gkountela et al., 2013; Gkountela et al., 2015; Li et al., 2017; Fernandez et al., 2018*). Importantly, functional sufficiency of both human and NHP natural PGCs has been demonstrated by xenotransplantation as early as Carnegie stage 23 (*Clark et al., 2017*).

1.2 PGC Specification in mammalian embryonic stem cells

Several studies over the past years have established that these stages of mouse PGC development can be reliably recapitulated through Bmp4- facilitated differentiation of mouse embryonic stem cells (mESCs) *in vitro* (*Ohinata et al., 2009; Vincent et al., 2011)*. These mESC derived PGCs have been functionally validated by *in vitro* fertilization and have resulted in live births (*Nayernia et al., 2006; Hayashi et al., 2012*). Though human PGCs follow an inductive

model of development similar to the mouse, recent work has shown crucial differences between the two species *in vivo*, calling into question the relevance of the mouse model in understanding human PGC specification. In recognition of the successes of mouse ESC models of PGC development, efforts to generate germ cells from human embryonic stem cells (hESCs) have been made in earnest. PGC-like cells can be derived from human ESCs at very low levels during simple undirected embryoid body differentiation or via BMP4 addition (*Kee et al., 2007; Julaton and Reijo Pera, 2011*). Following the successes of the mouse *in vitro* PGC specification field in the early 2010's, there has been a great amount of effort towards defining hPGC specification in an analogous manner. Recent work has demonstrated higher efficiency production of hPGCs by reverting cells to a naive state and described a new role for SOX17 as the initiator of human PGC specification in lieu of BRACHYURY (BRA) (*Tang et al., 2015; Irie et al., 2015)*. Other groups have managed efficient hPGC specification by induction through a transient mesendoderm induction (*Sasaki et al., 2015*). Notably, most of these protocols include additional signaling (e.g. SCF, LIF, EGF, ROCK inhibition) for stable specification of hPGCs. It remains an open question whether BMP4 alone is sufficient to generate hPGCs from hESCs, what role other signaling factors in the niche play in specification, and which PGC markers are truly useful for distinguishing primate PGC populations.

1.3 Validating *in vitro*-derived human PGCs

Though PGCs have been examined in human embryos for over 100 years, many of the molecular markers *in vivo* are not necessarily specific enough for use in *in vitro* derivation protocols. Canonically, hPGCs have been tracked using tissue nonspecific alkaline phosphatase

(TNAP), as well as OCT4, NANOG and SSEA-4, which are all known to be expressed in pluripotent hESCs (*De Felici, 2013*). Thus, they are not useful markers for the study of early hPGC specification from hESCs, and co-expression of many markers must be used to validate the identity of these cells by histological methods; a difficult task in precious human samples, especially as these cells undergo continuous morphologic and physiologic shifts throughout development.

As fertilization, the ultimate gold standard of germ cell fate, is a standard difficult to achieve within the scope of current guidelines for hPGCs, several other benchmarks have been proposed to ensure that *in vitro*-derived PGCs do faithfully model *in vivo* development[1]. Recently, these include a comparison with natural human, monkey (cynomolgus), and porcine PGCs; transcriptomic changes and epigenetic reprogramming signatures have been the most convincing arguments of parity with *in vivo* development (*Handel et al., 2014; Gkountela et al., 2013; Kobayashi et al., 2017*). Further developmental potential through the acquisition of a gonadal germ cell signature has also been investigated.

In a great boost to *in vitro* hPGC studies, a long-term culture study was recently published in which *in vitro* specified hPGCs cultured with somatic cells of the mouse embryonic ovary exhibited epigenetic reprogramming and molecular signatures characteristic of pre-meiotic

---

[1] This is also why many propose the term PGC-like cells when describing *in vitro*-derived human PGCs. As the term "PGC" *by definition* refers to a cell that is committed to developing into a germ cell, as long as this has been demonstrated, the distinction is one of semantics. Whether an *in vitro*-derived germ cell is capable of fertilization or not is a separate question, and would imply the use of the term "germ cell-like cells" for later stages of development instead. For clarity's sake, the term "PGC" will be used for both *in vivo* and *in vitro*-derived PGCs throughout this document, as a large and growing body of evidence has demonstrated the ability of *in vitro*-derived PGCs to develop along the same molecular, epigenetic, and behavioral axes as *in vivo* PGCs.

germ cells (*Yamashiro et al., 2018*). Other groups have demonstrated relevance of *in vitro*-derived hPGCs by pushing them to gonadal follicle-like cells, either through the exogenous overexpression of early PGC markers PRDM1, PRDM14, LIN28A, together with late markers DAZL, VASA and SYCP3 (*Medrano et al., 2016*), or through overexpression of DAZL and BOULE alone (*Jung et al., 2017*). These cells display morphological characteristics of meiotic germ cells, synaptonemal complex formation, and evidence of epigenetic reprogramming.

We propose that in addition to these measures, it is instructive in this burgeoning field to be able to evaluate the functionality of *in vitro*-derived hPGCs by other means. One of the key functions of PGCs is the ability to follow chemoattractant cues to the gonadal ridge – indeed, migratory phenotypes of *in vitro*-derived hPGCs in response to conserved SDF-1 – CXCR4 signaling has recently been demonstrated (*Mitsunaga et al., 2017*). By replicating the environments in which hPGCs specify, migrate, and mature, we can not only begin to understand the interactions between hPGCs and the various cell types making up the niche with greater clarity, but also provide a functional definition of hPGCs that, while not quite the standard of fertilization, will provide another axis along which to evaluate these cells. Clearly, the prospect of fully *in vitro*-derived human germ cells is within reach – a major milestone for the study and treatment of infertility.

1.4 Modeling the gastrulating embryo: investigating hPGC development within a complex signaling niche

hPGCs have a suggested origin in the posterior epiblast and/or amnion at weeks 2-3 of human development, after which they must migrate to the site of the developing gonad. However, due to the limited accessibility of early human embryos and species-specific differences across model organisms, the molecular and cellular basis of hPGC specification and migration remains obscure. We have shown previously that in response to BMP4, hESC colonies grown on micropatterned substrates self-organize to form "gastruloids" that harbor radially symmetrical patterns of all embryonic germ layers, as well as extra-embryonic fates (*Warmflash et al., 2014; Etoc et al., 2016*).

We have previously shown that the cascade of signaling involved in self-organization of human gastruloids begins, similarly to in the actual human embryonic disc, with an exogenous BMP4 signal, which in the embryo is provide by extra-embryonic tissues and in the gastruloid is provided in the media (*Warmflash et al., 2014*). In both systems, this BMP4 signal activates WNT signaling, thus inducing the NODAL/ACTIVIN pathway, which in turn can feed back to induce or maintain endogenous BMP4 signaling (**Figure 1.1**, *Arnold and Robertson, 2009*). One major benefit of micropatterned hESC differentiation culture is the ability to probe the effects of morphogen gradients on cell fate acquisition. Due to the greater number of neighbors a cell at the center of colony has when compared to one at the edge, cells in the center of a colony find their response to exogenous stimuli blunted by both endogenous secreted factors as well

as differences in receptor localization (*Etoc et al., 2016*). The reaction-diffusion gradient

generating mechanism also describes micropatterned responses to the WNT3A/DKK1 and

ACTIVIN/CER1 ligand-inhibitor pairs (*Martyn et al., 2019*).
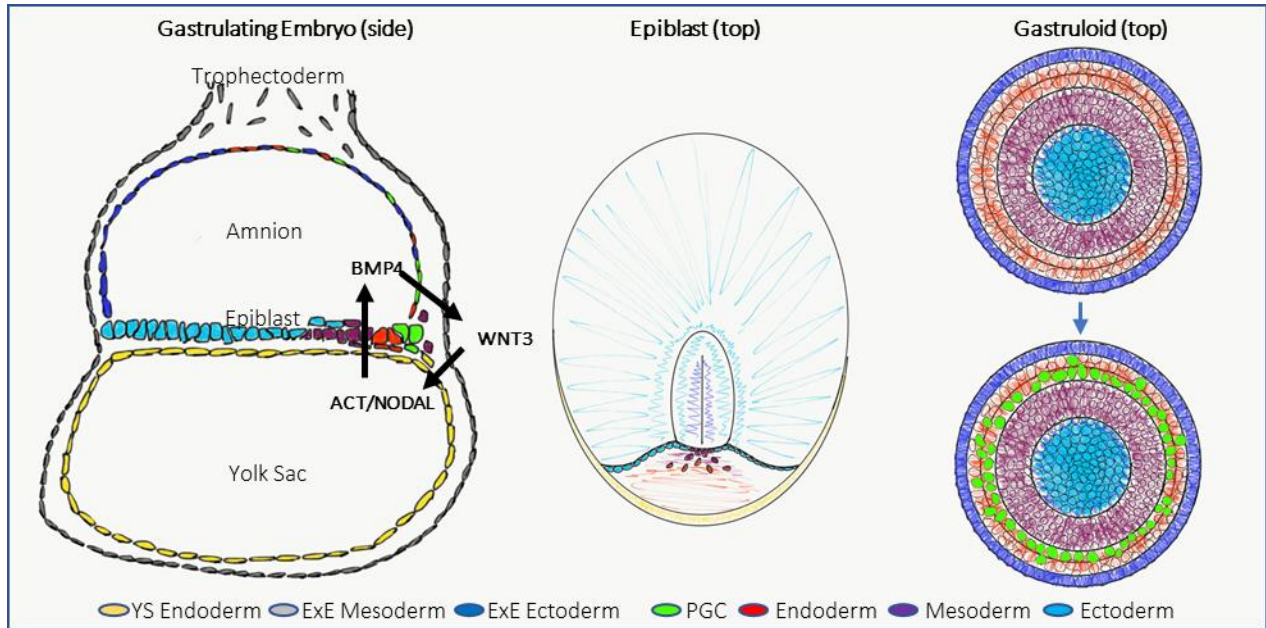
Figure 1.1. *The gastruloid culture system as a model for the gastrulating human embryo.*

Illustrated cross-section of the human embryo (left) and top-down view of the gastrulation

epiblast (center) provide a basis for in vivo correlation with spatial self-organization of the

gastruloid due to a BMP4-WNT3-ACTIVIN/NODAL cascade in tissues at the posterior end of the

epiblast. In the gastruloid system, a hydrophobic surface is treated with laminin to generate a

micropatterned hESC culture surface. When stimulated with exogenous BMP4, hESCs cultured in

this fashion undergo regularized fate changes due to the resultant BMP4 gradient, as well as

downstream WNT and ACTIVIN signaling (Warmflash et al., 2014; Etoc et al., 2016). The radial

patterning of fate change induces SOX2+ ectoderm in the center of colonies, with BRA+

mesoderm, SOX17+ endoderm, and GATA3+/CDX2+ extraembryonic fates sequentially more

distal (right). We show in Chapter 3 that the SOX17+ ring includes both distal definitive

endoderm and medial PGCs.

As indicated in **Section 1.1**, years of genetic work in the mouse has established a role for each

of these signaling cascades in PGC development. Evidence for a Bmp4 dose dependence exists,

as Bmp4-/- mice fail to establish PGCs, while Bmp4+/- mice have fewer PGCs that migrate

inefficiently (*Tremblay et al., 2001*). A permissive role for Wnt signaling in mouse PGCs was also

established (*Ohinata et al., 2009*). More recently, Nodal was found to be necessary, as Nodal-/-

mice were found to contain cells expressing Blimp1, but not co-expressing Blimp1 and AP2y

(*Senft et al., 2019*). The presence of these signaling ligands during PGC specification was also

validated by *in situ* hybridization in cynomolgus (*Macaca fascicularis)* NHP embryos (*Sasaki et*

al., 2016). Interestingly, overexpression of Nanog alone has been shown to be sufficient for *in*

vitro mouse PGC induction, independent from Bmp4, Wnt, and Nodal signaling (*Murakami et*

al., 2016), raising the prospect that multiple independent mechanisms may act to specify PGCs.

1.5 *In vivo* primate embryogenesis: towards a synthesis of *in vitro* and *in vivo* tools

Few modern (post-Carnegie collection) studies have had the good fortune of access to primary

human PGCs for study of specification, migration and maturation. These have been invaluable

not only because of their insight into hPGC biology in their own right, but in their capacity to

recast studies from model organisms and allow close evaluation of where human biology

diverges. From these, it is clear that an *in vivo* system for the interrogation of uniquely primate

biology is necessary. Recently, PGC specification has been examined in cynomolgus embryos

(*Sasaki et al., 2016*), and the combination of precise *in vitro* stem cell modeling and validation in

non-human primate embryos has borne much fruit, such as suggesting the origin of PGCs in the amnion of primate as opposed to the posterior yolk sac/allantois as inferred from Carnegie collection examination. These studies suggest the peculiarities of human PGC development may be shared across the primate order.

The common marmoset (*Callithrix jacchus*) provides an ideal model for investigating unique primate biology, particularly cognitive function and reproduction (*Li et al., 2005*). At a systems level, they provide a unique platform for neuroscience and behavioral studies shedding light unto the mechanism of cognition (*Miller et al., 2017; Miller et al., 2016*). From a developmental biology perspective, they provide a very useful system to study early primate embryogenesis, *in vivo*, a prospect that is ethically impossible to contemplate in humans, even if it were to be technically feasible *(Chambers & Hearn, 1985; Hearn et al., 1978; Kropp et al., 2017; National Academies of Sciences, Anestidou & Johnson, 2019*). Recent advances at the tissue and cellular level, including self-organizing stem cells into embryoids and organoids, combined with the advanced molecular tools of genome editing and transgenic technologies will illuminate the primate-specific aspects of embryonic development. With a gestation period of 140-150 days most often producing dizygotic twins, it possesses numerous benefits for developmental studies and the creation of transgenic primates (*Park & Silva, 2019; Park et al., 2016; Sasaki et al., 2015; Schiel & Souto, 2017; Tomioka et al., 2017*). There is an essential need to establish an accessible and scalable means by which to monitor embryonic and fetal development with demonstrated correlates to *ex vivo* references.

Previous studies have demonstrated key milestones and growth curves utilizing ultrasound (*Jaquish et al., 1995; Tardifet al., 1998*). They formed the foundational basis by which marmoset development could be tracked utilizing ultrasound, but they lacked the temporal resolution to capture all stages of development and thereby are limited in scope for examining particularly early developmental signatures. *Ex vivo* histological studies have created the standard by which marmoset embryos can be staged into Carnegie Stages, but lack a critical link to *in vivo* monitoring of the embryos at the same stages (*Enders & Lopata, 1999; Moore et al., 1985; Phillips, 1976; Smith et al., 1987*). These *ex vivo* studies revealed, however, a remarkable finding that marmoset embryos enter into a period of delayed development just after implantation and prior to gastrulation. Although, it is debated whether or not this constitutes a form of primate diapause, a state of either facultative or obligate embryonic suspension at the blastocyst stage (*Deng et al., 2018; Phillips, 1976; Renfree & Fenelon, 2017*). Marmosets do not show evidence of instantiating any delay in their pregnancies based on season, resource availability, or hormonal state *(Phillips, et al., 1976)*, but nevertheless, the embryos exhibit marked delay in progressing through the stages of embryonic development compared to all other mammals for which this has been measured (*Hill et al., 2020*). Still, studies to date have been unable to temporally resolve the earliest stages of embryonic development.

Specific organogenesis has been tracked grossly by *ex vivo* studies that measured the presence and weight of organs across development, but this offers little structural insight (*Chambers & Hearn, 1985)*. Other studies have focused on development of individual organs, such as the brain, and have been able to detail development both *in vivo* and *ex vivo* utilizing magnetic

resonance imaging (MRI) (*Hikishima et al., 2013; Sawada et al., 2014*). This imaging method

requires a significant investment in custom and costly equipment thereby making it infeasible

to scale across marmoset colonies as a means by which each individual fetus can be screened

and tracked for key developmental milestones. Further, although the spatial resolution in these

MRI studies proved to be exquisite, temporal resolution and movement artifact limits the

measure of functions such as the heart beat or blood flow which provide key markers of cardiac

development. There exists an opportunity to better match standard human measures for key

organ development including kidney, heart, and brain *(Anderson et al., 2003; Moorman et al.,*

2003; Sousa et al., 2017; van Vuuren et al., 2012)

1.6 Concluding remarks

In my thesis, I present three main approaches to modeling the intricacies of the earliest stages

of human reproductive development. They represent a synthesis of *in vitro* and *in vivo*

techniques, fusing quantitative techniques in stem cell culture, transcriptomic analysis, genetic

and cell imaging techniques with *in vivo* embryonic grafts, clinical ultrasound imaging, and

machine learning in order to help bridge the gap between basic science and clinical

presentation.

In the first, we examine hPGC specification in a stem cell model of a human epiblast. We

determine the hPGC niche and characterize specified hPGCs, finding a molecular signature that

indicates migratory potential via canonical and novel mechanisms, as well as raising several potential candidates for further investigation into the under-studied migratory phase of germ cell development. We seek to generate tests of hPGC function in an embryonic context by investigating migratory potential in grafts to the chick embryo.

In the second, we harness the power of precise geometrical confinement and genetic tools including CRISPR-Cas9 to probe further into the signaling environment of the hPGC niche in the gastruloid model. We elucidate the role of each leg of the BMP4 – WNT – ACTIVIN/NODAL signaling cascade in hPGC specification. These experiments not only demonstrate how BMP4 acts independently in addition to this cascade to directly specify hPGCs, but also how all three pathways work in harmony to generate self-organization of hPGCs within their gastruloid niche. By probing further into hPGC specification, we find that BLIMP1 alone, a canonical regulator of PGCs, is sufficient to induce later markers of hPGC fate, and surprisingly is also sufficient to downregulate SOX2 and upregulate SOX17.

In the third, we propose the marmoset as a good model of primate reproduction and embryogenesis. We systematically investigate marmoset embryonic development *in vivo* using serial, high temporal and spatial resolution ultrasound imaging. We elucidate morphological characteristics of implantation, gastrulation, neurulation, and organogenesis, as well as the curious marmoset phenotype of "diapause" which we suggest is an elongated peri-gastrulation window, incorporating slowed embryonic growth and rapid extraembryonic development. In

14

doing so, we generate an ultrasound atlas of embryogenesis, and train models to identify developmental stages and predict fertilization age from a single frame.

To complete our vision of a fully translational system, synthesizing human gastruloid development and marmoset embryonic development, we describe ongoing efforts towards controlled maturation of hPGCs as well as generation of induced pluripotent stem cell (iPSC) lines from marmosets in our colony. These lines possess the capacity to generate marmoset gastruloids upon micropatterned culture and stimulation with BMP4, providing us the invaluable opportunity to translate our findings in the hESC gastruloid system into the parallel marmoset system. This body of work provides a platform of molecular techniques, genetic tools, and imaging systems necessary to continue to explore early primate development.

CHAPTER 2

Materials and Methods

Cell Culture: hESCs were grown in HUES medium conditioned by mouse embryonic fibroblasts (CM) and supplemented with 20 ng/ml bFGF. Mycoplasma testing was carried out before beginning experiments and again at two-monthly intervals. For maintenance, cells were grown on GelTrex (1:40 dilution, Invitrogen)-coated tissue culture dishes (BD Biosciences). The dishes were coated overnight at 4 °C and then incubated at 37 °C for at least 20 minutes before the cells were seeded. Cells were passaged using Gentle Cell Dissociation Reagent (Stem Cell Technologies, 07174) and reseeded in CM with 10 µM ROCK inhibitor Y-27632 (Abcam). WIS24 wildtype hESCs and WIS-NANOS3-mCherry are published (*Irie et al., 2015*) and were a gift from Dr. Azim Surani.

Micropatterned cell culture: Micropatterned culture chips were made using coated coverslips from CYTOO, Inc (Cambridge, MA). Chips were coated with 5 µg/ml human laminin 521 (Biolamina) diluted in PBS with calcium and magnesium (PBS++) for 4 hours at 37°C. After two washes with PBS++, cells were dissociated from growth plates with StemPro Accutase (Life Technologies) for 5 minutes. Cells were washed with PBS and re-suspended in CM with 10 µM ROCK inhibitor Y-27632 (Abcam). Micropatterned chips were placed in 35-mm plastic tissue culture dishes, and 1×10^6 cells in 1.5 ml of medium were seeded for each chip. After 1 hour the ROCK inhibitor was removed and replaced with CM. Cells were treated with 50 ng/ml BMP4 12 hour after seeding.

Immunocytochemistry: Cells were fixed in 4% paraformaldehyde for 30 minutes at room temperature, washed twice with PBS, then blocked and permeabilized in a blocking buffer containing 2% donkey serum and 0.2% Triton X-100 in PBS for 30 minutes. Cells were incubated overnight with primary antibodies (Supplemental Table 1, 1:200 unless otherwise noted) in blocking buffer at 4 °C. The next day they were washed 3times with blocking buffer and incubated with secondary donkey antibodies (Alexa Fluor 488, Alexa Fluor 555, AlexaFluor594, Alexa Fluor 647, 1:500 dilution) and DAPI for 30 minutes. Cells were then washed 3times with PBS and mounted on glass slides for imaging using ProLong Antifade (Invitrogen). For experiments requiring staining with two sets of primary antibodies of the same species, cells were incubated with conjugated secondary Fab fragments after the first set of primary antibodies, and incubated for 2hours at room temperature with excess host species Fc fragments (1:100 dilution) prior to the second set of primary antibodies.

Cell Microscopy and Image Analysis: Images were acquired with either a Zeiss Axio Observer and a 20× 0.8-NA lens, or with a Leica SP8 inverted confocal microscope with a 40× 1.1-NA water immersion objective. Image analysis and stitching was performed with ImageJ (FIJI) using the Trainable WEKA segmentation plugin. Single cell analysis, including thresholding via Otsu's method and K-means clustering, was performed using custom Python scripts. Live OCT4-eGFP and CellTracker Videos generated for Supplemental Video 1 and 2 were acquired using the CellVoyager CV1000 spinning disk confocal system (Yokagawa/Olympus) with a 20x dry objective, deconvolved using Autoquant software, and analyzed in Imaris.

Single Cell RNA Sequencing and Analysis: Micropatterned 500 µm colonies were stimulated

with BMP4 for 4 days and collected as single cells using Accutase treatment as above. Cells

were passed through 20µm mesh filters to isolate single cells for analysis, and sequenced at The

Rockefeller University Genomics Resource Center using the 10X Genomics Chromium System.

UMI count matrices for RUES2 and 56CAG were loaded into R (v.3.5.2) and analyzed using

Seurat [20] (v.3.1) and SCANPY (v.1.4.4, *Haghverdi et al., 2016; Wolf et al., 2018*). The raw

matrices were filtered to have a minimum of 200 detected genes per cell and a gene was only

kept if expressed in at least 5% of cells. Cells with over 10% mitochondrial UMIs were discarded.

The top 2,000 variable genes for each dataset were found by the SCTransform method

implemented in Seurat. Principal component analysis (PCA) was performed after scaling the

data and the first 30 principal components stored. To unbiasedly find clusters we employed the

Louvain algorithm with multilevel refinement with a resolution parameter of 0.3 and using the

first 30 principle components.

Clustering, genetic background and gene expression (normalized UMI counts) were visualized in

a low-dimensional space using UMAP plots with parameters min.dist = 0.4, n.neighbors = 20

and using 20 PCA dimensions. Marker genes for each cluster were found using differential gene

expression analysis using Seurat's implementation of logistic regression testing, as well as

power ranking using receiver operating characteristic curve analysis. Genes were filtered using

three criteria: adjusted P value <0.05, >25% cells expressing the gene in the cluster of interest,

and average log fold change >0.1. To determine characteristic genes with high clustering power

in the hPGC cluster, more stringent criteria were used as follows: genes were filtered for an

AUC value>0.7, >50% of cells expressing the gene in the hPGC cluster, and <30% of cells outside

of the hPGC cluster expressing the gene. To directly compare between gastruloid and primary hPGC datasets, gastruloid data was integrated with primary human gonadal data[6] into a single Seurat object by first finding anchors in a shared low-dimensional space using canonical correlation analysis before PCA and clustering as above. For pseudotime analysis and trajectory inference within the hPGC cluster, gastruloid data was analyzed using the SCANPY library in Python using the Spyder3 environment (*Wolf et al., 2018*).

Partition-based graph abstraction was used to generate a topology-preserving map of the gastruloid dataset (*Wolf et al., 2019*). A root node was chosen via differential expression analysis of canonical pluripotency markers as above, and the diffusion pseudotime method was used for trajectory determination [31]. Single cell mouse gastrulation data for comparison with hPGCs was generously provided by the laboratory of John Marioni (https://github.com/MarioniLab/EmbryoTimecourse2018).

Transplantation of hESC-derived hPGC colonies into the developing chick embryo: Fertilized White Leghorn chicken eggs from Charles River Laboratories were incubated at 37-38°C and 50% humidity until reaching the desired stage of development (*Wodak et al., 2007*). Eggs were removed from the incubator and cooled for 30 minutes, then embryos were collected and set up in Early-Chick culture (*Chapman et al., 2001*) to be used as a host for transplantation. hESC (WIS24-NANOS3-mCHERRY) were grown on Arena A Cytoo coverslips for 4 days as described above, and 225μm diameter micropatterned colonies were selected for transplantation. 225μm diameter colonies were peeled off from the coverslip using tungsten needles (Fine Science Tools) and washed three times with Pannett-Compton solution, to remove culture growth

factors and ligands. Next, colonies were grafted along the anterior-posterior axis of the HH4 chick embryo. The grafted embryos were re-incubated for 24 hours to allow further development and were then fixed in 4% paraformaldehyde in PBS.

hESC-derived PGCs time-lapse: After the transplantation, hESC-derived PGCs were tracked using time-lapse video microscopy. The human cells were labeled with green CMFDA (5-chloromethylfluorescein diacetate) dye (Thermo Fisher Scientific) and were imaged using a spinning disk confocal microscope (CellVoyager CV1000, Yokogawa). Fluorescent images were acquired every 20 minutes for 24 hours. Next, the embryos were fixed in 4% paraformaldehyde in PBS and submitted to immunofluorescence.

Chick embryo immunofluorescence: Embryos were fixed in 4% paraformaldehyde in PBS for 1 hour at room temperature or overnight at 4°C. They were then washed 3 times with PBST (PBS+0.5% Triton X-100) for 1 hour each on nutator and blocked and permeabilized with 3% donkey serum, 1% bovine serum albumin in PBST for 2 hours, also at room temperature. Next, embryos were incubated overnight with primary antibodies diluted in blocking buffer at 4°C. In the following day, embryos were washed 3 times with PBST for 1 hour each on a nutator and then incubated overnight with secondary donkey antibodies, anti-human nuclear antigen conjugated with Alexa Fluor 647 (Novus Biologicals NBP2-34525AF647), and DAPI. When using antibodies raised from same species host, a sequential staining of each primary antibody was performed. To eliminate the non-specific binding between Fc portions of antibodies, Fab fragment secondary antibodies were used. Alternatively, after the first secondary antibody incubation, unconjugated Fab fragments (Jackson ImmunoResearch Laboratories, 711-007-003,

715-007-003) against the host species of the primary antibody were used to block the binding. Embryos were washed 3times with PBST for 1 hour each and mounted in glass slides with fluoromount to image. Primary antibodies used are listed in Supplemental Table 1.

Injection of hPGCs within the chick embryo vasculature: hESC micropatterned colonies stimulated for 4 days with BMP4 were dissociated and resuspended in Howard Ringer's salt solution before being transferred to the chick embryos. A small window was opened in the eggshell to access the embryo and approximately 1-2 µl containing approximately 5000 cells were injected into the dorsal aorta of HH16-17 embryos (2.5-3 days of development). Control embryos were injected with saline solution. After the procedure, embryos were hydrated with 1 ml of Howard Ringer's salt solution, re-sealed with tape and re-incubated to allow them for further growth. They were collected 2-3 days post-transplantation (HH26-28), when the gonads were already established, fixed in PBS/paraformaldehyde 4%. Embryos were divided into 2 groups: in one group the gonads were dissected, stained and imaged using a confocal microscope; the other group of embryos were cleared using the iDISCO method (*Renier et al., 2014*).

Chick Microscopy and Image Analysis: Chick embryos images were acquired using an inverted Zeiss Axio Observer Z1 microscope with a 20x/0.8 numerical aperture (NA) lens. Image analysis and stitching were performed with Zeiss Zen microscope software and Imaris. Images used in Supplemental Video 3 were deconvolved with Autoquant software and analysed in Imaris. hPGCs were visualized in the gonads of chimeric embryos through 3D imaging using the light sheet microscope after iDISCO clearing (LaVision BioTec). Surviving hESC-PGCs were identified

by staining with human nuclear antigen (HNA) specific antibody. 3D reconstruction was done using Imaris software.

Generation and Validation of Stem Cell Lines: OCT4-eGFP cells were generated by CRISPR-mediated homologous recombination from the parental RUES2 line. OCT4-eGFP-2A-Puro was a gift from Rudolf Jaenisch (Addgene plasmid # 31939). Human Q002 iPSCs were derived from a fibroblast sample donated anonymously using the non-integrating ReproRNA-OKSGM reprogramming vector (StemCell Technologies) in a feeder-free setup as per manufacturer's instructions. Cells were karyotyped and maintained for 20 passages after stemness was confirmed by expression of TRA-1-61, OCT4, and SOX2 and pluripotency was confirmed with 48 hour micropatterned gastruloids generating cells of all three germ layers as well as GATA3+ extraembryonic cells. All clones were selected and confirmed by sequencing and karyotyping.

Generation of co-culture systems. Granulosa-like cells were generated by a modified protocol adapted from *Lan et al., 2013*. hESCs were nucleofected with an ePiggyBac (ePB) transposon containing a dox-inducible FOXL2 overexpression cassette. 48 hours of induction with BMP4, ACTIVIN and CHIR (a WNT agonist) upon micropatterned cells caused an upregulation of mesodermal markers, after which doxicyclin induction of FOXL2 expression together with BMP4 and FST (follistatin) generate cells expressing some markers of granulosa, such as AMH receptor.

Supplemental Table 1. Primary Antibodies used in the study

Antigen	Antibody
SOX2	Cell Signaling (2579)
BRACHYURY	Santa Cruz (sc-17745)
GATA3	Invitrogen (MA1-028)
BLIMP1	Cell Signaling (9115S)
SOX17	R&D (AF-1924)
AP2-GAMMA	Santa Cruz (sc8977)
OCT4	Santa Cruz (sc-8628)
CDX2	Biogenex (MU392-UC)
V5 tag	Invitrogen (MA5-15253)
NANOG	BD Biosciences (560482)
FOXA2	Santa Cruz (sc-6554)
FOXF1	Abcam (ab168383)
DAZL	Abcam (ab34139)
Human Nuclear Antigen	Novus (NBP2-34525AF647) 1:300
Human Nuclear Antigen	Novus (NBP2-34342) 1:300

Ultrasound Procedures: We performed ultrasound measurements on 7 dams across 13 pregnancies totaling 34 fetuses. Each marmoset was kept in stable, bonded pairs on luetolytic cloprostenol (0.03 mL, 1x/28 days) until entering breeding. Upon commencing breeding, each marmoset was palpated weekly or by ultrasound exam if pregnancy was suspected. Once an open lumen in the uterus was observed, the dam underwent biweekly ultrasound examinations using the Vevo 2100 (FUJIFILM VisualSonics, Ontario, Canada) ultrasound. Probes used included MS700, MS550D, and MS250. For the last half of pregnancy, ultrasounds were conducted on a weekly basis. Ultrasounds were conducted without sedation. Animals were behaviorally trained with positive reinforcement to remain still while being lightly held in a recumbent position.

Ultrasound data collection and analysis: An independent scorer who did not partake in the ultrasound examination took key measurements of crown-rump length and parietal distance for each fetus. Notes were made on stages of organogenesis. Across all animals in the study, there were 104 imaging days consisting of 289 individual ultrasound exams with a total of 6159 recordings.

All recordings were exported as DICOM and converted into TIFF stacks using custom MATLAB code (MathWorks, Massachusetts, USA). Measurements were made in ImageJ (NIH, Maryland, USA) and notes were taken in Excel (Microsoft, Washington, USA). Data were collected and analyzed in Python (Python Software Foundation, Delaware, USA) and comparison charts based on notes were made in Excel. To calculate day post fertilization, we took each animal's date of birth and assumed a gestation period of 143 days (*Jaquish et al., 1995*). The day of each of ultrasound exam was then converted into a presumed date from fertilization. We plotted all data both in time from birth and presumed day post fertilization. For ease of comparison to other studies, all figures here are plotted only in day post fertilization.

Predictive modeling: Image classifier and regression models were built using the Keras deep learning API platform built upon TensorFLow 2.0 in Python (*Challot et al., 2015*). Resnet50 architecture was used to build all models. For binary image classification, a ResNet50 model was fitted with a binary crossentropy classifier and trained on a set of ~200 salient and non-salient frames that were sorted by hand. Following this, we trained deep convolutional dense

nets, adapted for usage with ultrasound frames and written in Keras for regression of input

frames along the axis of embryonic age, utilizing the time until birth as ground truth. Learning

rate and batch size parameters were optimized. We trained on a set of 1200 salient frames,

before generalizing the regression model by training across all 800,000 frames from the

ultrasound atlas dataset. The Adam optimization scheme was utilized, with learning rates

between 0.001 and 0.00001, with decreasing learning rates used after every 25 epochs. We

found that performance increased when two networks were used – one for regression during

the first half of gestation and another for the second. We found that the trained network

performed as well on random unseen frames as well as standard clinical measurements on

idealized frames.

Tissue Histology: The gross dissection of the ex vivo kidneys were conducted on kidneys taken

from a pup that did not survive the birthing process and from an adult that was humanely

euthanized for a related project.

Animal Care and Use: All animals were in normal health and were group housed with other

conspecifics at the facility. All procedures were approved by The Rockefeller University Animal

Care and Use Committee and were performed in accordance with the National Institute of

Health guidelines for care and use of laboratory animals.

CHAPTER 3

Specification and Self-organization of Migratory Primordial Germ Cells in Human Gastruloids

The molecular basis of hPGC specification and migration remains obscure, due to the limited accessibility of early human embryos and species-specific differences across model organisms. We have shown previously micropatterned gastruloids are a powerful tool for the study of self-organized cell fate acquisition due to endogenous signaling, downstream of an exogenous BMP4 signal (*Warmflash et al., 2014; Etoc et al., 2016; Morgani et al., 2018*). Here, we first demonstrate the specification and self-organization of migratory hPGCs within gastruloids, demonstrating the utility of gastruloid culture for studying hPGC development relative to other cells populations in the early embryo. Using scRNA-seq we capture the gene expression signature of hPGCs as they acquire migratory potential. Our analysis revealed 15 genes not previously associated with PGC development, including DOCK2 and its migratory machinery previously only associated with leukocyte migration. This signature was also present in a recently published dataset from human fetal PGCs (*Li et al., 2017*), suggesting a novel mechanism of migration for hPGCs. Finally, we show that when transplanted into chick embryos, hPGCs display migratory behaviors, providing an *in vivo* validation that gastruloid-derived hPGCs can follow functional and behavioral cues in an embryonic environment. Our molecular, cellular, and embryonic characterization provides a blueprint for functional hPGC assays that could be used in the future to dissect aberrant pathways involved in reproductive disorders.

3.1 Micropatterned generation of human gastruloid cultures contain a PGC niche

Micropattern-assisted self-organization of human gastruloids has proven to be an invaluable

tool for studying species-specific events in the early embryo with molecular detail (*Warmflash et al., 2014; Etoc et al., 2016; Morgani et al., 2018*). Since gastruloids self-organize in response to a single BMP4 signal to generate patterned embryonic and extra-embryonic tissues, we first asked whether they also contain hPGCs. Using standard gastruloid induction protocols (*Deglincerti et al., 2017*), RUES2 cells (XX line, NIH# 0013) were cultured on micropatterned colonies and stimulated with BMP4. As previously shown, cells displayed self-organized fate acquisition after 48 hours of differentiation, with ectodermal cells (SOX2+) in the center, extraembryonic tissue (GATA3+) at the edge, and mesodermal (BRA+) as well as endodermal fates (SOX17+) in between (*Warmflash et al., 2014*) (**Supplemental Figure 3.1a-b**). Remarkably, expression of the early hPGC markers BLIMP1 and AP2γ began at day 2 of stimulation with BMP4, within a subset of SOX17+ cells. (**Figure 3.1a-d, h**). After 4 days of differentiation, SOX17+/BLIMP1+/AP2γ+ cells that were initially confined to a diffuse ring on the edge of the colony are now found closer to the center (**Figure 3.1a, e-g**). These cells highly expressed OCT4 and NANOG providing the molecular signature (SOX17+/BLIMP1+/AP2γ+/OCT4+/NANOG+) of specified hPGCs (**Figure 3.1e-g, i, Supplemental Figure 3.1c;** *Fernandes et al., 2018*).

After specification *in vivo*, PGCs become motile and follow migratory directional cues in order to settle in the future gonadal region. Newly specified hPGCs in the gastruloids begin to express CXCR4 (**Figure 3.1a, f-g**), which regulates PGC migration and survival in mouse and chick embryos (*Molyneaux et al., 2003; Stebler et al., 2004*). The same induction and self-organization of hPGC fate was observed in a male hESC line, RUES1 (XY, NIH# 0012), and a female iPSC line (XX, Q002), indicating that hPGC patterning occurs similarly in both genders and regardless of genetic background (**Supplemental Figure 3.1d**).

Figure 3.1 *Specification and self-organization of PGCs in 500 μm human gastruloids.* **a,** *Schematics showing the PGC fate acquisition in human gastruloids. After 2 days of BMP4 stimulation, a subset of SOX17 population starts to express BLIMP1 and AP2Y. Specific germ cell markers are upregulated by day 4 and PGCs move radially inwards. b, Early PGC marker BLIMP1 is detected in radial patterns within the endodermal (SOX17+), and mesodermal (BRA+) territory after 2 days of BMP4 stimulation. c-d, BLIMP1+ cells also express the germ cell marker AP2γ and are located adjacent to BRA+ and constitute a subset of SOX17+ cells after 2 days of BMP4 stimulation (50ng/μl). e-g, Expression of BLIMP1, AP2γ, and OCT4 in a subset of SOX17+ cells after 4 days of BMP4 stimulation. Cells move towards the center and OCT4 expression becomes restricted to BLIMP1+/AP2γ+ cells. A subset of the newly specified PGCs (AP2γ+) start to express CXCR4. h, Single cell image quantification of 5 colonies (~10,000 cells) after 2 days of BMP4 stimulation displays a consistent spatial organization of AP2γ expression intensity as the medial subpopulation of SOX17+ cells (left). However, K-means clustering of the joint histograms (right) indicates only two clusters, SOX17- (blue) and SOX17+ (orange), and no distinctly separated PGC cluster. i, By 4 days of stimulation, AP2γ (left) demarcates a fully specified medial PGC population. Clustering (right) now demonstrates a PGC population (green) separate from the population of SOX17+/AP2γ- presumptive endoderm (orange). a.u. arbitrary unit*

Figure 3.1

a

b   BRA

BMP4 - 2 days

BLIMP1   SOX17

c

d
BRA
SOX17

BLIMP1
SOX17

AP2ɣ
BLIMP1

e

BMP4 - 4 days

BLIMP1

AP2ɣ   OCT4

f

g
AP2ɣ

AP2ɣ
SOX17

AP2ɣ
BLIMP1

AP2ɣ
OCT4

AP2ɣ
CXCR4

BMP4 - 4 days

SOX17 intensity (a.u.)

AP2ɣ intensity (a.u.)

AP2ɣ intensity (a.u.)
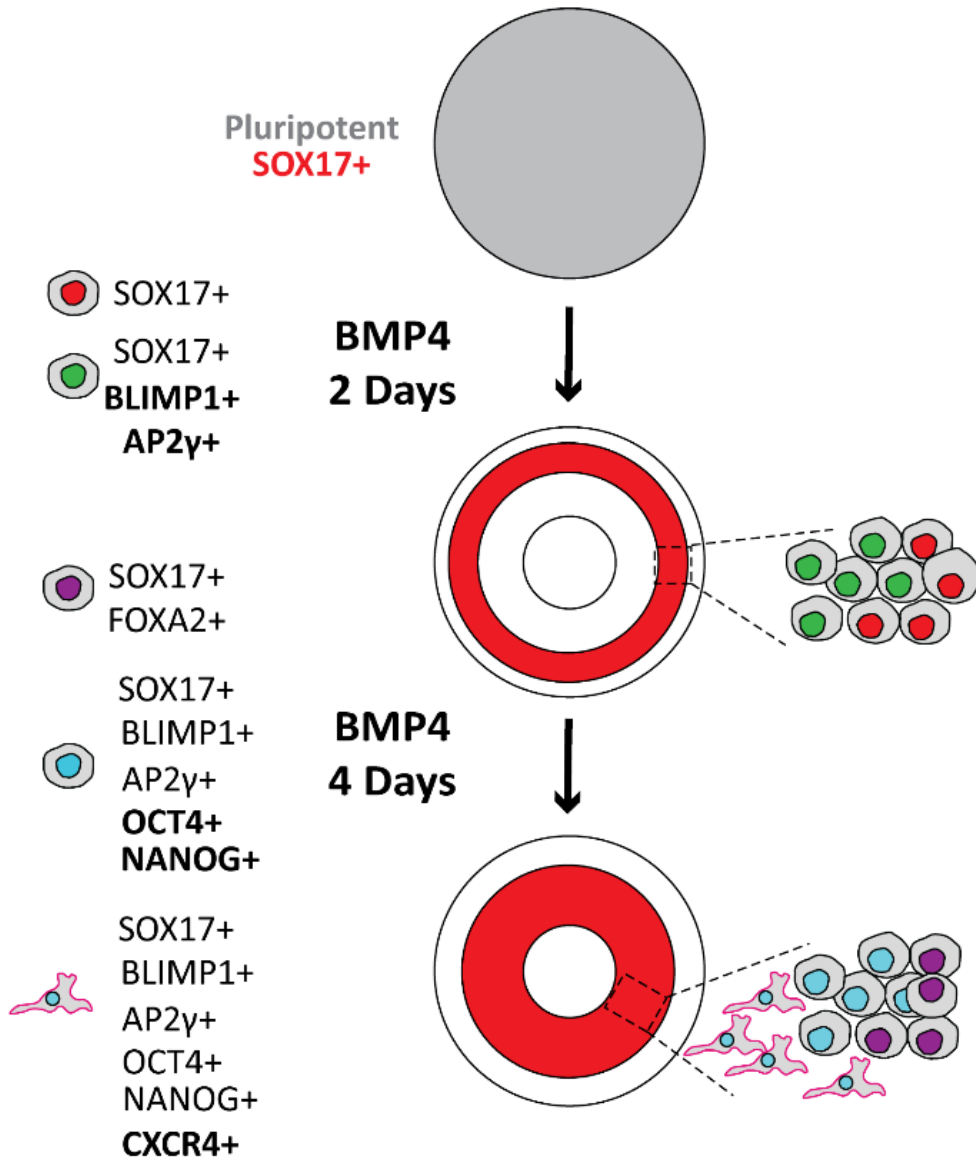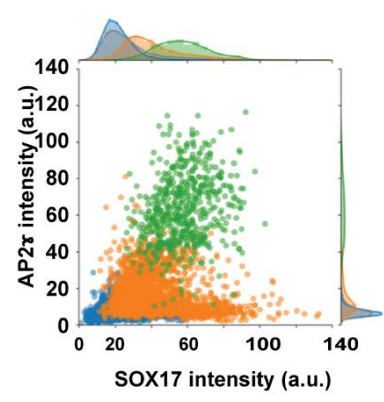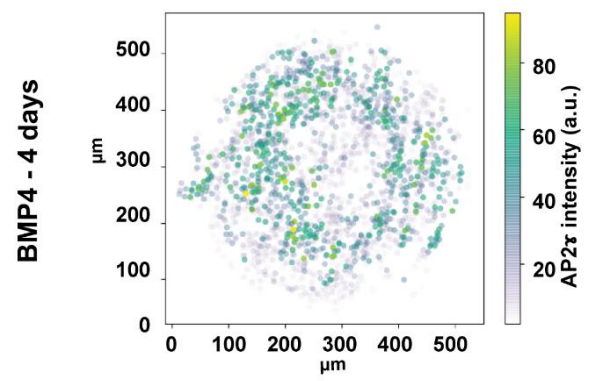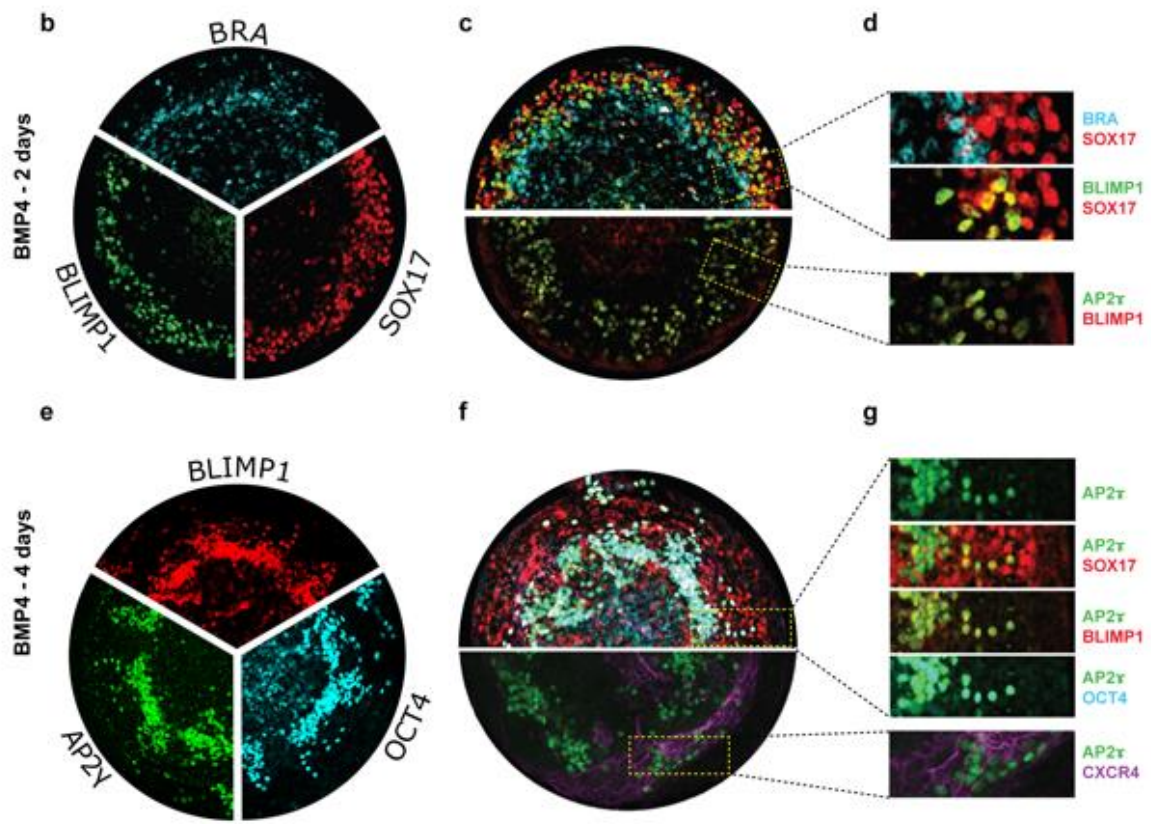
Micropatterned culture also provided an opportunity to optimize hPGC specification in response to BMP4 in the context of self-organization. hPGC specification was found to follow a dose-dependent response to BMP4 stimulation. As BMP4 dose was increased from 0 ng/ml to 100 ng/ml, the PGC population also increased proportionally and demonstrated greater central movement. In addition to concentration-dependent effects, PGC specification is also a function of duration of signal presentation. Four days of continuous BMP4 treatment provides a specific radial PGC fate acquisition (**Supplemental Figure 3.1e-f**). Our previous work demonstrated that reducing the colony diameter leads to an enrichment of the periphery population at the expense of central fates (*Warmflash et al., 2014*). To determine the colony size that maximize PGCs fate acquisition with minimum variability, we stimulated colonies ranging from 80μm to 1000μm with BMP4 for 2 and 4 days. The percentage of hPGCs in colonies increased as colony size decreased, though at the smallest diameters at the cost of significant intercolony variability (**Supplemental Figure 3.2a**). Therefore, we define an optimized method for hPGC enrichment by confining the hESCs in 225μm micropatterned colonies. To identify whether the shift of the radial distribution of hPGCs within the colony is due to cell movement or waves of transcriptional activation, we used CRISPR-Cas9 genome editing to generate an eGFP-tagged-OCT4 RUES2 line (*Hockemeyer et al., 2012*) (RUES2-OCT4-GFP). After initial patterning with 2 days of BMP4 stimulation, OCT4+ PGCs were tracked by time-lapse video microscopy for 16 hours with 5 minutes frame interval. While OCT4 expression was lost by all other cells in the colony, it was maintained and enriched in the hPGC population, and these cells began to display movement towards the center of the gastruloid (**Supplemental Figure 3.2b-d**). These results demonstrate that confinement of geometry and stimulation with a single ligand is sufficient to

31

induce hPGCs fate, thus providing a simplified alternative for the currently cumbersome

protocols involving a back-and-forth transition between an naïve and primed state and a large

cocktail of inductive factors (*Irie et al., 2015; Sasaki et al., 2015*).

3.2 Single cell transcriptomic analysis of hPGCs indicates a progression towards a migratory

molecular signature

Single cell RNA-seq analysis (scRNA-seq) was then used to identify the gene expression

signature of hPGCs in the context of self-organized 500 μm gastruloids after 4 days of BMP4

stimulation. We performed scRNA-seq on 4,984 individual cells with a mean of approximately

37,000 reads per cell. After unbiased clustering, uniform manifold approximation and

projection (UMAP) plots revealed a cluster of hPGCs clearly separated from embryonic and

extraembryonic mesodermal, ectodermal, and endodermal lineages (**Figure 3.2a-b**).

Differential expression analysis of the cluster revealed genes previously associated with hPGC

development such as SOX15, SOX17, TFAP2C, and NANOS3 as well as a group of non-canonical

genes with high clustering power (*Li et al., 2017; Chen et al., 2018; Yamashiro et al., 2018*;

AUC>0.7 by ROC test; **Figure 3.3a, Supplemental Figure 3.3a**). Comparison with published and

publicly available mouse scRNA-seq data revealed several genes highly associated with PGC

development that were not present in mouse datasets, and therefore of specific interest in

determining species-specific transcription signatures *(Pijuan-Sala et al., 2019)*. These genes

include morphogens such as WNT2, FGF13, FGF16; tyrosine kinase receptor SGK1; surface

membrane proteins PDPN, EPCAM, STOM; cilia-associated gene PIFO; anchoring protein

DOCK2, signal transducers RAB15, RASD1, PIM2, CRYM; and transcription factors SOX17, LBH,

MORC4. Strikingly, many of these genes are involved in the DOCK2-mediated signaling pathway,

working in concert to transduce chemoattractive signals during lymphocyte migration and

Figure 3.2. *scRNA-seq of gastruloids reveals a unique hPGC molecular signature.* ***a****, A UMAP plot generated from scRNA-seq analysis of day 4 gastruloids unveils 10 distinct clusters, as defined by differential gene analysis. Two of these clusters are subtypes of ectoderm (blue box), and 5 clusters are mesoderm (purple box), with an undifferentiated cluster separating the two germ layers. Both endoderm (red box) and PGCs (green box) consist of a single distinct cluster.* ***b****, Differential expression of select markers defining each cluster as a log-scaled dot-plot, in which dot size indicates the percent of cells in a cluster that express the given gene.*

Figure 3.2
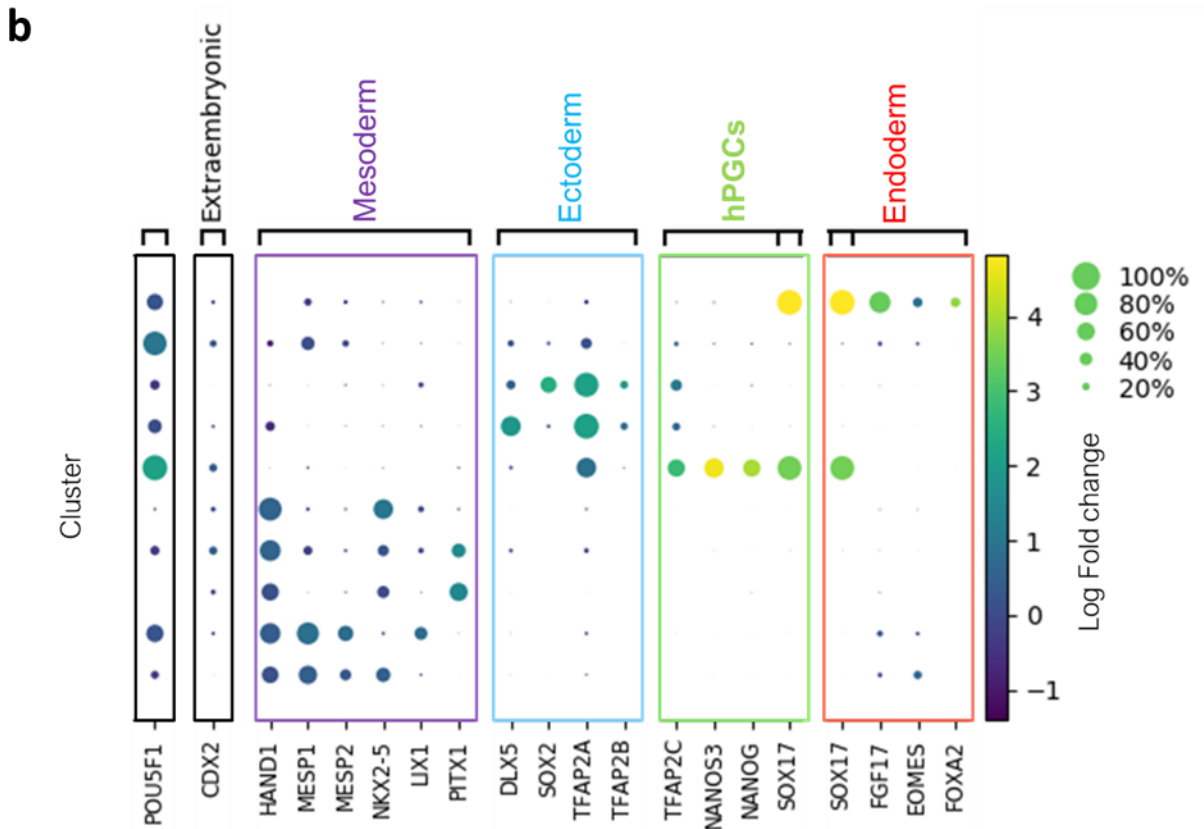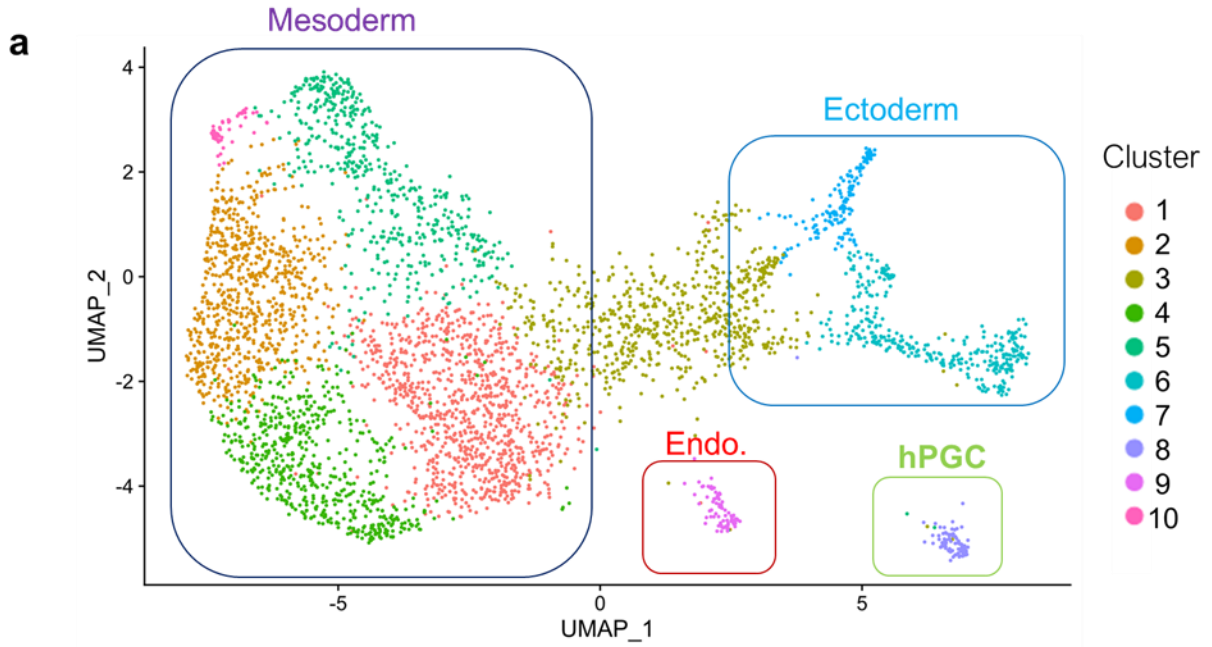
thymus colonization (**Supplemental Figure 3.3a**; *Fukui et al., 2001; Dobbs et al., 2015*). In order

to determine the extent to which our gastruloid-derived hPGCs recapitulate expression

patterns of hPGCs *in vivo*, we integrated our gastruloid dataset with a recently published

dataset containing primary hPGCs and gonadal stroma directly from human embryos (**Figure**

3.3b; *Li et al., 2017; Lee et al., 2010; Stuart et al., 2019*). 93% of gastruloid-derived hPGCs

(90/97) clustered with primary hPGCs identified as "migratory and mitotic" by expression of

markers such as POU5F1/OCT4 and NANOG, as well as strong expression of the newly identified

markers listed above (**Figure 3.3a**). The primary cells within this cluster were obtained from

predominantly week 4-10 embryos of both sexes, providing an estimated correlate for the

developmental stage of gastruloid-derived hPGCs.

Pseudotime analysis (*see Methods*) was performed on the entire gastruloid dataset, and the

hPGC cluster was isolated to investigate the acquisition of specification and migratory markers.

Cells ordered in pseudotime displayed a concurrent increase in expression of specification

markers (AP2γ) as well as migration-related genes (CXCR4 and DOCK2) (**Figure 3.3c**). This

demonstrates that the specified hPGCs in day 4 gastruloids express molecular machinery

necessary to respond to various guidance cues (*Molyneaux et al., 2001; Mitsunaga et al., 2017*)

Migration of PGCs via vascular mechanisms has not been previously described in mammals,

however in chicken embryos, the normal route of PGC migration as they home toward the

gonads is through the developing vasculature (*De Melo et al., 2012*). This suggests the intriguing

possibility that cues of PGC migration through the chick vasculature and mesonephros might be

evolutionary conserved in human PGC migration through the hindgut and mesonephros as well.

Figure 3.3. Comparison of gastruloids derived PGCs with fetal PGCs. ***a,*** Dot-plot and (***b***)

UMAP clustering of newly identified and canonical hPGC markers in an integrated

gastruloid-derived hPGC and primary hPGC dataset. Clusters are defined by cell type (see

Supplemental Figure 3) and split by primary (blue) or gastruloid (red) origin. ***c,***

Pseudotime analysis of single cells within the PGC node indicates a smooth progression

of developmental time within the cluster (left). Expression of novel hPGC marker DOCK2

(center) and canonical migratory PGC marker CXCR4 (right) display a similar increase in

expression in PGCs later in the pseudotime trajectory.

Figure 3.3
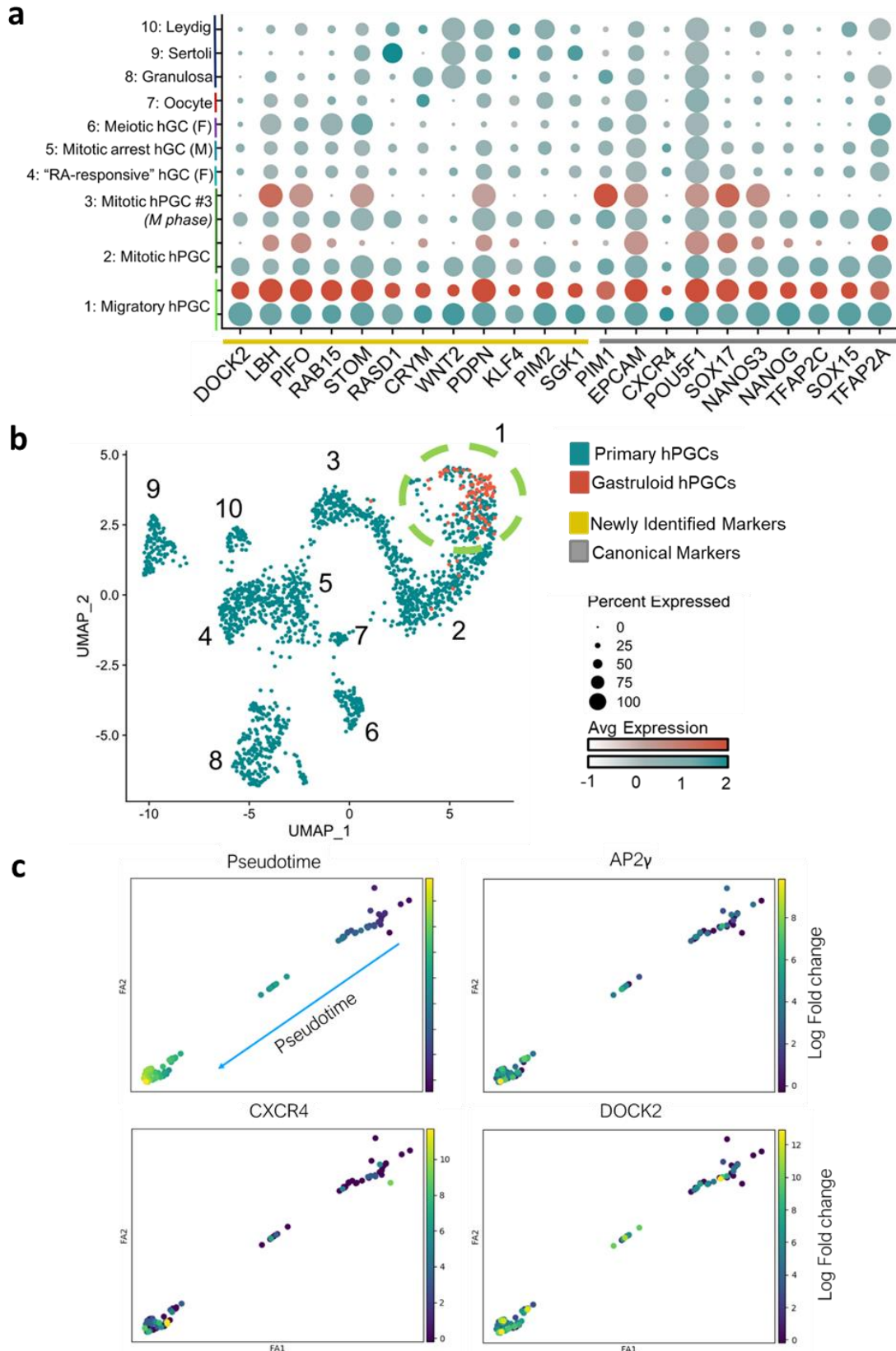
3.3 Gastruloid-derived hPGCs demonstrate migratory potential in a chick embryonic environment

In order to test whether this molecular signature truly indicated migratory potential, hPGCs were transplanted into chick embryos to validate fate and migration properties in an *in vivo* environment. Avian PGCs found in the anterior extraembryonic region (termed the germinal crescent (GC)) converge axially to enter the embryonic circulation. Then, avian PGCs use the bloodstream to move through the embryo exiting in the mesonephros at the genital ridge (De Melo et al., 2012; Kang et al., 2015). This migration path is known to require several of the same key genes identified in hPGCs in our scRNA-seq dataset, particularly CXCR4 (Li et al., 2017; Stebler et al., 2004). To investigate whether hPGCs could follow early chick migratory cues *in vivo*, a hPGC reporter line (WIS24-NANOS3-mCHERRY) and a control line H2B-tagged (RUES2-H2B-Venus) were cultured in 225 µm micropatterned colonies and stimulated with BMP4 for 4 days to enrich for hPGCs. Individual colonies were lifted from the micropatterns and transplanted in several locations along the anterior-posterior axis of the HH4 (~16 hour) chick embryo (*Wodak et al., 2007*). Examination of transplanted chick embryos after 24 hours revealed that <u>only</u> hPGCs specifically grafted in the GC displayed lateral displacement and are distributed in this region; in a manner similar to what has been previously described for chick PGC migratory behavior (**Figure 3.4a-k, Supplemental Figure 3.4,** *Kang et al., 2015; Nakamura et al., 2007; Lee et al., 2016*). Active movement of hPGCs, consistent with the acquisition of migratory machinery detected in our scRNA-seq dataset, is supported by the fact that neither untreated hESC colonies transplanted into the GC, nor hPGC colonies grafted in positions outside the GC, move significantly from the grafted area (**Supplemental Figure 3.4**).

Figure 3.4. hPGCs grafted into the developing chick embryo display migratory

characteristics. *a*, Schematic of hPGC micropatterned colony graft into different positions

of the anterior-posterior (AP) chick axis. HH4 stage (~16 hours of development) embryos

were maintained in EC culture and used as host for transplantation of 225μm WIS24-

NANOS3-mCHERRY micropatterned colonies treated with BMP4 for 4 days. *b*, Embryos

developed normally after the transplant, and *b'*, HNA+ hPGCs can be detected up to 600

μm from the original graft site (outlined) after 24 hours. *c-d* (white dotted area in *b'*),

DAPI+/HNA+ cells coexist with DAPI+/HNA- chick cells in the embryo. e, OCT4 is

specifically expressed in HNA+ cells in the region of the graft. *f-g*, DAZL is expressed in

the cytoplasm of OCT4+ hPGCs that move away from the initial graft site. *h-j,* Cells in the

core of the initial grafted site express NANOS3 and do not turn on DAZL.*, k,* Area of

displacement of hPGCs grafted along the axis. Colonies stimulated with BMP4 for 4 days

grafted at the GC displayed a significantly larger area of displacement after 24 hours

than cells grafted in the lateral or posterior positions. Pluripotent colonies transplanted

at the GC display no movement from the initial graft site. Results shown correspond to

the mean of at least 3 independent grafted embryos in each A-P position. Scale: *b* -

500μm, *b'* - 70μm, *c-e*, *h-j* - 50μm, *f-g*.
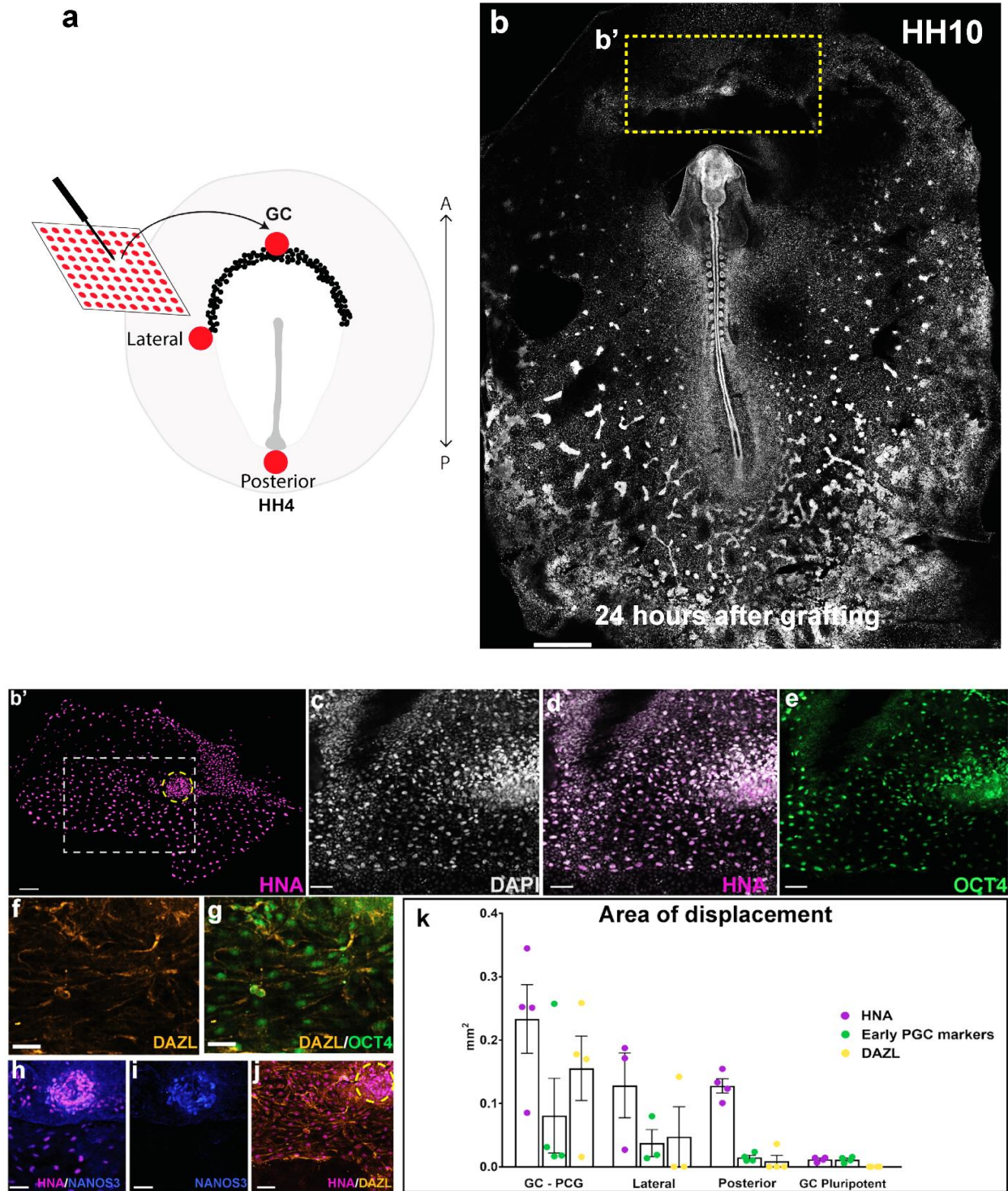
Figure 3.4

Additionally, expression of DAZL, which in chick embryos is observed in early migrating PGCs, is detected in hPGCs after 24 hours in the chick embryo (**Figure 3.4f-g**). DAZL is not turned on in cells in the original grafted site (**Figure 3.4j**, dotted area) nor in untreated colonies grafted at GC and colonies grafted in the posterior end of the embryo's axis (**Figure 3.4i, Supplemental Figure 3.4n**). One out of 3 embryos grafted with hPGCs in the lateral side expressed DAZL and few OCT4+ cells displayed movement towards the embryo (**Supplemental Figure3. 4b-k, p**). Interestingly, NANOS3 was only expressed in hPGCs that remained at the initial graft site, suggesting that expression of NANOS3 in hPGCs may not be continuous throughout all stages of PGC migration (**Figure 3.4h-i**). Indeed; we indicate in **Chapter 6** that PGCs that migrate away from the center of gastruloids upon long term Matrigel culture or co-culture express DAZL while downregulating OCT4 and NANOS3. Other analyses are required to demonstrate definitively which stage of PGC development these cells resemble most, though this OCT4-DAZL switch is described in human ovaries as they transition form first trimester (day 61) to second (14 wk) in Anderson et al., 2007. These results together demonstrate instructive crosstalk between chick and human embryonic cells. This has also been observed in our previous grafting experiments as previously described in human organizer transplanted into chick embryos (*Martyn et al., 2018*).

We next asked whether hPGCs could follow a later phase of chick PGC migration and use the

vasculature route to home to the gonads. hPGCs grown under the same condition were

dissociated and microinjected into the dorsal aorta of HH16-17 stage chick embryos (~56 hours,

Supplemental Figure 3.5a, a'). Survival rate and number of embryos analyzed are shown in

Supplemental Table 3.2. Examination of embryos 2.5 days after injection revealed the presence

of HNA+/DAZL+ cells in the chick gonads (**Figure 3.5l-r, Supplemental Figure 3.5d-f'**). SOX2

expression (which is maintained in chick and mouse PGCs but not in hPGCs during

differentiation), is not detected in our hPGCs, pointing to cell-autonomous behavior of

hPGCs[12,29] (**Supplemental Figure 3.5g-h**). Control embryos, injected in parallel with either

Ringer's saline solution or pluripotent RUES2 cells (RUES2-H2B-mCITRINE[30]), with no HNA+ in

the gonads or acquire DAZL expression (**Supplemental Figure 3.5i, i', 6**). Chick PGCs present in

the developing gonads of control embryo express DAZL (**Supplemental Figure3.5 j-k**). This

provides evidence for the specific migratory behavior of hPGCs and provides a convenient *in

vivo* xenograft platform upon which to perform functional migratory assays. Further

experiments will demonstrate whether the migratory machinery identified **in Figure 3.2** is

indeed responsible for this migratory behavior, and whether this occurs downstream of either

SDF-1 - CXCR4 or SCF – KIT chemoattractive signaling. These experiments demonstrate that

hPGCs can follow PGC migratory cues in two aspects of the chick embryonic environment, and

provide *in vivo* functional evidence for the ability of hPGCs to use the vasculature system to

home to their target, indicating that many of these cues may be conserved between the two

species - despite the superficial differences between migratory paths.

Figure 3.5. Human PGCs injected into the chick vasculature reach the chick gonad by

HH25. *l, Schematics showing the gonadal region of a chick embryo at stage HH25-26.*

*Gonads are positioned ventral to the mesonephros. **m-m'**, Ventral view of gonadal*

region of the chick embryo at stage HH25, 2.5 days after hPGC intravascular injection.

Clusters of HNA+ cells were observed within the gonad (dotted area in m'), and the

*surrounding mesonephric area (arrow in **m'**). **n-o**, Confocal image of dissected gonads*

shows nuclear expression of HNA in hPGCs. DAZL expression demarcates the cortical

*germ cell containing region of the gonad. **p**, Merged image highlights DAZL+ expression*

*in the HNA+ hPGCs (magnified in **q** and **r**). Scale: **n-o** - 30µm, **m** - 150µm, **m'**-50µm.*

Figure 3.5
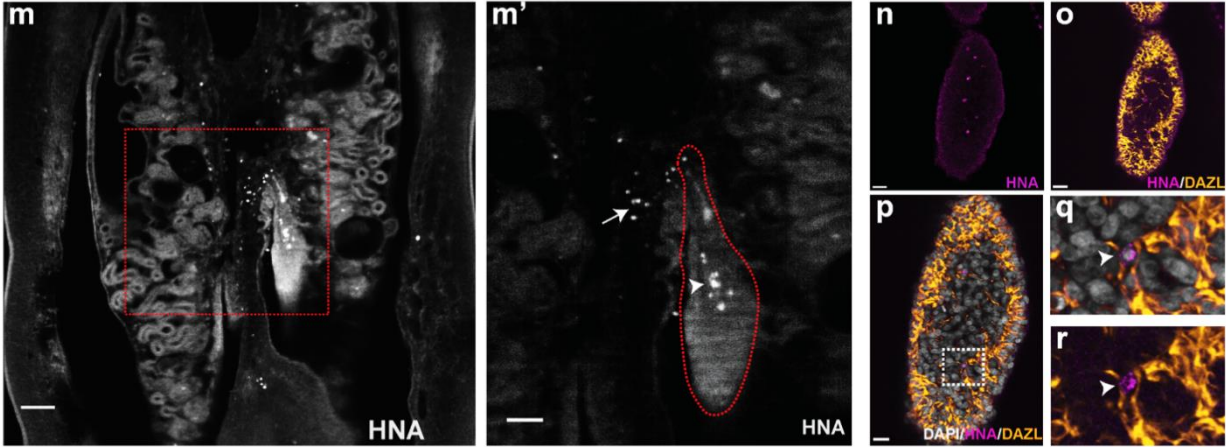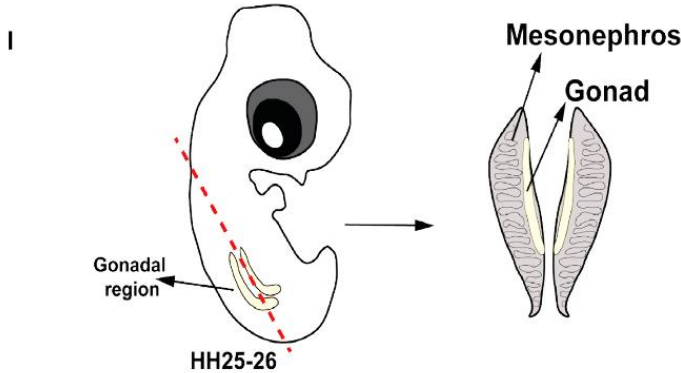
Taken together this study provides the molecular signatures of hPGCs, highlights genes previously not associated with their development including species-specific players, and *in vivo* evidence of their migratory and homing potential. The chick assays described here will be vital for interrogating mechanisms of hPGC migration, including the human players described in this study. The combination of molecular, cellular, and embryonic approaches could be used in the future to dissect aberrant pathways involved in reproductive disorders.

Supplemental Figure 3.1. *Additional characterization of PGC specification within human gastruloid culture. **a**, Normal gastruloid development was confirmed by confocal imaging, indicating radially organized expression of the germ layer markers SOX2 (blue), BRA (red) and GATA3 (green) within 500um micropatterned colonies cultured for 2 days with 50 ng/mL BMP4. Coexpression of NANOG (cyan) with SOX17 (red, left) and AP2γ (red, right) in 2 days gastruloids (**b**) in comparison with 4 days gastruloids (**c**) demonstrates complete specification by day 4. **d,** Confirmation of PGC specification in other genetic backgrounds is demonstrated by expression of SOX17 (red) and AP2γ (green) in 2 days gastruloids for male RUES1 (left) and female iPSC (right) lines. **e,** Timecourse of PGC specification between days 2 and 4 demonstrates both an increase in the number of PGCs stained for BLIMP1 (red), AP2γ (green) and OCT4 (cyan) and a movement towards the center of colonies (left). Single cell quantification and joint histogram (center) of AP2γ (x-axis) and OCT4 (y-axis) expression indicates that an increase AP2γ expressing cells is concurrent with the specification of OCT4 to AP2γ + cells and loss of the orthogonal OCT4+/AP2γ - population (Pearson 2d = 0.29, 3d = 0.68, 4d = 0.87; n=5 colonies ~10,000 cells/condition). The radial distribution of AP2γ + cells (right) displays a progressive decrease in mean radial distance and variance between day 2 of culture (blue), day 3 (orange), and day 4 (green). **f,** PGC specification after 4 days of culture as defined by colocalization of BLIMP1/AP2γ/OCT4 is dependent on BMP4 dosage (left). Single cell analysis (center) indicates a gradual increase in AP2γ/OCT4+ cells from 0 ng/mL to 100 ng/mL BMP4 exposure (Pearson (25ng/ul) = 0.39, (100ng/ul) = 0.87). The total number of AP2γ/OCT4+ cells per colony displays a roughly linear dependence on dose (right). Scale bar: 100μm.*

Supplemental Figure 3.1. *Additional Characterization of PGC Specification within Human Gastruloid Culture*
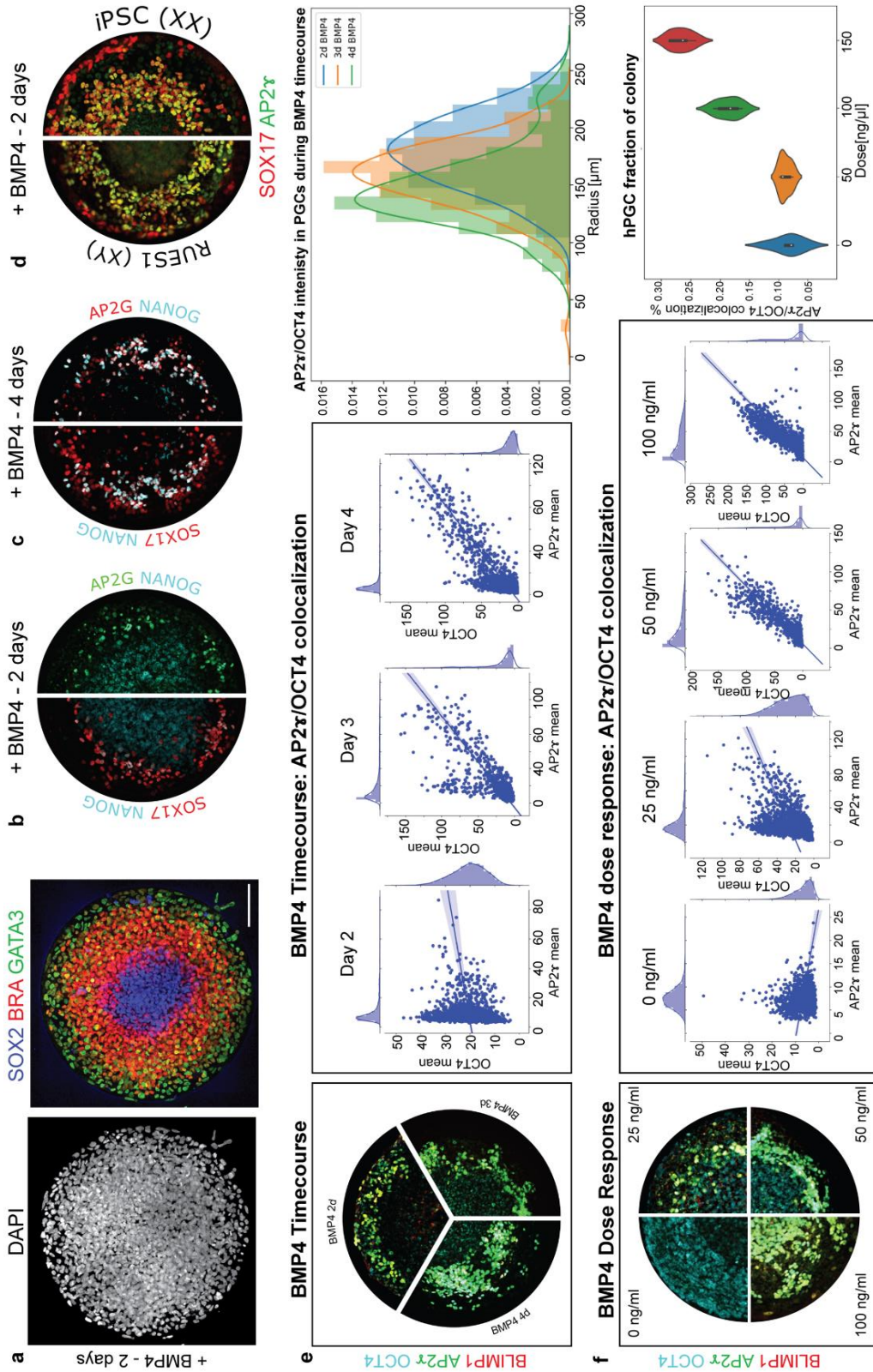
Supplemental Figure 3.2 Geometric confinement and hPGC movement during specification. *a,* Representative confocal imaging of SOX17 (red) and AP2γ (green) at day 2 (left slices) and 4 days (right slices) used to calculate optimum graft colony size as quantified in the right panel. Changing the micropatterned colony size allows optimization of the fraction of colony that acquires a PGC fate while minimizing inter-colony variability (n=10,000 cells/condition. K-means clustering across four markers of PGC fate: AP2γ, BLIMP1, SOX17, OCT4). *b,* Stills from OCT4-EGFP time-lapse imaging at 48 hours of BMP4 stimulation and after 63 hours. The vector motion of each specified OCT4-EGFP+ cell at the end of the video can be computed at each frame and used to generate a map of PGC mean radial distance (right), demonstrating a net movement of 29 μm towards the center over the 16 hour time-lapse, beginning at 53.8 hours of BMP4 stimulation. (Orange line indicates 20-frame moving average). *c,d,* The radial distributions of four PGC markers AP2γ (blue), SOX17 (orange), OCT4 (green), and BLIMP1 (red) in 500 um colonies display a progressive decrease in mean radial distance and variance between days 2 (*c*) and 4 (*d*) to define a tightly colocalized population of PGCs as the medial aspect of the SOX17 distribution (orange; n=5 colonies ~10,000 cells/condition).

Supplemental Figure 3.2 *Geometric confinement and hPGC movement during specification.*
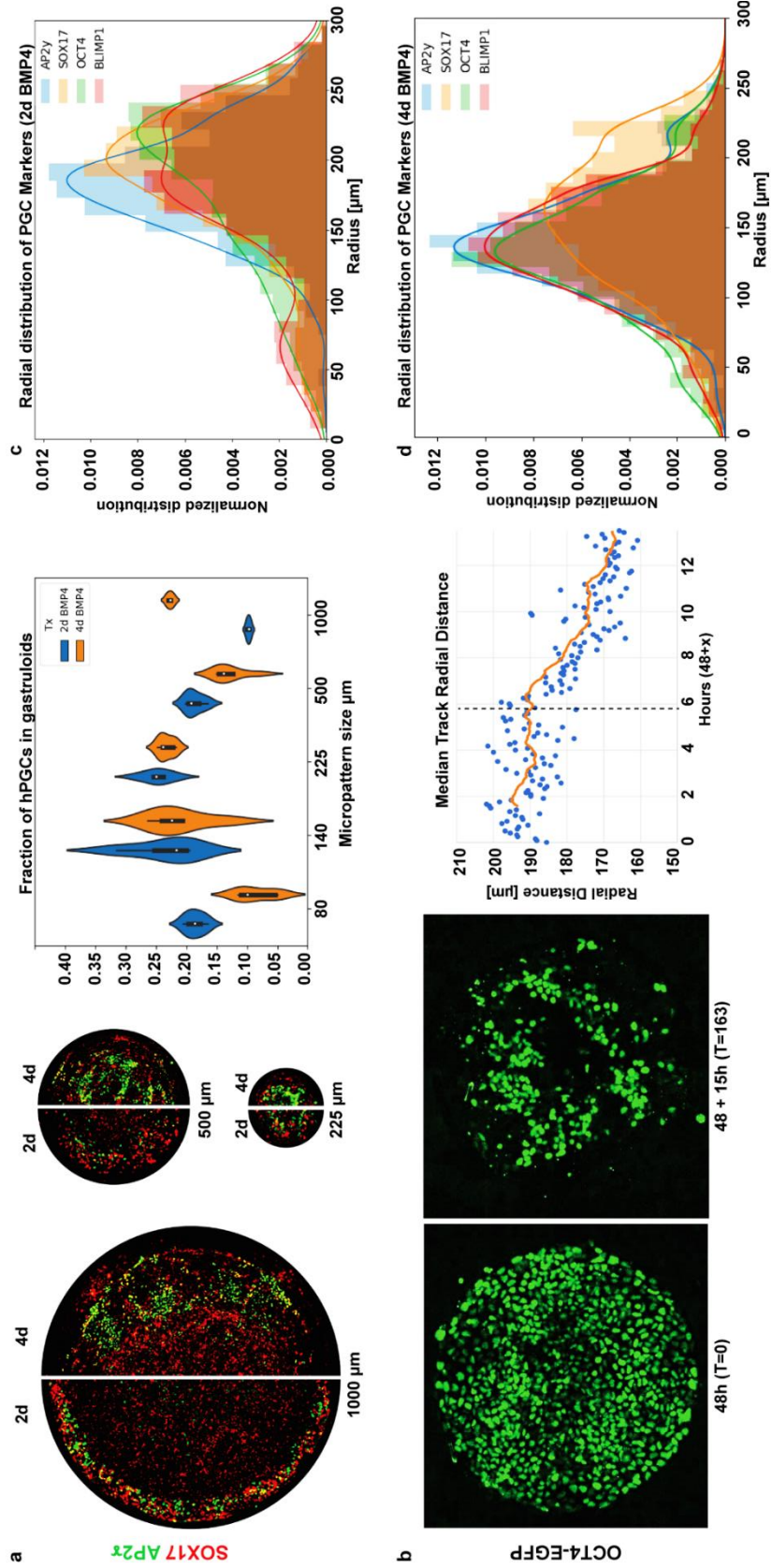
Supplemental Figure 3.3. Novel human PGC genes of interest with high clustering power and low mouse expression. a, Receiver Operating Characteristic (ROC) curve analysis was used to generate power rankings for genes expressed by PGCs, which together with percent population expression was used to generate a table of genes that are most highly specific and most characteristic of human PGCs. Known human specific marker SOX17 was excluded from this set. Genes were filtered for transcription factor and cell signaling function. This table includes canonical genes such as POU5F1, NANOG, TFAP2C, and NANOS3 (dark blue), as well as many genes with no described roles in human PGC development (yellow). Genes whose function is cell motility and migration are highlighted (dark yellow) b, Pseudotime analysis of 96 hours gastruloids, with PGCs (red box) and other local maxima numbered and labeled according to characteristic regulatory genes (right). c, Expression of markers used to stage hPGCs, hGCs and stromal cells in an integrated dataset comprised of gastruloid-derived hPGCs and primary human cells. Gastruloid-derived hPGCs cluster with migratory, mitotic primary hPGCs, but do not yet express late markers DDX4, DAZL, and DPPA3.

AUC – area under ROC curve, Avg. diff. expression – average differential expression, Avg log-fold change – average log-fold change, % of cells in PGC cluster– percent of cells in PGC cluster expressing gene X, % of cells outside PGC cluster – percent of cells outside PGC cluster expressing gene X.

Supplemental Figure 3.3 Single Cell RNA-seq Analysis of Day 4 Gastruloids reveals hPGC Genes of Interest

a

Gene	myAUC	Avg. diff. expression	Power	Avg. log-fold change	% of cells in PGC cluster	% of cells outside PGC cluster
PDPN	0.990	2.132	0.980	2.132	1.000	0.239
LBH	0.977	1.913	0.954	1.913	0.989	0.212
PIFO	0.951	1.492	0.902	1.492	0.955	0.158
RAB15	0.917	1.319	0.834	1.319	0.876	0.102
EPCAM	0.910	1.549	0.820	1.549	0.955	0.301
STOM	0.879	1.089	0.758	1.089	0.933	0.332
DOCK2	0.862	1.426	0.724	1.426	0.730	0.017
PIM2	0.825	1.299	0.650	1.299	0.697	0.085
RASD1	0.793	1.041	0.586	1.041	0.607	0.030
MORC4	0.788	0.613	0.576	0.613	0.640	0.079
CRYM	0.784	0.802	0.568	0.802	0.596	0.032
SGK1	0.767	0.846	0.534	0.846	0.573	0.052
FGF13	0.756	0.725	0.512	0.725	0.685	0.256
FGF16	0.740	0.626	0.480	0.626	0.483	0.003
WNT2	0.727	0.583	0.454	0.583	0.461	0.008
POU5F1	0.974	1.934	0.948	1.934	1.000	0.910
SOX17	0.955	1.693	0.910	1.693	0.944	0.044
NANOS3	0.876	1.675	0.752	1.675	0.764	0.028
NANOG	0.835	1.150	0.638	1.066	0.685	0.081
TFAP2C	0.819	1.066	0.638	1.066	0.685	0.081
SOX15	0.801	1.068	0.602	1.068	0.663	0.104
NKX1-2	0.751	0.590	0.502	0.590	0.528	0.032
TFAP2A	0.750	0.517	0.500	0.517	0.775	0.276
PIM1	0.734	0.515	0.468	0.515	0.764	0.357
ALPL	0.723	0.484	0.446	0.484	0.708	0.345
KLF4	0.711	0.712	0.422	0.712	0.438	0.018
CXCR4	0.657	0.512	0.314	0.512	0.348	0.035
PRDM1	0.635	0.368	0.270	0.368	0.303	0.038

b Pseudotime in gastruloid (UMAP1 / UMAP2)

Nodes	Characteristic Genes
* Origin Node	SOX2, POU5F1
1 PGC	TFAP2C, NANOG, NANOS3, POU5F1
2 Definitive Endoderm	SOX17, FOXA2, EOMES
3 Neuroectoderm	SOX2, OTX2, LHX5
4 Non-neural ectoderm	DLX5, TFAP2A
5 Neural Crest progenitors	TFAP2A, TFAP2B, FOXD3
6 Paraxial Mesoderm	FOXC1, MESP1
7 Intermediate Mesoderm	PAX2, LHX1
8 Cardiac Mesoderm	NKX2-5, HAND1
9 Lateral Mesoderm	CDX2, HAND1
10 Hemangiopoetic cells	FLT1, PODXL, GATA2

c

10: Leydig
9: Sertoli
8: Granulosa
7: Oocyte
6: Meiotic hGC (F)
5: Mitotic arrest hGC (M)
4: "RA-responsive" hGC (F)
3: Mitotic hPGC #3 (M phase)
2: Mitotic hPGC
1: Migratory hPGC

Percent expressed: 0, 25, 50, 75, 100
Avg Expression: -1, 0, 1, 2

*Supplemental Figure 3.4. Controls and quantification for hPGC grafts along the chick anterior-posterior axis. **a-c,** Schematic representation showing the strategy for hPGC-chick graft at different positions along the anterior– posterior axis. **d-f,** Untreated hESC colonies grafted at the GC area remain pluripotent (OCT4+) and display no movement nor DAZL expression. **h-j,** Colony transplanted in the lateral region present few cells moving towards the embryo and express cytoplasmic DAZL. Only one out of three embryos expressed DAZL. **l-n,** Colonies grafted at the posterior region (further from GC) do not express DAZL. **g, k** and **o,** Representative single cell quantification of the grafts in **a, b** and **c** demonstrating OCT4 expression in HNA+ cells as used to calculate area of graft displacement. **p,** The radial distributions of the untreated cells (red) and cells treated with BMP4 for 4 days grafted in the embryo's GC (blue), lateral (orange) and posterior positions (green). **q,** Single cell quantification of hPGC grafted at the GC showed in Figure 3. Scale bar: 50μm.*

Supplemental Figure 3.4. *Controls and quantification for hPGC grafts along the chick anterior-posterior axis.*
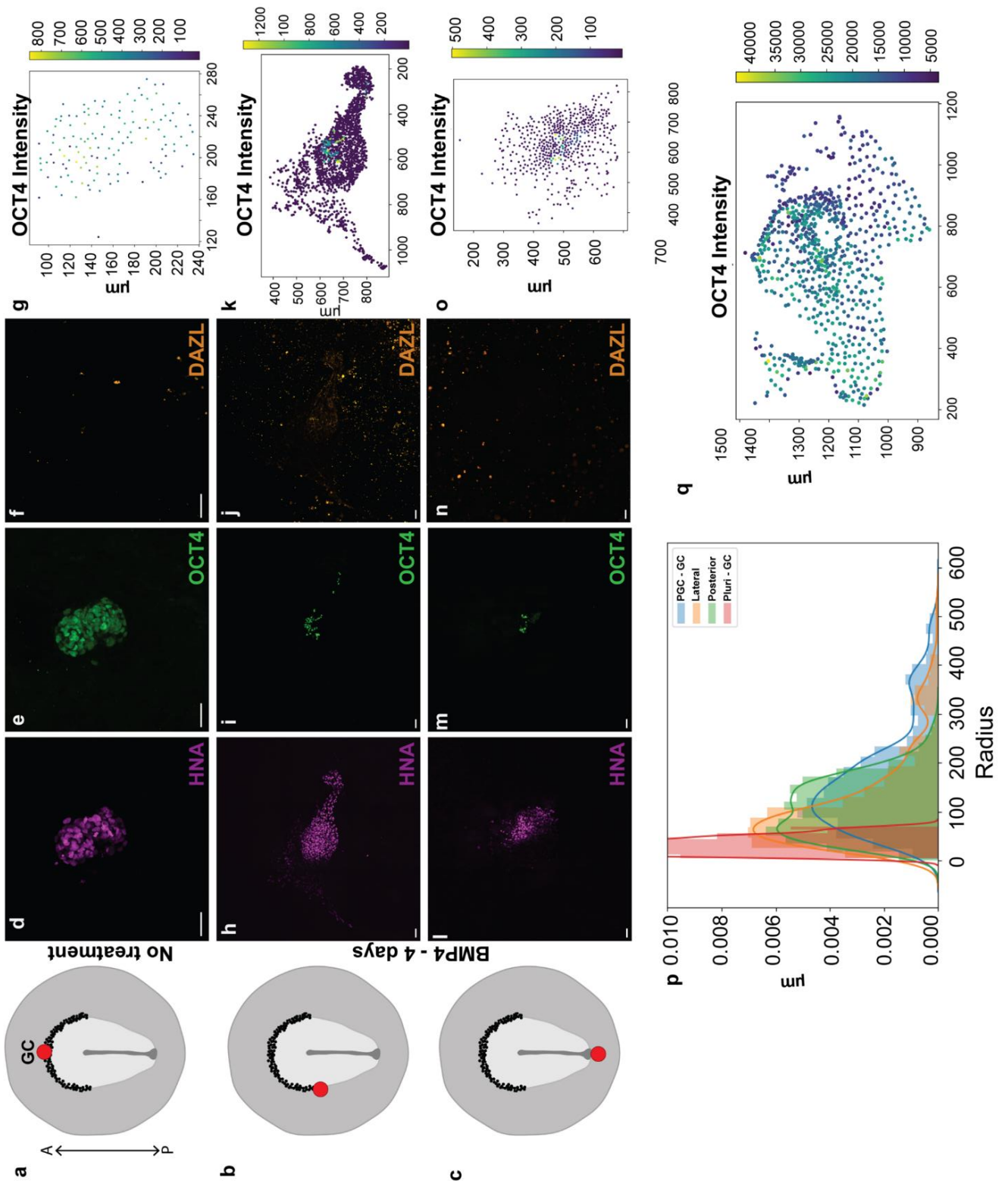
Supplemental Figure 3.5. hPGC injections and migration through the chick vasculature. *a-a'*, Still images of the injection site in the dorsal aorta of the HH16 (~56-60 hours) chick embryo immediately pre-injection (left) and post injection (right). The injection needle is visible on the left of both images and fast green was used to visualize injected fluid. *b*, Injected embryos display normal development 24 hours after hPGC injection. *c*, Embryos display normal morphology 2.5 days post-injection reaching stage HH25. *d*, 3D reconstruction of the chick embryo showing the gonads and expression of SOX2. *e, e'*, Human cells are found in the chick developing gonads (dotted yellow area) by the expression of HNA specific antibody. *f, f'*, Ventral view of a z-stack of the developing chick gonad (yellow dotted area in f and g) shows HNA+ cells in the mesonephric ducts and the gonad (arrow). *g-h*, These cells do not stain positive for SOX2, a marker of human pluripotency and chick PGCs (green). *i-i'*, Control embryos injected with Ringer's saline solution do not present any HNA expressing cells (gonad outlined and shown in inset). *j*, DAZL (magenta) is stained to demarcate the normally developing chick gonad (merge image in *k*). Scale bar: *d* - 400µm, *e* - 200µm, *e'*-100µm, *f-g* - 300µm, *i-k* - 200µm.

Supplemental Figure 3.5. *hPGC injections and migration through the chick vasculature*.
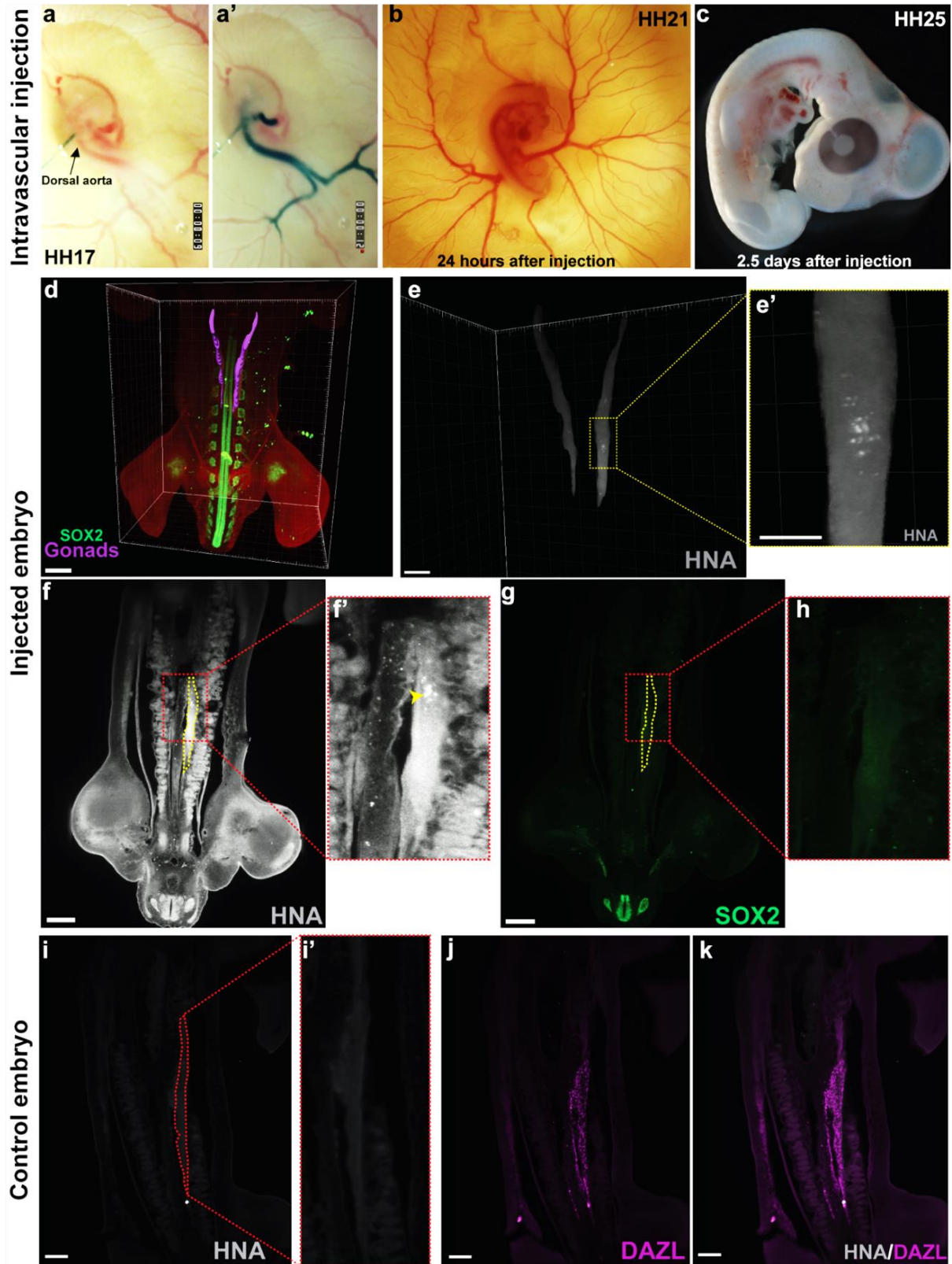
Supplemental Figure 3.6. *Controls for hPGC injections within the chick vasculature. iDISCO prepared embryos injected with pluripotent H2B-mCITRINE control cells were imaged using light sheet microscopy to ensure specificity of migratory behavior to hPGCs. **a,** Sagittal view of HH25 chick embryo 2.5 days after the injection. DAZL expression (grey) demarcates the gonad region. Arrow indicates the mesonephric region. **b-c,** There is no co-expression of HNA (magenta) and H2B-mCITRINE (green) in the DAZL+ cells in the gonad (inset shows gonad and surrounding tissue). **d-f,** Higher magnification of the gonad (outlined) region, dotted area in **a-c**. **g-i,** Confocal images demonstrate DAZL+ cells populating the gonad (yellow) and all nuclei marked with DAPI (grey). There is no HNA (magenta) staining in the gonad. j-l, Higher magnification of the outlined region in **g-i**. Scale bar: **a-f**: 200μm, **g-i**: 150μm.*

Supplemental Figure 3.6. *Controls for hPGC injections within the chick vasculature.*
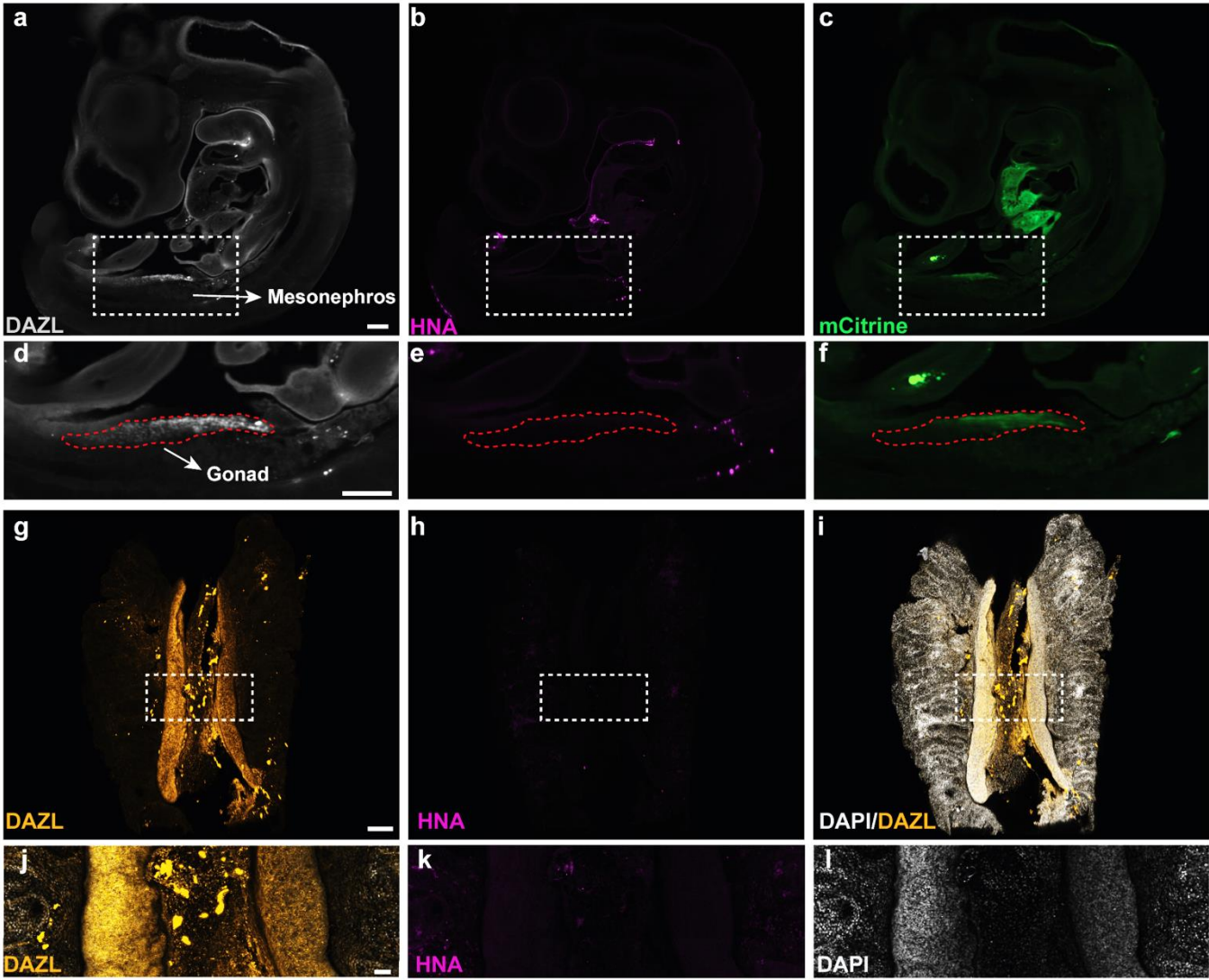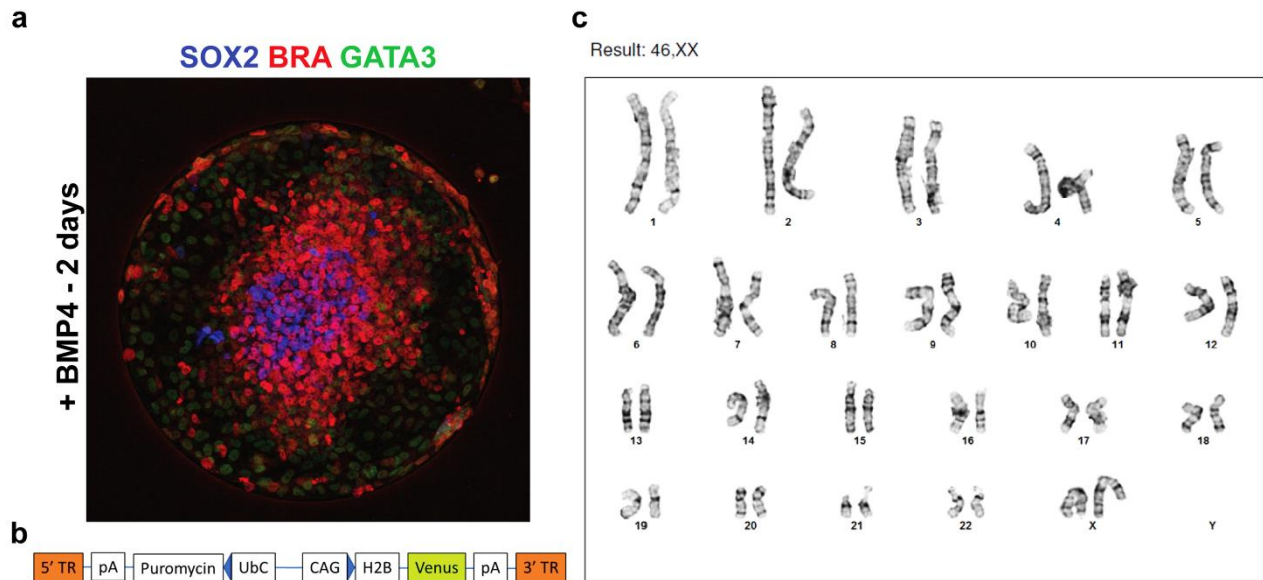
Supplemental Figure 3.7. *Design and Validation of Genetically Modified Cell Lines. **a,** Q002*

iPSCs generated using the ReproRNA-OSKGM reprogramming vector (StemCell Technologies)

form characteristic gastruloid structures with a central SOX2+ domain surrounded by BRA+ cells

and GATA3+ cells at the periphery of the colony (2 days of 50 ng/ml BMP4, 500 μm

*micropatterned colonies). **b,** This iPSC line has a normal human female karyotype. **c,** Design of*

the ePB Puromycin-H2B-Venus plasmid to generate the RUES2-H2B-Venus tagged line.

Supplemental Table 3.1. Summary of injection experiments

	Survival	Embryos expressing HNA
Control (Saline solution)	4/4	0/4
Control (Untreated colonies)	7/9	0/3
BMP4 4 days colonies	45/67	8/20

Survival column indicates the number of live embryos that developed normally after the microinjection. Many of these embryos were used to optimize the iDISCO protocol and final number of embryos analyzed are shown in HNA+ column.

CHAPTER 4

A signaling pathway analysis of the human PGC niche in a gastruloid embryo model

One major benefit of micropatterned hESC differentiation culture is the ability to probe the effects of morphogen gradients on cell fate acquisition. Due to the greater number of cell neighbors a cell at the center of colony has when compared to one at the edge, cells in the center of a colony find their response to exogenous BMP4 blunted by both endogenous secreted factors as well as differences in receptor localization (*Etoc et al., 2016*). These mechanisms underlie a powerful emergent phenotype of gastruloids: regularized, self-organized cell fate patterning, allowing for quantifiable analysis of endogenous signaling during peri-gastrulation events.

Here we harness the power of micropattern gradient-generated cell fate analysis to determine the exact roles of BMP4, WNT, and NODAL signaling in human PGC fate acquisition. Provocatively, a cascade of these same three signaling factors has been elucidated in the induction of mesendoderm cells, which share expression of SOX17 with hPGCs. In this cascade, BMP4 is an initiating factor inducing WNT expression, while WNT primes cells to become responsive to ACTIVIN and begin stable mesendodermal gene transcription (*Yoney et al., 2018*). We investigate which elements of this cascade are necessary and sufficient for human PGC specification, and determine their order and timing (**Figure 4.1**).
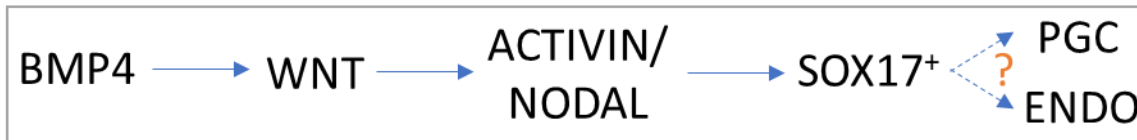
Figure 4.1. Morphogen cascade in micropatterned gastruloids acts to specify SOX17+ endoderm cells, as well as SOX17+ hPGCs. It is yet to be fully elucidated how SOX17+ cells in the same niche are instructed to become either endoderm or hPGCs.

4.1 Two roles of BMP4 in hPGC specification

We first tested whether – downstream of BMP4 - the WNT or ACTIVIN pathways acted as direct inducers of hPGCs. Gastruloids grown under the same conditions described in the previous chapter (500 um micropatterned colonies) were presented with either the WNT3A or ACTIVIN and hPGC fate markers were analyzed. We find that neither ligand alone is sufficient to induce hPGC fates (**Figure 4.2a**). WNT stimulation (in ACTIVIN-containing CM media) generates an edge-restricted SOX17+/BLIMP1+ population that does not express OCT4, NANOG, or AP2y and is thus presumptive early definitive endoderm, and not hPGC (**Figure 4.2a**). ACTIVIN alone is insufficient for differentiation and gastruloid self-organization. These results are in line with previous studies which show that cellular memory of WNT signaling is necessary for ACTIVIN morphogen activity and endodermal commitment, and demonstrate that neither pathway is sufficient for hPGC specification.

To address the requirement of the WNT and SMAD2 pathways for hPGC specification, micropatterned colonies were stimulated with BMP4 presented with or without inhibitors of WNT and ACTIVIN signaling (IWP2, and SB respectively). Both treatments inhibited hPGC specification, demonstrating that active WNT and ACTIVIN pathways are required downstream of BMP4 for hPGC specification (**Figure 4.2b**). Thus, while necessary, neither WNT nor ACTIVIN signaling is sufficient for hPGC specification.

Together, these results suggest that BMP4 must have another role in hPGC specification independent from its role as the initiator of the WNT-ACTIVIN cascade.

Figure 4.2
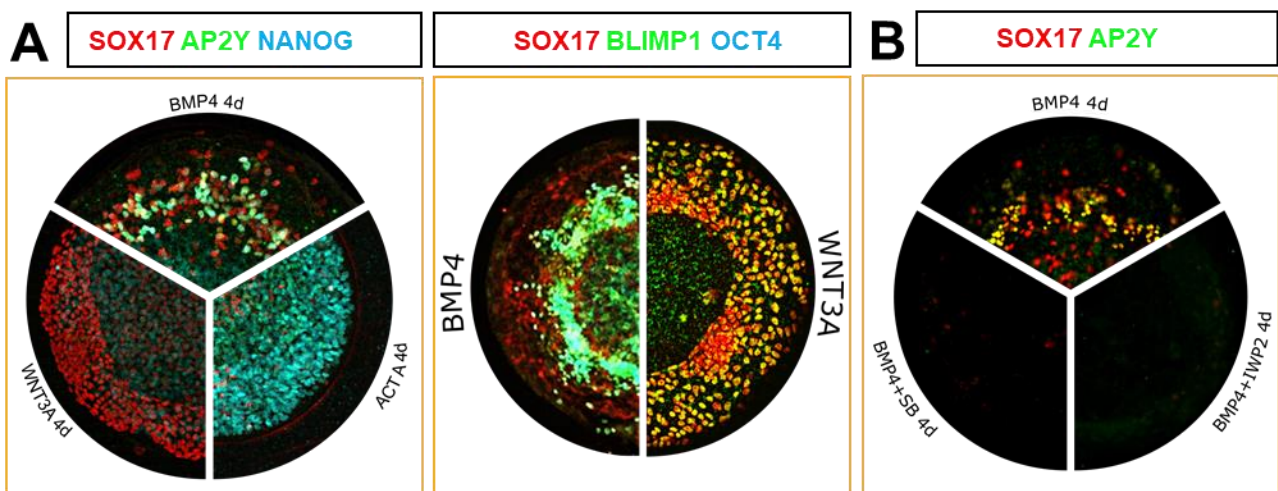
Figure 4.2. WNT and ACTIVIN are required but not sufficient for hPGC specification downstream of BMP4. *a,* hESCs grown on micropatterned 500 um colonies and treated with 4 days of either BMP4 (50 ng/ml), WNT3A (100ng/ml), or ACTIVIN (10ng/ml). Only with BMP4 addition do cells co-expressing SOX17, TFAP2C, and NANOG or SOX17, BLIMP1 and OCT4 exist. *b,* Blocking WNT signaling with IWP2 or ACTIVIN signaling with SB during BMP4 treatment prevents hPGC specification (as well as other SOX17+ populations), demonstrating WNT and ACTIVIN necessity but not sufficiency. Colony diameter = 500um).

4.2 hPGC specification requires BMP4 during the first 24 hours of differentiation

Taken together our signaling pathway analysis demonstrates that while the BMP4 - WNT - ACTIVIN signaling cascade acts to generate a niche of cells that can become hPGCs, there is an independent role for BMP4 as an instructive signal directly involved in hPGC specification. To determine the timing of signals required for hPGC specification in this niche, we changed the order and length of stimulations.

First, we induced the BMP4-WNT-ACTIVIN cascade and blocked BMP4 signaling after 24, 48, or 72 hours (**Figure 4.3a**). Withdrawal of BMP4 after 24 hours yields few to no hPGCs, while withdrawal after 48 hours yields a reduced number of hPGCs and withdrawal after 72 hours yields only slightly fewer hPGCs than standard 96 hour BMP4 stimulation. hPGC specification is therefore dependent on the duration of BMP4 signaling, in addition to the dose (as described in the previous chapter). It is also interesting to note that any hPGCs when BMP4 is withdrawn after 48 hours or earlier are always radially distal, and never found close to the center of the colonies – possibly suggesting a role for BMP4 in hPGC migration.

It is possible that hPGCs are not specified when BMP4 is withdrawn after 24 hours simply because 24 hours is an insufficient time to stably upregulate WNT and ACTIVIN signaling, and therefore a result of the cascade role of BMP4 rather than a result of an early independent role. To test this, colonies were directly stimulated with WNT and ACTIVIN during the first 24 hours of treatment *in the absence of external BMP4,* and then followed by 3 days of BMP4 stimulation (**Figure 4.3a**). These colonies generated hPGCs that exhibit a slightly more pronounced central movement than BMP4 stimulation alone, but maintain a similar hPGC population size (**Figure 4.3a**). Quantification of both cases indicates similar covariance and similar expression levels of hPGC markers (**Figure 4.3b**).

However, when BMP4 signaling is blocked by LDN presentation only during the first 24 hours of WNT/ACTIVIN treatment, followed by 3 days of BMP4 treatment, no hPGCs can be detected (**Figure 4.3E**, 9 o'clock). Conversely, when WNT signaling is blocked only during the first 24 hours, (in the presence of ongoing BMP4 signals) hPGC specification still occurs – only as long as WNT is present between 24-48 hours of differentiation (**Figure 4.3e**, 3 o'clock, **Supplemental Fig 4.1**). Thus, BMP4 signaling is required during the initial 24 hours of hPGC specification separately from its role in inducing WNT and ACTIVIN, while WNT signaling is not necessary during this time window and can be induced later.

Figure 4.3. BMP4 is required during the first 24 hours of hPGC specification. *a,* 500 um

gastruloid stimulation with BMP4 (50 ng/ml) was blocked by addition of LDN after 24

hours (12 o'clock), 48 hours (3 o'clock), or 72 hours (9 o'clock), and cells were stained for

BLIMP1 (red), AP2y (green), and OCT4 (blue). *b,* Quantification of cells coexpressing

BLIMP1, AP2y and OCT4 demonstrates that hPGC specification is correlated to the

duration of BMP4 exposure (n=5 colonies/10,000 cells). *c,* Staining for AP2y and OCT4 in

colonies treated with 4 days BMP4 and colonies where the first day of treatment is

replaced by direct stimulation with ACTIVIN (10 ng/ml) and CHIR (3 uM). Cross-section

(below) showing DAPI (grey), AP2y (green), and OCT4 (blue) demonstrates more cells at

the center of the colony in the ACTIVIN/CHIR pre-treated case than the control. *d,*

quantification of conditions shown in *c* demonstrate that equivalent numbers of hPGCs

are formed with similar intensity profiles (left: Pearson r = 0.86, p < 0.01; right: Pearson r

= 0.88, p < 0.01). *e,* Gastruloids treated on day 1 with either ACTIVIN/CHIR (12 o'clock),

BMP4/IWP2 to block early WNT signaling (3 o'clock), or ACTIVIN/CHIR/LDN to block

early BMP4 signaling, and then followed with 3 days BMP4. Only cells in which BMP4

signaling was blocked on day 1 were unable to generate hPGCs. (Colony diameter =

500um)

Figure 4.3

A — BLIMP1 AP2γ OCT4

BMP4 Duration

+LDN d2-4

+LDN d4

+LDN d3-4

B — LDN Treatment During PGC Induction

AP2γ/OCT4 Colocalized Cells

LDN d2-4   LDN d3-4   LDN d4

Treatment

C — BMP4 4d   CHIR/Act 1d→BMP4 3d

AP2γ OCT4   AP2γ OCT4

E — SOX17 AP2γ NANOG

D1: A/C

D1:A/C+LDN

D1:BMP+IWP2

D — PGCs – 4 days BMP4

OCT4 expression

AP2γ 21%
Oct4 20%

AP2y expression

PGCs – CHIR/ACT-BMP4

AP2γ 23%
Oct4 16%

AP2y expression
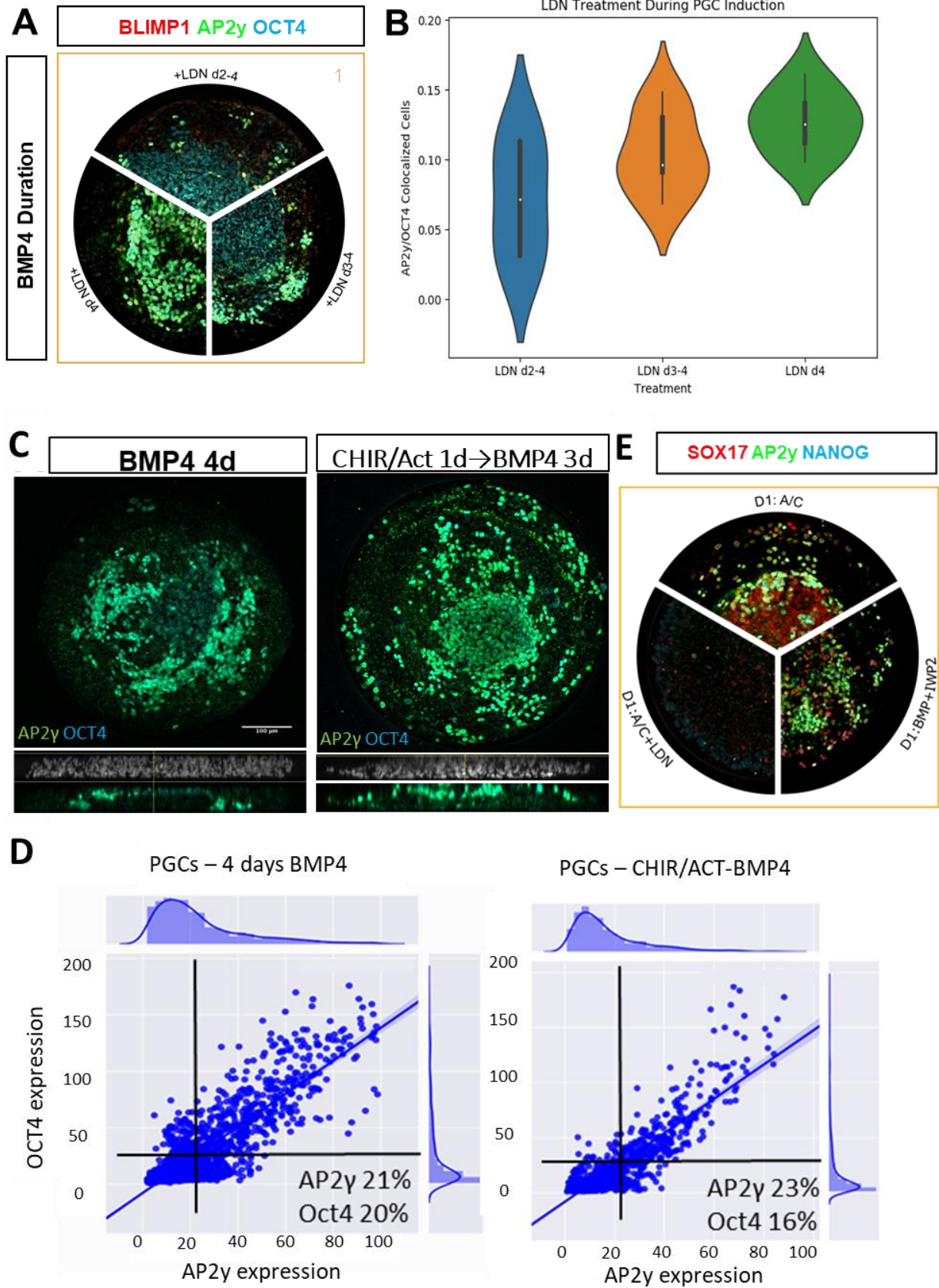
4.3 hPGC specification is not solely dependent on BMP4 gradients

We have shown that BMP4 signaling is both necessary and sufficient for hPGC specification, and that downstream WNT and ACTIVIN signaling (while necessary to generate a permissive niche for BMP4 -directed hPGC specification) are not sufficient for hPGC specification. This establishes an independent input of BMP4 on hPGCs in addition to its role as a cascade initiator.

We therefore investigated the molecular mechanisms underlying BMP4 directed hPGC specification dose and duration dependence. We have previously shown that a reaction-diffusion mechanism involving both BMP4 and its inhibitor NOGGIN is partially responsible for self-organization of gastruloids (*Etoc et al., 2016*). Isogenic NOGGIN null RUES2 gastruloids (RUES2-NOGGIN-/-) no longer self-organize into the three germ layers, and the colonies show aberrant proportions of cell populations. To address the role of NOGGIN in hPGC specification, we stimulated an RUES2-NOGGIN-/- line with BMP4, and find increased specification of both early endoderm (SOX17+/NANOG-) and hPGCs (SOX17+/NANOG+) at 2 days, while hPGC populations fall in line with WT levels by day 4 (**Figure 4.4**). This transient increase is consistent with the loss of an inhibitor leading to more activity for the ligand and highlights a direct role for BMP4 in early hPGC specification. NOGGIN secretion is therefore important to limit early hPGC specification during the first 48 hours, but may not play a role in later hPGC behavior. However, even after abrogation of the BMP4 gradient by elimination of NOGGIN secretion, hPGC specification is still restricted to the edge of the colony. Thus the hPGC niche may be patterned not only by BMP4 gradients across the gastruloid colony alone, but by endogenous signaling patterns downstream of BMP4 as well.

Figure 4.4
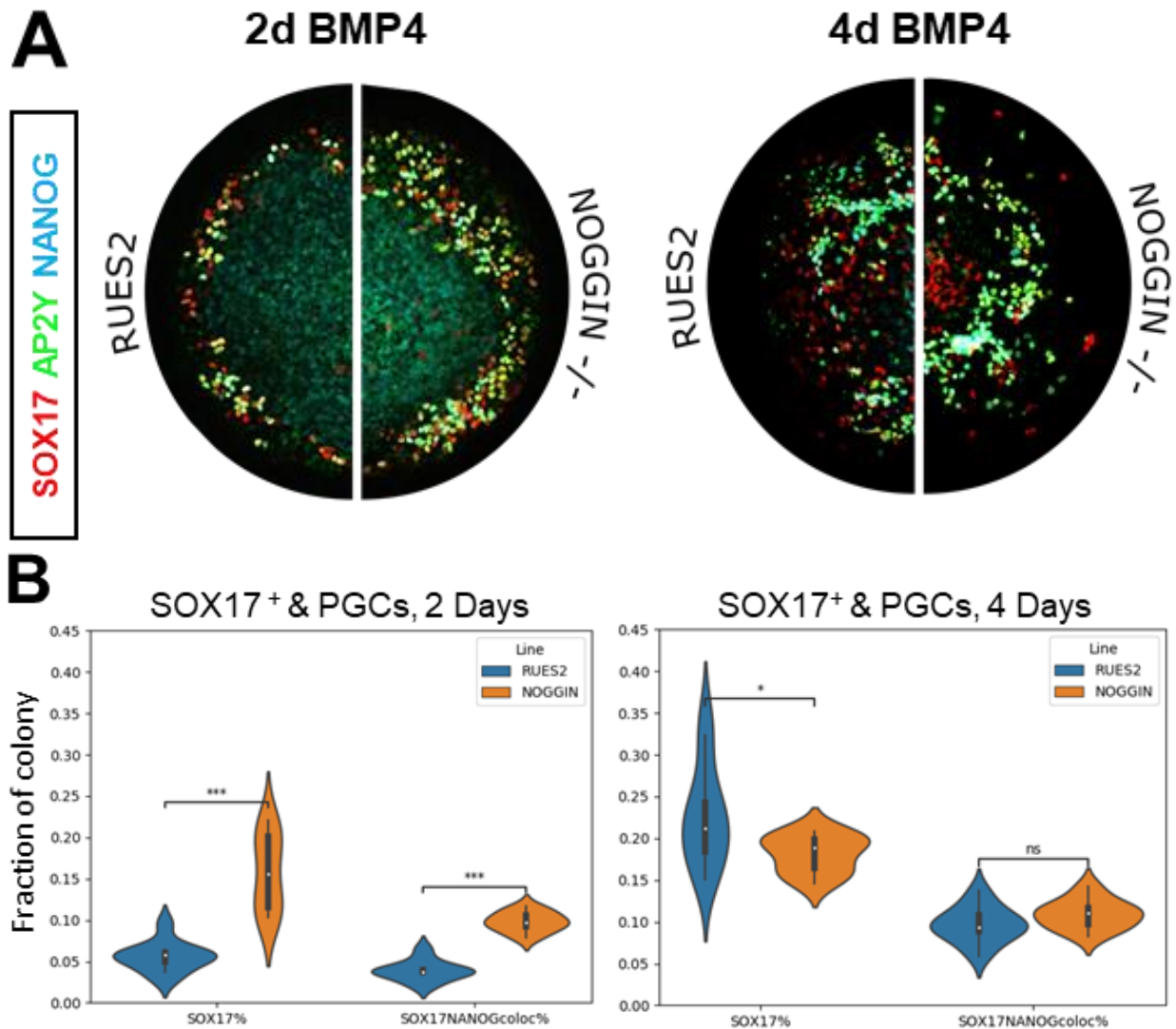
Figure 4.4. hPGC specification is a not solely dependent on BMP4 gradients. *a,* Wildtype RUES2 gastruloids (left half) and NOGGIN-/- gastruloids (right half) were stimulated with BMP4 for 2 days and 4 days. hPGCs were determined through coexpression of SOX17 (red), AP2y (green), and NANOG (blue). *b,* Quantification of these colonies indicated that a significant increase in the fraction of both SOX17+/NANOG- endoderm and SOX17+/NANOG+ hPGCs at two days (left) that returned to control levels by four days of stimulation (Colony diameter = 500um; n=5 colonies/10,000 cells. Mann-Whitney-Wilcoxon test two-sided, *** : p<=1.00e-3, * : p<= 5.00e-2, ns: 5.00e-02 < p <= 1.00e+00).

**4.4 Secreted WNT and ACTIVIN inhibitors pre-pattern an hPGC competent niche**

We next investigated how modulating endogenous WNT and ACTIVIN signaling affected the hPGC niche. Stimulation of confined RUES2 cultures by WNT3A also leads to self-organization that is distinct from that of the BMP4 gradient alone, mediated by a reaction-diffusion mechanism between WNT3A and DKK1 with "pre-patterning" by CDH1 (*Martyn et al., 2019*). WNT induces a wave of mesendodermal fate change beginning at the distal edge of the colony and reaching a steady state when halted by the diffusible, long-range WNT inhibitor DKK1 (CDH1 determines cell sensitivity to the initial WNT stimulus). As WNT signaling is necessary but not sufficient for hPGC specification, we postulated that the width of the WNT gradient might specify a hPGC niche in which BMP4 can induce hPGCs. We aimed to explore the molecular mechanism underlying this gradient response by first examining DKK1-/- and CDH1-/- RUES2 lines (*Martyn et al., 2019*).

RUES2-DKK1-/- and RUES2-CDH1-/- cells were stimulated for 2 or 4 days with BMP4 and evaluated for hPGC specification patterns. Loss of DKK1 in presence of upstream BMP4 signaling (and thus expansion of the WNT gradient) did not result in specification of significantly different numbers of SOX17+ or AP2y+ cells (**Figure 4.5b**), however, while SOX17 endoderm patterning was disrupted, AP2y patterning remained intact (**Figure 4.5a, 4.6**). This suggests that AP2y expression was *directly* a function of BMP4, and though WNT presence was necessary, a

Figure 4.5. hPGC specification in DKK1-/- and CER1-/- backgrounds revealed the effects

of WNT and ACTIVIN signaling gradients on the hPGC specification niche. *a,* Expression

of SOX17 (red), AP2y (green), and NANOG (blue) in wildtype RUES2 gastruloids (12

o'clock), DKK1-/- gastruloids (3 o'clock), and CER1-/- gastruloids (9 o'clock) after 2 days

and 4 days of BMP4 stimulation. *b,* Quantification of gastruloids from each genetic

background indicated a significantly larger population of both SOX17+ and

SOX17+/AP2y+ hPGCs in CER1-/- cells, with no significant change in population size in

DKK1-/- cells (Colony diameter = 500um; n=5 colonies/10,000 cells. Mann-Whitney-

Wilcoxon test two-sided, **** : p<=1.00e-4, ns: 5.00e-02 < p <= 1.00e+00).

Figure 4.5
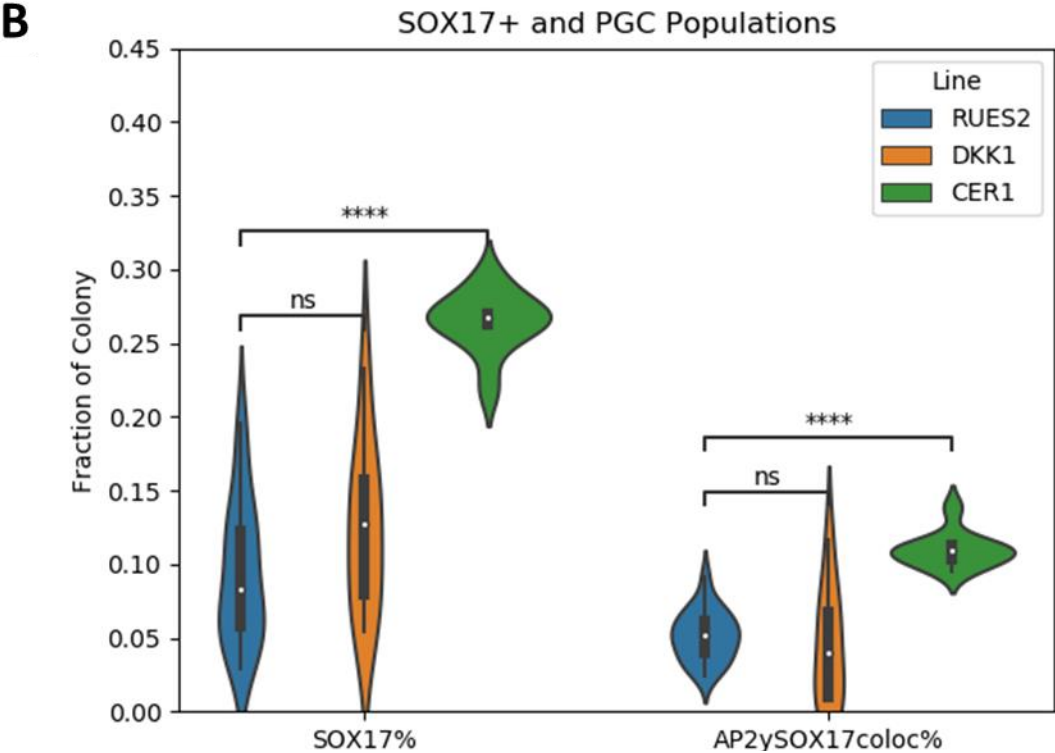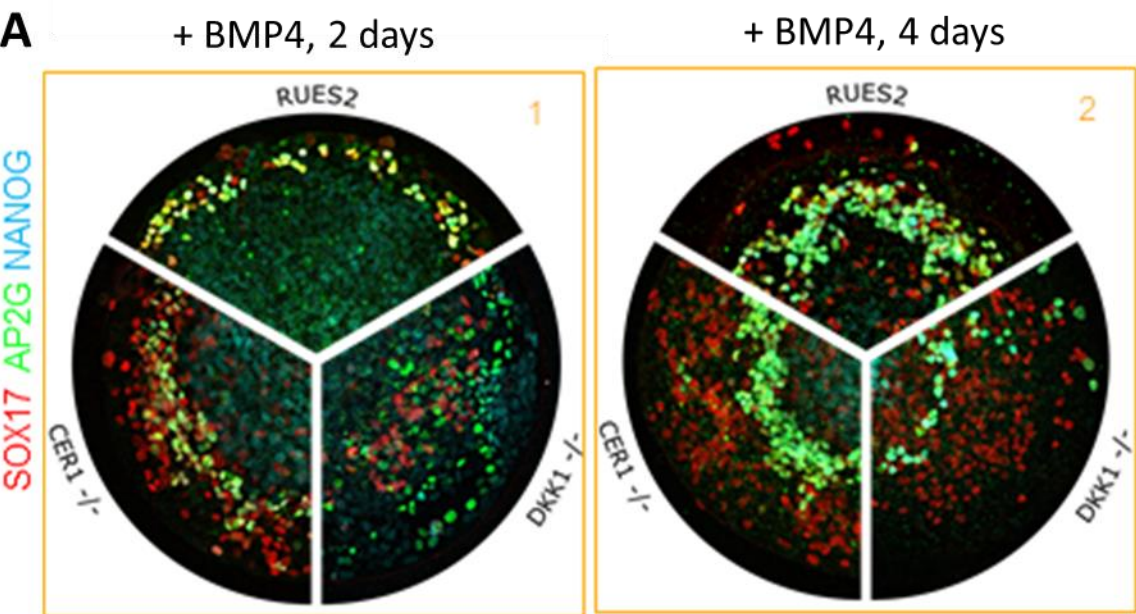
BMP4 threshold must be reached in order for a cell to acquire an hPGC fate. This conclusion

was supported by hPGC patterning in CDH1 -/- cells, in which a steeper WNT gradient results in

a greater number of specified hPGCs, also remaining at the distal edge of colonies

(**Supplemental figure 4.2**). Surprisingly, DKK1-/- colonies also display a marked bifurcation

between AP2y/NANOG and SOX17 expression, with few cells expressing high levels of both

markers (**Figure 4.6b**). While these AP2y$^+$/NANOG$^+$/SOX17$^{(low)}$ cells persist at least until culture

day 4 and display similar behavior as hPGCs in the wildtype case, further analysis is required to

determine if they remain hPGCs, or are transitioning to a different fate. In any case, these

experiments demonstrate that, as opposed to SOX17 expression, expression of hPGC specific

markers AP2y and NANOG is not dependent on WNT concentration.

We also sought to determine the role of ACTIVIN patterning in generating a hPGC niche. We

have previously demonstrated that ACTIVIN signaling in cells that have been previously exposed

to WNT mediates induction of mesendodermal transcription (*Warmflash et al., 2014; Yoney et

al., 2018*). In a micropatterned environment, ACTIVIN stimulation forms a gradient from edge to

center as demonstrated by SMAD2/3 expression, similarly to the cases of BMP4 and WNT

(*Yoney et al., 2018*). When CER1, an inhibitor of ACTIVIN/NODAL but not WNT signaling, is

knocked out, the ACTIVIN gradient is both steeper and extend further from the edge of the

colony (*Martyn et al., 2019*). In a CER1-/- background, mesendodermal cells are biased towards

an endodermal fate, whereas knockout of another secreted WNT inhibitor LEFTY1 did not

change fate decisions (*Martyn et al., 2019*).

In order to investigate whether secreted ACTIVIN inhibitors control the extent of hPGC

patterning, CER1 and LEFTY1 null lines (RUES2-CER1-/- and RUES2-LEFTY1-/-) were used to

disrupt ACTIVIN gradients (*Martyn et al., 2019*). RUES2-CER1-/- showed a proportional increase

in number of SOX17+ cells into proximal SOX17+/AP2y+/NANOG+ hPGCs and distal

SOX17+/AP2y-/NANOG- endodermal fate as well as a more prominent radial segregation of the

two fates (**Figure 4.5,4.6**). Therefore, NODAL was responsible for the width of the hPGC niche,

functioning similarly to a recent report in the mouse (*Senft et al., 2019*), but was not an

instructive cue for the hPGC fate decision. RUES2-LEFTY1-/- colonies displayed the same

phenotype as RUES2-CER1-/- after 2 days of BMP4 stimulation (**Supplemental figure 4.3**).

Interestingly, loss of LEFTY1, and thus enhanced NODAL signaling, led to the loss of both

endodermal and hPGC sub-populations sometime between days 2 and 4, suggesting that there

may be a specific role for LEFTY1 in maintenance of both populations (**Supplemental figure

4.3**).

One benefit of examining the effects of these signaling gradients within a gastruloid system is

that it gives us the ability to define the cell populations that produce these factors and thus

understand how they might arise in the embryo. We have begun analyzing the scRNA-seq

dataset of 96 hour BMP4-stimulated gastruloids generated in **Chapter 3** with this in mind, as

well as collecting preliminary data at earlier timepoints (**Supplemental figure 4.4**). BY the time

hPGCs are specified, they are moving towards the center of the colony alongside FOXA2+

Figure 4.6. Quantification and qualitative model of hPGC signaling niche. *a,* Dotplot representation demonstrating expression of SOX17 (red) and AP2y (green) across five independent gastruloids of each genetic background at 2 days BMP4 stimulation, quantified at the single cell level (n=5 colonies/10,000 cells). *b,* Scatter plots demonstrating correlations between AP2y and SOX17 intensity in each condition from *a.* DKK1-/- gastruloids demonstrated an AP2y+ population that is remarkably uncorrelated to SOX17 expression (center). *c,* Schematics of BMP4, WNT, and ACTIVIN signaling gradients under each of three genetic backgrounds. BMP4 is an instructive signal, as only cells exposed to levels of BMP4 within a defined range became hPGCs. WNT is a permissive signal for hPGCs (and an instructive signal for endoderm), as cells exposed to WNT did not necessarily become hPGCs. ACTIVIN is an instructive signal (for both hPGCs and endoderm) as increasing the ACTIVIN gradient increased both SOX17+/AP2y+ and SOX17+/AP2y- populations, within the niche defined by BMP4 (n=5 colonies/10,000 cells).

Figure 4.6
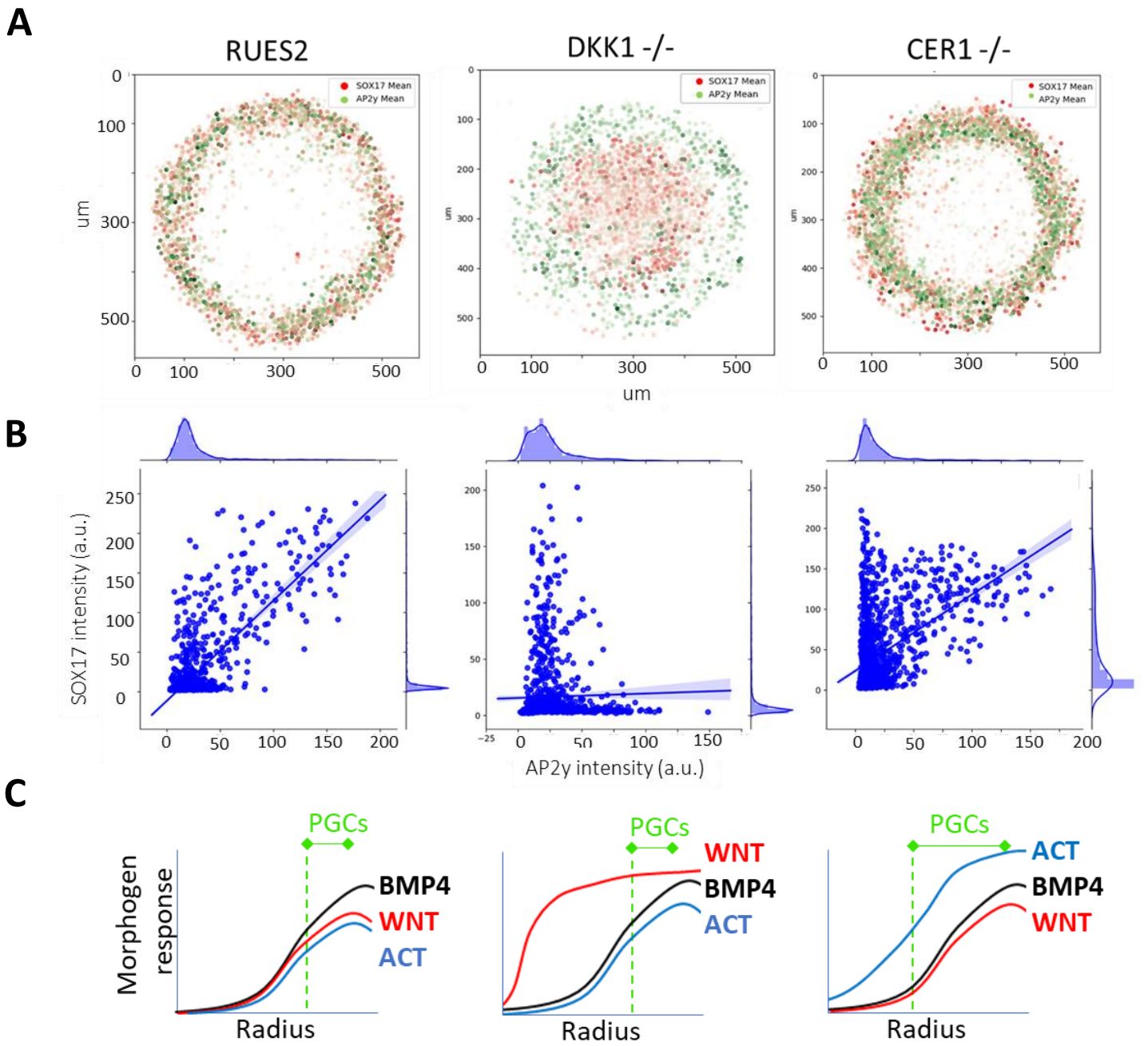
endoderm, with both populations expressing both CXCR4 and KIT (**Supplemental Figure 4.4**).

Endogenous BMP4 as well as the majority of WNT signaling is being produced by FOXF1+ mesodermal cells which remain towards the distal edges of the colony as endodermal and hPGC populations migrate centrally. CXCL12, the ligand for CXCR4, is mainly found in the endoderm while KIT ligand is not highly expressed in any population (**Supplemental figure 4.4**).

Preliminary analysis of morphogens and their inhibitors at earlier time points (24 and 48 hours after initiation of BMP4 treatment) indicates that at 24 hours, endogenous BMP4 production occurs primarily in GATA3+/HAND1+ extraembryonic mesodermal cells, and WNT and NODAL are being produced by BRA+/EOMES+ mesendoderm. At 48 hours, endogenous BMP4 NODAL and WNT are all transcribed in populations of mesoderm, while a SOX17+ population including both endoderm and hPGCs has begun to express CXCR4 and KIT. CER1 and LEFTY are produced by mesodermal populations, while NOGGIN is produced by GATA3+/HAND1+ extraembryonic mesoderm and DKK3 is highly upregulated in GATA3+ extraembryonic ectoderm. Though this dataset is incomplete, it provides some insight into the power of scRNA-seq methods coupled with the single cell spatial resolution of the gastruloid model of the hPGC niche[2]. This data offers the intriguing hypothesis that the gastruloid is a good model not only of gastrulation, but other peri and even post-gastrulation events when cultured for longer time periods.

---

[2] This preliminary scRNA-seq data was generated from a combination of wildtype and Huntington's disease RUES2 gastruloid populations, which for technical failures could not be analyzed independently. Yet they provide a fascinating insight into the complex webs of endogenous signaling downstream of exogenous BMP4 in the hPGC niche. A repetition of this experiment with scRNA-seq library production protocols that have now been well defined will yield an even clearer picture.

4.5 BLIMP1 is sufficient to induce hPGC markers, and itself induces SOX17 in a positive feedback loop

In the previous section, it was determined that the hPGC niche is governed by permissive WNT and NODAL signals downstream of BMP4, and that only cells in this niche that have reached a threshold of direct BMP4 exposure are specified as hPGCs. Following this, it is still unknown at which level in the hPGC transcription factor hierarchy this direct BMP4 input is required (**Figure 4.7**).
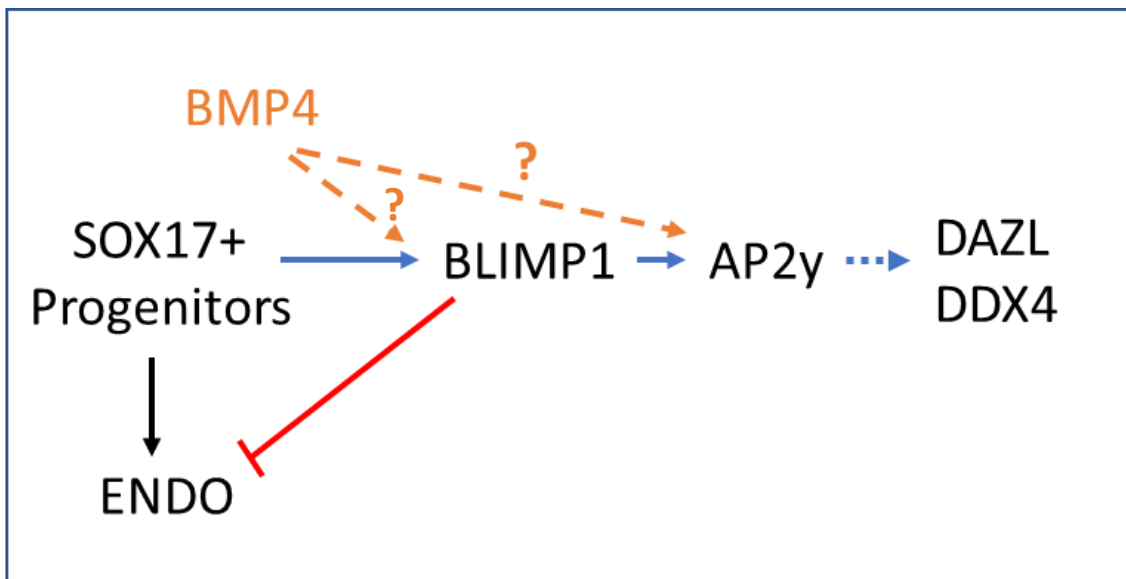
Figure 4.7. *Flow chart demonstrating rationale for knockout and overexpression studies. While we have demonstrated a direct role for BMP4 in PGC specification independent from its role in PGC niche patterning, it remains to be demonstrated whether BMP4 may be acting to induce BLIMP1 expression or whether its action is independent of BLIMP1 as well. If BMP4 acts upstream of BLIMP1, overexpression of BLIMP1 will generate PGCs, while if required downstream, BLIMP1 overexpression alone will be insufficient.*

77

SOX17 has been previously shown to be necessary for hPGC specification, acting upstream of

BLIMP1. SOX17-/- cells (WIS24-SOX17-/-) fail to induce BLIMP1+ cells and other germ cell

markers. BLIMP1 is a canonical regulator of PGC specification and loss of function

downregulates germ cell markers including NANOS3, NANOG, OCT4 and AP2y while

upregulating mesendodermal genes (*Irie et al., 2015*). In order to determine if the same is true

for our gastruloid-derived hPGCs, WIS24-SOX7 -/- and WIS24-BLIMP1 -/- cells were cultured

according to the protocols established using RUES2 cells above. Examination of hPGC and germ

layer ratios showed that SOX17-/- cells generate a BRA+ population that is shifted distally

compared with WT, and in line with previous observations BLIMP1 expression and PGCs

specification is lost (**Figure 4.8**). BLIMP1-/- cells express SOX17+ and BRA+ populations.
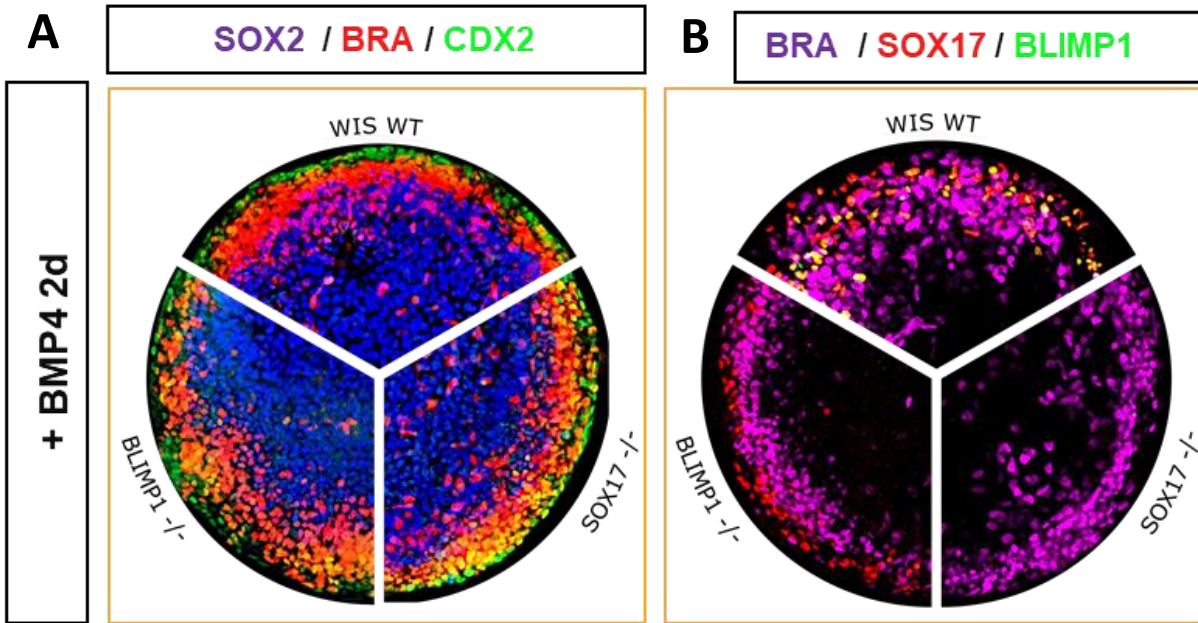
Figure 4.8

*Figure 4.8. SOX17 is required upstream of BLIMP1 in human PGC specification. **a**,*

Expression of SOX2 (blue), BRA (red), and CDX2 (green) in 48 hour gastruloids

demonstrates normal radial acquisition of fates in wildtype (12 o'clock), SOX17-/- (3

*o'clock), and BLIMP1-/- (9 o'clock) genetic backgrounds. **b**, All three genetic backgrounds*

are capable of producing normal BRA+ (purple) populations, while SOX17-/- cells express

neither SOX17 (red) nor BLIMP1 (green) and BLIMP1-/- cells produce a SOX17+

*population but no SOX17+/BLIMP1+ population.*

Given the prominent role for BLIMP1 in regulating hPGC specification, we wondered whether

directly overexpressing BLIMP1 would be sufficient for specifying hPGC fate in the absence of

external signals. We explored the epistatic relationship between SOX17 and BLIMP1 by

performing a gain-of-function experiment. V5-tagged-BLIMP1 (BLIMP1-V5) was introduced via

ePiggyBac into both wildtype and SOX17 -/- background. In the SOX17 -/- background, overexpression of BLIMP1-V5 did not result in expression of AP2y, and the cells maintained their state of pluripotency as marked by strong OCT4 expression (**Figure 4.9a**).

BLIMP1 overexpression in wildtype hESCs, however, surprisingly induced expression of SOX17 and downregulated SOX2 in a clonal population of widely spaced cells (minimizing secreted factors) (**Figure 4.9b**). OCT4 and AP2y expression are covariant with BLIMP1-V5 in a mixed population of clones expressing varying levels of BLIMP1-V5 (**Figure 4.9b**). Single cell quantification demonstrates that every BLIMP1-V5 expressing cell also expresses SOX17, AP2y, and OCT4 (**Figure 4.9b**). This signature is unique to human PGCs and demonstrates that inductive effects of BMP4 stimulation may be recapitulated by BLIMP1 overexpression alone. In order to demonstrate this definitively, this BLIMP1 overexpression experiment must be repeated in a BMP4-/- genetic background to account for any effects from endogenous signaling.

Figure 4.9. BLIMP1 expression is sufficient to induce both SOX17 and later PGC markers.

a, SOX17-/- cells transfected with an ePB transposon containing a dox-inducible BLIMP1-V5 overexpression cassette were stimulated with dox for 2 days. Expression of DAPI (grey), BLIMP1 (red), SOX17 (purple), OCT4 (blue), and AP2y (green) indicated that these cells remain pluripotent. *b,* Dox-induced expression of BLIMP1-V5 (red) in non-clonal RUES2 cells demonstrate selective downregulation of SOX2 (blue) and upregulation of SOX7 (green), even in sparse culture. Dox-induced clonal cells demonstrated expression of OCT4 (blue) and AP2y (green) that correlated with expression of BLIMP1 (red, bottom row). *c,* Quantification of *b* indicated significant positive correlation between markers, with Mander's coefficient of correlation of approximately 0.8 for expression of V5-tagged BLIMP1 within pixels expressing SOX17, AP2y, and OCT4 (represented by "V5/x" column) and slightly lower values between 0.4 -0.6 for the inverse coefficient (the number of V5+ cells expressing other markers, represented by "x/V5"). "Mixed" refers to a non-clonal culture that may contain cells with varying transposon insertion, while "Picked" experiments were performed in clonal cultures containing cells with exogenous BLMIP1-V5 expression approximately equal to endogenous BLIMP1 expression levels in PGCs by IF. No significant difference was seen between "mixed" and "picked" cultures. (n=5, bars indicate 95% CI *d,* Expression of DAZL (green), DDX4 (red) , and OCT4 (blue) in RUES2-BLIMP1-V5 cells induced with doxycycline for four days. (All scale bars: 100um)

Figure 4.9
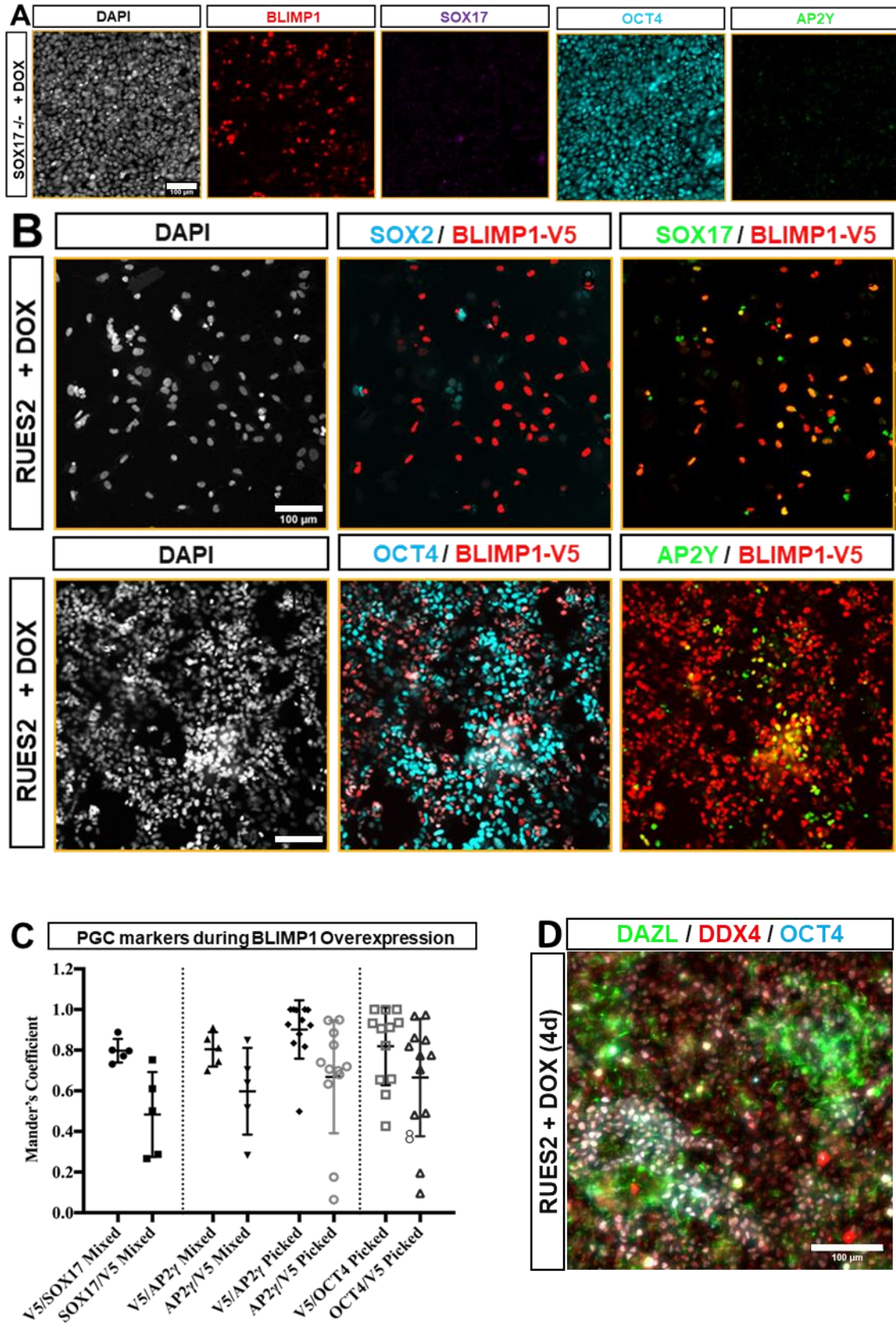
In addition to demonstrating a feedback loop upregulating SOX17, upregulating AP2y, and downregulating SOX2, we sought to determine whether BLIMP1 overexpression could induce expression of later hPGC markers. To our surprise, after four days of BLIMP1 overexpression, some patches of cells began to upregulate DAZL and downregulate OCT4, similar to hPGCs adopting a gonadal fate *in vivo* (**Figure 4.9d**; *Anderson 2007*). This data indicated that cells expressing hPGC markers derived from BLIMP1 overexpression have the capacity to develop later germ cell characteristics.
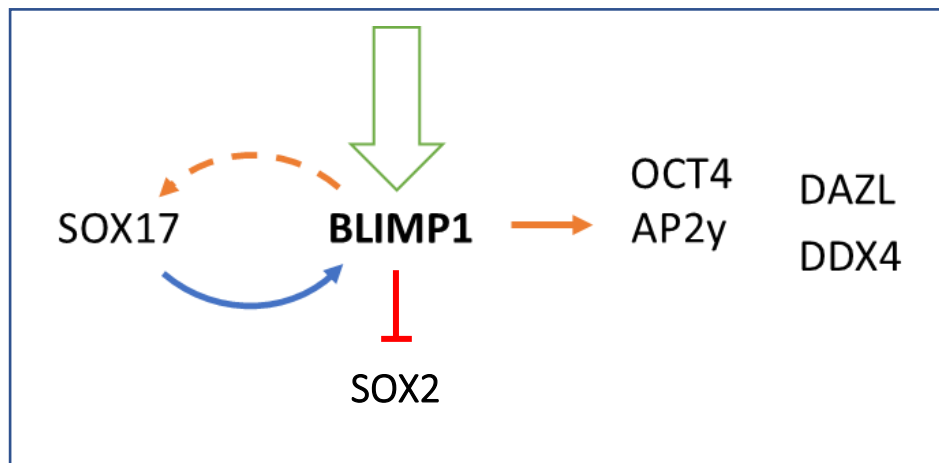
Figure 4.10. *Overview of BLMIP1 overexpression experiments. Direct expression of BLIMP1-V5 via a dox-inducible ePB cassette allowed us to investigate hPGC specification in the absence of exogenous signaling factors (green arrow). BLIMP1 expression is sufficient for cells in this experiment to downregulate SOX2, upregulate SOX17 (orange dotted arrow), and express OCT4 and AP2y (solid orange), as well as DAZL and DDX4. This indicates that BLIMP1 may be the effector through which exogenous BMP4 functions (though it does not preclude the possibility that endogenous BMP4 is playing a role).*

Recently, Kobayashi et al. (2017) also investigated the results of BLIMP1 overexpression in SOX17-/- cells either alone or in combination with SOX17 overexpression. In their system, only overexpression of both SOX17 and BLIMP1 in SOX17-/- cells resulted in expression of AP2y and NANOS3. Our results are in agreement, in that BLIMP1 alone was unable to induce hPGC markers in SOX17-/- cells. However, we additionally demonstrated that BLIMP1 alone is sufficient to downregulate SOX2, induce AP2y and SOX17, and moreover induce later hPGC markers DAZL and DDX4 in the absence of other exogenous stimuli (**Figure 4.10**).

In all, we demonstrated a requirement for BMP4, WNT and ACTIVIN in hPGC specification, and demonstrated specific roles for BMP4 and ACTIVIN gradients in the generation of the hPGC niche, while WNT signaling seems to be permissive but not instructive. We also demonstrated that overexpression of BLIMP1 downregulates SOX2 and induces expression of PGC markers – including SOX17; thus, recapitulating to some extent the effects of this signaling cascade. Though **Figure 4.2** indicates that exogenous WNT signaling induced SOX17+/BLIMP1+ cells that do not express other markers of hPGCs, this additional data indicated that in the absence of exogenous WNT, BLIMP1 may be sufficient to push cells towards an hPGC fate.

While examining a model of the self-organizing embryo has given insight into hPGC specification with spatial and temporal precision, *in vivo* validation is necessary to ensure that our models truly represent cell behaviors relevant to primate embryogenesis. The weaknesses of our model rest on the assumptions that it contains all cell populations present in the in vivo hPGC niche, and that differences in timing and spatial arrangement between the gastruloid and

the 3D embryo self-organization are not consequential. In order to provide a system in which

these assumptions may be tested, we develop *in vivo* imaging of the marmoset embryo in

Chapter 5.

Supplemental Figures

Supplemental Figure 4.1
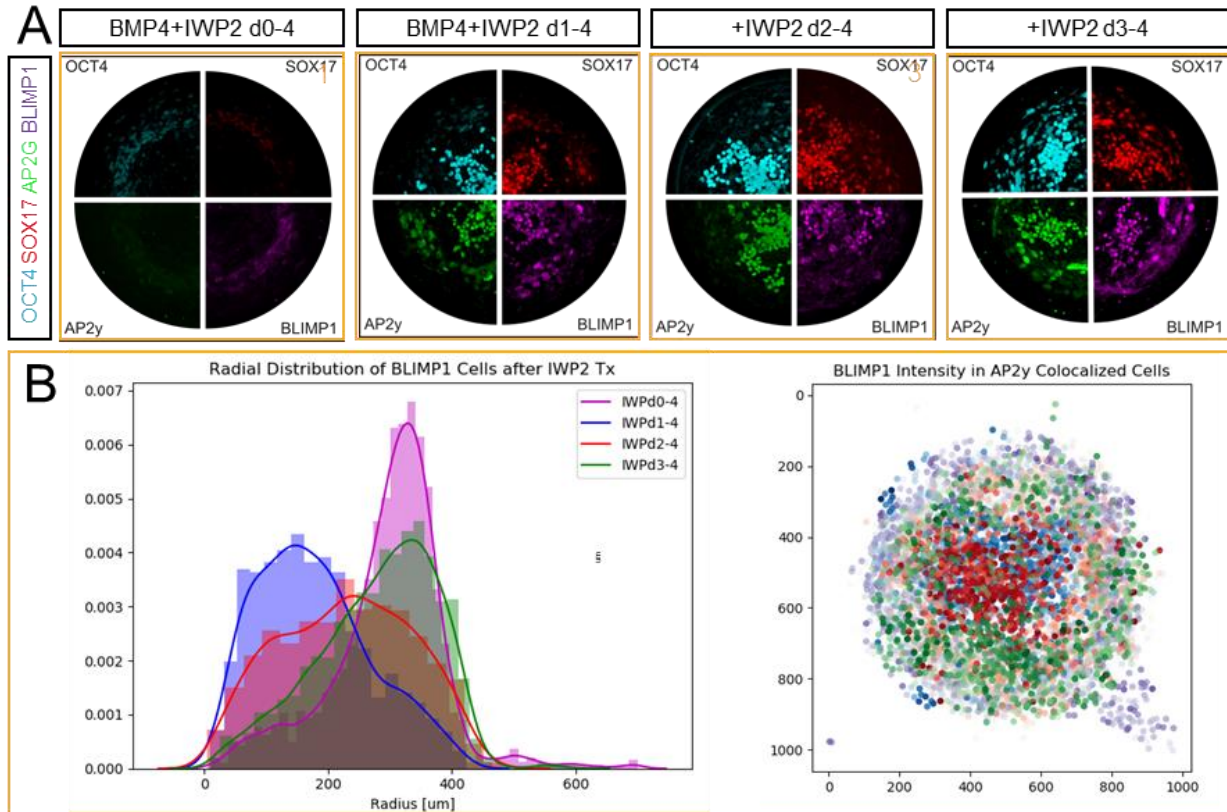
Supplemental Figure 4.1. *WNT signaling is necessary within the first 24 hours of treatment for hPGC specification. **a,** Expression of OCT4 (blue), SOX17 (red), AO2y (green) and BLMIP1 (purple) in gastruloids treated with 96 hours of BMP4 (50 ng/ml) with IWP2 to inhibit WNT secretion either (i) immediately, (ii) after 24 hours, (iii) after 48 hours, or (iv) after 72 hours. hPGCs only fail to specify when WNT secretion is blocked for the entire duration of the experiment. Radial intensity (**b**) and dotplot representations (**c**) of single cells expressing BLIMP1 in the conditions in **a** indicate that though population size is the same for conditions ii through iv, hPGCs do not migrate centrally to the same extent when WNT signaling is allowed during the first 72 hours of treatment. Thus, WNT may play a role in early hPGC migration after the patterning of the hPGC niche. Colony diameter: 500um.*

Supplemental Figure 4.2. *Disruption of WNT wave decouples endoderm from hPGC fates.* ***a,***

Expression of SOX17 (red), AP2y (green), and NANOG (blue) in representative gastruloids with

wildtype RUES2, DKK1-/-, CDH1-/-, and CDH1-/-DKK1-/- backgrounds at 2 days of BMP4

stimulation (left) and 4 days (right). In all cases, patterning of the hPGC niche resembles that of

wildtype, though the SOX17+/AP2y-/NANOG- endoderm niche is expanded in DKK -/- cells. ***b,***

Quantification indicates that knockout of CDH1 increases the size of the total SOX17+ niche,

while knockout of DKK1 does not change population sizes (n=5 colonies/10,000 cells). ***c,*** *Dotplot*

analysis of SOX17 and AP2y expression demonstrates the patterning of endoderm and hPGC

competent niches (n=5 colonies/10,000 cells). Only in the case of DKK1-/- is there a decoupling

of cells expressing AP2y and SOX17, indicating that DKK1 is required for cells to coexpress both

SOX17 and AP2y. (Colony diameter: 500um. Mann-Whitney-Wilcoxon test two-sided; ns: 5.00e-

02 < p <= 1.00e+00; *: 1.00e-02 < p <= 5.00e-02; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 < p

<= 1.00e-03; ****: p <= 1.00e-04).

Supplemental Figure 4.2

A | 2d BMP4 | 4d BMP4

SOX17 AP2G NANOG

RUES2   CDH1 -/-   1
CDH1 -/- DKK1 -/-   DKK1 -/-

RUES2   CDH1 -/-   2
CDH1 -/- DKK1 -/-   DKK1 -/-

B   SOX17⁺ & PGCs after WNT wave Disruption

C   RUES2   CDH1-/-

DKK1-/-   CDH1-/- DKK1-/-

88
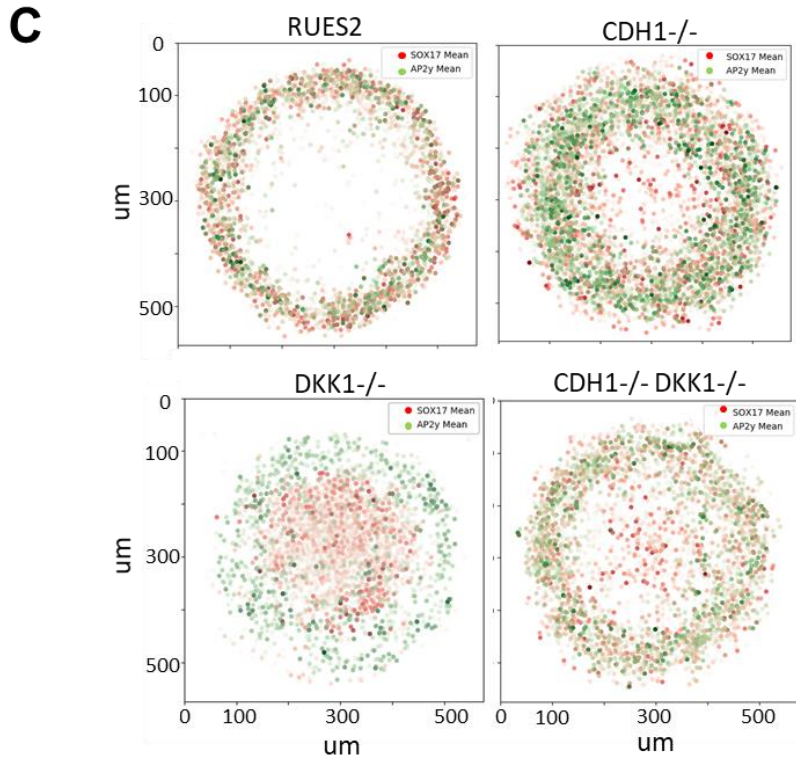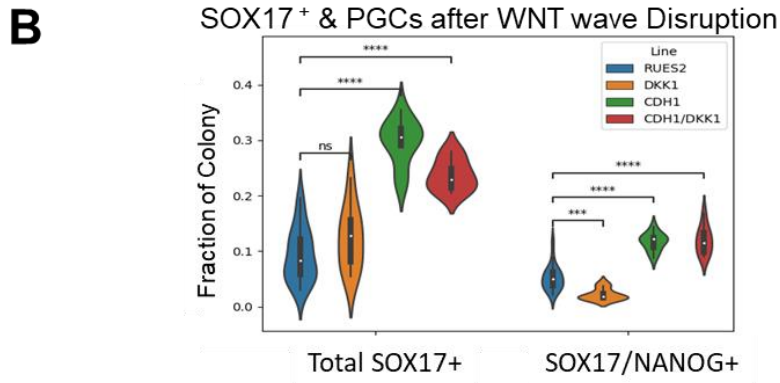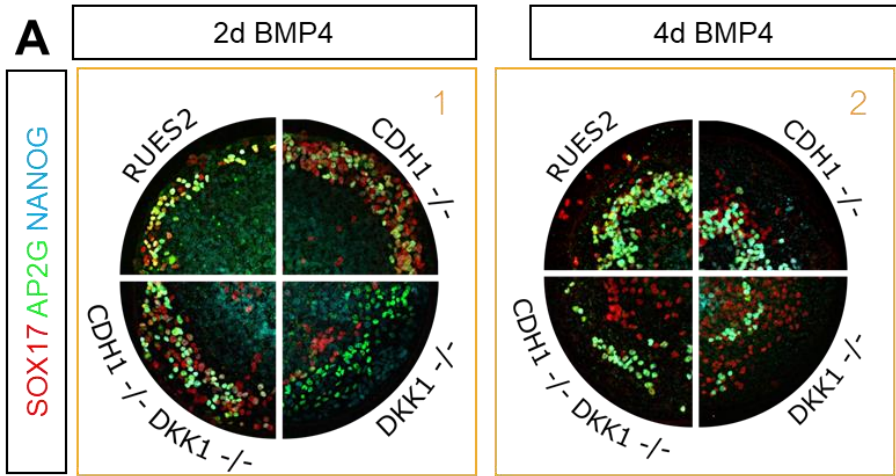
Supplemental Figure 4.3. *Knockout of endogenous ACTIVIN inhibitors increases the hPGC competent population. **a,** Expression of SOX17 (red), AP2y (green), and NANOG (blue) in representative gastruloids with wildtype RUES2, CER1-/-, LEFTY-/-, and CER1-/-LEFTY-/- backgrounds at 2 days of BMP4 stimulation (left) and 4 days (right). In all cases, patterning of the hPGC niche resembles that of wildtype at two days. After four days, however, SOX7+/AP2y+ hPGCs have migrated centrally, while SOX17+/AP2y- endoderm has not, demonstrating that early hPGC migration is not dependent on endoderm migration. LEFTY-/- completely abolishes both populations at by day 4, while the double knockout rescues this phenotype, indicating that CER1 and LEFTY may play opposing roles in SOX17+ population maintenance after early specification. **b,** Quantification indicates that knockout of both CER1 and LEFTY increases the size of the total SOX17+ niche including hPGCs (n=5 colonies/10,000 cells). **c,** Dotplot analysis of SOX17 and AP2y expression in these colonies demonstrates the patterning of endoderm and hPGC competent niches. (Colony diameter: 500um. Mann-Whitney-Wilcoxon test two-sided; ns: 5.00e-02 < p <= 1.00e+00; *: 1.00e-02 < p <= 5.00e-02; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 < p <= 1.00e-03; ****: p <= 1.00e-04).*

Supplemental Figure 4.3

A

| 2d BMP4 | 4d BMP4 |

SOX17 AP2y NANOG

RUES2 / CER1 -/-
CER1 -/- LEFTY -/- / LEFTY -/-

B

SOX17 $^+$ & PGCs after ACTIVIN/NODAL disruption

Line
- RUES2
- CER1
- LEFTY
- LEFTY/CER1

Fraction of Colony

Total SOX17+   SOX17/NANOG+

C

RUES2   CER1-/-

SOX17 Mean
AP2y Mean

LEFTY-/-   LEFTY-/- CER1 -/-

um   um

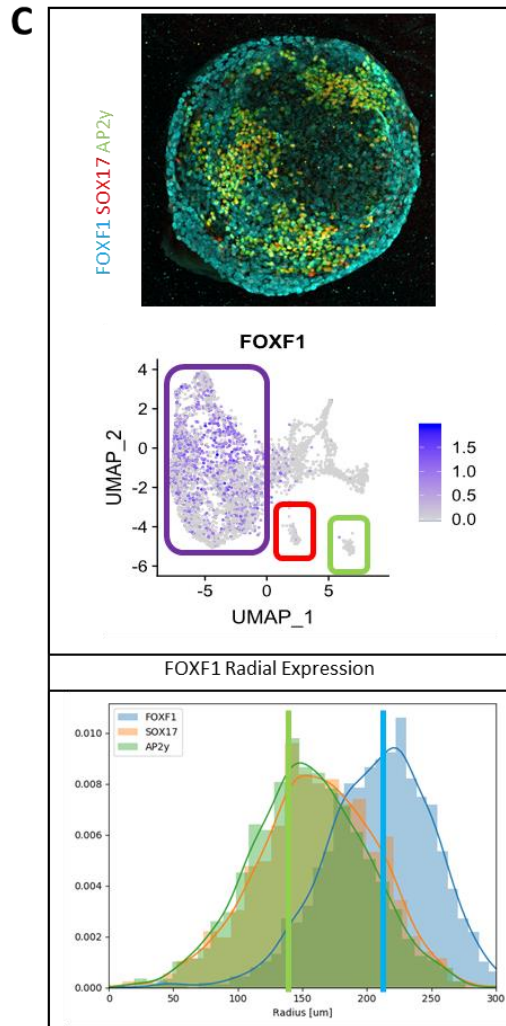um   um
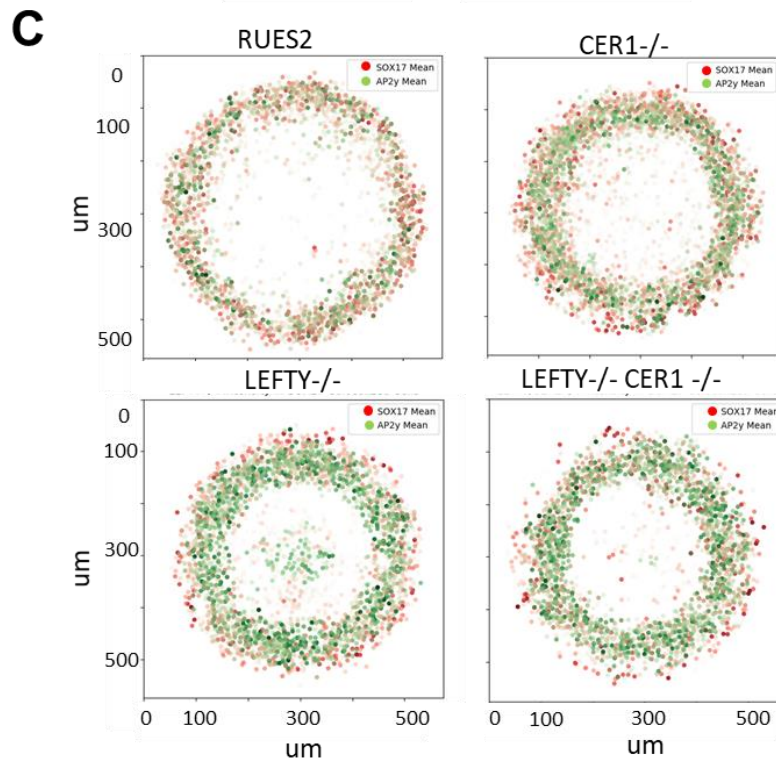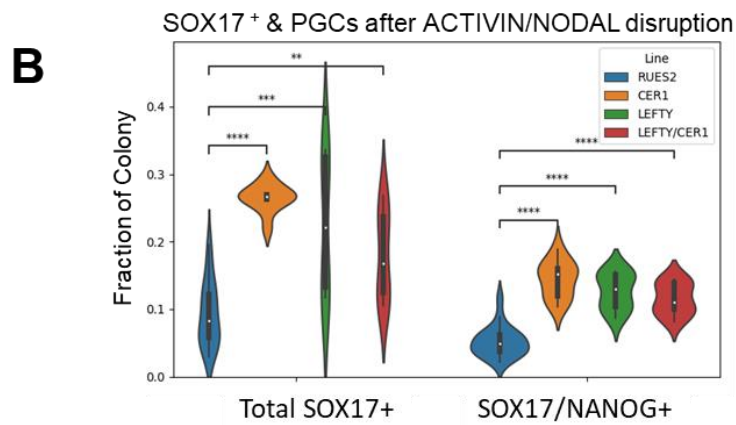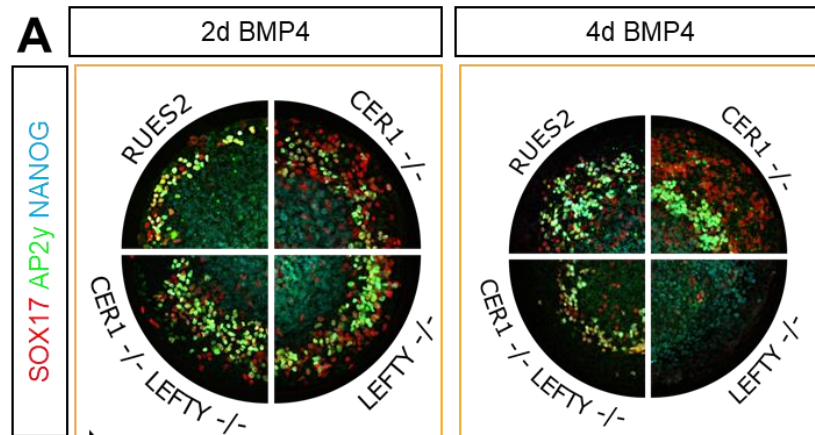
Supplemental Figure 4.4. *Analysis of morphogens present after 96 hours of BMP4 stimulation.*

a, Dotplot analysis of selected signaling factors, receptors, and inhibitors investigated within this chapter in clusters from the gastruloid scRNA-seq database generated in Chapter 3. Size of dots indicates the fraction of cells within a given cluster that express the gene, while color indicates expression level (blue/low to red/high). Clusters are grouped according to their germ layer origin. b, Expression of FOXA2 (purple), SOX17 (red), and AP2y (green) 96 hour BMP4 derived gastruloid validate FOXA2 as a specific marker of definitive endoderm (top). Expression of FOXA2 is solely within the endoderm cluster (red circle) in a UMAP representation of the scRNA-seq dataset (middle), while single cell analysis of immunofluorescence indicates that expression of AP2y and FOXA2 indicate two separate populations. c, Expression of FOXF1 (blue), SOX17 (red), and AP2y (green) within 96 hour gastruloids indicates that FOXF1+ cells lie distal to endoderm and hPGCs after 4 days (top, radial quantification at bottom). Expression of FOXF1 within a UMAP representation of scRNA-seq data indicates that at this time, FOXF1 is a pan-mesodermal marker (middle) and thus the origin of most endogenous BMP4, NODAL, WNT, and DKK1 is at the distal aspect of the gastruloid after 96 hours. This strategy can also be used at earlier timepoints to specify the spatial origins of endogenous signaling during hPGC specification. (Colony diameter: 500um)

Supplemental Figure 4.4

A — Morphogen Expression by Cluster

B

FOXA2 SOX17 AP2y

FOXA2

AP2y/FOXA2 Colocalization

C

FOXF1 SOX17 AP2y

FOXF1

FOXF1 Radial Expression
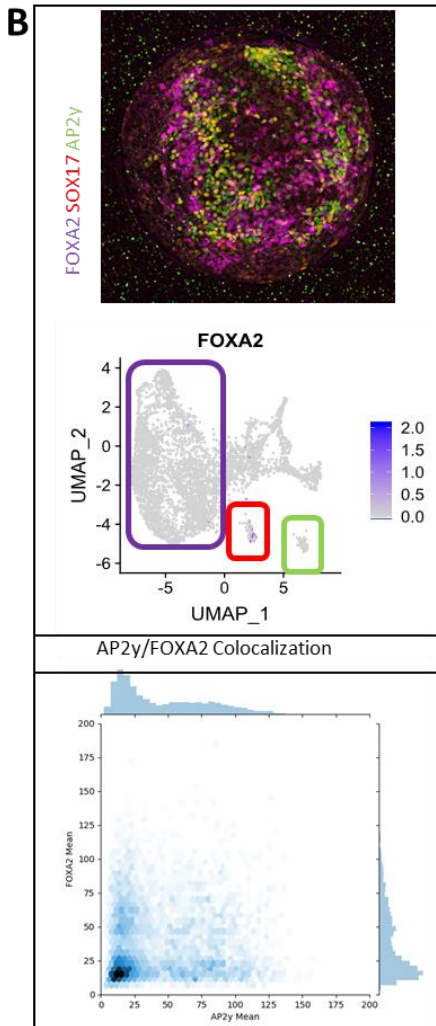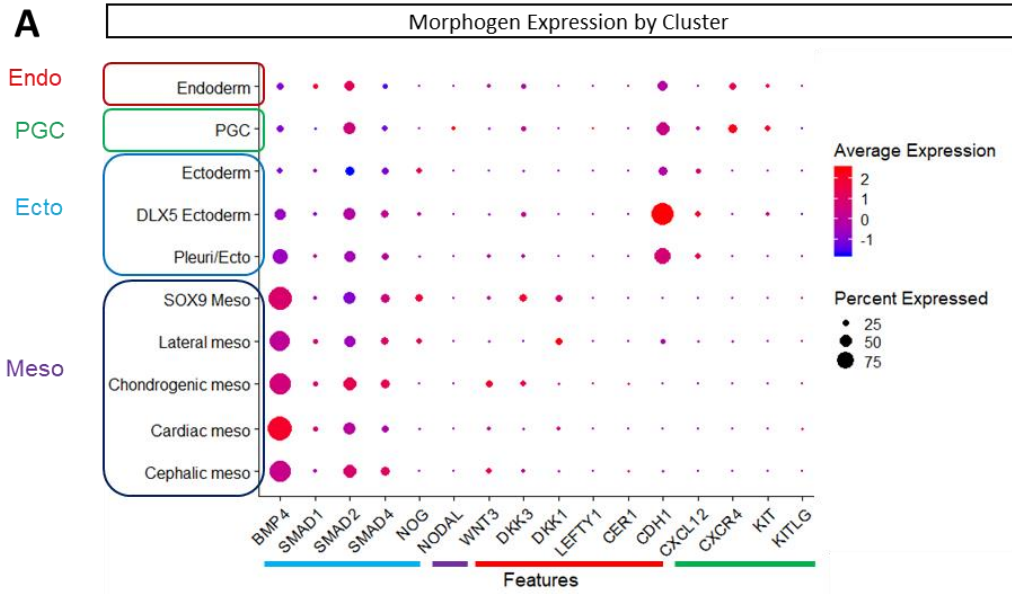
CHAPTER 5

A marmoset model of *in vivo* primate embryogenesis: Interactive, high-resolution, and accessible ultrasound atlas of embryonic and fetal development of the common marmoset

Recently, several groups have demonstrated that both NHP embryos and porcine embryos may faithfully represent elements of human PGC development (*Aeckerle et al., 2014; Sasaki et al., 2016; Kobayashi et al., 2017*). In order to define such a model in our own studies, not only of hPGC specification but also of human embryogenesis as a whole, we began to investigate marmoset embryogenesis with the goal of defining embryonic staging and establishing temporal windows for truly *in vivo* transgenic modeling of early embryonic events.

High-resolution, serial ultrasound scans of the common marmoset (*Callithrix jacchus*) were used to track 7 dams through 13 pregnancies for a total of 34 fetuses. After confirming pregnancy, each fetus was tracked with biweekly scans for the first half of the pregnancy and weekly thereafter for an average of 19 ultrasound sessions per fetus. The high frequency of serial imaging, particularly during the early stages of pregnancy just after implantation, unveiled the intricacies of a delayed development period unique to the common marmoset. Subsequent scans feature the development of major organs, with a specific focus on heart and brain development. Established obstetric and embryologic measurements were performed for each scan and used for comparison with the standardized Carnegie staging of human development and for the first time provides ultrasound correlates to *ex vivo* staging. This work establishes

ultrasound as an accessible and viable means of noninvasive monitoring for embryonic and fetal marmoset development. All scans were used to generate an interactive ultrasound atlas available for the monitoring of pregnancies globally and across many colonies. Finally, by providing an unprecedented temporal and spatial resolution, our study will inevitably improve primate embryonic description and staging, including that of our own development.

Here we present the high spatial- and temporal- resolution, serial ultrasound atlas for embryonic and fetal marmoset development, that verifies ultrasound as an accessible technique, reveals insights into the period of delayed development as well key features of organogenesis, and offers functionality to be queried by individuals to stage their own marmoset ultrasound images.

5.1 Serial scans reveal early embryonic development with protracted gastrulation

Prior to this study, the earliest insights into the earliest stages of marmoset development were accessible only with *ex vivo* histological studies (*Moore et al., 1985; Phillips, 1976; Smith et al., 1987)*. These studies revealed an intriguing developmental delay which was further confirmed with later stage ultrasound imaging *(Jaquish et al., 1995*). Utilizing the standard of Carnegie Stages as a common reference, it was noted that marmoset embryos take many more days to progress through the earliest stages (*Jaquish et al., 1995; Phillips, 1976*) Their observed measurements started later in pregnancy on days 33 and 24, respectively. The cause for this delay remains unknown. It was previously hypothesized that this delayed development in the common marmoset could not be termed diapause due to the lack of delay dependence on

neither seasonal nor lactational quiescence as is evident in most other mammals that exhibit diapause (*Renfree & Fenelon, 2017*). It was instead suggested that the delay was simply a matter of slower growth as a result metabolic demand for the twinned blastocysts to develop a fused chorion (*Phillips, 1976*).

In this study, we were able to capture our earliest measures at estimated day 12 post fertilization (**Figure 5.1c**). Pooling scans from all our subjects, we were able to capture 89% of the days of the pregnancy from day 12 to birth, defined as day 143. Using this serial, high resolution imaging, we were able to stage the ultrasounds using the Carnegie Staging method bridging the gap between histological and ultrasound measurements (**Figure 5.1**). In order to identify the exact stage in which the developmental delay occurs in marmosets compared to humans, **Figure 5.1a** shows the first day during which each Carnegie Stage was entered. Given that marmoset gestation is approximately half that of human gestation, we expected marmosets to enter each stage of embryonic development earlier than humans. Instead, not only is there a marked delay, but this delay is present during a few distinct stages of development. Plotting the percent of embryonic development spent in each stage (**Figure 5.1b**) revelaled that there is a single stage, Stage VII, in which significantly more time is spent during marmoset development than human as measured by a two-proportion z-test (p = 0.0038) (**Figure 5.1c**). **Figure 5.1c** depicts representative ultrasounds from notable Carnegie stages with features outlined. Individual ultrasounds for every day captured of embryonic development can be found in **Supplemental Figure 5.1** and **5.2** for pre and post neural tube closure, respectively.

Figure 5.1
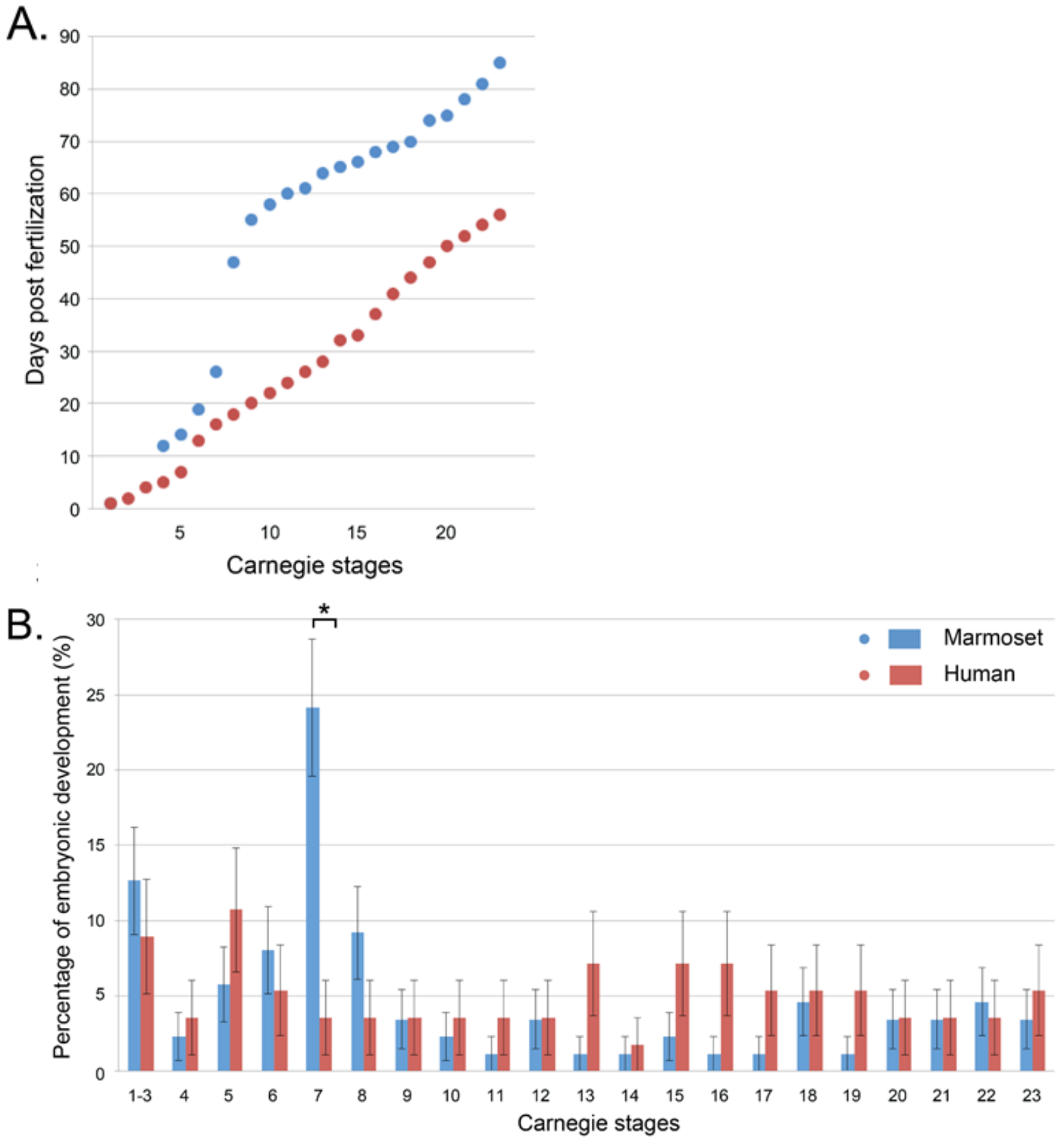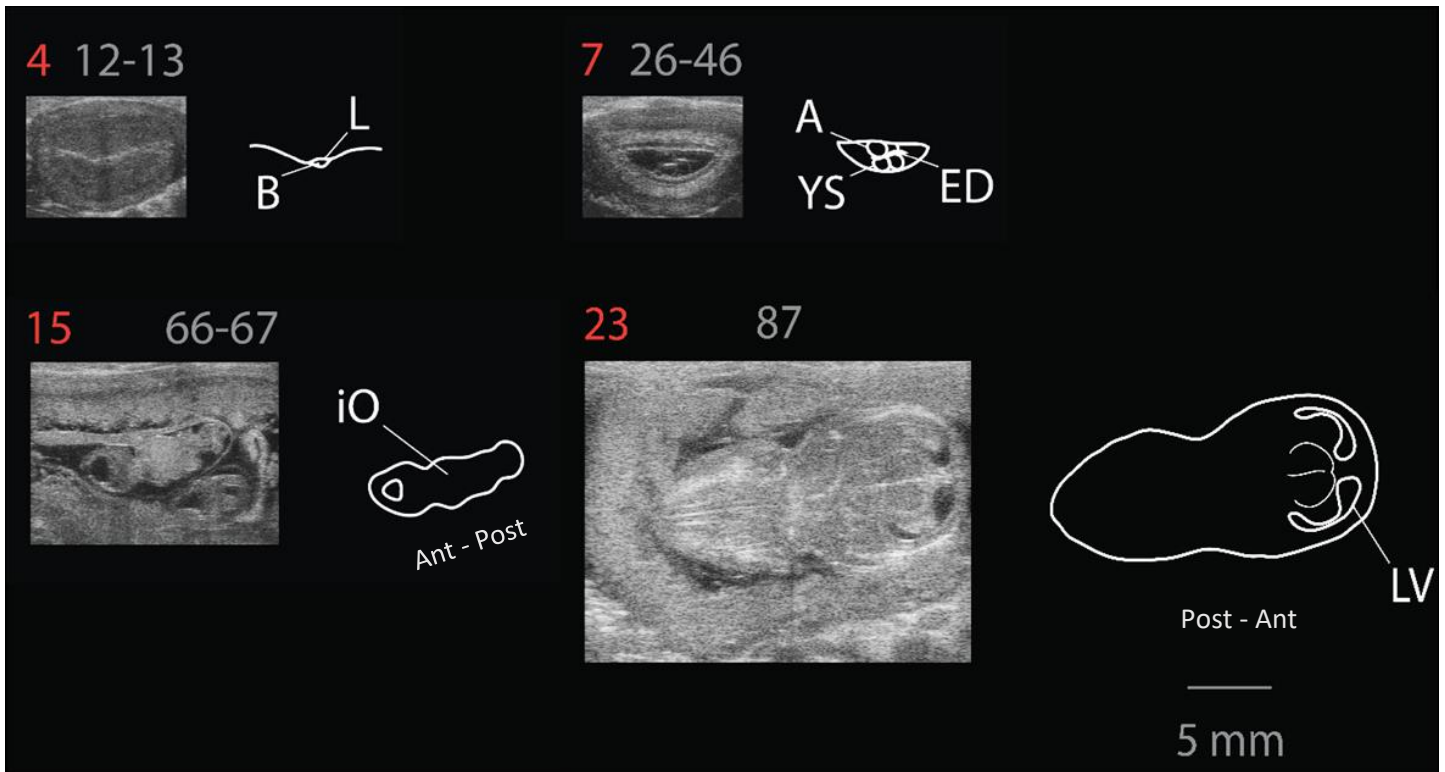
Figure 5.1. *Early ultrasound measurements reveal development through Carnegie Stages. **a,** The first day post fertilization during which the developing embryo enters into each Carnegie Stage (blue, marmoset; red, human). **b,** The percentage of the embryonic development (estimated to be 87 days marmoset, 56 days human) spent in each Carnegie Stage. **c,** Representative ultrasound images taken for each Carnegie Stage. Stage 4 is the earliest post-fertilization imaging we were able to acquire, stage 7 is the stage of longest marmoset dwell time, stages 15 and 23 (post-neural tube closure) are shown for comparison. Outlines in white detail relevant features in ultrasound. Yellow highlights key areas or orientations in figures. Red numbers indicate Carnegie Stage. Ant – Post indicates anterior-posterior axis. Gray numbers indicate range of days per stage. A, amnion; B, blastocyst; E, embryo; ED, embryonic disk; I, implantation; iO, internal organs; L, lumen; V, ventricle; YS, yolk sac.*

To further interrogate the patterns of marmoset early embryonic growth and development,

measurements of the embryonic disk (which correspond to the crown-rump length (CRL) after

day 50) and biparietal distance were first taken across all 35 fetuses (**Figure 5.2b**). The initial

segment in **Figure 5.2b** revealed a period in which the CRL changes little over nearly a 40-day

period, after which a rapid period of stable growth began close to the first discernable

heartbeat, at day 57 (**Figure 5.2a**). However, when individual ultrasound scans were examined,

it is clear that the embryos and uterus were not in a quiescent state even during these first 40

days, as evidenced by the dramatic extraembryonic growth and restructuring taking place

during this time. Measures of maximum chorionic cavity cross-sectional area and yolk sac

diameter were taken during days 14-62 (**Figure 5.2c,d**), confirming extraembryonic growth and

development. Diapause is strictly defined as a quiescent period of no growth. This period can

therefore not be considered diapause given the dramatic restructuring of the amnion. Shown in

Figure 5.2E are the outlines of the lumen, embryo, yolk sac, amnion. The original ultrasounds

from which these outlines were taken can be found in **Supplemental Figure 5.1**.

5.2 *Standard Embryonic Measures as predictors of fertilization age*

We next sought to determine which of these standard measures, if any, are reliable predictors

of fertilization age in isolation. Data from each measure was fit to a polynomial linear

regression (**Supp. Fig. 5.2a**). All measures prove useful predictors with 95% confidence intervals

of approximately 10 days, with chorionic cavity area as the best predictor of early embryonic age (days 13-53) and BPD the best overall predictor (days 87-birth) (Models summarized in **Table 5.1**). Previous studies suggest that BPD measurements diverge as an indicator of postnatal survival (*Tardif et al., 1998)*. We see no such separation within our BPD measures. In fact, due to the lower variance of measures of BPD, we find them to be a better predictor of gestational age and due date.

More excitingly, the evidence of this delayed growth period further reveals an opportunity to interrogate the molecular mechanism. Given that there is nearly a two week window during which gastrulation occurs in Stage VII (**Figure 5.1e**), it is possible to sample amniotic fluid during this stage to determine what signaling factors occur locally. Similarly, there is a window of opportunity for directed gene therapy interventions at this stage before the closure of the neural tube. This is an unprecedented opportunity in a primate species.

Table 5.1. Standard Embryonic Measures as predictors of fertilization age

Measure	R^2	95% CI (days)
Chorionic Cavity Area	0.88	9.0
Yolk Sac Diameter	0.65	13.8
Crown-Rump Length	0.90	10.8
Biparietal Distance	0.97	8.4

Figure 5.2. High resolution, serial ultrasounds reveal period of delayed embryonic development distinct from diapause during gastrulation. *a,* First measurable heart rates in embryos across seven pregnancies (mean =day 57, 95% CI = day 52-64). *b,* Classical clinical embryonic measurements for crown-rump length (CRL) from day 14-93 and biparietal distance (BPD) from 87 days to birth. Measurements are taken across all embryos during 7 pregnancies. Each colored line corresponds to one course of pregnancy; the width of shading represents the 95% confidence interval for each embryo in the pregnancy. Inset shows range of values for *c,d*. *c,* Yolk sac diameter measured from day 13-53 days before birth. *d,* Maximum chorionic cavity area was calculated from 13-53 days before birth as a correlate of chorionic volume. *e,* Relevant outlines of serial ultrasounds across all pregnancies to maximize coverage. Border of each outline represent the extent of the uterine lumen/gestational sac. Red numbers indicate Carnegie Stage. Gray numbers indicate range of approximate days post fertilization. Yellow indicates structures associated with Embryo 1, while blue indicates embryo 2. In outlines in which a third embryo is present, these tissues are denoted in red.

Figure 5.2

A. First Measurable Heart Rate

B.

C. Yolk Sac Diameter

D. Chorionic Cavity XS Area

E.
4 12-13
5 14-18
6 19-25
7 26-46
8 47-54
9 55-57

5 mm
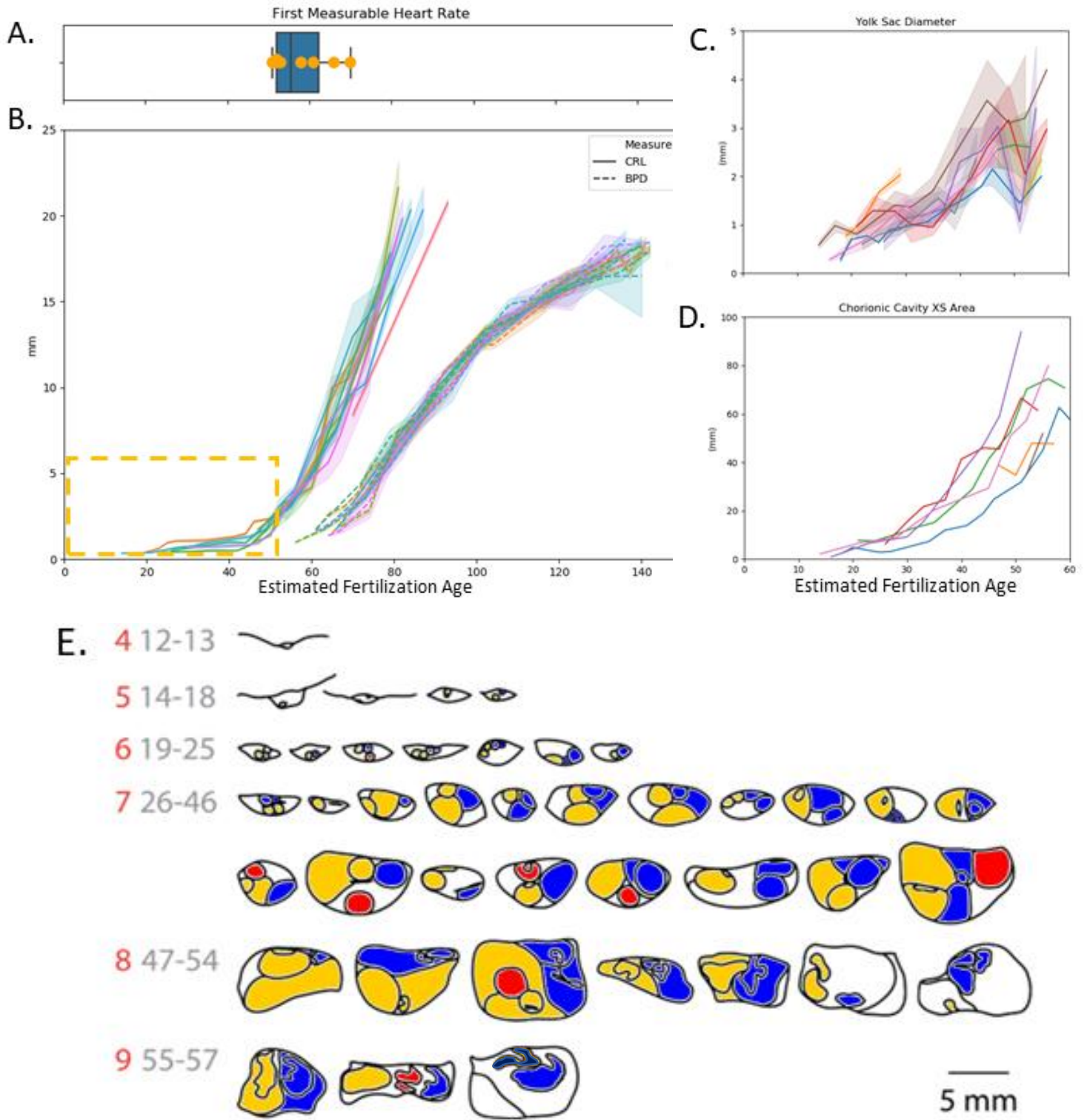
5.3 Fetal organogenesis milestones as viewed on ultrasound

Throughout this ultrasound atlas, we identified key stages of organogenesis that may be of particular interest in tracking normal development of organ systems. This is of importance when creating transgenic animals such that you need to track both the expected and unexpected results of the phenotype, particularly in light of the fact that single mutations can have wide ranging effects across systems (*Bakrania et al., 2008)*. A recent study has established an MRI fetal atlas for the common marmoset and demonstrate approximately similar findings between *ex vivo* and *in utero* MRI images (*Hikishima et al., 2013)*. This work lays the foundation for the level of detail that can be achieved not only at resolution but with different scanning sequences. Nevertheless, for individuals in need of tracking development of their common marmoset fetuses but without access to the specialized equipment needed for *in utero* MRI, we demonstrated here that key findings can be appreciated on ultrasound. Thus, used in concert with the MRI atlas, it is possible to achieve a much more comprehensive and accessible understanding. Further, it is possible to use real time ultrasound with reference MRI atlas images to facilitate procedures that must be done *in utero*. Interestingly, **Figure 5.3a** revealed the development of cortical layers through indirect cues. Both showed enlarged lateral ventricles and thin cortical mantle that gradually get smaller and larger, respectively, as the ventricular zone gives rise to the progenitor cells which differentiate into the layered cortex. These findings were also demonstrated in the Hikishima et al., 2013 study confirming that we are not observing an artifact of a distorted image when seeing these enlarged ventricles. In fact, this layer development can be appreciated with the widely accessible ultrasound technique.

In regard to other organ systems, one of the first signs of a viable fetus is the presence of a

heart tube that then allows flow of blood through the cardiovascular system as seen on Doppler

and through measuring the fetal heartbeat. **Figure 5.3b** shows the different stages of heart

development and the first point at which we were able to observe a heart beat in the common

marmoset fetus, at day 57 (+- 3.8). Finally **Figure 5.3c** reveals a surprising and tantalizing insight

into adrenal and kidney development in the common marmoset. Whereas the fetal

adrenal:kidney length ratio is approximately 2:7 in humans (*van Vuuren et al., 2012*) and is

consistent throughout gestation, the comparable marmoset ratio was approximately 1:1. This is

shifted dramatically later in life as is shown in the inset picturing a freshly dissected adult

kidney next to a freshly dissected infant kidney taken on postnatal day 0.

Figure 5.3. Organogenesis milestones tracked with ultrasound. *a,* Choroid plexus distinct within lateral ventricles as denoted by arrows. Large lateral ventricles (yellow bars) that gradually reduce in size as the cortical thickness increases over time (white bars). *b,* Development of the heart tube with subsequent inclusion into the thoracic cavity and development of distinct chambers. Yellow outlines indicate regions of heart development. *c,*. (i) Adrenal to kidney ratio on ultrasound reveals ratio of 1:1 shown to be accurate with (ii) ex vivo gross dissection. (iii) This ratio dramatically decreases as the kidneys and adrenals mature. All red text indicates Carnegie Stage for images still classified as embryos. All gray text indicates days post fertilization except for P0 (day of birth) and Adult. Scale = 10mm.

Figure 5.3
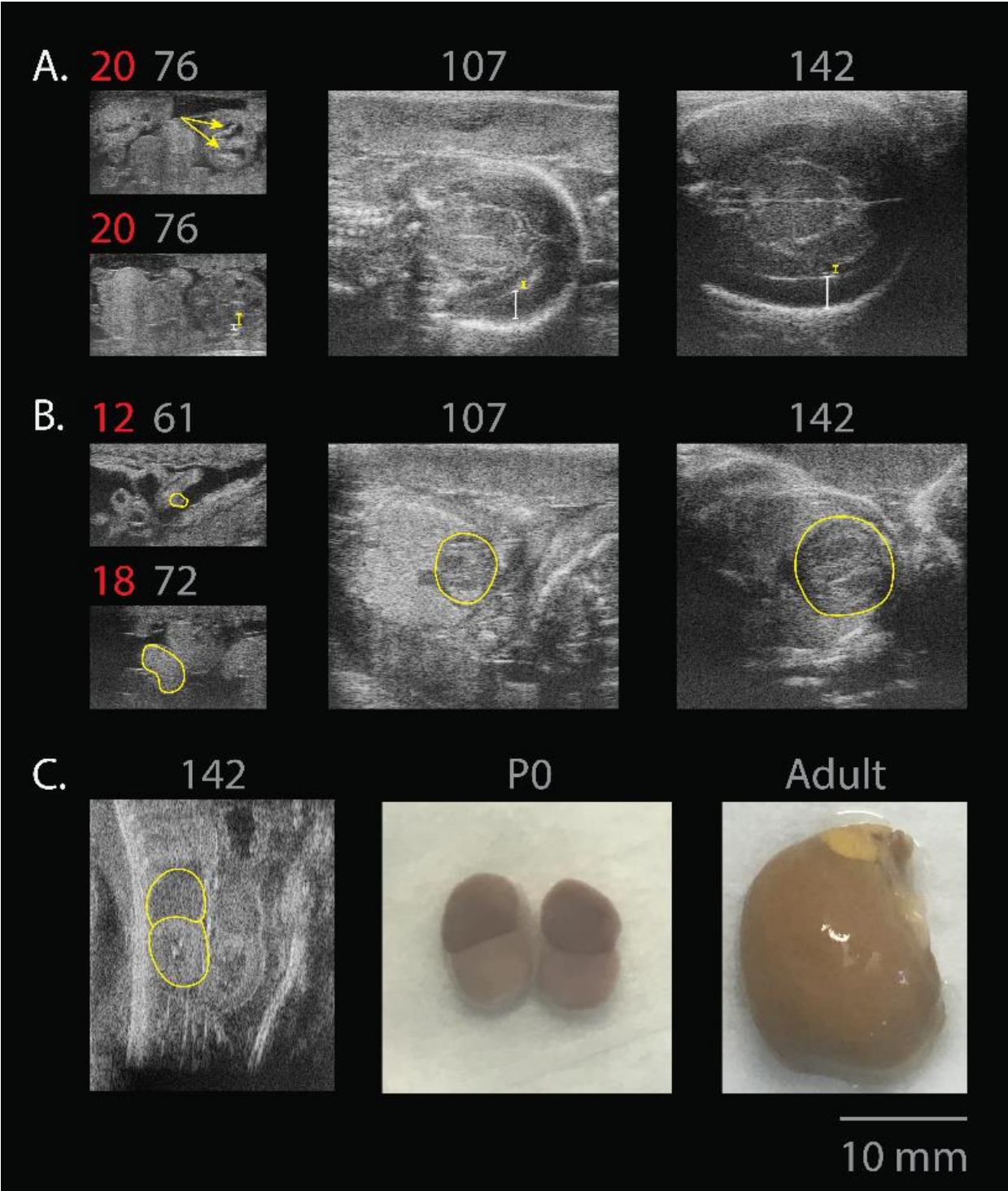
5.4 A residual network model of marmoset development predicts embryonic age

While ultrasound has proven to be an invaluable method of investigating the anatomy and physiology of embryonic and fetal development, there remain several caveats. Firstly, proper scanning and interpretation of imaging require a skilled imaging technician and analyst, in both the NHP experimental field as well as in clinical human radiology. Secondly, accurate and precise measures of embryonic age and stage of development, especially early in development, require multiple serial scans – which is inconvenient in the experimental setting and infeasible in the clinic. We determined to use the data collected as part of this ultrasound atlas to generate predictive models of marmoset development, allowing accurate extrapolation of embryonic age from single random frames. In doing so, we provide a tool that minimizes the technical expertise, time, and effort required to track marmoset development.

We utilized machine learning tools in the Keras deep learning API (*Chollet et al., 2015*) to define and train a set of residual neural networks that sort salient frames and assign an embryonic age in an automated fashion (**Figure 5.4a**). When evaluating the performance of this system on a random set of novel, unbiased ultrasound frames (i.e. drawn from any of the frames collected during any of the ultrasound sessions in our database), our models were able to estimate embryonic age with a median error of 5.44 days (95% CI +/- 8.84 days, **Figure 4b,c,d**). In comparison, clinical measurements of gestational age by a skilled analyst as described in Figure 2 achieved a maximum accuracy of between +/- 8-13 days, depending on the measure and timepoint analyzed.
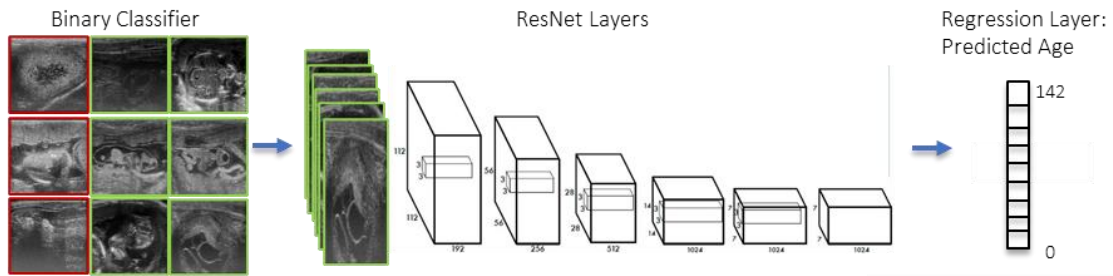
This tool also opens new windows for analysis, not only by predicting windows for intervention using a single scan, but also including the potential for predicting abnormal development and evaluating health in transgenic animals.

Figure 5.4. *A residual network model of marmoset development predicts embryonic age.*

a, *Schematic of residual network model. Images are sorted by a binary classifier built*

upon ResNet50 in step 1, after which salient images are input for a separate residual

network that abstracts features from the input images that are relevant to output

predictions (see Ch2 - Materials and Methods). The output of this network is resolved

into a regression layer that outputs a single vector prediction of embryonic age. *b,*

Performance of model using 180 randomly selected frames as input with predicted age

(orange) and actual age (blue) plotted according to time until birth. *c,* *Predictive*

accuracy of the model as a box-and-whisker plot based upon predicted age for the

*frames plotted in (**b**). Median error = 5.44 days, 95% CI = +/- 8.84 days.* *d,* *Sample frames*

for which the model performed best (top row) and worst (bottom). Clear anatomic

landmarks can be seen in the top frames, while the bottom frames are of poor quality.

Figure 5.4

A

Binary Classifier

ResNet Layers

Regression Layer:
Predicted Age

B Predictive performance of model

C Predictive Accuracy

D Best / Worst Model Performance
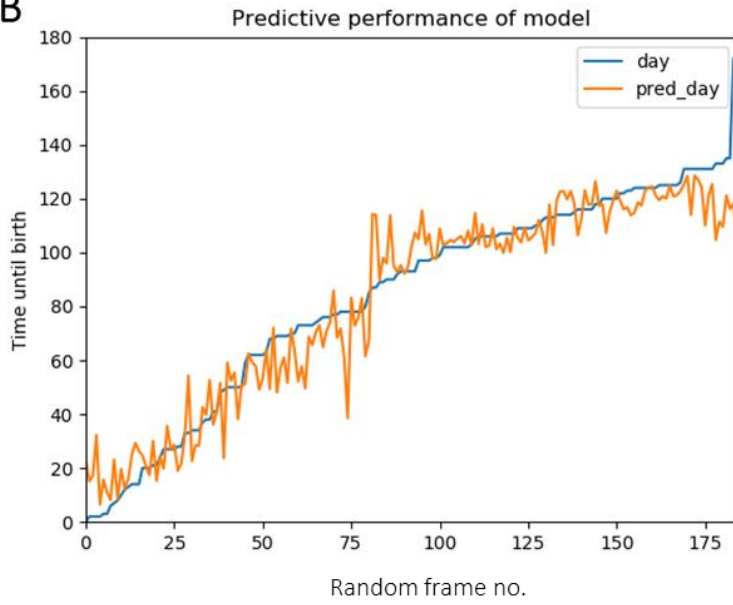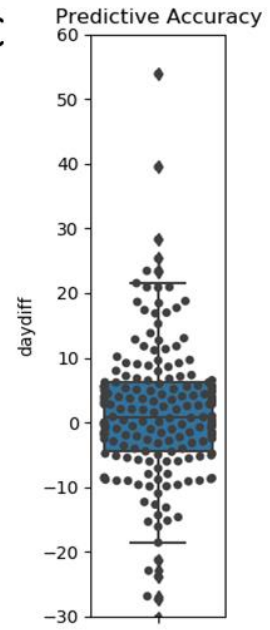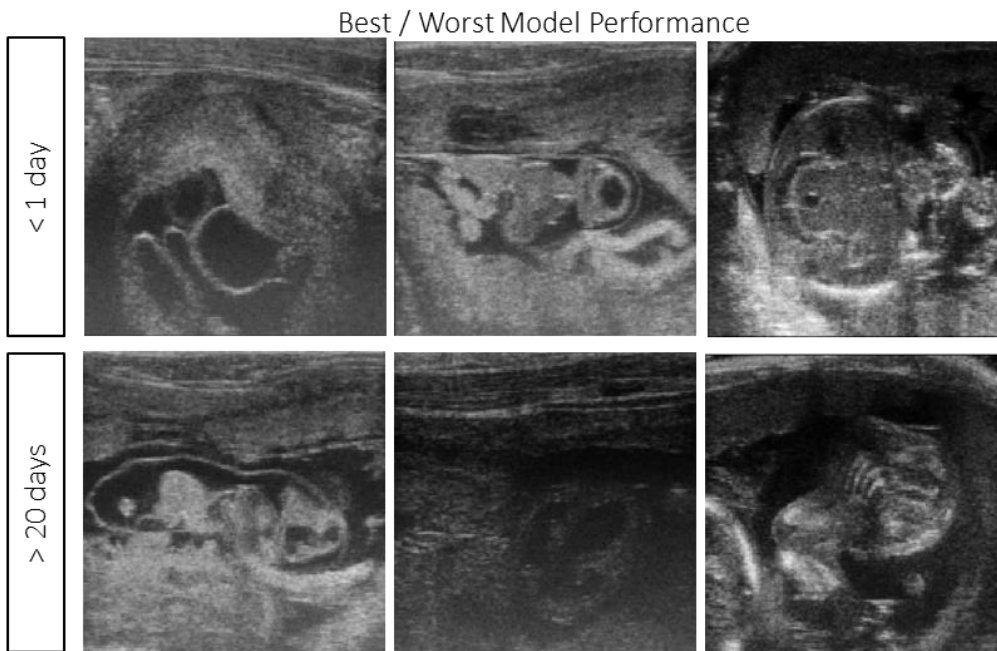
This high resolution, serial marmoset ultrasound atlas builds on the tremendous work done in this field from the earliest *ex vivo* studies of embryos through to the first ultrasound derived growth curves and *in utero* MRI atlas. With ease of data sharing and accessibility to ultrasound units in most facilities, it establishes the opportunity to pursue not only remaining enigmas of common marmoset development but also facilitates utilizing their unique physiology to rapidly push forward molecular and transgenic utility to probe our understanding of reproduction, cognition, and biomedical interventions.

Supplemental Figure 5.1a
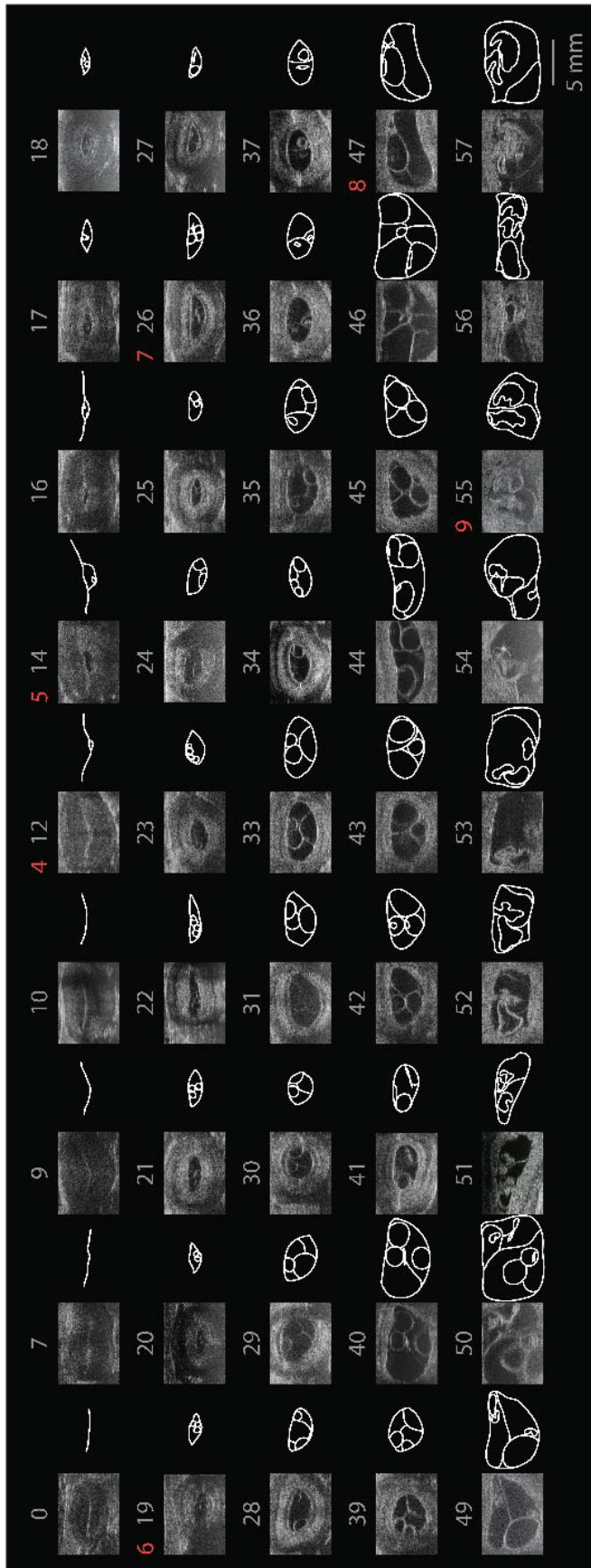
An example of interactive, annotated atlas data for the first 57 days of marmoset embryonic development. Annotated cartoons and selected frames will be made available in addition to scrollable z stacks and access to the predictive modeling system developed in Figure 5.4 on the Marmoset Ultrasound Atlas server after publication.

Supplemental Figure 5.1b – Further day-by-day examples of representative frames and

annotated cartoons from embryonic development post-neural tube closure.
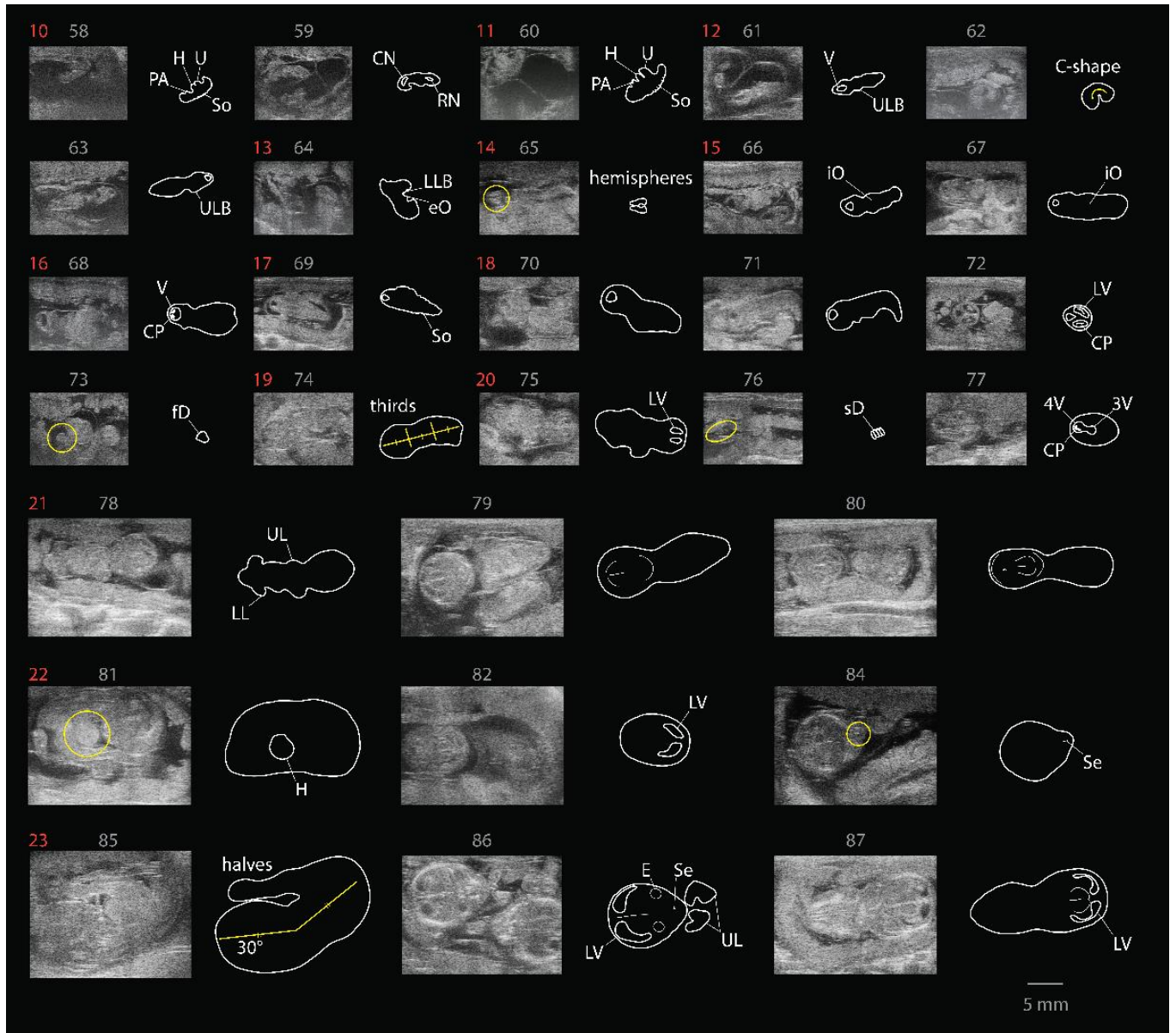
Outlines in white detail relevant features in ultrasound. Yellow highlights key areas or orientations in figures. Red numbers indicate Carnegie Stage. Gray numbers indicate range of days per stage. A, amnion; B, blastocyst; CP, choroid plexus; E, embryo; eO, external organs; ED, embryonic disk; fD, fused digits; H, heart; I, implantation; iO, internal organs; L, lumen; LL, lower limb; LLB, lower limb bud; PA, pharyngeal arches; sD, separated digits; Se, septum; So, somites; U, umbilicus; UL, upper limb; ULB, upper limb bud; V, ventricle; YS, yolk sac.

Supplemental Figure 5.2. Standard embryonic measures as predictors of fertilization

age. *a,* Polynomial linear regressions based upon each of CCA, YSD, CRL, BPD as

individual predictors of fertilization age. Dots represent each measurement aggregated

from collected data, and lines trace the curve of best fit. *b,* Model description and

statistics for linear regression model performance for each measure.

Supplemental Figure 5.2

a.
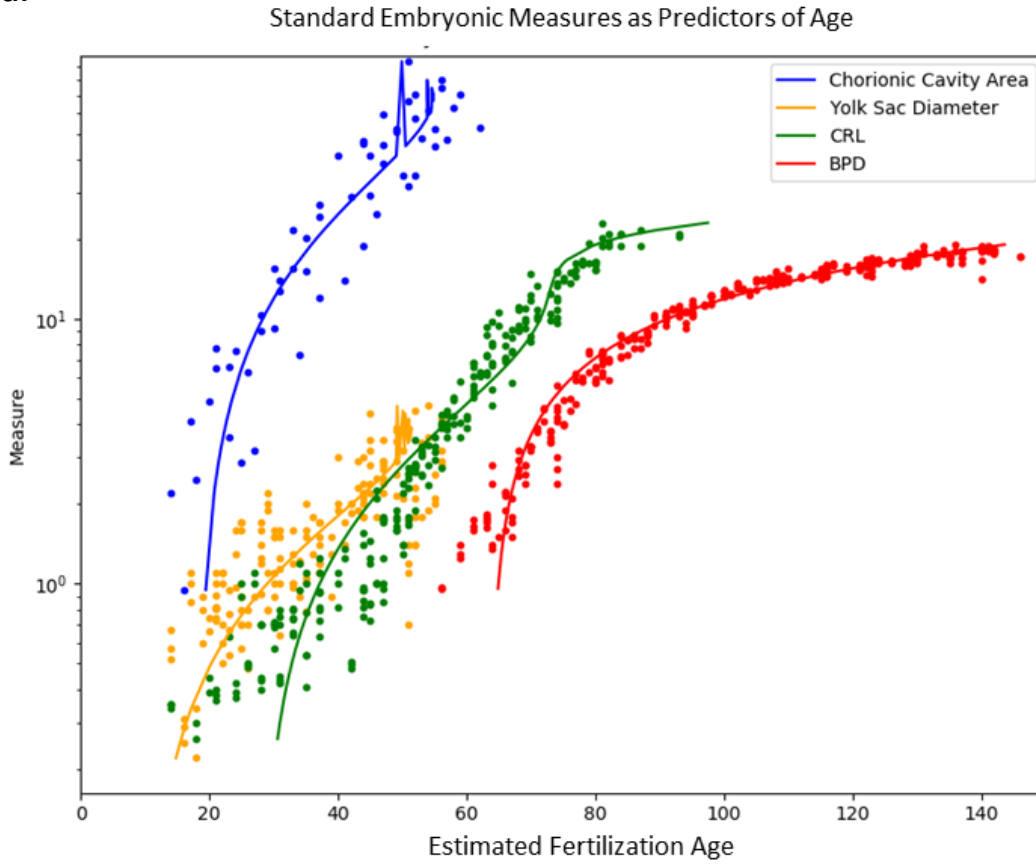
b.

Measure	R²	95% CI (days)	Best Fit Curve
Chorionic Cavity Area	0.88	9.0	$y = -125 + 1.06x - 0.008x^2$
Yolk Sac Diameter	0.65	13.8	$y = -133 + 21.6x - 2.83x^2$
Crown-Rump Length	0.90	10.8	$y = -115 + 9.72x - 0.728x^2 + 0.019x^3$
Biparietal Distance	0.97	8.4	$y = -79.4 + 1.18x - 0.157x^2$

CHAPTER 6

Discussion: Towards translation models of human germ cell development

6.1 Modeling the PGC niche

It is rather remarkable that, despite countless variations on a theme, germ cell specification by induction seems to follow a core pattern of BMP4 stimulation in a mesendodermal population throughout the animal kingdom. Despite drastic changes in body plan, embryo morphology, or gestational timing (or indeed viviparity vs oviparity), what remains constant is that extra-embryonic and embryonic cells work in concert to generate a signaling niche in which mesendodermal precursors are exposed to the right combination of BMP4, WNT, and ACTIVIN to balance the specification of protected germ cell progenitors with the need for defining somatic populations.

In **Chapter 3**, we demonstrate that the *gastruloid*, a powerful model of gastrulation which demonstrates self-organized somatic cell fates in response to an exogenous BMP4, also specifies hPGCs. By minimizing exogenous signaling and relying instead on endogenous signaling to self-organize cell fates, the gastruloid provides a platform in which quantitative dissection of molecular and cellular interactions is possible. scRNA-seq analysis coupled with immunofluorescent imaging demonstrated that a nascent SOX17+ population within the gastruloid is actually bipotentiate. Previously thought to consist of definitive endoderm, we

demonstrate that. 48 hours after BMP4 stimulation, indeed all SOX17+ cells coexpress BLIMP1, AP2y, and OCT4 and thus are capable of becoming PGCs, as validated by scRNA-seq clustering and pseudotime analysis. This assertion is validated by other observations, which demonstrate that AP2y functions upstream of SOX17 to determine whether TFAP2A+ cells generate hPGCs or undergo somatic differentiation (*Chen et al., 2019*).

These results are significant as they highlight a more precise timeline for hPGC specification, and demonstrate that SOX17+ cells in the gastruloid are derived from the epiblast in the absence of yolk sac endoderm (hypoblast). Our scRAseq analysis indicates that the extraembryonic cells present at the borders of gastruloids consist of mesodermal and ectodermal derivatives, and are thus a suggested correlate of amniotic tissues. No specific markers of syncytiotrophoblast or cytotrophoblast tissue (e.g. HCGA/B, HLA-G) are present at this timepoint. This is in agreement with a hypothesis that epiblast and amnion form a continuum, both derived from the inner cell mass and more closely related to each other than to the yolk sac endoderm or trophoblast, as suggested by embryo analysis and immunostaining of *in vitro* attached embryo culture (*Deglincerti et al., 2016*).

6.2 New functional assays for *in vitro*-derived hPGCs

The transcriptomic signature of gastruloid-derived hPGCs included known chemoattractant receptors CXCR4 and KIT, migratory cell surface markers including PDPN and EPCAM, as well as several other proteins involved in the transduction of migratory signals, including the highly specific and strong expression of DOCK2, a GEF family member that is known to activate chemotaxis programs in lymphocytes, but which has not been previously described in PGCs in model systems or *in vivo* (*Fukui et al., 2001; Dobbs et al., 2015*). The expression of chemoattractant receptors and the signal transduction machinery involved in active migration prompted us to test this ability, and thus functionally validate one of the signatures of hPGCs.

We demonstrate this migratory potential using the chick embryo as a model system, as this model system allows for efficient xenotransplantation and imaging access to the early developing embryo and has proven useful in the functional validation of other *in vitro*-derived cell types in the past (*Martyn et al., 2018*). PGC migration in the chick embryo is known to involve SDF-1 – CXCR4 signaling, however, chick PGCs take the unique step of migrating through the developing vasculature to the intermediate mesoderm to arrive at the gonadal ridge as opposed to mouse, NHP and human PGCs, which migrate through the hindgut epithelium into the intermediate mesoderm to arrive at their destination. In this xenotransplantation system, hPGCs move laterally towards the embryo when transplanted at the germinal crescent. Indeed, these cells also begin to express cytoplasmic DAZL as *in vivo* migratory cells do, though they

display a curious downregulation of NANOS3, which may represent a transient state of active migration in which NANOS3 may be downregulated.

It remains to be seen whether these transplanted cells take the same vasculature route as chick PGCs. These questions may be addressed in the future using CRISPR-Cas9 knockout of chemoattractant receptors such as CXCR4 and KIT within transplanted hPGCs. Our preliminary observations point to the ability of hPGCs to migrate towards a source of SDF-1 (*data not shown*).

We next asked whether xenotransplanted hPGCs could follow endogenous chick cues and home to the developing gonad. Experiments involving direct injection of either chick or quail PGCs into the vasculature of a chick have resulted in migration to the gonad and proper cell maturation, demonstrating the validity of this assay in closely related species (*Nakamura et al., 1991, 1992*). When we performed a similar experiment with hPGCs, we observed specific colonization of the gonad by a few HNA+/DAZL+ cells 24 hours later, with more cells found in the mesonephric area surrounding the gonad, but not anywhere else in the embryo. Though specific chemoattractant cues are yet to be demonstrated in this transplantation system, it provides both a path for functional *in vitro*-derived PGC validation that is independent from maturation to a fertilizable germ cell (a step impossible with human *in vitro*-derived PGCs) but also another intriguing window into how disparate animals conserve elements of PGC development throughout evolution.

6.3 Defining the hPGC signaling niche within the gastruloid

In **Chapter 4** we define the signaling requirements of the hPGC specification niche. BMP4, WNT, and ACTIVIN are all necessary, with WNT and ACTIVIN/NODAL potentially functioning to push cells toward a transient mesendodermal fate (*Irie et al., 2015, Sasaki et al., 2015*). Our conclusions, while in agreement with these studies, provide several important clarifications. We demonstrate that in a self-organizing system, BMP4 has two functions. First, to initiate the cascade of signaling that results in a SOX17+ precursor population that co-expressed BLMIP1, AP2y, OCT4 and NANOG, fated to become hPGCs as well as definitive endoderm. Secondly, BMP4 generates a gradient in which only SOX17+ within a tight window of BMP4 during the first 24 hours become hPGCs. Altering the WNT gradient (by knockout of DKK1) does not alter the size of the hPGC niche while it does increase the definitive endoderm niche; thus decoupling the two fates. Altering the ACTIVIN gradient (by knockout of LEFTY or CER1), on the other hand, increases the number of specified hPGCs within the same radially confined niche.

Thus, BMP4 and ACTIVIN act as instructive signals, while WNT signaling is permissive but not instructive for hPGC specification. Our DKK1-/- experiment also demonstrates that expression of AP2y and NANOG is downstream of BMP4 but not WNT, suggesting that BMP4 functions independently from the other two pathways to induce hPGCs. Our results are in agreement with recent studies demonstrating that AP2y is independent from SOX17 and BLIMP1

expression in humans (*Chen et al., 2019*), highlighting an hPGC specification pathway that stands in contrast to the mouse, where AP2y is induced downstream of BLIMP1.

In order to begin probing the roles of transcription factors vital to hPGC identity, we similarly generated hESCs that conditionally express BLIMP1. In a SOX17-/- background, we also find that BLIMP1 overexpression is not sufficient for PGC fate acquisition, but in wildtype cells, we found that BLIMP1 overexpression alone induces expression of SOX17 and downregulation of SOX2, indicating that a feedback loop exists between BLIMP1 and SOX17. These cells also express OCT4 and AP2y, which indicates that either BLIMP1 may function to directly upregulate these factors, or that in the presence of BLIMP1 endogenous signaling is able to maintain expression of these factors. Indeed, the direct effects of BMP4 stimulation on AP2y and NANOG expression demonstrated earlier in this chapter indicate that the latter may be the case, which can be investigated by overexpression of BLIMP1 in a BMP4-/- background. We further show that four days of BLIMP1 overexpression in the absence of exogenous signaling is sufficient to induce DAZL and DDX4 in a subset of cells, validating these cells as PGCs and providing a mechanism for investigating factors downstream of BLIMP1 which induce these later PGC markers.

**6.4 A perspective into modeling later hPGC development**

As we push towards a model of later hPGC development, we wondered how the acquisition of hPGC markers such as DAZL (observed within hPGC chick xenotransplants) could be modeled *in vitro*. We found that 225 um gastruloids enhanced for hPGC specification begin to express DAZL

119

when removed and cultured upon Matrigel in the presence of BMP4, though migrating cells still lose expression of NANOS3 (**Figure 6.1a,b**). In an effort to provide a gonadal niche for these cells, we cultured PGC enhanced gastruloids upon a monolayer of primary marmoset stromal cells, reasoning that a NHP host may be more suitable as a supportive layer than mouse gonadal tissue. Under these conditions, PGCs remain in tightly know nests within the stroma, and express DAZL, DDX4, and NANOS3 in addition to OCT4 in the absence of any exogenous signaling (**Figure 6.1c**).

In an attempt to model this environment in a xeno-free manner in a more controlled setting, we have also begun co-culture of FACS-sorted PGCs on micropatterned colonies of hESC-derived induced granulosa-like cells (**Figure 6.1d).** We are in the process of refining this approach, which was taken to more closely mimic the phenotype of early embryonic granulosa than using adult tissue, with the reasoning that mouse PGC maturation and gametogenesis has required co-culture with embryonic/fetal stromal tissue (*Ohinata et al., 2006; Vincent et al., 2011; Nayernia et al., 2006; Hayashi et al., 2012; Yamashiro et al., 2018*). Early observations indicate that not only do hPGCs survive in these "ova-noid" colonies, but these colonies self-organize (**Figure 6.1d**). In the future, this approach will bring some of the advantages of self-organized organoid culture to studies of later hPGC development.

Figure 6.1. Culture strategies for long-term PGC growth and maturation. *a,* 225 um gastruloids enriched for PGCs were dissociated and seeded onto micropatterned surfaces in the presence of BMP4 for 7 days. Cells maintain SOX17 expression (red) and acquire AMHR2 (green) expression (also NANOG+/DAZL+ - data not shown). Only cells seeded on 80 um patterns demonstrate formation of a central lumen containing one cell nucleus. *b,* Gastruloids enriched for PGCs and grown on a Matrigel plate for 7 days in the presence of BMP4 express NANOS3 (red) and OCT4 (purple) as well as DAZL (green) in the center of the colony, while cells that proliferate and migrate away from the center express cytoplasmic DAZL only. *c,* Gastruloids enriched for PGCs and seeded on a bed of primary adult marmoset gonadal stromal cells for 7 days express nuclear marker H2B-mCitrine (green), NANOS3 (red), and DDX4 (blue) and do not migrate from the colony. *d,* FACS-sorted PGCs co-cultured on a micropatterned layer of hESC-derived granulosa-like cells display radial self-organization. As opposed to the other culture methods, OCT4+ germ cells (green) are widely spaced in the center of the colony and surrounded by FOXL2+ cells (purple). Colonies display radially symmetric expression of AMHR2 (blue), while PGC-derived cells expressing H2B-mCardinal (red) but not OCT4 (fate identity to be determined) are distal. Preliminary experiments using PGCs expressing CRISPR-Cas9 encoded OCT4-EGFP and DDX4-tdTomato reporters indicate a switch from OCT4 expression to DDX4 expression after 2 weeks of co-culture (data not shown).

Figure 6.1
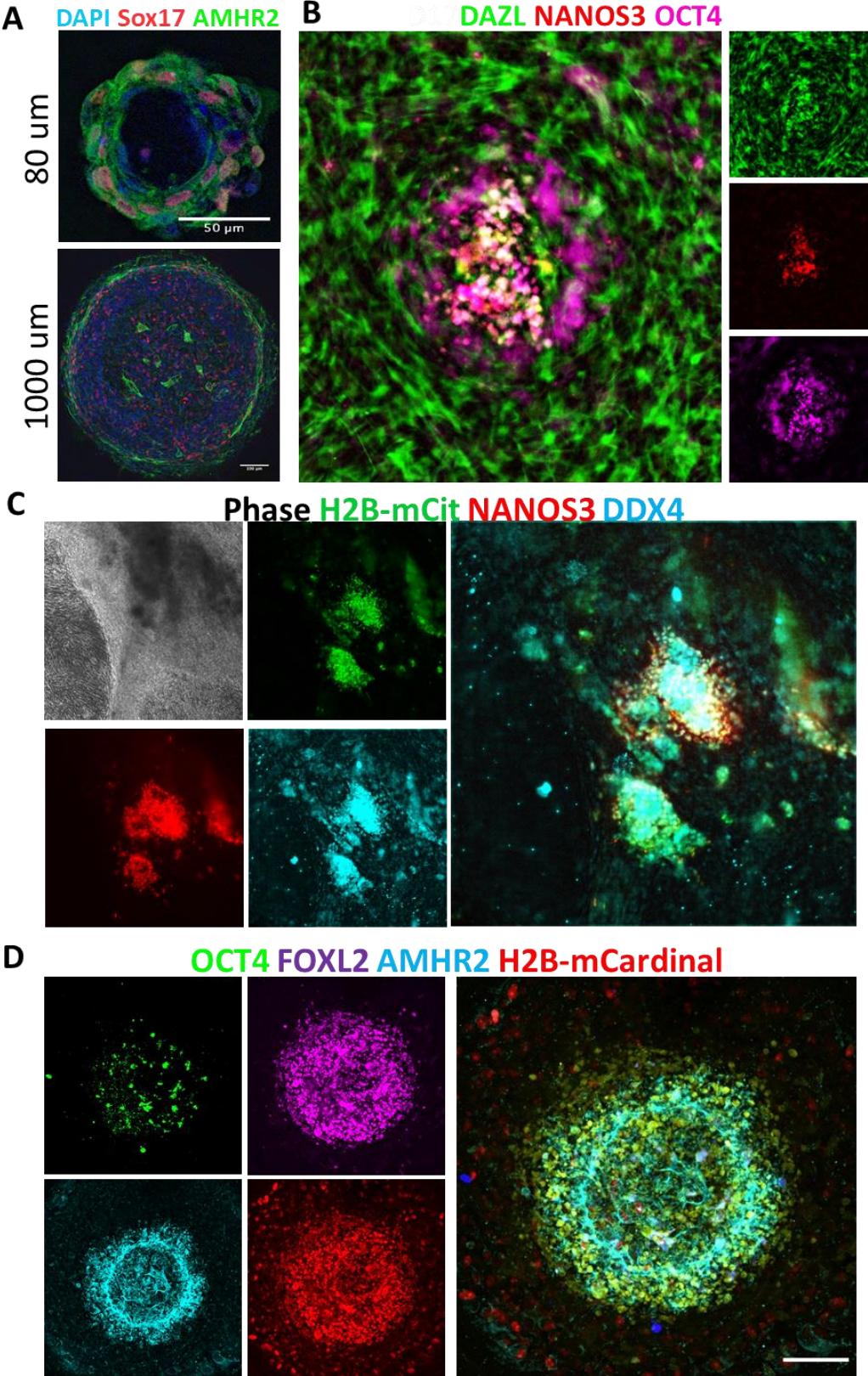
6.5 Towards translational models of Non-Human Primate embryology

Given the relative inaccessibility of the human embryo for basic science studies and the complete inaccessibility of the human embryo for transgenic modeling, we have worked to develop a bridge between NHP embryologic studies and our human stem cell culture models, with the ultimate goal of generating isogenic *in vivo* and in vitro correlates to our hESC studies. Several groups have demonstrated the potential of the macaque and the marmoset as useful model systems for the study of primate-specific elements of reproductive development (*Clark et al., 2017, Boroviak et al., 2017, Lin et al., 2015, Aeckerle et al., 2014, Ting et al., 2017; Fereydouni et al., 2016*). Other groups, particularly the RIKEN Institute, have paved the way in clinical marmoset embryology, establishing the marmoset as a convenient transgenic system - particularly for models of neurological disease (*Sasaki et al., 2009; Takahashi et al., 2014; Liu et al., 2016; Park et al., 2019; Park et al., 2016*).

We introduced the marmoset as a new model system in our laboratory in 2015, generating a relatively small breeding colony (10 animals, 5M/5F) and working to lay the groundwork for transgenic models of development, when existing transgenic marmoset experiments typically require many more animals and 300-500 embryos each to result in the few animals with germline transmission. The first step we took towards this goal was the methodical generation of the ultrasound atlas of marmoset development described in **Chapter 5**. The next step was the establishment of isogenic marmoset stem cell lines derived from animals in our colony, to serve as a bridge between *in vitro* gastruloid studies and *in vivo* development - as well as the lay

the groundwork for future generation of transgenic animals through somatic cell nuclear transfer, blastocyst injection, and allogenic transplantation (*data not shown*).

With these tools in hand, the foundation for *in vivo* NHP studies of PGC development utilizing the new frontiers of genome editing, allogenic transplantation, and lineage tracing has been established. Precise knowledge of gene targets from hESC *in vitro* studies and the ability to efficiently edit marmoset iPSCs using CRISPR/Cas9 systems allow us to propose fruitful transgenic experiments. Accurate ultrasound staging and automated analysis methods will not only simplify non-invasive tracking of transgenic embryo development, but also create windows of opportunity for amniotic injection of lineage tracers, cell recorders, and even transplantation of *in vitro*-derived marmoset PGCs. In all, we hope that in addition to germline transgenic models, an arsenal of new tools based upon *in vivo* NHP manipulations in the peri-gastrulation window will be built upon the foundations we provide here; laying the groundwork for studies that will be transformative for our understanding of primate development.

References

Aeckerle, N., Drummer, C., Debowski, K., Viebahn, C., & Behr, R. (2014). Primordial germ cell development in the marmoset monkey as revealed by pluripotency factor expression: Suggestion of a novel model of embryonic germ cell translocation. Molecular Human Reproduction, 21(1), 66–80. https://doi.org/10.1093/molehr/gau088

Arnold, S. J., & Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nature Reviews Molecular Cell Biology, 10(2), 91–103. https://doi.org/10.1038/nrm2618

Anderson, R. a, Fulton, N., Cowan, G., Coutts, S., & Saunders, P. T. (2007). Conserved and divergent patterns of expression of DAZL, VASA and OCT4 in the germ cells of the human fetal ovary and testis. BMC Developmental Biology, 7(February), 136. https://doi.org/10.1186/1471-213X-7-136

Anderson, R. H., Webb, S., Brown, N. A., Lamers, W., & Moorman, A. (2003). Development of the heart: (2) Septation of the atriums and ventricles. *Heart (British Cardiac Society)*, *89*(8), 949–958. https://doi.org/10.1136/heart.89.8.949Chapman SC, Collignon J, Schoenwolf GC, Lumsden A. Improved method for chick whole-embryo culture using a filter paper carrier. *Dev Dyn*. 2001;220(3):284-289. doi:10.1002/1097-0177(20010301)220:3<284::AID-DVDY1102>3.0.CO;2-5

Bakrania, P., Efthymiou, M., Klein, J. C., Salt, A., Bunyan, D. J., Wyatt, A., … Ragge, N. K. (2008).

Mutations in BMP4 cause eye, brain, and digit developmental anomalies: overlap between

the BMP4 and hedgehog signaling pathways. *American Journal of Human Genetics*, *82*(2),

304–319. https://doi.org/10.1016/j.ajhg.2007.09.023

Blaas, H. G. K. (1999). The examination of the embryo and early fetus: How and by whom?

Ultrasound in Obstetrics and Gynecology. https://doi.org/10.1046/j.1469-

0705.1999.14030153.x

Chambers, P. L., & Hearn, J. P. (1985). Embryonic, foetal and placental development in the

Common marmoset monkey (Callithrix jacchus). *Journal of Zoology*, *207*(4), 545–561.

https://doi.org/10.1111/j.1469-7998.1985.tb04951.x

Chen D, Liu W, Zimmerman J, et al. (2018). The TFAP2C-Regulated OCT4 Naive Enhancer Is

Involved in Human Germline Formation. *Cell Rep*. 2018;25(13):3591-3602.e5.

doi:10.1016/j.celrep.2018.12.011

Chollet, F. et al. (2015). Keras. https://github.com/fchollet/keras. (2015).

Clark, A. T., Gkountela, S., Chen, D., Liu, W., Sosa, E., Sukhwani, M., … Orwig, K. E. (2017).

Primate Primordial Germ Cells Acquire Transplantation Potential by Carnegie Stage 23. Stem

Cell Reports, 9(1), 329–341. https://doi.org/10.1016/j.stemcr.2017.05.002

Deglincerti A, Etoc F, Guerra MC, et al. (2016). Self-organization of human embryonic stem cells

on micropatterns. *Nat Protoc*. 2016;11(11):2223-2232. doi:10.1038/nprot.2016.131

De Melo Bernardo A, Sprenkels K, Rodrigues G, Noce T, De Sousa Lopes SMC. (2012). Chicken

primordial germ cells use the anterior vitelline veins to enter the embryonic circulation. *Biol*

Open. 2012;1(11):1146-1152. doi:10.1242/bio.20122592

Deng, L., Li, C., Chen, L., Liu, Y., Hou, R., & Zhou, X. (2018). Research advances on embryonic

diapause in mammals. *Animal Reproduction Science*, *198*, 1–10.

https://doi.org/10.1016/j.anireprosci.2018.09.009

Dickey, R. P., & Gasser, R. F. (1993). Ultrasound evidence for variability in the size and

development of normal human embryos before the tenth post-insemination week after

assisted reproductive technologies. *Human Reproduction*, *8*(2), 331–337.

https://doi.org/10.1093/oxfordjournals.humrep.a138046

Dobbs K, Conde CD, Zhang SY, et al. (2015). Inherited DOCK2 deficiency in patients with early-

onset invasive infections. *N Engl J Med*. 2015;372(25):2409-2422.

doi:10.1056/NEJMoa1413462

Enders, A. C., & Lopata, A. (1999). Implantation in the marmoset monkey: Expansion of the

early implantation site. *The Anatomical Record*, *256*(3), 279–299.

https://doi.org/10.1002/(SICI)1097-0185(19991101)256:3<279::AID-AR7>3.0.CO;2-O

Evans, T., Wade, C. M., Chapman, F. A., Johnson, A. D., & Loose, M. (2014). Acquisition of germ

plasm accelerates vertebrate evolution. Science (New York, N.Y.), 344(6180), 200–203.

https://doi.org/10.1126/science.1249325

Etoc F, Metzger J, Ruzo A, et al. (2016). A Balance between Secreted Inhibitors and Edge Sensing

Controls Gastruloid Self-Organization. *Dev Cell*. 2016;39(3):302-315.

doi:10.1016/j.devcel.2016.09.016

Fernandes MG, Bialecka M, Salvatori DCF, Chuva De Sousa Lopes SM. (2018). Characterization of migratory primordial germ cells in the aorta-gonad-mesonephros of a 4.5-week-old human embryo: A toolbox to evaluate in vitro early gametogenesis. *Mol Hum Reprod*. 2018;24(5):233-243. doi:10.1093/molehr/gay011

Fukui Y, Hashimoto O, Sanui T, et al. (2001). Haematopoietic cell-specific CDM family protein DOCK2 is essential for lymphocyte migration. *Nature*. 2001;412(6849):826-831. doi:10.1038/35090591

Gkountela, S., Li, Z., Vincent, J. J., Zhang, K. X., Chen, A., Pellegrini, M., & Clark, A. T. (2013). The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. Nature Cell Biology, 15(1), 113–122. https://doi.org/10.1038/ncb2638

Gkountela, S., Zhang, K. X., Shafiq, T. A., Liao, W. W., Hargan-Calvopiña, J., Chen, P. Y., & Clark, A. T. (2015). DNA demethylation dynamics in the human prenatal germline. Cell, 161(6), 1425–1436. https://doi.org/10.1016/j.cell.2015.05.012

Günesdogan, U., Magnúsdóttir, E., & Surani, M. A. (2014). Primordial germ cell specification: a context-dependent cellular differentiation event [corrected]. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 369(1657), 20130543-. https://doi.org/10.1098/rstb.2013.0543

Haghverdi L, Büttner M, Wolf FA, Buettner F, Theis FJ. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods*. 2016;13(10):845-848. doi:10.1038/nmeth.3971

Handel, M. A., Eppig, J. J., & Schimenti, J. C. (2014). Applying "gold standards" to in-vitro-derived germ cells. Cell, 157(6), 1257–1261. https://doi.org/10.1016/j.cell.2014.05.019

Harkness, L. M., Rodger, M., & Baird, D. T. (1996). *Morphological and molecular characteristics of living human fetuses between Carnegie stages 7 and 23: ultrasound scanning and direct measurements*. *Human Reproduction Update* (Vol. 3). Retrieved from https://academic.oup.com/humupd/article-abstract/3/1/25/593774

Hayashi, K., Ogushi, S., Kurimoto, K., Shimamoto, S., Ohta, H., & Saitou, M. (2012). Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. Science (New York, N.Y.), 338(6109), 971–975. https://doi.org/10.1126/science.1226889

Hayashi M, Kawaguchi T, Durcova-Hills G, Imai H. (2017). Generation of germ cells from pluripotent stem cells in mammals. *Reprod Med Biol*. 2017;(November):1-8. doi:10.1002/rmb2.12077

Hearn, J. P., Abbott, D. H., Chambers, P. C., Hodges, J. K., & Lunn, S. F. (1978). Use of the common marmoset, Callithrix jacchus, in reproductive research. *Primates in Medicine*, *10*, 40–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/417331

Hikishima, K., Sawada, K., Murayama, A. Y., Komaki, Y., Kawai, K., Sato, N., … Okano, H. (2013). Atlas of the developing brain of the marmoset monkey constructed using magnetic resonance histology. *Neuroscience*, *230*, 102–113. https://doi.org/10.1016/J.NEUROSCIENCE.2012.09.053

Hill, M. A. (2020). Embryology Carnegie Stage Comparison. Retrieved February 11, 2020, from

https://embryology.med.unsw.edu.au/embryology/index.php/Carnegie_Stage_Comparison

Hockemeyer D, Wang H, Kiani S, et al. (2012). Genetic engineering of human ES and iPS cells using TALE nucleases. *Nat Biotechnol*. 2012;29(8):731-734. doi:10.1038/nbt.1927.Genetic

Irie N, Weinberger L, Tang WWC, et al. (2015). SOX17 is a critical specifier of human primordial germ cell fate. *Cell*. 2015;160(1-2):253-268. doi:10.1016/j.cell.2014.12.013

Jaquish, C. E., Toal, R. L., Tardif, S. D., & Carson, R. L. (1995). Use of ultrasound to monitor prenatal growth and development in the common marmoset (Callithrix jacchus). *American Journal of Primatology*, *36*(4), 259–275. https://doi.org/10.1002/ajp.1350360402

Jung, D., Xiong, J., Ye, M., Qin, X., Li, L., Cheng, S., … Kee, K. (2017). In vitro differentiation of human embryonic stem cells into ovarian follicle-like cells. Nature Communications, 8(May 2016), 15680. https://doi.org/10.1038/ncomms15680

Johnson, A. D., & Alberio, R. (2015). Primordial germ cells: the first cell lineage or the last cells standing? Development, 142(16), 2730–2739. https://doi.org/10.1242/dev.113993

Julaton, V. T. A., Pera, R. A. R., & Reijo Pera, R. A. (2011). NANOS3 function in human germ cell development. Human Molecular Genetics, 20(11), 2238–2250. https://doi.org/10.1093/hmg/ddr114

Kang KS, Lee HC, Kim HJ, et al. (2015). Spatial and temporal action of chicken primordial germ cells during initial migration. *Reproduction*. 2015;149(2):179-187. doi:10.1530/REP-14-0433

Kee, K., Gonsalves, J.M., Clark, A.T., and Reijo Pera, R.A. (2006). Bone Morphogenetic Proteins Induce Germ Cell Differentiation from Human Embryonic Stem Cells. Stem Cells and

Development 2006 15:6, 831-837

Kobayashi, T., Zhang, H., Tang, W. W. C., Irie, N., Withey, S., Klisch, D., … Surani, M. A. (2017).

Principles of early human development and germ cell program from conserved model

systems. Nature, 546(7658), 416–420. https://doi.org/10.1038/nature22812

Kropp, J., Di Marzo, A., & Golos, T. (2017). Assisted reproductive technologies in the common

marmoset: an integral species for developing nonhuman primate models of human

diseases†. *Biology of Reproduction*, *96*(2), 277–287.

https://doi.org/10.1095/biolreprod.116.146514

Kuphal S, Bosserhoff A. (2009). Recent progress in understanding the pathology. *J Pathol*.

2009;(August):400-409. doi:10.1002/path

Lawson, K. A., Dunn, N. R., Roelen, B. A. J., Zeinstra, L. M., Davis, A. M., Wright, C. V. E., …

Hogan, B. L. M. (1999). Bmp4 is required for the generation of primordial germ cells in the

mouse embryo. Genes and Development, 13(4), 424–436.

https://doi.org/10.1101/gad.13.4.424

Lee G, Chambers SM, Tomishima MJ, Studer L. (2010). Derivation of neural crest cells from

human pluripotent stem cells. *Nat Protoc*. 2010;5(4):688-701. doi:10.1038/nprot.2010.35

Lee HC, Choi HJ, Lee HG, Lim JM, Ono T, Han JY. (2016). DAZL expression explains origin and

central formation of primordial germ cells in chickens. *Stem Cells Dev*. 2016;25(1):68-79.

doi:10.1089/scd.2015.0208

Li L, Dong J, Yan L, et al. (2017). Single-Cell RNA-Seq Analysis Maps Development of Human

Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell*. 2017;20(6):858-873.e4. doi:10.1016/j.stem.2017.03.007

Magnúsdóttir, E., & Surani, M. A. (2014). The instructive nature of PGC specification in mice. Development, 141, 245–252. https://doi.org/10.1242/dev.098269

Mamsen, L. S., Brøchner, C. B., Byskov, A. G., & Møllgard, K. (2012). The migration and loss of human primordial germ stem cells from the hind gut epithelium towards the gonadal ridge. International Journal of Developmental Biology, 56(10–12), 771–778. https://doi.org/10.1387/ijdb.120202lm

Marlow, F. (2015). Primordial Germ Cell Specification and Migration. F1000Research, 4. https://doi.org/10.12688/f1000research.6995.1

Martyn I., Kanno T.Y., Ruzo A., Siggia E.D., Brivanlou A.H. (2018). Self-organization of a human organizer by combined Wnt and Nodal signaling. *Nature*. 2018;558(7708):132-135. doi:10.1038/s41586-018-0150-y

Martyn, I., Brivanlou, A. H., & Siggia, E. D. (2019). A wave of WNT signalling balanced by secreted inhibitors controls primitive streak formation in micropattern colonies of human embryonic stem cells. Development, dev.172791. https://doi.org/10.1242/dev.172791

Medrano, J. V., Martínez-Arroyo, A. M., Míguez, J. M., Moreno, I., Martínez, S., Quiñonero, A., … Simón, C. (2016). Human somatic cells subjected to genetic induction with six germ line-related factors display meiotic germ cell-like features. Scientific Reports, 6. https://doi.org/10.1038/srep24956

Miller, C. T. (2017). Why marmosets? *Developmental Neurobiology*, *77*(3), 237–243.

    https://doi.org/10.1002/dneu.22483

Miller, C. T., Freiwald, W. A., Leopold, D. A., Mitchell, J. F., Silva, A. C., & Wang, X. (2016).

    Marmosets: A Neuroscientific Model of Human Social Behavior. *Neuron*, *90*(2), 219–233.

    https://doi.org/10.1016/j.neuron.2016.03.018

Mitsunaga S, Odajima J, Yawata S, et al. (2017). Relevance of iPSC-derived human PGC-like cells

    at the surface of embryoid bodies to prechemotaxis migrating PGCs. *Proc Natl Acad Sci*.

    2017;(15):201707779. doi:10.1073/pnas.1707779114

Molyneaux KA, Stallock J, Schaible K, Wylie C. (2001). Time-Lapse Analysis of Living Mouse Germ

    Cell Migration. 2001;498:488-498. doi:10.1006/dbio.2001.0436

Molyneaux KA, Zinszner H, Kunwar PS, et al. (2003). The chemokine SDF1/CXCL12 and its

    receptor CXCR4 regulate mouse germ cell migration and survival. *Development*.

    2003;130(18):4279-4286. doi:10.1242/dev.00640

Moore, H. D. M., Gems, S., & Hearn, J. P. (1985). Early implantation stages in the marmoset

    monkey (Callithrix jacchus). *American Journal of Anatomy*, *172*(4), 265–278.

    https://doi.org/10.1002/aja.1001720402

Moorman, A., Webb, S., Brown, N. A., Lamers, W., & Anderson, R. H. (2003). Development of

    the heart: (1) formation of the cardiac chambers and arterial trunks. *Heart (British Cardiac*

    Society), *89*(7), 806–814. https://doi.org/10.1136/heart.89.7.806

Morgani SM, Metzger JJ, Nichols J, Siggia ED, Hadjantonakis AK. (2018). Micropattern

differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning. *Elife*. 2018;7:1-35. doi:10.7554/eLife.32839

Murakami, K., Günesdogan, U., Zylicz, J. J., & Walfred, W. C. (2016). Europe PMC Funders Group Europe PMC Funders Author Manuscripts NANOG alone induces germ cells in primed epiblast in vitro by activation of enhancers, 529(7586), 403–407. https://doi.org/10.1038/nature16480.NANOG

Nakamuara, M., Maeda, H., Fujimoto, T. (1991). Behavior of chick primordial germ cells injected into the blood stream of quail embryos. Okajimas Folia Anat Jpn. 1991 Mar;67(6):473-7. DOI: 10.2535/ofaj1936.67.6_473

Nakamuara, M., Yoshinaga, K., Fujimoto, T. (1992). Histochemical identification and behavior of quail primordial germ cells injected into chick embryos by the intravascular route. J Exp Zool. 1992 Apr 1;261(4):479-83. 10.1002/jez.1402610415

Nakamura, T., Extavour, C. G. (2016). The transcriptional repressor Blimp-1 acts downstream of BMP signaling to generate primordial germ cells in the cricket gryllus bimaculatus. Development (Cambridge), 143(2), 255–263. https://doi.org/10.1242/dev.127563

Nakamura Y, Yamamoto Y, Usui F, et al. (2007). Migration and Proliferation of Primordial Germ Cells in the Early Chicken Embryo. *Poult Sci*. 2007;86(10):2182-2193. doi:10.1093/ps/86.10.2182

Nakamura, Y., Kagami, H., & Tagami, T. (2013). Development, differentiation and manipulation of chicken germ cells. Development, Growth & Differentiation, 55(1), 20–40.

https://doi.org/10.1111/dgd.12026

Nakano, M., Arisawa, K., Yokoyama, S., Nishimoto, M., Yamash- ita, Y., Sakashita, M., Ezaki, R., Matsuda, H., Furusawa, S. & Horiuchi, H. 2011. Characteristics of novel chicken embry- onic stem cells established using chicken leukemia inhibitory factor. J. Poult. Sci. 48,64–72.

National Academies of Sciences, E. and M. D. on E. and L. S. I. for L. A. R. R. on S. and W. in L. A. U., Anestidou, L., & Johnson, A. F. (2019). *Care, Use, and Welfare of Marmosets as Animal Models for Gene Editing-Based Biomedical Research*. *Care, Use, and Welfare of Marmosets as Animal Models for Gene Editing-Based Biomedical Research: Proceedings of a Workshop*. National Academies Press (US). https://doi.org/10.17226/25356

Nayernia, K., Nolte, J., Michelmann, H. W., Lee, J. H., Rathsack, K., Drusenheimer, N., … Engel, W. (2006). In Vitro-Differentiated Embryonic Stem Cells Give Rise to Male Gametes that Can Generate Offspring Mice. Developmental Cell, 11(1), 125–132. https://doi.org/10.1016/j.devcel.2006.05.010

Nikolic, A., Volarevic, V., Armstrong, L., Lako, M., & Stojkovic, M. (2016). Primordial germ cells: Current knowledge and perspectives. Stem Cells International. https://doi.org/10.1155/2016/1741072

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., … Surani, M. A. (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature, 436(7048), 207–213. https://doi.org/10.1038/nature03813

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., & Saitou, M. (2009). A Signaling

Principle for the Specification of the Germ Cell Lineage in Mice. Cell, 137(3), 571–584.

https://doi.org/10.1016/j.cell.2009.03.014

O'Rahilly, R., Müller, F., & Streeter, G. L. (1987). Developmental stages in human embryos :

including a revision of Streeter's "Horizons" and a survey of the Carnegie collection.

[Washington D.C.]: Carnegie Institution of Washington.

Park, J. E., & Silva, A. C. (2019). Generation of genetically engineered non-human primate

models of brain function and neurological disorders. *American Journal of Primatology*,

81(2), e22931. https://doi.org/10.1002/ajp.22931

Park, J. E., Zhang, X. F., Choi, S.-H., Okahara, J., Sasaki, E., & Silva, A. C. (2016). Generation of

transgenic marmosets expressing genetically encoded calcium indicators. *Scientific Reports*,

6(1), 34931. https://doi.org/10.1038/srep34931

Phillips, I. R. (1976). The embryology of the common marmoset (Callithrix jacchus). *Advances in*

Anatomy, Embryology, and Cell Biology, *52*(5), 3–47. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/827927

Pijuan-Sala B, Griffiths JA, Guibentif C, et al. (2019). A single-cell molecular map of mouse

gastrulation and early organogenesis. *Nature*. 2019;566(7745):490-495.

doi:10.1038/s41586-019-0933-9

Renfree, M. B., & Fenelon, J. C. (2017). The enigma of embryonic diapause. *Development*

(Cambridge, England), *144*(18), 3199–3210. https://doi.org/10.1242/dev.148213

Renier N, Wu Z, Simon DJ, Yang J, Ariel P, Tessier-Lavigne M. (2014). IDISCO: A simple, rapid

method to immunolabel large tissue samples for volume imaging. *Cell*. 2014;159(4):896-910. doi:10.1016/j.cell.2014.10.010

Saitou, M. (2003). A molecular programme for the germ line specification in mice. Seikagaku. The Journal of Japanese Biochemical Society, 75(1), 42–46.

Sasaki, E. (2015). Prospects for genetically modified non-human primate models, including the common marmoset. *Neuroscience Research*, *93*, 110–115. https://doi.org/10.1016/j.neures.2015.01.011

Sasaki K, Nakamura T, Okamoto I, Yabuta Y, Iwatani C, Tsuchiya H. (2016). The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion Article The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion. *Dev Cell*. 2016;39(2):169-185. doi:10.1016/j.devcel.2016.09.007

Sasaki K, Yokobayashi S, Nakamura T, et al. (2015). Robust In Vitro Induction of Human Germ Cell Fate from Pluripotent Stem Cells. *Cell Stem Cell*. 2015;17(2):178-194. doi:10.1016/j.stem.2015.06.014

Sawada, K., Hikishima, K., Murayama, A. Y., Okano, H. J., Sasaki, E., & Okano, H. (2014). Fetal sulcation and gyrification in common marmosets (Callithrix jacchus) obtained by ex vivo magnetic resonance imaging. *Neuroscience*, *257*, 158–174. https://doi.org/10.1016/J.NEUROSCIENCE.2013.10.067

Schiel, N., & Souto, A. (2017). The common marmoset: An overview of its natural history, ecology and behavior. *Developmental Neurobiology*, *77*(3), 244–262.

https://doi.org/10.1002/dneu.22458

Senft, A. D., Bikoff, E. K., Robertson, E. J., & Costello, I. (2019). Genetic dissection of Nodal and

Bmp signalling requirements during primordial germ cell development in mouse. Nature

Communications, 10(1), 1–11. https://doi.org/10.1038/s41467-019-09052-w

Smith, C. A., Moore, H. D., & Hearn, J. P. (1987). The ultrastructure of early implantation in the

marmoset monkey (Callithrix jacchus). *Anatomy and Embryology*, *175*(3), 399–410.

https://doi.org/10.1007/bf00309853

Sousa, A. M. M., Meyer, K. A., Santpere, G., Gulden, F. O., & Sestan, N. (2017). Evolution of the

Human Nervous System Function, Structure, and Development. *Cell*, *170*(2), 226–247.

https://doi.org/10.1016/j.cell.2017.06.036

Stebler J, Spieler D, Slanchev K, et al. (2004). Primordial germ cell migration in the chick and

mouse embryo: The role of the chemokine SDF-1/CXCL12. *Dev Biol*. 2004;272(2):351-361.

doi:10.1016/j.ydbio.2004.05.009

Strome, S., & Updike, D. (2015). Specifying and protecting germ cell fate. Nature Reviews

Molecular Cell Biology, 16(7), 406–416. https://doi.org/10.1038/nrm4009

Stuart T, Butler A, Hoffman P, et al. (2019). Comprehensive Integration of Single-Cell Data. *Cell*.

2019;177(7):1888-1902.e21. doi:10.1016/j.cell.2019.05.031

Tang, W. W. C., Dietmann, S., Irie, N., Leitch, H. G., Floros, V. I., Bradshaw, C. R., … Surani, M. A.

(2015). A unique gene regulatory network resets the human germline epigenome for

development. Cell, 161(6). https://doi.org/10.1016/j.cell.2015.04.053

Tardif, S. D., Jaquish, C. E., Toal, R. L., Layne, D. G., & Power, R. A. (1998). Estimation of gestational ages in the common marmoset (Callithrix jacchus) from published prenatal growth curves. *Journal of Medical Primatology*, *27*(1), 28–32. https://doi.org/10.1111/j.1600-0684.1998.tb00065.x

Tomioka, I., Nogami, N., Nakatani, T., Owari, K., Fujita, N., Motohashi, H., … Seki, K. (2017). Generation of transgenic marmosets using a tetracyclin-inducible transgene expression system as a neurodegenerative disease model†. *Biology of Reproduction*, *97*(5), 772–780. https://doi.org/10.1093/biolre/iox129

Tremblay, K. D., Dunn, N. R., & Robertson, E. J. (2001). Mouse embryos lacking Smad1 signals display defects in extra-embryonic tissues and germ cell formation. Development, 128(18), 3609–3621.

van Vuuren, S. H., Damen-Elias, H. A. M., Stigter, R. H., van der Doef, R., Goldschmeding, R., de Jong, T. P. V. M., … Pistorius, L. R. (2012). Size and volume charts of fetal kidney, renal pelvis and adrenal gland. *Ultrasound in Obstetrics & Gynecology*, *40*(6), 659–664. https://doi.org/10.1002/uog.11169

Verwoerd-Dikkeboom, C. M., Koning, A. H. J., van der Spek, P. J., Exalto, N., & Steegers, E. A. P. (2008). Embryonic staging using a 3D virtual reality system. *Human Reproduction*, *23*(7), 1479–1484. https://doi.org/10.1093/humrep/den023

Vincent, J. J., Li, Z., Lee, S. A., Liu, X., Etter, M. O., Diaz-Perez, S. V, … Clark, A. T. (2011). Single cell analysis facilitates staging of Blimp1-dependent primordial germ cells derived from mouse embryonic stem cells. PloS One, 6(12), e28960.

https://doi.org/10.1371/journal.pone.0028960

Warmflash A, Sorre B, Etoc F, Siggia ED, Brivanlou AH. (2014). A method to recapitulate early

embryonic spatial patterning in human embryonic stem cells. *Nat Methods*. 2014;11(8):847-

854. doi:10.1038/nMeth.3016

Weismann, A. (1893). The Germ-Plasm. Harvard Collection.

Witschi E. (1948). Migration of germ cells of human embryos from the yolk sac to the primitive

gonadal folds. Contr Embryol Carnegie Inst. 1948;209:67–80.

Wodak R, Chilton P, Nelson DN, Norman CRC, Fairclough L, Wilson (2007). J. Pro O F Pro.

Language (Baltim). 2007;x(x):1-11. doi:10.1002/aja.1001950404

Wolf FA, Angerer P, Theis FJ. (2018). SCANPY: Large-scale single-cell gene expression data

analysis. *Genome Biol*. 2018;19(1):1-5. doi:10.1186/s13059-017-1382-0

Wolf FA, Hamey FK, Plass M, et al. (2019). PAGA: graph abstraction reconciles clustering with

trajectory inference through a topology preserving map of single cells. *Genome Biol*.

2019;20(1):1-9. doi:10.1186/s13059-019-1663-x

Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., … Saitou, M. (2008). Critical

function of Prdm14 for the establishment of the germ cell lineage in mice. Nature Genetics,

40(8), 1016–1022. https://doi.org/10.1038/ng.186

Yamashiro C, Sasaki K, Yabuta Y, et al. (2018). Generation of human oogonia from induced

pluripotent stem cells in vitro. *Science 2018 (80-)*. 2018;362(6412):356-360.

doi:10.1126/science.aat1674

Yoney A, Etoc F, Ruzo A, et al. (2018). WNT signaling memory is required for ACTIVIN to

function as a morphogen in human gastruloids. *Elife*. 2018;7:1-28. doi:10.7554/eLife.38279