UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Heterogeneity in migrating primordial germ cells

Permalink
https://escholarship.org/uc/item/1kw6k2cm

Author
Jaszczak, Rebecca Garrett

Publication Date
2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Heterogeneity in migrating primordial germ cells

by
Rebecca Jaszczak

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Todd Nystul
B55FB5CFA606452...

Todd Nystul
Chair

DocuSigned by:
Diana Laird
D2C1E9F4D15F7...

Diana Laird

DocuSigned by:
Aaron Tward
D2A1E9F4D15F7...

Aaron Tward, MD, PhD

DocuSigned by:
Jingjing Li
8AE0E67221704E1...

Jingjing Li

Committee Members

Dedicated to my husband, Jacob, for his unwavering support and love.

And secondly to my grandmother, who taught me how to say research.

Acknowledgements

Firstly, to Diana, my esteemed mentor. Thank you for taking a chance on me. I turned up in your office, a motivated wet lab scientist, who told you I wanted to learn bioinformatics in a developmental biology context. You believed in my vision. Your trust has meant so much to me. You provided me many collaborative and independent ventures to develop my dry lab skill set. You believed in me and my abilities wholeheartedly, even as I was developing the skills our lab needed. Your patience through my training gave me confidence to take on bigger projects. Your trusting approach allowed me to get my 'bioinformatics legs' under me, and ready to tackle new problems. When I made mistakes, or have struggled, I count on you in my corner to cheer me on. When I have come up short, I have looked to your boundless optimism to pull me through a tough time. I am profoundly grateful for your mentorship approach, which provided a supportive environment but didn't shelter me from the opportunity to develop tenacity and grit in the face of difficulty. It is that grit that I have called on many times to surmount the trials of finishing a PhD. Thank you for all you invested in me. I hope to make you proud.

To Todd Nystul, my thesis chair, thank you for your leadership in committee meetings guiding me toward degree completion. I have appreciated your enthusiasm for my growth as a scientist, and your support as I started new techniques - even inviting me to your own lab to use your reagents during preliminary tests. To Aaron Tward, I am so grateful for the conversations we've had in your office about data analysis. I am so grateful for your straight shooter attitude, and your concise feedback which has tremendously lifted the level of my thesis work. To Jingjing Li, I am so grateful for our one on one meetings which have helped me overcome mental blocks in the ways I have thought about my data. My work has so benefitted from your innovating and insightful comments regarding future directions and new experiments.

To my Laird labmates, you have made this possible. Having a brilliant and strong team of scientists alongside me has synergized with my work. These results are more than simply the sum of their parts, and I am grateful to be a part of such a team. I could not have accomplished anything alone. To Ripla, who taught me to be a candid and polite scientist. To Bikem, who helps me not to pipette the pink solution, I have loved our collaborations. I'm so grateful for your example of hard work. To Dan, my late night work desk buddy and great happy hour friend. To Boris, who collaborated in retrotransposon analysis as well as helped create the Laird Lab fermentation club. To Suping, teaching us how to fold traditional dumplings. To Lina, who helped me start to make sense of an overwhelming list of gene candidates for follow-up, and also has the best Berkeley venue/food recommendations. To Jen and Lauren, whose ace sectioning skills built out an amazing slide bank for us to test markers from our single cell data, and who kept our mouse work going through their excellent PCR efforts. To Steven, my grad school year mate, and our much needed vent sessions over wine as well as excellent science conversations. To Noa, who has helped me troubleshoot *in situ*s and shared plugs to help me meet graduation timelines. To Caroline, Ernesto, Eliza, Mariko, Jing, Jocelyn, and JB - your passion, enthusiasm, and excitement coalesce into a beautiful energy you bring to the lab. The future of the Laird Lab is in your hands, and it looks so bright.

To my collaborators and co-authors, thank you for sharing my vision. Those of you at UCSF and beyond have helped me to raise my work beyond anything I could have accomplished alone. Thanks especially to Aparna Bhaduri, who taught me how to make single cell libraries and gave me my first shell script. You unlocked the world of coding for me, and I will forever be grateful for your time and efforts. I'm so thankful to have been a part of the cumulative effort from so many scientists along the way.

To my therapist Dr. Roehrig, who gave me the tools to climb out of the hole of despair that was my third and forth year. I still recite coping cards. To Dr. Ozores, who listened when no-one else would. I have never felt so understood by a psychiatrist. To Dr. Rajpoot, who helped me view the pressing anxiety of graduation as surmountable and gave me to the tools to do so. I am so grateful.

To my grad school cohort, especially Dan, Casey, Jess, Ariane, and Nate. Our Bitter Betty "wine down" happy hours sustained me towards the end of grad school. Transitioning to zoom during the pandemic was my social lifeline. I love you all and am so excited for your successes. Dan, I have enjoyed our friendship since we started studying Immuno together, and then discovered so many shared interests like DND and Rupaul's Drag Race. Without your support and companionship in learning how to code, I wouldn't be where I am. I admire you greatly and cannot wait to keep learning data science together. Nate, your insights on science life and real life have helped me navigate graduate school and beyond. I deeply admire your thoughtful scientific mind, and the way you care for those around you. Your loving friendship has been my rock these past few years. I could not have done this without you.

I am also so grateful to my DND friends, who I've played with in person here in SF (Nate, Dan, Natanya, Emily, Ashley, Bonnie, Tawny, Steph, Tanya, Marc, Allan, Tai) and online with international friends all over the world. Y'all have given me a breath of fresh, non scientific air. To my Dice Nation Zoom Crew, crafting together has kept me sane. To the Blob, your unwavering friendships in the toughest of times has sustained me. I cannot wait until our next West Marches adventure.

To my parents, I am so grateful for your lifelong support. I hope I make you proud. To my mom, who has always appreciated my deep and free thinking nature. To my father, who instilled in me ideal of "measure twice, cut once" a mantra for the woodshop or the science bench. To

my mother and father in law, who have celebrated me and encouraged me. To my siblings by blood and by marriage: Danny, Hannah, Pat, Andee, Becca, Eric, Ben, Meghan, David, Rana, Peter, Isabella, and Katie. I love you all so much.

Finally, to my most stalwart supporter. My email proofreader, my abstract editor, my "please help this sentence make sense" wordsmith. My amazing chef who makes the best dinners. My "I'm dropping you dinner off at lab, because it's 9pm and you need nourishment to be an amazing scientist." My craft buddy. My fisherman. The best choice I ever made. I am deeply thankful for the support of my husband Jacob. Dr. Jacob, you know firsthand the stress of getting a PhD, and I have leaned on your guidance every step of the way. I'm confident that my success also belongs to you. You are my ballast, my stability. Thank you for giving me a place of safety, so that I could conquer my PhD and beyond.

Contributions

The work in this dissertation was performed under the direct supervision of Dr. Diana Laird, Ph.D. Further guidance and valuable insight came from thesis committee members Dr. Todd Nystul Ph.D. (chair), Dr. Aaron Tward M.D. Ph.D, and Dr. Jingjing Li, PhD.

Chapter 1 of this work contains text and figures largely adapted from a published review article: Daniel H Nguyen*, **Rebecca G Jaszczak***, Diana J Laird. "Heterogeneity of primordial germ cells," Review Curr Top Dev Biol. 2019;135:155-201. doi: 10.1016/bs.ctdb.2019.04.009. Epub 2019 May 14.

*authors contributed equally to this work

DHN, RGJ, and DJL conceptualized, wrote, and did review and editing of the manuscript.

Chapter 2-4 of this work are largely adapted from a manuscript in preparation:

Rebecca G Jaszczak, Jocelyn Crawford, Diana J Laird. "Revealing intrinsic heterogeneity and niche-regulated development in migratory primordial germ cells with single cell transcriptomics."

RGJ and DJL designed the study and analyzed the data. RGJ performed the wet lab experiments. RGJ and JC performed the computational analysis. RGJ and DJL wrote the paper.

Chapter 5, Appendix B done in collaboration with Aidan S. Marshall and Dr. Nick S. Jones.

Heterogeneity in migrating primordial germ cells

Rebecca Garrett Jaszczak

Abstract

Mouse primordial germ cells (PGCs) are specified in the posterior embryo and migrate through hindgut endoderm and dorsal mesentery to the genital ridge. Although they move individually, PGCs do not migrate synchronously and are not all successful migrants. When migration ends at E11.5, we find approximately 6% of mouse PGCs outside the genital ridge; lacking survival cues, they are poised to undergo apoptosis. What heterogeneities distinguish a successful PGC, able to colonize the gonad, from a failed migrant? Are heterogeneities shared between clonally labeled mother and daughter cells? Our previous data show that mouse PGCs cultured ex vivo mount variable responses to migration cues, suggesting PGC behavioral variation may arise intrinsically; we also showed that proliferation is regulated by the somatic niche, increasing proliferation of PGCs as they progress along the migratory route. These data indicate that migratory speed and PGC-niche interactions may influence PGC gonadal colonization. To characterize intrinsic heterogeneities in migrating mouse PGCs, we conducted single cell RNA sequencing of purified PGCs and surrounding somatic niche for E9.5, E10.5, and E11.5 migratory timepoints. For E9.5 and E10.5, we generated libraries after splitting PGCs into anterior "leading" vs posterior "lagging" populations. Upon bioinformatic integration of all libraries, we analyzed differential gene expression probing positional differences as well as age related differences. We identified an E11.5 subcluster that may represent lagging PGCs transitioning medial/midline aorta-gonad-mesonephros tissues to the gonadal ridge. This cluster expresses imprinting genes. We have shown in vivo through HCR probes that cells expressing *Igf2*, a top imprinting gene marker of this cluster, are localized in the extragonadal space. We also probed receptor ligand interactions in

migratory PGCs, and found several known receptor-ligand pairs, like Cxcl12-Cxcr4 survival signals. Interestingly, we identified potentially novel regulators of migrating germ cells, including ephrins. Additionally, we reanalyzed some human fetal PGCs and found that Slit-Robo signaling was implicated in both mouse and human PGC-somatic cell interactions. Studying transcriptional profiles in migratory PGCs across different ages reveals signaling mechanisms associated with specific niches as well as the extent of heterogeneity between cells.

Additionally, we investigated whether mitochondrial heteroplasmies might contribute to migratory heterogeneities we observed. Using our single cell libraries, we were able to sequence some regions deeply enough to call variants. In general, we found that at either timepoint, anterior PGCs did not have different levels of heteroplasmy compared to posterior PGCs. However, we did observe that heteroplasmies increased from E9.5 compared to E10.5. In early human PGCs, it is known that nonsynonymous mutations decrease over time; in mouse, however, we saw that nonsynonymous mutations increased at E10.5 compared to E9.5.

Finally, using fluorescent lineage labeling, we can test whether these heterogeneities are clonal. Interestingly, cells labeled while in the hindgut still localize near their clonal family in the genital ridge. Studying transcriptional profiles in migratory PGCs across different ages is revealing signaling mechanisms associated with specific niches as well as the extent of heterogeneity between cells.

Table of Contents

Chapter 1 Heterogeneity of primordial germ cells 1

1.1 - Introduction to Sexual Reproduction at the Single Cell Level 1

1.2 - Heterogeneous Phenotypes of Primordial Germ Cells 3

Primordial germ cell specification and heterogeneity. 7

Primordial germ cell migration and heterogeneity 11

Proliferation of Primordial Germ Cells and Heterogeneity 15

Primordial Germ Cell Survival and Heterogeneity 17

1.3 - Determinants of Heterogeneity 24

Genetic Determinants 24

Extrinsic sources of heterogeneity: environment, niche, and timing 36

Buffers of heterogeneity 37

1.4 - Selection and Heterogeneity 38

Defining selection in the developing germline 39

How does PGC fitness relate to fitness of gametes or individuals? 41

1.5 - Chapter Conclusion 43

Chapter 2 Revealing intrinsic heterogeneity and niche-regulated development in migrating primordial germ cells 45

2.1 - Introduction to migration of primordial germ cells in mouse 45

2.2 - Migrating germ cell isolation and single cell RNA sequencing library creation of migratory timepoints 49

2.3 - Predicting cell signaling between migrating mouse primordial germ cells and their supporting niche 55

2.4 - Identifying PGC-niche receptor-ligand pairs shared between human and mouse 59

2.5 - Characterizing transcriptional differences in mouse E9.5 and E10.5 Anterior and Posterior migrants 67

2.6 - Imprinted genes mark a subset of germ cell bioinformatically, and are located extragonadally in the tissue 72

Cell cycle regression unmasks other heterogeneity 72

Identification of bioinformatic PGC outliers in scRNA seq data 77

Location of imprinted gene expressing PGCs in the extragonadal spaces of the AGM 77

2.7 - Chapter Conclusion 82

Chapter 3 Discussion of discovered PGC heterogeneities 84

3.1 - Summary of mouse single cell RNA sequencing field after our contribution 84

3.2 - Discussion of heterogeneities found and their implication 85

Mouse compared to human 85

The migratory niche 86

3.3 - The impact of imprinting and future directions 90

3.4 Conclusion 92

Chapter 4 Methods 93

Mouse Crosses 93

E9.5 dissociation 93

E10.5 dissociation....94
E11.5 aorta-gonad-mesonephros dissociation....94
Cell Sorting....95
Library Creation....96
scRNA Analysis....96
Seurat and SoupX....96
Somatic Labels....98
CellphoneDB....98
NicheNetR....99
HCR....99
Image analysis for whole-mount stained ovaries....100
Chapter 5 Appendices....101
Appendix A: Clonal heterogeneity in migrating mouse primordial germ cells....101
Introduction: Migration and Clonal Heterogeneity....101
Results: Clones migrate together....102
Discussion and Future Directions....103
Appendix B: Surveying mitochondrial heteroplasmy in migrating mouse primordial germ cells....107
Introduction to mitochondrial selection in germ cells....107
Results: Mouse PGC heteroplasmy increases across time....108
Discussion and Future Directions....112

References..... 115

List of Figures

Figure 1.1 Multiple cell state transitions are possible within a single cell type..... 4

Figure 1.2 Readouts of PGC Heterogeneity 6

Figure 1.3 Determinants of PGC Heterogeneity 25

Figure 2.1 Single cell RNA sequencing of migratory germ cells and their somatic niche across developmental time 51

Figure 2.2 (Supplemental for 2.1) E11.5 biologic replicates integrate well, and somatic identification scheme can be used to label somatic cell populations 54

Figure 2.3 Interrogation of the somatic expressed ligands originating from the migratory niche and the companion receptors in migrating mouse PGCs with CellPhoneDB 57

Figure 2.4 Cross-species comparison of clustering and receptor-ligand pairs in similarly-staged human migratory PGCs..... 60

Figure 2.5 (Supplemental for 2.4) Human migratory germ cells display temporally-specific gene expression 62

Figure 2.6 (Supplemental for 2.4) Interrogation of the somatic expressed ligands originating from the migratory niche and the companion receptors in migrating human germ cells with CellPhoneDB..... 64

Figure 2.7 In mouse and human, the state change between early migrating PGCs and gonadal ridge colonizers involves Slit-Robo signaling..... 66

Figure 2.8 Identification of spatially-specific genes in mouse PGCs from anterior and posterior sin the embryo 69

Figure 2.9 (Supplemental for 2.8) Cell cycle phase regression reveals additional mouse PGC heterogeneities 75

Figure 2.10 Elevated expression of imprinted genes marks extragonadal PGCs at E11.5..... 79

Figure 5.1 Clonally related PGCs localize together in clonal groups in the gonadal ridge 105

Figure 5.2 Assessing mitochondrial heterogeneity in E9.5 and E10.5 migrating PGCs 110

List of Tables

Table 2.1 Differential genes from Figure 2.8D-E sorted according to position and function.....72

Table 2.2 Top 30 Differentially Expressed Genes in E11.5 Cluster 3 and 4.....81

Chapter 1 Heterogeneity of primordial germ cells

1.1 - Introduction to Sexual Reproduction at the Single Cell Level

Sexual reproduction culminates with the fusion of two gametes to generate a zygote, briefly returning life to a fleeting, single-celled state. This simple cellular arithmetic, however, belies the momentous journeys undertaken by each gametic partner to enable this union. As we consider the epic reproductive development that precedes the production of this unicellular zygote, it is most appropriate that our focus begin, humbly, at the level of the single cell. Though the rise of metazoans introduced the dazzling complexity of diversified cell types and tissues, all multicellular systems are still fundamentally reducible to single cells - the smallest unit of biological structure.

Recent advances in single-cell technologies are revealing powerful insights about the identities of individual cells that compose a lineage or tissue, but historically, much of biology has begun by asking questions of single cells. From Anton van Leeuwenhoek at his microscope to Paul Ehrlich and his histological dyes, we have long analyzed differences that exist from cell to individual cell. The advantage of whole-mount or section immunofluorescence imaging lies in the ability to resolve multiple cells in a tissue while preserving individual differences for measurement (Levsky and Singer, 2003). Such powerful resolution permits a more comprehensive understanding of cellular interactions and a sensitivity to cell-to-cell variation. In contrast, analytical methods like Western blots, PCR, or bulk sequencing provide only population-level insights under the assumption that behavior is uniform. Even within single lineages, it is increasingly understood that significant variation can exist among seemingly identical cells (Altschuler and Wu, 2010; Raj and van Oudenaarden, 2008) and that this heterogeneity can have far-ranging functional consequences.

With this principle in mind, we can begin tracing the steps that generate gametes - themselves single-celled carriers of the genome. Each germ cell must endure a lengthy and complex development that begins shortly after fertilization when the germline is set aside from the somatic lineages with the formation of primordial germ cells (PGCs). The path from PGC to gamete is highly conserved, demonstrating the evolutionary importance of reproductive development (Nieuwkoop and Sutasurya, 1979). Early PGC development is especially dynamic and requires transition through differentiated states, interaction with diverse cellular environments, and processing a multitude of signals. This complexity can amplify differences among individual cells to yield varied cell fates. For germ cells that are tasked with creating gametes, those that survive this developmental crucible secure the ultimate biological prize: propagation of their genetic identity.

The concept of selective events acting on variation within a population has long been appreciated in biology, albeit on the level of organisms in Darwinian natural selection. These same principles are also applicable at the cellular level as well, and we argue that they can govern the fate of PGC heterogeneity during development (Buss, 2006; Laird et al., 2005b; Weissman, 2015). The diversity of challenges posed to PGCs in this period may function as developmental selection that acts upon cellular variation in the germline in a manner analogous to natural selection with organisms. This chapter will consider how early events in germ cell development with a particular emphasis on how heterogeneity manifest in germ cells, how heterogeneity arises, and its impact on the germline and reproductive fitness.

1.2 - Heterogeneous Phenotypes of Primordial Germ Cells

Heterogeneity is a fundamental property of biological systems consisting of multiple units such as cells. Even though cells may share a common lineage, it is increasingly recognized that substantial variation exists even within cell types. As our ability to measure cellular phenotypes and behaviors improves, we can distinguish rare cell subpopulations in tissues such the intestinal crypt (Grün et al., 2015) as well as canvass the broad diversity that exists in chaotic intratumoral settings (McGranahan and Swanton, 2017). In hematopoiesis, these subpopulations can represent increasingly smaller organizational units that possess distinctive behaviors and fates (Schroeder, 2010).

While heterogeneity has been studied in many homeostatic contexts such as the adult stem cells of the intestine or blood, its role in germline development is a unique and interesting consideration. The establishment of PGCs at the outset of embryogenesis, as the first lineage in many different organisms, mandates a lengthy process of development through which a subset of germ cells become gametes. The dynamic and multifaceted challenges encountered by PGCs during fetal development present numerous opportunities to assess heterogeneity that may exist in nascent PGCs. As PGCs go on to establish the germline and generate gametes, how heterogeneity is resolved by developmental selection can prove especially influential for transgenerational inheritance.

Definition of Cell Type

This section aims to establish a conceptual framework to define and identify heterogeneity as it pertains to the germline and reproductive fitness. Examination of evidence for heterogeneity in specific aspects of PGC development in several model organisms will uncover general principles that underlie the unique challenges and adaptations utilized by germ cells to address

variation. Thus, definition of common terms and how they will be used in this manuscript becomes necessary.

Although cell type has been defined by many criteria such as morphology, location, lineage history, and distinct transcriptional profile (Altschuler and Wu, 2010; Muller-Sieburg et al., 2012; Seoane, 2017) we treat the germ cell as a functional entity theoretically capable of giving rise to gametes of either sex (Figure 1.1). This definition encompasses changes in morphology, location, and gene expression involved in PGC development, as well as the different modes of specification and routes of PGC migration between organisms. We define PGCs as the committed but yet sexually undifferentiated embryonic precursor of the gamete. Although we acknowledge that PGC embryonic origins may have evolved multiple times, we draw comparisons between PGCs in different organisms based on their operational similarities and parallel constraints.

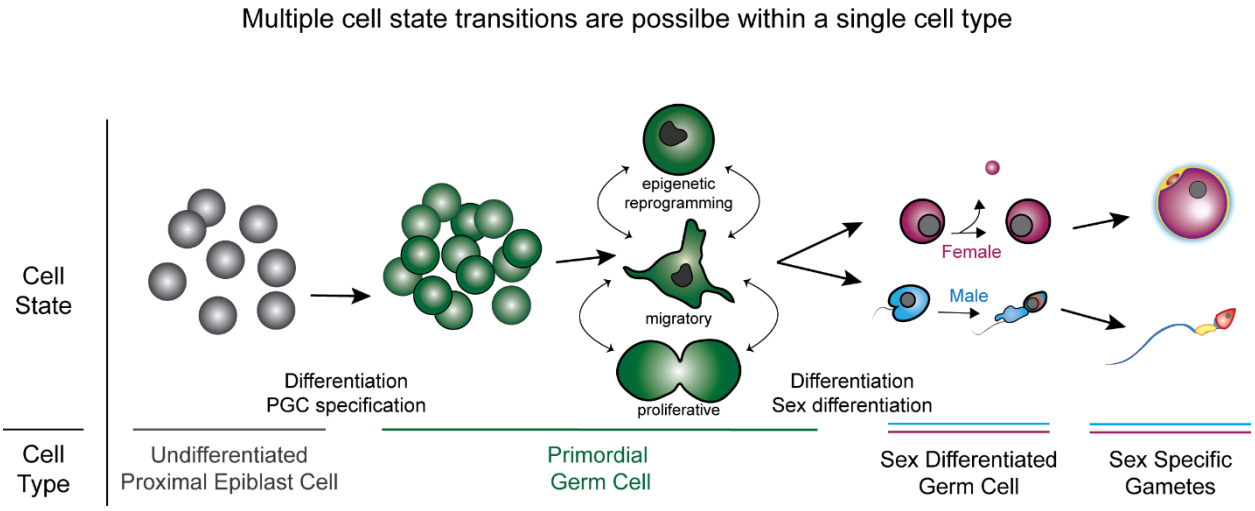


Figure 1.1 Multiple cell state transitions are possible within a single cell type

PGCs transit through multiple cell states during development but remain a single cell type defined by an uncommitted sexual identity. Cell states are not necessarily discrete nor mutually exclusive. We consider germ cells to encompass all cell types with gametogenic potential, including PGCs and sex-differentiated oogonia and prespermatogonia.

Definition of Cell State

A similarly nebulous term, we consider cell state to be a stable but transient set of activities, a potency, a distinct transcriptional profile, and/or a set of epigenetic marks that defines a subpopulation of cells within a particular cell type. Similar to the poised, active, and silenced states of embryonic stem cells associated with distinct sets of epigenetic marks and gene expression (Young, 2011), we consider that PGCs can occupy different states as well as transition back and forth between them. Cell state and cell type are interdependent; PGCs are a cell type, but PGCs in the embryo are a collection of cells in multiple states. PGC differentiation to male gonocyte or oocyte involves a change of cell type that may occur through a series of cell state transitions.

Definition of Heterogeneity (of Cells)

Synonymous with cell individuality and cell diversity, cell heterogeneity is treated here as a biological difference between cells of the same type (as opposed to noise that arises from experimental or technical factors). Heterogeneity can be detected as distinct cellular behaviors, including morphology or function (such as proliferation, capacity for differentiation, or migration); these behaviors may be reflected in transcriptional state or epigenetic profile. Determinants or causes of cellular heterogeneity may include genetic changes, epigenetic changes (which is relevant for PGCs more than other cells), non-genetic changes that may or may not be stochastic, or extrinsically-induced changes. These diversities are schematized in Figure 1.2.

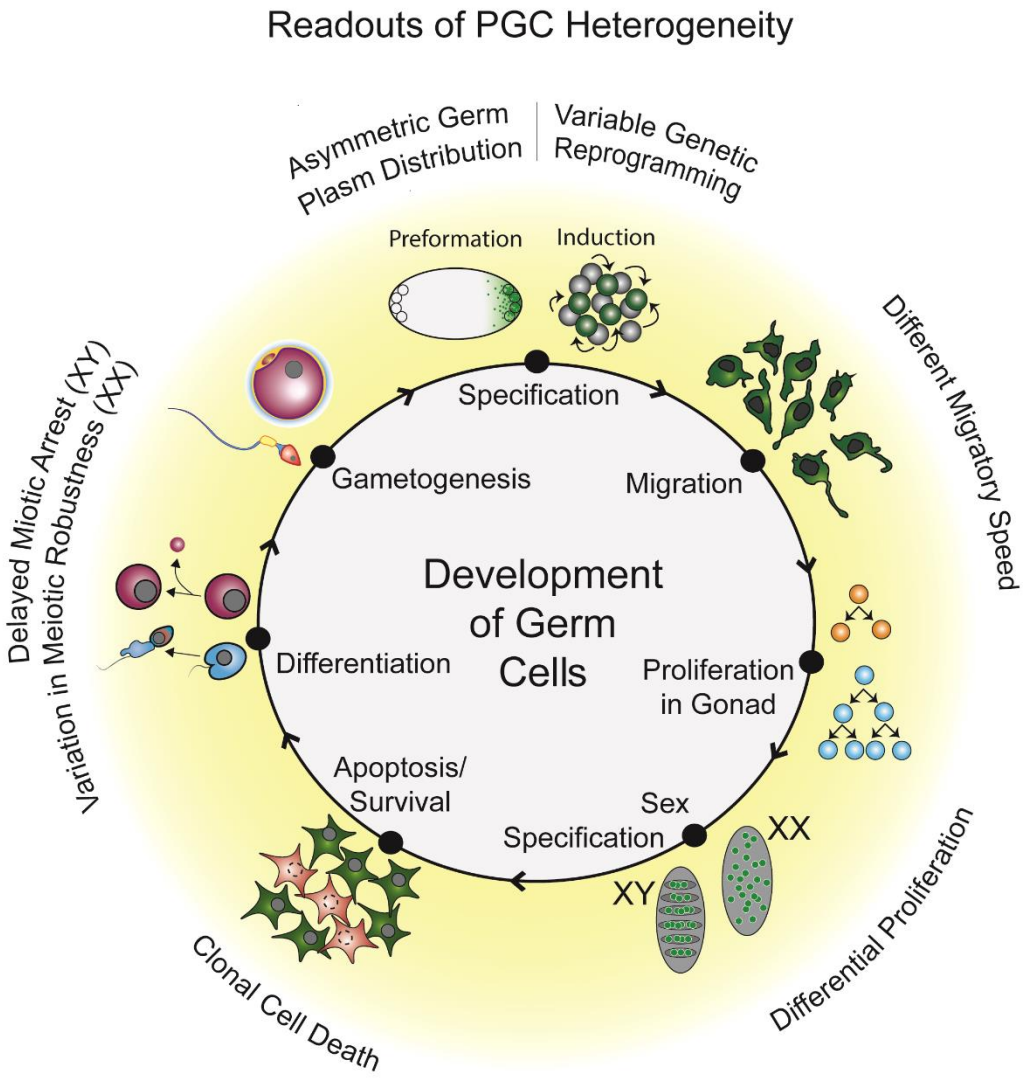


Figure 1.2 Readouts of PGC Heterogeneity

Snapshots of the reproductive life cycle evidence potential readouts of heterogeneity. These readouts can be broadly surveyed across many model organisms. Key phases of potential heterogeneity occur during specification, migration, proliferation, sex determination, survival, differentiation, and gametogenesis. Specific readouts may be applicable several cell states.

Primordial germ cell specification and heterogeneity.

This section considers the potential for heterogeneity to arise between PGCs during their specification, as well as specific examples. Among metazoans, there are two modes by which PGCs are specified: preformation and induction (Figure 1.2, top). Also termed epistasis, preformation confers a germline fate by cytoplasmic inheritance of maternally-derived organelles composed of proteins and RNAs termed germ plasm, germ granules, polar or P granules; mechanisms of germline preformation have been reviewed extensively by others (Extavour and Akam, 2003; Saitou et al., 2005; Solter, 2016). Of the model organisms commonly studied, Caenorhabditis elegans, Drosophila melanogaster, Xenopus laevis, and Danio rerio all generate PGCs through early preformation. By contrast, mammals, urodele amphibians and crickets specify germline by induction from somatic cells, which involves signaling from the surrounding niche that leads to a binary fate decision (Ewen-Campen et al., 2012; Lawson et al., 1999; Nieuwkoop, 1947; Tam and Zhou, 1996); underlying molecular mechanisms are well reviewed elsewhere (Solter, 2016).

Does the mode of specification limit the extent of heterogeneity that is possible between PGCs? It is not unreasonable to assume that the germ plasm promotes homogeneity from the inception of the PGC lineage by preformation. The inclusion of proteins and transcripts in the germ plasm regulates the levels of both in the cytoplasm, particularly since PGC specification precedes zygotic transcription in model organisms with preformation, and transcription remains suppressed in nascent PGCs even after beginning in the soma (Strome and Lehmann, 2007). It might be expected that germ plasm promotes PGC homogeneity on a continual basis given the capacity for movement through specialized cytoplasmic bridges between germ cells (Extavour and Akam, 2003). During PGC specification, it has been proposed that niche functions to shield

PGCs from somatic fate signals; this idea is reinforced by the observation that PGCs arise via preformation or induction paradigms in sequestered locations, such as the posterior pole in fly embryos (Eddy, 1975), near the extraembryonic tissues in mammals and birds (Ginsburg and Eyal-Giladi, 1987, 1986; Ginsburg et al., 1990), or on the periphery of zebrafish embryos (Raz, 2002). However, in the context of this discussion, we might speculate that the niche for specification instead protects PGCs from signals that would promote heterogeneity. On the other hand, the inductive mode is arguably more conducive to producing heterogeneous PGCs since the process of fate conferral by signaling can occur over a more protracted period of time in principle. The comparative effects of these two mechanisms of PGC fate determination may have relevance in an evolutionary context. Examining patterns of PGC specification across many species precipitated the notion that preformation is derived whereas induction is ancestral (Extavour and Akam, 2003). We might ask, then-- if germ plasm indeed promotes uniformity, whereas induction promotes greater diversity of PGCs-- what benefit does reduced heterogeneity confer upon those species that acquired preformation?

PGC specification by preformation

We discuss several examples of heterogeneous behaviors of PGCs associated with specification by preformation. In this mode of development, germline establishment requires asymmetrical distribution of germ plasm in the oocyte and zygote (Hird et al., 1996), followed by asymmetric inheritance during early divisions into the germline precursors and PGCs (Eddy, 1975; Mahowald, 2001, 1962). However, subsequent development of PGCs involves a switch from asymmetric to symmetric inheritance of germ plasm to daughter cells. A key question is whether minor asymmetries in either the allocation of germ plasm between founder PGCs or the subsequent allocation to PGC daughter cells have phenotypic consequences. A recent study in Drosophila suggests that indeed some founder PGCs are short-changed in their inheritance.

Quantifiable variation was observed in the levels of maternal mRNAs Nanos and Wunen associated with the germ plasm between individual PGCs. Importantly, the levels of both were not arbitrary, but highly correlated with the probability of subsequent germ cell death, suggesting that Wunen is an indicator of germ plasm quality (Slaidina and Lehmann, 2017). Similarly, in Caenorhabditis elegans, the asymmetric segregation of cytoplasmic determinants known as P granules occurs in the P4 ancestor cell of the two PGCs known as Z2 and Z3 (Sulston et al., 1983). Surprisingly, PGCs were specified and fertility was maintained in a mutant with symmetric distribution of P granules, suggesting that the P granules themselves are not sufficient for the worm germline (Gallo et al., 2010).

Further insight into the critical components of germ plasm for conferring a germ cell fate and potentially generating diversity of PGCs comes from recent work in Drosophila. The germ plasm component oskar is both necessary and sufficient for specifying PGCs in flies, and plays a deterministic role in the number of precursor pole cells (Ephrussi and Lehmann, 1992; Ewen-Campen et al., 2012; Smith et al., 1992). A surprising function of one isoform of oskar is to recruit mitochondria to nascent PGCs via the actin cytoskeleton; a reduction in the number of PGCs and diminished downstream oogenesis in mutants lacking this oskar long isoform implicates mitochondrial inheritance and number in the competence and normal development of PGCs (Hurd et al., 2016). In zebrafish, the maternal gene bucky ball functions in a parallel manner to oskar via interactions with microtubule networks to assemble the germ plasm, which involves aggregating mitochondria into a structure known as the Balbiani body. Ectopic PGCs are produced by ectopic expression of bucky ball, but how they may differ from properly localized germ cells is unclear (Bontems et al., 2009; Marlow, 2015). Similarly, it is not known how mislocalized PGCs that result from ectopic expression of oskar in flies might be different than those specified according to established patterning, but an avenue of inquiry will be

interrogate potential differences in mitochondria between individual PGCs as related to subsequent fate and function.

PGC specification by induction

Of the organisms that specify their germline through induction and signaling from somatic cells, Mus musculus is one of the most studied. Despite this fact, much remains unknown about how heterogeneity may impact murine reproductive potential. Early fate mapping studies of single cells from E4.5 the mouse embryo produced an important negative result: no single cell of the inner cell mass exclusively gave rise to PGCs (Gardner and Rossant, 1979). Mouse PGCs are induced in the epiblast by BMPs from the surrounding extraembryonic tissues (Fujiwara et al., 2001; Lawson et al., 1999; Ying et al., 2000). The earliest known marker of the PGC fate is Blimp1 (also known as Prdm1), which first appears in the epiblast at E6.25. Although other genes of PGC commitment do not come on until ~E7, lineage tracing with Blimp1-Cre suggests that all cells in the region of the allantois that have expressed Blimp1 become PGCs (Ohinata et al., 2005). This approach does not distinguish between cells that descend from the first to express Blimp1 and those derived from later onset of Blimp1 in subsequent recruits from the epiblast, hence the number of PGCs specified directly from epiblast in mice remains elusive. Several groups have attempted to identify the number of founder PGCs which actually contribute to germ cell lineages. Examination of the earliest Blimp1 transcript found four to eight Blimp1-positive cells in most proximal epiblast which grow into a tight cluster of approximately 20 cells at the mid-streak (MS) stage (Saitou et al., 2005). Approaches using retroviruses or generating chimeras with different-colored embryonic stem cells delivered to the blastocyst have estimated that all PGCs derive from a founder population of 3 or 4 cells (Soriano and Jaenisch, 1986; Ueno et al., 2009). It is unclear whether the growth of cell number from 4-5 to the observed BLIMP1 positive cells in the epiblast to the ~40 PGCs observed at e7 (Ginsburg et al.,

1990) occurs through induction of new cells into the germline from the epiblast, or proliferation of the previously induced 4-5 PGCs. The mouse germline undoubtedly derives from a narrow set of ancestral founders, and the yet unclear size of this founder population has potential implications for heterogeneity of PGCs (Zheng et al., 2005).

Further insight into the question of whether all PGC founders are equivalent will emerge from more sophisticated lineage tracing of the germline, such as genetic barcoding. A glimpse into potential heterogeneity comes from knockout studies. Genetic ablation of Blimp1 leads to the specification of PGCs with incomplete repression of somatic genes and activation of pluripotency genes, but this phenotype varies in severity depending on the mouse genetic background (Avilion et al., 2003; Yabuta et al., 2006). These results raise the possibility that variable levels of Blimp1 between early mouse PGCs could lead to differences in their capacity to fully realize their fate, which could potentially impact their behavior and function. More recently, the homeobox transcription factor Otx2 was identified as an earlier repressor of the PGC fate that turns off 12-24 hours before the appearance of Blimp1. Whereas overexpression of Otx2 decreases the number of PGCs, the absence of Otx2 augments the number of PGCs, indicating a negative correlation between PGCs and Otx2 expression (Zhang et al., 2018). Although these experiments lacked the sensitivity to detect low levels of Otx2, the question remains whether PGC precursors that incompletely extinguish Otx2 become PGCs with low levels of somatic gene expression that parlay into measurable differences in behavior (Laird, 2018).

Primordial germ cell migration and heterogeneity

In many organisms, PGCs are specified far from their permanent tissue niche, and must walk a long distance through growing and developing somatic compartments to find the gonad

primordia, which will furnish a niche that promotes their survival and differentiation. To accomplish this feat, PGCs must be responsive to external migration cues as well as able to navigate their way through temporary migratory niches, recognize the gonad niche, and stop migrating when they arrive. Navigation of these challenges by each individual germ cell may amplify initially small differences, such that the migratory process reveals heterogeneity.

From the single germline progenitor P4 in C. elegans emerges the two PGCs, Z2 and Z3 (Sulston et al., 1983). While Z2 and Z3 themselves do not migrate per se, their development includes two relevant processes. First, cells of the somatic niche derived from Z1 and Z4 (called the distal tip cells) undertake a migration to reach the PGCs. Distal tip cells signal through Notch pathway to control proliferation of the germline (Rohrschneider and Nance, 2013), although how migration shapes their signaling abilities is unknown. Second, development of the surrounding embryo pulls the PGCs from the surface to the center of the embryo in order to reach their destination (Rohrschneider and Nance, 2013). Analogous to PGCs in other species, it remains unknown which mechanisms signal cessation to the moving cells (in this case the somatic gonad progenitors Z1 and Z4), although the presence of PGCs is required for the final positioning of Z1 and Z4 (Rohrschneider and Nance, 2013).

D. melanogaster PGCs, or pole cells, are specified in the posterior extreme of the embryo. Their migration begins by passive conveyance to the interior of the embryo during gastrulation. Upon reaching the gut epithelium, the PGCs exhibit amoeboid migratory movements triggered by gut tissue (Grubb, 2006); it is not known whether these gut-derived signals may vary depending on location or intrinsic signaling capacities of PGCs. However, functional studies of individual PGCs and their ability to respond to cues or enact an amoeboid migratory phenotype could begin to address this question. Active migration through the mesoderm niche takes PGCs to the

forming gonads (Santos and Lehmann, 2004). Bifurcation from the midline toward the opposing gonads involves repulsive as well as attractive cues and takes individual PGCs on varying path lengths at varying speeds, but the extent or significance of this potential heterogeneity is not known.

The study of Zebrafish PGC migration reveals several interesting windows where heterogeneity may play a role in the development of the germline. After being specified in 4 random locations along the developmental axis, Zebrafish PGC morphology is smooth and round (Richardson and Lehmann, 2010). PGCs in Zebrafish must undergo a round of mitosis and a round of polarization before active migration initiates (Blaser et al., 2005), at which time they begin a directed, active movement under the guidance of chemokines (Raz and Reichman-Fried, 2006). This awakening occurs on an individual cell basis, emphasizing the potential for heterogeneity among the initial cell divisions and polarization processes (Raz and Reichman-Fried, 2006). Migratory initiation depends on the capability to mount an entirely new transcriptional program (Richardson and Lehmann, 2010), such that variations in transcription may correlate with variations in migration. Zebrafish PGCs exhibit a curious cycling of active “running” migration when they are polarized, and a “tumbling” pausing phase where they lose cell polarity and are able to make course corrections (Reichman-Fried et al., 2004). The ability of an individual PGC to cycle properly between these stages is a necessity for survival, as those failing to complete these exercises become ectopic and undergo apoptosis outside of the range of survival signals. Interestingly, Zebrafish PGCs switch between individualized migration in initial phases to movement toward the gonads in clusters following aggregation at concentrated locations of the chemoattractant Sdf-1. This later stage of group migration may serve to mitigate differences in chemokine responsiveness or motility that arise earlier in Zebrafish PGC development.

Mouse PGCs begin their migration from their site of specification on the posterior side of the epiblast near the base of the extraembryonic structure known as the allantois (Ginsburg et al., 1990). This 3-4 day journey takes PGCs from the allantois, through the primitive gut and surrounding mesentery, and into the gonadal ridge. Migration is not synchronous, but spread out in space and time; some PGCs "lead" migration while others "lag" behind. Time-lapse imaging captures the individualistic behavior of PGCs as well as the fraction that migrate improperly, failing to reach the next milepost or moving into ectopic locations (Molyneaux et al., 2001); McDole et al., 2018). Also evidenced in the live movies are differing levels of activity from E9.0-9.5, when germ cells are active and motile, but not yet exhibiting the directed migration which causes them to exit the gut epithelium. What distinguishes the earliest “pioneer” germ cells first exiting the gut from other germ cells that follow them is a heterogeneity profile as of yet unknown (Cantú et al., 2016; Gomperts et al., 1994). Another point at which heterogeneous PGC behaviors emerge is during the last stages of migration at E10.5, when approximately 10% of PGCs remain outside of the gonads, with little hope of completing their migration as determined by their persistence at E13.5 (Cantú et al., 2016; Nguyen et al., 2020). Motility of mouse PGCs is regulated by multiple signaling mechanisms, including the non-canonical Wnt signaling pathway via Wnt5a (Laird et al., 2011), which attenuates the mitogenic beta-catenin dependent arm of the pathway (Cantú et al., 2016). Although migratory mouse PGC behavior is influenced by signals from different niches, which certainly promotes heterogeneous migratory outcomes, there also appears to be intrinsic heterogeneity of PGCs.

Cultured ex vivo in defined conditions, mouse PGCs mounted variable signaling responses to cues such as Wnt5a (Cantú et al., 2016). PGCs from embryos have also been observed to be heterogeneous in expression of important signaling receptors such as cKit, which together with Kit ligand regulate migration and survival (Morita-Fujimura et al., 2009). Finally, it is also possible that as the mesentery niche changes and develops around them, the PGCs in the

lagging group lose contact with the proper migratory signals which would enable them to move forward. Even more intriguing is the idea that migration acts as a selection metric, to prevent germ cells with deleterious or detrimental heterogeneities from reaching the gonad and colonizing. This will be discussed more in the final section of this chapter.

Proliferation of Primordial Germ Cells and Heterogeneity

PGC development involves the expansion of a small handful of founder cells to a cohort of sex-undifferentiated germ cells in the gonads, which range in number from tens in flies and and worms to thousands in mice. The proliferative histories of individual PGCs have not been examined. However, discrepancies in the number of self-renewing divisions could potentially amplify any inherent heterogeneity between PGCs by over-representing or under-representing particular founders. Alternatively, unequal divisions could disproportionately increase the opportunity for replication-associated damage in subsets of germ cells, thereby introducing new heterogeneity.

We present examples in several organisms of heterogeneous behavior of PGCs during proliferation. The two C. elegans founder PGCs Z2 and Z3 defer proliferation until larval stages, and rely upon regulation by somatic gonad precursors (Kimble and White, 1981). Germline development is acutely responsive to nutritional cues in worms (Hubbard et al., 2013). Upon hatching, unless larvae eat, they enter a state of L1 developmental arrest called diapause, accompanied by an arrest in the proliferation of PGCs and their somatic support cells (Fukuyama et al., 2006). During L1 diapause, PGCs halt their cell cycle in G2 in a precisely controlled process under the auspices of Pten and the insulin/IGF-like signaling pathway. Release from L1 diapause and G2 arrest occurs upon feeding and involves cell cycle re-entry of only one of the two PGCs, which is believed to reflect communication between the PGCs via yet

unknown mechanisms (Fukuyama et al., 2006). Although it is further unknown which of the two PGCs proliferates and the basis for this decision, this proliferative asymmetry in the C. elegans germline could amplify any heterogeneity that exists between Z2 and Z3.

Similarly, the regulation of Drosophila PGCs during development involves differential proliferation that could introduce or augment pre-existing heterogeneity. The 12 PGCs that arrive at each of the gonads double their number every 24 hours during the first and second instar larval stages before curbing their expansion. Mutants possessing severely reduced numbers of embryonic PGCs -- such as oskar and germcell-less which average at 2 -- reach the middle of 3rd larval instar with the correct number of PGCs. These mutants revealed a cross-talk between PGCs and somatic cells of the ovary via EGF signaling that senses and adjusts the number of PGCs (Gilboa and Lehmann, 2006). As in other organs, the size of the PGC pool in the larval fly ovary relies upon Hippo signaling between PGCs and somatic cells that form their niche (Sarikaya and Extavour, 2015). These compensatory mechanisms for PGC proliferation indicate that the Drosophila germline can be built from different numbers of founder PGCs and through differing numbers of divisions.

Such mechanisms for detecting and adjusting cell numbers similarly govern the proliferation of PGCs in chicken, where compensatory expansion was observed after treatment with the anti-mitogen busulfan (Lee et al., 2013). In mice, PGCs were deemed to undergo approximately 8 divisions based on a population doubling time of 16 hours, with compensation for early losses in the number of PGCs occurring during the period of expansion in the fetal gonads (Tam and Snow, 1981). During their migration, the proliferation of PGCs in the mouse embryo increases with each successive niche through which they transit (Cantú et al., 2016). PGC mitosis is regulated by beta-catenin-dependent Wnt signaling pathway but instructed by the niche. This

extrinsic and niche-specific control of PGC proliferation, coupled with the asynchrony in PGC migration, raises the possibility that those migrating first or fastest would undergo more cycles of expansion, leading to a clonal dominance over those migrating last or slowest (Cantú et al., 2016). Although the number of cell divisions has been proposed to be deterministic in a number of processes-- including the decision to enter meiosis (McLaren and Southee, 1997; McLaren, 2003; Ohkubo et al., 1996), the probability of acquiring mutations (Ohno, 2019), and the quality of the oocyte (Henderson and Edwards, 1968) -- the functional impact of proliferative history on the germline and gamete remains to be determined. The advent of molecular 'flight recorders' that track cell divisions in vivo hold promise for revealing the extent and the consequences of heterogeneous proliferation in PGCs.

Primordial Germ Cell Survival and Heterogeneity

As PGCs progress through fetal development, many fail to complete the entire journey and are eliminated from contributing to the future germline. This elimination occurs mainly via developmentally programmed apoptosis at several stages in the fetal period. Importantly, this apoptosis is separate from PGC death in response to cytotoxic insults, such as irradiation or chemical exposure. Consistent PGC apoptosis during development suggests that apoptosis may act as a selective event. The PGCs that do survive fetal development represent the founding population that establishes the adult germline. Differences in PGC survival during this fetal period are therefore critical to determining the composition of gamete-producing cells in adulthood. Given that only a subset of the initial PGC population completes fetal development, survival represents a heterogeneous outcome that may reflect important differences between survivors and eliminated PGCs that are relevant to germ cell function.

Apoptosis is an invariant feature of C. elegans embryo development, with 131 out of 1090 somatic cells eliminated in a lineage-specific manner (Gartner et al., 2008). In the germline, apoptosis occurs exclusively in the adult during oogenesis and eliminates a variable number of potential oocytes that does not exceed 50% (Gumienny et al., 1999). Importantly, germline apoptosis differs from somatic apoptosis by virtue of the single oocyte lineage containing two variably sized but distinctly-fated populations: surviving versus eliminated cells. In contrast, somatic apoptosis reproducibly eliminates precisely the same 131 cells. Oocyte apoptosis can be classified into two categories: stress-induced and physiological. Stress-induced apoptosis occurs in response to insults such as environmental stress, DNA damage, or bacterial infection. In contrast, physiological apoptosis affects a fraction of cells during oogenesis in the absence of any stress and differs from stress-induced apoptosis by not relying on pro-apoptotic BH3 proteins. Physiological apoptosis is restricted to oocytes in late pachytene that occupy a particular region of the gonad. Inhibition of germline apoptosis results in poorer quality oocytes as characterized by decreased egg viability and difficulty producing full-sized oocytes (Andux and Ellis, 2008). These reproductive defects may be due to intrinsic faults in the normally-eliminated population that are not present in the oocytes that survive physiological apoptosis. Alternatively, apoptosis may not reflect intrinsic cellular differences between oocytes in the survivor versus dying populations but instead functions as a resource-management strategy. C. elegans germ cells are linked as a syncytia, allowing apoptotic cells to act as nurse cells and direct their cytoplasmic contents toward surviving cells (Wolke et al., 2007). These two models are non-exclusive; cellular differences stemming from oocyte heterogeneity can serve as the basis for determining whether an oocyte acquires a nurse cell or survivor fate.

Drosophila germ cell development shares many conserved elements with higher-order organisms such as migration. Drosophila PGCs, also termed pole cells, that fail to migrate

accurately undergo programmed cell death in the early embryo (Coffman et al., 2002). As many as half of the initial cohort of PGCs mismigrate and are ultimately eliminated (Underwood et al., 1980). Ectopic PGC removal does not require apoptosis genes grim, reaper, or hid and is caspase-independent (Sano et al., 2005). Several studies have demonstrated that elimination of ectopic PGCs instead requires a pathway involving both p53 and out (Coffman et al., 2002; Yamada et al., 2008). Overexpression of p53 induces programmed cell death only in ectopic PGCs, suggesting that the threshold for elimination varies across subpopulations of PGCs that differ by migratory success. Differentiation is another criterion that can distinguish surviving PGCs from eliminated ones. In studying Nanos (nos), an evolutionarily conserved protein critical for many aspects of germline development, Hayashi et. al. discovered that nos-mutant PGCs inappropriately acquire somatic identities and are subsequently removed through apoptosis (Hayashi et al., 2004). Interestingly, despite the central role of nanos in securing germline identity, nos-mutant PGCs only display a partially penetrant phenotype of somatic expression (Schaner et al., 2003), suggesting that heterogeneity exists among PGCs in regard to retaining somatic competence. Given the role of nanos in regulating histone modifications, heterogeneous chromatin states among individual cells may explain this variation in differentiation propensity. To summarize, programmed cell death in Drosophila operates through both an apoptosis-independent and dependent mechanism, with apoptosis being associated with PGCs that inappropriately differentiate to the somatic lineage.

In Zebrafish, PGC apoptosis also is linked to migration, with disrupted migration leading to apoptosis. Weidinger et al injected morpholinos targeting dead end mRNA, a component of germ plasm in zebrafish, and observed an increase in ectopic PGCs that would ultimately be eliminated by apoptosis (Weidinger et al., 2003). Apoptosis was preceded by downregulation of key germ cell maintenance genes such as nos-1, indicating that mismigration disrupts germ cell

identity in ectopic PGCs, leading to downstream apoptosis. Apoptosis is the eventual fate of ectopic PGCs in other migratory mutants (Köprunner et al., 2001; Ramasamy et al., 2006) and may also be due to loss of nos-1 in ectopic PGCs. Importantly, in these migratory mutants some PGCs do correctly localize to the gonads, maintain nos-1 and its protein product Nos1, and survive, suggesting that correct migration supports the maintenance of germ cell identity and provides a survival advantage.

PGC survival in mice follows a similar paradigm to that in Drosophila and Zebrafish. Correct PGC migration is tightly linked to PGC survival. This coupling is primarily accomplished through genes that have dual roles in chemotaxis as well as survival. In mice, mutants for Dead-end (Dnd) and KitL (Steel) result in increased numbers of ectopic PGCs, which subsequently undergo apoptosis (Runyan et al., 2006; Youngren et al., 2005). KitL is downregulated in the midline after PGCs migrate through the tissue by E10.5, enforcing a narrow pro-survival niche (Runyan et al., 2006). The stringent connection between survival and migratory signals resolves migratory heterogeneity in favor of PGCs that can accurately navigate to the gonad. Because ectopic PGCs can give rise to germ cell tumours, this linking of migratory success to survival can prevent deleterious differentiation from mismigrated PGCs (Runyan et al., 2008).

After reaching the gonads, mouse PGC apoptosis occurs primarily through a Bax-mediated mechanism peaking by E13.5 in both males and females (Coucouvanis et al., 1993; Wang et al., 1998). PGCs begin to upregulate transcripts of Bax, increasing their apoptotic potential between E10.5 through E13.5 (Stallock et al., 2003). The Bcl-2 family members such as Bcl-x (Rucker et al., 2000) further regulate PGC survival. Variation in the expression and balancing of these apoptotic and survival pathways can modulate cell death propensity and could give rise to the apoptotic fraction observed even in wild-type backgrounds. Interestingly, successful

differentiation is associated with increased survival. Knockout of Nanos3, a germline transcription factor important for germline differentiation, results in PGC depletion by Bax-dependent and -independent mechanisms (Suzuki and Saga, 2008). However, a subset of Nanos3 null PGCs are capable of surviving and differentiating in male and female adult gonads, suggesting that a subset of PGCs is more robust and capable of survival despite the loss of pro-differentiation components.

While studies of mutant PGCs have illuminated much of the differentiation and survival networks operating in the fetal period, the fraction of wild-type PGCs that undergo apoptosis indicate that not all PGCs are equivalent in their likelihood to contribute to gametogenesis. The identities of apoptotic germ cells on a wild-type background are beginning to be understood through techniques such as single-cell RNA sequencing and clonal labeling, both of which can elucidate subpopulation-level differences among germ cells that bulk analyzes cannot. In particular, clonal labeling allows for direct comparisons among potentially heterogeneous clonal subpopulations. Multicolor clonal labeling revealed that apoptosis in the fetal male germline is clonal, suggesting that mitotically heritable differences among subpopulations - particularly in the expression of epigenetically-regulated genes that control sex differentiation - can determine whether germ cells survive or are eliminated.

Differentiation of PGCs and heterogeneity

After specification, the next major differentiation event in PGC development is sex differentiation. Of the common model organisms, C. elegans is unique in featuring hermaphroditic gametogenesis, although general principles regarding somatic regulation of germline differentiation can apply. For example C. elegans somatic cells at the distal tips of the gonads maintain mitotic potential, and meiosis only occurs as germ cells move away (Ellis,

2008). In Drosophila and mice, the somatic gonadal sexual identity significantly influences the sex differentiation of resident germ cells (Casper and Van Doren, 2006; Ewen and Koopman, 2010). This section will primarily discuss differentiation of male and female germ cells in mice, with a focus on meiotic initiation and mitotic arrest.

Upon arriving in the gonad, PGCs undergo sexually dimorphic differentiation into the highly specialized male and female lineages. The signal for meiotic initiation in mice is retinoic acid (RA), which is secreted by somatic cells in the mesonephros (Bowles et al., 2006; Koubova et al., 2006). As RA must diffuse to germ cells in the ovary, the diffusion gradient could produce variation in both strength and timing of RA signaling. Meiosis in the mouse ovary is observed to occur in an anterior to posterior wave that may reflect asynchronous initiation (Menke et al., 2003). Interestingly, this anteroposterior wave does not align with the axis of diffusion between the mesonephros and the ovary, suggesting that other signals may modulate RA signaling within the ovary. RA has also been discovered to be produced by ovarian cells (Bowles et al., 2016; Childs et al., 2011) and the Aldh1a1 enzyme. Meiotic progression involves Stra8, which is activated by Dmrt binding to the Stra8 promoter (Matson et al., 2010). Stra8 mutants display greatly reduced follicle formation, but a small population of Dmrt1-null oocytes can continue to form functional follicles and express the appropriate markers of later meiosis, including Stra8, Sycp3, and Msy2. This meiotically competent subpopulation represents more robust oocytes that do not require Dmrt1 to complete meiosis, possibly due to a compensatory meiotic program that is only activated in a subset of oocytes. Double knockout mice that lack both Stra8 and Msx1/2, another activator of Stra8, also contain a subpopulation of oocytes that can still complete meiosis (Le Bouffant et al., 2011). The variation in meiotic robustness suggests that heterogeneity exists during female sex differentiation to endow certain oocytes with greater meiotic potential.

Male sex differentiation in mice is initiated at E12 by supportive somatic cells in the nascent testis called Sertoli cells (Ohta et al., 2012). Sertoli cell primarily regulate production of female-inhibitory and male-inducing factors such as Fgf9 and Cyp26b1 to suppress the female meiotic fate (Bowles et al., 2010, 2006). Male germ cells continue to proliferate in the testis but undergo mitotic arrest as early as E12.5. This timing is asynchronous, as some germ cells are still proliferating by E14.5 (Western et al., 2008). Mitotic arrest is accompanied by expression of Nanos2, a key male sex differentiation marker (Suzuki and Saga, 2008). Delayed mitotic arrest is associated with aberrant maintenance of pluripotency, decreased Nanos2, and an increased susceptibility for teratoma formation (Dawson et al., 2018; Heaney et al., 2012). Heterogeneity in the timing of mitotic arrest can expose later-arresting male germ cells to these deleterious fates. Mitotic arrest is part of many processes regulated by complex networks of secreted factors such as Tgfβ and Nodal (Spiller et al., 2017) active in germ cells during this period. Heterogeneous expression of these factors as well as of their receptors on germ cells can significantly modulate the differentiation states of male germ cells, leading to vastly asynchronous or divergent cell fates.

An overarching theme during sex differentiation is the variation in both position and timing of these events. Many secreted factors regulate this process in males and females, and the precise signaling environments remain to be defined. Heterogeneous progression through sex differentiation can also be due to PGCs themselves differing in their capacity to respond to differentiation cues. Sex differentiation has been shown to be facilitated by epigenetically-regulated genes that are activated by demethylation (Hill et al., 2018). Individual variation in the extent of demethylation particularly at these loci could lead to highly variable differentiation potentials. Thus, initial cellular differences in epigenetic state could be amplified as

epigenetically heterogeneous subpopulations of PGCs proceed through sex differentiation asynchronously, or fail to do so at all. The later impact of differentiation heterogeneity can be significant, affecting reproductive and tumorigenic potential (Hunt and Hassold, 2008).

1.3 - Determinants of Heterogeneity

In this section, we consider the potential causes for diverse behaviors among PGCs described above. We separate these causes into those which are genetically encoded, those which arise from traditional epigenetic marks such as DNA methylation, chromatin modifications, or small RNAs, and a third and separate category of non-genetic determinants (summarized in Figure 1.3). Among these sources, we distinguish between those which are permanent (and therefore heritable to all cellular progeny), those which are transient and do not cause similar behaviors in cellular progeny, and those determinants which are semi-permanent, or lasting in daughter cells but not for infinite cellular generations. Finally, we discuss mechanisms that mitigate variation between germ cells, or buffers of heterogeneity. For an outstanding taxonomy of the sources of heterogeneity, refer to Huang (Huang, 2009).

Genetic Determinants

Acquired genetic changes

De novo mutations (DNMs) in the DNA sequence typically originate during cell division. Although DNMs occur in all cells, only those in the germline are inherited to offspring. The Darwinian model of evolution by random mutation, causing a phenotypic change which undergoes natural selection in fact originates in germ cells, or potentially in the early embryo of animals such as mice which induce PGCs from somatic cells. Although the diversity of multicellular organisms is living testament to DNMs in the germline, the rate of mutation is low:

approximately 10^{-9} to 10^{-10} mutations per nucleotide pair per cell division in eukaryotic cells (Arana and Kunkel, 2010; Drake et al., 1998). Studies across mice and humans suggest that DNMs arise in the germline at an even lower frequency than in somatic cells (Kohler et al., 1991; Milholland et al., 2017; Walter et al., 1998).

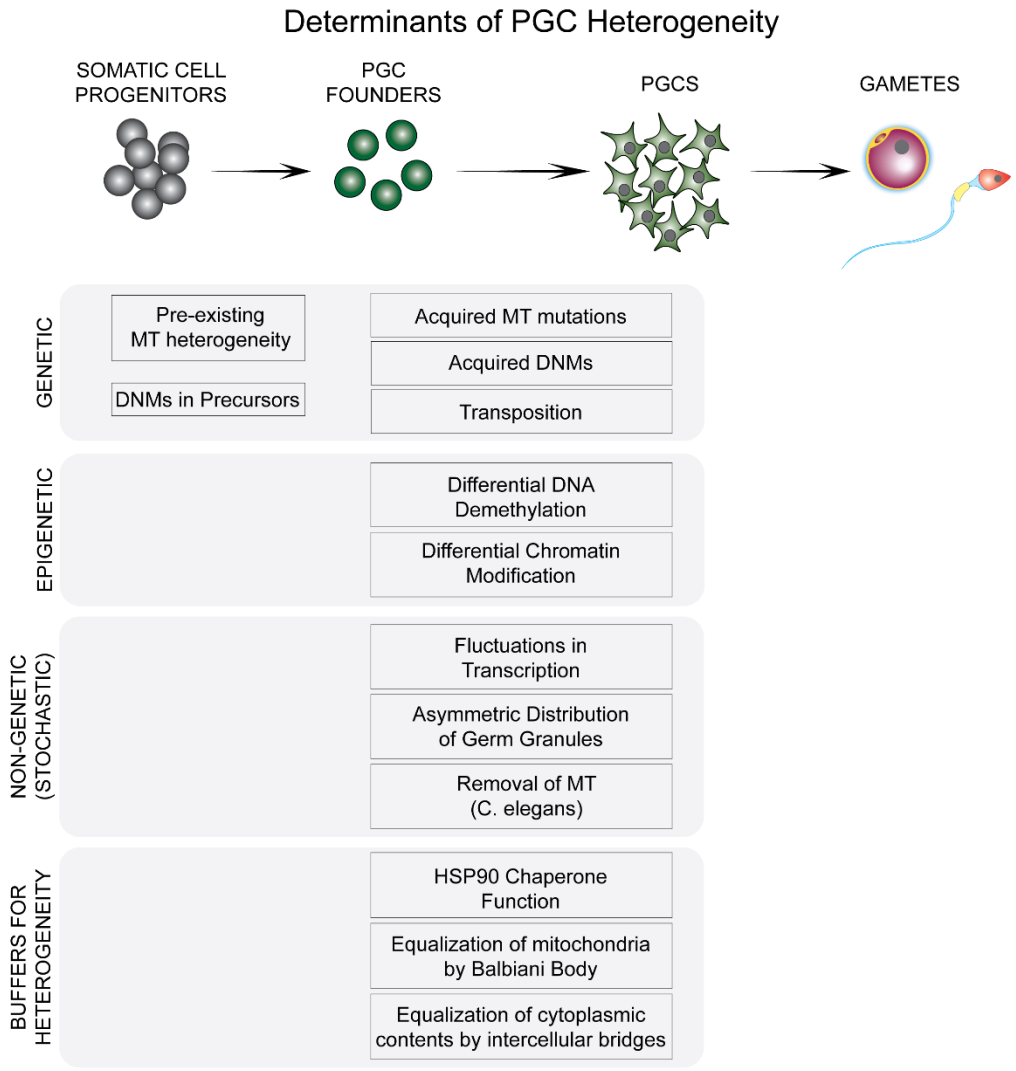


Figure 1.3 Determinants of PGC Heterogeneity

For the various cell types, different underlying cell states can be causes for heterogeneity, impacting cell type transition as well as impacting the fitness of progeny.

The observed rate of mutation is the output of an equation that includes incurred errors, DNA repair, and elimination of cells. For studies that infer germline mutations by examining progeny, gamete fitness is also included in the rate calculation. The recent generation of exome and whole genome data in human families has provided precise measurements of germline DNMs, with single-nucleotide variations (SNVs) arising at 1.2×10^{-8} per nucleotide per generation (Besenbacher et al., 2015; Campbell et al., 2012; Conrad et al., 2011; Jónsson et al., 2017; Kong et al., 2012; Lynch, 2010; Michaelson et al., 2012; Rahbari et al., 2016; Roach et al., 2010; Ségurel et al., 2014; Wong et al., 2016). The rate of germline mutation in humans as well as non-human primates (Pfeifer, 2017) is nearly an order of magnitude greater than that measured in mice (5.4×10^{-9} per nt/generation; (Uchimura et al., 2015)), birds (4.6×10^{-9}/nt/generation; (Smeds et al., 2016)), or Drosophila (2.8×10^{-9}/nt/generation; (Keightley et al., 2014)). This difference most likely stems from the greater number of cell divisions that occur in the human germline and specifically in the spermatogonial stem cells, which divide throughout life, in contrast to oocytes which cease replication in the fetus (Crow, 2000). Indeed, sequencing of parents and offspring revealed that the majority of human DNMs arise in the paternal germline and increase with paternal age by 1-2 per year (Campbell et al., 2012; Jónsson et al., 2017; Kong et al., 2012; Michaelson et al., 2012; Rahbari et al., 2016). However, with an estimated 10 cellular divisions from the zygote to PGC in mammals (Rahbari et al., 2016), the theoretical number of incurred mutations across the genome of each PGC ranges from ~2 to 15 for mice and ~4 to 38 for humans; hence DNMs do not account for the extent of heterogeneity in cellular behavior that is observed.

A second source of acquired mutations in PGCs is the mitochondrial genome. Separate from the nuclear genome, the circular mitochondrial DNA molecule produces 13 mRNAs, which all encode oxidative phosphorylation components, as well as 2 rRNAs and 22 tRNAs needed for

mitochondrial translation of these mRNAs (Falkenberg et al., 2007). Critically, the rate of mutation of the mitochondrial genome is 2 to 20-fold greater than the nuclear genome, due to replication errors (Allio et al., 2017). The presence of more than one mitochondrial haplotype within the same cell, or heteroplasmy, arises in cells because fidelity of the mitochondrial polymerase is poor and varies depending on biological condition; for example, the presence of reactive oxygen species (ROS) decreases the fidelity of the mitochondrial polymerase, while the presence of a proofreading subunit increases fidelity (Longley et al., 2001; Pinz et al., 1995). Although mitochondrial replication is paused in the early mouse embryo until the time of gastrulation (Cao et al., 2007; Wai et al., 2008), in the course of PGC development, estimates suggest the number of mitochondria per cell increases from 100 at E7.5 to over 6000 at E14.5 (Wai et al., 2008). Given the size of the mitochondrial genome, mutation rate, and rate of mitochondrial expansion, approximately 35 mitochondrial DNMs would accumulate by E14.5 in mice, although this does not account for unequal replication. However, as with nuclear mutations, this rate of acquired mutation in the mitochondrial genome is likely a very small contribution, at best, to observed heterogeneity of PGCs

A third potential source of acquired mutation in PGCs is through retrotransposition. Transposable elements comprise a significant fraction of the genome and many of the evolutionarily young Alu and autonomous L1 (or Long Interspersed, LINE-1) elements remain active and capable of re-inserting into new sites in the genome (Kazazian and Moran, 2017). During development, the period of epigenetic reprogramming in PGCs de-silences transposable elements, leading to their expression and potential for transposition in the window of development before transposons are silenced by piRNAs (Brennecke et al., 2007; Teixeira et al., 2017). In humans, new L1 and Alu integrations were detected in ~ 1/100 and 1/20 births, respectively (Hancks and Kazazian, 2012). This rate is even higher in mice, where new L1

insertions occur in 1/8 births (Richardson et al., 2017). However, on a cellular level, given its rate of occurrence, transposition again makes an exceedingly minor contribution to PGC heterogeneity.

Pre-existing genetic changes

A special case for animals that specify their germline by induction is the potential for prior genetic mutations as a source for heterogeneity. In mice, DNMs arising in the epiblast would lead to mosaicism among the founder PGCs as well as in somatic lineages. Cells in the mouse epiblast undergo rapid divisions -- measuring as little as 3 hours (Snow, 1977) -- hence providing opportunity for errors in replication. Furnishing evidence of the elevated potential for damage and repair during this period of rapid division in the early mouse embryo, PGCs and cells of the epiblast have been shown to be hypersensitive to genotoxic stress as measured by DNA breaks, Atm and p53-dependent apoptosis (Heyer et al., 2000). From genome sequencing of human families, it has been imputed that approximately 10 divisions occur between the zygote and the specification of PGCs (Ohno, 2019; Rahbari et al., 2016); and this number is probably similar in mice. Detection of DNMs in offspring that co-exist at low levels in the blood of one parent provides evidence that replication errors in cells of the early embryo before PGC specification are present in a subset of somatic and germline derivatives. Data in humans estimate that 0.2-0.6 DNMs accumulate per cell division in the pre-PGC embryo (Rahbari et al., 2016), which would suggest a theoretical maximum of 6 pre-existing mutations among PGC founders could contribute to heterogeneity.

A second type of pre-existing genetic mutation that could contribute to heterogeneity among PGCs is mitochondrial heteroplasmy (Mishra and Chan, 2014). Mitochondrial heteroplasmy and mutations in mitochondrial (mt) DNA arise in oocyte maturation due to the inherently higher

error rate in mitochondrial polymerase (Allio et al., 2017; Khrapko et al., 1997; Madsen et al., 1993). During early embryogenesis, the replication of mtDNA and fission of mitochondria ceases (~E1.0-E3.0 in mice) (Cao et al., 2007; Wai et al., 2008), so that each cell division segregates existing mitochondria and produces an increasingly purer pool of mtDNA. Known as a mitochondrial bottleneck, this segregation limits the diversity of mitochondrial genomes inherited from the oocyte (Mishra and Chan, 2014; Stewart and Chinnery, 2015). As PGC specification in mice occurs during this bottleneck, the founders inherit different mtDNA cohorts. Replication of mitochondria resumes after PGC specification, amplifying the heterogeneity between cells that was created during the bottleneck. Together these phenomena may reveal mtDNA mutations by altered PGC phenotypes.

The potential for revealing a mutation in a mitochondrial genome increases as heteroplasmy from the oocyte trends to homoplasmy (similarity in mtDNA) in PGCs. Preexisting or de novo mtDNA mutations would be inherited in a minority of PGCs during specification or could arise during PGC development; either way, those cells with a high concentration of mitochondria bearing deleterious mutations would exhibit altered phenotypes. Concomitant with the onset of PGC migration, metabolism within the embryo shifts from glycolysis to oxidative phosphorylation, which provides a potential source of selective pressure for cells unable to make the switch (Folmes and Terzic, 2014). Deleterious mitochondrial DNA mutations within PGCs may be revealed as compromised ability to initiate oxidative phosphorylation or the bioenergetic response required for migration. Little is known about the cellular or molecular mechanisms for mitochondrial organelle quality control and selection, but the migratory time-point may be particularly important for early embryonic metabolism. A new area of inquiry concerns the selective destruction of mitochondria in PGCs by mechanisms such as mitophagy. Recent work describes an unprecedented means for removing mitochondria from PGCs in C.

elegans. This unique cellular surgical procedure is carried out by somatic endoderm cells via extrusion of cytoplasmic lobes from the PGCs (Abdu et al., 2016). Although speculative, the effect could serve to concentrate other contents such as germ granules or to remove faulty mitochondria in order to increase cellular fitness, reduce oxidative damage, or alter metabolic capacity of PGCs.

Epigenetic determinants

Having focused on the genetic determinants of heterogeneity, we turn our attention to the determinants not directly linked to the genome itself (Brock et al., 2009). Our interest in the epigenetic determinants of heterogeneity encompasses both transcription regulation through modification of histone and chromatin structure, as well as the molecular states defined by covalent modifications of DNA. We note that epigenetic determinants have the capacity for altering PGC states semi-permanently, as many of the mechanisms below are preserved through cell divisions.

Reprogramming of the epigenome occurs in two main stages in the fetal germline: first during PGC migration during which methylation is maintained at specific regions and secondly after gonadal colonization when epigenetic marks are erased to create a hypomethylated epigenetic ground state (Hackett and Surani, 2013; Seisenberger et al., 2012). At day 8 of mouse development, PGCs begin a significant loss of genome-wide DNA methylation that is considered the first of three waves of demethylation before E13 (Bohacek and Mansuy, 2015; Hackett and Surani, 2013; Hajkova et al., 2002; Seisenberger et al., 2012). Changes in histone modifications of PGCs occur starting at E9.0 include loss of histone H3-lysine (K) 9 dimethylation and increase in H3K27 trimethylation (Seki et al., 2007, 2005). These epigenetic

changes occur during migration (Nikolic et al., 2016) and before sex differentiation, and are thought to carry out the following: ensure that the germline genome is ready to receive parent-of-origin specific epigenetic marks; erase potential epimutations incurred in early development; and confer a state of pluripotency for the embryo (Hajkova et al., 2002; Seki et al., 2007).

As epigenomic studies to date have analyzed PGCs in bulk, the variation in epigenetic states of single cells is not known. In principle, differences in the timing of epigenetic reprogramming or failure to lose cytosine methylation or histone marks -- epimutations-- could lead to PGC heterogeneities. Such differences could be manifested as altered expression of genes critical to differentiation. In support of this idea, PGC maturation genes such as Mvh (Vasa) and the meiotic component Sycp3 are regulated by methylation at their promoters (Maatouk et al., 2006) and would be sensitive to epi-mutations. Recent studies have a shown that efficient sex differentiation is dependent upon demethylation at the loci of 45 germline reprogramming-responsive genes (GRRs; (Hill et al., 2018)). Any individual cell differences in the extent of demethylation at these GRRs could stratify PGCs by differentiation capacity and produce greater population variation as individual PGCs heterogeneously differentiate.

Demethylation during epigenetic reprogramming can also relieve inhibition of transposon sequences that were previously repressed by methylation. Retrotransposition activity from these unsuppressed elements can lead to genomic disruption and inappropriate gene expression, further exacerbating initial variation among germ cell subpopulations. Accordingly, germ cells that fail to repress transposons (Kuramochi-Miyagawa et al., 2004; Malki et al., 2014; Soper et al., 2008) arrest prior to gametogenesis and are eliminated, indicating that dysregulated transposons are stringently selected against. Single-cell assessments of transposon expression have identified LINE-1 elements as differentially expressed within both mouse and human PGCs at individual timepoints, which could reflect cell-to-cell variability in repressive methylation (Nguyen et al., 2020; Reznik et al., 2019).

A second mechanism of non-genomic transcriptional regulation to be addressed as a potential source of PGC heterogeneity is small non-coding RNAs, including micro RNAs (miRNAs), endogenous small interfering RNAs (endo-siRNAs), and piwi-interacting RNAs (piRNAs). In D. melanogaster and D. rerio, miRNAs function to silence germline enriched mRNAs in cells which are destined to be a part of a somatic fate (Banisch et al., 2012). In the mouse, the suppression of the miRNA let-7 by lin-28 is necessary for expression of critical PGC specification genes like Blimp1/Prdm1, and without lin-28 the germline fails to form (West et al., 2009); this kind of binary signal requirement during specification does not invite much heterogeneity. However, it is not known whether incremental differences of line-28 can lead to phenotypic consequences. Recent work has begun to characterize miRNA expression in the mouse at later developmental times such as sex differentiation, finding different sexually and developmentally-regulated populations of miRNAs (Fernández-Pérez et al., 2018). Examining miRNA heterogeneity at a single cell level is becoming possible with new technology (Hagemann-Jensen et al., 2018), and variability in the expression of miRNAs has recently been speculated to lead to non-genetic cell-to-cell heterogeneity outside of the germline (Wang et al., 2019). Since dysregulation of miRNA expression has been linked to testicular cancer , the possibility of heterogeneous expression of miRNAs in the early germ cell populations merits exploration.

piRNAs are small RNAs that are nearly exclusive to the germ line which silence transposable elements in the genome (Banisch et al., 2012). Although transposition events are rare and therefore not a significant source of germ cell heterogeneity, heterogeneous expression of transposable elements such as intracisternal-A particle (IAP) family of endogenous retroviruses could lead to phenotypic readouts. Recently identified additional functions for IAPs in regulating the transcription of neighboring genes (Vasiliauskaitė et al., 2018) raise the possibility that the

de-repression of IAPs could lead to inappropriate gene expression in PGCs as well as their derivative gametes.

The wave of demethylation which germ cells undergo during their development potentially enables transposable elements relocation to deleterious effect. During this period, endo-siRNAs are considered a secondary defense to silence transposable elements by degrading transcripts. endo-siRNAs have been studied in C. elegans, D. melanogaster, and mouse, and also function to prevent viral infection (Banisch et al., 2012). By nature, endo-siRNAs are heterogeneous in their response to an infection or transposition event, as these occur on a cell-to-cell basis. However, the ability of an individual germ cell to mount an endo-siRNAs has not yet been investigated, as the contribution of endo-siRNA to the suppression of transposons is just beginning to be appreciated (Berrens et al., 2017).

Non-genetic determinants

Distinct from alterations to the DNA or epigenetic mechanisms discussed above, non-genetic determinants can lead to measurable differences between cells of a single type. Bacteriologists studying genetically-identical, clonal populations cultured in uniform conditions first described 'non-genetic cell individuality' based on cell-to-cell variations in viral yield (known as burst size; (Delbrück, 1945)) and chemotactic behavior (Spudich and Koshland, 1976). These heterogeneities were ascribed to the interaction of stochastic fluctuations in molecules within the cell (Spudich and Koshland, 1976). In the field of cancer biology, non-genetic causes of cellular heterogeneity have been parsed into ‘population noise’ and ‘temporal noise’. Temporal noise derives from the fluctuations of protein levels within a cell over short durations. Population noise, on the other hand, reflects different states of cells at a given moment. The broad

assumption is that for genetically identical cells, population noise merely reflects temporal noise, and that all cells experience a similar profile of fluctuations over time (Brock et al., 2009).

One such generator of fluctuation may be the very nature of transcription, which was revealed to be pulsatile, occurring in bursts that emerge at locus-specific rates (Chubb et al., 2006; Raj et al., 2006; Suter et al., 2011). In germ cells, stochastic variation due to transcription may be particularly relevant, given that an elevated rate of global transcription was recently reported in mouse fetal germ cells (Percharde et al., 2017). This result raises the possibility that non-deterministic effects due to timing and transcription may be magnified in the developing germline, although there are currently no studies that have rigorously compared the absolute levels of transcripts in germ cells across time to capture such cycles. It is not yet clear how variation in levels and rate of transcription relates to the level of proteins, although imminent developments in technologies for single cell proteomics will be revealing.

A key assumption of the stochastic processes that generate fluctuations in molecules is that they are not permanent; although a cell may remain in a particular state, its daughter cells will not remember that state. Remarkably, there are documented instances of cells in culture maintaining levels of particular proteins through mitosis. In one example, variations in the levels of apoptotic regulators were maintained through 1-2 cellular generations of human mammary epithelial cells, so that the response to extrinsic death signals was similar between mother and daughter cells (Spencer et al., 2009). In the early mouse embryo, the levels of cMYC are variable between cells, and inherited through divisions. As cMyc levels determine the capacity of pluripotent cells in the epiblast to outcompete their neighbors by inducing programmed cell death, this observation indicates that cell fitness states are remembered transiently through mechanisms that remain to be elucidated (Díaz-Díaz et al., 2017). We can conceive of

scenarios by which non-epigenetic, stochastic heterogeneity could lead to permanent genetic changes; for example, variability in the levels of DNA repair machinery could render some cells more susceptible or resistant to acquired mutations.

A final non-genetic source of heterogeneity we consider is the germ granule. A highly conserved feature of germ cells, these cytoplasmic granules consist of aggregates of mRNA, small RNA, and proteins (Seydoux and Braun, 2006). In C. elegans, Drosophila, and Xenopus, these granules (known as polar granules, P granules or germ plasm) segregate to PGCs during specification and are required for specification and function of germ cells (Gallo et al., 2010; Hay et al., 1988; Strome and Wood, 1983). In mammals, PGCs synthesize components of germ granules rather than inherit them prior to specification, although these germ granules are then passed on with each cell division. Mammalian germ granules resemble processing bodies in somatic cells, which contain mRNA decay machinery. By delivering granule-associated RNAs that can regulate gene expression, germ granules can produce heritable changes in recipient cells independent of changes to DNA sequence, DNA methylation, or chromatin state. Heterogeneity could be imparted by the composition of the germ granules that are transmitted. The contents of germ granules can be modulated according to different environmental conditions as a stress response mechanism (Snee and Macdonald, 2009, 2004). Alternatively, the number of germ granules can be asymmetrically divided between two daughter cells. In Drosophila, the amount of germ plasm as measured by wun2 mRNA levels can vary within a field of cells, with central cells receiving more germ plasm than peripheral cells (Slaidina and Lehmann, 2017). This asymmetry provides central cells with relatively higher amounts of maternal mRNA and is demonstrated to provide a survival advantage over peripheral cells.

Extrinsic sources of heterogeneity: environment, niche, and timing

The necessity of the tissue niche for proper development of cells has been widely recognized across biology, and the germline niche is no exception. Unlike the well characterized cooperative signaling and maintenance of cell fate of the adult Drosophila stem cell niche, the notion that PGCs receive instruction from their niche during specification, migration, and colonization is less established. Across many organisms, the developing embryo changes drastically from the time of PGC specification to the time that functional gametes are produced. The idea that the landscape with which the developing germ cells must interact changes constantly is not new. It is known that fly PGCs receive different levels of signaling from different niches (Stepanik et al., 2016). How the changing niche may contribute to heterogeneity is only just beginning to be appreciated.

PGC migration in mice exhibits spread over space and time. Egress of PGCs from the hindgut into the surrounding mesoderm occurs asynchronously, with 'pioneers' emerging first; chemotaxis toward the gonads occurs at varying speeds; and colonization of the gonads takes place over more than 24 hours, with late arrivals remaining in the mesentery (Molyneaux et al., 2001). A recent study found that the rate of proliferation of PGCs increases during migration, and is regulated by the successive niches through which they transit (Cantú et al., 2016). The progressive increase of mitotic signals along the migratory route implies that those PGCs that move through the series of niches first would gain a proliferative advantage. Upon reaching the gonads, the earlier or faster PGCs would have generated a larger number of clonal progeny as compared to the later or slower migrants. It is here that the heterogeneity of PGCs may be influenced by the environment, as cells which are capable of efficient migration are rewarded by enhanced clonal expansion.

Buffers of heterogeneity

In contrast to the determinants discussed above, which potentially generate and reinforce differences between individual cells, we now consider the opposite: those mechanisms that could mitigate differences between PGCs. Although we propose that a certain level of PGC heterogeneity could increase the overall fitness of the germline through selection (as discussed below), there are several ways that germ cell diversity could be unfavorable First, genetic mutations or epi-mutations in PGCs that potentially promote the initiation, growth, or metastasis of tumors by rendering PGCs more resistant to apoptosis, more proliferative, or able to migrate would be dangerous in gametes. Such PGC variants could out-compete other PGCs and even homogenize the gamete pool. Second, the phenotypes that increase the fitness of a PGC are not necessarily the same phenotypes that benefit a gamete or embryo. Finally, too much tolerance for heterogeneity may have consequences for the overall fertility of the organism over a lifetime, such as early loss of fertility or lack of consistency of gametes. In this light we point to evidence for mechanisms that buffer heterogeneity in PGCs.

Hsp90 is a highly conserved chaperone that assists with the folding of proteins. It has been dubbed an ‘evolutionary capacitor’ because its absence in a number of organisms leads to enhanced phenotypic variability that is believed to arise from cryptic protein variants that are normally degraded or chaperoned by Hsp90 (Rutherford and Lindquist, 1998; Specchia et al., 2010). Most cell types express Hsp90b and elevate expression in stressful conditions, but an A isoform is exclusive to germ cells (Nishikawa and Kinjo, 2018). Hsp90a was further linked to germ cell development through its interaction with the germ plasm component and piRNA biogenesis factor Piwi (Gangaraju et al., 2011). We posit that Hsp90a is a likely candidate to buffer heterogeneity in PGCs during their development.

Intercellular bridges are a deeply conserved cellular structure found in germ cells at various points in their development. Formed by incomplete cytokinesis, the bridges stably connect sister germ cells into ‘cysts’ and permit the passage of cytoplasm and organelles such as mitochondria (Greenbaum et al., 2011). Intercellular bridges coordinate differentiation and allow sharing of resources between cells, particularly through self-sacrifice, when one member of a cyst commits apoptosis (Lei and Spradling, 2016). Although the formation of intercellular bridges between sister cells limits the potential difference between cells of a cyst, we note their potential to buffer phenotypic consequences of heterogeneity that arises in PGCs such as acquired DNMs or epi-mutations. The increased geographic spread of cysts compared to single PGCs also raises the possibility that they buffer differences introduced by the niche.

1.4 - Selection and Heterogeneity

Natural selection is an elegant force of evolution, working slowly through time to shape and change organisms to better suit their needs in changing environments and pressures. This concept has implications for species that reproduce and share heritable characteristics with their offspring. As organisms reproduce, their offspring inherit different traits with stochastic variation; if a random trait is beneficial to the offspring, it will have a competitive advantage, and potentially pass this trait along through its germline. Through many of the causes of heterogeneity that we have already discussed in this review, changes can be wrought within an organism. Changes within somatic cells, while impactful for the organism itself, pale in comparison to the multigenerational effect that heterogeneous changes can have if the germline is altered.

Whether through a genetic or non-genetic transmission, discrete traits are brought to prominence over time if placed under enough positive selective pressure. Natural selection within a population results from interaction between organisms and their environment, as the

environment creates selection pressure among heterogeneous individuals, resulting in increased fitness of the population. Less fit animals are phased out of prominence as natural selection leads to increase in the frequency of certain alleles over time.

We propose that like natural selection acts at the organismal and population level, there is a similar kind of selection operating at the cellular level (Buss, 2014; Laird et al., 2005a). At both the population and cellular level, alleles subjected to positive selection increase in frequency. Over time, variants are selected due to a competitive advantage they convey to potentially produce an organism or cell with greater fitness. The diversity of PGC development is a special, comprehensive series of selection events during which multiple types of fitnesses can be examined. We will consider how the totality of these developmental events act selectively on PGCs and discuss their implications on reproductive fitness.

Defining selection in the developing germline

In developmental selection, each stage of germ cell development provides an opportunity for the emergence of underlying heterogeneity to produce different outcomes for subpopulations of germ cells. As previously discussed, both non-genetic and genetic variation among individual cells can present distinct cell identities or states for selection to act upon. We consider a selective process to be one which alters the proportions of PGC subpopulations during development -- by mechanisms such as apoptosis, proliferation, migration, or differentiation. Here, we will briefly explicate what we consider a “selected cell” and the processes by which this selection occurs.

Selection can be either negative or positive with respect to gametogenic potential. In negative selection, a germ cell subpopulation is functionally eliminated from further contributing to

gametogenesis. In positive selection, a germ cell subpopulation is increased relative to other subpopulations. For the developing germline, we will consider how key events during the fetal period can potentially select for certain subpopulations and discuss the impact on the germline. For example, migration is one of the first challenges a PGC faces after specification at E7.0. During migration, PGCs must respond to precise chemotactic cues to remain on course as they navigate to the gonads. Owing to the distinctive signaling environments presented by each niche along the migratory route, the precise timing of migration may also represent another form of selection. Damaged or unfit PGCs might migrate inefficiently, resulting in reduced colonization of the gonad by their cellular progeny, which corresponds to negative selection. If PGCs that arrive first in the gonad receive different signals than later arriving PGCs, this could alter differentiation on the basis of migratory speed. Additionally, the duration of these differentiation steps could vary as early migrants would be exposed to signals for a longer period of time compared to late arrivals.

In contrast to spatially-based selection during PGC migration, selection can also occur during differentiation. After reaching the gonad, PGCs continue maturation by epigenetic reprogramming as well as sex differentiation. The concomitant increase in the expression of apoptotic genes such as Bax sensitizes PGCs toward elimination. Differentiation downregulates these genes and protects successfully differentiated PGCs from apoptosis. PGCs are additionally eliminated based on clonal lineage, implying that apoptotic selection can work on heritable qualities such as epigenetic permissiveness for differentiation (Nguyen et al., 2020). The balancing of these two programs selects for differentiated PGCs and selects against PGCs that are incapable or inefficient in differentiating.

Selection can also occur through proliferation, rather than elimination. Germ cell subpopulations that are more proliferative will increase their proportional representation within the germline with each successive division. This may also induce competition for niches or access to signaling factors that continue the tightly regulated differentiation program essential for further germ cell development into optimal gamete-producing cells. Earlier studies utilizing clonal labeling have noted that a small subset of clonal populations in the fetal period is responsible for producing the small number of large monoclonal areas in testes that actually contribute to spermatogenesis (Ueno et al., 2009). Later spermatogenesis in the adult has also been observed to originate from a limited number of spermatogonial stem cell clones (Kanatsu-Shinohara et al., 2016). In examining fetal populations of germ cells, Lei and Spradling found clone sizes to be variable (Lei and Spradling, 2013), although a longitudinal study would clarify if larger clones resulted from or continued to enjoy a proliferative advantage relative to other germ cell clones.

How does PGC fitness relate to fitness of gametes or individuals?

While heterogeneities and selective pressure are seen during all phases of germline development, we have thus limited our discussion to competition among fellow cells of the same type. Considering how each competitive cell contributes to the gametes as a whole, providing variation and heterogeneity within a single organism, leads us to speculate on how PGC selection may have an impact on the rate of evolution. Heterogeneity within the germline may exert effects not only within one or two generations, but could also shape the population of a species on a longer evolutionary scale. The fine processes of generating and selecting upon PGC heterogeneity in the embryo may influence the processes of natural selection or genetic drift.

The relationship between processes of germline development and processes of evolution has been considered by others with PGC specification. Broad surveys of modes of germline formation across phylogeny have driven the conclusion that PGC preformation is not ancestral but rather has evolved convergently in many species as reviewed in (Extavour and Akam, 2003; Johnson et al., 2011). While one study found that species that form their germline through preformation enjoy a faster rate of sequence evolution (Evans et al., 2014), follow-up work found instead that species with PGC specification by preformation and PGC induction were equally represented among species with rapidly evolving genomes (Whittle and Extavour, 2016). Since the rate of evolution may be influenced by the heterogeneity of the germline, perhaps organisms with less constraint on their PGCs and a more stringent selection would promote variation within the germline and faster rate of evolution through generations.

If selection during the fetal period can have multi-generational effect, then the fitness of cells which will contribute to these population effects must be multifaceted in their ability to respond to changing environments. Tests of germ cell selection may occur at many developmental timepoints to ensure greatest cell competency; germ cells that excel in one characteristic might be less fit for others. Take, for example, cells which excel dramatically at migration. How might supermigrators such as these compete in later developmental tests? Previous work has shown that as oocytes mature, their order of entry into meiosis corresponds to the order of follicle maturation (Hirshfield, 1992; Polani and Crolla, 1991), and later differentiation follows a “first in, first out” order of follicle maturation and ovulation (Reizel et al., 2012). While supermigrators have not been directly linked to order of entry into meiosis, one might speculate that the earliest migrants will also be the first to enter meiosis, and thus the first to initiate folliculogenesis (Arora et al., 2016). However, many of the earliest follicles release before puberty occurs, causing them to die due to a lack of supportive hormones environment (McGee, 2000). In this instance,

a cell overly specialized for migration may fail later developmental checkpoints. The shifting selective pressures of germ cell development ensure that the fittest cells are the germ cells competent at facing many developmental challenges.

These selective processes that comprise germ cell development provide an opportunity for selfish cells to take advantage. Selection favoring a subpopulation of germ cells can eliminate aberrant cells to improve PGC fitness, but it also is vulnerable to cheating. For example, mutations in the FGF receptors can increase FGF ligand binding affinity to confer a proliferative advantage in spermatogenesis. This would lead to clonal expansion of FGFR mutants during development and increased representation in the gamete pool (Goriely et al., 2003). While advantageous for germ cell success, these mutations are detrimental to the organism and can lead to disorders in somatic tissues such as in Apert syndrome (Glaser et al., 2003). The potential for germ cell cheating perhaps illustrates the utility of a multipartite developmental examination of various germ cell behaviors and phenotypes. A single mutation that provides an advantage in one stage may be selected against in another. This cumulative developmental selection would secure diversity in the gamete pool, which is essential for variation at the organismal level to respond to environmental and evolutionary challenges.

1.5 - Chapter Conclusion

The arduous journey to establish the germline is defined by variation-- both in the challenges faced by PGCs as well as the heterogeneity of the PGCs themselves. The diversity of events during PGC development creates a gauntlet that is ideally positioned to assess heterogeneity in PGCs as they begin to expand and differentiate to establish the eventual germline. Applying the principles of natural selection to the level of germ cells, we see that multiple selective opportunities exist to alter the PGC population that continues onto gametogenesis. The ultimate

effect that PGC selection can have on reproductive fitness and even the fitness of organisms will be a fascinating area of study.

The advent of many single-cell analytical techniques now enables us to more comprehensively describe the variation within PGCs. Early attempts at morphological phenotyping are bolstered by single-cell transcriptomics, epigenomics, and single-molecule imaging. In addition to evaluating differences among PGCs at each stage of development, complex barcoding can also extend our understanding how these differences change over developmental time. This is especially relevant to the germline given the dynamic challenges faced by PGCs and the varied capabilities assessed throughout their entire journey. Combining these technologies with a longitudinal perspective will clarify how each event during PGC development contributes to improving gamete quality.

Heterogeneity is a ubiquitous property of biology, but this is especially amplified when a population is faced with making a highly regulated and complex transition such as PGCs in the early embryo. Recent advances in reproductive technology have enabled the generation of mouse germ cells in vitro that produce gametes capable of fertilization at exceedingly low efficiency (Hikabe et al., 2016; Ishikura et al., 2016). These germ cells do not undergo the same complex in vivo development and it remains to be seen how bypassing this lengthy process affects the fidelity of their development. Understanding the events and criteria of selection during early embryonic PGC development can improve how we evaluate in vitro-derived gametes to ensure the highest quality ones will be chosen.

Chapter 2 Revealing intrinsic heterogeneity and niche-regulated development in migrating primordial germ cells

2.1 - Introduction to migration of primordial germ cells in mouse

As the first cell in the reproductive lineage, primordial germ cells (PGCs) occupy an important place in reproductive development. These cells will ultimately generate the mature germ cells in the testis and ovary (Lawson and Hage, 1994), which contribute to future generations. Expansion of the ~40 mouse PGCs specified at E7.25 (Anderson et al., 2000; Canovas et al., 2017; Ginsburg et al., 1990) eventually gives rise to the whole population of the organism's reproductive cells called spermatocytes or oogonia. PGC specification from the epiblast is driven by bone morphogenetic protein (BMP) signaling from the extraembryonic tissue (Lawson et al., 1999; Ying et al., 2000). After specification, PGCs migrate from their site of specification to the forming gonadal ridges, the precursor to the gonads. Signaling cues such as Cxcr4 or KitL are known PGC chemoattractants, and contribute to this journey while also promoting PGC survival (Molyneaux et al., 2003; Pesce et al., 1993). Studies observing this migratory process characterize distinct behavioral patterns in the migratory population. Time lapse imaging across the period of mouse germ cell migration from E9.5-E11.5 shows that PGCs early in their migration are distinctively motile, but kinetic within the confines of the hindgut (Molyneaux et al., 2001). At E9.5, a behavioral change occurs, as PGCs begin exhibiting directed migration out of the hindgut and into the dorsal mesentery tissues. The first 'pioneer' PGCs exiting the hindgut, called migratory leaders, are contrasted with the PGCs remaining in the posterior hindgut, called migratory laggards. This targeted migration continues into E10.5, where leading PGC migrants home in on the developing gonadal ridges (Molyneaux et al., 2001). Throughout this process, PGCs traverse diverse tissue niches, with migratory leaders arriving in the gonadal ridges, actively migrating cells still in the dorsal mesentery, and migratory laggards remaining in the

hindgut. By E11.5, most PGCs have begun colonizing the gonadal ridges; however, a population of PGCs remains medial to the gonad, with some PGCs on the midline, and is considered ectopic (Runyan et al., 2008). Aside from the temporal and spatial differences in their location, why some PGCs migrate earlier than others remains unknown. Are there transcriptional changes associated with the behavioral states that we can observe during PGC migration? Here we sought to answer that question, drawing upon recent advances in single cell genomics.

Upon completion of migration, PGCs initiate the next phases of their reproductive development, making this a fascinating timepoint to study from a developmental context. As germ cells initiate differentiation in the gonad via signals from somatic support cells, differences between PGCs in the gonad and those in ectopic locations could provide insights into their differentiation cues and their sources. For example, the erasure of parent-specific epigenetic imprints in PGCs - which begins as early as E10.5 in mouse and progresses in the male development until E13.5 (Saitou et al., 2012) - could potentially contribute to heterogeneities in PGCs given its concurrence with migration. Genes typically silenced by imprinting are expressed from a single parent of origin fashion; some genes are maternally imprinted and express only from the paternal copy, while others are paternally imprinted and expressed only from the maternal copy (Szabó et al., 2002). This removal of silencing methyl marks in the loci of imprinted genes occurs as part of a larger process of genome-wide epigenetic reprogramming in PGCs that begins at E9.0 and continues through E11.5; it encompasses the loss of methylation marks at CpG islands as well as marks on histone tails (von Meyenn and Reik, 2015). Genome-wide DNA demethylation revokes silencing on portions of the genome such as retroviral elements and transposons (McCLINTOCK, 1950), which in can permit these mobile elements to jump around the genome, potentially destabilizing important genes or regulatory elements. (Transposition is also

discussion in Chapter 1, page 27, 31-33.) Since retrotransposition can be a dangerous process, this demethylation seems puzzling. However, demethylation is important for sex specific differentiation, as the male and female germline have their CpG methylation re-established at different rates (Greenberg and Bourc'his, 2019). In the male germline, remethylation begins at E14.5 and continues until birth; male germ cells maintain this sperm-specific methylation pattern until the formation of haploid spermatids and meiotic entry (Kota and Feil, 2010; Zeng and Chen, 2019). In the female germline, oocytes will arrest in meiosis II and will not reacquire methylation marks until after birth, in oocyte growth phase (Kota and Feil, 2010; Zeng and Chen, 2019). The rate of DNA methylation reacquisition varies when comparing male and female germ cells, as sperm genomes are almost fully methylated (90% of CpG dinucleotides), while oocyte genomes are much lower (~40% GpG dinucleotides) (Kobayashi et al., 2012; Zeng and Chen, 2019). Meiotic entry is closely linked to demethylations; if methyltransferase DNA methyltransferase 1 is lost in germ cells, germ cells lose their ability to maintain their methyl marks, leading to premature DNA demethylation and premature meiotic entry (Hargan-Calvopina et al., 2016). Tet1 enzyme actively demethylates promoters during demethylation; loss of this gene prevents active demethylation of promotors for important genes involved in gamete formation and meiosis, which leads to a delay in meiotic initiation (Hill et al., 2018). Thus, the wave of demethylation prepares germ cells for sex specification and reestablishment of the protective marks in a sex specific context.

Investigating transcriptional differences between individual PGCs during migration is difficult as the cell population, while expanding, starts out quite small with just ~200 cells present at E9.5 (McLaren, 2003), ~1,000 cells present at E10.5 (Saiti and Lacham-Kaplan, 2007), and ~2600 cells present at E11.5 (Laird et al., 2011). Some groups have begun using droplet- based sequencing methods to investigate early mouse development at the level of single cell

transcriptomes (He et al., 2020; et al., 2021; Nowotschin et al., 2019), however, these studies preclude in depth the study of migratory PGC behavior and transcriptional dynamics in any depth due to the relatively sparse number of migratory PGCs in their respective datasets.

By pooling many somite matched embryos, we created a large single cell population and interrogated rare subtypes of cells which might be missed in smaller libraries. In addition to profiling the transcriptional heterogeneity in the migrating germ cell population, we designed our sequencing libraries to contain both PGCs as well as representative cells from the migratory niches. By probing the receptor-ligand pairs of migrating PGCs and their somatic niche, we uncover new interactions which make migration a successful process, planting seeds for future generations.

A longstanding question is whether PGCs that successfully colonize the gonad are different from those that remain in the midline, are considered ectopic, and may be eliminated. In this study, we used single cell RNA sequencing to profile PGCs at unprecedented depth and resolution across migration. At three timepoints and in spatially-distinct parts of the embryo, we identify developmentally-regulated gene expression in PGCs and the lineages of nearby somatic cells, which include the relevant supporting cells for the PGCs in transit. We term this somatic conglomerate the ‘migratory niche’. By predicting the receptor-ligand pairs of migrating PGCs and their migratory niche, we uncover new interactions which make migration a successful process. We also investigate how the migratory niche contributes to the state change between an early migrant and a successful colonizer of the gonad. Finally, we reveal germ cell subpopulations at the conclusion of migration. We bioinformatically identify a signature of highly expressed imprinted genes in a subpopulation, and show by whole-mount *in situ* hybridization that these cells are located in extragonadal tissues. We hypothesize that these extragonadal

PGCs may not be able to complete migration. These data characterizing intrinsic and niche associated heterogeneity in PGCs provide a resource for an in depth understanding of differences in migrating PGCs.

2.2 - Migrating germ cell isolation and single cell RNA sequencing library creation of migratory timepoints

To ascertain how mouse PGCs might differ across the migratory time period, we performed droplet based single cell RNA sequencing to create libraries from early embryonic timepoints during which PGCs are migrating from their site of specification to the forming gonads (Molyneaux et al., 2001). We created libraries from E9.5, E10.5, and E11.5, to span the beginning of directed migration and into the period of when PGCs settle into the forming gonadal ridges (Figure 2.1A). Since PGCs are few in number at early migratory timepoints (~200 cells, E9.5 (McLaren, 2003); ~1,000 cells, E10.5 (Saiti and Lacham-Kaplan, 2007)) timepoints, we used the reporter Oct4-ΔPE-eGFP (MGI:3057158, (Szabó et al. 2002)) to enrich PGCs from multiple age-matched embryos (see Methods). During active migration, we sought to determine if cells leading the migration were distinct from cells lagging behind. To do this, we constructed libraries from E9.5 and E10.5 embryos bisected into an anterior half (leading migrants, Figure 2.1A anterior of split) and a posterior half (lagging migrants, Figure 2.1A posterior of split) using somite number to split embryos consistently (see Methods). At E11.5, 90-95% of PGCs have arrived at the forming gonadal ridge (Laird et al., 2011) with some remaining in the midline between the gonadal ridges (Figure 2.1A, arrowhead), others in the dorsal mesentery around the gonads and others in the tail mesoderm (Molyneaux et al., 2001). We dissected and dissociated the aorta-gonads-mesonephros (AGM) to enrich for both successful migrants, as well as PGCs remaining in the midline and mesentery proximal to the

gonads. While we did not split the libraries into gonadal ridge vs midline, we were able to bioinformatically identify some extragonadal PGCs later in analysis. A final component of our libraries was the inclusion of GFP -negative somatic cells to survey the supporting niche at each migratory timepoint.

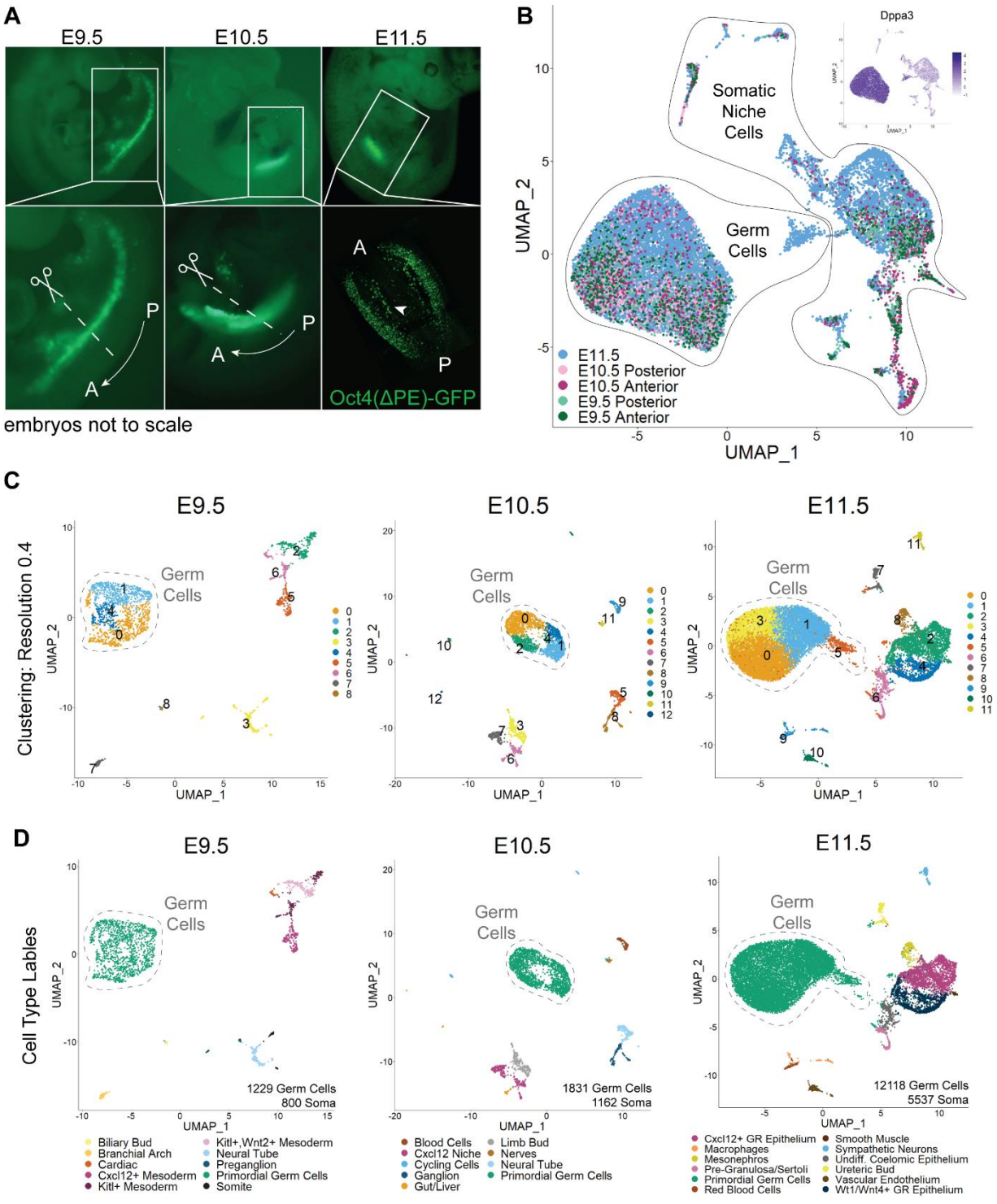


Figure 2.1 Single cell RNA sequencing of migratory germ cells and their somatic niche across developmental time

A) To assay transcriptional heterogeneity in migrating mouse primordial germ cells, single cell RNA sequencing libraries were created from E9.5, E10.5, and E11.5. *Oct4(ΔPE)-GFP* transgene expressing embryos were dissected, and GFP positive PGCs concentrated by FACS.

At E9.5, embryos were somite staged to 20-25 and kept. Anterior/Posterior split (dashed line) occurred at roughly somite 15.
At E10.5, embryos were somite staged to 34-38 and kept. Anterior/Posterior were split between somites 19/20. Curved arrows indicate direction of migration.
At E11.5, PGCs in the genital ridge and PGCs in the midline were included in the same library. Extragonadal, medial PGCs are indicated by the arrowhead. For each timepoint, we spiked back in GFP negative somatic cells, to enable probing of receptor-ligand pairs between PGCs and somatic cells. Embryos are not to scale.
B) UMAP dimensional reduction plot of all three timepoints, and five anatomical locations. PGCs were identified by expression of *Dppa3* (purple inset). E11.5 cells are labeled in blue. E10.5 Anterior cells are light pink, and E10.5 Posterior cells are dark pink. E9.5 Anterior cells are light mint green, and E9.5 are dark green. PGCs and somatic niche cells are circled and indicated on the UMAP plot.
C) Individual timepoints were bioinformatically clustered at a resolution of 0.4. Dashed outline identifies PGCs.
D) Cell type labels assigned to the somatic cells from each timepoint. PGCs are indicated by a dashed outline.

In total, we profiled 22,677 cells (15,178 PGCs and 7,499 somatic cells) from all migratory timepoints (Figure 2.1B). Results are visualized on Uniform Manifold Approximation and Projection (UMAP) dimension reductions. We identified PGCs based on expression of *Dppa3* (Figure 2.1B, inset) and other germ cell markers (see Methods). We sequenced 1229 E9.5 PGCs, 1831 E10.5 PGCs, 12118 E11.5 PGCs. We acquired an average of 1912 genes per cell at E11.5, 2550 genes per cell at E10.5, and finally 3942 genes per cell at E9.5. We next began to cluster the data and identify the germ cell and somatic cell components present. Unbiased clustering was done while varying the resolution value; each dataset was surveyed and we selected a resolution of 0.4 (Figure 2.1C) to use to identify somatic cell types present in the datasets, while also providing some germ cell subclusters (Figure 2.1C, dashed circles) to interrogate for heterogeneities. While the limited cell number of PGCs present in the tissue at E9.5 and E10.5 did not permit us to construct multiple libraries from multiple biologic replicates, we prepared the E11.5 library in biological duplicate. The samples for each replicate were dissected on and sequenced on different days can be well integrated (Figure 2.2A), and each replicate contributes cells to every cluster (Figure 2.2B).

To characterize the somatic cell populations, we utilized both bioinformatics-based approaches as well as publicly available *in situ* hybridization databases. The classification strategy used to identify cell types is depicted in the flowchart in Figure 2.2C. The genes used to classify somatic cells are depicted in Figure 2.2D, and somatic cell types identified are shown in a UMAP dimensionality-reduction plot for each time point (Figure 2.1D, Figure 2.2E). At E11.5, the most developmentally advanced phase, we also detected precursors to the gonadal support cells (like the Sertoli cells, granulosa cells, and Leydig cells) as pre-differentiated coelomic epithelium (Figure 2.1D, E11.5). From each timepoint, among the differentially-expressed genes that distinguished some somatic cell clusters, we identified known chemoattractants and survival signals for PGCs including *Sdf1*/*Cxcl12* and *KitL* in the *Cxcl12* positive Niche, or the *KitL* positive Niche named clusters. We used these known supporting niche markers to identify clusters from which supportive signals to the PGCs might be derived.

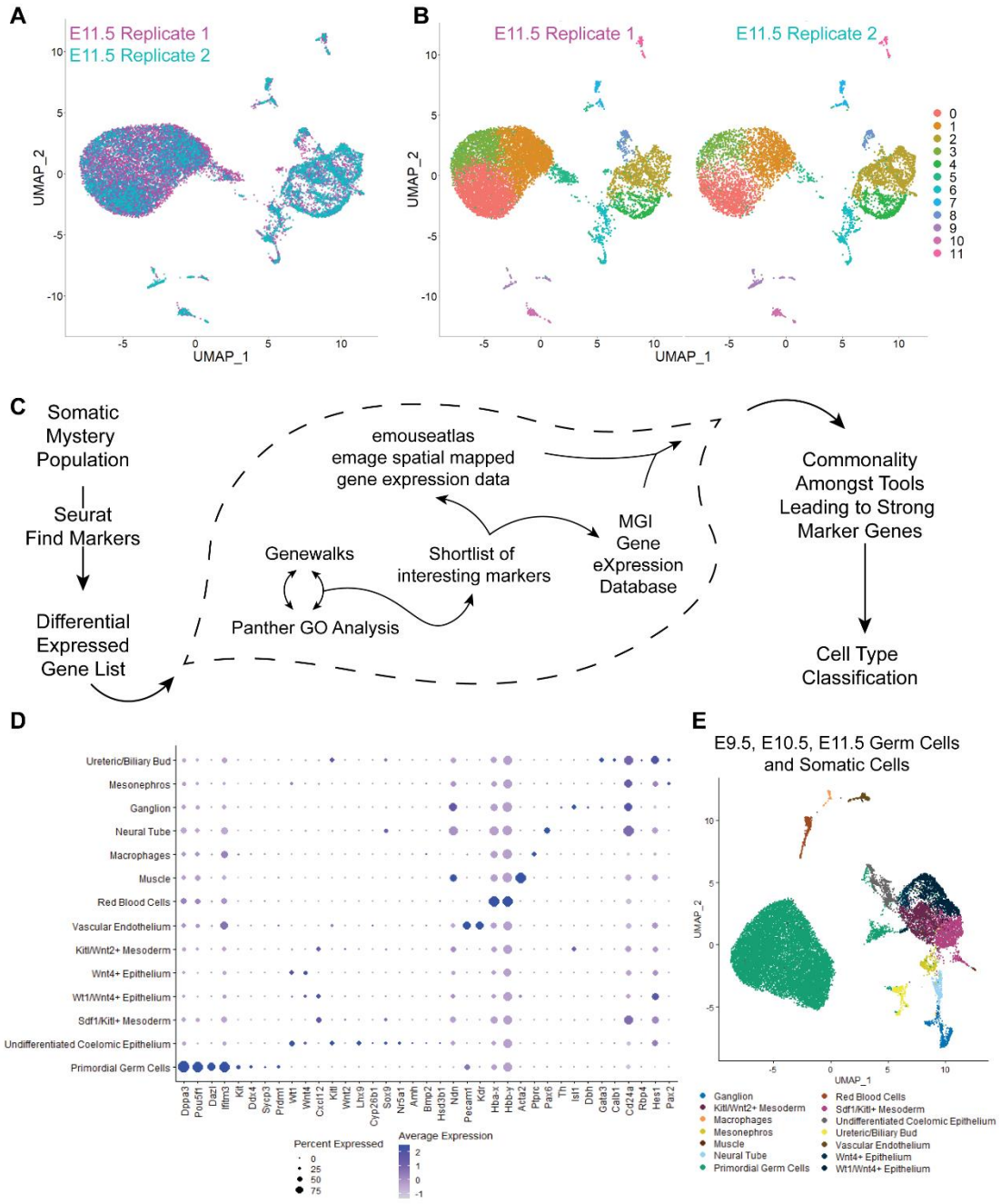


Figure 2.2 (Supplemental for 2.1) E11.5 biologic replicates integrate well, and somatic identification scheme can be used to label somatic cell populations

A) Larger numbers of migratory PGCs present at E11.5 permit the construction of libraries from biologic replicates (shown in purple and cyan, respectively). Cells from each replicate are found in all areas of the dimensional reduction, indicating the libraries integrate well with no obvious batch effect.

B) Replicate 1 and 2 each contribute to every bioinformatic cluster.
C) Strategy for annotating somatic clusters combined a bioinformatics based approach with RNA *in situ* databases. Beginning with differentially expressed genes for each cluster, we used a combination of GO Analysis (to classify biologic processes the cluster might be involved in) and the Genewalks program (to filter out 'promiscuous markers commonly found in many cell types) to identify a 'shortlist' of markers. Next, we checked expression of markers at E9.5-E11.5 in RNA *in situ* databases, seeking expression in the migratory niche to confirm our bioinformatic hypothesis. Finally, taking GO Analysis, Genewalks, and *in situ* results together, we classified the somatic cell type.
D) Markers used to identify somatic cell types.
E) UMAP of all migratory timepoints combined, with somatic cell types labeled.

2.3 - Predicting cell signaling between migrating mouse primordial germ cells and their supporting niche

With nearby somatic cell types identified, we sought to uncover potential interactions between migrating PGCs and these populations in order to identify their supportive niche. We hypothesize that different receptor-ligand mediated signaling may guide PGCs at different locations and phases of migration. With separate libraries that incrementally span the migratory time period, identification of receptor-ligand pairs unique to a particular time was possible. We used CellphoneDB (Efremova et al., 2020) to match ligands expressed by somatic sender cells with corresponding receptor expressed by PGC receiver cells. Although CellphoneDB was developed for human data, previous work shows that converting the human genes to their mouse orthologs (using a package like BioMart (Smedley et al., 2009)) gives valid results (Saatcioglu et al., 2019). We thus applied CellphoneDB to our datasets, thresholding resulting receptor-ligand pairs to identify the most promising candidates for further study (see methods).

We expect that receptor-ligand interactions shared between the E9.5 and E10.5 datasets would include long-range mechanisms of chemoattraction or repulsion, survival signals, adhesion or motility-promoting cues requiring cell-cell contact throughout the migratory route. A known

chemokine-receptor pair that we identified was Cxcl12-Cxcr4 at both migratory timepoints (Figure 2.3A-B, top pair). The chemokine Cxcl12, or Sdf1, is known to be a potent chemoattractant and survival cue for migrating mouse PGCs via the receptor Cxcr4 (Ara et al., 2003; Molyneaux et al., 2003). We expected the most meaningful interaction data to come from the Sdf1 positive niche population signaling to the migrating PGCs. While Sdf1 expression is also highly significant in the brachial arch and biliary bud categories, these tissue niches are outside the migratory route and likely not truly signaling to the PGCs. Similarly, the CellphoneDB analysis reported a significant Kit-Kitlg interaction, although the detection in a subset of PGCs and coupled with many different somatic cell types does not reflect prior expression and genetic data and might indicate dropouts; KitL (also known as SCF or Steel Factor) is a critical chemoattractant and survival cue for migrating PGCs (Besmer et al., 1993) (Figure 2.3A-B, second pair).

A novel suite of receptor-ligand pairs identified at E9.5 and E10.5 that has not been previously implicated in PGC migration is Eph-Efn (Figure 2.3, second category). Ephrins and their Eph receptors are involved in many migratory processes (Park and Lee, 2015; Pasquale, 2010; Santiago and Erickson, 2002), by mediating the stop and go of migration. Their role in PGC migration is previously unknown, but the presence of Ephrin receptor-ligand pairs at both migratory timepoints underscores their potential importance.

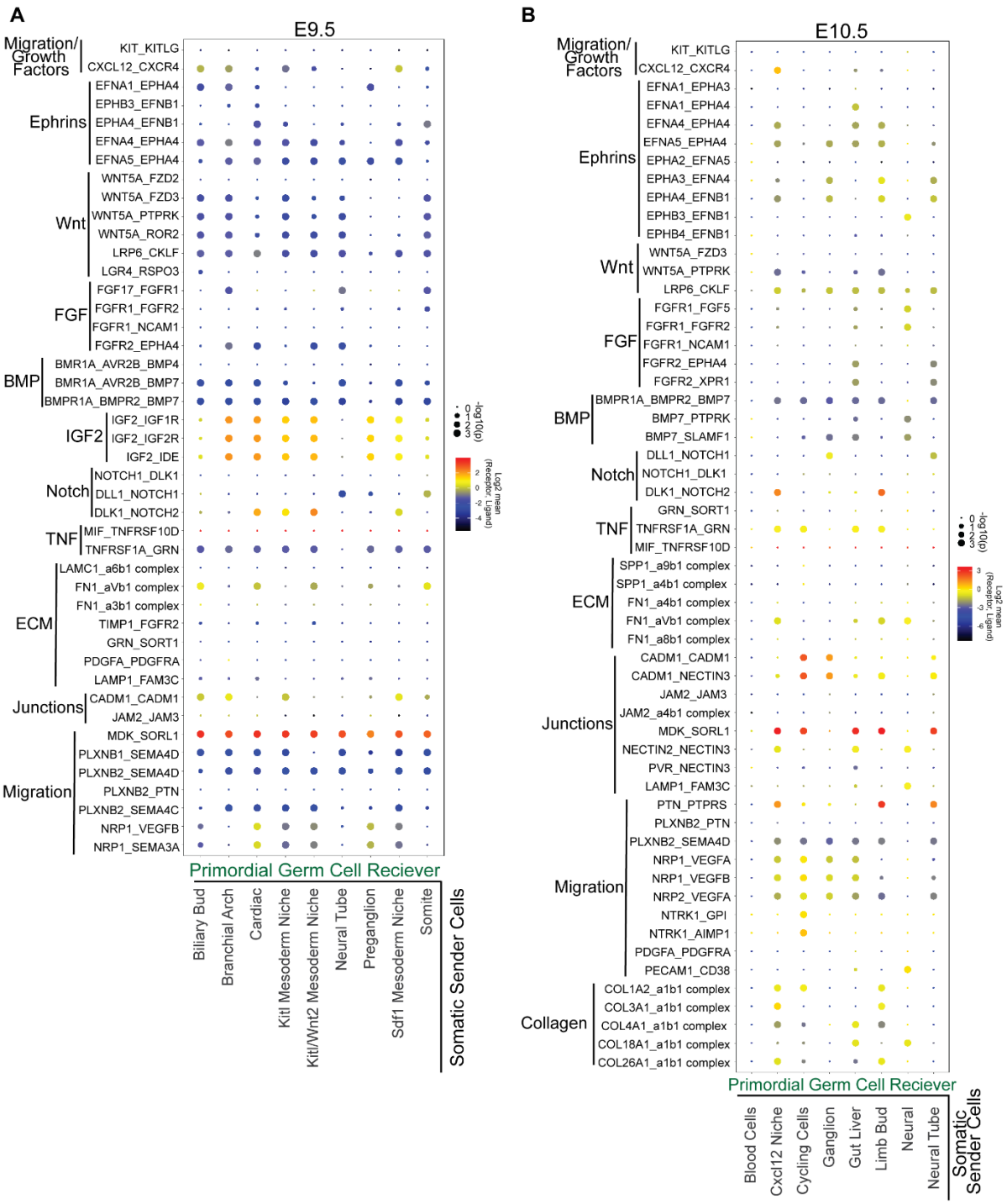


Figure 2.3 Interrogation of the somatic expressed ligands originating from the migratory niche and the companion receptors in migrating mouse PGCs with CellPhoneDB

Using CellPhoneDB, a package which queries receptor-ligand interactions, we categorized the potential interactions between migrating PGCs at the somatic niche through which they are migrating.

A) Significantly enriched receptor ligand interaction pairs identified between E9.5 mouse somatic cells (shown in columns), which send ligands, and migrating PGCs, which host the receptor. Receptor ligand pairs (rows) are categorized according to signaling pathway, or putative function. P values are indicated by circle size; mean of the average expression level of interacting molecule pairs in each cluster are indicated by the color scale.
B) Receptor ligand interaction pairs identified between E10.5 somatic cells and PGCs. At this time point, notable receptor ligand pairs include, but are not limited to, the increasing about of ECM/Collagen/Junction related receptor ligand pairs, perhaps reflecting the more advanced migratory germ cell population colonizing the gonad and transition out of a migratory profile and into a more attached profile. Note that at E10.5, there are still lagging cells in migratory phases, and we expected to and do see a Migration category here as well.

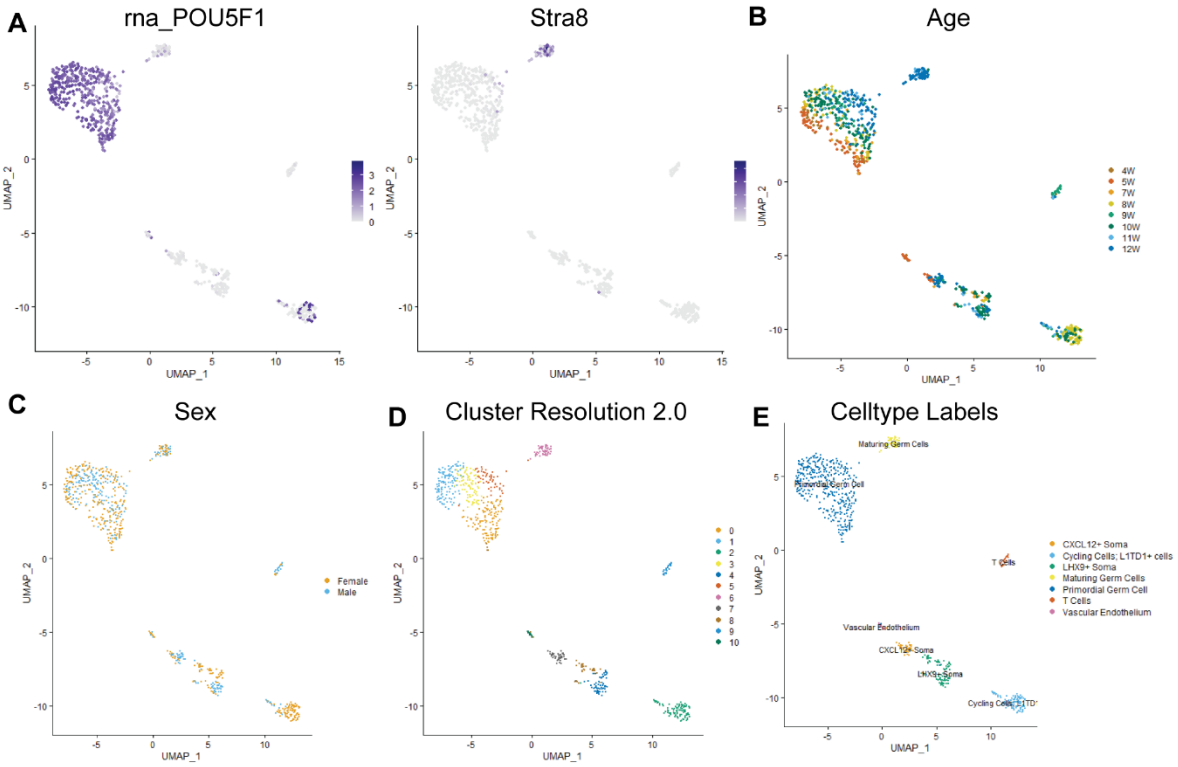
To identify receptor-ligand pairs unique to early migration, we focused first on the E9.5 timepoint. Wnt signaling emerged prominently, with Wnt5a ligand coupling to Ror2, Ptprk and Frizzled2/3 in E9.5 PGCs (Figure 2.3A, Wnt Category); Wnt5a-Ror2 signaling promotes motility of PGCs over proliferation in the hindgut (Cantú et al., 2016; Laird et al., 2011). Bmp7 and to a lesser extent, Bmp4 expressed by multiple somatic cells potentially interacts with Bmr1a/Avr2b expressed in the PGCs (Figure 2.3A, Bmp Category). Bmp signaling is known to be necessary for PGC specification (Lawson et al., 1999; Ying and Zhao, 2001; Ying et al., 2000), but less is known about what role it may play in migration.

The E10.5 dataset contains PGCs in the most migratory niches, with the most advanced cells already settling into the gonadal ridge, cells in the midst of migration in the dorsal mesentery, and lagging cells in the hindgut niche. Accordingly, the receptor-ligand pairs present at E10.5 reflect this varied physiologic positioning, with categories focused on migration, survival cues, and signaling pathways like Eph, Wnt, Bmp, and Notch (Figure 2.3B). We also found more extracellular matrix (ECM), Junction-related, and Collagen receptor-ligand pairs compared to E9.5. ECM molecules have been long appreciated for their contribution to cellular migration, and germ cell migration in particular (Alvarez-Buylla and Merchant-Larios, 1986; García-Castro et al., 1997; Soto-Suazo et al., 2004, 1999). Compared to E9.5, we observed more receptor-ligand

pairs which mediate establishment of adherens junctions (like the Nectin genes) and ECM related pairs (Figure 2.3B, ECM/Junctions/Collagen Categories). At this time point, we also see an increase in novel Eph/ephrin signaling receptor-ligands. These pairs may be linked to colonization of the gonad as well, since Eph/ephrin signaling pathways help coordinate adhesion and motility (Singh et al., 2012).

2.4 - Identifying PGC-niche receptor-ligand pairs shared between human and mouse

We hypothesized that some receptor-ligand pairs we identified in mice might also be found in analogous reproductive developmental time points in human samples. To test this, we turned to a previously published single cell RNA sequencing library (Li et al., 2017) containing PGCs and somatic cells from 4W-12W of human fetal development. We reanalyzed these data with Seurat and CellPhoneDB. We are able to identify the PGCs present at these timepoints using *OCT4* to mark PGCs, and *STRA8* to mark maturing germ cells (Figure 2.4A). Our reanalysis found that human early germ cells are arranged in a gradient of time, with a small outlier cluster of more mature 12W germ cells (Figure 2.4B). While these timepoints precede sex specification, inclusion of sex genotyping in the metadata revealed that germ cells from male and female do not segregate based on their sex; however, somatic tissue did somewhat group according to sex even at these early developmental timepoints (Figure 2.4C). Clustering divided the germ cells into 5 subclusters, while somatic cells were parsed into 5 clusters (Figure 2.4D), which were classified as previously described (Figure 2.4E).



F Pairs shared between E9.5 mouse, E10.5 mouse, and 4W-12W human

Interacting Pair	Receptor and Ligand Status		Ligand Secreted
KIT_KITLG	**KIT**	KITLG	TRUE
CXCL12_CXCR4	**CXCR4**	CXCL12	TRUE
EFNA1_EPHA4	**EPHA4**	EFNA1	TRUE
EFNA4_EPHA4	**EPHA4**	EFNA4	TRUE
WNT5A_FZD3	**FZD3**	WNT5A	TRUE
LRP6_CKLF	**LRP6**	CKLF	TRUE
FGFR1_FGFR2	**FGFR1**	**FGFR2**	TRUE
FGFR1_NCAM1	**FGFR1**	**NCAM1**	TRUE
TNFRSF1A_GRN	**TNFRSF1A**	GRN	TRUE
FN1_aVb1 complex	FN1		TRUE
GRN_SORT1	**SORT1**	GRN	TRUE
PDGFA_PDGFRA	**PDGFRA**	PDGFA	TRUE
LAMP1_FAM3C	**LAMP1**	FAM3C	TRUE
JAM2_JAM3	JAM2	JAM3	TRUE
MDK_SORL1	**SORL1**	MDK	TRUE
PLXNB2_PTN	**PLXNB2**	PTN	TRUE
EPHB3_EFNB1	**EPHB3**	EFNB1	FALSE
EPHA4_EFNB1	**EPHA4**	EFNB1	FALSE
EFNA5_EPHA4	**EPHA4**	EFNA5	FALSE
DLK1_NOTCH2	**NOTCH2**	DLK1	FALSE
CADM1_CADM1	CADM1	CADM1	FALSE

Bolded genes are listed as "Receptor" in the Database.

Purple gene is listed as "Integrin" in the Database.

Figure 2.4 Cross-species comparison of clustering and receptor-ligand pairs in similarly-staged human migratory PGCs

To draw cross species comparisons on the receptor-ligand pairs, we re-analyzed sequencing data on human fetal germ cells and their somatic niche from 4W, 5W, 7W, 8W, 9W, 10W, 11W, and 12W, the most developmentally similar times to the E9.5-E11.5 mouse migratory time period.
A) Identification of primordial germ cells (*POU5F1/OCT4*) and maturing germ cells (*STRA8*).
B) UMAP projection showing that the germ cell cluster arrays in a gradient of time, with earliest 4W/5W samples to the left of the cluster, and older timepoints to the right of the cluster.
C) At early timepoints, sex is not a differential factor driving clustering of germ cells, as male and female germ cells clustered together. Somatic cells, however show distinct sex groupings per cluster.
D) Unbiased clustering at resolution 2.0 reveals 5 germ cell sub-clusters and 6 somatic clusters.
E) Somatic populations identified and labeled based on differentially expressed genes as well as previously described methods.
F) CellphoneDB receptor ligand pairs that are shared by E9.5 mouse, E10.5 mouse, and 4W-12W human datasets.

To identify shared interactions between mouse and human PGCs and somatic cells during migration, we intersected values of E9.5, E10.5, and human W4-W12 receptor-ligand pairs. As shown in Figure 2.4F, we identified common signaling through established KIT and CXCR4 receptors as well as signaling mechanisms not previously described in migratory PGCs, including EPH-EPHRIN, NOTCH, and CADM1.

Distinct from mouse PGCs, human PGCs could be resolved by age-specific markers; we identified genes with bands of temporal expression that spanned the W4-5, W7-10, W11-12 (pluripotent), and W11-12 (maturing) time periods (Figure 2.5). In agreement with previously published work which reanalyzed some of these timepoints, our reanalysis of the earliest human PGC and somatic tissue also identified WNT, TGFB, and Notch signaling pathways in the early timepoints (Overeem et al., 2021). We were also able to identify many other categories labeled in detail in Figure 2.6.

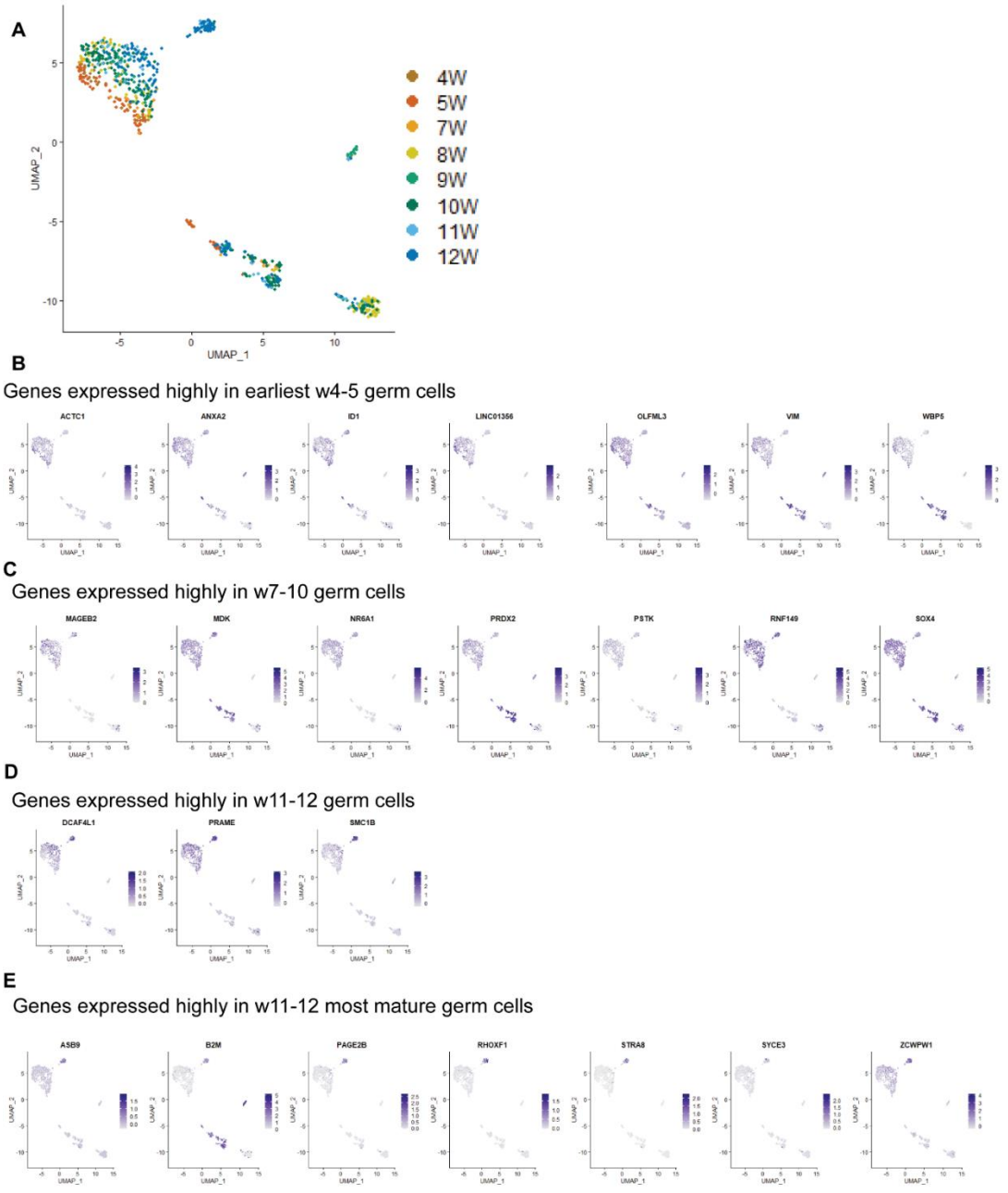


Figure 2.5 (Supplemental for 2.4) Human migratory germ cells display temporally-specific gene expression

Investigation of early human fetal germ cells reveals genes which are more highly expressed at different timepoints.
A) UMAP projection showing that the germ cell cluster arrays in a gradient of time, with earliest 4W/5W samples to the left of the cluster, and older timepoints to the right of the cluster.
B) Genes expressed most highly in earliest w4-5 PGCs.
C) Genes expressed most highly in w7-10 PGCs.
D) Genes expressed most highly in w11-12 PGCs.

E) Genes expressed most highly in w11-12 maturing germ cells, as defined by the expression of sex-specific genes.

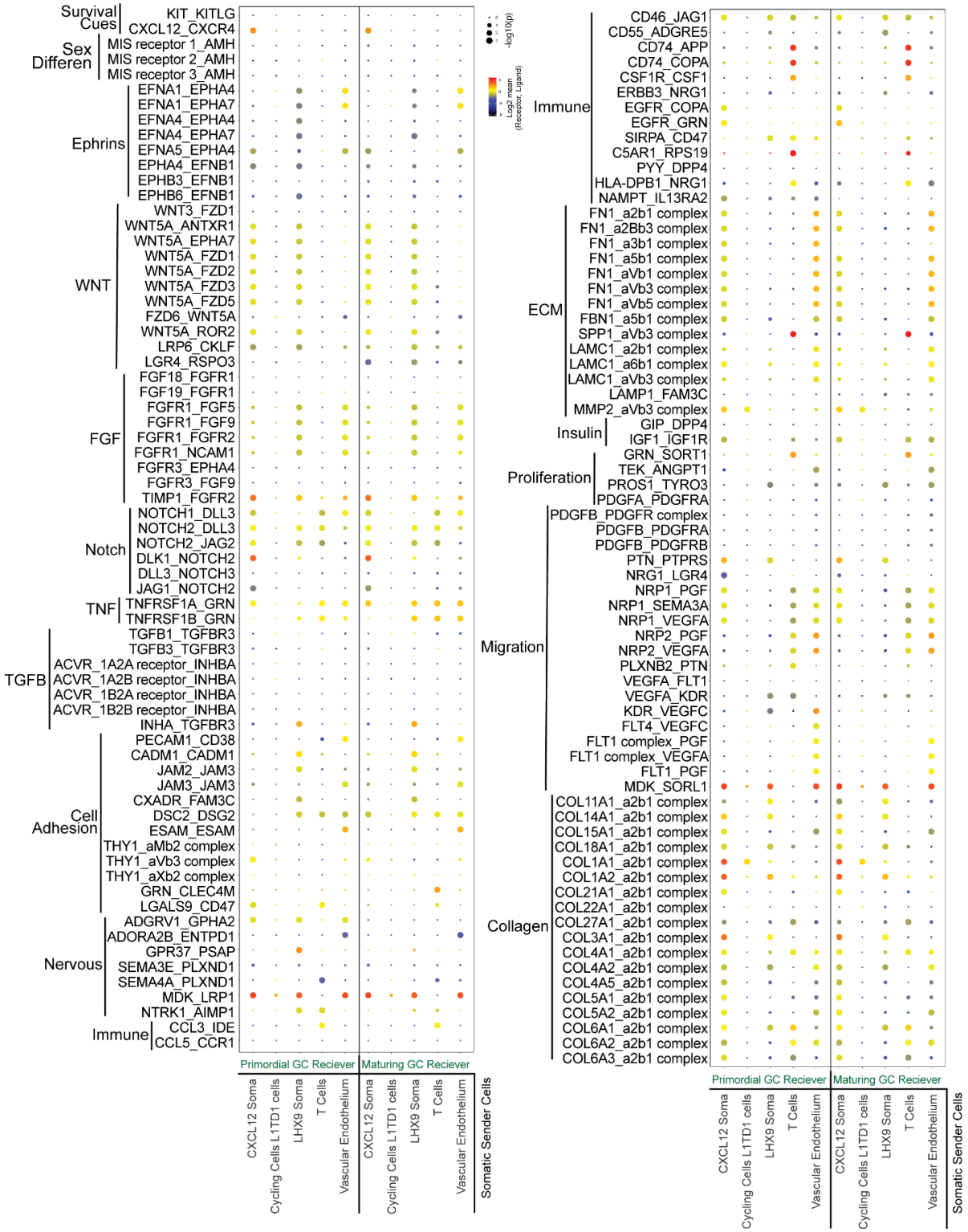


Figure 2.6 (Supplemental for 2.4) Interrogation of the somatic expressed ligands originating from the migratory niche and the companion receptors in migrating human germ cells with CellPhoneDB

Using CellPhoneDB, a package which queries receptor-ligand interactions, we categorized the potential interactions between human migrating germ cells at the somatic niche through which they are migrating. In general, more receptor ligand pairs are observed than in mice. We again categorized the pairs. Unlike mouse, human germ cells and their somatic niche began showing a sex differentiation category of receptor ligand pairs. Human germ cells and their somatic cells lack the BMP signaling found in mice. Additionally, due to the inclusion of T cells in the human library, there is a large Immune profile that is likely a technical artifact.

Finally, to investigate receptor-ligand pairs involved in specific developmental transitions, we turned to the NicheNetR package (Browaeys et al., 2020). While CellphoneDB examines all ligand-receptor interactions in a dataset, NicheNet goes one step further to probe receptor-ligand pairs which may mediate a state transition in a cell population (Figure 2.7A). We probed the developmental transition of mouse PGCs between E9.5 and E11.5, in order to compare an actively migrating population to one settled into the gonadal ridge. To set up the test, we generated a differential gene list by comparing an E9.5 PGCs (green cells, Figure 2.7B) to an E11.5 PGCs (orange cells, 8B). Then, any receptor-ligand interactions known to induce expression changes in those genes were queried in PGCs (receptor cells) and the somatic niche populations (coordinating ligand; Figure 2.7C). The ligands are also linked to predicted target genes which potentially mediate the E9.5 to E11.5 state change (Figure 2.7D). In the human dataset, we compared early PGCs (Figure 2.7E, dark blue cells) to later PGCs (Figure 2.7E, orange W11-12 cells; still expressing pluripotency markers). We similarly acquired a list of validated ligand-receptor pairs (Figure 2.7F), and then predicted target genes for top ligands in the human dataset (Figure 2.7G). We found that the state change from early to late PGC migrants revealed Slit-Robo signaling in both human and mouse (Figure 2.7C and F, bolded). Cldn11 was also shared between mouse and human (Figure 2.7C, F, italics).

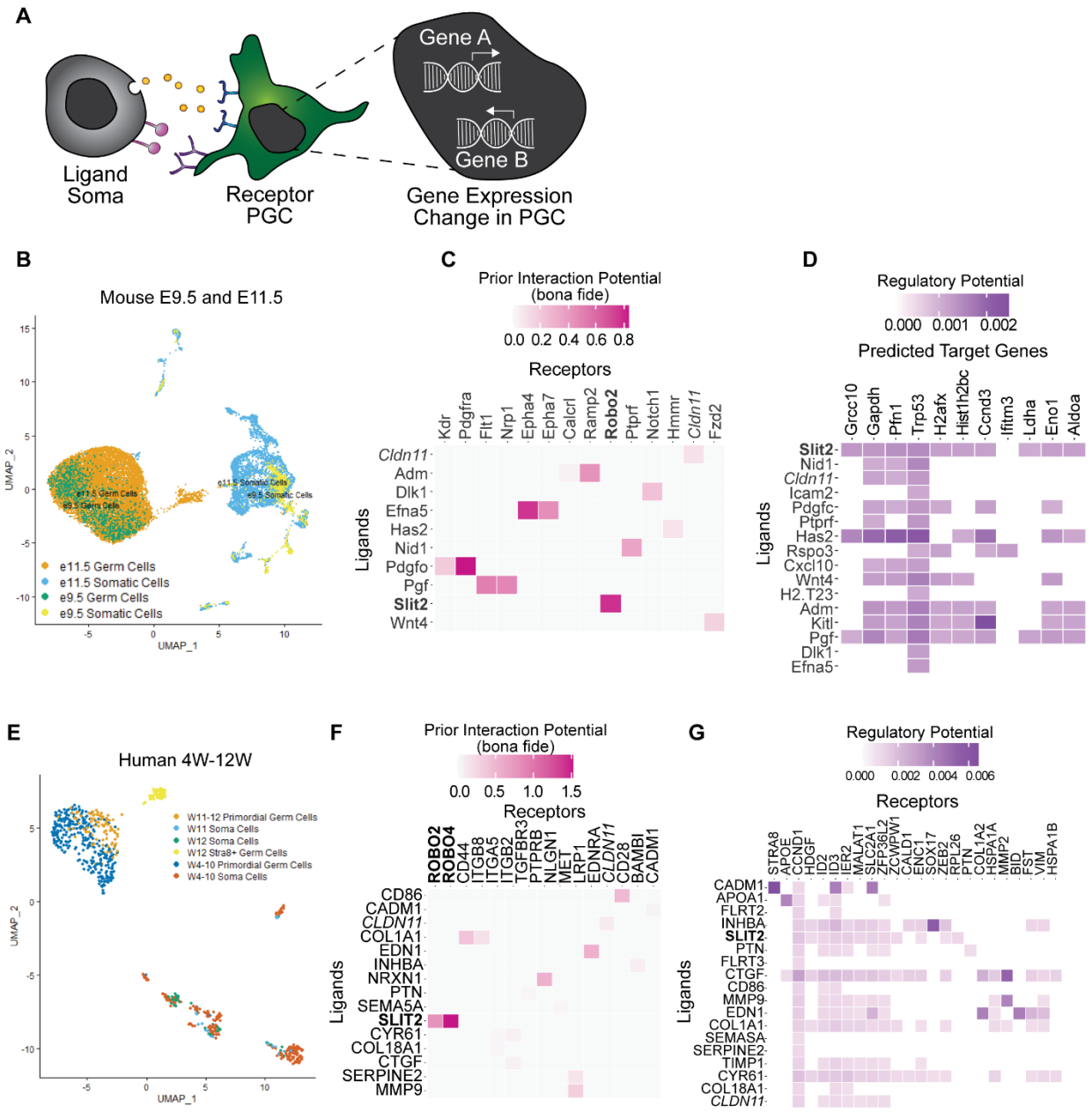


Figure 2.7 In mouse and human, the state change between early migrating PGCs and gonadal ridge colonizers involves Slit-Robo signaling

To further probe receptor ligand mediation of migrating germ cells, we investigated how receptor ligand pairs might be involved in mediating the state transition from an early migrant (E9.5 mouse, 4-10W human) VS a late migrant at the gonadal ridge (E11.5 mouse, 11-12W human.

A) Diagram detailing the analysis using NicheNet. Somatic migratory niche ligands and their corresponding receptor in the germ cell are identified, and linked to changes in gene expression in the germ cells.

B) In mice, UMAP identifying the state change to be probed by NicheNet: PGCs from E9.5 and E11.5. The corresponding niches were queried for ligands.
C) Somatic ligands and germ cell receptors identified in the transition from mouse E9.5 to E11.5 PGCs. Receptor ligand pairs were only considered if they came from the literature, from a curated ligand-receptor database.
D) Predicted target genes of somatic migratory niche ligand signaling to a receptor in mouse PGCs from E9.,5-11.5.
E) In humans, identifying the state change to be probed NicheNet:4-10W and 11-12 primordial germ cells and their corresponding niche cells.
F) Somatic ligands and germ cell receptors identified in humans. Receptor ligand pairs were only considered if they came from the literature, from a curated ligand-receptor database.
G) Predicted target genes of somatic migratory niche ligands signaling to a germ cell receptors from W4-10 to W11-12.

2.5 - Characterizing transcriptional differences in mouse E9.5 and E10.5 Anterior and Posterior migrants

We next investigated transcriptional heterogeneity within the migratory leaders and migratory laggards in the E9.5 and E10.5 datasets. When originally designing the libraries, we split the PGCs into anterior and posterior libraries based on PGC morphological location (Figure 2.1A, split indicator). While these morphological splits matched PGC activity during the migratory process, it was unknown whether this was reflected in transcriptional differences.

Importantly, at both E9.5 and E10.5, the Anterior and Posterior libraries integrate well without any batch effect (Figure 2.8A). To determine whether transcriptional differences matched the morphological location, we queried whether we could detect known genes which are canonically expressed in the posterior tip of the mouse tail in the posterior E9.5 somatic libraries. We identified somatic genes *Wnt5a*, *Wnt5b*, *Hoxb1*, *Rarg*, *Lfng*, and *T* (Gofflot et al., 1997) as highly enriched in the posterior somatic cells at E9.5 (Figure 2.8B). At E10.5, we took advantage of the well characterized relationship between Hox genes and somites (Wellik, 2007), to test the anterior/posterior binning strategy. Based on previous data, we expect expression of *Hoxd10*

(spanning somites 31-41; (Burke et al., 1995)) only in the posterior somatic cells; our anterior E10.5 library contained somites above 19/20, and our posterior E10.5 library contained somites greater than 21. A Hox gene like *Hoxb4* (which expresses from somite 6 all the way to 41), would be present in both the anterior and posterior somatic cells (Burke et al., 1995). Indeed, we see expression of *Hoxb4* in both anterior and posterior somatic cells, while expression of Hoxd10 is mostly restricted to the posterior library (Figure 2.8C), thus validating the split based on morphological location. Since the splits we introduced meaningfully split the somatic cells, we next sought to test whether anterior vs posterior PGCs showed transcriptional differences.

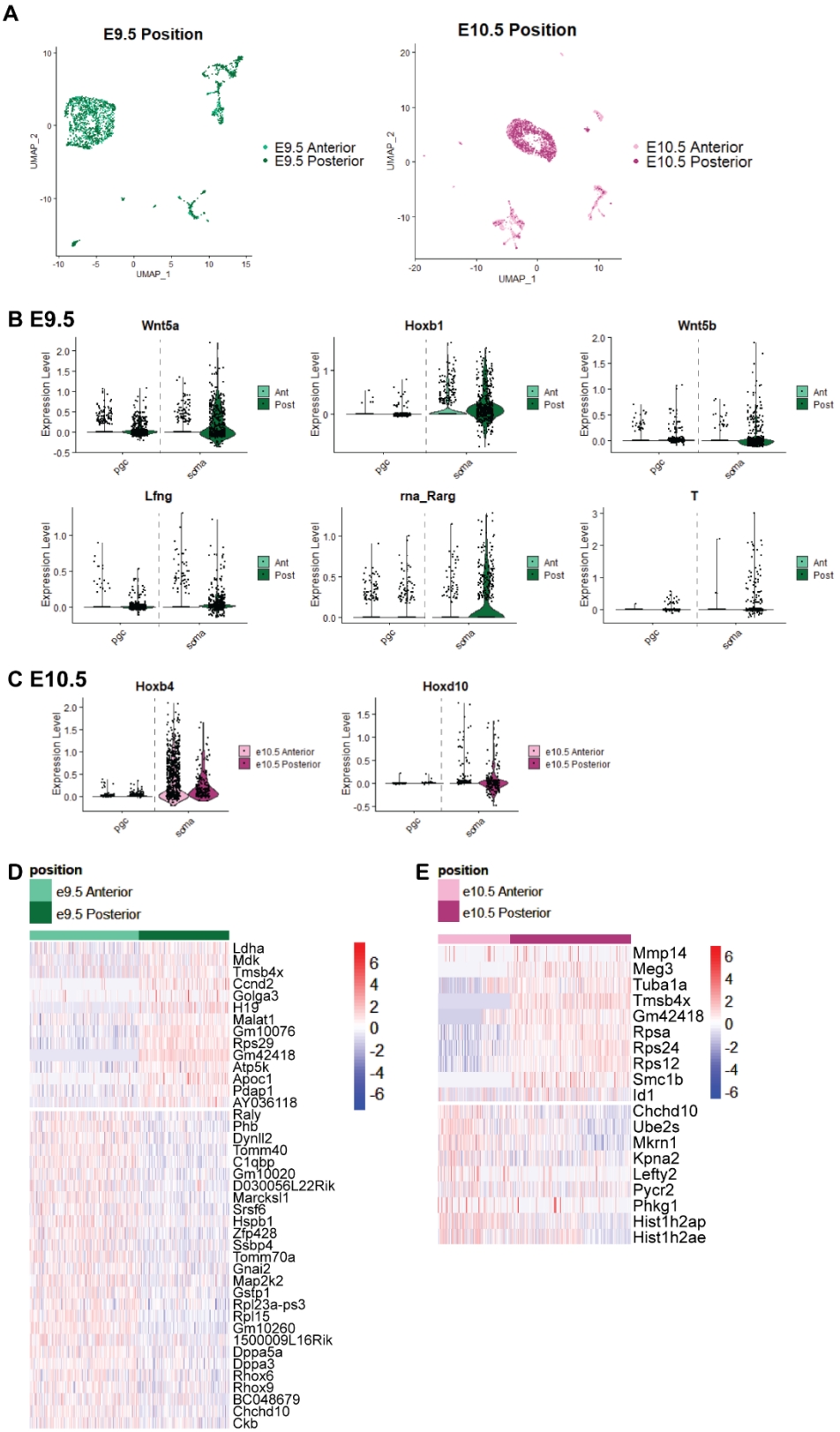


Figure 2.8 Identification of spatially-specific genes in mouse PGCs from anterior and posterior sin the embryo

A) E9.5 and E10.5 Anterior and Posterior libraries integrate well with no batch effect. At E9.5, migratory PGCs were bisected at approximately somite 15 to create the Anterior compared to Posterior libraries.
B) To validate the spatially-discrete libraries, we queried known markers of the embryo posterior in somatic cells. At E9.5, known posterior markers *Wnt5a, Hoxb1, Wnt5b, Lfng, Rarg,* and *T* were elevated in the posterior somatic cells, although expression was not detected in all posterior somatic cells, due to cell type diversity and dropouts.
C) At E10.5, *Hoxb4* expression is known to be restricted tosomites 6-41, and was found in equal abundance in Anterior and Posterior somatic cells, which were split at approximately somite 19/20. *Hoxd10*, expression is restricted to somites 31-41 and was elevated in the posterior soma.
D) Heatmap of differentially expressed genes in mouse E9.5 Anterior PGCs vs the E9.5 Posterior PGCs.
E) Heatmap of differentially expressed genes between mouse E10.5 Anterior PGCs vs E10.5 Posterior PGCs.

To determine if any genes were differentially expressed in either the anterior or posterior compartment, we set up a statistical test to compare total anterior vs total posterior PGCs. This analysis found 41 differentially expressed genes at E9.5 and 19 differentially expressed genes at E10.5 (Figure 2.8C and Figure 2.8D, respectively). At E9.5, a subset of Anterior-upregulated genes are involved in oxidative phosphorylation (*C1qbp*, *ChChd10*), perhaps based on increasing energy demands of active migration out of the hindgut.

Since we combined many embryos at times prior to sex specification, we have both male and female PGCs containing within our mouse datasets. At E10.5, Anterior cells express more *Lefty2* (Souquet et al., 2012) and *Pycr2* (Adra et al., 1987), two genes linked to male sex differentiation, perhaps because some of these leading migrants have already settled into the gonadal ridge and are progressing toward sex specification. However, we don't see a strong cell cycle signature at E10.5, which is something we expected to find based on previous data (Cantú et al., 2016).

To further categorize and understand the genes differentially expressed between anterior and posterior, we classified each gene according to position and function (Table 1). Both Anterior populations were marked by genes associated with metabolism, and *Chchd10* in particular was shared by both E9.5 and E10.5 Anterior. In both posterior migrant populations, imprinting genes came up as population markers. Based on these data, we conclude that PGCs from the Anterior vs Posterior libraries at E9.5 and E10.5 do have transcriptional differences.

Table 2.1 Differential genes from Figure 2.8D-E sorted according to position and function

Up in E9.5 Anterior:								
Germ Cell Pluripotency	Dppa3	Dppa5a						
Inflammation, Heatshock	Hspb1	Phb						
Metabolism	C1qbp	Chchd10	Ckb	Gnai2	Gstp1	Ssbp4	Tomm40	Tomm70a
Migration	Dynll2	Marcksl1	Map2k2					
Miscellaneous	BC048679	Gm10260	Srsf6	Zfp428	1500009L16Rik		D030056L22Rik	
Reproductive Homeobox	Rhox6	Rhox9						
Ribosome	Gm10020	Gm10260	Raly	Rpl15	Rpl23a-ps3			

Up in E9.5 Posterior:			
Cell Cycle	Ccnd2	Malat1	
Golgi	Golga3		
Growth Factor	Mdk		
Imprinting	H19		
Inflammation, Heatshock	Malat1	Pdap1	
Metabolism	Apoc1	Atp5k	Ldha
Miscellaneous	Gm10076	AY036118	
Ribosomal	Gm42418	Rps29	
Cytoskeleton	Tmsb4x		

Up in E10.5 Anterior:		
Histone	Hist1h2ae	Hist1h2ap
Male Differentiation	Lefty2	Pycr2
Metabolism	Chchd10	Phkg1
Nuclear Import	Kpna2	
Ubiquitin	Mkrn1	Ube2s

Up in E10.5 Posterior:				
Chromatid Cohesion	Smc1b			
Cytoskeleton	Tmsb4x	Tuba1a		
Extracellular Matrix Breakdown	Mmp14			
Imprinting	Meg3	Id1		
Ribosomal	Gm42418	Rps12	Rps24	Rpsa

2.6 - Imprinted genes mark a subset of germ cell bioinformatically, and are located extragonadally in the tissue

Cell cycle regression unmasks other heterogeneity

Next we isolated PGCs and reclustered them without somatic cells. Reclustering with just the PGCs removed the strong differential expression signatures between PGCs vs soma and revealed subtle differences between germ cell subtypes. Analysis of germ cell only subsets showed that cells' positionality and embedding in the UMAP space generally followed trends of

phases of cell cycle, mainly G2M vs S (Figure 2.9). Common cell cycle markers like Mki67 (Gerdes et al., 1983; Miyachi et al., 1978) and pHH3 (Hendzel et al., 1997) were not present in the top genes of the cell cycle associated principal component. Rather, we found these genes making up a principle component associated with cell cycle: *Cenpf, Prc1, Tpx2, Cdc20, Mis18bp1, Arl6ip1, Hmmr, Ube2c, Cenpe, Aurka, Malat1, Cks2, Nusap1, Plk1, Cenpa, Rps20, Rplp1, Rps29, Incenp, Zfp42*.

Migratory PGCs must decide whether to migrate or focus on proliferation (Cantú et al., 2016), and as such, perhaps this suite of cell cycle genes is permissive in the duality of rapidly switching between migration and proliferation during the development of the germline. Since some PGCs are replicating at this time point, this principal component was not surprising. However we were concerned that the strong phases of cell cycle signature we saw (Figure 2.9A, G'-I') might be masking other heterogeneities.

To probe the masking of cell cycle on heterogeneity, we scored each cell and assigned it a cell cycle phase (Figure 2.9A, G'-I'), and found that clustering of each migratory time point followed cell cycle phase (Figure 2.9B, G''-I''). Thus, we hypothesize that biologic heterogeneities are hidden by the strong cell cycle signal. To mitigate this, we performed a regression against the differentially expressed genes related to the G1, G2M, and S phases of cell cycle. We preserved true differences between a cycling cell and a non-cycling cell, as we know that cell cycle can be regulated by position along the migratory route (Cantú et al., 2016). After performing the regression, cell phase identity mixed across the UMAP (Figure 2.9D), and cells no longer clustered along the phases of cell cycle (Figure 2.9E). For all studied timepoints, we found that initial clustering (Figure 2.9G``, H``, I``) followed boundaries of cell cycle phase (Figure 2.9G`, H`, I`). Fortunately, regression against the phase of the cell cycle resulted in

more PGC sub-clusters (Figure 2.9E, G````, H````, I````), and clusters that did not fall along cell cycle phase boundaries anymore (Figure 2.9E, G```, H```, I```).

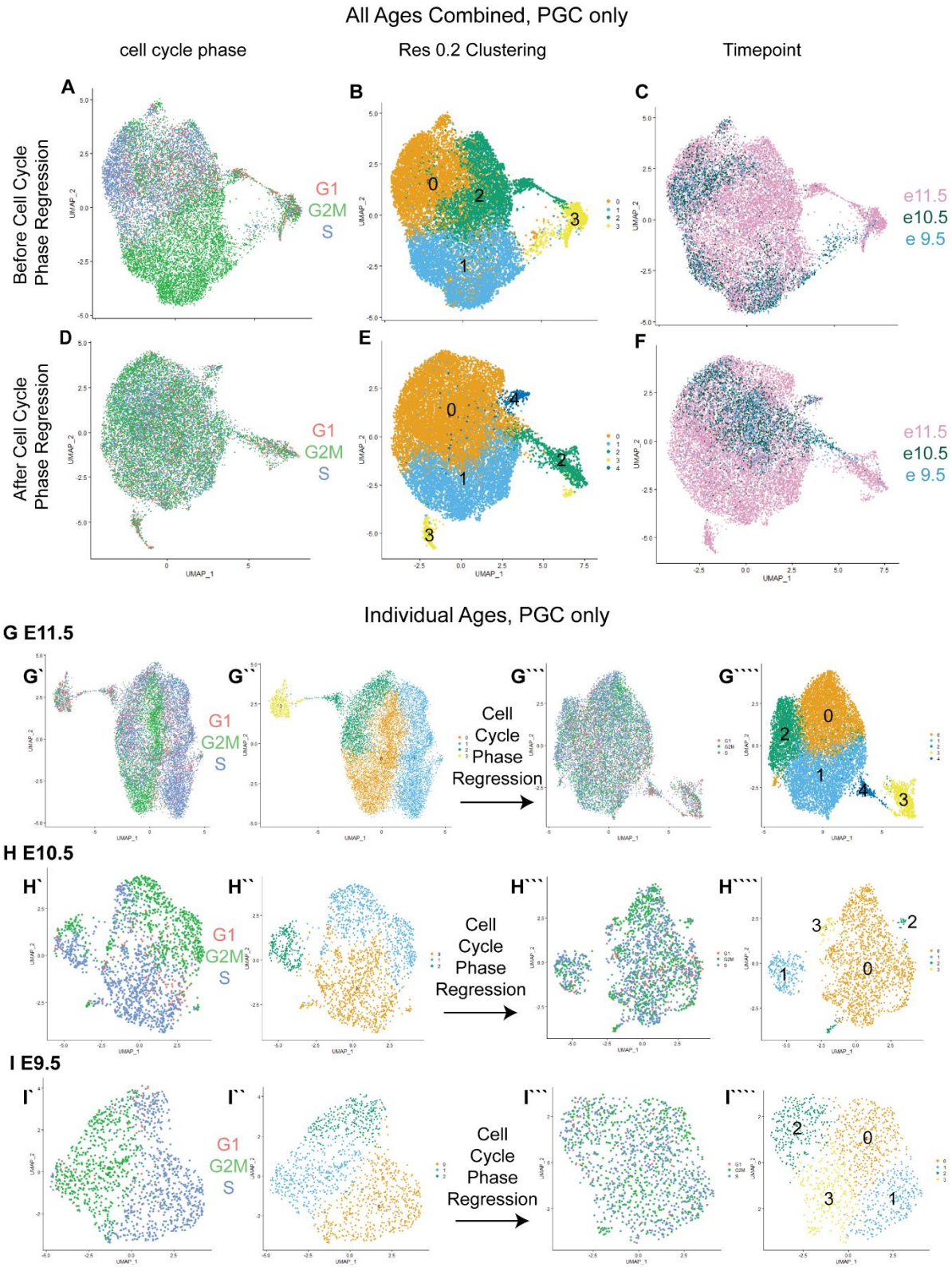


Figure 2.9 (Supplemental for 2.8) Cell cycle phase regression reveals additional mouse PGC heterogeneities

A) All mouse migratory timepoints combined, with the phases of cell cycle identified. PGC subclusters group according to G2M and S phase.
B) Subclustering patterns for all mouse timepoints combined follow cell cycle.
C) Cell age / timepoint identity shows interdigitation of PGCs at E9.5-E10.5 and E10.5-E11.5 .
D) After regression against the phases of cell cycle to remove that influence from the differentially expressed genes, PGCs no longer group according to G2M and S phase.
E) PGC subclusters are more discrete and no longer follows the same pattern as phases of cell cycle.
F) A comparison of cell age/time point at E9.5-E11.5 after regression.
G) Cell cycle regression for the E11.5 PGC library alone. Before integration, cell embedding and clustering followed the patterns of phases of cell cycle. After cell cycle phase regression, PGC subclusters were more distinct.
H) Cell cycle regression for the E10.5 PGC library alone. Before integration, cell embedding and clustering followed the patterns of phases of cell cycle. After cell cycle phase regression, clustering revealed more subtle subpopulations.
I) Cell cycle regression for the E9.5 PGC library alone increased separation of subclusters.

Identification of bioinformatic PGC outliers in scRNA seq data

When interrogating just E11.5 PGCs, unsupervised clustering revealed two spatially distinct populations of PGCs, Cluster 3 and 4 (Figure 2.10A). Clusters 3 and 4 had a differential gene expression profile unique enough to localize them away from the main body of PGCs. A list of markers for these cluster included several imprinted genes (Table 2). Maternally (*H19*, and *Meg3)* and paternally (*Igf2*) expressed genes marked Clusters 3 and 4 (Figure 2.10B), and were found with other markers for these cells in Table 2. This population of cells also expresses PGC maturation markers like Sox4 (Zhao et al., 2017) as well as Sox11, a marker classically associated with early embryonic development (Bhattaram et al., 2010). This sequencing library included both cells present in the gonadal ridge, and cells medial to the ridge and in the midline region. Thus, we hypothesized that the imprinting signature and differential gene expression profile present in Cluster 3 and 4 might reflect a gonadal vs extragonadal PGC location.

Location of imprinted gene expressing PGCs in the extragonadal spaces of the AGM

Since expression of imprinting genes is commonly found in somatic tissues, we were initially surprised by this result. However, later in reproductive development, germ cells are known to express *H19* and other imprinting genes (Lucifero et al., 2002). To validate the expression of these genes, we performed *in situ* HCR to localize *Igf2* in vivo. It strongly marked the somatic cells, as expected, so we quantified the expression of *Igf2* in Imaris. By selecting PGCs spots from our Oct4(ΔPE)-GFP positive AGM samples (Figure 2.10C-D), we were able to exclude somatic expression of *Igf2* and just consider expression of *Igf2* in the PGCs. We were able to split the GFP positive cells in extragonadal and gonadal populations (Figure 2.10D, Spots Panels), to then quantify the *Igf2* expression in either cell location. In samples from three

different litters, we found that *Igf2* marks the extragonadal cells (Figure 2.10D, right panel). When the top 10% of all PGC *Ifg2* was queried, 95.5% of that *Ifg2* expression came from extragonadal PGCs. The strong increase in *Igf2* expression in the extragonadal PGCs vs the gonadal PGCs is statistically significant (Figure 2.10F). Cluster 3 contains 5.4% of the total E11.5 PGCs population, and Cluster 4 contains 2.0% of the total E11.5 PGCs population.

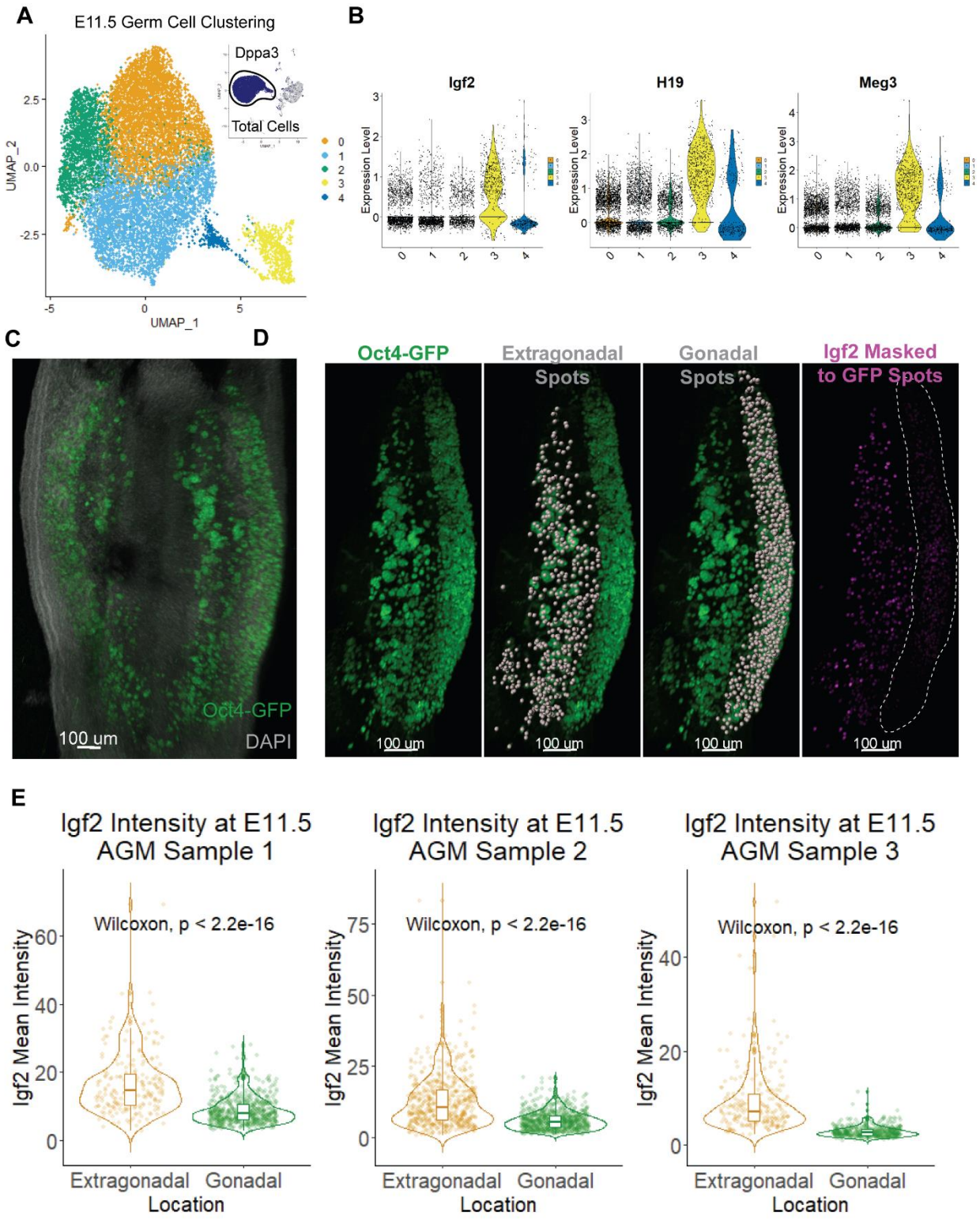


Figure 2.10 Elevated expression of imprinted genes marks extragonadal PGCs at E11.5

A) After clustering all E11.5 cells (inset) we isolated and reclustered E11.5 PGCs. Two spatially-

distinct clusters were observed (Cluster 3 and 4).
B) PGC subclusters 3 and 4, particularly cluster 3, defined by elevated expression of primarily maternally as well as paternally imprinted genes, with violin plots of *Igf2*, *H19*, and *Meg3* shown.
C) Whole-mount imaging of E11.5 aorta-gonad-mesonephros regions, where DAPI stain (grey) delineates tissue architecture, and Oct4-GFP marks PGCs.
D) Whole-mount i*n situ* hybridization for *Igf2* in E11.5 aorta-gonad-mesonephros from Oct4-GFP positive embryos. Since *Igf2* levels are high in somatic tissues, we analyzed transcript in PGCs only via the Imaris spot module to identify PGCs (green, Oct4-GFP and pink, *Igf2*). To quantify *Igf2* in PGCs by location, we separated “extragonadal” or “gonadal” spots based on their anatomical location using DAPI.
E) *Igf2* transcript levels were compared between gonadal and extragonadal PGCs in three separate aorta-gonad-mesonephross from three different age matched litters. In each aorta-gonad-mesonephros, *Igf2* was significantly increased in extragonadal PGCs.

Table 2.2 Top 30 Differentially Expressed Genes in E11.5 Cluster 3 and 4

Gene	P Value	Ave logFC	pct.1	pct.2	Adj P Value	Imprinted?	Expressed Allele
Tmsb4x	1.67E-186	1.25523371	0.819	0.499	3.34E-183	No	
H19	3.52E-165	1.228218952	0.734	0.307	7.04E-162	Yes	Maternal
Ptn	5.96E-251	1.202255789	0.744	0.212	1.19E-247	No	
Meg3	3.98E-139	1.160767459	0.706	0.314	7.96E-136	Yes	Maternal
Mdk	3.88E-247	1.151170114	0.861	0.53	7.77E-244	No	
Hmga2	6.78E-293	1.048985416	0.827	0.269	1.36E-289	No	
Cald1	3.62E-135	0.934550357	0.654	0.204	7.24E-132	No	
Ccnd2	3.51E-111	0.83765889	0.614	0.199	7.01E-108	No	
Malat1	1.03E-254	0.824052049	1	0.996	2.06E-251	No	
Actb	4.65E-121	0.76328151	0.999	0.999	9.29E-118	No	
Tshz2	7.91E-98	0.73401876	0.547	0.135	1.58E-94	No	
Gpc3	2.54E-87	0.669618666	0.538	0.156	5.08E-84	No	
Igf2	1.69E-65	0.631157137	0.518	0.115	3.37E-62	Yes	Paternal
Mest	7.00E-202	0.625559681	0.752	0.307	1.40E-198	No	
Meis2	4.39E-56	0.621901984	0.547	0.222	8.79E-53	No	
Tpm1	6.32E-58	0.582659942	0.768	0.636	1.26E-54	No	
Nnat	5.78E-77	0.580566521	0.526	0.154	1.16E-73	Yes	Paternal
Sparc	5.89E-78	0.568046144	0.851	0.741	1.18E-74	No	
Tsc22d1	3.15E-98	0.561615689	0.857	0.557	6.31E-95	No	
Igfbp5	9.32E-129	0.558317344	0.555	0.189	1.86E-125	No	
Peg3	1.04E-84	0.557844727	0.904	0.795	2.07E-81	Yes	Paternal
Col1a2	2.37E-121	0.552546414	0.55	0.154	4.75E-118	No	
Nr2f2	8.76E-56	0.547656728	0.452	0.113	1.75E-52	No	
Mfap2	1.34E-29	0.547552381	0.417	0.099	2.68E-26	No	
Sox11	6.42E-72	0.543836633	0.525	0.169	1.28E-68	No	
Cd63	1.91E-58	0.537220734	0.518	0.171	3.81E-55	No	
Serpinh1	4.00E-107	0.529125939	0.808	0.464	8.01E-104	No	
Dlk1	7.98E-13	0.527295085	0.391	0.151	1.60E-09	Yes	Paternal
Cdkn1c	6.25E-26	0.523518128	0.433	0.113	1.25E-22	No	
Sox4	1.34E-53	0.50402788	0.557	0.214	2.67E-50	No	

2.7 - Chapter Conclusion

Taken together, this work provides a comprehensive look at heterogeneities present in migrating PGCs of mouse and human. Studying the heterogeneities underlying migrating PGCs gives insights into the transcriptional makeup of the seed population of reproductive cells that colonize the gonad. The idea that some heterogeneities (eg, increased imprinting gene expression at E11.5) might make a cell less likely to reach the gonadal ridge has long reaching implications for the fate of the reproductive pool. In the case of imprinted genes which are typically silenced by methyl marks, the expression of imprinted genes calls into question the integrity of the genome methylation status in the extragonadal PGCs. If, as the expression of imprinted genes implies, extragonadal PGCs are lacking methyl marks, these cells may be at risk for a deleterious retrotransposition effect. Additionally, how a germ cell is able to interact with its somatic environment may be a fitness test; if a germ cell cannot stay on the guided path toward the gonad, what other signal cues might be missed during later developmental processes? Understanding migrating PGCs better will promote a more holistic understanding of germ cell fitness.

In this work, we have created the largest survey of migratory mouse PGCs thus far. By splitting our libraries into anterior leaders compared to posterior laggards, we could also probe for heterogeneity based on position in the migratory route. The many PGCs we queried also facilitated our receptor-ligand analysis, enabling robust identification of receptor-ligand pairs mediating the migratory process, as well as probing how the niche might contribute to a successful migrant. Finally, we bioinformatically uncovered a population of PGCs expressing imprinted genes. This exciting identification of extragonadal PGCs makes it possible to further follow these cells up for study in the future. We also are able to compare our findings to developmentally similar human PGCs. Together, these data provide a thorough account of the

similarities and differences present in migrating mouse and human PGCs, giving insights into the precursors to the reproductive lineages and the development of future generations.

Chapter 3 Discussion of discovered PGC heterogeneities

This work presents the deepest look into the individual transcriptional signatures of migrating PGCs in mouse and human while capturing the highest number of cells to date. Our sequencing libraries generated also contain positional information to further break down the heterogeneities present in migrating mouse PGCs while retaining their spatial context. Inclusion of somatic niche cells allowed for interrogation of receptor-ligand signaling involved in the regulation of migration. Finally, differential expression of genes at the bioinformatic level permitted identification of an extragonadal PGC population at the end of migration. How these extragonadal PGCs which express imprinting genes might factor into future gonadal development is unclear. We hypothesize that some of these extragonadal PGCs will remain ectopic as a result of their aberrant epigenetic reprogramming.

3.1 - Summary of mouse single cell RNA sequencing field after our contribution

The etiology of disease states for organs like the testis or ovary requires an understanding of the fundamental developmental biology of the organ; many organ-related diseases such as endometriosis and germ cell tumors co-opt these developmental programs during their progression. A key aspect of understanding the reproductive organs begins at their colonization after migration. Thus, more detailed description of how PGCs are moving through their normal migratory route is paramount to understanding how migration contributes to the greater cellular fitness of the reproductive system. Previous work has used movies and imaging to excellently characterize the phenotype of PGCs during their migration (Gomperts et al., 1994; Molyneaux et al., 2001). Now, using modern technologies to unbiasedly observe RNA transcripts from all

migratory PGCs, we provide a rich and promising companion data to those interested in studying the early PGCs. A recent study performed sequencing to characterize early germ cell fate and sex determination from matched time points included E9.0, E10.5, and E11.5 mouse cells (Mayère et al., 2021). Like us, this group found that early migratory timepoints did not differ strongly when all cells were integrated into a single object and their ages plotted. However, our work includes an order of magnitude more germ cells at each timepoint, ensuring that subtle phenotypes with small sample sizes are well represented. In addition, the design of our library includes the split of anterior migratory leaders compared to posterior migratory laggards at E9.5 and E10.5, allowing the retention of spatial and developmental contexts within our dataset.

3.2 - Discussion of heterogeneities found and their implication

Mouse compared to human

Comparisons of mouse PGC migratory libraries across each timepoint failed to find strong differential gene expression, suggesting an overall homogeneity within the PGC population during this process. The E9.5, E10.5, and E11.5 libraries integrated together well, implying that there are not large differential gene expression profiles distinguishing cells across migratory time. (Figure 2.9F). This is in contrast to the gene expression profiles found linked to similar developmental time points in the human fetal dataset (Figure 2.4B). The mouse as a model for human development is great, but differences between the two vertebrates are often found. For example, the length of time during which PGCs migrate in human compared to mouse development could contribute to both the heterogeneity of human migratory PGCs as well as the homogeneity of mouse migratory PGCs. Human PGCs are known to migrate over the course of weeks, from 4-14W of development (Mamsen et al., 2012). In contrast, the mouse PGC journey from the hindgut, through the dorsal mesentery, and into the gonadal ridge is a quick sprint lasting 2-2.5 days. The differences in time scale between human and mouse PGC

migration may permit greater RNA expression diversity and phenotypic variation. Additionally, the human PGC libraries contains fewer cells created from a well based library preparation sequenced very deeply, which is in contrast to our droplet-based libraries where we prioritized collection of many cells sequenced more shallowly. This technical difference in library preparation may also account for capturing rarer genes expressed across developmental time.

The migratory niche

While our analysis focused on the earliest developmental times present in the library, previous work (Overeem et al., 2021) also analyzed the Li et al library (Li et al., 2017) using CellphoneDB and characterized RTK, WNT, TGFβ/BMP, and NOTCH signaling pathways in germ cells during second trimester. When the early 4W-12W were reanalyzed, again the same pathways come up, indicating that RTK, WNT, TGFβ/BMP, and NOTCH have roles not just during the second trimester, but also during the earliest phases of PGC development and migration.

Differences in mouse and human migrating PGCs

Although human PGCs and mouse PGCs both undergo migration, the nature of and niche composition of that migration differs between the two species. Human germ cells have been shown previously to use nerve fibers as a migration substrate according to one report (Møllgård et al., 2010), while mouse germ cells showed no such co-localization with nerve fibers (Wolff et al., 2019). We saw some nervous system related receptor-ligand pairs upregulated in the human PGC receivers compared to the mature germ cell receivers. We do not see a nervous system category in the migrating mouse PGCs. For instance, the LXH9 positive somatic niche signaling to human migratory PGCs, nervous system receptor ligand pairs ADGRV1_GPHA2, GPR37_PSAP, MDK_LRP1, and NTRK1_AIMP1 were all significantly upregulated; enrichment for this LXH9 niche nervous system receptor-ligand signal was lost as PGCs developed into

mature germ cells. Understanding the functional significance of any of these receptor-ligand pairs and how they may impact PGC migration, survival, and development remains an open question. These data taken together underscore the need for a thorough understanding of both PGC intrinsic heterogeneity, as well as PGC interconnectedness within their migratory niches, including interaction with nervous and other somatic tissues.

Junctions at the end of migration

Considering both mouse and human migrating PGCs, the increased prevalence of cell junction related receptor-ligand pairs is perhaps indicative of PGCs migrating into and beginning to colonize the gonad, utilizing junctions to facilitate the end of their migration. Adherens junctions are required to anchor germline stem cells to their tissue niche in Drosophila (Song et al., 2002), and so the receptor-ligand pairs at E10.5 may be functioning similarly in vertebrates. Cadherins are known to have roles in coordinating phases of migration (Bendel-Stenzel et al., 2000), and indeed we find categories for junction related receptor-ligand interactions in all our timepoints, from both species.

Novel Eph/ephrin signaling found in mouse and human PGCs

We also see an increase in novel Eph/ephrin signaling receptor-ligands in both the human and mouse PGC datasets. To the best of our knowledge, the role of Eph/ephrin signaling in both human and mouse migrating PGCs has never been studied. The EFNA1-EPHA4 and EFNA4-EPHA4 receptor-ligand pairs, which are found in both mouse and human, may be linked to colonization of the gonad, as Eph/ephrin signaling pathways have established roles in coordination of adhesion and motility (Singh et al., 2012). Epha4 is known to be necessary for alignment of neuroblasts and organization of astrocytes in the developing rodent brain (Todd et al., 2017). Epha4 also regulates neural crest cell migration (Smith et al., 1997), another cell type

which undergoes a targeted migration. It is possible that Eph-Ephrin signaling modulates PGC migration in ways not yet known. It is exciting to find Eph-Ephrin signaling in migrating germ cells, and the importance of this novel finding is evident when considering how much is yet to be discovered about PGC migration. Heterogeneities in Eph-Ephrin signaling may set up a migration differential or colonization differential which impacts the seed population of germ cells which make it to the gonad. Future work to clarify what role Eph-Ephrin signaling has in relation to migrating germ cells will be of great value to the field.

A possible role for Bmp signaling in guiding migrating mouse PGCs

Although loss of Bmpr1a results in PGC migration along an incorrect route, exogenous BMP4 does not appear to be a PGC chemoattractant when added to gonadal ridge tissue slices (Dudley et al., 2007). A subsequent study (Dudley et al., 2010) proposed that BMPs do not directly act on PGCs, but rather influence the somatic cells along the migratory route, causing these somatic cells to express other cues which guide PGCs. Since we also see BMP7 ligand expressed in PGCs, it would be interesting to test whether BMP7 helps PGCs stay on the correct route, or if it plays a role as a PGC chemoattractant. Interestingly, we find BMP7 ligand expression in both mouse and human PGCs; whether BMP7 is one such chemoattractant or contributes to PGC migratory route guidance remains an open question. Whether Bmr1a or Avr2b receptors enable PGCs to home in on the migratory route is unknown. However, early embryo developmental requirements for BMP signaling impair straight-forward testing of this without inducing embryonic lethality.

Somatic contributions in the state change from E9.5 to E11.5 PGCs

Finally, we aimed to classify specific PGC-soma interactions involved in developmental state transition from E9.5 to E11.5, and early 4W-10W to 11-12W. NicheNet identified all potential

pairings present in the dataset regarding the developmental transition, but we knew that not all interactions are physically possible, based on where PGCs may be in migration. However, since Nichenet searched for receptor-ligand pairs involved in mediating the state change of our PGC specific differential expression list, the relevance of the receptor-ligand pairs returned is evidenced.

We found that Slit-Robo signaling is a shared receptor ligand pair associated with the developmental transition from an early migrant to a late migrant. We found these receptor ligand pairs in both mouse and human datasets; understanding how homologous/shared as well as heterogeneously expressed pathways between mouse and human in future work will help uncover the complex network of regulation controlling PGC migration and colonization of the gonadal ridges. Slit-Robo signaling has a known function in axon guidance (Tong et al., 2019), and may play a similar role in guiding migratory PGCs through the hindgut to the gonad. Another such receptor ligand pair mediating the developmental transition from E9.5 to E11.5 is Cldn11. This gene, which functions as its own receptor and ligand, is known to cause infertility later in male development (Mazaud-Guittot et al., 2011); it is unknown what role it might have in the maturation of migrating PGCs prior to sex specification. Since loss of Cldn11 has been associated with cancer invasiveness (Agarwal et al., 2009; Li et al., 2018), Cldn11 may be acting as a break on migration. Future work characterizing these receptor-ligand pairs and how they regulate PGC will help clarify the role niche interaction has on this migratory behavior.

The categories we have discovered in our receptor-ligand analyses confirm longstanding field knowledge, as well as contribute new avenues for exploration. Future directions following up any of these results will increase the niche-specific context of migrating PGCs. Additionally, it would be interesting to see if any genes which are expressed in a time specific manner affect

the ability of human PGCs to migrate. However, it is possible the banding patterns of gene expression serve some purpose beyond migration. Further study to characterize the time restricted markers should bear interesting results about their function.

3.3 - The impact of imprinting and future directions

The ability of germ cells to switch from active, directed migration to colonization has been a fascinating area of study, inspiring beautiful time lapse imaging and microscopy data to characterize germ cells as they complete their arduous journey to the gonad. PGCs during this time period balance many developmental programs simultaneously, including proliferation (Cantú et al., 2016), migration, and the start of epigenetic reprogramming (Ewen and Koopman, 2010). Epigenetic reprogramming of the cells which carry the potential for future generations is a potentially dangerous but necessary component of PGC further specification to mature spermatogonia or oogonia (De Felici, 2011). While the changing of methyl marks and histone modifications must occur to promote further sex specification of the germ cells, the opportunity for deleterious expression of retrotransposons suggests that there may be some quality control mechanism to ensure that PGCs complete their epigenetic reprogramming correctly and with minimal genomic damage.

Migration has been proposed to function, in part, as a filter for PCG fitness (Cantú and Laird, 2017), and our data demonstrating highest imprinting gene expression in extragonadal PGCs supports this theory. We also observe expression of imprinted genes in early migratory timepoints. Specifically, expression of H19 marked the E9.5 Posterior cells, and expression of Meg3 and Id1 marked the E10.5 Posterior cells (Table 1). The earliest waves of passive imprint erasure have been found to occur early in PGC migration (Kagiwada et al., 2013). As each PGC migrates into a more proliferation permissive niche (Cantú et al., 2016), it's is possible that the

replication-based removal of genome imprints described by Kagiwada and collogues is responsible for the increased expression of imprinted genes we saw in our PGC clusters. The expression of imprinted genes always occurred in the migratory laggards (the posterior libraries at E9.5 and E10.5 [Table 1], the extragonadal PGCs lagging in the medial/midline region of the aorta-gonad-mesonephros [Figure 2.10]). Thus, we hypothesize that imprint erasure which occurs too early leads to a decreased ability of PGCs to migrate to the gonadal ridge. A future direction for this work would be to clonally label extragonadal PGCs, either through a timed expression of an *Igf2* based inducible transgene, and measure whether cells with high imprinting gene expression can complete migration. If these cells remain in the extragonadal space and are eventually selected for death, something about their imprinting signature, or underlying methylation status, could have caused them to miss migratory cues.

We observed no imprinting signature observed in the reanalyzed human migratory PGCs (Figure 2.6). However, the comparison is complicated by the rare nature of these cells compared to the entire population. In the mouse, the number of germ cells which made up the imprinting cluster tallied at a modest 5.4% of all E11.5 PGCs surveyed. It is possible that imprinting gene expressing cells are a small enough population that when Li and collogues manually collected their ~2000 PGCs for sequencing (Li et al., 2017), they did not acquire a PGC with this aberrant profile. The advantage of droplet based sequencing approaches is the ability to capture many cells, and increase the chances of observing a rare cell population. In a head to head comparison in retinal bipolar neurons, more shallowly sequenced captured cells sequencing was better for classifying more cell types when compared to fewer cells deeply sequenced (Shekhar et al., 2016). Therefore, the possibility remains that droplet based sequencing of human PGCs in future may reveal the presence of imprinted loci.

How epigenetic reprogramming progresses differentially in gonadal PGCs compared to extragonadal PGCs remains an intriguing open question. Are the aberrant imprinting profile seen in the lagging migrants is a factor of their location? Or is the profile some key intrinsic heterogeneity that is magnified through the migratory process? A clonal label system which tracks their progress through epigenetic reprogramming, navigation through developmental checkpoints, and contribution to future niche colonization would provide important insight into differentiating these mechanisms.

3.4 Conclusion

In summary, this work provides a transcriptional survey of migrating mouse germ cells and migrating human germ cells. We interrogated migrating PGCs in the context of their somatic niche, and identified known and novel receptor ligand pairs which may be involved in regulation of migration or colonization of the gonadal ridge. We identified transcriptional differences between PGCs in different phases along the migratory route connecting their spatial heterogeneity to transcriptional heterogeneity. In the mouse, we assayed the difference between migratory leaders and laggards E9.5 and E10.5, and in the human characterized expression of time-restricted genes in the PGCs, finding that bands of genes expressed in a time specific manner. Finally, we distinguished gonadal and extragonadal germ cells bioinformatically at the end of migration, and validated these results using HCR to locate the *Igf2* expressing PGCs in the AGM. Understanding how migration may rank the fitness level between gonadal and extragonadal PGCs will promote better understanding of germ cell quality at one of the earliest developmental checkpoints.

Chapter 4 Methods

Mouse Crosses

CD1 females were crossed to males homozygous for Oct4-ΔPE-eGFP (MGI:3057158). Plug date was designated as 0.5 days post conception. Pregnant mice were euthanized and age matched embryos were dissected in 0.4% bovine serum albumin (BSA) in 1x phosphate-buffered saline (PBS) at E9.5, E10.5, and E11.5. Embryos were further staged by somite number, to tighten the range of developmental time for embryos used in library creation.

E9.5 dissociation

At E9.5, 6 litters were dissected and embryos with total somite counts of 20-25 were kept. To further split our E9.5 embryos into an Anterior Bin and Posterior Bin (containing migratory leaders and laggards, respectively), embryos were bisected right below where the PGCs have begun to push into the mesentery tissue, around somite 15. (Figure 2.1A, dashed line). Embryos were cut above the forelimb, and the posterior halves removed. Two to three heads were saved for a GFP negative control. Each litter was pooled into a tube and added 200uL of pre-warmed 0.25% Trypsin/ETDA, and incubated the tubes for 10 minutes in a 37°C water bath. After 10 minutes, a razor blade was used to cut a pipette tip to be slightly wider/gentler for sample mixing. For this and other timepoints, samples were triterated gently (the objective being to mix the tube, not use the force of pipetting to shear the embryos; sheared embryos resulted in poor digest/FACS yields). If after 10 minutes the digest still looked clumpy, then the samples were incubated another 10 minutes in the 200uL 0.25% Trypsin/ETDA. Next, 50uL of 1mg/ml DNAse was added, samples triterated gently, and incubated for 5 minutes at 37°C. After 5 minutes, samples were checked for consistency. If they were thin and homogenous, digests were quenched with a 1:1 volume of FBS. If samples were still a chunky or gluey, 50uL more

DNAse was added and samples were incubated for 3 more minutes at 37C, after which samples were quenched with a 1:1 volume of FBS and placed on ice.

E10.5 dissociation

At E10.5, 30 GFP positive embryos with total somite counts of 34-38 were kept. To separate E10.5 embryos into Anterior and Posterior Bins (containing migratory leaders and laggards, respectively), embryos were bisected caudal to the point where PGCs have begun to bifurcate toward left or right gonadal ridges, between somite 19 and 20 (Figure 2.1A, dashed line). Dissected embryos were cut in half just below the heart, and the posterior half of the embryo was dissociated to prepare for FACS germ cell concentration (GFP-negative heads were used as GFP-negative control tube). Embryo posterior halves were split into groups of 10, and placed in a 37C water bath with 400uL of pre-warmed 0.25% Trypsin/ETDA. An additional 200uL of pre-warmed 0.25% Trypsin/ETDA was added if the solution appeared too thick. Embryos were digested at 37°C for 15 more minutes, and digest progress aided by flicking and assessing solution viscosity every 5 minutes. 200uL of 1mg/ml DNAse was added, samples triterated, and then incubated for 5mins in 37C water bath. After a 5 minute digest, samples were assessed. If there are still clumps, an additional 5mins or additional 200uL of 1mg/ml DNAse can be added. Once the solution is homogenous and water-like, quench with a volume matched amount of FBS and place on ice.

E11.5 aorta-gonad-mesonephros dissociation

Aorta-gonad-mesonephros (AGM) regions of age matched embryos (45-49 somite range) were microdissected out from posterior embryo. AGMs were grouped in sets of 6, and placed in a 37°C water bath. GFP-negative heads were also processed as a control. 500uL of 0.25%

Trypsin/EDTA was added to embryos. Tubes were incubated for 20 minutes, triterating with a cut pipet tip at every 5 minute interval to gently mix the solution. After 20 minutes, 100uL of 1mg/ml DNase was added to each tube, and tubes were mixed by flicking. After 5 minutes of incubation with DNAse added, samples were assessed for homogeneity. Once the solution was waterlike and homogenous, digests were quenched with 1:1 volume FBS and placed on ice. For the first biological replicate, 23 embryos were dissociated. For the second replicate, 12 embryos were dissociated.

Cell Sorting

Once on ice, each sample was filtered through a 35nm strainer. Tip and tube and filter were rinsed with a small amount of 0.1% BSA to wash all possible GFP positive PGCs into the cell suspension. Sytox Blue live-dead indicator was added at the appropriate 1:1000 concentration to each digest. Samples were sorted on either a BD FACS Aria2, Aria3, or Aria Fusion, the following a gating pattern: 1) side scatter area vs forward scatter area to select for cells; 2) side scatter width vs side scatter height to select for cell singlets; 3) forward scatter width vs forward scatter height to select again for cell singlets; 4) Sytox Blue vs forward scatter area to select for Sytox Blue negative live cells; and finally 5) GFP vs forward scatter area. The range of GFP negative vs. positive gates was calibrated each experiment based on the GFP-negative digest sample, aiming to put GFP-negative cells in the 103 range and GFP positive cells in the 104 to 105 range. GFP positive cells were collected in 500uL of 0.2% BSA, and GFP-negative somatic cells were also collected in 500uL of 0.2% BSA. For each collection of GPF positive sorted cells, somatic cells from the respective position/niche were spiked back in the sample at a 1:1 ratio. Cells were concentrated into a 300-1500 cell/uL range, resuspending in 0.1% BSA.

Library Creation

Sorted GFP positive germ cells were pooled with an equal amount of GFP-negative somatic cells from each timepoint and anatomical location. Counted cells were loaded onto a 10X v2 chip to create a single cell emulsion. For E9.5 Anterior, we loaded 3400 cells, and for E9.5 Posterior, we loaded 2800 cells. For E10.5 Anterior, we loaded two technical replicates totaling 9300 cells. For E10.5 Posterior, we loaded two technical replicates totaling 11200 cells. For the two E11.5 biological replicates, we loaded 15,000 and 10,000 cells, respectively. Library creation was done in house following 10X ChromiumTM Single Cell 3' Reagent Kits v2 User Guide Rev A. Briefly, Single-Cell 3′ Reagent Version 2 Kit (10X Genomics) was used with fluidics to create individual cells in gel bead-in-emulsions (GEMS). Samples were prepared directly according to the protocol, with the following steps tailored to maximize library yield: at the cDNA amplification step used 14 cycles, and at the sample index PCR step used 14 cycles. Samples were tested for quality on an Agilent Bioanalyzer High Sensitivity chip before sequencing. E9.5 and E11.5 libraries were sequenced on HiSeq 4000 (Illumina) with paired-end sequencing parameters: Read1, 98 cycles; Index1, 14 cycles; Index2, 8 cycles; and Read2, 10 cycles. E10.5 libraries were sequenced on Novaseq (Illumina) with paired-end sequencing parameters: Read1, 150 cycles; i7 Index, 8 cycles; i5 index, 0 cycles; and Read2, 150 cycles.

scRNA Analysis

Seurat and SoupX

Mouse

To analyze the mouse datasets, count data was thresholded and analyzed in the package Seurat (v3.2.3) (Stuart et al., 2019). To stringently remove empty GEMS and keep only real cells, cells with fewer than 200 genes (E9.5, E11.5) or 500 genes (E10.5) were excluded. Cells

with greater than 4000 genes were excluded at E11.5. Cells with 5% or greater mitochondrial counts were also excluded. Libraries were normalized via NormalizeData, and 2000 top variable genes found with FindVariableFeatures. Libraries were integrated according to the Seurat Integration tutorial ("Introduction to scRNA-seq integration" Compiled: August 30, 2021) with the only change being 1:20 dims used when a dimension variable was called. RunPCA calculated principal components, and significant PCs were determined with ElbowPlot and DimHeatmap visualizations. RunUMAP was used to compute dimensional reduction on dims 1:20 and visualize the reduction. To acquire differentially expressed genes, FindMarkers was used on clusters or cell groupings (eg, Anterior vs Posterior) of interest. When aberrant hemoglobin gene expression in non-red blood cells was discovered, data was processed through SoupX (version 1.5.2) (Young and Behjati, 2020) to remove ambient 'soup' RNA contamination, and then rerun the above filtering/analysis. Soup parameters were either set by the function autoEstCont to automatically estimate ambient RNA contamination, or set by slowly increasing the threshold of filtering until Hbb* genes ceased to inappropriately be found as markers for non-red blood cell genes, as follows: E9.5 Anterior, 12%; E9.5 Posterior, autoEstCont; E10.5 Anterior, 19%, E10.5 Posterior, autoEstCont. To generate the PGC only analysis, cell names from clusters strongly expressing PGC markers (*Dppa3*, *Oct4*, *Dazl*, *Ifitm3*) were used to subset the SoupX cleaned count data, and PGC-only objects created per above. To compare the E9.5 anterior and posterior populations, dittoSeq was used to generate the heat maps.(Bunis et al., 2020)

Human

To reanalyze the human datasets, 4-12W count files were downloaded from GSE86146 (Li et al., 2017). Individual timepoints contained too few cells for Seurat 3.2.3 to integrate effectively, therefore, male and female matrices were created and integrated as described for mouse datasets.

Somatic Labels

To characterize a population, we first ran the Seurat FindMarkers function and output both UMAP plots of the top reported markers as well as Violin plots. Since hundreds of markers are often returned, we used two bioinformatic approaches to narrow candidate genes down (Figure 2.2C). First, we took highly differentially expressed and statistically significant markers for a population and ran GO analysis on the gene lists, using the online tool Panther. We were then able to group similar GO categories to determine what biologic processes are grouped with each cluster. Next, we used the Genewalk (Ietswaart et al., 2021) tool to aid in identifying genes which might be involved in relevant biological processes, vs "moonlighting" genes which encode proteins that function in more than one physiologically relevant process. Finally, we took genes which were strongly expressed in each cluster and queried them against existing mouse *in situ* databases, such as the one from MGI and emouse atlas. If we were able to find a trend amongst genes queried, we then made a cell type classification. dittoSeq (Bunis et al., 2020) was used to create the dot plot.

CellphoneDB

CellphoneDB (version 2.1.7) (Efremova et al., 2020) was run using somatic cluster labels as metadata, and counts data with default settings. Since CellphoneDB receptor-ligand databases were designed on human proteins, we first converted our mouse datasets to human orthologs using biomaRt (Durinck et al., 2009). However, this kind of analysis may miss mouse-specific receptor-ligand pairs. To generate dot plots, the file containing mean values for all the interacting partners was thresholded to p > 1, receptor_ligand pairs organized according to function, and relevant somatic populations signaling to germ cell populations were graphed using the cellphonedb dot_plot function.

NicheNetR

NichenetR (version 1.0.0) was run following the steps in the tutorial “Perform NicheNet analysis starting from a Seurat object Robin Browaeys 2019-11-08” or “Perform NicheNet analysis starting from a Seurat object: step-by-step analysis Robin Browaeys 2019-11-12”. In the human dataset analysis, we set our query up using a differentially expressed gene list comparing 4-10W primordial germ cells to 11-12W primordial germ cells, with the germ cells being the receiver and the somatic cells being the sender. In mouse, we first converted the human data model gene names to their mouse orthologs using biomaRt. While this kind of analysis may miss mouse-specific receptor-ligand pairs, the human dataset is fuller/more curated; thus, we took this approach. We set up our mouse query using a differentially expressed gene list comparing E11.5 PGCs to E9.5 PGCs, with the germ cells being the receiver and the somatic cells being the sender. We restricted the ligands we were interested in to those pairs which come from a validated curated ligand-receptor database, like KEGG and guide2Pharmacology (Browaeys et al., 2020).

HCR

Samples were generated per Mouse Crosses. *In situ* hybridization chain reaction (HCR) protocol was done as previously published (Choi et al., 2018) and following modifications as previously published (M. J. Anderson et al., 2020). Briefly, a custom HCR Probe Set (V3.0) for *Igf2* exons 2-5 was designed by Molecular Instruments, Inc. Following appropriate wash and preparatory steps as previously described, E11.5 AGMs were incubated in 2 pmol probe solution overnight at 37°C. After preparing samples for amplification as previously described, E11.5 AGMs were incubated in 60 pmol of each hairpin per 500uL amplification buffer overnight

at room temperature in the dark. E11.5 AGMs were then incubated with DAPI solution (0.5 µg/mL DAPI in 5x SSC with 0.1% TritonX-100, 1% Tween20) overnight at 4°C. Samples were cleared with Scale CUBIC Reagent 1 overnight at 4°C which preserved the endogenous Oct4-ΔPE-eGFP expression. To image a Fluotar VISIR 25x/0.95 water objective was used on a Leica TCS SP8 converted confocal microscope. Imaging parameters included a 0.75x optic zoom, 1024 x 1024 pixel resolution, and 2um distance between z stacks.

Image analysis for whole-mount stained ovaries

Using Imaris v9.3.0 (Bitplane), confocal z stacks of AGMs were analyzed. Surpass mode was used to import files. Spot detection module (cell diameter value 10um) was used on endogenous Oct4-ΔPE-eGFP expression to identify all fluorescently labeled germ cells in the AGM. Spots were thresholded to ensure only true Oct4-ΔPE-eGFP expression was labeled with a spot. Based on tissue morphology as shown by DAPI, GFP positive germ cell spots were classified as “gonadal” or “extragonadal.” Spots radius was set to 1.0, and detailed statistics for *Igf2* intensity were exported and analyzed in R using ggplot2 and ggpubr (Kassambara, 2020) to calculate statistics.

Chapter 5 Appendices

Appendix A: Clonal heterogeneity in migrating mouse primordial germ cells

Introduction: Migration and Clonal Heterogeneity

PGCs migrate individually and reach the gonad at different times. It is unknown how the order of migration into the gonadal ridge impacts the capacity of PGCs to colonize the gonadal ridge. The initial tissues through which the PGCs pass (hindgut and mesentery) promote migration over proliferation, and it is not until the PGCs reach the gonad that the local tissue niche promote more expansive proliferation (Cantú et al., 2016). Migrating PGC leaders arrive to the gonad as early as E10.5, and spend more time in the most proliferative-permissive niche, bolstering their colonization numbers. In contrast, laggard cells arrive at a gonadal ridge partially colonized by leaders. Laggard cells begin ramping up their expansion later, and thus their clonal progeny may be less represented in the gonad. It is un-clear whether leader PGCs already present in the gonadal ridge inhibit colonization by laggard PGCs. Thus, the reproductive system appears to be structured to promote dominance of clones which arrive early to the gonad. We are also curious whether migratory capability is a shared property among mother and daughter cells. To investigate heritability of heterogeneities present in migrating PGCs, we devised a genetic lineage label system to induce fluorescent clonal label in migrating germ cells. Investigating mother/daughter linages will reveal if shared traits affect PGC migratory capacity or ability to colonize the gonad.

Results: Clones migrate together

Using the inducible Cre mouse Pou5f1-creER (MGI:5049897) crossed with a Rainbow linage label mouse (Red-Horse et al., 2010), E8.5 migrating PGCs were sparsely labeled using tamoxifen before directed migration begins (Figure 5.1A). Since tamoxifen takes 24hrs to be metabolized and actively excise floxxed loci, clonal labeling was expected to initiate at E9.5, just as directed migration begins. We expect that clones of PGCs, based on the length of time it takes them to proliferate (13-16 hours), will number 8 daughter cells (Figure 5.1B). To test this, we injected tamoxifen at a very low dose, a 0.0125mg/g mouse, aiming to induce labeling in just one cell. Green tissue indicates a cell where recombination has not occurred, while a red, cyan, or yellow label indicates a clone/linage of that color. We found that when AGMs had just one cell labeled, the clonal progeny numbered 9 and 7 (Figure 5.1C). We increased our dose of tamoxifen to 0.01875mg/g mouse anticipating labeling of 5-25 cells to assess how cells of the same clonal linage migrated. We imaged 4 clonally label gonads (Figure 5.1D, sample image) at E11.5. This low density labeling allows red, yellow, and cyan clones to be discerned clearly from one another, although it was difficult to discern clones of the same color from each other when clone cell count numbered over ~8. A presumptive grouping 3-4 cyan clones number ~45 cells (dashed outline, Figure 5.1D) is one such instance. Across 4 AGM samples, cyan clones were made up of 10 +/- 7.4 cells, red clones were made up of 8.4 +/- 7.8 cells, and yellow clones were made up of 7.1 +/- 6.4 cells (Figure 5.1E). No differences in cell count per clonal color were significant. Next we defined the boundaries of the AGM using tissue morphology, and quantified whether clone size in the extragonadal mesentery differed from clone size in the gonadal ridge. We found a

significant increase in size for clone cell count in the gonadal ridge, although this is likely artificially inflated due to our inability to distinguish two-three clones of the same color next to one another in the tissue (Figure 5.1F). Finally, to see whether the left or right half of the gonad ridge ended up with more clones, we quantified number of cells per clone for each gonadal ridge half. There was no statistical difference between clone size in the left or right gonadal ridge.

Discussion and Future Directions

To investigate heritability of heterogeneities present in migrating PGCs, we devised a genetic lineage label system to induce fluorescent clonal lineage label in migrating germ cells. We investigated mother/daughter cells to see if migratory capability was shared within a clonal lineage. After precise, one clone labeling allowed us to deduce that clone cell number averaged 8, we increased our labeling frequency to test whether cells of the same clone seemed to cluster together upon reaching the migratory niche.

The idea that cells labeled just at the start of directed migration at E9.5 localize to the same general place in the gonadal ridge is quite striking. If migration were not a clonally shared property, we would expect to see an extremely mixed population of labeled cells sprinkled all throughout the gonadal ridges. Instead, where clones could be isolated away from other clones of the same color, clones clearly grouped together at the end of migration. Since PGCs migrate as individuals through the tissue, it is quite surprising that cells of the same clone end up so close to one another in the gonadal ridge.

Unfortunately, due to the randomness of fluorophore selection when the inducible Cre excises, presumable different clones of the same color were often found next to one another, making estimate of clone cell count very difficult. The replication rate of PGCs is well characterized during migration (Cantú et al., 2016), so it is extremely unlikely that PGCs could truly replicate to create a ~45 cell clone in the amount of time that tamoxifen is active. In the future, tamoxifen could be titrated to a dose in between 0.0125mg/g and to 0.01875mg/g mouse to lessen the chance that different clones of the same color run into one another. Additionally, as technologies evolve, lineage labeling through CRISPR barcodes (Bowling et al., 2020) will enable labeling of the entire germ cell population, from specification through migration, offering a fuller and completer picture of the clonality of migration. Understanding how clones migrate will give insight into whether being a migratory leader bestows a competitive colonization advantage to PGCs.

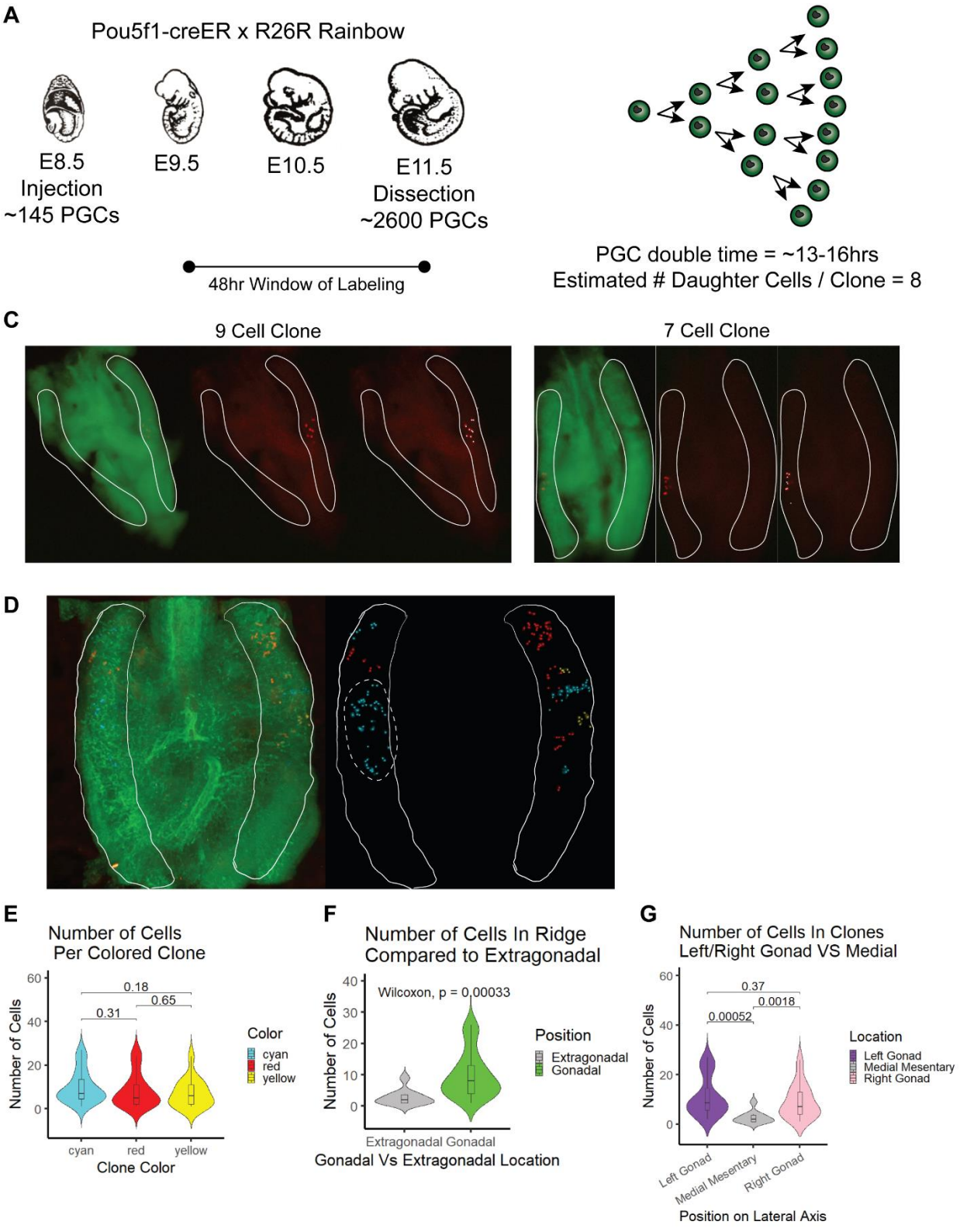


Figure 5.1 Clonally related PGCs localize together in clonal groups in the gonadal ridge

A) Cross scheme to generate clonal label aorta-gonad-mesonephros samples. Pou5f1-creER was crossed to R26R Rainbow. Samples were injected with tamoxifen on E8.5, and labeling induced 24hrs later at E9.5. Samples were dissected at E11.5 as PGCs were completing their migration.
B) Estimation of PGC replication cycle during the 48hr window of labeling.
C) Extremely sparse clonal labeling induced at E8.5 to gauge estimated clone size. Samples with just one clone induced showed 9 cells and 7 cells. Gonadal Ridges are outlined in white. Samples not to scale.
D) Example E11.5 aorta-gonad-mesonephros from Pou5f1-creER x R26R Rainbow cross; clones were induced at induced at E8.5, and visualized at E11.5. Imaris spots identify clones, and masks created to isolate clones. Clones are shown in red, cyan, and yellow. Dashed outline indicates presumptive 3-4 cyan clones intercalated. Samples not to scale.
E) Number of cells per clone, split according to color. Colors correspond with endogenous label.
F) Number of cells per clone, split according to whether the clone was in the gonadal ridge or extragonadal medial space.
G) Number of cells per clone, split according to whether the clone was in the left or right gonadal ridge or extragonadal medial space.

Appendix B: Surveying mitochondrial heteroplasmy in migrating mouse primordial germ cells

Done in collaboration with Aidan S. Marshall and Dr. Nick S. Jones.

Introduction to mitochondrial selection in germ cells

Diversity among PGCs indicates that there may be intrinsic differences within the PGC population resulting in a spectrum of fitness. Thus, different checkpoints throughout development to assess fitness of PGCs would promote survival of the highest quality PGCs, resulting in a population of high fitness to remain within the organism. It is currently unclear what checkpoints the developing organism uses to select cells of best reproductive potential. Since migration is one of the first challenges a PGC faces after specification at E7.0, migration emerges as a potential selector against cells with low fitness.

One candidate for distinguishing differential PGC fitness is mitochondrial heteroplasmy. Mutations in mitochondrial (mt) DNA arise in oocyte maturation due to the error rate in mitochondrial polymerase (A. P. Anderson et al., 2020; Brown et al., 1979). The replication of mtDNA and fission of mitochondria ceases during early mouse embryogenesis, so each cell division produces a smaller pool of mtDNA, limiting the diversity of mitochondrial genomes inherited from the oocyte (St John et al., 2010). PGCs are specified during this mitochondrial bottleneck, so founders inherit different mtDNA cohorts(Wai et al., 2008) (Mishra and Chan, 2014; Stewart and Chinnery, 2015). The mitochondrial bottleneck limits mtDNA diversity in developing PGCs, thus potentially affecting PGC fitness. We hypothesized that deleterious inherited mutations within PGCs may be revealed when PGCs are unable to mount the oxidative phosphorylation or bioenergetic response required for migration, resulting in less fit PGCs lagging in migration. Previous work has characterized the

extent of mitochondrial heteroplasmy in early human embryos (Floros et al., 2018). We sought to characterize the extent of heteroplasmies present in our mouse migratory model, particularly the difference between leading and lagging migrants.

Results: Mouse PGC heteroplasmy increases across time

Recent advances in single cell RNA sequencing technologies have enabled investigation of mitochondrial heteroplasmies at the single cell level (Marshall and Jones, 2021). To ascertain whether our E9.5, E10.5, and E11.5 single cell RNA sequencing libraries could be used to assay heteroplasmy in migrating mouse PGCs, we first tested depth of coverage along the mitochondrial chromosome. Only our E9.5 and E10.5 libraries were sequenced deeply enough for further analysis, and we report their depth of coverage at base log_{10} (Figure 5.2A). 10X libraries select for polyadenylated transcripts; thus deeper coverage corresponds with regions of the genome associated with transcribed mRNA. In the regions which met our coverage thresholds, we found that 59% (57-61% using a 95% confidence interval) of all SNPs were transitions such as A <-> G or C <-> T mutations (Figure 5.2B). This proportion of heteroplasmies is consistent with copying errors due to the fidelity of the mitochondrial polymerase (A. P. Anderson et al., 2020; Brown et al., 1979).

Next, we assessed heteroplasmy across developmental time. We found low levels of heteroplasmy in mice PGCs (Figure 5.2C), with heteroplasmies above 20% rarely found at E9.5, but greater heteroplasmy observed at E10.5. Since our libraries contained positional information about migratory leaders and laggards, we also tested whether migration might be acting as a quality control metric to select against cells with higher heteroplasmy. To probe whether anterior, migratory leaders might have fewer deleterious heteroplasmies, we assayed synonymous and non-synonymous mutations. We found that there was no difference in non-synonymous mutation

rate between anterior migratory leaders and posterior migratory laggards, at either timepoint (Figure 5.2D). Opposite our hypothesis, when E9.5 was compared to E10.5, the synonymous/non-synonymous mutation ratio increased over time.

Finally, to see if cells with multiple mutation were associated with PGC subclusters identified in the scRNA seq analysis (Figure 5.2, E',F'), we acquired the names of cells with more than one heteroplasmy and plotted them in our UMAP dimensional reduction. Cells with more than one heteroplasmy were not found preferentially in one cluster over another (Figure 5.2E''-F''). Additionally, there were not more cells in the lagging, posterior compartment at either timepoint, indicating that if heteroplasmies are selected against, it is not through a migration based filter.

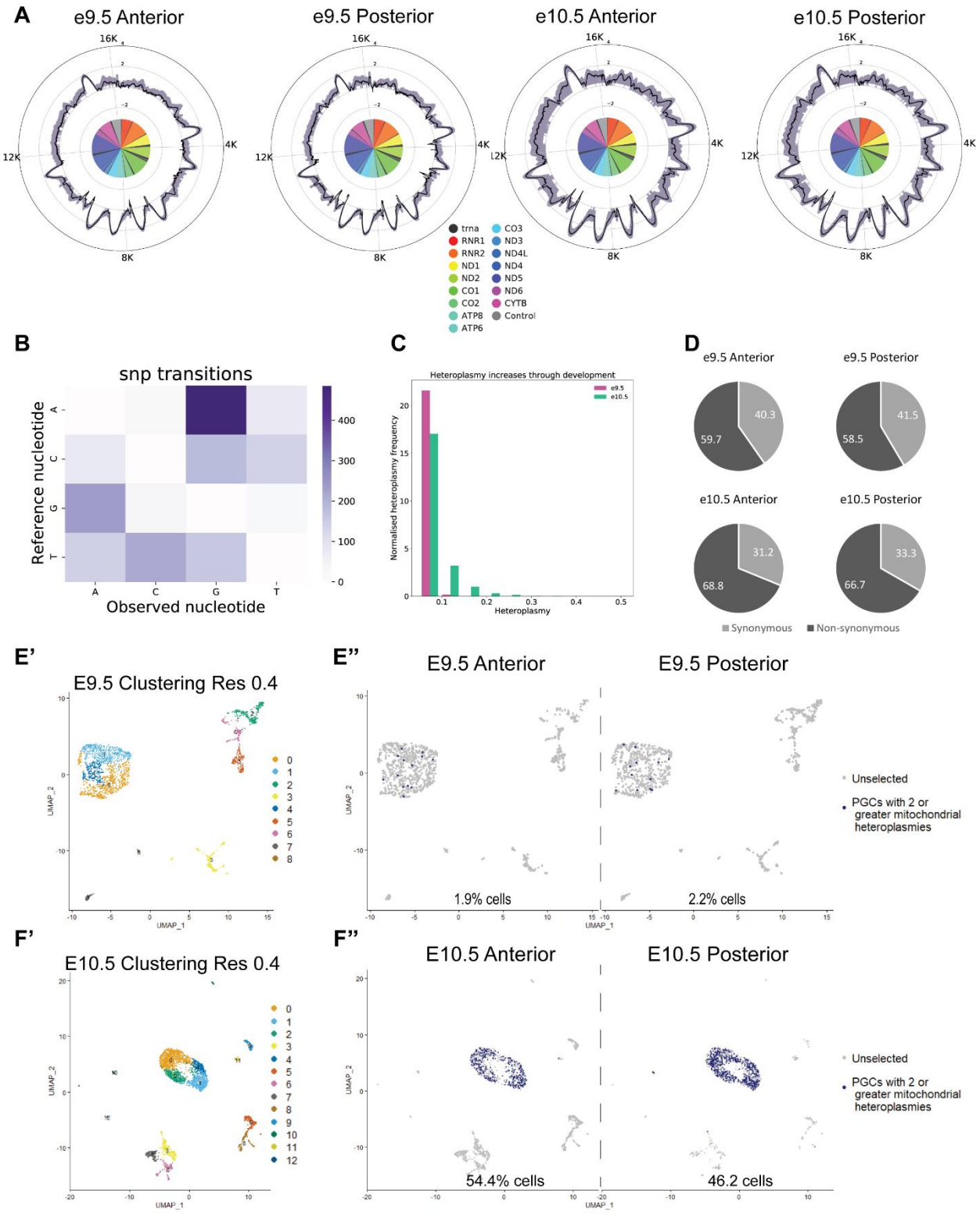


Figure 5.2 Assessing mitochondrial heterogeneity in E9.5 and E10.5 migrating PGCs

A) Depth and breadth of sequencing coverage across the mitochondrial genome in E9.5 Anterior, E9.5 Posterior, E10.5 Anterior, and E10.5 Posterior libraries. Regions of the

mitochondrial chromosome containing a mt-gene are indicated by color. Mean depth of coverage per cell is indicated by the black line, which is bounded by the grey bar indicating 90% depth of coverage interval.
B) Bulk SNP transitions observed compared to the reference nucleotide position.
C) Heteroplasmy increases through developmental time.
D) From either timepoint, anterior cells do not differ from posterior cells in level of heteroplasmy / rate of non-synonymous mutation. From E9.5 to E10.5, non-synonymous mutation rate increase.
E-F) Cells with two or more mutations identified in UMAP space at E9.5 and E10.5.

Discussion and Future Directions

In this study, we reanalyzed our scRNA seq libraries to quantify mitochondrial heteroplasmy in migrating mouse PGCs. Our E9.5 and E10.5 libraries were sequenced deeply enough to analyze further. We characterized the types of SNPs observed in the datasets. We found that heteroplasmy increased across developmental time, but did not vary depending on PGC position in migration. Finally, we found that the proportion of mutations that were non-synonymous actually increased over developmental time.

An important caveat to interpreting these data is that we can only quantify regions where polyadenylated RNAs were expressed. If mitochondrial heteroplasmies exist in non-transcribed regions of the genome, they would not be represented in our dataset. Additionally, the depth at which our libraries were sequenced also affected this analysis. The E10.5 libraries were sequenced more deeply than the E9.5 libraries, a technical caveat to be considered when interpreting these data. Additionally, many cells did not have enough sequencing reads to assess, so we were only able to assess heteroplasmy in certain cells of our dataset. This is particularly pertinent for Figure 5.2E-F, where we highlight individual cells containing greater than two mutations; we expect that other 'Unselected'' cells would have heteroplasmies which are unreported in our analysis due to a lack of coverage.

In addition to total sequencing coverage, a region from the mitochondrial chromosome must be transcribed to obtain coverage in a 10X 3' single cell library. Furthermore, only the first couple hundred base pairs of transcribed sequences are represented in the dataset due to the selection for polyadenylated transcripts. This effect can be see in Figure 5.2A, where there is a 'spike' in coverage at the 3' / clockwise position of each gene. It is important to remember that assessing

mitochondrial heteroplasmy in 3' polyadenylated enriched libraries is not representative of the entire mitochondrial genome.

That the level of heteroplasmies increases so much between E9.5 and E10.5 is quite curious. While it is possible that this reflects true biology, it is also possible that depth of coverage was not sufficient at E9.5 to accurately assess heteroplasmy. For example, at E9.5 only 75 cells met our quality control filtering (deemed high quality cells per Methods, and with deep enough sequencing to classify variants). However, at E10.5 878 cells passed quality control filtering. It is possible that the greater statistical power provided by having an order of magnitude more cells make the E10.5 heteroplasmy assessment more robust. MtDNA copy number only increases from roughly ~3500 copy number per cell at E9.5 to roughly ~3800 mtDNA per cell at E10.5 (Wai et al., 2008). Considering this with the known PGC doubling time (~13-16hrs, although mitochondria replication can happen independent of cell cycle), it is possible that depth of sequencing is contributing to the dramatic increase of heteroplasmy we see from E9.5 to E10.5. More validation is necessary to truly state that a large increase in heteroplasmy occurs over just one day of PGC migration.

Most interesting was the observed increase in non-synonymous mutations when comparing E9.5 to E10.5, which was opposite of the trend we hypothesized would be found. This indicates that if a quality control mechanism is working to filter out low quality PGCs with high mutational burden, it may not be migration. This result is particularly surprising when considering analogous timepoints in human reproductive development, where a decrease in non-synonymous mutations has been observed as age increases (Floros et al., 2018). Previous work quantified the level of mutational variance in mouse and human; although human datasets only were included for primary oocytes and offspring, compared to mouse the mutational level variance increased strongly. This previous work did not quantify human PGCs. However, if the trend observed in

older reproductive cells holds, mouse PGCs may experience a less severe selective pressure after mitochondrial bottleneck compared to human PGCs, due to a decreased mutational load. This would decrease the need for a quality control filter.

More work comparing mouse PGC and human PGC mitochondrial heteroplasmy will have implications for not only reproductive health, but also evolutionary insights into the mitochondrial bottleneck process.

References

Abdu Y, Maniscalco C, Heddleston JM, Chew T-L, Nance J. 2016. Developmentally programmed germ cell remodelling by endodermal cell cannibalism. *Nat Cell Biol* **18**:1302–1310. doi:10.1038/ncb3439

Adra CN, Boer PH, McBurney MW. 1987. Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene* **60**:65–74. doi:10.1016/0378-1119(87)90214-9

Agarwal R, Mori Y, Cheng Y, Jin Z, Olaru AV, Hamilton JP, David S, Selaru FM, Yang J, Abraham JM, Montgomery E, Morin PJ, Meltzer SJ. 2009. Silencing of claudin-11 is associated with increased invasiveness of gastric cancer cells. *PLoS ONE* **4**:e8002. doi:10.1371/journal.pone.0008002

Allio R, Donega S, Galtier N, Nabholz B. 2017. Large Variation in the Ratio of Mitochondrial to Nuclear Mutation Rate across Animals: Implications for Genetic Diversity and the Use of Mitochondrial DNA as a Molecular Marker. *Mol Biol Evol* **34**:2762–2772. doi:10.1093/molbev/msx197

Altschuler SJ, Wu LF. 2010. Cellular heterogeneity: do differences make a difference? *Cell* **141**:559–563. doi:10.1016/j.cell.2010.04.033

Alvarez-Buylla A, Merchant-Larios H. 1986. Mouse primordial germ cells use fibronectin as a substrate for migration. *Exp Cell Res* **165**:362–368. doi:10.1016/0014-4827(86)90590-2

Anderson AP, Luo X, Russell W, Yin YW. 2020. Oxidative damage diminishes mitochondrial DNA polymerase replication fidelity. *Nucleic Acids Res* **48**:817–829. doi:10.1093/nar/gkz1018

Anderson MJ, Magidson V, Kageyama R, Lewandoski M. 2020. Fgf4 maintains Hes7 levels critical for normal somite segmentation clock function. *eLife* **9**. doi:10.7554/eLife.55608

Anderson R, Copeland TK, Schöler H, Heasman J, Wylie C. 2000. The onset of germ cell migration in the mouse embryo. *Mech Dev* **91**:61–68. doi:10.1016/s0925-4773(99)00271-3

Andux S, Ellis RE. 2008. Apoptosis maintains oocyte quality in aging Caenorhabditis elegans females. *PLoS Genet* **4**:e1000295. doi:10.1371/journal.pgen.1000295

Arana ME, Kunkel TA. 2010. Mutator phenotypes due to DNA replication infidelity. *Semin Cancer Biol* **20**:304–311. doi:10.1016/j.semcancer.2010.10.003

Ara T, Nakamura Y, Egawa T, Sugiyama T, Abe K, Kishimoto T, Matsui Y, Nagasawa T. 2003. Impaired colonization of the gonads by primordial germ cells in mice lacking a chemokine, stromal cell-derived factor-1 (SDF-1). *Proc Natl Acad Sci USA* **100**:5319–5323. doi:10.1073/pnas.0730719100

Arora R, Abby E, Ross ADJ, Cantu AV, Kissner MD, Castro V, Ho H-YH, Livera G, Laird DJ. 2016. Meiotic onset is reliant on spatial distribution but independent of germ cell number in the mouse ovary. *J Cell Sci* **129**:2493–2499. doi:10.1242/jcs.189910

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**:126–140. doi:10.1101/gad.224503

Banisch TU, Goudarzi M, Raz E. 2012. Small RNAs in germ cell development. *Curr Top Dev Biol* **99**:79–113. doi:10.1016/B978-0-12-387038-4.00004-5

Berrens RV, Andrews S, Spensberger D, Santos F, Dean W, Gould P, Sharif J, Olova N, Chandra T, Koseki H, von Meyenn F, Reik W. 2017. An endosiRNA-Based Repression Mechanism Counteracts Transposon Activation during Global DNA Demethylation in Embryonic Stem Cells. *Cell Stem Cell* **21**:694-703.e7. doi:10.1016/j.stem.2017.10.004

Besenbacher S, Liu S, Izarzugaza JMG, Grove J, Belling K, Bork-Jensen J, Huang S, Als TD, Li S, Yadav R, Rubio-García A, Lescai F, Demontis D, Rao J, Ye W, Mailund T, Friborg RM, Pedersen CNS, Xu R, Sun J, Rasmussen S. 2015. Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. *Nat Commun* **6**:5969. doi:10.1038/ncomms6969

Besmer P, Manova K, Duttlinger R, Huang EJ, Packer A, Gyssler C, Bachvarova RF. 1993. The kit-ligand (steel factor) and its receptor c-kit/W: pleiotropic roles in gametogenesis and melanogenesis. *Dev Suppl* 125–137.

Bhattaram P, Penzo-Méndez A, Sock E, Colmenares C, Kaneko KJ, Vassilev A, Depamphilis ML, Wegner M, Lefebvre V. 2010. Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun* **1**:9. doi:10.1038/ncomms1008

Blaser H, Eisenbeiss S, Neumann M, Reichman-Fried M, Thisse B, Thisse C, Raz E. 2005. Transition from non-motile behaviour to directed migration during early PGC development in zebrafish. *J Cell Sci* **118**:4027–4038. doi:10.1242/jcs.02522

Bohacek J, Mansuy IM. 2015. Molecular insights into transgenerational non-genetic inheritance of acquired behaviours. *Nat Rev Genet* **16**:641–652. doi:10.1038/nrg3964

Bontems F, Stein A, Marlow F, Lyautey J, Gupta T, Mullins MC, Dosch R. 2009. Bucky ball organizes germ plasm assembly in zebrafish. *Curr Biol* **19**:414–422. doi:10.1016/j.cub.2009.01.038

Bowles J, Feng C-W, Miles K, Ineson J, Spiller C, Koopman P. 2016. ALDH1A1 provides a source of meiosis-inducing retinoic acid in mouse fetal ovaries. *Nat Commun* **7**:10845. doi:10.1038/ncomms10845

Bowles J, Feng C-W, Spiller C, Davidson T-L, Jackson A, Koopman P. 2010. FGF9 suppresses meiosis and promotes male germ cell fate in mice. *Dev Cell* **19**:440–449. doi:10.1016/j.devcel.2010.08.010

Bowles J, Knight D, Smith C, Wilhelm D, Richman J, Mamiya S, Yashiro K, Chawengsaksophak K, Wilson MJ, Rossant J, Hamada H, Koopman P. 2006. Retinoid signaling determines germ cell fate in mice. *Science* **312**:596–600. doi:10.1126/science.1125691

Bowling S, Sritharan D, Osorio FG, Nguyen M, Cheung P, Rodriguez-Fraticelli A, Patel S, Yuan W-C, Fujiwara Y, Li BE, Orkin SH, Hormoz S, Camargo FD. 2020. An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. *Cell* **181**:1410-1422.e27. doi:10.1016/j.cell.2020.04.048

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**:1089–1103. doi:10.1016/j.cell.2007.01.043

Brock A, Chang H, Huang S. 2009. Non-genetic heterogeneity--a mutation-independent driving force for the somatic evolution of tumours. *Nat Rev Genet* **10**:336–342. doi:10.1038/nrg2556

Browaeys R, Saelens W, Saeys Y. 2020. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* **17**:159–162. doi:10.1038/s41592-019-0667-5

Brown WM, George M, Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA* **76**:1967–1971. doi:10.1073/pnas.76.4.1967

Bunis DG, Andrews J, Fragiadakis GK, Burt TD, Sirota M. 2020. dittoSeq: Universal User-Friendly Single-Cell and Bulk RNA Sequencing Visualization Toolkit. *Bioinformatics*. doi:10.1093/bioinformatics/btaa1011

Burke AC, Nelson CE, Morgan BA, Tabin C. 1995. Hox genes and the evolution of vertebrate axial morphology. *Development* **121**:333–346. doi:10.1242/dev.121.2.333

Buss LW. 2014. The Evolution of Individuality. Princeton University Press.

Buss LW. 2006. The Evolution of Individuality (Princeton Legacy Library), 1st Ed. ed. Princeton, N.J: Princeton University Press.

Campbell CD, Chong JX, Malig M, Ko A, Dumont BL, Han L, Vives L, O'Roak BJ, Sudmant PH, Shendure J, Abney M, Ober C, Eichler EE. 2012. Estimating the human mutation rate

using autozygosity in a founder population. *Nat Genet* **44**:1277–1281. doi:10.1038/ng.2418

Canovas S, Campos R, Aguilar E, Cibelli JB. 2017. Progress towards human primordial germ cell specification in vitro. *Mol Hum Reprod* **23**:4–15. doi:10.1093/molehr/gaw069

Cantú AV, Altshuler-Keylin S, Laird DJ. 2016. Discrete somatic niches coordinate proliferation and migration of primordial germ cells via Wnt signaling. *J Cell Biol* **214**:215–229. doi:10.1083/jcb.201511061

Cao L, Shitara H, Horii T, Nagao Y, Imai H, Abe K, Hara T, Hayashi J-I, Yonekawa H. 2007. The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nat Genet* **39**:386–390. doi:10.1038/ng1970

Casper A, Van Doren M. 2006. The control of sexual identity in the Drosophila germline. *Development* **133**:2783–2791. doi:10.1242/dev.02415

Childs AJ, Cowan G, Kinnell HL, Anderson RA, Saunders PTK. 2011. Retinoic Acid signalling and the control of meiotic entry in the human fetal gonad. *PLoS ONE* **6**:e20249. doi:10.1371/journal.pone.0020249

Choi HMT, Schwarzkopf M, Fornace ME, Acharya A, Artavanis G, Stegmaier J, Cunha A, Pierce NA. 2018. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**. doi:10.1242/dev.165753

Chubb JR, Trcek T, Shenoy SM, Singer RH. 2006. Transcriptional pulsing of a developmental gene. *Curr Biol* **16**:1018–1025. doi:10.1016/j.cub.2006.03.092

Coffman CR, Strohm RC, Oakley FD, Yamada Y, Przychodzin D, Boswell RE. 2002. Identification of X-linked genes required for migration and programmed cell death of Drosophila melanogaster germ cells. *Genetics* **162**:273–284.

Conrad DF, Keebler JEM, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV, Zilversmit M, Cartwright R, Rouleau GA, Daly M, Stone EA, Hurles ME, Awadalla P, 1000 Genomes Project. 2011. Variation in genome-wide mutation rates within and between human families. *Nat Genet* **43**:712–714. doi:10.1038/ng.862

Coucouvanis EC, Sherwood SW, Carswell-Crumpton C, Spack EG, Jones PP. 1993. Evidence that the mechanism of prenatal germ cell death in the mouse is apoptosis. *Exp Cell Res* **209**:238–247. doi:10.1006/excr.1993.1307

Crow JF. 2000. The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet* **1**:40–47. doi:10.1038/35049558

Dawson EP, Lanza DG, Webster NJ, Benton SM, Suetake I, Heaney JD. 2018. Delayed male germ cell sex-specification permits transition into embryonal carcinoma cells with features of primed pluripotency. *Development* **145**. doi:10.1242/dev.156612

Delbrück M. 1945. The Burst Size Distribution in the Growth of Bacterial Viruses (Bacteriophages). *J Bacteriol* **50**:131–135.

Díaz-Díaz C, Fernandez de Manuel L, Jimenez-Carretero D, Montoya MC, Clavería C, Torres M. 2017. Pluripotency Surveillance by Myc-Driven Competitive Elimination of Differentiating Cells. *Dev Cell* **42**:585-599.e4. doi:10.1016/j.devcel.2017.08.011

Drake JW, Charlesworth B, Charlesworth D, Crow JF. 1998. Rates of spontaneous mutation. *Genetics* **148**:1667–1686.

Dudley B, Palumbo C, Nalepka J, Molyneaux K. 2010. BMP signaling controls formation of a primordial germ cell niche within the early genital ridges. *Dev Biol* **343**:84–93. doi:10.1016/j.ydbio.2010.04.011

Dudley BM, Runyan C, Takeuchi Y, Schaible K, Molyneaux K. 2007. BMP signaling regulates PGC numbers and motility in organ culture. *Mech Dev* **124**:68–77. doi:10.1016/j.mod.2006.09.005

Durinck S, Spellman PT, Birney E, Huber W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* **4**:1184–1191. doi:10.1038/nprot.2009.97

Eddy EM. 1975. Germ plasm and the differentiation of the germ cell line. *Int Rev Cytol* **43**:229–280. doi:10.1016/S0074-7696(08)60070-4

Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. 2020. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**:1484–1506. doi:10.1038/s41596-020-0292-x

Ellis RE. 2008. Sex determination in the Caenorhabditis elegans germ line. *Curr Top Dev Biol* **83**:41–64. doi:10.1016/S0070-2153(08)00402-X

Ephrussi A, Lehmann R. 1992. Induction of germ cell formation by oskar. *Nature* **358**:387–392. doi:10.1038/358387a0

Evans T, Wade CM, Chapman FA, Johnson AD, Loose M. 2014. Acquisition of germ plasm accelerates vertebrate evolution. *Science* **344**:200–203. doi:10.1126/science.1249325

Ewen-Campen B, Srouji JR, Schwager EE, Extavour CG. 2012. Oskar predates the evolution of germ plasm in insects. *Curr Biol* **22**:2278–2283. doi:10.1016/j.cub.2012.10.019

Ewen KA, Koopman P. 2010. Mouse germ cell development: from specification to sex determination. *Mol Cell Endocrinol* **323**:76–93. doi:10.1016/j.mce.2009.12.013

Extavour CG, Akam M. 2003. Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* **130**:5869–5884. doi:10.1242/dev.00804

Falkenberg M, Larsson N-G, Gustafsson CM. 2007. DNA replication and transcription in mammalian mitochondria. *Annu Rev Biochem* **76**:679–699. doi:10.1146/annurev.biochem.76.060305.152028

Fernández-Pérez D, Brieño-Enríquez MA, Isoler-Alcaraz J, Larriba E, Del Mazo J. 2018. MicroRNA dynamics at the onset of primordial germ and somatic cell sex differentiation during mouse embryonic gonad development. *RNA* **24**:287–303. doi:10.1261/rna.062869.117

Floros VI, Pyle A, Dietmann S, Wei W, Tang WCW, Irie N, Payne B, Capalbo A, Noli L, Coxhead J, Hudson G, Crosier M, Strahl H, Khalaf Y, Saitou M, Ilic D, Surani MA, Chinnery PF. 2018. Segregation of mitochondrial DNA heteroplasmy through a developmental genetic bottleneck in human embryos. *Nat Cell Biol* **20**:144–151. doi:10.1038/s41556-017-0017-8

Folmes CDL, Terzic A. 2014. Metabolic determinants of embryonic development and stem cell fate. *Reprod Fertil Dev* **27**:82–88. doi:10.1071/RD14383

Fujiwara T, Dunn NR, Hogan BL. 2001. Bone morphogenetic protein 4 in the extraembryonic mesoderm is required for allantois development and the localization and survival of primordial germ cells in the mouse. *Proc Natl Acad Sci USA* **98**:13739–13744. doi:10.1073/pnas.241508898

Fukuyama M, Rougvie AE, Rothman JH. 2006. C. elegans DAF-18/PTEN mediates nutrient-dependent arrest of cell cycle and growth in the germline. *Curr Biol* **16**:773–779. doi:10.1016/j.cub.2006.02.073

Gallo CM, Wang JT, Motegi F, Seydoux G. 2010. Cytoplasmic partitioning of P granule components is not required to specify the germline in C. elegans. *Science* **330**:1685–1689. doi:10.1126/science.1193697

Gangaraju VK, Yin H, Weiner MM, Wang J, Huang XA, Lin H. 2011. Drosophila Piwi functions in Hsp90-mediated suppression of phenotypic variation. *Nat Genet* **43**:153–158. doi:10.1038/ng.743

García-Castro MI, Anderson R, Heasman J, Wylie C. 1997. Interactions between germ cells and extracellular matrix glycoproteins during migration and gonad assembly in the mouse embryo. *J Cell Biol* **138**:471–480. doi:10.1083/jcb.138.2.471

Gardner RL, Rossant J. 1979. Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. *J Embryol Exp Morphol* **52**:141–152.

Gartner A, Boag PR, Blackwell TK. 2008. Germline survival and apoptosis. *WormBook* 1–20. doi:10.1895/wormbook.1.145.1

Gerdes J, Schwab U, Lemke H, Stein H. 1983. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int J Cancer* **31**:13–20. doi:10.1002/ijc.2910310104

Gilboa L, Lehmann R. 2006. Soma-germline interactions coordinate homeostasis and growth in the Drosophila gonad. *Nature* **443**:97–100. doi:10.1038/nature05068

Ginsburg M, Eyal-Giladi H. 1987. Primordial germ cells of the young chick blastoderm originate from the central zone of the area pellucida irrespective of the embryo-forming process. *Development* **101**:209–219.

Ginsburg M, Eyal-Giladi H. 1986. Temporal and spatial aspects of the gradual migration of primordial germ cells from the epiblast into the germinal crescent in the avian embryo. *J Embryol Exp Morphol* **95**:53–71.

Ginsburg M, Snow MH, McLaren A. 1990. Primordial germ cells in the mouse embryo during gastrulation. *Development* **110**:521–528.

Glaser RL, Broman KW, Schulman RL, Eskenazi B, Wyrobek AJ, Jabs EW. 2003. The paternal-age effect in Apert syndrome is due, in part, to the increased frequency of mutations in sperm. *Am J Hum Genet* **73**:939–947. doi:10.1086/378419

Gofflot F, Hall M, Morriss-Kay GM. 1997. Genetic patterning of the developing mouse tail at the time of posterior neuropore closure. *Dev Dyn* **210**:431–445. doi:10.1002/(SICI)1097-0177(199712)210:4<431::AID-AJA7>3.0.CO;2-H

Gomperts M, Garcia-Castro M, Wylie C, Heasman J. 1994. Interactions between primordial germ cells play a role in their migration in mouse embryos. *Development* **120**:135–141. doi:10.1242/dev.120.1.135

Goriely A, McVean GAT, Röjmyr M, Ingemarsson B, Wilkie AOM. 2003. Evidence for selective advantage of pathogenic FGFR2 mutations in the male germ line. *Science* **301**:643–646. doi:10.1126/science.1085710

Greenbaum MP, Iwamori T, Buchold GM, Matzuk MM. 2011. Germ cell intercellular bridges. *Cold Spring Harb Perspect Biol* **3**:a005850. doi:10.1101/cshperspect.a005850

Greenberg MVC, Bourc'his D. 2019. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* **20**:590–607. doi:10.1038/s41580-019-0159-6

Grubb BJ. 2006. Developmental Biology, Eighth Edition. Scott F. Gilbert, editor. *Integr Comp Biol* **46**:652–653. doi:10.1093/icb/icl011

Grün D, Lyubimova A, Kester L, Wiebrands K, Basak O, Sasaki N, Clevers H, van Oudenaarden A. 2015. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**:251–255. doi:10.1038/nature14966

Gumienny TL, Lambie E, Hartwieg E, Horvitz HR, Hengartner MO. 1999. Genetic control of programmed cell death in the Caenorhabditis elegans hermaphrodite germline. *Development* **126**:1011–1022.

Hackett JA, Surani MA. 2013. DNA methylation dynamics during the mammalian life cycle. *Philos Trans R Soc Lond B Biol Sci* **368**:20110328. doi:10.1098/rstb.2011.0328

Hagemann-Jensen M, Abdullayev I, Sandberg R, Faridani OR. 2018. Small-seq for single-cell small-RNA sequencing. *Nat Protoc* **13**:2407–2424. doi:10.1038/s41596-018-0049-y

Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA. 2002. Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* **117**:15–23. doi:10.1016/S0925-4773(02)00181-8

Hancks DC, Kazazian HH. 2012. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* **22**:191–203. doi:10.1016/j.gde.2012.02.006

Hargan-Calvopina J, Taylor S, Cook H, Hu Z, Lee SA, Yen M-R, Chiang Y-S, Chen P-Y, Clark AT. 2016. Stage-Specific Demethylation in Primordial Germ Cells Safeguards against Precocious Differentiation. *Dev Cell* **39**:75–86. doi:10.1016/j.devcel.2016.07.019

Hayashi Y, Hayashi M, Kobayashi S. 2004. Nanos suppresses somatic cell fate in Drosophila germ line. *Proc Natl Acad Sci USA* **101**:10338–10342. doi:10.1073/pnas.0401647101

Hay B, Ackerman L, Barbel S, Jan LY, Jan YN. 1988. Identification of a component of Drosophila polar granules. *Development* **103**:625–640.

Heaney JD, Anderson EL, Michelson MV, Zechel JL, Conrad PA, Page DC, Nadeau JH. 2012. Germ cell pluripotency, premature differentiation and susceptibility to testicular teratomas in mice. *Development* **139**:1577–1586. doi:10.1242/dev.076851

Henderson SA, Edwards RG. 1968. Chiasma frequency and maternal age in mammals. *Nature* **218**:22–28.

Hendzel MJ, Wei Y, Mancini MA, Van Hooser A, Ranalli T, Brinkley BR, Bazett-Jones DP, Allis CD. 1997. Mitosis-specific phosphorylation of histone H3 initiates primarily within

pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* **106**:348–360. doi:10.1007/s004120050256

Heyer BS, MacAuley A, Behrendtsen O, Werb Z. 2000. Hypersensitivity to DNA damage leads to increased apoptosis during early mouse development. *Genes Dev* **14**:2072–2084.

Hikabe O, Hamazaki N, Nagamatsu G, Obata Y, Hirao Y, Hamada N, Shimamoto S, Imamura T, Nakashima K, Saitou M, Hayashi K. 2016. Reconstitution in vitro of the entire cycle of the mouse female germ line. *Nature* **539**:299–303. doi:10.1038/nature20104

Hill PWS, Leitch HG, Requena CE, Sun Z, Amouroux R, Roman-Trufero M, Borkowska M, Terragni J, Vaisvila R, Linnett S, Bagci H, Dharmalingham G, Haberle V, Lenhard B, Zheng Y, Pradhan S, Hajkova P. 2018. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature* **555**:392–396. doi:10.1038/nature25964

Hird SN, Paulsen JE, Strome S. 1996. Segregation of germ granules in living Caenorhabditis elegans embryos: cell-type-specific mechanisms for cytoplasmic localisation. *Development* **122**:1303–1312.

Hirshfield AN. 1992. Heterogeneity of cell populations that contribute to the formation of primordial follicles in rats. *Biol Reprod* **47**:466–472.

Huang S. 2009. Non-genetic heterogeneity of cells in development: more than just noise. *Development* **136**:3853–3862. doi:10.1242/dev.035139

Hubbard EJA, Korta DZ, Dalfó D. 2013. Physiological control of germline development. *Adv Exp Med Biol* **757**:101–131. doi:10.1007/978-1-4614-4015-4_5

Hunt PA, Hassold TJ. 2008. Human female meiosis: what makes a good egg go bad? *Trends Genet* **24**:86–93. doi:10.1016/j.tig.2007.11.010

Hurd TR, Herrmann B, Sauerwald J, Sanny J, Grosch M, Lehmann R. 2016. Long Oskar Controls Mitochondrial Inheritance in Drosophila melanogaster. *Dev Cell* **39**:560–571. doi:10.1016/j.devcel.2016.11.004

Ietswaart R, Gyori BM, Bachman JA, Sorger PK, Churchman LS. 2021. GeneWalk identifies relevant gene functions for a biological context using network representation learning. *Genome Biol* **22**:55. doi:10.1186/s13059-021-02264-8

Ishikura Y, Yabuta Y, Ohta H, Hayashi K, Nakamura T, Okamoto I, Yamamoto T, Kurimoto K, Shirane K, Sasaki H, Saitou M. 2016. In Vitro Derivation and Propagation of Spermatogonial Stem Cell Activity from Mouse Pluripotent Stem Cells. *Cell Rep* **17**:2789–2804. doi:10.1016/j.celrep.2016.11.026

Johnson AD, Richardson E, Bachvarova RF, Crother BI. 2011. Evolution of the germ line-soma relationship in vertebrate embryos. *Reproduction* **141**:291–300. doi:10.1530/REP-10-0474

Jónsson H, Sulem P, Kehr B, Kristmundsdottir S, Zink F, Hjartarson E, Hardarson MT, Hjorleifsson KE, Eggertsson HP, Gudjonsson SA, Ward LD, Arnadottir GA, Helgason EA, Helgason H, Gylfason A, Jonasdottir Adalbjorg, Jonasdottir Aslaug, Rafnar T, Frigge M, Stacey SN, Stefansson K. 2017. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**:519–522. doi:10.1038/nature24018

Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M. 2013. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J* **32**:340–353. doi:10.1038/emboj.2012.331

Kanatsu-Shinohara M, Naoki H, Shinohara T. 2016. Nonrandom germline transmission of mouse spermatogonial stem cells. *Dev Cell* **38**:248–261. doi:10.1016/j.devcel.2016.07.011

Kassambara A. 2020. ggpubr: “ggplot2” Based Publication Ready Plots. CRAN.

Kazazian HH, Moran JV. 2017. Mobile DNA in health and disease. *N Engl J Med* **377**:361–370. doi:10.1056/NEJMra1510092

Keightley PD, Ness RW, Halligan DL, Haddrill PR. 2014. Estimation of the spontaneous mutation rate per nucleotide site in a Drosophila melanogaster full-sib family. *Genetics* **196**:313–320. doi:10.1534/genetics.113.158758

Khrapko K, Coller HA, André PC, Li XC, Hanekamp JS, Thilly WG. 1997. Mitochondrial mutational spectra in human cells and tissues. *Proc Natl Acad Sci USA* **94**:13798–13803. doi:10.1073/pnas.94.25.13798

Kimble JE, White JG. 1981. On the control of germ cell development in Caenorhabditis elegans. *Dev Biol* **81**:208–219. doi:10.1016/0012-1606(81)90284-0

Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, Suzuki Y, Kono T. 2012. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet* **8**:e1002440. doi:10.1371/journal.pgen.1002440

Kohler SW, Provost GS, Fieck A, Kretz PL, Bullock WO, Sorge JA, Putman DL, Short JM. 1991. Spectra of spontaneous and mutagen-induced mutations in the lacI gene in transgenic mice. *Proc Natl Acad Sci USA* **88**:7958–7962. doi:10.1073/pnas.88.18.7958

Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir Aslaug, Jonasdottir Adalbjorg, Wong WSW, Sigurdsson G, Walters GB, Steinberg S, Helgason H, Thorleifsson G, Gudbjartsson DF, Helgason A, Magnusson OT, Thorsteinsdottir U, Stefansson K. 2012. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**:471–475. doi:10.1038/nature11396

Köprunner M, Thisse C, Thisse B, Raz E. 2001. A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes Dev* **15**:2877–2885. doi:10.1101/gad.212401

Kota SK, Feil R. 2010. Epigenetic transitions in germ cell development and meiosis. *Dev Cell* **19**:675–686. doi:10.1016/j.devcel.2010.10.009

Koubova J, Menke DB, Zhou Q, Capel B, Griswold MD, Page DC. 2006. Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc Natl Acad Sci USA* **103**:2474–2479. doi:10.1073/pnas.0510813103

Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W, Lin H, Matsuda Y, Nakano T. 2004. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* **131**:839–849. doi:10.1242/dev.00973

Laird DJ, Altshuler-Keylin S, Kissner MD, Zhou X, Anderson KV. 2011. Ror2 enhances polarity and directional migration of primordial germ cells. *PLoS Genet* **7**:e1002428. doi:10.1371/journal.pgen.1002428

Laird DJ, Chang W-T, Weissman IL, Lauzon RJ. 2005a. Identification of a novel gene involved in asexual organogenesis in the budding ascidian Botryllus schlosseri. *Dev Dyn* **234**:997–1005. doi:10.1002/dvdy.20583

Laird DJ, De Tomaso AW, Weissman IL. 2005b. Stem cells are units of natural selection in a colonial ascidian. *Cell* **123**:1351–1360. doi:10.1016/j.cell.2005.10.026

Laird DJ. 2018. How to lose your inheritance. *Nature* **562**:497–498. doi:10.1038/d41586-018-06849-5

Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, Wright CV, Korving JP, Hogan BL. 1999. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* **13**:424–436.

Lawson KA, Hage WJ. 1994. Clonal analysis of the origin of primordial germ cells in the mouse. *Ciba Found Symp* **182**:68–84; discussion 84. doi:10.1002/9780470514573.ch5

Le Bouffant R, Souquet B, Duval N, Duquenne C, Hervé R, Frydman N, Robert B, Habert R, Livera G. 2011. Msx1 and Msx2 promote meiosis initiation. *Development* **138**:5393–5402. doi:10.1242/dev.068452

Lee Hyung, Kim S, Park T, Rengaraj D, Park K, Lee Hong, Park SB, Kim S, Choi SB, Han J. 2013. Compensatory proliferation of endogenous chicken primordial germ cells after elimination by busulfan treatment. *Stem Cell Res Ther* **4**:136. doi:10.1186/scrt347

Lei L, Spradling AC. 2016. Mouse oocytes differentiate through organelle enrichment from sister cyst germ cells. *Science* **352**:95–99. doi:10.1126/science.aad2156

Lei L, Spradling AC. 2013. Mouse primordial germ cells produce cysts that partially fragment prior to meiosis. *Development* **140**:2075–2081. doi:10.1242/dev.093864

Levsky JM, Singer RH. 2003. Fluorescence in situ hybridization: past, present and future. *J Cell Sci* **116**:2833–2838. doi:10.1242/jcs.00633

Li H-P, Peng C-C, Wu C-C, Chen C-H, Shih M-J, Huang M-Y, Lai Y-R, Chen Y-L, Chen T-W, Tang P, Chang Y-S, Chang K-P, Hsu C-L. 2018. Inactivation of the tight junction gene CLDN11 by aberrant hypermethylation modulates tubulins polymerization and promotes cell migration in nasopharyngeal carcinoma. *J Exp Clin Cancer Res* **37**:102. doi:10.1186/s13046-018-0754-y

Li L, Dong J, Yan L, Yong J, Liu X, Hu Y, Fan X, Wu X, Guo H, Wang X, Zhu X, Li R, Yan J, Wei Y, Zhao Y, Wang W, Ren Y, Yuan P, Yan Z, Hu B, Qiao J. 2017. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell* **20**:858-873.e4. doi:10.1016/j.stem.2017.03.007

Longley MJ, Nguyen D, Kunkel TA, Copeland WC. 2001. The fidelity of human DNA polymerase gamma with and without exonucleolytic proofreading and the p55 accessory subunit. *J Biol Chem* **276**:38555–38562. doi:10.1074/jbc.M105230200

Lynch M. 2010. Evolution of the mutation rate. *Trends Genet* **26**:345–352. doi:10.1016/j.tig.2010.05.003

Maatouk DM, Kellam LD, Mann MRW, Lei H, Li E, Bartolomei MS, Resnick JL. 2006. DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133**:3411–3418. doi:10.1242/dev.02500

Madsen CS, Ghivizzani SC, Hauswirth WW. 1993. In vivo and in vitro evidence for slipped mispairing in mammalian mitochondria. *Proc Natl Acad Sci USA* **90**:7671–7675. doi:10.1073/pnas.90.16.7671

Mahowald AP. 2001. Assembly of the Drosophila germ plasm. *Int Rev Cytol* **203**:187–213. doi:10.1016/s0074-7696(01)03007-8

Mahowald AP. 1962. Fine structure of pole cells and polar granules inDrosophila melanogaster. *J Exp Zool* **151**:201–215. doi:10.1002/jez.1401510302

Malki S, van der Heijden GW, O'Donnell KA, Martin SL, Bortvin A. 2014. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev Cell* **29**:521–533. doi:10.1016/j.devcel.2014.04.027

Marlow F. 2015. Primordial germ cell specification and migration. [version 1; peer review: 3 approved]. *F1000Res* **4**. doi:10.12688/f1000research.6995.1

Marshall AS, Jones NS. 2021. Discovering Cellular Mitochondrial Heteroplasmy Heterogeneity with Single Cell RNA and ATAC Sequencing. *Biology (Basel)* **10**. doi:10.3390/biology10060503

Matson CK, Murphy MW, Griswold MD, Yoshida S, Bardwell VJ, Zarkower D. 2010. The mammalian doublesex homolog DMRT1 is a transcriptional gatekeeper that controls the

mitosis versus meiosis decision in male germ cells. *Dev Cell* **19**:612–624. doi:10.1016/j.devcel.2010.09.010

Mazaud-Guittot S, Gow A, Le Magueresse-Battistoni B. 2011. Phenotyping the claudin 11 deficiency in testis: from histology to immunohistochemistry. *Methods Mol Biol* **763**:223–236. doi:10.1007/978-1-61779-191-8_15

McCLINTOCK B. 1950. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA* **36**:344–355. doi:10.1073/pnas.36.6.344

McGee EA. 2000. The regulation of apoptosis in preantral ovarian follicles. *Biol Signals Recept* **9**:81–86. doi:10.1159/000014626

McGranahan N, Swanton C. 2017. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* **168**:613–628. doi:10.1016/j.cell.2017.01.018

McLaren A, Southee D. 1997. Entry of mouse embryonic germ cells into meiosis. *Dev Biol* **187**:107–113. doi:10.1006/dbio.1997.8584

McLaren A. 2003. Primordial germ cells in the mouse. *Dev Biol* **262**:1–15. doi:10.1016/S0012-1606(03)00214-8

Menke DB, Koubova J, Page DC. 2003. Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. *Dev Biol* **262**:303–312. doi:10.1016/S0012-1606(03)00391-9

Michaelson JJ, Shi Y, Gujral M, Zheng H, Malhotra D, Jin X, Jian M, Liu G, Greer D, Bhandari A, Wu W, Corominas R, Peoples A, Koren A, Gore A, Kang S, Lin GN, Estabillo J, Gadomski T, Singh B, Sebat J. 2012. Whole-genome sequencing in autism identifies hot

spots for de novo germline mutation. *Cell* **151**:1431–1442. doi:10.1016/j.cell.2012.11.019

Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg J. 2017. Differences between germline and somatic mutation rates in humans and mice. *Nat Commun* **8**:15183. doi:10.1038/ncomms15183

Mishra P, Chan DC. 2014. Mitochondrial dynamics and inheritance during cell division, development and disease. *Nat Rev Mol Cell Biol* **15**:634–646. doi:10.1038/nrm3877

Miyachi K, Fritzler MJ, Tan EM. 1978. Autoantibody to a nuclear antigen in proliferating cells. *J Immunol* **121**:2228–2234.

Molyneaux KA, Stallock J, Schaible K, Wylie C. 2001. Time-lapse analysis of living mouse germ cell migration. *Dev Biol* **240**:488–498. doi:10.1006/dbio.2001.0436

Molyneaux KA, Zinszner H, Kunwar PS, Schaible K, Stebler J, Sunshine MJ, O'Brien W, Raz E, Littman D, Wylie C, Lehmann R. 2003. The chemokine SDF1/CXCL12 and its receptor CXCR4 regulate mouse germ cell migration and survival. *Development* **130**:4279–4286. doi:10.1242/dev.00640

Morita-Fujimura Y, Tokitake Y, Matsui Y. 2009. Heterogeneity of mouse primordial germ cells reflecting the distinct status of their differentiation, proliferation and apoptosis can be classified by the expression of cell surface proteins integrin α6 and c-Kit. *Dev Growth Differ* **51**:567–583. doi:10.1111/j.1440-169X.2009.01119.x

Muller-Sieburg CE, Sieburg HB, Bernitz JM, Cattarossi G. 2012. Stem cell heterogeneity: implications for aging and regenerative medicine. *Blood* **119**:3900–3907. doi:10.1182/blood-2011-12-376749

Nguyen DH, Soygur B, Peng S-P, Malki S, Hu G, Laird DJ. 2020. Apoptosis in the fetal testis eliminates developmentally defective germ cell clones. *Nat Cell Biol* **22**:1423–1435. doi:10.1038/s41556-020-00603-8

Nieuwkoop PD, Sutasurya LA. 1979. Primordial Germ Cells in the Chordates: Embryogenesis and phylogenesis (Developmental and Cell Biology Series). Cambridge [Eng.]: Cambridge University Press.

Nieuwkoop PD. 1947. Investigations on the regional determination of the central nervous system. *J Exp Biol* **24**:145–183.

Nikolic A, Volarevic V, Armstrong L, Lako M, Stojkovic M. 2016. Primordial germ cells: current knowledge and perspectives. *Stem Cells Int* **2016**:1741072. doi:10.1155/2016/1741072

Nishikawa K, Kinjo AR. 2018. Mechanism of evolution by genetic assimilation : Equivalence and independence of genetic mutation and epigenetic modulation in phenotypic expression. *Biophys Rev* **10**:667–676. doi:10.1007/s12551-018-0403-x

Ohinata Y, Payer B, O'Carroll D, Ancelin K, Ono Y, Sano M, Barton SC, Obukhanych T, Nussenzweig M, Tarakhovsky A, Saitou M, Surani MA. 2005. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**:207–213. doi:10.1038/nature03813

Ohkubo Y, Shirayoshi Y, Nakatsuji N. 1996. Autonomous regulation of proliferation and growth arrest in mouse primordial germ cells studied by mixed and clonal cultures. *Exp Cell Res* **222**:291–297. doi:10.1006/excr.1996.0037

Ohno M. 2019. Spontaneous de novo germline mutations in humans and mice: rates, spectra, causes and consequences. *Genes Genet Syst* **94**:13–22. doi:10.1266/ggs.18-00015

Ohta K, Yamamoto M, Lin Y, Hogg N, Akiyama H, Behringer RR, Yamazaki Y. 2012. Male differentiation of germ cells induced by embryonic age-specific Sertoli cells in mice. *Biol Reprod* **86**:112. doi:10.1095/biolreprod.111.095943

Overeem AW, Chang YW, Spruit J, Roelse CM, Chuva De Sousa Lopes SM. 2021. Ligand-Receptor Interactions Elucidate Sex-Specific Pathways in the Trajectory From Primordial Germ Cells to Gonia During Human Development. *Front Cell Dev Biol* **9**:661243. doi:10.3389/fcell.2021.661243

Park I, Lee H-S. 2015. EphB/ephrinB signaling in cell adhesion and migration. *Mol Cells* **38**:14–19. doi:10.14348/molcells.2015.2116

Pasquale EB. 2010. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. *Nat Rev Cancer* **10**:165–180. doi:10.1038/nrc2806

Percharde M, Wong P, Ramalho-Santos M. 2017. Global hypertranscription in the mouse embryonic germline. *Cell Rep* **19**:1987–1996. doi:10.1016/j.celrep.2017.05.036

Pesce M, Farrace MG, Piacentini M, Dolci S, De Felici M. 1993. Stem cell factor and leukemia inhibitory factor promote primordial germ cell survival by suppressing programmed cell death (apoptosis). *Development* **118**:1089–1094.

Pfeifer SP. 2017. Direct estimate of the spontaneous germ line mutation rate in African green monkeys. *Evolution* **71**:2858–2870. doi:10.1111/evo.13383

Pinz KG, Shibutani S, Bogenhagen DF. 1995. Action of mitochondrial DNA polymerase gamma at sites of base loss or oxidative damage. *J Biol Chem* **270**:9202–9206. doi:10.1074/jbc.270.16.9202

Polani PE, Crolla JA. 1991. A test of the production line hypothesis of mammalian oogenesis. *Hum Genet* **88**:64–70. doi:10.1007/BF00204931

Rahbari R, Wuster A, Lindsay SJ, Hardwick RJ, Alexandrov LB, Turki SA, Dominiczak A, Morris A, Porteous D, Smith B, Stratton MR, UK10K Consortium, Hurles ME. 2016. Timing, rates and spectra of human germline mutation. *Nat Genet* **48**:126–133. doi:10.1038/ng.3469

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4**:e309. doi:10.1371/journal.pbio.0040309

Raj A, van Oudenaarden A. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**:216–226. doi:10.1016/j.cell.2008.09.050

Ramasamy S, Wang H, Quach HNB, Sampath K. 2006. Zebrafish Staufen1 and Staufen2 are required for the survival and migration of primordial germ cells. *Dev Biol* **292**:393–406. doi:10.1016/j.ydbio.2006.01.014

Raz E, Reichman-Fried M. 2006. Attraction rules: germ cell migration in zebrafish. *Curr Opin Genet Dev* **16**:355–359. doi:10.1016/j.gde.2006.06.007

Raz E. 2002. Primordial germ cell development in zebrafish. *Semin Cell Dev Biol* **13**:489–495.

Red-Horse K, Ueno H, Weissman IL, Krasnow MA. 2010. Coronary arteries form by developmental reprogramming of venous cells. *Nature* **464**:549–553. doi:10.1038/nature08873

Reichman-Fried M, Minina S, Raz E. 2004. Autonomous modes of behavior in primordial germ cell migration. *Dev Cell* **6**:589–596. doi:10.1016/s1534-5807(04)00074-7

Reizel Y, Itzkovitz S, Adar R, Elbaz J, Jinich A, Chapal-Ilani N, Maruvka YE, Nevo N, Marx Z, Horovitz I, Wasserstrom A, Mayo A, Shur I, Benayahu D, Skorecki K, Segal E, Dekel N, Shapiro E. 2012. Cell lineage analysis of the mammalian female germline. *PLoS Genet* **8**:e1002477. doi:10.1371/journal.pgen.1002477

Reznik B, Cincotta SA, Jaszczak RG, Mateo LJ, Shen J, Cao M, Baskin L, Ye P, An W, Laird DJ. 2019. Heterogeneity of transposon expression and activation of the repressive network in human fetal germ cells. *Development* **146**. doi:10.1242/dev.171157

Richardson BE, Lehmann R. 2010. Mechanisms guiding primordial germ cell migration: strategies from different organisms. *Nat Rev Mol Cell Biol* **11**:37–49. doi:10.1038/nrm2815

Richardson SR, Gerdes P, Gerhardt DJ, Sanchez-Luque FJ, Bodea G-O, Muñoz-Lopez M, Jesuadian JS, Kempen M-JHC, Carreira PE, Jeddeloh JA, Garcia-Perez JL, Kazazian HH, Ewing AD, Faulkner GJ. 2017. Heritable L1 retrotransposition in the mouse primordial germline and early embryo. *Genome Res* **27**:1395–1405. doi:10.1101/gr.219022.116

Roach JC, Glusman G, Smit AFA, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ. 2010.

Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* **328**:636–639. doi:10.1126/science.1186802

Rohrschneider MR, Nance J. 2013. The union of somatic gonad precursors and primordial germ cells during Caenorhabditis elegans embryogenesis. *Dev Biol* **379**:139–151. doi:10.1016/j.ydbio.2013.03.019

Rucker EB, Dierisseau P, Wagner KU, Garrett L, Wynshaw-Boris A, Flaws JA, Hennighausen L. 2000. Bcl-x and Bax regulate mouse primordial germ cell survival and apoptosis during embryogenesis. *Mol Endocrinol* **14**:1038–1052. doi:10.1210/mend.14.7.0465

Runyan C, Gu Y, Shoemaker A, Looijenga L, Wylie C. 2008. The distribution and behavior of extragonadal primordial germ cells in Bax mutant mice suggest a novel origin for sacrococcygeal germ cell tumors. *Int J Dev Biol* **52**:333–344. doi:10.1387/ijdb.072486cr

Runyan C, Schaible K, Molyneaux K, Wang Z, Levin L, Wylie C. 2006. Steel factor controls midline cell death of primordial germ cells and is essential for their normal proliferation and migration. *Development* **133**:4861–4869. doi:10.1242/dev.02688

Rutherford SL, Lindquist S. 1998. Hsp90 as a capacitor for morphological evolution. *Nature* **396**:336–342. doi:10.1038/24550

Saatcioglu HD, Kano M, Horn H, Zhang L, Samore W, Nagykery N, Meinsohn M-C, Hyun M, Suliman R, Poulo J, Hsu J, Sacha C, Wang D, Gao G, Lage K, Oliva E, Morris Sabatini ME, Donahoe PK, Pépin D. 2019. Single-cell sequencing of neonatal uterus reveals an Misr2+ endometrial progenitor indispensable for fertility. *eLife* **8**. doi:10.7554/eLife.46349

Saitou M, Kagiwada S, Kurimoto K. 2012. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139**:15–31. doi:10.1242/dev.050849

Saitou M, Payer B, O'Carroll D, Ohinata Y, Surani MA. 2005. Blimp1 and the emergence of the germ line during development in the mouse. *Cell Cycle* **4**:1736–1740. doi:10.4161/cc.4.12.2209

Sano H, Renault AD, Lehmann R. 2005. Control of lateral migration and germ cell elimination by the Drosophila melanogaster lipid phosphate phosphatases Wunen and Wunen 2. *J Cell Biol* **171**:675–683. doi:10.1083/jcb.200506038

Santiago A, Erickson CA. 2002. Ephrin-B ligands play a dual role in the control of neural crest cell migration. *Development* **129**:3621–3632.

Santos AC, Lehmann R. 2004. Germ cell specification and migration in Drosophila and beyond. *Curr Biol* **14**:R578-89. doi:10.1016/j.cub.2004.07.018

Sarikaya DP, Extavour CG. 2015. The Hippo pathway regulates homeostatic growth of stem cell niche precursors in the Drosophila ovary. *PLoS Genet* **11**:e1004962. doi:10.1371/journal.pgen.1004962

Schaner CE, Deshpande G, Schedl PD, Kelly WG. 2003. A conserved chromatin architecture marks and maintains the restricted germ cell lineage in worms and flies. *Dev Cell* **5**:747–757. doi:10.1016/s1534-5807(03)00327-7

Schroeder T. 2010. Hematopoietic stem cell heterogeneity: subtypes, not unpredictable behavior. *Cell Stem Cell* **6**:203–207. doi:10.1016/j.stem.2010.02.006

Ségurel L, Wyman MJ, Przeworski M. 2014. Determinants of mutation rate variation in the human germline. *Annu Rev Genomics Hum Genet* **15**:47–70. doi:10.1146/annurev-genom-031714-125740

Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol Cell* **48**:849–862. doi:10.1016/j.molcel.2012.11.001

Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. 2005. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol* **278**:440–458. doi:10.1016/j.ydbio.2004.11.025

Seki Y, Yamaji M, Yabuta Y, Sano M, Shigeta M, Matsui Y, Saga Y, Tachibana M, Shinkai Y, Saitou M. 2007. Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development* **134**:2627–2638. doi:10.1242/dev.005611

Seoane J. 2017. Cancer: Division hierarchy leads to cell heterogeneity. *Nature* **549**:164–166. doi:10.1038/nature23546

Seydoux G, Braun RE. 2006. Pathway to totipotency: lessons from germ cells. *Cell* **127**:891–904. doi:10.1016/j.cell.2006.11.016

Shekhar K, Lapan SW, Whitney IE, Tran NM, Macosko EZ, Kowalczyk M, Adiconis X, Levin JZ, Nemesh J, Goldman M, McCarroll SA, Cepko CL, Regev A, Sanes JR. 2016. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* **166**:1308-1323.e30. doi:10.1016/j.cell.2016.07.054

Singh A, Winterbottom E, Daar IO. 2012. Eph/ephrin signaling in cell-cell and cell-substrate adhesion. *Front Biosci (Landmark Ed)* **17**:473–497.

Slaidina M, Lehmann R. 2017. Quantitative Differences in a Single Maternal Factor Determine Survival Probabilities among Drosophila Germ Cells. *Curr Biol* **27**:291–297. doi:10.1016/j.cub.2016.11.048

Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, Kasprzyk A. 2009. BioMart--biological queries made easy. *BMC Genomics* **10**:22. doi:10.1186/1471-2164-10-22

Smeds L, Qvarnström A, Ellegren H. 2016. Direct estimate of the rate of germline mutation in a bird. *Genome Res* **26**:1211–1218. doi:10.1101/gr.204669.116

Smith JL, Wilson JE, Macdonald PM. 1992. Overexpression of oskar directs ectopic activation of nanos and presumptive pole cell formation in Drosophila embryos. *Cell* **70**:849–859. doi:10.1016/0092-8674(92)90318-7

Snee MJ, Macdonald PM. 2009. Dynamic organization and plasticity of sponge bodies. *Dev Dyn* **238**:918–930. doi:10.1002/dvdy.21914

Snee MJ, Macdonald PM. 2004. Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components. *J Cell Sci* **117**:2109–2120. doi:10.1242/jcs.01059

Snow MHL. 1977. Gastrulation in the mouse: Growth and regionalization of the epiblast. *Development* **42**:293–303. doi:10.1242/dev.42.1.293

Solter D. 2016. Preformation versus epigenesis in early mammalian development. *Curr Top Dev Biol* **117**:377–391. doi:10.1016/bs.ctdb.2015.11.006

Soper SFC, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, Bortvin A. 2008. Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev Cell* **15**:285–297. doi:10.1016/j.devcel.2008.05.015

Soriano P, Jaenisch R. 1986. Retroviruses as probes for mammalian development: allocation of cells to the somatic and germ cell lineages. *Cell* **46**:19–29. doi:10.1016/0092-8674(86)90856-1

Soto-Suazo M, Abrahamsohn PA, Pereda J, Zorn TM. 1999. Distribution and space-time relationship of proteoglycans in the extracellular matrix of the migratory pathway of primordial germ cells in mouse embryos. *Tissue Cell* **31**:291–300. doi:10.1054/tice.1999.0041

Soto-Suazo M, San Martin S, Zorn TMT. 2004. Collagen and tenascin-C expression along the migration pathway of mouse primordial germ cells. *Histochem Cell Biol* **121**:149–153. doi:10.1007/s00418-004-0621-0

Souquet B, Tourpin S, Messiaen S, Moison D, Habert R, Livera G. 2012. Nodal signaling regulates the entry into meiosis in fetal germ cells. *Endocrinology* **153**:2466–2473. doi:10.1210/en.2011-2056

Specchia V, Piacentini L, Tritto P, Fanti L, D'Alessandro R, Palumbo G, Pimpinelli S, Bozzetti MP. 2010. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. *Nature* **463**:662–665. doi:10.1038/nature08739

Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK. 2009. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**:428–432. doi:10.1038/nature08012

Spiller C, Koopman P, Bowles J. 2017. Sex determination in the mammalian germline. *Annu Rev Genet* **51**:265–285. doi:10.1146/annurev-genet-120215-035449

Spudich JL, Koshland DE. 1976. Non-genetic individuality: chance in the single cell. *Nature* **262**:467–471. doi:10.1038/262467a0

Stallock J, Molyneaux K, Schaible K, Knudson CM, Wylie C. 2003. The pro-apoptotic gene Bax is required for the death of ectopic primordial germ cells during their migration in the mouse embryo. *Development* **130**:6589–6597. doi:10.1242/dev.00898

Stepanik V, Dunipace L, Bae Y-K, Macabenta F, Sun J, Trisnadi N, Stathopoulos A. 2016. The migrations of Drosophila muscle founders and primordial germ cells are interdependent. *Development* **143**:3206–3215. doi:10.1242/dev.134346

Stewart JB, Chinnery PF. 2015. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nat Rev Genet* **16**:530–542. doi:10.1038/nrg3966

Strome S, Lehmann R. 2007. Germ versus soma decisions: lessons from flies and worms. *Science* **316**:392–393. doi:10.1126/science.1140846

Strome S, Wood WB. 1983. Generation of asymmetry and segregation of germ-line granules in early C. elegans embryos. *Cell* **35**:15–25. doi:10.1016/0092-8674(83)90203-9

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**:1888-1902.e21. doi:10.1016/j.cell.2019.05.031

St John JC, Facucho-Oliveira J, Jiang Y, Kelly R, Salah R. 2010. Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. *Hum Reprod Update* **16**:488–509. doi:10.1093/humupd/dmq002

Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol* **100**:64–119. doi:10.1016/0012-1606(83)90201-4

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**:472–474. doi:10.1126/science.1198817

Suzuki A, Saga Y. 2008. Nanos2 suppresses meiosis and promotes male germ cell differentiation. *Genes Dev* **22**:430–435. doi:10.1101/gad.1612708

Szabó PE, Hübner K, Schöler H, Mann JR. 2002. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech Dev* **115**:157–160. doi:10.1016/S0925-4773(02)00087-4

Tam PP, Snow MH. 1981. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J Embryol Exp Morphol* **64**:133–147.

Tam PP, Zhou SX. 1996. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev Biol* **178**:124–132. doi:10.1006/dbio.1996.0203

Teixeira FK, Okuniewska M, Malone CD, Coux R-X, Rio DC, Lehmann R. 2017. piRNA-mediated regulation of transposon alternative splicing in the soma and germ line. *Nature* **552**:268–272. doi:10.1038/nature25018

Tong M, Jun T, Nie Y, Hao J, Fan D. 2019. The role of the slit/robo signaling pathway. *J Cancer* **10**:2694–2705. doi:10.7150/jca.31877

Uchimura A, Higuchi M, Minakuchi Y, Ohno M, Toyoda A, Fujiyama A, Miura I, Wakana S, Nishino J, Yagi T. 2015. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res* **25**:1125–1134. doi:10.1101/gr.186148.114

Ueno H, Turnbull BB, Weissman IL. 2009. Two-step oligoclonal development of male germ cells. *Proc Natl Acad Sci USA* **106**:175–180. doi:10.1073/pnas.0810325105

Underwood EM, Caulton JH, Allis CD, Mahowald AP. 1980. Developmental fate of pole cells in Drosophila melanogaster. *Dev Biol* **77**:303–314. doi:10.1016/0012-1606(80)90476-5

Vasiliauskaitė L, Berrens RV, Ivanova I, Carrieri C, Reik W, Enright AJ, O'Carroll D. 2018. Defective germline reprogramming rewires the spermatogonial transcriptome. *Nat Struct Mol Biol* **25**:394–404. doi:10.1038/s41594-018-0058-0

von Meyenn F, Reik W. 2015. Forget the parents: epigenetic reprogramming in human germ cells. *Cell* **161**:1248–1251. doi:10.1016/j.cell.2015.05.039

Wai T, Teoli D, Shoubridge EA. 2008. The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nat Genet* **40**:1484–1488. doi:10.1038/ng.258

Walter CA, Intano GW, McCarrey JR, McMahan CA, Walter RB. 1998. Mutation frequency declines during spermatogenesis in young mice but increases in old mice. *Proc Natl Acad Sci USA* **95**:10015–10019. doi:10.1073/pnas.95.17.10015

Wang N, Zheng J, Chen Z, Liu Y, Dura B, Kwak M, Xavier-Ferrucio J, Lu Y-C, Zhang M, Roden C, Cheng J, Krause DS, Ding Y, Fan R, Lu J. 2019. Single-cell microRNA-mRNA co-sequencing reveals non-genetic heterogeneity and mechanisms of microRNA regulation. *Nat Commun* **10**:95. doi:10.1038/s41467-018-07981-6

Wang RA, Nakane PK, Koji T. 1998. Autonomous cell death of mouse male germ cells during fetal and postnatal period. *Biol Reprod* **58**:1250–1256. doi:10.1095/biolreprod58.5.1250

Weidinger G, Stebler J, Slanchev K, Dumstrei K, Wise C, Lovell-Badge R, Thisse C, Thisse B, Raz E. 2003. dead end, a novel vertebrate germ plasm component, is required for zebrafish primordial germ cell migration and survival. *Curr Biol* **13**:1429–1434. doi:10.1016/s0960-9822(03)00537-2

Weissman IL. 2015. Stem cells are units of natural selection for tissue formation, for germline development, and in cancer development. *Proc Natl Acad Sci USA* **112**:8922–8928. doi:10.1073/pnas.1505464112

Wellik DM. 2007. Hox patterning of the vertebrate axial skeleton. *Dev Dyn* **236**:2454–2463. doi:10.1002/dvdy.21286

Western PS, Miles DC, van den Bergen JA, Burton M, Sinclair AH. 2008. Dynamic regulation of mitotic arrest in fetal male germ cells. *Stem Cells* **26**:339–347. doi:10.1634/stemcells.2007-0622

West JA, Viswanathan SR, Yabuuchi A, Cunniff K, Takeuchi A, Park I-H, Sero JE, Zhu H, Perez-Atayde A, Frazier AL, Surani MA, Daley GQ. 2009. A role for Lin28 in primordial germ-cell development and germ-cell malignancy. *Nature* **460**:909–913. doi:10.1038/nature08210

Whittle CA, Extavour CG. 2016. Refuting the hypothesis that the acquisition of germ plasm accelerates animal evolution. *Nat Commun* **7**:12637. doi:10.1038/ncomms12637

Wolke U, Jezuit EA, Priess JR. 2007. Actin-dependent cytoplasmic streaming in C. elegans oogenesis. *Development* **134**:2227–2236. doi:10.1242/dev.004952

Wong WSW, Solomon BD, Bodian DL, Kothiyal P, Eley G, Huddleston KC, Baker R, Thach DC, Iyer RK, Vockley JG, Niederhuber JE. 2016. New observations on maternal age effect on germline de novo mutations. *Nat Commun* **7**:10486. doi:10.1038/ncomms10486

Yabuta Y, Kurimoto K, Ohinata Y, Seki Y, Saitou M. 2006. Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling. *Biol Reprod* **75**:705–716. doi:10.1095/biolreprod.106.053686

Yamada Y, Davis KD, Coffman CR. 2008. Programmed cell death of primordial germ cells in Drosophila is regulated by p53 and the Outsiders monocarboxylate transporter. *Development* **135**:207–216. doi:10.1242/dev.010389

Ying Y, Liu XM, Marble A, Lawson KA, Zhao GQ. 2000. Requirement of Bmp8b for the generation of primordial germ cells in the mouse. *Mol Endocrinol* **14**:1053–1063. doi:10.1210/mend.14.7.0479

Ying Y, Zhao GQ. 2001. Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse. *Dev Biol* **232**:484–492. doi:10.1006/dbio.2001.0173

Youngren KK, Coveney D, Peng X, Bhattacharya C, Schmidt LS, Nickerson ML, Lamb BT, Deng JM, Behringer RR, Capel B, Rubin EM, Nadeau JH, Matin A. 2005. The Ter mutation in the dead end gene causes germ cell loss and testicular germ cell tumours. *Nature* **435**:360–364. doi:10.1038/nature03595

Young MD, Behjati S. 2020. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**. doi:10.1093/gigascience/giaa151

Young RA. 2011. Control of the embryonic stem cell state. *Cell* **144**:940–954. doi:10.1016/j.cell.2011.01.032

Zeng Y, Chen T. 2019. DNA Methylation Reprogramming during Mammalian Development. *Genes (Basel)* **10**. doi:10.3390/genes10040257

Zhang J, Zhang M, Acampora D, Vojtek M, Yuan D, Simeone A, Chambers I. 2018. OTX2 restricts entry to the mouse germline. *Nature* **562**:595–599. doi:10.1038/s41586-018-0581-5

Zhao L, Arsenault M, Ng ET, Longmuss E, Chau TC-Y, Hartwig S, Koopman P. 2017. SOX4 regulates gonad morphogenesis and promotes male germ cell differentiation in mice. *Dev Biol* **423**:46–56. doi:10.1016/j.ydbio.2017.01.013

Zheng C-J, Luebeck EG, Byers B, Moolgavkar SH. 2005. On the number of founding germ cells in humans. *Theor Biol Med Model* **2**:32. doi:10.1186/1742-4682-2-32

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Rebecca Jaszczak
C9613266B97240C...
Author Signature

12/7/2021
Date