To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of FRIEDA K. CHAN find it satisfactory and recommend that it be accepted.

-------------------------------------------------------------------------
Patricia Hunt, Ph.D Chair

-------------------------------------------------------------------------
Jennifer Watts, Ph.D

-------------------------------------------------------------------------
John Wyrick, Ph.D

-------------------------------------------------------------------------
Jim Pru, Ph.D
ACKNOWLEDGEMENT

I give my sincerest gratitude to my current advisor, Dr. Patricia Hunt, for her mentoring and encouragement me during my graduate career. I want to thank my former advisor Dr. Jon Oatley and the rest of the Oatley laboratory, both current and former members, for their support and assistance in my thesis project. I wish to acknowledge my current and former committee members, Dr. Jennifer Watts, Dr. John Wyrick, Dr. Jim Pru, Dr. Michael Griswold, and Dr. Ping Ye for their intellectual input and guidance throughout my training. I will like to recognize current and past members of the graduate advising committee for helping navigate through all the academic hurdles. I desire to express my appreciation to Dr. Grant Trobridge for his guidance in finding a postdoctorate position. Finally, I am forever indebted to my amazing boyfriend, Justin Bahrami, and my wonderful family for their constant love, support, and encouragement.
Chair: Patricia Hunt

Spermatogonial stem cells (SSCs) are a rare population of cells in the male germ line that self-renew and differentiate into progenitor spermatogonia for the initiation of spermatogenesis. The production of viable gametes is vital for transmitting genetic and epigenetic information to the next generation. Dysregulation of SSCs interferes with this process and causes infertility, which is detrimental to the survival and evolutionary fitness of sexually reproductive species.

In the US, infertility is suffered by ~10% of all reproductive age couples, and half of these cases are due to problems in the male, factors that interfere with the production of sperm. The severity of male infertility ranges fromazoospermia to reproductive tract abnormalities that prevent the delivery of sperm. Azoospermia has been associated with genetic disorders and cancer therapies that result in dysregulation of SSCs. Thus, understanding the molecular mechanisms that regulate SSC maintenance could lead to improved diagnostic procedures, advances in assisted reproductive technology, and the development of gene therapy techniques to treat infertility patients.

Prior to this dissertation, the field of reproductive biology had no viable means of isolating pure populations of SSCs. As a former member of the Oatley lab, I participated in
projects to identify and study SSCs. In published data, we identified inhibitor of DNA binding 4 (Id4) as a putative SSC marker. Subsequently, an Id4-Gfp transgenic mouse line was generated as a tool to isolate purified populations of SSCs and confirmed that SSC capacity is found only in the ID4-GFP+ cells. Consequently, the ID4-GFP+/SSCs and ID4-GFP-/progenitors were isolated for transcriptome analysis that identified 11 putative genes for SSC maintenance.

Undifferentiated transcription factor 1 (Utf1) was among the 11 genes suggested to regulate SSC state. This gene is particularly interesting because it has been implicated in controlling ESC pluripotency and differentiation by regulating the epigenome. Our studies demonstrated that UTF1 localizes to SSCs and a subpopulation of progenitor spermatogonia and showed that transient reduction of Utf1 decreased SSC abundance. Taken together, the studies summaries here have provided tools to study the transcriptome and epigenome that regulates SSC self-renewal and differentiation.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER 1 MOLECULAR PATHWAYS THAT INFLUENCE SPERMATOGONIAL STEM CELLS</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Spermatogenesis and spermatozoa maturation</td>
<td>2</td>
</tr>
<tr>
<td>Prevailing and alternative models of SSC self-renewal and differentiation</td>
<td>3</td>
</tr>
<tr>
<td>Molecular hallmark for spermatogonia</td>
<td>5</td>
</tr>
<tr>
<td>Understanding SSCs: a pathway to diagnose and treat infertility</td>
<td>7</td>
</tr>
<tr>
<td>2. RESEARCH AIMS</td>
<td>9</td>
</tr>
<tr>
<td>3. FIGURES</td>
<td>11</td>
</tr>
<tr>
<td>4. REFERENCES</td>
<td>18</td>
</tr>
<tr>
<td>CHAPTER 2 FUNCTIONAL AND MOLECULAR FEATURES OF THE ID4+ GERMLINE STEM</td>
<td></td>
</tr>
<tr>
<td>CELL POPULATION IN MOUSE TESTES</td>
<td></td>
</tr>
<tr>
<td>1. ABSTRACT</td>
<td>24</td>
</tr>
<tr>
<td>2. INTRODUCTION</td>
<td>25</td>
</tr>
<tr>
<td>3. RESULTS</td>
<td>28</td>
</tr>
<tr>
<td>Generation of an <em>Id4-Gfp</em> reporter mouse line</td>
<td>28</td>
</tr>
<tr>
<td>Expression of ID4-GFP selectively labels $A_{single}$ cells of the spermatogonial population</td>
<td>29</td>
</tr>
<tr>
<td><em>Id4</em> expressing cells represent a subset of the $A_{single}$ spermatogonial population</td>
<td>30</td>
</tr>
</tbody>
</table>
Stem cell capacity of the undifferentiated spermatogonial population resides in the ID4-GFP+ fraction.................................................................33

Unique transcriptome features of the ID4-GFP+ spermatogonial population...........35

4. DISCUSSION ........................................................................................................37

5. MATERIALS AND METHODS ........................................................................43

6. FIGURES.............................................................................................................54

7. REFERENCES.....................................................................................................71

CHAPTER 3 THE RESTORATIVE CAPACITY OF SPERMATOGONIAL STEM CELL IS INFLUENCED BY UTF1

1. ABSTRACT .......................................................................................................75

2. INTRODUCTION ...............................................................................................75

3. RESULTS ..........................................................................................................77

Utf1 and Sox2 are upregulated in SSCs .................................................................77

UTF1 is present in SSCs and a subpopulation of progenitor spermatogonia ..........78

Transient reduction of Utf1 diminishes the colonizing potential of SSCs ..........78

SOX2 activates Utf1 expression in SSCs ..............................................................79

4. DISCUSSION ...................................................................................................80

5. MATERIALS AND METHODS .......................................................................82

6. FIGURES ..........................................................................................................89

7. REFERENCES.....................................................................................................95

CHAPTER 4 SUMMARY

1. INTRODUCTION ............................................................................................97

2. FUTURE DIRECTIONS ....................................................................................98

What is the methylome that defines an SSC? .........................................................101
What are the genes that direct SSC differentiation? ......................................................... 103

3. REFERENCES ......................................................................................................................... 106
LIST OF TABLES

1. Table S1. Antibodies used for immunostaining analyses ..............................................47
2. Table S2. Primers used for Sybr green based qRT-PCR analyses ..................................49
3. Table S3. Oligonucleotide sequences used for capture an ld4 genomic fragment and insertion into the vector pClasperA. ................................................................................53
4. Table S4. Primer sequences for eGfp-Ura3 cassette insertion into ld4 exon 1 ........53
5. Primers used for Sybr green based qRT-PCR analyses ..............................................84
LIST OF FIGURES

1. Figure 1. Various stages of spermatogenesis in mammalian male ......................... 11
2. Figure 2. Undifferentiated and differentiating spermatogonia .................................. 12
3. Figure 3. Irreversible and reversible SSC models ...................................................... 13
4. Figure 4. Ret, Gfra1, and Gdnf are important for spermatogenesis in mice .............. 14
5. Figure 5. Plzf is important for spermatogenesis in mice ........................................... 15
6. Figure 6. GDNF signaling inhibits STAT3 phosphorylation to regulate Ngn3 transcription .................................................................................................................................. 16
7. Figure 7. Etv5 is important for spermatogenesis in both mice and men ................. 17
8. Figure 1. Generation of a transgenic Id4-Gfp reporter mouse line .......................... 54
9. Figure 2. Identification of ID4-GFP- expressing spermatogonia in testes of LT-11B6 mice at multiple stages of postnatal life .................................................................................. 56
10. Figure 3. Proportionality of ID4-GFP- expressing cells in the undifferentiated spermatogonial population of testes from LT-11B6 mice at multiple stages of postnatal life ................................................................................................................................. 57
11. Figure 4. Frequency and distribution of ID4-GFP- expressing spermatogonia in testes of LT-11B6 mice during postnatal development and adulthood ....................... 59
12. Figure 5. Regenerative capacity of ID4-GFP+ and ID4-GFP- spermatogonial populations from testes of hybrid LT-11B6/RosaLacZ mice ........................................ 61
13. Figure 6. Transcriptome analysis of ID4-GFP+ and ID4-GFP- spermatogonial populations from testes of LT-11B6 mice .............................................................................................................................. 63
14. Figure 7. The hierarchy for subsets of spermatogonia in the undifferentiated population of mouse testes and expression profile for intracellular molecular markers .................................................................................................................................................65

15. Figure S1. Representative image of ID4-GFP expressing cells within seminiferous tubules from testes of adult mice ........................................................................................................................................66

16. Figure S2. Representative images of co-immunofluorescent staining for the differentiating spermatogonial markers STRA8 and KIT (red) and ID4-GFP (green) in cross-sections of testes from adult LT-11B6 mice ........................................................................67

17. Figure S3. Representative images depicting tubular/blood vessel, tubular/tubular, and tubular/interstitial localization of *Id4-Gfp* expressing cells within cross-sections of testes from adult LT-11B6 mice ..........................................................68

18. Figure S4. Representative images of cross-sections from testes of *W/Wv* recipient mice three months after transplantation with ID4-GFP+ spermatogonia isolated from testes of adult LT-11B6 donor mice using FACS .................................................................69

19. Figure S5. Validation of differential abundance using qRT-PCR analysis for 11 transcripts determined to be significantly greater in ID4-GFP+ cells compared to ID4-GFP- cells using RNA-Seq analysis ............................................................................................................70

20. Figure 1. *Utf1* and *Sox2* are upregulated in SSCs ........................................................................................................89

21. Figure 2. *Utf1* is present in SSCs and a subpopulation of progenitor spermatogonia .................................................................................................................................................................90

22. Figure 3. Transient reduction of *Utf1* diminishes the colonizing potential of SSCs ........................................................................................................................................................................91

23. Figure 4. SOX2 activates *Utf1* expression in SSCs ..............................................................................................................93
Dedication

To the late Dr. Simon Chan: For inspiring me to pursue my love for science and music.

To Justin Bahrami: For motivating me to fight for my dreams.
CHAPTER ONE

MOLECULAR PATHWAYS THAT INFLUENCE SPERMATOGONIAL STEM CELLS

Introduction

The survival and evolutionary fitness of a sexually reproductive species depends on the production of viable gametes that transmit genetic and epigenetic information to the next generation. Male gametes or spermatozoa are produced continuously from the onset of puberty throughout the lifetime of most mammalian males, and an average human male is thought to produce approximately $10^{12}$ spermatozoa during the course of his life (Griswold and Oatley 2013).

A specialized developmental pathway called spermatogenesis produces haploid spermatozoa from diploid germ cells. Continual sperm production in the adult testis is made possible by a population of spermatogonial stem cells (SSCs). The ‘stemness’ of SSCs is defined by their ability to (1) self-renew and proliferate, contributing to tissue homeostasis, and (2) restore spermatogenesis in germ cell depleted testes. Misregulation, disruption, or damage to the SSC population can cause infertility or cancer.

In the rodent testis, SSCs are thought to comprise only 0.03% of all germ cells (Nagano 2003; Aponte, van Bragt et al. 2005). Although SSCs are rare, they have the unique ability to give rise to a large number of sperms. To understand the molecular characteristics of SSCs requires the ability to separate them from all other germ cells. Thus, the absence of a unique molecular marker to enrich for SSCs presented a critical barrier to investigation.

The studies in Chapter 2 describe the identification and testing of an SSC-specific marker that represents a significant advance in the field. Specifically, a molecular marker for SSCs will facilitate research on the establishment and maintenance of SSCs, provide new tools
for fertility treatment, aid in the refinement of *in vitro* germ cell culture techniques, and allow for the development of methods to protect SSCs during cancer treatment. It is also important to note that, prior to the publication of the findings detailed in Chapter 2, the only means of evaluating SSC potential was by a transplantation assay where putative SSCs from donor mice restore spermatogenesis in a germ cell depleted recipient testis.

**Spermatogenesis and spermatozoa maturation**

Mammalian spermatogenesis is a complex process regulated by endocrine, paracrine, and autocrine pathways. Endocrine factors are secreted by the hypothalamic-pituitary axis and testis (Holdcraft and Braun 2004; Oates 2012). Paracrine factors are produced by the somatic cells of the testis, including the myoid, Leydig, and Sertoli cells (Toppari, Kangasniemi et al. 1991; Holdcraft and Braun 2004). Sertoli cells are the only cells that have direct contact with the germ cells during the entire process of sperm production (Griswold 1995). While endocrine and paracrine factors are very important for spermatogenesis, this research focuses on the autocrine pathways within the germ cell population, more specifically, the molecular pathways that maintain SSC self-renewal. Spermatogenesis occurs within the seminiferous tubules of the mature testis and begins with the SSCs that reside along the basement membrane. Through mitotic cell divisions, SSCs self-renew to produce daughter cells that maintain the stem cell pool and produce progenitor cells (known as $A_{single}$ or $A_s$) that proceed with spermatogenesis. As progenitor cells amplify in number through multiple cycles of incomplete mitosis, producing daughter cells that remain connected by intercellular bridges. These cells are known as $A_{pair}$ and $A_{align}$ progenitors (De Rooij and Lok 1987; Meistrich and Van Beek 1993; Tegelenbosch and
Collectively, the heterogeneous population of SSCs and progenitor cells are known as undifferentiated spermatogonia (Figure 1,2).

When progenitor cells are exposed to retinoic acid, they transition into a differentiating state, and this class of differentiating spermatogonia is known as A1 cells. A1 cells undergo a series of mitotic divisions, producing A2, A3, A4, intermediate, and B cell populations (de Rooij 2001; Chiarini-Garcia and Russell 2002) (Figure 2). The competency to undergo meiosis is a property of B differentiating cells, thus, they mark the beginning of the meiotic process. During meiosis, diploid germ cells undergo one round of DNA synthesis and two rounds of cell division to generate four haploid gametes. At the end of meiosis, the haploid cells undergo the maturation process of spermiogenesis, during which the DNA becomes highly condensed, the Golgi apparatus becomes the acrosome cap, and one of the two centrioles is reorganized to form the sperm tail. The mature elongated spermatids are released into the lumen of the seminiferous tubule as non-motile spermatozoa (Oud, de Jong et al. 1979; Kluin, Kramer et al. 1984) (Figure 1). The entire process beginning with SSCs and ending with the release of non-motile spermatozoa into the lumen occurs over a period of 35 days in mice (Brinster 2007) (Figure 1). After leaving the seminiferous tubules, the non-motile spermatozoa pass through the epididymis where they gain motility and are stored prior to ejaculation.

**Prevailing and alternative models of SSC self-renewal and differentiation**

SSCs are characterized by maintaining slow cell cycle, unique gene expression, and distinct epigenetics. They also give rise to a population of rapidly dividing progenitor cells that undergo spermatogenesis to produce sperm. Disruption in the SSC cell cycle has been suggested to deplete the SSC pool, resulting in infertility (Yang, Gwost et al. 2013).
Interruption in gene expression has been shown to cause testicular cancer (Singh, Burnicka-Turek et al. 2011). Interference in SSC epigenetics has been indicated in transgenerational genetic disease (Guerrero-Bosagna and Skinner 2014; Skinner 2014; Vrooman, Oatley et al. 2015). Therefore, by understanding the distinctive features and unique mechanistic profile has the potential to develop new diagnostic and therapeutic techniques for infertility.

In mammals, the prevailing SSC differentiation model is where the A\textsubscript{s} progenitors are committed irreversibly to spermatogenesis and have lost self-renewal ability (Figure 3). Although this is the prevailing model for germline stem cell biology, some have questioned the model because evidence for it is based largely on morphological analyses of fixed specimens. The criticisms warrant further investigation through functional evaluation of SSCs in comparison to progenitors (Yoshida 2010).

The alternative model for mammalian SSCs is similar to the Drosophila SSC model. In Drosophila, transit-amplifying progenitors (A\textsubscript{pr} and A\textsubscript{al}) have the ability to dedifferentiate through activation of the Jak-STAT signaling pathway or withdrawal of Bam differentiation factor (Brawley and Matunis 2004; Sheng, Brawley et al. 2009). These dedifferentiated cells can return to the stem cell niche and resume self-renewal (Cheng, Tiyaboonchai et al. 2011). According to this model, the mouse A\textsubscript{s} progenitors have the ability to revert back to SSC state. It has been further postulated that A\textsubscript{pr} and A\textsubscript{al} cells could regain SSC function by losing their intercellular bridge to form A\textsubscript{s} that can revert back to SSCs (Barroca, Lassalle et al. 2009; Yoshida 2009; Spradling, Fuller et al. 2011)(Figure 3). Under this model, progenitor cell plasticity contributes to the maintenance of the SSC pool and restores
spermatogenesis in germ cell depleted testes (Nakagawa, Sharma et al. 2010; Hara, Nakagawa et al. 2014).

Investigations into the mechanisms of SSCs have been hampered because there was no definitive way to enrich SSCs. In 2011, Oatley et al. identified inhibitor of DNA binding 4 (ID4) as a putative marker for SSCs because its expression was restricted to a subpopulation of A₅ spermatogonia (Oatley, Kaucher et al. 2011). Chapter 2 of this dissertation describes the production of a transgenic mouse line expressing a Gfp-tagged Id4 transgene to further define the role of ID4 in germ cell biology. Using this transgene, ID4-GFP+ and ID4-GFP- spermatogonial populations were isolated and transplanted into germ-cell-depleted recipient testes. Post-transplantation analyses demonstrated a major difference between the two populations: ID4+ cells have enhanced colony formation by contrast there were fewer to no colony formation by ID4- cells. These results provide further evidence that ID4 is the first marker for SSCs and support the prevailing model where only SSCs have stem cell potential (Chan, Oatley et al. 2014).

**Molecular hallmark for spermatogonia**

The research performed in this thesis was based on groundwork laid by previous studies of many investigators. Their discovery of various molecular markers expressed by both SSCs and progenitors provided insights into the intricate molecular network of spermatogonial cells allowed for identification of ID4 as an SSC marker in Chapter 2. The generation of ID4-GFP will serve as an important tool to investigate the mechanisms that control SSC identity. One of the first important signaling molecules identified in germ cell maintenance was glial cell line-derived neurotrophic factor (GDNF). This paracrine factor secreted by Sertoli cells was shown to regulate SSCs in a dose-dependent manner by binding to a receptor complex
consisting of Glial cell line-derived neurotrophic factor family receptor alpha 1 (GFRa1) and the ret proto-oncogene (RET) (Hofmann, Braydich-Stolle et al. 2005; Naughton, Jain et al. 2006; He, Jiang et al. 2007). In vitro and in vivo experimentation showed that overexpression of GDNF led to an accumulation of SSCs and progenitors, whereas depletion resulted in a reduction in spermatogonia (Meng, Lindahl et al. 2000). Targeted ablation of Gfra1 and Ret receptor genes in mouse testes caused severe depletion in SSCs and infertility (Zhou and Griswold 2008) (Figure 4).

Based on these evidence, Oatley et al., investigated the genes that were upregulated in SSCs via GDNF signaling (Oatley, Avarbock et al. 2006; Oatley, Avarbock et al. 2007; Oatley, Kaucher et al. 2011). These studies resulted in the identification of Id4 as a candidate gene for maintenance of SSC state (Oatley, Avarbock et al. 2006; Oatley, Avarbock et al. 2007; Oatley, Kaucher et al. 2011). Subsequent Id4 expression and functional studies suggested that it was limited to a subset of Aₘ spermatogonia (Oatley, Avarbock et al. 2006; Oatley, Avarbock et al. 2007; Oatley, Kaucher et al. 2011). Taken together, these studies suggested ID4 as a marker of SSCs and prompted the investigations performed in Chapter 2 of this dissertation.

In addition to Id4 expression, GDNF is thought to be an important player for germ cell maintenance because its depletion caused infertility. Previous research showed that GDNF upregulated the expression of the POZ-Krüppel family (POK) transcription factor (Plzf or Zbtb16), which is expressed in both SSCs and progenitor spermatogonia (Filipponi, Hobbs et al. 2007). Recent studies suggest that GDNF promotes SSCs and progenitor proliferation through the upregulation of Plzf via the mTOR signaling pathway (Hobbs, Seandel et al. 2010). Consistent with this, Plzf ablation in the mouse testes caused germ-cell depletion
and infertility (Buaas, Kirsh et al. 2004; Costoya, Hobbs et al. 2004) (Figure 5). Based on these results, PLZF has become the gold standard in the field for marking the heterogeneous population of SSCs and progenitors.

In addition to upregulation \( Plzf \), recent evidence showed that GDNF represses the expression of another important player for germ cell maintenance: neurogenin 3 (\( Ngn3 \)), which is expressed by SSCs and progenitors (Yoshida, Takakura et al. 2004). \textit{In vitro} experiments suggested that GDNF inhibited STAT3 phosphorylation to activate \( Ngn3 \) transcription, which promotes SSC self-renewal (Kaucher, Oatley et al. 2012) (Figure 6). Reduction of \( Ngn3 \) expression prevented progenitors from transitioning to the A1 differentiating state, thereby implicating NGN3 in the priming progenitor spermatogonia for differentiation (Kaucher, Oatley et al. 2012). Using NGN3 as a marker in life cell imaging, Nakagawa et al., found evidence that NGN3+/A\(_{1}\) cells could disassociate from their intercellular bridge into NGN3+/A\(_{5}\) progenitors. Based on these findings, they postulated that these NGN3+/A\(_{5}\) cells dedifferentiate and reenter the SSC pool (Nakagawa, Nabeshima et al. 2007; Hara, Nakagawa et al. 2014). Although their proposed model is plausible, but because NGN3 doesn’t specifically mark SSCs, their results warrant further investigation.

**Understanding SSCs: a pathway to diagnose and treat infertility**

In the United States, \(~10\%\) of reproductive age human couples suffer from infertility and of these \(~50\%\) of the cases are attributable to the male factors (Chandra, Copen et al. 2013; Chandra, Copen et al. 2014). Although cases of most male infertility are idiopathic, both cancer treatments and genetic disorders are known to dysregulate the SSC population and cause infertility (Goldfarb, Mulhall et al. 2013; Aston 2014; Brannigan 2014; Hotaling 2014). Thus, understanding the characteristics of SSCs and the mechanisms that regulate
their differentiation is relevant to infertility caused by genetic disease and cancer treatment.

An example of a gene mutation that causes infertility is the ETS variant 5 (Etv5) gene. ETV5 is thought to regulate SSC establishment, self-renewal, and differentiation (Morrow, Hostetler et al. 2007). In both mice and men, homozygous carriers of Etv5 mutations have germ cell depletion and infertility resulting from SSCs dysregulation (Schlesser, Simon et al. 2008; Tyagi, Carnes et al. 2009; Simon, Ekman et al. 2010; O’Bryan, Grealy et al. 2012) (Figure 7).

In addition to genetic disorders, cancer therapies are one of the leading causes of male infertility. High-dose chemotherapy, whole-body radiation, and radiation to the gonads destroy both cancer cells and tissue-specific stem cells in the human body. Thus, most male cancer survivors are infertile (Meistrich 2009), although post-pubertal male patients can preserve fertility with techniques such as cryopreservation of sperm and in vitro fertilization pre-embryo transfer procedure (Hsiao, Stahl et al. 2011), but this is not an option for prepubertal male patients who have not produced sperms.

In the United States, there are ~5000 boys under the age of 15 diagnosed with cancer yearly and ~83% of these patients are expected to survive (Howlader, Noone et al. 1975-2008). With the improvement in cancer therapies improve there is a growing population of long-term cancer survivors who are infertile. Thus, markers for SSC could provide researchers an additional tool to improve techniques in pre-cancer germ cells preservation. These methods would include spermatogonia cryopreservation, SSC transplantation, testicular tissue grafting, testicular tissue organ culture, and induced pluripotent stem cell differentiation into gametes or transplantable male germ line stem cells (Dobrinski,
SSCs self-renewal and differentiation is important for continual spermatogenesis. Depletion of the SSC pool via genetic mutations and cancer therapies causes infertility, a prevalent disease that detriments the survival and evolutionary fitness of a reproductive species. Identification and isolation of SSCs will be critical in studying the molecular mechanisms that regulate their “stemness”, thereby providing novel methods in diagnosis and treatment of infertility.

**Research Aims**

The focus of my dissertation research was the molecular mechanisms that regulate SSC self-renewal and differentiation. The chapters that follow describe the identification of a novel SSC marker and the gene that regulate SSC state.

**Chapter 2:**

The goal of this project was to identify the transcriptional differences between the SSCs and progenitor spermatogonia. In previous studies, ID4 was identified as a putative SSC marker. An *Id4-Gfp* transgenic mouse was generated and allowed for isolation for purified population of SSCs for gene expression studies. Transcriptome analyses showed that 11 genes were upregulated by SSCs implicating these genes as candidate for the maintenance of SSC self-renewal and differentiation.
Chapter 3:

Undifferentiated transcription factor 1 (Utf1) was one of the 11 genes upregulated in SSCs. The goal of this project was to investigate the expression and function of UTF1 in mouse SSCs. *In vitro* and *in vivo* experiments showed that Utf1 expression was upregulated in SSCs. Additionally, the results provided evidence that SOX2 activation of Utf1 expression is conserved in SSCs, ESCs, and iPSCs. Taken together, these results underscore the importance of UTF1 in SSC maintenance.
Figure 1. Various stages of spermatogenesis in mammalian male. Diagram that depicts spermatogenesis in mammalian male. A) Along the basement membrane is the spermatogonia population which includes SSCs, progenitor spermatogonia, and differentiating spermatogonia. B) After gaining competency for meiotic initiation, the germ cells enter the spermatocyte stage which proceeds through two meiotic divisions. During this period, they become primary and secondary spermatocytes. C) By the end of the spermatocyte stage, are germ cells develop into haploid round spermatid that will undergo spermiogenesis and become mature elongated spermatids. D) The mature spermatids are then released into the lumen of the seminiferous tubules as spermatozoa. Spermatogenesis within mice occurs over a period of 35 days.
Figure 2. Undifferentiated and differentiating spermatogonia. In the established rodent model of spermatogenesis, the spermatogonial population consists of two differentiate subtypes: undifferentiated and differentiating spermatogonia. Undifferentiated spermatogonia include two different subsets of cells: SSCs and progenitors. Additionally, the undifferentiated spermatogonial population can be categorized into $A_{\text{single}} (A_s)$, $A_{\text{paired}} (A_{\text{pr}})$, and $A_{\text{aligned}} (A_{\text{al}})$. SSCs are found within a subset of $A_s$ spermatogonia and they self-renew to maintain a foundational pool of cells that give rise to the progenitor cell population. Progenitors transiently amplify via incomplete mitosis to form $A_{\text{pr}}$ and $A_{\text{al}}$ spermatogonia which are connected by intercellular bridges. Upon exposure to retinoic acid, progenitors transition into a differentiating state without undergoing cell division. Differentiating spermatogonia begin at $A_1$ stage and proceed through a series of mitotic division. This process subsequently produce $A_2$, $A_3$, $A_4$, intermediate, and $B$ spermatogonia.
Figure 3. Irreversible and reversible SSC models. (Top) Diagram shows that SSC transition to progenitor state is an irreversible process. (Bottom) Diagram shows that $A_{pr}$ and $A_{al}$ progenitors could regain SSC function by losing their intercellular bridge to form $A_s$ that can revert back to SSC state.
Figure 4. Ret, Gfra1, and Gdnf are important for spermatogenesis in mice. Representative images of hematoxylin and eosin stained mouse testes cross-section from (Top left) wild type, (Top right) Ret germ cell specific ablation, (Bottom left) Gfra1 germ cell specific ablation, and (Bottom right) Gdnf germ cell-specific ablation.
Figure 5. *Plzf* is important for spermatogenesis in mice. (Top left) Diagram shows the role of PLZF in spermatogenesis of wild type mouse. (Top right) Representative image of hematoxylin and eosin stained testes cross-section from wild type mouse. (Bottom left) Diagram showing *Plzf* germ cell-specific ablation results in the depletion of the SSC pool. (Bottom right) Representative image of hematoxylin and eosin stained testes cross-section from mouse with *Plzf* germ cell-specific ablation.
Figure 6. GDNF signaling inhibits STAT3 phosphorylation to regulate Ngn3 transcription. (Left) Diagram shows that GDNF signaling regulates STAT3 phosphorylation to regulate Ngn3 transcription. (Middle) Diagram shows that GDNF withdrawal increase STAT3 phosphorylation and increase Ngn3 transcription thus decrease SSC abundance. (Left) Reduction of Ngn3 expression inhibits progenitor spermatogonia differentiation.
Figure 7. *Etv5* is important for spermatogenesis in both mice and me. Representative images of hematoxylin and eosin stained testes cross-sections from (Left) wild type and (Right) *Etv5* germ cell-specific ablation mice.
References


CHAPTER TWO

FUNCTIONAL AND MOLECULAR FEATURES OF THE ID4+ GERMLINE STEM CELL POPULATION IN MOUSE TESTES

FRIEDA K. Chan,1 Melissa J. Oatley,1 Amy V. Kaucher,1 Qi-En Yang,1 Charles J. Bieberich,2 Cooduvalli S. Shashikant,3 and Jon M. Oatley1,4

1School of Molecular Biosciences, Center for Reproductive Biology, College of Veterinary Medicine, Washington State University, Pullman, Washington 99164, USA; 2Department of Biological Sciences, University of Maryland at Baltimore County, Baltimore, Maryland 21250, USA; 3Department of Animal Science, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Abstract

The maintenance of cycling cell lineages relies on undifferentiated subpopulations consisting of stem and progenitor pools. Features that delineate these cell types are undefined for many lineages, including spermatogenesis, which is supported by an undifferentiated spermatogonial population. Here, we generated a transgenic mouse line in which spermatogonial stem cells are marked by expression of an inhibitor of differentiation 4 (Id4)-green fluorescent protein (Gfp) transgene. We found that ID4-GFP+ cells exist primarily as a subset of the type A single pool, and their frequency is greatest in neonatal development and then decreases in proportion during establishment of the spermatogenic lineage, eventually comprising ~2% of the undifferentiated spermatogonial population in adulthood. RNA sequencing analysis revealed that expression of 11 and 25 genes is unique for the ID4-GFP+/stem cell and ID4-GFP−/progenitor fractions,
respectively. Collectively, these findings provide the first definitive evidence that stem cells exist as a rare subset of the Asingle pool and reveal transcriptome features distinguishing stem cell and progenitor states within the mammalian male germline.

**Introduction**

The functionality of tissues is provided by differentiated cells that are continually lost as a result of terminal differentiation or cytotoxic insult. Constant cycling of cell lineages to provide differentiated cells relies on the actions of undifferentiated subpopulations consisting of stem cells and progenitors. The tissue-specific stem cells self-renew to maintain a constant pool from which progenitors periodically arise and transiently amplify in number before committing to a pathway of terminal differentiation. Thus, mechanisms controlling these actions are critical for tissue homeostasis.

Identification of characteristics that distinguish stem cells from progenitors is of paramount importance for understanding the regulatory mechanisms governing their actions. For example, identification of *Lgr5* expression in the intestinal epithelium has allowed for discovery of key mechanisms that sustain the crypt stem cell population (Barker, van Es et al. 2007). In addition, identification of *Sox2* expression in a multitude of cell lineages has allowed for comparison of conserved stem cell and progenitor properties (Arnold, Sarkar et al. 2011). Despite these advancements, characteristics that distinguish stem cells and progenitors are undefined for many cell lineages.

For the mammalian spermatogenic lineage, the actions of spermatogonial stem cells (SSCs) provide the basis for continual generation of millions of genetically unique spermatozoa daily (de Rooij and Russell 2000; Oatley and Brinster 2008; Oatley and Brinster 2012). In testes of rodents, germline regenerative activity is a property of the type
A undifferentiated spermatogonial population that consists of A\textsubscript{single} (individual cells), A\textsubscript{paired} (cohorts of two cells), and A\textsubscript{aligned} (cohorts of four, eight, and 16 cells) subtypes. Traditionally, A\textsubscript{single} spermatogonia have been considered the SSC population that self-renews to maintain a foundational pool from which A\textsubscript{paired} spermatogonia arise (Huckins 1971; de Rooij and Russell 2000; Oatley and Brinster 2012). Proliferation of A\textsubscript{paired} generates A\textsubscript{aligned} cohorts, and together these represent the progenitor spermatogonial population. During steady-state conditions, the majority of progenitor spermatogonia transition periodically into differentiating type A spermatogonia that are committed to terminal differentiation, eventually yielding spermatozoa. This transition occurs every 8.6d in mice to sustain continual spermatogenesis (de Rooij 1998). Due to staggered timing in the initiation of terminal differentiation, spermatogenesis occurs in stages that make up a cycle of the seminiferous epithelium. In mice, 12 stages constitute a spermatogenic cycle, which takes 35 d to complete (Oakberg 1956). The actions of SSCs and progenitors during the cycle provide the foundation for continuity of spermatogenesis and are required for male fertility. However, despite their importance, little is known about the defining characteristics of SSC and progenitor spermatogonial subtypes in the mammalian male germline.

A widely accepted model is that, upon division, A\textsubscript{single} spermatogonia have the capacity for two different fates. The cells can separate and form two A\textsubscript{single}, thereby undergoing symmetric self-renewal to expand the SSC pool, or remain connected by an intercellular bridge due to incomplete cytokinesis, thereby forming A\textsubscript{paired} spermatogonia (Oakberg 1971; de Rooij and Russell 2000; Oatley and Brinster 2012). In this model, formation of A\textsubscript{paired} spermatogonia represents the initial step in commitment to terminal differentiation.
Recent studies have begun to challenge this paradigm by suggesting that chained undifferentiated spermatogonia (Apaired and Aaligned) fragment to produce smaller chains and even single cells that contribute to the SSC pool (Nakagawa, Sharma et al. 2010). While these findings are based solely on time lapse imaging of select seminiferous tubules, they suggest that the stem cell population in mouse testes may be more flexible than the traditional Asingle model implies. However, whether the newly formed putative Asingle or smaller cohorts produced by fragmentation fulfill the definition of a stem cell remains to be clarified.

In the most basic sense, stem cells can be defined by a functional ability to regenerate and sustain a cycling cell lineage. Thus, the defining feature of SSCs is the capacity to regenerate and sustain spermatogenesis. In comparison, the progenitor population contributes to steady-state conditions but lacks regenerative capacity. Transplantation methodology is an unequivocal means to determine stem cell capacity of a cell population based on regenerative ability. Germ cell transplantation methodology is a well-established tool for mice to determine the SSC capacity of a cell population based on regenerated colonies of continual spermatogenesis (Brinster and Avarbock 1994; Brinster and Zimmermann 1994). Previous studies suggest that testes of adult mice contain ~3000 cells with regenerative capacity (Nagano 2003), yet the number of Asingle spermatogonia has been calculated to be ~35,000 (Tegelenbosch and de Rooij 1993); thus, SSCs likely represent a subset of Asingle spermatogonia. Importantly, preparation of testis cell suspensions for transplantation involves disassociation of spermatogonial chains to generate a single-cell suspension, but only a small portion (~3000) of the population has regenerative capacity. Thus, fragmentation in itself does not induce stem cell capacity for
the mouse male germ cell lineage. At present, distinguishing morphological (chain identity) or molecular (expression of specific proteins) characteristics of the SSC and progenitor spermatogonial populations are undefined, and the only unequivocal measure is a functional ability for regeneration of spermatogenesis after transplantation.

In previous studies, we found that expression of the transcriptional repressor inhibitor of differentiation 4 (Id4) is selective for A\textsubscript{single} cells within the spermatogonial population of mouse testes and plays an important role in the regulation of SSC self-renewal (Oatley, Kaucher et al. 2011). Here, we succeeded in labeling the SSC population specifically by generating a novel transgenic mouse model in which Id4-expressing cells are marked by green fluorescent protein (Gfp). Using this model, we discovered that regenerative capacity is restricted to the Id4-expressing subpopulation of undifferentiated spermatogonia, supporting the traditional model of A\textsubscript{single} representing the SSC pool. In addition, we describe for the first time other distinguishing features of this population, including (1) relative abundance during neonatal development and adulthood, (2) localization within regions of seminiferous tubules, and (3) unique aspects of the transcriptome.

**Results**

**Generation of an Id4-Gfp reporter mouse line**

In previous studies, we used immunostaining to localize ID4 in testes of mice, which revealed selective expression by rare A\textsubscript{single} spermatogonia (Oatley, Kaucher et al. 2011). To label this population as live cells for in-depth analysis, we created a transgenic mouse line in which a Gfp reporter transgene is driven by Id4 regulatory elements (Fig. 1A). A 17-kb fragment containing all exons and introns of the mouse Id4 gene, including 59 (7512-base-}
pair [bp]) and 39 (9214-bp) flanking regions, was captured from a BAC clone. A yeast recombinase strategy was then used to insert an enhanced Gfp (eGfp)-Ura3 cassette in-frame of exon 1, deleting base pairs 13–29. The transgene construct was then used with pronuclear injection to generate founder males with an FVB genetic background. A PCR strategy was devised to identify mice containing the Id4-Gfp transgene (Fig. 1B). Thus, expression of Gfp faithfully represents ID4 protein expression. We identified five transgenic founder lines, and all possessed a similar pattern and level of Gfp expression in testes. One line was then chosen for expansion and backcrossed on the C57BL/6J genetic background to create a new line, referred to hereafter as LT-11B6 (Fig. 1A). Mice of this line containing the Id4-Gfp transgene are of normal viability and fertility, and transgene inheritance occurs at an expected Mendelian ratio. Expression of the transgene is detectable in numerous tissues, including the testes, brain, and kidney (Fig. 1C). Also, presence of the transgene does not result in excessive Id4 transcript abundance in testes of LT-11B6 mice (Fig. 1D).

Expression of Id4-Gfp selectively labels A single cells of the spermatogonial population

In mammalian testes, the cycling spermatogenic lineage is established during a defined period of postnatal development, encompassing 0 and 35 d of age in mice (Drumond, Meistrich et al. 2011). At postnatal day (PD) 0, the germ cell population consists of quiescent gonocyte precursors that were formed during late embryonic development. The undifferentiated spermatogonial population materializes from these precursors during PD 3–8. In adulthood, at PD 35 and beyond, dynamics of the undifferentiated spermatogonial population are identical to that established during PD 3–8 (Drumond, Meistrich et al. 2011). To assess expression of Id4-Gfp in the germline during postnatal development, we used immunostaining of testis cross-sections with an antibody
recognizing GFP and whole-mount imaging for GFP in live tissue. Immunostaining of cross sections revealed expression in most gonocytes at PD 0 and then a subset of type A spermatogonia at PD 3, 6, 8, 10, and 12 (Fig. 2A). In addition, Gfp expression was detected in both a subset of type A spermatogonia and some pachytene spermatocytes at PD 20 and 35 but not in other spermatogonial subtypes, other spermatocytes, or spermatids. Next, we aimed to determine whether Id4-Gfp expression within the spermatogonial population is localized to A single and/or chains of A paired/A aligned. To accomplish this, whole-mount imaging of dissected seminiferous tubules for live GFP+ cells was performed at PD 6 and 35. At both ages, individual GFP+ cells were observed at the periphery of seminiferous tubules (Fig. 2B). Scanning along the length of a tubule revealed numerous ID4-GFP + cells at PD 6, but the occurrence declined upon aging, with only a few cells per seminiferous tubule fragment observed at PD 35. Also, in rare incidences (four observations out of >300 Gfp+ cells), a cohort of two ID4-GFP + cells was observed (Supplemental Fig. S1). Whether these were true A paired or false pairs sometimes observed when A single have not separated following division could not be determined. Taken together, these findings indicate that expression of Id4 is restricted to some, if not all, A single spermatogonia within the undifferentiated spermatogonial population of mouse testes.

**Id4-expressing cells represent a subset of the A single spermatogonial population**

To explore further the expression of Id4-Gfp within the undifferentiated spermatogonial population specifically, we used coimmunofluorescent staining for the pan undifferentiated marker PLZF and GFP (Fig. 3A). While these analyses revealed overlap in expression of both markers, only a subset of the PLZF+ population was ID4-GFP + at all ages examined (Fig. 3B). At PD 3, when gonocyte precursors have transitioned to a
postnatal spermatogonial state, the ID4-GFP + cells comprised 92.9% ± 0.5% of the PLZF+ undifferentiated spermatogonial population, decreasing to 46.1% ± 1.1% at PD 6, 19.6% ± 0.6% at PD 12, and 8.5% ± 0.4% at PD 20 (n = 3 different mice and 50–60 tubules for each age). In adulthood, the ID4-GFP + population comprised only 1.9% ± 0.3% of the PLZF+ undifferentiated spermatogonial population (n = 3 different mice and 50 tubules). Overlap of staining for markers of differentiating spermatogonia, including KIT and STRA8, was not observed at any age examined (Supplemental Fig. S2), confirming that expression of Id4 is restricted to spermatogonia in the undifferentiated state of development. Next, we examined ID4-GFP + cells within the GFRa1-expressing undifferentiated spermatogonial population, which has been reported to be restricted to A_{single}, A_{paired}, and short chains (i.e., four cell cohorts) of A_{aligned} spermatogonial subsets (Suzuki, Sada et al. 2009; Nakagawa, Sharma et al. 2010). Results of coimmunofluorescent staining revealed that, beginning at PD 6, the ID4-GFP + population represents a subset of individual cells within the more abundant Gfra1+ spermatogonial population (Fig. 3C). Taken together, these findings demonstrate that the A_{single} spermatogonial population is heterogeneous, supporting a notion that expression of Id4 selectively marks the SSC pool.

Relative abundance and distribution of ID4-GFP + spermatogonia in mouse testes during postnatal development and adulthood

Next, we determined changes in the relative abundance of ID4-GFP + spermatogonia in testes during postnatal development and their distribution within seminiferous tubules during steady-state spermatogenesis in adulthood. To achieve this, immunostaining for GFP within cross-sections of testes was used, and the number of ID4-GFP + cells was quantified (Fig. 4A). At PD 0 and 3, most seminiferous tubule cross-sections were found to
contain Id4-Gfp+ spermatogonia at a rate of 98.9% ± 1.1% and 88.1% ± 1.8%, respectively. During advancing postnatal development, the frequency decreased to 70.8% ± 4.8%, 61.6% ± 9.6%, 29.9% ± 7.8%, and 6.8% ± 4.4% of seminiferous tubule cross-sections containing ID4-GFP + spermatogonia at PD 6, 12, 20, and 35, respectively. Interestingly, the number of ID4-GFP + spermatogonia within tubule cross-sections was relatively constant at all postnatal ages examined, with 2.5 ± 0.2, 2.7 ± 0.3, 2.0 ± 0.1, 1.7 ± 0.2, 1.3 ± 0.1, and 1.1 ± 0.1 positive spermatogonia at PD 3, 6, 12, 20, and 35, respectively (Fig. 4B). These findings indicate that the ID4-GFP + spermatogonial population is of greatest abundance in neonatal life, decreases in both frequency and number as the spermatogenic lineage becomes established, and is then maintained in a cycling manner during adulthood.

To assess further the ID4-GFP + spermatogonial population during steady-state conditions in adulthood, we examined the relative abundance of these cells at different stages of the seminiferous epithelial cycle. In mammalian testes, spermatogenesis occurs in a staggered manner along the length of seminiferous tubules to provide constant generation of spermatozoa. This staggering is reflected by different stages of germ cell development, and all stages collectively constitute a cycle of the seminiferous epithelium. At present, the behavior of SSCs during these stages is undefined. While there are 12 clearly defined stages in testes of adult mice, accurate identification of each relies on fixation and staining of nuclear material that is not compatible with immunostaining. Thus, to improve accuracy, we evaluated stages as early (II–VI), middle (VII–VIII), and late (IX–I) groupings (Fig. 4C). Results of these analyses revealed even distribution of the ID4-GFP+ spermatogonial population across the groupings (Fig. 4D). These findings suggest that the
SSC pool is distributed equally during steady-state conditions of a constant cycling spermatogenic lineage.

Last, we aimed to determine whether ID4-GFP+ spermatogonia are localized preferentially within select regions of seminiferous tubules of adult testes during steady-state spermatogenesis. Results of previous studies suggest that undifferentiated spermatogonia and possibly SSCs are positioned preferentially within regions juxtaposed to interstitial tissue and the vasculature (Chiarini-Garcia, Hornick et al. 2001; Nakagawa, Nabeshima et al. 2007). Quantification of Id4-Gfp+ spermatogonia in cross-sections of seminiferous tubules revealed that 49.6% ± 0.9% and 46.9% ± 1.3% (n = 3 different mice and 20 cross-sections) of the population was located in regions adjacent to other tubules or interstitial tissue not containing blood vessels, respectively (Fig. 4E; Supplemental Fig. S3). Importantly, only 3.6% ± 0.6% of the Id4-Gfp+ spermatogonial population was observed in tubular regions juxtaposed to interstitial tissue containing blood vessels (Fig. 4E; Supplemental Fig. S3). These findings suggest that the SSC pool in testes of adult mice is distributed randomly in regions of seminiferous tubules that are avascular.

**Stem cell capacity of the undifferentiated spermatogonial population resides in the ID4-GFP+ fraction**

While expression of Id4-Gfp by a subset of A single spermatogonia is indicative of stem cell specificity, regenerative capacity for derivation of the cycling spermatogenic lineage is the only unequivocal measure. Thus, we isolated ID4-GFP+ cells from testes of adult mice (2 mo of age) using fluorescence-activated cell sorting (FACS) followed by transplantation into seminiferous tubules of sterile W/Wv recipient mice. Examination of cross-sections from recipient testes 3 mo after transplantation revealed seminiferous tubules containing
normal spermatogenesis (Supplemental Fig. S4). Because W/Wv mice lack endogenous germ cells (Handel and Eppig 1979; Brinster and Zimmermann 1994), the spermatogenesis must have derived from transplanted donor cells. These findings demonstrated that at least some ID4-GFP+ spermatogonia function as SSCs.

Next, we assayed whether regenerative capacity in the undifferentiated spermatogonial population resides with the ID4-GFP + fraction, the ID4-GFP - fraction, or both. Using the LT-11B6 line, this comparison could not be made by isolation of cells from testes directly because the ID4-GFP - fraction was a mix of several germ cell and somatic cell types. Thus, to examine fractions in the context of the undifferentiated spermatogonial population specifically, primary cultures established from F1 double-transgenic Id4-Gfp;RosaLacZ hybrid mice were used (Fig. 5A). Primary spermatogonial cultures are comprised of SSCs with regenerative capacity (~20% of the population) and non-stem cell progenitors (Kubota, Avarbock et al. 2004; Oatley, Avarbock et al. 2006). Also, all cells of the hybrid donors possessed the LacZ transgene, including the various subtypes of germ cells, but the Gfp transgene was expressed selectively in ID4+ cells. Thus, colonies of donor-derived spermatogenesis could be detected in testes of recipient mice after transplantation by incubation with the substrate X-gal, which stained the donor cells blue. First, examination of the cultures using live imaging revealed that the GFP+ cells comprised a subset of the population (Fig. 5A). Quantification using flow cytometric analysis revealed that 18.7% ± 5.3% (n = 3 different cultures) of the population was GFP+. Second, the GFP+ and GFP-fractions were isolated using FACS and transplanted as separate cell populations into seminiferous tubules of 129XC57 recipient mice pretreated with busulfan to eliminate the endogenous germline. The FACS approach allowed for isolating fractions that were
highly enriched for the cell type of interest (Fig. 5B); however, some ID4-GFP+ cells were still present in the ID4-GFP- fraction and vice versa. Examination of the recipient testes 2 mo after transplantation for colonies of donor derived spermatogenesis by X-gal staining revealed extensive colonization by the ID4-GFP+ cell population in all testes (Fig. 5C). In contrast, the ID4-GFP- fraction generated only a small number of colonies in some recipient testes (Fig. 5C). In total, the ID4-GFP+ population generated 516.6 ± 98.1 colonies per 105 cells transplanted, whereas, the ID4-GFP- population produced 35.6 ± 7.9 colonies per 105 cells transplanted (Fig. 5D). Thus, ~95% of the SSC pool was captured in the FACS-isolated ID4-GFP+ fraction, and ~5% of the total SSC population was present in the ID4-GFP- fraction. The possibility that colonies from the ID4-GFP- fraction were produced from the contaminating ID4-GFP+ cells could not be ruled out. However, the extent of colonization by ID4-GFP+ cells suggests potent SSC capacity. Collectively, these findings indicate that most, if not all, SSCs of the undifferentiated spermatogonial population are present in the ID4-GFP+ fraction, and the ID4-GFP- fraction represents progenitors that lack stem cell capacity.

**Unique transcriptome features of the ID4-GFP+ spermatogonial population**

A major bottleneck in defining stem cell and progenitor states in the mammalian spermatogonial lineage has been an inability to examine the cell populations separately. Based on results of our transplantation analyses, we were confident that a means had been devised to overcome this limitation using FACS separation of ID4-GFP+ and ID4-GFP- populations from primary cultures of undifferentiated spermatogonia. Using this approach, we aimed to define the mRNA transcriptome of SSC and progenitor spermatogonial populations with RNA sequencing (RNA-seq) technology. Matched samples of RNA were
collected from FACS-isolated Id4-Gfp+ and Id4-Gfp- populations (n = 3 different cultures) and cDNA libraries generated from polyA mRNA. The Illumina HiSeq system was used to generate 24 million to 30 million paired-end reads of 100 bp in length followed by mapping to the mouse genome (mm9 build) with the TopHat alignment tool (Trapnell, Roberts et al. 2012). For each sample, transcript assembly of confidently mapped reads was performed with the Cufflinks bioinformatics program, and fragments per kilobase of transcript per million mapped reads (FPKM) values were calculated to provide a quantitative value of abundance using the Cuffdiff program (Trapnell, Roberts et al. 2012). Differential abundance of transcripts between ID4-GFP+ and ID4-GFP- populations was then determined statistically with the Cuffdiff program. The outcome of these analyses revealed that most genes are not differentially expressed between the two cell populations (Supplemental Data Set S1). Examination of genes known to be essential for primordial germ cell specification and germ cell licensing (Lesch and Page 2012) revealed no difference in expression between the two cell populations (Fig. 6A). Moreover, examination of a specific subset of genes previously reported to be expressed by the undifferentiated spermatogonial population revealed that transcript abundance for only Id4 and Utf1 was significantly different between the two populations (Fig. 6A). In total, 11 transcripts were determined to be of significantly (Q-value < 0.05) greater abundance in the Id4-Gfp+ population compared with the Id4-Gfp- population (Fig. 6B). Also, the abundance of 25 transcripts was significantly greater in the ID4-GFP- population compared with the Id4-Gfp+ population (Fig. 6B). Last, we used quantitative RT–PCR (qRT–PCR) analysis to validate differential abundance of the 11 transcripts in the ID4-GFP+ cell population (Supplemental Fig. S5). In all cases, outcomes of both experimental approaches were in
agreement. Collectively, these findings reveal attributes of the mRNA transcriptome that distinguish the stem cell and progenitor states in the mammalian spermatogonial lineage.

**Discussion**

The undifferentiated spermatogonial population in mammalian testes is heterogeneous, containing both stem cells and transiently amplifying progenitors (de Rooij and Russell 2000; Oatley and Brinster 2012; Griswold and Oatley 2013). For decades, the only means to unequivocally identify the SSC component has been retrospective determination of regenerative capacity following transplantation (Brinster and Zimmermann 1994; Brinster and Nagano 1998; Brinster 2002; Brinster 2007; Oatley and Brinster 2008; Oatley and Brinster 2012). Until now, molecular markers that distinguish SSC and progenitor spermatogonial subsets were undefined. Thus, studying the SSC population specifically to define molecular mechanisms that influence fate decisions and characteristics of these cells in vivo or during in vitro maintenance has been challenging. For the first time, the SSC population can be examined specifically in a range of in vivo conditions, including postnatal development of the spermatogenic lineage, steady-state conditions in adulthood, and regeneration of spermatogenesis following cytotoxic damage or transplantation, by using the LT-11B6/Id4-Gfp mouse model. Furthermore, the relationship between SSCs and somatic support cells can be investigated specifically rather than relying on studying the heterogeneous undifferentiated spermatogonial population as a whole. Moreover, the in vitro behavior of SSC and progenitor populations can be compared with defined proliferation kinetics and responses to a variety of environmental conditions.
A limitation of the LT-11B6/Id4-Gfp mouse model is apparent expression of the transgene in pachytene spermatocytes beginning at PD 20 and persisting into adulthood. Expression in other germ cell types, including differentiating spermatogonia, other spermatocytes, and round/elongate spermatids, was not observed. The expression in pachytene spermatocytes may be nonspecific. Indeed, expression of reporter transgenes consisting of regulatory elements from genes normally not expressed in spermatocytes or even the male germ cell lineage in postnatal life, including Nanog, Oct4, and Blimp-1, is observed at the pachytene stage of spermatocyte development (Wang, Zhuang et al. 2008; Kuijk, de Gier et al. 2010). Furthermore, immunostaining analysis for expression of endogenous ID4 in previous studies did not reveal the presence of the protein in pachytene spermatocytes (Oatley, Kaucher et al. 2011). Regardless, findings of the current study demonstrate that expression of the Id4-Gfp transgene in the undifferentiated spermatogonial population specifically is restricted to a rare subset of cells. Importantly, the ID4-GFP+ spermatogonia possess regenerative capacity, thereby classifying the population as bona fide male germline stem cells.

Results of quantitative analyses using cross-sections of seminiferous tubules and imaging of whole-mount preparations indicate that ID4-GFP+ spermatogonia represent a subset of the A single spermatogonial population. During postnatal development, the proportion of the undifferentiated spermatogonial population characterized as ID4-GFP+ was found to decline to the point of being a rare subset of cells in adulthood. Also, ID4-GFP+ spermatogonia were found to represent a subset of the GFRα1+ pool from PD 6 to adulthood. These findings confirm those of previous studies indicating heterogeneity of the Asingle spermatogonial population in testes of mice (Suzuki, Sada et al. 2009; Zheng, Wu et
Taken together, these observations suggest that the bona fide stem cell population of the mammalian male spermatogenic lineage is a subset of the A<sub>single</sub> spermatogonial population. Moreover, these findings indicate that a hierarchy of A<sub>single</sub> spermatogonia exists (Fig. 7). We propose a model in which a subset of the A<sub>single</sub> pool is self-renewing SSCs, and a second subset are progenitors poised to become A<sub>paired</sub> spermatogonia upon the next cell division. Importantly, these subsets can be distinguished based on expression of Id4 and possibly other unique markers that await validation.

During steady-state conditions in adulthood, the entire A<sub>single</sub> population must be replenished throughout a cycle of the seminiferous epithelium. Considering that there are ~35,000 A<sub>single</sub> in testes of adult mice (Tegelenbosch and de Rooij 1993), but the regenerative SSC population has been estimated to be only ~3000 in number (Nagano 2003), homeostasis of the pool requires vigorous regulation of division kinetics. In general, the undifferentiated spermatogonial population is estimated to undergo two to three doublings during a cycle of the seminiferous epithelium (Lok and de Rooij 1983). Obviously, this level of division by a small number of SSCs would not be sufficient to replenish the entire A<sub>single</sub> pool, and there are likely several possible mechanisms that could provide the infrastructure for replenishment. One possibility is that some A<sub>single</sub> progenitors (i.e., ID4- spermatogonia) retain limited self-renewal potential similar to short-term repopulating cells of the hematopoietic lineage but lack regenerative capacity, whereas most of the population forms A<sub>paired</sub> spermatogonia upon the next division (Fig. 7). Another potential mechanism is that the regenerative SSC pool (i.e., ID4+ spermatogonia) is larger than has been estimated in previous studies, thus comprising >10% of the A<sub>single</sub> population. For both of these scenarios, two to three divisions during a cycle of the
spermatozoa. Clearly, the cell division kinetics and heterogeneous nature of the A\textsubscript{single} population is quite dynamic and warrants further investigation. The molecular phenotypes that distinguish SSC and progenitor spermatogonial subtypes have been difficult to define due to the lack of conclusive markers and the apparent heterogeneity in A\textsubscript{single}, A\textsubscript{paired}, and A\textsubscript{aligned} pools. For example, expression of Neurog3 in spermatogonia is associated with the progenitor state (Nakagawa, Nabeshima et al. 2007; Kaucher, Oatley et al. 2012), but a subset of Neurog3+ spermatogonia possess SSC capacity (Nakagawa, Nabeshima et al. 2007). In addition, expression of Nanos2 and Gfra1 appears to be heterogeneous in the A\textsubscript{paired} and A\textsubscript{aligned} spermatogonial pools (Suzuki, Sada et al. 2009). Furthermore, until now, a molecular marker of A\textsubscript{single} spermatogonia only had not been described. The results of transcriptome analyses in the current study indicate that the SSC pool possesses a fixed and distinguishable molecular phenotype. This signature includes expression of many markers that are shared by progenitor spermatogonia, such as Neurog3, Nanos2, Plzf, Lin28, Sall4, Foxo1, and Gfra1, but also expression of distinguishing markers such as Id4 and Utf1 (Fig. 7). Indeed, while not yet characterized in mice, the expression of Utf1 is restricted to a rare subset of undifferentiated spermatogonia in rat testes, similar to Id4 expression in the spermatogonial lineage of mouse testes.

A defining feature of stem cells is the capacity for regeneration of a cycling cell lineage. For the male germline, stem cells are defined functionally by the ability to regenerate and sustain spermatogenesis. An unequivocal measure of this ability is generation of donor-derived colonies of spermatogenesis in testes of recipient mice following transplantation (Brinster and Avarbock 1994; Brinster and Zimmermann 1994).
In the present study, a major goal was to assess the regenerative capacity of ID4-GFP+ spermatogonia to determine whether these cells fulfill the defining criteria of an SSC. Results of transplanting ID4-GFP+ cells isolated from donor testes directly revealed this capacity. Furthermore, results of transplantation analysis with sorted ID4-GFP+ and ID4-GFP- fractions from primary cultures of undifferentiated spermatogonia demonstrated that most, if not all, regenerative capacity resides within the ID4-GFP+ cells. Considering that all cells in primary cultures express markers of undifferentiated spermatogonia, but only a subset possesses regenerative capacity, these findings indicate the ID4-GFP+ population represents the SSC pool, whereas the ID4-GFP- population represents transiently amplifying progenitors. A recent study suggests that the colonizing ability of SSCs in testes of adult recipient mice but not immature pup recipients varies based on the phase of the cell cycle (Ishii, Kanatsu-Shinohara et al. 2014). These findings indicate that cell cycle status does not alter the core ability of an SSC to regenerate spermatogenesis but does impact the efficiency of establishing an SSC niche unit via interaction with mature somatic support cells (i.e., Sertoli cells). It is possible that expression of the Id4-Gfp transgene varies by phase of the cell cycle, thereby influencing regenerative capacity following transplantation into testes of adult recipient mice. Regardless, the outcomes of transplantation analyses in the present study demonstrate that expression of Id4-Gfp marks spermatogonia that possess the core ability to function as an SSC.

Stem cells in most, if not all, tissues are thought to reside within a specialized microenvironment that influences their activities (Oatley and Brinster 2012). A stem cell niche in mammalian testes has been difficult to define due to the tubular configuration and lack of a polarized nature for the seminiferous epithelium in which SSCs reside as well as
undefined markers that distinguish SSCs from other undifferentiated germ cell types. Results of several previous studies suggest that SSCs reside within regions of seminiferous tubules juxtaposed to interstitial tissue and blood vessels (Chiarini-Garcia, Hornick et al. 2001; Chiarini-Garcia, Raymer et al. 2003; Nakagawa, Nabeshima et al. 2007). These observations have led to the postulation that the SSC niche is associated with vascular architecture (Nakagawa, Nabeshima et al. 2007). However, it is important to note that these observations were made by investigating the heterogeneous undifferentiated spermatogonial population as a whole (Chiarini-Garcia, Hornick et al. 2001; Chiarini-Garcia, Raymer et al. 2003) or examining a subpopulation marked by expression of a transgene predominately associated with progenitor spermatogonia (Nakagawa, Nabeshima et al. 2007). Results of our previous studies revealed that the overall number of a particular testis somatic cell population, Sertoli cells, is correlated to the number of niches accessible for colonization by transplanted SSCs (Oatley, Racicot et al. 2011). Interestingly, we found that increased niche accessibility did not coincide with a change in seminiferous tubular area associated with the vasculature or interstitial tissue (Oatley, Racicot et al. 2011). In the present study, we aimed to address this discrepancy by determining whether ID4-GFP+ spermatogonia are localized preferentially in certain regions of seminiferous tubules. Results from scoring of these cells within cross sections from testes of adult mice indicated no bias for regions adjacent to other tubules or juxtaposed to interstitial tissue. However, assessment of the ID4-GFP+ spermatogonia residing in tubule regions juxtaposed to interstitial tissue specifically revealed limited association with areas containing blood vessels. Based on these observations, we postulate that the SSC niche is likely avascular.
Many putative markers of SSCs have been described in recent years, but all are expressed by multiple subsets of spermatogonia, and some are even expressed by testicular somatic cells. Until now, a definitive marker that is expressed selectively by a single and labels a spermatogonial population that contains most, if not all, of the germline regenerative cells has been lacking. At present, the LT-11B6 mouse model developed in this study allows for arguably the most accurate investigation of SSCs. Using this model, we made seminal observations of defining features for the regenerative spermatogonial population. These include labeling by \textit{Id4-Gfp} expression, existing predominantly as single spermatogonia, comprising a rare subset of the undifferentiated spermatogonial population in adulthood, localizing in avascular regions of seminiferous tubules, and possessing a unique transcriptome defined by differential expression of a small subset of genes. In future studies, utilization of the LT-11B6 mouse model will undoubtedly reveal many more features of the SSC population in mammalian testes. In addition, \textit{Id4} is known to be expressed by stem cells of other lineages (Yun, Mantani et al. 2004; Jeon, Jin et al. 2008; Du and Yip 2011). Thus, comparative analysis to define commonalities among various tissue-specific stem cell populations and differences that dictate lineage specificity is likely possible using the LT-11B6 mouse model.

\textbf{Materials and methods}

\textit{RNA-seq analyses}

Total cellular RNA from FACS-isolated ID4-GFP+ and ID4-GFP- spermatogonial populations of three different primary cultures was isolated using a combination of Trizol reagent (Invitrogen) and RNeasy columns (Qiagen) as described previously (Oatley et al. 2006, 2009). Libraries of cDNA were then generated using oligo d(T) priming. The samples were
subjected to Illumina HiSeq2000 analysis (Otogenetics, Inc.), which generated ~24 million to 30 million paired-end reads of 100 bp in length for each sample. Sequencing reads were mapped against the mouse genome (mm9 build) using TopHat version 2.0.5. Confidently mapped reads for each sample were analyzed by Cufflinks.Cuffdiff for transcript assembly and generation of a FPKM value for each transcript that was directly proportional to abundance. Significant differences in FPKM values for individual transcripts between Id4-Gfp+ and Id4-Gfp- populations were determined statistically using Cuffdiff. A Q-value of <0.05 was considered significantly different. The entire data set can be accessed through the NCBI-supported Sequence Read Archive (SRA) database, accession SRR1291311.

**Transgenic mouse production**

A mouse Id4-Gfp reporter gene construct was generated using a yeast-based recombineering strategy described previously (Bentley, Shashikant et al. 2010). Briefly, a sequence comparison of the human and mouse Id4 locus was carried out using VISTA plot analysis to identify conserved regions surrounding the Id4 coding region. A 17-kb genomic fragment containing all exons, introns, and 59 (7512-bp) and 39 (9412-bp) flanking regions of high similarity (~80%) was identified. The similarity tapered off on either side of the locus, suggesting that major regulatory elements controlling Id4 expression were present within this genomic fragment. A 192-kb BAC clone (RP32 344L21) containing the mouse Id4 locus was purchased from the BACPAC consortium. Two 100-bp double-stranded oligonucleotides were designed to capture 17 kb of the Id4 genomic fragment (Supplemental Table S3). The oligonucleotides contained regions of homology with linearized pClasperA vector and 59 and 39 regions of the 17-kb Id4 genomic fragment to be captured. The clone pClasper-Id4 was further modified to insert an eGfp-Ura3 cassette in-
frame of the Id4 coding region by recombineering in yeast. The resulting \textit{Id4-Gfp} genomic fragment was used to generate transgenic mice in the FVB strain as described previously (Shashikant, Bieberich et al. 1995). Mice carrying the \textit{Id4-Gfp} transgene were identified by Southern hybridization performed on genomic DNA isolated from tail biopsies. Five founder lines carrying transgenes were identified, and males from F1 generation were examined for \textit{Gfp} expression in the testis. All five lines showed clear, bright GFP activity, and one was chosen for crossing with C57BL/6J females. Backcrossing of hybrid animals containing the Id4-Gfp transgene with C57BL/6J mice was then carried out for another six generations to produce a line denoted as LT-11B6. A PCR-based strategy was then designed to more readily detect transgenic mice.

\textbf{Animals}

All animal procedures were approved by the Washington State University Institutional Animal Care and Use Committee (IACUC). \textit{Id4-Gfp} transgenic founder mice were produced on a FVB background and then backcrossed onto a C57BL6/J background. Mice with ubiquitous expression of a \textit{LacZ} transgene from the Rosa26 locus (B6;129S-Gt (ROSA)26Sor/J) were obtained from Jackson Laboratories. Recipients for spermatogonial transplantation were C57BL/6J × 129S1/svImJ F1 hybrids or \textit{W/Wv} mutant mice.

\textbf{Fluorescence activated cell sorting (FACS)}

Single cell suspensions generated from testes of adult LT-11B6 mice or primary cultures of spermatogonia established from LT-11B6/RosaLacZ mice were subjected to sorting based on \textit{Gfp} intensity using a BD FACS Aria instrument. Both GFP+ and GFP- populations were collected followed by washing in mSFM using centrifugation at 600xg for 6 min. Cells were
then suspended in mSFM for transplantation into testes of recipient mice or processed for RNA isolation in order to conduct RNA-Seq analysis.

**Immunostaining of Testis Cross-Sections**

Testes from LT-11B6 mice were fixed in Bouin’s solution for 2 hr (PD 0-6 donors) or overnight (PD 10-35 donors) at 40°C followed by dehydration in a graded series of ethanol washes and embedded in paraffin. Cross-sections of 5-μm thickness were adhered to glass slides followed by de-paraffinization and rehydration. Antigen retrieval was achieved by incubation in boiling sodium citrate buffer (10mM citric acid, 0.05% Tween-20, pH 6.0) for 20 min followed by rinsing in PBS. Non-specific antibody binding was blocked by incubation in PBS containing 0.5% bovine serum albumin, 0.1% Triton X-100 and 10% normal serum from host species of the secondary antibody for 1 hr at room temperature. Sections were then incubated with primary antibodies diluted in binding buffer (PBS containing 0.5% BSA and 0.1% Triton X-100) for 2 hr at room temperature followed by extensive washing in PBS and incubation with secondary antibodies for 2 hr at room temperature. For colorimetric staining, sections were again washed in PBS and developed with a horseradish peroxidase (HRP)-conjugated streptavidin kit (Vector Labs) followed by hematoxylin counterstaining. For immunofluorescent staining, sections were washed in PBS and coverslips mounted with Pro Long Gold antifade reagent containing DAPI (Life Technologies, OR). Primary and secondary antibodies utilized in this study and working dilutions are listed in Table S1.
Table S1. Antibodies used for immunostaining analyses.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Source</th>
<th>Working Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goat anti-mouse GFR alpha1</td>
<td>R&amp;D Systems (AF560)</td>
<td>1:400</td>
</tr>
<tr>
<td>FITC conjugated goat anti-GFP</td>
<td>Abcam (ab6662)</td>
<td>1:300</td>
</tr>
<tr>
<td>Rabbit anti-GFP</td>
<td>Abcam (ab290)</td>
<td>1:400</td>
</tr>
<tr>
<td>Rabbit anti-mouse PLZF</td>
<td>Santa Cruz (H-300)</td>
<td>1:50</td>
</tr>
<tr>
<td>Alexa Fluor 555 Donkey Anti-Goat IgG</td>
<td>Life Technologies (A21432)</td>
<td>1:1000</td>
</tr>
<tr>
<td>Alexa Fluor 546 Donkey Anti-Rabbit IgG</td>
<td>Life Technologies (A10040)</td>
<td>1:1000</td>
</tr>
</tbody>
</table>

Immunofluorescent Staining of Whole Mount Seminiferous Tubules

Testes from LT-11B6 mice were collected in ice cold PBS and the tunica albuginea removed followed by gently teasing apart of seminiferous tubules with fine forceps. Samples were then fixed in 4% PFA at 4°C for 2 hr, briefly dried at room temperature for 5 min, and washed in PBS containing 0.2% NP40 for 20 min at 4°C. Tubules were then dehydrated in a graded series of methanol washes for 5 min each at 4°C. For immunostaining, tubules were rehydrated in PBST for 5 min at 4°C and non-specific antibody binding was blocked by incubating in blocking buffer (PBST containing 1% BSA and 10% normal serum of the host species for the secondary antibody) for 1 hr at 4°C. Tubules were then incubated with primary antibodies diluted in blocking buffer overnight at 4°C followed by washing extensive washing in PBST and incubated with AlexaFluor-conjugated secondary antibodies diluted in blocking buffer overnight at 4°C. Samples were washed again in PBST at 4°C before spreading on glass microscope slides and mounting of coverslips with ProLong Gold antifade reagent containing DAPI. Primary and secondary antibodies utilized in this study and working dilutions are listed in Table S1.
Microscopic Imaging

Fluorescent microscopy and digital images were captured using an Olympus IX51 inverted microscope and DP72 digital color microscope camera (Olympus Inc., USA). All images were captured using CellSense acquisition software (Olympus Inc., USA). For quantification of cell number within cross-sections of testes, at least 50 seminiferous tubules in at least 5 different cross-sections were scored for various types of immunostained cells.

Primary Cultures of Undifferentiated Spermatogonia

Cultures of undifferentiated spermatogonia were established from LT11-B6 or LT-11B/RosaLacZ donor mice at PD 6-8 as described previously (Kubota et al., 2004; Oatley and Brinster, 2006). Briefly, the Thy1+ testis cell fraction was isolated using magnetic activated cell sorting (MACS) and co-cultured with mitotically inactivated STO feeders cell monolayers. Cultures were maintained in mSFM supplemented with 20 ng/ml recombinant human GDNF (PeproTech, NJ) and 1 ng/ml recombinant human FGF2 (BD Biosciences) at 37°C in an atmosphere of 5% CO2. At 7 day intervals, spermatogonial clumps were disassociated into single cell suspensions and sub-cultured at a ratio of 1:2 or 1:3 onto fresh STO feeders. All experiments utilizing primary cultures were conducted 1-3 months after initial establishment.

Quantitative RT-PCR Analysis

Total cellular RNA from FACS isolated ID4-GFP+ and ID4-GFP-spermatogonial populations of primary cultures was isolated using Trizol Reagent (Invitrogen, USA). Samples were subjected to DNaseI treatment and reverse transcription was carried out using Superscript III reverse transcriptase and oligo (d)T priming to generate cDNA. Both Taqman and Sybr green assays were conducted using validated primers and probes for Id4, Utf1, Eno1, Eno2,
Piwil4, Pde2a, Pak6, Nlrp4, Sgpp2, Chga, T, and Rps2 (see Supplemental Table S2). All reactions were performed with an ABI ABI 7500 Fast Sequence Detection system (Applied Biosystems). Abundance of transcript for the constitutively expressed gene Rps2 was used for normalization and to generate relative expression values for genes-of-interest in each sample using the $2^{-\Delta\Delta CT}$ formula as described previously (Oatley et al., 2011; Yang et al., 2013).

Table S2. Primers used for Sybr green based qRT-PCR analyses.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amplicon Size</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utf1</td>
<td>135</td>
<td>TGTCCCGGTGACTACGTCT</td>
<td>CCCAGAAGTAGCTCCGTCTCT</td>
</tr>
<tr>
<td>Eno2</td>
<td>8</td>
<td>AGGTGGATCTCTATACTGCAAAGT</td>
<td>CCCATCCCTTAGTTTCAG</td>
</tr>
<tr>
<td>Piwil4</td>
<td>101</td>
<td>GGGAAAAAGTTGACCCAGTAA</td>
<td>CCAATCTGCGGCAAATCTCAAAAT</td>
</tr>
<tr>
<td>Pde2a</td>
<td>128</td>
<td>TGGCGTTGTGACGATGAG</td>
<td>CGCGATAGAAACGGGATGG</td>
</tr>
<tr>
<td>Pak6</td>
<td>116</td>
<td>GTGGGGTGAAAGTTCGCAA</td>
<td>TTCTGTGTTCCTAGGGCTAGG</td>
</tr>
<tr>
<td>Nlrp4</td>
<td>130</td>
<td>CAGCAAGTGCGAGGAAGTTCT</td>
<td>ACTGAGGGAAACGTGGGACAG</td>
</tr>
<tr>
<td>Sgpp2</td>
<td>135</td>
<td>TCCACCCACTGGAATATGACC</td>
<td>AAGTCTCAACCGGAGGAAA</td>
</tr>
<tr>
<td>Chga</td>
<td>129</td>
<td>CCAAGGTGATGAAAGTCCGTC</td>
<td>GTGTGCGAGGAGAGAGAAGGA</td>
</tr>
<tr>
<td>T</td>
<td>117</td>
<td>GCTTCAAGGAGCTAACTAACGAG</td>
<td>CAGCAAGAAAGAGTACATGAGG</td>
</tr>
<tr>
<td>Eno1</td>
<td>118</td>
<td>TCGTCCACTGGCATCCTAC</td>
<td>CAGAGCAGGCGCAATAGTTTTA</td>
</tr>
</tbody>
</table>

**RNA-Seq Analyses**

Total cellular RNA from FACS isolated ID4-GFP+ and ID4-GFP- spermatogonial populations of 3 different primary cultures was isolated using a combination of Trizol reagent (Invitrogen, USA) and RNeasy columns (Qiagen, USA) as described previously (Oatley et al., 2006; Oatley et al., 2009). Libraries of cDNA were then generated using oligo d(T) priming to represent the mRNA transcriptome. The samples were then sent to Otogenetics Inc. for Illumina HiSeq2000 analysis which generated ~24-30 million pair-end reads of 100 bp in length for each sample. Sequencing reads were mapped against the mouse genome (mm9 build) using the Tophat v2.0.5 program. Confidently mapped reads for each sample were
analyzed by the Cufflinks.Cuffdiff program for transcript assembly and generation of a Fragments Per Kilobase per Million mapped (FPKM) value for each transcript which is directly proportional to abundance. Significant differences in FPKM values for individual transcripts between ID4-GFP+ and ID4-GFP- populations was determined statistically using the Cuffdiff program. A q-value of <0.05 was considered significantly different.

**Spermatogonial Transplantation**

To examine regenerative capacity of ID4-GFP+ and ID4-GFP- spermatogonial populations, cells were transplanted into the seminiferous tubules of immunologically compatible recipient mice utilizing methodology described previously (Oatley, Avarbock et al. 2006). Briefly, single cell suspensions were diluted in mSFM to a concentration of 1X106 cells/ml and ~10 μl was microinjected into the seminiferous tubules of adult recipient mice. Depending on the experiment, W/Wv mutant recipient mice that lack endogenous germ cells or 129XC57 F1 hybrid mice pretreated with busulfan (60 mg/kg of body weight) to eliminate the germline were used as recipients. For comparison of Id4-Gfp+ and Id4-Gfp- populations, one testis of each recipient received Id4-Gfp+ cells and the contralateral testis received Id4-Gfp- cells. Testes of recipient mice were then examined ~2-3 mo post-transplantation for colonies of donor-derived spermatogenesis. For testes of W/Wv recipients that were transplanted with cells from LT-11B6 mice, donor-derived spermatogenesis was determined in cross-sections by hematoxylin and eosin staining. For testes of 129XC57 mice that were transplanted with cells from LT-11B6/RosaLacZ cells, donor-derived spermatogenesis was determined by staining with X-Gal. A stereo zoom dissecting microscope was used to quantify LacZ stained donor colonies within each recipient testis.
Statistical Analyses

All quantitative data are presented as mean±SEM for at least 3 different biological replicate samples. Differences between means for data other than that generated by RNA-Seq analysis were determined using the general linear model one-way ANOVA function of SAS software. Multiple comparisons analysis was conducted using Tukey post hoc test.

Transgenic Mouse Production

A mouse Id4-Gfp reporter gene construct was generated using a yeast based recombineering strategy described previously (Bentley, Shashikant et al. 2010). Briefly, a sequence comparison of human and mouse Id4 locus was carried out using VISTA Plot analysis to identify conserved regions surrounding the Id4 coding region. A 17 Kb genomic fragment containing all exons, introns, 5’ (7,512 bp) and 3’ (9,412 bp) flanking regions contained regions of high similarity (~80%). This similarity tapered off on either side of the locus suggesting that major regulatory elements controlling Id4 expression were present within this genomic fragment. A 192 Kb BAC clone (RP32 344L21) containing mouse Id4 locus was purchased from the BACPAC consortium. To capture 17 Kb of the Id4 genomic fragment, two 100 bp double stranded oligonucleotides were designed as described previously (Bentley, Shashikant et al. 2010). Two 100 bp double stranded oligonucleotides were designed to capture 17 Kb of the Id4 genomic fragment (Table S3). The oligonucleotides contained regions of homology to linearized pClasperA vector and 5’ and 3’ regions of the 17 kb Id4 genomic fragment to be captured. Linearized pClasper, Id4-BAC, and the oligonucleotides were mixed and transformed into Saccharomyces cerevisiae strain Y274 (gift, Mike Snyder, Stanford University). The colonies containing recombinant Id4-ClasperA clones were selected on leucine dropout medium on SC agar. Genomic DNA
from transformant yeast colonies was isolated and tested for the presence of *Id4* genomic locus by PCR using primers specific to mouse *Id4*. DNA from two positive colonies were used to transform E. coli DH10B strain by electroporation. Colonies containing pClasperA-*Id4* clones were screened by PCR. Plasmid DNA was then prepared from positive clones and digested with several restriction endonucleases for analysis by gel electrophoresis. The restriction digestion pattern was compared with the predicted pattern from known genomic sequence. Both experimentally generated and predicted restriction digestion patterns were identical suggesting no significant alteration in the sequences during recombineering procedures.

The clone pClasper-*Id4* was further modified to insert an *eGfp-Ura3* cassette in-frame of the *Id4* coding region by a similar recombineering method. The recombinogenic cassette containing *Gfp-Ura3* was amplified by PCR (Advantage, Clontec kit) using primers that contained regions of homology to *Id4* regions flanking the designed point of insertion (Table S4). The yeast strain was transformed with pClasperA-*Id4* and PCR generated recombinogenic *Gfp-Ura3* and colonies containing recombinant clones were identified by selection of SC agar medium lacking in leucine and uracil. DNA from positive yeast colonies were shuttled to DH10B as described above. Positive clones were extensively analyzed by restriction digestions and sequencing across the point of insertion to confirm integrity of clones. DNA isolate from *Id4*-Gfp pClasperA vector was completely digested with I-Sce I and the insert fragment isolated free of the vector fragment using a sucrose density gradient method described previously (Shashikant, Bieberich et al. 1995). The resulting *Id4-Gfp* genomic fragment (5ng/ul) was used to generate transgenic mice in FVB strain as described previously (Shashikant, Bieberich et al. 1995). Mice carrying the *Id4-Gfp*
transgene were identified by Southern hybridization performed on genomic DNA isolated from tail biopsy. Five founder lines carrying transgenes were identified and males from F1 generation were examined for GFP expression in the testis. All five lines showed clear, bright GFP activity and one was chosen for crossing with C57BL/6J females. Backcrossing of hybrid animals containing the Id4-Gfp transgene with C57BL/6J mice was then carried out for another 6 generations to generate a line denoted as LT-11B6.

Table S3. Oligonucleotide sequences used for capture an Id4 genomic fragment and insertion into the vector pClasperA.

<table>
<thead>
<tr>
<th>Region</th>
<th>Nucleotide Sequence1</th>
</tr>
</thead>
</table>
| 5’ Capture Oligo 1 | 5’TATCCCTAGG CCACTCGAGGC CGGCCGCGCCA AGCTTTGTCGA CACGCGTTTCG CATGACTATTGTGTAAGAAGG TTAAAAAT
TGGCCTCTCTTCAAGCACAAG3’ |
| Oligo 2 | 5’CGTTGTCAATGACACATCTCAGACCTCAATTCTTCCATACCAAT AGTACATTTCGAACCGCGTGACGCAAAGGTGTTGGGC CGGCCGGCGCTTAAATTA
CGAGGACTAAGGTTGTTG3’ |
| 3’ Capture Oligo 1 | 5’CCTACCATGGGACTACTGAGTACTAACACACTAACCGTGTATA GAGTTGGCGACCGCGGGGATCGCGTTTAAACCGCGCGCCGCTTAATTA ATTACGGATACA3’ |
| Oligo 2 | 5’GTATACCTCTAATTTAATTAGCGGCCGCGTTTAAACGGATCCCCG CGGTCGCGAAGCCTCTATACACCGGTCTAGTGTAGCTCAGTAGC CCCCATTGTAGG3’ |

Red denotes Id4 genomic sequence and black denotes pClasperA sequence.

Table S4. Primer sequences for eGfp-Ura3 cassette insertion into Id4 exon 1.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Id4-eGfp</td>
<td>5’CTCTACCGCTTTGTCGCCGCTCTCTCGCGACGAAAGCGCGATGAAAGCGGCGCCTGCGGATTGAGCAAGGGGCGAGG3’</td>
</tr>
<tr>
<td>Id4-Ura3</td>
<td>5’CAGGCCAGTGGGTAGTGGATGGATCGATCGATAAGGTGCTCAG GATCTCCGCGCGCCACCAGACTTTTCAATTCTTCAATT3’</td>
</tr>
</tbody>
</table>

Black denotes Id4 exon 1 genomic sequence, red denotes eGfp sequence, and blue denotes Ura3 sequence.
A

BAC-ld4 (wild-type)  5' Exon 1 Exon 2 Exon 3  3'

Exon 1

eGfp-Ura3

BAC-ld4-Gfp  5' Exon 1

5' Founder Males (FVB)

Pronuclear Injection (FVB Donor)

LT-11 Founder (FVB) X C57BL/6J Female

LT-11;FVB/C57BL/6J

C57BL/6J Backcross (6 Generations)

LT-11B6 (ld4-Gfp Line)

B

MW (100 bp)
ld4-Gfp Positive
ld4-Gfp Positive
ld4-Gfp Negative

C

MW (100 bp)
Pancreas Liver Kidney Brain Testis

D

Relative ld4 Transcript Abundance

Wild-Type LT-11B6
Figure 1: Generation of a transgenic Id4-Gfp reporter mouse line. (A) Schematic of the recombineering approach and breeding strategy used for generating the LT-11B6 mouse model that contains a Gfp reporter transgene representing expression of Id4. (B) Representative image of an agarose gel from genotyping analysis designed to identify LT-11B6 mice containing the Id4-Gfp transgene. MW is 100-bp ladder. The image displays products of PCR reactions with DNA samples from three transgenic littermates and one nontransgenic littermate. (C) Expression of the Id4-Gfp transgene in several different tissues of LT-11B6 mice as determined by RT-PCR analysis. (D) Quantitative comparison of Id4 transcript abundance in testes of adult LT-11B6 mice and wild-type littermates using qRT–PCR analyses.
Figure 2: Identification of Id4-Gfp-expressing spermatogonia in testes of LT-11B6 mice at multiple stages of postnatal life. (A) Representative images of immunostaining for Id4-Gfp-expressing cells in cross-sections of seminiferous tubules from testes of mice at PD 0, 3, 6, 12, 20, and 35. Arrows indicate ID4-GFP+ spermatogonia. Bars: PD 0, 3, and 6, 50 μm; PD 12, 20, and 35, 100 μm. (B) Representative pictures of live imaging for ID4-GFP+ cells in whole-mount seminiferous tubules from LT11-B6 mice at PD 6 and 35. Arrows indicate single ID4-GFP+ spermatogonia. Bars, 50 μm.
Figure 3: Proportionality of *Id4-Gfp*-expressing cells in the undifferentiated spermatogonial population of testes from LT-11B6 mice at multiple stages of postnatal life. (A) Representative images of coimmunofluorescent staining for PLZF (red) and *Id4-Gfp*-expressing cells in cross-sections of seminiferous tubules from testes of mice at PD 3, 6, 12, 20, and 35. DAPI (blue) was used to stain the nuclei of cells. At each age, the right image is a magnified view of the left image. Arrows indicate PLZF+/ID4-GFP+ spermatogonia, and stars indicate Plzf+/ID4-GFP- spermatogonia. Note that four images are provided for PD 35 to demonstrate multiple different spermatogonia. Bars: left images, 100 μm; right images, 50 μm. (B) Quantification of the proportion of the undifferentiated spermatogonial population that is ID4-GFP+ at multiple stages of postnatal development. Data are presented as mean ± SEM percentage of the Plzf+ undifferentiated spermatogonial population and were generated from immunostained cross-sections of three different mice at each age point. (C) Representative image of coimmunofluorescent staining for spermatogonia expressing the markers Gfra1 (red) and *Id4-Gfp* (green) within whole-mount seminiferous tubules from mice at PD 6. DAPI (blue) was used to stain the nuclei of cells. Arrows indicate single spermatogonia that are GFRa1+/ID4-GFP+ and stars indicate spermatogonia that are GFRa1+/ID4-GFP+ Bar, 100 μm.
Figure 4: Frequency and distribution of Id4-Gfp-expressing spermatogonia in testes of LT-11B6 mice during postnatal development and adulthood. (A) Quantification of seminiferous tubule cross-sections from testes of mice at PD 0, 3, 6, 12, 20, and 35 containing at least one ID4-GFP+ spermatogonia. Data are presented as the mean ± SEM percentage of total seminiferous tubule cross-sections. (B) Quantification of ID4-GFP+ spermatogonial numbers in cross-sections of seminiferous tubules from testes of mice at PD 0, 3, 6, 12, 20, and 35. Data are presented as the mean ± SEM number of positive cells per seminiferous tubule cross-section. (C) Representative images of immunostaining for Id4-Gfp-expressing cells in cross-sections of seminiferous tubules from testes of adult mice grouped as early (II–VI), middle (VII–VIII), or late (IX–I) stages of the seminiferous epithelial cycle. Arrows indicate Id4-Gfp+ spermatogonia. Bars, 50 μm. (D) Quantification of seminiferous tubule cross-sections at early, middle, and late stages of the seminiferous epithelial cycle containing at least one ID4-GFP+ spermatogonia. Data are presented as mean ± SEM percentage of the total ID4-GFP+ spermatogonial population. (E) Quantification of regions within seminiferous tubule cross-sections where ID4-GFP+ spermatogonia reside. Regions were defined as being juxtaposed to other tubules, interstitial tissue not containing blood vessels (−BV), and interstitial tissue containing blood vessels (+BV). Data are presented as mean ± SEM percentage of the total ID4-GFP+ spermatogonial population.
Figure 5: Regenerative capacity of Id4-Gfp+ and Id4-Gfp− spermatogonial populations from testes of hybrid LT-11B6/RosaLacZ mice. (A) Representative image of overlaid fluorescent and phase-contrast views of primary cultures of undifferentiated spermatogonia established from LT-11B6/RosaLacZ mice. Black arrows indicate clumps of spermatogonia containing ID4-GFP+ cells, and white arrows indicate clumps not containing ID4-GFP+ cells. Note that all spermatogonia (i.e., GFP+ and GFP−) express LacZ. (B) Representative post-sort scatter plots of ID4-GFP+ * and ID4-GFP- populations isolated from primary cultures of undifferentiated spermatogonia using fluorescence-activated cell sorting (FACS). (C) Representative images of testes from recipient mice 2 mo after transplantation with FACS-isolated ID4-GFP+ and ID4-GFP- subpopulations that have been stained for LacZ-expressing colonies of donor-derived spermatogenesis. Note that each blue segment is clonally derived from a single SSC. (D) Quantitative comparison of donor-derived colonies of spermatogenesis in testes of recipient mice that were generated from 10⁵ FACS-isolated ID4-GFP+ or ID4-GFP- cells 2 mo after transplantation. Data are mean ± SEM for three different cultures. (*)Significantly different at $P < 0.05$. 
Figure 6: Transcriptome analysis of ID4-GFP+ and ID4-GFP- spermatogonial populations from testes of LT-11B6 mice. (A) Quantitative comparison of transcript abundance for genes involved in specification of primordial germ cells (PGCs) and germ cell licensing and several genes reported previously to be expressed by undifferentiated spermatogonia. (B) Quantitative comparison of transcripts with significantly ($Q < 0.05$) different abundance. For both data sets, ID4-GFP+ and ID4-GFP- fractions were isolated from primary cultures of undifferentiated spermatogonia, and transcript abundances were determined using RNA-seq analysis. Data are presented as $\log_{10}$ of the mean FPKM generated from three different cultures. Differential abundance in B was determined statistically using Cuffdiff analysis. Note that dashed lines in A and B are set at 1, which was considered the cutoff of a transcript being considered present or absent.
Figure 7: The hierarchy for subsets of spermatogonia in the undifferentiated population of mouse testes and expression profile for intracellular molecular markers. The A_{single} spermatogonial pool consists of SSC and transitional progenitor subtypes. The SSC subtype possesses regenerative capacity and unique self-renewal properties to remain as A_{single} following division, thereby sustaining a pool from which new transitional A_{single} progenitor spermatogonia arise. In contrast, the transitional subtype attains enhanced propensity to form initial chained spermatogonia (A_{paired}) upon the next division and potentially retains limited capacity for self-renewal but lacks regenerative capacity. The SSC and transitional A_{single} subtypes express genes that are also expressed by chained (A_{paired-aligned}) progenitors but are distinguished by expression of specific genes such as Id4 and Utf1. Gene expression that distinguishes the transitional A_{single} progenitor spermatogonia is undefined but may include reduced levels of some genes also expressed by A_{single} SSCs.
Figure S1: Representative image of *Id4-Gfp* expressing cells within seminiferous tubules from testes of adult mice. Arrows indicate cohorts of two ID4-GFP+ cells and arrow heads indicate single ID4-GFP+ cells. Bar is 50μm.
Figure S2. Representative images of co-immunofluorescent staining for the differentiating spermatogonial markers STRA8 and KIT (red) and ID4-GFP (green) in cross-sections of testes from adult LT-11B6 mice. Stars indicate cells stained for Stra8 or Kit expression and arrows indicate cells stained for expression of Id4-Gfp.
Figure S3: Representative images depicting tubular/blood vessel, tubular/tubular, and tubular/interstitial localization of *Id4-Gfp* expressing cells within cross-sections of testes from adult LT-11B6 mice. Arrows indicate spermatogonia immunostained for expression of *Id4-Gfp*. Interstitial tissue, seminiferous tubules, and blood vessels are indicated by I, T, or BV, respectively.
Figure S4: Representative images of cross-sections from testes of W/Wv recipient mice three months after transplantation with ID4-GFP+ spermatogonia isolated from testes of adult LT-11B6 donor mice using FACS. Stars indicate seminiferous tubules containing donor-derived spermatogenesis. Arrowheads indicate seminiferous tubules completely devoid of germ cells that is typical of W/Wv mice.
Figure S5. Validation of differential abundance using qRT-PCR analysis for 11 transcripts determined to be significantly greater in ID4-GFP+ cells compared to ID4-GFP- cells using RNA-Seq analysis. Data are presented as mean±SEM fold-difference for three different cultures. The relative abundance for Id4-Gfp- cells is set at 1. Note that the error bars for most samples do not extend beyond the top of the data bar.
References


de Rooij, D. G. and L. D. Russell (2000). "All you wanted to know about spermatogonia but were afraid to ask." J Androl 21(6): 776-798.


CHAPTER THREE

THE RESTORATIVE CAPACITY OF SPERMATOGONIAL STEM CELL IS INFLUENCED BY UTF1 EXPRESSION

FRIEDA K. K. Chan, Melissa J. Oatley, Jon M. Oatley

School of Molecular Biosciences, College of Veterinary Medicine, Washington State University, Pullman, WA 99164

Abstract

Continual spermatogenesis is dependent on SSC self-renewal to maintain the stem cell pool and differentiation to initiate spermatogenesis. At present, the molecular mechanisms that regulate SSC state and the transition to progenitor spermatogonia are undefined. In recent studies, we utilized ID4-GFP as a tool to isolate enriched populations of SSCs and progenitors for transcriptome analysis and defined a set of candidates for regulating SSC state. Among these genes, undifferentiated transcription factor 1 (Utf1) was particularly interesting because of its role in maintenance of ESC pluripotency. Here, we set out to determine whether UTF1 is important for the regulation of SSC state. Our results showed localization of UTF1 to SSCs, demonstrated its expression is important for SSC restoration of spermatogenesis, and determined that SOX2 binds directly to the upstream element of Utf1 for activation of transcription. Taken together, our results suggested that UTF1 plays a vital role in SSC maintenance.

Introduction

In mammals, spermatogenesis originates from a rare population of spermatogonial stem cells (SSCs) and culminates in the production of millions of spermatozoa every day.
Continual spermatogenesis is possible because SSCs slowly divide to produce one daughter cell that retains the stem cell state and a second that initiates spermatogenesis by transitioning into a progenitor spermatogonium (de Rooij 1973). Currently, the molecular mechanism(s) that maintain the SSC state and regulate the transition to progenitor spermatogonia are unknown. There are several models to describe SSC differentiation, which can be broken down to irreversible and reversible. According to the irreversible model, progenitors are committed to spermatogenesis and can't revert back to SSC (de Rooij 1998). In the reversible model, progenitors have the plasticity to dedifferentiate and re-enter the SSC pool (Hara, Nakagawa et al. 2014). Understanding the mechanisms necessary for SSC maintenance enable the improvement in diagnoses for infertility and development of gene therapy to reestablish the SSC pool, thereby restoring fertility.

In previous studies we provided evidence that, upon cell division, daughter cells that retained stem cell status expressed inhibitor of DNA binding 4 protein (ID4) but cells that became progenitors did not (Oatley, Kaucher et al. 2011). By generating a fluorescently tagged protein, ID4-GFP, we were able to isolate purified populations of SSCs. Through transcriptome analysis of ID4-GFP+/SSCs and ID4-GFP-/progenitors, we found 11 genes whose expression was limited to ID4-GFP+ cells and thus are candidates for regulating stem cell state (Chan, Oatley et al. 2014). Among these, undifferentiated transcription factor 1 (Utf1) seemed particularly interesting because of its importance for the maintenance of embryonic stem cell (ESCs) pluripotency (Okuda, Fukushima et al. 1998). UTF1 has been shown to be essential for maintenance of ESCs through its function as a chromatin-associated protein that regulates poised gene expression and degradation (Jia,
Zheng et al. 2012). It has been implicated in regulating the transition from ESCs to differentiated cells because it is rapidly degraded upon ESC differentiation and \textit{Utf1} knockdown showed delay or block in differentiation (van den Boom, Kooistra et al. 2007). Conversely, upon the dedifferentiation of somatic cells into induced pluripotent stem cells (iPSCs), \textit{Utf1} expression has been reported to be activated by the stem cell transcription factor, sex-determining region Y-box2 (SOX2) (Takahashi and Yamanaka 2006). Finally, as in ESCs, UTF1 appears to contribute to the dedifferentiation and maintenance in iPS cells. Taken together, the available evidence suggests that UTF1 plays a role in stem cell maintenance, making it a strong candidate for playing a similar role in the maintenance of SSCs in the testis.

In the rat testes, UTF1 has been localized in SSCs and a subpopulation of progenitor spermatogonia, but its functional role in SSC maintenance remains unknown (van Bragt, Roepers-Gajadien et al. 2008). In this study, we investigated the expression and function of UTF1 in mouse SSCs. We found that UTF1 was upregulated in SSCs both \textit{in vivo} and \textit{in vitro}. Further, our data provide evidence for conserved mechanisms of function since, as in ESCs and iPSCs, SOX2 binds directly upstream of \textit{Utf1} for activation of transcription. Thus, our results suggest that UTF1 plays a vital role in SSC maintenance.

\textbf{Results}

\textit{Utf1} and \textit{Sox2} are upregulated in SSCs

\textit{Utf1} has been proposed to regulate SSC state, and SOX2 has been postulated to be an activator of \textit{Utf1} transcription in ESCs (Arnold, Sarkar et al. 2011; Chan, Oatley et al. 2014). To determine whether \textit{Utf1} and \textit{Sox2} are upregulated in SSCs, we prepared mRNA and protein samples from purified populations of SSCs and progenitor spermatogonia. We used
the novel SSC marker, ID4-GFP, to isolate purified populations of each cell type from primary germ cell cultures generated from LT-R mice (Chan, Oatley et al. 2014). Quantitative real-time PCR (qRT-PCR) analysis of three biological replicates demonstrated that \textit{Utf1} and \textit{Sox2} transcript abundance in SSCs was increased by 5.51±0.44 fold and 7.27±0.44 fold (p<0.05), respectively, by comparison to progenitors (Figure 1A). Western blot analysis of three biological replicates demonstrated that UTF1 and SOX2 protein abundance in SSCs was increased by 2.65±0.15 fold and 5.51±0.21 fold (p<0.05), respectively (Figure 1B, C).

**UTF1 is present in SSCs and a subpopulation of progenitor spermatogonia**

To localize UTF1 protein in neonatal to adult LT-R testes, it was co-immunostained with the SSC marker ID4-GFP in testes cross-sections. Results of three biological replicates showed that UTF1 was localized in all ID4-GFP+/SSCs and a subpopulation of ID4-GFP- cells (Figure 2). To further characterize this subpopulation, we co-immunostained testis cross-sections with UTF1 and the progenitor spermatogonia marker, lin-28 homolog A (LIN28). The results from three biological replicates showed that UTF1 localized to a subset of LIN28+ progenitor spermatogonia (Figure 2).

**Transient reduction of \textit{Utf1} diminishes the colonizing potential of SSCs**

To investigate the effect of UTF1 on SSC function, we used \textit{Utf1} siRNA on LT-R primary germ cell cultures followed by transplantation analysis to assess recolonization. Reduction efficiency measured by qRT-PCR analysis showed that \textit{Utf1} transcript abundance in treatment cells was reduced 0.45±0.06 fold (p<0.05) (Figure 3A) and Western blot analysis showed that UTF1 protein abundance was decreased to 0.59±0.16 fold (p<0.05) when compared to control (Figure 3B, C). To assess the potential impact on progenitor cell
population, we analyzed the expression of Plzf, Ngn3, Sohlh1, and Sohlh2, which was
0.97±0.02, 1.24±0.19, and 1.04±0.25 fold, respectively (Figure 3D), suggesting that Utf1
siRNA treatment did not have significant effect on the progenitor population.

To assess the impact of Utf1 transient reduction on SSCs, control or Utf1 siRNA treated cells
were transplanted into F1 hybrid germ cell-depleted recipient 7 days after transfection.
Two months post transplantation, the recipient testes were collected, exposed to X-Gal, and
LacZ positive donor colonies were quantified. The analysis of three biological replicates
showed Utf1 siRNA treated cultures had 61.14±5.07 colonies/10^5 cells injected and control
siRNA treated cultures had 112.60±10.30 colonies/10^5 cells injected (Figure 3E, F),
suggesting that Utf1 transient reduction reduced SSC number.

**SOX2 activates Utf1 expression in SSCs**

To determine if SOX2 binds to Utf1 promoter/enhancer region in SSCs,
immunoprecipitation with a validated SOX2-specific antibody was performed to isolate
SOX2-DNA complexes from 3 biological replicates of ID4-GFP purified SSCs. PCR analysis
with primers spanning an 102-bp segment of the Utf1 promoter/enhancer region that
contains the SOX2 binding site (canonical sequence ACATGA) showed an amplicon (NCBI
RefSeq accession no. NC_000073.6) (Figure 4E, F). The results of SOX2 ChIP analysis
performed on ESCs produced an amplicon of the same size (Figure 4F). For negative
control, normal rabbit IgG was used in place of primary antibody and the results did not
produce an amplicon (Figure 4F) suggesting that SOX2 binding to the enhancer/promoter
region of Utf1 was conserved from ESCs to SSCs.

To assess the effect of SOX2 on Utf1 expression, we conducted Sox2 siRNA studies with
primary germ cell cultures. Transcript abundance measured by qRT-PCR analysis in three
biological replicates showed a $0.61\pm0.02$ fold and $0.68\pm0.06$ fold reduction (p<0.05), in Sox2 and Utf1, respectively in treated cells by comparison with control (p<0.05; Figure 4A). Western blot analysis of three biological replicates demonstrated that SOX2 and UTF1 protein abundance in treatment cells decreased by $0.73\pm0.06$ fold and $0.31\pm0.08$ fold (p<0.05), respectively (Figure 4 B, C). To assess the impact on progenitor cell population, we analyzed gene expression of Plzf, Ngn3, Sohlh1, and Sohlh2, which was $1.02\pm0.12$, $1.41\pm0.28$, $1.37\pm0.57$, and $1.04\pm0.44$, suggesting that Sox2 siRNA treatment did not have significant effect on the progenitor population (Figure 4D).

**Discussion**

Our current understanding of the molecular mechanisms that regulate SSC self-renewal and differentiation is rudimentary. In previous research, Utf1 was implicated as a candidate gene for regulating SSC state (Chan, Oatley et al. 2014). In this study, our experimental results provide evidence that Utf1 is important for SSC maintenance and that its expression is activated by SOX2. Taken together, our results indicate that the role of UTF1 is conserved from ESCs and iPSCs to SSCs.

Our data support previous findings from studies in the rat (van Bragt, Roepers-Gajadien et al. 2008) that reported that UTF1 localizes to SSCs and a subpopulation of progenitors, since we see an identical localization pattern in the mouse testis. Although we can’t precisely identify the progenitor type in which Utf1 expression is down regulated, we hypothesize that Utf1 is expressed by both SSCs and Aα progenitor spermatogonia, but is down-regulated in Aαr and Aα1 progenitors. If this is correct, it suggests that the UTF1 positive progenitors are a population of transitional cells that possess epigenetic and genetic properties of both SSCs and progenitors.
Our previously published data suggested that, under normal conditions, SSCs but not progenitors are capable of restoring spermatogenesis (Chan, Oatley et al. 2014). However, data published by Shosei Yoshida’s laboratory suggest that progenitor spermatogonia have the capacity dedifferentiate and contribute to the SSC pool, but the precise molecular mechanism regulating this occurrence is unknown (Nakagawa, Sharma et al. 2010; Hara, Nakagawa et al. 2014). This type of dedifferentiation is clearly a feature of progenitor cells in Drosophila, where studies have shown that manipulation of the Jak-STAT signaling pathway or Bam withdrawal causes progenitors to regain stem cell capacity. By studying the epigenetic landscape and the gene expression of the SSC, UTF1+ progenitor, and UTF1- progenitor populations, we can understand the process of germ cell differentiation in the testis and answer the important question of whether progenitor cells have the capability to dedifferentiate and contribute to the SSC pool.

Support for Utf1 as an important gene for SSCs comes from the results of our transplantation analysis which showed that transient reduction of Utf1 decreased SSC number. However, further experiments are required to find out why SSCs decreased in number and to decipher the regulatory role of UTF1 in SSCs. In ESCs and iPSCs, UTF1 has been implicated as a chromatin-associated protein that binds to “poised” bivalent genes that are expressed during stem cell differentiation, i.e. genes that have both transcriptionally active histone 3-lysine 4 trimethylation and transcriptionally repressive histone 3-lysine 27 trimethylation epigenetic markers (Jia, Zheng et al. 2012). Comparison of the epigenetic landscape between SSCs and progenitors will provide regulatory mechanisms for the initiation of spermatogenesis, thereby, we can manipulate the epigenome and genetic properties of progenitor cells to revert them back to SSC state.
Dysregulation of SSC epigenome has been implicated in studies reporting transgenerational inheritance of genetic disease and infertility as a result of exposure to environmental stimuli (e.g., Guerrero-Bosagna and Skinner 2014; Skinner 2014; Vrooman, Oatley et al. 2015). However, in order to determine the disease causing epigenetic modifications, we need to understand the normal epigenic changes that occur during the course of spermatogenesis. By comparing the epigenome of SSCs and progenitors, we will understand the epigenetic modifications that occur during differentiation and how environmental factors impact the epigenome. This information will provide novel methods to diagnose and treat male infertility.

**Materials and Methods**

**Animals**

*Id4-Gfp* transgenic mice (LT-11) were maintained on a C57BL6/J background. Mice expressing the LacZ transgene on the Rosa26 locus (B6;129S-Gt (ROSA)26Sor/J) were obtained from Jackson Laboratories. LT-11 x B6;129S-Gt (ROSA)26Sor/J (LT-R) mice were used for generating primary germ cell cultures. Recipient mice for spermatogonial transplantation were C57BL/6J x 129S1/SvImJ (F1 hybrids). Washington State University is fully accredited by the American Association for Accreditation of Laboratory Animal Care, and all mice were maintained in a pathogen-free facility, housed in ventilated rack caging, and provided food and water ad libitum.

**Primary Germ Cell Cultures**

Germ cell cultures were established as described previously (Kubota, Avarbock et al. 2004) from LT-R mice at PD 6-8. Magnetic-activated cell sorting (MACS) was used to collect the Thy1+ testis cell fraction, which was cultured at 37°C on mitotically inactivated STO feeder
cell monolayer in defined mouse serum free media (mSFM) with growth factors GDNF (PeproTech, NJ) and FGF2 (BD Biosciences). At 7 day intervals, germ cells were subcultured onto fresh STO feeders at a ratio of 1:2 or 1:3. All experiments were conducted with cultures 1-3 months after establishment.

**Fluorescence-Activated cell sorting (FACS)**

ID4-GFP positive and negative cells from LT-R primary germ cell cultures were isolated based on GFP intensity using a BD FACS Aria instrument. Cell fractions were washed in mSFM and pelleted by centrifugation at 600xg for 7 minutes. The cell pellet was resuspended in mSFM and the resuspended cells were processed for qRT-PCR, Western blot and ChIP analyses.

**Small Interfering RNA Transfection**

Single cell suspensions from primary germ cell cultures at $10^5$ cells per 1ml in mSFM with growth factors were transfected with 75 pmol of ON-TARGETplus non-targeting control small interfering RNAs (siRNAs) (Dharmacon L-001810-01-05), SMARTpool: ON-TARGETplus Sox2-specific siRNAs (Dharmacon L-058489-00-0005), or SMARTpool: ON-TARGETplus Utf1-specific siRNAs (Dharmacon L-046858-01-0005) using Lipofectamine 2000 (Invitrogen 11668027). Transfected cells were incubated at 37°C in feeder free 24 well plates for 18 hours. After the incubation period, the transfected cells were collected by centrifugation at 600xg for 7 minutes and resuspended in mSFM with growth factors, cultured on STO feeders, and incubated at 37°C. For RNA and protein analyses, cells were collected 48 hours post siRNA transfection. For transplantation analyses, the cells were collected one week post siRNA transfection.
Quantitative RT-PCR Analysis

Total cellular RNA from experimental cell populations (ID4-GFP+ cell fractions; ID4-GFP- cell fractions; primary germ cell cultures treated with control, Sox2, or Utf1 siRNA) were isolated using Trizol Reagent (Invitrogen, USA). Samples were treated with DNase I and reverse transcribed with Superscript III reverse transcriptase and oligo (d)T priming to make cDNA. Taqman and Sybr green assays were conducted using validated primers and probes for Sox2, Utf1, Plzf, Ngn3, Sohlh1, Sohlh2 (Table below). All reactions were performed with an ABI 7500 Fast Sequence Detection system (Applied Biosystems). The constitutively expressed gene Rps2 was used for normalization. The 2^{-\Delta\Delta CT} calculation method was used to calculate relative transcript abundance for genes-of-interest in each sample, as previously described (Oatley, Kaucher et al. 2011; Yang, Gwost et al. 2013).

### Primers used for Sybr green based qRT-PCR analyses

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amplicon Size</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utf1</td>
<td>135</td>
<td>TGTCCCGGTGACTACGTCTCT</td>
<td>CCCAGAAGTAGCTCCGTCTCT</td>
</tr>
<tr>
<td>Sohlh1</td>
<td>167</td>
<td>AGAGGGCGTTGTCATTCAAGT</td>
<td>CATGGCGGAGATCCGGGGG</td>
</tr>
<tr>
<td>Sohlh2</td>
<td>154</td>
<td>GGGCAGGGGAGTAAATCTT</td>
<td>CAAACGAGTTAGCAGCAAAAG</td>
</tr>
<tr>
<td>Sox2</td>
<td>157</td>
<td>GCGGATGGAAGTTTGTGCC</td>
<td>CGGGAAGGGGTGACTTATCCTT</td>
</tr>
<tr>
<td>Rps2</td>
<td>106</td>
<td>CATGACCCCGCTGCAAAT</td>
<td>TCTCCCGGTACTCTGGTAG</td>
</tr>
</tbody>
</table>

### Primers used for Taqman based qRT-PCR analyses

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amplicon Size</th>
<th>Vendor</th>
<th>Catalogue numbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ngn3</td>
<td>73</td>
<td>Life Technologies</td>
<td>4331182</td>
</tr>
<tr>
<td>Plzf</td>
<td>63</td>
<td>Life Technologies</td>
<td>4331182</td>
</tr>
<tr>
<td>Rps2</td>
<td>62</td>
<td>Life Technologies</td>
<td>4331182</td>
</tr>
</tbody>
</table>

Western Blot Analyses

Protein lysate from experimental cell populations (ID4-GFP+ cell fractions; ID4-GFP- cell fractions; primary germ cell cultures treated with control, Sox2, or Utf1 siRNA) were isolated using radioimmunoprecipitation lysis buffer (Santa Cruz Biotechnology sc-24948).
Eighty µg of protein was resolved in NuPAGE Novex 4-12% Bis-Tris gels (Life Technologies NP0321BOX) using an XCell SureLock Mini-Cell electrophoresis system (Life Technologies E10001) at 200V for 60 minutes. Resolved proteins were transferred from the gel to nitrocellulose membrane (Bio-Rad #162-0115) using the XCell II blot module (Life Technologies E19051) at 30V for 60 minutes. Non-specific antibody binding was prevented by incubating the membrane in blocking buffer (5% nonfat dry milk in Tris-buffered saline with 0.1% Tween-20, TBST) for 1 hour at room temperature with gentle oscillation. The membranes were incubated with primary antibodies (Rabbit anti-UTF1 antibody, Cell Signaling #3909, 1:100; Rabbit anti-SOX2 antibody, Cell Signaling #2748, 1:50; Rabbit anti-Tubulin antibody, Santa Cruz Biotechnology sc-53646, 1:10,000) diluted in blocking buffer for 48 hours at 4°C with gentle oscillation. After primary antibody incubation, the membranes were washed for 3 x 10 minutes in TBST in room temperature. The membranes were incubated with secondary goat anti-rabbit IgG antibody conjugated to horseradish peroxidase (HRP) (Santa Cruz Biotechnology sc-2030, 1:1000) diluted in blocking buffer for 2 hours at room temperature with gentle oscillation. At the end of secondary antibody incubation, the membranes were washed for 6 x 10 minutes in TBST at room temperature. HRP abundance was detected using West Pico chemiluminescence detection reagent (Pierce Protein Research 34080). Digital images of the membranes were captured using the LAS-3000 Imaging System from Fuji. Protein abundance was measured using the ImageJ software (imagej.nih.gov/ij/list.html).

**Chromatin Immunoprecipitation**

One million ID4-GFP+ SSCs and one million ESCs were fixed in PBS with 1% formaldehyde for 10 minutes at room temperature to cross-link protein-DNA complexes. The crosslinking
process was quenched by adding glycine and incubated for 5 minutes at room temperature. The fixed cells were washed in ice cold PBS, collected by centrifugation, and resuspended and lysed in ice-cold 100µl of lysis buffer (R&D Systems 890713) containing protease inhibitor for 10 minutes at 4°C with rotation. The cell lysate was sonicated with a Bioruptor Plus Sonication System (B01020001) for 30 cycles of 30 seconds “on” and 30 seconds “off” at 4°C to generate sheared chromatin with an average length of approximately 1 kilobase. To collect the sheared chromatin, sonicated cell lysates were centrifuged at 12,000xg for 10 minutes at 4°C. The supernatant was collected and placed in a clean 1.5ml tube. To begin immunoprecipitation, 10µl of Rabbit anti-SOX2 (ChIP Formulated) antibody (Cell Signaling #2748) or 10µl of Normal Rabbit IgG (Cell Signaling #2729) diluted in 400µl of ChIP buffer was added to samples and incubated overnight at 4°C with rotation. To precipitate protein-DNA complexes, 50µl of streptavidin-conjugated agarose beads was added to the lysate and incubated for 2 hours at 4°C with rotation. To collect immunoprecipitated chromatin, the lysate was centrifuged for 1 minute at 6,000 rpm, the supernatant removed, and the pellet was washed 4 times with wash buffers provided in the ExactaChIP Human/Mouse STAT3 chromatin immunoprecipitation kit (ExactaChIP ECP1799). After the last wash, 100µl of Chelating Resin Solution (ExactaChIP ECP1799) was added to resuspend the pellet. To dissociate the DNA from protein, the lysate/resin solution was boiled for 10 minutes using a heatblock at 100°C. The samples were centrifuged at 6,000 rpm for 1 minute at room temperature and resulting supernatant containing the DNA fragments was collected into a clean microcentrifuge tube. The DNA fragments were purified using DNeasy DNA purification kit (Qiagen 28204). The purified DNA fragments were used for PCR analyses with primers (Figure 4E) designed to recognize a 102-bp segment of Utf1.
promoter/enhancer region (NCBI RefSeq accession no. NC_000073.6) that contains SOX2 binding site. PCR products were separated on a 1% agarose gel and imaged using the ChemiDoc XRS molecular imager system (Bio-Rad Laboratories).

**Fluorescence Immunostaining of Testes Cross-Sections**

Testes of LT-R mice from PD 3, 6, 12, 20, and adult were collected and fixed in 4% paraformaldehyde solution at 4°C for 2 hours to overnight (depending on the age), dehydrated in a graded series of ethanol washes, and embedded in paraffin. Five-µm thick cross-sections were adhered to positively charged glass slides, de-parafinized and rehydrated. The rehydrated sections were incubated in boiling EDTA buffer (1mM EDTA, 0.05% Tween20, pH8.0) for 20 minutes for antigen retrieval. To prevent non-specific antibody binding, the sections were incubated in blocking buffer (PBS containing 2% bovine serum albumin, 0.3% Triton X-100, and 5% donkey normal serum) for 1 hour at room temperature. Following blocking, the sections were incubated in primary antibodies (Goat anti-LIN28 antibody, R&D Systems AF3757, 1:100; Rabbit anti-UTF1 antibody, Abcam ab105090, 1:50; FITC conjugated goat anti-GFP, Abcam ab6662, 1:300) diluted in binding buffer (PBS containing 2% BSA and 3% Triton X-100) for 48 hours at 4°C. The sections were washed 3 x 10 minutes in PBS before incubation in secondary antibodies (Alexa Fluor 488 Donkey anti-goat IgG, Invitrogen Cat.A11055, 1:1000; Alexa Fluor 546 Donkey anti-rabbit IgG, Invitrogen Cat.A10040, 1:1000) for 4 hours at 4°C. The sections were washed 3 x 10 minutes in PBS before they were mounted with ProLong Gold antifade reagent containing 4′,6-diamidino-2-phenylindole (DAPI) (Invitrogen, P-36931).
**Fluorescence Microscopic Imaging**

Digital fluorescence microscopy images were captured with an Olympus IX51 inverted microscope and DP72 digital color camera (Olympus Inc., USA). Images were captured with CellSense acquisition software (Olympus Inc., USA).

**Transplantation Analysis**

To examine the regenerative capacity of germ cell populations, control and *Utf1* siRNA transfected cells were transplanted into the seminiferous tubules of busulfan treated immunologically compatible F1 recipient mice (60mg/kg of body weight) utilizing methodology described previously (Oatley and Brinster 2006). Ten µl of single cell suspension diluted in mSFM to a concentration of 1X10^6 cells/ml was microinjected into the seminiferous tubules of germ-cell-depleted adult recipient mice. Testes of recipient mice were examined at approximately 2 months post-transplantation for colonies of donor-derived spermatogenesis by staining with X-Gal (G-Biosciences). LacZ stained donor colonies within each recipient testis were counted on a dissecting microscope.

**Statistical Analyses**

All quantitative data are presented as mean±SEM for at least 3 different biological replicate samples. Differences between means for data were determined using the general linear model one-way analysis of variance function of SAS software.
Figure 1. Utf1 and Sox2 are upregulated in SSCs. (A) qRT-PCR analysis of Utf1 and Sox2 in ID4-GFP+/SSC and ID4-GFP-/progenitors. Data represent mean ± SEM of three biological replicates; * denotes significant difference (P ≤ 0.05). (B) Western blot analysis of UTF1 and SOX2 in ID4-GFP+/SSCs and ID4-GFP-/progenitors. TUBB1 was used as a loading control. (C) Quantification of Western blot data represented as mean ± SEM of three biological replicates; * denotes significant difference (P ≤ 0.05).
Figure 2. Utf1 is present in SSCs and a subpopulation of progenitor spermatogonia. (A) Representative images of co-immunofluorescence staining in testes cross-sections at PD 3, 6, 12, 20, and adult males. Arrows indicate UTF1 (red) /ID4-GFP colocalization. (B) Co-immunofluorescence staining of UTF1 (red) /LIN28 (green) in testes cross-sections at PD 3, 6, 12, 20, and adult. * denotes colocalization. DAPI (blue) was used to stain the nuclei of cells.
Figure 3. Transient reduction of *Utf1* diminishes the colonizing potential of SSCs. (A) qRT-PCR analysis of *Utf1* in control or *Utf1* siRNA transfected primary germ cell cultures. Data represent mean ± SEM of 3(?) biological replicates; * denotes significant difference (P ≤ 0.05). (B) Western blot analysis of UTF1 protein abundance in control or *Utf1* siRNA transfected primary germ cell cultures. (C) Quantification of Western blot. Data represent mean ± SEM of three biological replicates; * denotes significant difference (P ≤ 0.05). (D) qRT-PCR analysis for *Plzf, Ngn3, Sohlh1,* and *Sohlh2* in control or *Utf1* siRNA transfected primary germ cell cultures. Data represent mean ± SEM of three biological replicates. (E) Representative images of recipient mouse testes 2 months after transplantation with primary germ cell cultures 7 days after control or *Utf1* siRNA transfection. LacZ positive (Blue) colonies are spermatogenesis clonally derived from a single SSC. (F) Quantification of transplantation data represented as mean ± SEM where donor derived colonies were normalized to $10^5$ cells injected. * denotes significant difference (P ≤ 0.05).
**A**

![Graph A](image)

Relative Transcript Abundance (Fold Difference)

<table>
<thead>
<tr>
<th>Sox2</th>
<th>Utf1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control siRNA</td>
<td>*</td>
</tr>
<tr>
<td>Sox2 siRNA</td>
<td>*</td>
</tr>
</tbody>
</table>

**B**

![Graph B](image)

Relative Protein Abundance (Fold Difference)

<table>
<thead>
<tr>
<th>UTF1</th>
<th>Sox2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control siRNA</td>
<td>*</td>
</tr>
<tr>
<td>Sox2 siRNA</td>
<td>*</td>
</tr>
</tbody>
</table>

**C**

![Graph C](image)

Relative Transcript Abundance (Fold Difference)

<table>
<thead>
<tr>
<th>SOX2</th>
<th>UTF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control siRNA</td>
<td>*</td>
</tr>
<tr>
<td>Sox2 siRNA</td>
<td>*</td>
</tr>
</tbody>
</table>

**D**

![Graph D](image)

Relative Transcript Abundance (Fold Difference)

<table>
<thead>
<tr>
<th>Id4</th>
<th>Plzf</th>
<th>Ngn3</th>
<th>Sohlh1</th>
<th>Sohlh2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control siRNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sox2 siRNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**E**

5' AGCTGCTGCGGCGACTTCACTAGCATAGACA
ATGACGCGGCTATCCTGAGGCTCCAATTAC
AAATTCTGGGCCTCCGCGCTTACTGCT
AACCCACAGACTAAGTCTTGCCCTGAGGGAT
GTCCTGAGGAGTTATCCAGATGGTCCCAA 3'

**F**

![Experiment F](image)

**Utfl Promoter**

<table>
<thead>
<tr>
<th>SSCs</th>
<th>Input DNA</th>
<th>IgG SOX2</th>
<th>IP DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ESC**

<table>
<thead>
<tr>
<th>Input DNA</th>
<th>IgG SOX2</th>
<th>IP DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4. SOX2 activates *Utf1* expression in SSCs (A) qRT-PCR of *Sox2* and *Utf1* in control or *Sox2* siRNA transfected primary germ cell cultures. Data represent mean ± SEM of three biological replicates; * denotes significant difference (P ≤ 0.05). (B) Western blot analysis of SOX2 and UTF1 in control or *Sox2* siRNA transfected primary germ cell cultures. (C) Quantification of Western blot data represented as mean ± SEM of three biological replicates. * denotes significant difference (P ≤ 0.05). (D) qRT-PCR analysis of *Plzf*, *Ngn3*, *Sohlh1*, and *Sohlh2* in control or *Sox2* siRNA transfected primary germ cell cultures. Data represented as mean ± SEM of three biological replicates. (E) Nucleotide sequence of the *Utf1* promoter/enhancer region that contains the canonical SOX2 binding site in the mouse genome (NCBI RefSeq accession no. NC_000073.6)(Sequence bolded and in purple). Primer sequences for PCR analysis are bolded, underlined, and italicized. (F) Representative image of PCR analysis for *Utf1* promoter sequence that was immunoprecipitated from SSCs and ESCs using SOX2-specific antibody. Normal IgG instead of primary antibody was used as negative control.
References


CHAPTER FOUR
SUMMARY

Introduction

Prior to this dissertation, the reproductive biology field lacked a method to enrich for a purified SSC population. My participation in studies in the Oatley laboratory led to the publication of the studies detailed in Chapter 2, describing the generation of an \textit{ld4-Gfp} transgenic mouse and providing evidence that ID4-GFP can be used as a novel tool to isolate SSCs from progenitor spermatogonia. Subsequent transcriptome analysis of SSCs provided 11 putative genes for regulating SSC maintenance, including \textit{Utf1}, a chromatin-associated protein (Chan, Oatley et al. 2014). To further investigate the importance of \textit{Utf1} in SSC maintenance and differentiation, as detailed in Chapter 3, we transiently reduced \textit{Utf1} expression and assessed the effect on SSC maintenance. Transplantation analysis demonstrated that \textit{Utf1} reduction decreased SSC restoration of spermatogenesis, therefore indicating that \textit{Utf1} is important for SSC maintenance. Additionally, our studies suggest that SOX2 activation of \textit{Utf1} expression is conserved from ESCs and iPSCs to SSCs, leading us to hypothesize that UTF1 functions as a chromatin-associated protein to regulate the SSC epigenome. Taken together, the studies in this dissertation have provided novel methods to examine the epigenome and transcriptome during SSC self-renewal and differentiation, which, has the potential to contribute to our understanding of male infertility.

The severity of male infertility ranges from blockages that prevent the delivery of sperm, abnormal sperm function, low sperm production, to azoospermia. Research and clinical improvements over the past four decades have provided male infertility patients that produce viable sperm with the option of using assisted reproductive technology (ART)
to have biological children. Despite these advances, many causes of infertility remain idiopathic and ART cannot provide assistance for patients suffering from azoospermia. Improving diagnostic capabilities and developing new infertility treatments requires in depth knowledge of the molecular mechanisms that regulate spermatogenesis. The tools ID4-GFP and UTF1 could provide a means of studying SSC maintenance and pave the way to the development of new gene therapy approaches for the treatment of infertility through the re-establishment of spermatogenesis.

**Future Directions**

In 1978, the collaboration of Patrick Steptoe and Robert Edwards led to the first baby, Louise Brown, born as a result of *in vitro* fertilization (IVF) (Brinsden and Brinsden 2009). In 2006, at 27 years old, Louise Brown was able to conceive naturally and had a baby of her own. During the past four decades, assisted reproductive technology (ART) has become a standard clinical procedure to treat infertility, and today in the United States 1.5% of children are conceived through IVF (Society for Assisted Reproductive Technology, www.sart.org). Of these infertility cases, approximately half are associated with male infertility (Chandra, Copen et al. 2013).

Assisted reproductive technology (ART) is constantly evolving as scientists and clinicians strive to provide the safest and most ethical methods to produce a healthy child. A few studies have inferred low birth weight, preterm birth, epigenetic disorders, cancer, and poor metabolic health as side effects in ART children (Feuer, Camarano et al. 2013), but there is yet to be direct scientific data to support these claims. Studies will need to be performed to investigate whether ART causes these disorders. Further, because human lifespan is significantly longer than laboratory mice and the first IVF babies are only now
reaching reproductive age, we have yet to determine whether ART is associated with an increase in reproductive disorders and transgenerational genetic diseases. Additionally, the current ART procedures do not provide an option for patients without sperm production to have biological children. Therefore it is important to study the epigenome and transcriptome that regulate SSC self-renewal and differentiation in order to diagnose ART side effect and develop gene therapy to reestablish the germ cell pool in azoospermic patients.

To treat germ cell-depleted patients and provide them with an option to have biological children, sperm production needs to be restored via an in vivo method where the SSC pool is reestablished, an in vitro approach where methods are derived for ex vivo production of sperm, or a combination of both. Studies have been performed to understand the pathways that regulate male gametogenesis in vivo. This information was translated to in vitro culture conditions to generate haploid germ cells from diploid cells ex vivo (Toyooka, Tsunekawa et al. 2003; Geijsen, Horoschak et al. 2004). The results showed haploid spermatocyte-like cells generated from ESCs, but no offspring were produced when eggs were fertilized with the derived cells via intracytoplasmic sperm injection (ICSI) (Toyoka et al., 2003; Geijsen et al., 2004).

In order to improve in vitro derivation of sperm, the complex process that differentiates ESCs to haploid germ cells was broken down to smaller developmental stages. In vivo, embryonic cell differentiates into three cell types: epiblast, endoderm, and trophoectoderm. The epiblast cells have the competency to give rise to germ cells and all somatic cells. The primordial germ cells (PGCs) migrate to the genital ridge and become
sex-specific gametes. The male gametes, SSCs, self-renew and differentiate for initiation spermatogenesis (McLaren and Lawson 2005).

By focusing on the important signaling and transcription regulators for the differentiation of embryonic cells to epiblast and the transition from epiblast cells to primordial germ cells (PGCs) (Ohinata, Ohta et al. 2009), various research groups were able to derive epiblast-like stem cells (EpiSCs) from ESCs and iPSCs, thus differentiating the EpiSCs to PGC-like cells (Takahashi and Yamanaka 2006; Okita, Ichisaka et al. 2007; Wernig, Meissner et al. 2007; Yu, Vodyanik et al. 2007). In 2010, the Saitou research group was able to generate PGC-like cells from ESCs and iPSCs for reinitiation of spermatogenesis in germ cell-depleted testes and the resulting sperms were able fertilized eggs via ICSI to produce fertile offspring (Hayashi, Ohta et al. 2011). This is an important breakthrough in the reproductive field because it showed the possibility of taking somatic cells of azoospermic patients to generate iPSC-sperms, thus providing them a method to have biological children.

Despite the ability to generate iPSC-sperm and produce offspring, the Saitou lab study demonstrated that not all iPSC-PGCLCs have capability to restore spermatogenesis, testes injected with iPSC-PGCLCs could result in teratoma formation, embryos fertilized from iPSC-sperms have lower rate of implantation, and live birth offspring from iPSC-sperms died prematurely due to tumor formation in the neck region (Hayashi, Ohta et al. 2011). These results indicated that it is important to compare the molecular mechanisms of iPSC-gametes and normal gametes to identify the cause of these side effects mentioned above.
The new understanding of UTF1 and novel tool of ID4-GFP provided through my studies will aide in the effort to study the molecular mechanisms that regulate SSC self-renewal and differentiation. By using ID4-GFP as a tool to isolate purified populations of SSCs and progenitor, it will be possible to map the epigenome for these respective cells. UTF1 can be used to immunoprecipitate protein-DNA complexes to study genes that initiate SSC differentiation. Taken together, the identification of DNA methylation pattern and poised chromatin structures in SSCs could provide an epigenetic map to investigate the molecular basis that regulates SSC maintenance and initiates spermatogenesis.

**What is the methylome that defines SSCs?**

Every cell in the body has the same genome, but epigenetic marks program DNA to reflect developmental history, regulate current transcriptome, and guide future gene expression. Epigenetic modifications include DNA methylation and histone modifications, and characterizing these modifications in SSCs will increase our understanding of the molecular mechanisms that govern the unique ability of these cells to self-renew and provide a population of progeny cells that initiate spermatogenesis.

To understand the importance of epigenetic marks for gametogenesis, bisulfite sequencing has been done on cells at different stages of gametogenesis (Li and Sasaki 2011; Messerschmidt, Knowles et al. 2014). The results from this analysis indicated that immediately after fertilization the entire genome in embryonic cells is globally demethylated, with the exception for imprinted genes and transposons. The DNA of blastocyst cells is *de novo* remethylated which coincides with cell lineage specification. A second round of genome wide demethylation occurs in the germ cell lineage, primordial germ cells (PGCs). The DNA remethylation of male and female germ cells occurs after PGC
differentiation, at different times, and in a sex-specific manner. For female germ cells, maternal specific methylation is established after birth, during the oocyte growth phase prior to ovulation (Lucifero, Mertineit et al. 2002). For male germ cells, paternal methylation is initiated in mitotically arrested prospermatogonia prior to birth (Davis, Trasler et al. 1999; Li, Lees-Murdock et al. 2004).

The generation of ID4-GFP mice as described in Chapter 2 provided a tool to isolate purified populations of SSCs. This will facilitate investigators to define the epigenome of SSCs. Using ID4-GFP, it will be possible to isolate ID4-GFP+/SSCs and ID4-GFP−/progenitors for whole genome bisulfite sequencing. These high-resolution methylome maps of SSCs and progenitors will make possible the identification of DNA methylation patterns that define SSC identity, mediate SSC gene expression, maintain imprinted genes, and guide unipotent differentiation during spermatogenesis.

By comparing DNA methylation maps obtained for SSCs and progenitor cells with the previously published data from embryonic germ cells (Seisenberger, Andrews et al. 2012; Kobayashi, Sakurai et al. 2013), it will be possible to understand the dynamic epigenetic changes that occur during the transition from PGC to SSC and during the process of spermatogonial cell differentiation. This knowledge will facilitate our understanding on how methylation patterns guide gene expression during gametogenesis. Additionally, by manipulating these methylation marks, it will be possible to examine the consequences of methylome dysregulation.

One of the known consequences of DNA methylation error is imprinting disorders. ART techniques such as in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) has been associated with aberrant methylation of imprinted genes, though at which
point of the procedures these methylation mistakes occur remains unknown (Bowdin, Allen et al. 2007; Doornbos, Maas et al. 2007; Denomme and Mann 2012). Genomic imprinting is important for normal developmental processes such as embryogenesis, placental functions, and neurogenesis (McGrath and Solter 1984; Surani, Barton et al. 1984). The dysregulation in DNA methylation that leads to aberrant imprinted genes expression has been implicated as the cause of various human diseases, developmental disorders, and malignant tumors (Paulsen and Ferguson-Smith 2001). Recent studies have correlated an increased number of human imprinting disorders, such as Beckwith-Wiedemann syndrome, Angelman syndrome, Silver-Russell syndrome, and retinoblastoma, in association with ART procedures (Cox, Burger et al. 2002; Weksberg, Shuman et al. 2002; DeBaun, Niemitz et al. 2003; Gicquel, Gaston et al. 2003; Maher, Brueton et al. 2003; Orstavik, Eiklid et al. 2003; Halliday, Oke et al. 2004).

ART embryos are more vulnerable to methylation dysregulation because they are handled and cultured ex vivo during the procedures. Research has demonstrated that exposure to environmental stimuli during embryonic development could dysregulate the DNA methylation of imprinted genes (Denomme and Mann 2012). Methylome maps for cell lineage-specification, germ line-specification, and sex-specification could provide the basis to compare ART embryo DNA methylation. By understanding the cause and timing of ART related methylation errors, clinicians could improve their protocols to diminish the occurrence of ART associated imprinting diseases.

**What are the genes that initiate SSC differentiation?**

In addition to DNA methylation, chromatin modification is an important epigenetic mark for cell programing. In embryonic stem cells (ESCs), chromatin with both H3K4me3
for activation and H3K27me3 for repression regulate transcriptionally “silent” poised genes that are expressed upon ESC differentiation (Bernstein, Mikkelsen et al. 2006; Mikkelsen, Ku et al. 2007; Hattori, Niwa et al. 2013). These poised genes are particularly interesting because they have been shown to be important for developmental regulation (Bernstein, Mikkelsen et al. 2006).

In ESCs, undifferentiated transcription factor 1 (UTF1), a chromatin associated protein, was demonstrated to bind to the poised genes to maintain transcriptional silencing (Okuda, Fukushima et al. 1998; van den Boom, Kooistra et al. 2007; Jia, Zheng et al. 2012). Utf1 knockdown was shown to delay or block in ESC differentiation, suggesting that it plays a role in regulating the initiation of ESC differentiation (van den Boom, Kooistra et al. 2007).

It is important to understand the molecular basis for SSC differentiation. The SSC transcriptome analysis described in Chapter 2 implicated Utf1 as a putative gene in SSC maintenance. This is reinforced by the unpublished data in Chapter 3 that provides evidence that UTF1 localizes only to SSCs and a subpopulation of progenitor spermatogonia and that a transient reduction of Utf1 decreases SSC numbers. These data suggest that UFT1 functions to regulate poised genes in SSCs. To further understand the SSC differentiation process, poised genes are likely to be involved since they have been implicated in ESC differentiation.

An intricate molecular map of the epigenetic and gene expression changes for SSC differentiation could identify the poised genes that are expressed during the transition from SSCs to progenitors. An epigenomic map could be generated by performing ChIP-seq analysis for H3K4me3, H3K27me3, and UTF1 with the SSCs and progenitor spermatogonia. Comparison of these ChIP-seq data will show the overall chromatin methylation changes
that occur during SSC differentiation. Utilizing the chromatin map, it will be possible to identify the poised genes that are expressed during the transition from SSC to progenitor state. Additionally, these bivalent chromatin and poised genes could be manipulated to determine the exact regulatory mechanisms that initiate spermatogenesis.

Through experimental manipulation of these regulatory elements, it will be possible to determine the whether the transition from SSCs to progenitors is a simple on-off process or a complex multi-step gradual process. The identification of the precise molecular steps in this transition process, it will be possible to manipulate these regulatory steps to determine whether progenitor spermatogonia could dedifferentiate and revert back to SSC state. The ability to manipulate the molecular mechanisms that regulate SSC fate will be important for the understanding, diagnosing, and treating male infertility.
References


