6 Discussion

6.1 Gene expression analysis of mouse spermatogenesis

The apparent asynchronous behavior of mammalian spermatozoa coupled with the paucity of female gametes (eggs) prompted many laboratories to use invertebrate animal models to understand the molecular basis of fertilization (Garbers, 1992; Mengerink et al., 2002). Functional bioassays allowed identification of important signaling components in the invertebrate sperm cells, and the subsequent identification of homologous genes in the mammal seemed to be a strong approach for unraveling fertilization mechanisms in the mammal. However, in many or most cases, homologues of the signaling components specifically important for fertilization in the invertebrate were not found in mammalian germ cells (Garbers, 1992). In retrospect this is not surprising given the rapid evolution of fertilization-specific proteins (Swanson and Vacquier, 2002). These negative findings indicated that direct experimental approaches in the mammal were required. The continued use of functional bioassays to understand mammalian fertilization also appeared an antiquated and generally unsuccessful means of approach. Of various more global and unbiased approaches, genetic screens for fertilization-defective mutations could provide the strongest method. However, the implementation of a saturation-mutagenesis screen in the mouse to search for such mutations appears untenable at this point in time, at least within individual laboratories, given the number of personnel and costs associated with such a labor-intensive screen.

Amongst alternatives are the use of proteomics, *in silico* subtraction, microarrays, differential display, serial analysis of gene expression, and signal peptide trapping to identify candidate gene targets. Such targets could then be disrupted to determine whether or not they are critical for fertilization. A number of underlying basic assumptions would reduce the list of gene candidates to palatable numbers. First, it is a reasonable assumption that sperm-specific gene products play an important role in male germ cell function. Second, transcripts that markedly increase in the haploid cell are more likely to play a significant role in the fertilization process. Third, proteins found on the cell surface of gametes are more likely to play a role in germ cell interactions with their environment and thus in fertilization. The use of the yeast-based signal peptide trap method (Klein et al., 1996) offers the opportunity to produce spermatid (haploid cell)-enriched cDNA libraries from which mRNA-encoded cell surface proteins can be identified. Through the use of such libraries it was estimated that the number of sperm-specific membrane proteins is greater than 200 (Quill et al., 2001). This seemingly large number of sperm-specific proteins led us to initiate searches for sperm-specific transcripts using DNA microarrays to examine testis expression profiles after subtraction of Sertoli and interstitial cell culture transcripts.

In order to analyze gene expression during spermatogenesis, we originally acquired several mouse testis cDNA libraries and used them to make cDNA microarrays. The goal was to generate a microarray covering a great number of genes expressed in the testis. However, when performing initial experiments with these microarrays, we encountered several technical problems and observed high variability between experiments. We were able to acquire promising preliminary data, but we eventually realized that using these arrays efficiently would require a lot of fine-tuning. Even if we had been able to overcome all the problems with the custom cDNA microarrays, we still would not have been able to acquire the high quality and the quantity of data that we could obtain with the Affymetrix GeneChips. Furthermore, the Affymetrix microarrays contained a surprisingly large number of testis-specific genes, providing us with an ideal platform to examine gene expression in both mouse and rat male germ cells. Several recent studies of gene expression in spermatogenesis using custom-made cDNA microarrays proved the superiority of the Affymetrix platform, as none of these studies were able to define comprehensive gene expression profiles and only identified few regulated genes (Sha et al., 2002; Tanaka et al., 2002; Pang et al., 2003; Guo et al., 2004; Maratou et al., 2004).

We analyzed gene expression in the developing mouse testis from day 1 *post partum* to adult using the Affymetrix Mouse U74v2 oligonucleotide array set, generating profiles for roughly 20,000 genes expressed in the testis. The results provide an invaluable depository of gene transcripts that could be specifically involved in spermatogenesis and fertilization. In mining the microarray data, we subtracted all genes expressed by testicular somatic cells and identified 3486 different genes that were regulated at least 3-fold compared to day one. Interestingly, 2245 of these genes are still uncharacterized.

By performing cluster analysis on these genes we identified more than 1000 genes (clusters 6-8, see Fig. 11) whose relative abundance consistently decreased with animal age. The fact that pre-meiotic germ cells are out-numbered rapidly with the beginning of meiosis suggests that most of these genes are expressed by pre-meiotic germ cells. Genes dropping earliest (cluster 6) are candidates to be highly enriched in spermatogonial stem cells. Genes in clusters 7 and 8 are probably expressed in spermatogonial stem cells as well, but their expression is likely maintained through the early stages of spermatogonial differentiation (Fig. 24). Investigation of the tissue-specificity of the genes in clusters 6-8 based on the Unigene database showed that less than 0.5% of the genes are restricted to the testis, suggesting that only very few of the genes in spermatogonial stem cells or spermatogonia are unique to the testis.

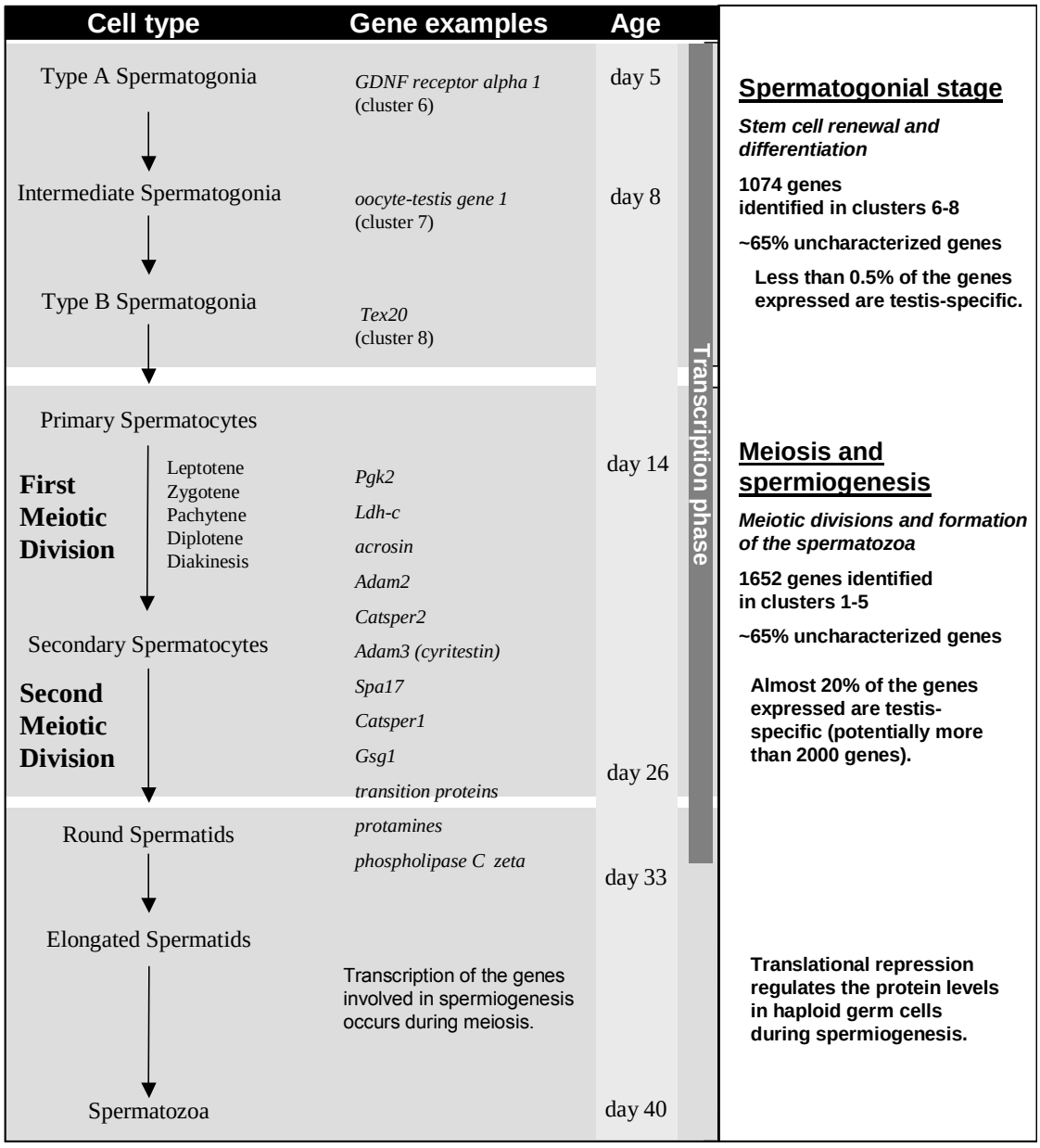

Figure 24. Genes identified to be expressed during spermatogenesis
Genes expressed in the spermatogonial stage are listed in the order in which a drop of expression was detected during mouse testis development (top to bottom). Genes expressed during meiosis and spermiogenesis are listed in order of their appearance. For certain cell types it is indicated at what age they are first formed in the mouse.

Cluster analysis also revealed 1652 genes whose expression increased coincident with or after meiosis (clusters 1-5, see Fig. 11). Genes in clusters 2 and 4 show the strongest increase in expression, corresponding with the formation of pachytene spermatocytes and round spermatids, respectively (Fig. 25). The list of genes in the five clusters contains numerous germ cell-specific genes known from the literature, but more than two

thirds of these genes are yet to be named and characterized. Remarkably, almost 20% of these genes are expressed exclusively in the testis and are therefore expressed only in the male meiotic or post-meiotic germ cell. Through further analysis of the Unigene database and through confirmation by quantitative real-time PCR we estimate that greater than 2300 genes are dedicated to male germ cell-specific transcripts, greater than 99% of which are first expressed during or after meiosis.

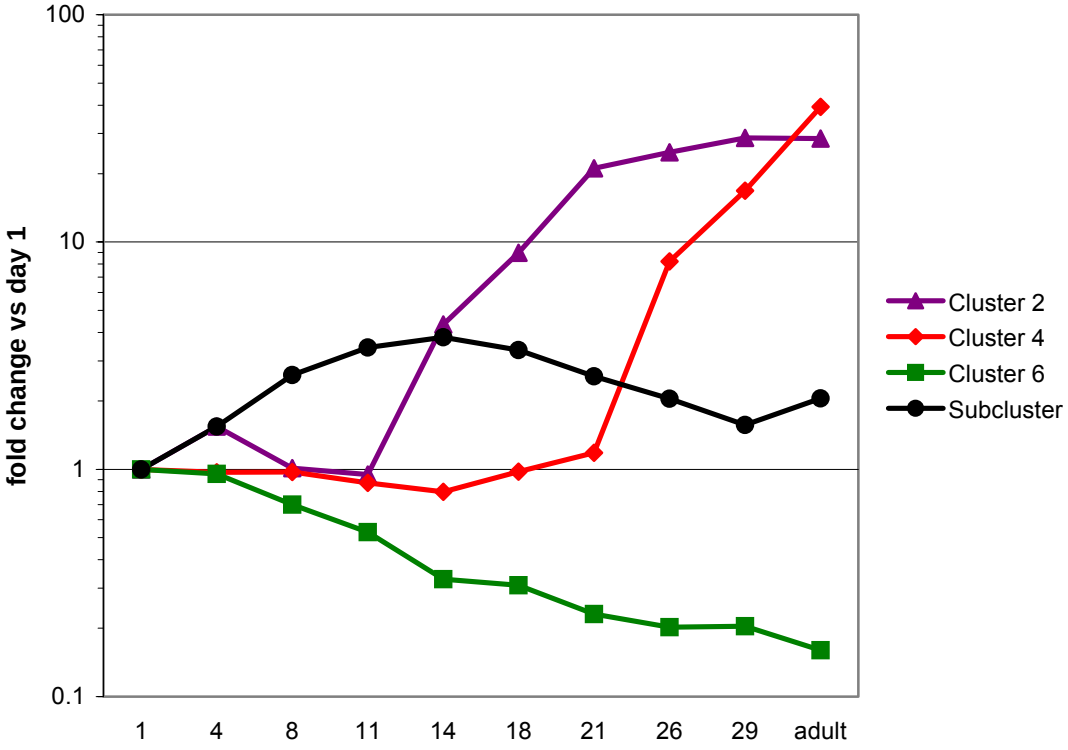

Figure 25. Important clusters of gene expression during mouse spermatogenesis
Expression patterns of the two most distinct clusters of genes expressed in spermatocytes and round spermatids (clusters 2 and 4), the gene cluster containing genes potentially enriched in type A spermatogonia (cluster 6), and a cluster derived by a second round of clustering of the unclassified genes, containing genes potentially involved in the prophase of meiosis (subcluster).

In order to estimate the actual percentage of these genes in relation to the mouse genome, one has to consider the predicted number of genes encoded by the mouse genome. While it is assumed that more than 99% of human genes have a mouse homologue and vice versa (Waterston et al., 2002; Sands, 2003), current estimates for the total number of genes differ widely. The Mouse Genome Sequencing Consortium predicts 30,000 genes for mouse and human (Lander et al., 2001; Waterston et al., 2002), and Celera Genomics predict 26,000-38,000 human genes (Venter et al., 2001). Both approaches are based on genomic sequencing. Combining data from these two projects leads to a slightly larger number of predicted genes, between 40,000 (Xuan et al., 2003)

and 44,000 (Parra et al., 2003). The use of EST sequence data produces even higher estimates of 57,000-81,000 genes (Liang et al., 2000). Based on a collection of 246 full-length cDNA libraries, a recent study by the RIKEN Genome Exploration Research Group predicts the existence of approximately 70,000 'transcriptional units' in the mouse (Okazaki et al., 2002). A significant proportion of this number is comprised of non-coding genes (about 15,000), a group of genes not considered by the genomic sequencing approaches. The predictions above do not include alternative splicing and other post-transcriptional modifications, which probably account for the majority of the variety of the mammalian proteome and could increase the number of different proteins by as much as three- or four-fold (Waterston et al., 2002).

Our estimate of 2300 germ cell-specific genes was derived from the total number of mouse Unigene clusters with at least two sequence entries (62,692). The Unigene database is based on EST sequences, and the total number of entries is similar to the number of genes predicted through EST sequencing approaches. Even if some EST entries will represent identical full-length transcripts and the actual number of genes is lower, there is no apparent reason that the testis would have more such redundant entries than other tissues. This justifies our conclusion that about 4% of the mouse genome is reserved for expression specifically in meiotic and post-meiotic male germ cells. That so many germ cell-specific genes are expressed late in development provides a vast number of potential participants in the process of fertilization.

It is noteworthy that some of the genes in the unclassified cluster can be further grouped by their expression patterns. The most interesting cluster emerging would be one containing the genes that play an early role in the process of meiosis but whose transcripts are not retained in round spermatids. Expression levels of these genes peak at day 14 and then decrease in abundance as the round spermatids begin to populate the testis (see Fig. 25). Examples for genes found in the unclassified cluster that match this pattern are *Sycp1*, *Ott*, and *Tex12*, all of which have previously been identified to be expressed in early meiotic cells. Synaptonemal complex protein 1 (Sycp1) is one of the components involved in synapsis during the prophase of meiosis and is normally expressed in both male and female meiotic germ cells. Transcription starts as early as the leptotene stage (Kerr et al., 1996b). Ott (ovary testis transcribed) is an x-linked gene that has been shown to be specifically expressed during male and female meiotic prophase (Kerr et al., 1996a). Testis expressed gene 12 (Tex12) is first expressed in spermatogonia (Wang et al., 2001), but its expression pattern and the fact that we find it to enriched in lam_{NB} cells over lam_{B} cells suggest that its expression level increases during the early stages of meiosis. There are more than 200 genes matching this expression pattern, many of which could play important roles at the start of meiosis.

With such a large number of genes showing similar expression patterns, it is very likely that the transcription of some of these genes is regulated by common mechanisms. The identification of these genes can potentially be achieved by mining the mouse genome sequence (Waterston et al., 2002). The full sequence and annotation of the genome will allow us to answer the following questions: Which of these genes are located in close proximity to each other in the genome and are potentially regulated by the same transcription factors? Which of the genes in different locations in the genome share common regulatory elements (promoters and enhancers)? Questions like these can already be answered in yeast and other fully sequenced and annotated organisms, and it is only a matter of time until we can get comprehensive answers for the mouse.

It is reasonable to assume that some or most of the genes identified here are intimately involved in fertilization, but are they critical components? Do these genes serve as potential targets for male germ cell-directed contraception? A search of the literature coupled with our own studies using targeted disruption of germ cell-specific genes suggests a high number of these genes will be required for fertility. We have disrupted three genes expressed late in germ cell development, a sperm-specific apparent sodium hydrogen exchanger (NHE; (Wang et al., 2003)), a cation channel (CatSper2 (Quill et al., 2001; Quill et al., 2003)) and an aminophospholipid transporter (Wang et al., 2004); in all cases normal sperm numbers were found in the epididymis, and sperm morphology was indistinguishable from that of wild-type spermatozoa, and yet two of the three genes (CatSper 2; NHE) appeared absolutely required for fertility. These results coupled with a search of the literature for other genes in clusters 1-5 shows that 18 of 19 germ cell-specific genes eliminated by homologous recombination are essential for normal male fertility (Table 7).

The genes in the table include proteases or proteins that interact with proteases (Adham et al., 1997; Cho et al., 1998; Shamsadin et al., 1999; Uhrin et al., 2000), protein kinases or proteins that interact with protein kinases (Miki et al., 2002), transcription factors (Pearse et al., 1997), proteins associated with chromatin (Yu et al., 2000; Adham et al., 2001; Cho et al., 2001), channels or transporters (Ren et al., 2001; Quill et al., 2003; Wang et al., 2003; Wang et al., 2004), mitochondrial-associated proteins (Narisawa et al., 2002; Nayernia et al., 2002), adhesion proteins (Baba et al., 2002), RNA polymerases (Kashiwabara et al., 2002) and genes involved in the interaction between spermatids and Sertoli cells (Mannan et al., 2003). More than 50% of these genes result in a complete loss of fertility when disrupted.

Table 7. Genes present in clusters 1-5 or known to be expressed late in germ cell development that have been disrupted by homologous recombination cause infertility or subfertility

Gene name	Symbol	Phenotype	Present in Clusters 1-5	Reference
A kinase anchoring protein 4	Akap4	infertile	+	Miki et al., 2002
Acrosin	Acr	subfertile	+	Adham et al., 1997
Cation channel of sperm 1	Catsper1	infertile	+	Ren et al., 2001
Cation channel of sperm 2	Catsper2	infertile	+	Quill et al., 2003
Cyritestin	Adam3	infertile	+	Shamsadin et al., 1999
Cytochrome c, testis	Cyct	subfertile	+	Narisawa et al., 2002
Fertilin β	Adam2	infertile	+	Cho et al., 1998
Mitochondrial capsule selenoprotein	Mcsp	infertile		Nayernia et al., 2002
Poly (A) polymerase β (testis-specific)	Papolb	infertile	+	Kashiwabara et al., 2002
Protamine 1	Prm1	infertile	+	Cho et al., 2001
Protamine 2	Prm2	infertile		Cho et al., 2001
Serine (or cysteine) proteinase inhibitor, clade A, member 5 (PCI)	Serpina5	infertile		Uhrin et al., 2000
Sperm adhesion molecule	Spam	subfertile	+	Baba et al., 2002
Sperm 1 POU-domain transcription factor	Sprm1	subfertile		Pearse et al., 1997
Testicular haploid expressed gene	Theg	fertile	+	Mannan et al., 2003
Transition protein 1	Tnp1	subfertile		Yu et al., 2000
Transition protein 2	Tnp2	subfertile	+	Adham et al., 2001
Sperm-specific NHE		infertile		Wang et al., 2003
Sperm-specific aminophospholipid transporter		subfertile (lower litter size)		Wang et al., 2004

Although tissues such as heart contain transcripts common to many other tissues (Bortoluzzi et al., 2000), a previous analysis of 49 tissues demonstrated that testis contained more outliers than other tissues, supportive of a conclusion that testis might contain a higher number of cell-specific transcripts than found in most other tissues (Kadota et al., 2003). Olfactory tissue, like testis, also contains a large number of specifically expressed transcripts, in this case a large family of odorant receptors. It has been suggested that there are more than 1500 such receptors in the mouse and 900 in the human (Young and Trask, 2002). Twenty percent of these genes in the mouse are predicted to be pseudogenes, and 63% in human (Quignon et al., 2003).

Our analysis indicates that a rather staggering number of germ cell-specific genes exist as possibly important participants in fertilization. These genes are potential targets for male-directed contraception and are candidates to explain many cases of male infertility. Based on the genes already disrupted, the inhibition of many of these genes seems likely to lead to male infertility. A vast majority of germ cell-specific transcripts are expressed only late in development (during or after meiosis), suggesting that most of their protein products are likely present in mature spermatozoa. Because the genes are apparently expressed exclusively in germ cells, inhibitory drugs would be expected to have few side effects.

6.2 The spermatogonial stem cell

6.2.1 Spermatogonial stem cell genes

The introduction of a functional transplantation assay for spermatogonial stem cells (Brinster and Avarbock, 1994; Brinster and Zimmermann, 1994) (see Fig. 4) has greatly facilitated research on male germ stem cells in recent years. This technique now allows the transfer of genetically modified spermatogonial stem cells. The ability to produce transgenic animals through the germline has raised hope for successful gene targeting in mammals other than the mouse. Thus far, gene disruption can only be performed in the mouse, principally because of the ability to maintain pluripotency of embryonic stem cells in culture. The laboratory rat however is one of the most comprehensively studied mammals with more than a million publications in a wide range of medically relevant areas (Gill et al., 1989; Hedrich, 2000) and provides an important physiological model system. Targeted gene disruption in this animal could prove invaluable to biology and medicine. In the absence of pluripotent ES cells, one approach to generate a knock-out rat would be through genetic modification of male germ cells in culture, followed by differentiation into spermatids and the production of pups through egg intracytoplasmic injection. However, due to the lack of a culture system that can produce spermatids, modifying spermatogonial stem cells and subsequent testicular transfer appears to be the best option to generate a transgenic rat. Improved culturing techniques with stem cell enrichment and a more efficient assay system could help achieve this goal.

We defined conditions to maintain spermatogonial stem cells *in vitro* for long periods of time with no significant decrease in stem cell activity based on colonization of a recipient testis. We also defined culture conditions under which spermatogonial stem cell activity is lost in the face of retention of male germ cell character. This has allowed us to identify gene transcripts that disappear prior to or coincident with a loss in stem cell activity or to define transcripts that remain highly expressed when stem cell character is maintained (Fig. 26; see Table 6). We developed a stem cell index for both rats and mice (see Fig. 22) and used one of the robust index proteins, Egr3, to identify and characterize spermatogonial stem cells in culture (see Fig. 23).

Genes of particular interest in Table 6 include *Ret* and *Gfra1*, which encode membrane bound receptors previously reported as present in undifferentiated spermatogonia (Meng et al., 2000; Tadokoro et al., 2002). These receptors appear to signal spermatogonial stem cell fate in response to GDNF (Meng et al., 2000; Tadokoro et al., 2002).

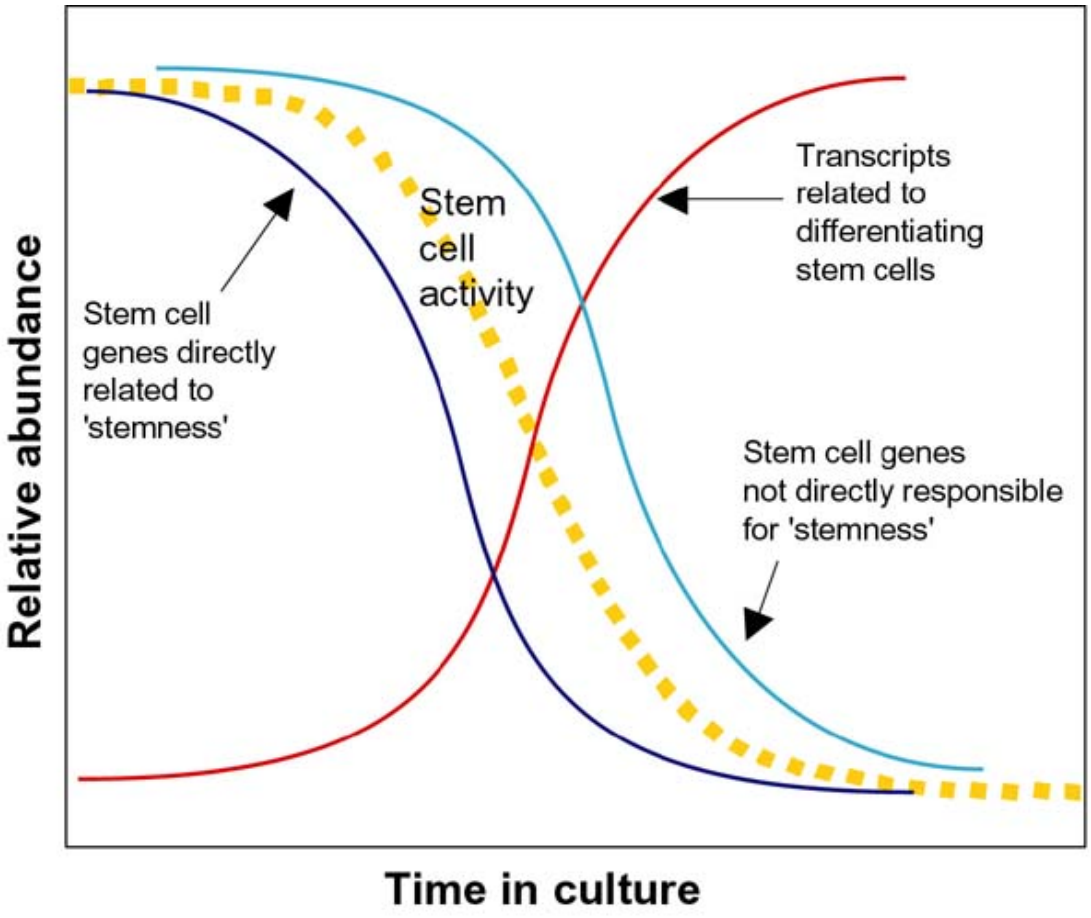

Figure 26. Gene expression related to loss of stem cell activity in culture.
Hypothetical relative transcript accumulation as a function of time after stem cells are placed in an environment where stem cell activity disappears coincident with differentiation. The model assumes that gene expression is the major determinant of 'stemness'. If stemness is determined principally by posttranslational modifications, the time course of the gene expression profile, or even the steady state level of transcripts, may have little relevance to maintenance of 'stemness'.

Several genes in Table 6 encode transcription factors, such as Hey-1, which regulates maintenance of neuronal precursors (Sakamoto et al., 2003), Egr3 and Egr4, which are members of an early growth response family (O'Donovan et al., 1999; Tourtellotte et al., 1999), Aebp1, a transcriptional repressor with carboxypeptidase activity that blocks differentiation of cultured preadipocytes (Kim et al., 2001), the forkhead transcription factor (Foxo1) that stimulates skeletal muscle satellite cell proliferation and blocks differentiation of cultured preadipocytes (Machida et al., 2003; Nakae et al., 2003), Foxa2, a second forkhead family member expressed coincident with endodermal differentiation of embryonic stem cells (pathway is repressed by Oct4 and FoxD3) (Guo et al., 2002), and Pou3f1 (Oct6/Tst-1/SCIP), an octamer-binding protein expressed by regenerating neuronal precursor cells (Kawasaki et al., 2003). Oct6 is expressed at high levels in cultures of undifferentiated embryonic stem cells or oligodendrocytes and is rapidly down-regu-

lated upon differentiation (Meijer et al., 1990; Collarini et al., 1992). In the mouse testis, Oct4 is a transcription factor known to be expressed by gonocytes and type A spermatogonia, but which is then down-regulated in more differentiated male germ cells entering meiosis (Pesce et al., 1998). *Oct4* was not detected as present in the germ cell cultures or in the testis using the rat and mouse arrays. In contrast, *Oct4* is detected as present at substantial levels in mouse embryonic stem cells using mouse Affymetrix microarrays (data not shown), suggesting that the *Oct4* transcript is more abundant in cultures of embryonic stem cells than in cultures of type A spermatogonia. Another transcription factor, Stra8, known to be expressed by mouse spermatogonial cells (Oulad-Abdelghani et al., 1996) was present at highest levels in cultures of mouse lam_B cells (signal intensities = lam_B cell, 1457; lam_{NB} cells 260, somatic cells, 44). An apparent rat homologue for Stra8 was not found in the NCBI database. Thus, it is not known if transcripts such as *Stra8*, which are more abundant in mouse lam_B cells than in lam_{NB} cells and which are not identified on the rat microarrays, would meet the criteria established for Table 6. In the mouse testis, the *Stra8* transcript peaks in relative abundance at day 11 and then dramatically drops, consistent with a peak expression of Stra8 protein in differentiated spermatogonia (Oulad-Abdelghani et al., 1996).

In contrast to other stem cells, spermatogonial stem cells in culture appear to express rather low levels of cell cycle progression transcripts, suggesting that these cells are relatively quiescent. *In vivo*, the duration of the cell cycle (~56 h) for rat type A_s spermatogonia is longer than for type A spermatogonia committed to differentiation (~42 h; types A_2 to A_4) (Huckins, 1971c; Huckins, 1971a). This is also the case in hamsters, where the cell cycle duration is ~90 h for type A_s spermatogonia, but is shortened to ~60 h in differentiating type A spermatogonia (types A_2 to A_4) (Lok and de Rooij, 1983a; Lok et al., 1983). More recent colonization assays have also demonstrated a relatively long doubling time of ~67 h for donor populations of rat spermatogonial stem cells (Ryu et al., 2003). However, due to the lack of molecular markers specific for authentic spermatogonial stem cells across different species, there is still debate about which populations of spermatogonia have the potential to function as a stem cell. The relative low abundance of transcripts involved in cell cycle kinetics listed in Table 6 further suggests similarities between type A spermatogonia within cultures of lam_B germ cells (prior to differentiation on STO cells) and the slow-cycling stem cell populations previously studied in the testis (Clermont and Bustos-Obregon, 1968; Huckins, 1971d; Lok et al., 1984).

Spermatogonial stem cells also express numerous genes involved in extracellular matrix production and attachment. Genes in this category may function to help spermatogonia colonize their niche on the basement membrane of the seminiferous tubules (Chiarini-Garcia et al., 2003; Ryu et al., 2003). Here (see Fig. 14) and in earlier studies (Shinohara et al., 1999; Shinohara et al., 2000a; Hamra et al., 2002; Orwig et al.,

2002b), the attachment of spermatogonial stem cells to the extracellular matrix protein, laminin, has been used for enrichment during isolation. In fact, the first defined surface antigens on functional spermatogonial stem cells were laminin receptor subunits encoding the integrins, *Itga6* and *Itgb1* (Shinohara et al., 1999). Similar to selection on laminin matrix, these and other cell surface antigens have been successfully used in cell sorting and subsequent transplantation assays to enrich spermatogonial stem cell activity. The cells isolated were characterized as side-scatter^{lo}, integrin-α6^{+}, integrin-β1^{+}, Thy-1^{+}, CD24^{+}, integrin-αv^{-}, MHC-Ia/β2M^{-}, Sca-1^{-}, CD34^{-} and Kit^{-} (Shinohara et al., 1999; Shinohara et al., 2000b; Kubota et al., 2003; Kanatsu-Shinohara et al., 2004). However, we find that transcripts encoding the prominent marker proteins Itga6 (integrin-α6), Itgb1 (integrin-β1) (Shinohara et al., 1999; Shinohara et al., 2000b) and Thy1 (thymus cell surface antigen) (Kubota et al., 2003) relatively abundant in cultures of differentiated germ cells and testicular somatic cells (see Fig. 21). Since stem cell activity is almost undetectable in our differentiated culture samples (lam_{NB} cells and lam_{B} cells cultured on STO feeder cells for 20 days), these transcripts do not fit our criteria for inclusion in Table 6.

It should also be noted that enrichment of stem cell activity through cell sorting using the above markers has in part relied on the generation of cryptorchid animals as a source of testicular cells. With this treatment plus two subsequent FACS to remove testicular somatic cells and differentiated germ cells, a preparation of mouse spermatogonial stem cells can be enriched 166-fold (Shinohara et al., 2000b) and potentially up to 600-fold (Kubota et al., 2003) relative to adult non-cryptorchid testis. However, with respect to cell surface antigens specifically expressed on spermatogonial stem cells in the testis, *Sema4D*, *Cacna1g*, *Ret* and *Gfra1* were the only such transcripts in Table 6 that appeared exclusively in mouse and rat lam_{B} cells; these transcripts were absent in cultures of mouse or rat somatic (interstitial and tubular) or differentiated germ cells. These gene products, therefore, could also be important environmental detectors within the spermatogonial stem cell niche.

Mouse spermatogonial stem cells also have been enriched by cell sorting, based on germ line-specific expression of transgenes. A population of GFP^{+}, Kit^{-}, integrin-α6^{+}, EE-2 antigen^{+} spermatogonial stem cells have been enriched by about 50-fold from testicular cords of 7 day old transgenic mice. In testes of these animals, GFP is expressed specifically in the germ line under control of a 18-kb region of the *Oct4* promoter (Ohbo et al., 2003). In comparison to adult animals, a 700-fold enrichment in spermatogonial stem cell activity was achieved through the use of transgenic mice that ectopically express a non-functional CD4 antigen on premeiotic and meiotic germ cells under control of a 400 bp region of the *Stra8* promoter (Giuili et al., 2002). Thus, germ line-specific transgene expression in the mouse is a means by which to dramatically simplify the isolation of large numbers of highly purified spermatogonial stem cells.

Thus, these studies help to define the spermatogonial stem cell in both mice and rats and provide a list of transcripts that not only correlate strongly with spermatogonial stem cell activity at steady state, but also identify gene transcripts that change coincident with or prior to a loss of stem cell activity. These results suggest it is now possible to study spermatogonial stem cells in culture without the primary need to initially perform testicular transfers (see Fig. 4). Such transfers could now become a secondary means by which to define stem cell activity.

By removing the obstacle of transplantation it will be possible to assess the stem cell character of germ cell populations more efficiently. This can help to further improve the culture and enrichment of spermatogonial stem cells. The expansion of spermatogonial stem cells will contribute to general stem cell research and can potentially enable targeted gene disruption in the rat.

6.2.2 Comparison with other stem cells

Stem cells exist in several parts of the body and at different times during the life of an organism. They range from unspecialized cells which can give rise to nearly all cell types in the body (embryonic stem cells) to specialized stem cells with a more limited repertoire of cells they can become. Unspecialized stem cells tend to exist in very early embryonic and fetal life. More specialized stem cells are found in adults. The ability of stem cells to self-renew and to give rise to a large number of cell types make them very special cells with enormous potential for regenerative medicine. Stem cells offer the ability to regrow cells and tissue destroyed by disease or aging process, e.g. nerve cells (neurons) in Parkinson's disease or heart muscle damaged by a heart attack.

Embryonic stem cells are relatively easy to derive and culture, and they have the potential to form most cell types and tissues in the body, making them an ideal system to work with. However, there has been much controversy about the ethicality of the use of human ES cells. Adult stem cells could be a less controversial alternative. Recent research has indicated that even these specialized stem cells may be more versatile than previously thought and give rise to multiple diverse cell types when transplanted, an ability often referred to as plasticity.

Stem cells differ from other kinds of cells in the body. All stem cells have three general properties: they are capable of dividing and renewing themselves for long periods; they are unspecialized; they can give rise to specialized cell types. It has been suggested for stem cells as diverse as neural, hematopoietic and embryonic, that there may be a set of

genes that is enriched in all stem cells and that could be responsible for the stem cell properties (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Evsikov and Solter, 2003; Fortunel et al., 2003; Ivanova et al., 2003). We tried to compare the results from these studies to the transcriptional profile of the spermatogonial stem cell.

Although previous work on other stem cells has not analyzed the kinetics of gene transcript changes relative to variations in stem cell activity, the gene products listed in Table 6 can be compared to gene transcripts that are enriched in cells such as neural, hematopoietic and embryonic stem cells, relative to a select number of enriched tissue fractions (Ramalho-Santos et al., 2002). Transcriptional regulators and signal transduction modulators seem more highly expressed in all of these stem cells, as well as in spermatogonial stem cells, however, only a few of the molecules within these functional categories are represented by the same gene transcripts. In fact, only five genes listed in Table 6 (*Elovl6*, *F2r*, *Rpl22*, *Sh3d19*, *5730599O09Rik*) are also enriched in neuronal, embryonic and hematopoietic stem cell populations (Fig. 27).

All five genes are in the functional groups of signal transduction and transcriptional regulation. *F2r* encodes a protease-activated receptor (PAR1; thrombin receptor) which has established roles in platelet aggregation, chemotaxis, proliferation and differentiation (Macfarlane et al., 2001). *Elovl6* encodes a recently discovered long chain fatty acid elongase postulated to convert C16:0 palmitic acids to C18:0 steric acids (Moon et al., 2001). The nucleolar, ribosomal protein encoded by *Rpl22* is postulated to modulate viral gene expression in eukaryotic hosts (Toczyski et al., 1994). The human homologue for the protein encoded by *5730599O09Rik*, TGIF2, interacts with histone deacetylase 1 and functions as a transcriptional silencer (Melhuish et al., 2001), while *Sh3d19*, based on conserved domains, is predicted to encode a protein that signals organization of the cytoskeleton (Shimomura et al., 2003).

When all spermatogonial stem cell genes in Table 6 are compared to genes enriched in individual populations of these other stem cells, only 25, 22, and 26 genes are enriched in neuronal, hematopoietic and embryonic stem cells, respectively (Ramalho-Santos et al., 2002).

These results imply that most genes functioning in spermatogonial stem cells that maintain or correlate with stem cell activity are different across the various types of stem cells. Such a hypothesis is supported by recent studies that identified genes commonly expressed by diverse stem cell populations of embryonic and somatic origin; these “stemness” transcripts displayed little overlap when data between different groups were compared (Evsikov and Solter, 2003; Fortunel et al., 2003; Ivanova et al., 2003; Vogel, 2003).

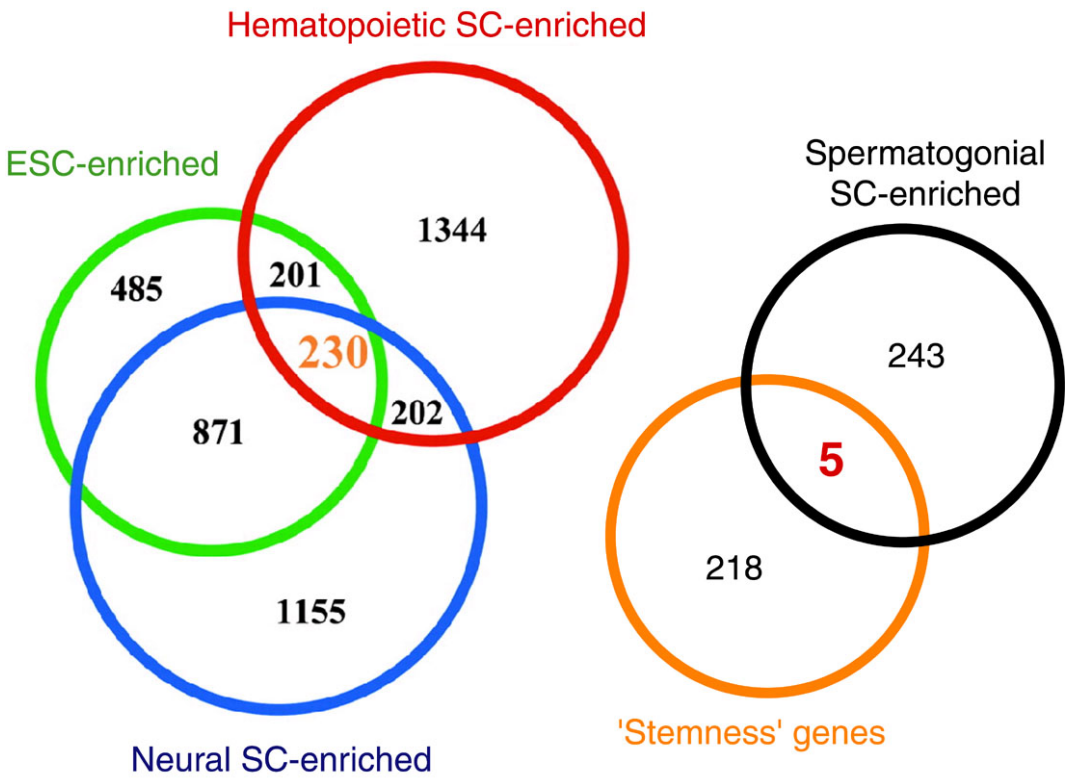

Figure 27. Comparison of genes enriched in different types of stem cells.
Ramalho-Santos and colleagues identified 223 genes in embryonic, hematopoietic and neural stem cells that were commonly enriched over their respective differentiated cell types *(Left)*. However, comparison of these 223 genes with the genes that we found enriched in spermatogonial stem cells reveals an overlap of only 5 genes *(Right)*.

Thus, steady state transcript abundance, although a measure of the character of stem cells, fails to specifically define those genes primarily critical to stem cell maintenance. However, these studies also indicated that differences in experimental design and data analysis complicate comparisons between groups (Fortunel et al., 2003; Ivanova et al., 2003). Here, we set forth a hypothesis that if specific genes are required for maintenance of stem cell activity, then in the process of moving from a stem cell to a differentiated cell, candidate genes directly linked to stemness should be represented by transcripts that change in abundance prior to or coincident with a loss of stemness. For example, based on these criteria, genes in clusters C3 and C4 of Fig. 20B are eliminated as stem cell maintenance candidates. In these groups most transcripts decrease in lam_B germ cells well after a loss of stem cell activity on STO feeder layers.