RNA Expression in Male Germ Cells During Spermatogenesis (Male Germ Cell Transcriptome)

8

Tin-Lap Lee, Albert Hoi-Hung Cheung, Owen M. Rennert, and Wai-Yee Chan

Abstract

Spermatogenesis is a key process in mammalian reproduction. This highly ordered process requires precise and well-controlled programs governed by dynamic patterns of gene expression. Some genes are exclusive to spermatogenic cells, while others are closely related to genes expressed in somatic cells. Although key genes in male germ cell development have been identified, the biological mechanisms and transcripts that govern the programs of spermatogonial stem cell renewal, germ cell differentiation during spermatogenesis, or fertilization remain largely unknown. This is partly due to the lack of information on the identity of genes involved. However, with the advent of various high-throughput genomic assays, it is now possible to obtain the whole-genome RNA expression. This chapter provides a brief account of current knowledge of the male germ cell transcriptome as revealed by studies using expression profiling platforms such as microarray and Serial Analysis of Gene Expression (SAGE). Major findings with regard to transcriptional regulation, transcript diversity, and chromatin-related regulation during male germ cell development are reviewed.

Keywords

RNA expression in germ cells • Spermatogenesis • Germ cell transcriptome • Male germ cell development • Serial analysis of gene expression

T.-L. Lee (✉)
Reproduction, Development and Endocrinology Program, School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

Laboratory of Clinical and Developmental Genomics, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA
e-mail: leetl@cuhk.edu.hk; leetl@mail.nih.gov

Spermatogenesis is a highly regulated developmental process occurring in the seminiferous tubules of the testis. The process begins with the asymmetric division of spermatogonial progenitor cells (spermatogonia), followed by meiosis to form spermatocytes, postmeiotic differentiation to form spermatids, and finally giving rise to mature spermatozoa (Fig. 8.1). Mouse male germ

A. Zini and A. Agarwal (eds.), *Sperm Chromatin: Biological and Clinical Applications in Male Infertility and Assisted Reproduction*, DOI 10.1007/978-1-4419-6857-9_8,

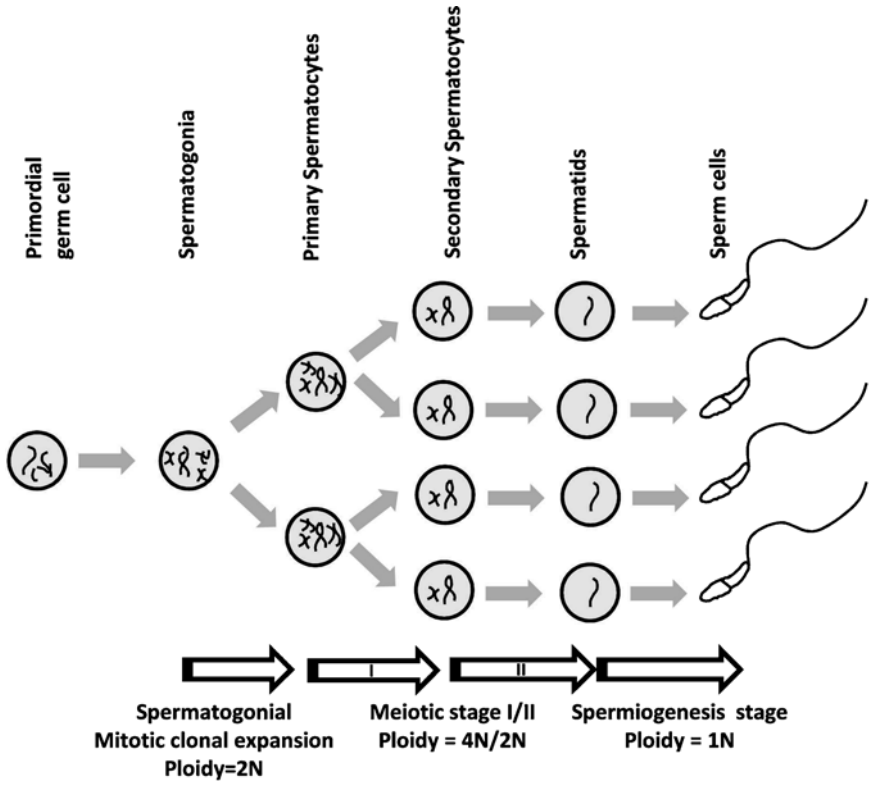

Fig.8.1 Overview of spermatogenesis. Spermatogenesis is the developmental process by which spermatogonial stem cells differentiate to pachytene spermatocytes, followed by formation of haploid spermatids by meiosis. Male germ cell genome displays several features unique to germ cells only. First, a subset of spermatogonia (type A spermatogonia) undergo mitosis during self-renewal, whereas committed spermatogonia (type B spermatogonia) undergo meiosis to generate haploid spermatids. Primary spermatocytes replicate their genomes during S-phase, followed by meiosis I to form secondary spermatocytes and subsequently meiosis II to form haploid spermatids. Meiosis only occurs in the germ line

cells provide an ideal model for studying the biology of spermatogenesis. This process has been well studied in the mouse with established developmental milestones starting at the derivation of primordial germ cells from the embryonal ectoderm. Embryonic staging of the developing male gonad is accomplished morphologically, or with genetic markers such as the Sex-determining Region Y (*Sry*) [1, 2]. Male germ cells, at different stages of development, have different density, distinct morphology, and stage-specific surface markers. These features serve as the basis of methods for preparation of relatively pure populations of germ cells at different stages of development. Relatively pure preparations of gonocytes can be obtained using laser capture techniques [3, 4]. All germ cells present in the testis of 6-day-old mice are type A spermatogonia (Spga). In adult mice, germ cells at all stages of development are present, and the different cell types can be separated based on their density using the STAPUT procedure [5].

Spermatogenesis consists of a number of hallmark developmental stages: germinal stem cells undergoing self-renewal, Spga progenitor cells at the juncture of renewal and proliferation, pachytene spermatocytes (Spcy) undergoing meiosis, and round spermatids (Sptd) undergoing postmeiotic differentiation. Therefore, studying genetic events occurring during these stages of spermatogenesis will permit a comprehensive look at the genetic events that underlie cellular proliferation and differentiation.

Spermatogenesis is a complicated process. Each step of spermatogenesis is precisely regulated. Studying the genetic programs controlling

proliferation and differentiation of male germ cells will provide an insight for understanding infertility and will allow for development of new approaches for male contraception. Demonstration of pathways specific for different stages of spermatogenesis would allow for identification of novel targets and diagnostic markers for intervention or enhancement of the male reproductive process. Knowledge of factors regulating cellular proliferation, as contrasted to differentiation, may be applied to study developmental regulation of other cell types, including stem cells.

The Transcription Landscape in Male Germ Cell Development

Little is known about the underlying mechanisms of stage-specific regulation of gene expression during spermatogenesis despite its biological importance in the genetic regulation of germ cell-specific transcripts during development [6, 7]. Limited knowledge of germ cell gene expression and the lack of a systematic approach for pathway discovery have hampered identification of biological pathways active in these cells. With the completion of various genome projects and the availability of high-throughput expression assays, a better understanding of the male germ cell transcriptome becomes a reality.

Expression Profiling of Male Germ Cell Transcriptome: Past, Present, and Future

From cDNA Library to High-Throughput Sequencing

The first attempt to characterize the male germ cell transcriptome was reported by McCarrey et al. [8]. A total of 23 cDNA libraries representing various developmental stages of mouse and rat testicular cells were constructed. Direct comparison between the cDNA populations in various cells provided the basis for the demonstration of differential gene expression. Though theoretically feasible, this approach is laborious in practice. "Deep sequencing" of cDNA libraries is required before a near-complete picture of the transcriptome can emerge. Because of these considerations, the use of cDNA libraries for transcriptome analysis was not popular.

Transcriptome profiling was a tedious job until the application of microarray platforms. A microarray is a solid support on which DNA probes of known sequence are deposited. The probes may take the form of oligonucleotides, cDNA, or DNA fragments. These probes are hybridized to sequences present in the sample. Depending on its resolution, a whole-genome human microarray chip may contain more than two millions probes. DNA microarrays were originally developed for high-throughput gene expression analysis. But they can also be applied in genetic analysis to detect single-nucleotide polymorphisms or gene copy number variation. Their fast, comprehensive, and flexible nature makes them an indispensable tool in the postgenomic era.

Another popular, widely adopted expression assay is Serial Analysis of Gene Expression (SAGE) [9]. It offers distinct advantages over other expression profiling methods by efficiently detecting polyadenylated transcript populations by sequencing short tags, usually 14–26 bp in length [8, 9]. The tags are first isolated from an anchoring enzyme restriction site (e.g., *NlaIII*) closest to the poly(A) tail of the transcripts. These tags are linked together to form long concatemers that are cloned into vectors generating a SAGE library. A SAGE library is sequenced to the desired depth. Expression of particular transcripts is quantified by the count of the associated SAGE tags in the SAGE library. Once the tags are extracted and counted, the identity of the transcript may be mapped with the SAGEmap database [9]. SAGE provides three important features over microarrays for transcriptome analysis. First, the absolute nature of tag counts allows direct comparison, without normalization and limitation of platform incompatibility, in microarray experiments. Second, since tag-to-transcript mapping in SAGE may be updated with the most current genome information, the transcriptome information provided by SAGE library is eternal. SAGE analysis allows identification of novel transcript species since prior knowledge of transcripts is not required. Finally, microarray analysis provides no

orientation information on the transcript, whereas SAGE can differentiate the sense and antisense population in the transcriptome.

While sequence-based transcriptome analysis provides additional advantages, it is cumbersome and slow, with relatively high performance cost ($0.10 per 1,000 bases). However, this is rapidly changing due to the continued improvement of sequencing technologies. The 454 sequencer was introduced in 2005 and was shortly followed by newer and faster sequencers such as Illumina and SOLiD. These technologies are referred to as "next-gen" sequencing [10]. They offer faster (up to 100×) and more cost-effective (up to 1/2,000 of the price) sequencing than conventional methods. Transcriptome analysis at single-base resolution, known as RNA-seq, is now possible [11–13]. Next-gene sequencing will be an important tool for transcriptome analysis in coming years. The huge quantity of generated data by these technologies poses great challenges to experimental biologists [14].

Overview of Germ Cell Transcriptome Studies

A list of male germ cell transcriptome studies is shown in Table 8.1. Most of these were performed on microarray platforms [15–31]. This is because sample preparation and experimental protocols are simpler when using the microarray platform. Additionally, less RNA is required as compared to SAGE [32], cDNA library [8] and differential display [33] methods. Oligonucleotide microarrays are more popular than cDNA microarrays partially because cDNA microarrays are more prone to variation in slide quality and experimental protocol. Another reason is that they are more affordable. Nevertheless, each expression platform has its strength and weakness. Renormalization against known references and correction by using statistical models are required to compare data from different studies.

Key Biological Findings and Implications

Based on the transcriptome data provided by the studies listed (Table 8.1), a number of conclusions about the dynamic changes of the transcriptome of developing male germ cells can be drawn:

Active genome-wide transcription during spermatogenesis. A major observation is that the genome is actively transcribed during germ cell development. It was previously suggested during testis development from birth to adulthood up to 58% of the mouse genome was transcribed [18, 22, 30]. Among the described transcripts, some were either male germ cell-specific or testis-predominant. About 2.3% of the rat testicular transcriptome was testis-specific [30], and ~4% of the mouse genome was only transcribed in male germ cells [18]. Many differentially expressed transcripts were unknown or uncharacterized. Examples include uncharacterized full-length cDNA transcripts, express sequence tags (ESTs), large open reading frames (ORFs), predicted transcripts of hypothetical proteins, and cross-species and predicted transcripts derived from orthologs and homologs. Depending on the cell preparation and experimental platform, the percentage of uncharacterized transcripts ranged from 40 to 60% [18, 32, 34]. Meta-analysis of these transcripts suggested that they demonstrated similar expression trends. These results imply that these transcripts were truly expressed at a particular germ cell stage.

Dynamic expression pattern in conjunction with specific developmental regulation. Transcript overexpression, as revealed by measurement of polyadenylated RNA levels in meiotic and post-meiotic male germ cells, was documented in an earlier study in rats [35]. Such phenomena might be a bystander effect occurring as a consequence of an open chromatin structure, which leads to overall activation of the transcriptional machinery in a specified cell type [7]. Alternatively, it may be a mechanism for maintaining transcript availability in response to cessation of gene transcription due to chromatin condensation during spermiogenesis [36]. Based on global gene expression analyses in various transcriptome studies, germ cell transcriptome exhibited three phases of change. The first phase, peak expression of testicular transcripts, occurred in the mitotic phase, from the day of birth to postnatal day 8, when spermatogonial proliferation predominated. The second phase occurred at the

Table 8.1 Overview of male germ cell transcriptome studies

Samples studied	Expression platform	Reference
Whole mouse adult testes, seminiferous tubule cells from adult testes, combined primary spermatocytes from 18-day-old mouse testes, type A and B spermatogonia, preleptotene, leptotene plus zygotene spermatocytes, juvenile and adult pachytene spermatocytes, round spermatids, Sertoli cells from 6, 8, 17, and 18–20-day-old mice, and peritubular cells from 18- to 20-day-old mice	cDNA library sequencing	McCarrey et al. [8]
Mouse type A spermatogonia, adult mouse wild-type testis, and W/W(v) mutant mouse testis	Differential display	Anway et al. [99]
Mouse and human testes	Microarray	Rockett et al. [15]
Human fetal and adult testes	Microarray (cDNA)	Sha et al. [16]
Mouse type A spermatogonia, pachytene spermatocytes, and round spermatids	Microarray (cDNA)	Pang et al. [17]
Mouse testes from days 1, 4, 8, 11, 14, 18, 21, 26, 29, and 60	Microarray (Oligo)	Schultz et al. [18]
Mouse type A and B spermatogonia, preleptotene and pachytene spermatocytes, round and elongating spermatids	Microarray (cDNA)	Yu et al. [19]
Mouse type A and B spermatogonia, preleptotene and pachytene spermatocytes, round and elongating spermatids	Microarray (cDNA)	Guo et al. [20]
Mouse sertoli cells, spermatogonia, spermatocytes, round spermatids	Microarray (Oligo)	Schlecht et al. [21]
Whole testes from neonates at Days 0, 3, 6, 8, 10, 14, 18, 20, 30, 35, and 56 postpartum	Microarray (Oligo)	Shima et al. [22]
Mouse adult and fetal testes	Microarray (cDNA)	Wang et al. [23]
Mouse type A spermatogonia, pachytene spermatocytes, and round spermatids	SAGE	Wu et al. [32]
Mouse sertoli cells, type A spermatogonia, spermatocytes, round spermatids	Microarray (cDNA)	Clemente et al. [24]
Testes from 17-day-old, 22-day-old, and adult mice	Microarray (Oligo)	Iguchi et al. [25]
Normal testis, patients with maturation arrest or Sertoli-cell-only syndrome	Microarray (cDNA)	Lin et al. [26]
Type A and type B spermatogonia, pachytene spermatocytes, and round spermatids	Microarray (Oligo)	Namekawa et al. [27]
Sertoli cells, spermatogonia, spermatocytes, round spermatids, seminiferous tubules, and total testis from human, rat, and mouse	Microarray (Oligo)	Chalmel et al. [28]
Testicular biopsies obtained from 289 men with azoospermia	Microarray (Oligo)	Feig et al. [29]
Rat seminiferous tubules at various stages, microdissection, sertoli cells, spermatogonia, spermatocytes, pachytene spermatocytes, and round spermatids	Microarray (Oligo)	Johnston et al. [30]
Testis samples of mice aged 4, 9, 18, 35, 54 days and 6 months	Microarray (Oligo)	Xiao et al. [31]

initiation of meiosis, on day 14, during early pachytene spermatocytes development. This was followed by entry into spermiogenesis on day 20 when round spermatids first appeared [37]. Comparison of these three phases showed increased transcript abundance in meiotic and postmeiotic stages. The number of unique genes expressed in these cells was significantly higher than that in spermatogonia [22] when up to 80% of differentially expressed genes, between meiotic and postmeiotic male germ cells, were absent or expressed at relatively low levels in type A spermatogonia [22, 34]. Increased expression of unique genes in meiotic and postmeiotic stages may imply a concomitant increase in the demand of specific gene activities for initiation and maintenance of meiosis-related events, as well as preparation for spermatozoon formation. It is noteworthy that most transcripts first expressed during or after meiosis tended to be testis- or male germ cell-specific [18, 28]. On the contrary, most genes active in spermatogonia (and Sertoli cells) were also expressed in nonreproductive tissues [28].

There was a preferential switch of active genetic loci at different stages of germ cell development. Genes related to meiotic and postmeiotic functions, and displaying higher expression level in testis, are mainly localized to autosomes [38]. By contrast, genes expressed at earlier stages of spermatogenesis are frequently localized to the X chromosome [28, 38, 39]. Similarly, many genes expressed in mitotic and somatic cells were localized on the X chromosome. A similar phenomenon was observed in a particular subset of genes, the X chromosome-derived autosomal retrogenes and their X-linked progenitor genes. Although not all testis-specific autosomal genes were X-derived retrogenes or retrogenes, the absence of X-linkage in general was believed to be a consequence of the selective force imposed by meiotic sex chromosome inactivation (MSCI) [40–42].

The dynamic and specific nature of the germ cell transcriptome was also associated with specific development and regulatory programs. Ontology analysis of the germ cell transcriptome data revealed different categories of biological processes distinctively associated with mitotic, meiotic, and postmeiotic male germ cells [28, 34, 37, 43]. For instance, processes such as integrin signaling, ribosome biogenesis and assembly, carbohydrate metabolism, protein biosynthesis, RNA processing, cell cycle, DNA replication, chromosome organization and biogenesis, and germ cell development were preferentially associated with type A spermatogonia. Surprisingly, genes involved in embryonic development and gastrulation were also found to be prevalent in these cells. On the other hand, biological processes associated with spermatogenesis and reproduction were commonly seen in meiotic and postmeiotic male germ cells. Biological processes such as meiotic cell cycle, chromatin structure and dynamics, chromosome segregation, cytoskeleton and protein degradation (ubiquitin cycle) were overrepresented in pachytene spermatocytes. Genes involved in protein turnover, signal transduction, energy metabolism, intracellular transport, ubiquitin cycle, proteolysis, peptidolysis, and fertilization were more prevalent in round spermatids.

Conserved germ cell transcriptome between human and rodents. The universal features of gametogenesis among mammalian species led to the postulation that a conserved set of genes would be involved in this process. Indeed, recent cross-species whole-genome expression profiling studies of testicular and somatic tissues in human, mouse, and rat revealed hundreds of genes that display concordant meiotic and postmeiotic expression profiles, implying the existence of a "conserved" transcriptome of mammalian spermatogenesis [28, 37]. Conserved genes involved in specific biological transitions during male germ cell development were identified by analysis of gene ontology. For example, doublesex and mab-3 related transcription factor 1 (*Dmrt1*) was found to be essential for testis differentiation; aurora kinase C (*Aurkc*), cyclin A1 (*Ccna1*) and speedy homolog A (*Spdy1*) were associated with meiotic division, whereas genes like *Socs7*, *Ankrd5*, *Fscn3*, and *Spag4l* were involved in postmeiotic regulation. Such findings suggest that rodent models could be used to study aspects of human spermatogenesis. A similar differential expression pattern of testicular genes across species suggests the presence of comparable regulatory mechanism in the control of their transcription.

In addition to changes in expression pattern of protein-encoding genes, emerging evidence identified the prominent presence in testis of non-protein-coding transcripts, including antisense transcript, small and long noncoding RNAs. These novel transcript species have been implicated to play important roles in mammalian testis development [44, 45]. The complexity of the spermatogenic process led to the search for male germ cell-specific transcripts derived from alternative splicing of somatic genes. Additionally, many germ cell genes derived from sex-linked progenitor genes through retroposition to generate testis-specific isoforms of gene products were identified. The limitations of design and probe set information inherent in microarray analysis restrict its capacity to identify non-protein-coding and alternative spliced transcripts. This is a consequence of the need, when using microarray analysis, to have prior sequence knowledge of the

transcripts, and whether it is a coding or noncoding sequence to be identified. This problem could be resolved by using the nonstatic and unguided approach of SAGE [46].

Revealing Transcription Complexity of Male Germ Cell Development by Serial Analysis of Gene Expression

Using SAGE, we examined the transcriptomes of mouse Spga, Spcy, and Sptd. SAGE libraries were constructed and sequenced to a comparable depth (~150,000 SAGE tags). A total of 34,619 transcripts were identified among the germ cell libraries. Over 2,700 of them were novel. This represents the most comprehensive male germ cell transcriptome data available. The details and related data of this analysis can be accessed at http://nichddirsage.nichd.nih.gov/publicsage/. The data obtained by the SAGE studies provide a rich resource for germ cell transcriptome discovery. By developing various bioinformatics algorithms, we succeeded in exploiting the SAGE data to decode a number of complex regulatory mechanisms and transcript species that could not be archived by microarray analyses [47–51].

Alternative Splicing

The use of multiple promoters and transcription start sites is one mechanism to create gene diversity in spermatogenesis. Alternative promoter usage allows cells to generate isoforms as well as to establish tissue specificity [52]. A large number of testis-specific splicing variants have been reported. For example, GH-releasing hormone (*GHRH*) is expressed in hypothalamus and placenta of rat. The use of a spermatogenic-specific promoter and alternative transcription initiation allows testicular germ cells to express testis-specific isoforms [53]. Expression of the testis-specific *HEMGN* mRNA (*HEMGN-t*) is developmentally regulated and synchronized with the first wave of meiosis in prepuberal mice. *HEMGN-t* is transcribed by use of alternative promoters and polyadenylation sites, suggesting a role for this testis-specific isoform in spermatogenesis [54]. Calspermin is a Calcium/calmodulin-dependent protein kinase triggering a signaling cascade. A testis-specific isoform is expressed in postmeiotic germ cells, possibly controlled by binding of CREM to the CRE motifs [55].

We reported the global identification and analysis of transcript variants with alternative 3′ end usage based on analysis of SAGE libraries of Spga, Spcy, and Sptd [47]. Unique SAGE tags at each stage of spermatogenesis were mapped to the SAGEmap database to retrieve the unigene cluster. Tags sharing the same unigene cluster within or among the stages were compared against different alternative splicing resources and validated by real-time PCR. The number of genes with 3′ end alternative splicing variants (3′ AS) expressed in Spga, Spcy, and Sptd was 74, 58, and 62, respectively. Two hundred and seven genes with 3′ AS were expressed in both Spga and Spcy. The number of genes expressed in both Spga and Sptd was 249, and the number expressed in both Spcy and Sptd was 158. There were 73 genes with different 3′ AS in all three stages examined. Novel variants involved in developmental and transcriptional control were identified. Examples included heat shock protein 4 (*Hsp4*), H3 histone, family 3B (*H3f3b*), and ubiquitin protein ligase E3A (*Ube3a*). In summary, SAGE not only provides a rapid global survey of the gene expression profile in the germ cell transcriptome but also allows identification of novel alternative splicing variants that may contribute to the unique characteristics of spermatogenesis. Further functional studies of these variants will provide new insight into germ cell development during spermatogenesis.

Antisense Transcription

Though antisense transcription has been recognized in prokaryotes for many years, the widespread occurrence of antisense transcripts in humans and mice has only been recently documented. Most studies on antisense transcription used a computational approach to identify the global presence of antisense transcripts or focused on a single gene. Few reports document the mechanism by which an antisense transcript is generated. A number of processes in spermatogenesis such as genomic imprinting, translation

repression, and stage-specific alternative splicing are frequently associated with antisense transcripts [7]. A systematic search for antisense transcripts in spermatogenic cells has not previously been reported.

Utilizing the germ cell SAGE database, our laboratory, employing orientation specific RT-PCR and molecular cloning, demonstrated that a significant percentage (31.1%) of differentially expressed genes in spermatogenic cells are associated with antisense transcripts [48]. Nucleotide sequence analysis of orientation specific RT-PCR products of 19 genes, as well as cloned full-length antisense transcripts, showed that antisense transcripts could potentially arise through a wide spectrum of mechanisms, including reverse transcription of sense mRNA in the cytoplasm, transcription of the opposite strand of the sense gene locus, transcription of a pseudogene, as well as transcription of neighboring genes and the intergenic sequence. Some of the antisense transcripts underwent normal and alternative splicing, 5′ capping, and 3′ polyadenylation like their sense counterparts. There were also antisense transcripts that were not capped and/or polyadenylated in the testis. In all cases, the levels of the sense transcripts were higher than that of the antisense transcripts while the relative expression in nontesticular tissues was variable. Thus antisense transcripts have complex origins and variable structure. Sense and antisense transcripts could be regulated independently.

Noncoding RNA Transcription

Mammalian cells produce thousands of noncoding RNAs (ncRNAs) of unknown function [51, 56–63]. These non-protein-coding portions of the genome often were considered "junk," but present research has highlighted that ncRNAs can have a wide range of regulatory functions. Small ncRNAs such as microRNA (miRNA) [64], Small interfering RNA (siRNAs) [65] and Piwi-interacting RNA (piRNA) [66] have been widely reported to function in various regulatory processes, including male germ cell development [67–74]. Recently, a new class of ncRNAs known as long ncRNAs (>200 bp) has also been demonstrated to function in developmental regulation, such as mouse ESCs pluripotency and differentiation [75–78]. These observations suggest that long ncRNAs may be indispensible in male germ cell development.

To identify potential specific long ncRNA involved in male germ cell development, we searched the SAGE data for the presence of long ncRNA candidates. A computational algorithm was developed to blast, map, and compare the RNA secondary structure of these candidates against various ncRNA databases, including NRED [79], RNAdb [80], fRNAdb [81], and NONCODE [82]. A total of 50, 35, and 24 potential long ncRNA candidates were identified in Spga, Spcy, and Sptd, respectively. These long ncRNA transcripts could be classified based on their association with various genomic features, such as promoter, intronic, intergenic, and antisense. Preliminary functional analysis in a P19 differentiation cell model suggested some long ncRNAs decreased remarkably following induction of differentiation by retinoic acid. The decrease was more obvious in the comparison of testes from vitamin A deficient (VAD) and control animals (Boucheron et al., unpublished). Several ncRNAs exhibited more than a 1000-fold decrease when compared to control testis. These results suggested that long ncRNA might play an active role in male germ cell differentiation and were dependent on retinoic acid-related regulatory pathways.

Germ Cell Transcriptome Informatics

The integration of genome and transcriptome data provides a powerful approach for understanding transcription regulatory networks in germ cell biology. However, the magnitude of this genomic data is a challenge for wet-lab biologists, as they require efficient informatics skills in data handling and processing. Fortunately, an emerging number of online user-friendly tools are available that allow for analysis of transcriptome data from a variety of angles, including static retrieval of data from databases and dynamic analysis at a systems biology level through integration of different biological information.

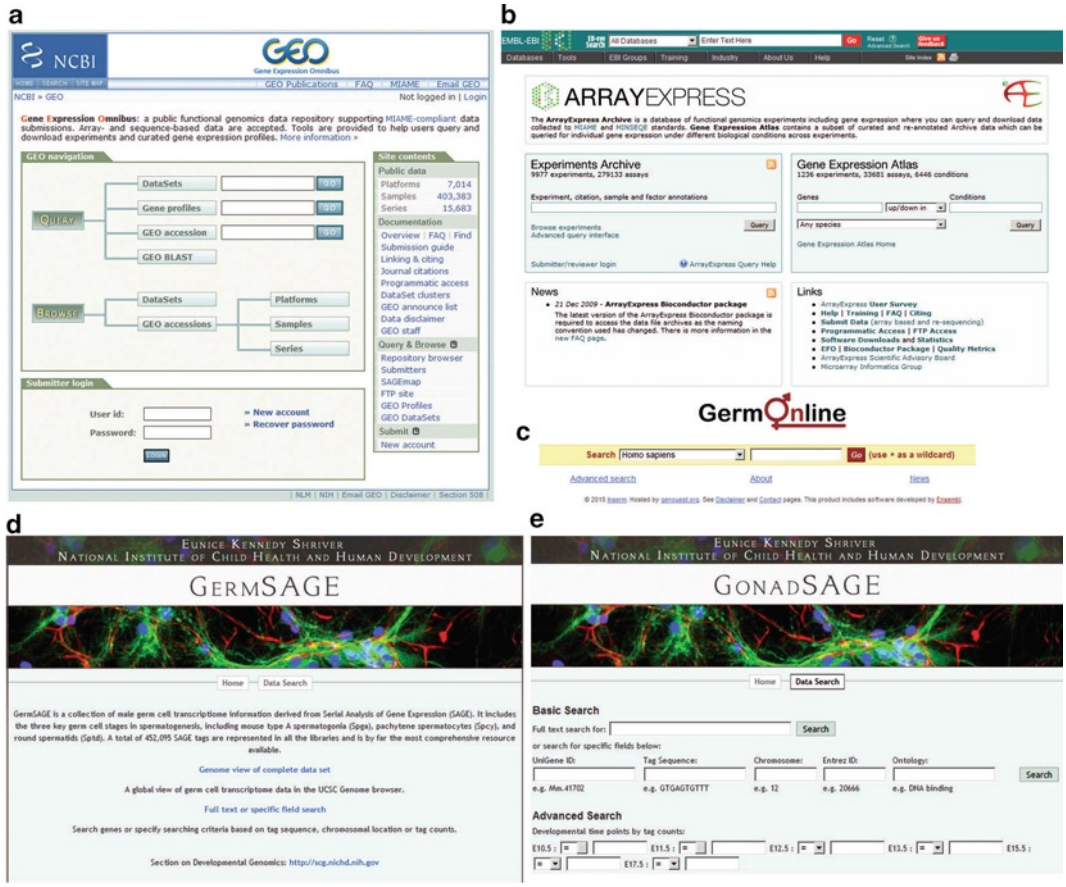

Fig. 8.2 Overview of transcriptome-related databases

Germ Cell Transcriptome Resources

The advent of various high-throughput technologies and completion of various genome projects in recent years have generated a huge amount of information. To allow effective data mining of these data in a standard format and facilitate the sharing of experimental setup and protocols, centralized public database resources were established. Currently, the most popular Web-based public repositories for transcriptome data are Gene Expression Ominbus (GEO, http://www.ncbi.nlm.nih.gov/geo/) at the National Center for Biotechnology Information (NCBI) and ArrayExpress Archive (http://www.ebi.ac.uk/microarray-as/ae/) at the European Bioinformatics Institute (EBI) (Fig. 8.2).

All data in GEO and ArrayExpress are either complied in Minimum Information About a Microarray Experiment [83] (MIAME) or Minimum Information about a high-throughput SeQuencing Experiment [84] (MINSEQE) format to enable the interpretation of experimental results in an unambiguous fashion and to potentially reproduce the experiment. A specific ID, in the form of GSE number (GEO) or E-GEOD-number (ArrayExpress), is assigned to an experiment. At the time of writing this review (Feb, 2010), the GEO and ArrayExpress database contain a total of 15,683 and 9977 experiments, respectively. In addition to GEO and ArrayExpress, online specialized resources on germ cell transcriptome are also available,

which can be divided in terms of platform as described in the next section.

Microarray-based transcriptome resource: GermOnline. GermOnline [85, 86] is a microarray expression database that focuses on mitosis-, meiosis-, and gametogenesis-related studies. It adopted a familiar Ensembl-database layout for data presentation. Currently, it covers published transcriptome data from eleven species. The data are presented in Ensembl genome browser format. A microarray information management and annotation system (MIMAS) and a comprehensive system for online editing of database entries (MediaWiki) were applied to describe experimental data from various microarray-based experiments, such as RNA expression levels, transcript start sites and lengths, and exon composition. The database also provides an open environment for scientists to maintain database entries on genes and gene products in a complete and accurate manner by submitting up-to-date curations. The database is accessible at http://www.germonline.org/.

Sequence-based transcriptome resource: GermSAGE and GonadSAGE. GermSAGE [87] (http://germsage.nichd.nih.gov) is a comprehensive Web-based database generated by SAGE. Data deposited represent major stages in mouse male germ cell development, with sequence tag coverage of 150k in each SAGE library. The database covers 452,095 tags derived from type A spermatogonia, pachytene spermatocytes, and round spermatids. It provides an array of easy tools for browsing, comparing, and screening male germ cell transcriptome data. The data can be exported or further analyzed by aligning it with various annotations available in the built-in genome browser of the database. This flexible platform is useful for gaining a better understanding of the genetic networks that regulate spermatogonial cell renewal and differentiation and allows for novel gene discovery.

GonadSAGE [88] (http://germsage.nichd.nih.gov) is another SAGE database on male gonad development. A total of six male mouse embryonic gonad stages were included (E10.5, E11.5, E12.5, E13.5, E15.5, and E17.5). The sequence coverage for each SAGE library is above 150K. A total of 908,453 SAGE tags are represented in all the libraries and is by far the most comprehensive resource available. Altogether, it contains 24,975 known and over 275,583 unannotated transcripts, including an extensive presence of antisense transcripts and splicing variants.

Chromatin Remodeling and Spermatogenesis

Background

Eukaryotic gene regulation is more complicated than prokaryotic gene regulation. Transcriptional regulation is tightly coordinated, determined not only by the genetic information stored in the DNA sequence but also by interactions between a diversity of modifiers on chromatin. The complexity of eukaryotic transcriptional control is reflected by the structure of chromatin, which is composed of small repeating units, the nucleosomes. A nucleosome consists of double-stranded DNA wrapped on histone proteins. Four histone proteins, namely, H3, H4, H2A, and H2B, respectively, form the octamer core of the nucleosome. During DNA packaging, DNA helix is deposited on the H3(2)/H4(2) tetramer, followed by incorporation of two sets of H2A/H2B dimmers. Such packaging allows the huge chromosome to be organized in a compact structure. It is postulated that the linker histone protein H1 further promotes chromatin packaging to a higher-order structure by potentially shielding the negative charge of DNA linking nucleosomes.

Histones are not solely for DNA packaging. The eukaryotic system has evolved another mechanism of gene regulation by changing the chemical nature of histone tails extruding from the nucleosomal core. This is achieved by several posttranslational modifications on the amino residues of the N-terminus of histones. Currently, known covalent modifications on histones include acetylation and ubiquitination of lysine, methylation of arginine or lysine, and phosphorylation of serine or threonine [89]. Together with methylation on cytosine of DNA, these modifications

form the epigenetic marks in mammalian genomes. By interacting with different chromatin modifiers, epigenetic marks provide an additional layer of gene regulation through establishing either an active or repressive state of chromatin. Epigenetic control of gene expression permits different cell types to express unique sets of genes despite having the same genome.

A large number of histone variants are found in male germ cells. Many of these variants are testis-specific. Expression of these testis-specific variants suggests the existence of a special nucleosomal architecture during spermatogenesis. There is nuclear reorganization in the chromatin of spermatids, where histone–protamine transition takes place. Postmeiotic haploid spermatids utilize protamine, an arginine-rich H1-like small protein, to replace histone. In addition, mammalian germ cells utilize another testis-specific nuclear protein, the transition proteins (TP1 and TP2), prior to protamine displacement. Transition proteins may not be essential for fertility since knockout of TP1 or TP2 did not result in infertility [90, 91]. It is generally believed that transition proteins replace histones, preparing the chromatin for protamine incorporation. Male germ cells use protamines to create highly compact nuclei, the size of which is about 5% of the somatic nucleus. Unlike transition proteins, protamines are essential for the development of mature sperm. Loss of protamine in mice resulted in male infertility [92]. The creation of a compact nucleus is not favorable for gene transcription. Indeed, HP1 is recruited to the heterochromatic chromocenter of spermatids after the histone–protamine transition, indicative of a silencing mechanism coupled with heterochromatin condensation (HP1 is a transcription repressor binding to methylated H3 in transcriptionally silenced heterochromatin).

Chromatin-Related Transcriptional Regulations in Spermatogenesis

Transcriptional regulation in male germ cells is different from somatic cells as evidenced by the use of histone variants, the expression of testis-specific homologs in the transcriptional machinery, and the use of alternative promoters in spermatogenesis.

For transcription to initiate, nucleosomes must be reorganized to allow access to promoters of transcription factors. Mechanisms include a transient unwrapping of the DNA from histone octamers or shifting nucleosomes along the length of DNA (nucleosome sliding). To accomplish this, chromatin remodeling complexes utilize ATP hydrolysis to disassemble the nucleosomal core, possibly by a mechanism of histone displacement. Chromatin remodeling complexes SWI/SNF, RSC and Pol II are responsible for histone displacement, with histone chapersones as the acceptor. Remarkably, different testis-specific histone variants, such as TH2A, TH2B, H2A.Z, TH3, H1t, H1t2, and HILS1, are incorporated during spermatogenesis and spermiogenesis. Incorporation of variants can change the nucleosomal structure (or the epigenetic modification on the variants' tails), thus influencing gene regulation. For instance, during spermatogenesis canonical H3 is displaced by variants H3.3A and H3.3B. H3.3 variants prominently replacing H3 at active genes [93] probably accounts for active transcription in spermatocytes. Although the role of testis-specific variants on chromatin structure and function of male germ cells is unclear, it is generally believed that the variants result in altered nucleosomal structure, creating a specialized nuclear organization that facilitates binding of chromatin remodeling factors and prepares the sperm genome for subsequent fertilization. Notably, a non-testis-specific H3 variant, CENP-A, localizes to the newly duplicated centromere of germ cells. CENP-A is not displaced during the histone–protamine transition. Its inheritance raises the speculation that it might function in fertilization.

Transcription activation involves three classes of proteins, namely, TATA-binding protein, DNA binding transactivator, and coactivator protein complex. Some transcription factors (TBP, TFIIB, RNA polymerase II) are constitutively expressed but at a much higher level in haploid germ cells [94]. Some are restricted to testis. Some, instead of expressing a regular form in somatic cells, are expressed in testes with a tissue-specific isoform.

A well-studied example is the expression of a testis-specific transcription activator CREM

(cyclic AMP-responsive element modulator). CREM is highly expressed in postmeiotic germ cells [8]. It is a homolog of CREB (cAMP response element-binding protein), an activator of cAMP-responsive promoter elements (CREs). In somatic cells, phosphorylation of CREB triggers transcription activation. However, CREB is not expressed in testes. Instead, CREM is actively expressed in haploid germ cells for transcriptional regulation of many genes critical for late spermatogenesis. CREM-mutant mice showed defective spermiogenesis and increased apoptosis of germ cells [95]. Unlike CREB, CREM is phosphorylation-independent, but activated by a coactivator ACT. Notably, ACT is also restricted to male germ cells.

The transcriptional initiation complex in germ cells contains TLF (TBP-like factor), which activates genes with TATA-less promoters. Expression of TLF is developmentally regulated in spermatogenesis. Knockdown of TLF caused complete arrest of late spermiogenesis and fragmentation of the chromocenter in early spermatids [96]. Male germ cells express homologs of other transcriptional factors in the transcriptional machinery. For example, a homolog of TFIIA (Transcription factor II A) is predominantly expressed in testes, the biochemical function of which is indistinguishable from its counterpart [97, 98]. TAF7L (TAF7-like RNA polymerase II), a paralog of TAF7 of the TFIID complex, is X-linked and testis-specific [39]. Since the X chromosome is silenced in spermatocytes and spermatids, an autosomal homolog may exist particularly for spermatogenesis.

Conclusions

Investigation into regulation of gene expression in spermatogenesis is hampered by the lack of a comprehensive understanding of gene expression in germ cells. This is further confounded by limitations of the traditional single gene–single pathway approach. In the past decade, many transcriptome studies have been conducted to examine the biology of germ cell development. With the availability of comprehensive germ cell transcriptome databases, identification and characterization of gene functions in male germ cell development become possible. It is now clear that the germ cell transcriptome is more complex than previously envisioned. It involves not only protein-encoding genes but also non-protein-coding transcripts such as antisense transcripts, small and long noncoding RNAs, etc. Dynamic regulation and usage of the germ cell transcriptome are also obvious. A significant number of male germ cell-specific transcripts undergo alternative splicing or are derived from sex-linked progenitor genes through retroposition to generate testis-specific isoforms, presumably to cope with the specific needs in the spermatogenic process. The application of genome-wide analysis and systems biology approaches should permit elucidation of more novel modes of transcription regulation and identification of biological pathways critical for male germ cell development.

Acknowledgments This work was supported in part by the Intramural Research Program of the National Institutes of Health (NIH), Eunice Kennedy Shriver National Institute of Child Health and Human Development, and the Chinese University of Hong Kong.

References

1. Berta P, Hawkins JR, Sinclair AH, et al. Genetic evidence equating SRY and the testis-determining factor. Nature. 1990;348:448–50.
2. Wallis MC, Waters PD, Graves JA. Sex determination in mammals–before and after the evolution of SRY. Cell Mol Life Sci. 2008;65:3182–95.
3. Gaskell TL, Robinson LL, Groome NP, Anderson RA, Saunders PT. Differential expression of two estrogen receptor beta isoforms in the human fetal testis during the second trimester of pregnancy. J Clin Endocrinol Metab. 2003;88:424–32.
4. Robinson LL, Gaskell TL, Saunders PT, Anderson RA. Germ cell specific expression of c-kit in the human fetal gonad. Mol Hum Reprod. 2001;7:845–52.
5. Bellve AR, Cavicchia JC, Millette CF, O'Brien DA, Bhatnagar YM, Dym M. Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J Cell Biol. 1977;74:68–85.
6. Eddy EM. Male germ cell gene expression. Recent Prog Horm Res. 2002;57:103–28.
7. Kleene KC. A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. Mech Dev. 2001;106:3–23.

8. McCarrey JR, O'Brien DA, Skinner MK. Construction and preliminary characterization of a series of mouse and rat testis cDNA libraries. J Androl. 1999;20: 635–9.
9. Lash AE, Tolstoshev CM, Wagner L, et al. SAGEmap: a public gene expression resource. Genome Res. 2000;10:1051–60.
10. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008;92:255–64.
11. Nagalakshmi U, Waern K, Snyder M. RNA-Seq: a method for comprehensive transcriptome analysis. Curr Protoc Mol Biol. 2010; Chapter 4:Unit 4 11 11–13.
12. Marguerat S, Bahler J. RNA-seq: from technology to biology. Cell Mol Life Sci. 2010;67:569–79.
13. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10:57–63.
14. Lee TL. Big data: open-source format needed to aid wiki collaboration. Nature. 2008;455:461.
15. Rockett JC, Christopher Luft J, Brian Garges J, et al. Development of a 950-gene DNA array for examining gene expression patterns in mouse testis. Genome Biol. 2001;2:RESEARCH0014.
16. Sha J, Zhou Z, Li J, et al. Identification of testis development and spermatogenesis-related genes in human and mouse testes using cDNA arrays. Mol Hum Reprod. 2002;8:511–7.
17. Pang AL, Taylor HC, Johnson W, et al. Identification of differentially expressed genes in mouse spermatogenesis. J Androl. 2003;24:899–911.
18. Schultz N, Hamra FK, Garbers DL. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. Proc Natl Acad Sci USA. 2003;100:12201–6.
19. Yu Z, Guo R, Ge Y, et al. Gene expression profiles in different stages of mouse spermatogenic cells during spermatogenesis. Biol Reprod. 2003;69:37–47.
20. Guo R, Yu Z, Guan J, et al. Stage-specific and tissue-specific expression characteristics of differentially expressed genes during mouse spermatogenesis. Mol Reprod Dev. 2004;67:264–72.
21. Schlecht U, Demougin P, Koch R, et al. Expression profiling of mammalian male meiosis and gametogenesis identifies novel candidate genes for roles in the regulation of fertility. Mol Biol Cell. 2004;15: 1031–43.
22. Shima JE, McLean DJ, McCarrey JR, Griswold MD. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. Biol Reprod. 2004;71:319–30.
23. Wang H, Zhou Z, Xu M, et al. A spermatogenesis-related gene expression profile in human spermatozoa and its potential clinical applications. J Mol Med. 2004;82:317–24.
24. Clemente EJ, Furlong RA, Loveland KL, Affara NA. Gene expression study in the juvenile mouse testis: identification of stage-specific molecular pathways during spermatogenesis. Mamm Genome. 2006;17:956–75.
25. Iguchi N, Tobias JW, Hecht NB. Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. Proc Natl Acad Sci USA. 2006;103:7712–7.
26. Lin YH, Lin YM, Teng YN, Hsieh TY, Lin YS, Kuo PL. Identification of ten novel genes involved in human spermatogenesis by microarray analysis of testicular tissue. Fertil Steril. 2006;86:1650–8.
27. Namekawa SH, Park PJ, Zhang LF, et al. Postmeiotic sex chromatin in the male germline of mice. Curr Biol. 2006;16:660–7.
28. Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, et al. The conserved transcriptome in human and rodent male gametogenesis. Proc Natl Acad Sci USA. 2007;104:8346–51.
29. Feig C, Kirchhoff C, Ivell R, Naether O, Schulze W, Spiess AN. A new paradigm for profiling testicular gene expression during normal and disturbed human spermatogenesis. Mol Hum Reprod. 2007;13:33–43.
30. Johnston DS, Wright WW, Dicandeloro P, Wilson E, Kopf GS, Jelinsky SA. Stage-specific gene expression is a fundamental characteristic of rat spermatogenic cells and Sertoli cells. Proc Natl Acad Sci USA. 2008;105:8315–20.
31. Xiao P, Tang A, Yu Z, Gui Y, Cai Z. Gene expression profile of 2058 spermatogenesis-related genes in mice. Biol Pharm Bull. 2008;31:201–6.
32. Wu SM, Baxendale V, Chen Y, et al. Analysis of mouse germ-cell transcriptome at different stages of spermatogenesis by SAGE: biological significance. Genomics. 2004;84:971–81.
33. Anway MD, Li Y, Ravindranath N, Dym M, Griswold MD. Expression of testicular germ cell genes identified by differential display analysis. J Androl. 2003;24:173–84.
34. Pang AL, Johnson W, Ravindranath N, Dym M, Rennert OM, Chan WY. Expression profiling of purified male germ cells: stage-specific expression patterns related to meiosis and postmeiotic development. Physiol Genomics. 2006;24:75–85.
35. Morales CR, Hecht NB. Poly(A)+ ribonucleic acids are enriched in spermatocyte nuclei but not in chromatoid bodies in the rat testis. Biol Reprod. 1994;50: 309–19.
36. Sassone-Corsi P. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science. 2002;296:2176–8.
37. Wrobel G, Primig M. Mammalian male germ cells are fertile ground for expression profiling of sexual reproduction. Reproduction. 2005;129:1–7.
38. Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nat Genet. 2004;36: 642–6.
39. Wang PJ, McCarrey JR, Yang F, Page DC. An abundance of X-linked genes expressed in spermatogonia. Nat Genet. 2001;27:422–6.
40. Turner JM. Meiotic sex chromosome inactivation. Development. 2007;134:1823–31.

41. Long M, Betran E, Thornton K, Wang W. The origin of new genes: glimpses from the young and old. Nat Rev Genet. 2003;4:865–75.
42. Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JM. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. Nat Genet. 2008;40:794–9.
43. Lee TL, Alba D, Baxendale V, Rennert OM, Chan WY. Application of transcriptional and biological network analyses in mouse germ-cell transcriptomes. Genomics. 2006;88:18–33.
44. Amaral PP, Mattick JS. Noncoding RNA in development. Mamm Genome. 2008;19:454–92.
45. Hayashi K, Chuva de Sousa Lopes SM, Kaneda M, et al. MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. PLoS ONE. 2008;3:e1738.
46. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science. 1995;270:484–7.
47. Chan WY, Lee TL, Wu SM, et al. Transcriptome analyses of male germ cells with serial analysis of gene expression (SAGE). Mol Cell Endocrinol. 2006;250: 8–19.
48. Chan WY, Wu SM, Ruszczyk L, et al. The complexity of antisense transcription revealed by the study of developing male germ cells. Genomics. 2006;87: 681–92.
49. Lee TL, Yeh J, Van Waes C, Chen Z. Epigenetic modification of SOCS-1 differentially regulates STAT3 activation in response to interleukin-6 receptor and epidermal growth factor receptor signaling through JAK and/or MEK in head and neck squamous cell carcinomas. Mol Cancer Ther. 2006;5:8–19.
50. Lee TL, Li Y, Alba D, et al. Developmental staging of male murine embryonic gonad by SAGE analysis. J Genet Genomics. 2009;36:215–27.
51. Lee TL, Pang AL, Rennert OM, Chan WY. Genomic landscape of developing male germ cells. Birth Defects Res C Embryo Today. 2009;87:43–63.
52. Ayoubi TA, Van De Ven WJ. Regulation of gene expression by alternative promoters. FASEB J. 1996;10: 453–60.
53. Srivastava CH, Monts BS, Rothrock JK, Peredo MJ, Pescovitz OH. Presence of a spermatogenic-specific promoter in the rat growth hormone-releasing hormone gene. Endocrinology. 1995;136:1502–8.
54. Yang LV, Heng HH, Wan J, Southwood CM, Gow A, Li L. Alternative promoters and polyadenylation regulate tissue-specific expression of Hemogen isoforms during hematopoiesis and spermatogenesis. Dev Dyn. 2003;228:606–16.
55. Sun Z, Sassone-Corsi P, Means AR. Calspermin gene transcription is regulated by two cyclic AMP response elements contained in an alternative promoter in the calmodulin kinase IV gene. Mol Cell Biol. 1995;15:561–71.
56. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding RNAs: regulators of disease. J Pathol. 2010;220:126–39.
57. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10:155–9.
58. Dinger ME, Amaral PP, Mercer TR, Mattick JS. Pervasive transcription of the eukaryotic genome: functional indices and conceptual implications. Brief Funct Genomic Proteomic. 2009;8:407–23.
59. Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454:766–70.
60. Carninci P, Non-coding RNA. transcription: turning on neighbours. Nat Cell Biol. 2008;10:1023–4.
61. Mehler MF, Mattick JS. Non-coding RNAs in the nervous system. J Physiol. 2006;575:333–41.
62. Johnson JM, Edwards S, Shoemaker D, Schadt EE. Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. Trends Genet. 2005;21:93–102.
63. Kapranov P, Drenkow J, Cheng J, et al. Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. Genome Res. 2005;15:987–97.
64. Vasudevan S, Tong Y, Steitz JA. Switching from repression to activation: microRNAs can up-regulate translation. Science. 2007;318:1931–4.
65. Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science. 1999;286:950–2.
66. Seto AG, Kingston RE, Lau NC. The coming of age for Piwi proteins. Mol Cell. 2007;26:603–9.
67. Wang J, Saxe JP, Tanaka T, Chuma S, Lin H. Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. Curr Biol. 2009;19: 640–4.
68. He Z, Kokkinaki M, Pant D, Gallicano GI, Dym M. Small RNA molecules in the regulation of spermatogenesis. Reproduction. 2009;137:901–11.
69. Chi YH, Cheng LI, Myers T, et al. Requirement for Sun1 in the expression of meiotic reproductive genes and piRNA. Development. 2009;136:965–73.
70. Marcon E, Babak T, Chua G, Hughes T, Moens PB. miRNA and piRNA localization in the male mammalian meiotic nucleus. Chromosome Res. 2008;16: 243–60.
71. Iguchi N, Xu M, Hori T, Hecht NB. Noncoding RNAs of the mammalian testis: the meiotic transcripts Nct1 and Nct2 encode piRNAs. Ann N Y Acad Sci. 2007;1120:84–94.
72. Ro S, Park C, Song R, et al. Cloning and expression profiling of testis-expressed piRNA-like RNAs. RNA. 2007;13:1693–702.
73. Carmell MA, Girard A, van de Kant HJ, et al. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell. 2007;12:503–14.
74. Grivna ST, Pyhtila B, Lin H. MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc Natl Acad Sci USA. 2006;103:13415–20.

75. Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS. Specific expression of long noncoding RNAs in the mouse brain. Proc Natl Acad Sci USA. 2008;105:716–21.
76. Dinger ME, Amaral PP, Mercer TR, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 2008;18:1433–45.
77. Mehler MF, Mattick JS. Noncoding RNAs and RNA editing in brain development, functional diversification, and neurological disease. Physiol Rev. 2007;87:799–823.
78. Mattick JS. A new paradigm for developmental biology. J Exp Biol. 2007;210:1526–47.
79. Dinger ME, Pang KC, Mercer TR, Crowe ML, Grimmond SM, Mattick JS. NRED: a database of long noncoding RNA expression. Nucleic Acids Res. 2009;37:D122–6.
80. Pang KC, Stephen S, Dinger ME, Engstrom PG, Lenhard B, Mattick JS. RNAdb 2.0–an expanded database of mammalian non-coding RNAs. Nucleic Acids Res. 2007;35:D178–82.
81. Mituyama T, Yamada K, Hattori E, et al. The Functional RNA Database 3.0: databases to support mining and annotation of functional RNAs. Nucleic Acids Res. 2009;37:D89–92.
82. He S, Liu C, Skogerbo G, et al. NONCODE v2.0: decoding the non-coding. Nucleic Acids Res. 2008;36: D170–2.
83. Brazma A, Hingamp P, Quackenbush J, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet. 2001;29:365–71.
84. Brazma A. Minimum information about a microarray experiment (MIAME)–successes, failures, challenges. ScientificWorldJournal. 2009;9:420–3.
85. Gattiker A, Niederhauser-Wiederkehr C, Moore J, Hermida L, Primig M. The GermOnline cross-species systems browser provides comprehensive information on genes and gene products relevant for sexual reproduction. Nucleic Acids Res. 2007;35:D457–62.
86. Wiederkehr C, Basavaraj R, Sarrauste de Menthiere C, et al. GermOnline, a cross-species community knowledgebase on germ cell differentiation. Nucleic Acids Res. 2004;32:D560–7.
87. Lee TL, Cheung HH, Claus J, et al. GermSAGE: a comprehensive SAGE database for transcript discovery on male germ cell development. Nucleic Acids Res. 2009;37:D891–7.
88. Lee TL, Li Y, Cheung HH, et al. GonadSAGE: a comprehensive SAGE database for transcript discovery on male embryonic gonad development. Bioinformatics. 2010;26:585–6.
89. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell. 2007;128:707–19.
90. Yu YE, Zhang Y, Unni E, et al. Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. Proc Natl Acad Sci USA. 2000;97:4683–8.
91. Zhao M, Shirley CR, Yu YE, et al. Targeted disruption of the transition protein 2 gene affects sperm chromatin structure and reduces fertility in mice. Mol Cell Biol. 2001;21:7243–55.
92. Cho C, Willis WD, Goulding EH, et al. Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nat Genet. 2001;28:82–6.
93. Mito Y, Henikoff JG, Henikoff S. Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet. 2005;37:1090–7.
94. Schmidt EE, Schibler U. High accumulation of components of the RNA polymerase II transcription machinery in rodent spermatids. Development. 1995;121:2373–83.
95. Nantel F, Monaco L, Foulkes NS, et al. Spermiogenesis deficiency and germ-cell apoptosis in CREM-mutant mice. Nature. 1996;380:159–62.
96. Martianov I, Brancorsini S, Gansmuller A, Parvinen M, Davidson I, Sassone-Corsi P. Distinct functions of TBP and TLF/TRF2 during spermatogenesis: requirement of TLF for heterochromatic chromocenter formation in haploid round spermatids. Development. 2002;129:945–55.
97. Upadhyaya AB, Lee SH, DeJong J. Identification of a general transcription factor TFIIAalpha/beta homolog selectively expressed in testis. J Biol Chem. 1999;274:18040–8.
98. Ozer J, Moore PA, Lieberman PM. A testis-specific transcription factor IIA (TFIIAtau) stimulates TATA-binding protein-DNA binding and transcription activation. J Biol Chem. 2000;275:122–8.
99. Anway MD, Johnston DS, Crawford D, Griswold MD. Identification of a novel retrovirus expressed in rat Sertoli cells and granulosa cells. Biol Reprod. 2001;65(4):1289–96.