Gene Expression Profile of 2058 Spermatogenesis-Related Genes in Mice

Ping XIAO, Aifa TANG, Zhendong YU, Yaoting GUI, and Zhiming CAI*
Shenzhen Key Lab of Male Reproduction and Genetics, Peking University Shenzhen Hospital; Shenzhen 518036, P.R. China. Received July 22, 2007; accepted October 21, 2007

The aim of this study was to characterize the gene expression profiles of mouse testis at different stages by performing large-scale complementary DNA (cDNA) analysis using DNA microarrays. The transcription profile of mouse testis was determined by comparing testis samples of mice aged 4, 9, 18, 35, 54 d and 6 months using Affymetrix Genechip mouse microarrays (430 v.2). Of the six comparative pairs of two neighboring spots examined, lots of differently expressed genes existed on day 18 versus day 9 and on day 35 versus day 18, indicating these genes may be involved in unique events during those periods. In mining the Genechip data, the gene expression profiles of 2058 genes represented in a gradually increased manner in the developmental stages of mice testis, and its validity was confirmed by the data of 21 known spermatogenesis-related testis-specific genes in our chip and the results of 10 random sample from 2058 genes produced by the sub-quantitative RT-PCR. So, we called these 2058 genes spermatogenesis-related genes. By informatics analysis, the gene expression clustering, gene ontology, KEGG pathway and domain regions were predicted for describing the potential roles that may play during mouse spermatogenesis. This study provides a molecular basis for the identity of spermatogenesis-related gene in mouse testis and leads to the elucidation of the molecular events underlying mammalian male reproduction.

Key words DNA microarray; spermatogenesis; testis-specific gene

Sperm development, termed Spermatogenesis, is characterized by a mitotic ( spermatogonia), a meiotic (spermatocytes) and a differentiative haploid (spermatids) phase. This complex process is orchestrated by the expression of thousands of genes encoding proteins that play essential roles during specific phases of germ cell development. The dissection of the mechanisms that regulate the mitotic and meiotic cell cycles in mammalian germ cells is useful for screening and characterizing key genes in sperm progress, and thus for better understanding of the molecular requirements for spermatogenesis to occur. The understanding of the biological function of key genes during spermatogenesis is helpful for the treatment of male sterility caused by abnormality of sperm development, protecting from withering of male sexual potency, and offering new contraceptive targets and health care drugs.

Testes are part of a selected set of organs necessary for the continued propagation of a species that reproduces sexuality. Specifically, testes are the sites of spermatogenesis, a complex process in which the Sertoli cells of the seminiferous epithelium support the differentiation of germ cells into functionally competent spermatooza. Compared to somatic cells, germ cells have some unique events including meiosis, genetic recombination, gene expression in haploid cells, chromatin remodeling and condensation, acrosome and flagellum formation, involve unique gene products. Identification of the genes involved in the production of sperm via spermatogenesis especially in those unique events that occur in germinal cells would shed light on the mechanism analysis of mammalian spermatogenesis. As reviewed in previous work, the unique events during spermatogenesis show that the expression of these testis genes is developmentally regulated. Both transcriptional and translation control mechanisms are responsible for temporal and stage-specific expression pattern. The identification of genes encoded proteins that play specific roles in different steps of germ cell development should shed fresh light on the mechanisms analysis of spermatogenesis. The list of genes that are involved in the mammalian spermatogenesis is rapidly growing. For example, CERM, tesmin, MSJ-1, Soggy, SP-10, mT-SARG3, SRG4, TESF-19 have been proved to be the spermatogenesis-related genes in previous study.

To know how genes are controlled to express at the exact time and which genes function uniquely in specific cells during spermatogenesis, an important step is taken to define gene profiles. DNA microarray(10) is a useful, high throughput method, which provides a platform to evaluate the abundance of genes in parallel, allowing monitoring changes in gene expression during developmental events. Sha et al. (2002)(11) have compared gene expression profiles between adult and fetal human testes by using of a self-made cDNA chip that comprised of 9216 genes. 731 different expressing genes have been characterized comprising of 54 known genes, in which 18.52% were exclusively expressed in spermatogenesis. Guo et al. (2004)(12) isolated six different types of spermatogenic cells (primitive type A spermatogonia, type B spermatogonia, preleptotene spermatocytes, pachytene spermatocytes, round spermatiss and elongating spermatids) from Balb/C mice testes. Atlas cDNA arrays containing 1176 known mouse genes were used to determine the gene expression profiles of the spermatogenic cells. The expressions of 260 genes were detected and 115 genes show differential expression in six different stages of the spermatogenic cells.

Recently, we isolated testis from 4, 9, 18, 35, 54 d and 6 months old Balb/C mice. cRNAs prepared from these testis samples have been hybridized with commercially available GeneChip Mouse Genome 430 2.0 Array (Affymetrix Inc.) chip, which contained ca. 34000 known mouse genes and 8000 unkown genes or ESTs (Expressed Sequence of Tags), and thus spanning the whole mouse genome. In mining the microarray data, we identified 2058 gradually up-regulated and 903 down-regulated genes (data not show) from 4 d to 6
months of testis samples. These genes, including some known and unknown genes, should be related to the mitosis of spermatogonia, the meiosis of spermatocyte, and the morphological transformation of spermatotde during mouse testes development and spermatogenesis. Herein, the characteristics of the 2058 up-regulated expressing genes were analyzed and the authenticity of the microarray data was confirmed by doing Semi-quantitative RT-PCR. Their possible correlations with mouse spermatogenesis are also discussed.

MATERIALS AND METHODS

Animal Care Male and female Balb/c mice (aged 4—6 weeks) were obtained from Laboratory Animals Center of South Medical University (Guangzhou, China) and maintained in a temperature- and humidity-controlled room. All the animals had free access to standard mouse chow and water. Male and female mice were mated naturally, and the date of birth was designated as day 1. Testes were individually collected from Balb/C mice aged 4, 9, 11, 14, 18, 21, 35, 54, 60 d and 6 months. Other organs including brain, heart, lung, liver, kidney, spleen, epididymis, ovary and uterus from adult mice were also collected. All samples were immediately frozen in RNAlater liquid (QIAGEN, Valencia, CA, U.S.A.). Animal experiments were approved by the Animal Test Center of China.

RNA Extraction and Affymetrix Genechip Analysis Total RNA was extracted from mouse organs by using the Trizol reagent (Invitrogen, San Diego, CA, U.S.A.) according to the manufacturer’s recommendations. The concentrations and the integrities of total RNA were assessed by measuring the 260 : 280 nm ratios and by fractionation in 1% denaturing agarose gel (formaldehyde). The purity of total RNA was made by using RNeasy Mini Kit (QIAGEN, Valencia, CA, U.S.A.). Total RNA from testis samples of Balb/C mice aged 4, 9, 18, 35, 54 d and 6 months with a 260 : 280-nm ratio of 1.8 or higher was used to generate biotinylated cRNA target for the Mouse whole genome 430 v2 GeneChip (Affymetrix), which contained 45000 pairs of probe and 39000 transcripts including 34000 well characterized mouse genes. All of these procedures were carried out as described by Affymetrix. After hybridization, the array was washed, stained with streptavidin phycoerythrin using the Affymetrix GeneChip Fluidics Workstation 400, and scanned on a Hewlett-Packard gene array scanner (Hewlett-Packard Co., Palo Alto, CA, U.S.A.). After the arrays were scanned, the signals generated were determined and analyzed by MAS 5.0 software. The absolute and comparison analyses were also performed by using MAS 5.0. After normalization of these data, the comparison analysis compares an experimental array to a baseline array so as to monitor changes in the expression of transcripts across the samples targeted to different arrays (refer to www.Affymetrix.com for details on the statistics of these analyses).

Semi-quantitative RT-PCR Semi-quantitative RT-PCR was carried out to analyze and confirm the expression of candidate genes from Affymetrix chip analysis. RT-PCR was used to generate cDNA from testes sample of 4, 9, 11, 14, 18, 21, 38 d and 6 months old. Total RNA (1 μg) was reverse-transcribed into cDNA in a reaction primed by oligo deoxynucleotide T (dT)12—15 primer using Superscript II reverse transcriptase (Invitrogen) according to the manufacturer’s instructions. Forward and reverse oligonucleotide primers, specific to the chosen candidate genes, were designed using Primer Express 2.0 software (Applied Biosystems, Foster City, CA, U.S.A.) as described by the manufacturer. The primer sequences described above were shown in Table 1. All samples from various dates of testes were plated in triplicate PCR reactions.

Bioinformatics Analysis Gene ontology was predicted by using AmiGO (http://www.godatabase.org/cgi-bin/amigo/go.cgi) and DAVID 2.1 Beta (http://david.niaid.nih.gov/david/ease.htm). Gene clustering was analyzed by using Cluster 3.0 and Eisensoftware-Treeview. The KEGG (Kyoto

<table>
<thead>
<tr>
<th>Table 1. Primers for Semi-quantitative RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selected gene</td>
</tr>
<tr>
<td>BF152877</td>
</tr>
<tr>
<td>BC024760</td>
</tr>
<tr>
<td>BB667208</td>
</tr>
<tr>
<td>AK016041</td>
</tr>
<tr>
<td>AV258142</td>
</tr>
<tr>
<td>BG070552</td>
</tr>
<tr>
<td>AV270194</td>
</tr>
<tr>
<td>A644442</td>
</tr>
<tr>
<td>AV266439</td>
</tr>
<tr>
<td>AW045458</td>
</tr>
<tr>
<td>Beta-actin</td>
</tr>
</tbody>
</table>
Encyclopedia of Genes and Genomes) pathway and Domain region was analyzed by DAVID 2.1 Beta for selected genes. The signal peptide gene predicted by SignalP 3.0 Server, while the analysis of transmembrane domain by TMHMM Server v. 2.0.

RESULTS

Expression Profiling of Mouse Testis of Postnatal Days

The Affymetrix GeneChip Mouse Genome 430 2.0 Array used in the present study consists of approximately ca. 34000 known mouse genes and 8000 unknown genes or ESTs. Comparative analysis of the gene expression profiles was performed between two neighboring stages of six dates in the developmental stages of mice testis including day 4, 9, 18, 35, 54, and 6 months (Fig. 1). Of the six comparative pair examined, pair 18/9, pair 35/18, and pair 6 months/4 d have widely scatters, indicating the existence of many differently expressed genes. Other three comparative pair including day 9/4, day 54/35, and 6 months/day 54 have centralized scatters. After comparing two neighboring spots of six dates described above with 4 times of spot represented change, 2058 genes were selected and their expression levels gradually increased according to the developmental stages of mice testis (Fig. 2).

Validation of Microarray Data

To identify the worthiness of 2058 genes in spermatogenesis of mice, two analyses were carried out to confirm the ranking of gene expression levels determined by microarray analysis. Firstly, the scaling signal intensities of 21 spermatogenesis-related, testis-specific genes searched from published journal in the tested Affymetrix chip were analyzed and the results showed that the gene expression levels of all these 21 genes were gradually increased in the developmental stages of spermatogenesis in mice (Table 2). Except for GRTH, 2P1, Trif, and Spata6, other 17 genes have no expression in the mice testis of day 4, 9 and/or 18, and then the gradually increased expression of these genes begins. Secondly, in order to verify the Affymetrix chip data of 2058 genes, the sub-quantitative RT-PCR was performed in the 10 candidates selected randomly from 2058 genes. As shown in Fig. 3, these 10 genes have no or weak expression in the mice testis of day 4, 9, 11, and 14, and their gradually increased expression begins in the mice testis of day 18 or at the later stages of testis (Fig. 3).
The total RNA was extracted from testis samples of Balb/C mice aged 4, 9, 11, 14, 18, 21, 38 and 60 d. RT-PCR was then performed to detect the expression of 10 candidate genes randomly selected in the Affymetrix chip data. As shown in the figure, the numbers below each lane indicates the postnatal day. β-Actin was used as RT-PCR control. In 10 candidates, BC024760, BG070552, AV270194 have weak expression in ages of days 4, 9, 11, 14 and their high expression begins at 18 d in mouse testes; BF152877, BB667208 begin their expression at 14d; KO045458, AF644422, AV258142 begin their expression at 18 d and AK016041, AV266439 begin their expression at 38 d. The expression profiles of the 10 selected genes determined by RT-PCR has a consistency with the data of these candidates in Affymetrix chip analysis data.

Gene ontology was predicted by using AmiGO and DAVID 2.1 Beta. 2058 genes were classified by their biological process, molecular function and cellular component respectively (Fig. 5). 2058 genes in biological process, 434 in physiological, 241 in cellular process, 82 in development, 52 in unknown-process, 6 in regulation, 5 in behavior, 1151 in unclassified. 2058 genes in molecular function, 344 in binding, 250 in catalytic activity, 68 in signal transducer activity, 62 in transporter activity, 53 in transcription regulator activity, 48 in molecular-function unknown, 25 in structural molecule activity, 22 in chaperone activity, 22 in enzyme regulator activity, 15 in motor activity, 3 in antioxidant activity, 3 in translation regulator activity, 1439 in unclassified. 2058 genes in cellular component, 411 in cell, 124 in extracellular, 43 in cellular-component unknown, 6 in unlocalized, 1544 in unclassified. The KEGG pathway and Domain region of 2058 genes were also analyzed by DAVID 2.1 Beta (data not shown).

**DISCUSSION**

Microarray analysis was used to examine gene expression in the mouse testis with the goal of identifying genes that have gradually increased expression in the developmental stages of mice testis during its spermatogenesis. Primitive A spermatagonia is the enriched cell population in ages of day 4 mouse testes; Primitive B spermatagonia, Pachytene spermatocytes, Round spermatids, Elongating spermatids are...
the main enriched cell populations in day 9, 18, 35, and 54 mouse testes, respectively; elongated spermatides and immature sperm are specifically presented in day 70 and adult mouse testes. At the ages of 11—14 d, first meiosis begins in mouse testes, and the second meiosis is completed by 21 d. Comparative analysis of the neighboring two dates between days 4, 9, 18, 35, and 54, and 6 months indicates that pair 18/9 and pair 35/18 represent large amounts of differently expressed genes. From postnatal days 9 to 35 in mice testis, unique events including meiosis, gene expression in haploid cells, chromatin remodeling and condensation, and acrosome and flagellum formation happened during its spermatogenesis. These differently expressed genes detected in the comparative analysis of day 18/9 and day 35/18 pairs may be involved in these unique events. To identify and characterize those differently expressed genes described above, 2058 genes were selected by comparing two neighboring spots of six-postnatal-day mice testis with 4-fold difference. These 2058 genes have no or weak expression in the early postnatal days of mouse testis, and their gradually increased expression begins on postnatal days 9, 18, or 35 when unique events happened during its spermatogenesis. These 2058 genes are therefore called spermatogenesis-related genes in mice.

To determine the relevance and reliability of these data, we selected 21 known spermatogenesis-related, testis-specific genes as candidates and determine their expression patterns in our Affymetrix chip data. Interestingly, all of these 21 genes including CERM, tesmin, MSJ-1, Soggy, SP-10, mTSARG3, SRG4, TESF-1 and HILs1 were among the 2058 genes, and accordingly their expression gradually increased by the developmental stages of mouse testis. To confirm the validity of the Affymetrix chip data of 2058 genes, we analyzed 10 selected genes using subquantitative RT-PCR assay and the results showed that all 10 candidates have a gradually increased expression pattern on days 4, 9, 11, 14, 18, 21, 38 and 6 months in Balb/C mouse testes, implying their biological roles during spermatogenesis. For example, the BF152877 and BB667208 genes begin to be expressed on day 14 and might be involved in meiosis of spermatocytes, while AK016041 and AV266439 might be in-

---

**Fig. 4. Expression Clustering of the 2058 Spermatogenesis-Related Genes**

**Fig. 5. Gene Ontology Profile of the 2058 Spermatogenesis-Related Genes**

The gene ontology terms of these genes were grouped into a number of categories that belong to the three main classes: a) biological process, b) molecular function, and c) cellular component.
volved in the morphological transformation of round spermatids to elongating spermatids. Bioinformatics analysis of AK016041 gene shows that the AK016041 protein has a molecular weight of 22.685 kDa, a signal peptide cleavage site at position 18—19 in amino acids predicted by SignalP 3.0 Server, two transmembrane domains at position 1—20, and 40—60 as predicted by TMHMM Server v. 2.0. Iguchi et al. (2006) analyzed gene expression in testis from 17-d-old, 40—60 as predicted by TMHMM Server v. 2.0. Iguchi et al. (2006)13) analyzed gene expression in testis from 17-d-old, 22-d-old, and adult mice using the Affymetrix MOE430A chip containing a total of 22690 probe sets to assess total gene expression, and found 657 upregulated and 86 down-regulated expressing probe sets. In 15 polypeptide upregulated genes confirmed by Iguchi et al.’s study, Spata6 and Pgk-2 are the two candidates in the 21 known genes listed in Table 2. Spata6 is specifically expressed in haploid germ cells.14) Pgk-2 is one of the few genes transcribed during meiosis and known to be translationally regulated after meiosis.13) Beissbarth et al. (2003)13) analyzed gene expression in testis samples of mouse aged 1, 4, 8, 11, 14, 18, 21, 26, and 29 d to 60 d using the Affymetrix Mouse U74v2 chip, which contains ca. 12500 known mouse genes or ESTs, and thus spanning approximately 1/3rd of the mouse genome. After comparing two neighboring dates of ten postnatal days of mice testis with a 2-fold difference, they estimated that >2300 genes (ca. 4% of the mouse genome) are dedicated to male germ cell-specific transcripts, >99% of which are first expressed during or after meiosis. We analyzed the gene expression in mouse testis aged 4, 9, 18, 35, 54 d and 6 months using the Affymetrix GeneChip Mouse Genome 430 2.0 Array chip, which contained ca. 34000 known mouse genes and 8000 unknown genes or ESTs, and thus spanning the whole mouse genome. From our results in the present study, 2058 genes with gradually increasing expression were selected by comparing two neighboring spots of six postnatal days of mice testis with a 4-fold difference, this amount is slightly less than 2300. Considering the different chip density and the fold discrepancy of neighboring spots used in Beissbarth’s study and ours, the two sets of the data are consistent in predicting the number of total up-regulated expression in mouse spermatogenesis.

Characterizing these 2058 genes and mapping of their signal transduction pathways will shed light on mechanism of spermatogenesis of mammals. Based on bioinformatics analysis, the expression of 2058 genes were divided into eight clusters. Genes in one cluster have similar expression profiles in a synergistic expression manner, indicating that genes in each cluster may depend on one cellular signal transduction pathway. Based on gene ontology analysis, the 2058 genes were divided into biological process, molecular function, and cellular component. However, the majority of genes were unclassified and further analysis is needed to describe the characteristics of these genes. The KEGG pathway and domain region were also analyzed for these 2058 genes (data not shown). Given the vast amount of information acquired by microarray analysis, it is difficult to discuss in great details the roles that these 2058 genes may play. However, by examining groups of genes with associated biological functions, new insight is gained regarding signaling pathways and other regulatory mechanisms that may mediate those unique events during mouse spermatogenesis. From these data, some key expression genes can be selected for future study in our lab. For example, two novel testis-specific genes, TSC2113) and TSC24,16) have been detected that have gradually increasing expression in the developmental stages of spermatogenesis in mice. New genes have been identified that may be important for spermatogenesis, and the data presented here open the door to the elucidation of the molecular events underlying mammalian male reproduction.

Acknowledgments This work was supported by National Natural Science Foundation of China (No. 30700824) and Natural Science Foundation of Guangdong Province (No. 7008952).

REFERENCES