Comparative expression profiling of testis-enriched genes regulated during the development of spermatogonial cells

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1371/journal.pone.0175787

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:33029848

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Abstract

The testis has been identified as the organ in which a large number of tissue-enriched genes are present. However, a large portion of transcripts related to each stage or cell type in the testis still remains unknown. In this study, databases combined with confirmatory measurements were used to investigate testis-enriched genes, localization in the testis, developmental regulation, gene expression profiles of testicular disease, and signaling pathways. Our comparative analysis of GEO DataSets showed that 24 genes are predominantly expressed in testis. Cellular locations of 15 testis-enriched proteins in human testis have been identified and most of them were located in spermatocytes and round spermatids. Real-time PCR revealed that expressions of these 15 genes are significantly increased during testis development. Also, an analysis of GEO DataSets indicated that expressions of these 15 genes were significantly decreased in teratozoospermic patients and polyubiquitin knockout mice, suggesting their involvement in normal testis development. Pathway analysis revealed that most of those 15 genes are implicated in various sperm-related cell processes and disease conditions. This approach provides effective strategies for discovering novel testis-enriched genes and their expression patterns, paving the way for future characterization of their functions regarding infertility and providing new biomarkers for specific stages of spermatogenesis.
Introduction

The testis has been identified by RNA sequencing as the organ in which the largest number of tissue-enriched genes is expressed among various organs. It has been estimated that expressions of more than 1000 genes are enriched in the testis [1]; whereas, on average, there are approximately 200 signature genes in each tissue [2]. Tissue-enriched or tissue-specific genes are essential for the growth and development of specific cells and organs [3]. Thus, characteristic processes that occurred in germlinal cells in the testis, including meiosis, genetic recombination, spermatogenesis, and spermiogenesis may largely be attributed to a number of differential gene expressions. Spermatogenesis is a complex process that is orchestrated by expression of multiple genes at various stages containing particular cell types, such as spermatogonial stem cells, spermatogonia, spermatocytes, and spermatids [4]. In addition to germlinal cells, the somatic Sertoli cells play a role in testis formation and provide an essential environment for spermatogenesis [5], and Leydig cells produce androgen, which plays a key role in the regulation of spermatogenesis and undergo changes in gene expression [6, 7]. However, a large portion of transcripts and proteins related to each stage or cell type as well as their functions still remains unknown.

Investigation of gene expression and function during spermatogenesis has been hampered by a lack of immortalized cell lines for each stage [8]. Alternatively, testis transcriptome microarray analysis based on Gene Expression Omnibus (GEO) repository (www.ncbi.nlm.nih.gov/geo) followed by protein profiling using immunohistochemical data from the Human Protein Atlas portal (www.proteinatlas.org) is a useful tool for discovering highly expressed genes in each stage of spermatogenesis in the testis. Furthermore, gene expression profiles under various developmental, disease, and knockout conditions produced in GEO microarray datasets offer a platform for functional genomic studies of spermatogenesis stage-specific gene expression.

Using these sources combined with confirmatory gene expression measurements and pathway analysis, in this study, protein localization and signaling pathways of 15 testis-enriched genes were analyzed. The objectives of this study were to identify novel testis-enriched genes using gene expression profiles and analyze protein localization, developmental regulation and biological implications of testis-enriched genes in humans and mice. The current approach provides an effective strategy for discovering novel testis-enriched genes and their unique stage-specific expression, paving the way for future studies of normal development of the testis and associated diseases.

Materials and methods

Microarray data mining

The microarray-based, high-throughput gene expression data were obtained from the GDS DataSet (GDS) of the GEO repository in the National Center for Biotechnology Information (NCBI) archives (www.ncbi.nlm.nih.gov/geo). For analyzing tissue distribution pattern of gene expression in 12 male mouse tissues and 10 man tissues, GDS3142 for mice and GDS596 for humans were downloaded and sorted (Tables 1 and 2) as described in our previous reports [9, 10]. Also, gene expression patterns in mouse sperm cells (GDS2390), developing mouse testis (GDS605, GDS606 and GDS607), semen samples collected from 14 teratozoospermic individuals aged 21–57 (GDS2697), and polyubiquitin knockout mice (GDS3906) were examined.

Animal use and sample preparation

All animal care and procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at The Ohio State University. Mice were raised under ad libitum feeding
conditions in a mice housing facility at The Ohio State University. Mice were euthanized by carbon dioxide inhalation followed by cervical dislocation. For isolation of total RNAs, testis,
Table 2. Human testis-enriched genes based on GDS596.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fold</th>
<th>Testis</th>
<th>Ovary</th>
<th>Muscle</th>
<th>Liver</th>
<th>Brain</th>
<th>Lung</th>
<th>Kidney</th>
<th>Adipocyte</th>
<th>Thymus</th>
<th>Heart</th>
<th>P value</th>
<th>Testis enrichment</th>
<th>Location in human testis</th>
<th>Location in mouse Testis (Table 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRM2</td>
<td>861</td>
<td>12387</td>
<td>6</td>
<td>55</td>
<td>13</td>
<td>12</td>
<td>8</td>
<td>12</td>
<td>7</td>
<td>4</td>
<td>13</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>[64]</td>
<td>[46]</td>
</tr>
<tr>
<td>PRM1</td>
<td>800</td>
<td>37144</td>
<td>42</td>
<td>135</td>
<td>24</td>
<td>11</td>
<td>63</td>
<td>47</td>
<td>15</td>
<td>11</td>
<td>68</td>
<td>&lt; .0001</td>
<td>[65]</td>
<td>[64]</td>
<td>[45]</td>
</tr>
<tr>
<td>TNP1</td>
<td>186</td>
<td>17960</td>
<td>27</td>
<td>284</td>
<td>66</td>
<td>22</td>
<td>47</td>
<td>135</td>
<td>94</td>
<td>21</td>
<td>174</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>[66]</td>
<td>–</td>
</tr>
<tr>
<td>ODF2</td>
<td>181</td>
<td>4179</td>
<td>6</td>
<td>67</td>
<td>27</td>
<td>13</td>
<td>12</td>
<td>18</td>
<td>16</td>
<td>11</td>
<td>37</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>[67]</td>
<td>[52]</td>
</tr>
<tr>
<td>SPATA6</td>
<td>107</td>
<td>1742</td>
<td>8</td>
<td>44</td>
<td>15</td>
<td>8</td>
<td>7</td>
<td>22</td>
<td>13</td>
<td>9</td>
<td>20</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>Fig 2B</td>
<td>[57, 58]</td>
</tr>
<tr>
<td>PHF7</td>
<td>77</td>
<td>7891</td>
<td>82</td>
<td>417</td>
<td>47</td>
<td>77</td>
<td>47</td>
<td>62</td>
<td>53</td>
<td>46</td>
<td>86</td>
<td>&lt; .0001</td>
<td>[17]</td>
<td>Fig 2A</td>
<td>[17]</td>
</tr>
<tr>
<td>CRISP2</td>
<td>68</td>
<td>4526</td>
<td>53</td>
<td>396</td>
<td>43</td>
<td>25</td>
<td>8</td>
<td>10</td>
<td>4</td>
<td>14</td>
<td>14</td>
<td>&lt; .0001</td>
<td>[68]</td>
<td>[69]</td>
<td>–</td>
</tr>
<tr>
<td>SPINK2</td>
<td>63</td>
<td>7706</td>
<td>52</td>
<td>257</td>
<td>134</td>
<td>86</td>
<td>106</td>
<td>67</td>
<td>63</td>
<td>125</td>
<td>215</td>
<td>&lt; .0001</td>
<td>[70]</td>
<td>Fig 2A</td>
<td>[20]</td>
</tr>
<tr>
<td>NEK2</td>
<td>47</td>
<td>1427</td>
<td>28</td>
<td>71</td>
<td>13</td>
<td>29</td>
<td>7</td>
<td>16</td>
<td>7</td>
<td>75</td>
<td>28</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>[71]</td>
<td>[14, 15]</td>
</tr>
<tr>
<td>LDHC</td>
<td>46</td>
<td>3625</td>
<td>40</td>
<td>254</td>
<td>70</td>
<td>22</td>
<td>38</td>
<td>19</td>
<td>134</td>
<td>25</td>
<td>102</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>Fig 2A</td>
<td>[14, 15]</td>
</tr>
<tr>
<td>SMCP</td>
<td>39</td>
<td>2082</td>
<td>25</td>
<td>172</td>
<td>53</td>
<td>29</td>
<td>15</td>
<td>64</td>
<td>31</td>
<td>18</td>
<td>76</td>
<td>&lt; .0001</td>
<td>[72]</td>
<td>[72]</td>
<td>[48–50]</td>
</tr>
<tr>
<td>YBX2</td>
<td>37</td>
<td>8120</td>
<td>67</td>
<td>719</td>
<td>224</td>
<td>258</td>
<td>70</td>
<td>337</td>
<td>108</td>
<td>46</td>
<td>160</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>Fig 2C</td>
<td>–</td>
</tr>
<tr>
<td>ZPBP</td>
<td>35</td>
<td>3012</td>
<td>87</td>
<td>128</td>
<td>50</td>
<td>64</td>
<td>73</td>
<td>74</td>
<td>24</td>
<td>31</td>
<td>236</td>
<td>&lt; .0001</td>
<td>[26]</td>
<td>Fig 2B</td>
<td>[56]</td>
</tr>
<tr>
<td>ACTL7A</td>
<td>35</td>
<td>1938</td>
<td>17</td>
<td>245</td>
<td>38</td>
<td>27</td>
<td>21</td>
<td>59</td>
<td>28</td>
<td>16</td>
<td>55</td>
<td>&lt; .0001</td>
<td>[73]</td>
<td>Fig 2B</td>
<td>[54]</td>
</tr>
<tr>
<td>TCP11</td>
<td>33</td>
<td>7549</td>
<td>37</td>
<td>632</td>
<td>227</td>
<td>160</td>
<td>122</td>
<td>293</td>
<td>196</td>
<td>78</td>
<td>340</td>
<td>&lt; .0001</td>
<td>[74]</td>
<td>Fig 2A</td>
<td>[53]</td>
</tr>
<tr>
<td>ZMYN10</td>
<td>32</td>
<td>1505</td>
<td>10</td>
<td>165</td>
<td>42</td>
<td>18</td>
<td>29</td>
<td>66</td>
<td>22</td>
<td>8</td>
<td>60</td>
<td>&lt; .0001</td>
<td>[75]</td>
<td>Fig 2C</td>
<td>–</td>
</tr>
<tr>
<td>ACTL7B</td>
<td>30</td>
<td>1876</td>
<td>41</td>
<td>184</td>
<td>79</td>
<td>73</td>
<td>17</td>
<td>38</td>
<td>62</td>
<td>37</td>
<td>27</td>
<td>&lt; .0001</td>
<td>[73]</td>
<td>Fig 2B</td>
<td>[54]</td>
</tr>
<tr>
<td>ODF1</td>
<td>25</td>
<td>5549</td>
<td>183</td>
<td>493</td>
<td>223</td>
<td>130</td>
<td>80</td>
<td>300</td>
<td>145</td>
<td>72</td>
<td>361</td>
<td>&lt; .0001</td>
<td>[76]</td>
<td>Fig 2D</td>
<td>–</td>
</tr>
<tr>
<td>AKAP3</td>
<td>24</td>
<td>1166</td>
<td>23</td>
<td>120</td>
<td>59</td>
<td>39</td>
<td>33</td>
<td>61</td>
<td>18</td>
<td>16</td>
<td>69</td>
<td>&lt; .0001</td>
<td>[77–79]</td>
<td>[77]</td>
<td>[47]</td>
</tr>
<tr>
<td>GAPDHS</td>
<td>21</td>
<td>3357</td>
<td>29</td>
<td>250</td>
<td>216</td>
<td>192</td>
<td>86</td>
<td>96</td>
<td>120</td>
<td>73</td>
<td>401</td>
<td>&lt; .0001</td>
<td>[80]</td>
<td>Fig 2D</td>
<td>–</td>
</tr>
<tr>
<td>DDX4</td>
<td>19</td>
<td>447</td>
<td>7</td>
<td>115</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>33</td>
<td>5</td>
<td>4</td>
<td>27</td>
<td>&lt; .0001</td>
<td>[81]</td>
<td>[81]</td>
<td>–</td>
</tr>
<tr>
<td>TCFL5</td>
<td>14</td>
<td>2611</td>
<td>11</td>
<td>259</td>
<td>79</td>
<td>562</td>
<td>109</td>
<td>251</td>
<td>216</td>
<td>182</td>
<td>57</td>
<td>&lt; .0001</td>
<td>[82]</td>
<td>Fig 2A</td>
<td>[51]</td>
</tr>
<tr>
<td>STAG3</td>
<td>13</td>
<td>1885</td>
<td>44</td>
<td>216</td>
<td>140</td>
<td>130</td>
<td>92</td>
<td>217</td>
<td>57</td>
<td>224</td>
<td>141</td>
<td>&lt; .0001</td>
<td>[62]</td>
<td>Fig 2C</td>
<td>–</td>
</tr>
<tr>
<td>EFHC1</td>
<td>13</td>
<td>407</td>
<td>18</td>
<td>98</td>
<td>53</td>
<td>16</td>
<td>10</td>
<td>32</td>
<td>13</td>
<td>29</td>
<td>14</td>
<td>&lt; .0001</td>
<td>Fig 1</td>
<td>Fig 2A</td>
<td>[23]</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0175787.t002

muscle, liver, brain, lung, kidney, adipose tissue, thymus, spleen, and small intestine were collected from 3-month-old FVB mice (n = 3) using Trizol reagent (Invitrogen, Carlsbad, CA,
USA) [11]. Total RNAs from the adult human kidney, liver, lung, heart, muscle, testis, thymus, and brain were purchased from Agilent Technologies (Santa Clara, CA, USA) and adult human RNA from adipose tissue was purchased from Clontech Laboratories (Mountain View, CA, USA). For RNA isolation from mouse testis at 10 days postpartum (dpp), 21 dpp, and 91 dpp (three months postpartum), C57BL/6 mice (n = 4) were euthanized and both testes were harvested.

Reverse transcription PCR (RT-PCR)

To measure the quantity of RNA, a Nanodrop spectrophotometer (Thermo Scientific, Wilmington, DE) was used. The RNA samples were stored at -80˚C until use. Approximately 1 μg of RNA was reverse-transcribed in a 20 μL total reaction to cDNA using Moloney murine leukemia virus (M-MLV) reverse transcriptase (Invitrogen). The thermal cycle of the reverse transcription was 65˚C for 5 min, 37˚C for 52 min, and 70˚C for 15 min. Exactly 1 μL of cDNA samples was used as a template for PCR in a 25 μL total reaction with AmpliTaq Gold DNA polymerase (Applied Biosystems, Carlsbad, CA). The conditions for this reaction were 95˚C for 1 min 30 s, 33 cycles of 94˚C for 30 s, 55˚C for 1 min, 72˚C for 1 min, with an additional extension step at 72˚C for 10 min. PCR products were separated by using 1% agarose gel electrophoresis. Forward and reverse primers for both humans and mice listed in supporting information were designed on different exons for multi-exon genes to avoid genomic DNA contamination.

Analysis of protein expression profiles from the Human Protein Atlas

Data visualizing immunohistochemically the expression patterns of selected proteins in human testes were obtained from the Human Protein Atlas portal (www.proteinatlas.org). A total of 15 testis-enriched proteins were analyzed for their localization in the human testis: ten of them have been published in terms of their localization in mouse testis, but not in human testis, and the rest of them have not been published regarding their localization in both human and mouse testis.

Real-time PCR

Quantitative real-time PCR (qPCR) was performed on an ABI 7300 Real-Time PCR instrument (Applied BioSystems, Foster City, CA) by using AmpliTaq Gold polymerase (Applied BioSystems) with SYBR green detection dye. Cyclophilin (CYC) was used as a housekeeping gene. Reactions were performed in duplicate 25μL volumes and conditions for the qPCR were 95˚C for 10 minutes followed by 40 cycles of 94˚C for 15 seconds, 60˚C for 40 seconds, 72˚C for 30 seconds, and 82˚C for 33 seconds. Relative quantification of gene expression was determined by using the 2^ΔΔC_T method [12].

Signaling pathway analysis

Signaling pathways of spermatocyte- or spermatid-enriched proteins were analyzed using Pathway Studio (v 11.2.5.9, Elsevier, Amsterdam, Netherland). A list of 15 testis-enriched proteins was entered into Pathway Studio. The resulting pathways were verified through the PubMed/Medline hyperlink embedded in each node.

Statistical analysis

For comparison of gene expression in testis versus other tissues, one-way ANOVA followed by a Fisher’s protected least significant difference test was performed using SAS version 9.2 (SAS
Institute Inc., Cary, NC). A Student’s t test was conducted to compare the difference between two means. Comparison of multiple means was conducted by one-way ANOVA followed by a Tukey’s post hoc test. The significance level was set at $p < 0.05$.

Results

Microarray analyses identified common testis-enriched genes for the mouse and human

Comparative analysis of GEO DataSets (GDS3142 for mice and GDS596 for humans), a public microarray repository, revealed that expressions of 24 genes in both the mouse and human testis are more than 10-fold higher than an average expression value of other tissues (Tables 1 and 2). For example, murine Tnp1 and human TNP1 expressions are 276- and 186-fold greater in the mouse and human testis, respectively, than an average value of other tissues. In addition, these 24 genes are expressed at very low levels in the ovary, showing that they are male-specific genes. Our literature search revealed that some, but not all, genes were reported for testis enrichment and protein cellular location in testis. For instance, testis-specific expression of murine Ldhc gene and cellular protein location of LDHC were reported in the mouse testis [13–15], but not in human testis. In this study, testis enrichment of selected genes was confirmed by RT-PCR (Fig 1) and their localization profiles in humans were explored through the Human Protein Atlas (Fig 2).

RT-PCR confirmed testis enrichment of selected genes

To validate the microarray data, RT-PCR was performed for murine Smcp, Odf1, Crisp2, Phf7, Gapdhs, Ddx4 and Zmynd10 and human PRM2, TNP1, SPATA6, NEK2, LDHC, YBX2 and

![RT-PCR of mouse and human testis-enriched genes](https://doi.org/10.1371/journal.pone.0175787.g001)
EFHC1, which have not been reported previously for expression in the testis. To prevent PCR saturation effects during amplification, the number of PCR cycles was reduced until the saturation no longer occurs. These genes showed testis-enriched expression patterns among various tissues (Fig 1), which is consistent with the GEO DataSets (Tables 1 and 2).

Analysis of immunohistochemical data showed protein expression in specific stages of human testis

With the very latest version of the Human Protein Atlas, cellular location of several testis-enriched genes in humans was analyzed. When these locations were reported in mouse, they were grouped in Fig 2A and 2B; otherwise they were grouped in Fig 2C and 2D. As shown in Fig 2, PHF7, SPINK2, LDHC, TCP11, EFHC1, and TCFL5 proteins were located in earlier stage cells (Fig 2A) than cells expressing ZPBP, ACTL7A, ACTL7B, and SPATA6 (Fig 2B). PHF7, SPINK2, LDHC, TCP11, EFHC1, and TCFL5 were localized in pachytene spermatocytes (PC) and round spermatids (RS). In detail, SPINK2 was expressed strongly in the cytoplasm of pachytene spermatocytes, LDHC showed expression in the tail of spermatozoa (SZ), and EFHC1 was highly expressed in Sertoli cells (ST) (Fig 2A). ZPBP was uniquely detected in the developing acrosomal granules (AG) of round spermatids. ACTL7A was expressed in round spermatids and exclusively in the acrosome granules, with a lesser degree in spermatozoa tails. ACTL7B showed a stronger expression than SPATA6 in round spermatids (Fig 2B).
Table 3. Differential gene expression in four types of spermatic cells according to GDS2390.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Type A spermatogonia</th>
<th>Type B spermatogonia</th>
<th>Pachytene spermatocytes</th>
<th>Round spermatids</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Testis-enriched genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phf7</td>
<td>62.3 ± 4.3^a</td>
<td>385.0 ± 327.7^ab</td>
<td>2877.8 ± 316.3^ab</td>
<td>3073.8 ± 14.9^a</td>
</tr>
<tr>
<td>Spink2</td>
<td>36.2 ± 4.7^c</td>
<td>414.7 ± 392.7^ab</td>
<td>2576.3 ± 181.8^a</td>
<td>2527.2 ± 45.7^a</td>
</tr>
<tr>
<td>Ldhc</td>
<td>34.5 ± 5.1^b</td>
<td>2462.1 ± 2433.8^ab</td>
<td>6269.1 ± 294.1^a</td>
<td>6783.0 ± 205.1^a</td>
</tr>
<tr>
<td>Tcp11</td>
<td>87.1 ± 4.6^c</td>
<td>435.2 ± 356.3^bc</td>
<td>2659.4 ± 144.6^p</td>
<td>4347.4 ± 19.3^k</td>
</tr>
<tr>
<td>Efhc1</td>
<td>66.1 ± 5.1^p</td>
<td>288.3 ± 242.1^ab</td>
<td>2012.3 ± 19.5^ab</td>
<td>2476.6 ± 69.3^a</td>
</tr>
<tr>
<td>Tcfl5</td>
<td>635.4 ± 68.0^c</td>
<td>943.9 ± 247.0^a</td>
<td>2967.8 ± 81.8^a</td>
<td>2408.3 ± 77.9^d</td>
</tr>
<tr>
<td>Zpbp</td>
<td>24.1 ± 1.4^bc</td>
<td>204.4 ± 164.8^c</td>
<td>1629.2 ± 224.3^ab</td>
<td>3194.8 ± 49.0^e</td>
</tr>
<tr>
<td>Actl7a</td>
<td>15.2 ± 4.0^d</td>
<td>48.8 ± 36.0^c</td>
<td>131.7 ± 22.6^b</td>
<td>2097.9 ± 28.8^c</td>
</tr>
<tr>
<td>Actl7b</td>
<td>58.8 ± 11.2^c</td>
<td>253.1 ± 180.5^bc</td>
<td>1120.6 ± 67.4^d</td>
<td>3373.9 ± 223.4^a</td>
</tr>
<tr>
<td>Spata6</td>
<td>47.2 ± 1.8^e</td>
<td>133.3 ± 68.9^c</td>
<td>784.1 ± 14.7^b</td>
<td>2767.5 ± 143.3^a</td>
</tr>
<tr>
<td>Ybx2</td>
<td>131.7 ± 39.9^b</td>
<td>651.6 ± 508.2^ab</td>
<td>3982.7 ± 107.3^a</td>
<td>4034.0 ± 130.1^a</td>
</tr>
<tr>
<td>Zmynd10</td>
<td>2.4 ± 0.6^c</td>
<td>323.0 ± 321.6^bc</td>
<td>2457.2 ± 121.3^a</td>
<td>997.2 ± 44.6^g</td>
</tr>
<tr>
<td>Stag3</td>
<td>725.9 ± 34.6^b</td>
<td>984.0 ± 124.2^d</td>
<td>2231.6 ± 124.7^a</td>
<td>1762.4 ± 169.8^ab</td>
</tr>
<tr>
<td>Odf1</td>
<td>3.7 ± 0.4^p</td>
<td>62.0 ± 58.0^c</td>
<td>275.9 ± 69.5^c</td>
<td>3558.2 ± 185.0^a</td>
</tr>
<tr>
<td>Gapdhs</td>
<td>71.3 ± 8.8^d</td>
<td>122.1 ± 71.8^bc</td>
<td>270.3 ± 0.3^b</td>
<td>3604.4 ± 141.2^d</td>
</tr>
</tbody>
</table>

**Non-testis-enriched genes**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Type A spermatogonia</th>
<th>Type B spermatogonia</th>
<th>Pachytene spermatocytes</th>
<th>Round spermatids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col1a2</td>
<td>2743.6 ± 314.6^a</td>
<td>3338.4 ± 173.2^a</td>
<td>1110.0 ± 18.6^b</td>
<td>146.4 ± 1.0^p</td>
</tr>
<tr>
<td>Zbtb16</td>
<td>304.7 ± 42.8^a</td>
<td>178.9 ± 56.9^ab</td>
<td>21.0 ± 1.2^p</td>
<td>33.8 ± 13.0^p</td>
</tr>
</tbody>
</table>

Means ± SEM are shown. Different superscript letters indicate significant differences between types of cells.

https://doi.org/10.1371/journal.pone.0175787.t003

Other testis-enriched proteins shown in Fig 2C and 2D, such as YBX2, ZMYND10, STAG3, ODF1, and GAPDHS, have not been published regarding their localization in both human and mouse testis. YBX2 and ZMYND10 were strongly localized in the cytoplasm of pachytene spermatocytes (PC) and, to a lesser degree, the nucleus of pachytene spermatocytes in the case of ZMYND10. STAG3 was expressed in the nucleus of pachytene spermatocytes. These proteins were also present in round spermatids (RS) except for low expression of ZMYND10 in round spermatids. In addition, YBX2 was detected in the tail of spermatozoa (SZ), and ZMYND10 showed expression in Leydig cells (LD) (Fig 2C). ODF1 and GAPDHS were expressed in round spermatids (RS). Also, GAPDHS showed expression in the tail of spermatozoa (SZ). ODF1 was expressed in elongating spermatids (ES) with developing tails (Fig 2D).

In summary, most testis-enriched proteins selected in this study are expressed after the spermatogonial stage. Their gene expression profiles curated in GDS2390 also showed a significant increase of expression during the stages of pachytene spermatocytes and round spermatids (Table 3). On the other hand, two non-testis-enriched proteins, COL1A2 and ZBTB16, were mainly expressed in the early-stage cells such as type A and type B spermatogonia (Fig 2E) and similarly, their gene expression profiles showed significantly higher mRNA expression in spermatogonia stages (Table 3). Therefore, whether expression of testis-enriched genes is regulated during testis development was further analyzed.

Expression of testis-enriched genes showed an increasing pattern during normal testis development

qPCR revealed that expression of selected testis-enriched genes is significantly increased during testis development. To investigate stage-specific expression patterns, 10 days postpartum (dpp) with mostly spermatogonia, 21 dpp when spermatocytes are the most abundant cell
type, and 91 dpp representing an adult stage with spermatids were selected. Compared to 10 days postpartum (dpp), expression of these selected genes was significantly increased at 21 dpp after weaning and/or at 91 dpp (three months postpartum) when sexual maturation occurs (Fig 3A–3D). These data are consistent with the microarray database in GDS605, GDS606, and GDS607, which also shows an increasing pattern of these genes during the period of 0 through 35 dpp (Table 4). It suggests that expression of these testis-enriched genes is up-regulated during testis development and plays a role in later stages of spermatogenesis. In contrast, expression of both Col1a2 and Zbtb16 was significantly decreased at 21 dpp (Fig 3E), and this pattern was also shown in GDS605, GDS606, and GDS607 (Table 4), suggesting that these non-testis-enriched genes may be involved in early stages of spermatogenesis rather than the later stages.

In addition, our further data analysis showed that, compared to fertile normal males, expression of these testis-enriched genes was significantly decreased in teratozoospermic patients with abnormal sperm morphology according to GDS2697 (S1 Table). It appears that those testis-enriched genes are involved in normal testis development without morphological defects and may serve as a biomarker for teratozoospermic condition. Moreover, according to GDS3906, polyubiquitin knockout resulted in a decreased expression pattern of testis-enriched genes at 28 dpp compared to wild-type (S1 Table).
Table 4. Differential gene expression during testis development based on GDS605, GDS606 and GDS607.

<table>
<thead>
<tr>
<th>Gene</th>
<th>0–3 dpp</th>
<th>8–10 dpp</th>
<th>18–20 dpp</th>
<th>30–35 dpp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testis-enriched genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phf7</td>
<td>151.7 ± 22.5&lt;sup&gt;b&lt;/sup&gt;</td>
<td>153.2 ± 7.2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5045.2 ± 421.8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6208.0 ± 450.3&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Spink2</td>
<td>122.5 ± 8.7&lt;sup&gt;c&lt;/sup&gt;</td>
<td>99.1 ± 6.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1708.5 ± 221.6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2506.2 ± 596.5&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ldhc</td>
<td>48.4 ± 45.4&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5.4 ± 1.5&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5935.9 ± 669.2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>8665.4 ± 572.5&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tcp11</td>
<td>23.9 ± 1.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>26.9 ± 1.3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1256.9 ± 243.3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4906.3 ± 307.0&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Etfc1</td>
<td>46.7 ± 10.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>40.3 ± 4.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>829.2 ± 175.9&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1299.2 ± 58.4&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tcf15</td>
<td>68.8 ± 6.0&lt;sup&gt;c&lt;/sup&gt;</td>
<td>139.0 ± 14.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>845.1 ± 201.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1272.3 ± 50.8&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Zpbb</td>
<td>8.8 ± 2.3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>7.5 ± 3.1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>265.0 ± 91.9&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1004.4 ± 33.0&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Actl7a</td>
<td>29.0 ± 6.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>32.3 ± 6.5&lt;sup&gt;b&lt;/sup&gt;</td>
<td>40.6 ± 18.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1705.5 ± 183.3&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Actl7b</td>
<td>18.0 ± 4.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>12.0 ± 0.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>427.2 ± 152.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1892.4 ± 96.4&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Spata6</td>
<td>96.7 ± 10.8&lt;sup&gt;b&lt;/sup&gt;</td>
<td>98.0 ± 8.4&lt;sup&gt;b&lt;/sup&gt;</td>
<td>525.0 ± 103.2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>687.2 ± 28.6&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ybx2</td>
<td>10.7 ± 1.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>19.5 ± 4.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1957.4 ± 601.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4059.9 ± 158.9&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Zmynd10</td>
<td>6.1 ± 1.5&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5.3 ± 0.9&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1574.3 ± 553.7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>989.3 ± 70.0&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Stag3</td>
<td>6.6 ± 0.9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>177.8 ± 30.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>611.0 ± 93.2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>560.8 ± 11.0&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Odf1</td>
<td>62.7 ± 9.4&lt;sup&gt;c&lt;/sup&gt;</td>
<td>51.8 ± 6.3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>72.0 ± 9.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5039.2 ± 517.1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gapdhs</td>
<td>37.4 ± 4.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>33.9 ± 5.2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>38.0 ± 3.8&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4990.9 ± 434.5&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Non-testis-enriched genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Col1a2</td>
<td>3456.8 ± 240.3&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2132.2 ± 104.3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>604.7 ± 72.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>280.0 ± 29.2&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Zbtb16</td>
<td>98.2 ± 19.2&lt;sup&gt;c&lt;/sup&gt;</td>
<td>254.2 ± 18.0&lt;sup&gt;c&lt;/sup&gt;</td>
<td>61.7 ± 16.1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>32.8 ± 2.9&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Means ± SEM are shown. Different superscript letters indicate significant differences between developmental time points.

https://doi.org/10.1371/journal.pone.0175787.t004

Testis-enriched genes are associated with various biological pathways in sperm

Signaling pathway analysis was conducted to identify corresponding pathways related to sperm-related biological functions and disease conditions. Schematic illustration was drawn to identify cellular and metabolic processes regulated by testis-enriched genes and showed that at least 12 out of 15 spermatocyte- or spermatozoa-enriched proteins were putatively associated with various sperm-related cell processes, clinical parameters, and disease conditions (Fig 4).

Discussion

In this study, testis-enriched genes in human and mouse were arranged based on microarray based-GEO database, and 15 genes that have not been published regarding their localization in human (Table 2) were selected to analyze their protein expression in testis using immunohistochemical data from the Human Protein Atlas portal.

Several proteins expressed in pachytene spermatocytes and round spermatids of human testis were analyzed

Proteins localized in pachytene spermatocytes (PC) and round spermatids (RS) were shown in Fig 2A and 2C. PHF7 is a male-specific transcription factor for germ cell development and sexual identity [16, 17]. SPINK2 is a Kazal-type serine protease inhibitor or an acrosin-trypsin inhibitor that is synthesized in the testis [18–20]. LDHC is an enzyme related to aerobic glycolysis in spermatozoa for energy production, and it regulates the sperm motility and capacitation [21]. Based on its localization, we postulated that LDHC is associated, not only with ATP generation in mature spermatozoa, but also with development of germ cells. TCP11 is a receptor...
of a fertilization promoting peptide that regulates sperm capacitation in the mouse [22]. EFHC1 has been found in mouse sperm flagella and is present in motile cilia, but not in immotile cilia [23]. In this study, EFHC1 was expressed in cytoplasmic regions of testicular cells. It suggests that EFHC1 may be associated with germ cell development and sperm motility. TCFL5 has been found in the testis during spermiogenesis, and it is associated with spermato genesis and the formation of sperm flagellum in the mouse [24]. YBX2, also known as contrin, is a germ cell specific protein and required for the formation of functional spermatozoa and has been implicated as a potential cause of azoospermia [25, 26]. ZMYND10 has been found in motile cilia of Drosophila, and it is associated with male fertility [27]. STAG3 is the meiosis-specific cohesion subunit and is associated with meiotic division of gametes [28].

Some proteins expressed in the acrosome or cytoplasmic region of spermatids of human testis were presented

The acrosome reaction is required for zona pellucida penetration and fertilization with oocytes [29], and four proteins, ZPBP [30], ACTL7A (T-ACTIN2), ACTL7B (T-ACTIN1), and SPATA6, were localized in the acrosome or cytoplasmic region of spermatids (Fig 2B), implicating their roles in fertility. Other proteins that may also be involved in spermatogenesis during the spermatid phase are shown in Fig 2D. ODF1 is one of the heat shock proteins that play an important role as molecular chaperones in spermatozoa, and it is located in the sperm tails and supports the flagella motility [31, 32]. GAPDHS is a testis-specific glycolytic enzyme and generally known to be present in the principal piece of spermatozoa, and it is associated with ATP production and flagella motility and capacitation [33, 34].

Expression of testis-enriched genes increased during normal testis development

Those selected testis-enriched proteins were mostly expressed in cells in the late spermatogenesis stages. The stage-specific mRNA expression of these genes showed similar patterns as
shown in the GDS2390 dataset (Table 3). Their expression was further analyzed during the testis development. During postnatal testicular growth, the proportion of germ cell types in seminiferous tubules changes. Before 10 dpp, testes of the mouse (Mus musculus) contain mostly spermatogonia. Between 21 and 24 dpp (weaning ages), spermatocytes become the most abundant cell type, round spermatids develops as the most advanced germ cells, and in adults, spermatids are a predominant cell type stage [35, 36]. Based on the changes in types of spermatogenic cells, qPCR was performed at the stage of spermatogonia (10 dpp), spermatocytes and round spermatids (21 dpp), and spermatids (91 dpp). Our qPCR data showed that expression of these testis-enriched genes was significantly increased around weaning ages when spermatocytes and round spermatids are present in testis (Fig 3). Some of them were further increased at 91 dpp when the mouse is sexually matured and spermatids are the predominant form of spermatogenic cells. These expression patterns were consistent with gene expression profiles in GDS605, GDS606, and GDS607 (Table 4). These results suggest that testis-enriched genes may be involved in advanced stages of spermatogenesis when spermatocytes and spermatids are dominant types of spermatogenic cells.

Testis-enriched genes tend to be repressed in diseases associated with male infertility

Expression of all of these testis-enriched genes was decreased in teratozoospermic patients compared to normal individuals (S1 Table). Spermatogenic cells are susceptible to impairment which causes spermatogenic cells to become arrested at a certain developmental stage. For example, spermatogenic arrest at spermatogonia leads to total germ cell depletion and Sertoli cell only (SCO) syndrome with a lack of germ cells, arrest at spermatocytes gives rise to azoosperma (no spermatozoa) and oligozoosperma (a reduced number of spermatozoa), and arrest at spermatids results in teratozoosperma (an abnormal shape of spermatozoa) [37]. Thus, decreased expression of genes enriched in spermatids in this study could be used as biomarkers for spermatid arrest, teratozoosperma, and subsequent infertility. In addition, polyubiquitin knockout mice showed a decreased expression of these testis-enriched genes compared to wild-types (S1 Table). The ubiquitin-proteasomal pathway (UPP) has been regarded to be a critical process for the successful maturation of spermatids into spermatozoa by tagging and degrading proteins related to morphological defects [38]. In addition, post-testicular presence of ubiquitin plays a role in disposal of defective mature spermatozoa [39, 40]. It has been reported that total knockout of the polyubiquitin gene in mice resulted in a developmental arrest of spermatogenesis followed by infertility [41]. Therefore, the decreased expression of testis-enriched genes in polyubiquitin knockout models can be used as an indicator of failure in sperm maturation. A recent study has shown that knockout mice lacking several testis-enriched genes were fertile [42]; however, the relationship between these genes and normal testis development remains to be explored. Genes presented in the current study that are related to testis development may provide appropriate targets for future knockout studies.

Various signaling pathways in sperm are linked to testis-enriched genes

Pathway analysis, in this study, provided comprehensive insight into the underlying biological functions and diseases involved in spermatocyte- or spermatozoa-enriched expression. As such, most of spermatocyte- or spermatozoa-enriched proteins being analyzed were implicated in a variety of sperm functions, including motility and capacitation, and multiple disease conditions such as infertility. On the other hand, three proteins (ZMYND10, ACTL7B, and TCFL5) out of those 15 spermatocyte- or spermatozoa-enriched proteins were not implicated in the biological conditions possibly due to incomplete functional annotations (Fig 4).
In conclusion, testis-enriched genes were found based on GEO profiles, and among them, protein localization of 15 genes was identified using the Human Protein Atlas. Mostly, these testis-enriched proteins were expressed in spermatocytes and/or round spermatids, and their expression significantly increased during testis development. In testicular disease conditions, expressions of these genes were significantly decreased suggesting their relation to normal spermatogenesis and testis development. Moreover, in our pathway analysis, most of these proteins exhibited multiple biological implications related to sperm function. Future studies should ascertain the potential involvement of these testis-enriched genes in male infertility.

Supporting information

S1 Table. Microarray analysis of testicular transcriptome. Samples were derived from normal and teratozoospermic individuals aged 21–57 (GDS2697) and from wild-type and polyubiquitin knockout mice at 28 dpp (GDS3906).

S2 Table. Primer sequences for RT-PCR and real-time PCR.

Acknowledgments

We are grateful to the staff members of University Laboratory Animal Resources (ULAR) at The Ohio State University for their care and effort for the proper use of animals.

Author Contributions

Conceptualization: JA KL.
Formal analysis: W-SK M-GP JA KL.
Funding acquisition: SH SSL.
Investigation: JA YS PC.
Methodology: C-HK.
Resources: TJL Y-JJ CC SH SSL.
Supervision: KL.
Visualization: JA W-SK M-GP.
Writing – original draft: JA Y-JP.
Writing – review & editing: M-GP SSL KL.

References


