A single-nucleotide polymorphism of the DAZL gene promoter confers susceptibility to spermatogenic failure in the Taiwanese Han

Yeng-Ni Teng1,†, Yi-Ping Chang2,†, Joseph T. Tseng3, Po-Hsiu Kuo4, I-Wen Lee5, Maw-Sheng Lee6, and Pao-Lin Kuo5,*

1Department of Biological Sciences and Technology, National University of Tainan, Tainan, Taiwan 2Institute of Molecular Medicine, College of Medicine, National Cheng-Kung University 3Institute of Bioinformatics, College of Bioscience and Biotechnology, National Cheng-Kung University 4Department of Public Health & Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan 5Department of Obstetrics and Gynecology, College of Medicine, National Cheng-Kung University, Tainan, Taiwan 6Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan; Division of Infertility Clinic, Lee Women’s Hospital, Taichung, Taiwan

*Correspondence address. Division of Genetics, Department of Obstetrics and Gynecology, National Cheng Kung University Hospital, 138 Sheng-Li Road, Tainan 704, Taiwan. Fax: +886-6-276-6185; E-mail: paolink@mail.ncku.edu.tw

Submitted on January 8, 2012; resubmitted on May 11, 2012; accepted on May 22, 2012

BACKGROUND: Deleted in AZoospermia-like (DAZL) is an autosomal homologue of Y chromosome-linked DAZ gene located on chromosome 3p24. DAZL is only expressed in the gonads and is critical to germ cell development in different species. However, the regulation of DAZL has not been explored.

METHODS: Reporter assays, electrophoretic mobility shift assays, supershift assays and bisulfate sequencing were used to identify the core promoter region of DAZL. Sequence analysis was used to identify single-nucleotide polymorphisms (SNPs) in the promoter region. A total of 337 infertile men with abnormal semen parameters and 203 fertile men with normal semen parameters were subjected to sequence analysis of the DAZL promoter region.

RESULTS: The DAZL gene core promoter is located 1 kb upstream of the transcription start site. Three SNPs (−792G>A, −669A>C and −309T>C) were identified in our population. Of these three SNPs, −792G>A was more prevalent in the infertile men (P = 0.0005). Quantitative analysis revealed that genotypes of −792G>A had effects on sperm concentration (P = 0.0025) and motility (P = 1.5 × 10−7). The G to A substitution was associated with decreased binding of the nuclear respiratory factor-1 (NRF-1) to the promoter region and decreased reporter gene activity.

CONCLUSION: We have identified the core promoter of the human DAZL gene. We also provide preliminary evidence for the role of a novel SNP of the DAZL gene promoter in human spermatogenic failure.

Key words: single-nucleotide polymorphism / DAZL / gene promoter / spermatogenic failure

Introduction

The World Health Organization (WHO) has defined infertility as inability of the couple to conceive after 2 years of unprotected sexual activity. It affects 15% of couples in Western countries and has become an equally important problem in the Eastern countries (De Kretser, 1997; Evers, 2002). In roughly half of the infertile couples, the defect can be traced to the man (Poongothai et al., 2009). In men, spermatogenic failure, including oligozoospermia, asthenozoospermia, teratozoospermia and azoospermia, is the major cause of infertility (World Health Organization, 1987; Shah et al., 2003).

Many genetics factors have been implicated in human spermatogenic failure (Ferlin et al., 2007; Tuttelmann et al., 2007). Screening
with markers on the long arm of the human Y chromosome has detected Yq microdeletions in 5–15% of males with spermatogenic failure. On the euchromatic Y chromosome, the DAZL-Deletion in AZoospermia (DAZ) gene is the first to be identified as a Y-chromosomal fertility gene (Reijo et al., 1996; Vogt et al., 1996; McElreavey and Krausz, 1999). The DAZ gene has an autosomal homolog, DAZL (DAZ-Like), on chromosome 3p24. It is believed that the DAZ gene arose 40 million years ago from the transposition, repeat amplification and pruning of an ancestral autosomal gene DAZL during primate evolution (Saxena et al., 1996; Shan et al., 1996; Yen et al., 1996).

DAZL is critical to germ cell development in different species. In Xenopus, Xdazl is required for early primordial germ cell differentiation and is indirectly necessary for the migration of primordial germ cells through the endoderm (Houston and King, 2000). In mice, the loss of function of the DAZL homologue leads to the loss of germ cells in both sexes (Ruggiu et al., 1997). In humans, we have identified a single-nucleotide polymorphism (SNP) of DAZL, 386 A>G, which confers susceptibility to severe spermatogenic failure (sperm concentrations <2 × 10^6/ml) in Han Taiwanese men (Teng et al., 2002). Later, we found that specific DAZL haplotypes were related to severe spermatogenic failure (Teng et al., 2006). Given the absence of 386 A>G SNP in other ethnic groups, 386 A>G seems to be restricted to the Taiwanese Han population (Becherini et al., 2004; Tschanter et al., 2004; Yang et al., 2005; Thangaraj et al., 2006).

DAZL is only expressed in the gonads, but how this gene is regulated at the transcriptional level remains obscure. In this study, we set out to identify the core promoter region of DAZL and to identify regulatory SNPs which may be related to male infertility. Several novel SNPs in the promoter region of DAZL were identified. Of these SNPs, −792G>A was more prevalent in infertile men with spermatogenic failure than in fertile men with normal semen parameters. Quantitative trait analysis showed that the −792A allele is associated with decreased sperm concentrations and poor sperm motility. We also found that the G to A substitution may interfere with binding of nuclear respiratory factor-1 (NRF-1) to the promoter region. To the best of our knowledge, this is the first report on the association of a DAZL promoter SNP with spermatogenic failure.

### Materials and Methods

#### Subjects

The study was approved by the Institutional Review Board (IRB) of the National Cheng-Kung University Hospital and Lee Women’s Hospital. From August 2006 to July 2008, a total of 337 infertile men presenting with azoospermia or oligozoospermia (sperm concentration <20 × 10^6/ml) were recruited to the study. We also enrolled 203 fertile men with normal semen parameters as control subjects. All of the control subjects had fathered at least two children without assisted reproductive technologies. Informed consents were obtained from all enrollees. All study and control subjects belonged to Han Taiwanese population, the major ethnic group in Taiwan (making up more than 95% of the country’s population). All patients underwent comprehensive surveillance according to the protocols described in our previous studies, including a detailed history taking, physical examination, semen analyses, endocrinology profiles testing (luteinizing hormone, follicular-stimulating hormone, prolactin and testosterone), karyotyping and a molecular test for Y-chromosome microdeletions (Lin et al., 2002; Teng et al., 2002, 2006, 2007). Semen analysis was performed according to the protocol recommended by the WHO using a modified Neubauer chamber (World Health Organization, 1992). The semen samples were collected after a minimum of 2 days and a maximum of 7 days of sexual abstinence. Molecular analysis of Y-chromosome microdeletions included a combination of five gene-based markers as described previously (Teng et al., 2007). Subjects with recognizable causes of male infertility were excluded from the screening; these included varicocele, hypogonadotropic hypogonadism, toxic habits (tobacco, alcohol or drug use), occupational exposures, previous treatments (immuno-therapy, cancer therapy or hormone therapy), unilateral orchidectomy, vascetomy, cryptorchidism and chronic genital infections. Cases with an abnormal karyotype or Y chromosomal microdeletion were also excluded.

#### Prediction of the promoter region and transcription factor binding

The following websites were used to predict the regulatory element of DAZL: Promoter Scan (P2) (http://htr.cit.nih.gov/molbio/proscan/); Promoter 2.0 Prediction Server (PP) (http://www.cbs.dtu.dk/services/Promoter/); Berkeley Drosophila Genome Project (McP) (http://www.fruitfly.org/seq_tools/promoter.html); DBTSS (http://dbtss.hgc.jp/index.html); and Prediction of CpG island (CPG) (http://www.ebi.ac.uk/emboss/cpgplot/). The Transcription Element Search System (http://www.cbi.upenn.edu/cgi-bin/tess/tess) was used to predict the binding of transcription factors (TFs).

#### Genotyping

Genomic DNA was extracted from peripheral blood samples using a Puregene DNA isolation kit (Genta, Minneapolis, MN, USA). Genomic DNA samples were subjected to sequence analysis for the promoter region of DAZL. Polymerase chain reactions (PCRs) were used to amplify the DAZL promoter region. The primer sequences were as follows: human DAZL forward 5′-GGGGTACCGTGATGGCGCTAACCCTGTGCCTTAG-3′ and reverse 5′-GAAGATCTCCACACCGAGGCAGGCCCACAT-3′; GAPDH forward 5′-TGAAGGTGGAGTCAACGGAATT-3′ and reverse 5′-CCTGGAAGATGTGATGGAATT-3′. The PCR consisted of 35 cycles of gene amplification as follows: 95°C for 30 s, 65–68°C for 30 s and 72°C for 1 min. The PCR products were resolved using agarose gels and the appropriate band was excised and purified for the subsequent steps of cloning or sequence determination. Sequence analysis was performed with an automatic sequencer (ABI 377, Applied Biosystems/PE).

#### Sodium bisulfite treatment and sequence analysis

Reactions for sodium bisulfite treatment were carried out according to standard conditions of the EZ DNA methylation kit (ZYMO RESEARCH, Irvine, CA, USA). DNA was dissolved in 40 μl water and used immediately for PCR amplification or stored at −20°C. Primers were designed using Methprimer (Li and Daihya, 2002) to distinguish methylated and unmethylated DNA after bisulfite treatment of the template. The sequence of the forward primers was 5′-TTTGTTTTATGAGAAGTTTTTGGAAT-3′; reverse primer was 5′-ACCAGTTCAAAACTACTATAATAACCC-3′. The PCR conditions included initial denaturing at 95°C for 5 min, followed by 40 cycles of amplification consisting of denaturation at 95°C for 30 s, annealing at 55°C for 90 s and elongation at 72°C for 2 min. The Gel-M Gel Extraction Miniprep System (Viogene, Cat No. EG1001) was used for extraction of the DNA fragment, which was then cloned into PGEM T easy vector (Promega Corp., Madison, WI, USA). At least 20 colonies were randomly chosen for sequencing.
Plasmid construction

Reporter plasmids were constructed by inserting various lengths of the 5′-upstream region of the human DAZL gene between the Nhel and HindIII sites of the pGL3-Basic Vector (Promega Corp.). An ~1 kb fragment of the DAZL promoter spanning nucleotides −993 to +52 (−993/+52) was subcloned into the firefly luciferase reporter pGL3-Basic vector lacking the promoter (Promega Corp.). The deletion constructs (−993/+52, −733/+52, −473/+52, −204/+52) were generated by PCR. The fidelity of all constructs created was confirmed by nucleotide sequence analysis.

Site-Directed Mutagenesis

The QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) was used to introduce mutations into the −792G-LUC and −792A-LUC DAZL promoter construct according to the manufacturer’s instructions. The negative control was the DAZL gene promoter region with the NRF-1-binding site deleted. The presence of the mutations was confirmed by both restriction enzyme digest and sequencing analysis.

Transient transfection and luciferase reporter gene assay

For the reporter gene assay, HeLa and GC1 cells (1.5 × 10⁵ cells) were seeded into 2 ml medium containing 7% (v/v) FBS in each well of a 6-well culture plate, and then cultured for 24 h. Lipofectamine reagent (Invitrogen, Carlsbad, CA, USA) was used for transfection (duplicate) of plasmid DNA [reporter plasmid (0.5 mg) and pRL-TK vector (0.5 mg) per well], according to the manual supplied by the manufacturer, except that the volume was quadrupled. Lysis buffer (90 ml) was added to the phosphate-buffered saline [10 mM sodium phosphate buffer (pH 7.2), 137 mM NaCl, 3 mM KCl] (PBS)-washed plate 48 h after the start of transfection, and a cell lysate was prepared (12 000 g, 10 min at 4 °C). An aliquot (20 and 5 ml) of the supernatant was used for measurement of the luciferase activities. The pRL-TK vector (Promega Corp.) containing the Renilla luciferase gene under control of the thymidine kinase promoter was transfected into HepG2 cells. The transfection efficiency was determined by the Renilla luciferase activity, which was used as a control. Reporter activity was normalized by calculating the ratio of firefly/Renilla values.

Electrophoretic mobility shift assay (EMSA)

Bandshift assays were carried out according to standard protocols. Briefly, a reaction mixture containing 5 mg of HeLa cell or GC1 cell nuclear extract, 1–1.5 mg/ml of poly(dI-dC), 100 mg/ml of bovine serum albumin, 16 mM HEPES (pH 7.9), 120 mM KCl, 4 mM EDTA, 0.8 mM DTT, 0.06% NP-40 and 6% Ficoll 400, with or without non-labeled competitor oligonucleotides at a 25–200-fold molar ratio against the labeled probe, was incubated for 10 min at room temperature. The labeled probe of 1 × 10⁶ cpm was then added to the mixture, and the mixture was further incubated at room temperature for 30 min. DNA–protein complexes formed in the mixture were separated in a 4.5% polyacrylamide gel containing 0.25 × TBE and visualized by autoradiography. For a supershift experiment, a plasmid containing the Myc-tagged full-length NRF-1 was transfected into HepG2 cells. The nuclear extract was first incubated with various amounts of anti-Sp1 antibody, anti-NRF-1 antibody or serum IgG at 0 °C for 60 min, and then the reaction was carried out as described above after the mixture without the nuclear extract was added back to this mixture. Nucleotide sequences of the upper strand of NRF-1 and mutant-NRF-1 oligonucleotides were 5′-GTAGGGTACGCACCG[C]GT GTCCGACGGTCC-3′ and 5′-GTAGGTACGCACCG[T]GTGTGCGC AGGTCCC-3′, respectively.

Statistics

Association tests for allelic and genotypic frequencies were performed using the χ² test or Fisher exact test by GraphPad Prism 5.0. The risk of spermatogenic failure was estimated using logistic regression to report odds ratios (ORs) and 95% confidence interval (CI).

Among the four markers, only −792 G>A exhibited significant association with spermatogenic failure. We further conducted quantitative trait analyses to evaluate the distributions of the three sperm parameters in relation to this polymorphic locus. Log-transformed sperm concentration (× 10⁶/ml), motility and morphology were analyzed using PLINK (Purcell, http://pngu.mgh.harvard.edu/~purcell/plink/) for the −792 G>A polymorphism. In addition to allelic test, we employed genotypic tests with an additive model and reported the mean (± SD) of the three sperm parameters by genotype. We used regression-based models, which were implemented in PLINK to calculate the effect size and significance level for each marker with quantitative traits. Because the results from the regression models with or without age adjustment were very similar, we only presented quantitative analysis results with age adjustment. Violation of Hardy–Weinberg equilibrium (HWE) was also tested via the PLINK with a threshold of 0.05. A nominal P value of <0.01 was set as the threshold for claiming statistical significance in both binary trait and quantitative analysis for sperm parameters.

Results

Semen parameters

The age was matched between infertile men and fertile controls (31.31 ± 3.98 vs 31.30 ± 3.92, P > 0.05). Table I shows the ages as well as the three major semen parameters of patients and fertile controls.

Identification of the core promoter and CpG islands of human DAZL

In-silicon searches showed that the DAZL promoter was located about 1 kb upstream of the transcription start site (TSS). Two DAZL transcripts have been identified in the National Center for Biotechnology Information.

Table I Semen parameters of fertile and infertile men.

<table>
<thead>
<tr>
<th>Semen parameter</th>
<th>Age (years)</th>
<th>Sperm concentration (×10⁶/ml)</th>
<th>Sperm motility (%)</th>
<th>Normal spermatozoa (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>95% CI</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Control (n = 203)</td>
<td>31.30 ± 3.92</td>
<td>108.9 ± 64.61</td>
<td>99.9–117.8</td>
<td>67.17 ± 11.56</td>
</tr>
<tr>
<td>Spermatogenic failure (n = 337)</td>
<td>31.31 ± 3.98</td>
<td>12.07 ± 10.19</td>
<td>10.98–13.17</td>
<td>33.29 ± 15.73</td>
</tr>
</tbody>
</table>

Nucleotide sequences of the upper strand of NRF-1 and mutant-NRF-1 oligonucleotides were 5′-GTAGGGTACGCACCG[C]GT GTCCGACGGTCC-3′ and 5′-GTAGGTACGCACCG[T]GTGTGCGC AGGTCCC-3′, respectively.
associated with increased luciferase activities. The results of the dual luciferase reporter assay are shown in Fig. 1B and C. We found that all four fragments were associated with increased luciferase activities. The $-993/\pm 52$ and $-808/\pm 52$ fragments caused the highest activities; in comparison the promoter sequence up to $-473$ resulted in reduced luciferase activity. Interestingly, a truncation of the promoter at $-808$ instead of $-993$ resulted in a substantial increase in the promoter activity in GC1 cells. We suggest that the DAZL core promoter is located in the region from $-808$ to $-473$ from the TSS. The CpG islands are located approximately between positions $-865$ and $-461$ from the TSS (Supplementary Fig. 1A). A total of 40 CpG sites were located within the predicted CpG islands (total length 405 bp) (Supplementary Fig. 1B). We also found that the DAZL promoter is frequently hypermethylated in normal blood cells but hypomethylated in the spermatozoa (Supplementary Fig. 2).

Regulatory SNPs conferring susceptibility to spermatogenic failure in Taiwanese men

We systematically sequenced up to 1 kb upstream of the DAZL TSS. Three SNPs ($-792G>A$, $-669A>C$ and $-309T>C$) were identified in both fertile men and infertile men (Fig. 2). Of these three SNPs, two ($-669A>C$ and $-309T>C$) have not been reported in databases or in the literature. The allelic and genotypic frequencies of these three SNPs are shown in Table II. Of the three SNPs, A was significantly associated with spermatogenic failure ($P = 0.0005$) with an OR of 8.13 (Table II). Even using the more stringent Bonferroni correction ($\alpha = 0.05/3 = 0.0167$), the significant finding still holds for $-792G>A$ ($P = 0.0005$). We further conducted 5000 permutations to obtain an empirical $P$-value for this marker, and the $P$-value was 0.0006, which is still significant. Fisher’s exact test showed that all SNPs were in HWE. The OR for each genetic homozygote variants is shown in Table III. The OR of $-792A$ variants was estimated to be 8.43 (95% CI = 1.98–35.90). Considering the bias or lack of precision in the OR estimate due to the small sample size, we also used the bootstrap method to resample 2000 times with 500 individuals in each sample to calculate a bootstrap OR and 95% CI. For $-792G>A$, the bootstrap allele OR was 8.98 (95% CI = 2.58–19.57), and the genotypic OR was 9.33 (95% CI = 2.64–20.62).

SNP $-792G>A$ decreases DAZL transcription by affecting NRF-1 binding

To investigate whether the nucleotide substitution of G by A would modify binding affinity of the promoter with nuclear proteins, EMSA was performed by incubating nuclear protein extracts of HeLa and GC1 cells with double-stranded oligonucleotide probes containing either wild-type (G) or mutant (A) allele. Using the DNA–EMSA, we found that $-792G>A$ (rs2241533) appeared to affect the binding of nuclear proteins from the GC1 cells (Fig. 3A). GC1 is a mouse type B spermatogonia cell line and was found to have endogenous dazl expression (data not shown).

Binding by NRF-1 was confirmed by the supershift assay using HepG2 cells over-expressing the Myc-tagged NRF-1. The signal for the DNA–protein complex containing the NRF-1-myc fusion protein was reduced by the monoclonal anti-Myc antibody. Considering NRF-1 usually binds to the GC-rich region in the promoter and usually interacts with other TFs, such as Sp1, for transcriptional regulation, we used unlabeled NRF-1 and Sp1 antibodies for the competition experiments. The DNA–protein complex was competed by NRF-1 antibody (Fig. 3B, lane 4), but not by serum IgG and Sp1 antibody (Fig. 3B, lane 3 and 5).

Figure 1 Transcriptional activity in the upstream region of the DAZL gene. (A) The deletion constructs of an approximately 1 kb fragment spanning the DAZL transcription start site (TSS) (positions $-993\text{ to }+52$) were subcloned into the pGL3-Basic luciferase reporter vector. Nucleotide numbering starts with $+1$ corresponding to the TSS which was determined according to NM_001190811.1. Cells (B, HeLa; C, GC1) were transfected with 200 ng of either empty vector or the DAZL promoter reporter vector in conjunction with 20 ng of control Renilla luciferase vector for normalization. Extracts were taken 24 h after transfection and the luciferase activity was determined. The values shown are the mean levels of luciferase expression after normalization for transfection efficiency by using control Renilla luciferase expression and are quoted in arbitrary units. Error bars represent the standard deviation from the mean ($n = 5$). Relative luciferase activity of each construct is shown in the right panel. Results are expressed as a percentage of the luciferase activity.
Functional significance of \(-792A\)

The HeLa (Fig. 4) and GC1 (Fig. 4B) cells transfected with \(-792G\) allele had a 1.5- to 2-fold increase in reporter gene activity as compared with the \(-792A\) allele, respectively. The result indicates that the G to A substitution decreases transcriptional activity of \(DAZL\).

We therefore further explored the association between \(-792G>A\) genotypes and three major semen parameters. Quantitative trait analysis showed that the \(-792A\) allele is negatively associated with sperm concentrations and sperm motility in both control subjects and patients. The two groups were pooled for analysis and the results are shown in Table IV.

**Discussion**

NRF-1 is a TF that acts on nuclear genes encoding respiratory subunits and components of the mitochondrial transcription and replication machinery (Gopalakrishnan and Scarpulla, 1995; Moraes, 2001). The NRF-1-binding site is one of the most abundant binding sites specific to the spermatogonia stage by transcriptional and biological network analysis. Although previous studies do not directly show that NRF-1 is involved in regulating germ-line-specific genes, it is reasonable to speculate NRF-1 may play an important role in spermatogenesis (Lee et al., 2006). In this report, we provide direct evidences that NRF-1 is involved in transcriptional regulation of \(DAZL\), a human sterility gene.

There is a growing consensus that gene expression changes are likely to play an important role in human phenotypic variation and in disease predisposition (Pastinen and Hudson, 2004; Chen et al., 2006; Khaitovich et al., 2006). Consistent with this point, experimental evaluation of the effects of genetic variation on gene transcription plays an important role in the effort to dissect the genetic bases of human

**Table II** Allelic frequencies of three sequence variants of \(DAZL\) in infertile men and control subjects.

<table>
<thead>
<tr>
<th>Allele</th>
<th>Control ((N = 406))</th>
<th>Spermatogenic failure ((N = 674))</th>
<th>(P) (odds ratio)</th>
<th>Genotype</th>
<th>Control ((N = 203))</th>
<th>Spermatogenic failure ((N = 337))</th>
<th>(P) (odds ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-792G&gt;A) G</td>
<td>404 (99.51%)</td>
<td>648 (96.14%)</td>
<td>0.0005 (8.13)</td>
<td>GG</td>
<td>201 (99.01%)</td>
<td>311 (92.28%)</td>
<td>0.0004 (8.429)</td>
</tr>
<tr>
<td>A</td>
<td>2 (0.49%)</td>
<td>26 (3.86%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-669A&gt;C) A</td>
<td>405 (99.75%)</td>
<td>673 (99.85%)</td>
<td>1 (0.60)</td>
<td>AA</td>
<td>202 (99.51%)</td>
<td>336 (99.70%)</td>
<td>1 (0.6030)</td>
</tr>
<tr>
<td>C</td>
<td>1 (0.25%)</td>
<td>1 (0.15%)</td>
<td></td>
<td>AC</td>
<td>1 (0.49%)</td>
<td>1 (0.30%)</td>
<td></td>
</tr>
<tr>
<td>(-309T&gt;C) T</td>
<td>401 (98.77%)</td>
<td>671 (99.55%)</td>
<td>0.16 (0.36)</td>
<td>TT</td>
<td>198 (97.54%)</td>
<td>334 (99.11%)</td>
<td>0.16 (0.3568)</td>
</tr>
<tr>
<td>C</td>
<td>5 (1.23%)</td>
<td>3 (0.45%)</td>
<td></td>
<td>TC</td>
<td>5 (2.47%)</td>
<td>3 (0.89%)</td>
<td></td>
</tr>
</tbody>
</table>

Allele and genotype frequencies were compared between infertile men and control subjects.

**Table III** Odds ratios for each sequence variant of \(DAZL\).

<table>
<thead>
<tr>
<th>Sequence variants</th>
<th>Tests for association</th>
<th>(\text{Odds ratio})</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-792G&gt;A)</td>
<td>G/G vs A/G</td>
<td>8.43</td>
<td>1.98–35.90</td>
</tr>
<tr>
<td>(-669A&gt;C)</td>
<td>A/A vs A/C</td>
<td>0.60</td>
<td>0.04–9.69</td>
</tr>
<tr>
<td>(-309T&gt;C)</td>
<td>T/T vs T/C</td>
<td>0.36</td>
<td>0.08–1.51</td>
</tr>
</tbody>
</table>
complex disease. Spermatogenic failure is a complex disease with multiple environmental and genetic factors contributing to its pathogenesis (Shah et al., 2003; Ferlin et al., 2007; Tuttelmann et al., 2007; Poongothai et al., 2009). The information available on how human genetic variation affects expression levels and how genetic information is translated into phenotypic variation in spermatogenesis is very limited. In this study, we found three SNPs located in the DAZL promoter region in our population. Of these three SNPs, −792G>A was significantly associated with spermatogenic failure. We demonstrated that −792A affects transcriptional activity and binding of nuclear proteins. It is noteworthy that −792G>A is located within the CpG islands. We reason that the G to A substitution can reduce DNA methylation, which in turn regulates SPAG16L, the sperm axoneme central apparatus protein necessary for flagella motility and sperm motility (Zhang et al., 2006, 2008; Zeng et al., 2008).

DAZL encodes a germ-cell-specific RNA-binding protein, which may be involved in post-transcriptional regulation of many substrate genes during spermatogenesis (Reynolds and Cooke, 2005). Although DAZL expression is highly regulated in the germ-cell lineages, the element that contributes to this exquisite expression has not been fully elucidated. In this study, about 500 bp sequences (−473 to −993) upstream of the human DAZL TSS was identified as having strong promoter activity. In the mouse, a 1.7 kb sequence in the 5′ flanking region of the mouse Dazl gene has been shown to be sufficient to drive robust germ cell-specific expression of green fluorescent protein in adult testis, with lower levels of expression in the adult ovary and fetal and newborn gonads of both sexes (Nicholas et al., 2009). Comparison of genomic regions upstream of the TSS in mouse and human Dazl genes showed poor conservation of DNA.

**Figure 3** The binding activity of −792G>A of DAZL promoter was confirmed by electrophoretic mobility shift analysis (EMSA). (A) Electrophoretic mobility shift analysis of −792G>A binding activity in HeLa and GC1 cells. Samples of 3 μg of HeLa (lanes 2, 5, 6) or GC1 (lanes 3, 7, 8) nuclear extracts were incubated with the −792G (lanes 4, 5, 7) DAZL–NRF-1, −792A (lanes 6, 8) DAZL–NRF-1, or SP1-labeled (lanes 1–3) oligonucleotides at 4°C for 30 min and then electrophoresed on a 4% Tris borate–EDTA polyacrylamide gel. The Sp1 consensus probe incubated with HeLa (lane 2) or GC1 (lane 3) nuclear proteins was used as a positive control. Mixtures containing no nuclear extract were used as negative controls (lanes 1, 4). The shifted band of the SP1 probe is indicated by the arrowhead; the shifted band of the DAZL–NRF-1 probe is indicated by the arrows. (B) Supershift analysis of the Myc-tagged fusion proteins. Plasmids expressing the Myc-tagged full-length NRF-1 were transfected into HepG2 cells. Transfected HepG2 cells nuclear proteins could bind to the radiolabeled −792G probe (lane 1). When 2 μg of anti-myc polyclonal antiserum were incubated with 30 μg of nuclear extracts prior to the addition of the −792G probes, the band was supershifted (lane 2). Normal mouse IgG was used as control (lane 3). Unlabeled NRF-1 (lane 4) or SP1 antibodies (lane 5) were incubated with nuclear proteins. The arrowhead indicates the supershifted band. The arrow indicates the NRF-1 binding.
sequences (data not shown). Although ultra-conserved non-coding DNA may suggest selective constraints or purifying selection, divergent DNA sequences do not necessarily imply functional dispensability or redundancy (Dermitzakis et al., 2005; McLean and Bejerano, 2008; Visel et al., 2009; Pennacchio and Visel, 2010). It remains to be explored whether human DAZL and mouse Dazl are subjected to similar regulatory mechanisms.

Genome-wide analysis suggests that tissue-specific differentially methylated regions may play a role in cellular identity and the regulation of tissue-specific function (Song et al., 2005, 2009; Ohgane et al., 2008; Rakyan et al., 2008; Chatterjee et al., 2011). Genome-wide analysis has also shown a distinctive methylation status of testicular DNA in germ, when compared with somatic, cells (Oakes et al., 2007). Methylated DNA may also be bound by proteins known as methyl-CpG-binding domain (MBD) proteins. MBD proteins then recruit additional proteins to the locus, such as histone deacetylases and other chromatin remodeling proteins thereby to form compact, inactive chromatin (Robertson, 2005). Germ cell-specific genes (e.g. Mvh, Dazl and Scp3) are usually silenced by DNA methylation in both pre-migratory germ cells and somatic cells, but not in post-migratory germ cells (Maatouk et al., 2006). In this study, we found that the CpG islands of the DAZL promoter is hypermethylated in the somatic cells, a finding in accord with previous reports (Chai et al., 1997; Navarro-Costa et al., 2010). The 5' end of DAZL and DAZ genes are hypomethylated in spermatozoa but not in leukocytes or placenta, consistent with the expression pattern of the genes (Chai et al., 1997). In vitro methylation of a reporter construct corresponding to these 149 bp resulted in complete suppression of porcine DAZL promoter activity in primordial germ cells, further supporting a role for methylation in regulating porcine DAZL expression (Linher et al., 2009). It has yet to be explored how CpG islands of the DAZL promoter are methylated in somatic cells while de-methylated in the germline. By using ejaculated sperm from olio-asthenoteratozoospermic (OAT) men, Navarro-Costa et al. showed there are 31 CpGs for DAZL, a number slightly lower our study. They also found disrupted DNA methylation of DAZL in OAT men. Their finding provided a link between incorrect epigenetic marks of sterile genes and male gametogenic defects. However, aberrant DNA methylation of the DAZL promoter may be a phenomenon secondary to abnormal spermatogenesis, instead of the cause of the spermatogenic defect considering that many genes may be dys-regulated in the poor-quality sperm (Garrido et al., 2009; García-Herrero et al., 2010; Navarro-Costa et al., 2010). We tend to believe that −792G>A is a bona fide regulatory SNP, which predispose patients with the A allele to spermatogenic failure by decreasing DAZL expression. Two DAZL transcripts have been identified in the database: NM_001190811.1 for transcript variant 1 and NM001351.3 for transcript variant 2. The distance between the TSS’s of these two transcripts is 709 bp. In this study, TSS and SNP locations were determined according to the transcript variant 1. Thus functional validation of the −792A allele could be only applicable to the transcript variant 1. Nevertheless, these two transcripts seem to share the same CpG islands, considering no difference in the CpG islands’ location between Navarro-Costa’s study (Navarro-Costa et al., 2010) and ours (Supplementary Fig. 3). The differences in expression patterns, regulatory mechanisms and functions of the two DAZL transcripts await further investigation.

Our study provides strong evidence for an association of a DAZL promoter polymorphism with spermatogenic failure. We also provide the first clue about regulatory mechanisms of DAZL. Of 49 DAZL promoter SNPs in the database (http://www.ncbi.nlm.nih.gov/SNP/), only one SNP (−792A>G) was identified in the Taiwanese Han population. We also found three novel SNPs, making the total number of DAZL promoter SNPs add to 51. It would be necessary to replicate our findings in different ethnic groups. Our report also adds −792G>A to the growing list of regulatory SNPs in spermatogenic failure. Indeed, only a limited number of regulatory SNPs have been reported to be associated with male infertility, including protamine genes, human follicle-stimulating hormone receptor gene, estrogen receptor alpha \xF7\x80\xF7\x80\x80gene and tumor necrosis factor-alpha gene (Wunsch et al., 2005; Guarducci et al., 2006; Gazquez et al., 2008; Tronchet et al., 2008). Our findings may shed light on

### Table IV Quantitative trait analysis of −792A genotype and semen parameters in study subjects.

<table>
<thead>
<tr>
<th>−792G&gt;A Genotype</th>
<th>Genotype counts (n%)</th>
<th>Mean ± SD</th>
<th>Model</th>
<th>β*</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0 (0%)</td>
<td>–</td>
<td>Total samples</td>
<td>−36.09</td>
<td>2.50 × 10^{-3}</td>
</tr>
<tr>
<td>A/G</td>
<td>28 (5.2%)</td>
<td>14.21 ± 13.18</td>
<td>Cases</td>
<td>−0.25</td>
<td>0.809</td>
</tr>
<tr>
<td>G/G</td>
<td>512 (94.8%)</td>
<td>50.01 ± 62.66</td>
<td>Controls</td>
<td>−58.17</td>
<td>0.205</td>
</tr>
<tr>
<td>Motility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0 (0%)</td>
<td>–</td>
<td>Total samples</td>
<td>−21.98</td>
<td>1.46 × 10^{-7}</td>
</tr>
<tr>
<td>A/G</td>
<td>28 (5.2%)</td>
<td>25.21 ± 12.96</td>
<td>Cases</td>
<td>−11.05</td>
<td>5.39 × 10^{-4}</td>
</tr>
<tr>
<td>G/G</td>
<td>512 (94.8%)</td>
<td>47.19 ± 21.57</td>
<td>Controls</td>
<td>−13.84</td>
<td>0.092</td>
</tr>
<tr>
<td>Morphology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0 (0%)</td>
<td>–</td>
<td>Total samples</td>
<td>1.17</td>
<td>0.443</td>
</tr>
<tr>
<td>A/G</td>
<td>28 (5.2%)</td>
<td>21.96 ± 4.89</td>
<td>Cases</td>
<td>2.12</td>
<td>0.255</td>
</tr>
<tr>
<td>G/G</td>
<td>512 (94.8%)</td>
<td>20.80 ± 7.98</td>
<td>Controls</td>
<td>1.94</td>
<td>0.566</td>
</tr>
</tbody>
</table>

*aThe effect size and significance results were based on regression analysis with additive genetic model adjusted for age.*
the regulatory mechanism of human DAZL. Of 51 DAZL promoter SNPs available at the present time, two SNPs (rs2241532 and rs36004482) are expected to be located in the SPI-binding sites (the Transcription Element Search System, http://www.cbil.upenn.edu/cgi-bin/tess/tess). NRF-1 usually corroborates with other TFs for transcriptional regulation. FMR-1 is mainly expressed in the brain and germ-line and NRF-1 and Sp1 have been shown to be strong, synergistic activators of an unmethylated human FMR1-driven reporter (Smith et al., 2004). Herein, we provide preliminary evidence for the role of a novel SNP in the DAZL promoter in human spermatoogenic failure. Further studies are required to explore whether NRF-1 corroborates with Sp1 or other TFs to regulate DAZL.

**Supplementary data**


**Acknowledgement**

We thank the ‘National Translational Medicine Resource Center’ for statistical consultation in the data analysis.

**Authors’ roles**

P.-L.K., M.-S.L. and I.-W.L. played roles in clinical care and sample collections. J.T.T. and Y.-P.C. participated in the measurement of SNPs and EMSA. P.-H. K. contributed with statistical consultation in the data analysis. Y.-N.T. contributed with some ideas, manuscript revisions and critical discussion. P.-L.K. was the leader of the research.

**Funding**

This study was supported by grants from the National Science Council of the Republic of China, NSC-95-2314-B-006-031-MY2 and NSC-97-2628-B-006-012-MY3.

**Conflict of interest**

None declared.

**References**


Chen JM, Ferenc C, Cooper DN. A systematic analysis of disease-associated variants in the 3′ regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3′ UTR variants. *Hum Genet* 2006; **120**:301 – 333.


Hsu CC, Kuo PH, Lee IW, Su MT, Tseng JT, Kuo PL. Quantitative trait analysis suggests human DAZL may be involved in regulating sperm counts and motility. *Reprod Biomed Online* 2010; **21**:77 – 83.


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.

Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 2002; **18**:1427 – 1431.


Maatouk DM, Kellam LD, Mann MR, Lei H, Li E, Bartolomei MS, Resnick JL. DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* 2006; **133**:3411 – 3418.


McLean C, Bejerano G. Dispensability of mammalian DNA. *Genome Res* 2008; **18**:1743 – 1751.


