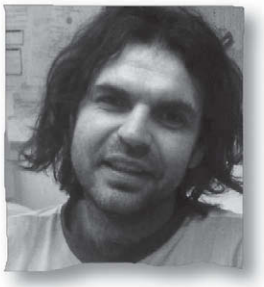
Article

TSPY gene copy number as a potential new risk factor for male infertility



Radek Vodicka PhD obtained his degree in biology from the Faculty of Science, Masaryk University, Brno, Czech Republic, specialising in Genetics and Molecular Biology. He joined the DNA laboratory of the Department of Medical Genetics and Foetal Medicine, the University Hospital, Olomouc where he obtained a post-graduate qualification in the Laboratory Method of Medical Genetics (2002). He defended his PhD thesis 'Y-chromosomal sequences in Turner syndrome patients' in 2004. His main research and clinical interest is a quantitative analysis of gonosomal sequences in relation to gonosomal aberrations and male infertility.

Dr Radek Vodicka

Radek Vodicka[1,6], Radek Vrtel[1], Ladislav Dusek[2], Arvind R Singh[1], Katerina Krizova[1], Veronika Svacinova[1], Vera Horinova[3], Jiri Dostal[4], Ivana Oborna[4], Jana Brezinova[4], Ales Sobek[5], Jiri Santavy[1]
[1]Department of Medical Genetics and Fetal Medicine, University Hospital Olomouc, IP Pavlova 6, 775 20 Olomouc; [2]Centre of Biostatics and Analyses, Masaryk University Brno, Kamenice 126/3, Brno; [3]Clinic of Reproductive Medicine and Gynaecology, U lomu 5, 760 01 Zlin; [4]Centre for Assisted Reproduction, Department of Obstetrics and Gynecology, University Hospital, IP Pavlova 6, 775 20 Olomouc; [5]Fertimed, Infertility Centre, Boleslavova 246/2, 772 00 Olomouc, Czech Republic
[6]Correspondence: Tel: +42 585854461; Fax: +42 585414906; e-mail address: vodickar@fnol.cz

Abstract

The human *TSPY* (testis-specific protein, Y-linked) gene family (30–60 copies) is situated in the *MSY* (male-specific) region of the Y chromosome. Testis-specific expression indicates that the gene plays a role in spermatogenesis. Refined quantitative fluorescence PCR (polymerase chain reaction) was applied to evaluate the relative number of *TSPY* copies compared with *AMELY/X* (amelogenin gene, Y-linked) genes in 84 stratified infertile men and in 40 controls. A significantly higher number of *TSPY* copies was found in infertile men compared with the controls ($P = 0.002$). The diagnostic discrimination potential of the relative number of *TSPY* copies was evaluated by receiver operating characteristic curve analysis. *TSPY/AMELY* was unambiguously found to be powerful in the diagnostic separation of both the control samples and the infertile men, reaching a good level of specificity (0.642) and sensitivity (0.732) at a cut-off point of 0.46. The findings were supported by independently repeated studies of randomly selected positive samples and controls. Evaluation of the *TSPY* copy number offers a completely new diagnostic approach in relation to the genetic cause of male infertility. The possible effect of the copy number of *TSPY* genes on spermatogenesis may explain indiscrete pathological alterations of spermatid quality and quantity.

Keywords: *capillary electrophoresis, male infertility, multicopy gene, quantitative fluorescence PCR,* TSPY *gene, Y chromosome*

Introduction

The human genome obviously contains numerous genes that could be involved in male infertility (Vogt, 2005). In addition, non-coding regions may also be involved in reproduction failure (Chatzimeletiou *et al*., 2006). Genetic-based spermatogenesis impairment or failure might be influenced by mutations in about 150 testis-specific expressed genes (Wang *et al*., 2004). The search for Y chromosome-specific male infertility candidate genes and diagnostics is mostly concentrated on the *AZF* (azoospermia factor region of the human Y chromosome) mapped to Yq11 (Vogt *et al*., 1996; Chai *et al*., 1997; Huynh *et al*., 2002; Ferlin *et al*., 2003, 2005; Yen, 2004).

The role of the *TSPY* (testis-specific protein, Y-linked) gene in spermatogenesis and tumourigenesis has recently been intensively investigated, mainly regarding RNA and protein concentrations (Kersemaekers *et al*., 2005; Yin *et al*., 2005). The majority of *TSPY* gene copies are arranged in 20.4 kb tandem repeat units in a 700-kb cluster ampliconic region on Yp (*DYZ5*) proximally near the centromere (Tyler-Smith *et al*., 1988; Manz *et al*., 1993; Skaletsky *et al*., 2003). Two *TSPY* copies are also localized on Yq (Dechend *et al*., 2000; Ratti *et al*., 2000). The overall number of *TSPY* copies, including pseudogenes, is estimated to range from 30 to 60 copies

(Tyler-Smith *et al.*, 1988; Manz *et al.*, 1993). The approximately 2.8 kb gene unit (six exons and five introns) comprises a 924-bp major prototypic coding sequence (Schnieders *et al.*, 1996; Vogel *et al.*, 1998). Sequence polymorphism of *TSPY* was found to be up to 10% (Manz *et al.*, 1993). Sequence differences in coding and promoter regions of functional *TSPY* members were reported to be up to 1% (Vogel *et al.*, 1998). Relatively wide transcript heterogeneity is a noticeable feature of the *TSPY* gene. Three different transcripts were found by Manz *et al.* (1993), two functional transcript variants produced by alternative splicing at the intron 4 site (Krick *et al.*, 2003), and two transcripts caused by cryptic splicing in donor/acceptor sites within exon 1 (Vogel *et al.*, 1998) were also described. Other rare splice variants are assumed with respect to sequence polymorphisms. Native expression (testes-specific) and tumour-specific expressions (gonadoblastoma, seminoma, prostate cancer, hepatocellular carcinoma) have been described and confirmed by cDNA and immunoassay analyses (Vogel *et al.*, 1998; Lau, 1999; Lau *et al.*, 2003; Kersemaekers *et al.*, 2005; Yin *et al.*, 2005). Low level expression was also observed in epithelial cells of the prostate (Lau *et al.*, 2003). *TSPY* belongs to the protein family (with conserved *SET/NAP* domain), which includes *SET* oncoprotein, nucleosome assembly protein-*1 (NAP-1), TSPY,* differentially expressed nucleolar transforming growth factor β1□ (TGF-β1□) target (*DENTT)* (Ozbun *et al.*, 2003, 2005), cell division autoantigen-1 *(CDA-1)* (Chai *et al.*, 2001) and *TSPX* (Delbridge *et al.*, 2004). Amino acid similarities to the *SET/NAP* protein family, and particular gene expression, strongly indicate that the *TSPY* gene is involved in spermatogenesis (Vogel *et al.*, 1998). The exact determination of the number of *TSPY* copies in association with male infertility has not yet been fully described. In this study the hypothesis that the copy number of the *TSPY* gene influences the quality of spermatogenesis was tested by refined quantitative fluorescent polymerase chain reaction (RFQ-PCR) (Vodicka *et al.*, 2004). The multi-copy *TSPY* gene locus was combined with a single copy gene from the Y and X chromosomes to determine the relative quantity of *TSPY* copies.

Materials and methods

Patients and control families were informed about the study aims and consented to taking part in the study. The project and sample acquirement were approved by the ethical committee at the University Hospital Olomouc under the regulations of the Czech Republic and European Union. In all of the infertile cases investigated, endocrine disorders, trauma, chronic diseases, cryptorchidism and obstructive azoospermia were ruled out and karyotyping was undertaken. Only 46,XY normal karyotype patients were analysed. Patients were divided into azoospermic (n = 24), spermatozoa concentration ≤5 × 10^6/ml (n = 36), spermatozoa concentration ≥5 × 10^6/ml (n = 15) groups, according to their spermiogram, and an *AZF* deletion (n = 9) group was also included. Six cases had deletions in *AZFc*, one in *AZFa*, one in *AZFb* and one in *AZFb,c*. All of them were azoospermic. The mean age of the patients was 33.09 years, ranging from 25 to 47 years. Selection of controls was strictly based on spermiogram normality, proven paternity and Y chromosome genotypic heterogeneity. The paternity of 40 control samples with normal spermiogram ranging from 18 × 10^6/ml to 155 × 10^6/ml with arithmetic average 59 × 10^6/ml (mean age 33.27 years, ranging from 28 to 43 years) was proven using 15 informative STR (short tandem repeats) loci in mother, father and child. It is very important to define and prove Y haplogroup heterogeneity and non-consanguinity in infertile men and controls as the Y chromosome structure and gene content varies among the different Y haplogroups. Y chromosomal haplotyping was performed using AmpF*l*STR Yfiler kit (Applied Biosystems, USA) in 17 STR loci. Fourteen control and 11 patient samples from the second independent test were chosen for the evaluation of Y-haplogroup differences. Individual profiles of haplotypes were processed both by univariate analysis of molecular variance (AMOVA) and by multivariate (cluster analysis) methods. Both methods proved non-significant differences between control and patient samples. The within-group variance of haplotype counts masked any between-group differences in all examined loci (P = 0.621–0.856) and cluster analysis mixed patients and healthy people even at the lowest level of linkage distance. All DNA samples were isolated from 10 ml of peripheral blood preserved in EDTA using the salting out method (Miller *et al.*, 1988). After spectrophotometrical measuring of the DNA concentration (Beckman Coulter DU 530; Beckman Coulter, USA), the working DNA solutions were diluted to 100 ng/μl. Quantitative potential and method sensitivity of RQF-PCR were previously thoroughly evaluated on gonosomal mosaic assessment in the author's department (Vodicka *et al.*, 2004). In principle, RQF-PCR assesses the amount of fluorescent-labelled PCR products from multi-copy genes and compares it with a single copy gene PCR product in the exponential phase of PCR cycles. The *TSPY* specific locus from intron 1 was chosen due to intronless pseudogene sequences that might influence functional gene quantification. In addition, another set of primers from *TSPY* (intron 1 and exon 2) and *AMELY/X* was chosen for the second confirmation independent experiment (2CIE). The primers used for *TSPY* (forward 5′-CGGGGAA GTGTAAGTGACCGATGGG-3′ and reverse 5′-CTGCTCTT CAAAAAGATGCCCCAAA-3′; second experiment forward 5′-GAGGTGCTCTCGGGGAAGTGTAAGTG-3′ and reverse 5′-GAGGGTGTATGATTCTGAGGCTGACTG-3′) and for *AMELY/X* (forward 5′-CTGATGGTTGGCCTCAAGCCT-3′ and reverse 5′-ATGAGGAAACCAGGGTTCCA-3′; second experiment forward 5′-CCCTGGGCTCTGTAAAGAA-3′ and reverse 5′-ATCAGAGCTTAAACTGGGAAGCTG-3′) were fluorescently labelled by TET (phosphoramidite) and 6-FAM (phosphoramidite) respectively. *TSPY* amplification specificity was proven by sequencing the PCR product and by including a female negative control. The *TSPY* and *AMELX/Y* primer mix was in a molar concentration ratio of 1:1. Multiplex PCR procedure was tested for the usual set of chemicals (25 μl of PCR mixture contains: 100 ng DNA, 15 pmol of each primer, 200 μmol/l dNTPs, 1.5 IU of polymerase (Gibco BRL, USA), 15 mmol/l $MgCl_2$, 2.5 μl of 10× reaction buffer, Gibco BRL). A different set of chemicals was used in the first independently repeated test; Restorase polymerase with supplied chemicals and Big Band primers (Sigma-Aldrich, Germany). This first independent confirmation test was performed by another worker. The second confirmation approach, primer independent quantification, was performed using Combi PPP Master Mix (Top-Bio, Czech Republic). The PCR protocol consisted of 10 μl of Master Mix, 7.5 pmol of each primer and 100 ng of DNA. Optimal quantifiable amplification was tested in each of the cycles, from 15 to 45. The exponential RFU (relative fluorescent unit) growth path for all loci together was determined from 22 to 28 cycles and from 19 to 30 cycles for 2CIE. RQF-PCR was performed in four independent tubes