WILEY-BLACKWELL

Asian J Androl 2008; 10 (3): 461–466
DOI:10.1111/j.1745-7262.2008.00386.x

AJA

·*Original Article*·

Novel *UBE2B*-associated polymorphisms in an azoospermic/oligozoospermic population

Ivan Huang[1], Benjamin R. Emery[2], Greg L. Christensen[2], Jeanine Griffin[2], C. Matthew Peterson[1], Douglas T. Carrell[2]

[1]*Division of Reproductive Endocrinology,* [2]*Andrology and IVF Laboratories, University of Utah Health Sciences, Salt Lake City, Utah 84106, USA*

Abstract

Aim: To assess whether abnormalities exist in the *UBE2B* gene in a population of infertile human males, and to establish biologic plausibility of any discovered mutations. **Methods:** We carried out polymerase chain reaction (PCR) amplification and sequence analysis of the 5'-untranslated region and six exons of the *UBE2B* gene, including flanking intronic regions, in a group of fertile and infertile men. Following the identification of a putative promoter region that contained single or dual triplet deletions within a 10-CGG repeat island, we evaluated the binding affinity of these identified polymorphisms as compared to the wild-type sequence to transcription factor SP1 using a DNA–protein gel shift assay. **Results:** There was a novel exonic single nucleotide polymorphism (SNP) noted in exon 4 in 5% of infertile men. *In silico* 3D modeling of the altered protein showed an innocuous isoleucine for valine substitution. There were no mutations noted within any of the other exons. Three novel intronic SNPs were identified within the fertile group, and seven novel intronic SNPs identified in the infertile group. The DNA–protein gel shift assay noted that both single CGG deletion and double CGG deletion bands had approximately twice the binding affinity compared to the wild-type for SP1. The negative control confirmed no non-specific protein binding. **Conclusion:** By themselves, a single or double CGG deletion is unlikely to pose biologic significance. However, such deletions in this suspected promoter region are associated with increased binding affinity for SP1, and might represent one of several factors required for alteration of *UBE2B* gene expression. *(Asian J Androl 2008 May; 10: 461–466)*

Keywords: *UBE2B*; SP1; spermatogenesis; male infertility; single nucleotide polymorphism

1 Introduction

UBE2A and *UBE2B* are mammalian homologs of *RAD6* in yeast, and have 96% and 100% amino acid (a.a.) identity to *RAD6*, respectively. *UBE2B* encodes a key protein in a pathway essential for post-replicative DNA repair (*RAD6* pathway) and has been shown in the mouse knockout to lead to infertile male mice [1]. This infertility has been shown to occur through disruption of primary prophase and to increase apoptosis of primary spermatocytes [2]. Of interest, the *UBE2A* male mouse knockout is fertile but the *UBE2A* female is not. *UBE2A*

Correspondence to: Dr Ivan Huang, 30 North Medical Drive, University of Utah Health Sciences, Salt Lake City, Utah 84106, USA.
Tel: +1-801-581-3834 Fax +1-801-585-2388
E-mail: fertilitydr@hsc.utah.edu
Received 2007-10-18 Accepted 2007-12-16

null oocytes fertilized with wild-type (WT) sperm fail to pass the two-cell block [3].

The *UBE2B* gene is located on the long arm of chromosome 5 (5q31.1), has six exons spanning a length of approximately 21 000 bp, and encodes a protein of approximately 150 a.a. It is one of a number of candidate genes for male infertility and thought to be critically involved with spermatogenesis [1]. It is commonly referred to as *HR6B* or the *RAD6* yeast homolog. In humans, there are currently 147 known polymorphisms of the *UBE2B* gene [4], none of which have been reported to be associated with male infertility.

Our goal for this study was to evaluate the candidacy of the *UBE2B* gene as a factor in human male infertility. We planned to sequence the *UBE2B* gene and its promoter region in a group of infertile human males with severe oligozoospermia or non-obstructive azoospermia, and, following comparison of any identified mutations to a reference population of fertile men, carry out appropriate studies to establish biologic plausibility.

2 Materials and methods

In the population of infertile males attending the Utah Center for Reproductive Medicine (Salt Lake City, UT, USA), 96 men with severe oligozoospermia (less than 5×10^6/mL) or non-obstructive azoospermia were identified. Men were identified as azoospermic if the ejaculate was confirmed to be void of any sperm by centrifugation of the ejaculate and evaluation of the seminal pellet by at least two trained technicians. The diagnosis of severe oligozoospermia or azoospermia was conducted in a College of American Pathologists certified andrology clinic following American Society of Reproductive Medicine guidelines. Individuals diagnosed with any condition or treatment connected with infertility (obstructive azoospermia, cystic fibrosis, varicocele, or prior exposure to chemotheraputics) were excluded. Genomic DNA was prepared from peripheral blood and stored until analysis was ready to be carried out.

Genomic DNA of 96 men with proven fertility was obtained from the Utah Genetic Reference Project (UGRP) and served as fertile controls. To ensure paternity of the fertile controls from the UGRP, men were chosen from generations of the UGRP where there are records of, and genetic verification of, a subsequent generation. Men from both study and control groups were of northern European descent.

2.1 Polymerase chain reaction (PCR) analysis of 5'-untranslated region (UTR) and exons

Promoter identification software was used to identify a putative promoter region in the 5'-UTR of the *UBE2B* gene [5, 6]. PCR primers (Table 1) were developed and optimized to amplify the region 500 bp upstream of the *UBE2B* coding region encompassing this putative promoter region. Six sets of PCR primers were also developed and optimized to amplify the six exons of *UBE2B* with a range of 70–200 bp flanking the upstream and downstream sequences of the exons. Unless specified, forward or reverse amplification primers were also used as sequencing primers, otherwise internal sequencing primers were used for sequencing.

Sequencing was conducted on an ABI 3700 capillary sequencer (Applied Biosystems; Foster City, CA, USA). Sequence trace files generated by the ABI 3700 were assembled using Phrap program software (www.phrap.org) and evaluated for alterations using both Phred and Consed sequence analysis programs (www.phrap.org).

2.2 In silico protein analysis of exonic single nucleotide polymorphism (SNP)

Table 1. Sequences of oligonucleotides used for polymerase chain reaction (PCR) amplification in this study. *Internal sequencing primers required: 5'-untranslated region (UTR)-i (GAACACAGAGAACACAAATT).

	Forward primer sequence	Reverse primer sequence	Product size (bp)
5'-UTR*	CATCCAGGCGAGGCGCTCAC	CGCTTGAAATCCCGCATGAG	627
Exon 1	CAGGGTTGTTTGTCAGTCTCGGCGG	GAACGAGCGGCAAAGCTTAT	360
Exon 2	CTGTAGGTGTTTACGTGGCCT	CTCACTGTCTCACTATATCGCCC	399
Exon 3	GGGTCTATGAAAATCCAGGGCT	AGGTGACCAAGCTGCTGTGG	371
Exon 4	TAAATCTTGGTAGCTAGGGCC	GGCCCATTCCTGTAATCACAG	1 119
Exon 5	GGGAACATTCAAACCATAGCAGG	TGTCCAAGCTCTGCCATTGG	374
Exon 6	TCCATATCTGACCCCTGTTGG	GCAGAATCCTTAAAACCTGTGGCA	300

The altered *UBE2B* a.a. sequence was generated from exonic SNP data in-frame with the WT *UBE2B* a.a. sequence. *In silico* models of tertiary protein structures were then created using Swiss-Model [7] using these linear a.a. sequences as templates. Next, comparison of tertiary protein structures was carried out using the Local–Global Alignment method of protein structure analysis [8].

2.3 Analysis of DNA–protein binding affinity

An *in silico* model [9] was used to identify several transcription factors applicable to the region of interest, and narrowed to transcription factor SP1 given its prior description in published reports as being active in spermatogenesis [10, 11]. Recombinant SP1 was purchased from Promega (Madison, WI, USA).

Three oligonucleotides representing the most common length CGG repeat and two polymorphic variants identified in gene sequencing were purchased through the Oligonucleotide/Peptide Synthesis facility located at the University of Utah (Salt Lake City, UT, USA), and consisted of a 10-CGG repeat island without deletion (WT), a single CGG deletion (SD), and a double CGG deletion (DD).

The gel mobility shift assay was carried out using a DIG Gel Shift Kit, 2nd Generation, purchased from Roche Diagnostics (Indianapolis, IN, USA). Pairs of the single-stranded synthetic oligonucleotides were annealed and labelled with 1 mmol digoxigenin (DIG)-11-ddUTP following the protocol outlined by the kit. Each 10 μL binding assay contained 35 fmol DIG-labelled oligonucleotide, 300 ng rhSP1 in a 1 × binding buffer (20 mmol/L HEPES [pH 7.6], 1 mmol ethylenediaminetetraacetic acid (EDTA), 10 mmol $(NH_4)_2SO_4$, 1 mmol dithiothreitol (DTT), 0.2% Tween-20, and 30 mmol KCl). The competition assays contained a 10-fold molar excess of unlabelled double-stranded oligonucleotide competitor. The binding reactions were incubated at 23ºC for 20 min. The samples were loaded on a 7.5% non-denaturing polyacrylamide gel, pre-run in 0.5 × Tris-borate EDTA (TBE) buffer, pH 8.0 using a PhastSystem apparatus (GE HealthCare, Piscataway, NJ, USA). The protein-bound probe and free probe were transferred to a positively-charged nylon membrane (Roche Diagnostics, Indianapolis, IN, USA) for 30 min by contact blotting as described in the kit protocol. The membrane was washed in buffer containing 0.1 mol/L maleic acid, 0.15 mol/L NaCl (pH 7.5), and 0.3% Tween-20, then transferred to blocking solution and incubated for 30 min. The membrane was then incubated with the primary antibody solution, anti-digoxigenin-alkaline phosphatase conjugate 1:10 000 dilution in blocking solution, for an additional 30 min. The antibody solution was removed and the membrane was washed twice for 15 min in washing buffer. The membrane was labelled with the secondary conjugate, washed twice and exposed to X-ray film. The blocking solution, anti-digoxigenin-alkaline phosphatase conjugate, disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro) tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate and all reagents for labelling of oligonucleotides were supplied with the kit. All reactions were carried out at room temperature unless otherwise stated.

The gel shift was digitized and analyzed for shift distance and band intensity using NIH Image J (version 1.38; National Institute of Health, Bethesda, MD, USA).

3 Results

A list of identified novel SNPs (all heterozygous) is summarized in Table 2. Briefly, we identified three novel SNPs in the fertile group and seven novel SNPs in the infertile group. Of these, two SNPs were present in both fertile and infertile groups, making for one novel SNP unique to the fertile population, and five novel SNPs unique to the infertile population. We also identified a known SNP in the 5'-UTR of infertile men at a frequency of 5.2% (tttgtcT/Gattact, rs17167484). This is considerably lower than the previously reported frequency of 20% [4]. There was one novel exonic SNP identified in exon 4 in 5% of infertile men consisting of a G/A conversion (aaactaG/Ataatag, freq. 5%) that resulted in an isoleucine for valine substitution (both non-polar, neutral a.a.) at position 56 of the 152 a.a. UBE2B protein. *In silico* protein modeling of the original and altered UBE2B proteins revealed no differences in protein tertiary structures.

The 5'-UTR was notable for a 10-CGG variable length repeat sequence located at position 133735064–133735093 on chromosome 5, within the identified promoter area approximately 100 bp upstream of the coding region (Figure 1). A single CGG deletion (SD) was noted in both infertile and fertile men with a frequency of 20%. A dual CGG deletion (DD) was noted only in infertile men with a frequency of 4.8%. Such dual CGG deletions have not been reported previously.

DNA–protein gel shift analysis of the WT, SD, and DD sequences found that both SD and DD sequences

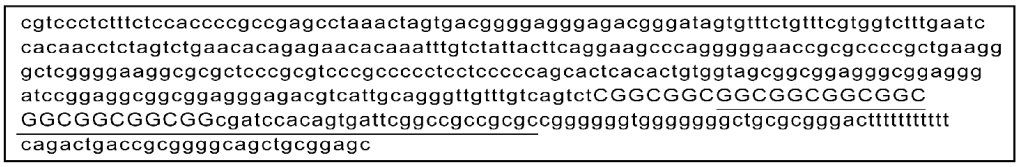

Figure 1. 5'-Untranslated region (UTR), CGG repeat sequence (capitalized), and putative promoter location (underlined). The 10 CGG variable length repeat (position 133735064–133735093 on chromosome 5) is shown in capital letters. The identified promoter region is underlined.

Table 2. Summary of novel single nucleotide polymorphisms (SNP) identified in 96 men with severe oligozoospermia or non-obstructive azoospermia and controls. Freq., frequency; UTR, untranslated region.

Location	Control group	Study population
Exon 1	No mutation	No mutation
Intron 1–2	No mutation	No mutation
Exon 2	No mutation	No mutation
Intron 2–3	tcttgtT/Attcagacc	tcttgtT/Attcagacc
	freq. 48.9 %	freq. 50%
		ttaatcA/Cgaaat
		freq. 1%
Exon 3	No mutation	No mutation
Intron 3–4	No mutation	catatctA/Gcttgaag
		freq. 1%
Exon 4	No mutation	aaactaG/Ataatag
		freq. 5%
Intron 4–5	cttttaG/Actcttt	cttttaG/Actcttt
	freq. 6.5%	freq. 19.5%
Exon 5	No mutation	No mutation
Intron 5–6	ttgtcaG/Atattct	
	freq. 1%	No mutation
Exon 6	No mutation	No mutation
3'-UTR	No mutation	gccaggC/Aagcac
		freq. 2.1%
		caaattT/Ctaagt
		freq. 2%

have twice the binding affinity to transcription factor SP1 compared to the WT sequence (Figure 2). The relative shift of these DNA–protein complexes were similar. Addition of unlabelled oligonucleotide in excess of labelled oligonucleotide showed competitive binding to SP1 and subsequent loss of signal. The negative control using bovine serum albumin (BSA) noted no binding to WT, SD, or DD oligonucleotides, and ruled out non-specific DNA–protein binding. The exact molecular size of these DNA–protein complexes was not determined as denaturation of the SP1 protein required for size evaluation would have disrupted native secondary and tertiary structures, and was not critical to the evaluation of binding affinity or relative shift.

4 Discussion

The underlying etiologies of abnormal sperm parameters remain largely unknown [12–14]. Although characterization of genes that cause male factor infertility have helped clarify the genetic component of some cases of male infertility [15–17], the majority of men with oligospermia or non-obstructive azoospermia are normal on karyotype evaluation and free from known endocrine pathologies or pathophysiologic abnormalities, highlighting the idiopathic nature of the diagnosis [18].

Of the 147 known sequence mutations within the *UBE2B* gene, four are SNPs located within the 5'-UTR, one is within the start codon of exon 1, and the remainder are within the intronic regions of the *UBE2B* gene [4]. None of these have been associated with human male infertility.

The identification of one novel intronic SNP unique to fertile males, and five novel intronic SNPs unique to infertile males, is of interest. However, given the limited numbers available to this study, designation of these SNPs as associated with male fertility or infertility is problematic, and would require further evaluation with larger study populations. Assessment of our novel exon 4 SNP revealed no significant differences in tertiary protein structures compared to WT, and it is unlikely to alter the UBE2B protein or be a factor in male infertility.

Variations in the CGG repeat sequence have been previously reported in 41 unrelated individuals of non-identified gender or ethnicity [19]. Reported were: a sextuplet deletion (2%); single deletion (15%); no deletion (78%); and a CGG insertion (5%). The frequency of

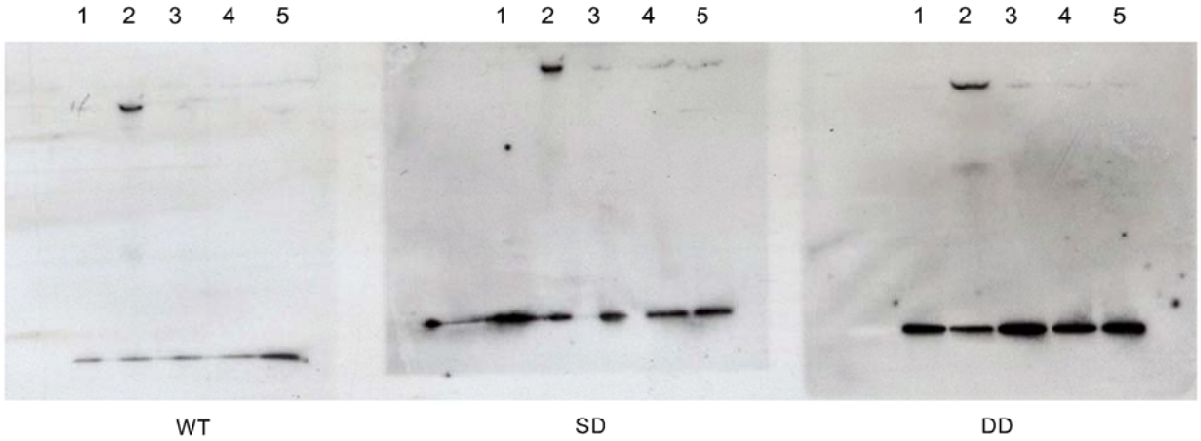

Figure 2. Gel shift analysis of CGG deletion sequences and transcription factor SP1 demonstrating 2 × binding affinity of single CGG deletion (SD) and double CGG deletion (DD) sequences as compared to wild-type (WT). This gel shift indicates equal shift of DNA–protein complexes for WT, SD and DD sequences bound to SP1, indicating complexes of equal sizes. Lane 1, labelled oligonucleotide. Lane 2, labelled oligonucleotide with SP1. Evaluation of band density shows a 2 × increased band density for SD and DD compared to WT, indicating 2 × binding affinity of SD and DD to SP1 compared to WT. Lane 3, labelled + excess unlabelled oligonucleotide with SP1. Loss of signal, as seen in this lane, with use of excess unlabelled oligonucleotide as a competitive inhibitor to labelled oligonucleotide, confirms specificity for SP1 binding. Lane 4, labelled oligonucleotide with bovine serum albumin (BSA). No signal with labelled oligonucleotide and BSA confirms no non-specific DNA–protein interaction. Lane 5, labelled + excess unlabelled oligonucleotide with BSA. Verification of no non-specific binding.

the single CGG deletion sequence was comparable to our study, however the fertility status of these 41 individuals was not reported.

As the frequency of the single CGG deletion was similar in both fertile and infertile men, at approximately 20%, the possibility exists that such findings might not be related to male infertility. However, the determination of increased binding affinity to SP1 of both single and double CGG deletions suggests that such deletions are not biologically insignificant. SP1 is a ubiquitous transcription factor present during multiple steps of human male spermatogenesis [11] and has been shown to act as both a promoter and repressor of transcription [20–22]. Thus, altered binding affinity of SP1 to these single and double CGG deletion sequences provides an enticing theory of disease cause and mechanism with regard to such deletions and subsequent male infertility.

Nevertheless, it is clear that such deletions alone are not the sole or definitive cause of human male infertility. Such deletions might represent one of several risk factors for acquisition of male infertility. In fact, notation of multiple SNPs as a predictor of disease has been shown in a number of disease states such as multiple sclerosis [23] and type 2 diabetes [24], and also have correlated with disease severity as seen with stroke risk in sickle cell patients [25]. Similarly, the presence of an as-yet undefined SNP (external to the *UBE2B* gene) might be required for these CGG deletions to affect male fertility. The identification of such SNPs would require multigene analysis and is best approached with current gene-chip technology.

The significance of this study should be interpreted within the context of the study population, that is, these findings might only be applicable to males of northern European ancestry. Furthermore, in addition to more extensive SNP evaluation, additional study is indicated to confirm altered *UBE2B* mRNA and protein expression concomitant with 5'-UTR CGG deletions in populations of men with severe oligospermia or non-obstructive azoospermia.

Although defects in exons 1–6 of the *UBE2B* gene have not been linked to human oligospermia nor non-obstructive azoospermia, CGG deletions noted in the 5'-UTR region could affect normal transcription and expression of the *UBE2B* gene product through altered binding of transcription factor SP1 and an as-yet unidentified co-factor. This in turn might lead to abnormal spermatogenesis and subsequent male infertility. Further in-

vestigations are needed to clarify whether such CGG deletions do ultimately affect *UBE2B* expression through appropriate translation studies.

References

1 Roest HP, van Klaveren J, de Wit J, van Gurp CG, Koken MH, Vermey M, *et al*. Inactivation of the *HR6B* ubiquitin-conjugating DNA repair enzyme in mice causes male sterility associated with chromatin modification. Cell 1996; 86: 799–810.

2 Baarends WM, Wassenaar E, Hoogerbrugge JW, van Cappellen G, Roest HP, Vreeburg J, *et al*. Loss of *HR6B* ubiquitin-conjugating activity results in damaged synaptonemal complex structure and increased crossing-over frequency during the male meiotic prophase. Mol Cell Biol 2003; 23: 1151–62.

3 Roest HP, Baarends WM, de Wit J, van Klaveren JW, Wassenaar E, Hoogerbrugge JW, *et al*. The ubiquitin-conjugating DNA repair enzyme HR6A is a maternal factor essential for early embryonic development in mice. Mol Cell Biol 2004; 24: 5485–95.

4 National Center for Biotechnology Information. Single Nucleotide Polymorphism database available at URL: http://www.ncbi.nlm.nih.gov/SNP/

5 Reese MG. Neural Network Promoter Prediction server available at URL: http://www.fruitfly.org/seq_tools/promoter.html

6 Reese MG. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Comput Chem 2001; 26: 51–6.

7 Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006; 22: 195–201.

8 Zemla A. LGA: a method for finding 3D similarities in protein structures. Nucleic Acids Res 2003; 31: 3370–4.

9 Kel A. Match and TRANSFAC 6.0. Available at URL: http://www.gene-regulation.com

10 Assumpcao JG, Ferraz LF, Benedetti CE, Maciel-Guerra AT, Guerra G Jr, Marques-de-Faria AP, *et al*. A naturally occurring deletion in the SRY promoter region affecting the Sp1 binding site is associated with sex reversal. J Endocrinol Invest 2005; 28: 651–6.

11 Thomas K, Wu J, Sung DY, Thompson W, Powell M, McCarrey J, *et al*. SP1 transcription factors in male germ cell development and differentiation. Mol Cell Endocrinol 2007; 270: 1–7.

12 Ferlin A, Raicu F, Gatta V, Zuccarello D, Palka G, Foresta C. Male infertility: role of genetic background. Reprod Biomed Online 2007; 14: 734–45.

13 Sadeghi-Nejad H, Farrokhi F. Genetics of azoospermia: current knowledge, clinical implications, and future directions. Part I. Urol J 2006; 3: 193–203.

14 Krausz C, Giachini C. Genetic risk factors in male infertility. Arch Androl 2007; 53: 125–33.

15 Zhang J, Qiu SD, Li SB, Zhou DX, Tian H, Huo YW, *et al*. Novel mutations in ubiquitin-specific protease 26 gene might cause spermatogenesis impairment and male infertility. Asian J Androl 2007; 9: 809–14.

16 Tsujimura A, Fujita K, Komori K, Tanjapatkul P, Miyagawa Y, Takada S, *et al*. Associations of homologous RNA-binding motif gene on the X chromosome (RBMX) and its like sequence on chromosome 9 (RBMXL9) with non-obstructive azoospermia. Asian J Androl 2006; 8: 213–8.

17 Gu AH, Liang J, Lu NX, Wu B, Xia YK, Lu CC, *et al*. Association of *XRCC1* gene polymorphisms with idiopathic azoospermia in a Chinese population. Asian J Androl 2007; 9: 781–6.

18 Brugh VM 3rd, Matschke HM, Lipshultz LI. Male factor infertility. Endocrinol Metab Clin North Am 2003; 32: 689–707.

19 Lench NJ, Thompson J, Markham AF, Robinson PA. (CGG) trinucleotide repeat polymorphism in the 5' region of the *HHR6B* gene: the human homolog of the yeast DNA repair gene *RAD6*. Hum Genet 1995; 96: 369–70.

20 Li R, Hodny Z, Luciakova K, Barath P, Nelson BD. Sp1 activates and inhibits transcription from separate elements in the proximal promoter of the human adenine nucleotide translocase 2 (*ANT2*) gene. J Biol Chem 1996; 271: 18925–30.

21 Zaid A, Hodny Z, Li R, Nelson BD. Sp1 acts as a repressor of the human adenine nucleotide translocase-2 (*ANT2*) promoter. Eur J Biochem 2001; 268: 5497–503.

22 Zaid A, Li R, Luciakova K, Barath P, Nery S, Nelson BD. On the role of the general transcription factor Sp1 in the activation and repression of diverse mammalian oxidative phosphorylation genes. J Bioenerg Biomembr 1999; 31: 129–35.

23 International Multiple Sclerosis Genetics Consortium, Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, Jager PL, *et al*. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007; 357: 851–62.

24 Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, *et al*. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007; 316: 1341–5.

25 Hoppe C, Klitz W, D'Harlingue K, Cheng S, Grow M, Steiner L, *et al*. Confirmation of an association between the TNF(-308) promoter polymorphism and stroke risk in children with sickle cell anemia. Stroke 2007; 38: 2241–6.

Edited by Dr Gordon Baker