Mitochondrial ND Genes: Relevance of Codon Usage to Semen Quality in Men

A thesis
submitted in partial fulfillment
of the requirements for the Degree
of
Master of Science in Cellular and Molecular Biology
in the
University of Canterbury
by
SADIA JIHAN KHAN

University of Canterbury
2006

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The last one and half year of my academic life has been a true learning period. This one page is not adequate to thank all those people who have been supporting me during this time.

Firstly, I would like to thank my supervisor Associate Professor Frank Sin, who gave me the opportunity to do this research and his continuous assistance, guidance, support, suggestions and encouragement have been invaluable to me. I honestly appreciate the time and effort he has given me to conduct my research. I would also like to thank my associate supervisors, Dr. Steven Gieseg and Dr. Drusilla Mason for supporting my research project.

Secondly, I would like to give a special thank to Associate Professor H. Khris Mahanty for steering me into the right path. Without his recommendation, encouragement and direction I wouldn't even think of studying in NZ. I'm indebted for his right supervision at the right time on choosing my higher studies destination.

I would like to express my gratitude to my lab mates and friends who had made my university life more enjoyable and pleasant by sharing their invaluable time. I also thank Angie and Lucy for unraveling the mystery of the world of my silly questions regarding PCR analysis. I specially thank Claudia, Jacquie and Melani for their priceless help and encouragement and also for proofreading of my thesis.

Finally, and most importantly, I would like to acknowledge my beloved husband Kamal's role in my academic endeavor. I sincerely thank him for his immense mental support and invaluable love. I remember those moments when I was overwhelmed by glitches and difficulties; this is the person who patiently boosted me with his amorousness. I particularly thank my parents, who kept faith on me and encouraged me all the time to accomplish my further higher

studies in NZ. In my mind, without their guidance since my childhood I could not have been able to carry out my academic study. I thank my father-in-law, mother-in-law, my brother, my sister and my family for their unconditional love, inspiration and encouragement throughout the last few years of my academic work.

The combined efforts from all of the people mentioned above helped me to achieve my goals. THANKS!

ABSTRACT

Studies have discovered higher frequencies of single nucleotide polymorphisms (SNPs) in different mitochondrial genes are associated with subnormozoospermia. However, the frequencies of SNPs in *ND1* and *ND2* are not unknown. The present research was aimed to determine the frequencies of SNPs in *ND1* and *ND2* genes of the mitochondrial genome in fertile and subfertile men and whether changes in codon usage was associated with fertility phenotypes. Total genomic DNA from 157 semen samples was extracted using the proteinase K/SDS digestion procedure, followed by phenol/chloroform purification and ethanol precipitation. *ND1* and *ND2* genes were amplified respectively from 80 and 92 DNA samples from different fertility groups. Each PCR product was sequenced to identify mutations. Codon change resulting from a nucleotide substitution was determined by comparison with a reference mtDNA sequence obtained from the NCBI database. The frequency of codon usage in the reference mtDNA was determined by the computer program MEGA version 2.1. Eleven synonymous nucleotide substitutions and two non-synonymous substitutions were found in this study. Four SNPs were previously characterized; all SNPs were homoplasmic. None of the SNPs were likely to affect the function of the proteins on the basis of the hydrophobicity plots or secondary structure predictions. Sixty two percent of synonymous mutations were found to change from a high to a low relative codon usage values; 37% of synonymous mutations changed from a low to a high relative usage value. Chi-square (χ^2) test (χ^2= 0.067 with 1 d.f.) showed that there was no significant difference at the 5% level between these changes. Thus, change in codon usage was not related to semen quality in men. Further, there were no statistically significant differences in the observed frequencies of SNPs of fertile and subfertile men. However, the sample size was small and this study was only focused on a single NZ Caucasian population. Further study including larger and more diverse population samples may provide further

insight into the functional importance of codon usage and its relevance to fertility status in men.

CHAPTER ONE - INTRODUCTION

INTRODUCTION

1.1 THE MITOCHONDRION

The mitochondrion is the "power house of a cell" and was first described by Altman in 1890 (cited in Cummins, 2001). This thread-like organelle is responsible for cellular metabolism and programmed cell death (Grossman, 1995; Cummins, 2001). Mitochondria contain their own DNA. Mutations in the mitochondrial genome play a major role in some human genetic diseases (Wallace *et al*., 1999; DiMauro and Schon, 2001; Wallace, 2005).

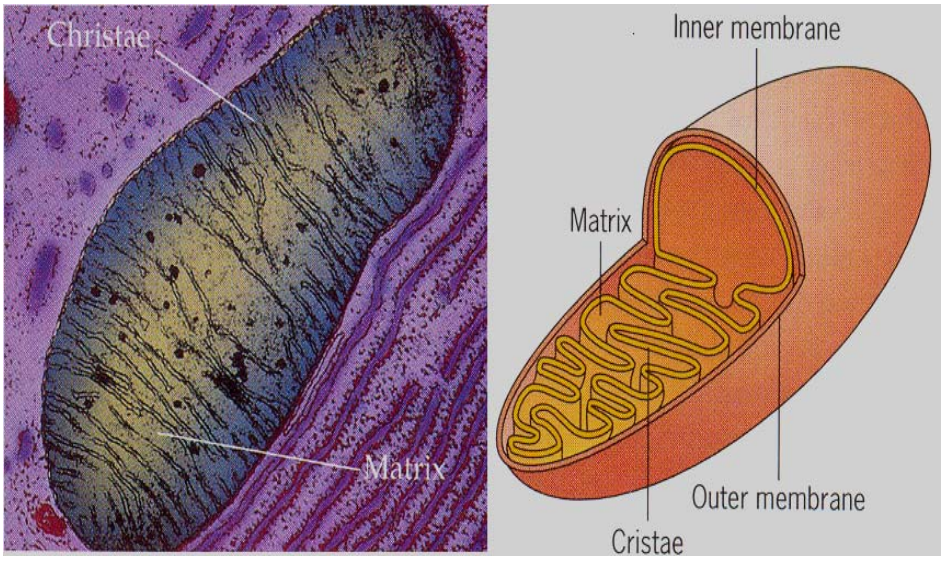


Fig 1.1 The human mitochondrion. A mitochondrion under electron microscopy (left) and schematic diagram of a surface cut mitochondrion with its internal structure (right).Taken from Snustad and Simmons (2000).

1.1.1. STRUCTURE OF MITOCHONDRIA

Mitochondria are found in the cytoplasm of eukaryotic cells and are therefore completely separated from the nucleus and nuclear DNA. With the exception of red blood cells, all eukaryotic cells contain many mitochondria. These semiautonomous organelles vary greatly in size and shape (Darley-Usmar *et al.,* 1994). The structures of mitochondria are highly conserved in different organisms. Each mitochondrion has an outer and an inner membrane (Figure 1.1). The outer membrane contains the original host membrane and forms a permeability barrier to cations (Cummins *et al*., 1998). The inner membrane consists of finger-like projections called cristae. Cristae are impermeable (Cummins *et al.,* 1998) and contain essential enzymes for metabolic functions such as ATP synthesis and cell respiration. ATP is produced by oxidative phosphorylation (OXPHOS).

1.1.2. FUNCTION OF MITOCHONDRIA

The main function of mitochondria is oxidative phosphorylation. This process is carried out by an electron transport chain, which is mediated by different protein complexes. The five complexes are nicotinamide adenine dinucleotide coenzyme Q reductase (NADH-CoQ reductase) (complex I), succinate CoQ reductase (complex II), ubiquinol cytochrome b reductase (complex III), cytochrome c oxidase (complex IV) and ATP synthase (complex V). Figure 1.2 is a schematic diagram of the OXPHOS pathway in a mitochondrion. Complex I uses the energy in NADH to pump protons into the inter-membrane space of the mitochondrion. Complex I receives electrons from succinate. These electrons are delivered to complex III by coenzyme Q. Complex III

delivers electrons to complex IV and reduces oxygen atoms inside the matrix to H_2O. A proton motive force is then generated in the inter-membrane space by coupling electron transport to protein translocation at

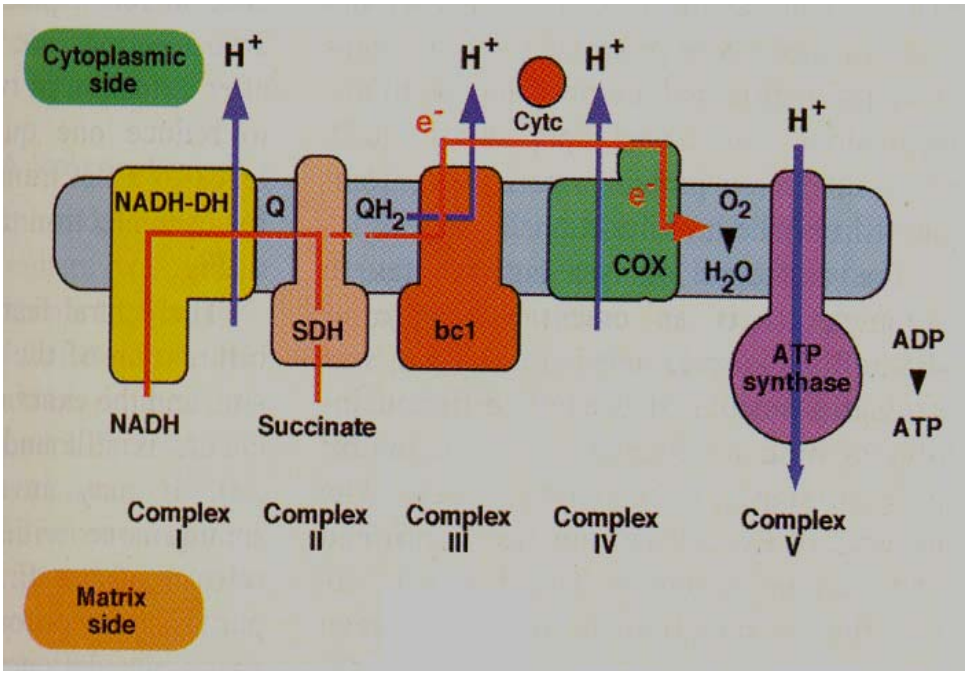


Fig. 1.2. **Schematic representation of oxidative phosphorylation.** The enzymes of the mitochondrial inner membrane involved in OXPHOS and protons are translocated across the membrane to generate ATP. Diagram was taken from Saraste (1999).

complex I, II, and IV (Moore *et al*., 2000; Wallace, 2001; Wallace, 2005; Schǎfer *et al*., 2006). The enzyme ATP synthase generates ATP when protons are transported down their concentration gradient into the matrix. Complex I, III and IV are partially encoded by mtDNA and crucial for generating the proton motive force (Moore *et al*., 2000). OXPHOS generates reactive oxygen species (ROS), an endogenous toxin as a by-product (Reha and Robinson, 2001).

ROS are generated in the form of $OH^{\bullet}$, O_2^{-}, and H_2O_2 by reduction of oxygen. ROS are highly mutagenic and/or carcinogenic (Cummins, 2001). In the human body each mitochondrion produces 1×10^7 ROS per day. Since mitochondria occupy up to 25% of the total cytosol, the cell is very susceptible to oxidative damage (Cummins, 2001). ROS can damage the mitochondrial and cellular proteins, lipids, nucleic acids, mtDNA, and interrupt energy production (Attardi *et al.,* 1988; Wallace *et al.,* 1999; Wallace 2005).

Excessive production of ROS may cause programmed cell death (apoptosis), which is another function of mitochondria. Apoptotic cascade is generated by the activation of mitochondrial permeability transition pore (mtPTP) (Wallace, 2001). Cytochrome c is released (Green and Reed, 1998). The release of cytochrome c and the latent form of specialized proteases, caspases, promote programmed cell death by activating the cytosolic protein degradation pathway and destroying the cytoplasm (Wallace *et al*., 1999). This interrelated system involved in apoptosis is assembled from roughly 1500 genes distributed in the nuclear and mitochondrial genomes (Wallace, 2005).

1.1.3. HUMAN MITOCHONDRIAL GENOME

The human mitochondrial genome (mtDNA) is 16,569 base pairs (bp) of nucleotides in length (Anderson *et al*., 1981). The genome is a double-stranded, covalently closed circular DNA molecule with little redundancy and rare forms of telomere like terminations (Cummins, 2001). The mtDNA is located within the mitochondrial matrix. The mitochondrial genome encodes 37 genes including 2 ribosomal RNAs (rRNAs), 22 transfer RNAs (tRNAs) and 12 polypeptides (Wallace *et al*., 1999; Shanske *et al*., 2001; Wallace, 2005). The 12 polypeptides include seven

subunits of NADH dehydrogenase (ND1, ND2, ND3, ND4, ND4L, ND5 and ND6) that make up complex I, a cytochrome b in complex III, three complexes of cytochrome c oxidase in complex I, II, III and complex IV, and two ATP subunits, ATPase6 and ATPase8, in complex V (Anderson *et al*., 1981) (Figure. 1.3). MtDNA has two strands, a guanine rich (heavy, H) strand and a cytosine rich (light, L) strand (Wallace, 1992). The heavy strand contains 12 of the 13 polypeptide encoding genes, 14 of the 22 tRNA encoding genes and both rRNA encoding genes (Wallace, 1992). Introns are absent in mtDNA and all of the coding sequence are contiguous (Anderson *et al*., 1981; Wallace, 1992).

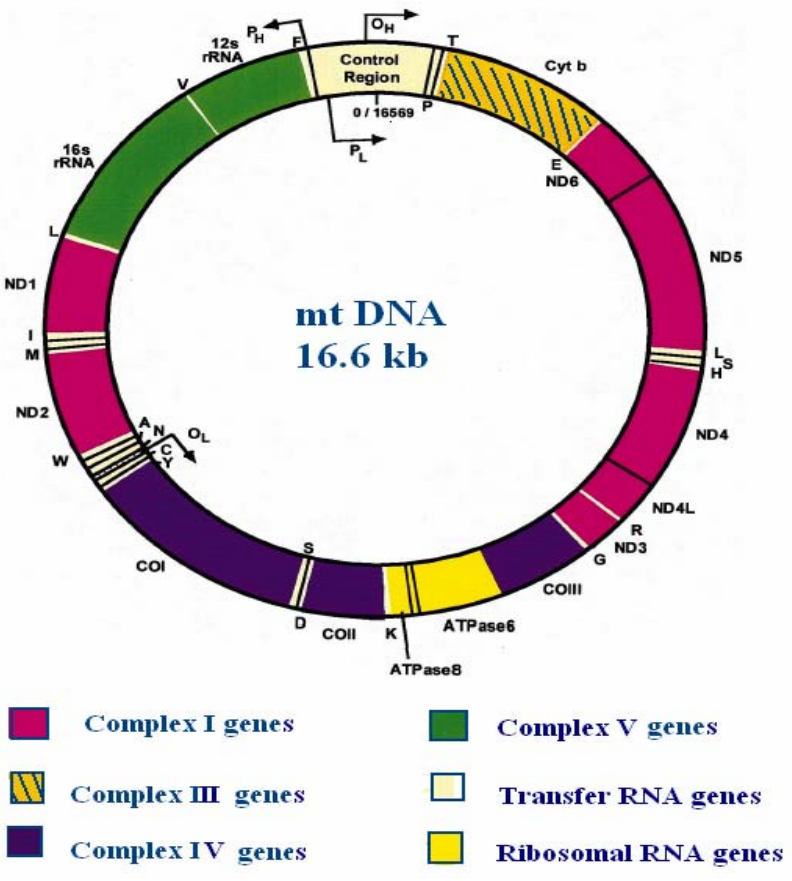


Fig.1.3. **Structure of a human mitochondrial genome**. Modified from Wallace *et al*. (1999).

The non-coding segment of mtDNA is the displacement loop (D loop), a region of 1121 bp that contains the origin of replication of the H strand (O_H), and the promoters for L and H stands (P_L and P_H) (Wallace, 1999). The mtDNA is replicated from two origins.

DNA replication is initiated at O_H using an RNA primer generated from the L strand transcript. H strand synthesis proceeds two-third of the way around the mtDNA, displacing the parental H strand until it reaches the L stand origin (O_L), situated in the cluster of five tRNA genes (Holt, 1994). The L strand synthesis starts at O_L and further proceeds back along the H strand (Wallace, 1992; Holt, 1994). MtDNA replication is bidirectional. Transcription of mtDNA is initiated from two promoters in the D loop, P_L and P_H. Most of genes are transcribed from P_H, which reads anti-clockwise and from P_L transcription which starts clockwise (Wallace, 1992).

1.1.4. MITOCHONDRIAL GENETICS

The location of mtDNA and the high copy number contribute a number of unique features to mitochondria (Wallace, 1999). MtDNA differs from nuclear DNA as it is maternally inherited; it has a higher mutation rate than nuclear DNA; it demonstrates replicative segregation and threshold expression, and is more susceptible to mutation.

Maternal inheritance is a unique feature of the mitochondrial genome. The mammalian egg contains about 100,000 molecules of mtDNA, while sperm cells contain 100-150 mtDNA (Chen *et al*., 1995). Sperm mitochondria enter into the egg and are discarded by the egg during fertilization (Cummins *et al*., 1997; 1999; Cummins, 2001; DiMauro *et*

al., 2001), although low level of paternally transmitted mtDNA representing ~10^4 of total mtDNA have been detected in mice (Wallace, 1994).

The high mutation rates observed in mtDNA is a result of the lack of efficient DNA repair mechanisms (Bogenhagen, 1999). Further, mtDNA is not bound to histone (Cummins, 2001). Thus, it is not protected from oxidative damage due to the highly toxic ROS, which are by-products of OXPHOS. MtDNA damage can be expected to have an impact on cell function as the probability of damaging coding sequences is higher. If damage accumulates, cellular function will be impaired because mitochondria are deficient in a nucleotide-excision repair (NER) and mismatch repair (MMR) system. This limitation of a repair system would partly explain the higher rate of mutations in mitochondria (Bogenhagen, 1999; Copeland *et al*., 2002). The mutation rate in mtDNA is 10 times higher than in nuclear DNA (Cummins, 2001).

The third distinctive feature of mitochondria is the replicative segregation of mtDNA during meiosis and mitosis (Wallace, 1999; Shanske *et al.,* 2001; Wallace, 2005). Each cell contains hundreds of mitochondria and each mitochondrion contains 10-100 copies of mtDNA. Normally all mtDNA in a cell are identical (homoplasmy) (DiMauro *et al.,* 2003). During cell division the mitochondria and their genomes are randomly distributed to the daughter cells by a process called replicative segregation (Wallace, 1999). The higher rate of mutation in mtDNA creates a mixture of mutant and wild type DNA molecules, which is known as heteroplasmy (DiMauro *et al*., 2003). Not all but most of the pathogenic mutations are heteroplasmic (Shanske *et al.,* 2001; DiMauro *et al*., 2003).

A minimal number of mutant mtDNAs has to be present within a tissue before oxidative dysfunction occurs. This is called the threshold effect (DiMauro, 2003). This threshold varies for different types of mtDNA mutations. When the threshold level is exceeded, changes occur in the phenotype (Shanske *et al*., 2001). Organs such as, heart, brain etc. with a high ATP demand will have the lowest regenerative capacity (Wallace *et al.*, 1999; Wallace, 2005).

The mitochondrial mutation rate is 10-20 times higher than nuclear DNA (Wallace, 1992). Since the mitochondrial genome does not contain introns and non-coding sequences, any mutation occurring in mtDNA would affect the coding regions. Thus, a variety of neutral and mildly deleterious mutations are found in human populations.

1.2. MITOCHONDRIAL DNA MUTATIONS AND DISEASE

Disease caused by mitochondrial dysfunction can be due to mutations in mtDNA. The first oxidative disorders were described by Luft *et al*. in 1958 (cited in Haushmand, 2003). The term mitochondrial disorder refers mainly to the disorders with abnormal morphological aspects of mitochondria in muscle (DiMauro *et al*., 2000). In 1988, the first pathogenic mutations associated with myopathy were discovered by Holt, who found large scale detections in mitochondria (cited in Haushmand, 2003). In the same year, Wallace *et al.*, reported a point mutation in the *ND6* gene associated with LHON (Leber's hereditary optic neuropathy). Houshmand (2003) stated that, in 1990, two new mutations were discovered, A8344G in $tRNA^{lys}$ (Shoffner *et al.*, 1990) and A3243G in $tRNA^{leu}$ (UUR) gene in MELAS syndrome (Goto *et al*., 1990). To date

more than 100 pathogenic mutations have been found to be associated with human diseases (DiMauro *et al.,* 2000). MtDNA mutations cause defects in brain, heart, muscle vision and hearing, and are also involved in other abnormal cellular states such as cancer (Jakupciak *et al*., 2005), diabetes mellitus (Thomas *et al*., 1996) and infertility (Ruiz-Pesini *et al*., 2000; Holyoake *et al*., 2001; Thangaraj *et al.,* 2003).

Three different types of mtDNA mutations have so far been reported to be involved with diseases. They are single nucleotide deletions in tRNA and rRNA genes and in protein-coding gene (Shanske *et al.,* 2001; DiMauro *et al*., 2001; 2003). These mtDNA mutations will affect protein synthesis in the mitochondria. Although in most cases, single nucleotide deletions in mitochondria are sporadic events (DiMauro *et al*., 2001; Shanske *et al*., 2001).

The most commonly observed point mutation is A8344G, located in MERRF disease in $tRNA^{lys}$ gene. These mutations are heteroplasmic and vary widely between different individuals and different tissues of the same individual (Wallace, 1992). Many pathogenic mutations have been found in the $tRNA^{leu}$ gene. High levels of A3243G mutations have been found in muscles and ragged red fibers (Wallace *et al.,* 1999). It was also associated with MELAS and KSS. Two other mutations (T3271C and T3291C) are reported to be associated with MELAS (Goto *et al*., 1994). T7512C, T9957C and T8357C mutations are also found with MERRF/MELAS. Some rRNA mutations are also associated with deafness, complex neurology and diabetes mellitus type II (DiMauro *et al.,* 2003; Haushmand, 2003). These mutants are A1155G, G1606A and T3200C.

The first mitochondrial point mutation discovered in a protein coding gene was associated with a maternally inherited disease called LHON (Wallace, 1988). Four main mutations in the complex I gene (*ND* gene); G11778A in *ND4*, G3460A in *ND1*, T14484C and G14459A in *ND6* are associated with LHON (DiMauro *et al*., 2003). Recently T4216C in the *ND1* gene was reported to be associated with LHON (www.mitomap.org, as for 18th may, 2006). LHON patients normally lose their eye sight rapidly during adolescent (Leonard *et al*., 2000). Houshmand (2003) reported that at least fifteen other LHON mutations were proposed by Howell *et al.* (1995) and Brown *et al.* (1995). G11778A, G3460A and T14484C cause reduction in NADH respiratory substrates and also partially reduce complex I activity (Wallace, 1999). Another T8993C heteroplasmic mutation in the ATPase6 gene in mtDNA was found to be associated with maternally inherited NARP syndrome (Shanske *et al.*, 2001; Wallace, 2005).

1.2.1. THE ASSOCIATION OF MITOCHONDRIAL DNA HAPLOGROUPS WITH DISEASE

MtDNA haplogroups analysis is one of the tools used to study human pathologies (Torroni *et al*., 1996; Ruiz-Pesini *et al*., 2000). MtDNA haplogroups were initially based on information provided by RFLPs and the hypervariable segments of the control region (Finnila *et al*., 2001) but are now commonly defined by whole mtDNA sequence analyses.

Recent studies have shown that mtDNA haplogroups play an important role in disease expression. Haplogroup J was associated with LHON (Torroni *et al*., 1996; Wallace *et al.,* 1999; Wallace 2005). Wallace (1999) showed that 47 of LHON patients of European descent contain three primary mutations: MTND4*LHON11778A, MTND1*LHON3460

and MT*ND6LHON14484C. The association between MTND6*14484C and haplogroup J was striking. Haplogroup J contains only 9% of European population and 37% LHON patients showed MTND4*LHON11778A mutation, while 80% were with MT*ND6LHON14484C. Association of haplogroup J with the LHON mutation at nt 9438 in COX III was also observed in Caucasian populations. The J haplogroup with mild MTND1*4216C variant might be also associated with LHON (Wallace, 1999; 2005).

1.3. HUMAN INFERTILITY

1.3.1 GENERAL ASPECTS OF HUMAN INFERTILITY

Reproduction is a natural process. Infertility is a reproductive health problem that affects many couples in the human population. About 13-18% of couples suffer from it, and one half of all cases can be traced to either partner (Seshagiri, 2001). Infertility is the inability to conceive after a period of 12 months of intercourse without using any contraception or inability to carry pregnancies to live birth (Cummins *et al.,* 1994).

1.3.2. HUMAN MALE INFERTILITY

Male infertility is much more difficult to describe than female infertility (Cummins, 1994). Recent evidence suggests that male infertility has a genetic cause (Affara, 2001; Holyoake *et al.,* 2001; Spiropoulos *et al*., 2002). Around 20-25% of infertility cases are due to male factors (De Krestser *et al*., 1999).

1.3.3. SPERMATOGENESIS

Spermatogenesis is a complex process. This biological process occurs in the seminiferous tubules of the testes (Purves *et al*., 2004). Spermatogenesis begins in the male germ cell line, which contains stem cells that produce spermatogonia. Repeated mitotic divisions produce diploid spermatocytes from spermatogonia (Holstein *et al*., 2003). Each

primary spermatocyte then divides by meiosis to produce secondary spermatocytes, which further divide to produce four haploid cells called spermatids (Holstein *et al*., 2003). Spermatids pass through the process of spermiogenesis and become free living spermatozoa. Figure 1.4 is a diagram showing the human spermatogenesis process.

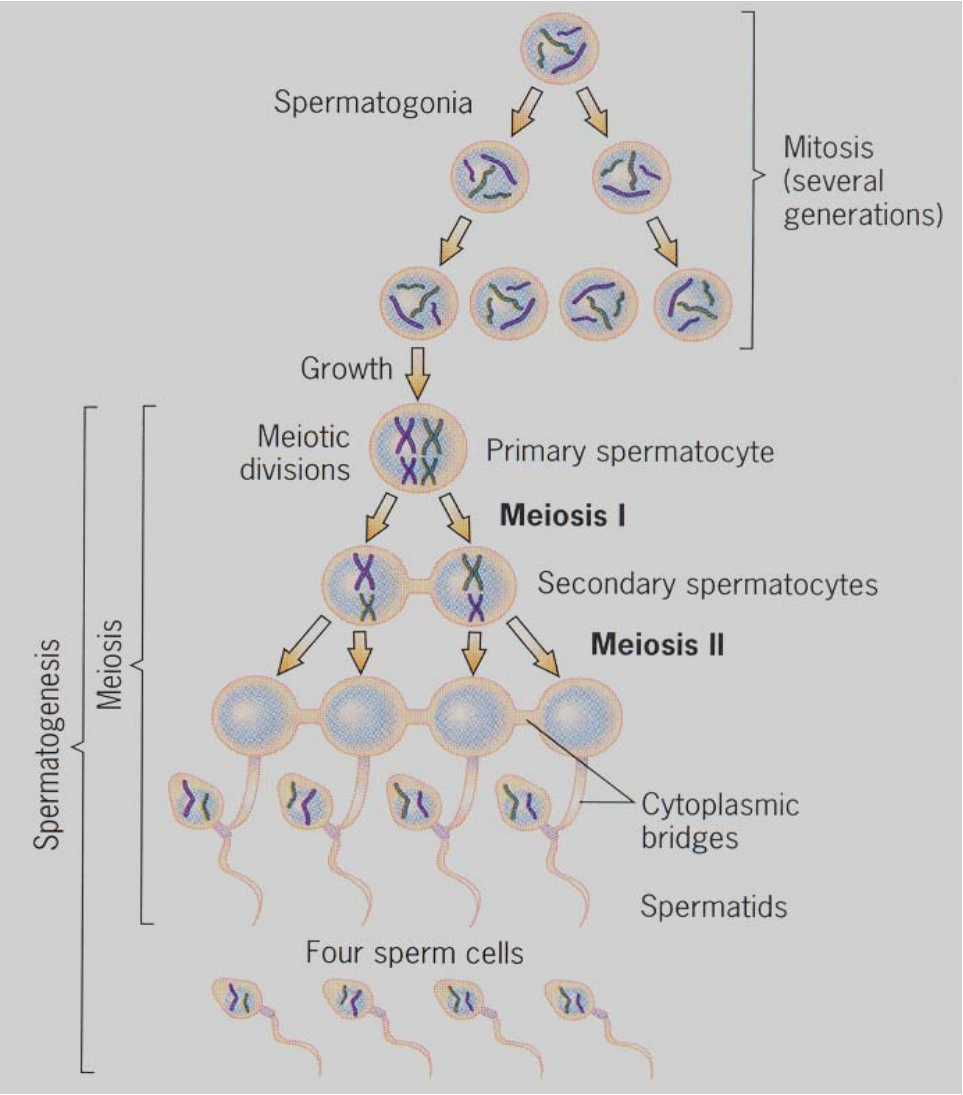


Fig 1.4. Human spermatogenesis process. Diagram was taken from Snustad and Simmons (2000).

1.3.4. HUMAN SPERM CELL STRUCTURE

The human sperm cell consists of three parts: head, midpiece and tail (Fig 1.5). The head contains a nucleus and has a cap like structure called the acrosome. The acrosome enables the sperm cell to penetrate the egg. Motility is the purpose of the sperm tail. The midpiece is tightly packed with mitochondria. A typical sperm cell contains 20-28 mitochondria (St Jones *et al*., 2001). These mitochondria generate ATP to provide energy for sperm to swim (Kao *et al*., 1998). Defects in the internal structure and organization of mitochondria in the midpiece are linked to poor semen quality in men (Kao *et al*., 1998) and low sperm motility (Thangaraj *et al.,* 2003).

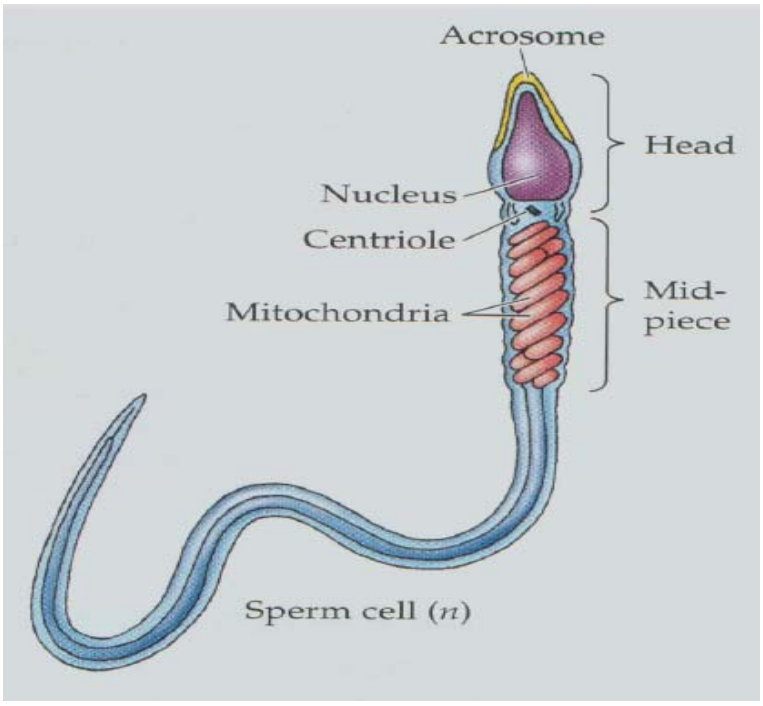


Fig. 1.5. A typical structure of a sperm cell. Taken from Purves *et al*. (2004).

1.3.5. CAUSES OF MALE INFERTILITY

Male infertility may be caused by obstructive lesions, spermatogenic failure, sexual dysfunction, endocrine disorder and idiopathic sperm dysfunction (Cummins *et al*., 1994). Approximately 7-15% of infertile men suffer from obstructive lesions and have ejaculatory problem. Spermatogenic failure disrupts spermatogenesis and its cause is still unknown (Namiki, 2000). Mutations in germ line genes may be a cause of defective spermatogenesis. Approximately 30% of infertile men tend to show defect in spermatogenesis (Cummins *et al*., 1998). Idiopathic sperm dysfunction is related to environmental factors (smoking, drugs, heavy metals, etc.). Many cases of idiopathic infertility have a genetic or molecular basis (Seshagiri, 2001).

1.3.6. GENETIC BASIS OF MALE INFERTILITY

Genetic disorders involved in male infertility can be due to chromosomal abnormalities, autosomal gene defect or Y chromosome mutation.

The 47, XXY karyotype variant of Klinefelter's syndrome, is the most common chromosomal disorder associated with male infertility (Singh *et al*., 2006). Eleven percent of azoospermic and 0.7% of oligozoospermic men tend to have this condition (Bhasin *et al*., 2000). Autosomal abnormalities in men may cause congenital bilateral absence of vas defense (CBVAD), which is associated with the cystic fibrosis transmembrane conducted regulator gene. This is responsible for 1-2% of infertility in men (De Krestser *et al*., 1999). Four percent of infertile men

tend to show sex-chromosomal abnormalities, while 1% shows autosomal abnormalities (Singh *et al.*, 2006).

The Y chromosome is mainly responsible for male infertility. Y-linked mutations have most effects on spermatogenesis (Hargreave *et al*., 2000; Singh *et al*., 2006). Y-chromosome microdeletions were detected in men with suboptimal semen parameters (7.4%) (Erasmuson *et al*., 2003). Microdeletions in the AZF factor, located in the long arm of Y chromosome are associated with male infertility (Vogt *et al*., 1996). The AZF factor consists of four regions: AZFa, AZFb, AZFc and AZFd (Brinton-Jones *et al*., 2000) (Fig 1.6). Three genes (DAZ, RBM and DFFRY) are also located in these regions.

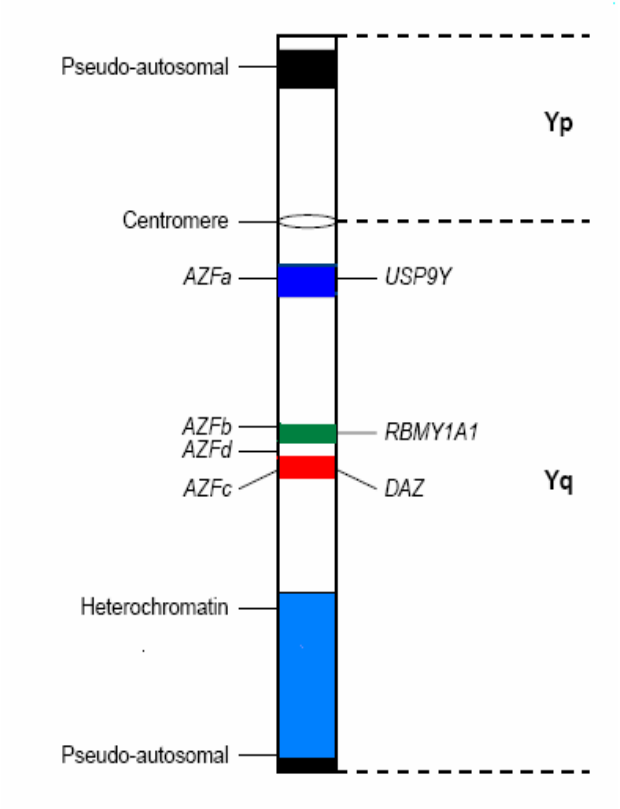


Fig 1.6. Schematic representation of the human Y chromosome containing AZF region. Modified from Brinton-Jones *et al*. (2000).

The DAZ (Deleted in azoospermia) gene located in the AZFc region of the Y chromosome and is found to be deleted in 10-15% oligozoospermic and azoospermic men (Moore *et al.,* 2000). The RBM (RNA binding motif) gene is in the AZFb region (Moore *et al.,* 2000). Deletions of this region are found in oligozoospermic and azoospermic men. Similarly, DFFRY (*Drosophila* fat facet related Y) gene located in the AZFa region is associated with male infertility. Another azoospermic factor AZFd is proposed by Kent-First *et al*. (1999), which is associated with mild oligozoospermia or normal sperm count with abnormal sperm morphology (Brinton-Jones *et al*., 2000).

X linked disorder, such as Kaltman's syndrome is also suggested to be associated with male infertility (mutations in a gene on $X_P22.3$) (Seshagiri, 2001). Infertile men tend to show higher incidence of translocation (Affara, 2001). But this contribution to infertile men has not properly understood (Huyanh *et al*., 2002). However, a CAG repeat at the amino terminal domain of the androgen receptor is believed to cause a defective spermatogenic disorder, but the role of this CAG repeats is controversial (De Krestser *et al*., 1999; Erasmuson *et al*., 2003). Men with poor semen quality showed similar CAG repeat number to men with normal semen quality (21.46 ± 0.30 vs. 20.99 ± 0.28, $p = 0.126$) (Erasmuson *et al*., 2003). Therefore, higher CAG repeats are not always associated with infertility in men.

1.3.7. MALE INFERTILITY AND MITOCHONDRIAL DYSFUNCTION

Sperm motility is one of the major factors of male infertility (Ruiz-Pesini *et al*., 2000). Sperm movements require large amount of ATP. A defect in mitochondrial respiratory function causes a decline in motility and infertility (Kao *et al*., 1998; Ruiz-Pesini *et al*., 2000). Mutations in the mtDNA of a sperm cell can result in either function-less or malfunctioning proteins that affect sperm motility (Thangaraj *et al*., 2003).

Ruiz-Pesini *et al*. (2000) showed that asthenozoospermia, not oligozoospermia, was associated with mtDNA haplogroups and spermatozoa dysfunction. Mitochondrial haplogroup H was more abundant in non-asthenozoospermia and haplogroup T was more abundant in asthenozoospermia. Three nucleotide positions 7025 (*COI*), 10394 (*ND3*) and 12308 ($tRNA^{Leu}$) were identified in haplogroup H and significantly more frequent in samples with high sperm motility. Five nucleotide positions 4216 (*ND1*), 4914 (*ND2*), 7025 (*COI*), 13366 (*ND5*) and 13704 (*ND5*) are identified in T haplogroup. The presence of these SNPs is frequently observed in semen samples of poor quality. But Thangaraj *et al*. (2003) established that mitochondrial mutations in sperm can cause low sperm motility that is not related to infertility. They obtained 8, 13, 5 and 10 SNPs in *COI*, *COII*, *ATPase8* and *ATPase6* genes, respectively, in the sperm DNA of an oligoasthenozoospermic man. These mutations was associated with low motile sperm but was fertile.

Spiropoulos *et al*., (2002), studied the semen sample of a man who was inherited with an A3243G mutation in mtDNA. Using a percol

fractionation method, these authors separated sperm into five groups. They observed that decreased sperm motility was correlated with an increase in mutant mtDNA from 42 to 64%.

Semen samples with a high number of multiple mtDNA deletions appear to be associated with male infertility (St-John *et al*., 2001). Kao *et al*. (1995) found 4977 bp deletions in mtDNA in normal semen parameters and this showed a negative correlation between base pair deletion and sperm motility. But Cummins *et al*. (1998) repeated this experiment and found this deletion in men with normal sperm parameters. He suggested that spermatozoa might contain huge number of mutations that can cause sperm dysfunction. He also proposed that defective respiration in sperm mitochondria will produce ROS and cause cellular damage. Most mutations in mtDNA that are associated with asthenozoospermia are caused by ROS (St-John *et al*., 2001). Male infertility is associated with oxidative stress and oxidative damage (Cummins *et al.*, 1994; Smith *et al*., 2006). Oxidative stress plays a major role in the etiology of defective sperm count (Sharma *et al*., 1999). Spermatozoa are vulnerable to lipid peroxide damage. Production of abnormal levels of ROS alters the spermatozoa membrane function, morphology, impaired metabolism, motility and fertility (Cummins *et al.*, 1994). Excessive ROS production will decrease sperm motility and viability. This will increase midpiece sperm defects and acrosome reaction (St-John *et al*., 2001)

Holyoake *et al*. (1999) found that severe oligozoospermic men with a T8821C point mutation in the ATPase6 gene possess immature spermatids. These authors also found a high incidence of nucleotide substitutions in the mitochondrial genome which were associated with poor semen quality (Holyoake *et al*., 2001). They showed that 8.4% of subnormozoospermic men contain SNPs, while 2.4% normozoospermic

men had an SNP. The two most common substitutions occur in nt 9055 and nt 11719. Eleven percent of men with poor semen quality showed nt 9055 substitution, but only 1.3% of men with normal fertility had the same substitution. Similarly for the nt 11719 substitution, 12% subnormozoospermic men are associated this substitution.

1.4. GENETIC CODE

1.4.1. GENETIC CODE VARIATION

The genetic code directing translation of mtDNA differs from the universal genetic code (Anderson *et al*., 1981). In the mammals' mtDNA, UGA encodes tryptophan instead of being a termination codon. AUG encodes methionine instead of isoleucine and AUG and AGG are termination codons, instead of arginine (Wallace, 2005).

In human, eight mitochondrial tRNAs recognize eight codon families with four fold degeneracy and 14 recognize the remaining codons (cited in Hoaushmand, 2003). The aminoacyl~tRNAs are activated with ATP and are the precursors of polypeptide synthesis on ribosomal RNA which recognizes the correct mRNA codons (Gardner *et al.*, 1991). This facilitates the formation of peptide bonds between new amino acids and growing peptide chains. Most amino acids have more than one codon. These codons specifying for the same amino acids are called synonymous codons.

1.4.2. CODON USAGE

An important characteristic of codon usage is that synonymous codons are used often with different frequencies (Bulmer, 1991). The factors that can influence codon usage are related to nucleotide base compositions and genome size. Mutational bias plays a major role in determining nucleotide base substitutions (Sharp, 1989). Replication errors, repair efficiencies are reflected from mutational bias. Third position transitions in a codon are more common and these mutations are usually synonymous.

Bulmer (1991) and Akashi (1994) demonstrated that several tRNAs recognize groups of synonymous codons. Codon recognized by the most abundant amino acids tends to be used preferentially. Among codons, those which recognized by the most abundant tRNAs are not as favored as the codons that form the natural Watson-crick pairing with tRNA anticodon. Strongly expressed genes are corresponding to most abundant tRNAs than weakly expressed genes. These biases give rise of two models (selective and neutral theory) to explain the preferential use of synonymous codons (Duret, 2003).

In humans, roughly 99.8% of DNA sequence variations do not alter the primary sequence (Venter *et al.*, 2001). However, there are controversies on the functional neutrality of synonymous codons. Not every synonymous site is neutral (Chamary *et al*., 2005). Chamary *et al.* demonstrated that synonymous mutations in mammalian exon can be selective and mediated by selecting mRNA stability. They showed the presence of cytosine at synonymous sites can be accounted for by selection and can affect the efficiency of translation via modification of the tertiary structure of mRNA. Knight *et al*. (2001) showed that codon

usage can be surprisingly biased in the nuclear genes of different species. They observed that the amino acid lysine has two codons, AAA and AAG. These codons were used equally in some organisms like *Lactobacillud acidophilus,* but *Sterptomyces venezuelae* uses AAA only 2.2% of the time, whereas *Buchneria aphidicola* uses it 91%. Similarly, Carlini *et al*. (2003) introduced three different unpreferred codons into the *Drosophila* alcohol dehydrogenase gene (*Adh*). They observed that there is a significant decrease in Adh protein production with those three unpreferred codons.

Thus, all recent researches reported that codon preference can influence translational efficiency. Most of the experiments so far are based on nuclear genes and are not related to human. The general concept may be applicable for mitochondrial gene but their origin and genetic code is not identical.

1.4.3. IS THERE ANY EFFCT OF SYNONYMOUS CODON ON MALE INFERTILITY

Synonymous mutations occur in both nuclear and mitochondrial genes. But their effects are difficult to detect. Different cells contain different number of mtDNA molecules. Further, different homoplasmic mutations can affect translation efficiency and upset the metabolic activity to different degrees. The oocyte contains thousands of mtDNA but is metabolically inactive. Sperm cells are most active and contain smaller number of mitochondria than oocytes (Cummins *et al*., 2001; Kao *et al.,* 1998; Ruiz-Pesini *et al*., 2000). Functional failure of sperm will affect the reproductive ability of men.

It is evident that mitochondrial mutations are associated with male infertility (Kao *et al.,* 1998; Holyoake *et al*., 1999; Ruiz-Pesini *et al*., 2000; Spiropoulos *et al.*, 2002; Thangaraj *et al*., 2003). Synonymous codon usage in nuclear genes has been shown to affect translation efficiency (Chamary *et al*., 2005; Knight *et al*. 2001). Can synonymous mitochondrial mutations affect sperm function?

1.5. HUMAN COMPLEX I

Human complex I consists of seven mitochondrial encoded subunits (ND1, ND2, ND3, ND4, ND4L, ND5 and ND6) and at least 36 nuclear gene coded proteins (Anderson *et al.*, 1981).The total molecular mass of complex I is 900 k Da. This is the largest protein complex of the electron transport chain. The proteins of complex I are traditionally divided into three groups. They are flavoprotein fraction (FPf), iron protein fraction (IPf) and hypophobic fraction (HPf) (Hatefi, 1985). The location of ND6 is still unknown (Tripeles *et al*., 2001), but the other six proteins are located in the HP fraction of complex I. ND1 is the terminal acceptor for quinone reduction within the membrane part of complex I as it has ubiquinone binding property. Complex I deficiency syndrome are also related with human diseases that include LHON (Wallace, 1999; 2005), MELAS and Parkinson disease, etc.

1.5.1. THE HUMAN *ND1* AND *ND2* GENES AND THEIR GENE PRODUCT

The human mitochondrial *ND1* gene is 956 bp long and situated between nt 3308 and nt 4263 of the mitochondrial genome. *ND2* is 1043 bp long

and located beside the *ND1* gene. *ND1* and *ND2* genes encode 316 and 347 amino acid polypeptides, respectively (www.mitomap.org).

ND1 and *ND2* genes are found to be associated with a large number of substitutions (Howell, 1996; Thomas *et al.,* 1996; Torroni *et al.*, 1996; Macaulay *et al*., 1999; Rose *et al.,* 2001; Harris *et al.,* 2006). Twenty published polymorphism that are listed for the *ND1* gene are associated with different diseases (www.mitomap.org, as for 18th may, 2006) (Appendix 4). These diseases occur due to primary mutations (homoplasmic mutations) (e.g.G3460A) or secondary mutations (heteroplamic mutations) (e.g. T4216C). Similarly, six published polymorphisms are listed for the *ND2* in the NCBI database that are associated with different diseases (LHON, Alzheimer's disease or Parkinson's disease) (www.mitomap.org) (Appendix 4).

1.6. AIMS OF THE STUDY

Recent studies have discovered numerous single nucleotide polymorphisms (SNPs) in the coding regions of the mitochondrial genome (Holyoake, *et al*. 2001; Brandstatter *et al.,* 2003; DiMauro *et al*., 2003; Torroni *et al*., 2003; Harris *et al*., 2006). Furthermore, different frequencies of SNPs in different mitochondrial genes can be associated with subnormozoospermia (Holyoake *et al*., 2001). However, the frequencies of SNPs in *ND1* and *ND2* are not unknown.

The present research is aimed to

1. Determine the frequencies of SNPs in the *ND1* and *ND2* genes of the mitochondrial genome in fertile and subfertile men, as this

region is highly polymorphic and is associated with many diseases (Wallace, 1999; 2005).

2. Test the hypothesis that codon usage is associated with fertility phenotypes. Single nucleotide substitutions may be synonymous or non-synonymous. Studies on nuclear genes have shown that synonymous mutations may affect the efficiency of translation (Knight *et al*. 2001; Chamary *et al*., 2005). Thus, knowing how a single nucleotide substitution will change a codon may provide insight into the importance of codon usage and whether codon usage can explain the associated phenotype. In the present study, codon changes caused by single nucleotide substitutions will be determined and the frequency of usage of the codons will be determined by comparison with the reference gene sequence obtained from the NCBI database.

Experimental approach include

1. Isolation of genomic DNA from human semen samples with normal and poor semen parameters.
2. Screening of mutations in the *ND1* and *ND2* genes in the mitochondrial genome using the polymerase chain reaction (PCR) and sequencing.
3. Determining the frequency of the *ND1* and *ND2* mutations.
4. Determine whether codons are changed by nucleotide substitutions by comparing with a reference mtDNA sequence obtained from the NCBI database.
5. Determine the frequency of usage of these changed codons by comparison with the frequency of usage in the reference sequence determined by MEGA version 2.1 (Kumar *et al*., 2001).
6. Determine whether a relationship exist between fertility status and codon usage.

CHAPTER TWO- MATERIALS AND METHODS

MATERIALS AND METHODS

2.1. MATERIALS

All the chemicals used in this study were of molecular grade (Appendix 2). The water used was autoclaved, deionized and ultra filtered (NANO pure ultra pure water system, Barnstead, 10, USA) and referred to as ddH_2O.

2.2. COLLECTION OF SEMEN SAMPLES

Semen samples were collected from patients of the New Zealand Centre for Reproductive Medicine in Christchurch, NZ. These samples were provided to the fertility clinic for routine semen analysis because of suspected fertility problem. Semen parameters, such as sperm count and sperm motility, were measured and recorded by staff at the NZ Centre for Reproductive Medicine. Since a suspected fertility problem may be attributed to either partner, many of the samples had normal sperm count and motility. All semen samples were stored at -80ºC in 200 µL aliquots to reduce DNA degradation by repeated freezing and thawing.

Based on the semen parameters, the samples were classified into seven groups according to the World Health Organization (WHO) guideline from 1999 tabulated in 2.1.

Table: 2.1 Classification of Semen Quality based on WHO (1999).

Classification	SpermCount (millions/mL)	Motility (%)	Samples used in this study		
			Total extracted Genomic DNA	*ND1*	*ND2*
Normozoospermic	>20	>50	75	40	46
Asthenozoospermic	> 20	< 50	43	28	28
Oligozoospermic	5 – 20	>50	10	5	6
Oligoasthenozoospermic	5 – 20	< 50	11	3	5
Severely Oligozoospermic	0 – 5	>50	6	2	3
Severely Oligoasthenozoospermic	0 – 5	>50	5	1	2
Azoospermic	0	0	3	1	2
Total			157	80	92

2.3. DNA PREPARATION FROM SEMEN SAMPLES

2.3.1. DNA EXTRACTION

157 genomic DNA was extracted from semen samples using a standard phenol/chloroform method Kao *et al.* (1995). Semen samples (approximately 200 µL) were thawed on ice. To separate spermatozoa from seminal plasma centrifugation was carried out at 4500 × g for 10 min at 4^{o}C in an Eppendorf bench centrifuge. The supernatant was discarded. Then the pellet was resuspended in 100 µL of lysis mix/extraction buffer (0.1 M Tris-HCl (pH 8.0), 0.1 M NaCl, 0.5 M EDTA (pH8.0), 1% (w/s) SDS, 10 mM dithiothreitol and 100 µg/ml

proteinase K) by gently flicking the tube. Samples were then incubated at 50^{o}C for 16-24 h. Proteins were extracted by adding an equal volume of phenol saturated TE8 (pH 8) buffer [10 mM Tris-HCL (pH 8), 1 mM EDTA]. Samples were gently mixed by inverting the tubes. To separate the aqueous and organic phases the solution was centrifuged at 2400 × g for 5 min at room temperature. The aqueous phase was transferred to a new tube using a large bore glass pipette. The remaining interphase and organic layer were back extracted with an equal volume (100 µL) of TE (pH 8.0) to remove residual DNA. Both aqueous phase were combined and extracted with equal volume of phenol: chloroform: isoamylalcohol (25:24:1). Then the aqueous phase was transferred to another tube and extracted with chloroform: isoamylalcohol (24:1). The final aqueous phase was transferred to a fresh tube and DNA was precipitated out by adding 0.1 volume of 3 M sodium acetate (pH 7) and two volumes of cold absolute ethanol (stored at -20^{o}C). DNA was pelleted by centrifugation at 13,800 × g at 4^{o}C for 20 min. The DNA pellet was washed in 500 µL of cold 70% alcohol. Then, the pellet was air dried for 10-15 min, resuspended in 50 µL TE8 and stored at 4^{o}C.

2.3.2. VISUALIZATION OF EXTRACTED DNA

The presence of DNA was checked by agarose gel electrophoresis. Three µL of the extracted DNA and 3 µL of orange G (30% glycerol and 0.35% orange G) gel loading buffer was prepared and. electrophoresis was done on an agarose gel (1.2%) using 1 × TBE buffer. Electrophoresis was carried out at 5 V/cm for 30 min. Gels were stained with ethidium bromide (0.5µg/ml) and then visualized by using a ultra-violet (UV) transilluminator (300 nm).

2.3.3. QUANTIFICATION OF EXTRACTED DNA

Purity and concentration of the extracted DNA was determined by the ratio of absorbance at 260 nm and 280 nm using a UV spectrophotometer. The concentration was calculated using the 260 nm absorbance reading (1 optical density (OD) unit at 260 nm corresponds to 50 µg/ml double-stranded DNA) (Sambrook *et al.*, 1989).

$$\text{DNA concentration } (\mu g/ml) = A_{260} \times 50 / 2_{(\text{volume factor})} \times 2_{(\text{dilution factor})}$$

The purity of DNA extraction was determined by the ratio of OD at 260 nm to 280 nm.

$$\text{Purity} = A_{260} / A_{280}$$

According to Sambrook *et al.* (1989), a ratio of 1.8 is the estimation for pure DNA quality.

2.4. POLYMERASE CHAIN REACTION (PCR) ANALYSIS

To amplify *ND1*and *ND2* mitochondrial genes, 2 sets of PCR primers were designed, which were located in the flanking regions of each gene.

Primer Design: Primers were based on the human mitochondrial sequence from the Entrez database; accession number NC_001807, provided by the National Centre of Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov). The computer program Primer 3

(http://www.path.com.ac.uk./cgi.bin/primer3.cgi) was used to design the primer sets to amplify the *ND1* and *ND2* genes. The criteria for the primer design were the product size, the primer size, GC content and melting temperature (T_m). The oligonucleotide primers were chemically synthesized by Invitrogen (Auckland, NZ). The properties of the primers are shown in table 2.2.

Table 2.2: Primer pairs used to amplify *ND1* and *ND2* gene.

Primer name	***Genomic location (bp)***	***Sequence (5'-3')***	***GC content***	***T_m (1 m Na^+)***	***Product size***
ND1 F	3245-3264	ccccggtaatcgcataaaact	45%	66°C	1125bp
ND1 R	4370-4351	ttttggattctcagggatgg	45%	66°C	
ND2 F	4382-4402	cctatcacaccccatcctaaa	47%	69°C	1183bp
ND2 R	5565-5545	tgcaacttactgagggctttg	47%	69°C	

F= Forward primer, R= Reverse primer, bp= base pair position related to the reference human mitochondrial sequence (accession number NC_001807).

2.4.1. PCR REACTION

To amplify the *ND1* and *ND2* genes PCR reactions were optimized for each primer pair with different annealing temperature as shown in Table 2. Approximately 160 ng of total sperm DNA was used in 50 µL of reaction mixture. The PCR reaction mixture contained 1.5 µM Mg^{+2}, 200 µM of deoxynucleotide triphosphate (dATP, dCTP, dTTP, dGTP), 1 × Taq buffer, 0.6 µM of each primer and 1 unit of Taq DNA polymerase (Roche). A standard PCR cycle was used initial denaturing at 94º C for 2 min, followed by 35 cycles of denaturing at 94º C for 30 sec, 58º C of annealing for 45 sec and extension of primers at 72º C for 2 min. The final extension was at 72º C for 5 min and reaction products were stored at 4º C.

2.4.2. VISUALIZATION OF PCR PRODUCTS

To confirm the presence of PCR product 3µL of the PCR reaction mixture was analyzed by electrophoresis on a 1.2% agarose gel using 1 × TBE buffer. Electrophoresis was carried out at 5 V/cm for 30 min. Gels were stained with ethidium bromide (0.5µg/ml) and then DNA was visualized using a ultra-violet (UV) transilluminator (300 nm).

2.4.3. PURIFICATION OF PCR PRODUCTS

To remove unused primers and dNTPs, PCR products were purified by gel band excision. The purified PCR products were separated by electrophoresis and stained with ethidium bromide. The bands were excised from the agarose gel and purified using a GFX-PCR DNA and

gel band purification kit (Amersham Bioscience) according to the manufacturer's instruction.

2.4.4. VISUALIZATION OF PURIFIED PCR PRODUCTS

Three microlitres of each purified PCR products were electrophoresed on 1.2% agarose gel using 1 × TBE buffer. Electrophoresis was carried out at 5V/cm for 30 min. Gels were then stained with ethidium bromide and visualized using an UV transilluminator (300nm).

2.5. DNA SEQUENCING

Automated DNA sequencing was performed using the Applied Bio System Big Dye^{TM} termination V 3.1 cycle sequencing kit. The mtDNA samples of different fertility groups were sequenced for both *ND1* and *ND2* gene.

2.5.1. ADDITIONAL PRIMERS FOR SEQUENCING

To get the whole sequence for every sample some additional primers were designed with reduced length and are shown in Table 2.3.

Table 2.3: Additional primers for sequencing for *ND1* Gene.

Primer name	***Genomic location (bp)***	***Sequence (5'- 3')***	***GC content***	***T_m (1 m Na^+)***
ND1 FI	3777-3796	tggctcctttaacctctcca	50%	66°C
ND1 R2	3668-3648	caccctgatcagaggattga	50%	68°C
ND1 R3	3948-3929	attcgatgttgaagcctgaga	42%	67°C
ND2 R2	4884-4865	tgggggctagtttttgtcat	45%	66°C
ND2 R3	5318-5289	gtggggatgatgaggctatt	50%	68°C

F= Forward primer, R= Reverse primer, bp= base pair position related to the reference human mitochondrial sequence (accession number NC_001807).

2.5.2. AUTOMATED SEQUENCING

All sequencing reactions were carried out in 20 µl of sequencing reaction mix containing 0.5µl of BDT (Big Dye™), 1.75 µl of 5X sequencing buffer, 1 µl of primer (stock concentration: 3.2 µM), 2.5 µl of purified PCR products and 10 µl of ddH_2O.

Thermal cycle sequencing reactions were carried out on the thermal cycle (mastercycle@gradient.Eppendroff) with the following cycle parameters: 96^0C of initial denaturing temperature of 10s, 50^0C of annealing temperature of 10s and 60^0C of primer extension temperature for 1.5 min. This thermal cycle was repeated for 25 times.

2.5.3. SEQUENCING REACTION CLEAN UP

Sequencing products were purified using sequencing cleanup plates (Whatman unifilter filtration GF/C 800 # 1 long drip detector microplates (Global)). The clean up plate wells contained a column of sephadex resin beads. Sequencing reactions were applied to each well according to the protocol. The plate was centrifuged at 750 × g for 5 min for several times. Then the plate was then sealed with PCR sealing film and ready to go for automated sequencing in ABI PRISM, 3100 genetic analyzer using Big Dye^{TM} termination V 3.1 cycle sequencing kit (Applied Bio System).

2.5.4. EVALUATION OF DNA SEQUENCE

All the data obtained from automated sequence was edited with a computer based program Sequencher^{TM}. These sequences were compared with the reference sequence (NCBI database, accession number: NC_00187) to determine the nature of mutations.

Codon usage analysis for silent mutations

The synonymous mutations were analyzed for codon usage. The results from this study and the data from Harris *et al*. 2006 were analyzed to predict the amino acid changes by the computer program MEGA version 2.1 (Kumar *et al*., 2001).

Mutation analysis of non-silent mutations

To determine the effect of changed amino acids on protein properties, the proteins were analyzed for hydrophobicity and polarity with computer based program (http://www.roselab.jhu.edu/~raj/MISC/hphobh.html) and ProtScal (http://download.invitrogen.com/evergreen/support_files/a_protscale_expasy.htm) respectively.

Statistical analysis of different mutations

The Z value was determined for testing for significant differences between the mutants/polymorphisms in the sets of fertile and subfertile men which were identified in at least 3 samples. The Z value was calculated from Daniel (1999).

The Chi-square (χ^2) test was used to determine the significant difference between the lower and higher relative codon usage value of the synonymous mutations at 5% level of significance. The Chi-square (χ^2) test was calculated from Walpole (1982).

CHAPTER THREE- RESULTS

RESULTS

3.1. DNA EXTRACTION

Genomic DNA was extracted from 157 semen samples with various fertility statuses (Appendix 3). The table below shows the total number of samples used in this study. These samples were classified to different groups according to the World Health Organization (1999). Figure 3.1 shows the DNA extracted from samples with different fertility status analyzed by electrophoresis.

Table 3.1: Number of semen samples used in this study.

Samples	Total extracted samples	Used in PCR and Sequencing	
		ND1	*ND2*
Normozoospermic	75	40	46
Asthenozoospermic	43	28	28
Oligozoospermic	10	5	6
Oligoasthenozoospermic	11	3	5
Severe oligozoospermic	6	2	3
Severe oligoasthenozoospermic	5	1	2
Azoospermic	3	1	2
Total number of non-normozoospermic	78	40	46

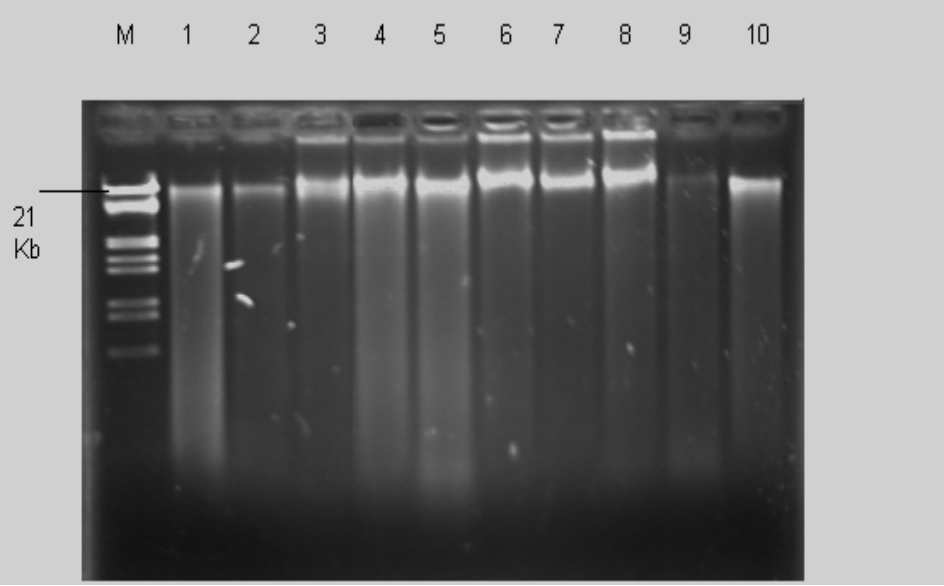


Fig 3.1: Genomic DNA from semen samples. Lane M: Lamda DNA marker digested with restriction enzymes, *EcoRI* and *HindIII*, lane 2-3: normozoospermic, lane 4-5: asthenozoospermic, lane 6: oligozoospermic, lane 7: oligoasthenozoospermic, lane 8: severe oligozoospermic, lane 9: severe asthenoligozoospermic, lane 10: azoospermic.
Electrophoresis was done on an agarose gel (1.2%) using 1 × TBE buffer and visualized using an ultra-violet (UV) transilluminator (300 nm).
The expected sizes of the lambda DNA marker are in descending order, 21kb, 5148bp, 4268bp, 3530bp 2027bp, 1904bp, 1584bp, 1375pb, 941bp, 831bp and 564bp.

3.2. AMPLIFICATION OF mtDNA

PCR reactions were performed to amplify the *ND1* and *ND2* genes from the extracted DNA samples. Among them, 80 samples showed the proper PCR product size (1125 bp) with *ND1* gene and 92 samples showed the correct PCR product size (1183 bp) with *ND2*. Figure 3.2 shows the 1125 bp PCR product of *ND1* for 8 samples Similarly, Figure 3.3 shows *ND2* PCR products for 7 samples.

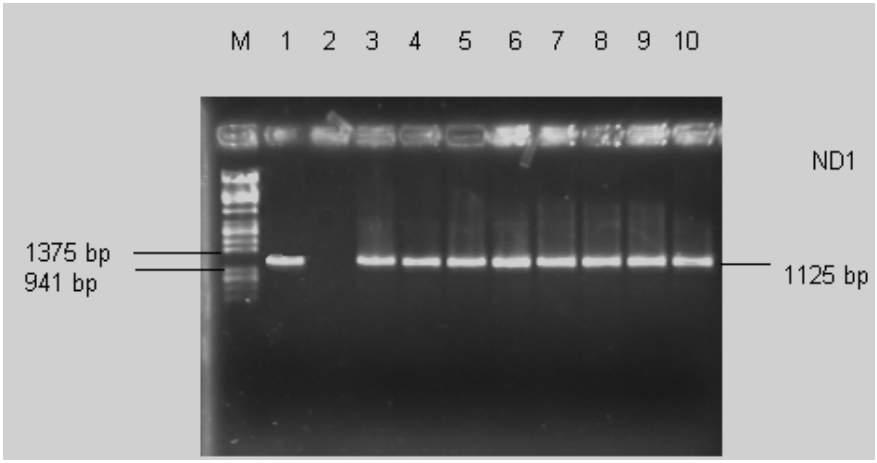


Fig 3.2: PCR products of *ND1*. Lane M: Lambda DNA marker digested with restriction enzymes, EcoRI and HindIII,, lane 1: Positive control (sample containing extracted DNA of correct PCR product size for *ND1*), lane 2: negative control (sample without extracted DNA), lane 3: normozoospermic, lane 4-5: asthenozoospermic, lane 6: oligozoospermic, lane 7: oligoasthenozoospermic, lane 8 severeoligozoospermic, lane 9: severeasthenoligozoospermic, lane 10: azoospermic.
Electrophoresis was done on an agarose gel (1.2%) using 1 × TBE buffer and visualized by using an ultra-violet (UV) transilluminator (300 nm). The expected sizes of lambda DNA marker are in descending order, 21kb, 5148bp, 4268bp, 3530bp, 2027bp, 1904bp, 1584bp, 1375pb, 941bp, 831bp and 564bp.

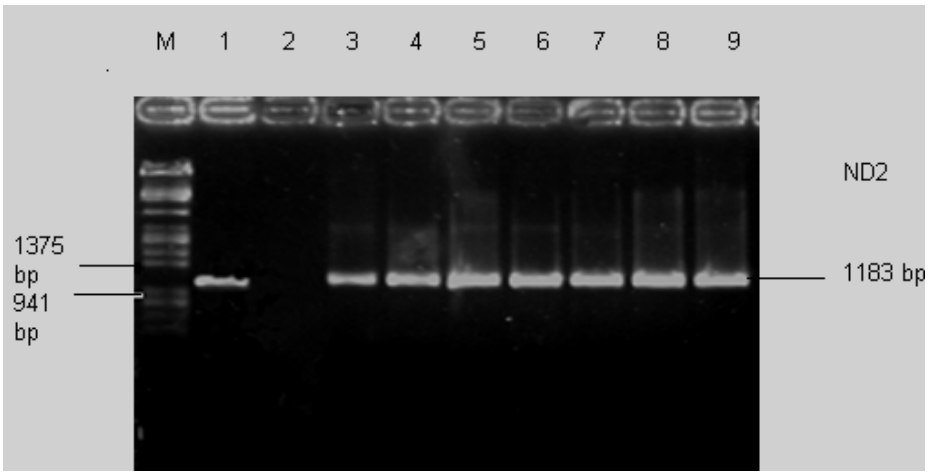


Fig 3.3: PCR products of *ND2*. Lane M: Lambda DNA marker digested with restricted enzymes, *EcoRI* and *HindIII*,, lane 1: Positive control (sample containg extracted DNA of correct PCR product size for *ND2*), lane 2: negative control (sample without extracted DNA), lane 3: normozoospermic, lane 4: asthenozoospermic, lane 5: oligozoospermic, lane 6: oligoasthenozoospermic, lane 7: severe oligozoospermic, lane 8: severe asthenoligozoospermic, lane 9: azoospermic.
Electrophoresis was done on an agarose gel (1.2%) using 1 × TBE buffer and visualized by using an ultra-violet (UV) transilluminator (300 nm). The expected sizes of lambda DNA marker are in descending order, 21kb, 5148bp, 4268bp, 3530bp, 2027bp, 1904bp, 1584bp, 1375pb, 941bp, 831bp and 564bp.

3.3. PCR PRODUCT PURIFICATION

Figure 3.4 shows 10 purified PCR products of *ND1*. Similarly, 10 purified *ND2* PCR products were shown in Fig 3.5.

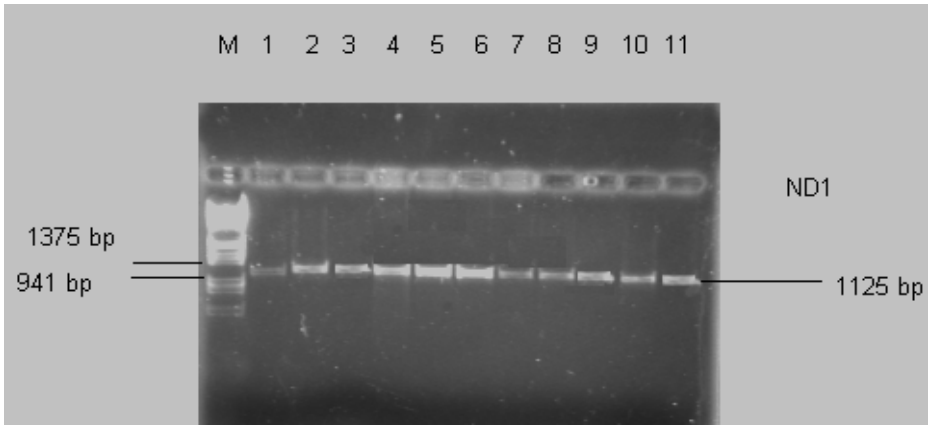


Fig 3.4: Purified *ND1* PCR products. Lane M: Lambda DNA marker, lane 1: Positive control (sample containg extracted DNA of correct purified PCR product size for *ND1*), lane 2: normozoospermic, lane 3-4: asthenozoospermic, lane 5-6: oligozoospermic, lane 7: oligoasthenozoospermic, lane 8-9: severeoligozoospermic, lane 10: severeasthenoligozoospermic, lane 11: azoospermic.
Electrophoresis was done on an agarose gel (1.2%) using 1 × TBE buffer and visualized by using an ultra-violet (UV) transilluminator (300 nm). The expected sizes of lambda DNA marker are in descending order, 21kb, 5148bp, 4268bp, 3530bp, 2027bp, 1904bp, 1584bp, 1375pb, 941bp, 831bp and 564bp.

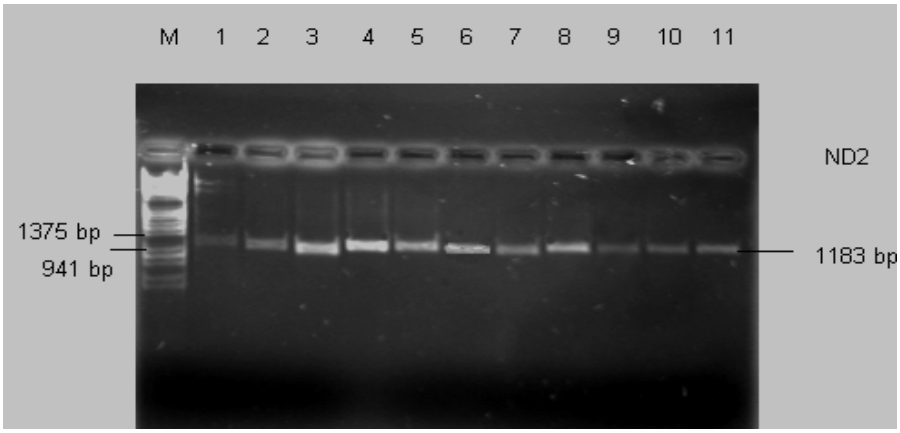


Fig 3.5: Purified *ND2* PCR products. Lane M: Lambda DNA marker, lane 1: Positive control (sample containg extracted DNA of correct purified PCR product size for *ND2*), lane 1-2: Normozoospermic, lane 3-4: asthenozoospermic, lane 5-6: oligozoospermic, lane 7: oligoasthenozoospermic, lane 8-9: severeoligozoospermic, lane 10: severeasthenoligozoospermic, lane 11: azoospermic.
Electrophoresis was done on an agarose gel (1.2%) using 1 × TBE buffer and visualized by using an ultra-violet (UV) transilluminator (300 nm). The expected sizes in lambda DNA marker are in descending order, 21kb, 5148bp, 4268bp, 3530bp, 2027bp, 1904bp, 1584bp, 1375pb, 941bp, 831bp and 564bp.

3.4. DNA SEQUENCE ANALYSIS

3.4.1. *ND1* gene

After purification of PCR products automated DNA sequencing was performed using the Applied Bio System Big Dye^{TM} termination V 3.1 cycle sequencing kit. Eighty samples of different fertility groups were sequenced for the *ND1* gene. These sequences were compared with the reference sequence obtained from the NCBI database (accession number: NC_001807). A computer software program Sequencher^{TM} was used to edit the sequences and determine the nature of mutations. Several mutations in different nucleotide position in *ND1* region were found in different groups of the samples. The base substitutions were located at nucleotides (nts) 3396, 3450, 3480, 3567, 3591, 3594, 3666, 3693, 3992, 4024 and 4216 (Table 3.3).

SNP T3396C

A T to C transition was identified at nt 3396 in the *ND1* gene. This SNP was found in a normozoospermic man (sample code, 010573). This SNP is a synonymous substitution in the third position of a tyrosine codon, changing it from TAT to TAC (Fig 3.6A and 3.6B) (Table 3.3).

SNP C3450T

A C to T transition was identified at nt 3450 in *ND1* region. This transition was observed in an oligoasthenozoospermic (sample code, 6445) sample (figure not shown). This SNP, a synonymous substitution, occurred in the third position of the proline codon, changing the codon CCC to CCT (Table 3.3).

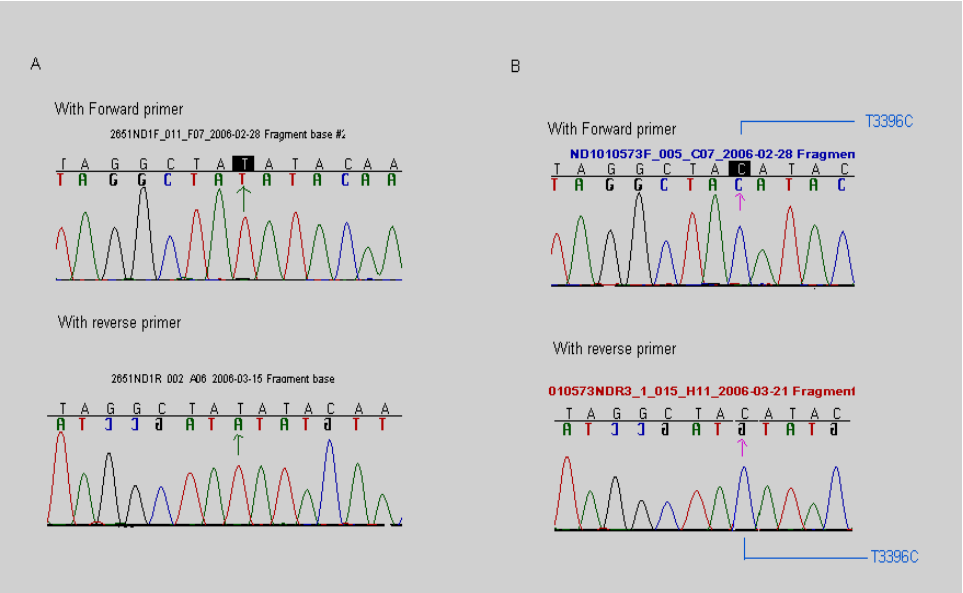


Fig 3.6. A Chromatogram of control sample at nt 3396. B shows T3396C in sample 010573 a normozoospermic sample.

SNP A3480G

An A to G transition was identified at nt 3480 in the *ND1* gene. This SNP was found in four samples in different fertility groups. This transition was observed in 1 normozoospermic (sample code, 010081), 2 in asthenozoospermic (sample code, 010480, 2495) and 1 in oligozoospermic (sample code, 2902) samples (Fig 3.7). This SNP was a synonymous substitution occurring in the third position of the lysine codon, changing the codon from AAA to AAG (Table 3.3).

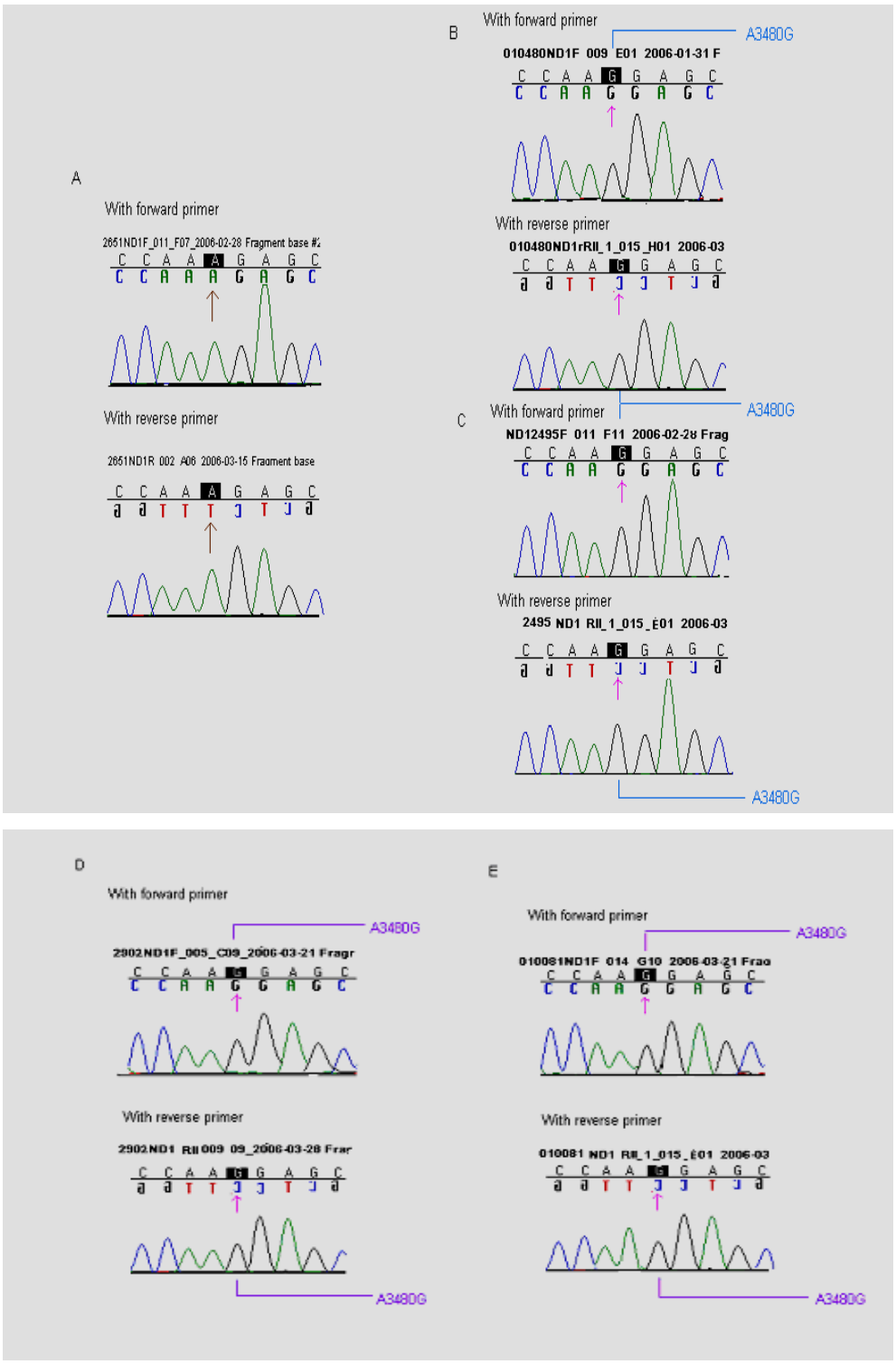


Fig: 3.7. A shows Chromatogram of control sample at nt 3480. B and C are from asthenozoospermic samples (010480, 2495), D from oligozoospermic (2902) sample and E from normozoospermic (010081) sample show transition of A to G at nt 3480.

SNP G3591A

A G to A transition was identified at nt 3591 in the *ND1* region. This transition was observed in an oligozoospermic sample (sample code, 010208, figure not shown). This SNP is a synonymous substitution that occurred in the third position of leucine codon, changing the codon CTG to CTA (Table 3.3).

SNP C3594T

A C to T transition was identified at nt 3594 in the *ND1* gene. This SNP was found in a normozoospermic (sample code, 010573) sample. This SNP is a synonymous substitution occurring in the third position of the valine codon, changing the codon from GTC to GTT (Fig 3.8C and 3.8D) (Table 3.3).

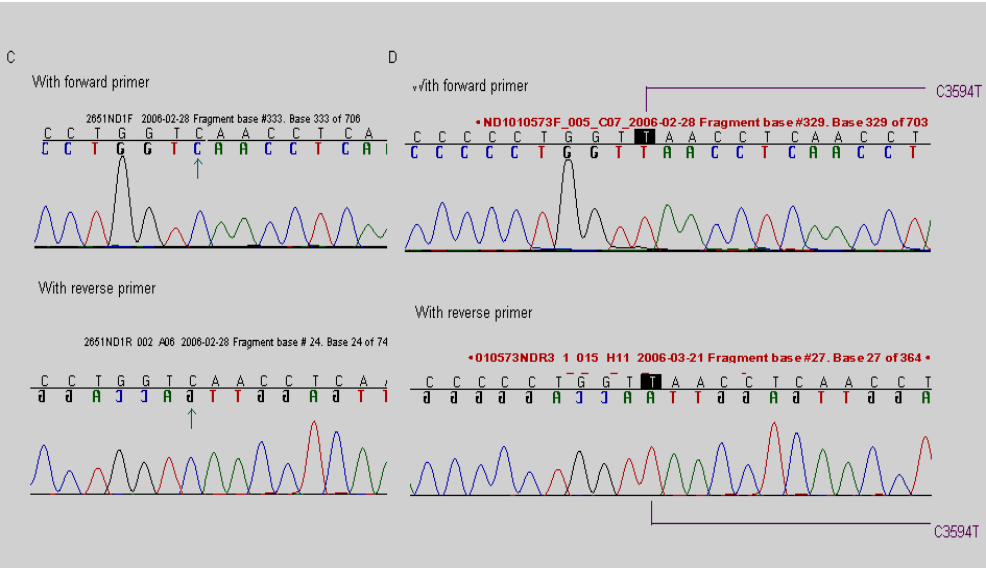


Fig: 3.8. C shows chromatograms of control sample at nt 3594 respectively D shows C3594T respectively in sample 010573, a normozoospermic sample.

SNP G3666A

A G to A transition was identified at nt 3666 in the *ND1* region. This SNP was found in a normozoospermic sample (sample code, 010573) and was a synonymous substitution occurring in the third position of the glycine codon, changing the codon from GGG to GGA (Fig 3.9) (Table 3.3).

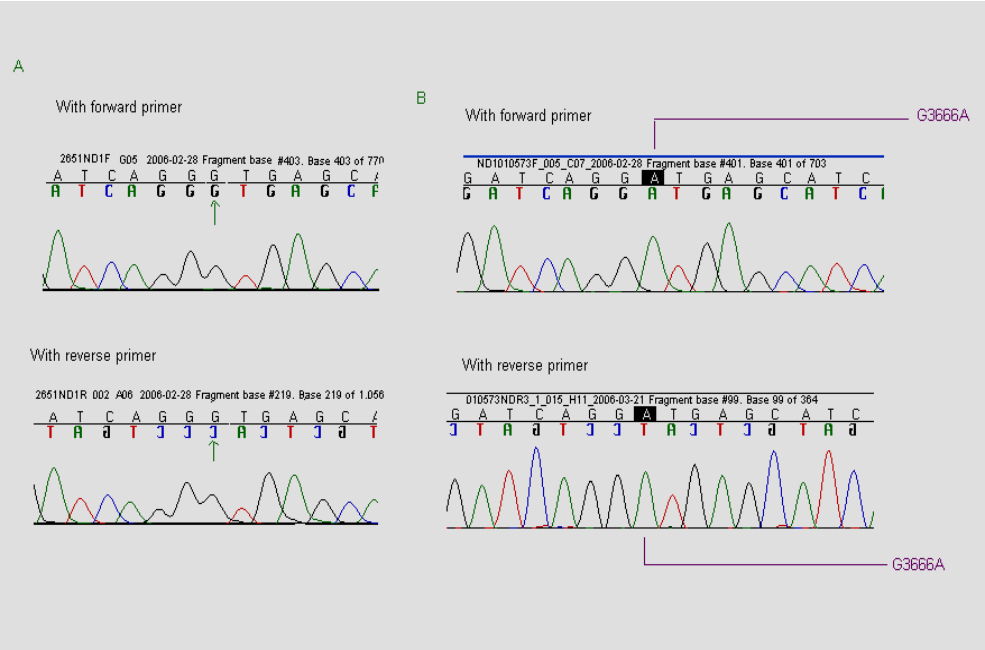


Fig 3.9: A Chromatogram of control sample at nt 3666. B shows G3666A in a normozoospermic sample 010573.

SNP G3693A

A G to A transition was identified at nt 3693 in the *ND1* gene. This SNP was found in a normozoospermic sample (sample code, 010573). This SNP is a synonymous substitution occurring in third position of the leucine codon, changing the codon from CTG to CTA (figure not shown) (Table 3.3).

SNP C3992T

A C to T transition was identified at nt 3992 in *ND1* gene. This SNP was found in a sample of a severely oligozoospermic man (sample code, 010639) (Fig 3.11). This SNP is a non-synonymous substitution that occurred in the second position of the threonine codon, changing the codon from ACA to ATA. A C to T transition at nt 3992 changed methionine (ATA) to threonine (ACA) (Table 3.4). The secondary structures of each residue in an amino acid sequence were predicted by computer program nnpredict (http://www.cmpharm.ucsf.edu/~nomi.nnpredict.).

nt 3992

↓

Normal Sequence: G C C G A A T A C A C A A A C A T T A T T A T A

Variant Sequence: G C C G A A T A T A C A A A C A T T A T T A T A

Secondary structure predicting

↓

Normal Sequence: - - - - - - H H H H H H H H H – H H H – H - - -

Variant Sequence: - - - - - - H H H H H H H H H – H H H – H - - -

(H = helix, E = strand, - = no prediction, tertiary structure class: alpha/beta)

Fig: 3.10. nnpredict of human ND1 protein at nt 3992.

nnpredict (http://www.cmpharm.ucsf.edu/~nomi.nnpredict.) is a computer program that predicts the secondary structure type of each residue in an amino acid sequence. However, when change the parameter to use alpha/beta tertiary structure, the structure remains almost the same.

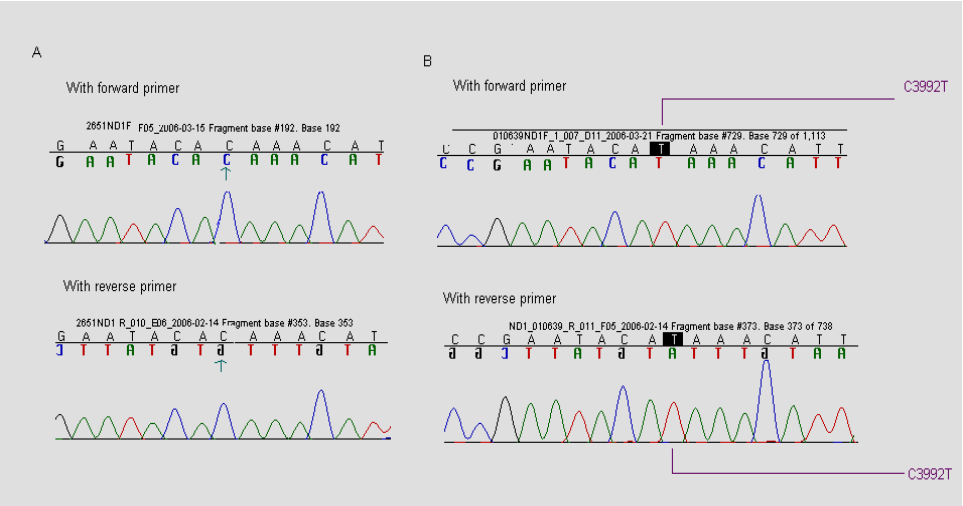


Fig 3.11. A: Chromatogram of control sample at nt 3992. B shows C3992T transition in, a severely oligozoospermic (sample code, 020639) sample.

SNP A4024G

An A to G transition was identified at nt 4024 in the *ND1* region. This transition was observed in a severely oligozoospermic (sample code, 010639) sample (Fig 3.12). This SNP is a synonymous substitution occurred in the third position of leucine codon, changing from the codon CTA to CTG (Table 3.3).

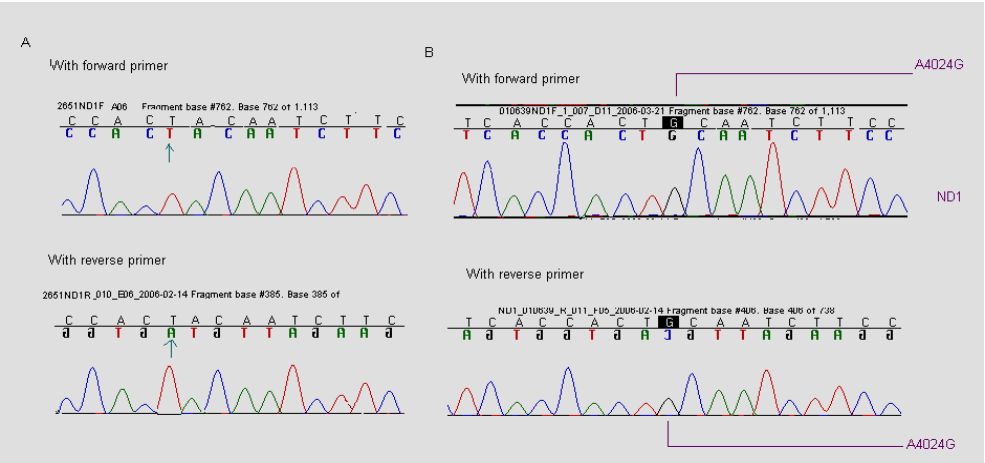


Fig 3.12. A: Chromatogram of control sample at nt 4024. B shows A4024G transition in 010639, a severely oligozoospermic sample.

SNP T4216C

A T to C transition was identified at nt 4216 in the *ND1* region. This transition is observed in one normozoospermic (sample code, SF 96-229) and 2 asthenozoospermic (sample code, 010545, 2409, figure not shown) (Fig 3.14) samples. This non-synonymous substitution occurred in the first position of tyrosine codon, changing the codon from TAT (tyrosine) to CAT (histidine) (Table 3.4). The secondary structures of each residue in an amino acid sequence were predicted by computer program nnpredict (http://www.cmpharm.ucsf.edu/~nomi.nnpredict.).

nt 4216

↓

Normal Sequence: T T A C T T A T A T G A T A T G T C T C C A T A

Variant Sequence: T T A C T T A C A T G A T A T G T C T C C A T A

Secondary structure predicting

↓

Normal Sequence: - - - - - - H H – H - - - - - - - - - - - - - -

Variant Sequence: - - - - - - H H - - - - - - - - - - - - - - -

(H = helix, E = strand, - = no prediction, tertiary structure class: alpha/beta)

Fig: 3.13. nnpredict of human ND1 protein at nt 4216.

nnpredict (http://www.cmpharm.ucsf.edu/~nomi.nnpredict.) is a computer program that predicts the secondary structure type of each residue in an amino acid sequence. However, when change the parameter to use alpha/beta tertiary structure, the structure remains almost the same.

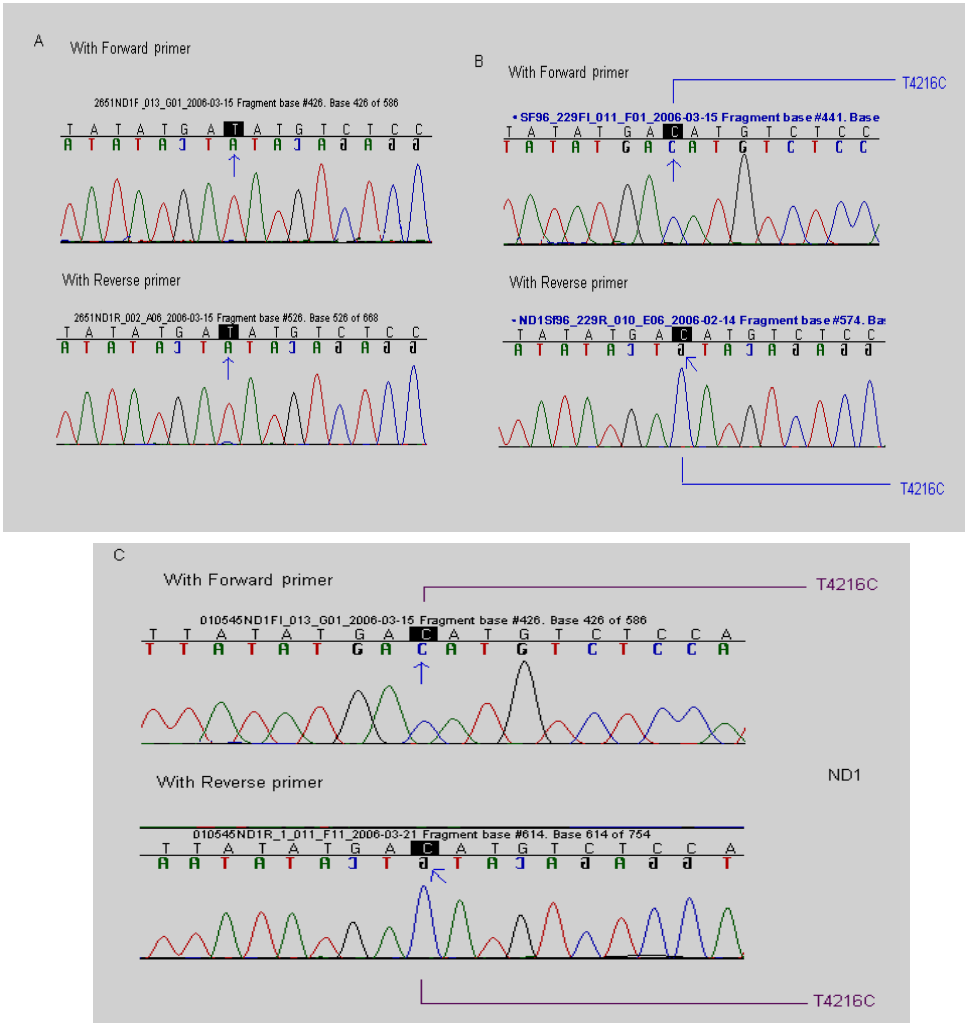


Fig 3.14. A: Chromatogram of control sample at nt 4216. B shows T4216C transition in SF96-229, a normozoospermic. C shows T4216C transition in 010545 an asthenozoospermic sample.

3.4.2. *ND2* gene

Ninety two samples from different fertility groups were sequenced for the *ND2* gene. These sequences were compared to the reference sequence, in the NCBI database (accession number: NC_001807). A computer software program Sequencher^{TM} was used to edit the sequences and

determine the nature of mutations. Two mutations were found in different fertility groups of the samples at different nucleotide position in this gene. The base substitutions were located at nucleotides (nts) 4514 and 4580 (Table 3.3).

SNP A4514G

An A to G transition was identified at nt 4514 in the *ND2*gene. This transition was observed in a normozoospermic (sample code, 6383) (figure not shown), and in asthenozoospermic, sample code, 2409 and 5097(figure not shown) sample (Fig 3.15). This SNP is a synonymous substitution occurring in the third position of an alanine codon, changing the codon GCA to GCG (Table 3.3).

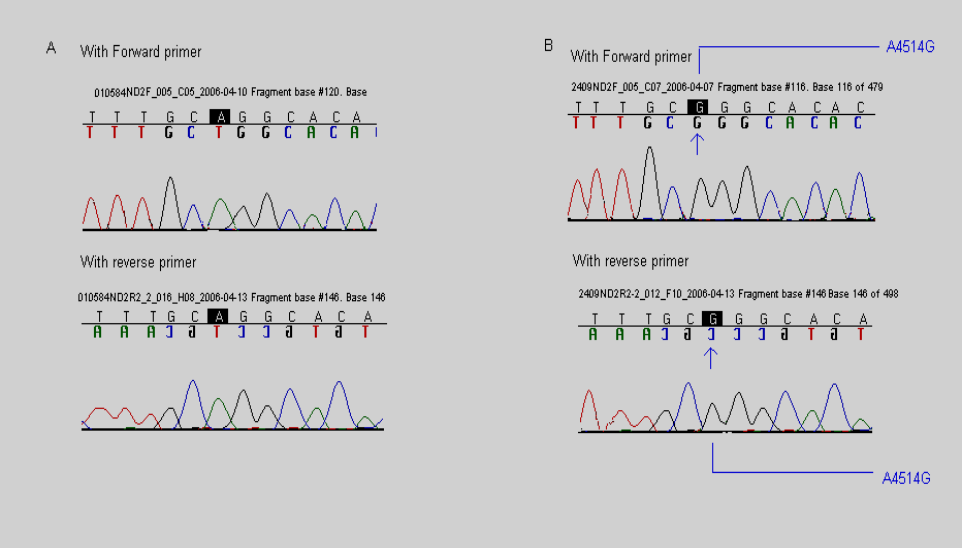


Fig 3.15. A: Chromatogram of a control sample at nt 4514 in ND2 region. B shows A4514G transition in 2409, an asthenozoospermic sample.

SNP G4580A

A G to A transition was identified at nt 4580 in the *ND2* region. This transition was observed in three normozoospermic sample (sample code, 010081, (figure not shown), 010526 (figure not shown) and 6750), in one oligoasthenozoospermic sample (sample code, 009593) and in one oligozoospermic (sample code 010830) sample (Fig 3.16). This SNP is a synonymous substitution that occurred in the third position of methioninie codon, changing the codon from ATG to ATA (Table 3.3).

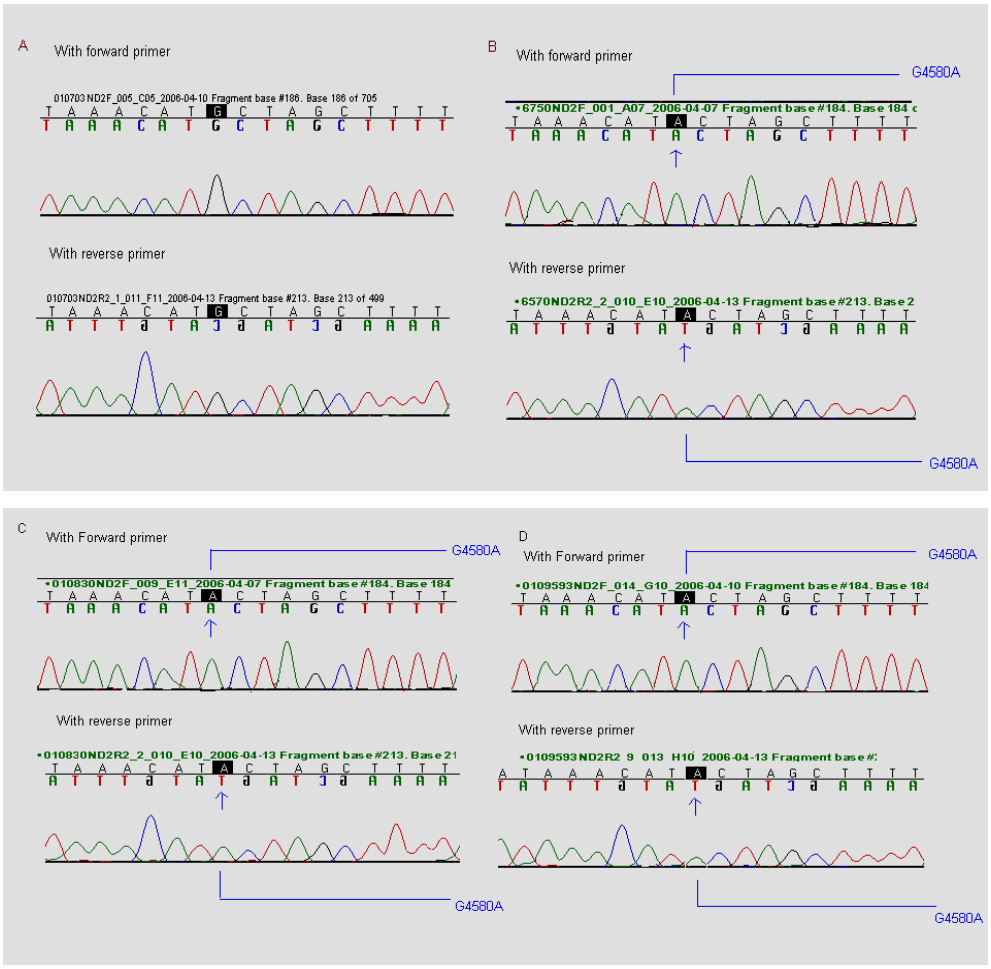


Fig 3.16. A: Chromatogram of a control sample at nt 4580 in the *ND2* region. B shows nt 4580 transition in normozoospermic (6750) sample. C and D show the G4580A transition in an oligozoospermic (010830) and oligoasthenozoospermic (009593) sample.

3.5. STATISTICAL ANALYSIS

A statistical analysis was performed on individual SNPs that were identified in at least 3 samples of the total fertile and subfertile population. The Z value was determined for testing for significant differences between the mutants/polymorphisms in the sets of fertile and subfertile men (Table 3.2).

The Z value was calculated from $Z = | P_1 - P_2 | / \sqrt{} (P (1-P)/n)$ (Daniel, 1999).

Where, $P_1 = x_1/n_1$, $P_2 = x_2/n_2$, $n=n_1+n_2$, $P = (n_1P_1 + n_2P_2)/ (n_1+n_2)$, x_1= number of mutants in fertile group, x_2 = number of mutants in subfertile group, n_1= total number of fertile men, n_2= total number of subfertile men.

Table 3.2. Statistical analysis of different mutants in both *ND1* and *ND2* genes.

SNPs	Gene	P_1(Fertile Men)	P_2(Infertile Men)	Z-Value
A3480G	*ND1*	1/40	3/40	0.2236
T4216C	*ND1*	1/40	2/40	0.1290
G4580A	*ND2*	3/46	2/46	0.0932
A4514G	*ND2*	1/46	2/46	0.1203

In all the four cases, z-values were less than the 5% value (z=1.96 at 5% level of significance). Therefore the observed differences in all the four mutants were not significant.

3.6. ANALYSIS OF CODON USAGE IN SYNONYMOUS MUTANTS IN MITOCHONDRIAL DNA

The frequency of codon usage for each mitochondrial gene was determined using the computer program MEGA version 2.1 (Kumar *et al*., 2001). The amino acid and their frequency change rate were determined. Data obtained from a previous study (Harris *et al.,* 2006) has also been analyzed for codon usage. Analyses of codon usage in synonymous mutants from this study and from Harris *et al.* (2006) are tabulated in Table 3.3.

Table 3.3: DNA sequence analysis and the codon usage of synonymous mutations in different groups of semen samples in mt DNA.

Classification	*gene*	*Samples*	*Spm Cnt (M/ml)*	*Motility %*	*Mutation*	*Codon Change*	*Amino Acid change*	*Codon frequency change*	*Codon frequency change rate*
N	*ND1*	010573*	93	62	T3396C*	TAT→TAC	Silent	1.63→1.60	Lower
					C3594T*	GTC→GTT	Silent	2.00→0.80	Lower
					G3666A*	GGG→GGA	Silent	1.00→.033	Lower
					G3693A*	CTG→CTA	Silent	0.48→1.57	Higher
		010081*	109	53	A3480G*	AAA→AAG	Silent	1.71→0.29	Lower
As	*ND1*	010480*	70	24	A3480G*	AAA→AAG	Silent	1.71→0.29	Lower
		2495*	24	31	A3480G*	AAA→AAG	Silent	1.71→0.29	Lower
So	*ND1*	010639*	1	4	A4024G*	CTA→CTG	Silent	2.57→0.48	Lower
O	*ND1*	010208*	4.6	22	G3591A*	CTG→CTA	Silent	0.48→2.57	Higher
		2902*	3.2	35	A3480G*	AAA→AAG	Silent	1.71→0.29	Lower
		6768*	3.8	38	C3567T*	ACC→ACT	Silent	2.29→0.11	Lower
OAs	*ND1*	6445*	9.3	39	C3450T*	CCC→CCT	Silent	3.09→0.36	Lower
N	*ND2*	010081*	109	53	G4580A*	ATG→ATA	Silent	0.17→1.83	Higher
		010526*	109	636	G4580A*	ATG→ATA	Silent	0.17→1.83	Higher
		6570*	75	79	G4580A*	ATG→ATA	Silent	0.17→1.83	Higher
		6383*	70	56	A4514G	GCA→GCG	Silent	1.60→0.60	Lower
As	*ND2*	2409*	55	28	A4514*	GCA→GCG	Silent	1.60→0.60	Lower
		5097*	65	37.2	A4514G*	GCA→GCG	Silent	1.60→0.60	Lower
O	*ND2*	010830*	46	15	G4580A*	ATG→ATA	Silent	0.17→1.83	Higher
OAs	*ND2*	009593*	21	50	G4580A*	ATG→ATA	Silent	0.17→1.83	Higher
Az	*ND1*	2473	0	0	A3480G	AAA→AAG	Silent	1.71→0.29	lower
Az	*ND1*	4071	0	0	A3480G	AAA→AAG	Silent	1.71→0.29	lower
SOAs	*ND1*	4110	5	18	A3480G	AAA→AAG	silent	1.71→0.29	lower

O	*ND1*	5597	16.8	53	A3480G	AAA→AAG	Silent	1.71→0.29	lower
Az	*COXI*	2473	0	0	G6260A	GAG→GAA	Silent	0.6→1.40	Higher
As	*COXI*	2772	24	30	T6911G	GCT→GCG	Silent	1.20→0.10	lower
As	*COXI*	5415	161	41	T6911G	GCT→GCG	Silent	1.20→0.10	lower
As	*COXII*	CCR	42	33	G7853A	GTC→ATC	Silent	2.15→1.36	Lower
As	*COXII*	QD747V	34	29	G8251A	GTC→ATC	Silent	1.45→1.45	-
As	*COXII*	QR509H	222	57	G8251A	GTC→ATC	Silent	1.45→1.45	-
As	*COXII*	94-150	26	19	G8251A	GTC→ATC	Silent	1.45→1.45	-
N	*COXII*	SF95-122	115	62	G8251A	GTC→ATC	Silent	1.45→1.45	-
As	*COXII*	SF96-279	65	30	C7819A	CTC→CTA	Silent	2.40→1.78	lower
As	*COXII*	QD788L	85	50	G7789A	CTG→CTA	Silent	0.55→2.18	Higher
SOAs	*COXIII*	CD	0.65	8	G9380A	TGG→TGA	Silent	0.67→1.33	Higher
O	*COXIII*	AM472M	19.80	57	G9380A	TGG→TGA	Silent	0.67→1.33	Higher
N	*COXIII*	AM235Q	190	51	G9380A	TGG→TGA	Silent	0.67→1.33	Higher
Az	*ND6*	4449	0	0	G14581A	ATG→ATA	Silent	1.6→0.4	Higher Lower
As	*ND6*	AT665c	51	29	G14581A	ATG→ATA	Silent	1.6→0.4	Higher Lower
Az	*ND6*	2473	0	0	G14581A	ATG→ATA	Silent	1.6→0.4	Lower
As	*ND6*	6169	70.40	56	G14581A	ATG→ATA	Silent	1.6→0.4	Higher
Az	*hND2*	3008	0	0	C5360G	ATC→ATG	Silent	1.50→0.17	Lower
OAs	*ND5*	CP	4.6	32	G13368A	CGG→CGA	Silent	0.0→1.50	Higher
SO	*ND5*	2578	2.5	70	G12618A	TTG→TTA	Silent	1.50→0.78	Lower
OAs	*ND5*	2703	10.6	15	G12618A	TTG→TTA	Silent	1.50→0.78	Lower
N	*h ND1*	SF 94_97	86	72	C3414G	GCC→GCG	Silent	2.07→2.67	Higher
N	*h ND1*	QR 509H	75	82	C3414G	GCC→GCG	Silent	2.07→2.67	Higher
N	*ATPase6*	SF96_305	52	55	A9150G	TTA→TTG	Silent	0.55→0.14	Lower
As	*ATPase6*	94-149	34	40	A9150G	TTA→TTG	Silent	0.55→0.14	Lower
OAs	*ATPase6*	SF362	15.5	26	A9150G	TTA→TTG	Silent	0.55→0.14	Lower
As	*ATPase6*	QD747V	34	41	G8856A	GCG→GCA	Silent	0.19→0.76	Higher
0	*ND4L*	5378	8	16.80	A10551G	ATA→ATG	Silent	1.80→0.20	Lower
	ND4				C10874T	CTA→TTA	Silent	2.63→0.50	Lower
					C11018T	AGA→AGT	Silent	1.26→0.14	Lower
					A11468G	AAA→AAG	Silent	1.64→0.36	Lower
					A12309G	AAA→AAG	INTRON		
	ND4/ND5								
	ND5				G12373A	CTG→CTA	Silent	0.46→2.0	Higher
					T12706C	ATT→ATC	Silent	0.72→1.28	Higher

O	*ND4L*	JN	3.4	18	A10551G	ATA→ATG	Silent	1.80→0.20	Lower
	ND4				C10874T	CTA→TTA	Silent	2.60→0.40	Lower
					C11018T	AGA→AGT	Silent	1.26→0.14	Lower
					A11468G	AAA→AAG	Silent	1.64→0.36	Lower
					A12309G	AAA→AAG	INTRON		
	ND4/ND5								
	ND5				G12373A	CTG→CTA	Silent	0.46→2.0	Higher
Az		3008	0	0					
					C10874T	CTA→TTA	Silent	2.60→0.40	Lower
	ND4				C11018T	AGA→AGT	Silent	1.26→0.14	Lower
									Lower
SOAs	*ND4L*	TI	4.8	32	A10551G	ATA→ATG	Silent	1.80→0.20	Lower
	ND4				C10874T	CTA→TTA	Silent	2.60→0.40	Lower
					A11468G	AAA→AAG	Silent	1.64→0.36	Lower
					G11915A	ACG→ACA	Silent	0.8→1.40	Higher
					A12309G	AAA→AAG	INTRON		
	ND4/ND5								
	ND5				G12373A	CTG→CTA	Silent	0.46→2.0	Higher
					T12706C	ATT→ATC	Silent	0.72→1.28	Higher
					T13327C	TGT→TGC	Silent	0.33→1.67	Higher
N	*ND3*	2905	135	64	A10290G	TGA→TGC	Silent	1.83→0.17	Lower
	ND4L				A10551G	ATA→ATG	Silent	1.80→0.20	Lower
	ND4				C10874T	CTA→TTA	Silent	2.60→0.40	Lower
					A12309G	AAA→AAG	INTRON		
	ND4/ND5				G12373A	CTG→CTA	Silent	0.46→2.0	Higher
	ND5				T12706C	ATT→ATC	Silent	0.72→1.28	Higher
					G12502A	ATG→ATA	Silent	0.23→1.77	Higher

* Indicates the mutations obtained from different groups of semen samples from this study. Other samples were analyzed from Harris *et al* 2006. As= Asthenozoospermic, Az= Azoospermic, OAs= Oligoasthenozoospermic, SOAs= Severely oligozoospermic, So= Severely oligozoospermic, O= Oligozoospermic and N= Normozoospermic samples.

3.7. ANALYSIS OF NON-SYNOMYMOUS MUTATIONS

Two non-synonymous (T4216C and C3992T) mutations were found in this study (Table 3.4). These mutations were also analyzed for hydropathicity and polarity value. These hydropathicity and polarity values were compared with that of the normal amino acid sequence. According to Kyte and Doolittle (1982), the hydropathicity readings are the measurement of the ability of polypeptides to interact with water. If the value was below zero, then the amino acid was hydrophilic, otherwise it was hydrophobic. The polarity reading was assessed by the change of the amino acid. If the reading was above ten, then the amino acid was basic or acidic. If the reading was in between five to seven then it was nonpolar (Granthum, 1974). The hydrophathicity and polarity values for all the non-synonymous mutations from this study are tabulated in Table 3.5.

Table 3.4: **DNA sequence analysis of non-silent mutants in different groups of semen samples.**

Classification	***Gene***	***Samples***	***Spm Cnt (M/ml)***	***Motility %***	***Mutation***	***Codon change***	***Amino Acid change***
N	*ND1*	SF 96-229	75	50	T4216C	TAT→CAT	Tyr→His
As	*ND1*	2409	55	28	T4216C	TAT→CAT	Tyr→His
		010545	74	19	T4216C	TAT→CAT	Tyr→His
So	*ND1*	010639	1	4	C3992T	ACA→ATA	Thr→Met

Here, As= Asthenozoospermic, So= Severe oligozoospermic and N= Normozoospermic samples.

Table: 3.5. Hydropathicity and polarity analysis of non-silent mutants.

Classification	***Gene***	***Sample***	***Mutation***	***Codon change***	***Amino acid change***	***Hydropathicity***		***Polarity***	
						Normal	**variant**	**Normal**	**variant**
N	*ND1*	SF 96-229	T4216C	TAT→CAT	Tyr→His	1.222	1.011	6.211	6.678
As		2409 010545							
So	*ND1*	010639	C3992T	ACA→ATA	Thr→Met	0.622	0.911	7.300	7.622

Here, As= Asthenozoospermic, So= Severe oligozoospermic and N= Normozoospermic samples.

CHAPTER FOUR- DISCUSSION

DISCUSSION

4.1. MUTATION DETECTION BY PCR AND AUTOMATED DNA SEQUENCING

Possible mutations within a nucleotide sequence can be detected by various methods. Polymerase chain reaction (PCR) is one of the unique, uncomplicated and rapid techniques that can amplify millions of copies of a gene from just one starting molecule (Brown *et al.,* 1995). The importance of PCR can be demonstrated by the successful amplification of DNA from an isolated sperm cell, containing just one haploid copy of a human genome or a few mitochondria (Brown *et al.,* 1995; Holyoake *et al*., 1999). Thus, its ease of application and reliability allows it to be widely used in genetic screening and research (Klug and Cummings, 1997).

In this study, *ND1* and *ND2* gene were amplified from 80 and 92 semen samples, respectively, from different fertility groups.

For this study, all the PCR products were sequenced with forward and reverse primers, and the mutations were confirmed by at least two or more sequencing reactions.

4.2. MITOCHONDRIAL DNA REPLICATION ERROR

The mutation rate of mitochondrial DNA has been estimated to be 10 fold higher than that of nuclear DNA (Copeland *et al.,* 2002). The presence of mutations are associated with various human diseases (Wallace, 1992;

1994; 1999; Torroni *et al*., 1996; Wallace *et al*., 1999; Shanske *et al*., 2001; Copeland *et al*., 2002; DiMauro *et al.,* 2003; Houshmand, 2003; Leob *et al*., 2005). Fidelity of mtDNA replication is important. Mitochondria appear to lack an efficient DNA repair system compared with the nuclear DNA replication system. Mitochondrial DNA also lack of protection of histone protein. Thus, mutations may result from DNA damage caused by external factors (Copeland *et al.,* 2002). Mitochondrial DNA polymerase (poly γ) is the only DNA polymerase that resides in mitochondria (Leob *et al.,* 2005). Poly γ is responsible for mtDNA synthesis including replication of mitochondrial genome and repair of DNA damage (Leob *et al.,* 2005). Expression of a mutated version of the poly γ gene results in the accumulation of mutations in mtDNA (Spellbrink *et al.,* 2002). The error rates for the human mitochondrial DNA poly γ are higher than those of chicken and pig DNA polymerase γ (Kunkel *et al*., 1988; Kunkel *et al*., 1989). Mice bearing mutant poly γ showed respiratory dysfunction in the absence of enhanced ROS production (Leob *et al.,* 2005). Leob *et al.* (2005) suggested that mitochondrial mutations were generated by mutant polymerase γ during mtDNA replication. However, a recent study by Harris *et al.* (2006) has established that there was no relationship between the poly γ genotypes and SNP frequencies in normozoospermic and non-normozoospermic men. Therefore, mutations in *ND1* and *ND2* genes of the semen samples, which were found in this study, might not be caused by mtDNA replication error.

4.3. MUTATION ANALYSIS

4.3.1. *ND1* MUTATION SCREENING

From this research 11 different mutations were identified in the *ND1* gene. Only two variants were present in more than one sample. An A to G transition at nt 3480 was found in the *ND1* gene in both normozoospermic and non-normozoospermic (asthenozoospermic and oligozoospermic) men. Only 1 out of 40 (2.5%) normozoospermic men showed this mutation. But 3/40 (7.5%) subfertile men displayed the same A to G transition at nt 3480. The sperm motility for these three subfertile patients was around 30-35%, while that of the normozoospermic patient was 53%. A comparison with sperm motility of other normozoospermic men found that they all had a higher sperm motility of around 70% or greater. Could the A3840G mutation be the cause for low sperm motility? This A3480G transition was a silent mutation, which did not change the amino acid. Examination of the codon usage revealed that relative frequency of usage of the AAA codon is higher than AAG (1.71→0.29). Further, there are only 22 tRNAs in mitochondria. A change from a commonly used codon to a less commonly used codon may affect translation efficiency. This has been established by Duan *et al.* (2003). They demonstrated that a synonymous mutation in the human domain receptor D2 affects the folding of DRD2 mRNA. This mutation destabilized the mRNA structure and lowered the translation efficiency. Similarly, Smith *et al*. (2001) found in *E. coli* that less commonly used synonymous codons were maintained within the gene because of the ribosomal binding motif at the start of a gene affecting translation efficiency.

A T to C transition at nt 4216 in the *ND1* gene found in different fertility groups in this study. One normozospermic man out of forty (2.5%) showed this transition. On the other hand 2/40 subfertile men (5%) were with T4216C mutation. This is a non-silent mutation which changes the amino acid from tyrosine to histidine. Both amino acids are hydrophobic and nonpolar. Secondary structure prediction of the proteins showed that the structure was still similar despite the change. Hydrophobicity plots of proteins with changed amino acids as a result of nucleotide substitution could not affect the function of a protein (Holyoake *et al*., 2001). So the T4216C mutation might not have a severe effect on the function of ND1 protein but it is still doubtful.

Only one normozoospermic sample showed 4 different mutations in 4 nucleotide positions (3396, 3594, 3666, and 3693) in the *ND1* gene. No other normal semen samples used in this study showed these mutations. It is possible that these mutations were transmitted from his mother. But as his mother's mtDNA genome was not available for examination it is not clear whether these mutants were inherited from his mother or *de novo* mutations. All these mutants are silent and their codon frequency change rates are lower except in nt 3693 for a G to A transition. The sperm motility of this man was 62%.

A severely oligozoospermic sample had an A to G silent transition at nt 4024, and a C to T non-silent transition at nt 3992 with an amino acid change from threonine to methionine. However, this mutation did not change the predicted secondary structure. The sperm count and sperm motility of this man were 1M/ml and 4%, respectively.

Three silent mutants were also found in three different nucleotide positions (G3591A, C3567T and C3450T) in the *ND1* gene in three different subfertile patients. Except for a G3591A transition mutation in

an oligozoospermic sample, the others displayed a reduced level of codon usage frequency.

Overall the *ND1* gene mutations occurred more in subfertile men than fertile men. All of the mutations appear to be associated with reduced sperm motility. Nine were synonymous and two (T4216C, C3992T) were non-synonymous (Table 3.3; Table 3.4). Of these, eight non-silent mutations from both fertile and sufertile groups were associated with reduced level of relative codon usage frequency (Table 4.2). However, it is not known whether change in codon usage will affect translation efficiency.

4.3.2. *ND2* MUTATION SCREENINIG

Two different mutations were identified in the *ND2* gene from this research. A G to A transition was observed at nt 4580 in different samples of different fertility status. This mutation occurred in 6.5% of fertile men and 4.3% of subfertile men. The samples, included subnormozoospermic and normozoospermic men with this mutation had a motility range of 15-79%. As the occurrence rate of this mutation in the two groups was similar, this mutation might not be responsible for poor semen quality. Further this was a silent mutant with codon change from ATG to ATA. The relative frequency of codon usage rate for ATG and ATA is 0.17 and 1.83, respectively.

A silent mutation in *ND2* at nt 4514 was found in 2/46 (4.3%) subfertile men and 1/46 (2.2%) fertile men. The sperm motility for these two subfertile patients was 28% and 37%, while that of the normozoospermic patient was 56%. The sperm motility of other normozoospermic men tended to be higher, greater than 70%. The A4514G mutation could be a

cause for low sperm motility. But this mutation might not be solely responsible for poor semen production in men. The relative frequency of codon usage rate was reduced from 1.60 to 0.60.

4.4. CHARACTERIZATION OF IDENTIFIED MUTATIONS

A total of 13 SNPs were identified in *ND1* and *ND2*. Four of these SNPs have previously been characterized (A3480G, C3594T, G4580A, T4216C) (Howell *et al.,* 1993; 1996; 1998; Chen *et al.,* 1995; Thomas *et al.,* 1996; Torroni *et al.,* 1996; Macaulay *et al.,* 1999; Rose *et al*., 2001; Finnilă, 2001; Harrenstadt *et al.,* 2002; Wong *et al.,* 2004; Harris *et al.,* 2006).

The A3480G transition was characterized by Howell *et al.* (1993; 1996), Thomas *et al.* (1996) and Harris *et al.* (2006). In the present study, 7.5% of subfertile patients contained this mutation while 2.5% of fertile men had this mutation. Harris *et al.* (2006) found 8% of subfertile men contained this mutation, while for subfertile men it was only 1.4%. Thomas *et al*. (1996) studied 124 non-insulin dependent diabetes mellitus patients. He observed a population frequency of 1.22% with only one patient and with one control patient with this SNP. From all these studies it can be suggested that the subfertile population tend to show this SNP more than the normal fertile population.

A mutation at nt 4216 was characterized in different studies as a secondary mutation associated with LHON (Leber's hereditary optical neuropathy) (Howell *et al.,* 1995; 1998; Rose *et al*., 2001; Harrenstadt *et al.,* 2002; Wong *et al.,* 2004; also in the published polymorphism data

obtained at www.mitomap.org as of 18th may, 2006). Ruiz-Pesini *et al.* (2000) also utilized this polymorphic site as a marker to differentiate the T haplogroup which was present in 27% of Danes and Swedes population. Harrenstadt *et al.* (2002) showed that the occurrence of the highest number of polymorphism (46) in T haplogroup was at nt 4216. From this study it was found that 5% of subfertile men had the T4216C mutation. This percentage was just double what was found in normal men. Even though statistically it is not significant it has been established that this mutation was responsible for cause of LHON disease, so there might be a chance that this mutation may also affect human male fertility.

The silent mutation C3594T was also characterized by Chen *et al.* (1995); Bandelt *et al*. (2002) and Harrenstadt *et al.* (2002). Harrenstadt *et al.* (2002) studied 56 African human gene sequences and found the presence of C3594T polymorphism in L1 and L2 haplogroups. The occurrences of this polymorphism in these haplogroups were 13 and 23, respectively, which was very common. The L haplogroup is defined for African origin populations (Torroni *et al.,* 1996). Bandelt *et al*. (2002) and Chen *et al.* (1995) also observed this SNP in an African population. From the present study, only 2.5% normal men showed the presence of this mutation while no subfertile men had this SNP. The low rate of occurrence of this mutation suggested that this mutation might not be very common in the NZ Caucasian population. It is unlikely that it plays a role in causing poor semen parameters.

Another silent mutant at nt 4580 in the *ND2* gene was characterized by Torroni *et al.* (1996), Macaulay *et al.* (1999), Finnilǎ *et al.* (2001) and Harrenstadt *et al.* (2002). The polymorphism at G4580A determines the haplogroup V, one of the European haplogroups (Finnilǎ *et al.,* 2001). 44% of 192 Finnish Caucasian samples belonged to haplogroup V with

this mutation (Finnila *et al.,* 2001). Harrenstadt *et al.* (2002) stated that this mutant was previously established by Torroni *et al.* (1996) and Macaulay *et al.* (1999). These two studies were also done on European and Eurasian Caucasian populations. In the present study, 6.5% of fertile men and 4.3% of subfertile men contained this mutant. Therefore, this mutant might be common to NZ Caucasians but not significant to produce poor semen quality.

4.5. CHARACTERIZATION OF SYNONYMOUS CODON USAGE

Synonymous or silent mutations have traditionally been considered neutral as the sequence of the corresponding protein is unchanged and there is no phenotypic difference (Novella *et al.*, 2004), However, recently studies have shown that not every synonymous site is neutral (Chamary *et al*., 2005; Duret, 2002). Codon usage reflects a balance between selective and neutral evolutionary forces (Bulmer, 1991; Semon *et al.* 2006). Novella *et al*. (2004) proposed that protein expression, control, splicing, temperature stability, tRNA availability and resistance to tRNA interference provide increasing evidence of differences in the selective value of a particular codon. So far, most of the experimental studies have been done with synonymous mutants of nuclear genes. The mitochondrial and nuclear genomes employ similar mechanisms of translation, however their origin and genetic code, are not identical. According to the, MITOMAP data (www.mitomap.org) a high percentage of synonymous substitutions in coding region have been associated with diseases. For example, in *ND1*, 49 (60%) of the SNPs are synonymous and 21 (42%) are associated with disease include LHON, MELAS, breast cancer, type 2 diabetes, etc.

Synonymous mutations in the mtDNA genome are more common in subfertile men than in fertile men (Ruiz-Pesini *et al.,* 2000; Thangaraj *et al*., 2003; Harris *et al.,* 2006). Gomas (2002) found 8% of subfertile men had the A3480G mutation in the *ND1* gene. Similarly, from this study, 7.5% of subfertile men have this mutation. Their codon was changed in the third position from AAA to AAG with a reduced relative codon usage frequency (1.71 to 0.29). However, a few other mutations found in this study also showed a change from a lower relative codon usage to a high relative codon usage (eg. G4580A, 0.17 to 1.83). Synonymous mutations (C7024T in *COXI* and C7891T in *COXII*) which were identified as T polymorphic site were also found in *COX* gene in men with reduced sperm mobility (Ruiz-Pesini *et al.,* 2000).

The synonymous mutations found by Harris *et al*. (2006) also showed 58% (28 out of 48) of subfertile men showed reduced relative codon usage and only 33% (16 out of 48) with higher relative codon usage (Table 4.1). Four synonymous mutations out of 48 did not show any change in their relative codon usage (Table 4.1). Similarly, from this study, more synonymous mutations in subfertile groups were associated with reduced relative codon usage (Table 4.2). All these studies suggest that synonymous codons with low frequency rate are more frequent in subfertile men.

4.6. EFFECT OF IDENTIFIED MUTATIONS ON HUMAN MALE INFERTILITY

Mutations in mtDNA have been found to be associated with a large number of diseases in humans. More than 100 different point mutations have been related to different diseases (Wallace, 1999; Thangaraj *et al.,* 2003). It has been suggested that abnormal mitochondrial function might

cause male infertility. Mitochondrial DNA mutations were found to be responsible for sperm dysfunction (Spiropoulos *et al*., 2002). Mutations in mtDNA genome could alter the performance of ATP production in spermatozoa, which can lead to a decline of motility of sperm (Ruiz-Pesini *et al*., 2000). Reduced fertility may be associated with deletions of a few thousand base pairs of nucleotides in the sperm mitochondrial genome (Kao *et al.,* 1998).

Generally, if a mutation was solely responsible for a disease, it should be found in the patient with the disease, it should also be heteroplasmic and be able to affect the function of polypeptide in which the change occurs (Riordan-Eva *et al.,* 1995). None of the mutations from this study fulfilled the above rules. As some mutations were present in normal and subfertile men, these mutations indicate that they might not be associated with male infertility. Most of the mutations were neutral. Only two mutations (T4216C and C3992T) change the amino acid in the polypeptide sequence. But their hydrophobicity plot, polarity value and predicted secondary structure suggested that there was no affect in the function of the protein product. These identified SNPs did not directly affect the function of ND1 and ND2 protein but the mutant codon might have an effect on protein translation efficiency machinery (Holyoake *et al*., 2001). If the tRNA showed a codon preference then it might lead to a change in the levels of protein production, leading to a loss of efficiency in energy production (Holyoake *et al*., 2001). When mutations have an effect on translation efficiency, then they would be responsible for sperm dysfunction.

All the mutants in this study were homoplasmic. It is not known whether these mutations were transmitted maternally.

In the present study, statistical analysis showed that the Z values of all the mutations found in this study was less than the 5% value (z=1.96 at 5% level of significance). For example, the Z values in different mutations were: nt 3480 (z= 0.2236), nt 4216 (z= 0.129), nt 4580 (z= 0.932) and nt 4514 (z=0.1203). Therefore the observed differences in all the four cases are not significantly different between fertile and subfertile men.

However, it is evident that in some cases there might be an association between mitochondrial mutations and male infertility. But this study could only be stated that more mutations were found in subfertile men than in fertile men. In this study, the sample size was small and the study was only focused on the NZ Caucasian population. Based on a small study, it is difficult to establish the effect of mitochondrial mutations in male infertility.

4.7. COMPARISON WITH PREVIOUS RESEARCH

Eighty normozoospermic and 124 non-normozoospermic sample were analyzed by Harris *et al.* (2006) for nucleotide substitutions in mitochondrial genes by PCR and DNA sequencing. The entire region of the mtDNA genome except the cytochrome C and the D loop was analyzed. Several synonymous and non-synonymous mutants were found in different region of the mitochondrial genome. Sixty synonymous mutants of different fertility groups from the study of Harris *et al.* (2006) were compared with the human mitochondrial genetic codes and their frequencies of codon usage for each mitochondrial gene were determined (Table 3.3). The number of relative codon usage for each codon usage is tabulated in Table 4.1. Among the subfertile men, 28 of the 48 codon changes were from a high relative usage value to a low usage value.

Among the normozoospermic samples, 6 of the 12 codon changes were from a low to a high relative codon usage and 5 of the 12 change from high to a low value. More silent mutants were change from a high to a low relative codon usage value. Similarly, synonymous mutants from this study showed more mutations that change to a low relative codon usage value (Table 4.2).

Table 4.1. Number of mutations (silent) with reduced and higher codon usage rate against different fertility groups (Harris *et al*. 2006).

Classification of Mutants	Total number of mutants	Codon frequency change rate		
		Higher	Lower	No change
Normozoospermic	12	6	5	1
Asthenozoospermic	12	4	5	3
Oligozoospermic	15	4	11	-
Oligoasthenozoospermic	2	1	1	-
Severely Oligozoospermic	1	-	1	-
Severely Oligoasthenozoospermic	10	5	4	1
Azoospermic	8	2	6	-
Total number of non-normozoospermic	48	16	28	4

Six out of ten synonymous mutations (60%) in subfertile groups showed reduced relative codon usage value while, four out of ten synonymous mutations (40%) were with higher relative codon usage value (Table 4.2).

A total of 48 synonymous mutations were found in subfertile groups by Harris *et al*. (2006). For the calculation of Chi-square (χ^2) value between the lower and higher relative codon usage, a total number of non-normozoospermic synonymous mutations were 44, as 4 synonymous mutations did not show any change in their codon usage. All together including the results from Harris *et al*. (2006), a total of 34 (28+6) synonymous mutations out of 54 (44+10) subfertile samples (62%) showed a change to a lower relative codon usage value. Similarly, a total of 20 (16+4) synonymous mutations out of 54 (44+10) subfertile samples (37%) showed a change to higher relative codon usage values. Chi-square (χ^2) value (χ^2=0.067 with 1 d.f.; critical value = 3.841) shows that there is no significant difference at 5% level of significance between the lower and higher codon usage value of these synonymous mutations. These high frequencies of the synonymous mutations found in a population might be the result of a lack of selective pressure on the alleles. Although these mutations may provide a genetic background that renders the population more susceptible to influence by other mutations.

Table: 4.2. Mutations (silent) in *ND1* and *ND2* gene with reduced and higher codon usage rate against different fertility groups.

Classification of Mutants	Mutants	Codon frequency change rate		
		Higher	Lower	No change
Normozoospermic	T3396C	-	+	-
	T3594C	-	+	-
	G3666A	-	+	-
	G3693A	+	-	-
	A3480G*	-	+	-
	G4580A*	+	-	-
	A4514G*	-	+	-
Asthenozoospermic	A3480G*	-	+	-
	A4514G*	-	+	-
Oligozoospermic	G3591A	+	-	-
	A3480G*	-	+	-
	G4580A*	+	-	-
	C3567T	-	+	-
Oligoasthenozoospermic	C3450T	-	+	-
	G4580A*	+	-	-
Severely Oligozoospermic	A4024G	-	+	-
	G4580A*	+	-	-

* indicates that these mutants were found more than once in different fertility groups. To determine the Chi-square (χ^2) value, these mutants in subfertile groups were considered as different mutations.

Seven non-synonymous mutants from the research of Harris *et al*. (2006) were also analyzed. Their hydrophobic plot, polarity value and amino acid change were determined (Table 4.3). Change of amino acids due to

non-silent mutants suggested that codon usage was not relevant to semen quality in men.

Table: 4.3. Hydrophathicity and Polarity analysis of non-silent mutants.

Classification	***Gene***	***Sample***	***Mutation***	***Codon change***	***Amino acid change***	***Hydropathicity***		***Polarity***	
						Normal	**variant**	**Normal**	**variant**
As	*ND2*	Am433B	T5278C	TTC→TCC	Phe→Ser	-0.756	-1.156	9.256	9.700
AzAs	*ND6*	4449 AT665c 6169	C14213T	ACC→AAT	Thr→Asn	1.833	1.489	6.600	6.889
OAsAs	*ND6*	4850 5004 5360 3099 5310	C14167T	CTC→TTC	Leu→Phe	1.600	1.411	7.167	7.167
AsOAs	ND6	RL506B 5585	T14182C	TAC→CAC	Tyr→His	2.322	1.467	6.756	7.333
O SOAsN	*ND4*	5378 JN TM 2905	C11724T	ACA→ATA	Tre→Met	0.811	1.100	7.689	7.367
O	*ND5*	5378	T14115A	TTC→TAC	Phe→Tyr	2.433	1.978	5.856	5.967
O	*ND5*	2565	T13740C	CTA→CCA	Leu→Pro	1.489	0.889	7.444	7.789

Here, As= Asthenozoospermic, Az= Azoospermic, OAs= Oligoasthenozoospermic, SOAs= Severely oligozoospermic, O= Oligozoospermic and N= Normozoospermic samples.

Most of the non-synonymous mutants were hydrophobic (Kyte and Dolittle, 1982) and nonpolar. One mutation in the ND2 protein had a changed amino acid from phenylalanine, a nonpolar hydrophilic amino

acid to serine, a polar negatively charged amino acid. This mutant might have an affect on the ND2 polypeptide. The secondary structure of this mutant protein is also varied, which gave an idea that there was a change in ND2 polypeptide (Gomas, 2002).

4.8. FUTURE DIRECTIONS

Mutations were detected in the *ND1* and *ND2* of the mitochondrial genome by automated DNA sequencing. Most of the previous studies used SSCP analysis for preliminary mutation screening. The suspected mutations were confirmed by DNA sequencing. However, positives and negatives were detected frequently. In the present study, PCR amplification of the genes of interest and DNA sequencing was employed throughout. In order to determine the relevance of various mutations the mitochondrial genome in relation to semen quality, biochemical analyses must be carried out to establish the stability of these mutant proteins and to determine the function of ND1 and ND2 polypeptides. Detailed analysis of possible influence of different codons on translation efficiency should be performed.

4.9. CONCLUSIONS

This study has identified 13 SNPs in the *ND1* and *ND2* genes. Eleven of these were synonymous and two were non-synonymous. Four of them were previously characterized and all of them were homoplasmic. None of the identified SNPs are likely to affect the function of polypeptides on the basis of the hydrophobicity plots or secondary structure predictions. More mutations were synonymous. A higher frequency of SNPs in *ND1* and *ND2* were found in subfertile men then in fertile men, a trend consistent with previous studies of other mitochondrial genes. Combined results obtained from Harris *et al.* (2006) and the present study showed that more synonymous mutations (62%) in subferile groups changed their codon usage value from a high to a low relative codon usage which was not significantly different (χ^2= 0.067 with 1 d.f.) at 5 % level of significance. However, there were some mutations (37%) in subfertile samples that changed the relative codon usage value from low to high. Thus, change in codon usage was not related to semen quality in men. Statistical analysis also revealed insignificant differences in observed differences in the frequencies of SNPs of fertile and subfertile men. However, this small study included only a NZ Caucasian population; an extended population survey may provide insights into the role of SNPs in relation to male infertility.

REFERENCES

Affara N. A. (2001). **The role of the Y chromosome in male infertility**. *Expert Reviews in Molecular Medicine.* ISSN 1462-3994: 1-16. *Cambridge University press.*

Akashi H. (1995). **Inferring weak selection from patterns of polymorphism and divergence at silent sites in *Drosophila* DNA**. *Genetics*. 139. 1067-1076.

Attardi G. and Schatz G. (1988). **Biogenesis of mitochondria.** *Annual Reviews Cell Biology.* 4: 289-333.

Anderson S, Bankier A. T, Barrell B. G, de Bruijn M. H, Coulson A. R, Drouin J, Eperon I. C, Nierlich D. P, Roe B. A, Sanger F, Schreier P. H, Smith A. J, Staden R, Young I. G. (1981).**Sequence and organization of the human mitochondrial genome.** *Nature* 290: 457-465.

Bandelt H. J., Alves-silva. J., Guimarease P.E., Santon M. S., Brehm A., Pereira L., Coppa A., Larruga J.M., Scozzari R (2001). **Phylography of human mitochondrial haplogroup L3e: a snapshot of African prehistory and atlantic slave trade.** *Annals of Human Genetics*. 65(6): 549-563.

Brandstatter A., Parsons T. J., Parsons W. (2003). **Rapid screening of mtDNA coding region SNPs for identification of west European Caucasian haplogroups.** *International Journal of Legal Medical.* 117: 291-298.

Bogenhagen D. F (1999). **Repair of mitochondrial DNA**. *American Journal of Human Genetics*. 64:1276-1281.

Brinton-Jones C, Haines C.J. (2000). **Microdeletions on the long arm of the Y chromosome and their association with male-factor infertility.** *Hong Kong Medical Journal.* 6:184-9.

Brown M. D., Torroni A., Reckord C. L. and Wallace D. C. (1995). **Phylogenetic analysis of Leber's hereditary optical neuropathy mitochondrial DNA's indicates multiple independent occurances of common mutations.** *Human mutation*. 6(4). 311-325.

Brown T. A. (1995). **Gene cloning an Introduction.** Third Edition. *Stanley Thornes (Publisher) ltd.*

Bulmer M. (1991). **The selection-mutation-drift theory of synonymous codon usage.** *Genetics*. 129: 897-907.

Carlini D. B. and Stephan W. (2003**). In *Vivo* introduction of unpreferred synonymous codons into the *drosophila Adh* genes results in reduced levels of ADH protein.** *Genetics.* 163. 239-243.

Chamary J. V. and Hurst J. V. (2005).**Evidence of selection of synonymous mutations affecting stability of mRNA secondary structures in mammals**. *Genome Biology*. 6(9): Art no. R57.

Chen Y. S., Torroni A., Excofier L., Santachira A. S., Wallace. D. C. (1995). **Analysis of mtDNA variation in african populations reveals the most ancient of all human continent-specific haplogroups.** *American Journal ofHhuman Genetics*. 57(1): 133-149.

Copeland. W.C., Paromarev. M. V, Nguyen. D. Kunket. T.A. and Longley. M. J. (2002) **Mutation in DNA polymerase gamma cause error prone DNA synthesis in human mitochondrial disorder.** *Acta Biochemica Polonica* (50): 155-167.

Cummins J. M., Jequier A. M. and Kan R. (1994). **Molecular biology of human male infertility: links with aging, mitochondrial genetics, and oxidative stress?** *Molecular Reproduction and Development.* 37: 345-362.

Cummins J. M., Wakayama. Y. and Yanagimachi. R.(1997). **Fate of microinjected sperm components in the mouse oocyte and embryo.** *Zygote*. 5(4): 301-308.

Cummins J. M., Jaquier A. M., Martin R., Mehmet D., Goldblatt J. (1998). **Semen levels of mitochondrial DNA deletions in men attending an infertility clinic do not correlate with phenotype.** *International Journal of Andrology*. 21(1): 47-52.

Cummins J. M., Mehmet D., Kishikawa H. and Yanagimachi. R. (1999). **Fate of genetically marked mitochondrial DNA from spermatocytes microinjected into mouse zygotes**. *Zygote*. 7(2): 151-156.

Cummins J. M. (2001**). Mitochondria: potential roles in embryogenesis and nucleocytoplasmic transfer**. *Human Reproduction Update* 7: 217-228.

Darley-Usmar V., Ragan I., Smith P. and Wilson M. (1994). **The proteins of the mitochondrial inner membrane and their role in**

oxidative phosphorylation. In Mitochondria: DNA, Proteins and Disease. 1-25. *Eds Darley-Usmar, V. and Schapira, A. H. V. Portland Press Ltd. London.*

Daniel W. W. (1999). **Biostatistics: Foundation for analysis in the health sciences**. Seventh edition. *John Wiley and sons, Inc. NY.*

De Kretser D. M and Baker H. W. G. (1999). **Infertility in men: Recent advances and continuing controversies.** *The Journal of Clinical Endocrinology & Metabolism*. 84 (10): 3443-3450.

DiMauro S and Schon E. A. (2001). **Mitochondrial DNA mutations in Human disease**. *American Journal of Medical Genetics* (Semin. Med. Genet.) 106: 18-26.
DiMauro S and Schon, E. A. (2003). **Mitochondrial respiratory chain disease.** *The New England Journal of Medicine.* 348: 2656-2668.

Duan J., Wainwright M. S., Joseph M., Saitou N., Sanders A. R., Gelernter J and Gejman P. V. (2003). **Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor.** *Human Molecular Genetics.* 12(3):205-216.

Durent L. (2002). **Evolution of synonymous codon usage in metazoans.** *Current opinion in genetics & development.* 12(6):640-649.

Erasmuson T., Sin, I. L. and Sin F. Y. T. (2003) **Absence of association of androgen receptor trinucleotide expansion and poor semen quality**. *International Journal of Andrology*. 26, 46-51.

Finnila¨S., Lehtonen M. S., and Majamaa K (2001). **Phylogenetic Network for European mtDNA.** *American Journal of Human Genetics*. 68:1475–1484.

Gradner E. J., Simmons. M. J. and Snustad D. P. (1991). **Principles of genetics**. Eighth Edition. *John Wiley and sons, Inc. NY.*

Grossman L. I. (1995). **Mitochondrial mutations and human disease.** *EnvironmentalMolecular Mutagenesis.* 25 (Supple 26): 30-37.

Goto Y., Tsugane K., Tanabe Y., Nonaka I. and Horai S. (1994). **A new point mutation at nucelotide pair 3291 of the tRNALeu(UUR) gene in a patient with mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS).** *Biochemical and Biophysical Research Communications*. 202(3): 1624-1630.

Gomas K. (2002).**The association of single nucleotide polymorphism in mitochondrial *ND1* and *ND2* genes with human male fertility potential**. *Unpublished observation*.

Grantham R (1974). **Amino acid difference formula to help explain protein evolution.** *Science.* 185 (4154): 862-864.

Green D. R and Reed J. C. (1998). **Mitochondria and Apoptosis**. *Science.* 281: 1309-1312.

Hatefi Y. (1985).**The mitochondrial electron transport and oxidative phosphorylation system**. *Annual review of Biochemistry.* 54. 1015-1069.

Hargreave T. B. (2000). **Genetic basis of male infertility**. *British Medical Bulletin*. 56. 650-671.

Harris T. P., Gomas K. P., Weir F., Holyoake A. J., McHugn.P., Wu. M, Sin A., Sin I. L., and Sin F. Y. T. (2005). **Molecular analysis of polγ and mitochondrial polymorphism in fertile and subfertile man**. *International Journal of Andrology*. 29 (3): 421-432.

Herrnstadt C., Elson J. L., Fahy E., Preston G., Turnbull D. M., Ghosh S.S., Olefsky J. M., Beal F., Davis R. E. and Howell N. (2002).**Reduced-median-network analysis of complete mitochondrisl DNA coding-region sequences for major African, asian, and European haplogroups.** *American Journal of Human Genetics*. 70. 1152-1171.

Holstein A., Schulze W. and Davidoff M. (2003). **Understanding spermatogenesis is a prerequisite for treatment**. *Reproductive Biology and Endocrinology.* 1(1): 107.

Holt I. J. and Jacobs H. T. (1994). **Structure and expression of mitochondrial genome**. *In Mitochondria: DNA, Proteins and Disease*. *Eds Darley-Usmar, V. and Schapira, A. H. V. Portland Press Ltd. London.*

Holyoake A. J., Sin I. L., Benny P. S. and Sin F. Y. T. (1999). **Association of a novel human mtDNA ATPase6 mutation with immature sperm cells.** *Andrologia* 31(6): 339-345

Holyoake A. J., McHugh P., Wu M, O'Carroll S., Benny P., Sin I. L. And Sin F. Y. T. (2001). **High incidence of single nucleotide substitutions in the mitochondrial genome is associated with poor semen parameters in men**. *International Journal of Andrology*. 24: 175-182.

Houshmand M. (2003). **Mitochondrial DNA mutations, pathogenicity and inheritance**. *Iranian Journal of Biotechnology*. 1(1): 1-18.

Howell N., Kubacka I., Halvoronson S., Howell B., Mccullough D.A. (1993). **Leber's hereditary optic neurophathy: the etiological role of mutation in the mitochondrial cytochrome b gene.** *Genetics.* 133(1): 133-136.

Howell N., Kubacka I., Howell B., Mccullough D.A. (1995). **Phylogenetic analysis of mitochondrial genome from leber's hereditary optic neuropathy pedigrees**. *Genetics*.140 (1) 285-302.

Howell N., Bogolin C., Jamieson R., Marenda D.R., Mackey D.A. (1998). **MtDNA that cause optic neurophathy: how do we know?** *American Journal of Human Genetics.* 62(1): 196-202.

Hyuanh T., Mollard R. and Trounson A.(2002). **Selective genetic factors associated with male infertility.** *Human Reproduction Update*. 8. 183-198.

Jakupciak J. P., Wang W., Markowitz M. E., Ally D., Coble M., Srivastava S., Maitra A., Barker P. E., Sidransky D. and O'Connell C. D. (2005). **Mitochondrial DNA as a cancer biomarker.** *The Journal of Molecular Diagnostics*. 7(2):258-67.

Kao. S, Chao. H. T and Wei. Y. H (1995). **Mitochondrial deoxyrido nucleic acid 4977 bp deletion is associated with diminished fertility and motility of human Sperm** *Biology of Reproduction.* 52 (4) : 729-736

Kao S. H., Chao H. T. and Wei Y.H. (1998). **Multiple deletions of mitochondrial DNA are associated with the decline of motility and fertility of human spermatozoa**. *Molecular Human Reproduction.* 4: 657-666.

Klug W. S. Cummings M. R. (1997). **Concepts of Genetics.** Fifth Edition. *Prentice-Hall, Inc. New Jersey.*

Knight R. D., Freeland S. H. and Landweber L. F. (2005). **A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes.** *Genome Biology.* 2(4): research 0010.1– research0010.13.

Kumar S., Tamura K., Jakobsen I. B., and Nei M. (2001) MEGA2: **Molecular Evolutionary Genetics Analysis software**. *Bioinformatics.*

Kunkel T.A. and Soni. A (1988) **Exonucleolytic proofreading enhances the fidelity of DNA synthesis by chick embryo DNA polymerase gamma.** *Journal of Biological Chemistry.* 263: 4450-59.

Kunkel T.A. and Mosbangh D.W. (1989) Exonucleolytic **proofreading by mammalian DNA polymerases** Biochemistry. 28: 988-995.

Kyte J. and Doolittle. R. F. (1982). **A Simple Method for displaying the hydropathical character of a protein.** *Journal of Molecular Biology.* 157(1): 105-132.

Leonard J. V. and Scapira A. H. (2000). **Mitochondrial respiratory chain disorder II: Neurodegenerative disorders and nuclear gene defects.** *Lancet*. 355(9201): 389-394.

Loeb L. A., Wallace. D. C. and Marta. G. M. (2005).**The mitochondrial theory of aging and relationship to reactive oxygen spercies damage and somatic mtDNA mutations**. *Proceedings of National Academy of Science. USA.* 102 (52): 18769-18770.

Macaulay V., Richards M., Hickey E., Vega E., Cruciani F.,Guida V., Scozzari R., Bonne B and Torroni A. (1999). **The Emerging Tree of West Eurasian mtDNAs: A Synthesis of Control-Region Sequences and RFLPs**. *American Journal of Human Genetics.* 64:232–249.

Moore F. L. and Reijo-Pera R. A. (2000). **Male sperm motility dictated by mother's mtDNA.** *American Journal of Human Genetics*. 67: 543-548.

Namiki M. (2001) **Genetic aspects of male infertility.** *World Journal of Surgery*. 24(10): 1176-1179.

Novella I.S., Zarate S., Metzgae D., and Ebendick-corpus B. E. (2004). **Positive selection of synonymous mutations in vesicular stomatitis virus.** *Journal of Molecular Biology.* 342. 1415-1421.

Purves W. K., Orians G. H., Craig H. and Sadava D. (2004). **Life: The science of biology.** Seventh edition. *Sinauer Associates Inc. Massachuesetts*.

Reha S and Robinson B. (2001). **Mitochondria, Oxygen free radicals, and apoptosis.** *American Journal of Medical Genetics (Semin. Med. Genet.)* 106: 62-70.

Riordan-Eva P. and Harding A. E. (1995). **Leber's hereditary optic neuropathy: The clinical relevance of different mitochondrial mutations.** *Journal of Medical Genetics*. 32(2): 81-87.

Rose G., Passarino G., carrieni G., Altomare K., Bertolini G., and Benedictis G. (2001). **Paradoxes in longevity: sequence analysis of mtDNA haplogroup J in centenarians**. *EuropeanJjournal of Human Genetics*. 9(9). 701-707.

Ruiz-Pesini E., Lapena, A.C., Diez-Sanchez C., Perez-Martos A., Montoya J., Alvarez E., Diaz M., Urries A., Montoro L., Lopez-Perez M.J. and Enriquez J.A. (2000). **Human mtDNA haplogroups associated with high or reduced spermatozoa motility.** *American Journal of Human Genetics.* 67: 682-696.

Sambrook. J., Fritsch. E. F. and Maniatis. T (1989). **Molecular cloning. A Labrotory Manual.** Second edition*. Cold Spring Harbor laboratory press. Cold Spring Harbor, New York.*

Saraste M. (1999). **Oxidative phosphorylation at the fin de siecle**. *Science.* 283(5407): 1488-1493.

Schafer K., Seelert H. Reifschneider N. H., Krause. F., Dencher N. A. and Vonck J. (2006). **Architecture of active mammalian respiratory super complexes.** *The Journal of Biological Chemistry.* 281(22): 15370-15375.

Schon E. A (2000). **Mitochondrial genetics and disease**. *Trends in Biochemical Sciences*. 25: 555-560.

Seshagiri P. B. (2001). **Molecular insights into the causes of male infertility**. *Journal of Biosciences*. 26(4): 429-435.

Semon M., lobry J. R., and Duret L. (2006).**No evidence for tissue-specific adaptation of synonymous codon usage in humans.** *Molecular Biology of Evolution*. 23(3): 523-529.

Sharma R. K., Pasqualotto F. F., Nelson D. R., Anthony J., Thomas Jr. and Agaewal A. (1999). **The reactive oxygen species-total antioxidant capacity score is a new measure of oxidative stress to predict male infertility.** *Human Reproduction*. 14(11): 2801-2807.

Shanske A. L, Shanske S. and DiMauro S. (2001). **The other human genome**. *Archives of Pediatrics and Adolescent Medicine.* 155: 1210-1217.

Sharp P. M. (1989). **Evolution at silent sites in DNA and animal breeding**. *Reviews in Molecular and Quantitative Approaches*. 23-32.

Shoffner J. M and Wallace D. C. (1990).**Oxidative phosphorylation diseases. Disorders of two genomes.** *In Advances in Human Genetics.*19. 267-330.

Singh A. R., Vrtel R., Vodicka. R., Dhaifalah I., Konovalinka D. and Santavy J. (2006). **Genetic factors in male infertility and their implications.** *International journal of human genetics*. 6(2): 163-169.

Smith N.G. and Eyre-Walker A. (2001) **Why are translationally sub-optimal synonymous codons used in *Escherichia coli*?** *Journal of Molecular Evolution* **53**. 225–236.

Smith R., Kaune H., Parodi D., Madariga. M., Rios R., Morales I. and Castro A. (2006). **Increased sperm DNA damage in patients with vericocele: relationship with seminal oxidative stress.** *Human Reproduction*. 21(4): 986-993.

Snustad D. P. and Simmons M. J. (2000). **Principles of genetics.** Second edition*. John Wiley and sons, Inc. NY.*

Spellbrink J.N, Toivenon J.M Kurkala J.M. Cooper. H.M, Lehtinen S. K .etal (2000). **In Vivo functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cell.** *Journal of Biological Chemistry* 275: 24818-28.

Spiropolous J., Turnbull D. M. and Chinnery P. F. (2002).**Can mitochondrial DNA mutations cause sperm dysfunction?** *Molecular Human Reproduction*. 8(8): 719-721.

St. John J., Sakkas D. and Baratt C. L. R. (2000). **A role of mitochondrial DNA and sperm survival.** *Andrology*. 21: 189-199.

Thangaraj K., Joshi M. B., Reddy A. A., and Singh L. (2003). **Sperm mitochondrial mutations as a cause of low sperm motility**. *Journal of Andrology.* 24. 388-392

Thomas. A. W., Edward A., Sherratt E. J., Majid. A., Alcardo. J. C. (1996). **Molecular scanning of candidates mitochondrial tRNA genes in type 2 (non-insulin dependent) diabetes mellitus.** *Journal of Medical Genetics*. 33(3): 253-255.

Torroni A., Huoponen K., Francalacci P., Petrozzi M., Morelli L., scozzari R., Obinu D., Savontaus M. and Wallace. D. C. (1996). **Classification of European mtDNAs from an analysis of three European populations.** *Genetics*. 144: 1835-1850.

Torroni A., Campos Y., Rengo C., Sellotto D., Achilli A., Magri C., Semino O., Garcia A., Jara P., Arenas J. and Scozzari R. (2003). **Mitochondrial DNA haplogroups do not play a role in the variable phenotypic presentation of the A3243G mutation.** *American Journal of human Genetics.* 72: 1005-1012.

Triepels R. H. Van Den Heuvel L. P., Trijbels J. M. and Smeitink J. A. (2001). **Respiratory chain complex I deficiency.** *American Journal of Medical Genetics*. 62(2): 262-268.

Venter J. C., Adams. M. D., Myers. E. W., Li P. W., Mural R. J. Smith H. O., Yendell M. and Holt R. A. *et al*.(2001). **The sequence of the human genome.** *Science*. 291. 1304-1351.

Vogt P. H. (1996). **Molecular basis of male (in)fertility.** *Advances in Andrology*. ESHRE Campus Course, March, 1996, Freiburgi. Britain. *International Journal Of Andrology*., 20(3): 2-10.

Wallace D.C. (1992).**Diseases of mitochondrial DNA**. *Annual Review of Biochemistry*. 61: 1175-1212.

Wallace D.C. (1994). **Mitochondrial DNA sequence variation in human evolution and disease.** *Proceedings of National academy of Science, USA.* 91(19): 8739-46

Wallace D. C., Singh. G., Lott M.T., Hodge, J. A., Schurr T.G., Lezza A. M., Elsas L. J. and Nikoskelainen, E. K. (1988**). Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy.** *Science.* 242(4884): 1427-1430.

Wallace D.C. (1999). **Mitochondrial diseases in man and mouse**. *Science.* 283(5407): 1482-1488.

Wallace D. C., Brown M. D. and Lott M.T. (1999). **Mitochondrial DNA variation in human evolution and disease.** *Gene.* 238. 211-230.

Wallace, D. C. (2001). **Mouse models for mitochondrial disease**. *American Journal of Medical Genetics (Semin. Med. Genet.)* 106: 71-93.

Wallace D. C. (2005).**The mitochondrial genome in human adaptive radiation and disease: On the road to therapeutics and performance enhancement**. *Gene.* 354: 169-180.

Walpole R. E. (1982). **Introduction to statistics.** 3rd edition. *Macmillan Publishing Co.,Inc.,New York, Collier Macmillan Pub., London.*

WHO (1999). **Laboratory Manual for the Examination of Human Semen & Sperm-cervical Mucus Interactions** *Cambridge University Press, Cambridge.*

Wong L. C., Chen T. and Tan D (2004)**. Detection of mitochondrial DNA mutations using temporal temperature gradient gel electrophoresis.** *Electrophoresis.* 25 (15): 2602 – 2610.

APPENDIX 1: ABBREVIATIONS

A	Adenine or adenosine; alanine	**nt**	Nucleotide
A_{260}	Absorbance at 260 nm	**Nacl**	Sodium chloride
A_{280}	Absorbance at 280 nm	**PCR**	Polymerase chain reaction
ATP	Adenosine 5′ triphosphate	**s**	seconds
bp	Base pair	**SDS**	Sodium dodecyl sulfate
C	Cytosine or cytidine; cystine	**ss**	Single stranded
dATP	Deoxyadenosine triphosohate	**T**	Thymine or thymidine; threonine
dCTP	Deoxycytidine triphosphate	***Taq***	*Thermus aquaticus* DNA
dGTP	Deoxyguanosine triphosphate	**TBE**	Tris borate electrophoresis buffer
dNTP	Deoxyribonucleotiside triphosphate	**TEMED**	Tetramethylethyldiamine
dTTP	Deoxythymidine triphosphate	**T_m**	Melting temperature
ddH_2O	Deionized distilled water	**Tris**	Tris (hydroxymethyl) aminomethane
DNA	Deoxyribonucleotide	**U**	unit
ds	Double stranded	**UV**	Ultraviolet
DTT	Dithiothreiotol	**V**	Volt
EDTA	Ethylenediaminetetra acetic acid	**W**	Watts
G	Guanine or guanosine; glycine	**g**	Grams
h	Hours	**l**	Liters
kDa	Kilo Dalton	**m**	meters
M	Molar		
mA	Milliamps		
min	Minutes		
$MgCl_2$	Magnesium chloride		

APPENDIX 2: CHEMICAL SUPPLIERS

PRODUCT	SUPPLIER	PRODUCT	SUPPLIER
Agarose	Invitrogen	SDS	Invitrogen
Ammonium acetate	BDH	Sodium carbonate	Scharlau (Scientific Supplies ltd.)
Big Dye^{TM} termination V 3.1 cycle sequencing kit	Applied Bio System	Sodium hydroxide	Scharlau (Scientific Supplies ltd.)
Chloroform	BDH	Sodium cloride	Scharlau (Scientific Supplies ltd.)
Deoxynucleotidetriphosphates	Eppendrof	TEMED	Sigma
Ethanol	BDH	Tris base	Roche
Ehilium Bromide	Sigma	*Taq* Plymerase	Roche
EDTA	Scharlau (Scientific Supplies ltd.)		
Formaldehyde	BDH		
Gel band purification kit	Amersham Biosciences		
Isoamyl alcohol	BDH		
Isopropanol	BDH		
Magnesium acetate	BDH		
Orange G	Sigma		
Phenol	Scharlau (Scientific Supplies ltd.)		
Proteinase K	Roche		
Sodium acetate	BDH		

APPENDIX 3: SEMEN SAMPLES

Normozoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and Sequencing*	
			ND1	*ND2*
1751	76	54	+	+
1929	63	77	+	+
2151	43	72	+	+
2169	55	55	+	+
2223	57	56	+	+
2430	265	52	+	+
2446	87	67	+	+
2450	113	70	+	+
2521	108	72	+	+
2551	70	55	-	-
2595	96	53	-	-
2619	61	54	-	-
2407	95	54	+	+
2752	109	61	+	+
2945	110	61	-	-
3999	348	66	+	+
4038	56	53	-	-
4510	51.3	74	+	+
4937	112	67	+	+
5697	69	70	+	+
5390	131	67	+	+
5769	64	78	-	-
6570	77	68	+	+
6635	99	80	+	+
6732	89	65	-	-
4948	104	77	-	-
6357	58	72	+	+
6661	66	71	-	-
6682	26	88	+	+
6750	75	79	+	+
7028	141	61	+	+
3945	57	79	+	+
5250	30	56	-	-
7009	105	52	+	+
6619	109	74	+	+
AM 257b	55	66	+	+

Sample	Sperm count (M/ml)	Motility %	*Used for PCR and Sequencing*	
			ND1	*ND2*
QD201B	67	60	+	+
QN222J	44	72	-	-
SF 94-97	86	72	+	+
SF 96-220	67	47	-	+
SF 96-229	75	50	+	+
SF 95-79	107	53	-	+
SF 148	27	52	+	-
010136	67	71	+	+
009807	61	67	-	-
009810	74	57	+	+
010381	138	60	-	-
010294	313	60	-	-
010004	60	66	-	+
010641	72	67	-	+
009820	150	58	-	-
010526	109	63.6	+	+
009940	96	49	-	-
010772	108	62	-	-
010707	56.5	55	-	+
010797	62	57	-	-
010830	98	37	-	-
010745	127	89	+	+
010286	141	57	+	+
010558	111	229	+	+
009949	66	62	+	+
019249	79	83	+	+
009912	110	19	-	-
010032	110	31	-	-
009950	133	50	-	-
010495	98	64	+	+
010525	49	23	-	-
010557	140	66	-	-
010388	109	53	-	-
010653	79	66	-	-
2660	55	70	-	-
010573	93	62	+	+
010081	109	53	+	+
6383	70	56	+	+
AT 635c	148	80	-	-

Asthenozoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and sequencing*	
			ND1	*ND2*
1874	46	27	+	+
2409	55	28	+	+
2651	36	37	+	+
2655	52	41	-	-
2660	72	26	+	+
2675	32	32	+	+
2758	67	38	-	-
2806	42	48	-	-
2902	32	35	+	+
3312	74	42	-	-
4932	176	35	+	+
5453	65	33	-	-
5937	34	74	-	-
5097	65	37	+	+
4170	27	28	+	+
6740	81	40	-	-
5535	65	33	+	+
6112	21	38	+	+
94-149	34	40	-	-
RS	36	23	+	+
SF 94-104	21	74	-	-
2261	71	23	+	+
0109393	23	56	+	+
009884	17	71	+	+
010449	49	23	+	+
010749	71	46	+	+
010324	54	27	+	+
010140	33	41	-	-
010471	21	44	-	-
009831	32	40	-	-
010447	21	43	+	+
010703	524	27	+	+
010620	30	40	+	+
010480	70	24	+	+
010749	71	46	-	-
010036	36	60	-	-
010525	45	36	+	+
010545	74	19	+	+
010439	44	39	-	-
010703	72	32	+	+
010527	71	46	+	+

010756	21	44	+	+
2495	24	31	+	+

Oligozoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and sequencing*	
			ND1	*ND2*
2703	13.4	42	-	-
6768	3.8	38	+	+
2902	3.2	32	+	+
1699	10.70	67	-	-
6994	14	57	-	+
SF 94-99	4	50	-	-
010830	4.6	15	+	+
3535	11	51	+	+
QR497E	15	57	-	-
010208	4.6	22	+	+

Oligoasthenozoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and sequencing*	
			ND1	*ND2*
5292	15.5	26	-	-
5593	13	18	-	-
6471	6.7	47	-	-
6445	9.3	39	+	+
5916	9.3	39	-	-
AT 665c	17	26	-	+
AM470T	10.6	15	-	-
020527	5.4	23	-	+
010406	8.9	60	+	+
1754	15.1	4	-	-
009593	21	50	+	+

Severely Oligozoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and sequencing*	
			ND1	*ND2*
010596	4.6	70	+	+
5293	1.41	52	-	-
6344	3	72	-	-
010716	0.7	34	-	-
009884	0.7	51	-	+
010639	1	4	+	+

Severely Oligoasthenozoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and sequencing*	
			ND1	*ND2*
010583	3	1	-	-
5080	1	8	-	+
6745	3.8	8	+	+
2598	2.3	8	-	-
010494	3.2	5	-	-

Azoospermic group:

Sample	*Sperm count (M/ml)*	*Motility %*	*Used for PCR and sequencing*	
			ND1	*ND2*
2762	0	0	-	+
010729	0	0	+	+
4071	0	0	-	-

Samples with identified mutations are shown in **bold** format.

APPENDIX 4

MITOMAP: Reported Mitochondrial DNA Base Substitution Diseases: Coding and Control Region Point Mutation

Last edited May 18, 2006

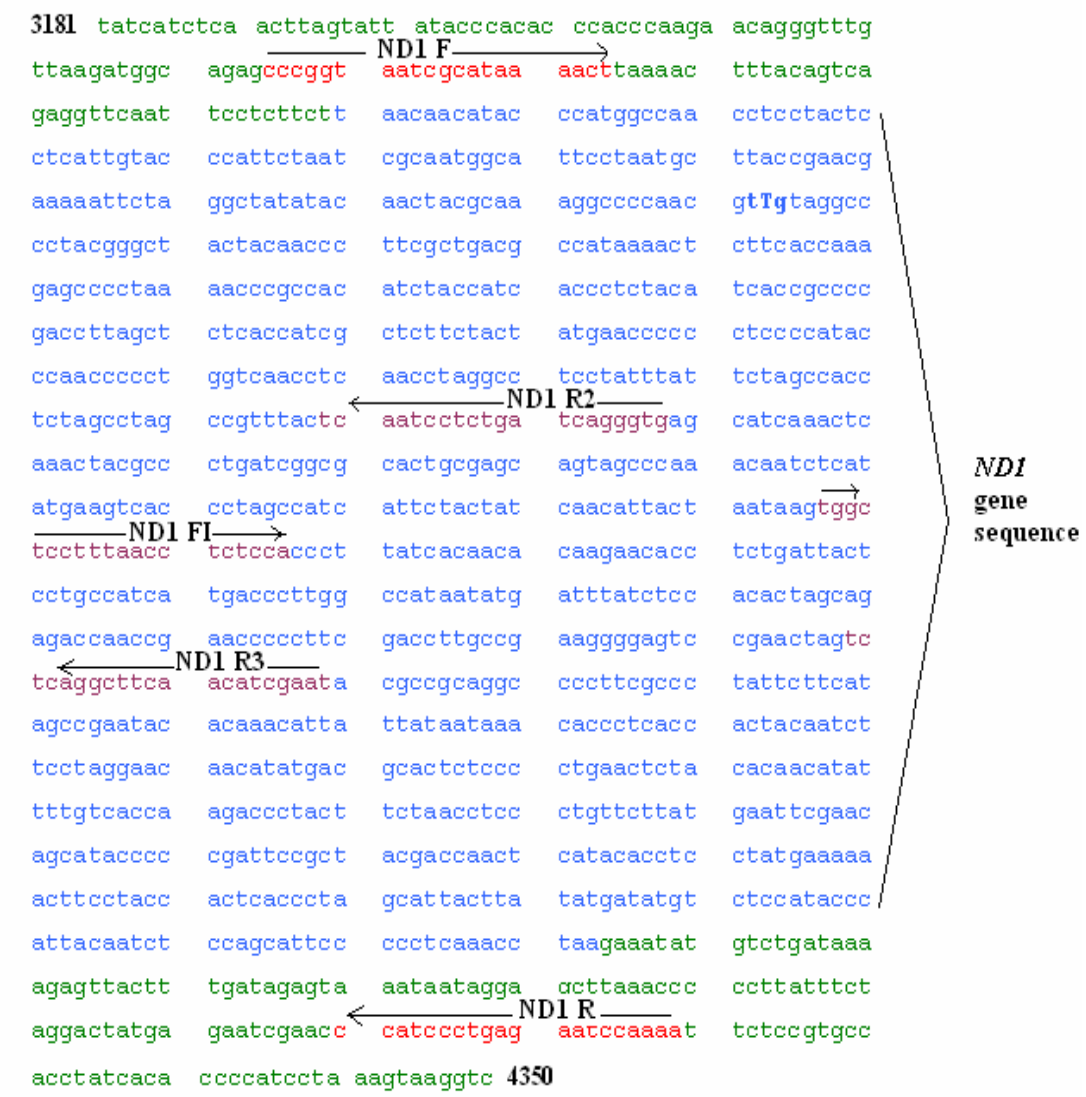
Locus	*Disease*	*Allele*	*Nucleotide Position*	*Nucldotide Change*	*Amino Acid Change*	*Homo-plasmy*	*Hetero-plasmy*
MT-ND1	MELAS	T3308C	3308	T-C	M-T	-	+
MT-ND1	NIDDM; LHON; PEO	G3316A	3316	G-A	A-T	+	-
MT-ND1	LHON/MELAS overlap	G3376A	3376	G-A	E-K	+	+
MT-ND1	LHON; NIDDM; CPT deficiency	T3394C	3394	T-C	Y-H	+	-
MT-ND1	ADPD	A3397G	3397	A-G	M-V	+	-
MT-ND1	LHON	G3460A	3460	G-A	A-T	+	+
MT-ND1	LHON	G3496T	3496	G-T	A-S	+	-
MT-ND1	LHON	C3497T	3497	C-T	A-V	+	-
MT-ND1	LHON	G3635A	3635	G-A	S-N	+	-
MT-ND1	MELAS	G3697A	3697	G-A	G-S	-	+
MT-ND1	LHON	G3700A	3700	G-A	A-T	+	-
MT-ND1	LHON	G3733A	3733	G-A	E-K	+	+
MT-ND1	Adult-Onset Dystonia	A3796G	3796	A-G	T-A	-	+
MT-ND1	MELAS	G3946A	3946	G-A	E-K	+	+
MT-ND1	MELAS	T3949C	3949	T-C	Y-H	-	+
MT-ND1	LHON	A4136G	4136	A-G	Y-C	+	-

MT-ND1	LHON	T4160C	4160	T-C	L-P	+	-
MT-ND1	LHON	C4171A	4171	C-A	L-M	+	+
MT-ND1	LHON	T4216C	4216	T-C	Y-H	+	-
MT-ND2	LHON	C4640A	4640	C-A	I-M	+	-
MT-ND2	LHON	A4917G	4917	A-G	N-D	+	-
MT-ND2	Longevity	C5178A	5178	C-A	L-M	+	-
MT-ND2	LHON	G5244A	5244	G-A	G-S	-	+
MT-ND2	AD; PD	G5460A	5460	G-A	A-T	+	+
MT-ND2	AD	G5460T	5460	G-T	A-S	+	+
MT-CO1	Prostate Cancer	C5911T	5911	C-T	A-V	+	-
MT-CO1	Prostate Cancer	G5913A	5913	G-A	D-N	+	-
MT-CO1	Myoglobinuria; Exercise Intolerance	G5920A	5920	G-A	W-Ter	-	+
MT-CO1	Prostate Cancer	A5935G	5935	A-G	N-S	+	-
MT-CO1	Prostate Cancer	G5973A	5973	G-A	A-T	+	-
MT-CO1	Motor Neuron Disease	6020del5	6020	5bp deletion-	ACLGG-AGPATer	-	+
MT-CO1	Prostate Cancer	G6081A	6081	G-A	A-T	+	-
MT-CO1	Prostate Cancer	G6150A	6150	G-A	V-I	+	-
MT-CO1	Prostate Cancer	T6253C	6253	T-C	M-T	+	-
MT-CO1	Prostate Cancer	G6261A	6261	G-A	A-T	+	-
MT-CO1	Prostate Cancer	G6267A	6267	G-A	A-T	+	-
MT-CO1	Prostate Cancer	G6285A	6285	G-A	V-I	+	-
MT-CO1	Prostate Cancer	C6340T	6340	C-T	T-I	+	-
MT-CO1	Prostate Cancer	G6480A	6480	G-A	V-I	+	-

MT-CO1	Therapy-resistant Epilepsy	C6489A	6489	C-A	L-I	-	+
MT-CO1	Prostate Cancer	A6663G	6663	A-G	I-V	+	-
MT-CO1	MM & Rhabdomyolysis	G6708A	6708	G-A	G-Ter	-	+
MT-CO1	Acquired Idiopathic Sideroblastic Anemia	T6721C	6721	T-C	M-T	-	+
MT-CO1	Acquired Idiopathic Sideroblastic Anemia	T6742C	6742	T-C	I-T	-	+
MT-CO1	Multisystem Disorder	G6930A	6930	G-A	G-Ter	-	+
MT-CO1	Prostate Cancer	G7041A	7041	G-A	V-I	+	-
MT-CO1	Prostate Cancer	T7080C	7080	T-C	F-L	+	-
MT-CO1	Prostate Cancer	A7083G	7083	A-G	I-V	+	-
MT-CO1	Prostate Cancer	A7158G	7158	A-G	I-V	+	-
MT-CO1	Prostate Cancer	A7305C	7305	A-C	M-L	+	-

APPENDIX 5: *ND1* AND *ND2* GENES

ND1 Gene sequence and Primer Binding sites



ND2 Gene sequence and Primer Binding sites