SNP ANALYSIS IN NUCLEAR SUBUNITS OF RESPIRATORY CHAIN ENZYME COMPLEX I AS NUCLEAR MODIFIER GENE (S) IN THE MITOCHONDRIAL DISEASE: LEBER HEREDITARY OPTIC NEUROPATHY (LHON)

BUSSARAPORN KUNHAPAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE (BIOCHEMISTRY)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2006

ISBN 974-04-6939-6

Thesis
Entitled

SNP ANALYSIS IN NUCLEAR SUBUNITS OF RESPIRATORY CHAIN ENZYME COMPLEX I AS NUCLEAR MODIFIER GENE (S) IN THE MITOCHONDRIAL DISEASE: LEBER HEREDITARY OPTIC NEUROPATHY (LHON)

Bussaraporn Kunhapan
Miss.Bussaraporn Kunhapan
Candidate

Assist.Prof.Chatchawan Srisawat,
M.D.,Ph.D.(Biological Chemistry)
Major-Advisor

Assoc.Prof.Patcharee Lertrit,
M.D.,Ph.D.(Biochemistry)
Co-Advisor

Wanicha Chuenkongkaew
Assoc.Prof.Wanicha Chuenkongkaew,
M.D.
Co-Advisor

Prof.M.R.Jisnuson Svasti,
Ph.D.
Dean
Faculty of Graduate Studies

Vorapan Sirivatanauksorn
Assist.Prof.Vorapan Sirivatanauksorn,
M.D.,Ph.D.
Chair
Master of Science
Programme in Biochemistry
Faculty of Medicine, Siriraj Hospital

Thesis
Entitled

SNP ANALYSIS IN NUCLEAR SUBUNITS OF RESPIRATORY CHAIN ENZYME COMPLEX I AS NUCLEAR MODIFIER GENE (S) IN THE MITOCHONDRIAL DISEASE: LEBER HEREDITARY OPTIC NEUROPATHY (LHON)

was submitted to the Faculty of Graduate Studies, Mahidol University
For the degree of Master of Science (Biochemistry)
on
22 February, 2006

Bussaraporn Kunhapan
Miss.Bussaraporn Kunhapan
Candidate

Assist.Prof.Chatchawan Srisawat, M.D.,Ph.D.(Biological Chemistry)
Chair

Wanicha Chuenkongkaew
Assoc.Prof.Wanicha Chuenkongkaew, M.D.
Member

Patcharee Lertrit
Assoc.Prof.Patcharee Lertrit, M.D.,Ph.D.(Biochemistry)
Member

Chayanon Peerapittayamongkol
Mr.Chayanon Peerapittayamongkol, M.D.,Ph.D.
Member

T. Sura
Prof.Thanyachai Sura, M.D.
Member

Prof.M.R.Jisnuson Svasti, Ph.D.
Dean
Faculty of Graduate Studies
Mahidol University

P. Sakolsatayadorn
Prof.Piyasakol Sakolsatayadorn, M.D.
Dean
Faculty of Medicine, Siriraj Hospital
Mahidol University

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Associate Professor Dr. Patcharee Lertrit and Assistance Professor Dr. Chatchawan Srisawat for being my advisor. They always give me many encouragements and lots of valuable advice throughout my thesis working. I deeply appreciate for the kindness of my two best advisors. Again, I would like to grateful and with deepest regret to all of my troubles to both advisors.

My sincere appreciation is extended to my co-advisor, Associate Professor Dr. Wanicha Chuenkongkaew, Department of Ophthalmology, Faculty of Medicine, Siriraj Hospital for her useful suggestion on LHON clinical and her kindness trough our field trip including all ophthalmologists who came with us all the field trips.

I would like to express my grateful to the members of thesis dissertation committees, Professor Dr. Thanyachai Sura, Division of Medical and Molecular Genetic, Department of Medicine, Faculty of Medicine, Ramatibodi Hospital for his valuable suggestion and more useful genetic comments.

My sincere gratitude is extended to Dr. Chayanon Peerapittayamongkol from our department for his valuable advices and discussions especially for the genetic statistical analysis throughout my research study.

I am deeply grateful and indebted to Associate Professor Dr.Klai-Upsorn Pongrapeeporn, the one of my best teacher here who gives me the best opportunity to live in this laboratory.

Special thank is given to Miss Rungnapa Suphavilai for her best lab teaching and look after me for a long time. Many thanks are also given to the members in my laboratory, Mr. Komon Luangtrakool, Miss Thitima Sanpachudayan, Miss Pattamon Tharaphan, Miss Benjamas Intharabut, Miss Treenud Suntisiri, Miss Supannee Kaewsutthi, Miss Preeyaphan Meesap, Miss Hathaichanoke Boonyarit, and Mr. Yutthana Joijinda for their great assistance, cheerfulness and warm friendship throughout my study here. Also, I am more thankful to Mr. Nopasak Phasukkijwatana for the LHON information, more useful advice, his friendship and assistance in many ways. I am indebted to all of them.

Many thanks are also expressed to graduate students and staffs in Department of Biochemistry, Faculty of Medicine, Siriraj hospital, especially to Dr. Toni Suthammarak, Dr. Bhoom Suktitipat, and Dr. Patchara Nantasri for their friendship and assistance in my lab work and in many ways, and to Miss Bantita Kasemsuk for her good care and beside me through our friendship. I am also many thanks all individuals surrounded me which seem to be some parts of my life as well.

This thesis is supported in part by Commission on Higher Education and Faculty of Graduate Studies, Mahidol University, in the Academic Year of 2004.

Finally, the extreme gratitude is expressed to my father, my mother, my brother, and my aunt for their support, encouragement and trusts in me throughout this thesis working. I am lucky to receive the best life from my beloved family.

Bussaraporn Kunhapan

SNP ANALYSIS IN NUCLEAR SUBUNITS OF RESPIRATORY CHAIN ENZYME COMPLEX I AS NUCLEAR MODIFIER GENE (S) IN THE MITOCHONDRIAL DISEASE: LEBER HEREDITARY OPTIC NEUROPATHY (LHON)

BUSSARAPORN KUNHAPAN 4537301 SIBC/M

M.Sc.(BIOCHEMISTRY)

THESIS ADVISORS: CHATCHAWAN SRISAWAT, M.D.,Ph.D.(BIOLOGICAL CHEMISTRY), PATCHAREE LERTRIT, M.D.,Ph.D.(BIOCHEMISTRY), WANICHA CHUENKONGKAEW, M.D.

ABSTRACT

Leber hereditary optic neuropathy (LHON) is a maternal inherited disorder of the optic nerve characterized by bilateral visual failure, central vision loss and the color perception impairment mostly in young males. This disease is primarily caused by mutations in mitochondrial DNA (mtDNA). The most common mtDNA mutation in LHON case worldwide is G11778A. This mutation changes amino acid arginine to histidine in the ND4 subunit of complex I in the respiratory chain. However, an incomplete penetrance and male predominance of this disease cannot be explained by maternal inheritance solely. Additional genetic or environmental factors are required for the disease expression. *NDUFS4*, *NDUFS8* and *NDUFV2* genes involving electron transfer are nuclear subunits of complex I. SNP in the coding and promoter region of these genes was analyzed for the modifier of this disease in our population. Total 10 SNPs in these 3 candidate genes were detected, 4 in *NDUFS4*; -447 T→A, +12 C→G, +198 C→A and +312 G→A, and 4 in *NDUFV2*; -602 G→A, -288 C→T, -233 T→C and +86 T→C and 2 in *NDUFS8*; -496 C→G and -45 A→C respectively. 6 validated SNPs were analyzed as the susceptibility markers closed to disease loci in 76 LHON patients, 205 of their relatives and 134 normal individuals. All validated SNPs genotyping of affected, unaffected and normal control groups were in the Hardy-Weinberg equilibrium. Two SNPs in coding region of *NDUFS4*, the +12 C→G and +312 G→A, and 2 SNPs in coding and promoter region of *NDUFV2*, the -602 G→A and +86 T→C, were associated with LHON phenotypic in comparison between affected and normal healthy individuals (P<0.005, Chi-square test) but was not significantly different between affected and unaffected LHON (P>0.05, Chi-square test). Neither single SNP nor haplotype susceptility loci were found to be associated with LHON survival function using family-based association test (FBAT). The Kaplan-Meier survival curve and Cox proportional hazard regression model indicated that sex, secondary mutation, and heteroplasmy had influenced the LHON manifestation. The relative risk of sex, secondary mutation, and heteroplasmy to LHON are 4.54, 3.52 and 3.59 times (P=0.001) respectively.

KEY WORDS: LEBER HEREDITARY OPTIC NEUROPATHY / MITOCHONDRIAL RESPIRATORY CHAIN COMPLEX I / NUCLEAR MODIFIER GENE / FAMILY-BASED ASSOCIATION STUDY

122 P. ISBN 974-04-6939-6

การวิเคราะห์ความสัมพันธ์ของ SNP ในนิวเคลียร์ยีน ที่ถอดรหัสให้โปรตีนหน่วยย่อยของเอนไซม์คอมเพล็กซ์ที่ 1 ในขบวนการ OXPHOS กับการแสดงออกของโรคไมโทคอนเดรีย LHON (LEBER HEREDITARY OPTIC NEUROPATHY) (SNP ANALYSIS IN NUCLEAR SUBUNITS OF RESPIRATORY CHAIN ENZYME COMPLEX I AS NUCLEAR MODIFIER GENE (S) IN THE MITOCHONDRIAL DISEASE: LEBER HEREDITARY OPTIC NEUROPATHY (LHON))

บุศราภรณ์ คุณหพันธ์ 4537301 SIBC/M

วท.ม. (ชีวเคมี)

คณะกรรมการควบคุมวิทยานิพนธ์ : ชัชวาลย์ ศรีสวัสดิ์, พ.บ., Ph.D. (Biological Chemistry), พัชรีย์ เลิศ-ฤทธิ์, พ.บ., Ph.D. (Biochemistry), วณิชา ชื่นกองแก้ว, พ.บ.

บทคัดย่อ

Leber hereditary optic neuropathy (LHON) เป็นโรคความผิดปกติทางพันธุกรรมที่ถ่ายทอดทางมารดาทำให้เกิดมีเส้นประสาทตาเสื่อม ลักษณะอาการของโรคคือมีตามัวเฉียบพลัน มองภาพตรงกลางไม่ชัดและมักมีตาบอดสีร่วมด้วย สาเหตุของโรคเกิดจากการกลายพันธุ์ของยีนไมโทคอนเดรีย ในผู้ป่วยชาวไทยพบการกลายพันธุ์ที่ตำแหน่ง G11778A มากที่สุด และเป็นตำแหน่งที่ถอดรหัสให้โปรตีนหน่วยย่อยของเอนไซม์คอมเพล็กซ์ที่ 1 ในขบวนการ OXPHOS อย่างไรก็ตาม พบว่าสมาชิกในครอบครัวของผู้ป่วยบางคนมีความผิดปกติทางพันธุกรรมของโรคนี้ แต่มิได้แสดงอาการของโรคแต่อย่างใด สันนิษฐานได้ว่าอาจมีปัจจัยทางด้านพันธุกรรมหรือสิ่งแวดล้อมอื่นที่มีผลต่อการเลือกแสดงอาการ การศึกษานี้มีวัตถุประสงค์เพื่อหาปัจจัยทางด้านพันธุกรรมในยีนนิวเคลียร์ที่น่าจะมีผลต่อการเลือกแสดงออกของโรคนอกเหนือจากการกลายพันธุ์ของยีนไมโทคอนเดรียเพียงอย่างเดียว ในการศึกษานี้ครอบคลุมบริเวณที่ถอดรหัสและส่วนควบคุมการถอดรหัสโปรตีนของยีน NDUFS4, NDUFS8 และ NDUFV2 ซึ่งถอดรหัสให้โปรตีนหน่วยย่อยของเอนไซม์คอมเพล็กซ์ที่ 1 และหาตำแหน่งความผันแปรทางพันธุกรรมเพื่อใช้เป็นเครื่องหมายในการศึกษาความเกี่ยวข้องของยีนนั้นต่อการแสดงออกของโรค LHON จากการศึกษาพบว่ามี 6 ตำแหน่งที่เลือกใช้เป็นเครื่องหมายทางพันธุกรรม คือ -447 T→A, +12 C→G, และ +312 G→A ในยีนนิวเคลียร์ NDUFS4 ตำแหน่ง -496 C→G ในยีน NDUFS8 และตำแหน่ง –602 G→A, +86 T→C ในยีน NDUFV2

ผลการศึกษาพบว่า ปัจจัยเรื่องเพศ ปริมาณความผิดปกติทางพันธุกรรมในเลือด และการมี/ไม่มีการกลายพันธุ์ของยีนไมโทคอนเดรียตำแหน่งรอง มีอิทธิพลต่อการแสดงออกของโรค LHON แต่ไม่พบความสัมพันธ์ระหว่างความผันแปรทางพันธุกรรมทั้ง 6 ตำแหน่งดังกล่าว บนยีนนิวเคลียร์ NDUFS4, NDUFS8 และ NDUFV2 กับการแสดงออกของโรคนี้เมื่อทดลองโดยใช้ FBAT

122 หน้า ISBN 974-04-6939-6

CONTENTS

CONTENTS (continued.)

LIST OF TABLES

LIST OF TABLES (continued.)

Page

LIST OF FIGURES

Page

LIST OF FIGURES (continued.)

Page

LIST OF ABBREVIATIONS

bp	Base pair
cDNA	Complementary deoxyribonucleic acid
°C	Degree Celsius
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
dGTP	Deoxyguanosine triphosphate
dTTP	Deoxythymidine triphosphate
dNTP	Deoxynucleotide triphosphate
DNA	Deoxyribonucleic acid
mg	Milligram
ml	Millilitre
mM	Millimolar
mRNA	Messenger ribonucleic acid
µg	Microgram
µl	Microliter
ng	Nanogram
nm	Nanometer
pmol	Picomole
PCR	Polymerase chain reaction
RCLB	Red cell lysis buffer
Taq	*Thermus aquaticus*

CHAPTER I
INTRODUCTION

Leber hereditary optic neuropathy (LHON) is a maternal inherited disorder of optic atrophy that primarily caused by mutations in the mitochondrial genome (mtDNA). The clinical phenotypes begin with vision failure that usually occurs in one eye and the other can be simultaneous or sequential progress. The visual acuity becomes drop to hand motion state together with the color blindness in less than a week or over a few months. The disease onset usually occur in young adults in the second to third decade of life while it can affect in children less than 10 years or adults over 50 years and predominantly affect in males.

The most primary LHON pathogenesis were mutations in mitochondrial genome in subunits of mitochondrial respiratory enzyme complex I (NADH-ubiquinone oxidoreductase). These mutations were the G3460A, G11778A, and T14484C mutation. All of them alter highly conserved amino acids in ND1 Ala52Thr, ND4 Arg340His, and ND6 Met64Val respectively. Although these primary mutations are necessary for disease development in LHON patients, not all individuals carrying these mutations become optic atrophy. The LHON characteristic shows an incomplete penetrance and has the predilection for males to lose vision, suggesting that not only the primary LHON mutations, but an additional genetic factor as well as the environmental influences also interact with the disease development.

Considering the mitochondrial background, secondary LHON mutations have been associated with increasing the risk of LHON expression or act in synergy with the primary mutations. More than 20 secondary LHON mutations have been reported (http://www.mitomap.org/cgi-bin/mitomap/tbl8gen.pl); these mutations create both conservative and non-conservative amino acid substitutions that can be found in normal population. Furthermore, the proportion of mutated mtDNA (heteroplasmy) and mitochondrial background (mitochondrial haplogroup) can also contribute to the LHON penetrance. There are reports of a minimal risk of blindness if the

heteroplasmy mutational load is less than 60%. Whereas, the mutation load more than 80% have the same risk of blindness as homoplasmic phenotype. Also, haplogroup J, which is one of nine specific mtDNA European haplogroups, shows a strong preferential association in LHON pedigrees. Additionally, the environmental evidences such as smoking, alcohol assumption, nutritional deprivation, psychological stress, or acute illness have also reported as the effective factors to LHON manifestation. However, the explanation on the risk of visual loss increasing by these additional factors is disputed and required more clarification.

Studies on genetic factors, either an X-linked susceptibility gene or nuclear modifier genes were in consideration. Since the majority of the LHON mtDNA mutations contribute in complex I subunit and its enzyme activities was found to be decreased significantly in LHON patients. In an attempt to elucidate this complex genetic-bases of LHON, the nuclear encoded genes for complex I subunits are the strong candidates for candidate genes approach.

In this study, 45 LHON pedigrees were analyzed in order to search for LHON modifying genes. Single nucleotide polymorphisms (SNPs) in candidate genes, *NDUFS4*, *NDUFS8*, and *NDUFV2* were used as markers for association study with LHON phenotype. These genes involve in electron transfer, G-protein, and cAMP dependent phosphorylation of the respiratory pathway which are necessary for complex I function. The parental allele transmitted to affected offspring based on the linkage disequilibrium was used to imply as risk variants to disease. Covariates of LHON manifestation (gender, mtDNA mutation proportion, and secondary mutation) were also assessed by Kaplan-Meier survival analysis and Cox proportional hazard regression model. Then, the family-based association study was analyzed with family-based association test (FBAT).

CHAPTER II
LITERATURE REVIEW

1. Leber Hereditary Optic Neuropathy (LHON)

Leber hereditary optic neuropathy (LHON) is a maternal inherited disorder of optic atrophy that was first described by the German ophthalmologist Theodor Leber in 1871 (1). The primary cause of LHON development is a mutation in the mitochondrial genome which involves a loss of central vision due to degeneration of the retinal ganglion cells and optic nerve axons. Typically pathology begins with presymtomatic subsequently to acute and become atrophic stage causing the loss of visual acuity. The initial symptom is a blurring or clouding of central vision that more commonly affected one eye and the other can progress within a delay of weeks or months. During the acute phase, the blindspot becomes enlargement and expands to absolute central scotoma, the failure of color perception also accompanies. The peripapillary nerve fibre layer disappears, disk becomes pale, and then to complete blindness in atrophic phase (2-5). Optic atrophy onset-age is usually between 15 and 35 years, but the visual deterioration can occur with a broad range from children under 10 to adult over 60 years (4-6).

1.1 Biochemical scheme for the optic neuropathy in LHON

The LHON pathogenic mutations are primary result of mutation in mitochondrial DNA (mtDNA), which involves genes encoding complex I subunits of the respiratory chain. An alteration of conserved amino acid and protein configuration in complex I might be either affect the function of respiratory chain or defect the role of mitochondria in retina leading to visual failure. There appear to be several possible pathways that link the respiratory chain dysfunction in LHON, between the pathogenic mutations and the degeneration of optic nerve. The primary function of mitochondria is known for the ATP synthesis by aerobic metabolism (7). Normally, both retinal ganglion cells and nerve fibre layer showed a high mitochondrial enzyme activity (8). It is more likely that optic nerve degeneration in LHON is affected since they lose a

sustained level of ATP for normal function. Mitochondria functions also serve as a critical role in redox balancing and the regulation of apoptotic pathway. The pathway that activates mitochondrial permeability transition and redox-sensitive NMDA (*N*-methyl-D-aspartate) channel, which is an early step in apoptotic pathways, is triggered by the superoxide anion (9-11). The accumulation of free radicals (superoxide anion) in mitochondria can damage mtDNA and then impair mitochondrial function. It can inward result to higher rates of superoxide production (12). Superoxide anions are particularly resulted from respiratory chain complex I which is remarkable sites for the most primary LHON mutations. Therefore, an abnormal respiratory chain in LHON is not only losing the clearance function of these radicals but also regulate the accumulating and triggering the retinal ganglion cells death by apoptotic pathway. Another intriguing factors in the regulation of apoptosis in ganglion cell layer are also involved in *bcl*-2 and the releasing of cytochrome *c* in mitochondria (13, 14). Either respiratory chain dysfunction or defect in mitochondrial metabolism due to the LHON mutation is challenged for optic atrophy.

1.2 Intriguing factors for LHON development

LHON is a mitochondrial disorder with complex traits and incomplete penetrance. This disease is primarily caused by mutation(s) in the mitochondrial genome (mtDNA) which is transmitted through maternal lineage. However, individuals with pathogenic LHON mutation have different genetic risk assessment in their family history. Over 50% of males and 85% of female harboring the LHON mtDNA mutation do not develop the disease. The proposed pathways lead to optic nerve degenerate involves both genetic and environmental factors. The pathophysiology of LHON are summarized in Figure 1 (15).

1.2.1 Mitochondrial factors associated with LHON manifestation

The human mitochondria have their own DNA (mtDNA) in circular and 16,569 base pairs in length. Mitochondrial genome is a double-stranded molecule consists of a heavy (H) and a light (L) chain without any histone coat. The non-coding region, called D (displacement) loop, is the promoter region for the replication and transcription of both heavy and light strand. There is no intron but sequence with a small separation of 1-2 bases between genes. 93% of mitochondrial genome contains coding region, whereas the nuclear genome is only 3%. There are 37 genes coding

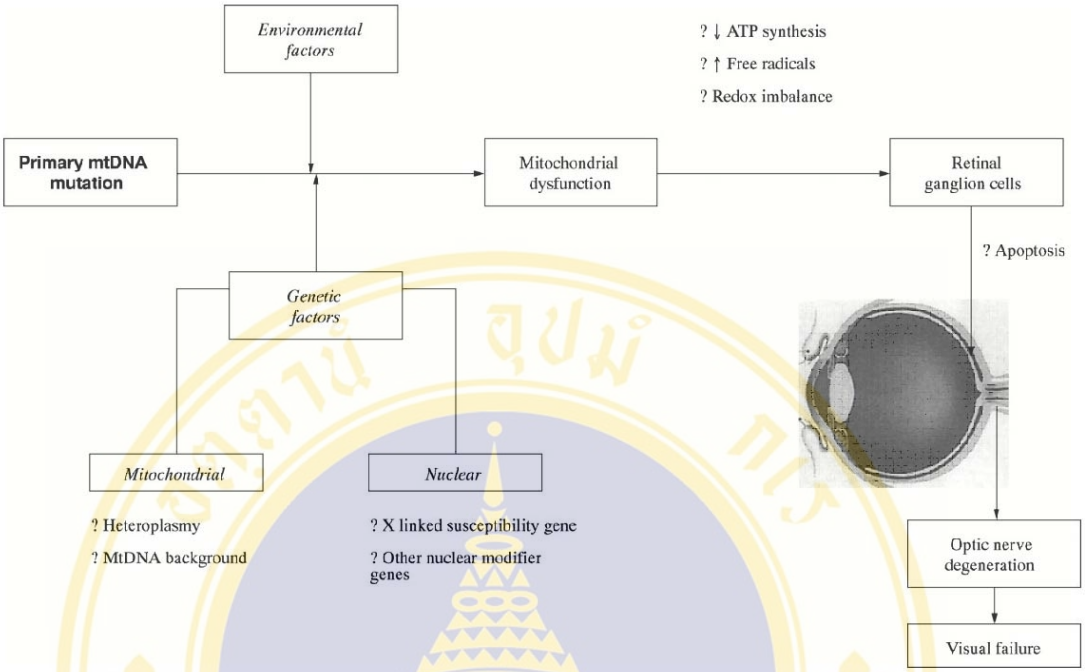


Figure 1. Schematic pathways leading to optic nerve degeneration in LHON (15).

region, whereas the nuclear genome is only 3%. There are 37 genes comprising 22 transfer RNAs (tRNA), 12S and 16S ribosomal RNA (rRNA), together with 13 proteins involved in mitochondrial oxidative phosphorylation pathway. Mutation in mtDNA has a high rate that probably due to the high rate of replication, an exposure to oxygen free-radicals produced during respiration, a lack of protective histones itself and less efficient in repair system. This high mutation rate causes a number of polymorphisms which accumulate along maternal lineages. Mutations in mtDNA do cause disease, since it is such a short sequence and contains very heavy information.

Many neurological and degenerative disorders are often linked to mitochondrial disorders because the brain is highly energy intensive. Some diseases occur as a result of mtDNA mutations or deletions. LHON was the optic degeneration that primarily caused by point mutations in mitochondrial genome. The G3460A, T14484C and G11778A are specifically pathogenic mutations of LHON which all alter the polypeptide of mitochondrial respiratory chain complex I.

1.2.1.1 Primary and secondary LHON mutation

The pathogenic mutations in mitochondrial genome (mtDNA) that cause LHON was first identified by Wallace and colleagues (16). One of three mutations, G11778A (ND4), G3460A (ND1) and T14484C (ND6) are known as the primary mutation of LHON cases worldwide. All of them affect the NADH dehydrogenase (ND) gene coded for complex I subunits of the mitochondrial respiratory chain (17-20). Its penetrance has arisen on different ethnic origins, but G11778A is account for the most common (21-23). Other rare primary mutations were identified in complex I especially in ND6 gene, this locus seems to be the hot spot for LHON mutations (20, 24). Additionally, Johns and Berman (1991) proposed that there might be another mitochondrial mutations that play a role in the etiology and/or pathogenesis of LHON (25). In contrast to primary LHON mutations, these are so-called secondary mutations and considered to be low-risk mutations for developing the disease. Secondary mutations can be found in normal population but at the lower frequencies than those in patients and they are usually accompanied with both primary and another secondary mutation to increase the probability of LHON expression. However, the pathogenic or etiological role of these secondary mutations remains to be elucidated (6, 26-32).

1.2.1.2 Heteroplasmy

Hundreds of mitochondria with thousands of mtDNA copies exist in highly metabolically active tissue throughout the body. The threshold for a mixture of normal and mutant mtDNA causing disease, which termed as heteroplasmy, might influence the expression and inheritance pattern of LHON (33-37). It has been reported that if the 11778 mutation load is less than 60%, the risk of developing optic atrophy is lesser (38, 39). Additionally, studies of inheritance pattern of the G11778A mutation observed that the risk of affected offspring born to mothers with >80% mutated mtDNA and mothers with homoplasmic mutated mtDNA was not significantly difference. In contrast, the risk of affected offspring born to mothers with ≤80% mutated mtDNA was significantly less than the risk to homoplasmic mothers (39). Further, heteroplasmy contributed a significant effect to the expression of LHON in families with the 3460 mutation (36). Levels of mutant (more than 50% and 75% mutated G3460A) showed a strong correlation with the disease phenotype. In addition, LHON cases who carried haplogroup J together with the heteroplasmic of 14484 mutation have low threshold for disease expression with the deficiency in respiratory activity (19). Heteroplasmy can determine the relative risk for disease expression in some pedigree. However, it should be concerned about an accuracy quantification of the heteroplasmy of mtDNA mutation existed in variable tissue (34, 40, 41).

1.2.1.3 MtDNA background

Since mutation rate of mtDNA is very high and mtDNA polymorphisms have evolved and accumulated along maternal lineages. A number of mtDNA polymorphisms which are silent mutations but more often occurred are used to organize mtDNA haplotype individuals. Groups of the specific related haplotypes are named as mitochondrial haplogroups. Mitochondrial genetic background that differ in each ethnic population may influence the specific prevalence of LHON mutations (20, 42). The association among mitochondrial haplogroup, primary and secondary mutation to LHON is still controversy. Based on phylogenetic analysis, it has been shown that some secondary mutations, T4216C, G13708A, G15257A, and G15812A, cluster on a specific mtDNA background haplogroup J in European (43). Also, the primary T14484C and G11778A mutations show a strong preferential

association with haplogroup J in this population (44, 45). However, the G3460A mutation is distributed randomly among all mtDNA haplogroups without any preferential association (44-46).

1.2.2 Nuclear genetic factors influencing the LHON development

Although, LHON were clearly associated with mtDNA point mutations which transmitted along maternal lineages. There were several characteristics in majority of LHON pedigrees that cannot be explained by mitochondrial inheritance alone. These included male predominance, late age of onset for females, incomplete penetrance and the expression limited to the optic nerve. The primary mutations are necessary for disease development, but other factors might have a role in influencing the clinical expression. In fact, oxidative phosphorylation enzyme subunits are encoded both by its own genome and by nuclear genome. The deleterious interactions of enzymes encoded by mtDNA and/or nuclear DNA could lead to abnormal activity or dysfunction of mitochondrial multi-enzyme systems. This possibility was analyzed on linkage and segregation analyses. It has been suggested that the interactions of mitochondrial locus and X-chromosome-linked disorder may contribute to the difference in LHON characteristics between males and females (15, 47).

1.2.2.1 X-linked susceptibility gene

Gender bias is an intriguing feature of LHON that males have more tendencies to lose vision than females, although the gender specific bias for disease development differs in each primary LHON mutation (48-52). Affected male to female ratio with G11778A carrying extends for 3.7-5.1:1. While, the wide range ratio occurs in males with T14484C mutation at 2.1 to 7.7 folds to females. A less marked gender bias in male cases with G3460A remains 2.3 to 4.3 folds to females. An X linked susceptibility locus was first suggested to play role in synergy with the mtDNA mutation (49, 53-55). From segregation analysis involving both mitochondrial and nuclear including the X chromosome-linked loci, showed that 60% of affected females are likely to be heterozygous at the X chromosome-linked (53, 55). They are thought to be affected by unfortunate X chromosome inactivation, these features also provide the explanation for the later age of onset and inclination in females to males. The primary evidence of linkage on the X chromosome was

suggested by Vilkki *et al* (54). They provided the evidence using 15 polymorphic markers spanning the X chromosome at 2-50 cM interval in six LHON pedigrees. The X susceptibility loci is placed closely linked to marker DXS7 with a maximum lod score (Z_{max}) of 2.48 and a recombination fraction (theta) of point zero. However, there are strong arguments for this susceptibility locus (47, 56-64). Recently, an attempt to explain the reduced penetrance and sex bias was performed (65). Nonparametric linkage (NPL) analysis and fine mapping considering by a dense microsatellite markers were utilized to confirm the previously mapping of X-chromosome susceptibility loci. This study had also evaluated the epigenetic mechanism to the incomplete penetrance of LHON phenotypic. The X-chromosome imprinting model was used to explain but this principal was not the underlying cause.

1.2.2.2 Nuclear modifier gene(s)

The existence of nuclear modifying loci often thought to be on the X chromosome, which could explain the male predominance. However, the incomplete penetrances had suggested that another nuclear gene(s) might play an important role on the expression of LHON. Studies of nuclear genetic background attributed a disease development were most analyzed in cybrid cells. Cybrids (cytoplasmic hybrid) were prepared from the fusion of enucleated fibroblasts cells with the cells lacking mtDNA (ρ^0) (66, 67). The 3460 primary LHON mutation was hold as a severe complex I defect compared to another primary LHON mutations (68). However, in cybrid cells of LHON carried 3460 containing a different nuclear background (between lung derived carcinoma ρ^0 and osteosarcoma ρ^0 cells), they showed the difference in biochemical defect. The clones of 3460 fibroblasts with osteosarcoma ρ^0 cells had shown a severe deficiency of complex I activity, but this defect was not apparent in the entire clones studied of lung carcinoma ρ^0 cells (66). The results might be inferred that the nuclear environment can influence the expression of the biochemical defect in LHON. Another genetic approach in 11778 LHON mutation was utilized, this mutation was an ambiguous studies in complex I activity. Genetic analysis with the comparative genomic hybridization (CGH) technique was used to screen chromosomal gains and losses throughout the whole genome in LHON patients with 11778 mutation. This study cannot point any specific chromosomal regions that might have the nuclear factors in the pathogenesis of the

disease (69). Nevertheless, such a LHON nuclear factor may influence with a little effect accumulating to mitochondrial genetic and non-genetic factors. Nuclear gene (s), whatever, associated partly with pathway induced neuronal defect and/or optic neuropathy might be in consideration.

1.2.3 Effect of environmental factors to the optic neuropathy in LHON

The LHON penetrance might be triggered by additional environmental, lifestyle, gender-based physiological or even anatomical differences between males and females (48, 70). Two models for the development of the optic neuropathy have been suggested. First, the threshold model, this model supposed that if individuals carried a primary LHON mutation that could increase the rate of neurodegeneration until it exceed the threshold. Then, the additional factors would further modify the neurodegeneration rate (6, 42). Another model depends on the primary LHON mutation which may impair the ganglion cell metabolism without increasing neuronal death. But when the additional conditions are at-risk, they can trigger neuronal death rapidly (5).

Many reports studied on monozygotic twin harboring primary LHON mutation in order to support the environmental effect. The affected and unaffected monozygotic twins show the vision loss in discordance whether unaffected sib lost vision later (49, 71-74). The existence of discordant twins is strongly suggested that the environmental factors also contribute to the penetrance. In addition, smoking, alcohol consumption, head injury, psychological stress, some chemical/drugs exposure or even the systemic illness can accelerate the onset of vision loss (75-81). However, some case-controls studies failed to confirm the association between these factors and the risk of visual loss (48, 82). Evaluating these potential environmental factors needs a proper study design, a typical clinical profile and a compatible family history in case of patients have lost vision for several years.

2. Mitochondria and dysfunction in LHON

Mitochondria are double membrane-bound organelles distributed through the cytosol of most eukaryotic cells. Their compartments contain an outer membrane, inner membrane which separated by an intermembrane space. The mitochondrial matrix is inside and surrounded by inner membrane. These two membranes are

phospholipids bilayer with embedded proteins. The outer membrane contains many complexes of integral membrane proteins that form channels for a variety of molecules and ions to move through. The inner membrane has many folds which are called cristae and contains electron carriers (coenzyme Q and cytochromes) and 5 complexes of integral membrane proteins. These complexes are essential to provide the energy via oxidative phosphorylation (OXPHOS) system. The matrix contains a high concentration of enzymes that catalyzes several steps in cellular respiration, and the site of mitochondrial DNA (mtDNA) and ribosomes suspended. Mitochondria have their own DNA (mtDNA) containing genes coded for 2 rRNAs, 22 tRNAs, and a number of protein complexes assembly which required the expression of both mitochondrial and nuclear genome.

Primary role of mitochondria is to generate ATP via OXPHOS system and serve a variety of functions within the cells such as urea cycle, the biosynthesis of heme, pyrimidines, amino acids, phospholipids and nucleotides. In addition, mitochondria also play key roles in cellular signaling through apoptosis and the generation of reactive nitrogen- and oxygen species (RNS/ROS).

2.1 OXPHOS in mitochondria

Mitochondria are known as the powerhouse of most eukaryotic cells since their primary function is to convert organic materials to high energy molecules via the process of oxidative phosphorylation (OXPHOS). The energy is stored in high energy phosphate bonds in a molecule called adenosine triphosphate (ATP). OXPHOS is the process that couples the removal of hydrogen ions and providing phosphate molecule to ADP to synthesize ATP. Nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide ($FADH_2$) producing Kreb's cycle are served as the hydrogen ions (or protons) with electrons carrier. Electrons are carried out the electron transport chain (mitochondrial respiratory chain) mediated by protein complexes (Complexes I-V, CoQ oxidoreductase, and ANT) and electron shuttle molecules (coenzyme Q and cytochrome *c*) in the mitochondrial inner membrane. Complex I, or NADH dehydrogenase:ubiquinone oxidoreductase, uses the energy from NADH for proton pumping into the intermembrane space and passing energy on to complex III (coenzyme Q or ubiquinone:cytochrome bc1 oxidoreductase) in the form of electron. Complex II (succinate dehydrogenase:CoQ oxidoreductase) accepts energy from

$FADH_2$ and passes electrons to complex III. Then, complex III pumps protons out and move electrons past to complex IV (cytochrome oxidase). Complex IV is the last site for proton pumping into the space. These protons are pumped against their concentration gradient into the intermembrane space and reentering to the matrix via the ATP pump of complex V (ATP synthase:proton translocase). Electrons transferred (from NADH and $FADH_2$) and protons that passed back are used to reduce oxygen molecules to form water. Complex V uses the energy from electrochemical potential of proton gradient to form bonding between ADP and phosphate group to ATP. The overview of OXPHOS system in mitochondria is shown in Figure 2.

2.2 Mitochondrial respiratory chain dysfunction in LHON

The LHON pathogenesis were associated with mutations in mitochondrial genome and caused the mitochondrial respiratory chain defect in sequences of complexes I, III and IV subunits (6, 27, 28). Most of these mutations had been demonstrated specifically in mitochondrial respiratory chain complex I.

2.2.1 Mitochondrial respiratory chain complex I

Complex I (NADH:ubiquinone oxidoreductase) is the largest complex in mitochondria that embedded in the inner mitochondrial membrane (IMM) and partly protruding into the matrix. This complex dehydrogenates NADH and shuttles electrons to coenzyme Q and then generates a proton gradient across the IMM. It provides the proton-motive force used for ATP synthesis. The human complex I consists of 46 protein subunits comprising 7 mitochondrial-encoded gene products (ND1, 2, 3, 4, 4L, 5 and 6) and the remainders were encoded by nuclear genome (83-85). To date, approximately 34 nuclear-encoded genes were well known. The complex I is an L-shape configuration including 3 fractions of 3 flavoproteins (FPs), 7 iron-sulpher proteins (IPs), and 24 hydrophobic proteins (HPs) and the remainders remain unknown (86). Both FP and IP fractions have a water-soluble peripheral arm protruding into the mitochondrial matrix, whereas the HP fraction, a water hydrophobic arm, is embedded in the inner mitochondrial membrane. Electrons are transferred from NADH to ubiquinone through these 3 fractions and simultaneously transported the protons toward the intermembrane space (Figure 3). The present functional properties of complex I subunits has not been well known (Table 1).

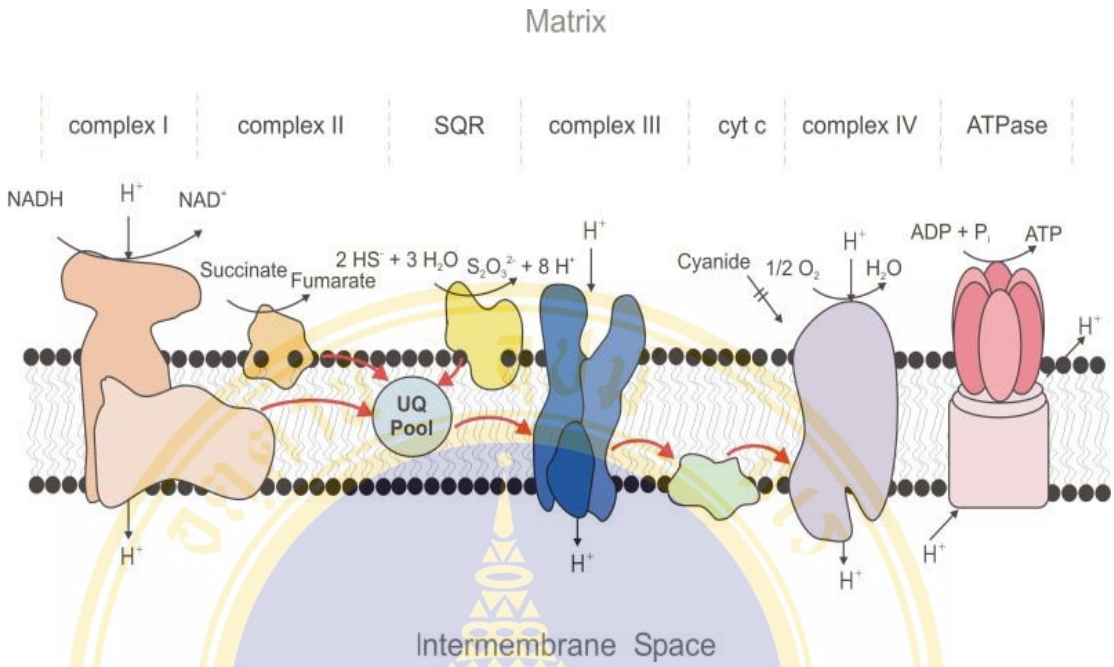


Figure 2. The OXPHOS system in mitochondria. The OXPHOS system of human mitochondria involves 5 enzyme complexes which located in the mitochondrial inner membrane. Complex I, II, III and IV comprise the chain of electron transport according to which NADH or succinate becomes oxidized. Then, an electrochemical potential of proton gradient is concurrently generated protons across the mitochondrial inner membrane to intermembrane space. Complex V utilizes this potential energy to synthesize ATP from ADP and inorganic phosphate (87).
(http://www.molevol.de/lab/nu04/ursula04.html)

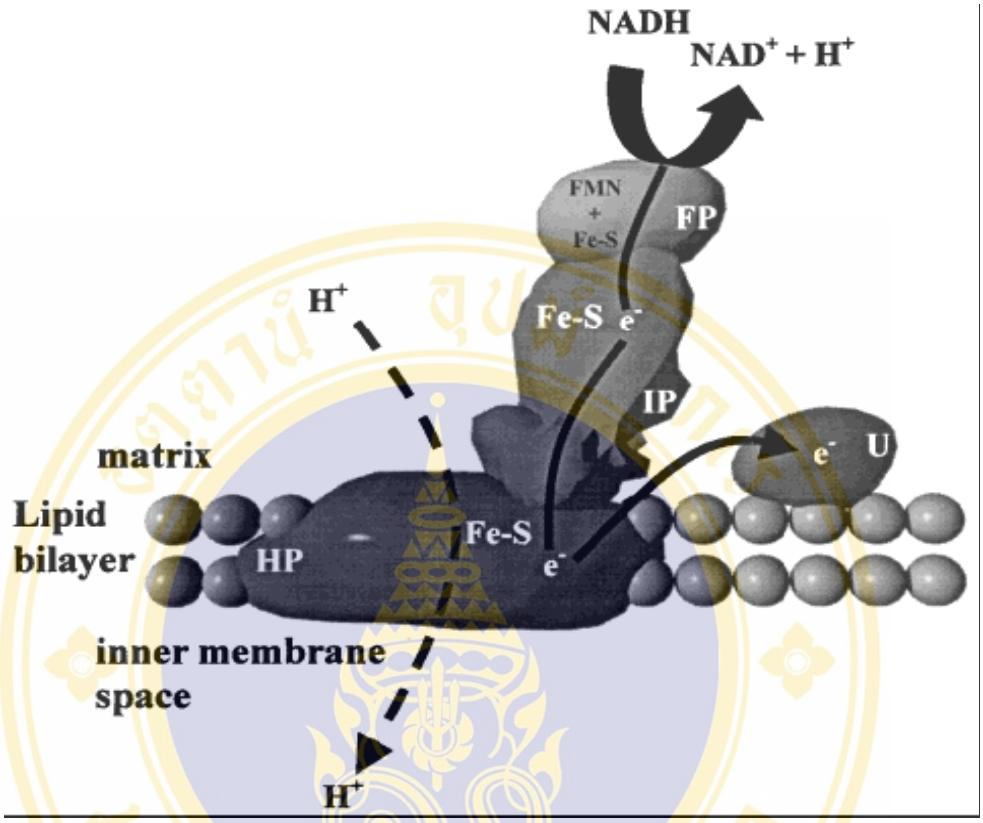


Figure 3. Schematic overview of human complex I. Electrons are transferred from NADH to ubiquinone through three distinct parts of complex I. Simultaneously with this electron transport, protons are pumped toward the intermembrane space (88).

Table 1. Human Complex I Nuclear Genes (89-100).

Gene Group and Gene	cDNA/nDNA/Leader Sequence	Chromosomal Localization	Biochemical Function
Flavoprotein:			
NDUFV1 (*NuoF*)	+/+/+	11q13	NADH binding, FMN binding, electron transfer (4Fe-4S cluster;N3)
NDUFV2 (*NuoE*)	+/+/+	18p11.2-p11.21	G-protein, Electron transfer (2Fe-2S cluster;N1a, 4Fe-4S cluster;N4)
NDUFV3	+/+/+	21q22.3	
Iron-sulfur:			
NDUFA5	+/-/…	7q31.33	
NDUFS1 (*NuoG*)	+/-/+	2q33-34	Electron transfer
NDUFS2 (*NuoD*)	+/-/+	1q23	Electron transfer
NDUFS3 (*NuoC*)	+/-/+	11p11.11	Electron transfer
NDUFS4	+/-/+	5q11.1	cAMP dependent Phosphorylation
NDUFS5	+/-/…	1p34.2-p33	
NDUFS6	+/-/+	5pter-p15.33	
Hydrophobic:			
NDUFA1	+/+/…	Xq24-25	
NDUFA2	+/-/…	5q31.2	
NDUFA3	+/-/…	19	
NDUFA4	+/-/…	7	
NDUFA6	+/-/…	22q13.1	
NDUFA7	+/-/…	19p13.2	Ubiquinone binding?
NDUFA8	+/-/…	9q33.2-34.11	
NDUFA9	+/-/+	12p13	NAD(P)H reductase/isomerase
NDUFA10	+/-/+	2	
NDUFAB1	+/-/…	16p12.3-12.1	Acyl carrier protein motif
NDUFB1	+/-/…	14q31.3	
NDUFB2	+/-/+	7q34-35	
NDUFB3	+/-/…	2	
NDUFB4	+/-/…	3	
NDUFB5	+/-/+	3	
NDUFB6	+/-/…	9p13.2	
NDUFB7	+/-/…	19p13.12-13.11	
NDUFB8	+/-/+	10q23.2-23.33	
NDUFB9	+/+/…	8p24.21	
NDUFB10	+/-/…	16p13.3	
NDUFS7 (*NuoB*)	+/-/+	19p13.3	Electron transfer (4Fe-4S cluster;N3)
NDUFS8 (*NuoI*)	+/+/+	11q13.1-13.3	Electron transfer
NDUFC1	+/-/…	4q28.2-28.3	
NDUFC2	+/-/…	11	
Unknown:			
17.2 kD	+/…/…	...	

2.2.2 LHON and dysfunction in complex I

Since most primary pathogenic LHON mutations, G3460A, T14484C, and G11778A, alter mtDNA encoded for ND1, ND6, and ND4 subunits of mitochondrial respiratory chain complex I (NADH dehydrogenase:ubiquinone oxidoreductase), an abnormality due to these mtDNA mutations affect the highly conserved mitochondrial polypeptides of this complex, thus affecting the enzyme activity in system.

Of most common primary etiologic mutations in LHON, mutation at nucleotide 3460 show markedly reduced in enzyme activity of rotenone-sensitive and ubiquinone-dependent electron transfer in complex I (18, 66, 68, 101-103). The 3460 mutation can reduce the maximal respiration rate about 20-28% of the control value and resulted in 79% reduction in specific complex I enzymatic activity assays (68, 101, 102). This mutation causes the substitution of threonine for alanine at amino acid position 52 of the ND1 subunit of complex I which was proposed to interact with rotenone and ubiquinone (104, 105). This mutation affects highly conserved amino acid in the inner membrane of mitochondria of this complex.

The mutation at nucleotide position 14484 of ND6 subunit of complex I, also effected respiratory chain function and complex I activity (68, 106-108). However, this mutation converted the moderately conserved methionine to valine in protein ND6. This might explain the mildest phenotype with visual recovery in up to 50% of patients with 14484 mutation (31, 50, 109). Whereas, LHON carried 11778 mutation showed visual recovery in 4-25% (31, 48, 49), and intermediary recovery rate at approximately 20% in cases with 3460 mutation (49, 51).

The biochemical studies in LHON cases with 11778 mutation were still ambiguous. Many reports revealed that the G11778A did not significantly decreased electron transfer activity of complex I (68, 101, 102, 110-113). However, deficiencies in respiratory chain function were believed to play vital roles in the pathogenesis of 11778 LHON. The measurements of the respiratory rate with NADH-linked substrates in the LHON carrying G11778A revealed lower activity of oxidation in complex I. This resulted in 30-50% reduction of complex I-linked NADH substrate (68, 110). Thus it was suggested that the 11778 mutation may defect the interactions of the dehydrogenases with complex I, without the activity alteration of complex I

itself (101, 110). The 11778 mutation was located in ND4 subunit of complex I which contains a number of charged amino acids that are evolutionarily conserved and embedded in mitochondrial transmembrane region. The replacement of charged amino acid arginine with histidine inducing by this mutation could affect the biochemical features (112). It might reduce the affinity of complex I for the ubiquinone substrate and increases resistance to rotenone, triggering a short circuit in electron transport of complex I. This phenomenon was observed by the decrease of respiration rate using NAD-linked substrates assay in LHON patients, while the effect in electron transfer properties conferred a little alteration (103, 112). Other studies on cytoplasmic hybrid (cybrid) cells and the ^{31}P-MRS spectra in brain and muscle of LHON patients were also supported the pathogenic role of 11778 mutation in impairing cell respiration (111, 113).

2.3 The role of ROS production and optic nerve degeneration in LHON

Alteration in mitochondrial energy generation, ROS production, and apoptosis can all contribute to pathophysiology of mitochondrial disease. Typical mtDNA mutations cause their mitochondria to generate more superoxide and other ROS than do normal mitochondrial DNA (114). Since, an optic neuropathy in LHON with mtDNA mutation might be involved either the loss of free radicals clearance function or the retinal ganglion cells death triggered by apoptotic pathway.

The respiratory chain dysfunction in LHON may cause ganglion cells and axon swelling through a pathway that activates the mitochondrial permeability transition and activation of redox-sensitive NMDA (*N*-methyl-D-aspartate) channel (5, 23). It was proposed to be an early step in apoptotic pathway (9-11). NMDA receptor channel is considered to spatial distribution on neurons and has proposed to induce the mitochondrial ROS formation via the excess of intracellular Ca^{2+}. In addition, NMDA receptor can also induce higher level of $Ca^{2+}$and another excitotoxic neuronal damage levels influx into neurons (115-118). This ability reflects the permeability of NMDA receptor-gated channel and the loss of mitochondrial membrane potential. Moreover, NMDA receptor also induced activation of phospholipase A which is a specific mediator of mitochondrial ROS production (119). Furthermore, NMDA receptor-mediated Na^+ loading and increased membrane Na^+, K^+-ATPase activity can enhance the mitochondrial ROS production. Consequently, an excitotoxic NMDA-receptor

activation in neuronal mitochondria are expected to affect neuronal cell function. Besides the effect of Ca^{2+} that can damage mitochondria, particularly superoxide anion can triggered the mitochondrial permeability transition. The accumulation of these radicals reflects the permeability of NMDA receptor-gated channel and the loss of mitochondrial membrane potential induced apoptotic cascade pathway and nerve cells injury. An optic neuropathy may then ensue.

Superoxide anions are resulted from respiratory chain dysfunction of complex I which is remarkable sites for the most primary LHON mutations (118). An increase levels of these radicals, not only affect the mitochondrial permeability transition but involved in the radicals clearance function of mitochondrial superoxide dismutase (SOD) (12, 120). The manganease superoxide dismutase (MnSOD) in mitochondria is encoded by *SOD2* gene located in the nucleus. This SOD catalyzes the dismutation of two superoxide radicals, producing H_2O_2 and molecular oxygen. In LHON, the flow of electrons that normally pass along the electron transport chain may be disrupted by the mutation in complex I gene. The freed electrons are now available to react with molecular oxygen, generating superoxide to levels beyond the capability of dismutation by endogenous levels of MnSOD that normally present within mitochondria. This oxidative stress may then damage cell and accumulate in cell death inducing the optic neuropathy (121, 122). Therefore, some polymorphisms that diminish the levels of mitochondrial SOD may reduce protection against superoxide. Although, no SOD2 polymorphism have yet been linked to LHON, at least two polymorphisms in the nuclear-encoded SOD2 may affect mitochondrial SOD activity (123-125).

2.4 Involvement of apoptosis to LHON

Other intriguing factors, *bcl*-2 and cytochrome *c*, are involved in regulation of apoptosis in the ganglion cell. It is attractive to suggest another possible step in LHON neurodegeneration. The respiratory chain dysfunction in LHON can initiate a pathway that involves mitochondrial swelling and release of cytochrome *c* prior to caspase activation. As an apoptosis is triggered, *bcl*-2 is associated with mitochondria-prevents cytochrome *c* releasing (13, 14, 118). Studies of cybrids with G11778A and G3460A mutations showed a higher sensibility for Fas-induced apoptosis (121). This pathway goes through mitochondria with the activation of Fas

receptor in cell membrane by Fas ligand and cause caspase-8 cleavage. Caspase-8 can cut proapoptotic protein causing its translocation to the outer mitochondrial membrane. Then the cytochrome *c* releases, this can regulate the apoptotic signal and activate caspase cascade. The respiratory chain dysfunction in LHON can initiate pathway that involves mitochondrial swelling and releasing of cytochrome *c* prior to caspase cascade activation (23). Or else the combination of complex I subunit mutations, partial deficiency of oxidative phosphorylation and increased of ROS production can impair mitochondria and activated cell death and optic neuropathy. Studies in cybrid cells of LHON patients revealed the propensity to undergo apoptosis and sensitized to apoptotic death than those controls (121, 126, 127).

3. Genetic studies approach for complex disease

Assessing the function of genetic variants in complex diseases is processed on the systematically scanning. Different molecular markers and methods have been developed to assay variation in DNA sequences (128). However, identification of the relevant regions contribute to disease was difficult. At present, genome-wide linkage analysis and the candidate-gene association studies have conferred as significant methodologies.

3.1 Linkage analysis

Linkage analysis is the method traditionally used to identify loci that underlies diseases by scanning the entire genome sequences of various members of families affected using a highly variable of genetic markers. Markers should be flanked the disease gene and segregate with disease in families study. The co-segragation with disease status can occur in markers within 10-20 cM (129). Linkage approach identify genetic region which is “in linkage” to disease by observing affected family members who share certain alleles located in the regions more frequent than would be expected by chance (130, 131). Although this method has been remarkably successful in identifying high relative risk gene, it has not been successful in identifying genes that are involved in the complex disease (129, 132-134). As for the most common diseases, linkage analysis has achieved only limited success. Whereas the genes discovered usually explain only a small fraction of the overall heritability of the disease (129, 132, 134-138). Nevertheless, this method is important for complex disorders that have no prior knowledge of the physiology or biology underlying the

disorder. This is obtained for extensive candidate gene studies exceeded to find the causal genes further.

The lack of success in linkage mapping might be the result of three main features of complex disease. First, such diseases typically vary in severity of symptoms and age of onset resulting in unproper defining an appropriate phenotype and selecting the best population to study. Second, they vary in their etiological mechanisms which might involve various biological pathways. Third, and perhaps most important, complex diseases are more likely to cause by several, and even numerous, genes, each with a small contribution and relative risk. Therefore, any individual genetic variant will generally have a relatively small effect on disease risk that could not be detected by linkage analysis (129).

3.2 Association study

In contrary, an association study using candidate-gene approach looks for a correlation between specific genetic variants and a disease. It is likely to be more effective tools than linkage for studying complex trait, because it has greater statistical power to detect several genes of small effect (139-141). In simplest form, association study compares the frequency of alleles or genotypes of a particular variant between disease cases and controls. It does not require large families with both affected and unaffected members, but can be performed either with unrelated cases and control subjects or with small families (only a proband and parents). Furthermore, the family-based association approach can alternatively be used to avoid the potential problem of population stratification. However, the selection of potential candidate genes relies on the basis of biological hypotheses. They should be evaluated for the involvement of specific genes to disease phenotype. In addition, linkage studies might provide information about candidate regions that can be explored further.

Molecular markers tested for association must be the causal allele or highly correlated (in linkage disequilibrium, LD) with the disease allele (129). Common variations in association studies have identified single nucleotide polymorphisms (SNPs) using haplotypes, single or multiple markers depending on association searching methods. The selection of validity SNPs should be adequate to provide functional effects in the genome (142). A priority for single polymorphism selection is shown in Table 2.

Table 2. Priorities for single-nucleotide-polymorphism selection (129).

Type of variant	Location	Functional effect	Frequency in genome	Predicted relative risk of phenotype
Nonsense	Coding sequence	Premature termination of amino-acid sequence	Very low	Very high
Missense/ non-synonymous (non-conservative)	Coding sequence	Changes an amino acid in protein to one with different properties	Low	Moderate to very high, depending on location
Missense/ non-synonymous (conservative)	Coding sequence	Changes an amino acid in protein to one with similar properties	Low	Low to very high, depending on location
Insertions/deletions (frameshift)	Coding sequence	Changes the frame of the protein-coding region, usually with very negative consequences for the protein	Low	Very high, depending on location
Insertions/deletions (in frame)	Coding or non-coding	Changes amino-acid sequence	Low	Low to very high
Sense/synonymous	Coding sequence	Does not change the amino acid in the protein — but can alter splicing	Medium	Low to high
Promoter/regulatory region	Promoter, 5′ UTR, 3′ UTR	Does not change the amino acid, but can affect the level, location or timing of gene expression	Low to medium	Low to high
Splice site/intron–exon boundary	Within 10 bp of the exon	Might change the splicing pattern or efficiency of introns	Low	Low to high
Intronic	Deep within introns	No known function, but might affect expression or mRNA stability	Medium	Very low
Intergenic	Non-coding regions between genes	No known function, but might affect expression through enhancer or other mechanisms	High	Very low

4. Search for LHON modifier in Thai families

In Thailand, 44 families with primary LHON mutation were recruited. Forty-two (93%) families carried the G11778A mutation and 3 (7%) families with T14484C mutation. Age of onset is higher in males. Mean age of onset (± SD) in male was 20.9 ± 9.3 years with the median at 19 years of age; while in female was 28.6 ± 14.6 years with the median at 30 years of age. Their visual deterioration occurs with a broad range from 6 to 44 years for male an 10 to 53 years for female respectively. The predilection for males to lose vision was observed. Of 79 LHON cases with G11778A mutation, 60 (76%) were of males and 19 (24%) were of females (male: female ratio is 3.2:1). The analyses of our G11778A LHON pedigrees showed some characteristics different from those of Caucasian and Japanease (Phasukkijwatana *et al*., unpublished data). Our studies found the higher prevalence of G11778A heteroplasmy which seemed to influence the disease manifestation. However, no association between the heteroplasmicity and the disease-onset was found. The estimated overall penetrance of our G11778A LHON population was 37% for males, and 13% for females. Further, our LHON pedigrees showed an incomplete penetrance as in majority of LHON pedigrees worldwide.

Whereas LHON has been far more complicated disease, it might be the result of multiple interactions from genetic factors, genetic susceptibility loci, environmental, and quantitative traits. Since the pathogenesis of LHON was reported to associate with the enzyme activity in complex I. Polymorphisms in complex I genes might contribute to the complex I dysfunction in LHON. Since complex I subunits were expressed by both mitochondrial and nuclear genome, the nuclear-encoded genes for subunits of complex I are strong candidate genes for LHON modifying factors

CHAPTER III
MATERIALS AND METHODS

MATERIALS

1. Subjects

1.1 Thai normal controls from different part of Thailand for SNPs database generation

Nineteen normal controls composed of 2 Northern, 8 Central, 2 Northeastern, 1 Southern and 6 of mixed parent hometown were included in this study for the generation of our SNPs database. Their ethnic origins were classified by the family information and could be traced back for at least three generations. Their details were shown in Table 3. All of them were healthy, had neither genetic diseases nor clinical phenotype of LHON which was examined by the physician. They had no primary LHON mutation at nucleotide position 11778 by routine molecular screening. All blood samples were taken with informed consent.

1.2 Thai normal controls for association study

The whole blood from another 115 Thai normal controls was collected with their informed consent. They were from Check up clinic, Siriraj Hospital and were mixed-urban Thai population. All of them were healthy, had no genetic diseases, and had no the clinical phenotype of LHON.

1.3 LHON pedigrees

Forty-four unrelated LHON pedigrees were included in this study. Forty-two pedigrees carried G11778A mtDNA mutation and were of Thai and/or Chinese origins except for one pedigree with was Indian origin. Another 2 pedigrees were T14484C LHON mtDNA mutation (only blood samples from probands were obtained). Total of 292 individuals from 42 LHON pedigrees of G11778A mutation, comprising of 79 patients (affected persons), 2 unknown status, 124 individuals who did not show clinical phenotype of LHON but carried the G11778A mutation (unaffected) and 8 unaffected of maternal lineage but no G11778A mutation and 79 of non-maternal

Table 3. The ethnic origin of nineteen Thai normal controls from different part of Thailand.

Code	Sex	Ethnic origin[a]	Age[b]
NR1	M	N	34
NR2	M	C	26
NR3	F	C	49
NR4	F	C	34
NR5	M	C	31
NR6	F	M	33
NR7	F	NE	37
NR8	M	M	36
NR9	F	M	28
NR10	F	M	34
NR11	M	M	26
NR12	M	M	25
NR13	M	NE	34
NR14	F	N	27
NR15	M	S	26
NR16	F	C	23
NR17	M	C	24
NR18	F	C	25
NR19	M	C	26

[a] Individual ethnic origins were classifiied by the family information that could be traced back for three generations.
N = Northern
C = Central
NE = Northeastern
S = Southern
M = Mixed parent hometown

[b] Age at 2005.

lineage from these families, were studied.

For association study, 76 LHONs and 132 unaffected were included. Unaffected males and females whose their age lowers than 30 and 45 years respectively were excluded. The cutoff-age was calculated from mean onset age in our LHON patients of each sex plus 1SD (mean+1SD). All of them carried mutant mtDNA in their blood for more than 95%.

2. Oligonucleotide primers

The oligonucleotide primers used in this study were synthesized by Invitrogen Custom Primers, Invitrogen Japan K.K. The primers sequences used were shown in Table 4.

3. Restriction enzymes

Six restriction endonuclease enzymes used in the PCR-RFLP analysis were *Acl*I, *Bsr*BI, *Bsm*AI, *Bsu*36I, *Hpy*CH4III, and *Scr*FI. They were purchased from New England Biolabs, USA.

4. Equipment

- Submarine Agarose Gel Unit, Minnie the Gel-Cicle™ HE 33, Hoefer, U.S.A.
- Submarine Agarose Gel Unit, Sub-Cell GT System, Bio-Rad, U.S.A.
- UV generator, TFX-40.M, Vilber Lourmat, France
- Bench-top centrifuge, Centrifuge 5417C, eppendorf, Germany
- Tabletop centrifuge, Sorvall TC6, DuPont, U.S.A.
- Thermal cycler, *Px2* Thermal Cycler (PCR License Registration), Thermo
- Hybaid, U.S.A.
- Gel documentation systems, Gel Doc™EQ, Bio-Rad, U.S.A.
- Pipette, Pipetman P, Gilson, France
- Pipet tips/microtubes, Axygen, U.S.A.
- Thin-walled PCR tubes, Sorenson, U.S.A.

Table 4. Primers used to determine the DNA sequence in *NDUFS4, NDUFS8,* and *NDUFV2* gene.

Gene name	Code	Primer sequences (5'-3')	Length (bp)
NDUFS4	PFS4U	GAAAATTTTCTCTGAGGCATCTCT	24
	PFS4C	CTATTGATTCTTCATATAA	19
	PFS4L	GAAAGGGCAGCTACAGCCACTG	22
	FS4U	GTGATCCGTCCTTTCATCCTG	21
	FS4L	ATACTGCACAGCTGACTTTATTCAC	25
	S4X1L	CTCCGGAACTCAGCCAGCAGA	21
	S4X3U	TTGCATGAATATAGGAAACA	20
	S4X3L	AGAAAAAGATAGCTGACAATA	21
NDUFS8	PFS8U	AGCCTGCCCTTGTCCTCGTC	20
	PFS8C	TTCGCTGGTAGGAAACATAG	20
	PFS8L	CGGCCAAAACGTGAACACC	19
	FS8U	CCAATGGCAGCGTCCTACAGTG	22
	FS8L	GGGGTCGGAGTGGTTTTATTGG	22
NDUFV2	PFV2U	AAGGAGTGGGAACCAAGGCATTTG	24
	PFV2C	TCACCTCCCGCTCCCCACTA	20
	PFV2L	CTTCCCCAGCCGCGCCAG	18
	FV2U	CGCGCTCGGGATTCTCG	17
	FV2L	CAAATAAGTTTACGTAGATTGGAAGTCC	28
	V2X2U	TAAGTGAAAGCTGGCATCCAAGTT	24
	V2X2L	TAAAGGACATAAGCCAAAAGAGGA	24

5. Software

The oligonucleotide primers and restriction endonuclease enzymes for PCR-RFLP analysis were designed by WinStar software package (DNASTAR, Madison, Wis.) and pDRAW32 version 1.0 (http://www.acaclone.com) respectively. Sequences were aligned to multiple sequences and single nucleotide polymorphisms (SNPs) were detected by GenalysWin program version 2.0 beta (http://software.cng.fr/) and BioEdit version 7.0.0.1 (http://www.mbio.ncsu.edu/Bioedit/bioedit.html). The error of genotyping data of all pedigrees members which violated Mendelian laws of inheritance were corrected using PedCheck program version 1.1 (http://watson.hgen.pitt.edu/register). This program was useful for single-locus checks and quickly analyzed many loci with large number of alleles. The GOLD program (http://www.sph.umich.edu/csg/abecasis/GOLD/) version 1.1.0.0 was used to prepare SimWalk2 input files for making analysis of disequilibrium patterns and haplotype estimation in family data on the SimRun/SimWalk 3.0b1 program (http://www.genetics.ucla.edu/software/). HaploPainter program version 024 beta (http://haplopainter.sourceforge.net) was used to paint with haplotypes generated from SimWalk2 program. For association study using family-based approach, the data was analyzed on FBAT program version 1.5.5 (free downloaded from http://www.biostat.harvard.edu/~fbat/default.html) which was available for linkage as well as association test. This program could test for association with either single marker locus or haplotype marker and was useful to avoid spurious associations caused by admixture of populations. MedCalc software version 7.6.0.0 and Intercooled Stata version 7.0 were used for standard statistical analysis.

METHODS

1. Generation of Thai SNP database of *NDUFS4*, *NDUFS8*, and *NDUFV2* gene

The whole blood drawn from 19 Thai normal controls were collected with EDTA-anticoagulant. Total RNA and genomic DNA were extracted and both promoter and coding region of *NDUFS4*, *NDUFS8*, and *NDUFV2* gene were amplified and sequenced.

1.1 Amplification of coding region (by RT-PCR)

1.1.1 RNA extraction by TRIZOL® Reagent

Ten milliliters of whole blood obtained from 19 normal controls was collected in 2 ml of 5% EDTA (Ethylenediaminetetraacetic acid) solution and diluted by the addition of an equal volume of PBS (Phosphate Buffered Saline). Diluted blood from one sample was slightly layered over 3 tubes of 4 ml Lymphoprep ™ (AXIS-SHIELD, Norway) before centrifugation at 800xg for 30 minutes at room temperature in a swing-out rotor. The lymphocyte layer at the interface from 3 tubes was then removed to a new centrifuge tube using Pasteur pipette. The lymphocyte fraction was diluted with 10 ml PBS and the lymphocyte pellet was collected after centrifugation for 10 minutes at 250xg. This step was done twice. The pellet was then lysed in 1 ml of TRIZOL® Reagent in 1.5 ml microcentrifuge tube. The homogenized sample was incubated for 5 minutes at room temperature. Two hundreds microliters of chloroform was added in sample tube and shaken vigorously by hand for 15 seconds and the tube was incubated at room temperature for another 3 minutes before centrifugation at 12,000xg for 15 minutes. The mixture was separated into 3 phases, a lower red which was a phenol-chloroform phase, an interphase, and a colorless upper aqueous phase. Total RNA was in the upper aqueous phase. The RNA phase was transferred to a fresh tube and precipitated by mixing with 500 µl isopropyl alcohol. The tube was incubated at room temperature for 10 minutes and then centrifuged at 12,000xg for 10 minutes. The RNA pellet was washed with 1 ml of 75% ethanol followed by centrifugation at 7,500xg for 5 minutes. The supernatant was discarded and the pellet was air-dried for 5-10 minutes at room temperature. The RNA pellet was dissolved with 10-50 µl of RNase-free water depended on the pellet size. The RNA solution was stored at -80 °C until use.

1.1.2 Determination of RNA concentration and purity

The dissolved RNA was diluted to 1:100-1:500 with RNase-free water. The concentration and purity of RNA sample was determined by spectrophotometric method using Shimadzu UV-160 spectrophotometer. A pure sample of RNA should have an $A_{260/280}$ ratio of 1.6-1.8 as the manufacturer's suggestion. An absorbance measurement at wavelength 260 nm could permit the

direct calculation of RNA sample in this manner:

$$[RNA]\mu g/ml = A_{260} \times dilution \times 40.0,$$

where A_{260} = absorbance (in optical densities) at 260 nm

dilution = dilution factor (usually 200-500)

40.0 = extinction coefficient of RNA

The integrity of RNA was determined by nondenaturing agarose gel. The RNA sample was heat to 65°C for 5 minutes just prior to electrophoresis in order to disrupt secondary structure. Two microliters of RNA was mixed with 1.2 µl loading dye and analyzed on 1% agarose gel electrophoresis in 1X TBE buffer comparing with λ DNA/*Hin*d III Fragments. The agarose gel was then stained with ethidium bromide and visualized by UV transilluminator. The bands of predominant 28S (~5kb), and 18S (~2kb) ribosomal RNA (rRNA) with a minimum of smearing above, between, and below these two rRNA bands were detected.

1.1.3 RT-PCR amplification

Total RNA sample was used as template for reverse transcription to cDNA. The coding region of each the *NDUFS4*, *NDUFS8*, and *NDUFV2* gene was amplified by QIAGEN® OneStep RT-PCR Kit (QIAGEN, Germany) with gene-specific primers (Table 5). The reaction in a total volume of 50 µl contained 0.5-1 µg of RNA, 10 µl of 1x QIAGEN OneStep RT-PCR Buffer, 2µl of 400 µM of dNTP Mix (containing 10 mM of each dNTP), 1µl of 0.6 µM of gene specific primer, and 2 µl of QIAGEN OneStep RT-PCR Enzyme Mix. The RT-PCR condition comprised a reverse transcription step at 50°C for 30 minutes, followed by an initial PCR activation step at 95°C for 15 minutes. The amplification was performed for 30 cycles of 1 minute at 94°C, 45 seconds at the corresponding annealing temperature of each gene, 1.5 minutes at 72°C and finished with a final extension at 72°C for 10 minutes (using Thermo Hybaid/*Px2* Thermal Cycler). Five microliters of the RT-PCR product was mixed with 1.2 µl loading dye and analyzed on 2% agarose gel electrophoresis in 1X TAE buffer comparing with 100 bp DNA Ladder. The agarose gel was then stained with ethidium bromide and visualized by UV transilluminator.

Table 5. Amplification primers for promoter and coding region of *NDUFS4, NDUFS8,* and *NDUFV2* gene.

Gene name	Primer name	Product size (bp)	Annealing temperature(°C)
Promoter region			
NDUFS4	PFS4U PFS4L	1122	65
NDUFS8	PFS8U PFS8L	1120	65
NDUFV2	PFV2U PFV2L	1101	65
Coding region			
NDUFS4	FS4U FS4L	621	55
NDUFS8	FS8U FS8L	769	62
NDUFV2	FV2U FV2L	875	57

1.2 Amplification of promoter region (by PCR)

1.2.1 DNA extraction by standard phenol-chloroform

Ten milliliters whole blood was collected in 2 ml of 5% EDTA solution and centrifuge at 1,000xg for 5 min. The plasma was discarded and the bottom layer, containing packed red cells and buffy coat, was kept at -20°C for 2 hours or overnight. The packed cell and buffy coat was thawed at room temperature and mixed with 10 to 12 ml of Red Cell Lysis Buffer (RCLB) prior to centrifugation at 3,000xg for 3 minutes. The supernatant was partially removed by Pasteur pipette, leaving 2 ml of the solution in the tube. The leukocytes pellet was packed at the bottom of tube. Ten milliliters of RCLB was again added and mixed vigorously by vortexing. The mixture was then centrifuged at 3,000xg for 3 minutes and the supernatant was discarded by pouring. The red cell lysis step was repeated until the clear leukocyte pellet was obtained. The pellet was suspended with 250 µl of distilled water and transferred to a new microcentrifuge tube. The leukocyte suspension was mixed with 250 µl of lysis buffer and 25 µl of 20 mg/ml proteinase K and then incubated at 65°C for 2 hours or 37°C for overnight. The leukocyte suspension was then mixed with 500 µl of saturated phenol and centrifuged at 12,000xg for 1 minute. The mixture was separated into a lower red which was phenol phase, an interphase, and a colorless upper aqueous phase. DNA remained in this aqueous phase. DNA supernatant phase was transferred to a fresh microcentrifuge tube and the extraction step using phenol was repeated. The supernatant was mixed with 250µl of saturated phenol and 250 µl of isoamyl:chloroform (1:24) followed by centrifugation of the mixture at 12,000xg for 1 minute. The supernatant was transferred to a new microcentrifuge tube and mixed with 500 µl of chloroform-isoamyl alcohol solution before being centrifuged at 12,000xg for 1 minute. Genomic DNA in the supernatant was precipitated with 1 ml of 100% ethanol and then kept at -20 °C overnight. The DNA pellet was obtained by centrifugation at 12,000xg for 20 minutes. The pellet was collected and washed twice with 1 ml of 70% ethanol followed by centrifugation as described. The DNA was air-dried at room temperature and dissolved in distilled water and stored at -20°C until use.

1.2.2 Determination of DNA integrity

To determine the integrity of DNA, 1 μl of DNA was mixed with 1.2 μl loading dye and analyzed on 1% agarose gel electrophoresis in 1X TAE buffer comparing with λ DNA/*Hin*d III Fragments. The agarose gel was then stained with ethidium bromide and visualized by UV transilluminator.

1.2.3 PCR amplification

Genomic DNA was used as template for the amplification of promoter region of each *NDUFS4*, *NDUFS8*, and *NDUFV2* gene with gene-specific primers (Table 3). The PCR reaction was carried out in a total volume of 50 μl containing 5-10 ng of DNA, 10 mM Tris-HCl pH 9.0, 50 mM KCl, 1% of Triton X-100, 1.25 mM of $MgCl_2$, 200 μM of each dNTP, 0.4 μM of each primer, and 1.25 U of *Taq* DNA polymerase. The PCR thermal profile consisted of an initial denaturation step at 94°C for 5 minutes, 30 amplification cycles of 1 minute at 94°C, 1 minute at the corresponding annealing temperature of each gene, 1.5 minute at 72°C, and finished with a final extension at 72°C for 10 minutes (using Thermo Hybaid/*Px2* Thermal Cycler). Five microliters of the PCR product was analyzed on 2% agarose gel electrophoresis in 1X TAE buffer as described in 1.1.3.

1.3 Direct sequencing

Promoter and coding region of three candidate genes, *NDUFS4*, *NDUFS8*, and *NDUFV2* gene were amplified and sequenced with gene-specific primer sets. The sequencing primers used were in Table 2. The nucleotide sequences were aligned and compared with the promoter and coding sequences database of each the *NDUFS4*, *NDUFS8*, and *NDUFV2* gene. GenBank accession numbers for coding region of *NDUFS4*, *NDUFS8*, and *NDUFV2* gene were NM_002495, NM_002496 and NM_021074 respectively. The accession numbers for 1.1 Kb promoter region of *NDUFS4*, *NDUFS8*, and *NDUFV2* gene from Eukaryotic Promoter Database (SIB-EPD) were EP73752, EP73753, and EP74377 respectively.

1.4 SNP detection and validation

Bi-directional sequences of our 19 Thai normal controls were read with both GenalysWin and BioEdit program in order to identify polymorphisms. SNPs were chosen toward non-synonymous for coding region and on the basis of their putative biological relevance for promoter SNPs. SNPs of 19 normal controls were validated

on another molecular genetic analysis platform and genotyping results were concordant with those determined from sequencing. The validated SNPs should have minor allele frequency at least 10% in our controls. These SNPs were genotyped in our LHON pedigrees members and normal controls for further association analysis.

2. SNP genotyping of 115 Thai normal controls and 292 individuals from LHON pedigrees

The validated SNPs chosen from 1.4 were genotyped in 292 LHON pedigrees members and 115 controls. The validated SNPs genotyping was performed by PCR and restriction fragment length polymorphism (RFLP).

2.1 DNA extraction by standard phenol-chloroform

Ten milliliters whole blood of all samples was collected in 2 ml of 5% EDTA solution and genomic DNA extraction was performed as described in 1.2.1.

2.2 SNP detection by PCR-RFLP

The nuclear DNA spanning validated SNPs in the promoter and coding region of each the *NDUFS4*, *NDUFS8*, and *NDUFV2* was amplified using gene-specific primers (Table 6). The PCR amplification was performed and run on Thermo Hybaid/*Px2* Thermal Cycler in a total volume of 50 µl with PCR-reactions as described in 1.2.3. The PCR thermal profile consisted of an initial denaturation step at 94°C for 5 minutes, 30 amplification cycles of 1 minute at 94°C, 30 seconds at the corresponding annealing temperature of each gene, 1 minute at 72°C, and finished with a final extension at 72°C for 10 minutes. Five microliters of the PCR product was analyzed on 2% agarose gel electrophoresis in 1X TAE buffer as described in 1.1.3. For RFLP analysis, 5-10 microliters of the PCR product was digested with 1-3 units of appropriate restriction enzymes reaction (New England Biolabs). The digestion patterns were shown in Table 4. The digested products were analyzed on 2% or 3% agarose gel electrophoresis in 1X TAE buffer.

3. Searching for nuclear modifier gene susceptibility to LHON manifestation using association study

The case-control study was used to identify the association between genetic variants that contribute to disease phenotype in our study. The association between

Table 6. PCR-RFLP analysis of *NDUFS4*, *NDUFS8*, and *NDUFV2* gene.

SNP position	Primer name	Product size (bp)	Annealing temperature (°C)	Restriction enzyme	Change of Restriction pattern (bp) results from SNP
Coding region					
NDUFS4					
SNP +12 C→G	FS4U S4X1L	217	61	*Bsm*A I	145/54/18→199/18
SNP +312 G→A	S4X3U S4X3L	303	51	*Scr*F I	217/86→303
NDUFV2					
SNP +86 T→C	V2X2U V2X2L	1114	62	*Hpy*CH4 III	441/395/278→441/673
Promoter region					
NDUFS4					
SNP -447 T→A	PFS4U PFS4L	1122	65	*Bsu*36 I	598/524→1122
NDUFS8					
SNP -496 C→G	PFS8U PFS8L	1120	65	*Acl* I	602/518→1120
NDUFV2					
SNP -602 G→A	PFV2U PFV2L	1101	65	*Bsr*B I	518/485/98→1003/98

genetic variants and disease phenotype could be test by the application of linkage disequilibrium (LD) which also known as allelic association. When genotype and allele frequencies at the marker SNPs were differ between cases and controls, the SNP was appeared as disease association.

In order to identify the nuclear modifier gene susceptibility to LHON expression, the association of the SNPs validated from 1.4 was tested using the family-based association analysis that was conducted to avoid possible confounding of gene-phenotype association due to inappropriately chosen controls or population stratification. By using this approach, parental alleles were classified into those transmitted and not transmitted to affected children. Parental alleles not transmitted were used as controls and they were able to evaluate the risk of disease arising from particular allelic markers. According to this approach, the FBAT (family-based association tests) program was used.

Prior to perform family-based association test, pedigrees genotyping data were assessed in order to eliminate confounding from incompatibility with Mendelian inheritance. Data was prepared for running with PedCheck program in order to check erroneous genotyping of each single-locus polymorphism. Likewise, the haplotype of polymorphisms was estimated in family data and used to determine founder haplotype by the Simwalk2 program. It was automated interface to the GOLD (graphical overview of linkage disequilibrium) package and made analysis of disequilibrium patterns by measurement pair-wise in a region. Finally, prepared data was run with FBAT program. Either single and haplotype marker locus were used to determine the parental transmission pattern and performed under the FBAT statistic with the null hypothesis of no association. Genotype and allele frequencies of our population including LHON individuals were confirmed by the Hardy-Weinberg equilibrium analysis. The schematic of SNP database generation and association study were shown in Figure 4 and 5 respectively.

3.1 Pedigree error checking with PedCheck program

Pedigree structure and genotypes were prepared as a pre-makeped linkage file referred to ‘pedigree.dat’. The additional file, ‘marker.dat’, listed all marker loci that used in the pedigree file. These two files were prepared on Microsoft notepad and then renamed from text file (*.txt) to dat file (*.dat) in dos command prompt before

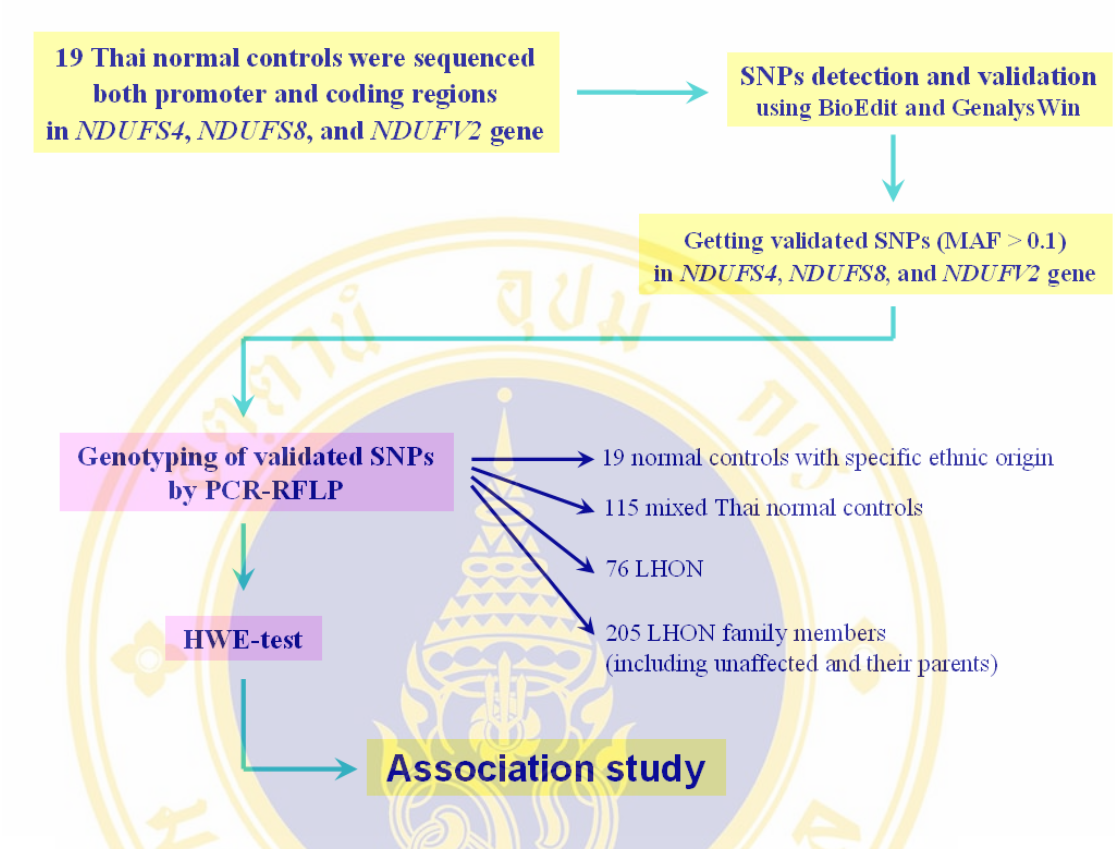


Figure 4. The schematic of Thai SNPs database generation for *NDUFS4*, *NDUFS8*, and *NDUFV2* gene.

MAF = Minor Allele Frequency

HWE = Hardy-Weinberg Equilibrium

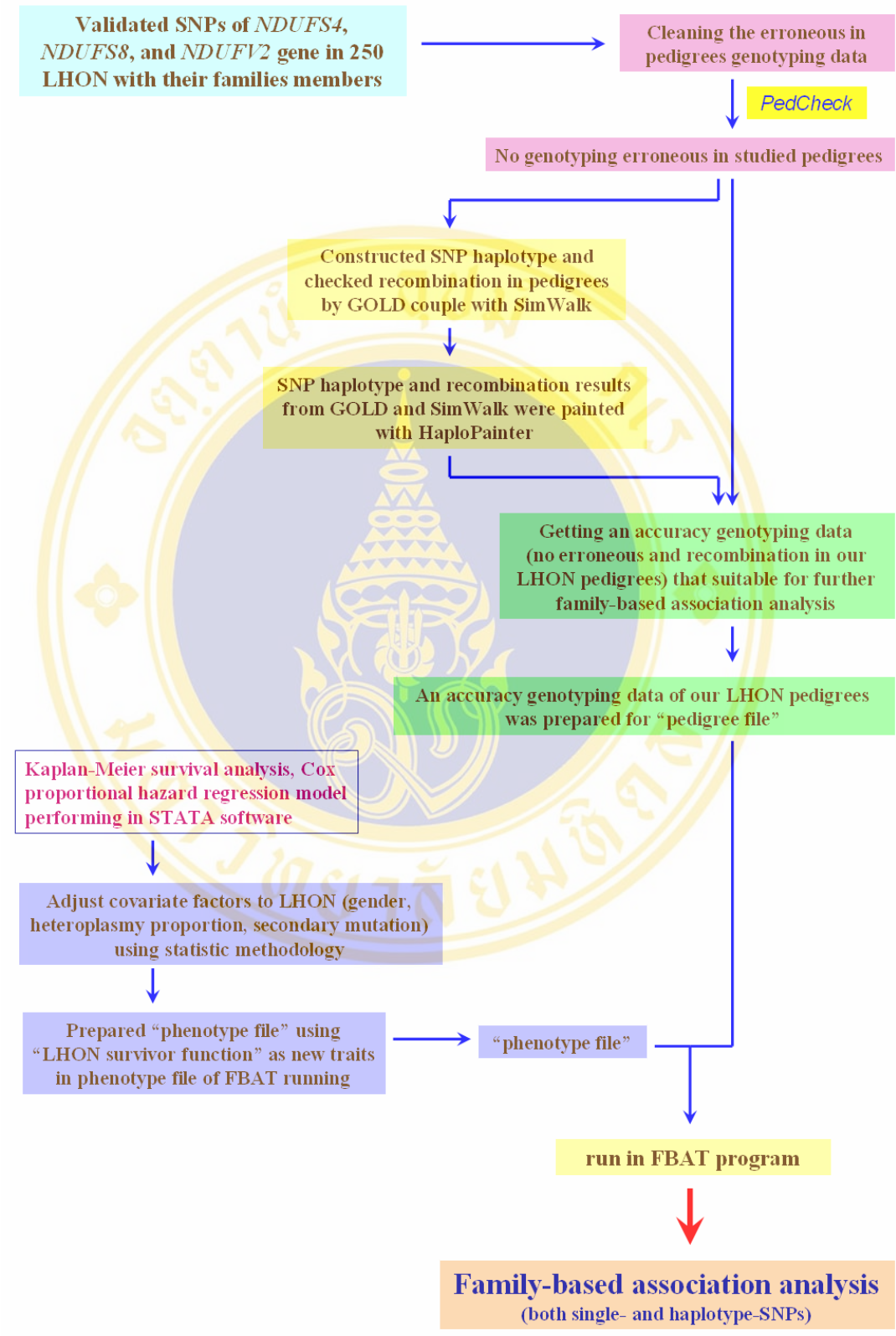


Figure 5. The schematic of family-based association study of LHON survival functions with validated SNPs in *NDUFS4*, *NDUFS8*, and *NDUFV2* gene.

ready to run PedCheck. The standard pedigree file format of the first line, which was referred to pid, id, fid, mid, sex, aff, A11 A12 A21 A22, and etc., was replaced with data format as shown in Table 7. To create a marker file format, the first line was preceded by an 'X' which told PedCheck that the first column after the pedigree structure and sex contained the affection status that was skipped. The remaining lines were the names of markers genotyped. Example of pedigree and marker file was shown in Figure 6a and 6b respectively. Genotyping errors in Mendelian inheritance and inconsistencies within pedigrees were checked for 4 levels as the following command. Begin with PedCheck in dos window, level 0 and level 1 checked on the nuclear family level. It was done with command "-n <marker>.dat -p <pedigree>.dat". Level 2 checking was continued when any inconsistencies in the pedigrees data was not found. It was produced with command "-n <marker>.dat -p <pedigree>.dat -2". Level 2 was done on genotype-elimination algorithm. If there were no level 2 errors detected, the pedigree was Mendelian consistent. Level 3 and level 4 checking were used to determine critical genotypes in the most likely person to be in error by the same command as level 2 but replacing '-2 with '-3' or '-4' respectively. The genotyping data was resolved and any genotyping errors at each level was corrected until the re-running PedCheck did not indicated any source of errors. The erroneous information was screened to a file named 'pedcheck.err' and overwritten each time PedCheck was run.

3.2 Disequilibrium analysis and haplotype estimation in LHON family

The linkage disequilibrium analysis was done on the GOLD program. Two files required for GOLD composed of the pedigree file in standard pre-makeped linkage and the map file format (*.dat file). The pedigree file was prepared the same as in 3.1, pid, id, fid, mid, sex, A11 A12 A21 A22, and etc., except the affected status data was no needed (Table 7). The map file format described the physical location of markers. Pedigree and map file was shown in Figure 7a and 7b respectively. The GOLD was called from dos command prompt and the pedigree and map file was read by the following command "setup-simwalk2 -p<pedigree>.dat -m <map>.dat". This command was used to set up input files for estimated haplotype on SimWalk2 running. Three files from GOLD preparing, BATCH.DAT, LOCUS.DAT and PEDIGREE.DAT, were automated interface to SimWalk2 program with editing of

Table 7. The standard pedigree data in lingkage format.

pid	pedigree ID
id	individual ID
fid	father ID Use 0 (zero) for founders or marry-ins (parents not specified) in a pedigree
mid	mother ID Use 0 (zero) for founders or marry-ins (parents not specified) in a pedigree.
sex	1 = male, 2 = female
aff	affection status 2 = affected, 1 = unaffected, 0 = unknown
A_{ij}	allele j of marker i (j = 1,2 ; i = 1,2,... Alleles are represented by positive integers. Use 0 (zero) for missing alleles.

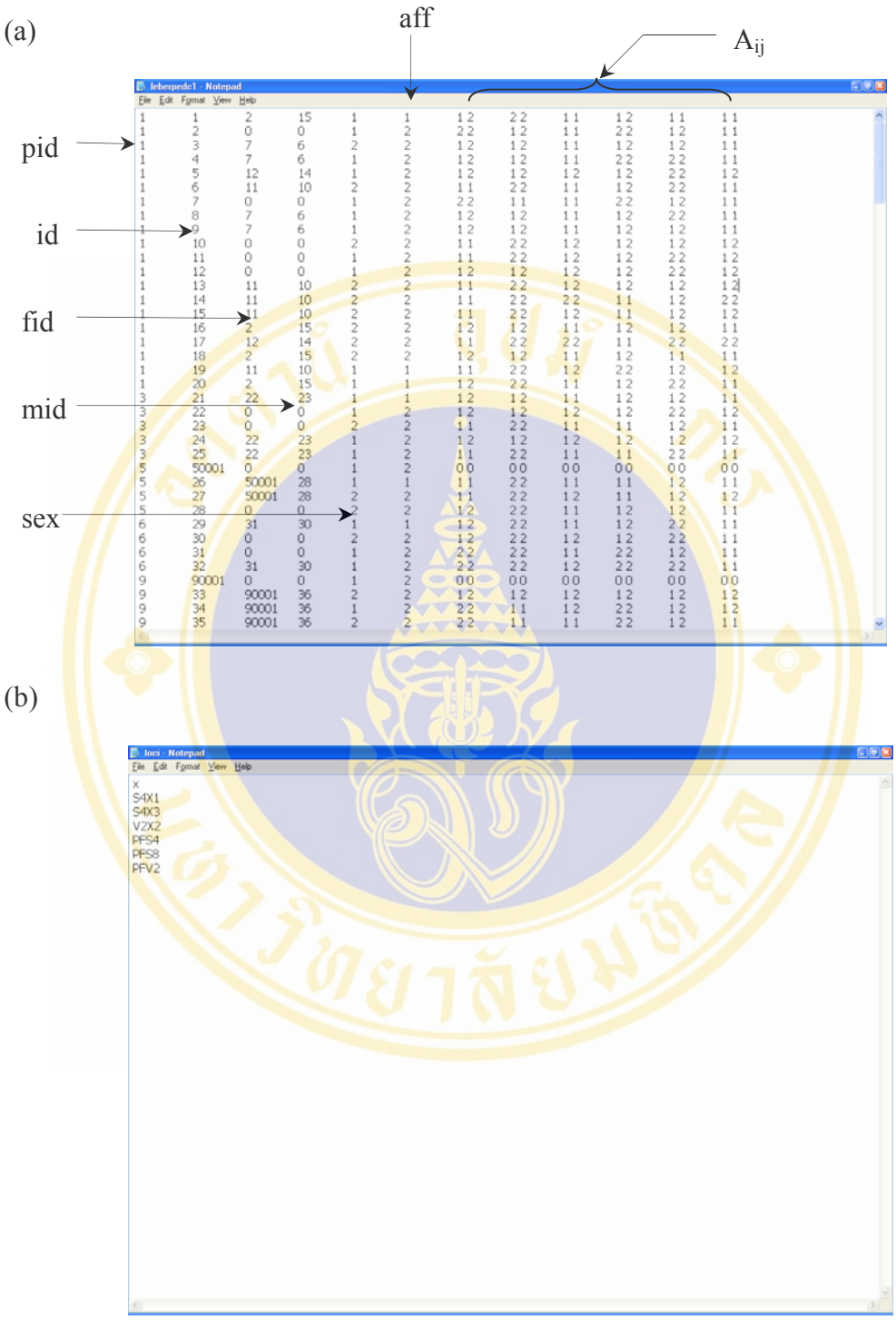


Figure 6. File prepared for PedCheck running. (a) Pedigree file in standard pre-makeped linkage format described pedigree 1,3,5,6, and 9 contained founders and their offspring which genotyped at 6 markers. (b) Marker file described all names of genotyped markers.

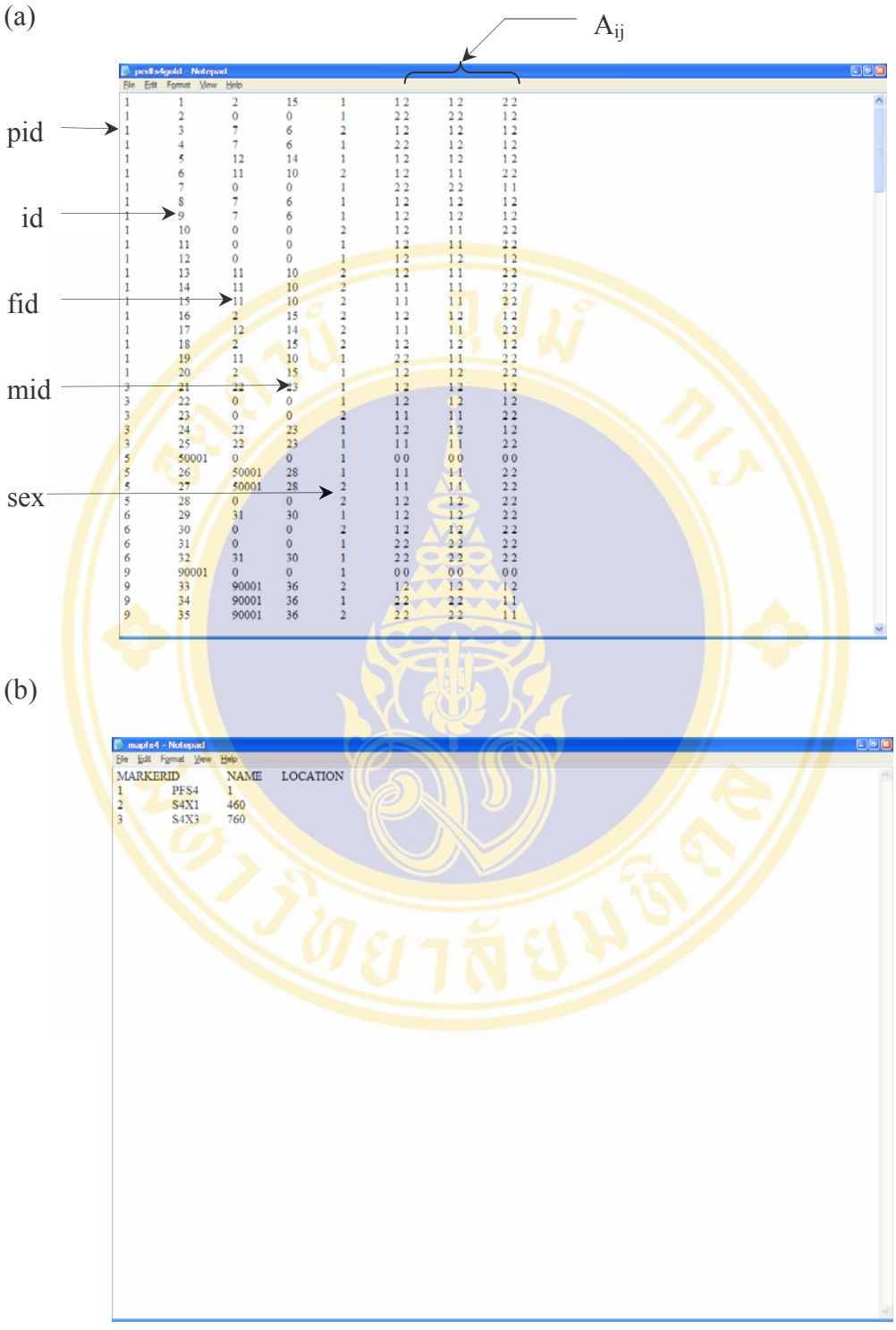


Figure 7. File prepared for GOLD software. (a) Pedigree file in standard pre-makeped linkage format with 3 genotyed for marker haplotyping. (b) Map file indicated the marker location in the pedigree file on the physical map, in base pairs.

BATCH file. The 000046 item was replaced N Y N with N N N and this file was overwritten. The next step, haplotype from pedigree data was estimated, the edited BATCH file was read by the "Read in Batch File" tool bar of SimWalk2 (SimRun/SimWalk 3.0b1). After the edited BATCH file was read, data was analyzed for haplotype by "Run SimWalk" bar. The haplotyping results from SimWalk2 program coupling with GOLD program were in text file format (Microsoft®Notepad). The "HAPLO-01" files generated from SimWalk2 program were used to paint pedigrees with their haplotype by HaploPainter program.

3.3 Family based association analysis with LHON phenotype using FBAT program

Family-based association analysis was performed as transmission/disequilibrium test (TDT) in order to detects linkage disequilibrium or causative association between the alleles of marker locus and a genetically influenced disease. The Family Based Association Test (FBAT) program based on this approach was performed and more powerful than a conventional TDT. The FBAT could allowed families with missing parental data for the analyses by either reconstructing parental genotypes from children (RC-TDT), or using unaffected sibs as controls (Sib-TDT). Multiallelic loci could be analyzed using either an overall multiallelic test or a biallelic test in which each allele is tested against all others. The FBAT approach was adjusted to use an empirical variance-covariance estimator for the statistic, which is consistent when sibling marker genotypes are correlated (e.g., when there was linkage and the analysis included multiplex families).

Data format for usage with the FBAT software was prepared into 2 files composing the pedigree and phenotype data file. The pedigree data file had been format to pre-makeped linkage file from 3.1 ("pedigree.dat" file). But FBAT file was "*.ped" file ("pedigree.ped" file) and the first line with the name of all markers of the genotype data was needed (Figure 8a). The additional phenotype data file was performed to test FBAT with a quantitative trait. Data file needed pid, id, trait1, trait2, and etc. which was replaced with data format as shown in Table 7. The first line of this file was included with the names of all traits in the phenotype file (Figure 8b). The phenotype data file was renamed to "*.phe" file. FBAT program was run in dos window. Both pedigree and phenotype data file were loaded by the command "load

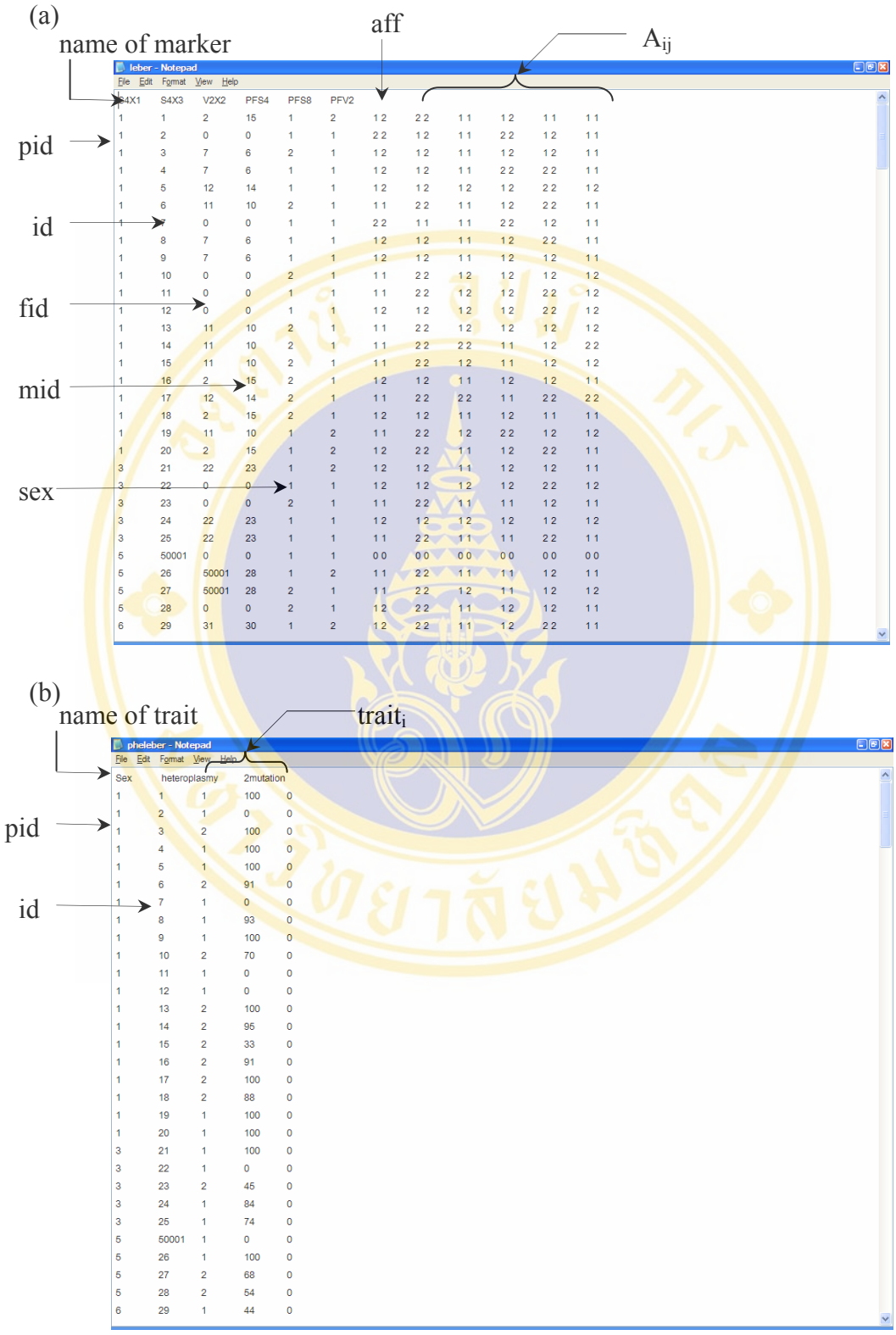


Figure 8. File prepared for FBAT software. (a) Pedigree file in standard pre-makeped linkage format and their offspring which genotyped at 6 markers. (b) Phenotype file described all traits in the phenotype.

*.ped" and "load *.phe" respectively. Data files were taken into FBAT calculation. The software command was shown in Table 8.

3.4 Hardy-Weinberg analysis

In this study, Hardy-Weinberg equilibrium was performed in order to test for population admixture. If a population composed of a recent admixture of different ethnic groups that differ in marker allele frequencies and disease phenotypes, spurious associations might be resulted between the marker genotypes and the complex traits. The HWE was tested at the candidate gene in association studies in order to reduce false positive finding of genes underlying complex traits. Most case-control association study, controls were reasonable to check for consistency with HW proportion, while genotypes of cases might not be in equilibrium. Additionally, the HWE test was also a general tool in population and evolutionary genetics in validating the assumption of the HWE.

The Hardy-Weinberg theorem was a state for using any set of allele frequencies from the observed genotype data of the population of each locus to calculate the expected genotype frequencies. The difference between expected and observed genotype frequencies would be compared with a statistical method of goodness-of-fit test (Chi-square). To calculate HWE in population, with dominant allele A and recessive allele B on observed genotype AB, p was let to be frequency of the dominant allele A, and q of the recessive allele B respectively. The HWE calculation was performed on the equation, $p^2 + 2pq + q^2 = 1$. The p and q values were calculated from the observed genotype data, where p^2 was probability of AA occurring, 2pq was of all heterozygote AB, and q^2 was of BB respectively. These p and q values were used to invert calculation (p^2 for AA, 2pq for AB and q^2 for BB genotype respectively) to each expected genotype frequencies by multiplication with the total number of population. These expected results were the actual numbers of population expected if the population were in HWE. Significant difference between number of actual and expected genotypes were determined using the statistical test known as Chi-square goodness of fit. If there was no difference between the observed and expected values; the population was in Hardy Weinberg equilibrium. On the other hand, if there *was* a significant difference, therefore the population was *not* in Hardy-Weinberg equilibrium. The significance of the difference were prepared into the

Table 8. The FBAT command.

Type	"FBAT" Command	Brief Description
General	? [command]	Help on specified command; "?" lists all available "FBAT" commands
	Quit	Exit "FBAT"
Data input/output	load [ped,phe] file_name	Loading pedigree and trait data
	log [log_file,on,off]	Logging inputs and outputs
	run script_file	Running batch commands from file
Descriptive / Diagnostic	afreq [marker(s)]	Estimating allele frequencies
	genotype pedigree_id marker1 [marker2 …]	Displaying raw genotype data
	hapfreq [-d] [marker(s)]	Estimating haplotype frequencies
	viewhap [-c] [-s] [-e] [marker(s)]	Viewing haplotype configurations
	viewmarker marker [pedigree_id]	Displaying info about marker genotype distribution
	viewstat [-e] [-o] [-c] [-s] marker	Displaying detailed information about FBAT statistics
FBAT test	displayp [p_value]	Supress p_values > specified value
	fbat [-e] [-o] [-c] [marker(s)]	Computing family-based association test statistic (FBAT)
	hbat [-c] [-e] [-o] [-p[value]] [marker(s)]	Computing haplotype version of FBAT statistic
	maxcmh [count]	Setting the maximum allowable number of compatible mating haplotypes
	minsize [size_value]	Specifying minimum number of informative families
	mode [b,m,a]	Selecting biallelic, multiallelic testing procedures or both (a)
	model [a,d,r,g]	Selecting the genetic model
	offset [offset_value]	Setting the trait offset
	sdt [marker]	Computing sibship disequilibruim test (SDT)
	setafftrait aff_t unaff_t unknown_t	Setting trait values
	trait [trait_name(s)]	Specifying the trait(s) of interest

equation:

$$\chi^2 = \sum \frac{(observed - expected)^2}{expected}$$

Chi-square test for HWE was tested at P-value 0.05 with one degree of freedom (Table 9). If the critical value was less than 3.841, the observed and expected values were not significantly difference. Then the population tested was indeed in Hardy-Weinberg equilibrium.

If the studied population was in Hardy-Weinberg equilibrium for a given locus, it indicated that this population was well-designed and reasonable to use for association study. LHON patients, all pedigrees members and normal controls in this thesis were tested for HWE. If the samples were in HWE, it could be concluded that our control did not undergo any evolution. They were from random mating with respect to genotype, as well as the allele frequencies were the same in males and females, and mutation, selection, and migration were negligible.

3.5 Standard statistical analysis

Association test between SNP either allele or genotype frequencies and case-control status were assessed using standard χ^2 test. The *P* values were determined via a χ^2 approximation and Fisher's exact test. The power of all statistical tests were accepted at $P<0.05$, a significant result of type 1 error. Odd ratio (OR) and 95% confidence interval (CI) for LHON were determined by MedCalc software. Covariates (gender, mtDNA mutation proportion or heteroplasmy, secondary mutation) were analyzed for an association with survival (LHON manifestation). Unadjusted survival analyses of each covariate were performed using Kaplan-Meier plots. The Cox proportional hazards regression model was used to assess the ability of covariates to predict survival, with goodness-of-fit assessed by martingale residuals. Cox–Snell residuals were used to predict the survival function and further interpret in genetic study. Analyses were performed using Intercooled Stata software.

Table 9. The Chi-square table of critical values at degree of freedom 1 to 5 with probability of exceeding the critical value.

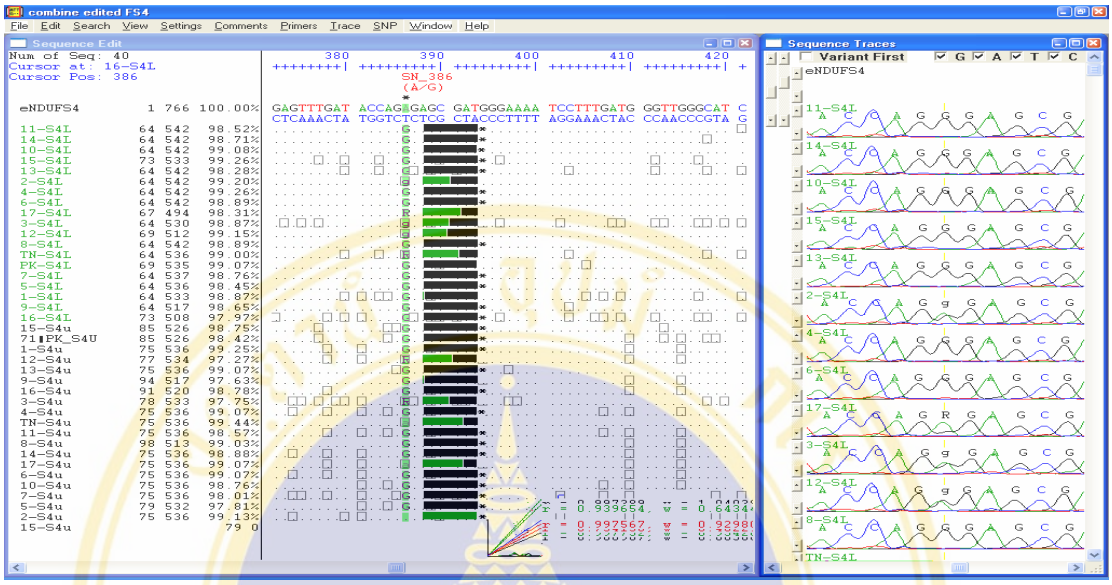
Degree of freedom	Probability of exceeding the critical value				
	0.1	0.05	0.025	0.01	0.001
1	2.706	**3.841**	5.024	6.635	10.828
2	4.605	5.991	7.378	9.21	13.816
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.07	12.833	15.086	20.515

CHAPTER IV
RESULTS

1. SNPs database of *NDUFS4*, *NDUFS8*, and *NDUFV2* gene from 19 Thai normal controls

Sequences of 19 Thai normal controls, approximately 1.1 kb of the promoter region and the coding region of *NDUFS4, NDUFS8* and *NDUFV2* gene were screened for single nucleotide polymorphisms (SNPs) by direct sequencing. SNPs were identified and validated by either GenalysWin or BioEdit program (Figure 9). Validated SNPs require the minor allele frequency (MAF) of more than 10% in our normal population, and either in the promoter region that could affect the level or regulatory region of gene expression, or in the coding region that could affect either amino acid that cause different protein properties. SNPs that do not change any amino acid but alter the splicing pattern or affect the mRNA stability expression are also validated. Total ten SNPs were discovered in these 3 genes were detected in our study, 4 each in *NDUFS4* and *NDUFV2*, and 2 in *NDUFS8*. For 4 SNPs in *NDUFS4* gene, one was in the promoter region, -447 T→A, and another 3 synonymous SNPs were at nucleotide position 12, 198 and 312 in the coding region (Table 10). The variant +12 C→G (Val4Val) located in exon 1, both +198 C→A (Gly66Gly) and +312 G→A (Arg104Arg) were in exon3. These 3 SNPs were previously reported in NCBI dbSNP cluster ID: rs2279516, rs31304, and rs31303, respectively. The -447 T→A in the promoter region of this gene corresponded to the NCBI dbSNP cluster ID: rs4147732 and also involved in the transcription factor binding site. Heterozygous TA of -447 T→A variant was found (n=11, Table 10) 45.46%, and both homozygous (TT and AA) were each found 27.27% in our population. The allele frequencies of both A and T allele at this position was equal to 50%. This polymorphism was then selected for further genotyping (MAF > 10%). The +12 C→G located in exon1 of *NDUFS4* gene, this polymorphism coded for valine at position 4 and did not alter any amino

(a)



(b)

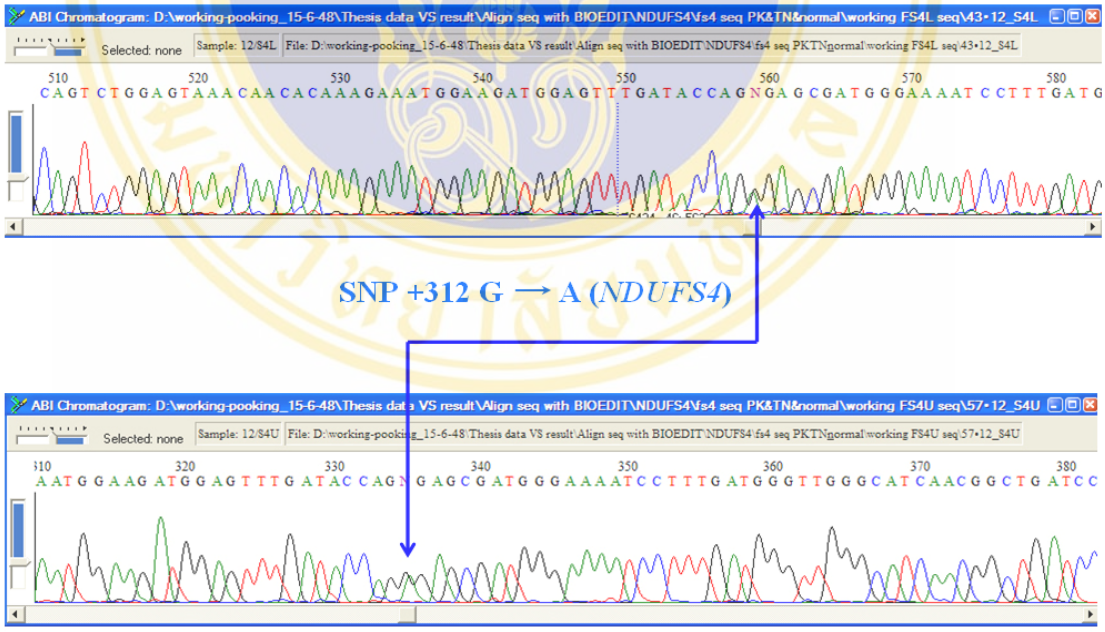


Figure 9. Identification of validated SNP at position 312 upstream from start codon of *NDUFS4* gene. (a) The alignment of multiple sequences and polymorphisms detection by GenalysWin program. (b) SNP +312 G→A was confirmed in both strand of *NDUFS4* gene using BioEdit program.

Table 10. SNPs found spanning the promoter and coding region of *NDUFS4*, *NDUFS8* and *NDUFV2* genes in 19 Thai normal controls.

Gene	Polymorphisms	Genotype distribution (%)			Allele frequency (%)	
NDUFS4	**-447 T→A** (n = 11)	**TT** 3 (27.27)	**TA** 5 (45.46)	**AA** 3 (27.27)	**T** 11 (50)	**A** 11 **(50)***
	+12 C→G (n = 19)	**CC** 9 (47.37)	**CG** 8 (42.1)	**GG** 2 (10.53)	**C** 26 (68.42)	**G** 12 **(31.58)***
	+198 C→A (n = 19)	**CC** 18 (94.74)	**CA** 1 (5.26)	**AA** 0 (0)	**C** 37 (97.37)	**A** 1 (2.63)
	+312 G→A (n = 19)	**GG** 14 (73.68)	**GA** 5 (26.32)	**AA** 0 (0)	**G** 33 (86.84)	**A** 5 **(13.16)***
NDUFS8	**-496 C→G** (n = 10)	**CC** 1 (10)	**CG** 4 (40)	**GG** 5 (50)	**C** 6 **(30)***	**G** 14 (70)
	-45 A→C (n = 10)	**AA** 9 (90)	**AC** 1 (10)	**CC** 0 (0)	**A** 19 (95)	**C** 1 (5)
NDUFV2	**-602 G→A** (n = 10)	**GG** 3 (30)	**GA** 3 (30)	**AA** 4 (40)	**G** 9 **(45)***	**A** 11 (55)
	-288 C→T (n = 10)	**CC** 9 (90)	**CT** 1 (10)	**TT** 0 (0)	**C** 19 (95)	**T** 1 (5)
	-233 T→C (n = 10)	**TT** 5 (50)	**TC** 5 (50)	**CC** 0 (0)	**T** 15 (75)	**C** 5 **(25)***
	+86 T→C (n = 19)	**TT** 4 (21.05)	**TC** 8 (42.11)	**CC** 7 (36.84)	**T** 16 **(42.11)***	**C** 22 (57.89)

* SNPs shown minor allele frequency more than 10%

acid in the protein. The +12 C→G had minor allele (G) frequency for 31.58% (n=19), the genotype frequencies of CC, CG and GG were 47.37, 42.1, and 10.53% respectively. This polymorphism was also selected for further study (MAF > 10%). Another 2 synonymous SNPs of this gene were on exon 3, the +198 C→A encoded glycine at amino acid position 66 and the +312 G→A coded for arginine at position 104. These two SNPs did not alter any amino acid in the protein. Minor allele (A) of +198 C→A had frequency of 2.63% (n=19), whereas the CC genotype was most found at 94.74% and CA genotype frequency was found at 5.26%. No AA genotype of this allele was observed. The +312 G→A had minor allele (A) frequency of 13.16% (n=19), the genotype distribution of GG and GA were 73.68% and 26.32% respectively. No AA genotype of this allele was observed. The +312 G→A polymorphism was chosen for further study since the minor allele frequency was in our criteria (MAF > 10%). Therefore, validated 3 SNPs form *NDUFS4* gene, -447 T→A in the promoter region, +12 C→G in exon 1 and the +312 G→A in exon 3 spanning the coding region, were used for further genotyping and generated the haplotype estimation.

For *NDUFS8* gene, no polymorphisms was found in the coding region; but 2 SNPs in the promoter region at downstream nucleotide position 496 and 45 from open reading sequence, were determined in 10 normal controls sequences. The -496 C→G corresponded to NCBI dbSNP cluster ID: rs4147776 while the -45 A→C was not found in public database. Minor allele frequency (C of -496 C→G) was 30% and the genotype frequency of GG, CG, and CC were 50%, 40%, and 10% respectively. The minor allele C of -45 A→C presented only 5% and the AA and AC genotype were found 90% and 10% respectively. None of the AA genotype was found. Therefore, only the polymorphism -496 C→G was taken for further genotyped (MAF > 10%).

For *NDUFV2*, sequences of ten normals were determined in the promoter region and in nineteen normals for coding region. Four SNPs were found in the *NDUFV2* gene. Three polymorphisms located in the promoter sequences, -602 G→A, -288 C→T, and -233 T→C. All these SNPs have never been reported in the SNP database. The frequency of minor allele G of SNP -602 G→A was 45%, while the GG and GA genotype distribution were 30% each and the remaining AA genotype was 40%. The

SNP -288 C→T showed its minor allele of T for 5% and its genotype distribution favored CC to CT (90% to 10% respectively). The -233 T→C SNP showed allele C frequency of 25% and genotype TT and TC had been found 50% each. One SNP detected in the coding region was +86 T→C in exon 2. This polymorphism was the non-synonymous SNP corresponding to NCBI dbSNP cluster ID: rs906807 which altered an amino acid from valine to alanine at position 29. The minor frequency of T allele was 42.11% and the genotype distribution of TC (42.11%) had been found higher than CC (36.84%) and TT (21.05%) respectively. Both -602 G→A and +86 T→C polymorphisms were selected for genotyping and haplotype estimating in further association study. The -233 T→C SNP was not included in our further study since it could not be detected by standard PCR-RFLP method.

Therefore, total 6 validated SNPs were chosen for further study. Three polymorphisms were from *NDUFS4* gene, -447 T→A in promoter region, +12 C→G and +312 G→A in the coding region, one polymorphism -496 C→G from *NDUFS8* gene and two SNPs from *NDUFV2* gene including the -602 G→A and +86 T→C located in the promoter and coding region of this gene respectively. These six selected SNPs were also genotyped with PCR-RFLP platform in all 19 normal controls again in order to confirm the genotype data from DNA sequencing. The results from both sequencing and PCR-RFLP method were concordant in all 19 normal controls. Allele frequency and genotype distribution of SNPs found in our 19 normal controls were summarized in Table 10.

2. SNPs detection and Hardy-Weinberg analysis of six validated SNPs genotyping

One hundred and fifteen Thai normal controls and 281 LHON family members were genotyped for 6 validated SNPs spanning promoter and coding region of *NDUFS4, NDUFS8* and *NDUFV2* gene. The frequency of each allele and genotype in each population were calculated. Our normal controls were separated into two groups, 19 individuals from four parts of Thailand which could be traced back for their ethnic origins. Another 115 controls were from mixed-urban Thai population participated from Check up clinic, Siriraj Hospital. For LHON family members, the phenotype

status of all family members were recorded (affected, unaffected, and unknown status). The HWE results as well as both genotype and allele frequencies of all SNPs loci were displayed in Table 11. The genotype distribution of both normal control groups 1 and 2 were in HW equilibrium for all SNPs loci tested (P>0.05). Similar to LHON pedigree members, both affected (n=76) and unaffected (n=203), were also in HWE for all SNPs loci. In addition, unaffected group can be divided into three subgroups, maternal lineage members with the LHON 11778 mutation, maternal lineage members without the LHON 11778 mutation and the non-maternal lineage members. Most of unaffected subgroups of six SNPs tested were in equilibrium, excepted for +12 C→G loci of *NDUFS4* gene. Subgroup of maternal lineage carrying 11778 mutation was apart from equilibrium (P<0.05). However, of all SNPs loci tested, our normal controls were reasonable to be used for association study and there was no selection operating at our locus and no violation process from equilibrium interfered.

3. Comparison of genotype and allele frequencies of 6 validated SNPs between normal population and LHON

Total 281 LHON family members and 115 individual controls were genotyped for 6 validated SNPs spanning promoter and coding region of *NDUFS4, NDUFS8* and *NDUFV2* gene. First of all, SNPs distribution was examined in 2 normal control groups which were 19 Thai normal with specific ethnic origin, and another 115 Thai mixed normal controls. The data showed no significant difference of all SNPs genotyped frequency in these two groups (Table 12).

The distribution of both genotype and allele frequency of -447 T→A, +12 C→G and +312 G→A of *NDUFS4*, -496 C→G of *NDUFS8*, -602 G→A and +86 T→C of *NDUFV2* in affected and normal controls were determined. Seventy-six LHON patients who carried both homoplasmic and heteroplasmic of 11778 mtDNA were compared to 115 normal controls. The genotypes of SNPs, -447 T→A, -496 C→G and -602 G→A, located in promoter region of *NDUFS4*, *NDUFS8* and *NDUFV2* gene conferred no significant advantage. The -447 T→A had P value for TT genotype versus AA genotype at 0.1318 (OR = 2.0093; 95%CI = 0.8358-4.8300). The CC versus GG genotype of SNP -496 C→G showed P value = 0.1946; OR=1.7968 and

Table 11. The distribution of genotype, allele frequency and HWE test in normal controls and individuals from LHON pedigrees.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			Allele frequency (%)		HWE-test (*P*-values)
-447 T→A (*NDUFS4*)	TT	TA	AA	T	A	
Thai normal controls						
Group 1 (n=19)[a]	7 (36.8)	6 (31.6)	6 (31.6)	20 (52.6)	18 (47.4)	2.5615 (0.1095)
Group 2 (n=115)[b]	27 (23.5)	57 (49.6)	31 (26.9)	111 (48.3)	119 (51.7)	0.0128 (0.9099)
LHON family members						
Affected (n=76)	21 (27.6)	43 (56.6)	12 (15.8)	85 (55.9)	67 (44.1)	1.6590 (0.1977)
Unaffected (n=203)	37 (18.2)	112 (55.2)	54 (26.6)	186 (45.8)	220 (54.2)	2.5140 (0.1128)
maternal lineage, 11778 positive (n=124)	24 (19.4)	71 (57.2)	29 (23.4)	119 (48)	129 (52)	2.6804 (0.1016)
maternal lineage, 11778 negative (n=8)	3 (37.5)	2 (25)	3 (37.5)	8 (50)	8 (50)	NA
not maternal lineage (n=71)	10 (14.1)	39 (54.9)	22 (31)	59 (41.5)	83 (58.5)	1.2194 (0.2695)
Unknown status (n=2)	1 (50)	1 (50)	0 (0)	3 (75)	1 (25)	NA
Total (n=281)	59 (21)	156 (55.5)	66 (23.5)	274 (48.8)	288 (51.2)	3.4733 (0.0624)

Table 11 (continued). The distribution of genotype, allele frequency and HWE test in normal controls and individuals from LHON pedigrees.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			Allele frequency (%)		HWE-test (*P*-values)
+12 C→G (*NDUFS4*)	CC	CG	GG	C	G	
Thai normal controls						
Group 1 (n=19)[a]	9 (47.4)	7 (36.8)	3 (15.8)	25 (65.8)	13 (34.2)	0.6308 (0.4271)
Group 2 (n=115)[b]	59 (51.3)	45 (39.1)	11 (9.6)	163 (70.9)	67 (29.1)	0.318 (0.5728)
LHON family members						
Affected (n=76)	20 (26.3)	42 (55.3)	14 (18.4)	82 (53.9)	70 (46.1)	0.9571 (0.328)
Unaffected (n=203)	77 (38)	103 (50.7)	23 (11.3)	257 (63.3)	149 (36.7)	1.7228 (0.1893)
maternal lineage, 11778 positive (n=124)	39 (31.5)	71 (57.2)	14 (11.3)	149 (60.1)	99 (39.9)	4.6458 **(0.0311)***
maternal lineage, 11778 negative (n=8)	5 (62.5)	0 (0)	3 (37.5)	10 (62.5)	6 (37.5)	NA
not maternal lineage (n=71)	33 (46.5)	32 (45.1)	6 (8.4)	98 (69)	44 (31)	0.2054 (0.6504)
Unknown status (n=2)	0 (0)	1 (50)	1 (50)	1 (25)	3 (75)	NA
Total (n=281)	97 (34.5)	146 (52)	38 (13.5)	340 (60.5)	222 (39.5)	2.1653 (0.1412)

Table 11 (continued). The distribution of genotype, allele frequency and HWE test in normal controls and individuals from LHON pedigrees.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			Allele frequency (%)		HWE-test (*P*-values)
+312 G→A (*NDUFS4*)	GG	GA	AA	G	A	
Thai normal controls						
Group 1 (n=19)[a]	14 (73.7)	5 (26.3)	0 (0)	33 (86.8)	5 (13.2)	0.4459 (0.5043)
Group 2 (n=115)[b]	88 (76.5)	23 (20)	4 (3.5)	199 (86.5)	31 (13.5)	2.5356 (0.1113)
LHON family members						
Affected (n=76)	39 (51.3)	34 (44.7)	3 (4)	112 (73.7)	40 (26.3)	1.8272 (0.1765)
Unaffected (n=203)	116 (57.1)	78 (38.4)	9 (4.5)	310 (76.4)	96 (23.6)	0.8437 (0.3583)
maternal lineage, 11778 positive (n=124)	66 (53.2)	54 (43.5)	4 (3.3)	186 (75)	62 (25)	3.2258 (0.1993)
maternal lineage, 11778 negative (n=8)	7 (87.5)	0 (0)	1 (12.5)	14 (87.5)	2 (12.5)	NA
not maternal lineage (n=71)	43 (60.6)	24 (33.8)	4 (5.6)	110 (77.5)	32 (22.5)	0.0877 (0.7671)
Unknown status (n=2)	0 (0)	1 (50)	1 (50)	1 (25)	3 (75)	NA
Total (n=281)	155 (55.2)	113 (40.2)	13 (4.6)	423 (75.3)	139 (24.7)	1.7964 (0.1801)

Table 11 (continued). The distribution of genotype, allele frequency and HWE test in normal controls and individuals from LHON pedigrees.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			Allele frequency (%)		HWE-test (*P*-values)
-496 C→G (*NDUFS8*)	CC	CG	GG	C	G	
Thai normal controls						
Group 1 (n=19)[a]	4 (21.1)	8 (42.1)	7 (36.8)	16 (42.1)	22 (57.9)	0.354 (0.5519)
Group 2 (n=115)[b]	17 (14.8)	56 (48.7)	42 (36.5)	90 (39.1)	140 (60.9)	0.0585 (0.8089)
LHON family members						
Affected (n=76)	16 (21.1)	38 (50)	22 (28.9)	70 (46.1)	82 (53.9)	0.0032 (0.9549)
Unaffected (n=203)	39 (19.2)	111 (54.7)	53 (26.1)	189 (46.6)	217 (53.4)	2.0074 (0.1565)
maternal lineage, 11778 positive (n=124)	29 (23.4)	70 (56.4)	25 (20.2)	128 (51.6)	120 (48.4)	2.1176 (0.1456)
maternal lineage, 11778 negative (n=8)	0 (0)	2 (25)	6 (75)	2 (12.5)	14 (87.5)	NA
not maternal lineage (n=71)	10 (14.1)	39 (54.9)	22 (31)	59 (41.6)	83 (58.4)	1.2194 (0.2695)
Unknown status (n=2)	0 (0)	2 (100)	0 (0)	2 (50)	2 (50)	NA
Total (n=281)	55 (19.6)	151 (53.7)	75 (26.7)	261 (46.4)	301 (53.6)	1.853 (0.1734)

Table 11 (continued). The distribution of genotype, allele frequency and HWE test in normal controls and individuals from LHON pedigrees.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			Allele frequency (%)		HWE-test (*P*-values)
-602 G→A (*NDUFV2*)	GG	GA	AA	G	A	
Thai normal controls						
Group 1 (n=19)[a]	4 (21.1)	7 (36.8)	8 (42.1)	15 (39.5)	23 (60.5)	1.0092 (0.3151)
Group 2 (n=115)[b]	10 (8.7)	47 (40.9)	58 (50.4)	67 (29.1)	163 (70.9)	0.0139 (0.9061)
LHON family members						
Affected (n=76)	3 (3.9)	19 (25)	54 (71.1)	25 (16.4)	127 (83.6)	0.6740 (0.4117)
Unaffected (n=203)	10 (4.9)	67 (33)	126 (62.1)	87 (21.4)	319 (78.6)	0.1283 (0.7202)
maternal lineage, 11778 positive (n=124)	6 (4.8)	39 (31.5)	79 (63.7)	51 (20.6)	197 (79.4)	0.1969 (0.6572)
maternal lineage, 11778 negative (n=8)	0 (0)	7 (87.5)	1 (12.5)	7 (43.7)	9 (56.3)	NA
not maternal lineage (n=71)	4 (5.6)	21 (29.6)	46 (64.8)	29 (20.4)	113 (79.6)	0.6109 (0.4344)
Unknown status (n=2)	0 (0)	2 (100)	0 (0)	2 (50)	2 (50)	NA
Total (n=281)	13 (4.6)	88 (31.3)	180 (64.1)	114 (20.3)	448 (79.7)	0.3167 (0.5736)

Table 11 (cont.) The distribution of genotype, allele frequency and HWE test in normal controls and individuals from LHON pedigrees.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			Allele frequency (%)		HWE-test (*P*-values)
+86 T→C (*NDUFV2*)	TT	TC	CC	T	C	
Thai normal controls						
Group 1 (n=19)[a]	4 (21.1)	8 (42.1)	7 (36.8)	16 (42.1)	22 (57.9)	0.3540 (0.5519)
Group 2 (n=115)[b]	10 (8.7)	52 (45.2)	53 (46.1)	72 (31.3)	158 (68.7)	0.3158 (0.5741)
LHON family members						
Affected (n=76)	3 (3.9)	19 (25)	54 (71.1)	25 (16.4)	127 (83.6)	0.6740 (0.4117)
Unaffected (n=203)	14 (6.9)	69 (34)	120 (59.1)	97 (23.9)	309 (76.1)	0.8646 (0.3524)
maternal lineage, 11778 positive (n=124)	9 (7.3)	40 (32.3)	75 (60.4)	58 (23.4)	190 (76.6)	1.2877 (0.2565)
maternal lineage, 11778 negative (n=8)	0 (0)	7 (87.5)	1 (12.5)	7 (43.7)	9 (56.3)	NA
not maternal lineage (n=71)	5 (7)	22 (31)	44 (62)	32 (22.5)	110 (77.5)	0.8923 (0.3448)
Unknown status (n=2)	0 (0)	2 (100)	0 (0)	2 (50)	2 (50)	NA
Total (n=281)	17 (6)	90 (32)	174 (62)	124 (22.1)	438 (77.9)	1.3338 (0.2481)

NA = not available

a = Thai normal controls with specific ethnic origin

b = The mixed Thai normal controls with non-system random sampling

Table 12. The comparison of genotype distributions and allele frequencies between 2 normal control groups in 6 available SNPs spanning 3 candidate genes.

Single Nucleotide Polymorphisms (SNPs)	Genotype distribution (%)			χ^2 (*P*-values)*	Allele frequency (%)		*P*-values**
-447 T→A (*NDUFS4*)	**TT**	**TA**	**AA**		**T**	**A**	
Normal with specific ethnic group (n=19)	7 (36.8)	6 (31.6)	6 (31.6)	2.396 (0.3018)	20 (52.6)	18 (47.4)	0.7266
Normal Controls (n=115)	27 (23.5)	57 (49.6)	31 (26.9)		111 (48.3)	119 (51.7)	
+12 C→G (*NDUFS4*)	**CC**	**CG**	**GG**		**C**	**G**	
Normal with specific ethnic group (n=19)	9 (47.4)	7 (36.8)	3 (15.8)	0.676 (0.7130)	25 (65.8)	13 (34.2)	0.5673
Normal Controls (n=115)	59 (51.3)	45 (39.1)	11 (9.6)		163 (70.9)	67 (29.1)	
+312 G→A (*NDUFS4*)	**GG**	**GA**	**AA**		**G**	**A**	
Normal with specific ethnic group (n=19)	14 (73.7)	5 (26.3)	0 (0)	0.989 (0.6098)	33 (86.8)	5 (13.2)	1.0000
Normal Controls (n=115)	88 (76.5)	23 (20)	4 (3.5)		199 (86.5)	31 (13.5)	
-496 C→G (*NDUFS8*)	**CC**	**CG**	**GG**		**C**	**G**	
Normal with specific ethnic group (n=19)	4 (21.1)	8 (42.1)	7 (36.8)	0.558 (0.7566)	16 (42.1)	22 (57.9)	0.7241
Normal Controls (n=115)	17 (14.8)	56 (48.7)	42 (36.5)		90 (39.1)	140 (60.9)	
-602 G→A (*NDUFV2*)	**GG**	**GA**	**AA**		**G**	**A**	
Normal with specific ethnic group (n=19)	4 (21.1)	7 (36.8)	8 (42.1)	2.678 (0.2620)	15 (39.5)	23 (60.5)	0.2536
Normal Controls (n=115)	10 (8.7)	47 (40.9)	58 (50.4)		67 (29.1)	163 (70.9)	
+86 T→C (*NDUFV2*)	**TT**	**TC**	**CC**		**T**	**C**	
Normal with specific ethnic group (n=19)	4 (21.1)	8 (42.1)	7 (36.8)	2.73 (0.2554)	16 (42.1)	22 (57.9)	0.1964
Normal Controls (n=115)	10 (8.7)	52 (45.2)	53 (46.1)		72 (31.3)	158 (68.7)	

* Chi-square test (d.f.=2)

** Fisher's exact test

95%CI = 0.7636-4.2277. The -602 G→A, *P* value of AA versus GG genotype was 0.1395 (OR = 3.1034; 95%CI = 0.8107-11.8808). The allele frequencies of either -447 T→A or -496 C→G conferred no significant advantage. Allele T of SNP -447 T→A showed OR = 1.3601; 95%CI = 0.9012-2.0527; and *P* = 0.1452 and allele C of SNP -496 C→G showed OR = 1.3279; 95%CI = 0.8773-2.0101; and *P* value = 0.2038. Although the SNP -602 G→A of *NDUFV2* gene conferred no significant in genotype distribution, allele A of this polymorphism showed highly significant difference between affected LHON and normal controls (*P* value = 0.0049; OR = 2.0881 and 95%CI = 1.2481-3.4933). The SNP +86 T→C of *NDUFV2* gene showed no significant difference in both CC and TC genotype versus TT genotype, but allele C showed highly significant difference (*P* value = 0.0011; OR = 2.3149 and 95%CI = 1.3881-3.8606).

Two SNPs, +12 C→G and +312 G→A located in *NDUFS4* gene, were significantly different in both allele and genotype distribution between 76 LHON patients and 115 individual controls. The +12 C→G, genotype distribution of both GG and CG versus CC was highly significant. *P* value for GG versus CC genotype was 0.0068 (OR = 3.7545 and 95%CI = 1.4688-9.5974). In addition, the CG genotype versus CC showed *P* value = 0.0024 (OR = 2.7533 and 95%CI = 1.4248-5.3208). Frequency allele G of SNP +12 C→G showed highly significant with P value = 0.001; OR = 2.0768; 95%CI = 1.3546-3.1842. Likewise, the +312 G→A, heterozygous GA versus GG genotype showed significant *P* value at 0.0003 (OR = 3.3356; 95%CI = 1.7418-6.3878) and *P* value at 0.002 (OR = 2.2926; 95%CI = 1.3589-3.8680) for allele A. The results were summarized on Table 13.

However, in order to study the association between marker loci (SNPs) and LHON phenotype, controls from unaffected members of LHON pedigrees were used to avoid population admixture. Twenty-nine unaffected and 67 affected were available for this study. Unaffected criteria are maternal lineage and older than 30 years of age for male, than 45 years of age for female. Both affected and unaffected were homoplasmic for the G11778A mutation (>95% mtDNA mutation load) in their blood. The result showed no significant difference between affected and unaffected group in either genotype or allele distribution of all 6 SNPs genotyped. The

Table 13. The comparison of genotype distributions and allele frequencies between affected LHON and normal control.

SNPs	Genotype distribution (%)			Allele frequency (%)	
-447 T→A (*NDUFS4*)	**TT**	**TA**	**AA**	**T**	**A**
Affected (n=76)	21 (27.6)	43 (56.6)	12 (15.8)	85 (55.9)	67 (44.1)
Normal Controls (n=115)	27 (23.5)	57 (49.6)	31 (26.9)	111 (48.3)	119 (51.7)
OR (95%CI)	2.0093 (0.8358 - 4.8300)	1.9488 (0.8977 - 4.2306)	1	1.3601 (0.9012 - 2.0527)	1
P-values*	0.1318	0.0959		0.1452	
+12 C→G (*NDUFS4*)	**CC**	**CG**	**GG**	**C**	**G**
Affected (n=76)	20 (26.3)	42 (55.3)	14 (18.4)	82 (53.9)	70 (46.1)
Normal Controls (n=115)	59 (51.3)	45 (39.1)	11 (9.6)	163 (70.9)	67 (29.1)
OR (95%CI)	1	2.7533 (1.4248 - 5.3208)	3.7545 (1.4688 - 9.5974)	1	2.0768 (1.3546 - 3.1842)
P-values*		**0.0024***	**0.0068***		**0.001***
+312 G→A (*NDUFS4*)	**GG**	**GA**	**AA**	**G**	**A**
Affected (n=76)	39 (51.3)	34 (44.7)	3 (4)	112 (73.7)	40 (26.3)
Normal Controls (n=115)	88 (76.5)	23 (20)	4 (3.5)	199 (86.5)	31 (13.5)
OR (95%CI)	1	3.3356 (1.7418 - 6.3878)	1.6923 (0.3614 - 7.9234)	1	2.2926 (1.3589 - 3.8680)
P-values*		**0.0003***	0.6774		**0.002***
-496 C→G (*NDUFS8*)	**CC**	**CG**	**GG**	**C**	**G**
Affected (n=76)	16 (21.1)	38 (50)	22 (28.9)	70 (46.1)	82 (53.9)
Normal Controls (n=115)	17 (14.8)	56 (48.7)	42 (36.5)	90 (39.1)	140 (60.9)
OR (95%CI)	1.7968 (0.7636 - 4.2277)	1.2955 (0.6695 - 2.5068)	1	1.3279 (0.8773 - 2.0101)	1
P-values*	0.1946	0.5055		0.2038	
-602 G→A (*NDUFV2*)	**GG**	**GA**	**AA**	**G**	**A**
Affected (n=76)	3 (3.9)	19 (25)	54 (71.1)	25 (16.4)	127 (83.6)
Normal Controls (n=115)	10 (8.7)	47 (40.9)	58 (50.4)	67 (29.1)	163 (70.9)
OR (95%CI)	1	1.3475 (0.3336 - 5.4423)	3.1034 (0.8107 - 11.8808)	1	2.0881 (1.2481 - 3.4933)
P-values*		1	0.1395		**0.0049***
+86 T→C (*NDUFV2*)	**TT**	**TC**	**CC**	**T**	**C**
Affected (n=76)	3 (3.9)	19 (25)	54 (71.1)	25 (16.4)	127 (83.6)
Normal Controls (n=115)	10 (8.7)	52 (45.2)	53 (46.1)	72 (31.3)	158 (68.7)
OR (95%CI)	1	1.2179 (0.3024 - 4.9051)	3.3962 (0.8851 - 13.0319)	1	2.3149 (1.3881 - 3.8606)
P-values*		1	0.08		**0.0011***

* Fisher's exact test

approximate *P*-value was estimated from Fisher's exact test and the odd ratio with 95% confidence interval (CI) for LHON was calculated by MedCalc software. Results were shown on Table 14.

4. Searching for nuclear modifier gene susceptibility to LHON manifestation using association study

Family-based association design was used to search for nuclear modifier gene for LHON expression in this study. Eighteen LHON pedigrees (from 42 pedigrees) containing 250 individuals including 70 nuclear families were analyzed. They were 53 affected (males = 38, females = 15), 112 unaffected with maternal lineage (males = 38, females = 74), 2 females unknown status with maternal lineage, 24 males with non-maternal lineage, and 59 individuals (12 with maternal and 47 with non-maternal lineage) generated with no DNA samples but necessary for the pedigree construction. Therefore, total 191 LHON with family members were genotyped for 6 validated SNPs and used for further family-based association study.

Our LHON characteristics were adjusted and classified for further covariate factors analyses. Considering the 179 members with maternal relatives (54 affected, 111 unaffected, 2 unknown status, and 12 whom no DNA), 160 individuals were positive, while 7 were negative for the G11778A mutation and 12 were not verified since they had no DNA (new generated for pedigree construct). In addition, of 167 maternal family members (excluded the 12 individuals because they had no DNA samples), 18 of them also carried the secondary LHON mutation that of 6 were C3497T and 12 were G3316A. This secondary mutation classified total 250 individuals into 3 groups; a group with the C3497T, with G3316A, and without any secondary mutation. Heteroplasmy of the G11778A mutation classified all the samples into 3 groups, a group which carried the mutation load less than 70%, between 71-95%, and more than 95% of mutation load in their bloods.

For the total of 53 affected (males = 38, females = 15), 5 patients (3 males and 2 females) carried the 11778 mtDNA mutation for 70-95%, and 47 patients (34 males and 13 females) carried the homoplasmic (>95%) of 11778 mutation. Only one affected male had 44% LHON mutation load. Unadjusted of 112 unaffected LHON were 38 males and 74 females with various 11778 mutation proportion. Six

Table 14. The comparison of genotype distributions and allele frequencies between affected and unaffected LHON.

SNPs	Genotype distribution (%)			Allele frequency (%)	
-447 T→A (*NDUFS4*)	**TT**	**TA**	**AA**	**T**	**A**
Affected (n=67)[+]	19 (28.4)	38 (56.7)	10 (14.9)	76 (56.7)	58 (43.3)
Unaffected (n=29)[++]	8 (27.6)	14 (48.3)	7 (24.1)	30 (51.7)	28 (48.3)
OR (95%CI)	1.6625 (0.4664 - 5.9256)	1.9 (0.6053 - 5.9639)	1	1.223 (0.6592 - 2.2691)	1
P-values*	0.5207	0.3632		0.5318	
+12 C→G (*NDUFS4*)	**CC**	**CG**	**GG**	**C**	**G**
Affected (n=67)[+]	18 (26.9)	36 (53.7)	13 (19.4)	72 (53.7)	62 (46.3)
Unaffected (n=29)[++]	11 (37.9)	12 (41.4)	6 (20.7)	34 (58.6)	24 (41.4)
OR (95%CI)	1	1.8333 (0.6780 - 4.9574)	1.3241 (0.3893 - 4.5033)	1	1.2199 (0.6541 - 2.2750)
P-values*		0.3051	0.7622		0.6357
+312 G→A (*NDUFS4*)	**GG**	**GA**	**AA**	**G**	**A**
Affected (n=67)[+]	31 (46.3)	33 (49.2)	3 (4.5)	95 (70.9)	39 (29.1)
Unaffected (n=29)[++]	17 (58.6)	11 (38)	1 (3.4)	45 (77.6)	13 (22.4)
OR (95%CI)	1	1.6452 (0.6668 - 4.0589)	1.6452 (0.1586 - 17.0658)	1	1.4211 (0.6910 - 2.9223)
P-values*		0.3652	1		0.3804

***** Fisher's exact test

[+] available affected from total 76 individuals

[++] available unaffected in the criterion of males >30, and female >45 years of age

Table 14 (continued). The comparison of genotype distributions and allele frequencies between affected and unaffected LHON.

SNPs	Genotype distribution (%)			Allele frequency (%)	
-496 C→G (*NDUFS8*)	**CC**	**CG**	**GG**	**C**	**G**
Affected (n=67)[+]	14 (20.9)	35 (52.2)	18 (26.9)	63 (47)	71 (53)
Unaffected (n=29)[++]	7 (24.1)	19 (65.6)	3 (10.3)	33 (56.9)	25 (43.1)
OR (95%CI)	1.0857 (0.3741 - 3.1508)	1	3.2571 (0.8496 - 12.4870)	1	1.4876 (0.7998 - 2.7671)
P-values*	1		0.0942		0.2712
-602 G→A (*NDUFV2*)	**GG**	**GA**	**AA**	**G**	**A**
Affected (n=67)[+]	3 (4.5)	15 (22.4)	49 (73.1)	21 (15.7)	113 (84.3)
Unaffected (n=29)[++]	0 (0)	8 (27.6)	21 (72.4)	8 (13.8)	50 (86.2)
OR (95%CI)	NA	1	1.2444 (0.4583 - 3.3788)	1.1615 (0.4819 - 2.7996)	1
P-values*	NA		0.7959	0.8288	
+86 T→C (*NDUFV2*)	**TT**	**TC**	**CC**	**T**	**C**
Affected (n=67)[+]	3 (4.5)	15 (22.4)	49 (73.1)	21 (15.7)	113 (84.3)
Unaffected (n=29)[++]	1 (3.5)	9 (31.0)	19 (65.5)	11 (19.0)	47 (81.0)
OR (95%CI)	1.8 (0.1618 - 20.0286)	1	1.5474 (0.5798 - 4.1295	1	1.2594 (0.5631 - 2.8166)
P-values*	1		0.4423		0.6735

***** Fisher's exact test

[+] available affected from total 76 individuals

[++] available unaffected in the criterion of males >30 and female >45 years of age

unaffected males and 26 unaffected females carried the 11778 mutation less than 70% in their bloods. Fourteen unaffected (males = 7, females = 7) carried the 11778 mtDNA mutation load for 70-95%, and 66 individuals (25 males and 41 females) carried the homoplasmic (>95%) of 11778 mutation. The mean age (± SD) of 112 unaffected was calculated from age at present of individuals. That was 29.3 ± 15.6 (n = 38, range: 8-70 years, median: 27 years) in males, and was 42.5 ± 19.2 (n = 74, range: 7-87 years, median: 44 years) in females.

Whereas, the mean age of affected onset were calculated from 68 of 79 LHON affected. The exclusion of 11 individuals in our study was due to their uninformative onset-age. Mean age of onset (± SD) in male was 20.9 ± 9.3 years (n = 54, range: 6-44 years, median: 19 years), and in female was 28.6 ± 14.6 years (n = 14, range: 10-53 years, median: 30 years). The disease predilection of male to female (male:female ratio) in our patients was 3.2:1 (n = 79, males = 60, females = 19). The LHON family characteristics were summarized in Table 15.

4.1 Pedigree error checking with PedCheck program

The genotype data in our LHON pedigree were checked for all erroneous prior to the family-based association analysis. All Mendelian inconsistencies in the pedigree data were eliminated by running on PedCheck program. This program provided four error-checking algorithms. The algorithms ranged from simple parent-offspring compatibility checks (level 0, 1, and 2) to a single-locus likelihood-based statistic (level 3 and 4). Two later algorithms could identify and rank the individuals most likely to be in error. Our LHON pedigree data sets were introduced to check these algorithms on PedCheck program. The inconsistencies in our LHON pedigrees data were corrected and re-run until no erroneous genotype of our data was detected (Figure 10-13).

4.2 Disequilibrium analysis and haplotype estimation in LHON family

Haplotype of 3 SNPs spanning *NDUFS4* and 2 SNPs of *NDUFV2* gene were estimated by Simwalk2 and GOLD program. The summary table for the haplotype analysis from SimWalk2, file "TABLE-01.ALL", could be read on Microsoft® Notepad. Figure 14 and 15 showed haplotype summarized analysis of marker loci on *NDUFS4* and *NDUFV2* gene respectively. The total number of pedigrees actually analyzed was 18 pedigrees with 322 of meiosis in these pedigrees.

Table 15. The characteristics of our LHON families in this study.

Characteristic	No.of subjects	Age, % Mutation load
Maternal lineage		
positive 11778	160	
negative 11778	7	
NA	12	
Non-maternal lineage (with/without DNA)	24/47	
Affected/Unaffected/Unknown status	54/111/2	
11778 mutation proportion (n)**,†		
< 70%		
Affected		
- male	1	18, 44%
- female	0	
Unaffected		
- male	6	15 to 54, 0-59%
- female	26	16 to 87, 0-68%
70-95%		
Affected		
- male	3	16 to 20, 72-93%
- female	2	10 to 13, 72-80%
Unaffected		
- male	7	22 to 43, 70-94%
- female	7	7 to 64, 70-95%
>95%		
Affected		
- male	34	7 to 44, 100%
- female	13	10 to 45, 100%
Unaffected		
- male	25	8 to 70, 100%
- female	41	9 to 79, 100%
Other mutation (yes/no)*	18/149	
Age onset		**mean±SD**
Affected		
- male (n=54)	6 to 44	20.9 ± 9.3
- female (n=14)	10 to 53	28.6 ± 14.6
Age		
Unaffected		
- male (n=38)	8 to 70	29.3 ± 15.6
- female (n=74)	7 to 87	42.5 ± 19.2

NA = not available (no DNA)
* maternal lineage
† age of affected = age of on-set age of unaffected = age at 2005

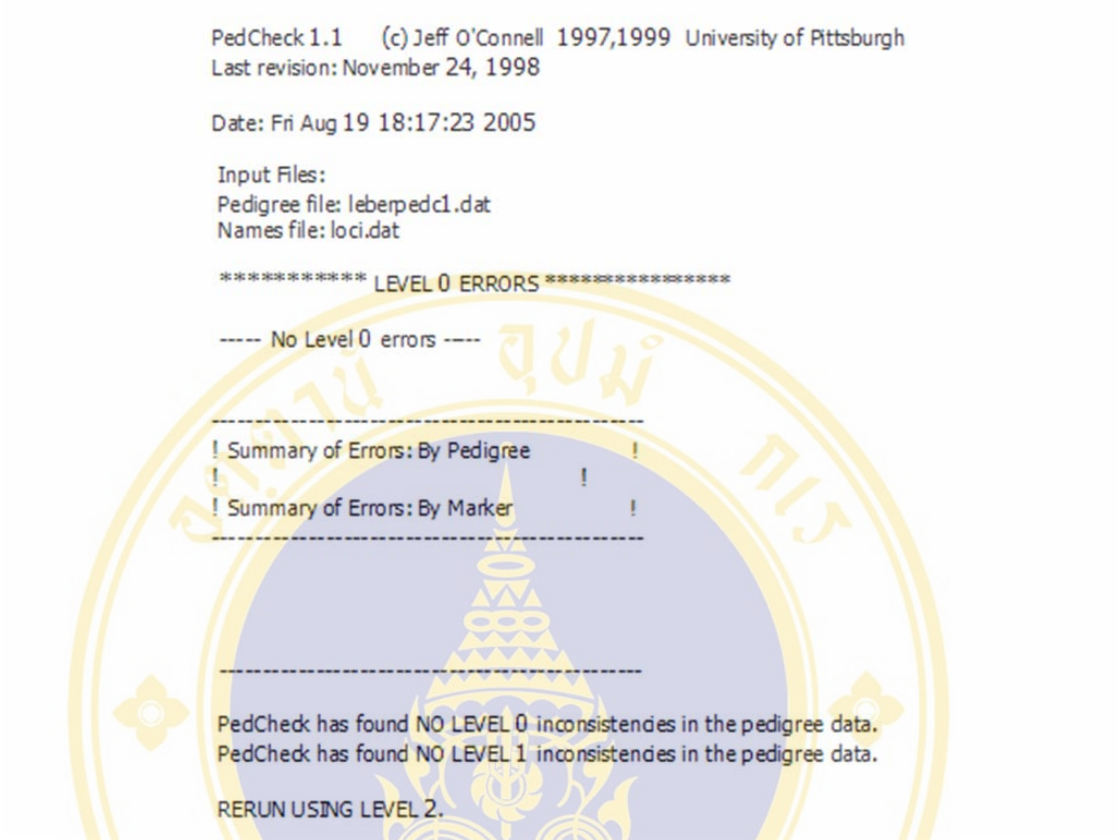


Figure 10. The result of error-checking algorithm on level 1 of genotype data of LHON pedigree with PedCheck program.

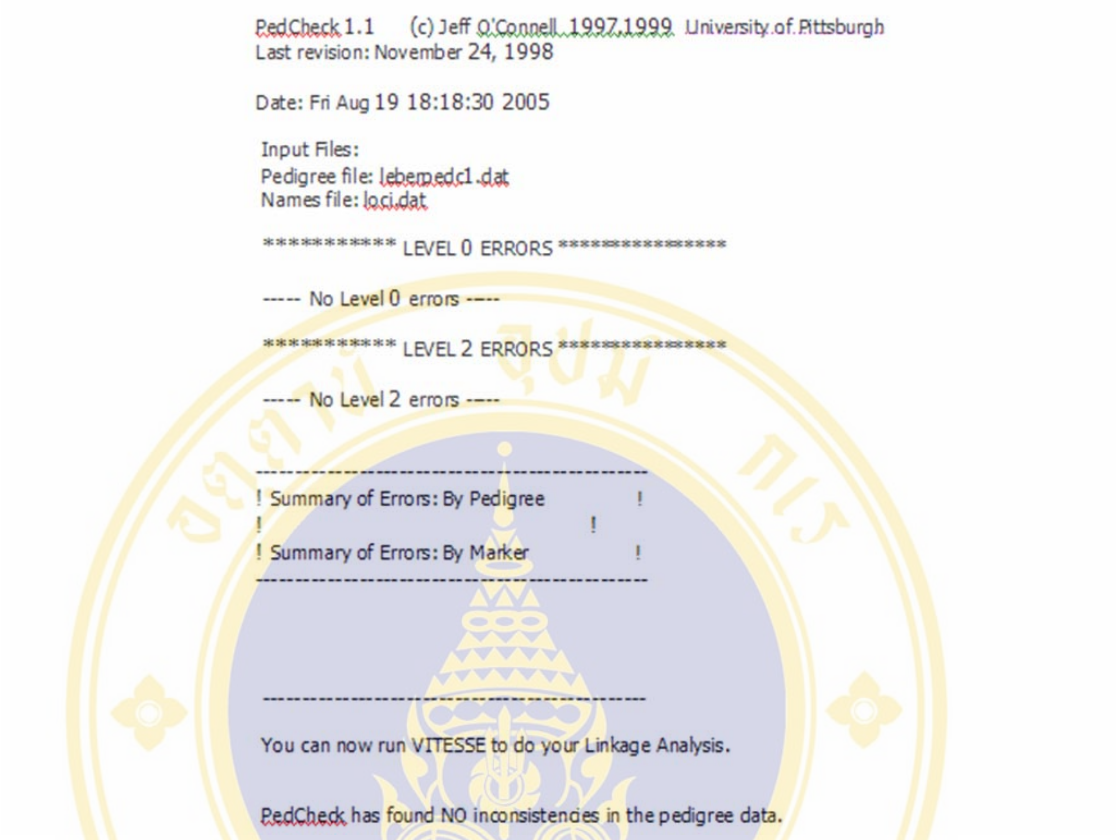
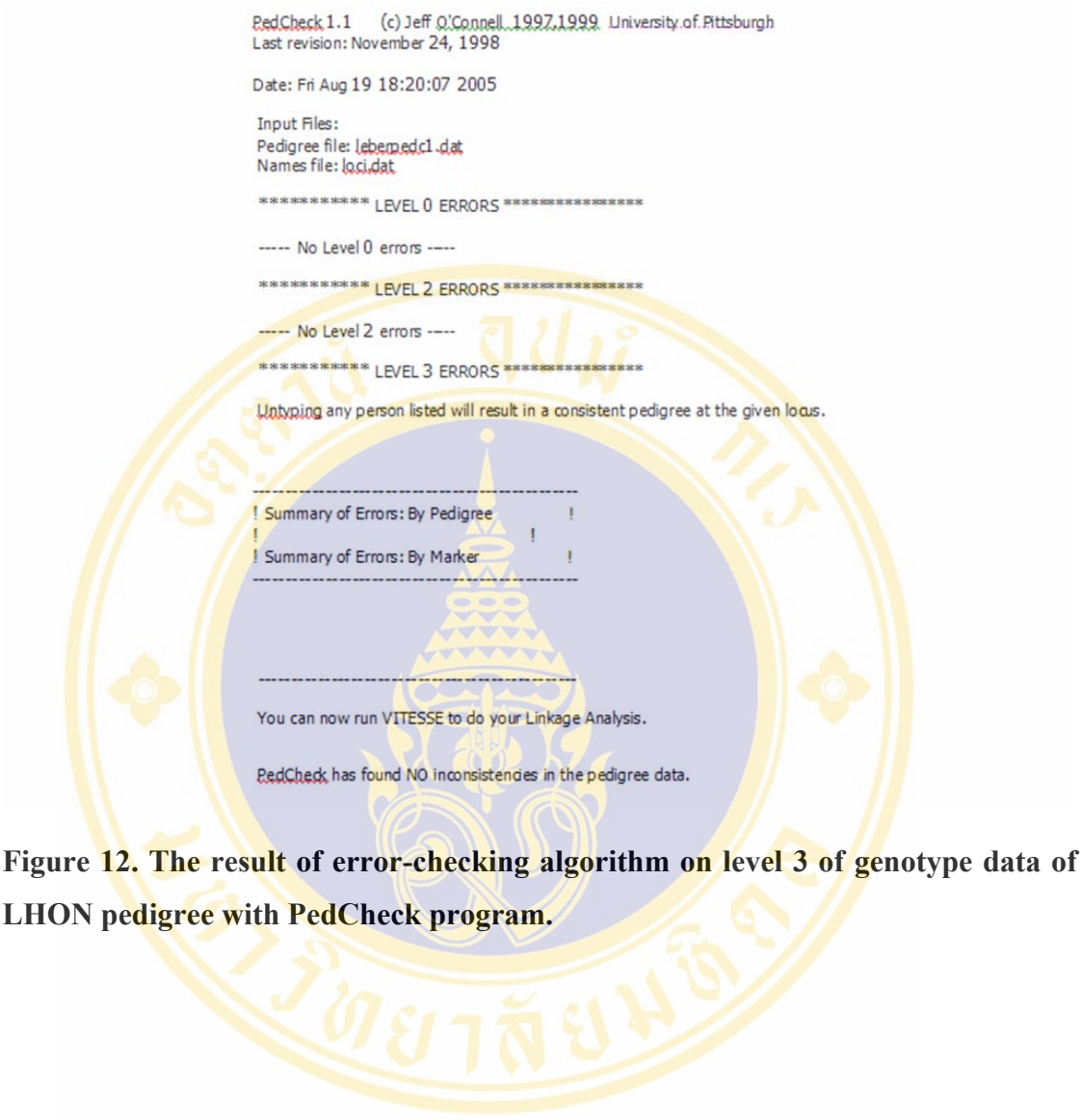


Figure 11. The result of error-checking algorithm on level 2 of genotype data of LHON pedigree with PedCheck program.



```
PedCheck 1.1     (c) Jeff O'Connell  1997,1999  University of Pittsburgh
Last revision: November 24, 1998

Date: Fri Aug 19 18:20:07 2005

Input Files:
Pedigree file: leberpedc1.dat
Names file: loci.dat

********** LEVEL 0 ERRORS **********

----- No Level 0 errors -----

********** LEVEL 2 ERRORS **********

----- No Level 2 errors -----

********** LEVEL 3 ERRORS **********

Untyping any person listed will result in a consistent pedigree at the given locus.

-----------------------------------------
! Summary of Errors: By Pedigree       !
!                                  !
! Summary of Errors: By Marker         !
-----------------------------------------


-----------------------------------------
You can now run VITESSE to do your Linkage Analysis.

PedCheck has found NO inconsistencies in the pedigree data.
```

Figure 12. The result of error-checking algorithm on level 3 of genotype data of LHON pedigree with PedCheck program.

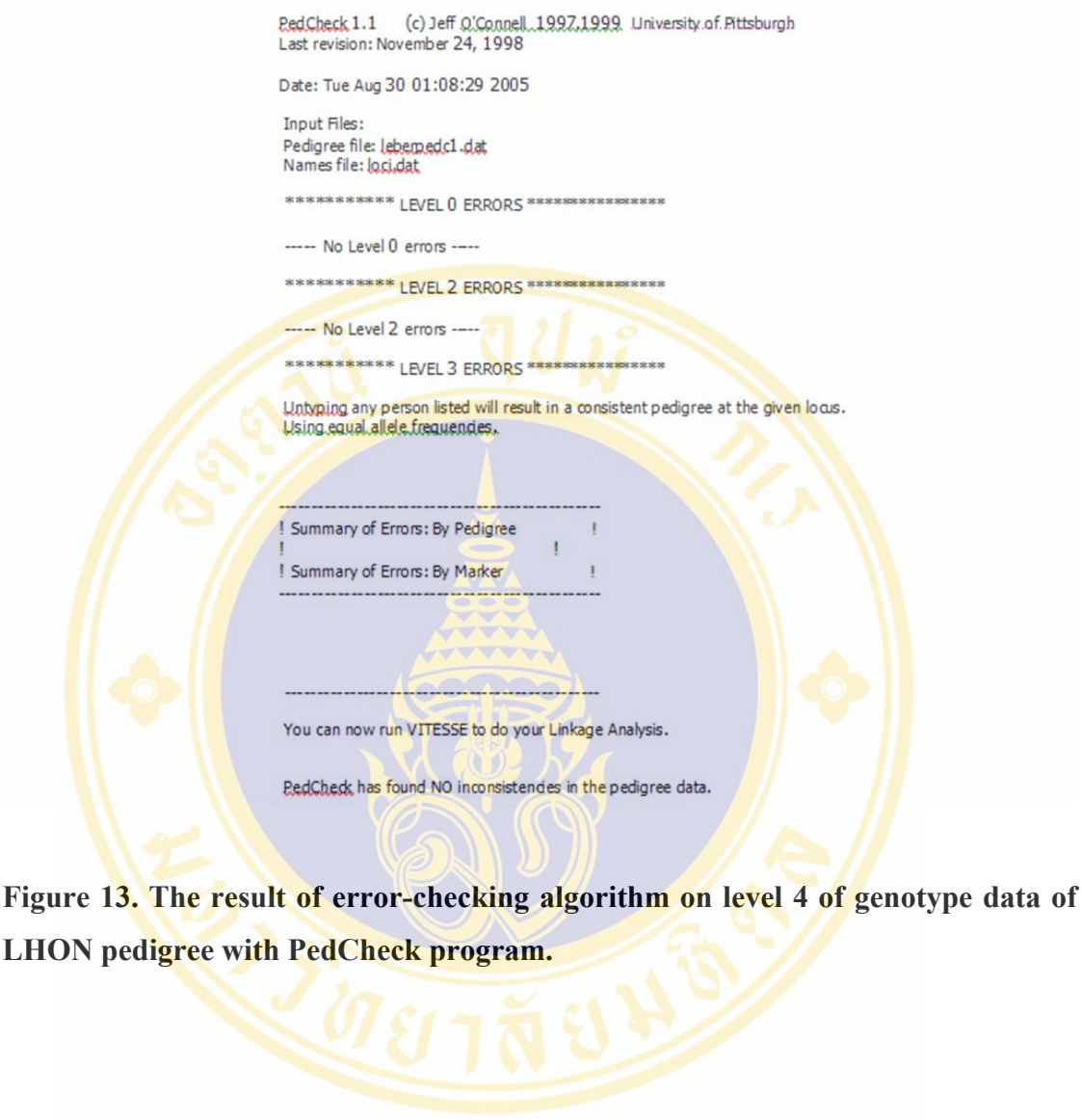


Figure 13. The result of error-checking algorithm on level 4 of genotype data of LHON pedigree with PedCheck program.

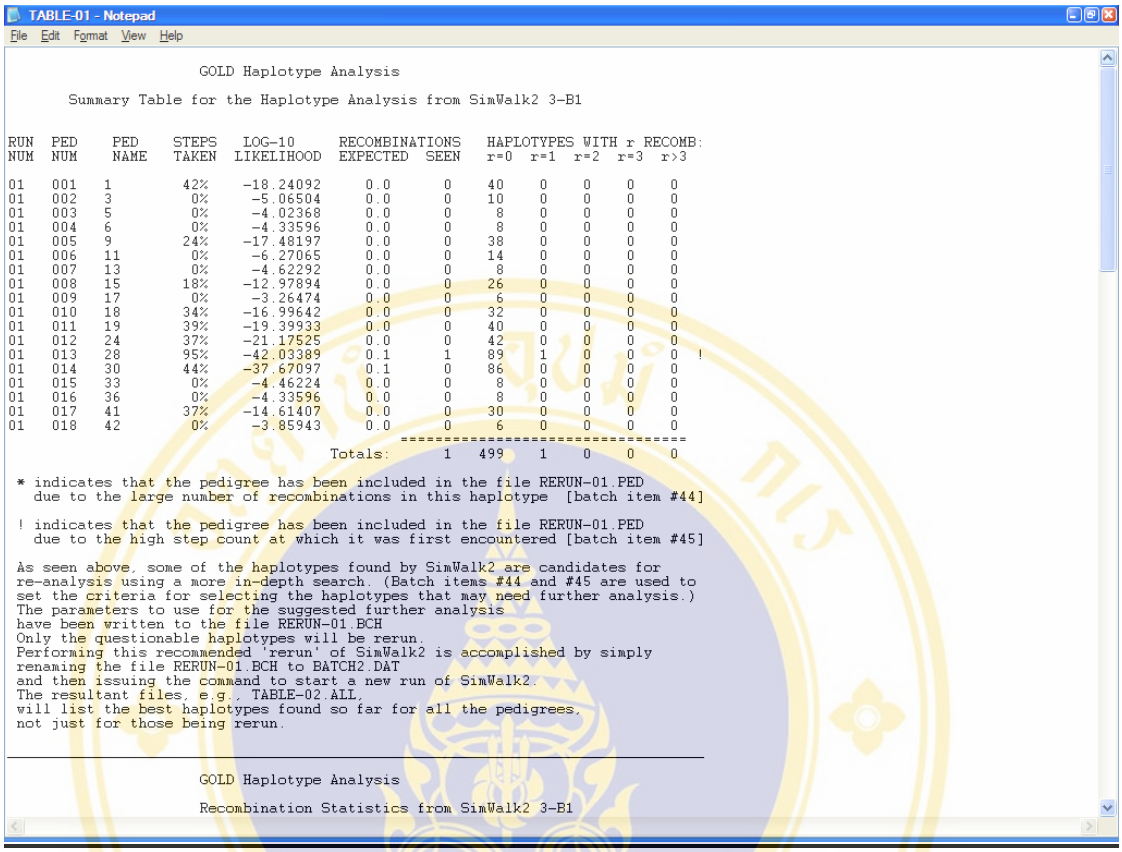


TABLE-01 - Notepad

File Edit Format View Help

```
                    GOLD Haplotype Analysis

      Summary Table for the Haplotype Analysis from SimWalk2 3-B1


RUN  PED    PED    STEPS   LOG-10     RECOMBINATIONS   HAPLOTYPES WITH r RECOMB:
NUM  NUM    NAME   TAKEN  LIKELIHOOD  EXPECTED  SEEN   r=0  r=1  r=2  r=3  r>3

01   001    1        42%    -18.24092     0.0       0     40    0    0    0    0
01   002    3         0%     -5.06504     0.0       0     10    0    0    0    0
01   003    5         0%     -4.02368     0.0       0      8    0    0    0    0
01   004    6         0%     -4.33596     0.0       0      8    0    0    0    0
01   005    9        24%    -17.48197     0.0       0     38    0    0    0    0
01   006    11        0%     -6.27065     0.0       0     14    0    0    0    0
01   007    13        0%     -4.62292     0.0       0      8    0    0    0    0
01   008    15       18%    -12.97894     0.0       0     26    0    0    0    0
01   009    17        0%     -3.26474     0.0       0      6    0    0    0    0
01   010    18       34%    -16.99642     0.0       0     32    0    0    0    0
01   011    19       39%    -19.39933     0.0       0     40    0    0    0    0
01   012    24       37%    -21.17525     0.0       0     42    0    0    0    0
01   013    28       95%    -42.03389     0.1       1     89    1    0    0    0  !
01   014    30       44%    -37.67097     0.1       0     86    0    0    0    0
01   015    33        0%     -4.46224     0.0       0      8    0    0    0    0
01   016    36        0%     -4.33596     0.0       0      8    0    0    0    0
01   017    41       37%    -14.61407     0.0       0     30    0    0    0    0
01   018    42        0%     -3.85943     0.0       0      6    0    0    0    0
                                                      ================================
                                        Totals:       1   499    1    0    0    0

 * indicates that the pedigree has been included in the file RERUN-01.PED
   due to the large number of recombinations in this haplotype  [batch item #44]

 ! indicates that the pedigree has been included in the file RERUN-01.PED
   due to the high step count at which it was first encountered [batch item #45]

 As seen above, some of the haplotypes found by SimWalk2 are candidates for
 re-analysis using a more in-depth search. (Batch items #44 and #45 are used to
 set the criteria for selecting the haplotypes that may need further analysis.)
 The parameters to use for the suggested further analysis
 have been written to the file RERUN-01.BCH
 Only the questionable haplotypes will be rerun.
 Performing this recommended 'rerun' of SimWalk2 is accomplished by simply
 renaming the file RERUN-01.BCH to BATCH2.DAT
 and then issuing the command to start a new run of SimWalk2.
 The resultant files, e.g., TABLE-02.ALL,
 will list the best haplotypes found so far for all the pedigrees,
 not just for those being rerun.

 ____________________________________________________________________________

                    GOLD Haplotype Analysis

                    Recombination Statistics from SimWalk2 3-B1
```

Figure 14. The summary table for the haplotype analysis of 3 SNPs in *NDUFS4* gene using SimWalk2 and GOLD program.

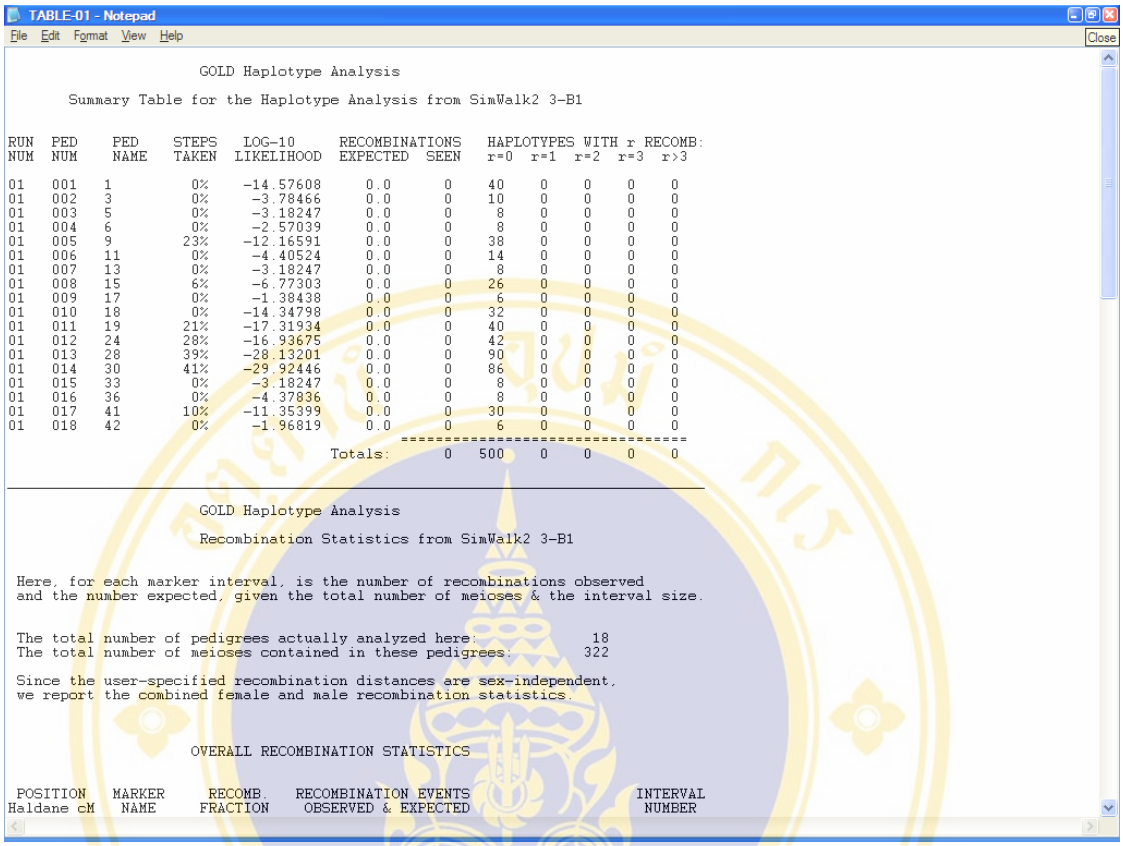


TABLE-01 - Notepad

File Edit Format View Help

Close

GOLD Haplotype Analysis

Summary Table for the Haplotype Analysis from SimWalk2 3-B1

RUN NUM	PED NUM	PED NAME	STEPS TAKEN	LOG-10 LIKELIHOOD	RECOMBINATIONS EXPECTED	RECOMBINATIONS SEEN	HAPLOTYPES WITH r RECOMB: r=0	r=1	r=2	r=3	r>3
01	001	1	0%	-14.57608	0.0	0	40	0	0	0	0
01	002	3	0%	-3.78466	0.0	0	10	0	0	0	0
01	003	5	0%	-3.18247	0.0	0	8	0	0	0	0
01	004	6	0%	-2.57039	0.0	0	8	0	0	0	0
01	005	9	23%	-12.16591	0.0	0	38	0	0	0	0
01	006	11	0%	-4.40524	0.0	0	14	0	0	0	0
01	007	13	0%	-3.18247	0.0	0	8	0	0	0	0
01	008	15	6%	-6.77303	0.0	0	26	0	0	0	0
01	009	17	0%	-1.38438	0.0	0	6	0	0	0	0
01	010	18	0%	-14.34798	0.0	0	32	0	0	0	0
01	011	19	21%	-17.31934	0.0	0	40	0	0	0	0
01	012	24	28%	-16.93675	0.0	0	42	0	0	0	0
01	013	28	39%	-28.13201	0.0	0	90	0	0	0	0
01	014	30	41%	-29.92446	0.0	0	86	0	0	0	0
01	015	33	0%	-3.18247	0.0	0	8	0	0	0	0
01	016	36	0%	-4.37836	0.0	0	8	0	0	0	0
01	017	41	10%	-11.35399	0.0	0	30	0	0	0	0
01	018	42	0%	-1.96819	0.0	0	6	0	0	0	0
				Totals:		0	500	0	0	0	0

GOLD Haplotype Analysis

Recombination Statistics from SimWalk2 3-B1

Here, for each marker interval, is the number of recombinations observed and the number expected, given the total number of meioses & the interval size.

The total number of pedigrees actually analyzed here: 18
The total number of meioses contained in these pedigrees: 322

Since the user-specified recombination distances are sex-independent, we report the combined female and male recombination statistics.

OVERALL RECOMBINATION STATISTICS

POSITION Haldane cM	MARKER NAME	RECOMB. FRACTION	RECOMBINATION EVENTS OBSERVED & EXPECTED	INTERVAL NUMBER

Figure 15. The summary table for the haplotype analysis of 2 SNPs in *NDUFV2* gene using SimWalk2 and GOLD program.

Then, "HAPLO-01" files generated from Simwalk2 program were used to paint pedigrees with their haplotype by HaploPainter program. From haplotype estimation of *NDUFS4* gene, one recombination occurring in pedigree 28 was detected (Figure 14). However, from *NDUFS4* haplotype drawing of "Family-28", the recombination detected could be accurate transmission from their parents (Appendix A). No recombination occurred in our pedigree data sets of haplotype estimation in *NDUFS4* and *NDUFV2* gene. Eighteen pedigrees with their *NDUFS4* and *NDUFV2* haplotype were displayed in appendix A. Haplotype estimated from SNP in *NDUFS8* gene was not performed since it had only one marker.

4.3 Family based association analysis with LHON phenotype using FBAT program

Family-based association test (FBAT) program was used to search the transmission and nontransmission of parental alleles to affected offspring which familiar to transmission disequilibrium test (TDT). Estimated haplotype in 18 LHON pedigrees was constructed from 250 individuals (191 with genotyping of 6 validated SNPs and 59 without genotyping) (Appendix A and B). All SNPs were in Hardy-Weinberg equilibrium; no Mendelian inconsistencies in genotyping of each SNPs, as well as no recombination occurred in haplotype estimation in *NDUFS4* and *NDUFV2* gene of our pedigree data sets.

Although, the LHON phenotypes (affected and unaffected) are clearly use as the general traits in FBAT running. The LHON affected status also depends on the effect of multiple confounding factors. Gender, secondary mutation, and heteroplasmic mutation load were significantly effect the age-dependent penetrance of our LHON pedigrees (Phasukkijwatana *et al*). Despites the effect of gender, heteroplasmy (mtDNA mutation loaded), and secondary mutation, the maternal lineage and the 11778 mutation were included in this analysis. These effects were assessed before performing family-based genetic study. Survival analysis was suitable for our data, while LHON development was the event of interest and the time until disease developed was the survival time. The Kaplan-Meier survival functions *S(t)*, were used to estimate the LHON-free surviving probability. This was the probability that a person still not develop LHON beyond survival time *t* which could obtain separately for each covariate factors. The incidence rates for each covariate to LHON

manifestation and the equality of survivor functions tested using a log-rank statistic were shown in figure 16-20. From figure 16, male patients developed disease faster than females with significant difference (P<0.0001). The 11778 mutation load also showed an effect to LHON-free survivor function (Figure 17). The homoplasmy of 11778 mutation was associated with the LHON expression. Most of the patients carried homoplasmic mutation (>95%). Individuals who carried the mutation load between 70-95% and lower than 70% showed low risk to develop the disease. The log-rank test found strongly significant difference (P<0.0001) among these 3 groups (Figure 17). In addition, the secondary mutation also influenced the LHON expression seen from Kaplan-Meier survival estimation (Figure 18). Patients with C3497T, with G3316A, without any secondary LHON mutation had a significant difference to disease-free survivor functions (P = 0.0001). The maternal lineage and the G11778A LHON mutation were also tested as the covariates. Neither of family members who are non-maternal lineage nor negative for 11778 mutation showed any risk to LHON expression (Figure 19, 20).

To regress survival outcome (not undergoing affected) against covariates to LHON, univariate and multivariate survival analyses were fitted to the Cox proportional hazard model. As illustrated, the 11778 mutation and maternal lineage were at high risk to LHON expression with the hazard ratio 2.09e+16 and 4.93e+16 respectively (Table 16). All patients were in maternal lineage and had positive for 11778 mutation. Whereas, the disease risk due to heteroplasmy was 4.04 and due to gender was 3.53, both factors were highly significance at P <0.001. There was no difference in the effect of secondary mutation group to LHON surviving (relative risk = 1.44, P = 0.095). The results were summarized in table 16. Further, multivariate analysis was performed with the risk of univariate testing at P<0.1. Interaction of 5 multiple confounding factors showed that the relative risk for sex and secondary mutation to LHON were 4.54 and 3.52 times respectively with a highly significant (P<0.001). Heteroplasmy remained the high risk for LHON manifestation (RR = 3.59, P = 0.001). The effect of 11778 mutation was still high risk to disease survival at 1.25e+16, whereas the hazard ratio of maternal lineage decreased to 0.60. The results were summarized in Table 17. Then, the Cox-Snell residual from Cox hazard regression model were used to estimate survivor function to be used in further FBAT

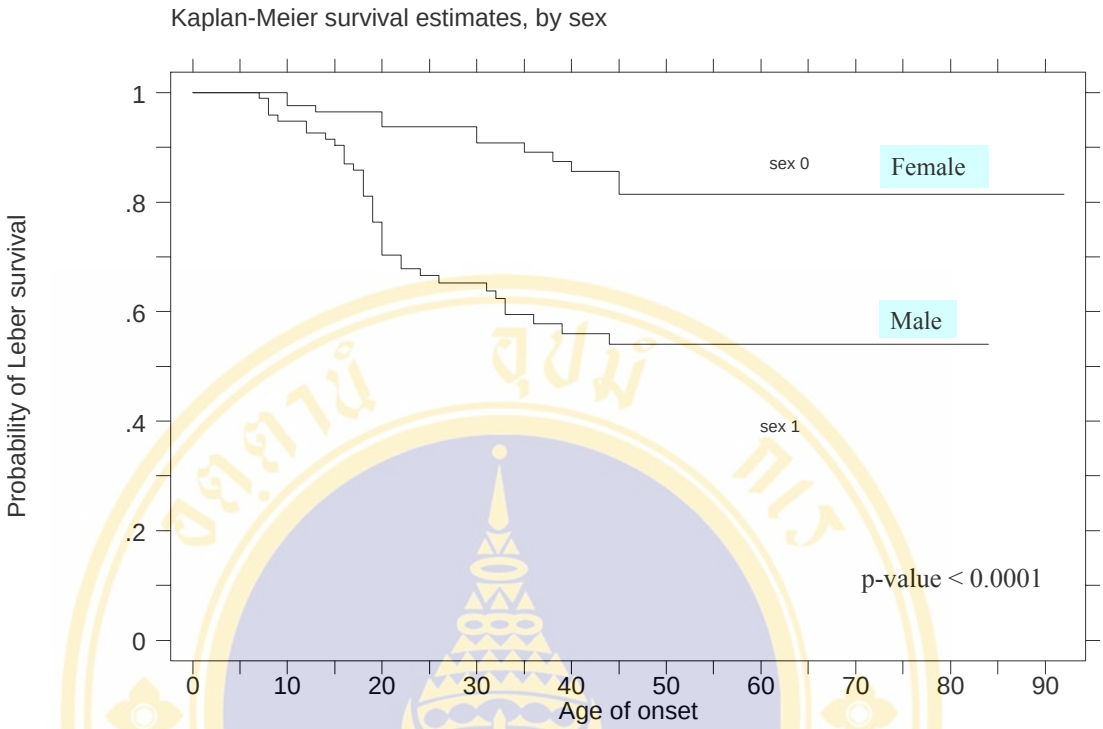


Figure 16. Kaplan-Meier estimation of LHON-free survivor functions subdivided by gender. Male patients developed disease faster than females with significant difference (*P*<0.0001).

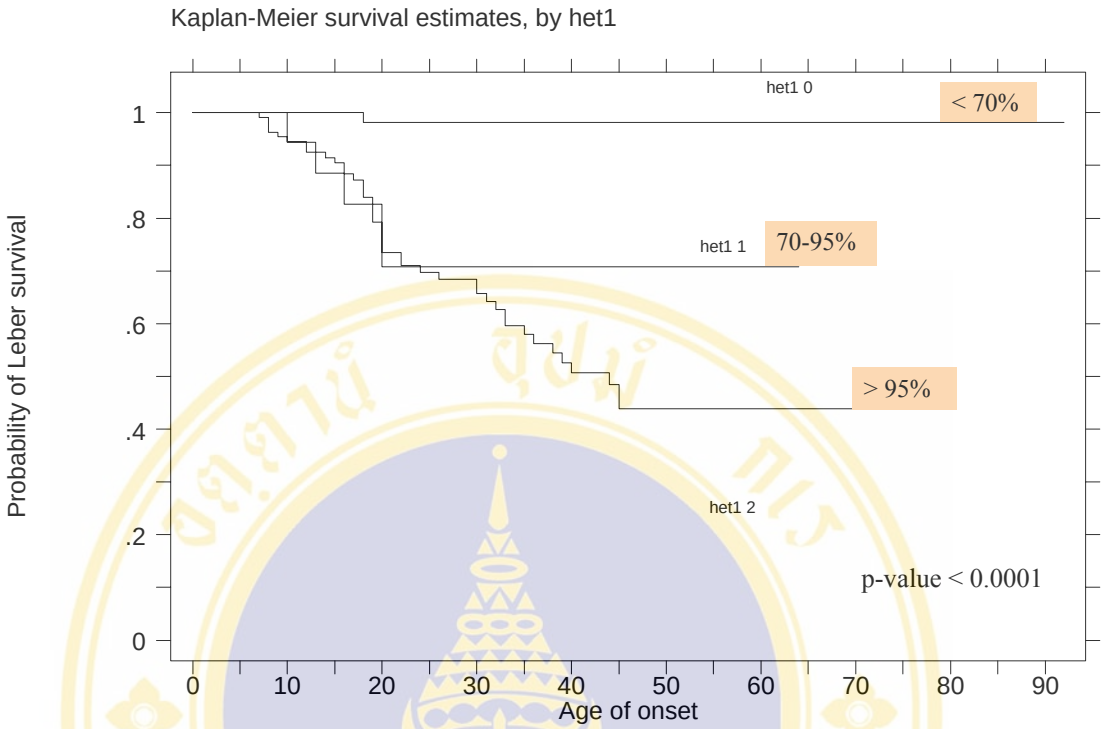


Figure 17. Kaplan-Meier estimation of LHON-free survivor functions subdivided by heteroplasmic proportion. Patients carried the 11778 mutation over 95% had a higher risk to develop the disease than the mutation load 70-95%. The mutation load less than 70% showed the lowest risk to develop disease (*P*<0.0001).

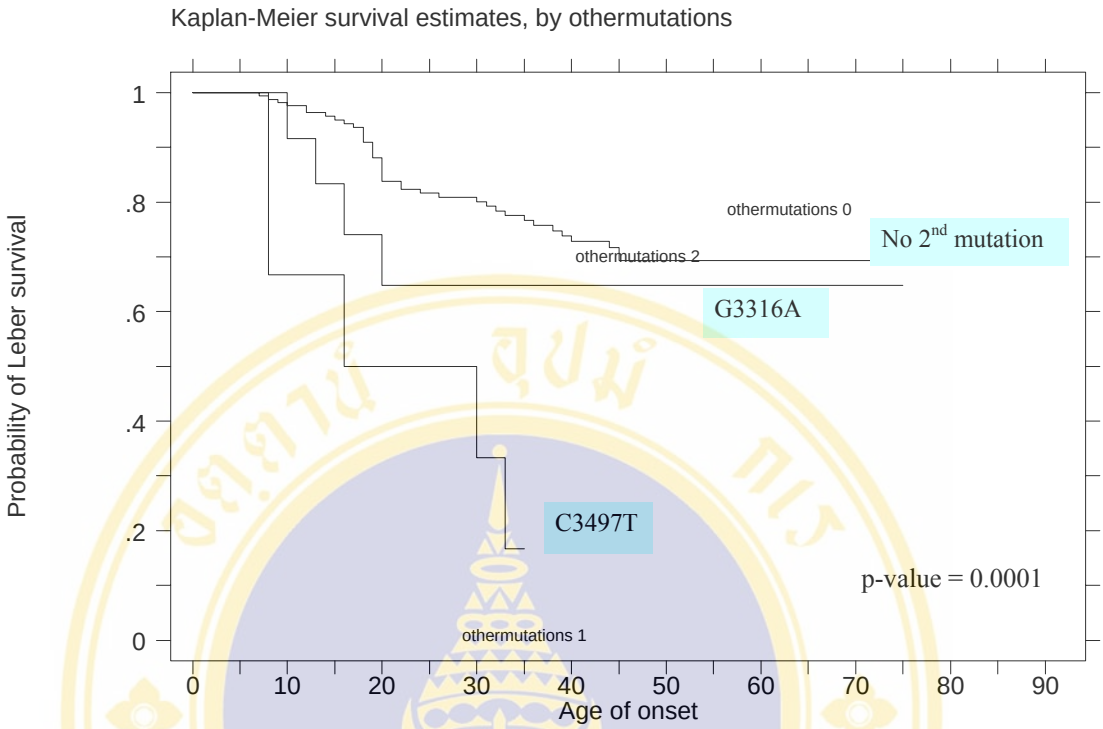


Figure 18. Kaplan-Meier estimation of LHON-free survivor functions subdivided by secondary mtDNA mutation. The C3497T and G3316A significantly influence the risk of LHON development (*P* = 0.0001).

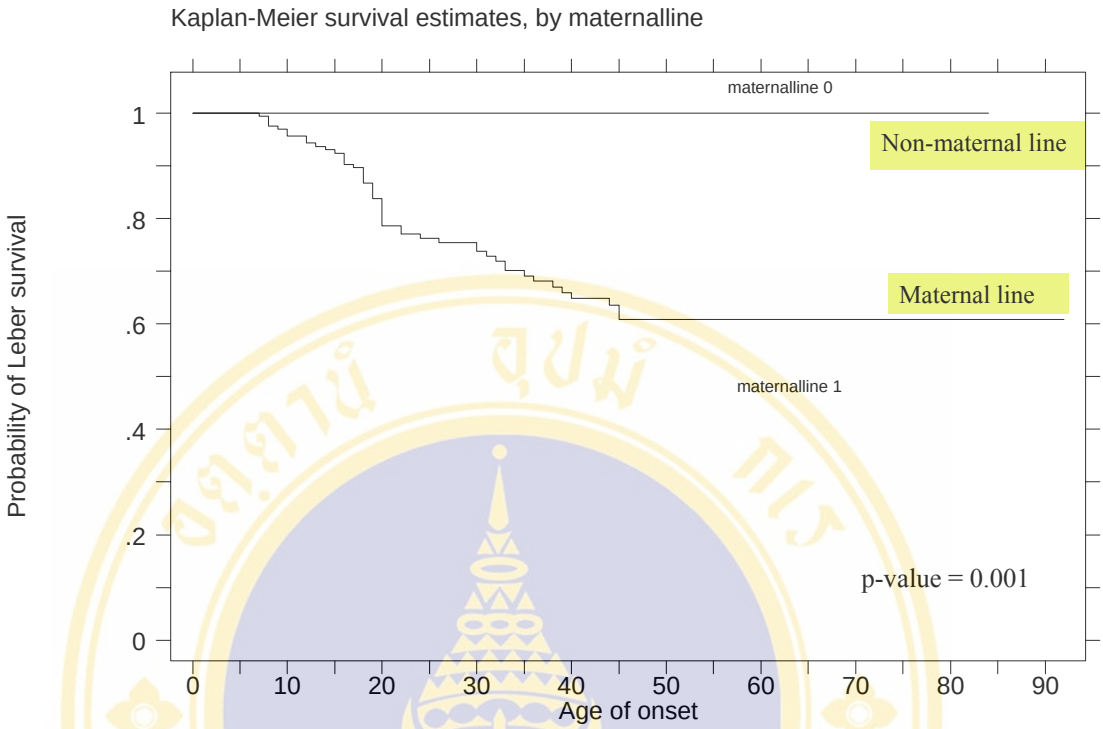


Figure 19. Kaplan-Meier estimation of LHON-free survivor functions subdivided by maternal lineage. None of non-maternal lineage members has any risk to develop disease (*P* = 0.001).

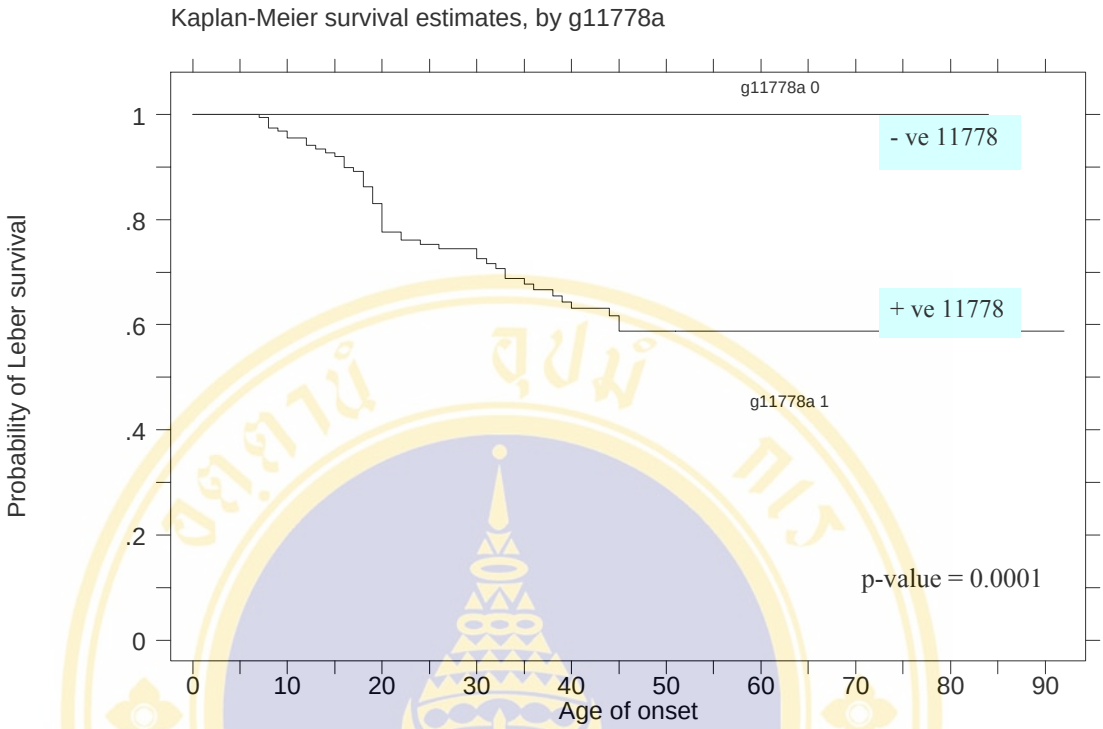


Figure 20. Kaplan-Meier estimation of LHON-free survivor functions subdivided by 11778 mtDNA mutation. None of LHON family members with negative of 11778 mutation has any risk to develop disease (*P* = 0.0001).

Table 16. Unicovariate Cox proportional hazards models.

Variable	Relative risk	Standard error	95% CI	*P*-value
sex	3.528075	1.175381	1.836367-6.778228	<0.001
heteroplasmy	4.038703	1.223296	2.230584-7.312489	<0.001
secondary mutation	1.437689	0.3121901	0.9393537-2.200397	0.095
G11778A	2.09e+16	NA	NA	NA
maternal lineage	4.93e+16	NA	NA	NA

NA = not available

Table 17. Multicovariate Cox proportional hazards models.

Variable	Relative risk	Standard error	95% CI	*P*-value
sex	4.542122	1.576886	2.300105 - 8.969536	< 0.001
heteroplasmy	3.589251	1.359841	1.708101 - 7.542132	0.001
secondary mutation	3.518501	1.012534	2.001722 - 6.1846	< 0.001
G11778A	1.25e+16	NA	NA	NA
maternal lineage	0.6044831	NA	NA	NA

NA = not available

analysis. The LHON survivor functions were taken to a phenotype data file for further FBAT running.

In the single-locus FBAT analysis, a total of 70 nuclear LHON families (53 patients and 197 members) with at least one LHON-affected offspring were identified. None of polymorphisms were associated with LHON manifestation (from FBAT command). The result was shown in Table 18. Then, haplotype family-based association analysis of *NDUFS4* and *NDUFV2* gene was observed. The 3 bi-allelic markers in *NDUFS4* gene had rise to 8 potential haplotypes, only 5 haplotypes can be derived from the marker phenotype data. While, the 2 bi-allelic markers in *NDUFS8* gene provided for all 4 potential haplotypes. No association between LHON survival function and *NDUFS4* as well as *NDUFV2* gene in haplotype FBAT detected (from HBAT command). Results were shown in Table 19 and 20.

Table 18. Single-locus family-based association analysis of LHON survival functions with SNPs in *NDUFS4*, *NDUFS8*, and *NDUFV2* gene.

Gene/SNP	Frequency of common allele	Informative families	S	E(S)	Z	*P* value
NDUFS4						
-447 T→A	0.546	30	72.17	78.777	-1.668	0.095
+12 C→G	0.567	25	65.055	61.917	0.94	0.347
+312 G→A	0.761	24	77.08	75.7	0.449	0.653
NDUFS8						
-496 C→G	0.62	32	76.196	76.926	-0.19	0.849
NDUFV2						
-602 G→A	0.748	24	75.313	72.285	1.129	0.259
+86 T→C	0.747	23	69.396	67.908	0.573	0.567

S is the FBAT statistic counting the number of a given allele in affected offspring
E(S) is the expectation of S under the null hypothesis of no association
P value, level of significance corresponding to Z ($P < 0.05$)

Table 19. Haplotype family-based association analysis of LHON survival functions with SNPs in *NDUFS*4 gene.

Haplotype	-447 T→A	+12 C→G	+312 G→A	EM estimates of frequency	Informative families	S	E(S)	Z	*P* value
H1	A	C	G	0.504	29	95.081	87.534	1.845	0.065025
H2	T	G	A	0.217	24	48.661	49.892	-0.366	0.714039
H3	T	G	G	0.177	14	22.279	24.355	-0.984	0.325348
H4	T	C	G	0.098	10	NA	NA	NA	NA
H5	A	G	G	0.004	0	NA	NA	NA	NA

S is the FBAT statistic counting the number of a given allele in affected offspring
E(S) is the expectation of S under the null hypothesis of no association
P value, level of significance corresponding to Z ($P < 0.00625$)
NA = not available (informative families < 15%)

Table 20. Haplotype family-based association analysis of LHON survival functions with SNPs in *NDUFV2* gene.

Haplotype	-602 G→A	+86 T→C	EM estimates of frequency	Informative families	S	E(S)	Z	*P* value
H1	A	C	0.781	24	84.922	81.883	1.077	0.2817
H2	G	T	0.189	21	27.727	29.205	-0.575	0.565114
H3	A	T	0.019	1	NA	NA	NA	NA
H4	G	C	0.011	2	NA	NA	NA	NA

S is the FBAT statistic counting the number of a given allele in affected offspring
E(S) is the expectation of S under the null hypothesis of no association
P value, level of significance corresponding to Z ($P < 0.0125$)
NA = not available (informative families < 15%)

CHAPTER V
DISCUSSION

The LHON pathogenesis was majority caused by mtDNA point mutations in genes that coded for respiratory chain complex I subunits. The G11778A, T14484C and G3460A were the most common mutations in LHON cases worldwide. Similar in Thailand, 93% of our LHON pedigrees carrying G11778A and the remainders (7%) were T14484C mutation. None of our case carried the G3460A mutation so far. However, an incomplete penetrance and male predominance in this disease could not be explained by maternal inheritance solely. In our LHON pedigrees harboring the G11778A mutation, not individuals in maternal lineage become affected; even though they were the same sex in the same generation and carried the same proportion of mutant mtDNA that transmitted from their mother. Only about 37% of males and 13% of females actually develop the optic neuropathy (1). This was different in each population. An intriguing aspect of LHON had pointed to the possible role of a nuclear factor(s) for the disease expression (2). The reduced in mitochondrial respiratory complex I activity in lymphoblast and cybrid cells of LHON cases was evidenced. This complex was the largest complex of respiratory chain enzyme complexes in mitochondria and played an important role in OXPHOS system. Out of 45, 38 subunits of complex I were coded from nuclear genome (nDNA) whereas the remainders were encoded by mtDNA. Thus the mutations in nuclear-encoded subunits of complex I could co-interact with pathogenic LHON mutation in mtDNA and deteriorate mitochondrial functions leading to optic neuropathy.

SNPs association study in nuclear-encoded genes for complex I subunits as nuclear modifier to LHON manifestation

In this study, *NDUFS4*, *NDUFS8* and *NDUFV2* gene, whose their function are necessary for complex I activity were analyzed as nuclear modifying factors in LHON. Coding and promoter regions of these 3 candidate genes were sequenced and

nucleotide polymorphisms were identified. Six validated SNPs, -447 T→A, +12 C→G and +312 G→A in *NDUFS4*, the -496 C→G in *NDUFS8*, the -602 G→A and +86 T→C in *NDUFV2* had their minor allele frequency of more than 10% in normal population, thus supposed to be common variants and available in association study as the disease susceptibility or marker loci. These polymorphisms could be genotyped by PCR-RFLP. All genotyping data of LHON family members and normal individuals were in Hardy-Weinberg equilibrium. It was trusted that all validated SNPs did not undergo any evolution and all samples could be studied without spurious association resulting from population admixture. 4 SNPs, the +12 C→G and +312 G→A in coding region of *NDUFS4*, and the other 2 SNPs, the -602 G→A in coding and +86 T→C in promoter region of *NDUFV2* showed their association with LHON phenotype compared to normal controls. Based on association study design, family-based controls were also performed. Covariate factors to LHON manifestation were analyzed within this family-based test in order to search for nuclear modifier gene(s). Our results did not show any association of these SNPs to LHON patients compared to unaffected family members who also carried the mtDNA mutation.

NDUFS4, *NDUFS8* and *NDUFV2* gene were necessary for complex I assembly and mitochondrial functions (3-6). Many polymorphisms in these genes had been associated with complex I deficiency (5, 7-10), neurological diseases (11-14) as well as the mitochondrial disorders (15). The *NDUFS4* gene coded for 18 kDa (IP) subunit of complex I localized at position 5q11.1 of chromosome 5, contained five exons intercalated by four large introns (3). The cDNA encoded for 175 amino acids which were highly conserved among many mammalial species. Human NDUFS4 protein contains two conserved phosphorylation sites. One site was located in the leader sequence (RSLS) of the protein, and the other was in the mature protein (RVS) (3, 16). This protein subunit was phosphorylated by cAMP-dependent protein kinase (PKA) and dephosphorylated by PP2Cγ-type phosphatase in mitochondrial inner membrane. The phosphatase was Mg^{2+} -dependent and was inhibited by Ca^{2+}. In *vivo* studies, elevation of intracellular cAMP level can promote the phosphorylation of the 18 kDa subunit that stimulates complex I activity and NAD-linked mitochondrial respiration (4). Variants that effected these phosphorylation sites had been proposed to suppress the cAMP activation of complex I (4), impair its normal function as well as the protein

importing to mitochondria (3, 4). In this study, 3 validated SNPs in *NDUFS4* gene comprising the -447 T→A, +12 C→G, and +312 G→A were tested as haplotype markers in the analysis of family-based association to LHON. The +312 G→A polymorphism had been reported in patients with complex I deficiency (9, 17). While in human sequences, SNP -447 T→A was in the essential site for transcription factor binding of HNF-4α, GR, RXR-α, and VDR proteins. These transcription factors were members of Zinc finger motif of nuclear receptor type that mediated DNA-binding (http://polly.wustl.edu/promolign/main.html). However, our family-based association study did not find any linkage disequilibrium between SNP haplotype markers in *NDUFS4* gene and the LHON susceptibility loci.

The human *NDUFS8* gene encoded 23 kDa (TYKY) subunit of complex I. This gene is located on chromosome 11q13 spanning about 6 kb with seven exons (6, 18). The TYKY protein sequences contained two four-cysteines that were presumed to be the ligand of two [4Fe-4S] ion-sulfur clusters. Thus it made the TYKY subunit a very critical subunit of complex I and played important role in complex I assembly (6, 18, 19). This protein was thought to participate in electron transfer and in the reduction of ubiquinone coupled with the proton pumping (20). Furthermore, the promoter region of this gene had several transcription factors binding (TFB) sites for YY1, Sp1, NRF-1 and NRF-2 protein (6, 21). The Sp1 and NRF1 binding site motifs were presented in the first intron of *NDUFS8* that might also be important factors controlling the efficiency of transcription of this gene (6, 21, 22). The nuclear respiratory factors, NRF-1 and NRF-2 were well characterized as regulators of oxidative phosphorylation (OXPHOS) gene transcription. In this study, the -496 C→G located in promoter region of *NDUFS8* gene, was determined. This polymorphism has never been reported to involve in any TFB sites of the *NDUFS8* gene. This variant did not show an effective loci associated with LHON phenotype either.

The last candidate gene in this study, the *NDUFV2*, coded for 24 kDa subunit of complex I spanning 31.5 kb and contained eight exons that interrupted by seven introns (23). Two distinct loci for this subunit were detected in the human genome, the functional gene (NDUFV2) transcribed from chromosome 18, while the pseudogene (NDUFV2P1) was an inactive locus on chromosome 19 (24). The 24 kDa subunit was constituted on flavoprotein fraction (FP) which played an important role

in the catalytic function of complex I (25, 26). This subunit contained one 2Fe-2S cluster involving in the electron transport and was suggested to be a G protein. Thus the protein was important in the enzyme function. There had been an association between bipolar disorder and the haplotype of polymorphisms in upstream region including the SNP -602 G→A of *NDUFV2* gene (12, 13). Additionally, the +86 T→C transition was found in association with Parkinson disease (11). This mutation altered an amino acid from valine to alanine at position 29 and caused a change in the hydropathy profile and secondary structure of signal peptide in this gene. Affected in signal peptide was critical for processing-recognition events by mitochondrial processing peptidase (MPP) that can influence the level of the mature protein 24 kDa subunit of complex I. Nevertheless, in our analysis, either -602 G→A or +86 T→C polymorphisms was not associated with the LHON phenotypes.

Although, the results presented in this study reported the strong significance difference (P<0.005) in allele frequency of 2 SNPs each in *NDUFS4* and *NDUFV2* gene between Thai normal population and LHON cases. Family-based controls were more suitable for association study design due to the pertaining of the same mitochondrial effects. Furthermore, the family-based association study could discard the effect of population admixture on the basis of LD application. Using FBAT analysis with the Cox proportional hazard model to predict multicovariate factors to LHON, our results found no disease parental alleles in these 3 candidate genes, *NDUFS4*, *NDUFS8* and *NDUFV2* gene, transmitted to affected offspring. It was suggested that these three candidate genes were unlikely to be the nuclear modifier genes for optic neuropathy in LHON.

On account of these three candidate genes, *NDUFS4*, *NDUFS8* and *NDUFV2* gene, were unlikely to be the modifying factor to LHON. Other nuclear-encoded genes for complex I subunits are still the good candidate genes for family-based association studies in LHON.

LHON survival functions and multivariate influence disease phenotypes

In our survival analysis with Kaplan-Meier curves, patients with homoplasmic of G11778A mutation (>95% mutant mtDNA) had the highest risk to develop disease. The moderate risk of disease expression was the mutation load between 70 and 95%.

The penetrance of this mutant level (70-95%) was proposed to be equal to the homoplasmic patients and there was no risk to develop this disease when the mutation load was less than 60% (27). This study also confirmed the predilection in males to lose vision than females. The secondary mutation also influenced the expression of LHON. These 3 intriguing factors, heteroplasmy, gender and secondary mutation showed the relative risk for the disease development analyzed by Cox-proportional hazard models. These factors were suggested to be covariance for LHON expression that should be concerned in the study of the effect of genetic association.

REFERENCES

1. Leber T. Ueber hereditare und congenital-angelegte Sehnervenleiden. Graefe's Archive for Clinical and Experimental Ophthalmology (Historical Archive) 1871;17(2):249-291.
2. Nikoskelainen E. The clinical findings in Leber's hereditary optic neuroretinopathy. Leber's disease. Trans Ophthalmol Soc U K 1985;104 (Pt 8):845-52.
3. Newman NJ, Wallace DC. Mitochondria and Leber's hereditary optic neuropathy. Am J Ophthalmol 1990;109(6):726-30.
4. Riordan-Eva P, Sanders MD, Govan GG, Sweeney MG, Da Costa J, Harding AE. The clinical features of Leber's hereditary optic neuropathy defined by the presence of a pathogenic mitochondrial DNA mutation. Brain 1995;118 (Pt 2):319-37.
5. Howell N. Leber hereditary optic neuropathy: mitochondrial mutations and degeneration of the optic nerve. Vision Res 1997;37(24):3495-507.
6. Brown MD, Voljavec AS, Lott MT, MacDonald I, Wallace DC. Leber's hereditary optic neuropathy: a model for mitochondrial neurodegenerative diseases. Faseb J 1992;6(10):2791-9.
7. Orth M, Schapira AH. Mitochondria and degenerative disorders. Am J Med Genet 2001;106(1):27-36.
8. Andrews RM, Griffiths PG, Johnson MA, Turnbull DM. Histochemical localisation of mitochondrial enzyme activity in human optic nerve and retina. Br J Ophthalmol 1999;83(2):231-5.
9. Marchetti P, Castedo M, Susin SA, Zamzami N, Hirsch T, Macho A, et al. Mitochondrial permeability transition is a central coordinating event of apoptosis. J Exp Med 1996;184(3):1155-60.
10. Nieminen AL, Petrie TG, Lemasters JJ, Selman WR. Cyclosporin A delays mitochondrial depolarization induced by N-methyl-D-aspartate in cortical

neurons: evidence of the mitochondrial permeability transition. Neuroscience 1996;75(4):993-7.

11. Schinder AF, Olson EC, Spitzer NC, Montal M. Mitochondrial dysfunction is a primary event in glutamate neurotoxicity. J Neurosci 1996;16(19):6125-33.
12. Raha S, Robinson BH. Mitochondria, oxygen free radicals, and apoptosis. Am J Med Genet 2001;106(1):62-70.
13. Kluck RM, Bossy-Wetzel E, Green DR, Newmeyer DD. The release of cytochrome c from mitochondria: a primary site for Bcl-2 regulation of apoptosis. Science 1997;275(5303):1132-6.
14. Yang J, Liu X, Bhalla K, Kim CN, Ibrado AM, Cai J, et al. Prevention of apoptosis by Bcl-2: release of cytochrome c from mitochondria blocked. Science 1997;275(5303):1129-32.
15. Man PY, Turnbull DM, Chinnery PF. Leber hereditary optic neuropathy. J Med Genet 2002;39(3):162-9.
16. Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza AM, et al. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science 1988;242(4884):1427-30.
17. Huoponen K, Vilkki J, Aula P, Nikoskelainen EK, Savontaus ML. A new mtDNA mutation associated with Leber hereditary optic neuroretinopathy. Am J Hum Genet 1991;48(6):1147-53.
18. Howell N, Bindoff LA, McCullough DA, Kubacka I, Poulton J, Mackey D, et al. Leber hereditary optic neuropathy: identification of the same mitochondrial ND1 mutation in six pedigrees. Am J Hum Genet 1991;49(5):939-50.
19. Johns DR, Neufeld MJ, Park RD. An ND-6 mitochondrial DNA mutation associated with Leber hereditary optic neuropathy. Biochem Biophys Res Commun 1992;187(3):1551-7.
20. Wallace DC, Brown MD, Lott MT. Mitochondrial DNA variation in human evolution and disease. Gene 1999;238(1):211-30.
21. Brown MD, Torroni A, Reckord CL, Wallace DC. Phylogenetic analysis of Leber's hereditary optic neuropathy mitochondrial DNA's indicates multiple independent occurrences of the common mutations. Hum Mutat 1995;6(4):311-25.

22. Mackey DA, Oostra RJ, Rosenberg T, Nikoskelainen E, Bronte-Stewart J, Poulton J, et al. Primary pathogenic mtDNA mutations in multigeneration pedigrees with Leber hereditary optic neuropathy. Am J Hum Genet 1996;59(2):481-5.

23. Howell N. Leber hereditary optic neuropathy: respiratory chain dysfunction and degeneration of the optic nerve. Vision Res 1998;38(10):1495-504.

24. Chinnery PF, Brown DT, Andrews RM, Singh-Kler R, Riordan-Eva P, Lindley J, et al. The mitochondrial ND6 gene is a hot spot for mutations that cause Leber's hereditary optic neuropathy. Brain 2001;124(Pt 1):209-18.

25. Johns DR, Berman J. Alternative, simultaneous complex I mitochondrial DNA mutations in Leber's hereditary optic neuropathy. Biochem Biophys Res Commun 1991;174(3):1324-30.

26. Johns DR, Neufeld MJ. Cytochrome b mutations in Leber hereditary optic neuropathy. Biochem Biophys Res Commun 1991;181(3):1358-64.

27. Brown MD, Voljavec AS, Lott MT, Torroni A, Yang CC, Wallace DC. Mitochondrial DNA complex I and III mutations associated with Leber's hereditary optic neuropathy. Genetics 1992;130(1):163-73.

28. Brown MD, Yang CC, Trounce I, Torroni A, Lott MT, Wallace DC. A mitochondrial DNA variant, identified in Leber hereditary optic neuropathy patients, which extends the amino acid sequence of cytochrome c oxidase subunit I. Am J Hum Genet 1992;51(2):378-85.

29. Howell N, Kubacka I, Halvorson S, Mackey D. Leber's hereditary optic neuropathy: the etiological role of a mutation in the mitochondrial cytochrome b gene. Genetics 1993;133(1):133-6.

30. Huoponen K, Lamminen T, Juvonen V, Aula P, Nikoskelainen E, Savontaus ML. The spectrum of mitochondrial DNA mutations in families with Leber hereditary optic neuroretinopathy. Hum Genet 1993;92(4):379-84.

31. Oostra RJ, Bolhuis PA, Wijburg FA, Zorn-Ende G, Bleeker-Wagemakers EM. Leber's hereditary optic neuropathy: correlations between mitochondrial genotype and visual outcome. J Med Genet 1994;31(4):280-6.

32. Nikoskelainen EK, Huoponen K, Juvonen V, Lamminen T, Nummelin K, Savontaus ML. Ophthalmologic findings in Leber hereditary optic neuropathy,

with special reference to mtDNA mutations. Ophthalmology 1996;103(3):504-14.

33. Zhu DP, Economou EP, Antonarakis SE, Maumenee IH. Mitochondrial DNA mutation and heteroplasmy in type I Leber hereditary optic neuropathy. Am J Med Genet 1992;42(2):173-9.
34. Barboni P, Mantovani V, Montagna P, Bragliani M, Cortelli P, Lugaresi E, et al. Mitochondrial DNA analysis in Leber's hereditary optic neuropathy. Ophthalmic Paediatr Genet 1992;13(4):219-26.
35. Smith KH, Johns DR, Heher KL, Miller NR. Heteroplasmy in Leber's hereditary optic neuropathy. Arch Ophthalmol 1993;111(11):1486-90.
36. Black GC, Morten K, Laborde A, Poulton J. Leber's hereditary optic neuropathy: heteroplasmy is likely to be significant in the expression of LHON in families with the 3460 ND1 mutation. Br J Ophthalmol 1996;80(10):915-7.
37. Tanaka A, Kiyosawa M, Mashima Y, Tokoro T. A family with Leber's hereditary optic neuropathy with mitochondrial DNA heteroplasmy related to disease expression. J Neuroophthalmol 1998;18(2):81-3.
38. Mashima Y, Saga M, Hiida Y, Oguchi Y, Wakakura M, Kudoh J, et al. Quantitative determination of heteroplasmy in Leber's hereditary optic neuropathy by single-strand conformation polymorphism. Invest Ophthalmol Vis Sci 1995;36(8):1714-20.
39. Chinnery PF, Andrews RM, Turnbull DM, Howell NN. Leber hereditary optic neuropathy: Does heteroplasmy influence the inheritance and expression of the G11778A mitochondrial DNA mutation? Am J Med Genet 2001;98(3):235-43.
40. Yen MY, Yen TC, Pang CY, Liu JH, Wei YH. Mitochondrial DNA mutation in Leber's hereditary optic neuropathy. Invest Ophthalmol Vis Sci 1992;33(8):2561-6.
41. Howell N, Xu M, Halvorson S, Bodis-Wollner I, Sherman J. A heteroplasmic LHON family: tissue distribution and transmission of the 11778 mutation. Am J Hum Genet 1994;55(1):203-6.
42. Wallace DC. 1994 William Allan Award Address. Mitochondrial DNA variation in human evolution, degenerative disease, and aging. Am J Hum Genet 1995;57(2):201-23.

43. Torroni A, Wallace DC. Mitochondrial DNA variation in human populations and implications for detection of mitochondrial DNA mutations of pathological
43. Torroni A, Wallace DC. Mitochondrial DNA variation in human populations and implications for detection of mitochondrial DNA mutations of pathological
significance. J Bioenerg Biomembr 1994;26(3):261-71.

44. Torroni A, Petrozzi M, D'Urbano L, Sellitto D, Zeviani M, Carrara F, et al. Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. Am J Hum Genet 1997;60(5):1107-21.

45. Brown MD, Sun F, Wallace DC. Clustering of Caucasian Leber hereditary optic neuropathy patients containing the 11778 or 14484 mutations on an mtDNA lineage. Am J Hum Genet 1997;60(2):381-7.

46. Man PY, Howell N, Mackey DA, Norby S, Rosenberg T, Turnbull DM, et al. Mitochondrial DNA haplogroup distribution within Leber hereditary optic neuropathy pedigrees. J Med Genet 2004;41(4):e41.

47. Bu X, Rotter JI. Leber hereditary optic neuropathy: estimation of number of embryonic precursor cells and disease threshold in heterozygous affected females at the X-linked locus. Clin Genet 1992;42(3):143-8.

48. Newman NJ, Lott MT, Wallace DC. The clinical characteristics of pedigrees of Leber's hereditary optic neuropathy with the 11778 mutation. Am J Ophthalmol 1991;111(6):750-62.

49. Harding AE, Sweeney MG, Govan GG, Riordan-Eva P. Pedigree analysis in Leber hereditary optic neuropathy families with a pathogenic mtDNA mutation. Am J Hum Genet 1995;57(1):77-86.

50. Macmillan C, Kirkham T, Fu K, Allison V, Andermann E, Chitayat D, et al. Pedigree analysis of French Canadian families with T14484C Leber's hereditary optic neuropathy. Neurology 1998;50(2):417-22.

51. Johns DR, Smith KH, Miller NR. Leber's hereditary optic neuropathy. Clinical manifestations of the 3460 mutation. Arch Ophthalmol 1992;110(11):1577-81.

52. Johns DR, Heher KL, Miller NR, Smith KH. Leber's hereditary optic neuropathy. Clinical manifestations of the 14484 mutation. Arch Ophthalmol 1993;111(4):495-8.

53. Bu XD, Rotter JI. X chromosome-linked and mitochondrial gene control of Leber hereditary optic neuropathy: evidence from segregation analysis for dependence on X chromosome inactivation. Proceedings of the National Academy of Sciences of the United States of America 1991;88(18):8198-202.

54. Vilkki J, Ott J, Savontaus ML, Aula P, Nikoskelainen EK. Optic atrophy in Leber hereditary optic neuroretinopathy is probably determined by an X-chromosomal gene closely linked to DXS7. Am J Hum Genet 1991;48(3):486-91.

55. Nakamura M, Fujiwara Y, Yamamoto M. The two locus control of Leber hereditary optic neuropathy and a high penetrance in Japanese pedigrees. Hum Genet 1993;91(4):339-41.

56. Chen JD, Cox I, Denton MJ. Preliminary exclusion of an X-linked gene in Leber optic atrophy by linkage analysis. Hum Genet 1989;82(3):203-7.

57. Carvalho MR, Muller B, Rotzer E, Berninger T, Kommerell G, Blankenagel A, et al. Leber's hereditary optic neuroretinopathy and the X-chromosomal susceptibility factor: no linkage to DXs7. Hum Hered 1992;42(5):316-20.

58. Sweeney MG, Davis MB, Lashwood A, Brockington M, Toscano A, Harding AE. Evidence against an X-linked locus close to DXS7 determining visual loss susceptibility in British and Italian families with Leber hereditary optic neuropathy. Am J Hum Genet 1992;51(4):741-8.

59. Juvonen V, Vilkki J, Aula P, Nikoskelainen E, Savontaus ML. Reevaluation of the linkage of an optic atrophy susceptibility gene to X-chromosomal markers in Finnish families with Leber hereditary optic neuroretinopathy (LHON). Am J Hum Genet 1993;53(1):289-92.

60. Chalmers RM, Davis MB, Sweeney MG, Wood NW, Harding AE. Evidence against an X-linked visual loss susceptibility locus in Leber hereditary optic neuropathy. Am J Hum Genet 1996;59(1):103-8.

61. Pegoraro E, Carelli V, Zeviani M, Cortelli P, Montagna P, Barboni P, et al. X-inactivation patterns in female Leber's hereditary optic neuropathy patients do

not support a strong X-linked determinant. Am J Med Genet 1996;61(4):356-62.

62. Oostra RJ, Kemp S, Bolhuis PA, Bleeker-Wagemakers EM. No evidence for 'skewed' inactivation of the X-chromosome as cause of Leber's hereditary optic neuropathy in female carriers. Hum Genet 1996;97(4):500-5.

63. Handoko HY, Wirapati PJ, Sudoyo HA, Sitepu M, Marzuki S. Meiotic breakpoint mapping of a proposed X linked visual loss susceptibility locus in Leber's hereditary optic neuropathy. J Med Genet 1998;35(8):668-71.

64. Pegoraro E, Vettori A, Valentino ML, Molon A, Mostacciuolo ML, Howell N, et al. X-inactivation pattern in multiple tissues from two Leber's hereditary optic neuropathy (LHON) patients. Am J Med Genet A 2003;119(1):37-40.

65. Hudson G, Keers S, Man PYW, Griffiths P, Huoponen K, Savontaus M-L, et al. Identification of an X-Chromosomal Locus and Haplotype Modulating the Phenotype of a Mitochondrial DNA Disorder. Am J Hum Genet 2005;77(6):000-000.

66. Cock HR, Tabrizi SJ, Cooper JM, Schapira AH. The influence of nuclear background on the biochemical expression of 3460 Leber's hereditary optic neuropathy. Ann Neurol 1998;44(2):187-93.

67. Bodnar AG, Cooper JM, Holt IJ, Leonard JV, Schapira AH. Nuclear complementation restores mtDNA levels in cultured cells from a patient with mtDNA depletion. Am J Hum Genet 1993;53(3):663-9.

68. Brown MD, Trounce IA, Jun AS, Allen JC, Wallace DC. Functional analysis of lymphoblast and cybrid mitochondria containing the 3460, 11778, or 14484 Leber's hereditary optic neuropathy mitochondrial DNA mutation. J Biol Chem 2000;275(51):39831-6.

69. Yen MY, Chen YJ, Lin CH, Wang AG, Wei YH. Genetic analysis in Leber's hereditary optic neuropathy using the comparative genomic hybridization technique. Clin Experiment Ophthalmol 2003;31(5):435-8.

70. Hollander H, Makarov F, Stefani FH, Stone J. Evidence of constriction of optic nerve axons at the lamina cribrosa in the normotensive eye in humans and other mammals. Ophthalmic Res 1995;27(5):296-309.

71. Nikoskelainen EK, Savontaus ML, Wanne OP, Katila MJ, Nummelin KU. Leber's hereditary optic neuroretinopathy, a maternally inherited disease. A genealogic study in four pedigrees. Arch Ophthalmol 1987;105(5):665-71.

72. Johns DR, Smith KH, Miller NR, Sulewski ME, Bias WB. Identical twins who are discordant for Leber's hereditary optic neuropathy. Arch Ophthalmol 1993;111(11):1491-4.

73. Biousse V, Brown MD, Newman NJ, Allen JC, Rosenfeld J, Meola G, et al. De novo 14484 mitochondrial DNA mutation in monozygotic twins discordant for Leber's hereditary optic neuropathy. Neurology 1997;49(4):1136-8.

74. Lam BL. Identical twins no longer discordant for Leber's hereditary optic neuropathy. Archives of Ophthalmology 1998;116(7):956-7.

75. DuBois LG, Feldon SE. Evidence for a metabolic trigger for Leber's hereditary optic neuropathy. A case report. J Clin Neuroophthalmol 1992;12(1):15-6.

76. Cullom ME, Heher KL, Miller NR, Savino PJ, Johns DR. Leber's hereditary optic neuropathy masquerading as tobacco-alcohol amblyopia. Archives of Ophthalmology 1993;111(11):1482-5.

77. Golnik KC, Schaible ER. Folate-responsive optic neuropathy. Journal of Neuro-Ophthalmology 1994;14(3):163-9.

78. Charlmers RM, Harding AE. A case-control study of Leber's hereditary optic neuropathy. Brain 1996;119 (Pt 5):1481-6.

79. Hwang JM, Park HW. Carbon monoxide poisoning as an epigenetic factor for Leber's hereditary optic neuropathy. Korean Journal of Ophthalmology 1996;10(2):122-3.

80. Tsao K, Aitken PA, Johns DR. Smoking as an aetiological factor in a pedigree with Leber's hereditary optic neuropathy. British Journal of Ophthalmology 1999;83(5):577-81.

81. Shaikh S, Ta C, Basham AA, Mansour S. Leber hereditary optic neuropathy associated with antiretroviral therapy for human immunodeficiency virus infection. American Journal of Ophthalmology 2001;131(1):143-5.

82. Kerrison JB, Miller NR, Hsu F, Beaty TH, Maumenee IH, Smith KH, et al. A case-control study of tobacco and alcohol consumption in Leber hereditary optic neuropathy. American Journal of Ophthalmology 2000;130(6):803-12.

83. Carroll J, Shannon RJ, Fearnley IM, Walker JE, Hirst J. Definition of the nuclear encoded protein composition of bovine heart mitochondrial complex I. Identification of two new subunits. J Biol Chem 2002;277(52):50311-7.

84. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. Nature 1981;290(5806):457-65.

85. Fearnley IM, Carroll J, Shannon RJ, Runswick MJ, Walker JE, Hirst J. GRIM-19, a cell death regulatory gene product, is a subunit of bovine mitochondrial NADH:ubiquinone oxidoreductase (complex I). J Biol Chem 2001;276(42):38345-8.

86. Galante YM, Hatefi Y. Purification and molecular and enzymic properties of mitochondrial NADH dehydrogenase. Arch Biochem Biophys 1979;192(2):559-68.

87. Hatefi Y. ATP synthesis in mitochondria. Eur J Biochem 1993;218(3):759-67.

88. Loeffen JL, Smeitink JA, Trijbels JM, Janssen AJ, Triepels RH, Sengers RC, et al. Isolated complex I deficiency in children: clinical, biochemical and genetic aspects. Hum Mutat 2000;15(2):123-34.

89. Schuelke M, Smeitink J, Mariman E, Loeffen J, Plecko B, Trijbels F, et al. Mutant NDUFV1 subunit of mitochondrial complex I causes leukodystrophy and myoclonic epilepsy. Nat Genet 1999;21(3):260-1.

90. Runswick MJ, Fearnley IM, Skehel JM, Walker JE. Presence of an acyl carrier protein in NADH:ubiquinone oxidoreductase from bovine heart mitochondria. FEBS Lett 1991;286(1-2):121-4.

91. Triepels RH, Van Den Heuvel LP, Trijbels JM, Smeitink JA. Respiratory chain complex I deficiency. Am J Med Genet 2001;106(1):37-45.

92. Hattori N, Suzuki H, Wang Y, Minoshima S, Shimizu N, Yoshino H, et al. Structural organization and chromosomal localization of the human nuclear gene (NDUFV2) for the 24-kDa iron-sulfur subunit of complex I in mitochondrial respiratory chain. Biochem Biophys Res Commun 1995;216(3):771-7.

93. Hegde R. The 24-kDa subunit of the bovine mitochondrial NADH:ubiquinone oxidoreductase is a G protein. Biochem Biophys Res Commun 1998;244(3):620-9.

94. Ohnishi T. Iron-sulfur clusters/semiquinones in complex I. Biochim Biophys Acta 1998;1364(2):186-206.

95. Emahazion T, Brookes AJ. Mapping of the NDUFA2, NDUFA6, NDUFA7, NDUFB8, and NDUFS8 electron transport chain genes by intron based radiation hybrid mapping. Cytogenet Cell Genet 1998;82(1-2):114.

96. Emahazion T, Beskow A, Gyllensten U, Brookes AJ. Intron based radiation hybrid mapping of 15 complex I genes of the human electron transport chain. Cytogenet Cell Genet 1998;82(1-2):115-9.

97. Loeffen JL, Triepels RH, van den Heuvel LP, Schuelke M, Buskens CA, Smeets RJ, et al. cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: human complex I cDNA characterization completed. Biochem Biophys Res Commun 1998;253(2):415-22.

98. Skehel JM, Fearnley IM, Walker JE. NADH:ubiquinone oxidoreductase from bovine heart mitochondria: sequence of a novel 17.2-kDa subunit. FEBS Lett 1998;438(3):301-5.

99. Smeitink JA, Loeffen JL, Triepels RH, Smeets RJ, Trijbels JM, van den Heuvel LP. Nuclear genes of human complex I of the mitochondrial electron transport chain: state of the art. Hum Mol Genet 1998;7(10):1573-9.

100. Lin X, Wells DE, Kimberling WJ, Kumar S. Human NDUFB9 gene: genomic organization and a possible candidate gene associated with deafness disorder mapped to chromosome 8q13. Hum Hered 1999;49(2):75-80.

101. Majander A, Huoponen K, Savontaus ML, Nikoskelainen E, Wikstrom M. Electron transfer properties of NADH:ubiquinone reductase in the ND1/3460 and the ND4/11778 mutations of the Leber hereditary optic neuroretinopathy (LHON). FEBS Lett 1991;292(1-2):289-92.

102. Smith PR, Cooper JM, Govan GG, Harding AE, Schapira AH. Platelet mitochondrial function in Leber's hereditary optic neuropathy. J Neurol Sci 1994;122(1):80-3.

103. Carelli V, Ghelli A, Ratta M, Bacchilega E, Sangiorgi S, Mancini R, et al. Leber's hereditary optic neuropathy: biochemical effect of 11778/ND4 and 3460/ND1 mutations and correlation with the mitochondrial genotype. Neurology 1997;48(6):1623-32.

104. Earley FG, Patel SD, Ragan I, Attardi G. Photolabelling of a mitochondrially encoded subunit of NADH dehydrogenase with [3H]dihydrorotenone. FEBS Lett 1987;219(1):108-12.

105. Friedrich T, Strohdeicher M, Hofhaus G, Preis D, Sahm H, Weiss H. The same domain motif for ubiquinone reduction in mitochondrial or chloroplast NADH dehydrogenase and bacterial glucose dehydrogenase. FEBS Lett 1990;265(1-2):37-40.

106. Cock HR, Cooper JM, Schapira AH. The 14484 ND6 mtDNA mutation in Leber hereditary optic neuropathy does not affect fibroblast complex I activity. Am J Hum Genet 1995;57(6):1501-2.

107. Oostra RJ, Van Galen MJ, Bolhuis PA, Bleeker-Wagemakers EM, Van den Bogert C. The mitochondrial DNA mutation ND6*14,484C associated with leber hereditary optic neuropathy, leads to deficiency of complex I of the respiratory chain. Biochem Biophys Res Commun 1995;215(3):1001-5.

108. Carelli V, Ghelli A, Bucchi L, Montagna P, De Negri A, Leuzzi V, et al. Biochemical features of mtDNA 14484 (ND6/M64V) point mutation associated with Leber's hereditary optic neuropathy. Ann Neurol 1999;45(3):320-8.

109. Mackey D, Howell N. A variant of Leber hereditary optic neuropathy characterized by recovery of vision and by an unusual mitochondrial genetic etiology. Am J Hum Genet 1992;51(6):1218-28.

110. Larsson NG, Andersen O, Holme E, Oldfors A, Wahlstrom J. Leber's hereditary optic neuropathy and complex I deficiency in muscle. Ann Neurol 1991;30(5):701-8.

111. Cortelli P, Montagna P, Avoni P, Sangiorgi S, Bresolin N, Moggio M, et al. Leber's hereditary optic neuropathy: genetic, biochemical, and phosphorus magnetic resonance spectroscopy study in an Italian family. Neurology 1991;41(8):1211-5.

112. Degli Esposti M, Carelli V, Ghelli A, Ratta M, Crimi M, Sangiorgi S, et al. Functional alterations of the mitochondrially encoded ND4 subunit associated with Leber's hereditary optic neuropathy. FEBS Lett 1994;352(3):375-9.

113. Vergani L, Martinuzzi A, Carelli V, Cortelli P, Montagna P, Schievano G, et al. MtDNA mutations associated with Leber's hereditary optic neuropathy: studies on cytoplasmic hybrid (cybrid) cells. Biochem Biophys Res Commun 1995;210(3):880-8.

114. de Grey AD. Reactive oxygen species production in the mitochondrial matrix: implications for the mechanism of mitochondrial mutation accumulation. Rejuvenation Res 2005;8(1):13-7.

115. Dugan LL, Sensi SL, Canzoniero LM, Handran SD, Rothman SM, Lin TS, et al. Mitochondrial production of reactive oxygen species in cortical neurons following exposure to N-methyl-D-aspartate. J Neurosci 1995;15(10):6377-88.

116. Reynolds IJ, Hastings TG. Glutamate induces the production of reactive oxygen species in cultured forebrain neurons following NMDA receptor activation. J Neurosci 1995;15(5 Pt 1):3318-27.

117. White RJ, Reynolds IJ. Mitochondria and Na+/Ca2+ exchange buffer glutamate-induced calcium loads in cultured cortical neurons. J Neurosci 1995;15(2):1318-28.

118. Skulachev VP. Why are mitochondria involved in apoptosis? Permeability transition pores and apoptosis as selective mechanisms to eliminate superoxide-producing mitochondria and cell. FEBS Lett 1996;397(1):7-10.

119. Hillered L, Chan PH. Role of arachidonic acid and other free fatty acids in mitochondrial dysfunction in brain ischemia. J Neurosci Res 1988;20(4):451-6.

120. Qi X, Lewin AS, Hauswirth WW, Guy J. Optic neuropathy induced by reductions in mitochondrial superoxide dismutase. Invest Ophthalmol Vis Sci 2003;44(3):1088-96.

121. Danielson SR, Wong A, Carelli V, Martinuzzi A, Schapira AH, Cortopassi GA. Cells bearing mutations causing Leber's hereditary optic neuropathy are sensitized to Fas-Induced apoptosis. J Biol Chem 2002;277(8):5810-5.

122. Wong A, Cavelier L, Collins-Schramm HE, Seldin MF, McGrogan M, Savontaus ML, et al. Differentiation-specific effects of LHON mutations introduced into neuronal NT2 cells. Hum Mol Genet 2002;11(4):431-8.

123. Rosenblum JS, Gilula NB, Lerner RA. On signal sequence polymorphisms and diseases of distribution. Proc Natl Acad Sci U S A 1996;93(9):4471-3.

124. Shimoda-Matsubayashi S, Matsumine H, Kobayashi T, Nakagawa-Hattori Y, Shimizu Y, Mizuno Y. Structural Dimorphism in the Mitochondrial Targeting Sequence in the Human Manganese Superoxide Dismutase Gene. Biochem Biophys Res Commun 1996;226(2):561-5.

125. Borgstahl GE, Parge HE, Hickey MJ, Johnson MJ, Boissinot M, Hallewell RA, et al. Human mitochondrial manganese superoxide dismutase polymorphic variant Ile58Thr reduces activity by destabilizing the tetrameric interface. Biochemistry 1996;35(14):4287-97.

126. Ghelli A, Zanna C, Porcelli AM, Schapira AH, Martinuzzi A, Carelli V, et al. Leber's hereditary optic neuropathy (LHON) pathogenic mutations induce mitochondrial-dependent apoptotic death in transmitochondrial cells incubated with galactose medium. J Biol Chem 2003;278(6):4145-50.

127. Battisti C, Formichi P, Cardaioli E, Bianchi S, Mangiavacchi P, Tripodi SA, et al. Cell response to oxidative stress induced apoptosis in patients with Leber's hereditary optic neuropathy. J Neurol Neurosurg Psychiatry 2004;75(12):1731-6.

128. Kwon JM, Goate AM. The candidate gene approach. Alcohol Res Health 2000;24(3):164-8.

129. Tabor HK, Risch NJ, Myers RM. Opinion: Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002;3(5):391-7.

130. Daly MJ, Rioux JD. New approaches to gene hunting in IBD. Inflamm Bowel Dis 2004;10(3):312-7.

131. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. Am J Hum Genet 2001;69(5):936-50.

132. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273(5281):1516-7.

133. Risch NJ. Searching for genetic determinants in the new millennium. Nature 2000;405(6788):847-56.

134. Cardon LR, Bell JI. Association study designs for complex diseases. Nat Rev Genet 2001;2(2):91-9.

135. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. Nature 2004;429(6990):446-52.

136. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med 2002;4(2):45-61.

137. Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet 1999;22(2):139-44.

138. Jorde LB. Linkage disequilibrium and the search for complex disease genes. Genome Res 2000;10(10):1435-44.

139. Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function. Genome Res 2002;12(3):436-46.

140. Sunyaev S, Ramensky V, Koch I, Lathe W, 3rd, Kondrashov AS, Bork P. Prediction of deleterious human alleles. Hum Mol Genet 2001;10(6):591-7.

141. Chasman D, Adams RM. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. J Mol Biol 2001;307(2):683-706.

142. Vincent AL. Searching for modifier genes. Clin Experiment Ophthalmol 2003;31(5):374-5.

143. Phasukkijwatana N, Chuenkongkaew WL, Suphavilai R, Suktitipat B, Pingsuthiwong S, Ruangvaravate N, et al. The Unique Characteristics of Thai Leber Hereditary Optic Neuropathy: Analysis of 30 G11778A Pedigrees. J Hum Genet 2006;In press.

144. Hofhaus G, Johns DR, Hurko O, Attardi G, Chomyn A. Respiration and growth defects in transmitochondrial cell lines carrying the 11778 mutation associated with Leber's hereditary optic neuropathy. J Biol Chem 1996;271(22):13155-61.

145. Budde SM, van den Heuvel LP, Smeitink JA. The human complex I NDUFS4 subunit: from gene structure to function and pathology. Mitochondrion 2002;2(1-2):109-15.

146. Papa S. The NDUFS4 nuclear gene of complex I of mitochondria and the cAMP cascade. Biochim Biophys Acta 2002;1555(1-3):147-53.

147. Antonicka H, Ogilvie I, Taivassalo T, Anitori RP, Haller RG, Vissing J, et al. Identification and characterization of a common set of complex I assembly intermediates in mitochondria from patients with complex I deficiency. J Biol Chem 2003;278(44):43081-8.

148. de Sury R, Martinez P, Procaccio V, Lunardi J, Issartel JP. Genomic structure of the human NDUFS8 gene coding for the iron-sulfur TYKY subunit of the mitochondrial NADH:ubiquinone oxidoreductase. Gene 1998;215(1):1-10.

149. van den Heuvel L, Ruitenbeek W, Smeets R, Gelman-Kohan Z, Elpeleg O, Loeffen J, et al. Demonstration of a new pathogenic mutation in human complex I deficiency: a 5-bp duplication in the nuclear gene encoding the 18-kD (AQDQ) subunit. Am J Hum Genet 1998;62(2):262-8.

150. Budde SM, van den Heuvel LP, Janssen AJ, Smeets RJ, Buskens CA, DeMeirleir L, et al. Combined enzymatic complex I and III deficiency associated with mutations in the nuclear encoded NDUFS4 gene. Biochem Biophys Res Commun 2000;275(1):63-8.

151. Petruzzella V, Vergari R, Puzziferri I, Boffoli D, Lamantea E, Zeviani M, et al. A nonsense mutation in the NDUFS4 gene encoding the 18 kDa (AQDQ) subunit of complex I abolishes assembly and activity of the complex in a patient with Leigh-like syndrome. Hum Mol Genet 2001;10(5):529-35.

152. Loeffen J, Smeitink J, Triepels R, Smeets R, Schuelke M, Sengers R, et al. The first nuclear-encoded complex I mutation in a patient with Leigh syndrome. Am J Hum Genet 1998;63(6):1598-608.

153. Hattori N, Yoshino H, Tanaka M, Suzuki H, Mizuno Y. Genotype in the 24-kDa subunit gene (NDUFV2) of mitochondrial complex I and susceptibility to Parkinson disease. Genomics 1998;49(1):52-8.

154. Washizuka S, Kakiuchi C, Mori K, Kunugi H, Tajima O, Akiyama T, et al. Association of mitochondrial complex I subunit gene NDUFV2 at 18p11 with bipolar disorder. Am J Med Genet B Neuropsychiatr Genet 2003;120(1):72-8.

155. Washizuka S, Iwamoto K, Kazuno AA, Kakiuchi C, Mori K, Kametani M, et al. Association of mitochondrial complex I subunit gene NDUFV2 at 18p11 with bipolar disorder in Japanese and the National Institute of Mental Health pedigrees. Biol Psychiatry 2004;56(7):483-9.

156. Washizuka S, Kakiuchi C, Mori K, Tajima O, Akiyama T, Kato T. Expression of mitochondria-related genes in lymphoblastoid cells from patients with bipolar disorder. Bipolar Disord 2005;7(2):146-52.

157. Benit P, Beugnot R, Chretien D, Giurgea I, De Lonlay-Debeney P, Issartel JP, et al. Mutant NDUFV2 subunit of mitochondrial complex I causes early onset hypertrophic cardiomyopathy and encephalopathy. Hum Mutat 2003;21(6):582-6.

158. Papa S, Sardanelli AM, Scacco S, Technikova-Dobrova Z. cAMP-dependent protein kinase and phosphoproteins in mammalian mitochondria. An extension of the cAMP-mediated intracellular signal transduction. FEBS Lett 1999;444(2-3):245-9.

159. Petruzzella V, Papa S. Mutations in human nuclear genes encoding for subunits of mitochondrial respiratory complex I: the NDUFS4 gene. Gene 2002;286(1):149-54.

160. Procaccio V, Depetris D, Soularue P, Mattei MG, Lunardi J, Issartel JP. cDNA sequence and chromosomal localization of the NDUFS8 human gene coding for the 23 kDa subunit of the mitochondrial complex I. Biochim Biophys Acta 1997;1351(1-2):37-41.

161. Dupuis A, Skehel JM, Walker JE. A homologue of a nuclear-coded iron-sulfur protein subunit of bovine mitochondrial complex I is encoded in chloroplast genomes. Biochemistry 1991;30(11):2954-60.

162. Albracht SP, Hedderich R. Learning from hydrogenases: location of a proton pump and of a second FMN in bovine NADH--ubiquinone oxidoreductase (Complex I). FEBS Lett 2000;485(1):1-6.

163. Lescuyer P, Martinez P, Lunardi J. YY1 and Sp1 activate transcription of the human NDUFS8 gene encoding the mitochondrial complex I TYKY subunit. Biochim Biophys Acta 2002;1574(2):164-74.

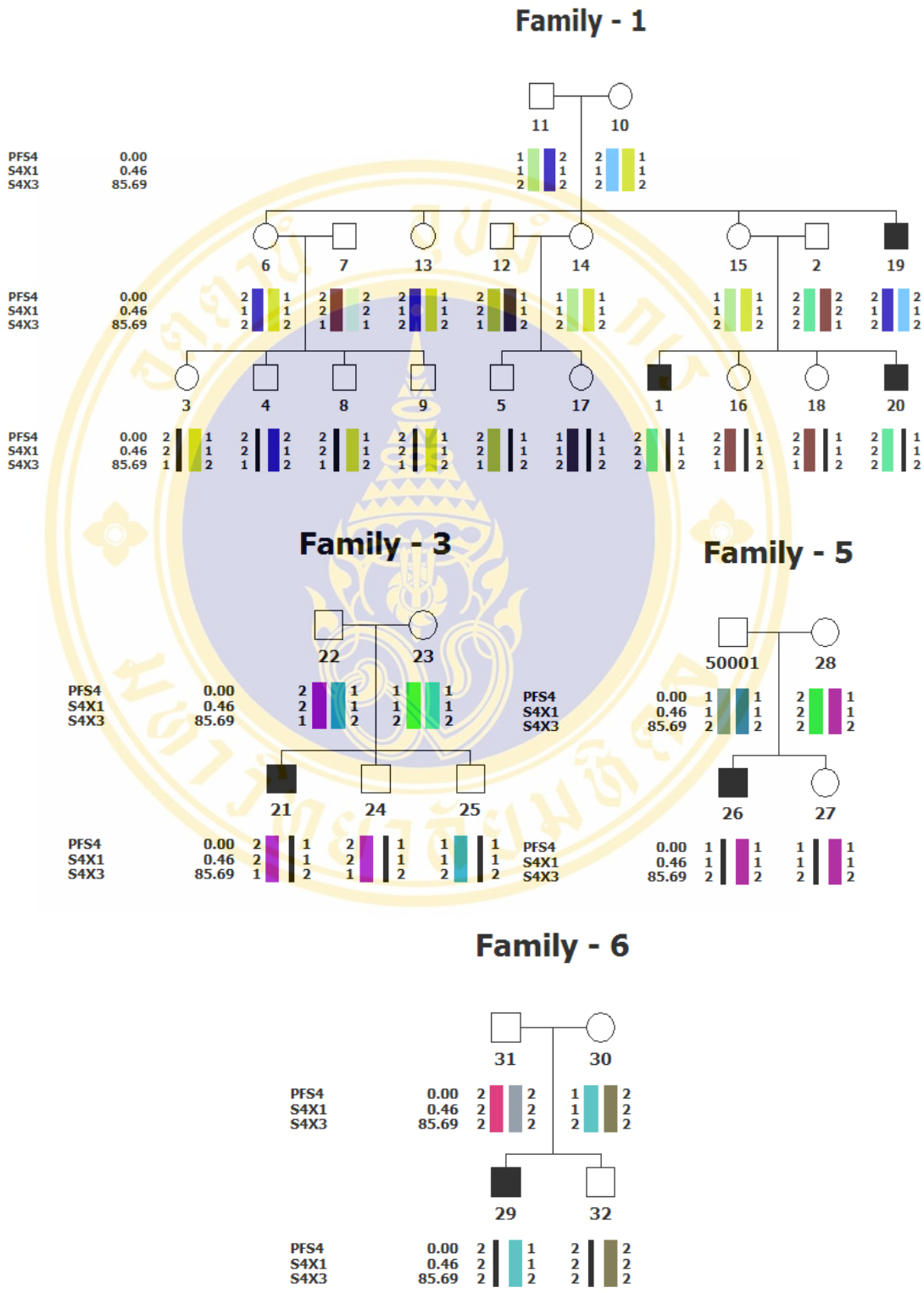
164. Scarpulla RC. Nuclear control of respiratory chain expression in mammalian cells. J Bioenerg Biomembr 1997;29(2):109-19.

165. de Coo R, Buddiger P, Smeets H, Geurts van Kessel A, Morgan-Hughes J, Weghuis DO, et al. Molecular cloning and characterization of the active human mitochondrial NADH:ubiquinone oxidoreductase 24-kDa gene (NDUFV2) and its pseudogene. Genomics 1995;26(3):461-6.

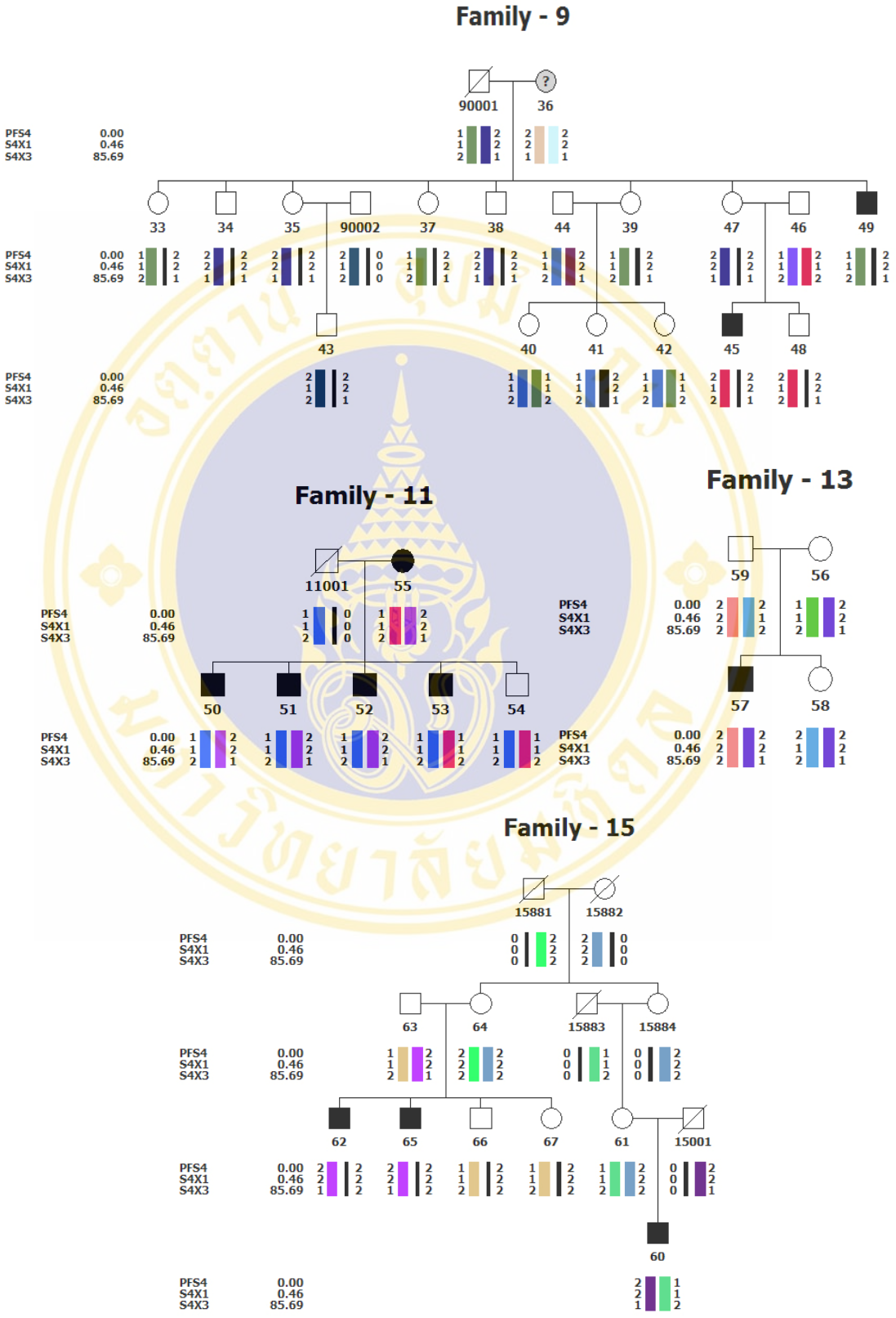
166. Yamaguchi M, Hatefi Y. Mitochondrial NADH:ubiquinone oxidoreductase (complex I): proximity of the subunits of the flavoprotein and the iron-sulfur protein subcomplexes. Biochemistry 1993;32(8):1935-9.

APPENDIX

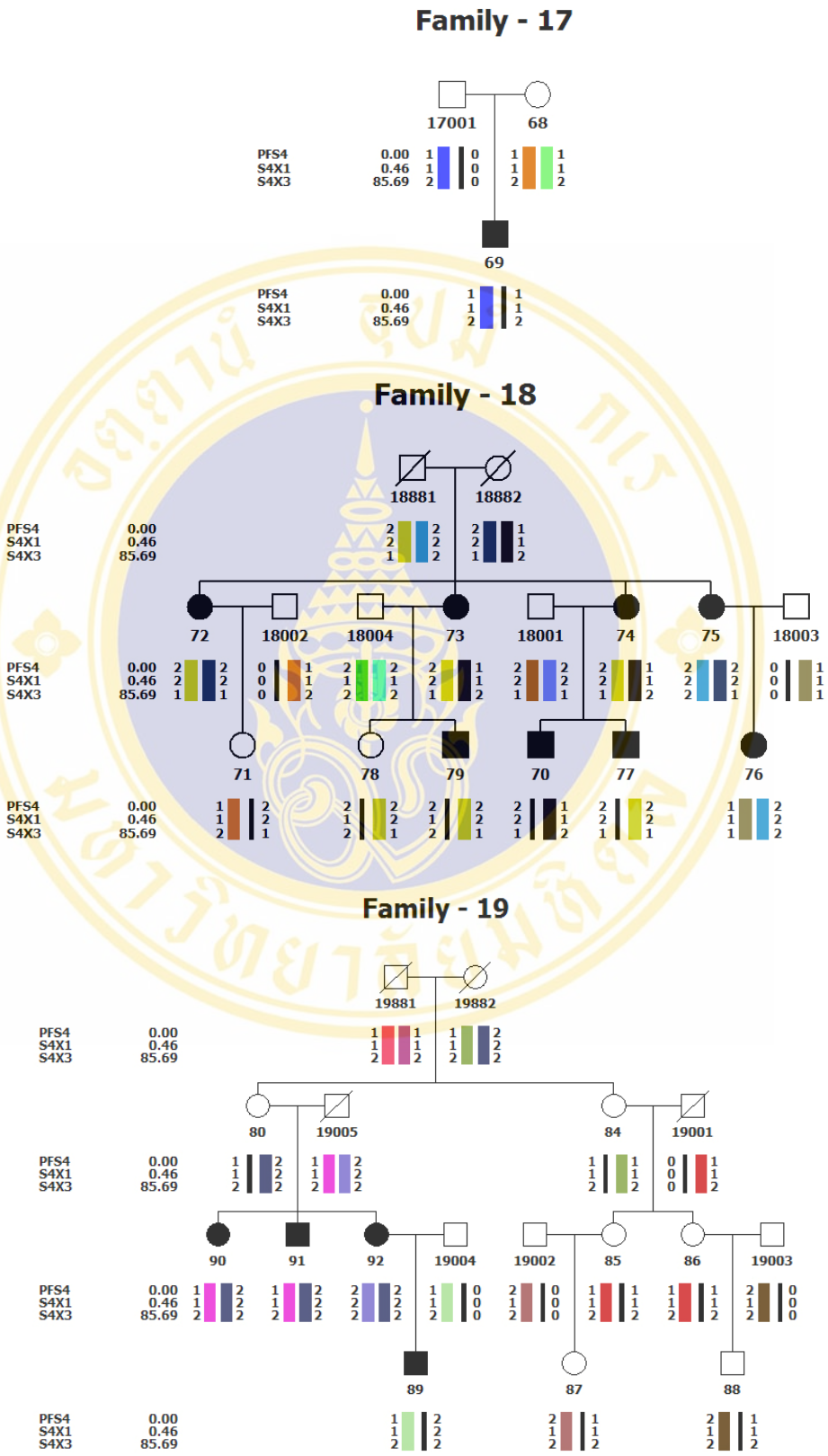
APPENDIX A



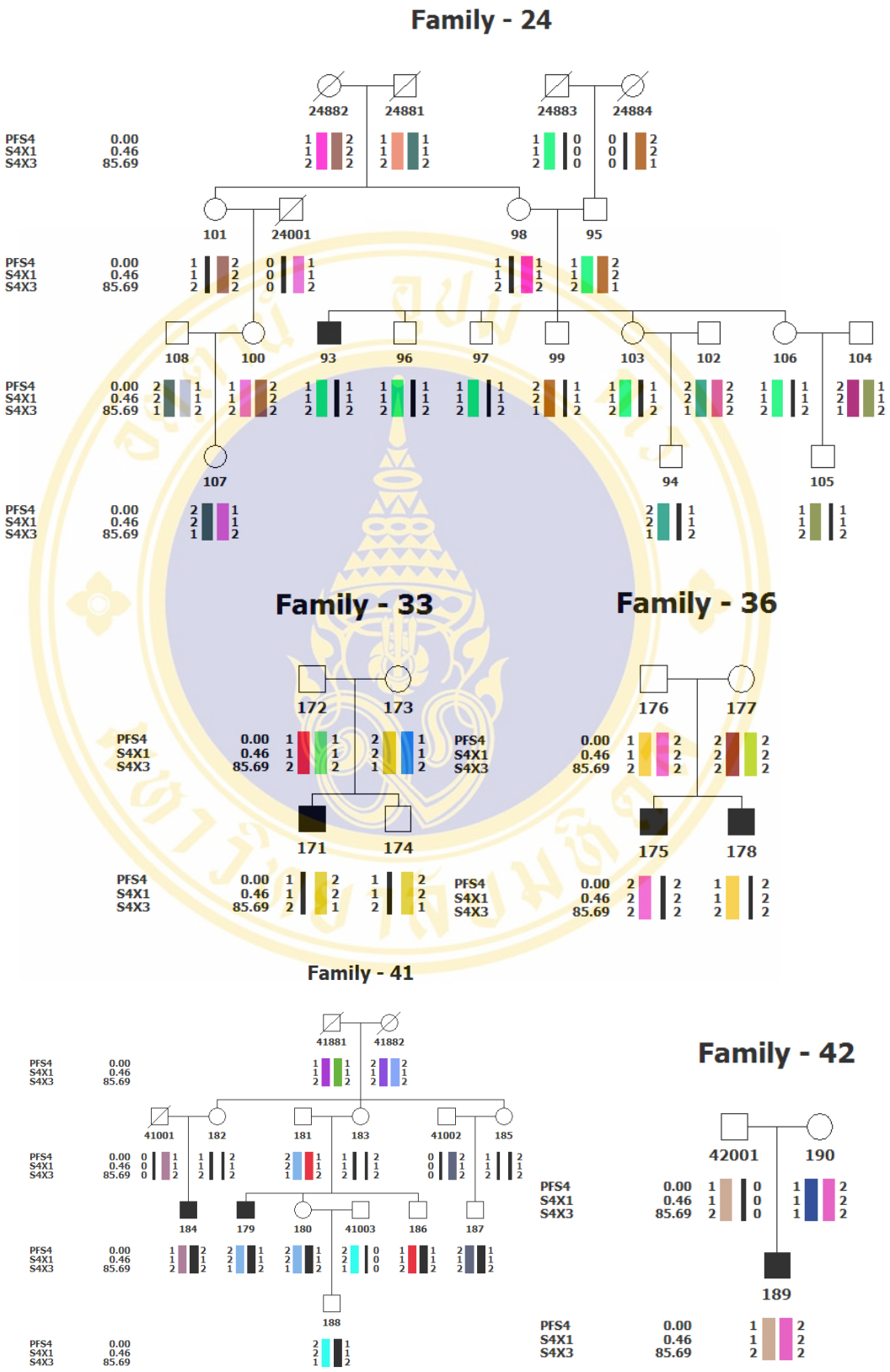
Appendix A. Haplotype estimation for *NDUFS4* gene in LHON family 1, 3, 5 and 6.



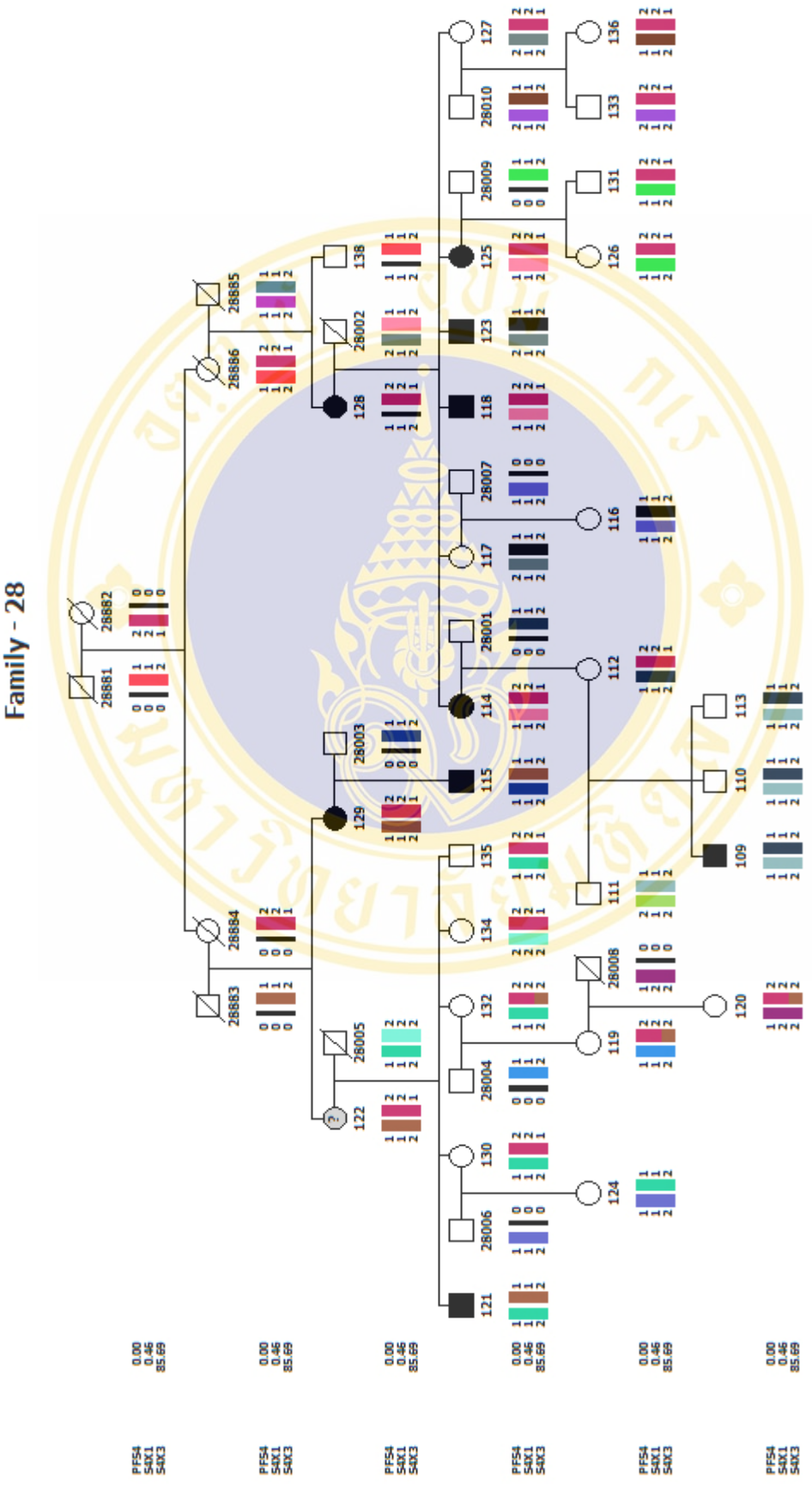
Appendix A (continued). Haplotype estimation for *NDUFS4* gene in LHON family 9, 11, 13 and 15.



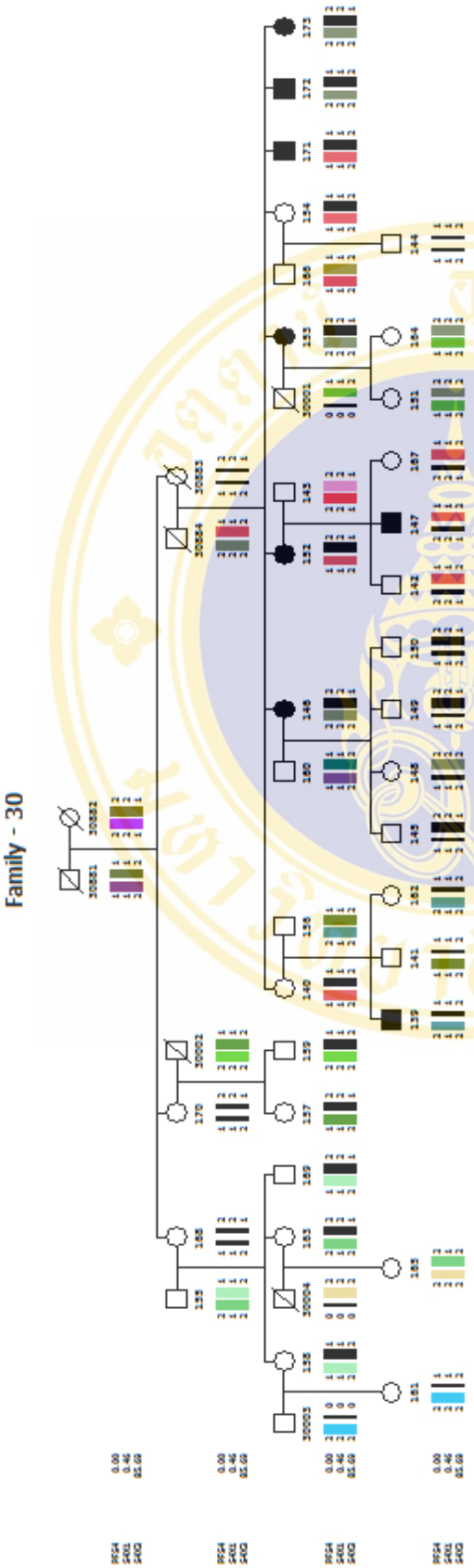
Appendix A (continued). Haplotype estimation for *NDUFS4* gene in LHON family 17, 18 and 19.



Appendix A (continued). Haplotype estimation for *NDUFS4* gene in LHON family 24, 33, 36, 41 and 42.

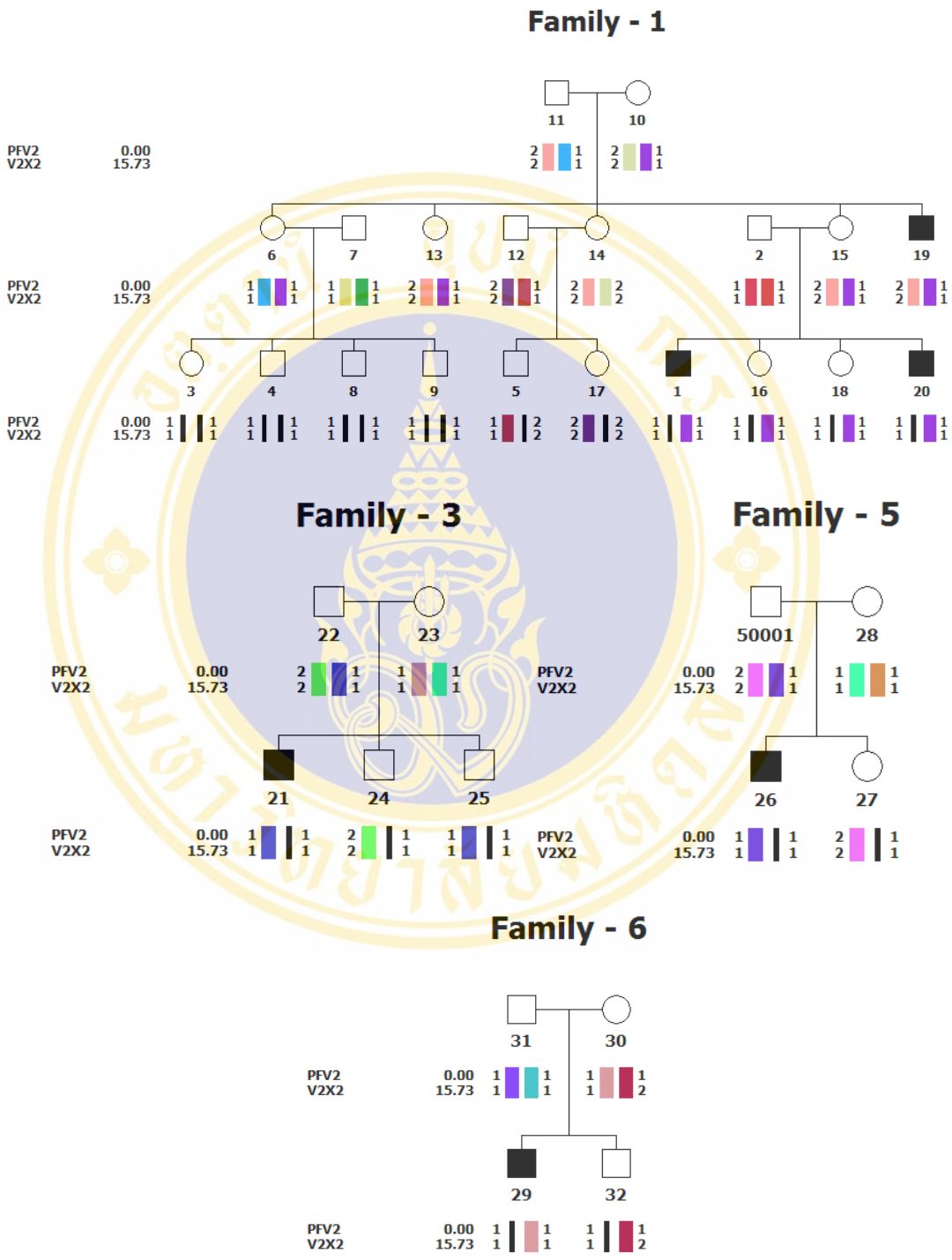


Appendix A (continued). Haplotype estimation for *NDUFS4* gene in LHON family 28.

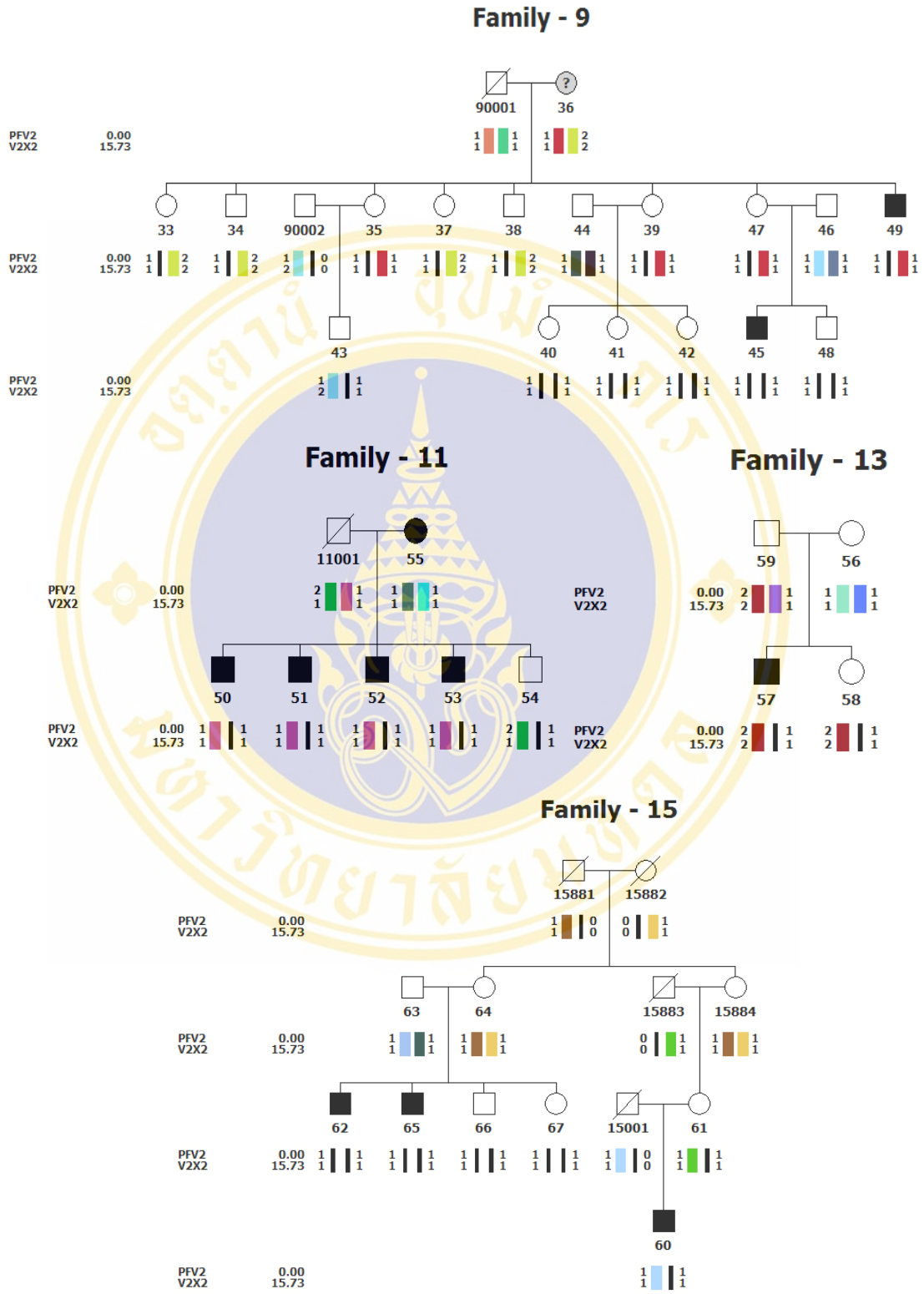


Appendix A (continued). Haplotype estimation for *NDUFS4* gene in LHON family 30.

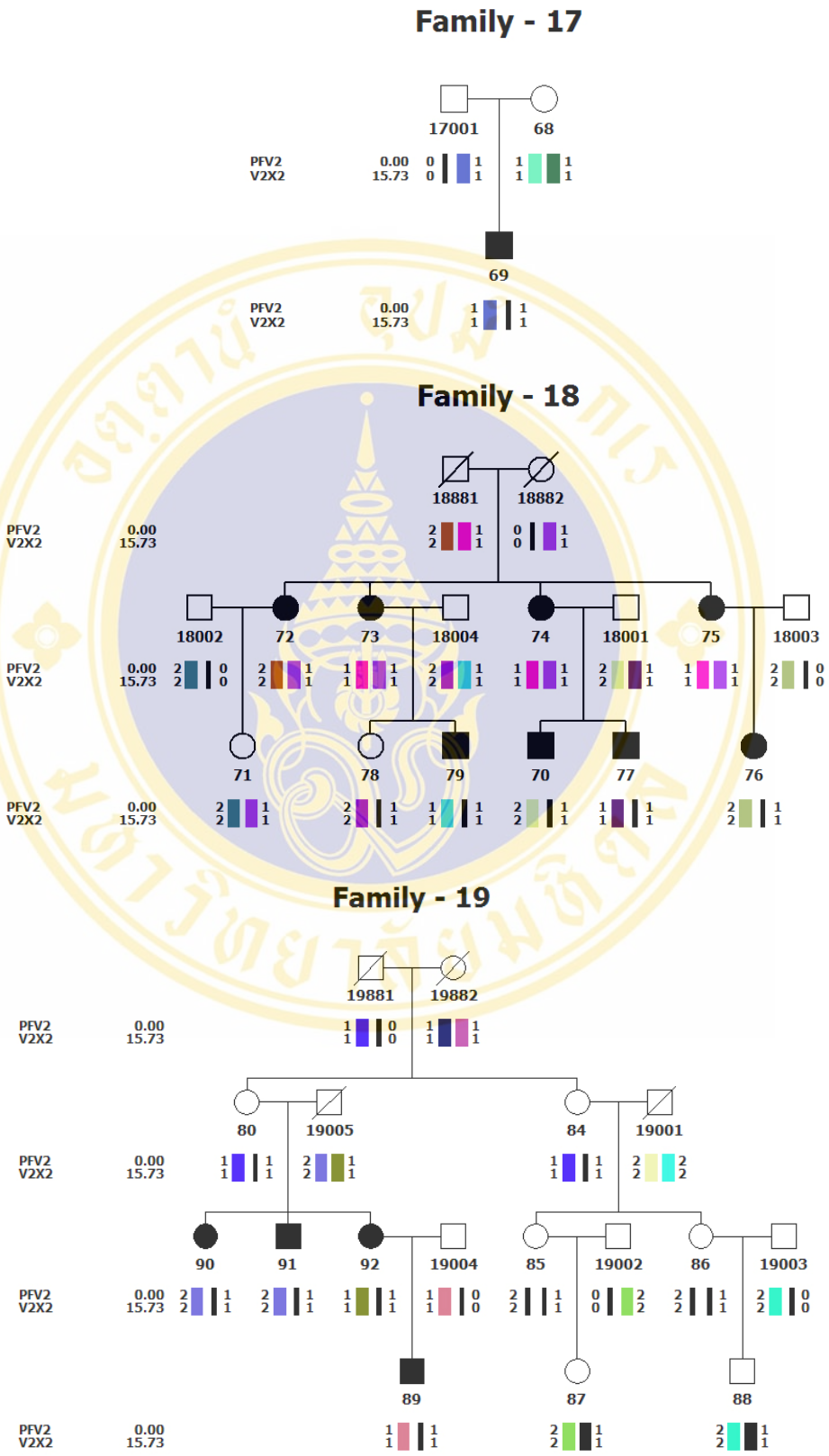
APPENDIX B



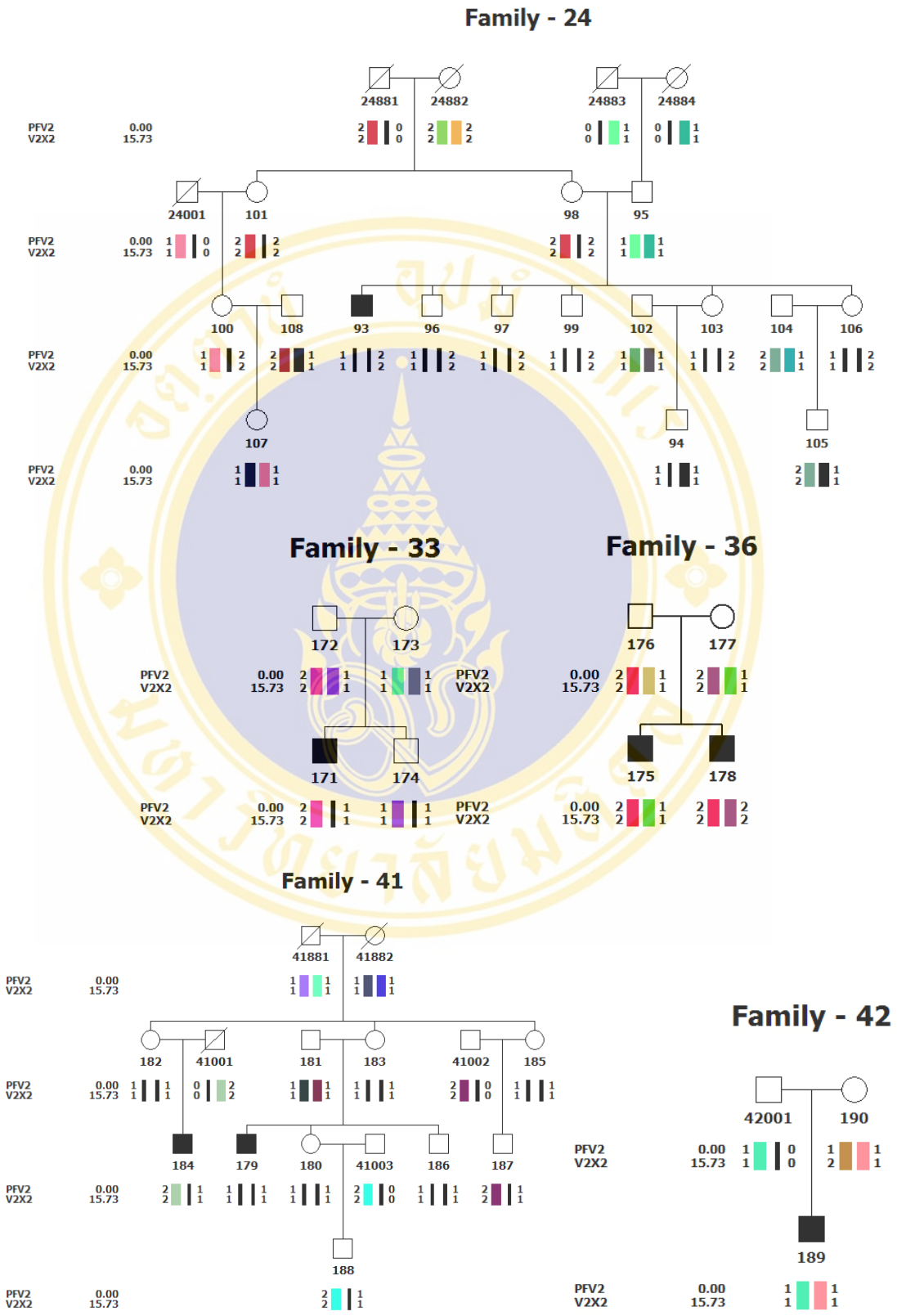
Appendix B. Haplotype estimation for *NDUFV2* gene in LHON family 1, 3, 5 and 6.



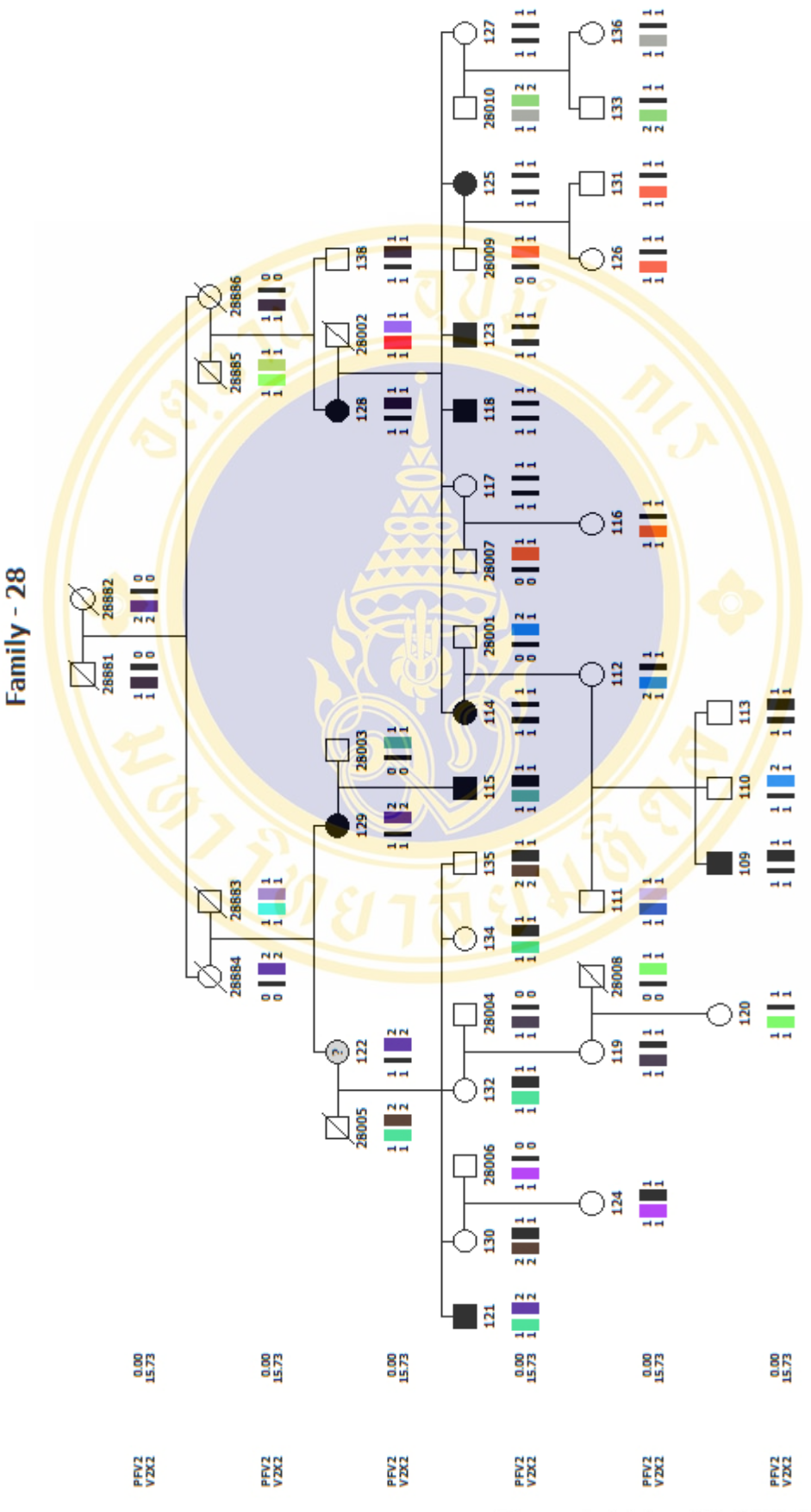
Appendix B (continued). Haplotype estimation for *NDUFV2* gene in LHON family 9, 11, 13 and 15.



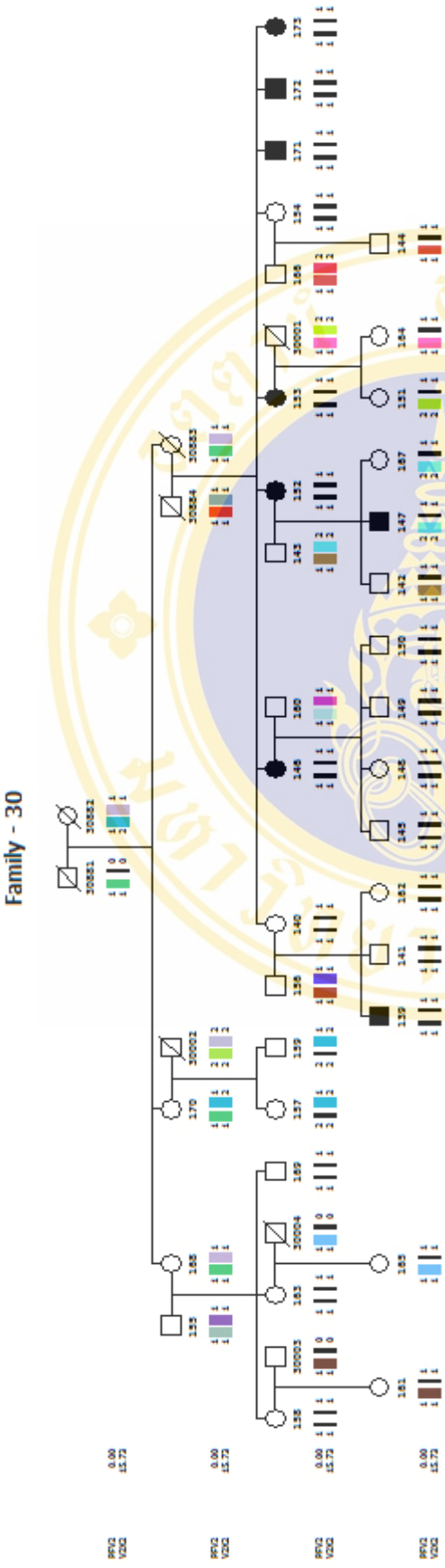
Appendix B (continued). Haplotype estimation for *NDUFV2* gene in LHON family 17, 18 and 19.



Appendix B (continued). Haplotype estimation for *NDUFV2* gene in LHON family 24, 33, 36, 41 and 42.



Appendix B (continued). Haplotype estimation for *NDUFV2* gene in LHON family 28.



Appendix B (continued). Haplotype estimation for *NDUFV2* gene in LHON family 30.

BIOGRAPHY

NAME	Miss.Bussaraporn Kunhapan
DATE OF BIRTH	24 December 1980
PLACE OF BIRTH	Nakhonpathom, Thailand
INSTITUTIONS ATTENDED	Chulalongkorn University, 2001 : Bachelor of Science (Biochemistry) Mahidol University, 2006 : Master of Science (Biochemistry)
RESEARCH GRANT	Supported in part by the Thesis Grant, Faculty of Graduate Studies, Mahidol University Supported in part by Commission on Higher Education and Faculty of Graduate Studies, Mahidol University, in the Academic Year of 2004
HOME ADDRESS	9/154 Moo 8, Petchakasem Rd., Omyai, Sampran, Nakhonpathom, 73160 Thailand E-mail: kunhapan_bp@yahoo.com