MOLECULAR GENETIC INVESTIGATION OF MITOCHONDRIAL DYSFUNCTION IN RELATION TO MIGRAINE SUSCEPTIBILITY

Shani Stuart

BSc Genetics 2009 (University of Pretoria)
BSc Human Genetics Honours 2010 (University of the Witwatersrand)

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Institute of Health and Biomedical Innovation (IHBI)

School of Biomedical Sciences

Faculty of Health

Queensland University of Technology

2015

Keywords

12s rRNA, Bioinformatics, Energy Metabolism, Gene Expression, Genomics, Inheritance, Genotyping, Haplogroups, Heteroplasmy, Human Disease, Ion Torrent, Mitochondria, Migraine Pathophysiology, Migraine Susceptibility, Mitochondrial Dysfunction, Molecular Genetics, Next Generation Sequencing, Nuclear Encoded Mitochondrial Proteins, Oxidative Phosphorylation, Polymerase Chain Reaction, Quality Control, Sanger Sequencing, Sequence Alignment, Signalling Pathways, Variant Calls

Research Outputs

Publications arising during candidature

1. **Stuart, S**., Cox, H. C., Lea, R. A. & Griffiths, L. R. The role of the MTHFR gene in migraine. *Headache: The Journal of Head and Face Pain* **52**, 515-520 (2012).

2. **Stuart, S**. & Griffiths, L. R. A possible role for mitochondrial dysfunction in migraine. *Molecular Genetics and Genomics* **287**, 837-844 (2012).

3. **Stuart, S**. *et al.* Detection of a Novel Mutation in the CACNA1A gene. *Twin Research and Human Genetics* **15**, 120-125 (2012).

4. Maher, B., **Stuart S**. *et al.* Analysis of 3 common polymorphisms in the KCNK18 gene in an Australian Migraine Case-control cohort. *Gene* **528**, 343-346 (2013).

5. Oikari, L. E., **Stuart S**. *et al.* Investigation of lymphotoxin α genetic variants in migraine. *Gene* **512**, 527-531 (2013)

6. **Stuart, S**. *et al.* Investigation of APOE isoforms and the association between APOE E3 and E4 with migraine in the Australian Caucasian population. *NeuroReport* **24**, 499-503 (2013).

7. **Stuart, S**., Maher, B. H., Oikari, L. & Griffiths, L. R. Molecular Genetics of Migraine. *eLS* **DOI**: 10.1002/9780470015902.a0022493 (2013).

8. **Stuart, S**. *et al.* Genetic Variation in Cytokine-Related Genes and Migraine Susceptibility. *Twin Research and Human Genetics* **16**, 1079-1086 (2013).

9. Roos-Araujo, D., **Stuart, S**., Lea, R. A., Haupt, L. M. & Griffiths, L. R. Epigenetics and migraine; complex mitochondrial interactions contributing to disease susceptibility. *Gene* **543**, 1-7 (2014).

10. An S. Tan, James W. Baty, Lan-Feng Dong, Ayenachew Bezawork-Geleta, Berwini Endaya, Jacob Goodwin, Martina Bajzikova, Jaromira Kovarova, Martin Peterka, Bing Yan, Elham Alizadeh Pesdar, Margarita Sobol, Anatolyj Filimonenko, **Shani Stuart**, Magdalena Vondrusova, Katarina Kluckova, Karishma Sachaphibulkij, Jakub Rohlena, Pavel Hozak, Jaroslav Truksa, David Eccles, Larisa Haupt, Lyn Griffiths, Jiri Neuzil, and Michael V. Berridge. Mitochondrial Genome Acquisition Restores Respiratory Function and Tumorigenic Potential of Cancer Cells without Mitochondrial DNA. *Cell Metabolism* **21,** 81-94 (2015).

Under Review

11. **Shani Stuart**, Miles C. Benton, David Eccles, Heidi Sutherland, Larisa M. Haupt, Rod A. Lea and Lyn R. Griffiths. Association of Nuclear Encoded Mitochondrial Protein genes provides evidence implicating the PCDH cluster with migraine susceptibility. *Human Molecular Genetics* (2015).

12. Miles C Benton, **Shani Stuart**, Claire Bellis, Donia Macartney-Coxson, Joanne Curran, Geoff Chambers, John Blangero, Rod A Lea and Lyn R Griffiths. Mutiny on the Bounty': The Genetic History of Norfolk Island reveals extreme gender biased admixture. *Scientific Reports* (2014).

Manuscripts in Progress

13. **Shani Stuart**, David Eccles, Miles C. Benton, Bridget H. Maher, Heidi Sutherland, Larisa M. Haupt, Rod A. Lea and Lyn R. Griffiths. Full Mitochondrial Genome Sequencing Reveals that the 12S rRNA mitochondrial sub-unit is Involved in Migraine Susceptibility in the Genetically Isolated Norfolk Island Population. *Nature Genetics.*

Conference Presentations

Oral

Interactive Oral Poster Presentation at the 2011 GHI Seaworld Conference.

Variation in Mitochondrial DNA shows an Association with Migraine

S Stuart, RA Lea, LM Haupt, LR Griffiths

3 Minute Thesis Competition, GHI 2012

Migraine Mechanisms: Mitochondrial Dysfunction

S Stuart, B Maher, M Benton, D Eccles, H Sutherland, RA Lea, LM Haupt, LR Griffiths

Oral presentation for HGSA Workshop 5 December 2012

The role of mitochondrial dysfunction and migraine susceptibility

S Stuart, RA Lea, LM Haupt, LR Griffiths

Talk for HDR Griffith University Workshop 26th July 2013

The role of mitochondrial dysfunction in relation to migraine susceptibility

S Stuart, RA Lea, LM Haupt, LR Griffiths

Oral presentation at the QUT Systems Biology and Bioinformatics Symposium 26th November 2013

Mitochondrial genome sequencing using a custom ion torrent analysis pipeline

S Stuart, MC Benton, D Eccles, B Maher, H Sutherland, RA Lea, LM Haupt, LR Griffiths

Invited talk Genetics World Tour 2014 Life Technologies, Mariott Hotel Brisbane 20th May 2014

Applications of next generation sequencing sequencing technology: full mitochondrial genome sequencing in a genetically isolated population

S Stuart, M Benton, D Eccles, RA Lea, LM Haupt, LR Griffiths

Oral presentation at IHBI Inspires 2014

Full Mitochondrial Genome Sequencing Reveals that the 12S rRNA Mitochondrial sub-unit is Involved in Migraine Susceptibility in the Genetically Isolated Norfolk Island Population

Stuart, Shani; Eccles, David; Benton, Miles; Maher, Bridget; Sutherland, Heidi; Haupt, Larisa; Lea, Rod; Griffiths, Lyn

Finalist Australian Society for Medical Research (ASMR) Post Graduate Student Award 2015

Molecular Genetic Investigation of Mitochondrial Dysfunction in relation to Migraine Susceptibility

Shani Stuart, Miles C Benton, David A Eccles, Heidi Sutherland, Larisa M Haupt, Rod A Lea, Lyn R Griffiths

Poster

Poster at 2012 ASMR Student Conference

Mitochondrial dysfunction and migraine

S Stuart, RA Lea, LM Haupt, LR Griffiths

Poster at 2012 Gene Mappers

Mitochondrial dysfunction and migraine: Initial clues to find the links

S Stuart, RA Lea, LM Haupt, LR Griffiths

Poster presentation at the Human Genetics Society of Australasia 37th Annual Scientific Meeting 2013

Next generation sequencing: investigating mitochondrial variation in migraine susceptibility through high throughput full mitochondrial genome sequencing

S Stuart, M Benton, B Maher, D Eccles, H Sutherland, R Allcock, RA Lea, LM Haupt, LR Griffiths

Poster presentation at the QUT Systems Biology and Bioinformatics Symposium 25th November 2013

Mitochondrial genome sequencing using a custom ion torrent analysis pipeline

S Stuart, MC Benton, D Eccles, B Maher, H Sutherland, RA Lea, LM Haupt, LR Griffiths

Poster presentation at the Human Genetics Society of Australasia 38th Annual Scientific Meeting 2014

Identifying functional mitochondrial genome variants associated with familial migraine susceptibility

S Stuart, M Benton, B Maher, D Eccles, H Sutherland, RA Lea, LM Haupt, LR Griffiths

Professional Memberships and Involvement

- Human Genetics Society of Australasia (HGSA) student representative from 2012 until present
- Australian Society for Medical Research (ASMR) member

Abstract

Migraine is a common neurological disorder, affecting approximately 12 % of the Australian Caucasian population. It is ranked as one of the top twenty most debilitating disorders in developed countries by the World Health Organisation (WHO). This common neurological disease poses a significant personal and economic burden, with the cost of absenteeism from the work place far exceeding the cost of effective treatment. Current therapies are only effective for a proportion of sufferers and new therapeutic targets are desperately needed to alleviate this burden. Genetic studies aim to identify regions of susceptibility in the human genome and reveal new targets for therapeutic interventions.

Genes involved in neurological, vascular or hormonal pathways have all been implicated in predisposition towards developing migraine. Given the clear gender bias of migraine with a 2:1 ratio of affected females compared to males, X-linked and/or mitochondrial inheritance could be involved. Furthermore, given the role of mitochondria in a number of neurological disorders and in energy production it is possible that mitochondrial variants may play a role in the pathogenesis of this disease. Few variants in the mitochondrial genome have so far been investigated in migraine and this is the first comprehensive molecular genetic study aimed towards investigating the role of mtDNA in this common disorder.

The aim of this study was to investigate the role of mitochondrial dysfunction in relation to migraine susceptibility by using samples and technology available within the Genomic Research Centre (GRC) located within the Institute for Health and biomedical Innovation (IHBI), Queensland University of Technology (QUT). It was the aim of this study to investigate both mitochondrial variants and nuclear encoded variants affecting mitochondrial

function, thereby comprehensively addressing genetic variance which influences mitochondrial dysfunction. The specific aims of this study included investigation of full mitochondrial genome sequences in the genetically isolated Norfolk Island population, investigation of Nuclear Encoded Mitochondrial Proteins (NEMPs) in the Norfolk Island population and replication of significant findings in large Australian Caucasian migraine case-control populations. Techniques utilised to achieve these aims included full mitochondrial genome sequencing using the latest next generation sequencing technologies and genotyping selected variants using the high throughput Sequenom platform. In addition to these main aims, the study aimed to define mitochondrial haplogroups of those individuals that underwent sequencing and to interpret these findings in relation to human migration as well as disease burden.

Norfolk Island is a genetically isolated population situated off the East coast of Australia and is useful for studying complex disease due to the reduced genetic and environmental heterogeneity. Samples from this population were selected from the most related individuals, coalescing back to the original founders for full mitochondrial genome sequencing on the Ion Torrent platform. Initially a pilot project was undertaken to develop a cost effective method for sequencing and to prove the feasibility of the experiment. Following on from this proof of concept, the main project involving 315 Norfolk Island individuals was undertaken.

In the pilot project mt sequence variation was identified in 48 NI individuals and found to differ from the Reconstructed Sapiens Reference Sequence (RSRS) at 296 variable sites. Of these variant sites, 29 variants were common in the 48 NI individuals (>5 %). Many of these common variants are the defining markers of mitochondrial haplogroup B, and its further substructure (haplotypes) such as B4a1a1, to which Polynesians belong. There were 136

singleton variants identified and a further 160 variants were shared by two or more individuals. Variants found in 3 or less samples (less than 5 % of the co-hort) were defined as rare variants and 51.2 % of the variants identified within migraine samples were found to be rare compared to 45.8 % in the control group.

High quality sequence information was generated at a sufficient depth of coverage to detect heteroplasmic variants up to a 10 % threshold based on the calculations that sequencing the 19 kb mitochondrial genome at a minimum of 100 x coverage would be achievable by multiplexing 6 samples on a 314 chip. The manufacturers' specifications (10 Mb per 314 chip) were far exceeded and on average between 250-550 x coverage of the entire mitochondrial genome was obtained for each of the 48 samples. After the initial pilot project 48 samples were plexed on 316 chips in order to achieve high throughput sequencing with maximum coverage obtained in excess of 10 000 x.

Importantly, preliminary data identified 6 novel (undocumented) mtDNA variants in the NI sample. An extended database search of mtDB (34) and Mitomap (35) for these variants returned no hits, thus these positions are deemed to be novel. Interestingly, 5 out of 6 of the novel variants were found in individuals who are migraine sufferers. Also of note**,** the majority of rare variants found only in migraine sufferers are clustered in the genes encoding for the ATPase8 protein and the NADH dehydrogenase subunits, which are core components of the oxidative phosphorylation pathway and essential for the production of ATP. These may well be of functional importance to migraine pathophysiology.

The main part of the study which involved full mitochondrial genome sequencing in n=315 individuals from the Norfolk Island population identified 3 homoplasmic and 11

heteroplasmic variants to be significantly associated with migraine susceptibility. Logistic regression analysis showed that mt 930 G>A located within the 12S rRNA subunit is significantly associated with migraine. A Fisher's exact test identified two rare variants associated with migraine susceptibility. We hypothesize that mt 6480 G>A situated within the *COX1* gene presents genetic evidence that there could be a shared pathogenic mechanism involved in common migraine and stroke. This is the first genetic evidence corresponding to documented biochemical evidence showing a link between *NADH* dehydrogenase and migraine. NADH reductase deficiency has been successfully treated with riboflavin and this presents a new therapeutic avenue. Interestingly based on the available data it appears that heteroplasmic variants which may be acquired during the lifetime may play a more significant role than inherited homoplasmic variants and this idea needs to be explored further.

The three SNPs found to be significantly associated with migraine in the Norfolk Island population were investigated in a replication study as specified by the aims of the project to replicate significant findings. The mt 11930 A>G variant was not detected in any of the Australian outbred samples, suggesting that this novel variant is Norfolk specific. The other two SNPs identified to be significantly associated with migraine in Norfolk Island, mt 930 G>A and mt 6480 G>A were detected in very few case-control samples and were not found to be significantly associated with migraine, again suggesting that these variants are specific to Norfolk Island and do not play an important role in the Australian Caucasian population. However this study has still identified key mitochondrial regions that should be investigated further as other variants in these regions could play a role in migraine pathogenesis in the Australian population. It has also provided conclusive evidence for the first time that mitochondrial variation is linked to migraine susceptibility.

The majority of proteins involved in mitochondrial function are encoded by nuclear DNA and imported into the matrix. In order to comprehensively investigate the role of mitochondrial dysfunction in relation to migraine susceptibility, nuclear encoded mitochondrial protein (NEMP) gene variants needed to be tested for association with migraine susceptibility in addition to mitochondrial encoded variants. In the discovery phase 16820 SNPs in 315 individuals from the large multigenerational pedigree from NI were tested and it was found that 667 NEMP SNPs were significantly associated with migraine. Of these, 21 SNPs significantly associated with migraine with $p<0.0001$ were selected for replication in a large outbred migraine case-control cohort (544 cases, 584 controls). Replication analysis identified 9 SNPs to be significantly associated with migraine located in the genes *ELOVL6* (p=2.55E-05), *SARDH* (p=0.000248), *CSNK1G3* (p=0.007141) and *PCDHG* (p=0.008661).

Interestingly**,** four variants out of the 9 were found to be located within the *PCDHG* cluster, providing further evidence implicating this locus in migraine susceptibility. Variants within the large *PDHG* gene region could alter the way in which connections are established and maintained in the brain, making an individual more susceptible towards developing migraines. An association with multiple variants within this region strengthens the evidence that this gene cluster is a key component in migraine pathogenesis. Haplotype analysis showed some level of LD between these variants and a strong association between haplotypes containing these variants and migraine susceptibility. Given the discovery of this gene cluster being involved in migraine susceptibility in the genetically isolated Norfolk Island population and the very clear replication in a large outbred population, this gene should be investigated further in great detail. Variants playing a key role in MA were also identified in the NEMP genes *ELOVL6*, *SARDH*, and *CSNK1G3*.

In line with the project aims, haplogroups were defined in all individuals that underwent mitochondrial genome sequencing. Mainly Polynesian and European haplogroups were identified, confirming historical accounts which detail maternal Polynesian and paternal European population origins. Of further interest was the discovery that haplogroup K significantly increases an individual's risk of developing migraine. This finding supports the theory of adaptation according to geographical region and temperature zones where specific mitochondrial changes alter an individual's vulnerability to energy reliant processes.

Through full mitochondrial genome sequencing on the Ion Torrent platform we identified 3 homoplasmic and 11 heteroplasmic variants to be significantly associated with migraine susceptibility in the Norfolk Island population. Haplogroup K was also found to be associated with migraine in the Norfolk Island pedigree. We further investigated NEMPs as specified by our aims and found the *PCDHG* gene region to play a particularly important role in migraine pathogenesis. All of the aims set out by this project were achieved and new avenues of research to pursue in future studies were identified. It would be valuable to further investigate full mitochondrial genome information in an Australian outbred Caucasian population, specifically with regard to heteroplasmic variants. It would also be useful to further explore the role of NEMPs in migraine through deep sequencing, genotyping and gene expression studies in case-control populations, the Norfolk Island pedigree and migraine family samples.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Note: Human gene symbols are shown in italics as per convention followed by a brief expansion of the gene name

Abbreviation	Full Definition
ATP1A2	P type Na^+/K^+ ATPase
PGR	progesterone receptor
SCN1A	Na^+ channel α subunit
µL	micro litre
µM	micro molar
ACE	angiotensin converting enzyme
ADARB2	adenosine de-aminase, RNA specific B2
ADP	adenosine di-phosphate
ASTN	astrotactin 1
ATP	adenosine tri-phosphate
ATP1A	ATP synthase subunit A
B2	riboflavin
BAM	binary alignment
Ca	calcium
CACNA1A	P/Q calcium channel
CADASIL	cerebral autosomal dominant arteriopathy
CEU	european
CGRP	calcitonin gene related peptide
CHB	chinese

cM	centi morgan
CMOS	complementary metal-oxide semiconductor
CNS	central nervous system
COVAR	covariate
COX	cytochrome c oxidase
CPEO	progressive external opthalmoplegia
CSD	cortical spreading depression
CSF	cerebral spinal fluid
CSNK1G3	casein kinase I, gamma 3
CT	computerised tomography
CVD	cardiovascular disease
Cytb	cytochrome b
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
ds	double stranded
ELOVL6	fatty acid elongase 6
ESR1	estrogen receptor 1
EtBr	ethidium bromide
ETC	electron transport chain
FAD	flavin adenosine dinucleotide
Fam	family
FHM	familial hemiplegic migraine
FMN	flavin mononucleotide

GABA	glutamate
GP	general practioner
GRC	genomics research centre
GRM7	glutamate receptor metabotropic
GTP	guanosine tri-phosphate
GWAS	genome wide association studies
HDACs	histone de-acetylases
HEPH	hephaestin
HTR7	5-hydroxytryptamine (serotonin) receptor 7
HWE	Hardy Weinberg equilibrium
IAF	inconsistent allele frequency
ID	identification
IGV	integrated genome viewer
INSR	insulin receptor
ISPs	ion sphere particles
kb	kilobase
KCCN3	potassium channel gene
KSS	Kearns-Sayre syndrome
LD	linkage disequilibrium
LD	Leigh's disease
LDLR	low density lipoprotein receptor
LHON	Leber hereditary optic neuropathy
LRP1	low density lipoprotein receptor 1
LS	Leigh syndrome

LVNC	left ventricular non-compaction cardiomyopathy
MA	migraine with aura
MAF	minor allele frequency
MALDI-TOF	matrix assisted desorption/ionisation time-of-flight
MAO	mono amine oxidase
Mb	megabase
MEF2D	myocyte enhancer factor 2D
MELAS	mitochondrial encephalomyopathy, lactic acidosis and stroke like episodes
MERF	myoclonic epilepsy with ragged red fibres
MM	mitochondrial myopathy
MO	migraine without aura
MRI	magnetic resonance imaging
MS	multiple sclerosis
mt	mitochondrial
mtDNA	mitochondrial DNA
MTHFR	methylenetetrahydrofolate reductase
NADH	nicotinamide dinculeotide dehydrogenase
NaOH	sodium hydroxide
NARP	neuropathy ataxia and retinitis pigmentosa
ND1	NADH dehydrogenase subunit 1
NEB	New England Biolabs
NEMP	nuclear encoded mitochondrial protein
NGS	next generation sequencing
NI	Norfolk Island

NO	nitric oxide
OH	hydroxide
OMIM	online Mendelian inherited disorders of man
OXPHOS	oxidative phosphorylation chain
PCDHG	protocadherin gamma subfamily C
PCR	polymerase chain reaction
PCr	phosphocreatine
PED	pedigree
PFO	patent foramen ovale
PGCP	plasma glutamate carboxypeptidase
pH	parts hydrogen
PHACTR1	phosphatase and actin regulator 1
Pi	phosphate
pM	pico moles
PRDM16	PR domain containing 16
QOL	quality of life
QUT	Queensland University of Technology
RFF	ragged red fibres
RFLP	restriction fragment length polymorphism
ROS	reactive oxygen species
RP	retinitis pigmentosa
rRNA	ribosomal ribonucleic acid
Rxn	reaction
SAP	shrimp alkaline phosphatase

SARDH	sarcosine dehydrogenase
SCN1A	sodium channel voltage gated type 1
SLC6A4	solute carrier family 6 member 4
SNHL	sensorineural hearing loss
SNP	single nucleotide polymorphism
T.E	Tris-EDTA
TDF	template dilution factor
TGFBR2	transforming growth factor beta receptor 2
TNC	trigeminal nucleus caudalis
TRESK	potassium channel subfamily K
tRNA	transfer ribonucleic acid
TRPM	transient receptor potential cation channel
V	volts
VSIG4	Vset and immunoglobulin domain containing 4
WHO	World Health Organization

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

QUT Verified Signature

Signature:

Date: 17 June 2015

Acknowledgements

I would like to thank my supervisors Professor Lyn R Griffiths, Dr Larisa M Haupt and Dr Rod A Lea for their invaluable input and support throughout my PhD. Without their many hours of thought and effort I would not have been able to achieve any of the research outcomes presented in this thesis. Special thanks have to go to Lyn for being willing to give me that first chance and a foot in the door which will never be forgotten. I also thank Larisa for being a shoulder to cry on and a pillar of support when either experiments or life was going a bit pear shaped. Her compassion and empathy tided me through many difficult times. Thanks to Rod for helping to make numbers make sense and for navigating the most challenging topic of statistical analysis and bioinformatics. Thank you also for your patience during this difficult learning curve for me.

I would like to thank all the members of the GRC for the help, chats and encouragement that I received during my studies. This has truly been a fantastic group to work in and special mentions go to Heidi Sutherland, Miles Benton, David Eccles, Bishaka Roy, Nesli Avgan, Cassie Albury, Bridget Maher and Rachel Okolicsanyi. Rachel and Cassie have been special friends to me and eternal coffee buddies at all hours to discuss research and all life matters. We spent many late nights in the lab together and I thank them both for the companionship, advice and strength to go on when I felt like giving up. Thanks to both Miles Benton and David Eccles for their invaluable bioinformatics support which made this project possible.

Lastly, but most importantly I would like to thank my family Scot McKenzie, Zephne Stuart and Alistair Stuart for encouraging me and supporting me during this time. Credit for this

work belongs with them. Without their belief in me I would never have set foot out the door. Thank you to my mother who encouraged academic brilliance from the very first moment and never missed a prize giving or Science project. Thank you for helping to mould me into the woman I am today and for supporting my schooling and university with the enthusiasm and dedication that you did. Thank you to my brother Alistair Stuart for being my best friend growing up and for all the visits to Australia which encouraged me greatly to continue with my PhD. I can never thank my partner Scot McKenzie enough and I fear no words will do justice here to the role that he has played during this time. There is no amount of driving, bringing dinners, cleaning the house, listening to me and fiercely defending my side that he has not done. I feel like Scot deserves a PhD in support and that without him I never would have gotten this far. Thank you for the sacrifice on your part which made this project possible and kept me on a semi sane level.

.

Chapter 1: Introduction

The first sections of this chapter outline the background and context of the research, and its purposes. The next section describes the significance and scope of the research. Finally, the last section includes an outline of the remaining chapters of the thesis.

Aims and Significance

Migraine is a common neurological disorder characterised by debilitating head pain and an assortment of additional symptoms which can include nausea, emesis, photophobia, phonophobia and occasionally visual sensory disturbances. Migraine is a complex disease caused by interplay between predisposing genetic variants and environmental factors. Genes involved in neurological, vascular or hormonal pathways have all been implicated to play a role in predisposition towards developing migraine. All of these are nuclear encoded genes, but given the role of mitochondria in a number of neurological disorders and in energy production it is possible that mitochondrial variants may play a role in the pathogenesis of this disease. Few variants in the mitochondrial genome have so far been investigated in migraine and this is the first study to comprehensively asses the molecular genetic role of mitochondrial dysfunction in relation to migraine susceptibility.

This study hypothesises that development of migraine is influenced by mitochondrial dysfunction. The aim of this project was to conduct a complete mitochondrial genome scan to:

- identify the full spectrum of mtDNA variation in a selection of the Norfolk Island pedigree
- determine whether the variants are associated with risk of migraine and
- whether these mtDNA variants or mitochondrial influencing variants from nuclear encoded genes modify migraine risk associated with key environmental factors.

Rationale, Hypotheses and Objectives

This study was aimed at identifying variants in the mitochondrial genome which contribute towards migraine susceptibility. Subsequent to conducting full mitochondrial genome sequencing utilising Next Generation Sequencing (NGS) technology to identify these variants, this project aimed to examine nuclear encoding genes involved in mitochondrial pathways which modify migraine risk. To date a number of variants within the broad categories of neural, vascular and hormonal class genes have been identified to play a role in predisposing individuals towards developing migraine. Given the role of mitochondria in a number of neurological disorders and in energy production, mitochondrial dysfunction may lower the threshold for migraine attacks through insufficient adenosine triphosphate (ATP) production which is critical for correct neural functioning. Mitochondrial dysfunction in relation to migraine has only been previously researched in a limited number of small studies and remains a largely unpursued area.

While a proportion of variants attributing to the heritability of migraine have been accounted for, common variants with large effect sizes have yet to be identified necessitating novel approaches to understand the genetic causes of migraine. The role of rare genetic variants in familial migraine is becoming increasingly apparent given the results of recent Genome-Wide

Association Studies (GWAS) and candidate gene studies that have thus far had limited power to detect variants which play a significant role in common migraine. This suggests that the missing heritability in migraine is in part contributed to by rare variants. Families with multiple affected individuals are likely to be enriched for rare genetic variants that contribute to disease and therefore provide an ideal avenue for their discovery. Genetically isolated populations, such as the Norfolk Island (NI) population are also a valuable resource for discovering variants which contribute towards complex disease susceptibility. Heritability for migraine has been shown to be significantly higher in the Norfolk Island population than in the local outbred Australian Caucasian population, making genetic studies utilising the Norfolk DNA stocks valuable in identification of disease causative variants. The reduction in phenotypic and genetic diversity observed in migraine as a result of geographic isolation reduces the heterogeneity of this complex disorder and increases the likelihood of identifying true susceptibility variants.

This project provided complete mitochondrial genome sequence information for a selection of the Norfolk Island Core pedigree. The Norfolk Core pedigree consists of the most related individuals which coalesce back to the original founders. Ultimately this allowed identification of genetic variants contributing to disease within Norfolk Island and migraine families which will subsequently provide insight into the molecular pathology of migraine. Current therapeutics are only effective for a proportion of sufferers, making the identification of novel therapeutic targets of paramount importance.

Hypothesis

The pathophysiology of migraine is influenced by mitochondrial dysfunction.

Specific Objectives:

a) Obtain entire mitochondrial genome sequence information from selected samples using the Ion Torrent platform.

b) Align the mtDNA sequence information with independent worldwide samples to identify unique variants and determine frequencies of known variants as well as phylogenetic haplogroups.

c) Align sequence information with previously obtained data from the genetically isolated Norfolk Island population and identify regions of similarity.

d) Collect genotype data from nuclear encoded mitochondrial influencing genes.

e) Statistically model the association of mtDNA variants and haplogroups with heritable migraine traits.

f) Statistically model the effects of key environmental factors on mtDNA and ancestry associations.

g) Perform validation studies in independent cohorts.

Study Subjects:

For this study the mtDNA of the Norfolk Island pedigree samples was sequenced and case-control cohorts already available at the Genomics Research Centre (GRC) were used for validation. These subjects were selected on the basis of being affected with migraine and were all diagnosed according to International Headache Society criteria [1]. Medical information and informed consent as well as ethical approval was obtained for all patient samples.

Mitochondrial Genome Sequencing:

The Ion Torrent platform available in the GRC was used to perform full mitochondrial sequencing for all samples. The Sequenom instrument was used for collection of genotype data from nuclear encoding genes which influence mitochondrial function.

Statistical Analysis:

Statistical modelling of the association of mitochondrial variants, ancestry and environmental variables with heritable migraine traits was performed using linkage-based multivariate regression methods. Plink version 1.07 was the main statistical tool used.

Expected Outcomes

This study will contribute to current knowledge about migraine pathogenesis and identify genetic variants which affect the risk of developing migraine. New information identifying novel genomic regions involved in the pathogenesis of this disease may also lead to drug targets and the potential for new treatments.

BACKGROUND

Migraine is a common neurological disorder characterised by severe head pain and an assortment of additional symptoms which can include nausea, photophobia, phonophobia and for some subtypes of migraine additional neurological symptoms. Migraine is classified according to the International Headache Society into two broad categories namely migraine without aura (MO) and migraine with aura (MA) [2, 3]. Most patients suffer from MO, with only 20 % of sufferers experiencing an aura before the onset of a migraine attack. Approximately 12 % of the Caucasian population suffers from this debilitating disease with

almost 2/3 of sufferers being female. Migraine is classified by the World Health Organisation (WHO) as one of the top twenty most debilitating diseases in the developed world and poses a significant personal and economic burden [4].

During a migraine all areas of a patient's life (work, home and leisure) can be considerably disrupted [5]. A number of studies have shown a significant inverse correlation between migraine disability and quality of life (QOL) [6-8]. The disease also incurs substantial economic costs with an early survey by the Australian Bureau of Statistics estimating that the total cost of migraine was ~$721 M pa (ABS, 1990). Recent studies have shown that most of the annual financial burden stems from lost productivity and/or reduced occupational effectiveness [9, 10], however the costs associated with healthcare utilisation are also sizable, with millions spent on medical consultation, hospital beds and treatment [9]. Similar trends have been observed in Europe with reports in 2010 estimating that headache disorders in Europe cost an estimated €43.5 billion per year [11]. These studies concurred that the cost of continuous absenteeism from the work place is actually higher than the direct cost of treatment. It was also found that the total percentage of costs attributed to loss of work place productivity caused by chronic disease is dominated by migraine with 81 % attributed to migraine and only 19 % for other chronic conditions [12]. Current therapies are only effective for a proportion of sufferers and new therapeutic targets are desperately needed to alleviate this burden.

The burden of migraine is further compounded by the high comorbidity of this disorder with other neurological and vascular conditions that are associated with poor health-related outcomes and a decreased QOL. Studies have shown a high prevalence, and increased risk of depression and other psychiatric disorders in migraineurs [13, 14]. Similarly research has shown that women who suffer from the subtype Migraine with Aura (MA) demonstrate a 1.7-

fold increase risk for ischemic stroke [15], an effect that is independent of common risk factors such as smoking and contraceptive use [16, 17]. Although the mechanisms leading to these comorbidities remains unclear, there are several lines of evidence, including the effectiveness of some medications in treating migraine and depression [18], that suggest a bidirectional aetiology [19]. In addition the strong familial occurrence of these disorders supports the proposed hypothesis of a shared genetic basis. Overall the burden of migraine is substantial, encompassing social and economic costs caused by reduced productivity, healthcare utilisation and compounded by high comorbidity with other disorders that similarly cause significant disability and reduced QOL. Therefore investigations into migraine aetiology will provide direct benefits in diagnosis, treatment and management options as well as potentially providing insights into a number of other highly comorbid disorders.

Migraine Diagnosis

Migraine presents with variable clinical phenotypes which can be heterogeneous in the population. In the absence of laboratory based diagnostics, migraine is currently clinically diagnosed based on the International Classification of Headache Disorders 2nd Edition (ICHD-II). The ICHD-II [20]formally classifies migraine into two main subtypes; Migraine with/without Aura (MA and MO, respectively). These subtypes have substantial symptomatic overlap; however MA sufferers also experience distinguishing neurological disturbances that usually precede the headache phase of an attack. Overall, MO and MA sufferers account for ~70 % and 20-30 % of migraineurs, respectively. There are a number of other rare sub-types of migraine that are accompanied by distinctive neurological symptoms, including Familial Hemiplegic Migraine (FHM), in which headache is accompanied by prolonged hemiparesis [20].

Migraine Pathophysiology

There is strong evidence that migraine pathophysiology is caused by neuronal dysfunction. Within this broad theory are subdivisions based on the mechanism which primarily activates the trigeminal vascular system. The earliest ideas pertaining to a neuronal mechanism go back to Leao's work originating in the 1940's. Leao proposed that cortical spreading depression (CSD) activates the trigeminal nerves and is the phenomenon responsible for migraine attacks. Cortical spreading depression is a wave of neuronal and glial depolarization/ neuronal hyperexcitability followed by a long lasting suppression of neural activity [21]. This electrophysiological event has been linked to aura in the human visual cortex and is thought to be partly responsible for the sensory and motor disturbances experienced during MA attacks.

New studies have challenged this theory and propose that the importance of CSD in generating a migraine attack has been overstated and may only account for a proportion of cases. Lambert have put forward the view that migraine pain and trigeminovascular activation are caused by a central mechanism which doesn't require primary sensory input [22, 23]. The most recent theory behind migraine pathogenesis describes migraine as a dysfunction of the subcortical brain structures including the brainstem and diencephalic nuclei which are involved in modulating sensory inputs. The theory suggests that aura is triggered by dysfunction of these nuclei and that the same mechanism is responsible for the pain and other symptoms experienced during migraine attacks [24].

Several lines of evidence exist to suggest that mitochondrial dysfunction may also contribute to the pathogenesis of at least some sub types of migraine. The hypotheses are based on the idea that an impaired mitochondrial oxidative metabolism may contribute to the pathogenesis

of migraine by interrupting proper nervous functioning. Brain and muscle are highly dependent on oxidative metabolism and are therefore the most severely affected tissues in the mitochondrial disorders. A variety of morphological, biochemical, imaging and genetic studies have provided evidence that mitochondrial dysfunction may play a role in migraine susceptibility [25].

Mitochondrial Dysfunction and Migraine

Genes involved in neurological, vascular or hormonal pathways have all been implicated in predisposition towards developing migraine. All of these are nuclear encoded genes, but given the role of mitochondria in a number of neurological disorders and in energy production it is possible that mitochondrial variants may play a role in the pathogenesis of this disease. Mitochondrial DNA has been a useful tool for studying population genetics and human genetic diseases due to the clear inheritance shown through successive generations. Given the clear gender bias found in migraine patients it is of importance to investigate X linked inheritance and mitochondrial related variants in this disorder. Few variants in the mitochondrial genome have so far been investigated in migraine and new studies should be aimed towards investigating the role of mitochondrial DNA in this common disorder [26].

Generating an action potential is a process that requires large amounts of energy as Adenosine tri-phosphate (ATP) is used to restore ion gradients after the generation of synaptic and action potentials. Even though the brain only makes up 2 % of our body mass, it accounts for at least 20 % of our energy expenditure each day [27, 28] therefore an adequate supply of blood glucose and oxygen is essential to sustain neuron function. Based on this information it is possible that an insufficient energy supply caused by mitochondrial dysfunction could predispose individuals to migraine attacks by lowering the cortical spreading depression

threshold. By-products of energy metabolism including adenosine and lactate have been shown to directly induce increased blood flow to the brain [29, 30]. A change in blood flow such as has been documented during CSD is clearly an important factor of migraine attacks. Thus energy metabolism could influence migraine attacks via an accumulation or deficit of by-products.

Other neuropathic diseases which may share some pathways with migraine pathogenesis have been associated with mtDNA variants in a large number of studies. These diseases include encephalomyopathies such as MELAS, MERRF and Kearns-Sayre syndrome [31]. Given the co-morbidity of MA with stroke, the study of genes involved in MELAS are especially promising starting points. Other neurological diseases with mitochondrial associations including major depression, bipolar disorder and schizophrenia have also been shown to be co-morbid with migraine, strengthening the idea of shared pathways and susceptibility variations.

A rare form of migraine, familial hemiplegic migraine type 1 (FHM1) is caused by mutations in the *CACNA1A* gene. Electrophysiological studies have shown that either an increased or decreased influx of Ca^{2+} ions into cells caused by mutations in the *CACNA1A* gene has an effect on the depolarised state of cells [32]. It has been shown that FHM1 mutations can produce gain-of-function Ca (V) 2.1 channels and as a result initiate cortical spreading depression which is the phenomenon thought to underlie migraine aura.

Other functional studies have examined the role calcium ions play in increased blood flow to the brain and have shown that an increased calcium concentration within astrocytes causes vasoconstriction during cortical spreading depression. This process is mediated by a

phospholipase A2 derivative (an arachidonic acid metabolite) [33, 34]. Since mitochondria are involved in calcium ion homeostasis, a process essential for the normal functioning of neurons, then it is logical that an imbalance of calcium ions could lead to an increased susceptibility to migraines. Hence variants found in mitochondrial DNA which affect calcium homeostasis could show association with migraine sufferers.

Examination of muscle tissue has shown a correlation between abnormal mitochondria and migraine sufferers. Ragged red fibres and cytochrome-c-oxidase fibres have been found in the skeletal muscle of some migraine patients [35]. This association was only observed for the more severe cases such as those patients affected by migraine with prolonged aura and FHM patients [36-38]. Ragged red fibres (RRFs) have an abnormal number of sacrcolemmal mitochondria [39], while cytochrome-c-oxidase (COX) negative fibres often have increased fat concentrations [40, 41]. COX negative fibres are found in most patients suffering from mitochondrial encephalomyopathies and are considered to be histologically characteristic of the disease [31, 39, 42, 43].

Examination by electron microscope has revealed clusters of giant mitochondria with paracrystalline inclusions in migraine sufferers [36, 44]. Additional ultra-structural changes of the mitochondria in affected individuals have also been recorded and may reflect alterations caused by impaired oxidative metabolism. With particular reference to ultra-structural changes, accumulation of subsarcolemmal mitochondria have been found in muscle fibres of patients suffering from migraineous stroke (MS) [45].Despite many promising morphological associations, pathogenic mtDNA variations remain to be detected. A more detailed investigation into this area is needed in order to produce significant results. Most

studies have been limited by small number of study participants and examination of larger cohorts in combination with DNA testing may produce significant results.

A number of studies dating as far back as 1973 have shown that there is evidence supporting the hypothesis that a deficit in the muscular and/or cerebral mitochondrial energy metabolism is significantly associated with migraine attacks. Analysis of intermediates in the oxidative phosphorylation and Krebs cycle pathways have demonstrated this relationship. Physiological studies have shown that lactate levels in the cerebral spinal fluid (CSF) are increased compared to normal levels during migraine attacks [44, 46, 47].

Elevated levels of lactate in the CSF is considered to be an indicator of a defective oxidative metabolism as lactic acidosis is indicative of an impaired utilisation of pyruvate in the Krebs cycle [48]. This finding has led to further study of other metabolite intermediates and enzymes involve in the oxidative phosphorylation pathway and also the Krebs cycle. These compounds include pyruvate, monoamine-oxidase (MAO), succinate-dehydrogenase, NADH cytochrome-c-reductase, succinate-cytochrome-c-reductase, NADH-dehydrogenase and citrate synthetise [36, 49, 50]. In each study the findings indicated a correlation between impaired mitochondrial metabolism and migraine attacks. A more recent study further showed a correlation between carnitine deficiency and MA. In these patients treatment by carnitine replacement resulted in lessening of headache severity and frequency [51, 52].

Phosphorus magnetic resonance spectroscopy is an imaging technique which allows researchers to non-invasively investigate brain energy metabolism *in vivo* [53]. It has proven to be a very useful tool in examining altered oxidative phosphorylation metabolism in migraine sufferers. The functionality of mitochondria are assessed by measuring intracellular

levels of phosphocreatine (PCr), inorganic phosphate (Pi) and adenosine diphosphate (ADP). Under anaerobic conditions or when mitochondrial enzymes are near saturation, additional energy is generated by transferring Pi from PCr to ADP to form ATP. This reaction is catalysed by creatine kinase and the efficiency at which it takes place is considered to be a measure of mitochondrial functionality. Thus by calculating the PCr/Pi ratio a measurement of the energy status of the cell is obtained. The lower the ratio, the less energy is available in the cell [54-56]. An additional measure of oxidative metabolism is the V/Vmax ratio. This is indicative of the velocity of oxidative metabolism in relation to its maximum capability. A further method of monitoring metabolism is the measurement of intracellular pH [57]. As mentioned previously lactic acidosis is indicative of inefficient pyruvate metabolism.

A large number of studies investigating mitochondrial metabolism in MA, FHM, MPA, MS and the related disorders CPEO, MELAS, LD, MERRF, LOHN, MS, MM and RP have all identified similar patterns of metabolic abnormalities. In almost all cases a low PCr/Pi ratio indicating low availability of free energy was observed. Increased ADP indicating a lower energy reserve in the brain cells and an increased V/Vmax ratio were also recorded. An increase V/Vmax ratio is indicative of a defective respiratory chain that supplies insufficient energy to meet the cell's demands. This pattern has suggested to be typical of mitochondrial disorders [58-62]. However whether these changes are caused by a primary mitochondrial dysfunction or are merely a side effect of brain hyperexcitability remains unclear [56, 63].

The Genetic Basis of Migraine

For human diseases and other complex traits, heritability can be estimated from the concordance rate between monozygotic and dizygotic twins [64]. More complex models which examine the correlation of offspring and parental phenotypes can be used to estimate

heritability via complex statistical methods. These are employed when phenotypic measures are available on individuals with a mixture of relationships, both within and across multiple generations, or when there are unequal numbers of observations per family [65]. A number of twin studies in migraine have shown that the heritability of migraine ranges from 40-60 % and that the contribution of non-shared environmental factors is 35-55 % [66-68]. Based on this information it is clear that both genetic factors and environment play an approximately equally important role in the pathogenicity of migraine. Previously, much emphasis has been placed on genetic studies and conducting large Genome Wide Association Study (GWAS) experiments, but with the knowledge that environmental factors may directly trigger migraine attacks, more inclusive models which take into account environmental factors are needed [69].

To date, a number of causative genes have been identified in a rare, severe form of migraine – FHM that shows autosomal dominant inheritance. These genes include the *CACNA1A*, *SCN1A* and *ATP1A2* genes that can contain any of numerous identified mutations that independently cause the disorder. This rare subtype exemplifies the significant genetic and allelic heterogeneity in the aetiology of the disorder. Despite this knowledge however, the number and identity of genes involved in the MA and MO subtypes of migraine that are considered in this project, is not clear, suggesting an even more complex aetiology of these subtypes involving a multifactorial interplay of both environmental and genetic factors.

GWAS have had some success in identifying variants that contribute to MA and MO. The *MTDH* gene, involved in glutamate homeostasis, was the first to be implicated in a large two-stage study conducted by the International Headache Genetics Consortium. In subsequent studies the *TRPM8* gene, and the lipoprotein receptor have also been implicated [70].Variants

in these genes all demonstrated genomewide significance and are plausibly integrated into the current understanding of migraine pathophysiology. However the effect size of each variant is small to moderate supporting the proposed effects of rare genetic variants.

Genomewide linkage studies using migraine pedigrees also support a heterogeneous genetic basis for both MA and MO. Numerous linkage studies have been undertaken with over 15 migraine susceptibility loci mapped to date on chromosomes 1-6, 9-11, 14,15,17-19 and the X chromosome. However few of these have been replicated or had causal genes identified. This is possibly a consequence of the current methods for gene identification in linkage regions that have largely consisted of candidate gene association studies using large unrelated population cohorts. This approach is not designed to identify rare variants and generally does not detect allelic heterogeneity within a gene, thus true susceptibility variants are often undetected.

Genetically Isolated Norfolk Island Population

Following a mutiny aboard the British Royal Navy ship HMS Bounty on 28 April 1789, mutineers settled Pitcairn Island in an effort to avoid detection by the British Navy [71]. When the population grew too large for this small island 193 people from Pitcairn Island, all descended from 9 ‘Bounty mutineers’ and 12 Tahitian women, moved to the uninhabited Norfolk Island in 1856 [72]. Due to geographical isolation this population has largely remained a genetically isolated population making it an ideal population for the study of complex multi-factorial diseases including migraine [73].

Extensive family histories have been documented and maintained by Norfolk Island inhabitants and more recently genealogists. A database exists which contains detailed

information about almost the entire Norfolk Island population and in 2005 the pedigree included 6379 individuals comprising 2185 families dating back to the original founders over an 11 generation period [71]. The database is maintained in a popular genealogy program Brother's Keeper (Version 6.0, Rockford, Mich, USA) [74].

Due to the many inbreeding loops found in Norfolk's early generations along with the size and complexity of the pedigree it has necessitated the need to split the larger 6500 individual pedigree into smaller branches. A core pedigree containing the lower generations and consisting of members originating directly from the population founders has been constructed using a peeling algorithm in the pedigree database management system PEDSYS. This pedigree has been used in previous studies examining risk traits for complex disease [75, 76]. The most up to date core pedigree structure includes 1388 individuals [77]. Contained within this pedigree is a large migraine affected family (n=21 individuals, 7 migraine sufferers) which is of particular interest because the family members are expected to have Polynesian mtDNA haplotypes inherited from the original founders and the family spans four generations.

Significance

Migraine is a highly prevalent disorder with clear social and economic burdens and significant negative impact on quality of life. Despite this, the underlying mechanisms that contribute to migraine remain largely unknown resulting in poor management and ineffectual treatment for many sufferers. Epidemiological and genetic studies have clearly demonstrated that migraine has a genetic aetiology however to date few causal genes have been conclusively identified. This is largely attributed to the phenotypic diversity that migraine

displays and a polygenic genetic basis with significant allelic heterogeneity. GWAS have shown limited success in unravelling the genetic basis of this disease, indicating that rare variants and allelic heterogeneity have a stronger influence on migraine. Here we propose a new approach to identify these rare variants. This approach takes advantage of our large migraine pedigree resources and novel NGS technologies that have significantly lowered the costs of sequencing.

This study aimed to develop and use highly cost effective in-house methods to conduct full mitochondrial genome sequencing for the entire Norfolk Island Core pedigree samples. The results of this study will identify new migraine genes and provided new candidates for functional studies to investigate migraine aetiology. The identification of causal variants can also be immediately translated into lab-based diagnostics within our laboratory. The ultimate goal of this research is to translate the outcomes to clinicians and patients.

Thesis Outline

Chapter 2 provides a detailed and critically reviewed account of the current migraine literature followed by experiments and results in Chapters 3, 4, 5 and 6. Parts of Chapter 2 have been published as 4 literature reviews as follows:

Stuart, S., Cox, H. C., Lea, R. A. & Griffiths, L. R. The role of the MTHFR gene in migraine. *Headache: The Journal of Head and Face Pain* **52**, 515-520 (2012).

Stuart, S. & Griffiths, L. R. A possible role for mitochondrial dysfunction in migraine. *Molecular Genetics and Genomics* **287**, 837-844 (2012).

Stuart, S., Maher, B. H., Oikari, L. & Griffiths, L. R. Molecular Genetics of Migraine. *eLS* **DOI**: 10.1002/9780470015902.a0022493 (2013).

Roos-Araujo, D., **Stuart, S**., Lea, R. A., Haupt, L. M. & Griffiths, L. R. Epigenetics and migraine; complex mitochondrial interactions contributing to disease susceptibility. *Gene* **543**, 1-7 (2014).

Chapter 3 outlines initial genotyping experiments carried out in an outbred migraine case-control population investigating mitochondrial variants in relation to migraine susceptibility. This initial work was critical in terms of molecular genetic training, providing the basic knowledge and skills to continue with more complex aims and remains unpublished.

In Chapter 4 the optimisations carried out to develop a cost effective method for full mitochondrial genome sequencing are described as well as the initial pilot project undertaken in n=48 samples.

Chapter 5 describes the largest portion of the project where full mitochondrial genome sequencing was undertaken on n=315 individuals comprising the Norfolk Island core pedigree. It also details analysis undertaken so far and draws on conclusions based on these findings. Chapter 5 has been written as a manuscript in preparation and therefore some portions such as the introduction contain repetitive elements.

In Chapter 6 one of the final project aims is addressed by assessing the role of Nuclear Encoded Mitochondrial Proteins (NEMPs) in migraine susceptibility. This chapter describes the methods and results using a Sequenom mass array study approach to investigate mitochondrial findings in relation to migraine susceptibility. This chapter has also been set out as a manuscript in preparation and contains some repetitive elements.

Due to the nature in which this thesis has been set out with an introduction, literature review, preliminary findings and two results chapters formatted as papers in preparation, some repetitive elements are found throughout the thesis. The final chapter provides an overview of planned future research leading on from this project and recommends a number of further avenues to pursue.

Chapter 2: Literature Review

Migraine

Migraine is currently classified by the World Health Organisation (WHO) as being one of the top twenty most debilitating diseases [78]. It is a common neurological disorder characterised by severe head pain and an assortment of additional symptoms which can include nausea, emesis, photophobia, phonophobia and occasionally visual sensory disturbances [79]. Migraine is a complex disease caused by a combination of predisposing genetic variants and environmental factors. It is hypothesised that variants of small to medium effect occur in a combination which when exposed to particular environmental triggers bring about migraine attacks [80].

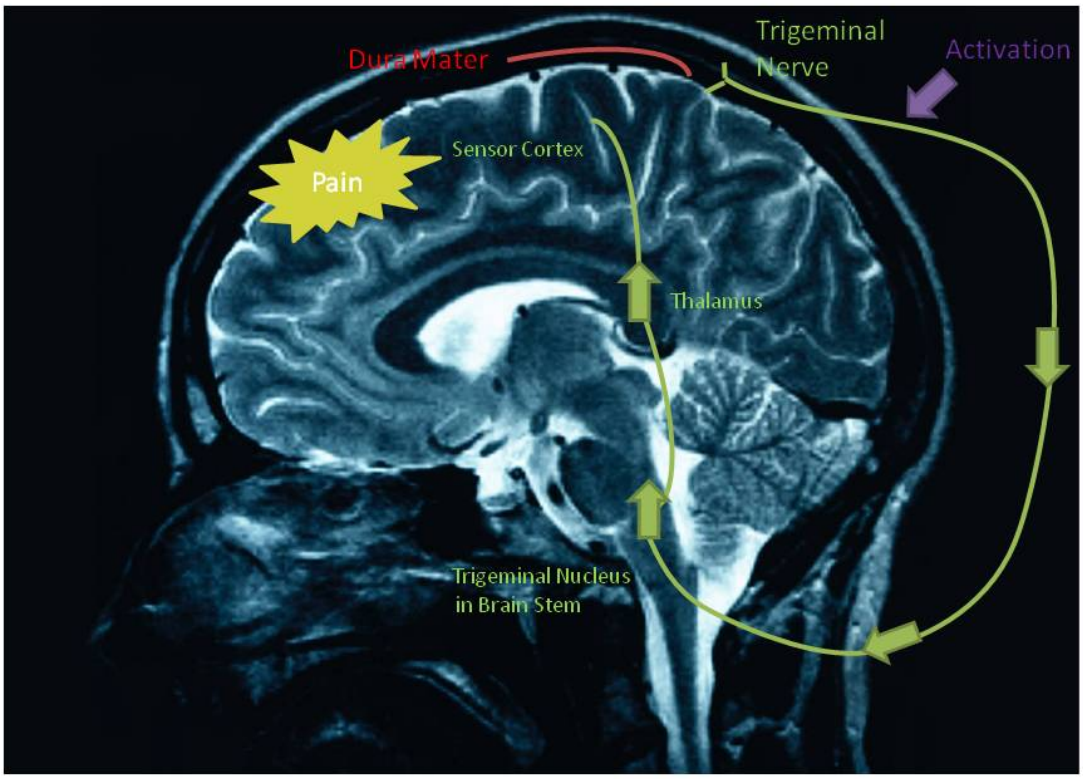
This disease affects approximately 12 % of the Caucasian population, with two thirds of affected cases being female [81]. Migraine has been shown through twin and family studies to have a significantly heritable component, which has driven genetic studies aimed at identifying causal variants [82]. Specific causal mutations have been identified in both rare familial forms of migraine such as familial hemiplegic migraine (FHM) as well as common types of migraine. However the inheritance patterns of rare familial migraine types are more clearly defined than for common types of migraine, making it easier to identify causal variants [83].

The underlying pathophysiology of migraine is still poorly understood. Current ideas are based on the theory that activation of the trigeminal nerve system by a neural, vascular or neurovascular trigger leads to a migraine. The trigeminal nerves carry pain signals from the

meninges and blood vessels infusing the meninges to the trigeminal nucleus in the brain stem which in turn sends signals to the sensor cortex via the thalamus. The sensor cortex processes pain signals and other senses, thus leading to the sensation of pain experienced during migraine attacks. This concept is illustrated in the Figure 1 below.

The exact processes which activate the trigeminal nerve is still one that is being debated, but it is thought that neuronal over excitability and/or neuronal dysfunction caused by various physiological malfunctions play an important role [84]. Ion channels regulate the release of essential ions during the generation of an action potential and therefore any dysfunction of genes involved in this process could lead to the propagation of a migraine attack. This has led many researchers to study FHM and other ion channelopathies in the hopes of identifying common causal variants involved in migraine pathogenesis. FHM is discussed in further detail in section 1.2.

Figure 1: Upon activation of the trigeminal nerve, pain signals are carried to the trigeminal nucleus and then to the sensor cortex via the thalamus



Neural, vascular or chemical triggers can activate the trigeminal nerve. The trigeminal nerve innervates the meninges and activation results in a cascade effect with pain signals being carried to the sensor cortex where they are interpreted as pain. This process is thought to cause the throbbing pain associated with a migraine attack.

Migraine Mechanisms

The neuronal hypothesis is based on Leao's ideas originating as early as the 1940's. Leao proposed that cortical spreading depression (CSD) activates the trigeminal nerves and is thereby the phenomenon responsible for migraine attacks [85]. Cortical spreading depression is a wave of neuronal and glial depolarization/ neuronal hyperexcitability followed by a long lasting suppression of neural activity [21]. This electrophysiological event has been linked to

aura in the human visual cortex and is thought to be partly responsible for the sensory and motor disturbances experienced during MA attacks. Even though only 20 % of migraine attacks include an aura, a role for CSD has been suggested by the neurovascular hypothesis in the pathogenesis of MO as well.

The neurovascular theory is based on the concept that both vascular and neural stimuli leading to CSD are responsible for migraine attacks. The hypothesis is that meningeal inflammation occurs as a consequence of CSD and local mediators such as calcitonin, gene-related peptide (CGRP) and substance P activate meningeal sensory neurons. It is hypothesised that these neuropeptides rather than vasodilation are directly responsible for activation of the pain pathway [86]. This theory contradicts the vascular theory upheld by Wolf, which states that the pain sensation is caused directly by the vasodilation of intracranial and extra cranial vessels [87]. Furthermore, according to the neurovascular hypothesis neuro-inflammation, resulting from the release of inflammatory mediators such as cytokines and mast cells following CSD, may further promote and sustain the activation and sensitisation of meningeal nociceptors, inducing the persistent throbbing headache characterised in migraine.

New studies have challenged this theory and propose that the importance of CSD in generating a migraine attack has been overstated and may only account for a proportion of cases. Recently it has been suggested that migraine pain and trigeminovascular activation are caused by a central mechanism which doesn’t require a primary sensory input [22, 23]. The most recent theory behind migraine pathogenesis describes migraine as a dysfunction of the subcortical brain structures including the brainstem and di-encephalic nuclei which are involved in modulating sensory inputs. The theory suggests that aura is triggered by

dysfunction of these nuclei and that the same mechanism is responsible for the pain and other symptoms experienced during migraine attacks [24].

Regardless of which theory is a more accurate description of migraine initiation, it is well known that the brain needs a continuous and sufficient supply of energy in the form of ATP in order to function efficiently. It is therefore highly feasible that any interruption in energy production could result if neuronal dysfunction and lower the threshold for initiation of a migraine attack. The majority of the body's energy supply is produced via the oxidative phosphorylation pathway or the electron transport chain which is contained within the mitochondria [88]. Any disruption in mitochondrial function which adversely affects the production of energy is likely to affect the tissues which are most heavily reliant on a sufficient source of energy and in some cases cause a pathological state. To date very few migraine studies have examined the association between mitochondrial variation and migraine susceptibility, making this an area which sorely needs to be addressed.

While FHM studies and large migraine GWAS studies have significantly contributed towards our understanding of the possible mechanisms involved in migraine pathogenesis, all of the causal variants and how they interact with each other still remain to be fully elucidated. A large number of loci have been shown to have significant association with migraine and a number of genes/pathways are also known to play a role [81]. Continuous research in this complex arena is needed to improve our understanding further, aid in more accurate diagnosis and contribute to the development of ever improving therapeutics.

Familial Hemiplegic Migraine

Familial hemiplegic migraine (FHM; OMIM 141500) is a rare hereditary subtype of migraine with aura. It is typically accompanied by hemiparesis and a variety of neurological symptoms which may include both visual and motor sensory disturbance [89, 90]. The prevalence of this disease has been estimated to be around 0.01 % in European populations, with familial FHM accounting for half the incidence rate. A sex ratio of three affected females to every one male has also been recorded, hence women are 3 times more likely to suffer from FHM than their male counterparts [91]. FHM is autosomally dominantly inherited but shows variable expressivity and genetic heterogeneity with 70 % to 90 % penetrance [92]. In most cases the phenotype is severe as a result of neural disturbances and in rare cases can be fatal after minor head traumas [93, 94]. Due to the strong genetic component of FHM it has become a favourable target for the study of genes involved in migraine pathogenesis [95].

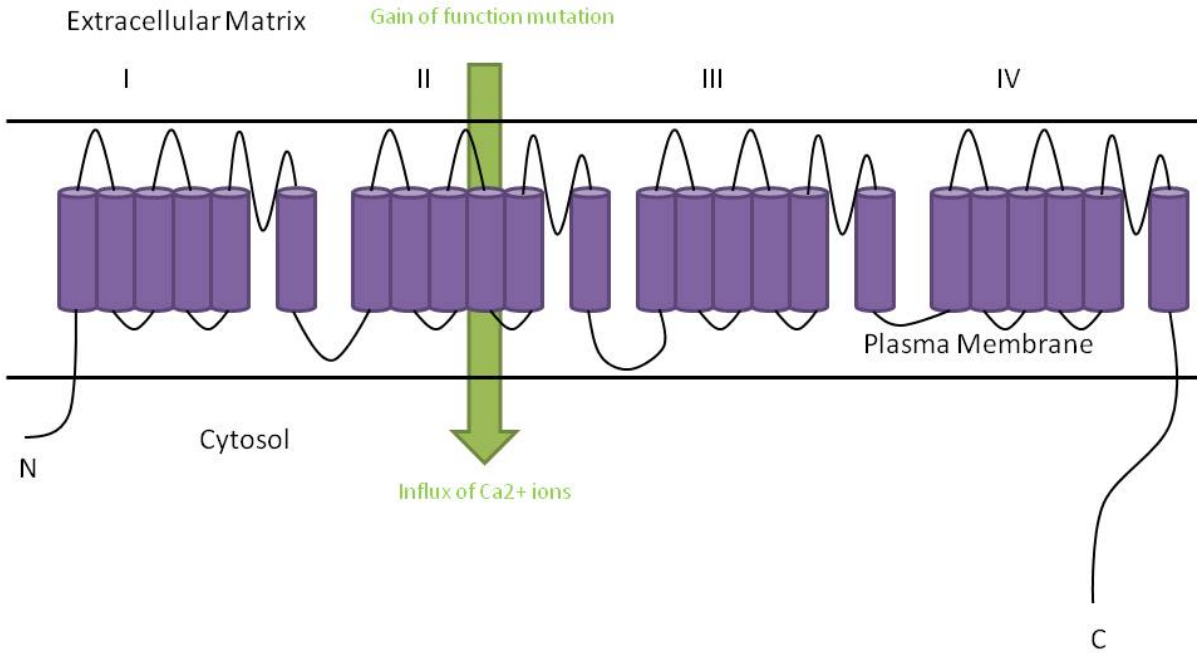
Three main genes have been implicated in the pathogenesis of FHM namely *CACNA1A* (P/Q calcium channel), *ATP1A2* (P type Na^+/K^+ ATPase) and *SCN1A* (Na^+ channel α subunit), with a possibility of a fourth locus at 14q32 [96]. Mutations in the calcium channel gene *CACNA1A* located on chromosome 19p13 have been linked to FHM1 in a number of family studies [97]. Currently 21 different missense mutations which cause FHM1 have been identified in the *CACNA1A* gene as summarised in Table 1 below [98]. It is estimated that *CACNA1A* mutations account for 50 % of FHM patients and that it is this gene that is responsible for the majority of FHM cases [32]. *CACNA1A* encodes for the alpha 1A subunit of the neuronal voltage dependant P/Q-type calcium channel.

Table 1: Summary of known non-pathogenic polymorphisms in CACNA1A

Exon	DNA change	Amino Acid Change	Frequency	rs Number	Study First Reported In
1	c.1-53_1-54insGGC	No	0.04	-	[99]
4	c.579G>A	No	0.02	rs41276894	[100, 101]
6	c.876A>G	No	0.07	rs16006	[101, 102]
9	c.1199-31A>G	No	0.6	rs16008	[103]
11	c.1360G>A	No	0.02	rs41276886	[101, 104]
16	c.2094G>A	No	0.12	Rs16016	[101, 102]
19	c.2978A>T	No	0.62	Rs16023	[103]
19	c.3060G>A	No	0.62	Rs16025	[103]
20	c.3313G>A	No	0.25	Rs16027	[102]
20	c.3549C>T	No	0.02	Rs16029	[103]
39	c.5742C>T	No	0.006	Rs16044	[103]

Ion channel genes play a critical role in normal functioning of the central nervous system and neuromuscular pathways. Functional studies have shown that FHM mutations can lead to both gain- and loss-of function of P/Q-type calcium channels, all of which affect the physiological functioning of the channels in a variety of ways [83] Electrophysiological studies have shown that either an increased or decreased influx of Ca^{2+} ions into cells caused by mutations in the *CACNA1A* gene has an effect on the depolarised state of cells as shown in Figure 2 [32]. It has been shown that FHM1 mutations can produce gain-of-function Ca (V) 2.1 channels and as a result initiate cortical spreading depression which is the phenomenon thought to underlie migraine aura. The increased activity of the Ca(V) 2.1 channel facilitates increased Ca(V) 2.1 dependant neurotransmitter release from cortical neurons, in particular glutamate [105]. A large amount of phenotypic heterogeneity of clinical symptoms is still observed which is indicative of the complexity and variability of FHM1 [32].

Figure 2: The transmembrane topology of the Cav2.1-α1 protein with its four membrane complexes is illustrated in this figure. Gain of function mutations within the CACNA1A transmembrane protein can cause an increased influx of calcium ions resulting in neurotransmitter release. Adapted from [106]



Gain of function mutations can result in an increased uptake of calcium ions as shown in the green arrow. The increased flow of calcium into the cytosol results in neurotransmitter release and neuronal activation. Purple bars indicate the four sub-complexes of the CACNA1A protein, situated within the phospholipid membrane of cells.

Current diagnostic protocol for FHM involves analysis of patient information by a clinician according to criteria specified by the Headache Classification Subcommittee of the International Headache Society [89]. Suspected cases are then referred for molecular diagnosis by mutation analysis. Treatments such as administration of Triptans or β-blockers are commonly used approaches to treat common migraine and can also be used to treat FHM

sufferers but have limited success depending on the individual [107, 108]. Other drugs used to treat hemiplegic migraines include verapamil, acetazolamide, ibuprofen and calcium channel blockers. Most of these treatments were originally designed to target different disorders like epilepsy, high blood pressure, cardiac arrhythmia and glaucoma. Any medications which constrict arteries need to be avoided as this increases the risk of stroke. FHM and MA sufferers are already at an elevated risk for stroke and drugs targeting high blood pressure can aggravate this predisposition [109]. For the most part treatments are still predominantly focused on alleviating symptoms and the development of new therapeutics targeting specific pathways, especially those involving FHM are needed.

Common forms of Migraine

Rare forms of migraine such as familial hemiplegic migraine (FHM) have better understood genetic causes with well-defined causal variants [96, 97, 110] as previously discussed. On the contrary underlying genetic mechanisms for common forms of migraine are still poorly understood. Both MA and MO exhibit genetic heterogeneity with large phenotypic variance, confounding effects to identify causal variants resulting in increased susceptibility towards developing common forms of migraine.

To try and investigate the most plausible candidate genes involved in common migraine pathogenesis, criteria utilising a combination of physiological functionality in conjunction with regions of genomic association are used. Thus far three different groups of genes have been identified and investigated on this basis. These are genes involved in neurological, vascular or hormonal pathways [111, 112]. Under the broad category of neurological genes include those involved in expressing or control of ion channels (*CACNA1A, KCCN3, K Na-*

ATPase), the synthesis and/or release and binding of neuropeptides (*CGRP* protein, glutamate) and the dopamine and serotonergic pathways.

The second category of genes includes those involved in blood pressure regulation, expression of endothelial cells, vasoconstriction and vasodilation (*MTHFR, ACE*) [82]. Many of the vascular type genes associated with migraine also exhibit an overlap with genes playing a role in elevated risk of stroke and heart disease. This is of particular relevance to migraine with aura which has been shown to have a significant co-morbidity with risk for stroke and depression. The third sets of genes seem to be most relevant to females which in some part explains the sex biased distribution of affected patients. Genes associated with females and menstrual migraine include those governing oestrogen and progesterone and fall into the bracket of hormonal pathways (*ESR1, PGR*) [113]. The increased preponderance of female suffers also suggests an X-linked or mitochondrial mode of inheritance.

Neurological Genes

Given the important role that the trigeminovascular system in conjunction with neuropeptides is thought to play during a migraine attack, genes encoding for neurological functions have been extensively studied. Signalling systems involving serotonin, glutamate, dopamine and GABA have all been implicated in migraine onset and genes encoding for products involved in these pathways as well as other ion channels expressed in the CNS have been the centre of many migraine studies [112].

Serotonin (5-HT) is a neurotransmitter and a number of genes within the serotonergic system have been studied for associations with migraine susceptibility. These include the serotonin transporter gene *SLC6A4* as well as tryptophan hydroxylase, monoamine oxidase and

calcitonin-gene-related peptide. However studies have produced conflicting results for many of the variants identified [114-118] and a clear role for serotonergic genes is yet to be defined.

The dopaminergic system has also been extensively studied however, the mechanism through which it influences migraine remains unclear despite the fact that dopamine antagonists are known to be effective in relieving migraine symptoms. SNPs within the dopamine beta hydroxylase gene as well as an insertion/deletion polymorphism have shown significant association with MA in a number of studies. One SNP in particular namely, rs161115, is known to be responsible for 31-52 % of dopamine beta hydroxylase's enzymatic functioning and is significantly associated with migraine in Caucasian populations [119-121].

The role of glutamate in migraine is becoming increasingly supported with evidence accumulating from GWAS [122] .Glutamate acts as a neurotransmitter and, as previously described, is thought to influence CSD susceptibility and activation of the trigeminovascular system [123].Several genes outside those identified in GWAS have also been considered including the GRIA receptors with associations identified at the *GRIA1* and 3 genes [124].

Finally, dysfunction of other ion channels involved in neurotransmitter release and neuronal excitability are becoming a focus of many studies. This has been particularly exemplified by the *KCNK18* gene that encodes the TRESK protein – a potassium channel involved in cellular excitability. This protein is expressed primarily in the dorsal root ganglion and trigeminal ganglia, key regions of the CNS thought to be involved in pain processing thus supporting a role for this gene in migraine. A frameshift mutation in this gene was identified as segregating perfectly with MA in an Australian family as well as a number of other mutations that have been identified only in migraine cases and not controls [125]. However, recent functional

investigations of these variants have demonstrated that a mutation causing a non-functional copy of the gene is not sufficient alone to cause migraine and therefore interactions with other genetic and/or environmental factors must be involved [126].

Overall, the study of neurological genes have demonstrated a strong contribution to migraine susceptibility however further studies are needed to examine and fully characterise the role of neurological genes in migraine pathogenesis.

Vascular Genes

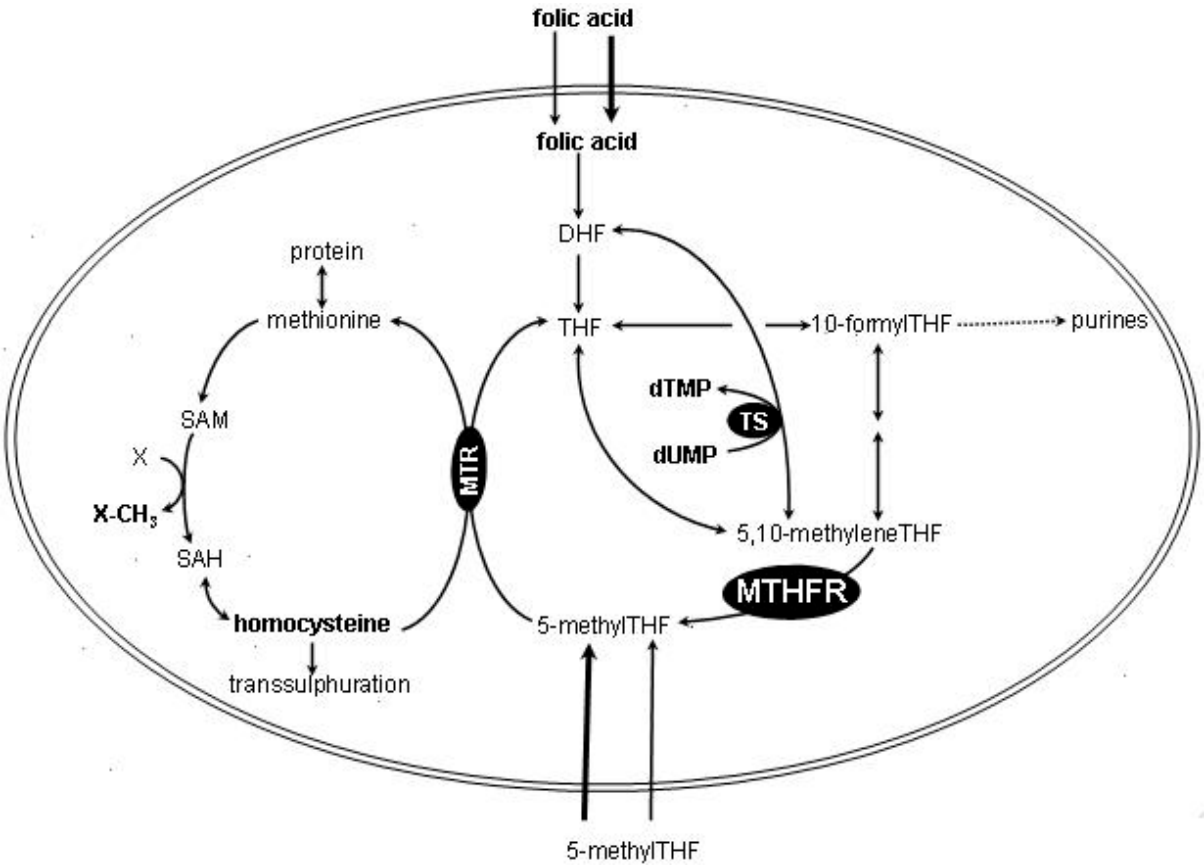
Migraine is known to be co-morbid with other disorders many of which are vascular related diseases including stroke, hypertension and PFOs (Patent Foramen Ovale) [127]. This connection, combined with a known efficacy of vasoactive drugs in migraine treatment, has prompted researchers to investigate genes related to these disorders which may overlap with migraine pathogenesis. Two such key genes include the *MTHFR* gene encoding methylenetetrahydrofolate reductase and *ACE* encoding for Angiotensin I converting enzyme [82].

MTHFR

The *MTHFR* gene maps to chromosomal location 1p36.3 and has been studied extensively to try and identify associations between variants and increased risk for disease [128]. Numerous associations have been made for a variety of diseases, many of which are neural and/or vascular and may therefore be involved in pathways overlapping with migraine pathogenesis. The methylenetetrahydrofolate reductase (*MTHFR*) gene encodes for the MTHFR enzyme which is involved in the amino acid and purine biosynthesis pathway as illustrated in Figure 3. The MTHFR enzyme catalyses the reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate which is needed for the conversion of homocysteine to methionine

[129, 130]. Folate is also needed to drive the pathway therefore a lack of dietary folate and/or reduced MTHFR enzymatic activity can result in an increase of homocysteine levels in the blood plasma. Hyperhomocysteinemia has been associated with a variety of metabolic disorders and increased risk for complex diseases including heart disease and stroke as well as migraine with aura [131-133].

Figure 3: Flow diagram illustrating the production of homocysteine as part of the amino acid and purine biosynthesis pathway. Included in the illustration is the role of a key enzyme in this process, MTHFR.



Abbreviations: dihydrofolate (DHF), deoxythymidine 5'-monophosphate (dTMP), deoxyuridine monophosphate (dUMP), S-adenosyl homocysteine (SAH), S-adenosyl methionine (SAM), tetrahydrofolate (THF), thymidylate synthase (TS). MTHFR catalyses the reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate which is needed for the conversion of homocysteine to methionine. A reduction in MTHFR activity results in an accumulation of homocysteine.

The clinical consequences of elevated homocysteine plasma levels include endothelial cell injury, spontaneous trigeminal cell firing and alterations in coagulant properties of blood [134]. Spontaneous trigeminal cell firing leading to inflammation in the meninges and dilation of cerebral vessels is thought in part to cause the pain associated with migraine [135]. Thus homocysteine dysfunction can clearly increase patient propensity for developing migraine. Oxidative damage to the vascular endothelium via formation of superoxide anions (auto-oxidation of homocysteine) may also increase the likelihood of migraine and other vascular disorders such as stroke [136]. Physiological studies have demonstrated these relationships, making it clear that genetic variants altering enzyme activity or substrate pathways can increase the risk for developing migraine and other vascular diseases.

Variants within the *MTHFR* gene which result in decreased enzyme activity may therefore be associated with migraine and should be studied. Two variants in particular within this gene have been examined in previous studies. These are an A>C change occurring at position 1298 of the *MTHFR* gene and a C>T change at position 677 [82]. The 1298 A>C variant results in decreased MTHFR activity to a somewhat lesser degree than the 677 C>T variant. It has been associated with neural tube defects and cardiovascular disease, but its role in migraine pathogenesis remains unexamined [137, 138].

The question of whether or not the 677 C>T variant is associated with migraine generated some controversy. Six previous studies found a significant association between the *MTHFR* 667 C>T variant and migraine with aura [139-144], while two conflicting studies found no association [145, 146]. Hinsanori Kowa conducted a study in a Japanese patient cohort in 2000 to examine the association between the *MTHFR* 667 C>T variant and migraineous headaches (the T/T genotype has been found to increase homocysteine levels in the blood).

The study found a significant association between the T/T genotype and migraine sufferers. The homozygous transition was found in 20.3 % of migraine sufferers compared to 9.6 % in controls, and occurred in a remarkably high frequency in individuals suffering from MA (40.9 %). Findings were concluded to be highly significant with an odds ratio of 6.5 [139].

A different study group examining Spanish patients suffering from migraine with aura also found a significant association with the T/T genotype. Patients were recorded as having elevated homocysteine blood plasma levels and were therefore also considered to be at risk of stroke and other vascular anomalies. While the odds ratio in this cohort was only 2.34, findings were still considered to be significant and authors concluded that the homozygous mutation is associated with MA in the population group they were studying [143]. Similarly Scher *et al*. and other studies found significant association between the T/T genotype and migraine with aura (odds ratios were in the region of 2.05). A subsequent large meta-analysis which pooled data from all previous studies regarding the association between 677 C>T found that migraine but only MA and not MO is associated with the C to T transition [147]. This significant study has provided compelling evidence that the MTHFR gene plays a critical role in MA pathogenesis.

The Effects of Vitamin Supplementation and MTHFR

The results of a recent clinical trial verifies that the 677 C>T *MTHFR* variant which results in reduced enzyme levels does in fact have a direct correlation with homocysteine levels in the blood and pathogenesis of migraine with aura. The results strengthen the hypothesis pertaining to the role of homocysteine in MA susceptibility by showing that vitamin supplementation leading to decreased homocysteine levels also brought about a decrease in

the frequency and severity of migraines in MA sufferers [148]. Homocysteine can be alternatively metabolised through a B6 dependent pathway [149] and therefore supplementation with folic acid, B6 and B12 can decrease homocysteine levels. The pilot study by R Lea *et al* showed that a combination of folic acid, B6 and B12 supplementation reduced the homocysteine levels in all patient's blood samples and that this correlated with a reduction in the frequency and severity of MA attacks. Furthermore response to treatment was directly correlated to *MTHFR* 677 C>T genotype, with T/T homozygotes showing the smallest response to supplementation. It was suggested that dose should be dependent on genotype and that T/T individuals should receive the highest doses [148]. This is what is referred to as personalised medication, a concept that has been growing in popularity.

Angiotensin I Converting enzyme

Angiotensin I-converting enzyme catalyses the conversion of angiotensin I to angiotensin II, which acts as a vasoconstrictor. *ACE* also plays a role in the inactivation of bradykinin, a strong vasodilator, thus having an overall powerful vasoconstrictory effect on blood vessels [150]. A modest risk factor for vascular disease involving the angiotensin I-converting enzyme (*ACE*) I/D polymorphism has also been implicated as a risk factor for migraine [151]. It has been suggested that the D/D genotype could have a synergistic effect in individuals carrying the *MTHFR* T/T genotype and result in a much greater propensity towards developing migraine. As with the *MTHFR* gene, *ACE* seems to play an important role in the pathogenesis of migraine with aura, rather than in migraine without aura. Studies by Paterna first suggested a role for the *ACE* D/D genotype in migraine and later an investigation in a Japanese cohort showed a relationship between the D allele and MA [152-154]. A recent study in an Australian cohort confirmed an over-representation of the D/D genotype in

patients experiencing MA and furthermore suggested that the *MTHFR* T/T genotype acts in combination with the D allele in increasing migraine susceptibility [155].

While several other vascular genes have been investigated with migraine susceptibility the results have been largely conflicting or negative. Calcitonin-Gene Related Peptide (*CGRP*) [156] and the *NOTCH3* [157] genes are a few examples where studies may have been hindered by variability in ethnic cohorts, migraine subtypes or generally under-powered studies that have difficulty replicating associations especially when the effect size of the variant is minimal.

Hormonal Genes

It has been well documented that migraine affects three times as many females as male patients. The abrupt increase in female patients occurs at puberty and migraine incidence has been correlated with other hormonal changes such as menstruation, pregnancy and menopause [158, 159]. Therefore genes involved in hormonal processes during female development and the menstrual cycle have been a key area of migraine genetic studies. In particular genes governing oestrogen and progesterone have received significant attention.

It has been proposed that the oestrogen receptor (*ESR1*) may play a significant role in migraine. The receptor is expressed in many parts of the brain including in neurons in the trigeminal ganglion [160] and there is evidence to show that oestrogen acts as a regulator of neurotransmitters involved in migraine including CGRP and serotonin. Overall it has been proposed that oestrogen plays a role in sensitization of the trigeminal neurons although the exact mechanisms through which this is achieved are yet to be defined. There is evidence to suggest that abrupt oestrogen withdrawal precipitates onset of a migraine episode [161] or

that oestrogen is involved in CSD susceptibility. At a genetic level associations at the 594 G>A and 325 G>C SNPs in the *ESR1* gene have been reported in several migraine populations [162-165], however contradicting results have also been observed in a number of studies.

Other hormonal genes that have been investigated to varying degrees, and similarly with varying levels of validation [166], include the Progesterone receptor (particularly an Alu insertion at intron 7) [113, 167, 168], Androgen receptor [113], and Follicle Stimulating Hormone receptor [164]. While there is evidence to propose a role for these hormonal genes in migraine - such as expression levels in areas of the CNS involved in pain pathways, and/or interactions with known neurotransmitters that may also act in these pathways; the exact mechanisms through which these genes exert an influence on migraine susceptibility, and the extent of that influence, requires further investigation.

Migraine with Aura

Approximately 20 % of migraine attacks involve a preceding neurological disturbance known as an aura [169]. Numerous studies have linked cortical spreading depression with the visual scintillations typically experienced during aura and recently MRI has been used to confirm these findings in addition to describing underlying physiological mechanisms [170]. Cortical spreading depression (CSD) is a wave of neuronal and glial depolarization/ neuronal hyperexcitability followed by a long lasting suppression of neural activity [85]. CSD is thought to activate the trigeminal nerves thus resulting in the pain associated with migraines. This electrophysiological event has been linked to aura in the human visual cortex and is thought to be partly responsible for the sensory and motor disturbances experienced during

MA attacks. Furthermore a change in blood flow in the occipital lobe has been shown to correspond with the aura attack. This could be as a result of increased and decreased neuron firing and merely a side effect of energy metabolism or alternatively the altered blood flow may be a contributing factor towards triggering attacks. This in conjunction with the fact that MA patients also have an elevated risk for stroke has made genes involved in vascular pathways good candidates for genetic studies [171-173].

Given the role of neuronal hyperexcitability in CSD any genes which may alter the electrophysiological signalling of neurons are also good candidate genes to examine. This idea has led to an interest in ion channel genes involved in channelopathies. Previous studies have shown that FHM mutations are not found in typical migraine with aura, suggesting that other ion channels may be involved. A recent and exciting study by Lanfreniere has highlighted the role that the *TRESK* potassium channel gene may play in common migraine with aura [125]. As discussed previously, *TRESK* encodes for K2P channels which are expressed throughout the central nervous system, including the trigeminal ganglion neurons [125]. They play an important role in controlling resting membrane potential and neuronal excitability [174], and therefore an alteration in expression could feasibly lower the threshold for CSD.

A large cohort of both case-control individuals as well as families were studied, which identified a number of variants in the *TRESK* gene. The most notable variant identified was a frame shift mutation (F139WfsX24) which segregated perfectly in a family affected with typical MA. All eight affected individuals carried the mutation while all eight unaffected individuals did not. The mutation was shown to suppress wild type channels and was thus classified as being a dominant-negative mutation inherited in a dominant fashion. Drug

induced inhibition of *TRESK* activity has previously been associated with an increase in headache frequency and/or intensity [175]. This adds to the growing body of evidence that *TRESK* plays a role in MA susceptibility. Lanfreniere and colleagues have highlighted the need to investigate other variants in the *TRESK* gene which may be associated with MA and have also suggested the potential role of agonist mediated drugs as a therapeutic target [125].

Previous Molecular Genetic Studies of Migraine

GWAS

A number of large GWAS studies aiming to identify associations between genetic variants and migraine have been conducted, with the first being published by the International Headache Genetics Consortium (IHGC) in 2010 [122]. This initial study examined 2,731 MA migraineurs in the Finish, German and Netherland population compared to 10,747 controls from the same population group. The study was replicated using an additional 3,202 cases and 40,062 controls including patients suffering from MO. After analysing 429,912 SNPs only one reached genome wide significance ($p<5x10^{-8}$), with another 11 SNPs showing threshold significance levels. The discovery of rs1835740 gave support to the hypothesis that glutamate plays a role in migraine pathophysiology as this variant is located between the genes *MTDH* and *PGCP* which are both involved in the glutamate homeostasis pathway [176]. Earlier studies have shown that glutamate plays a role in trigeminovascular pain processing and may modulate the threshold for CSD [123]. This GWAS therefore strengthened theories relating to a neurological mechanism involved in the propagation of migraine attack.

Two subsequent GWAS studies conducted by the Dutch-Icelandic migraine genetics consortium (DICE) and the Women's Genome Health Study (WGHS) failed to replicate the results from the first GWAS and only identified modest associations [177]. However a large meta-analysis of all the cohorts identified three new SNPs showing significance at the genome wide level namely rs2651899, rs10166942 and rs11172113 located in intron 1 of *PRDM16*, at 2q37 near *TRPM8* and intron 1 of *LRP1* respectively [178]. The role that *PRDM16* may play in migraine is unknown, but *TRPM8* encodes a cold burning pain sensor and has been found to be expressed in sensory neurons and the dorsal root ganglion [179]. Animal models have shown that *TRPM* may be a target of neuropathic pain and since migraine and neuropathic pain share some characteristics, it is biologically plausible for there to be a link between *TRPM* and migraine [180]. *LRP1* is a lipoprotein receptor and is of particular interest as it interacts with NMDA glutamate receptors on neurons providing further support for a role of the glutamatergic system in migraine [176].

The most recent GWAS study has identified susceptibility loci for MO [181]. This was the first migraine GWAS which focused on MO as a pose to MA. The study examined 2,326 migraineurs from headache centres in Germany and the Netherlands, with 4,580 population matched controls. The top 12 loci were then repeated in a replication study involving 4 independent clinic-based European MO samples in a total of 2,508 cases and 2,652 controls. The MA susceptibility loci *TRPM8* and *LRP1* identified in the previously mentioned GWAS were replicated and therefore found to also be significantly associated with MO [178]. In addition SNPs located in the *MEF2D* gene (1q22) and near the *TGFBR2* gene (3p24) showed highly significant association in both the discovery and replication datasets. Weaker association was also found for SNPs at the *PHACTR1* and *ASTN2* loci.

The *MEF2D* gene encodes for myocyte enhancer factor 2D which is a transcription factor highly expressed in brain tissue and responsible for regulation of neuronal differentiation. Activation of this transcription factor has also been shown to restrict the number of excitatory synapses lending evidence to the theory that dysregulation may affect neuronal excitatory neurotransmission [182]. *TGFBR2* is also involved in regulation of cell proliferation and differentiation and encodes for a serine-threonine kinase. In a large family it was found that a missense mutation in *TGFBR2* (p.Arg460His) caused monogenic familial aortic dissection along with migraineous headaches in 11 of out the 14 affected probands [183].

In summary genome wide studies in large case-control cohorts have provided evidence largely for neural and vascular mechanisms in the pathogenesis of migraine.

Genomewide Pedigree Studies

Families showing strong heritability of migraine provide unique advantages that cannot be replicated in outbred populations. In particular it is expected that families will have reduced heterogeneity in the variants that are functionally relevant to migraine susceptibility within their family. Although the caveat to this is that due to the relatively high prevalence of migraine in the general population (~12 %) it is often difficult to obtain a pedigree for study where migraine inheritance has not been moderately complicated by migraine affected spouses marrying into the family potentially contributing different susceptibility variants. Another potential advantage is that families allow for a small degree of control over environmental factors such as lifestyles and diet particularly where the family unit has been maintained in a defined geographical locale [184].

Linkage studies of multigenerational pedigrees have been successful in identifying several migraine susceptibility loci across the genome. Key loci that have received considerable attention include 19p13 [185, 186], which encompasses a region that codes biologically plausible genes including *CACNA1A* (FHM1), *INSR*, *NOTCH3* and *LDLR*; 15q11-q13 [187], which contains a cluster of genes encoding GABA receptor subunits; and Xq24-28 [186, 188, 189] which contains another cluster of genes encoding GABA subunits, as well as a number of GRIA subunits. In addition several of these loci have been identified using a number of independent families and some, such as the regions on Chromosome 4 – 4q21 [190] and 4q24 [191], have been identified as overlapping regions. However, perhaps due to the genetic heterogeneity of migraine as well as its diverse clinical presentation none of these loci have had causative genes conclusively identified within them.

A final genomewide pedigree approach that has been employed is analysis of the Norfolk Island genetic isolate where migraine prevalence is observed at 25 % and heritability has been estimated as 0.53. A pedigree based GWAS implicated a SNP in the Zinc Finger Protein 555 gene of unknown function. Further analysis involving biological prioritisation implicated SNPs within the neurotransmitter-related *ADARB2*, *GRM7* and *HTR7* genes [192]. In addition risk haplotypes at the Xq12 region encoding the *HEPH* and *VSIG4* genes were also identified in this genetic isolate and showed replication in a large independent migraine cohort [193].

Overall pedigree based studies have provided a means to direct research to targeted regions of the genome. While identification of causative genes are still pending, this is likely a result of genetic and clinical heterogeneity in the samples being used to replicate these linkages and identify the genes. However, as the era of sequencing advances, it will be possible to reconsider these regions in smaller populations that display greater homogeneity in migraine

presentation using an approach that can interrogate the region at greater depth and coverage than current candidate gene studies.

Current methods for diagnosis and treatment of migraine

For individuals suffering from migraine attacks or migraine like symptoms it is recommended to see a general practioner (GP) who can make an initial assessment. Based on the patient's symptoms and family history as well as the severity of the presented condition the GP will either prescribe some basic drugs or make a referral to a neurologist as necessary. In some cases it may be deemed necessary to rule out other possible causes by making use of diagnostic tests such as a computerised tomography (CT) or magnetic resonance imaging (MRI) scan and in some circumstances a spinal tap. Sub-types of migraine are diagnosed according to criteria as set out by the International Headache Society.

Once a diagnosis has been made there are a plethora of drugs available for the treatment of migraines. There are two main categories of drugs namely preventative and pain-relieving. Some of the most common drugs prescribed to curb a migraine attack once the pain has already begun (preventative drugs) include anti-algesics such as ibuprofen, acetaminophen, paracetamol, triptans, ergotamine, anti-nausea medications, opiates and dexamethosone. In cases where a patient suffers from severe migraines on a regular basis preventative medication may be prescribed. These drugs are taken on a daily basis and include cardiovascular drugs such as beta blockers and calcium channel blockers, anti-depressants, anti-seizure drugs, cyproheptadine and botulinum toxin type A. Unfortunately many of these drugs are only effective for a percentage of patients. Many individuals who suffer from migraine attacks do not receive adequate treatment and therefore development of new

therapeutics is of importance. Another downfall of many medications is their side effects which can be severe for some individuals. Commonly prescribed medications are listed in Table 2 below.

Table 2: Commonly prescribed drugs used to treat migraine

CATEGORY	Drug Name	Side Effects
Pain Relieving		
Anti-algesics	Ibuprofen Acetaminophen Indomethacin Excedrin (acetaminophen, aspirin and caffeine combination)	Stomach ulcers Gastrointestinal bleeding Re-bound headaches
Triptans*	Sumatriptan Rizatriptan Almotriptan Naratriptan Zolmitriptan Frovatriptan Eletriptan Treximet (sumatriptan and haproxen sodium combination)	
Ergots	Migergot Cafergot Dihydroergotamine	
Anti-nausea	Metoclopramide Prochlorperazine	
Opiates	Codeine	Addictive
Dexamethasone		Steroid toxicity
Preventative		
Cardiovascular	Propranolol (Beta blockers) Verapamil (Calcium channel blocker) Lisinopril (Anti-hypertensive)	Dizziness, Drowsiness Light headedness
Anti-depressants	Amitriptyline Nortriptyline Protriptyline Venlafaxine	
Anti-seizure	Valproate Topiramate Gabapentin Lamotrigine	nausea and vomiting, diarrhea, cramps, hair loss, and dizziness
Cyproheptadine		
Botulinum toxin type A		

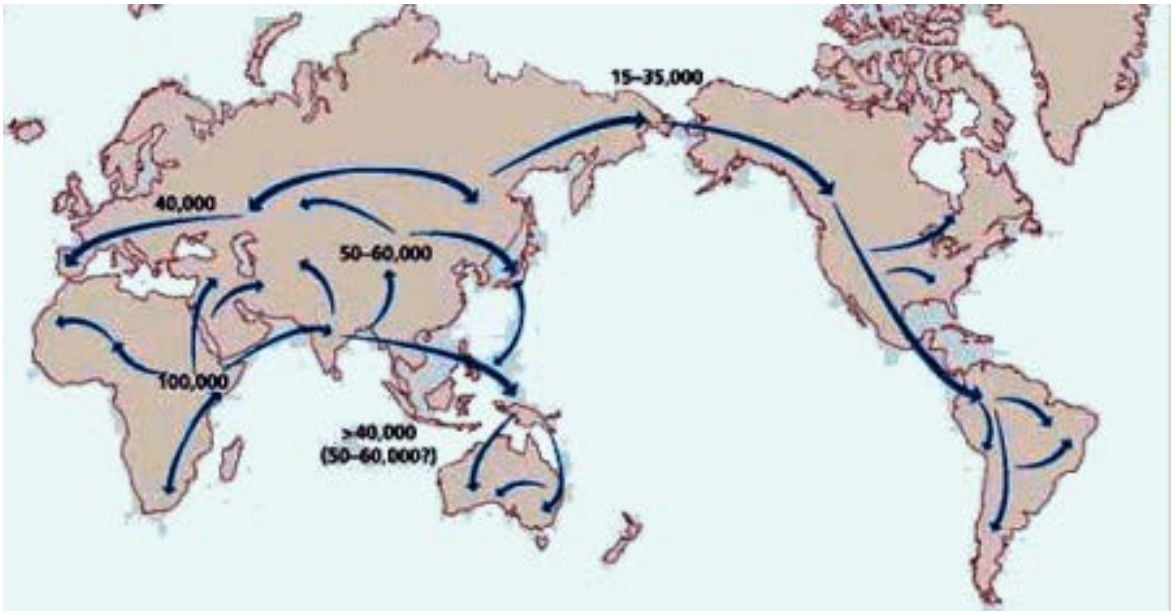
*not suitable for patients who are at risk for stroke or heart disease

The Mitochondrial Genome

Mitochondria are organelles found in eukaryotic cells, including all mammalian cells and function primarily to produce ATP via the oxidative phosphorylation chain. Human mitochondrial DNA occurs as a distinct set of DNA, apart from the nuclear DNA found within the nucleus of each cell. Each organelle contains its own genetic system capable of replication, transcription and translation [194]. Mitochondria range from 0.5 to 10um in size and are thought to originate from an endosymbiotic relationship with a bacterium in our distant evolutionary history[35]. The genome is a multi-copy, circular dsDNA molecule similar in structure to the bacterial genome. The number of copies found in each cell depends on the cell type, with liver and muscle cells containing the highest densities of mitochondria, to cater for their high energy needs. As many as several thousand mitochondria have been found per cell in muscle and neural tissues.

The mitochondrial genome is significantly smaller than the nuclear genome, containing only 16.6 kb of DNA which encodes for a total of 37 genes [195]. Of these genes, 13 are protein coding and encode polypeptides of the oxidative phosphorlylation pathway. The other genes encode RNAs involved in translational functions with 2 genes encoding for rRNAs and 22 for tRNAs. A control region or D loop enables the start of transcription and mutations within this region have been used to track human migration patterns through haplogroup combinations found along our ancestral lineages as shown in Figure 4 below.

Figure 4: World map showing proposed human migration patterns according to the "Out of Africa theory"



Taken from http://mrbarlow.wordpress.com/2011/05/page/2/

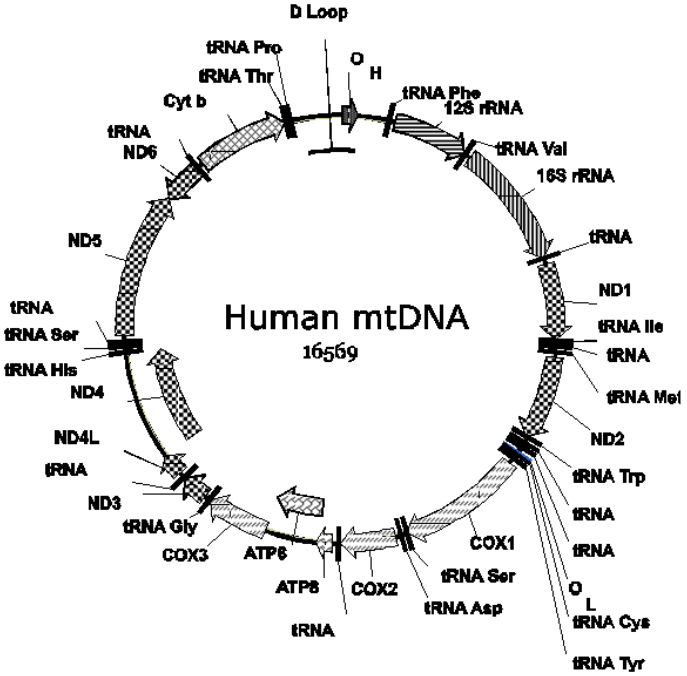
Arrows indicate the initial migration thought to have occurred out of Africa approximately 100 000 years ago, followed by subsequent migrations into Europe and Asia. Arrows indicate the direction of human migratory patterns and time estimates for each major migration event.

Since mitochondrial DNA does not undergo recombination and is inherited exclusively along the maternal lineage, any changes that occur in the DNA sequence can be tracked through thousands of generations. By connecting maternal lines in living people from all over the world, we can all trace our ancestry back to a mitochondrial Eve who lived approximately 150 000 years ago in Africa [194, 196]. Furthermore, all European lineages can be traced back to seven women who lived in a period dated to the first major human migration out of Africa [197]. This molecular evidence in conjunction with archaeological finds has made the "Out of Africa" theory the one that most scientists subscribe to. This theory states that all humans originated from Africa and that two major migrations, one 70 000 years ago and another more recently 35 000 years ago resulted in population of the rest of the globe.

Evolution from these time points onwards explains much of the genetic variability currently found between different ethnic groups [198]. While this is the most widely accepted theory concerning human evolution, opposition does still exist within some groups of scientists who have alternative views.

The basic structure and organisation of the mitochondrial genome is shown below in Figure 5.

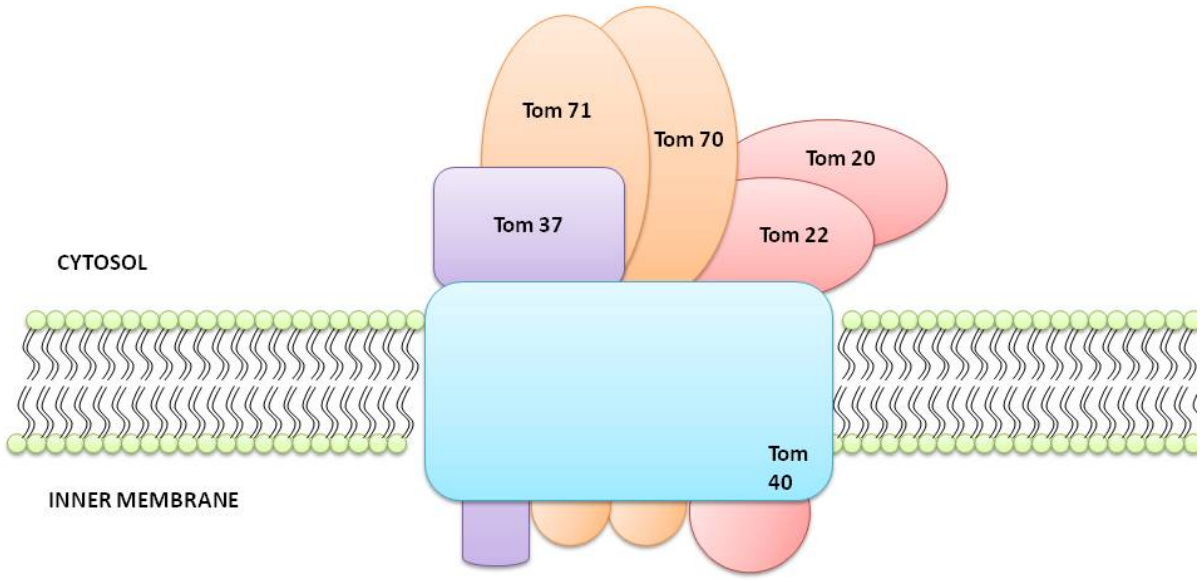
Figure 5: Structure of the Mitochondrial Genome



From: Inna Shokolenko, Susan LeDoux, Glenn Wilson and Mikhail Alexeyev (2011). Mitochondrial DNA Damage, Repair, Degradation and Experimental Approaches to Studying These Phenomena, DNA Repair - On the Pathways to Fixing DNA Damage and Errors, Francesca Storici (Ed.), ISBN: 978-953-307-649-2, InTech, Available from: http://www.intechopen.com/books/dna-repair-on-the-pathways-to-fixing-dna-damage-and-errors/mitochondrial-dna-damage-repair-degradation-and-experimental-approaches-to-studying-these-phenomena Each of the functional elements of the mitochondrial genome are shown above with arrows indicating the direction of transcription. Shown are 37 genes, 13 protein coding, 2 encoding rRNAs and 22 encoding for tRNAs.

Despite the fact that the mitochondrial genome encodes for so few proteins, thousands of proteins are still found within the mitochondria. The rest of the proteins are encoded by the nuclear genome and are transported into the mitochondria via membrane receptor proteins. MitoProteome is a database which records all known nuclear encoded mitochondrial proteins and is curated on an ongoing basis. Currently 780 known nuclear encoded mitochondrial proteins are listed, with an additional 492 listed as putative mitochondrial proteins [199, 200]. These 1000+ proteins are targeted to the mitochondria and sorted to the different mitochondrial sub compartments following translocation through the mitochondrial membranes. Separate translocases in the outer (TOM complex) and inner (TIM complex) membrane assist in identifying pre-proteins and concurrently transporting them across the two membranes. Factors in the cytosol as well as molecular chaperones in the matrix assist in this process [201].

Figure 6: Illustration of the TOM Complex involved in import of nuclear encoded proteins into the mitochondria. Adapted from [201]



The phospholipid membrane found surrounding all cells is depicted by a green circle representing the phosphate head and double lines representing the hydrophobic tails. Each sub-unit of the TOM complex is shown with geometric shaded shapes and labelled accordingly. The TOM complex assists with identification of pre-proteins that need to be imported into the mitochondrial matrix for further processing and downstream pathways. This is essential for mitochondrial function and importing nuclear encoded proteins in this way allows the majority of functional molecules to perform their designated duties.

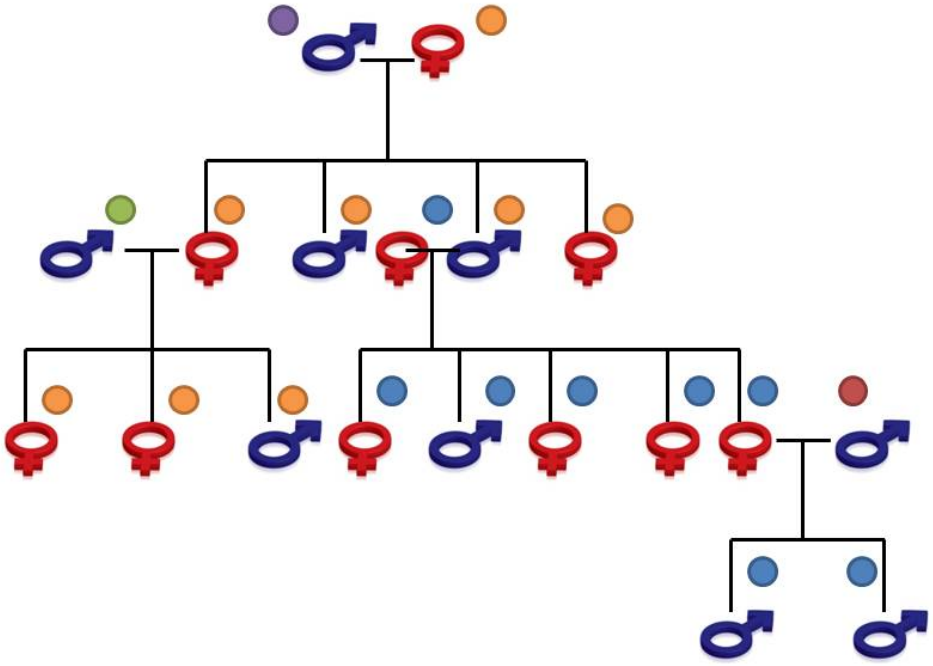
New evidence suggests that regulators of gene expression including miRNAs are also imported into the mitochondria. A recent study showed that miRNAs which are already known to be transcribed within the nucleus and undergo processing and maturation in the cytosol also localise to the mitochondria. The study found 13 miRNAs to be significantly

enriched in the mitochondria and the authors hypothesize that post-transcriptional regulation via miRNAs could be a plausible mechanism to alter expression in response to metabolic conditions of the cell [202]. Two of the miRNAs namely miR-1974 and miR-1977 mapped to the mitochondrial tRNA and rRNA genes, suggesting that not all mitochondrial miRNAs originate from the nucleus. As is the case for proteins, it is likely that a combination of mitochondrial and nuclear genes are involved in the production of miRNAs which are involved in regulation of mitochondrial function.

Inheritance of Mitochondria

Mitochondrial DNA is passed exclusively along maternal lineages as illustrated in Figure 7 and previously discussed in the introduction [203]. This clear mode of inheritance has aided researchers in identifying causal variants involved in mitochondrial disorders through family linkage studies. However, as is the case for nuclear inherited variants, mitochondrial diseases can be Mendelian or complex. Most complex mitochondrial diseases involve nuclear encoded genes in addition to mitochondrial variants and can show variable expressivity with different levels of penetrance, even in the same family. Some of the variable expressivity can be explained by heteroplasmy (discussed in detail in section 2.2) and an age related accumulation of DNA changes [194].

Figure 7: Mode of inheritance for mitochondrial DNA



Coloured circles show the pattern of mitochondrial inheritance, with mitochondria passed directly down the maternal lineage. In each generation it can be seen that all offspring from the same female line exhibit identical mitochondria which are then successively passed down to both sons and daughters. Married in individuals present new mitochondrial lineages, but are only passed on through the female line.

As we age the number of somatic mtDNA mutations increases. This is largely due to the higher mutation rate of the mitochondrial genome which is estimated to be around 100 times higher than the nuclear genome [204]. However as the number of mtDNA mutations increases, so does the rate at which free radicals are generated which further damages the mtDNA and can lead to clonal proliferation of the mutated mitochondria over time. Mitochondria exist as heterogeneous populations sometimes even within the same cell. Once

a ‘threshold’ level of mutant mitochondria is reached, mitochondrial disease may become clinically apparent depending on the functional effect of the mutation [194, 205].

Heteroplasmy

Each mitochondrion contains an average of five to ten mitochondrial genomes and up to several thousand mitochondria can be found in an individual cell [194]. This may give rise to a heterogeneous mixture of mitochondria and concurrently any particular variant may be present in only a proportion of the mitochondrial population. This is also true for disease causing mutations and the phenomenon is termed heteroplasmy. The percentage of mitochondria which contain a particular pathogenic mutation depends on both the proportion of wildtype versus mutant copies inherited upon fertilisation and also the accumulation of somatic mutations with age. Usually a threshold proportion of mutant versus wild type mitochondria needs to be exceeded in order for an individual to present with clinical disease symptoms [31]. This value varies depending on the mutation and functional effect.

Due to this genetic heterogeneity of complex mitochondrial diseases; variable expressivity, variable penetrance and large phenotypic heterogeneity are observed even within families. This is further complicated by environmental interactions making it difficult to identify causative genetic variants and to provide accurate phenotypic predictions for complex mitochondrial diseases. Thus far some of the methods used to identify mitochondrial genetic variation include electrospray mass spectrophotometry, PCR based amplification, restriction fragment length polymorphism (RFLP) analysis and sequencing [206].

Due to heteroplasmy it can be difficult to make a definitive diagnosis using some of these more traditional methods. Next generation sequencing allows for the generation of multiple sequencing reads covering the same region. With an increase in coverage depth, heteroplasmy can be determined more accurately. MV Zaragoza showed that 10-20 x coverage is sufficient to detect homoplasmic variants, while 200 x coverage is needed to detect up to 10 % heteroplasmy [207]. To detect heteroplasmic variants at a lower frequency much deeper sequencing is required. Coverage needed to detect heteroplasmy >5 % and as low as 2 % is estimated to be 1500 x and 15 000 x respectively [208, 209].

The full length mtDNA consensus sequence is currently available and is used to help assemble random reads generated by mass parallel sequencing and to further aid in variant identification. Just like for nuclear DNA studies, mtDNA studies require large cohorts of cases and controls to detect small to moderate associations between mtDNA variants and complex disease.

Function of Mitochondria

The known primary function of mitochondria is to produce ATP via the electron transport chain. ATP is used to power all cellular processes that require energy. Other main functions include production of reactive oxygen species (ROS), regulation of apoptosis and calcium homeostasis [35]. Neurons are heavily reliant on mitochondria to produce sufficient ATP and also to help regulate intracellular calcium levels [210]. Any pathogenic mutations affecting mitochondrial function can impair energy metabolism and ion homeostasis in neurons thus resulting in a range of downstream abnormalities, the full extent of which is not characterised or fully understood yet. Similarly accumulating damage from oxidative stress during the

aging process can render neurons vulnerable to neuronal degenerative diseases such as Parkinson's disease and Alzheimer's [211, 212].

One of the largest groups of mitochondrial disorders and perhaps the best defined are the mitochondrial myopathies and encephalomyopathies. Some of the most common syndromes are listed in Table 3 below. The mitochondrial myopathies comprise those diseases which cause muscle weakness and wasting with severity ranging from progressive weakening to death [194]. As mentioned previously muscle and neural tissues are the most sensitive to mitochondrial dysfunction, as they have the highest energy requirements out of all the body's tissues. The combined effects of energy shortage and toxin accumulation give rise to many of the symptoms of mitochondrial myopathies which include muscle weakness, exercise intolerance, heart failure, movement disorders and droopy eyelids [213].

Table 3: Summary of the common mitochondrial myopathies and encephalomyopathies

Disease	Abbreviation	Confirmed Causative Mutations	Reference
Kearns-Sayre syndrome	KSS	5 kb deletion, 12315 G>A	[214, 215]
Leigh syndrome	LS	3243A>G, 5537insT, 8363G>A	[216-218]
Mitochondrial encephalomyopathy, lactic acidosis and stroke like episodes	MELAS	583G>A, 3243A>G, 3256C>T, 3271T>C, 3291T>C, 4332G>A, 12147G>A	[216, 219-225]
Myoclonus epilepsy with ragged red fibres	MERRF	8344A>G, 8356T>C, 8363G>A, 12147G>A	[225-228]
Neuropathy, ataxia and retinitis pigmentosa	NARP	8993G/C	[229]
Progressive external opthalmoplegia	PEO	3243A>G, 4298G>A, 4308G>A, 5703G>A, 12315G>A	[215, 230-233]
Leber hereditary optic neuropathy	LHON	11778A, 14484C, 14459A, 3460A	[234]
Deafness	DEAF	1494C>T, 1555A>G, 8363G>A	[235-237]

Mitochondrial encephalomyopathies include a myopathy component in addition to a neurological aspect to the disease. Additional symptoms include hearing impairment, migraine-like headaches, seizures and in severe cases stroke like episodes [238]. Some of these symptoms can be controlled through prescription medication, but many individuals present with mitochondrial disease before the age of 20 and due to the progressive nature of the disease it can be fatal in many cases [239]. As briefly mentioned earlier, a reduction in oxidative phosphorylative activity can cause other severe neurological diseases later on in life, a process which is likely linked to aging and an increase in reactive oxygen species over time. Alzheimer's and Parkinson's disease affect many elderly people worldwide with devastating phenotypic characteristics. The risk of developing this disease is much greater if the proband has an affected mother, thus strengthening the hypothesis that mitochondrial dysfunction plays a role in susceptibility [194, 211].

Other common diseases that mtDNA variants have been associated with include cancer, diabetes, stroke, cardiomyopathy, mental retardation, migraine and male infertility [195]. Studies have found that there is an increase in the production of reactive oxygen species during neoplastic transformation for many types of cancer, a process which is known to damage the mitochondrial genome. This is thought to increase the rate at which somatic mutations occur within mitochondria promoting tumorigenesis. Given this finding some of these mutations have been proposed as being informative markers in early cancer detection. Mitochondrial mutations have been implicated in a wide range of cancers including breast, colon, oesophageal, endometrial, head and neck, hepatocellular, kidney, leukaemia, lung, melanoma, oral, prostate and thyroid cancer [194].

Mitochondrial Dysfunction and Migraine

Several lines of evidence exist to suggest that mitochondrial dysfunction contributes to the pathogenesis of at least some sub types of migraine. The hypotheses are based on the idea that an impaired mitochondrial oxidative metabolism may contribute to the pathogenesis of migraine by interrupting proper nervous functioning. Brain and muscle are highly dependent on oxidative metabolism and are therefore the most severely affected tissues in the mitochondrial disorders. A variety of morphological, biochemical, imaging and genetic studies have provided evidence that mitochondrial dysfunction may play a role in migraine pathogenesis [35]. The female preponderance of migraine affected patients further strengthens the idea that mitochondrial genomic aberrations could increase migraine susceptibility.

Generating an action potential is a process that requires large amounts of energy as Adenosine tri-phosphate (ATP) is used to restore ion gradients after the generation of synaptic and action potentials. Even though the brain only makes up 2 % of our body mass, it accounts for at least 20 % of our energy expenditure each day [27, 28] therefore an adequate supply of blood glucose and oxygen is essential to sustain neuron function. Based on this information it is possible that an insufficient energy supply caused by mitochondrial dysfunction could predispose individuals to migraine attacks by lowering the cortical spreading depression threshold. By-products of energy metabolism including adenosine and lactate have been shown to directly induce increased blood flow to the brain [29, 30]. A change in blood flow such as has been documented during CSD is clearly an important factor of migraine attacks. Thus energy metabolism could influence migraine attacks via an accumulation or deficit of by-products.

Other neuropathic diseases which may share some pathways with migraine pathogenesis have been associated with mtDNA variants in a large number of studies. These diseases include encephalomyopathies such as MELAS, MERRF and Kearns-Sayre syndrome [31]. Given the co-morbidity of MA with stroke, the study of genes involved in MELAS are especially promising starting points. Other neurological diseases with mitochondrial associations including major depression, bipolar disorder and schizophrenia have also been shown to be co-morbid with migraine, strengthening the idea of shared pathways and susceptibility variations.

Mitochondria and Calcium

A rare form of migraine, familial hemiplegic migraine type 1(FHM1) is caused by mutations in the *CACNA1A* gene. Electrophysiological studies have shown that either an increased or decreased influx of Ca^{2+} ions into cells caused by mutations in the *CACNA1A* gene has an effect on the depolarised state of cells [32]. It has been shown that FHM1 mutations can produce gain-of-function Ca (V) 2.1 channels and as a result initiate cortical spreading depression which is the phenomenon thought to underlie migraine aura.

Other functional studies have examined the role calcium ions play in increased blood flow to the brain and have shown that an increased calcium concentration within astrocytes causes vasoconstriction during cortical spreading depression. This process is mediated by a phospholipase A2 derivative (an arachidonic acid metabolite) [33, 34]. Since mitochondria are involved in calcium ion homeostasis, a process essential for the normal functioning of neurons, then it is logical that an imbalance of calcium ions could lead to an increased susceptibility to migraines. Hence variants found in mitochondrial DNA which affect calcium homeostasis could show association with migraine sufferers.

Morphological Studies

A simplistic approach to identifying causes of migraines is to examine neuronal tissue of affected patients in order to try and identify any histological differences which may be pathogenic. Thus far ethical reasons have somewhat limited histological studies to accessible tissues such as blood. It has been shown that non-actively diving tissues are most affected by structural mitochondrial abnormalities. These tissues include skeletal muscle and neurons [39, 240]. Examination of muscle tissue has shown a correlation between abnormal mitochondria and migraine sufferers. Ragged red fibres and cytochrome-c-oxidase fibres have been found in the skeletal muscle of some migraine patients [35]. This association was only observed for the more severe cases such as those patients affected by migraine with prolonged aura and FHM patients [36-38]. Ragged red fibres (RRFs) have an abnormal number of sacrcolemmal mitochondria [39], while cytochrome-c-oxidase (COX) negative fibres often have increased fat concentrations [40, 41]. COX negative fibres are found in most patients suffering from mitochondrial encephalomyopathies and are considered to be histologically characteristic of the disease [31, 39, 42, 43].

Examination by electron microscope has revealed clusters of giant mitochondria with paracrystalline inclusions in migraine sufferers [36, 44]. Additional ultra-structural changes of the mitochondria in affected individuals have also been recorded and may reflect alterations caused by impaired oxidative metabolism. With particular reference to ultra-structural changes, accumulation of subsarcolemmal mitochondria have been found in muscle fibres of patients suffering from migraineous stroke (MS) [45]. Despite many promising morphological associations, pathogenic mtDNA variations remain to be detected. A more detailed investigation into this area is needed in order to produce significant results. Most

studies have been limited by small number of study participants and examination of larger cohorts in combination with DNA testing may produce significant results.

Biochemical Evidence

A number of studies dating as far back as 1973 have shown that there is evidence supporting the hypothesis that a deficit in the muscular and/or cerebral mitochondrial energy metabolism is significantly associated with migraine attacks. Analysis of intermediates in the oxidative phosphorylation and Krebs cycle pathways have demonstrated this relationship. Physiological studies have shown that lactate levels in the cerebral spinal fluid (CSF) are increased compared to normal levels during migraine attacks [44, 46, 47].

Elevated levels of lactate in the CSF is considered to be an indicator of a defective oxidative metabolism as lactic acidosis is indicative of an impaired utilisation of pyruvate in the Krebs cycle [48]. This finding has led to further study of other metabolite intermediates and enzymes involve in the oxidative phosphorylation pathway and also the Krebs cycle. These compounds include pyruvate, monoamine-oxidase (MAO), succinate-dehydrogenase, NADH cytochrome-c-reductase, succinate-cytochrome-c-reductase, NADH-dehydrogenase and citrate synthetise [36, 49, 50]. In each study the findings indicated a correlation between impaired mitochondrial metabolism and migraine attacks. A more recent study further showed a correlation between carnitine deficiency and MA. In these patients treatment by carnitine replacement resulted in lessening of headache severity and frequency [51, 52].

Phosphorus magnetic resonance spectroscopy studies

Phosphorus magnetic resonance spectroscopy is an imaging technique which allows researchers to non-invasively investigate brain energy metabolism *in vivo* [53]. It has proven

to be a very useful tool in examining altered oxidative phosphorylation metabolism in migraine sufferers. The functionality of mitochondria is assessed by measuring intracellular levels of phosphocreatine (PCr), inorganic phosphate (Pi) and adenosine diphosphate (ADP). Under anaerobic conditions or when mitochondrial enzymes are near saturation, additional energy is generated by transferring Pi from PCr to ADP to form ATP. This reaction is catalysed by creatine kinase and the efficiency at which it takes place is considered to be a measure of mitochondrial functionality. Thus by calculating the PCr/Pi ratio a measurement of the energy status of the cell is obtained. The lower the ratio, the less energy is available in the cell [54-56]. An additional measure of oxidative metabolism is the V/Vmax ratio. This is indicative of the velocity of oxidative metabolism in relation to its maximum capability. A further method of monitoring metabolism is the measurement of intracellular pH [57]. As mentioned previously lactic acidosis is indicative of inefficient pyruvate metabolism.

A large number of studies investigating mitochondrial metabolism in MA, FHM, MPA, MS and the related disorders CPEO, MELAS, LD, MERRF, LOHN, MS, MM and RP have all identified similar patterns of metabolic abnormalities. In almost all cases a low PCr/Pi ratio indicating low availability of free energy was observed. Increased ADP indicating a lower energy reserve in the brain cells and an increased V/Vmax ratio were also recorded. An increase V/Vmax ratio is indicative of a defective respiratory chain that supplies insufficient energy to meet the cell's demands. This pattern has suggested to be typical of mitochondrial disorders [58-62]. However whether these changes are caused by a primary mitochondrial dysfunction or are merely a side effect of brain hyperexcitability remains unclear [56, 63].

Genetic Studies

Molecular genetic studies investigating the potential role of mitochondrial variation in migraine susceptibility have primarily focused on the mtDNA mutations that are strongly implicated in the mitochondrial encephalomyopathies [35]. These include the mitochondrial encephalopathy, lactic acidosis and stroke-like episodes or MELAS mutations (mt 3243 A>G and mt 3271 T>C), the myoclonus epilepsy with ragged red fibres or MERFF mutation (mt 8344 G>A), Kearns-Sayre syndrome common 4977 bp deletion and the Leber hereditary optic neuropathy or LHON mutations [38, 45, 241-247]. No significant associations were found in any of these studies except for one false positive finding in a Japanese study. Given that the largest patient sample size used was 47 this may have been a limiting factor. Similar to nuclear DNA studies, mtDNA studies require large cohorts of cases and controls to detect small to moderate associations between mtDNA variants and complex disease. Very large sample sizes are needed in order to obtain statistically significant results, especially for mitochondrial variants which can be rare and occur at very low frequencies of the population. Therefore future investigation of these already studied variants may be required.

One previous study showed a positive association between one of the MELAS causative mutations, mt 11084 A>G and migraine in a Japanese population, but this was later shown to be an artefact of ethnicity. The mt 11084 A>G change was further investigated by the same research group and then determined to be a common polymorphism in the Japanese population and not associated with migraine [241, 248]. Since this study other research has shown significant association between mitochondrial variants and migraine susceptibility.

Finnila *et al* found that patients with a rare mt 4336 A>G change, are at a significantly increased risk of developing sensorineural hearing loss and/or migraine in later life [249]. A

total of 13 individuals out of a patient cohort of 1072 were found to have the rare variant, while only 3 control samples out of 575 included the study of neurodegenerative disorders as well as diabetes, epilepsy, hearing loss, stroke, white matter disease and ataxia. Upon statistical analysis it was found that only migraine and sensorineural hearing loss were significantly associated with the rare mt variant at a 95 % confidence interval. However of the 13 only 2 patients with the rare variant had migraine and of the 575 cases tested, only 42 were defined as migraine sufferers. Thus the rare variant was detected twice in a small cohort of migraine sufferers. Full mitochondrial genome sequencing was performed for 10 of the patient samples and all 3 controls that had the mt 4336 A>G change and it was found that the only variant that all 10 patients had in common was the mt 4336 A>G change.

Recently Zaki showed significant association between the mitochondrial variants mt 16519 C>T and mt 3010 G>A in haplogroup H cyclic vomiting syndrome patients compared to controls. The first variant identified, mt 16519 C>T, was further found to be significantly associated with migraine, with 26 % of migraineurs carrying the variant compared to 1.6 % of controls. The corresponding odds ratio of 15 illustrates a very high disease association [250]. The chronic vomiting syndrome cohort consisted of children recruited from the chronic vomiting syndrome (CVS) database of North America, whereas the migraine cohort was made up of adult affected migraine patients. A further study found that mt 16519 C>T and mt 3010 G>A are only associated with CVS in paediatric onset cases, with adult onset cases showing no association [251].

Given the biological plausibility and presented evidence it is likely that other mitochondrial variants play a role in migraine susceptibility. Many studies have been limited by sample size and future studies which comprehensively assess mitochondrial variation in combination with

haplogroup status would be useful in advancing this field. It has been shown previously that entire haplogroups may harbour mutations which are disease causing. Majamaa found that the U mtDNA haplogroup constitutes a risk genome for migraine associated stroke. It was found that 83 % of Majamaa's patient cohort had the U haplogroup [45]. Interestingly haplogroup U is more than 30 fold more common among the Finnish population as found by Finnila *et al* than for any other European population [252]. Further examination of haplogroup status and migraine susceptibility may prove useful.

Norfolk Island Population

Following a mutiny aboard the British Royal Navy ship HMS Bounty on 28 April 1789, mutineers settled Pitcairn Island in an effort to avoid detection by the British Navy [71]. When the population grew too large for this small island 193 people from Pitcairn Island, all descended from 9 'Bounty mutineers' and 12 Tahitian women, moved to the uninhabited Norfolk Island in 1856 [72]. Due to geographical isolation this population has largely remained a genetically isolated population making it an ideal population for the study of complex multi-factorial diseases such as cardiovascular disease and migraine [73].

Bellis conducted a study which involved recruiting 602 individuals participating in the Norfolk Island Health Study and obtaining consent for participation in genetic research studies. Ethical clearance was initially obtained from the Griffith University Human Research Ethics Committee for the study and subsequently ethical clearance was granted by the Queensland University of Technology (QUT). Blood samples were collected from each individual and analysed for plasma chemistry and additional samples were stored at -80 ^{0}C

for subsequent DNA extraction [253]. Phenotype information was also collected for each individual including anthropometric measures such as weight, height, waist circumference, hip, body fat and blood pressure.

In addition to this information participants were required to fill out detailed medical history questionnaires and provide information about family medical histories, lifestyle choices and existing medical history. Multiple complex disorders were assessed for each individual including cardiovascular disease and migraine. Each participant was allocated a number which corresponds to medical questionnaires, blood samples, DNA and other relevant information pertaining to each individual. A database was set up containing all phenotype information for each study participant which matches the respective DNA sample [72]. Since this original study additional participants have been recruited and included in the Norfolk Island DNA population at the Genomics Research Centre (GRC) situated within the Institute for Health and Biomedical Innovation (IHBI), Queensland University of Technology (QUT). For many samples multiple collections have also been conducted allowing for retrospective studies.

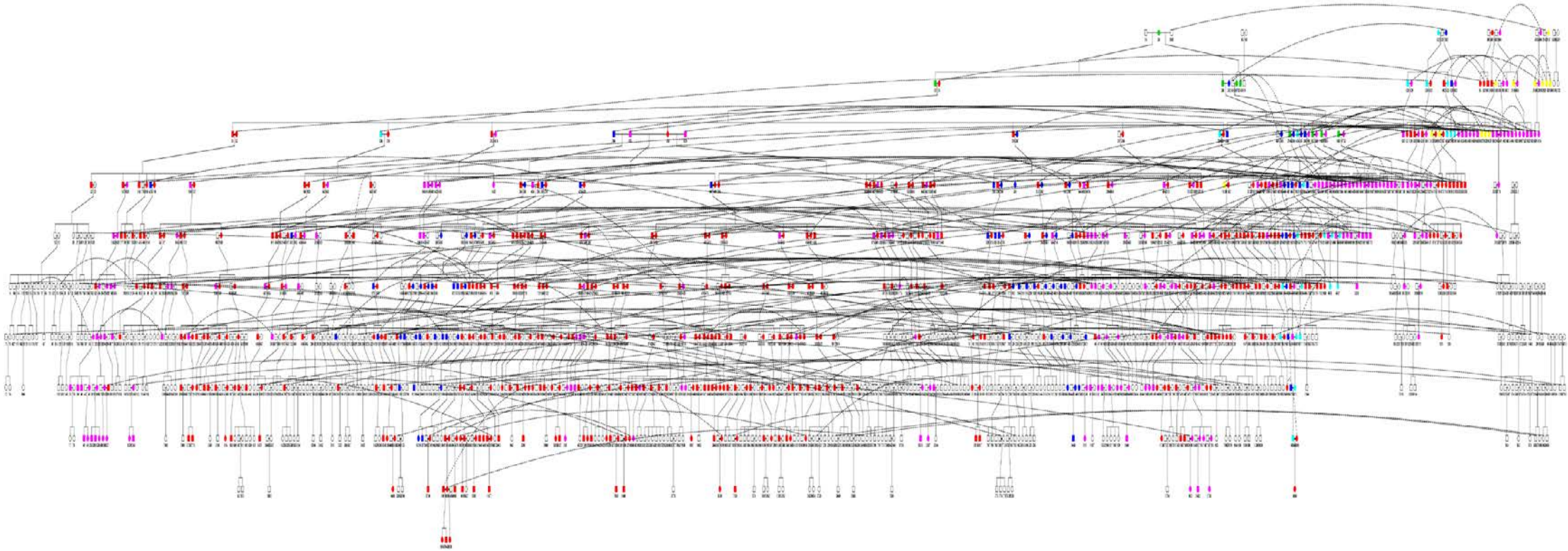
Pedigree Structure

Extensive family histories have been documented and maintained by Norfolk Island inhabitants and more recently genealogists. A database exists which contains detailed information about almost the entire Norfolk Island population and in 2005 the pedigree included 6379 individuals comprising 2185 families dating back to the original founders over an 11 generation period [71]. The database is maintained in a popular genealogy program

Brother's Keeper (Version 6.0, Rockford, Mich, USA) [74]. Other studies involving analysis of complex pedigrees found statistical analysis to be challenging with demanding computer resources required [254]. Due to the many inbreeding loops found in Norfolk's early generations along with the size and complexity of the pedigree it has necessitated the need to split the large 6500 individual pedigree into smaller branches. A core pedigree containing the lower generations and consisting of members originating directly from the population founders has been constructed using a peeling algorithm in the pedigree database management system PEDSYS. This pedigree has been used in previous studies examining risk traits for complex disease [75, 76].

The most up to date core pedigree structure includes 1388 individuals as can be seen in Figure 8. In this figure the original Polynesian mtDNA lineages are coloured and the filled symbols represent migraine sufferers [77]. Contained within this pedigree is a large migraine affected family (n=21 individuals, 7 migraine sufferers) which is of particular interest because the family members are expected to have Polynesian mtDNA haplotypes inherited from the original founders and the family spans four generations.

Figure 8: Norfolk Island Core Pedigree Structure



The large and complex Norfolk Island pedigree is depicted above with individuals represented by coloured pixels. Relationships between individuals within the pedigree are shown by connecting lines. As can be seen this is a multi-generational, large and complex pedigree originating from a small number of founding individuals. Multiple connections between individuals demonstrate the high degree of relatedness within the pedigree and clearly demonstrate a unique genetic admixture with reduced genetic heterogeneity.

Heritability

Heritability is the proportion of a trait or disease phenotype which can be attributed to genetic variation. The official definition of heritability is the "proportion of phenotypic variation (V_P) that is due to variation in genetic values (V_G)." Genetic values (V_G) include the combined effect of all loci as well as interactions within (dominance) and between (epistasis) loci. Two different basic heritability values can be calculated namely broad-sense and narrow-sense heritability. Broad-sense heritability, or H^2 is defined as the proportion of phenotypic variation due to genetic values which include effects of dominance and epistasis ($H^2 = V_G/V_P$) while narrow-sense heritability only considers genetic variation due to additive genetic values ($h^2 = V_A/V_P$) [74].

For human diseases and other complex traits heritability can be estimated from the concordance rate between monozygotic and dizygotic twins [64]. More complex models which examine the correlation of offspring and parental phenotypes can be used to estimate heritability via complex statistical methods. These are employed when phenotypic measures are available on individuals with a mixture of relationships, both within and across multiple generations, or when there are unequal numbers of observations per family [65]. Heritability is a population parameter and is population specific which means that one value cannot be used to predict heritability across all population groups. Population specific factors such as allele frequencies, the effects of gene variants, and variation due to environmental factors effect heritability and as a result heritability needs to be calculated for each specific pedigree [255].

A number of twin studies in migraine have shown that the heritability of migraine ranges from 40-60 % and that the contribution of non-shared environmental factors is 35-55 %. However, little significance has been found between migraine phenotype and shared environmental factors [66-68]. Heritability for migraine as well as for cardiovascular disease and metabolic syndrome has been shown to be significantly higher in the Norfolk Island population than in the local outbred Australian Caucasian population. This makes genetic studies utilising the Norfolk DNA stocks valuable in identification of disease causative variants.

Epigenetics

Epigenetics refers to partially heritable alterations which influence gene expression, not due to changes in DNA sequence but rather as a result of higher structural modifications. There are three main systems involved in epigenetic structuring namely: methylation, histone modification and RNA-associated silencing [256]. It is well known that epigenetics plays a crucial role in gene regulation, growth and especially in development [257]. Any alterations in epigenetic state may result in inappropriate expression or silencing of genes and consequently lead to a myriad of downstream problems. These alterations are often chemical additions or removals of methyl or acetyl groups that result in a change to the chromatin conformation and consequently a change in gene expression. Alterations can be caused by either environmental factors or trigged by ageing processes. Research has shown that epigenetics plays a crucial role in the development of many cancers, but more recently the role of epigenetics in the development of other complex diseases including migraine has become an emerging topic [258, 259].

Methylation

Methylation is the main form of epigenetic modification and occurs by the addition of a methyl group to the 5' cytosine of CpG groups. This chemical reaction is initiated and maintained by DNA methyl transferases. The addition of methyl groups promotes a closed chromatin conformation and prevents transcription from taking place. The mechanistic action blocking the binding of transcription factors and other regulatory sequences such as enhancers occurs through the recruitment of MECP2 proteins which bind to methylated cytosines and attract histone de-acetylases (HDACs). HDACs in turn function to promote a closed chromatin formation. This closed conformation also prevents the transcriptome, a critical complex involved in DNA polymerase binding, from associating with the promoter site which further down regulates transcription. In this way methylation in conjunction with other intricate pathways partly controls gene expression [260-262].

Histone Modifications

Histones form the basis of chromatin modelling, a process which is intricately linked with gene regulation. Chemical modifications which alter histone structure include acetylation, ubiquination, phosphorylation, sumoylation and methylation. Acetylation occurs when the H atoms of the free amino group on particular lysine residues at the N terminus of histone molecules are substituted with acetyl groups, thus reducing the positive charge of the histone and allowing the chromatin to unwind. DNA is negatively charged due to the phosphate group constituting the phosphate backbone and therefore when in close proximity to positively charged histones, the DNA is strongly attracted to the charge and is kept in a tight conformation. The reduction of this positive charge by the addition of acetyl groups allows the DNA to take on a looser conformation and thus makes it more accessible to transcription [263].

These reactions are catalysed by histone acetyl transferase enzymes. Activators and inhibitors are also known to be involved in gene regulation by manipulating these chemical reactions. Histones can be acetylated by activators promoting chromatin structure from a closed to an open configuration. The reverse can occur where inhibitors direct deacetylation of histones with the opposite effect.

Ubiquitination is another type of chemical modification that can alter histone structure. Ubiquitin, a small protein consisting of 76 amino acids forms an internal branched chain with lysine of histone 2A in position 119. This reduces the positive charge of the histone, which promotes a more open conformation and allows more transcription to take place. Phosphorylation is yet another modification which produces an open chromatin structure and allows transcription to take place. The negative charges of the phosphate groups make the histones less positive, and repel the negatively charged DNA [264].

Combinations of chemical changes can also affect chromatin structure such as phosphorylation of histone residue H3 on serine 10 which promotes acetylation on the adjacent lysine 14 residue. De-methylation of lysine 9 is accompanied by phosphorylation of serine 10 and acetylation of lysine 14 which produces an open chromatin conformation. All of these modifications play a major role in genome regulation and are important for normal cellular functions like X chromosome inactivation and genomic imprinting. It has been contemplated that a histone code exists which can be read and interpreted by different cellular factors. Recent research is placing more emphasis on the role that these processes play not just in normal cell functioning, but also in the development of complex diseases [259, 263, 265].

RNA Associated Silencing

It has been experimentally proven that transcription is widespread throughout the genome and is many fold higher than originally predicted. It is now clear that non-coding transcripts are abundant throughout many cell types and that they play an essential role in gene regulation. Non-coding RNAs are divided into several sub-classes and have been associated with important functions, most of which are regulatory in nature. Functions which have been associated with ncRNAs include transcriptional activation, gene silencing, imprinting, dosage compensation, translational silencing, modulation of protein function and binding as riboswitches to regulatory metabolites. More specifically it has been well established that anti-sense RNAs can bind to complementary regions in the genome, mediating RNA degradation and effectively silencing gene expression. RNA associated silencing and the role that this regulatory mechanism plays in disease is becoming an important research focus [266, 267].

Epigenetic therapy

The initial thinking behind developing epigenetic therapies is that if it is possible to chemically manipulate factors such as methylation, acetylation etc., then it may be possible to alter regions where aberrant changes have taken place in order to try and restore the original state. Many agents capable of altering both methylation and acetylation have been discovered, and the applications of these are currently being tested [256]. Agents include 5-azatine, 5-aza-2-deoxycytidine, procainamide, and tea extracts. The compounds 5-azacytine and 5-aza-2-deoxycytidine were initially used as cytotoxic agents, but their dual functionality has since been exploited by use in inhibition of methylation [268]. By inhibiting methylation in regions where aberrant hyper methylation has occurred, appropriate gene expression may be restored

[269]. The mechanism by which this happens is the conversion of nucleoside analogues to deoxynucleotide triphosphates which are substituted into replicating DNA in place of cytosine. DNA methyltransferases get trapped on DNA containing modified bases such as azacytosine, 5-fluorocytosine, pseudoisocytosine or zebularine, resulting in the formation of heritably de-methylated DNA.

Targeting epigenetic mechanisms other than methylation could also be an avenue to pursuing novel drug targets. By understanding the importance of acetylation in epigenetic modification, it then follows that this modification process is also a good target to manipulate epigenetic status. The association between silencing and histone deacetylation, which is catalysed by histone de-acetylases (HDACs), has been well established. A growing number of small molecules have been designed to inhibit HDACs and thereby activate gene expression in regions where aberrant silencing has taken place.

An immediate and logical potential problem associated with taking such an approach is the cytotoxic properties of the above mentioned compounds. However clinical trials indicate that low doses of these compounds may procure benefits which far surpass any cytotoxic effects. A new trial has shown that myeloid dysplastic syndrome and other leukaemias can be somewhat effectively treated using this approach [270]. It may even be possible to make use of tea and sponge extracts in place of harsh chemicals to reverse methylation [271]. Current trials are underway to test the efficacy of such compounds.

Different HDAC inhibitors are being used intravenously or orally in several phase I and II cancer clinical trials, in which changes in histone acetylation have been documented. Depending on the outcomes of this research, it may be that epigenetic therapy could become the novel therapeutic so desperately needed for the treatment of many complex diseases,

including migraine. However some caution must also be used with these approaches as treatment must be specific to the target gene region. Global de-methylation would be far more detrimental than any benefits of treatment. It will therefore be important in the future to design therapies which can target individual enzymes/ genes and thus increase the precision of this type of approach. Despite the promise of miracles, there remains much work to be done before epigenetic therapies can become mainstream treatment plans [256, 272].

Mitochondrial Methylation

Research interests in the relationship between nuclear DNA methylation, environmental exposures and disease outcome are well established. Epigenetic profiling has already become integrated into clinical practise for early diagnosis of cancer and as a molecular tool for determining cancer stages [273, 274]. Bisulphite sequencing and methylated DNA immune precipitation in peripheral blood have been used to demonstrate the presence of methylated cytosines in the human D-loop of mtDNA, proving that the mitochondrial genome is methylated. For the first time, it has been shown that there is a difference in mtDNA epigenetic status between healthy controls and those with disease, especially for neurodegenerative and age related conditions [275-277]. Changes in the level of 5-methyl cytosine have been detected in mitochondria isolated from neurons of patients with amyotrophic lateral sclerosis (ALS) compared to healthy controls [278]. While still remaining to be investigated it has been suggested that mtDNA methylation could play an important role in the aetiology of Alzheimer's disease, Parkinson's disease and dementia [277, 279]. It may not be such a stretch of the imagination then to conclude that investigating mitochondrial methylation in relation to migraine could be a novel and useful avenue.

A Multi-layered Approach

For complex diseases where both a genetic and environmental component play an integral role in pathogenesis, it is becoming more important to develop models which factor in both of these components. Epigenetic changes which are so heavily influenced by the environment have been intricately studied in the nuclear genome. Recent evidence directly suggests a link between nuclear epigenetic changes and migraine and indirectly suggests that the emerging field of mitochondrial methylation could provide a piece of the answer for the complex question: what causes migraine? The systems biology approach, where multiple layers of information are integrated is becoming more important in complex disease modelling. Merging genomic, epigenetic, transcriptomic, proteomic and metabolomic data in order to provide a complete model is becoming a focus in biomedical research. Creative thinking and new approaches are needed to develop better treatment strategies for diseases such as migraine which have such a profound personal and economic impact.

Chapter 3: Investigation of Mitochondrial Variants by RFLP

Chapter 3 outlines initial genotyping experiments carried out in an outbred migraine case-control population investigating mitochondrial variants in relation to migraine susceptibility. This initial work was critical in terms of molecular genetic training, providing the basic knowledge and skills to continue with more complex aims and remains unpublished. This study follows on from a published study by Finnila *et al* which identified two variants within the mitochondrial genome which were found to be associated with migraine susceptibility. Given the limited sample size presented by Finnila *et al*, the aim of this study was to provide a more conclusive link between these variants and migraine by utilising a large migraine case-control population with sufficient statistical power to detect variants with a small to moderate effect on disease susceptibility.

Introduction

Migraine is a common neurological disorder characterised by debilitating head pain and an assortment of additional symptoms which can include nausea, emesis, photophobia, phonophobia and occasionally visual sensory disturbances. Migraine is a complex disease caused by an interplay between predisposing genetic variants and environmental factors. Genes involved in neurological, vascular or hormonal pathways have all been implicated to play a role in predisposition towards developing migraine. All of these are nuclear encoded genes, but given the role of mitochondria in a number of neurological disorders and in energy production it is possible that mitochondrial variants may play a role in the pathogenesis of this disease.

Few variants in the mitochondrial genome have so far been investigated in migraine, however one mitochondrial variant, an A to G substitution occurring at position 4336 of the mitochondrial genome has been associated with migraine in a previous study investigating a variety of disorders in relation to mitochondrial sequence changes [280]. The aim of this study was to further investigate this variant and determine whether the mt4336A>G is significantly associated with migraine in a large Australian migraine case-control population. The mt 4336 A>G variant is found within the tRNA Q gene and is transcribed to produce the transfer RNA for glutamine. Any change in the sequence could have downstream effects on translation of mRNA to protein and may cause disease symtpoms. Studies have found that the G allele significantly increases an individual's risk for developing both Alzheimer's and Parkinson's disease, making this an interesting marker to study in other neurological diseases such as migraine [211, 212, 281-283].

The second SNP selected for this study was a T>C substitution occurring at position 4216 of the mitochondrial genome. Mt 4216 T>C is used as a marker to identify individuals belonging to haplogroups T and J. Given that haplogroups have been shown to harbour certain disease mutations, it may be valuable to test for association between haplogroups T and J and migraine [45]. Research has found that haplogroup J variants have a functional effect and are not just markers of human evolution. Individuals possessing haplogroup J have partially uncoupled OXPHOS which reduces the efficacy of ATP output in favour of heat production. As a result of this pre-existing tendancy, individuals belonging to haplogroup J who have very mild complex I mtDNA missense mutations have exacerbated symptoms associated with LHON compared to individuals belonging to other haplogroups [284].

Materials and Methods

Sample Selection

Migraine cases and controls were recruited for the local South East Queensland region as previously described [162]. All collected samples were of Caucasian origin, and diagnosed as having MA or MO based on criteria specified by the International Headache Society. An unaffected control group with no family history of migraine was matched for age (+/- 5 years), sex and ethnicity. Blood samples obtained from patients were collected through the Genomics Research Centre (GRC) patient clinic (Ethics Approval Number 1300000484). This study was approved by the QUT ethics committee for experimental work on human samples. Two independent case control populations were genotyped for two mitochondrial SNPs.

Molecular Analysis

DNA was extracted and purified from blood using standard salting out procedures. A restriction fragment length polymorphism (RFLP) assay was then used to genotype each sample for the mitochondrial variant. First DNA was amplified by polymerase chain reaction (PCR) using forward 5' GATTCCGCTACGACCAACTC 3' and reverse 5' GCACGGAGAATTTTGGATTC 3' primers. A final reaction volume of 20 uL containing 4 uL buffer, 0.4 uL of forward and reverse primers, 0.4 uL dNTPs, 1.4 uL $MgCl_2$, 0.2 uL Amplitaq Gold and 11.2 uL water was made up. Final reaction concentrations are given in Table 4 below.

Table 4: Reaction conditions for PCR

Reagent	Stock concentration	Final reaction concentration	1x Volume
Buffer	5x	1x	4 uL
Primers	5uM	100nM	0.4 uL
dNTPs	5uM	100nM	0.4 uL
$MgCl_2$	25mM	1.75mM	1.4 uL
AmpliTaq	5 U/uL	1 U	0.2 uL

Cycling conditions were as follows: 94 ^{0}C for 10 min followed by thirty cycles of 94 ^{0}C for 1 min, 55 ^{0}C for 1 min and 72 ^{0}C for 1 min followed by a final extension at 72 ^{0}C for 10 min. An 8 h digest using 6 U of the restriction enzyme *NlaIII* and 7 uL of PCR product made up to a final volume of 12 uL with water was then performed. RFLP products were electrophoresed on a 4 % (w/v) agarose gel at 5 V/cm and viewed under UV light. *NlaIII* cuts at the recognition site CATG and will produce different size fragments when either mt 4336 A>G

or mt 4216 T>C is present. For each SNP, when the mutant allele is present the recognition cut site for *NlaIII* is introduced as shown in red on the FASTA sequence below. Each genotype was confirmed by sequencing using an AB 3130 genetic analyser using BigDye® Terminator v3.1 chemistry (Catalogue Number 4337456).

FASTA Sequence

GCATTACTTATATGA**[T/C]ATG**TCTCCATACCCATTACAATCTCCAGCATTCCCCC
TCAAACCTAAGAAATATGTCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAG
GAGCTTAAACCCCCTTATTTCTAGGAC**CAT[A/G]**AGAATCGAACCCATCCCTGAG
AATCCAAAATTCTCCGTG

Statistical Analysis

The genotype counting method was used to determine the allele frequencies in both case and control populations. A Chi Square analysis was performed to test for significant allele frequency differences between the case and control population and to determine whether alleles were significantly associated with migraine. A Fisher's exact test was also utilised for the mt 4336 A>G change to correct for low detection rate. A significance threshold of $p<0.05$ was used. Secondary protein folding structure prediction software was used to predict any folding changes caused by the mt 4336 A>G substitution.

Results

When the wild type allele is present for both SNPs a single 323 base pair band can be seen, when mt 4336 A>G is present a 209 and 119 base pair band can be seen and when mt 4216 T>C is present a 249 and 89 base pair band can be seen as shown in Figure 11. All samples containing the mt 4336 A>G nucleotide substitution were sequenced as shown in Figure 9.

Migraine Population 1

mt 4336 A>G

The nucleotide change at position 4336 of the mitochondrial genome was found in 1 case and 7 controls (Table 5). Pearson's Chi Square test indicated a significant difference between these values, with a p value of p=0.0031. When a Fisher's exact test was used to adjust for small detection rate a p value of p=0.0367 was obtained. The odds ratio at a 95 % confidence interval was 0.14, implying that the G allele has a protective function.

mt 4216 T>C

This SNP is a diagnostic SNP used in haplogroup analysis. The presence of a 4216 T>C change indicates that the person belongs to either haplogroup T or J and is of European ancestry [285]. It was found that 23.1 % of cases and 21.0 % of controls belonged to haplogroup T or J. The Chi Square test indicated that there was no significant difference between cases and controls (p=0.3906). A summary of the above results can be seen in Table 5 and 6.

Table 5: Summary of Genotyping results for mt 4336 A>G in migraine population 1

	Number of Samples Detected in		Allele Frequency	
	A allele	G allele	A	G
Cases	298	1	99.7 %	0.3 %
Controls	302	7	97.7 %	2.3 %
				p Value=0.0031

Table 6: Summary of Genotyping results for mt 4216 T>C in migraine population 1

	Number of Samples Detected in		Allele Frequency	
	T allele	C allele	T	C
Cases	230	69	76.9 %	23.1 %
Controls	244	65	79.0 %	21.0 %
				P Value=0.3906

Migraine Population 2

mt 4336 A>G

When we repeated our study in an independent population we found no significant association between mt 4336 A>G and migraine susceptibility. A Chi Square analysis indicated no significant differences observed between the case and control sample frequencies (p=0.3148) in the second population. A Fisher's exact test gave a p value of 0.3263 which also falls below the significance threshold.

mt 4216 T>C

A similar distribution of alleles was found in both patient cohorts. Chi Square analysis revealed no significant differences between allele frequency distributions with a p value of p=0.0714. Tables 7 and 8 summarise the results from our second population.

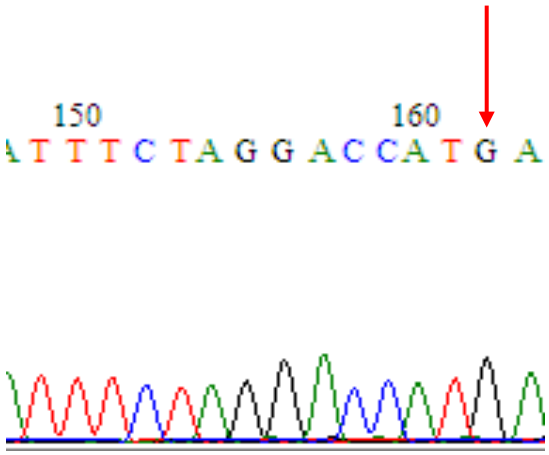
Table 7: Summary of Genotyping results for mt 4336 A>G in migraine population 2

	Number of Samples Detected in		Allele Frequency	
	A allele	G allele	A	G
Cases	230	2	99.1 %	0.9 %
Controls	250	3	98.8 %	1.2 %
				p Value=0.3263

Table 8: Summary of Genotyping results for mt 4216 T>C in migraine population 2

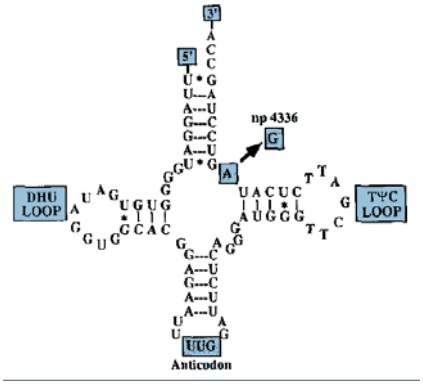
	Number of Samples Detected in		Allele Frequency	
	T allele	C allele	T	C
Cases	175	57	75.4 %	24.6 %
Controls	203	50	80.2 %	19.8 %
				P Value=0.0714

Figure 9: Validation of the 4336A>G mitochondrial variant in a positive control



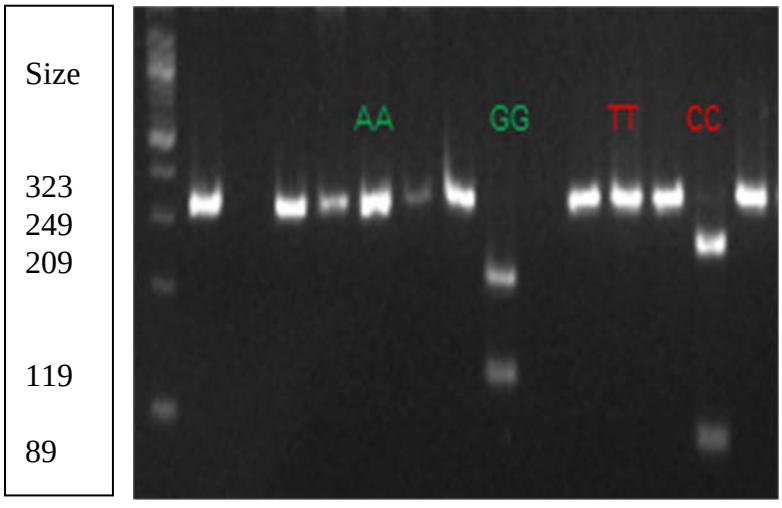
The Sanger sequence trace/electropherogram illustrates that the G allele is present in samples with the mt 4336 A>G variant and confirms the PCR/RFLP assay for the variant.

Figure 10: Position of nucleotide change in tRNA Q molecule



Predicted secondary structure of the tRNA molecule, showing no predicted change when A is substituted to G at position 4336. The overall secondary structure remains intact with all functional domains preserved.

Figure 11: The gel below shows genotype analysis of mt variants by size fractionation of *Nla*III digested PCR products from migraine case/control cohort. Genotypes are visualised on 4 % (w/v) agarose gel and can be determined according to differential banding patterns.



Banding patterns on a 4 % (w/v) agarose gel indicate the genotype of each sample. Every lane on the gel represents a unique sample. Wildtype and mutant banding patterns are shown by the green and red letters for both SNPs mt 4336 A>G and mt 4216 T>C respectively. When mt 4336 A>G is present a 209 and 119 base pair band can be seen and when mt 4216 T>C is present a 249 and 89 base pair band can be seen. A 100 base pair ladder is used as a size standard in the first lane.

Discussion

Limitations of mitochondrial studies which may distort the results include heteroplasmic conditions where mutant mtDNA may be present in undetectable amounts in blood cells. Mutated DNA may also be present in sufficient amounts only in non-actively dividing cells such as neural tissue which can't be examined for ethical reasons [286-289]. Therefore negative results obtained using blood samples may not be true negative results at all. Furthermore genes which are nuclear encoded may be the ones having an effect on mitochondrial function. Many proteins are transported into mitochondria and are not manufactured there. It may be necessary to study a group of nuclear encoded genes in addition to the mitochondrial genome in order to investigate the role of mitochondrial variants in migraine [290-292].

The mt 4336 A>G variant is found within the tRNA Q gene and is transcribed to produce the transfer RNA for glutamine. Nucleotide 4336 is positioned in the region connecting the amino acid acceptor stem with the TψC stem and has been found to be moderately conserved across vertebrates [211, 283]. When the substitution is present, RNA secondary structure prediction software utilised by the mt SNP database [293] indicates that no structural changes occur as can be seen in Figure 10. As with all bioinformatic applications, this outcome is predictive only and is thus limited. Functional studies would need to be undertaken to investigate potential translational effects of this variant.

Studies which have examined the role that mt 4336 A>G may play in contributing towards risk of developing Alzheimer's and/or Parkinson's disease are suggestive that the G allele is mildly deleterious. Despite some negative results, the majority of studies have found mt 4336 A>G to be positively associated with both Alzheimer's and Parkinson's [294, 295]. A meta-

analysis of all previous studies has indicated that the G allele does in fact significantly increase an individual's risk for developing both Alzheimer's and Parkinson's disease [211, 212, 281-283]. However a recent full mitochondrial genome sequencing project has suggested that when considered as a whole the contribution of rare mitochondrial variants towards developing Alzheimer's or Parkinson's is not significant and that *APOE* remains the single high risk gene [296]. The study further states that some mitochondrial variants may actually be protective.

The only study which has previously examined the potential role of mt 4336 A>G in contributing towards the risk of developing migraine, rather than a neurodegenerative disease, found the G allele to be positively associated with migraine and sensineuronal hearing loss [249]. The minor allele frequency of the mt 4336 A>G variant is very low, which means that a very large sample size is needed to accurately genotype a select population. Skewing of allele frequencies in smaller cohorts due to chance can occur easily [297]. Allele frequencies of the minor allele in healthy Caucasian control populations have been estimated to be between 0.63 and 3.8 % depending on sample size and nationality [281-283, 298-302].

When data from eight studies was pooled the average minor allele frequency in a total of 2751 Caucasian subjects at a 95 % confidence interval was 0.98 % [249]. In our study the frequency of the G allele is similar to reports in other healthy control populations thus suggesting that the G allele doesn't play a significant role in migraine susceptibility. A Chi Square test indicated that there is a significant difference between case and control frequencies in one of our population cohorts, but given the lack of significance in our second cohort this may have been a sampling bias due to chance.

Contrary to Finnila *et al*, we found the mt 4336 A>G variant to have a mildly protective effect in one migraine case-control population but no significant effect in the second population genotyped. There may be a number of reasons why our results contradict the results found by Finnila *et al*, which could be attributed to differences in population, study design and analysis or simply insufficient sample size. Our study examined the mt 4336 A>G variant in an exclusive migraine patient cohort, while Finnila *et al* looked at a group of patients with an assortment of diseases simultaneously.

Finnila *et al* studied only 42 patients out of the cohort of 1072 case samples were migraine sufferers, with the rest belonging to other disease groups. The mt 4336 A>G variant was detected in 2 migraine sufferers, while it was found in 3 healthy control individuals. The remaining 12 individuals with the mt 4336 A>G variant belonged to other disease groups. When interpreted in this manner it is clear that the mt 4336 A>G variant was found in more control samples than in migraine case samples. This outcome is consistent with our findings and it is only the analysis of the results which differ. Furthermore we had a much larger sample size of 299 migraine patients in our first cohort and 232 in our second, providing greater statistical power and decreasing the chance of sampling bias. The method of association analysis was not mentioned in Finnila *et al*, so it may be possible that differences in analysis could also produce slightly different results. In our study we performed a Chi square test as well as a Fisher's exact test to correct for low detection rate of mt variants. Patient samples used by Finnila *et al* were all of Finnish origin, while our patient samples comprised a larger selection of European lineages, and thus there may also have been ethnic differences.

The second SNP, mt 4216 T>C, was not significantly associated with migraine in our patient cohort. While many samples were found to belong to haplogroups T and J, this result was expected due to the fact that all samples used in the study are from individuals of European descent. While it is known that entire haplogroups may harbour mutations which are disease causing, haplogroups T and J did not show any significant association in our study. Majamaa *et al* found that the U mtDNA haplogroup constitutes a risk genome for migraine associated stroke. It was found that 83 % of Majamaa's patient cohort had the U haplogroup [45]. Interestingly haplogroup U is greater than 30 fold more common among the Finnish population in which Finnila *et al* conducted the study than for any other European population [252]. This may further explain our contradicting results and attribute the positive association found in Finnila et a to a variant other than mt 4336 A>G located within the U haplogroup.

Conclusion

According to our study in an Australian Caucasian population, the mt 4336 A>G change is not significantly associated with migraine susceptibility, with the mt 4336 A>G change occurring in similar frequencies in case and control individuals. It is unclear if this variant has an effect on final translation and protein function but due to a rare (<1%) minor allele frequency, investigation in a larger patient cohort may produce more accurate frequency data. Our data also showed that there is no association between haplogroups T or J and migraine in Caucasian populations. A study which further analyses the mitochondrial haplogroups, especially haplogroup U may prove useful.

Chapter 4: Use of semiconductor sequencing technology to investigate mitochondrial variants in a subset of the Norfolk Island Core Pedigree

In Chapter 4 the optimisations carried out to develop a cost effective method for full mitochondrial genome sequencing are described as well as the initial pilot project undertaken in n=48 samples. The aim was to test the Ion Torrent technology which was new at the time of undertaking these experiments in order to prove the feasibility of using this approach in a large number of samples. A small sample selection was made in order to conduct a pilot-project before undertaking the main part of the study. Developing a method which was as efficient as possible and highly cost effective was essential before undertaking a large study. This work has not been published yet.

Introduction to Semiconductor Sequencing Technology

The launch of the Human Genome project propelled development of parallel DNA sequencing at an ever faster and more efficient rate. Next Generation Sequencing technologies arising from this era have had a profound impact on research and medicine [303]. Easy and cheap access to large amounts of sequence data has allowed us to progress in the field of molecular medicine. In this short period of time many sequencing applications developed from this explosion of genetic information are already changing the way that cancer and other heritable diseases are diagnosed [304-306]. Human Genetics is not the only field benefitting from the development of better sequencing technologies, with fields as diverse as ecology, conservation and the study of ancient DNA progressing as more information is accumulated [307-309]. Even though sequencing costs have dropped dramatically over the last ten years, there is a continuous demand and desire to continue dropping the cost of sequencing to further democratise the availability of sequence information. The goal has long been to reach the $1000 genome which has sparked much competition between biotechnology companies [310].

To date the most expensive component of sequencing has been orientated around costly nucleotides and other reagents which are excited upon laser exposure to release a light or fluorescent signal which is picked up by complex imaging technologies. This requirement for electromagnetic intermediates in the form of either X-rays or light has limited the reduction in cost of DNA sequencing [311-313]. The recent development of non-optical sequencing has shifted this paradigm and sparked a new cost reduction at an exponential rate consistent with Moore's law. The use of CMOS processes which have already undergone decades of evolution through the computer and cell phone industry have largely made this possible [314]. Using highly scalable integrated circuits to detect the release of hydrogen ions upon

incorporation of natural nucleotides is the concept upon which Ion Torrent technology is based [315].

Jonathon Rothberg leveraged the advances made in recent years in the semiconductor industry to produce an electronic sensor capable of detecting the release of hydrogen ions during sequencing by synthesis [315]. Ceramic beads or ion sphere particles bound to a single strand of template DNA and a complementary sequencing primer which provides the first OH group for the addition of natural nucleotides are loaded into single wells. The addition of DNA polymerase then allows for base pairs complementary to the template DNA to be incorporated through sequential flows of nucleotides [316]. During each known flow, a nucleotide is either incorporated or washed away according to the template sequence. Each time a nucleotide is incorporated, a hydrogen bond is formed between the preceding OH group and the incorporated nucleotide, resulting in the release of an H ion. Since H+ is acidic and lowers the pH of a given solution, this change in pH can be detected using a sophisticated pH meter. This change is then used to directly generate a digital signal which is used to determine the sequence of each template strand [314].

Materials and Methods

In order to sequence either DNA or RNA samples, a lengthy sample preparation procedure must be followed to reach a suitable end product. Starting material varies according to individual application, therefore only DNA sequencing will be included in this discussion with workflows referring specifically to this research project. As little as 10 ng of whole genomic DNA extracted from any tissue of choice is sufficient for this application, however slightly higher DNA inputs are desirable. DNA was extracted from whole blood samples

collected during a previous study using salting out and was suitably stored for research purposes [72].

Sample Selection

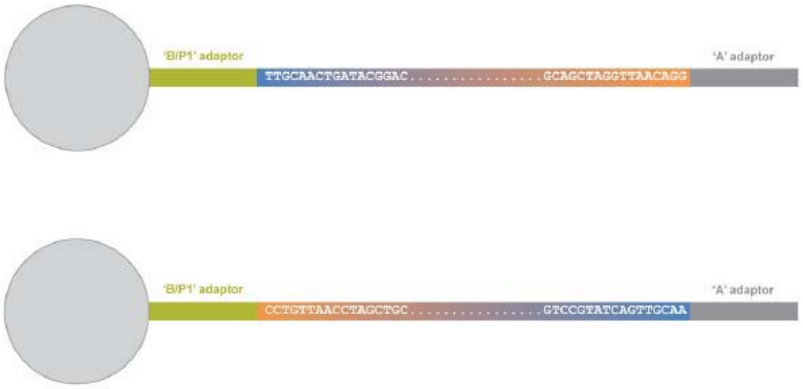
Initially as a pilot project, a subset of 48 individuals were chosen from the Norfolk Island Core Pedigree (n=48) for full mitochondrial genome sequencing. The selection criteria used included selection of only female individuals, selection of individuals from each generation of the extended migraine family with all four individuals being migraine affected, selection of one sample from each original founder Polynesian lineage and selection of samples for which DNA is available within the GRC. Other factors taken into consideration included attempting to obtain an even mixture of migraine to non-migraine affected individuals with the final selection consisting of n=23 migraine (48 %) and n=25 non-migraine (52 %) individuals. Also a filter was applied for selection of larger families where possible and samples were further filtered based on available genotype information from previous studies. In total n=43 (90 %) of individuals with genotype info, and n=35 (73 %) with expression data were chosen. This will allow for correlation between nuclear SNPs and mtDNA.

Ethical clearance was obtained from the Griffith University Human Research Ethics Committee for the collection and utilisation of all DNA samples to be used in this study as required for research on human material.

Molecular Techniques

Any DNA or RNA target of sufficient quality can be sequenced through unique workflows on this platform. Library preparation is purposed to prepare your target sequence of choice such that by the end of this workflow single strands of appropriately sized target regions ligated to adapter sequences are ready to be attached to ion sphere particles (ISPs). During emulsion amplification the aim is for single template strands to attach to a single ISP. The purposes of the adapter sequences on either side of the target templates are to aid in positive selection of ISPs as well as to provide unique identifiers for each sample allowing for multiplexing. The goal following positive selection of ISPs is for a single ISP, attached to a single template strand, to be loaded into a single well of a sequencing chip. Each well sits above a single sensor plate capable of detecting small pH changes. As discussed in the introduction, when H^+ ions are release after the incorporation of a natural nucleotide, the pH decreases due to the acidic nature of H^+. This drop in pH results in a voltage change which is in turn directly interpreted as a digital signal or ionogram.

Figure 12: Following library preparation, target sequences of an appropriate size (200bp-400bp depending on sequencing chemistry) should have adapters ligated onto each end. During template preparation library products attach to ISPs.



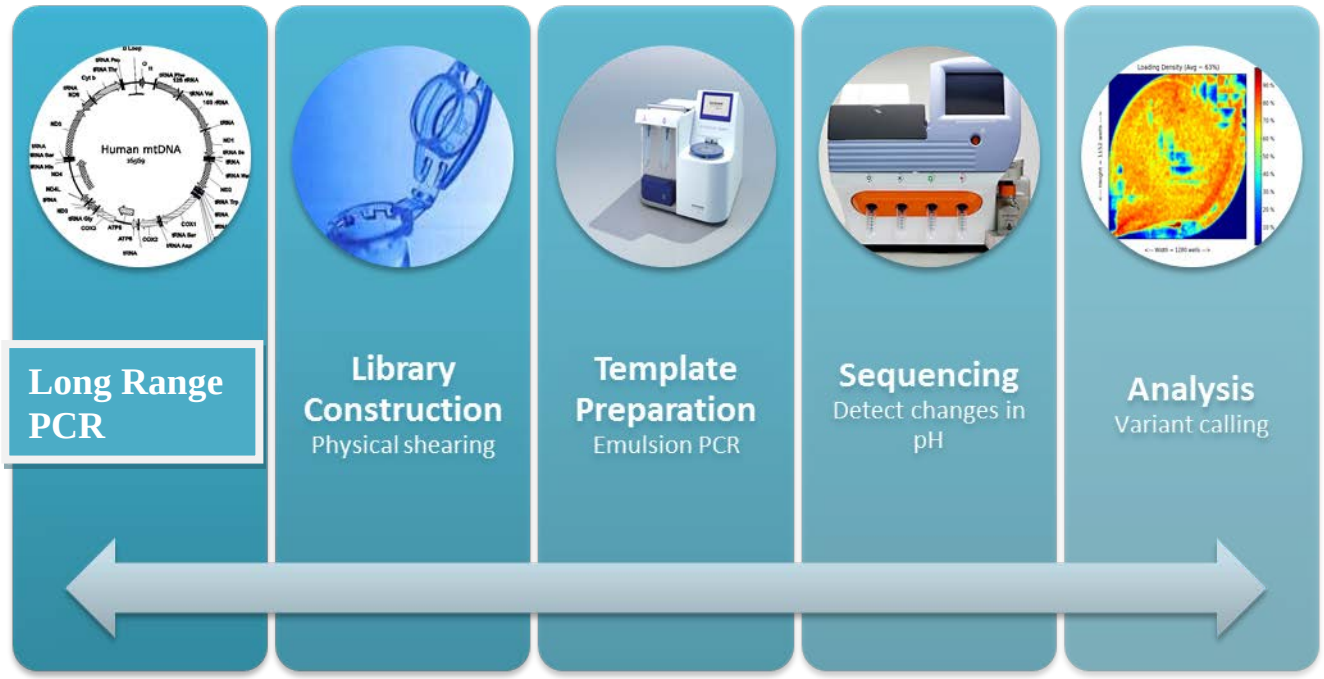
Ceramic ion sphere particles shown in grey, should ideally be ligated to a single strand of template following emulsion amplification. P1 adapters are shown in green and provide the link between ISPs and the target fragment of DNA to be sequenced. A adapters are represented by solid grey bars and contain biotin to facilitate positive selection of ISPs which are successfully ligated to template DNA.

As shown in Figure 12 above, each template strand is sheared and size selected to the correct size. Adapters are then ligated onto each end. The P1 adapter can also contain a unique barcode made up of DNA sequence which provides a label specific to a given sample. The P1 adapter also contains a library key sequence, complementary to a sequencing primer which provides the necessary OH to initiate sequencing by synthesis. The A adapter contains a biotin label which facilitates selection of only those ISPs which are attached to template through the use of streptavidin beads.

Overall Workflow

Following long range amplification using two overlapping primers, samples are quantified and then undergo library prep, followed by emulsion PCR and sequencing. The overall workflow is illustrated in Figure 13 below.

Figure 13: Overall workflow to conduct full mitochondrial genome sequencing on the Ion Torrent Platform



Samples underwent long range PCR, using overlapping primer pairs which allowed for full mitochondrial genome coverage. Following purification, mitochondrial fragments underwent library preparation and emulsion amplification before undergoing sequencing on the Ion Torrent platform. Sequence data was then aligned to a reference genome and variants were called relative to the reference sequence.

Long Range PCR

Samples were amplified by long range PCR, utilising two primer pairs which produce overlapping fragments covering the entire mitochondrial genome. Primer sequences as well as basic thermocycling conditions were obtained from Dr Richard Allcock at the University of Western Australia. PCR conditions were then optimised to maximise non-specific amplification. Parameters which were tested and accordingly altered include final primer concentration, DMSO concentration, final reaction volume, input DNA, annealing temperature and thermocycling conditions. We found that a higher DNA input (100 ng starting material) and larger reaction volumes resulted in better amplification. The optimum DMSO concentration was found to be 3 % and after a number of optimisations a final primer concentration of 200 nM was chosen. Also we found that the most effective way to address non-specific amplification was to break the thermocyling into two sets of cycling, with an initial lower annealing temperature followed by a second set of cycling at a higher temperature. Final reaction conditions are outlined in Table 9, 10 and 11 below.

PCR products were run on 1 % (w/v) agarose gels to accommodate for large fragment sizes at 2 V/cm and visualised under UV light. Ethidium bromide was used for staining at a 4 % concentration. Negative controls were included in each PCR run to check for contamination.

Table 9: Primer Pairs used for long range PCR

Amplicon	Amplicon Position in Genome	Primer Sequences	Final Concentration
mt_Frag1	569 (forward)	5' AAC CAA ACC CCA AAG ACA CC 3'	200 nM
	9819 (reverse)	5' GCC AAT AAT GAC GTG AAG TCC 3'	
mt_Frag2	9611 (forward)	5' TCC CAC TCC TAA ACA CAT CC 3'	200 nM
	626 (reverse)	5' TTT ATG GGG TGA TGT GAG CC 3'	

Table 10: Final Reaction Conditions

Reagent	1x Reaction Volume uL	Final Concentration	Stock Concentration
Nuclease free water	28.3	-	-
Buffer	10	1 x	5 x
dNTPs	2.5	0.5 mM	10 mM
F/R primer	2	200 nM	5 uM
DMSO	1.5	3 %	100 %
Enzyme Mix*	0.7	0.05 U/uL	3.75 U/uL

*Roche Expand Long Range Kit (Catalogue Number 04829042001)

Table 11: Final Thermocycling Conditions

Temperature	Time	Cycles
92 ^{0}C	2 min	1
92 ^{0}C	10 s	9
55 ^{0}C	15 s	
68 ^{0}C	9 min	
92 ^{0}C	15 s	19
60 ^{0}C	20 s	
68 ^{0}C	9 min	
68 ^{0}C	7 min	1
8 ^{0}C	∞	

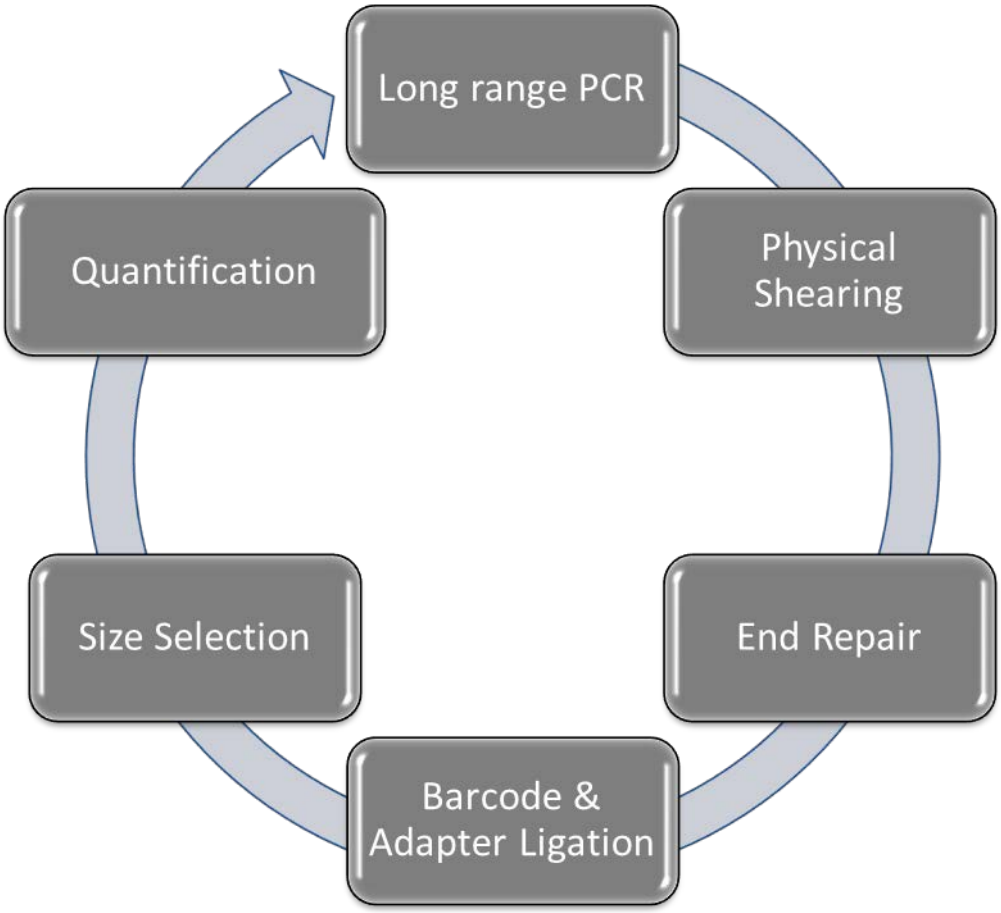
Post PCR

Upon confirmation of successful amplification, samples were cleaned using QIAquick post PCR cleanup columns from QIAGEN (Catalogue Number 28106) to remove any excess dNTPs or other reagents. Once only the pure mitochondrial fragments remained, they were accurately quantified using Agilent DNA 12000 chips (Catalogue Number 5067-1508) on the bioanalyser. At this time point, each sample was divided into two fragments, corresponding with the two primer pairs used. Using the concentration in ng/uL obtained from the bioanalyser runs, the two overlapping fragments produced during the long range PCR process were pooled together in equimolar amounts for each sample. In order to achieve accurate equimolar pooling and thus even representation of each fragment upon DNA sequencing, both fragments were diluted to the same concentration and then equal volumes of each was added to a single tube.

After pooling and thorough mixing, the required amount of input DNA for library preparation was measured out. Initially 100 ng was used for library preparation, but it was found that larger amounts of input DNA were more desirable. 1ug was found to be sufficient for the Life Technologies Ion Express Library Kit (Catalogue Number 4471269) and later one it was found that as little as 250 ng of input DNA was highly effective when utilising NEB ion torrent compatible library kits (Catalogue Number 4474178). The quality of DNA greatly influences the efficacy of library preparation and ultimately sequence output.

Library Preparation

Figure 14: Major steps involved in library preparation



Following long range amplification, samples were purified, fragements were pooled in equimolar amounts to ensure even representation of both sides of the mitochondrial genome and were then sheared using sonication. After physical shearing, fragments were end repaired and ligated to adapter and barcode sequences. Libraries then underwent a size selection process to ensure optimal sequence output and maximum utilisation of reagents at the sequencing stage. Each library was quantified for accurate dilution and pooling thus ensuring even representation between barcoded samples.

Life Technologies Ion Express Library Preparation (used for first 29 samples)

A number of different library preparation kits are commercially available. Each one makes use of similar molecular techniques and only vary slightly. To begin with the Life Technologies Ion Express Library Preparation kit (Catalogue Number 4471269) was used as proof of principal. Library preparation was the part of the workflow where our target products were processed to desirable sequence compatible fragments. This involved a number of detailed steps, each of which will be discussed. Following amplification, products of approximately 9000 base pairs were generated. Fragments of this size are too large to sequence directly on the Ion Torrent platform and were sheared to a target size of 200 base pairs. Both an enzyme shearing method and physical shearing were tested for efficacy.

After shearing, fragments were repaired to produce blunt ends which are suitable for ligation of barcodes and adapter sequences. Adapters are a necessary component of the sequencing chemistry used and provide a library key sequence as well as a preceding nucleotide for Taq to add onto during sequencing by synthesis. The barcodes are unique DNA sequences which allow for multiplexing of samples. After successful adapter ligation, size selection was applied to select for a smaller size range of fragments thus maximising sequencing efficiency. Finally each library is accurately quantified and diluted to maximize template preparation.

Physical shearing with the BioRuptor Sonication System

The biorupter shears DNA by ultrasonic fragmentation and can be used for a wide concentration range of starting material. Shearing time can be adjusted to compensate for DNA concentration and desired fragmentation profile. Since we were making use of 200 base pair sequencing chemistry, samples were sheared for long time periods. The biorupter was operated in a sound proof box to compensate for the high frequency noise produced. 1 ug of

each pooled sample was pipetted into a 1.7 ml safe lock lo bind epindorf tube and carefully placed into the rotor. The lid of each tube was sealed to prevent any water from leaking in. Parameters were set as follows; high intensity, 30 s sonication followed by 30 s cool-down time for 15 cycles. Ice water was added to just above the fill line and replenished with ice after every cycle. Each sample was sheared for 15 cycles, 5 times. Following shearing, the DNA fragmentation profile was assessed using the bioanalyser system and Agilent DNA 1000 chips (Catalogue Number 5067-1504). A fragment size range between 50 and 500 base pairs, with a peak at 200 base pairs is most desirable. Any samples which were fragmented insufficiently underwent a second round of shearing.

End repair and purification of sheared DNA

After checking the fragmentation profile of each sample, they were end-repaired and purified. Materials in the Ion Plus Fragment Kit were used to do this including 5 x end repair buffer and end repair enzyme as well as additional materials including nuclease-free water, LoBind tubes, Agencourt AMPure beads (Catalogue Number A63881) and a magnetic rack. Firstly all components were defrosted and pulse-spun. Also nuclease-free water was added to the sheared DNA to a total final volume of 158 uL. Since the initial pooled samples were aliquoted out according to concentration, the volume of each one differed, hence the need to bring each sample to a uniform volume. Then 40 uL of 5 x end repair buffer and 2 uL of end repair enzyme was added, followed by a 20 min incubation at room temperature.

For the purification step which follows on from end repair, freshly prepared 70 % ethanol was used. A higher percentage of ethanol causes inefficient washing of smaller sized molecules while a lower percentage could cause sample loss. Therefore ethanol was carefully and accurately prepared before each use. Before the ethanol washing step, 360 uL of Agencourt

beads (Catalogue Number A63881) were added to each end repaired sample, pipetted up and down five times and incubated at room temperature for 5 min. Samples were then pulse spun and placed in the sample tube of a DynaMag magnetic rack for 3 min or until the solution cleared completely. The supernatant was removed without disturbing the pellet and discarded. Leaving the open tubes in the magnetic rack, 500 uL of 70 % ethanol was added followed by a 30 s incubation. The tube was turned around 180 ^{0} twice to move the beads around and maximize the ethanol wash. After the solution cleared, the supernatant was removed again. This washing step was repeated a second time. To remove residual ethanol, each tube was closed, pulse spun and placed back on the rack. A smaller 20 uL pipette was then used to remove any remaining supernatant. Keeping the tubes on the magnet, samples were air dried for 5 min. Then tubes were removed from the magnet and 25 uL of low TE was added to each sample and thoroughly mixed. After a further 1 min incubation tubes were placed back on the magnet and this time the clear supernatant containing the eluted DNA was transferred to a new Eppendorf tube.

After purification the first optional stopping point was reached. At this step samples could be stored at -20 ^{0}C before commencing with the next part of library preparation.

Barcode and Adaptor Ligation

Once all sheared samples were repaired to have blunt ends, barcodes and adapters were ligated on. This was done by transferring each sample into a 0.2 ml PCR tube and combining all reagents as indicated in Table 12 below into each tube and mixing well by pipetting up and down

Table 12: Table showing barcode reaction components and volumes

Component	Volume for 1ug input in uL
DNA	25
10 x ligase buffer	10
Ion P1 adapter	10
Ion Express barcode	10
dNTP Mix	2
Nuclease free water	31
DNA ligase	4
Nick repair polymerase	8

After mixing the tubes were placed in a thermal cycler and run according to the following program:

25 ^{0}C 15 min

72 ^{0}C 5 min

4 ^{0}C ∞

Then all 100 uL were transferred to a clean Eppendorf tube and purified using freshly prepared 70 % ethanol and Agencourt beads (Catalogue Number A63881) as described in the section "end repair and purification of sheared DNA". However only 140 uL of beads were used per sample for this clean up. At this point there was another optional stopping point.

Size Selection

The E-gel size select system (Catalogue number G6610-02) was used to size select the unamplified library according to desired read length. For 200 base pair sequencing chemistry the library was selected for a target peak of 330 base pairs. After installing the iBase unit on top of the safe imager transilluminator and connecting the power source, a 2 % (w/v) agarose e-gel was loaded into the iBase unit. The 2 % SizeSelect program was selected from the menu

and the timer was left at the default 12 min setting. Before loading sample onto the gel, 20 uL of low TE buffer was added to each purified, ligated DNA sample, bringing the total volume to 40 uL. Two adjacent top row wells were used for each sample and 20 uL was loaded into each well. The 50 bp ladder was diluted to 25 ng/uL which is a 1 in 40 dilution and 10 uL of diluted ladder was then added into the middle well. All empty wells including the bottom row of wells were filled with 25 uL of nuclease free water.

The amber filter was then placed over the E-gel iBase unit and the Size Select program was initiated by pressing Go. After the 12 min program was complete, the collections wells or bottom row wells were re-filled with 10 uL of sterile water. After re-filling the same program was run again, but this time with constant supervision. As soon as the 350 base pair mark as shown by the ladder reached the top of the collection well, the run was stopped and samples were recovered from the collection wells using a pipette. Wells were refilled with 10 uL of nuclease free water and recovered again to ensure maximum recovery for each sample. Used gels were then disposed of as hazardous waste.

Quantification

Using the bioanalyser each library was quantified using an Agilent DNA 1000 chip (Catalogue Number 5067-1504). Samples were then grouped into pools according to previous barcoding and according to experimental design. It was decided to pool 6 samples on each chip and library preparation was carried out accordingly. For each pool, all samples were diluted to the same concentration and then equal volumes were added to a single tube. From the stock pool, a dilution was made for optimum template preparation. This was calculated according to the manual where:

Template Dilution Factor (TDF) = Library pool concentration in pM/26 pM

After optimising template preparation it was found that for the Life technologies library preparation kit, the best results were obtained from doubling the recommended input library.

NEB Library Preparation Kit (Catalogue Number 4474178)

This kit makes use of the same major steps and principals as described with use of the Life Technologies library preparation kit and was used for all remaining samples. There are some slight variations and for that reason, I will describe the major steps and include reaction setups in this section. Also this process involves an additional compulsory amplification step. The advantage of using this kit is related to the massive cost saving. The generic kits cost only a fraction of the price, enabling more samples to be processed for the same budget. Please refer to the user manual for detailed reaction setup and steps. Approximately 250 ng of input DNA was used for each sample with this protocol. Where applicable the 100 ng protocol was used.

Physical shearing with the BioRuptor Sonification System

Shearing was conducted exactly the same as described in the previous section.

End repair

The following components were mixed in a sterile Eppendorf tube for each sample. Starting material from 100ng to 1ug can be used, with reagent volumes remaining the same regardless of input DNA. Samples were kept on ice throughout.

Table 13: NEB End Repair Protocol

Component	Volume
Fragmented DNA	1-51 uL
NEB End Repair Reaction Buffer	6 uL
NEB End Repair Enzyme Mix	3 uL
Nuclease free water	Variable
Total Volume	60 uL

After mixing with a pipette samples were incubated in a thermal cycler for 25 min at 25 ^{0}C and for a further 10 min at 70 ^{0}C. Then samples were pulse spun and returned to ice, ready for preparation of Adapter Ligated DNA.

Barcode and Adaptor Ligation

The following was added to each tube from the previous step:

Table 14: NEB Protocol for barcode and adapter ligation

Component	Volume
Nuclease free water	16 uL
T4 DNA Ligase Buffer	10 uL
NEBNext DNA Library P1 Adapter	4 uL
Barcode X	4 uL
T4 DNA Ligase	6 uL
Total Volume	40 uL

After adding the above reagents, the volume for each sample was 100 uL. Contents were mixed using a pipette and then incubated in a thermal cycler for 15 min at 16 ^{0}C.

Purification

Here only 180 uL of AMPure beads (Catalogue Number A63881) were added to each sample. A pipette was then used to thoroughly mix samples, followed by a 5 min incubation at room temperature. After a quick pulse spin, each tube was placed in a DynaMag magnetic rack for 1 min or until the solution cleared. The supernatant was removed, without disturbing the pellet and then 400 uL of 80 % ethanol was added while keeping the tubes on the magnetic rack. Each tube was rotated twice, for proper washing of beads. Then supernatant was removed again. This wash step was repeated a second time. Residual ethanol was removed by pulse-spinning each tube, returning the sample to the magnetic rack and using a P20 pipette to remove any remaining ethanol. Samples were air dried on the rack for 5 min and beads were resuspended in 25 uL of nuclease free water. The remaining beads were then spun in the tube, the sample was placed back on the magnet and the supernatant containing the purified sample was collected into a clean tube.

Size Selection

Size selection was carried out as described in the Life Technologies protocol.

PCR Amplification of Adapter Ligated DNA

The following components were mixed in a sterile tube on ice:

Table 15: NEB PCR Amplification Mix

Component	Volume
Adapter Ligated DNA	40 uL
Primers	4 uL
Nuclease free water	6 uL
One Taq Hot Start 2x Master Mix	50 uL
Total Volume	100 uL

Samples were placed on the thermocycler using the following cycling conditions:

Table 16: PCR Cycling Conditions

Step	Temperature	Time
Nick Translation	68 ^{0}C	20 min
Initial Denaturation	94 ^{0}C	30 s
4-8 Cycles*	94 ^{0}C	30 s
	58 ^{0}C	30 s
	68 ^{0}C	1 min
Hold	4 ^{0}C	∞

*Avoid over-amplification to optimize the number of unique molecules. For 100 ng input DNA use 6-8 cycles

Purification

140 uL of AMPure beads (Catalogue Number A63881) were used. All other steps were the same as described in "Purification" above.

Quantification

Same as described in Ion Torrent library section.

Template Preparation

The Ion One Touch instrument was installed and set up according to the user manual. Before each use the amplification plate, reagent oil, recovery solution and plastic ware was set up as described in the user manual. Each time a new kit was used, all plastic ware and reagents were changed to correspond with the kit currently in use. Importantly the waste container was emptied before every run, as any backwash would cause permanent contamination of the entire instrument. An amplification solution was prepared and installed for each sample pool (6 samples). Before setting up the reaction, the Ion OneTouch 2 x Reagent Mix was thawed and subsequently kept at 4 ^{0}C. Also the enzyme mix and pre-prepared library were vortexed just before use. In an Eppendorf tube the following components were added in the designated order:

1. 280 uL Nuclease free water
2. 500 uL Ion OneTouch 2x reagent mix
3. 100 uL Ion OneTouch enzyme mix
4. 20 uL Diluted library pool

The Ion Sphere Particles (ISPs), included in the template preparation kit, were then vortexed at maximum speed for 1 min and immediately 100 uL was added to the 900 uL of amplification solution. The filter assembly was then prepared and installed by using a pipette to add the 1 ml prepared solution and 1.5 ml of OneTouch oil to the reaction filter. The filter was inverted and inserted into the OneTouch as described in the manual. An assisted run was

selected from the dropdown menu and after following each prompt, the run was started. The OneTouch was cleaned according to instruction after each use.

Recovery of template positive ISPs

Both of the recovery tubes were removed immediately after the last centrifugation step in the OneTouch and all but 50 uL of recovery solution was removed using a pipette. Care was taken not to disturb the pellet and the template-positive ISPs were resuspended in the remaining recovery solution. The ISPs were then washed in 1 ml of Ion OneTouch wash solution. At this stage ISPs could be stored for up to 3 days at 4 ^{0}C. Just before sequencing ISPs were centrifuged for 2.5 min at 15500 x g, and all but 100 uL of supernatant removed. The pellet was then resuspended by vortexing. Each sample underwent this process and from here the sequencing part of the protocol was commenced.

Sequencing

The Personal Genome Machine (PGM) has been maintained according to manufacturer's recommendations including weekly chlorite washes when the machine is in use and water washes after each run. Before each run the PGM was initialised in strict accordance with the user manual using a Life Technologies 200 base pair sequencing reagents kit (Catalogue Number 4482006). Before each run, a run plan was generated using the Ion Torrent browser on the server. The run plan specified sequencing settings, number of flows, kit type used in sample preparation, barcodes used and corresponding samples, sample type i.e. DNA and the reference file to be used. We used the revised Cambridge mitochondrial genome as our reference sequence. This allows the server to automatically assemble sequence fragments against this reference sequence. All torrent software was also updated regularly.

Template Positive ISP Preparation

For quality assessment of each sequencing run, 5 uL of control ISPs were added to half of the 100 uL volume of ISPs previously stored in the fridge. The other 50 uL were kept as reserve stores. The “A” and “D” test fragments found in this control mixture provided insight about the quality of sequencing runs and could be found in the run report generated from each run. After adding control fragments, 100 uL of annealing buffer was added and mixed in by pipetting up and down. The tube was centrifuged for 2 min at 15 500 x g and all but 3 uL of supernatant was removed. The sequencing primer was then annealed to each target fragment by adding 3 uL of thawed sequencing primer followed by a 95 ^{0}C 2 min and 37 ^{0}C 2 min thermal cycling step.

314 Chip Loading

The chip was placed on the PGM system via the grounding plate and “experiment” was selected on the main menu to initiate a chip check. Following a successful chip check, the chip was washed once with 50 uL of 100 % isopropanol and twice with 50 uL of annealing buffer. Sequencing polymerase was then added to the ISPS, to bring the total volume to 7 uL. After a 5 min incubation period, ISPs were loaded onto the chip and equally distributed through a series of centrifugations. The run was then started. Each run took approximately 2 and half hours.

Statistical Methods

The FASTQ sequences generated from each run were aligned using bowtie2 and SAMTOOLS making use of the following command prompts:

>bowtie2 -q hs_mtDNA_ref IonXpress_002_R_2012_06_07_23_47_49_user_SN1-9-Shani_Run1_Auto_SN1-9-Shani_Run1_9.fastq,IonXpress_002_R_2012_06_13_23_43_59_user_SN1-11-Shani_Repeat_Run1_Auto_SN1-11-Shani_Repeat_Run1_11.fastq -S Shani_002_align.sam --qc-filter

>samtools view -b -o Shani_002_raw.bam -S Shani_002_align.sam

>samtools sort Shani_002_raw.bam Shani_002_sorted

>samtools index Shani_002_sorted.bam

This part of the process was performed by Miles Benton from our bioinformatics division. The aligned FASTA sequence was then run through MitoTool, using the revised Cambridge sequence as a comparison to each generated sequence. The variants found in each sample were then compared to MitoMap, a database which stores information pertaining to mitochondrial variants which are associated with disease risk.

As an additional analysis undertaken by myself, raw sequence reads were aligned to the RSRS (Revised Cambridge Reference Sequence) mitochondrial genome sequence using default settings on the torrent server. Variants were then called in each sample relative to the reference sequence using a Java based pluggin Variant Caller on the torrent server. All variants were typed into an excel spreadsheet for comparison between case and control samples. Using a filtering process it was noted which variants were present only in cases and not control samples and vice versa. These variants were plotted on a circular plot according to

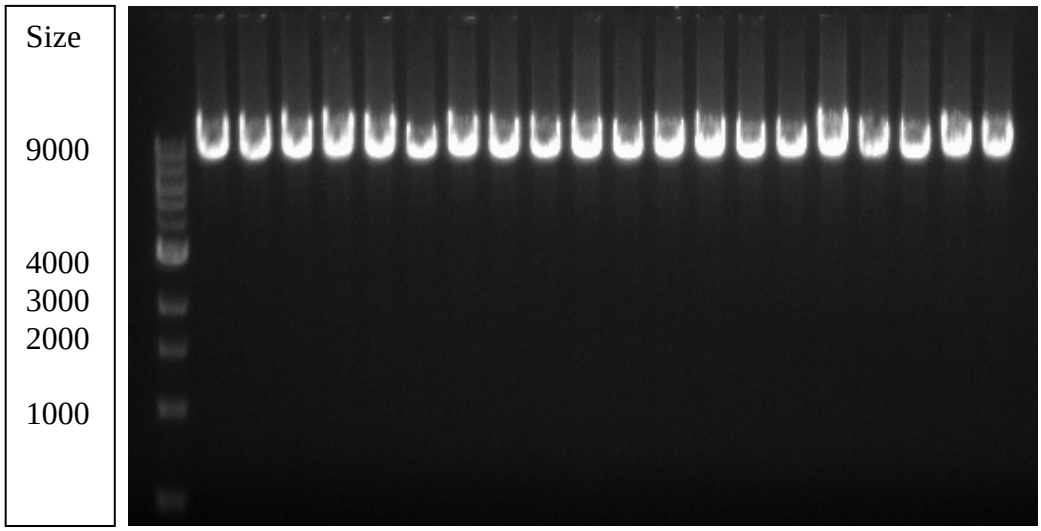
position in the mitochondrial genome to provide insights into functional effects. The variant frequencies were further assed and grouped according to common and rare variants.

Results

Long Range PCR and Sample Pooling

After several optimisations using the conditions described in the methods section, amplified product was obtained for each sample, with no non-specific bands. A single, bright ~9000 base pair fragment was obtained for each sample (Figure 15). For the samples which didn't work well on the first attempt, new dilutions were made from the original TE stocks and the PCR was repeated until sufficient amplification was obtained. Negative controls were set up with each run and checked on an agarose gel for contamination of any reagents. All runs were found to be contamination free.

Figure 15: Size fractionation of long range PCR products to be used for sequencing library preparation.



Samples were run on an agarose gel to confirm amplification and to check negative controls for contamination. The bright bands as seen under UV light indicate that all samples amplified very well, producing more template than what was required for library preparation. The expected fragment sizes were obtained as shown when comparing samples to a 1 kb ladder run in the first lane. All negative controls were contamination free.

Once it was confirmed that the PCR worked for each sample, samples were quantified using the bioanalyser as previously described. Each sample underwent two PCR reactions, one for each mitochondrial fragment. As shown in Table 17 below, both fragments were pooled for each sample, by diluting fragments to the same concentration and adding equal volumes of each diluted product. The required input for library preparation was then measured out according to concentration for each sample as shown below.

Table 17: Example of pooling calculations

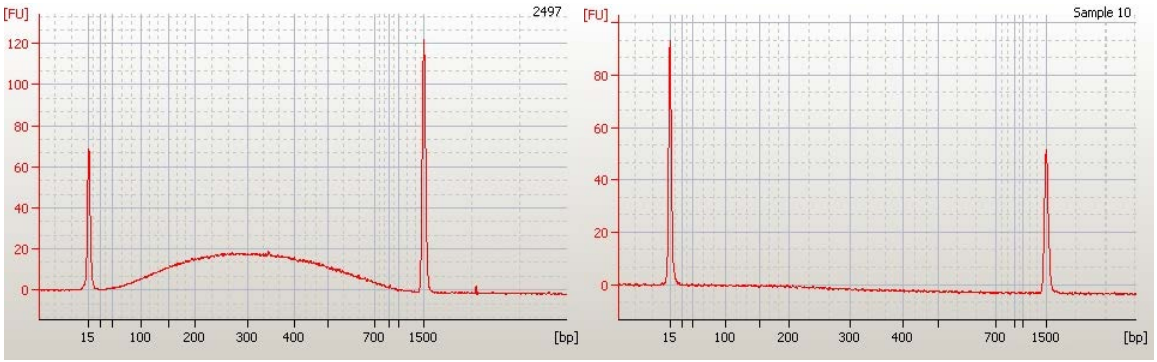
Sample	Lower Con*	Higher Con*	Volume Higher Con* to dilute	Volume water to be added	Volume for 100 ng	Volume for 1 ug
1146	17.34	23.4	37.1	12.9	5.8	57.7
6885	19.64	27.74	35.4	14.6	5.1	50.9
6919	15.99	36.97	21.6	28.4	6.3	62.5
2476	23.16	49.25	23.5	26.5	4.3	43.2

***Con=Concentration**

Library Preparation

Shearing

Figure 16: Efficacy of physical shearing versus enzymatic shearing

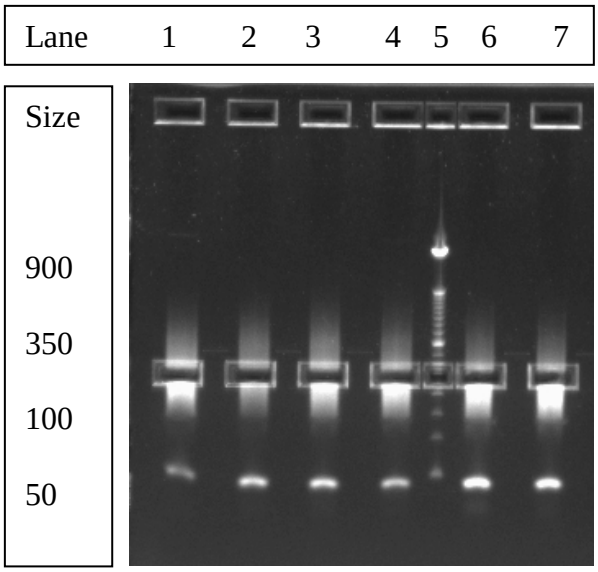


As shown in Figure 16 above on the right side, enzymatic shearing was found to be ineffective. Physical shearing as seen on the left side was found to be a successful approach. Both the biorupter and covaris systems were found to be satisfactory, but due to cost effectiveness the biorupter was chosen to carry out shearing for all samples.

Size Selection

Following barcode and adapter ligation and all clean up steps, size selection was performed using an E-gel size select system (Catalogue number G6610-02). A manual approach of running products on an agarose gel, cutting the band out and selecting the target using a purification kit was also attempted but found to be unsuccessful. As can be seen in the image below the E-gel system was highly effective and allowed for accurate size selection according to the marker in the middle well. The first bright band that can be seen is the 350 base pair mark.

Figure 17: UV image of E-gel used for size selection of library products

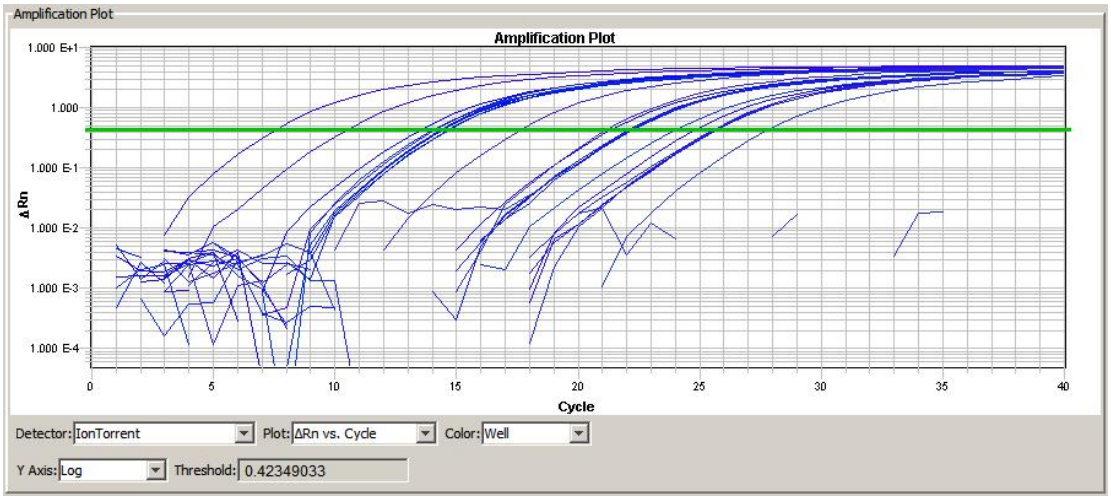


Each lane contains a unique library in the process of construction. The lower well facilitates capture of the target library. Lane 5 contains a 50 bp marker for accurate fragment size estimation. Once the first bright marker reached the top of the well, electrophoresis was halted and the target library was collected. The first bright line in the marker series indicates the 350 bp mark.

Library Quantification

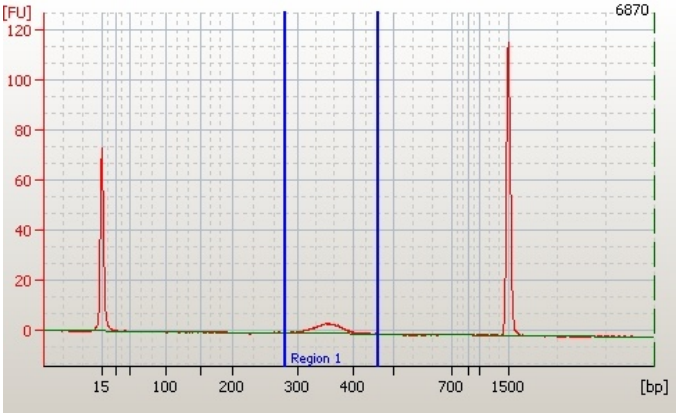
Two methods were tested for the final library quantification. A real time PCR, using a serial E-coli dilution with a known concentration was used to construct a standard curve for comparison to each sample. This was found to be the most accurate method of quantification and is illustrated in Figure 18 below. However this approach can be costly and time consuming. Quantification using Agilent DNA 1000 bioanalyser chips (Catalogue Number 5067-1504) was also performed and was found to be a satisfactory method for library quantification. A typical electropherogram utilising a Life Technologies kit is shown in Figure 19. As can be seen in Figure 20 the NEB library preparation kits (Catalogue Number 4474178) were much more effective despite being more cost effective.

Figure 18: Library quantification by real time PCR



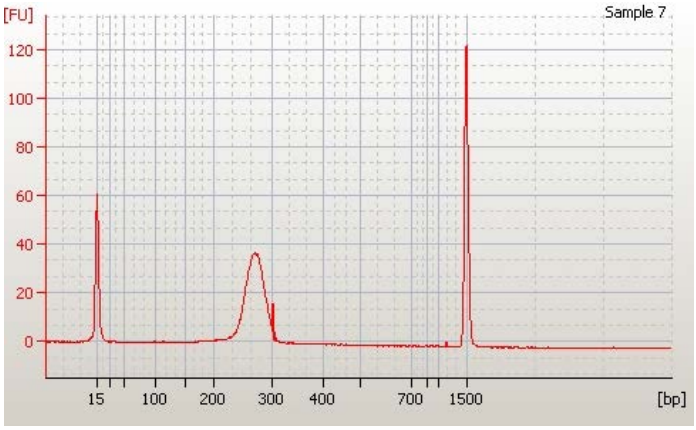
A serial E.coli dilution was used to construct a standard curve. Each sample was compared to the known standard curve in order to determine concentration. Cycles of amplification are depicted on the x axis with relative fluorescence shown on the y axis.

Figure 19: Library quantification by bioanalyser (Life Tech)



The region tool was used to accurately quantify each sample using a bioanalyser instrument. The blue lines show the defined region quantified. As shown the library produced was in the expected size range of 300 to 400 bp which is ideal for 200 bp sequencing chemistry. This approach allowed for accurate quantification of all libraries to ensure accurate dilution and pooling of samples.

Figure 20: Library quantification by bioanalyser (NEB)



A much higher peak was obtained when using the NEB kits, indicating a greater library yield and overall more efficacious library preparation. The NEB kit was found to produce a superior result in comparison to the brand name Life Technologies kits.

Template Preparation

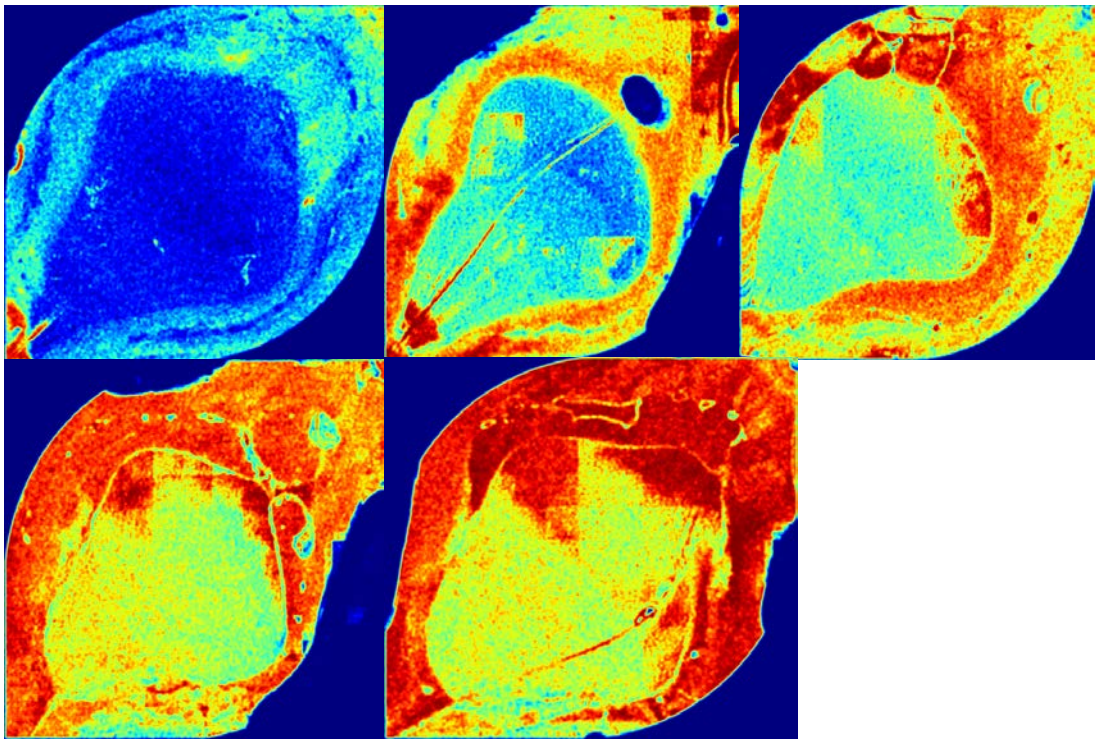
After several E.coli verification runs and a number of optimisations, it was found that for the Life Technologies kits, double the recommended amount of pooled library was optimum. For the NEB library kit (Catalogue Number 4474178), a much smaller input was sufficient. Too much library input resulted in a high percentage of polyclonal amplification and data loss.

Sequencing

Below in Figure 21, a series of heat maps are shown which correspond to the density of template positive beads sequenced on each chip. Blue depicts empty wells, while red is indicative of a high density of template positive ISPs. Poor template preparation or insufficient library input into template preparation results in low sequence coverage, as seen in the blue far left heat map. A higher recovery of template positive ISPs following template

preparation and good chip loading techniques results in a high yield of sequence information as shown in the second row.

Figure 21: Heat maps depicting bead loading and template density



The density of ISP deposition is shown with higher density levels in red and lower levels in blue. The more densely positive ISPs are deposited, the greater the sequence data output. The top row shows poor chip loading with large blue and green areas, indicating few positive ISPs. The bottom row illustrates acceptable chip loading metrics which corresponds to high sequence output.

Preliminary Analysis

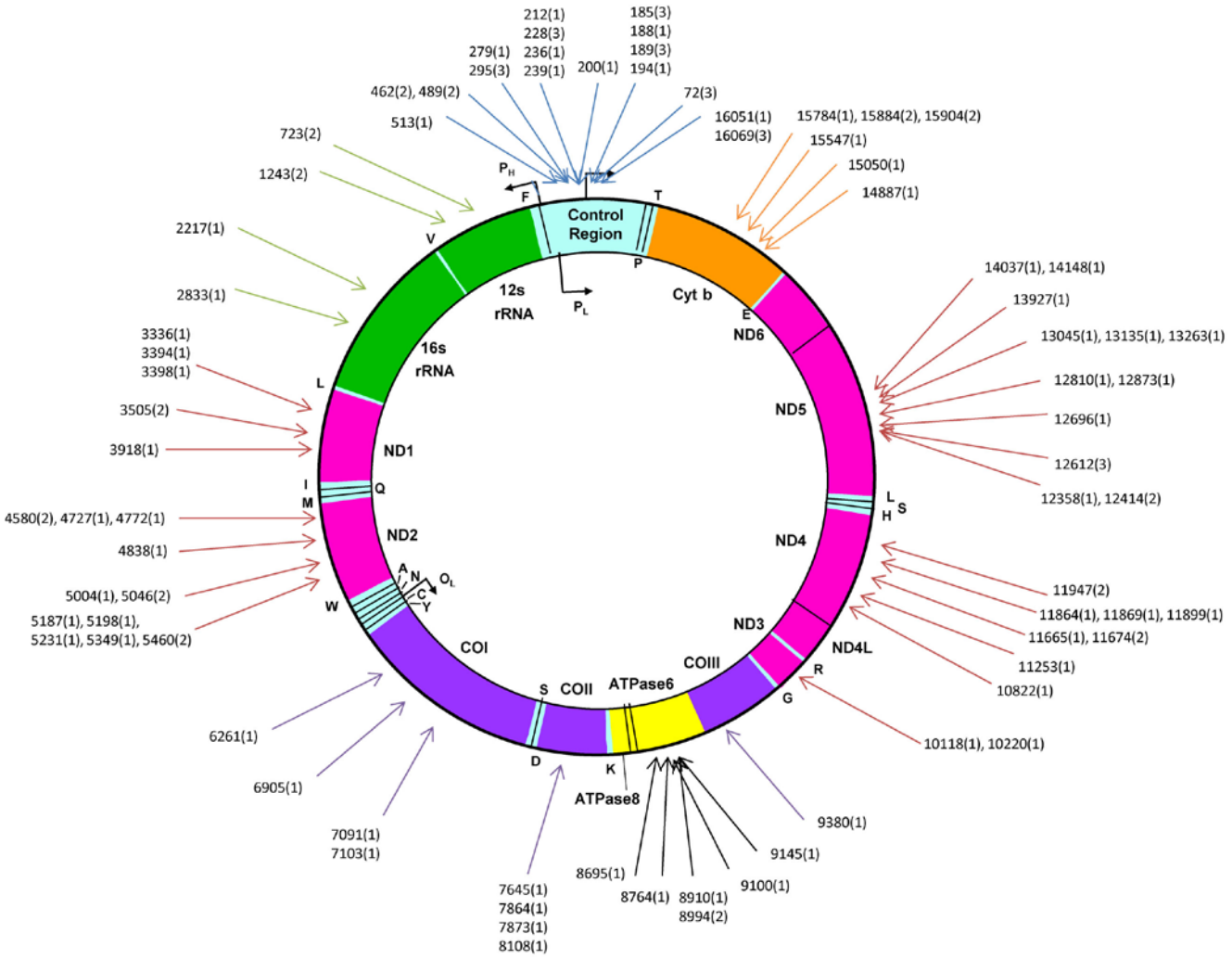
After performing a preliminary analysis as described in the statistical methods section, a number of disease susceptibility variants were identified in both migraine cases and unaffected individuals as shown in Table 18 below. It is interesting to note that many variants associated with metabolic disorders were found, as Norfolk Islanders are known to have a higher rate of these type of disorders within their population [317]. Also the two migraine affected cases both carried the same two variants and none of the others.

Table 18: Variants found in first 6 samples sequenced according to MitoMap

Sample (LAB ID)	Migraine Status	Variants	Disease Associations
2573	Not affected	195	Bipolar disorder, schizophrenia, depression
		9055	Parkinson's Protective
		11467	Altered brain pH
		12372	Altered brain pH
		16189	Cancer, Metabolic syndrome, endometrial cancer, diabetes, cardiomyopathy, respiratory chain defects
2525	Not affected	195	Bipolar disorder, schizophrenia, depression
		3736	LHON
		10398	Invasive breast cancer, Alzheimer's, Parkinson's, Bipolar lithium response, Type 2 diabetes
321	Not affected	10398	Invasive breast cancer, Alzheimer's, Parkinson's, Bipolar lithium response, Type 2 diabetes
		11467	Altered brain pH
		12372	Altered brain pH
		15693	Possible LVNC cardiomyopathy associated
		16189	Cancer, Metabolic syndrome, endometrial cancer, diabetes, cardiomyopathy, respiratory chain defects
6920	Not affected	195	Bipolar disorder, schizophrenia, depression
		6480	Prostate Cancer
		15043	Major Depressive disorder, schizophrenia, bipolar
		16189	Cancer, Metabolic syndrome, endometrial cancer, diabetes, cardiomyopathy, respiratory chain defects
2558	Migraine	195	Bipolar disorder, schizophrenia, depression
		10398	Invasive breast cancer, Alzheimer's, Parkinson's, Bipolar lithium response, Type 2 diabetes
1146	Migraine	195	Bipolar disorder, schizophrenia, depression
		10398	Invasive breast cancer, Alzheimer's, Parkinson's, Bipolar lithium response, Type 2 diabetes

After the raw sequence reads were filtered and aligned to the Revised Cambridge Reference Sequence, variants were called relative to the reference and summarised according to migraine status. For each possible variant identified, the number of migraine sufferers who were found to carry the variant were summed and compared to the number of healthy control cases carrying the variant. We found a large number of rare variants which were only present in migraine sufferers and were not found in any controls. These are shown in Figure 22 which also shows the distribution of variants across the mitochondrial genes and the number of migraine sufferers identified who carry each rare variant.

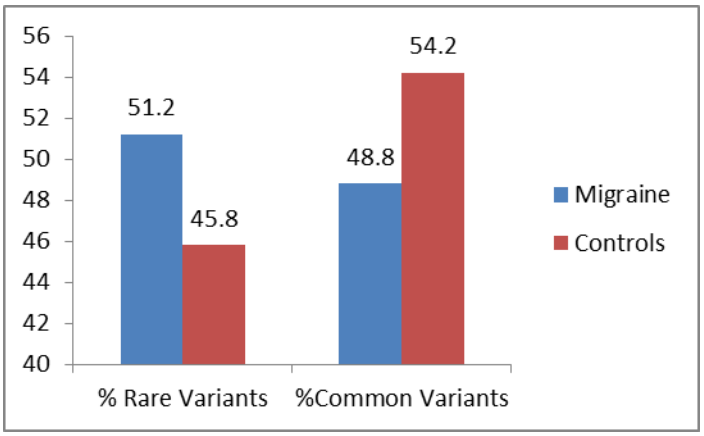
Figure 22: Distribution of rare variants identified only in migraine sufferers and not found in any controls. The location of each identified variant in the mitochondrial genome is shown along with the number of migraine sufferers harbouring each variant indicated in brackets.



Protein coding genes are depicted by solid coloured bars. Each arrow represents an identified variant with the location in the mitochondrial genome shown by the arrow. This diagram clearly illustrates the distribution of variants and clustering according to gene location. In brackets are the number of migraine suffers which were found to carry the variant.

The mt sequence variation identified in the 48 NI individuals differed from the RSRS at 296 variable sites. Of these variant sites, 29 variants were common in the 48 NI individuals (>5 %). Many of these common variants are the defining markers of mitochondrial haplogroup B, and its further substructure (haplotypes) such as B4a1a1, to which Polynesians belong. There were 136 singleton variants identified and a further 160 variants were shared by two or more individuals. As can be seen the majority of rare variants found only in migraine sufferers are clustered in the genes encoding for the ATPase 6 protein and the NADH dehydrogenase subunits, which are core components of the oxidative phosphorylation pathway and essential for the production of ATP. Variants found in 3 or less samples(less than 5 % of the cohort) were defined as rare variants and 51.2 % of the variants identified within migraine samples were found to be rare compared to 45.8 % in the control group. This trend is shown in Figure 23.

Figure 23: Comparison between the percentage of rare variants identified in migraine sufferers compared to controls



The blue bar represents migraine sufferers, with healthy controls shown in red. As can be seen, more rare variants were identified in migraine sufferers compared to controls.

We were able to generate high quality sequence information at a sufficient depth of coverage to detect heteroplasmic variants up to a 10 % threshold and calculated that sequencing the 19 kb mitochondrial genome at a minimum of 100 x coverage would be achievable by multiplexing 6 samples on a 314 chip. We far exceeded the manufacturers' specifications (10Mb per 314 chip) and obtained on average between 250-550 x coverage of the entire mitochondrial genome for each of the 48 samples.

Importantly, our preliminary data identified 6 novel (undocumented) mtDNA variants in the NI sample (see Table 19). An extended database search of mtDB (34) and Mitomap (35) for these variants returned no hits, thus these positions are deemed to be novel. Interestingly, 5 out of 6 of the novel variants were found in individuals who are migraine sufferers. Also of note**,** the majority of rare variants found only in migraine sufferers are clustered in the genes encoding for the ATPase8 protein and the NADH dehydrogenase subunits, which are core components of the oxidative phosphorylation pathway and essential for the production of ATP. These may well be of functional importance to migraine pathophysiology.

The next aim which was achieved and is described in Chapter 5 was to conduct a full project using the cost effective in-house method already developed during the pilot project to sequence the remainder of the identified Norfolk Island Core pedigree individuals (n=306). We already showed the ability to sequence mitochondrial genomes with 100 % coverage in this pilot stage of the project and identified several novel variants only present in migraineurs. The next aim in line with the project objectives was to acquire additional mitochondrial sequences with the goal of identifying functional variants associated with migraine and

determining whether variants in the Norfolk pedigree are found more widely in outbred migraine populations.

Table 19: Novel mtDNA variants identified in the Norfolk Island population

Gene	Nucleotide Change	Homoplasmic	Amino Acid Change	No. Ind	Migraine Sufferer
ND1	3833T>C	NO	YES	1	NO
ND2	5349C>T	YES	NO	1	YES
CO1	7103C>T	YES	NO	1	YES
ATPase8	8695A>T	YES	YES	1	YES
ND5	13045A>C	NO	YES	1	YES
Cytb	15050C>T	YES	NO	1	YES

Chapter 5: Mitochondrial Genome Sequencing in the Norfolk Island Core Pedigree

This chapter is formatted and presented as a manuscript in preparation for peer review. The appropriate headings and sub-headings have been used to correspond to scientific journal format so that this chapter can be submitted to an appropriate journal. Additional details have been included here for the purposes of the thesis and a condensed version will be used for publication. Analysis has been conducted in the most comprehensive manner possible and will be further expanded on to improve on the genotype, phenotype correlation. The continued use of sophisticated bioinformatics tools will further improve visualisation and understanding of this large body of data. Next generation sequencing technologies produce vast amounts of data which presents a challenge to all researchers for adequate visualisation and interpretation. As the field of bioinformatics evolves, the tools available will enable more comprehensive interpretation, making these large volumes of sequence information more useful in mapping the function of the human genome. This study produced terabytes of sequencing data which presents enormous analytical challenges. Through the use of commercially available software in conjunction with customised in-house bioinformatics tools, many of these challenges were overcome with the results described here.

Abstract

Migraine is a complex disease and multifactorial in nature with both genetic and environmental components contributing to the overall phenotype. Molecular genetic studies have identified a number of susceptibility regions in the genome which contribute towards disease risk, but a large proportion of the genetic variance still remains to be elucidated and novel avenues of exploration are needed to address this area. Biochemical, morphological and therapeutic studies provide strong evidence that mitochondrial dysfunction could be involved in migraine pathogenesis. The aim of this study is to explore the role of mitochondrial dysfunction in relation to migraine susceptibility from the molecular genetics point of view. Norfolk Island is an ideal population for identification of complex disease traits as the genetic heterogeneity typical of complex disease is reduced. In total 315 individuals were selected from the most related individuals with direct ancestry to the original founders for full mitochondrial genome sequencing on the Ion Torrent platform. Logistic regression analysis showed that 1 SNP, mt 930 G>A located within the 12S rRNA subunit is significantly associated with migraine. A Fisher's exact test identified two rare variants associated with migraine susceptibility. We hypothesize that mt 6480 G>A situated within the *COX1* gene presents genetic evidence that there could be a shared pathogenic mechanism involved in common migraine and stroke. Further we present the first genetic evidence showing a link between *NADH* dehydrogenase and migraine. NADH reductase deficiency has been successfully treated with riboflavin and this presents a new therapeutic avenue for migraine sufferers.

Introduction

Migraine is a common neurological disorder characterised by debilitating head pain and an assortment of additional symptoms which can include nausea, vomiting, photophonia and phonophobia [318]. Migraine has been classified by the International Headache Society into two main subtypes namely migraine with aura (MA) and migraine without aura (MO). Sufferers who experience additional neurological symptoms ranging in severity from visual scintillations to numbness and hemiplegia are categorised as MA sufferers, while patients who do not experience these additional symptoms fall into the MO classification [2, 319]. The World Health Organisation (WHO) ranks migraine as one of the top twenty most debilitating diseases in the developed world [4]. Migraine poses a significant personal burden to sufferers, but also has a major economic impact caused by lost productivity due to absenteeism from the workplace. It has been shown that the indirect costs associated with lost productivity far exceed the direct medical costs of treatment. However current treatments are only effective for a proportion of sufferers and new therapeutic targets are needed to alleviate the burden of migraine [5, 9-11].

Migraine is a complex disease and multifactorial in nature with both genetic and environmental components contributing to the overall phenotype. Twin studies have shown a high concordance rate between monozygotic and dizygotic twins, with heritability estimates ranging between 40-60 % [320]. A clear tendency for migraine to occur in families has been observed, further strengthening the evidence that there is a strong genetic contribution to the overall phenotype [321, 322]. Molecular genetic studies have identified a number of susceptibility regions in the genome which contribute towards disease risk, but a large proportion of the genetic variance still remains to be elucidated. Strategies have included large GWAS studies, family linkage studies, genotyping case control cohorts and more

recently sequencing genes or even entire genomes [323-327]. Scientific studies in conjunction with clinical investigations have improved our understanding of the pathophysiology of migraine, but the complete physiological process is still not fully understood.

The current most widely accepted theory is that activation of the trigeminal nervous system through a variety of vascular and/or neurological mechanisms results in generation of a pain signal within the sensory cortex [328]. Cortical spreading depression (CSD) which is a wave of neuronal and glial depolarization or neuronal hyperexcitability followed by a long-lasting suppression of neuronal activity [329] is also thought to play a role in trigeminal activation, especially in the pathophysiology of MA. CSD has been linked with some of the additional neurological symptoms experience by MA sufferers such as visual scintillations [330]. An alternative proposal is that dysfunction of the diencephalic nuclei situated within the brain stem causes a signalling disturbance resulting in a migraine attack [331]. Given the complex multifactorial nature of migraine and highly variable phenotype it is likely that multiple pathways and different signalling cascades could result in the same phenotype, but differ between individuals.

Thus far three main groups of genes have been found to be involved in migraine namely vascular, neurological and hormonal [332]. Given the vascular and neurological component of migraine, it is unsurprising that variants within genes which control these functions have been found to play an important role. Also given the female preponderance of migraine sufferers and the clearly defined sub-classification of menstrual migraine it is logical that hormonal genes influence disease state. Even with these discoveries, a large proportion of the genetic variance responsible for migraine remains unexplained. Novel avenues of exploration are needed to address this area of research. The link between mitochondrial dysfunction and

migraine was first suggested by scientists in the 1980's and since then a growing body of evidence has strengthened this hypothesis [26, 333, 334]. The aim of this study is to explore the role of mitochondrial dysfunction in relation to migraine susceptibility.

Mitochondria are the power houses of the cell and provide energy to the body's tissues in the form of ATP via the oxidative phosphorylation (OXPHOS) chain. They also function critically in stabilising intracellular Ca^{2+} levels, protecting the cell from damaging Reaction Oxygen Species (ROS) and controlling vascular tone through Cytochrome C Oxidase (COX) and Nitrogen Oxide (NO). Mitochondria provide energy for the body through other metabolic pathways including production of ketones from fatty acids and anaerobic ATP synthesis. In mitochondrial disorders the tissues which are most adversely affected are those with the highest energy requirements namely muscle and nervous tissue. Strong evidence from biochemical, morphological and therapeutic studies show a link between mitochondrial dysfunction and migraine pathogenesis [47, 280, 333, 335-338]. Genetic studies have been limited by sample size and molecular data making this an area that needs to be addressed. Four mitochondrial variants have thus far been associated with an increased risk for developing migraine including mt 4336 A>G [280], mt 16519 C>T, mt 3010 G>A [339] and mt 15699 G>C [340] in a limited number of patient samples.

This project is the first molecular genetic study to comprehensively examine the full mitochondrial genome in a large genetic isolate in relation to migraine susceptibility. Norfolk Island is a genetically isolated population situated off the cost of Australia, best known from the "Mutiny on the Bounty" historical account. This isolate is an ideal population for identification of complex disease traits as the genetic heterogeneity typical of these diseases is reduced. Geographical isolation further reduces environmental heterogeneity and

environment-gene interactions, increasing the chance of identifying true causal variants [341, 342]. More than 1000 nuclear genes encode products which are imported into the mitochondria and thereby directly affect mitochondrial function [343]. This study further addresses the role of Nuclear Encoded Mitochondrial Proteins (NEMPs) and migraine susceptibility, results outlined in Chapter 6.

Subjects and Methods

Subjects

Norfolk Island is a volcanic island situated 1600 kms east, north-east from Sydney Australia and has a unique history which has been well documented over time. Fletcher Christian aboard the British "Bounty" led a mutiny that resulted in Pitcairn Island being settled by 9 sailors and 12 Tahitian women. Once the population exceeded carrying capacity almost everyone relocated to Norfolk Island on June 8, 1856. Since then geographic isolation and strict immigration policy has resulted in a unique genetically isolated population, with the majority of current residents being related to the original founders [342]. Genetic isolates such as Finish, Sardinian, Icelandic and other Scandinavian populations have been sampled and utilised in genetic studies to map genes causing disease [344, 345]. Given the population structure of Norfolk Island, it is a useful genetic isolate for complex disease mapping. Traits which are very prominent in this group include cardiovascular disease (CVD), diabetes, metabolic syndrome and migraine [341].

In the year 2000 collections were taken from 602 individuals residing on Norfolk Island as part of the Norfolk Island Health Study. Extensive anthropometric measurements were recorded for each individual including weight, height, waist circumference, hip

circumference, body fat percentage and blood pressure. A blood sample was also taken for full plasma chemistry analysis and later DNA extraction. Each participant completed detailed questionnaires describing family medical histories, lifestyle choices and current medical conditions. DNA was extracted from whole blood samples using standard salting out and stored appropriately. Stock samples were diluted to a working concentration of 20 ng/uL for use in the lab. All participants consented to the collection and use of information as part of an ongoing genetic research study. Ethical clearance was obtained initially from the Griffith University Human Research Ethics Committee, and later from the Queensland University of Technology Ethics Committee, for the collection and utilisation of all DNA samples included in this study [342].

Accurate and detailed historical accounts have been used by genealogists to create and maintain a well-documented database of the entire Norfolk Island population, spanning all the way back to the original founders. This pedigree has been drawn up and is maintained in a genealogy program known as Brother's Keeper. The pedigree includes ~5700 individuals coalescing over 11 generations or 200 years back to the original 9 European sailors and 12 Tahitian women [346]. The Norfolk Island Health Study sampled individuals from the lower four generations of the pedigree and included 386 (64 %) individuals possessing lineages back to the founders and 216 individuals (36 %) who were considered to be new founders and did not show direct ancestral links [342]. An updated core pedigree was constructed using this information and genetic information as it became available through genetic studies. Currently the core pedigree structure contains those individuals that are most closely related and coalesce directly back to the original founders [347]. For this study 315 individuals were selected from the core pedigree for full mitochondrial genome sequencing on the Ion Torrent platform.

Methods

Molecular Techniques

For this study mitochondrial DNA was enriched and purified before undergoing library preparation. Samples were amplified by long range PCR, utilising two primer pairs which produce overlapping fragments covering the entire mitochondrial genome. Parameters which were tested and accordingly altered include final primer concentration, DMSO concentration, final reaction volume, input DNA, annealing temperature and thermocycling conditions. Final reaction conditions are shown in Tables 20, 21 and 22. It was found that quarter reaction volumes were sufficient to produce quality fragments suitable for library preparation. PCR products were run on 1 % (w/v) agarose gels to accommodate for large fragment sizes at 2 V/cm and visualised under UV light. Ethidium bromide was used for staining at a 4 % concentration. Negative controls were included in each PCR run to check for contamination and ensure adequate quality control.

Table 20: Primer pairs used for long range PCR

Amplicon	**Amplicon Position in Genome**	**Primer Sequences**	**Final Concentration**
mt_Frag1	569 (forward)	5’ AAC CAA ACC CCA AAG ACA CC 3’	200 nM
	9819 (reverse)	5’ GCC AAT AAT GAC GTG AAG TCC 3’	
mt_Frag2	9611 (forward)	5’ TCC CAC TCC TAA ACA CAT CC 3’	200 nM
	626 (reverse)	5’ TTT ATG GGG TGA TGT GAG CC 3’	

Table 21: Final reaction conditions

Reagent	1 x Reaction Volume uL	Final Concentration	Stock Concentration
Nuclease free water	28.3	-	-
Buffer	10	1 x	5 x
dNTPs	2.5	0.5 mM	10 mM
F/R primer	2	200 nM	10 uM
DMSO	1.5	3 %	100 %
Enzyme Mix*	0.7	0.05 U/uL	3.75 U/uL

Table 22: Thermocycling conditions for long range PCR

Temperature	Time	Cycles
92 °C	2 min	1
92 °C	10 s	9
55 °C	15 s	
68 °C	9 min	
92 °C	15 s	19
60 °C	20 s	
68 °C	9 min	
68 °C	7 min	1
8 °C	∞	

Mitochondrial PCR fragments were cleaned using QIAquick post PCR cleanup columns from QIAGEN (Catalogue Number 28106) to remove any excess dNTPs or other reagents. Once only the pure mitochondrial fragments remained, they were accurately quantified using Agilent DNA 12000 chips (Catalogue Number 5067-1508) on a bioanalyser instrument. The two overlapping fragments produced during the long range PCR process were pooled together in equimolar amounts for each sample. After pooling and thorough mixing, 100 ng was aliquoted for library preparation. The NEB NextFlex Fast DNA Library Prep Set for Ion Torrent (Catalogue Number 4474178) was utilised for all remaining samples not sequenced during the pilot stage (Chapter 4) and manufacturer instructions were followed at all times. The aliquoted 100 ng of DNA was topped up to final volume of 51 uL using deionised water as a larger volume allows for more efficient DNA shearing. Samples were sheared using the NGS Biorupter system for 15 cycles of 30 sec on, 30 sec off on the high power setting at 4 °C.

Immediately following the shearing process samples were end repaired using the enzyme mix and reaction buffer provided by NEB in the tubes labelled with green caps. This process ensured that fragments were blunt ended and compatible with the adapter and barcode sequences which were ligated in the next step. The repair process also prevented any degradation of sheared DNA fragments. In total 6 uL of NEBNext End Repair Reaction Buffer and 3 uL of NEBNext End Repair Enzyme Mix was added to each sample, bringing the final volume up to 60 uL. An incubation step consisting of 20 min at 25 °C followed by 10 min at 70 °C was applied at this stage. Samples were returned to ice before continuing to adapter and barcode ligation.

Unique barcodes were added to each sample to allow multiplexing and high throughput batching. Batches of 48 samples per chip were processed with each batch containing samples labelled with 48 unique barcodes. A P1 adapter was also ligated to the end repaired fragments and acted as a sequencing primer once attached to the Ion Sphere Particles. The following reagents were added to each sample: 19 uL sterile water, 10 uL T4 DNA Ligase Buffer for Ion Torrent, 1 uL Bst DNA Polymerase, 2 uL P1 adapter (Bioo Scientific), 2 uL barcode X (Bioo Scientific) and 6 uL T4 DNA Ligase in a total reaction volume of 40 uL. Bioo Scientific NEXTflex™ DNA Barcodes for Ion PGM™ (Catalogue Number 401004) were chosen as they were the most cost effective way of labelling samples for multiplexed batches. Samples were mixed by pipetting and incubated in a thermal cycler for 15 min at 25 °C, followed by 5 min at 65 °C. It was found that a faulty batch of DNA polymerase caused primer dimer reducing the efficacy of sequencing, and a replacement batch from the manufacturer corrected this problem without causing any significant data loss.

Post adapter and barcode ligation, samples were cleaned using AMPure XP Beads (Catalogue Number A63881) from Agencourt. Exactly 180 uL of beads were added to each sample followed by a 5 min incubation at room temperature. Tubes were pulse spun and placed on a magnetic rack for 3 min after which the supernatant was carefully removed. Two ethanol washes were then performed using 80 % ethanol freshly prepared on the day of use. 500 uL of 80 % ethanol was added to each tube, incubated for 30 sec and then removed using a pipette. This step was repeated to perform a total of 2 washes. Samples were then air dried for 5 min and re-suspended in 25 uL of 0.1 x T.E buffer. 20 uL of clear supernatant was transferred to a clean tube for the next step. It was critical not to transfer any of the beads and so 5 uL was discarded at this step.

A size selection process was undertaken to ensure maximal data output. A range of 290-350 bp was recommended to be the most desirable target size and it was the aim to capture these fragments. An E-Gel size select system (Catalogue number G6610-02) was used to capture the desired fragments with a 2 % (w/v) precast agarose gel. Samples were loaded into the top well with 10 uL of marker into the middle well and run for 12 min. All empty wells were filled with 25 uL of deionised water and toped up with 15 uL after the initial 12 min run. Samples were then run for another approximately 4 min while being constantly viewed. As soon as the 350 bp marker entered the top of the collection well the run was stop and samples were collected into a fresh tube using a pipette. A second washing step using 15 uL of deionised water was undertaken to ensure maximum sample recovery. Each time water was added to the well, pipetted up and down and then transferred to the respective sample tube.

Following size selection samples underwent library amplification. Here 10 uL of NEB Library Primers were added to each sample as well as 50 uL of NEBNext High-Fidelity 2 x PCR Master Mix. Each sample was mixed carefully using a pipette. PCR cycling conditions are given in Table 23 below.

Table 23: PCR amplification of adaptor ligated DNA

STEP	TEMPERATURE	TIME
Initial Denaturation	98 °C	30 s
8 cycles	98 °C	10 s
	58 °C	30 s
	72 °C	30 s
1 cycle	72 °C	5 min
Hold	4 °C	∞

A final cleaning step was undertaken using 100 uL of AMPure Agencourt beads (Catalogue Number A63881). The procedure was identical to the initial cleaning step and only differed with the initial volume. The ratio of beads to sample is imperative to ensure collection of the correct sized fragments in an efficient manner.

After the final amplification and purification step all libraries were quantified on a Bioanalyser using Agilent DNA 1000 chips (Catalogue Number 5067-1504). Samples for each multiplex were pooled in equimolar amounts and diluted to 26 pM. Initially 6 samples were pooled per 314 chip and later to increase thoroughput 48 samples were plexed into a single sequencing reaction on 316 chips.

Analysis

Raw sequence reads were aligned to the revised Cambridge Reference Sequence (rCRS) using SAMTOOLs to produce Binary Alignment (BAM) files which were subsequently indexed. A custom script was used to call all variants relative to the reference genome. This alignment and variant calling was undertaken by David Eccles and all subsequent analysis

were based on this custom pipeline developed by our bioinformatics team. As my own contribution all sequences were aligned using default settings on the torrent server and variants called using Java based pluggins. Interestingly there was good concordance between my own analysis and that undertaken by David, showing that tools available to wet lab users are of sufficient quality for sequence analysis.

Variants which passed quality control were analysed using a logistic regression model in Plink [348]. In total 265 variants met all quality thresholds across 315 individuals. Individuals included 80 migraine sufferers, 235 healthy controls and an even gender ratio of 152 males and 163 females. Variants with a MAF of <0.01 were excluded from the final regression model to avoid skewing of results. In total 201 common SNPs were included in the regression model and 64 rare variants were analysed separately using a Fisher's exact test in Plink v1.09. The regression test factored in for the covariates age, gender and kinship. RNAfold was used to predict secondary structural changes to 12S rRNA.

The sequencing depth obtained with a maximum of 16000 x coverage allowed us to examine heteroplasmic variants, present in a proportion of sequencing reads. A specialised script was used to call heteroplasmic variants and to determine the percentage of mitochondria the variants are found in by using the percentage of sequence reads containing the variant as a direct estimate. As an initial analysis heteroplasmic variants were coded as 1 (heteroplasmy present) and 0 (no heteroplasmy) and a logistic regression model factoring in for the covariates age, gender and relatedness was used to test for association between heteroplasmic variants and migraine susceptibility.

Haplogroups were defined for all samples and an association test was undertaken comparing haplogroup to migraine status utilising a similar logistic regression model as described before. The Norfolk Island ancestry was explored using the haplogroup information obtained and a summary of haplogroups was generated.

Quality Control

Due to technical problems with the one touch/emulsion amplification equipment, several multi-plexes were repeated. This produced in excess of 240 duplicate sequences which were pure technical replicates. Having access to so many duplicate sequences enabled a very detailed quality control process to take place, where the sequencing accuracy of the torrent platform could be comprehensively assessed. A cut off QUAL score of 999 was used. Only sequences which exceeded a QUAL of 999 were considered for variant calling. This measure assesses the quality of an individual sequence read and the likelihood that the sequence is correct. A second quality control measure was then implemented, enabled by the duplication of samples. This primarily examined the variant calling process and looked for any inconsistencies between duplicate samples. Only variants which had an Inconsistent Allele Frequency (IAF) of 1 % or lower were considered to be true variants. This rigorous quality control process ensured that any sequencing errors were filtered out and that logistic regression analysis was based on true variant calls. As with all massive parallel sequencing platforms, some sequencing error bias was observed. It was found that long homopolymer stretches presented a challenge to the Ion Torrent platform and that some sequencing slippage occurred in regions containing long homopolymer stretches. This was found to be especially true for a common 9 bp Polynesian deletion, and manual curation of this region was required. Calling indels also presented a bioinformatic challenge due to sequencing error bias, requiring

a higher level of manual input. It is important to note here that all sequencing platforms are prone to their own types of sequencing error corresponding to the chemistry used [349].

Results

Sequencing

Due to the large volume of data generated, it is not feasible to present all of the sequencing data in a thesis. Each sequencing run would fill hundreds of pages, therefore a representative run has been selected and summarised here as an example of the data generated. The full run report for run 14_repeat is shown in Appendix D.

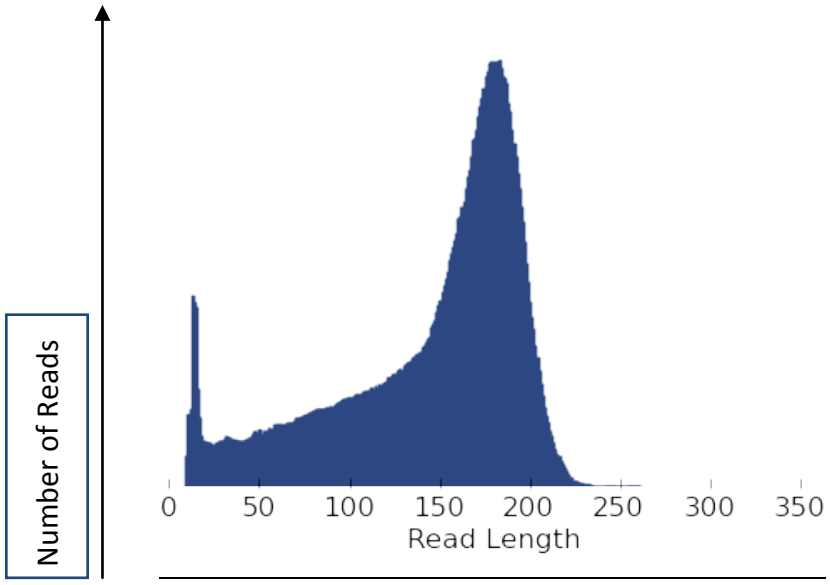
As seen in Table 24 below, sequencing run 14_repeat produced in excess of two million sequencing reads, with a total data output of 294.37 Mb and 260.49 Mb exceeding Q20 (Phred score) quality. Phred quality scores were originally developed to assist in the automation of DNA sequencing in the Human Genome Project and have become widely accepted to characterize the quality of DNA sequences. The Pred (Q) score is calculated as a probability of a base being called incorrectly. A Pred (Q) score of 20 means that 1 in 100 base pairs are likely to be an error which seems relatively high, but is the currently accepted threshold for accurate sequence data [350]. Approximately 34 Mb of data was filtered out using this quality threshold, fortunately only a small fraction of the total data generated. In this example the average read length was 142 bp which is within the expected range for 200 bp sequencing chemistry and the longest read was 376 bp indicating that the size selection process was successful and high quality libraries were generated.

Table 24: Per base quality scores for a representative sequencing run

Quality Score	Measurement
Total Number of Bases (Mbp)	294.37
Number of Q20 Bases (Mbp)	260.49
Total Number of Reads	2,060,080
Mean Length (bp)	142
Longest Read (bp)	376

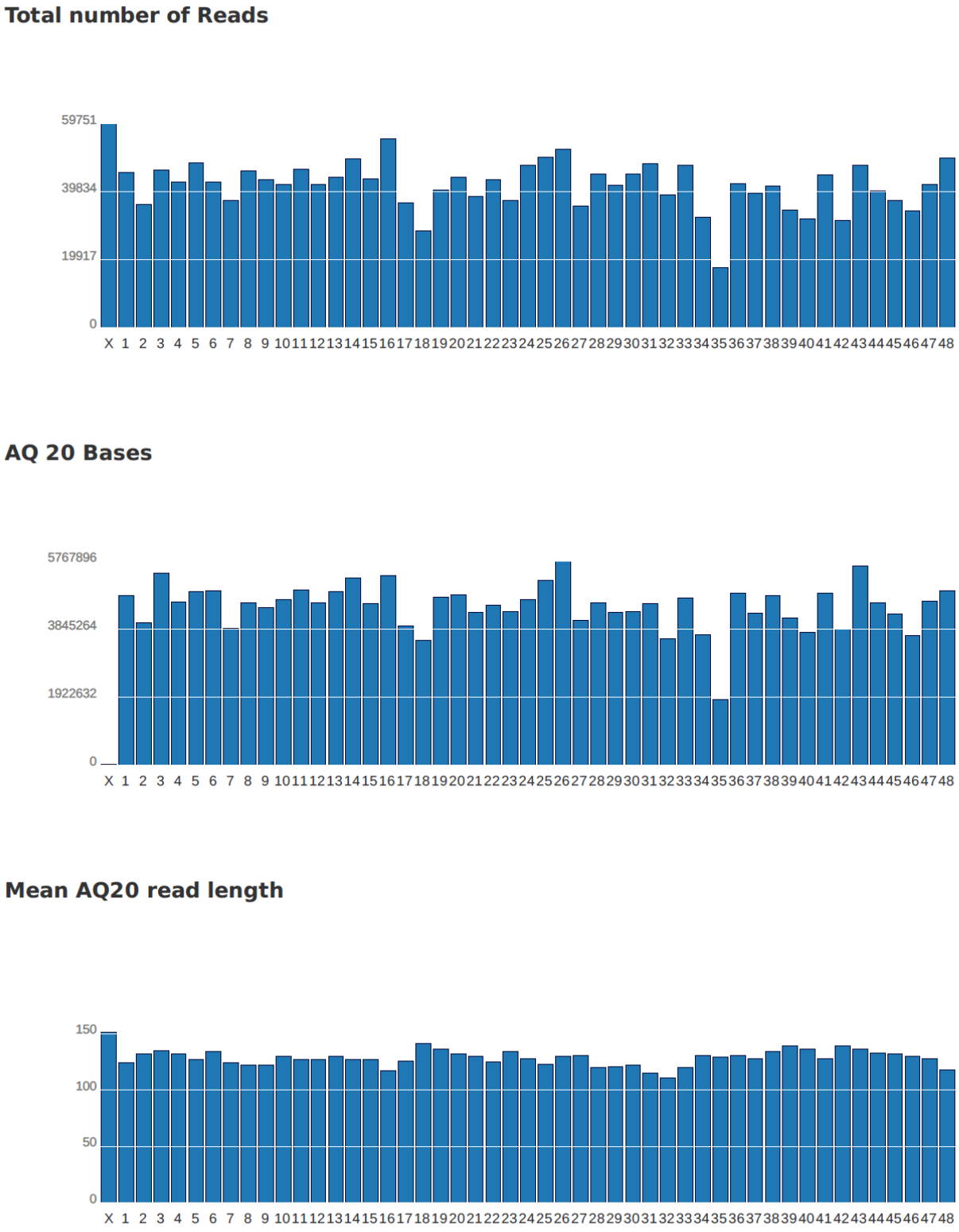
Figure 24 below graphically represents the read length of each sequence generated during the run with base pair length shown on the x axis and number of reads represented on the y axis.

Figure 24: Read length of each sequence generated during the sequencing run



The size of each fragment sequenced is shown in bp on the x axis with number of reads shown on the y axis. Most fragments sequenced were between 150 and 210 base pairs as expected when using 200 bp sequencing chemistry. While a number of smaller reads were found, the majority of sequences were in the target size range as shown by the clear peak in the diagram.

Figure 25: Distribution of barcoded samples



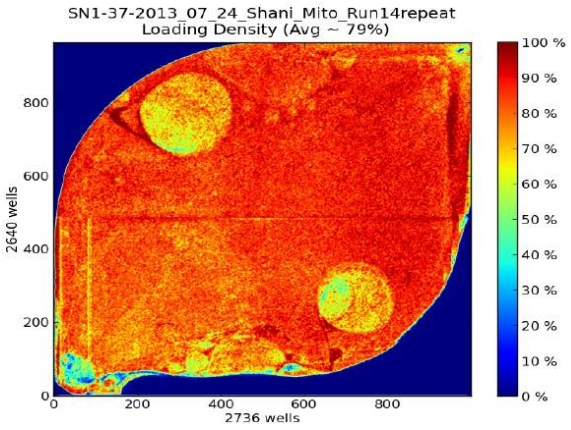
On average each barcoded sample is evenly represented. This was achieved through accurate library quantification and dilution prior to pooling.

The distribution between barcoded samples is represented in Figure 25 above with each barcode 1-48 shown on the x axis. As previously described, plexes of 48 were utilised. As can be seen, a mostly even distribution was obtained between samples ensuring that each sample was sequenced at an adequate depth of coverage. This demonstrates the accuracy of library quantification using the Bioanalyser as described in the methods section. While some variability can be seen between samples with barcodes 18 and 35 being slightly under-represented, adequate depth of coverage was still obtained for all samples.

Figure 26: Ion Sphere Particle Summary

	Count	Percentage
Total Addressable Wells	6,348,235	
▸ Wells with ISPs	5,008,543	79%
▸ Live ISPs	5,005,983	100%
▸ Test Fragment ISPs	21,477	<1%
▸ Library ISPs	4,984,506	100%

	Count	Percentage
Library ISPs / Percent Enrichment	4,984,506	100%
▸ Filtered: Polyclonal	2,052,872	41%
▸ Filtered: Primer dimer	391,251	8%
▸ Filtered: Low quality	480,179	10%
▸ **Final Library Reads**	**2,060,204**	**41%**



Chip loading and sequencing metrics are given for an example run. The red colour indicates dense loading of positive ISPs which corresponds to high sequence output. The percentage of polyclonal reads, where multiple templates are attached to a single ISP, is 41 %. This is slightly higher than manufacture recommendations and was corrected by decreasing the amount of library used during template preparation. Overall metrics indicate a successful sequencing run.

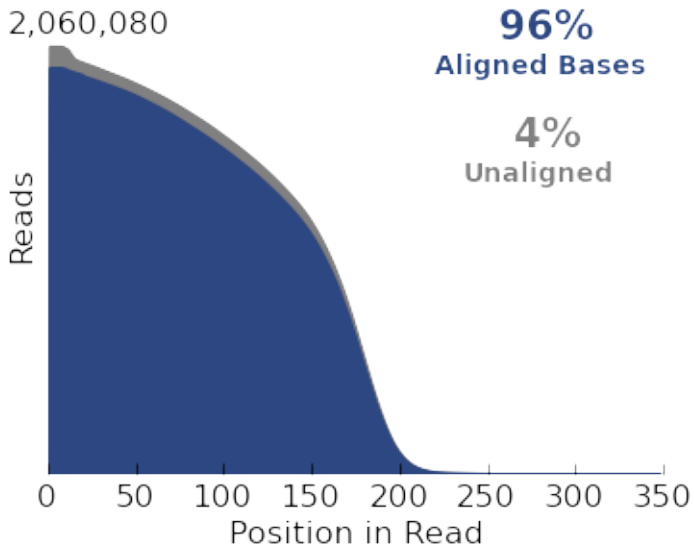
Figure 26 illustrates the summary properties of ion sphere particle (ISP) distribution. As explained in the methods section, ideally a single strand of template DNA is attached to each ceramic bead or ion sphere particle to produce an optimal sequencing run. Metrics such as these are useful for trouble shooting and provide information about the efficacy of library preparation, template preparation, chip loading and sequencing. Values falling outside of expected thresholds provide insight into the source of a potential problem. As shown above, chip loading was optimal with red indicating a high density of bead disposition. The red colour further indicates a high proportion of live ISPs, meaning that they are attached to amplified library fragments and this is ideal. Manufacturers recommend that at least 60 % of wells should contain live ISPs and with 79 % chip loading in the above example this threshold was exceeded. Test fragments are included in each sequencing run as a positive control and should constitute less than 1 % of fragments which was achieved in every run. These are known *Escherichia coli* sequences which serve to provide a basis for sequencing efficacy and accuracy of each run. If the Ion Torrent platform is functioning optimally, at least 85 % accuracy of the test fragments must be obtained. Any deviations from this threshold indicates that there could be an instrument malfunction and a field application specialist would need to assess the equipment.

These metrics provide further quality control thresholds and provide useful information for optimisations in future runs. Ideally polyclonal reads, meaning more than one template strand of DNA is attached to a single ISP, should be between 20-30 %. The template dilution factor calculation and accurate quantification is critical to achieve this. Higher percentages of polyclonal amplification, as can be observed here with 41 % of the reads being polyclonal, results in excessive data loss. Reducing the amount of template used in the emulsion PCR

process would correct this problem and this action was taken to ensure maximum data output. An acceptable percentage of primer dimer and low quality libraries resulted in further data loss in the above example. This is expected for all sequencing runs due to the nature of massive parallel sequencing. The advantage of this technology is that despite numerous errors, the amount of data produced is so vast that even after the filtering process, excess coverage is available across all regions for all samples.

Figure 27 shows the alignment of sequence data to the reference genome (RSRS), with 96 % of sequenced bases matching the reference and 4 % differing. This 4 % which differs to the reference, represents the variation which individuals possess. The aim of this project was to identify sequence variation and determine the correlation with migraine status.

Figure 27: Sequencing data aligned to the RSRS



96 % of sequenced bases aligned to the RSRS, with 4 % of sequenced bases not matching the reference sequence. This 4% are the variants identified through sequencing which are different to the reference sequence and possibly involved in modulating disease susceptibility

Figure 28 shows a snap shot from Integrated Genome Viewer (IGV), a bioinformatics platform which aids in visualisation of sequence data. Each grey bar represents a sequencing read, with variants relative to the reference genome shown as coloured bars. Samples are represented as a panel on the left hand side, with each sample shown in its own panel. IGV allows viewing of multiple samples relative to the reference genome simultaneously and also provides a good indication of average read depth obtained across the genome.

Figure 28: FASTQ sequence observed in IGV

Sequence data was visualised in Integrated Genome Viewer (IGV) with each grey bar representing a sequence read. The coloured points indicate sites that differ from the reference sequence. These may be variants or possible mis-alignments.

Analysis

Raw sequence reads were aligned to the RSRS genome using both SAMTOOLS and the default torrent server suite settings. Overall alignments obtained were similar regardless of the method used. In-house bioinformatic tools were given preference over commercial software and variant calls were based on the alignments produced by SAMTOOLS. A custom pipeline was used for variant calling after stringent quality control methods were implemented. As a comparison, Java based pluggins were also used to call variants using the torrent server and again the variant calls obtained were similar to those produced by custom bioinformatics tools.

In total 265 variants met all quality thresholds across 315 individuals. Individuals included 80 migraine sufferers, 235 healthy controls and an even gender ratio of 152 males and 163 females. Variants with a MAF of <0.01 were excluded from the final regression model to avoid skewing of results. In total 201 common SNPs were included in the regression model and 64 rare variants were analysed separately using a Fisher's exact test in Plink v1.09. The regression test factored in for the covariates age, gender and kinship.

Common Variants

Logistic regression analysis showed that 1 SNP is significantly associated with migraine susceptibility namely a G>A substitution at position 930 of the mitochondrial genome. This SNP is located within the 12S rRNA subunit and is transcribed to produce part of the ribosomal complex which translates transcripts from all the coding regions of the mitochondrial genome into proteins. Table 25 shows the top twenty hits ranked according to p value.

Figure 29 depicts all 201 SNPs tested and shows the location of detected variants throughout the mitochondrial genome. As seen in this circle plot, the only variant which exceeds the significance threshold $p<0.05$ illustrated by the blue line is 930 G>A. This substitution has been reported 556 times previously in MitoMap and currently has no known disease associations [351]. The RNAfold webserver [352] was used to predict the secondary structural changes caused by this substitution as seen in Figure 30. According to the predictive algorithms no secondary structural changes occur as a result of this substitution. However these results are only best estimates and functional studies would be needed to determine the true outcome in terms of functional effect.

Table 25: Results of logistic regression analysis of common variants factoring in for all covariates

Variant	Base Pair Position	Test	Odds Ratio	Rank	P Value
930.G	930	Additive	3.967	1	0.0233
16239.C	16239	Additive	2.389	2	0.06759
189.A	189	Additive	1.881	3	0.07081
16224.T	16224	Additive	1.781	4	0.07441
9380.G	9380	Additive	2.777	5	0.1093
11253.T	11253	Additive	2.777	6	0.1093
4727.A	4727	Additive	2.776	7	0.1095
239.T	239	Additive	2.763	8	0.1109
9055.G	9055	Additive	1.633	9	0.114
7028.T	7028	Additive	0.7352	10	0.12
1189.T	1189	Additive	1.859	11	0.1218
2706.G	2706	Additive	0.7397	12	0.1272
9698.T	9698	Additive	1.6	13	0.1272
10550.A	10550	Additive	1.6	14	0.1272
3480.A	3480	Additive	1.592	15	0.1312
11299.T	11299	Additive	1.592	16	0.1312
14167.C	14167	Additive	1.592	17	0.1312
16129.A	16129	Additive	0.4517	18	0.132
497.C	497	Additive	1.785	19	0.1416
11719.A	11719	Additive	0.7746	20	0.1592

Rare Variants

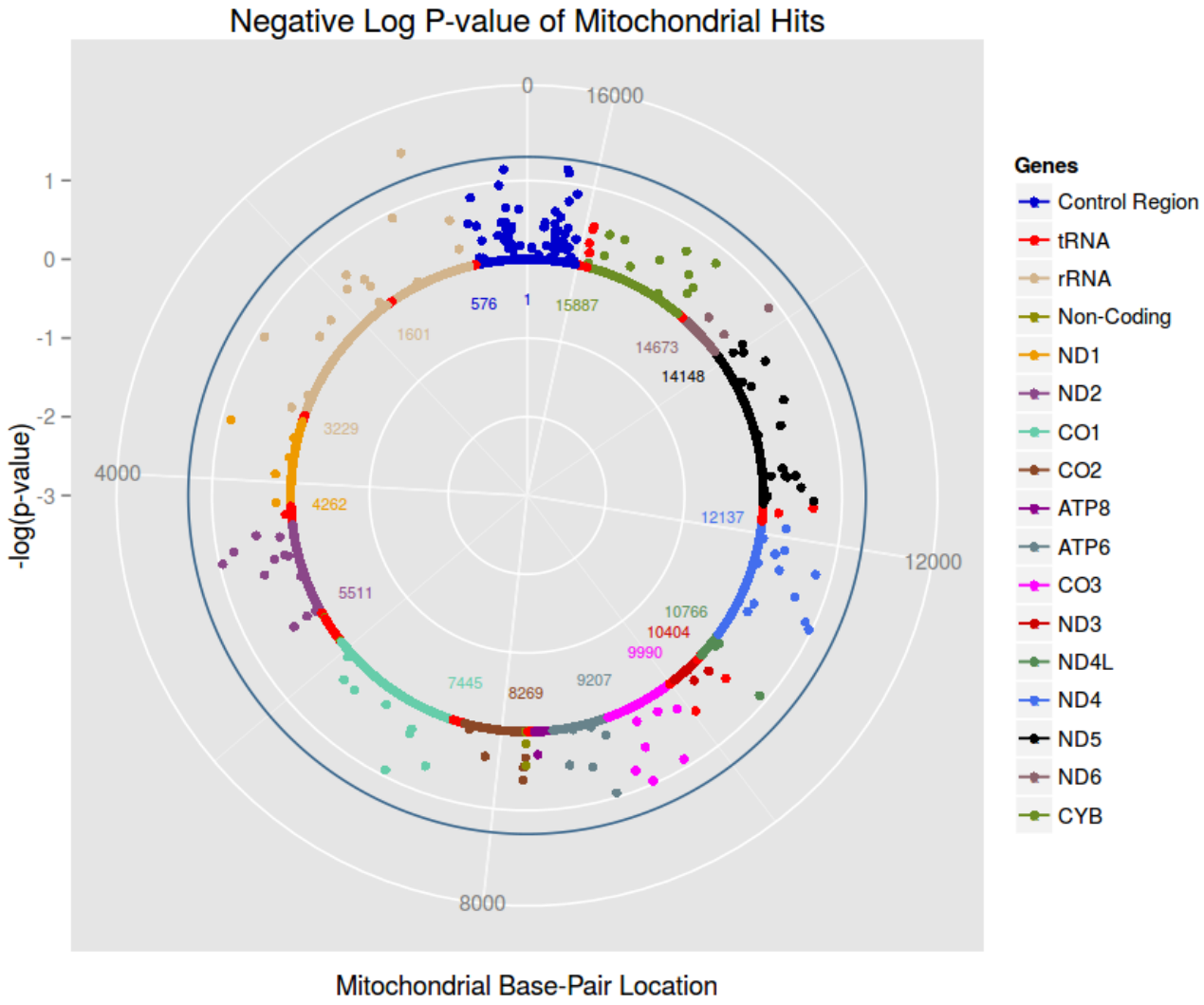
A Fisher's exact test was performed in Plink v1.09 to analyse the remaining rare variants. This showed that 2 SNPs are significantly associated with migraine and given that they were only found in healthy controls they may confer a slightly protective effect. Table 26 illustrates the results of this analysis for all variants which were polymorphic. Those variants which differed from the reference genome, but were present in all 315 individuals and therefore had a MAF of 0 are not shown in Table 26. The most significantly associated rare variant, mt 11930 A>G is a non-synonymous amino acid substitution, with the affected codon causing a change from isoleucine to valine. This variant is novel and has never been previously reported

in MitoMap or mtDB [353]. It is situated in the mitochondrial gene encoding for NADH dehydrogenase, subunit 4 which forms a critical component of the oxidative phosphorylation chain. The second hit, mt 6480G>A is also a non-synonymous amino acid substitution, situated in the *COX1* gene. It too causes a change from isoleucine to valine and has previously been associated with prostate cancer [354] .

Table 26: Fisher's exact test, identifying association between rare variants and migraine

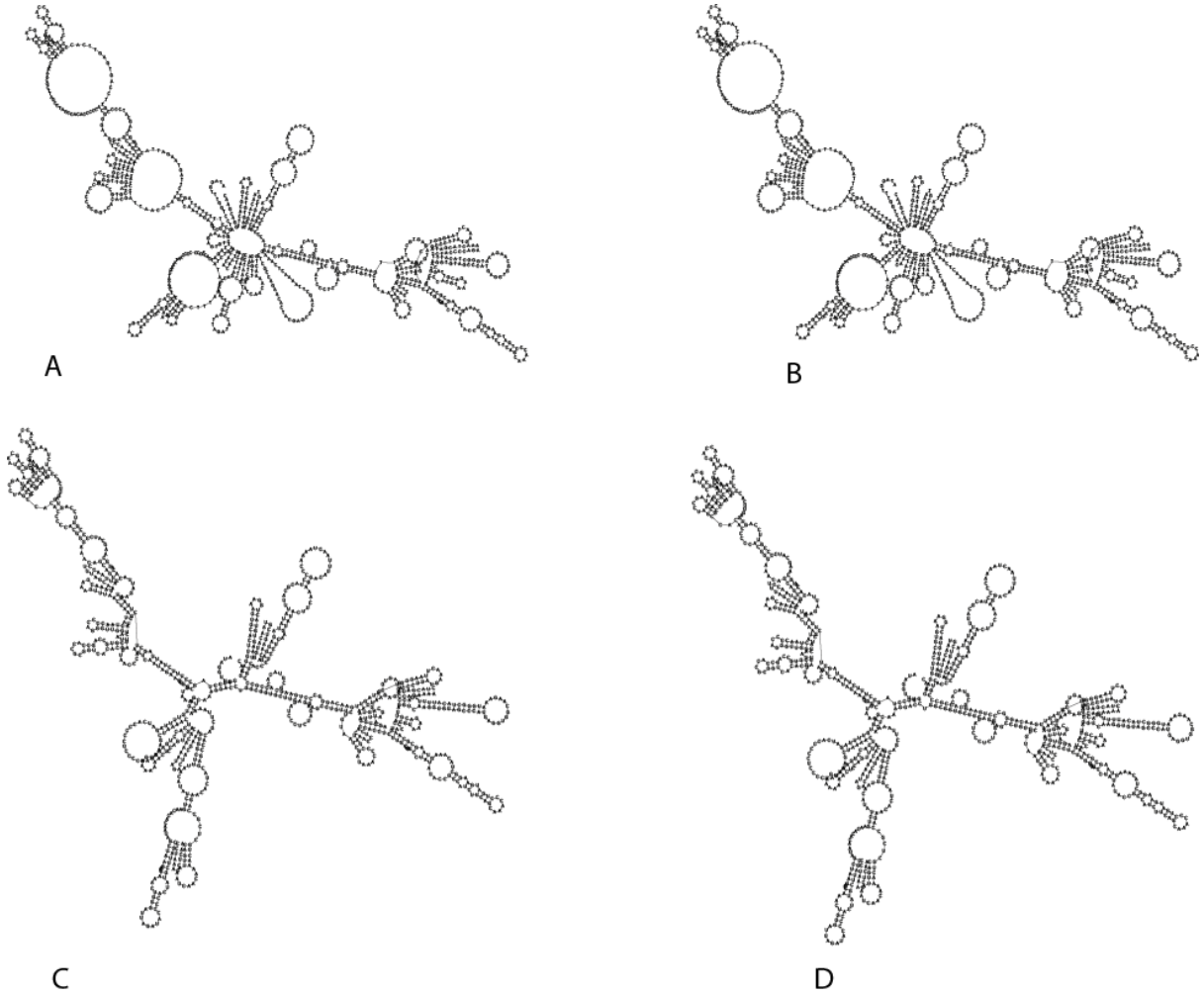
Variant	Base Pair Position	MAF* Cases	MAF* Controls	P Value	Odds Ratio
11930.A	11930	0	0.02991	0.02628	0
6480.G	6480	0	0.02979	0.0263	0
279.T	279	0.0125	0.03863	0.1232	0.315
10822.C	10822	0.0125	0.03863	0.1232	0.315
5187.C	5187	0.0125	0.03846	0.1236	0.3165
16169.C	16169	0.0125	0.03846	0.1236	0.3165
16257.C	16257	0.0125	0.03846	0.1236	0.3165
3918.G	3918	0.0125	0.0383	0.1239	0.3179
6261.G	6261	0.0125	0.0383	0.1239	0.3179
7873.C	7873	0.0125	0.0383	0.1239	0.3179
12358.A	12358	0.0125	0.0383	0.1239	0.3179
12873.T	12873	0.0125	0.0383	0.1239	0.3179
12795.G	12795	0.0125	0.004274	0.2695	2.949
9145.G	9145	0.025	0.01277	0.2842	1.983
6905.A	6905	0.025	0.01282	0.2856	1.974
1240.ACC	1240	0	0.01064	0.3391	0
16280.A	16280	0.0125	0.03004	0.3821	0.4087
4796.C	4796	0.025	0.01717	0.5138	1.468
14893.A	14893	0.025	0.01717	0.5138	1.468
14971.T	14971	0.025	0.01717	0.5138	1.468
12103.C	12103	0.025	0.01724	0.5149	1.462
14003.C	14003	0.025	0.01724	0.5149	1.462
3736.G	3736	0.0125	0.02564	0.5356	0.481
6782.T	6782	0.0125	0.02564	0.5356	0.481
16271.T	16271	0.0125	0.02564	0.5356	0.481
318.T	318	0.0125	0.02575	0.5358	0.4789
10688.A	10688	0	0.008511	0.5769	0
16356.T	16356	0.025	0.03617	0.6162	0.6833
7559.A	7559	0.0125	0.008511	0.647	1.475
198.C	198	0.025	0.0279	1	0.8935
499.G	499	0.025	0.02778	1	0.8974
3841.TGA	3841	0.007042	0.008969	1	0.7837
5999.T	5999	0.025	0.02778	1	0.8974
6047.A	6047	0.025	0.02778	1	0.8974
8818.C	8818	0.025	0.02778	1	0.8974
11332.C	11332	0.025	0.02766	1	0.9014
12937.A	12937	0.025	0.02778	1	0.8974
14620.C	14620	0.025	0.03017	1	0.8242
15244.AGGC	15244	0	0.002262	1	0

Figure 29: Logistic Regression Analysis of 201 common variants depicted as a Circular Plot showing each variant according to mitochondrial genome position



The circular plot shows each identified common variant as a coloured dot according to location within the mitochondrial genome. Each gene is colour coded according to the key given on the right of the plot. The solid blue line represents the cut off for significance where p<0.05. Only a single common variant exceeds the solid blue threshold line. It is a G>A substitution at position 930, located in the 12S rRNA subunit.

Figure 30: Predicted secondary structural changes to 12s rRNA at position 930



According to RNAfold the predicted secondary structure of 12S rRNA does not change with mt 930G>A substitution. A and C show the centroid and the predicted base pair probability pairing respectively where the wildtype G allele is present. B and D show the same structures when the A substitution occurs. There is no change in the predicted secondary structure.

Heteroplasmic Variants

Logistic regression factoring in for the covariates age, gender and relatedness identified 11 variants to be significantly associated with migraine susceptibility. These results are presented in Table 27 below.

Table 27: Results of heteroplasmic variant analysis

CHR	SNP	BP	A1	TEST	OR	STAT	P
26	8697{A}	8697	A	ADD	0.3639	-2.56	0.01047
26	16148{T}	16148	A	ADD	0.4923	-2.559	0.0105
26	10400{T}	10400	T	ADD	2.029	2.551	0.01075
26	279{C}	279	A	ADD	2.11	2.489	0.0128
26	10398{A}	10398	A	ADD	0.5052	-2.376	0.0175
26	8655{A;C}	8655	A	ADD	1.96	2.293	0.02187
26	16239{T}	16239	A	ADD	0.2414	-2.254	0.02419
26	12937{G}	12937	A	ADD	1.926	2.232	0.02559
26	1240-1246{D1-2}	1240	T	ADD	2.032	2.188	0.02865
26	9716{C}	9716	T	ADD	1.72	2.01	0.04447
26	8818{T}	8818	A	ADD	0.5084	-1.963	0.04963
26	16362{A;C;G}	16362	T	ADD	0.5574	-1.919	0.055
26	4529{T}	4529	A	ADD	1.712	1.892	0.05855
26	12103{A}	12103	T	ADD	1.641	1.816	0.06931
26	3505{G}	3505	A	ADD	3.067	1.815	0.06951
26	14022{C;G;T}	14022	A	ADD	1.753	1.809	0.07042
26	508-532{D1-2}	508	T	ADD	1.728	1.755	0.07924
26	239{C}	239	A	ADD	0.3353	-1.707	0.08781
26	9123{A}	9123	A	ADD	0.5763	-1.704	0.08833
26	9145{A}	9145	A	ADD	1.597	1.69	0.09096

As described in the methods section heteroplasmic variants were defined and called using a custom script and an in-house developed bioinformatics pipeline. This analysis incorporated the depth of sequence and percentage of reads containing the identified variant. It is interesting that heteroplasmic variants were found to play a more significant role than homoplasmic variants in migraine susceptibility. This data suggests that acquired variants accumulating over time and reaching a threshold may be an important factor in migraine pathogenesis and other neurological diseases. This topic should definitely be explored further in future studies.

Haplogroups

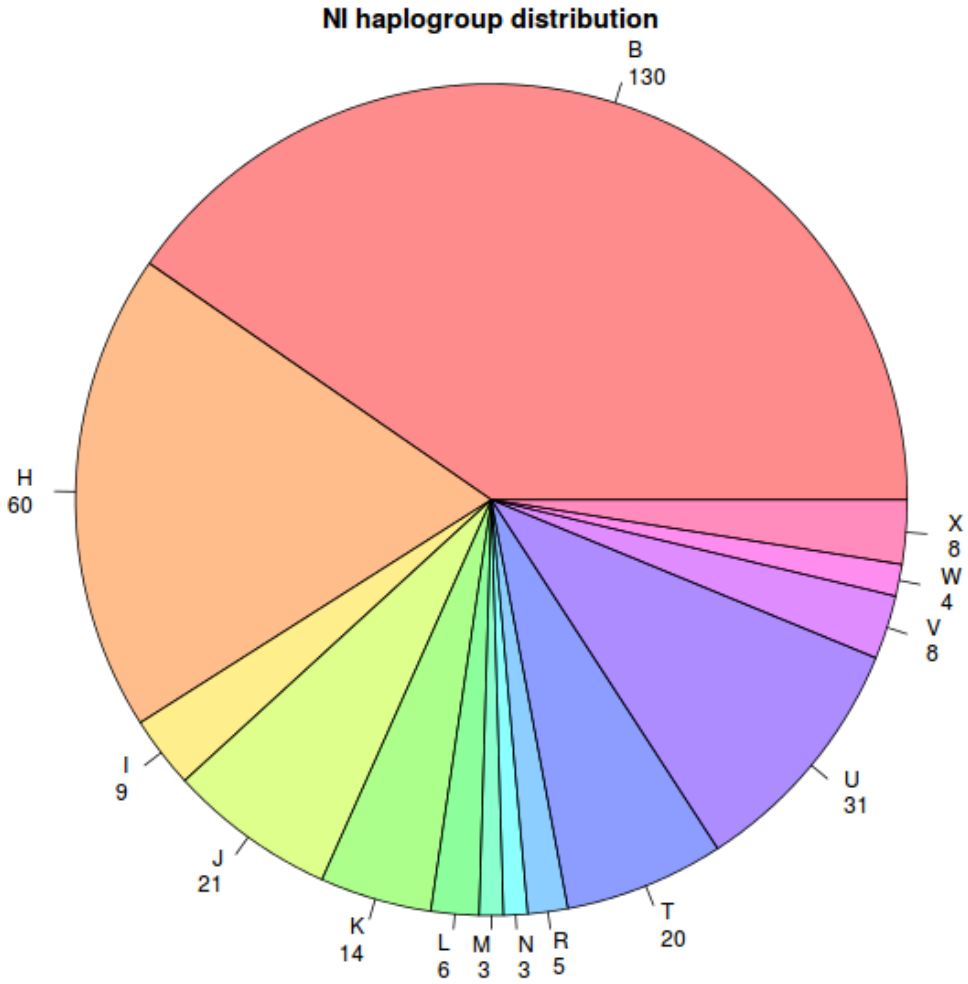
Table 28: Haplogroup analysis

CHR	SNP	BP	A1	TEST	OR	STAT	P
26	K	12	A	ADD	4.749	2.119	0.03408
26	R	10	A	ADD	4.478	1.364	0.1725
26	B	2	A	ADD	0.5751	-1.326	0.185
26	W	13	A	ADD	5.966	1.191	0.2336
26	X	6	A	ADD	2.088	0.7939	0.4273
26	H	5	A	ADD	1.452	0.7855	0.4322
26	U	1	A	ADD	0.7129	-0.4162	0.6773
26	J	3	A	ADD	0.7201	-0.4069	0.6841
26	T	7	A	ADD	1.271	0.2791	0.7802
26	I	9	A	ADD	3.27E-09	-0.00184	0.9985
26	L	11	A	ADD	2.59E-09	-0.00165	0.9987
26	V	8	A	ADD	1.95E-09	-0.0015	0.9988
26	N	4	A	ADD	3.24E-09	-0.00113	0.9991

As shown in Table 28 above, haplogroup K is significantly associated with migraine susceptibility. This finding is consistent with other studies which have found haplogroups to be significantly associated with complex disease [284]. This data suggests that individuals belonging to haplogroup K have an increased risk of developing migraine.

The haplogroups identified were mostly of European (H, I, J, K, U) and Asian (B) origin which is in line with historical records and supports the reports of European paternal and Polynesian maternal origins. Figure 31 shows the distribution of defined haplogroups in the sequenced individuals, with 130 individuals belonging to haplogroup B and 60 belonging to Haplogroup H. Given the maternal Polynesian origins of the Norfolk Island population the fact that the majority of individuals carry haplogroup B, which is defining of Asian and subsequently Polynesian lineages, corresponds to historical accounts. The European and other non-Asian haplogroups most likely originate from married-in individuals who have introduced new mitochondrial lineages into the population.

Figure 31: Distribution of haplogroups in NI sequenced individuals



The majority of sequenced individuals present with haplogroup B which is defining of the Polynesian lineage and corresponds to historical records. The European and additional non-European haplogroups most likely originate from married-in individuals.

Key Findings and Replication

The overall aim of this study was to obtain full mitochondrial genome sequence information from selected Norfolk Island individuals in order to identify mitochondrial variants associated with migraine susceptibility. The Norfolk Island pedigree is an ideal cohort for identifying variants involved in complex disease due to the reduced genetic and environmental heterogeneity. The power of advancing technology was utilised to achieve this goal through the development of a cost effective method capable of producing large volumes of sequence data quickly. In total 315 individuals from the Norfolk Island pedigree underwent full mitochondrial genome sequencing, producing a high level of coverage across the entire genome for each individual sequenced. These sequences were aligned to a reference genome (RSRS) and variants were called relative to this reference. Careful quality control and filtering processes were used to ensure that only true variants were called.

A series of statistical tools were used to compare identified variants found in migraine sufferers with unaffected individuals with the aim of identifying variants involved in migraine susceptibility. A logistic regression model factoring in for gender, age and relatedness was used to undertake this comparison. It is important to correct for the above mentioned covariates, to prevent bias and skewing of results. In order for the regression model to be valid, only variants with a MAF>0.01 were considered. These were defined as common variants. Out of the 201 common variants tested, only one variant showed a significant allele frequency difference between migraineurs and controls. This was a G>A substitution at position 930 of the mitochondrial genome. This SNP is located within the 12S rRNA subunit and is transcribed to produce part of the ribosomal complex which translates transcripts from all the coding regions of the mitochondrial genome into proteins.

The identified variant, mt 930 G>A (p=0.0233) passed the unadjusted significance threshold of $p<0.05$. If a Bonferroni correction were to be applied or the significance threshold was to be adjusted to account for multiple testing, then this variant would not meet the significance threshold. Many researchers question the stringency of correcting for multiple testing and it has been suggested that doing so unnecessarily invalidates potentially real results [355]. The logistic regression model indicates an odds ratio of 3.967 for mt 930 G>A which is very high. This means that an individual who carries the G allele is almost four times more likely to develop migraine than an individual that carries the A allele. Given the location in 12S rRNA and the strong association of this region with sensineuronal hearing loss, a neurological condition, it is biologically plausible for this mitochondrial region to be strongly involved in migraine pathogenesis.

The remaining 64 rare variants were analysed using a Fisher's exact test which is a more statistically valid approach to take when comparing small numbers of alleles. This test does still contain limitations as with all other statistical measures and the rarer a variant is, the more difficult it becomes to meaningfully gauge whether it has an effect on disease outcome [356]. Using this statistical measure it was found that two rare variants are associated with migraine susceptibility namely mt 11930 A>G (p=0.026) and mt 6480 G>A (p=0.026). Both of these variants were only identified in unaffected controls and were not found in migraine sufferers, suggesting that the non-ancestral allele confers a slightly protective effect. Again these variants would not meet the significance threshold if applying correction for multiple testing.

In order to better understand the role of these variants in migraine susceptibility, a replication study was undertaken on a large Caucasian migraine case control population. The methods for this study are described in detail in Chapter 6. All three mitochondrial variants which were identified as potentially contributing to migraine susceptibility were genotyped using a Sequenom plex and analysed using an appropriate logistic regression model which corrected for gender bias. The results are shown below in Table 29.

Table 29: Logistic Regression model of genotyped mitochondrial SNPs in a large migraine case control population

CHR	SNP	BP	A1	TEST	NMISS	STAT	P
26	rs41352944	930	A	ADD	1051	0.3023	0.7625
26	rs199476128	6480	A	ADD	1070	0.00088	0.9993
26	11930A>G	11930	0	ADD	1076	NA	NA

NMISS stand for non-missing and indicates the number of individuals successfully genotyped. As can be seen in Table 29, the genotyping assay was largely successful and over 1000 individuals were genotyped for each SNP. For the variant mt 11930 A>G which is a novel variant and has never been reported previously, no individuals were found to carry the G allele. All 1076 individuals typed carried the A allele, suggesting that this variant is either extremely rare or is a Norfolk specific variant. Given the unique genetic architecture specific to the Norfolk Island pedigree, it is likely that mt 11930 A>G is only found in the Norfolk pedigree. Studying other variants situated in the same region as mt 11930 A>G in an outbred population would be useful and provide evidence that *NADH* dehydrogenase is critical in migraine pathogenesis.

Both mt 930 G>A and mt 6480 G>A were found not to be significantly associated with migraine susceptibility in the tested population. However when considering how few individuals were found to be carrying the alternate allele as shown in Table 30 below, it is not surprising that the results were not significant. F_A represents the MAF in cases and F_U for controls. Similarly these variants may be very rare or Norfolk specific and further investigation of the 12S rRNA region and the *COX1* gene would be useful.

Table 30: Allele frequencies of mitochondrial variants tested in a large outbred migraine case control population

CHR	SNP	BP	A1	F_A	F_U
26	rs41352944	930	A	0.05273	0.04808
26	rs199476128	6480	A	0.001869	0
26	11930A>G	11930	0	0	0

In line with the aims of this study, sequence data was further assessed according to identified heteroplasmic variants and defined haplogroups. Calling heteroplasmic variants was more challenging and required a higher level of manual curation. Complex bioinformatic approaches were used to define heteroplasmic levels and assign appropriate calls. Heteroplasmic variants were coded categorically as either heteroplasmy present or absent and a logistic regression model was applied to assess differences between migraineurs and unaffected controls. This qualitative approach has limitations and a quantitative model of analysis would provide more accurate insights into the role of heteroplasmy in migraine. Using the available methods, 11 heteroplasmic variants were identified as being associated with migraine susceptibility.

The most significant heteroplasmic variant was found to have a protective effect against developing migraine with p=0.0104 and an odds ratio of 0.363. An odds ratio smaller than one indicates relative reduced disease risk. Individuals carrying the A allele at position 8697, even when only present in a proportion of mitochondria, are protected from developing migraine (Table 27). It was an unexpected finding that heteroplasmic variants play a greater role in migraine susceptibility than their homoplasmic counterparts. This could be explained by an accumulation of mutations reaching a threshold level within the mitochondrial cell population and causing disease once the threshold is exceeded. The idea of accumulated mutations contributing to neurological disease corresponds to the observation of the severe mitochondrial neuropathies, where disease progression accelerates as time goes on [357]. Further investigation in an outbred population would provide much needed insight.

With regards to haplogroup analysis, haplogroup K was found to be significantly associated with migraine susceptibility. Theories regarding this finding are discussed in Chapter 7. Overall 3 mitochondrial variants were found to be associated with migraine susceptibility, 1 common variant mt 930 G>A (p=0.0233) and 2 rare variants; mt 11930 A>G (p=0.026) and mt 6840 G>A (p=0.026). Unfortunately none of these were replicated in an outbred population indicating that investigation of other variants in the identified gene regions would be useful. In addition to the initial findings, 11 heteroplasmic variants were identified as being associated with migraine and further deep sequencing studies in other populations would be useful to see if the same effect is observed.

Discussion

Migraine and Stroke, Shared Pathogenic Mechanisms

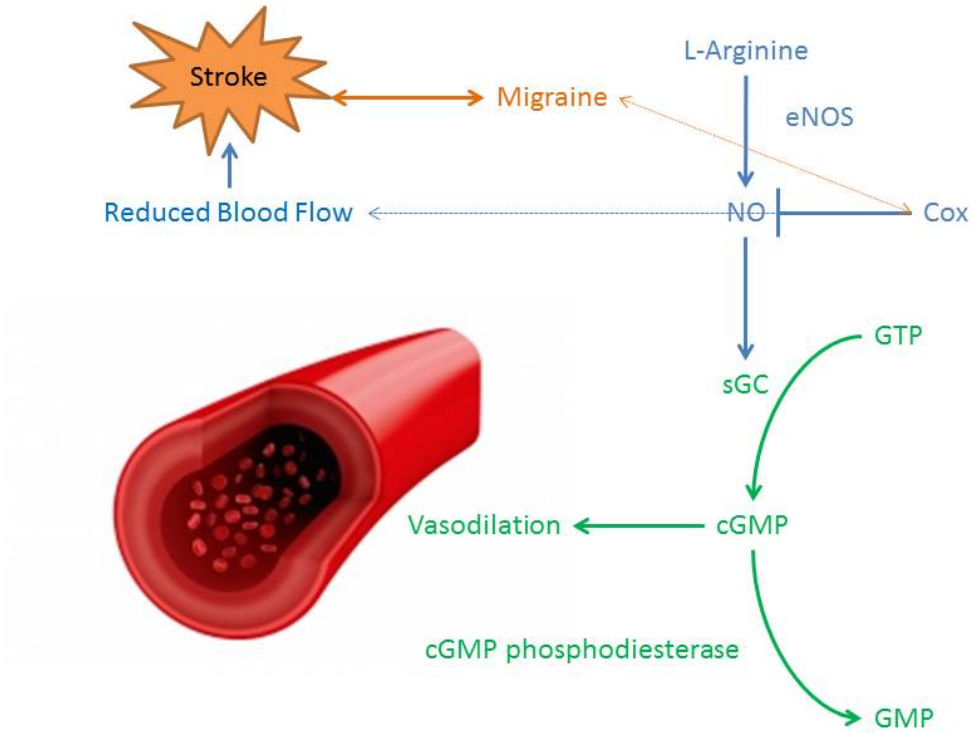
It has been well documented that migraine and stroke are co-morbid disorders. Recently it has been suggested that there may be common underlying mechanisms for the pathogenesis of both diseases and evidence is accumulating which strengthens the link between these conditions and suggests a shared pathway [358]. Data from epidemiologic studies have shown a high prevalence of migraine sufferers among stroke victims leading to the now established fact that these diseases are co-morbid [173, 359]. The finding of silent infarct-like brain lesions in patients suffering from migraine has strengthened the link between migraine and stroke, but has also raised the hypothesis that migraine could cause long lasting damage and in severe cases, especially in those individuals suffering regular attacks, result in stroke [360]. This hypothesis is further supported by the finding that stroke can occur during migraine with aura attacks, usually in more rare familial forms of migraine such as Familial Hemiplegic Migraine (FHM). This demonstrates a causal relationship between migraine and stroke and points to the idea of multiple MA attacks predisposing an individual towards having a stroke later in life [358, 361].

It has been suggested that migraine and ischemic stroke may actually be caused by a common pathogenic mechanism resulting in a spectral phenotype ranging from migraine all the way to stroke. It has also been suggested that migraine could be a progressive disorder and have long term effects [361]. Further evidence of a common pathogenic mechanism is the shared symptoms of migraine and stroke in certain neurological diseases such as CADASIL (cerebral autosomal-dominant arteriopathy with subcortical infarcts and leukoencephalopathy) and MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes)

[362, 363]. In these diseases and other similar neurological conditions, patients experience both severe migraine attacks as well as stroke. The overlap of symptoms in mitochondrial disorders such as MELAS and others, further suggests an underlying mitochondrial pathogenic mechanism.

As mentioned in the introduction mitochondria play an important role in ion homeostasis, protecting the cell from damaging ROS and controlling vascular tone through COX and NO. NO is a potent vasodilator and has been shown to bind to COX and through this binding action regulate vasodilation [364, 365]. Studies have shown that in vessels rich with COX or where there is an alteration in COX functionality, there may be sequestration of NO, preventing vasodilation during strokes suffered by patients who present with MELAS [333, 365]. Our results show a significant association with mt 6480 G>A situated within *COX1* and for the first time suggests that this could be a pathogenic mechanism in common migraine sufferers and not just individuals suffering rare mitochondrial disorders. In this study those individuals carrying the variant are slightly protected from migraine rather than being predisposed. However Norfolk Island has a very unique genetic architecture and further genetic studies examining other *COX1* variants in a large outbred population are warranted. Figure 32 below illustrates the proposed shared pathway between COX, NO, stroke and migraine.

Figure 32: Illustration of the interaction between COX, NO and resultant blood flow which affects both stroke and migraine.



Nitric oxide is a potent vasodilator. COX inhibits NO activity by binding to free NO in the blood, resulting in reduced blood flow. This COX, NO mechanism and resultant reduction in blood flow has been shown to occur in stroke. We hypothesize that this could be a common underlying pathogenic mechanism in both stroke and migraine attacks, which are known to be co-morbid. Our study provides genetic evidence that there is a link between COX and migraine, lending support to the hypothesis that genetic variation which alters the binding of NO to COX (either increase or decrease) results in a pathogenic mechanism common to both migraine and stroke. Abbreviations cGMP: cyclic guanosine monophosphate, GMP: guanosine monophosphate, GTP: guanosine triphosphate, sGC: Soluble guanylyl cyclase, eNOS: endothelial nitric oxide synthase, NO: nitric oxide

Mitochondrial Dysfunction and Migraine

Biochemical studies have shown that migraine sufferers have deficient levels of NADH dehydrogenase, citrate synthetase and COX compared to healthy controls. Research also found that migraine sufferers have impaired function in addition to lower levels of NADH dehydrogenase, citrate synthetase and COX leading to intravascular platelet dysfunction and vulnerability to oxidative stress [333, 336, 364, 366]. Our most significantly associated rare variant, mt 11930 A>G is situated in the mitochondrial gene encoding for *NADH* dehydrogenase, subunit 4 which forms a critical component of the oxidative phosphorylation chain. This is the first genetic evidence linking up with the already well documented biochemical evidence showing a link between NADH dehydrogenase and migraine.

Strong therapeutic evidence strengthens this link and presents the possibility of a new treatment tailored specifically to individuals with deleterious mutations in the *NADH* gene. NADH reductase deficiency has been successfully treated with riboflavin (vitamin B2), which has also been shown to be an effective prophylactic treatment for migraine sufferers [367]. Riboflavin is an essential compound that is necessary for the Electron Transport Chain (ETC) to function. It is a precursor to flavin adenine dinucleotide, which functions by transferring electrons to donors such as NAD and NADP [368]. Research has shown that patients with non H haplogroups respond better to treatment [338, 369], making a comprehensive genetic diagnosis essential for predicting the efficacy of treatment.

Biochemical, morphological and therapeutic studies all show that there is a relationship between mitochondrial dysfunction and migraine. Examination of muscle fibres in migraine sufferers has identified ragged red fibres, fatty COX^- fibres, paracrystalline inclusions and subsarcolemmal mitochondria which are all hallmark characteristics of defective OXPHOS

function and overall impaired mitochondrial function [334]. Biochemical studies suggest that dysregulation of Ca^{2+} ions due to mitochondrial dysfunction alters the signalling properties of neurons and this phenomenon has been linked to CSD and peripheral pain mechanisms. This pathology has been compared to the dysfunction of calcium channel genes resulting in FHM [26, 333].

Other biochemical studies have shown through P-MRS (phosphorus magnetic resonance studies) that migraine sufferers have a reduction in OXPHOS function [370]. Lactic acidosis has also been described in migraineurs, which is another hallmark characteristic of mitochondrial dysfunction. Perhaps the strongest evidence of all is response to therapy which targets different deficient mitochondrial pathways. Co-enzyme Q10, vitamin B2, Mg^{2+}, niacin, carnitine, topiramate and thioctic acid have all been shown to be effective therapeutics in the treatment of migraine [367, 368, 371, 372]. All of these molecules target mitochondrial mechanisms. Further study of nuclear encoding mitochondrial genes and regulators of mitochondrial function would be valuable [373].

12S rRNA and Migraine

12S rRNA forms part of the ribosomal machinery within the mitochondria for translation of transcripts into proteins. Mutations in this region are most commonly associated with sensorineural hearing loss (SNHL). The cochlea requires important mechanisms such as membrane depolarization, ion transport and transmitter release to take place in order to function correctly. These mechanisms are highly dependent on the ATP produced by mitochondrial oxidative phosphorylation and any impairment in function can result in deafness. Five rRNA mutations have been found to cause SNHL in the 12S rRNA gene, while no deafness-associated mutations in the 16S rRNA gene have been found thus far [374].

Given the importance of mitochondrial function for this sensitive neuronal process, it is possible that variants in the 12S rRNA region could affect other neuronal processes including lowering the threshold for a migraine attack. Further studies are needed to examine this area in more detail.

Conclusions

The three SNPs found to be significantly associated with migraine in the Norfolk Island population were investigated in a replication study described in Chapter 6, following on from this chapter. Mt 11930 A>G was not detected in any of the Australian outbred samples, suggesting that this novel variant is Norfolk specific. The other two SNPs identified to be significantly associated with migraine in Norfolk Island, mt 930 G>A and mt 6480 G>A were detected in very few case-control samples and were not found to be significantly associated with migraine, again suggesting that these variants are specific to Norfolk Island and do not play an important role in the Australian Caucasian population. However this study has still identified key mitochondrial regions that should be investigated further as other variants in these regions could play a role in migraine pathogenesis in the Australian population. It has also provided conclusive evidence for the first time that mitochondrial variation is linked to migraine susceptibility.

Significantly it appears that heteroplasmic variants which may be acquired during the lifetime may play a more significant role than inherited homoplasmic variants. We found 11 heteroplasmic variants to be significantly associated with migraine in the Norfolk Island population and as part of the future research aims, further investigation into this finding is warranted. Recent literature explores the idea of acquired mitochondrial variants which

accumulate and reach a threshold later in life causing complex disorders. It has been suggested that these acquired age related variants play a more critical role in complex disease than previously thought [357]. This is the first study which produced mitochondrial genome sequencing at a depth of coverage deep enough to address this question of heteroplasmy and we present valuable findings here which should be explored further.

Chapter 6: Nuclear Encoded Mitochondrial Proteins (NEMPs) and Migraine Association in the Norfolk Island Population

Similarly to Chapter 5, this results chapter is presented as a manuscript in preparation. A condensed version of the results presented here has been submitted to American Journal of Human Genetics for peer review.

Abstract

Migraine is a complex disease and multifactorial in nature with both genetic and environmental components contributing to the overall phenotype. Molecular genetic studies have identified a number of susceptibility regions in the genome which contribute towards disease risk, but a large proportion of the genetic variance still remains to be elucidated and novel avenues of exploration are needed to address this area. Biochemical, morphological and therapeutic studies provide strong evidence that mitochondrial dysfunction could be involved in migraine pathogenesis. The aim of this chapter is to explore the role of mitochondrial dysfunction in relation to migraine susceptibility from the molecular genetics point of view.

Norfolk Island is an ideal population for identification of complex disease traits as the genetic heterogeneity typical of complex disease is reduced. In our discovery phase using this unique population it was found that 667 NEMP SNPs are significantly associated with migraine in the Norfolk Island population. Of these SNPs, 21 were carried forward to a replication study utilising a large migraine case-control cohort. Overall findings identified 9 SNPs to be significantly associated with migraine in the genes *ELOVL6, SARDH, CSNK1G3* and the *PCDHG* family in both populations. Four variants out of the 9 significantly associated SNPs were found in to be in the *PCDHG* gene cluster, suggesting a particularly important role for this locus in migraine pathogenesis. Haplotype analysis shows that multiple risk alleles in the *PCDHG* gene region further increases the risk of developing migraine and this strengthens the link between migraine susceptibility and this locus. This is the first molecular genetic study to comprehensively investigate the role of NEMPs in migraine susceptibility and to conclusively identify a new link between mitochondrial dysfunction and migraine.

Introduction

Mitochondria are thought to be ancient eukaryotes which formed a symbiotic relationship with our ancestors. As a result these organelles have evolved as a key part of all mammalian physiology. The mitochondrial genome encodes for 37 genes, of which 13 polypeptides are translated by the mitochondrial ribosomes to form subunits of the OXPHOS chain. The remainder of the proteins, enzymes and signalling molecules are all encoded by the nuclear genome and imported into the mitochondrial matrix. In total 99.995 % of all the molecules involved in mitochondrial function are imported from the cytosol. To comprehensively investigate the role of mitochondrial dysfunction in migraine susceptibility, the Nuclear Encoded Mitochondrial Proteins (NEMPs) genes must also be considered. In this chapter the role of NEMPs in relation to migraine are investigated.

Mitochondria are the power houses of the cell and provide energy to the body's tissues in the form of adenosine tri-phosphate (ATP) via the oxidative phosphorylation (OXPHOS) chain. They also function critically in stabilising intracellular Ca^{2+} levels, protecting the cell from damaging Reaction Oxygen Species (ROS) and controlling vascular tone through Cytochrome C Oxidase (COX) and Nitrogen Oxide (NO). Mitochondria provide energy for the body through other metabolic pathways including production of ketones from fatty acids and anaerobic ATP synthesis. In mitochondrial disorders the tissues which are most adversely affected are those with the highest energy requirements namely muscle and nervous tissue. Strong evidence from biochemical, morphological and therapeutic studies show a link between mitochondrial dysfunction and migraine pathogenesis [47, 280, 333, 335-338]. Genetic studies have been limited by sample size and molecular data making this an area that needs to be addressed.

The mitochondrial genome encodes for 13 polypeptides which are translated on the mitochondrial ribosomes and form structural subunits of the OXPHOS enzyme complexes. Complex I is composed of 46 polypeptides of which 7 are mitochondrially encoded, complex III includes just 1 mitochondrially encoded polypeptide our of 11 components and complex IV has 3 out of 13. Similarly only 2 out of 16 proteins in complex V are encoded by the mitochondria [357]. All the other components of OXPHOS including the entire subunit II are encoded by nuclear genes. In addition to structural components of OXPHOS, all the mitochondrial metabolic enzymes, transcriptions factors and other regulatory molecules which govern mitochondrial function are nuclear encoded [375]. Current estimates are that more than 1000 proteins are encoded by the nuclear DNA and imported into the mitochondrial matrix [343]. The vast majority of active molecules involved in mitochondrial function are imported into the matrix from the nucleus in this way with less than 1 % of mitochondrial function attributed to the mitochondrial genome itself [376]. The aim of this study is to investigate these nuclear encoded mitochondrial proteins (NEMPs) which are critical components of mitochondrial dysfunction in relation to migraine susceptibility.

Norfolk Island is a genetically isolated population situated off the coast of Australia, best known from the “Bounty on the Mutiny” historical account. This isolate is an ideal population for identification of complex disease traits as the genetic heterogeneity typical of these diseases is reduced. Geographical isolation further reduces environmental heterogeneity and environment-gene interactions, increasing the chance of identifying true causal variants [341, 342]. A large Australian outbred migraine case-control population provides a further valuable replication cohort. This is the first molecular genetic study to comprehensively investigate the role of NEMPs in migraine susceptibility.

Norfolk Island Population

Sample Selection

Genotyping data from previous studies were selected for all known NEMP genes in n=315 individuals from the Norfolk Island pedigree. In total 16280 SNPs in known NEMP genes were genotyped across the entire cohort of individuals. The selected individuals also underwent full mitochondrial genome sequencing on the ion torrent platform, making up a full subset of samples where full mitochondrial information was available in conjunction with NEMP data. This study design was selected to comprehensively answer the research question posed which is: what is the role of mitochondrial dysfunction in relation to migraine susceptibility. Mitochondrial dysfunction involves both the mitochondrial genome and the 1000+ proteins which are encoded by the nuclear genome and imported into this organelle.

Data Files

Data were formatted according to tped and tfam file specification for analysis in Plink v1.07. These are similar to ped and map files, but contain some transposed elements in the opposite orientation. The tfam file contained columns for family ID, individual ID, genomic co-ordinates of each SNP and phenotype. For this analysis the phenotype included was migraine where 1=unaffected and 2=affected. The tped file had columns for each SNP tested and the corresponding genotype for all 315 samples selected. A covariate file was utilised to correct for gender, age and relatedness. The covariate file included columns for family ID and individual ID which matched the tfam file. Columns for gender, relatedness and age were also included corresponding to COVAR1, COVAR2 and COVAR3 respectively in the results.

Analysis

Plink v1.07 was used for all analyses. The selected NEMP SNPs were tested for association with migraine using a logistic regression model which factored in for the covariates gender, relatedness and age. The final additive test value was considered for significant association with migraine. Results were ranked according to p value from smallest to largest.

Replication Study in Migraine Association Population

Sequenom Plex Design

The 32 previously selected NEMP SNPs (Table 40, Appendix B) in conjunction with 3 mitochondrial SNPs were grouped together for a plex design using Sequenom *Assay Design Suite Software 1.0.* The 3 mitochondrial SNPs were prioritised to be included in the plex during the design process as they were found to be significantly associated with migraine susceptibility (see previous mitochondrial genome sequencing Chapter 5). Due to primer interactions only a subset of the desired SNPs can be included in the final design. In total 21 NEMP SNPs and 3 mitochondrial SNPs were selected for the final assay. These are shown in Appendix B. Known SNPs were included in the assay design using their designation rs number. One SNP, namely an A>G change at position 11930 of the mitochondrial genome, was novel and the convention used is given in (Table 41) Appendix B.

Sample Selection

Migraine cases and controls were recruited from the local South East Queensland region as previously described [377]. They were all of Caucasian origin, and diagnosed as having MA or MO based on criteria specified by the International Headache Society. An unaffected control group with no family history of migraine was matched for age (+/- 5 years), sex and

ethnicity. Blood samples obtained from patients were collected through the Genomics Research Centre clinic and DNA was extracted using a salting out method. Approval for the study protocol was acquired from QUT's Ethics Committee. This Australian Caucasian migraine case-control population was genotyped for all 24 selected SNPs. In total 1128 individuals comprising 544 cases and 584 controls were genotyped. Migraine sufferers included both MA and MO subtypes, with 381 MA cases and 163 MO cases. As is typical of migraine, samples were skewed in a 3:1 gender ratio with 294 males and 834 females included.

Molecular Techniques

Matrix-assisted desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry was used to genotype each sample in a multiplexed reaction [378]. The Sequenom instrument and accompanying software was used to carry out all genotyping work as well as the design of amplification and genotyping primers. Unique primer pairs were designed as described previously to amplify each region of interest such that only one region would be amplified for each primer pair. A genotyping primer was then designed to anneal directly adjacent to the SNP of interest, allowing for enzymatic extension of a single dideoxy base pair following PCR amplification. The extended primer was then robotically dispensed onto a silicon chip preloaded with matrix, ionised and released. The time of flight was recorded for each fragment. The density of each genotype differs slightly and can be interpreted as a scatter plot.

Dilution of Annealing and Extension Primers

Primers were ordered and arrived resuspended at a stock concentration of 100 uM. All PCR forward and reverse primers were diluted and pooled to a 0.5 uM concentration in a final volume of 2000 uL. Using the formula:

C1* C2 = C2* V2

(100 uM)*V1=0.5 uM*2000 uL

V1=1000/100

=10 uL of each primer

Where C1=initial concentration

C2=final concentration

V1= initial volume

V2=final volume

A total of 24 forward primers and 24 reverse primers were included bringing the total volume of primer up to 480 uL. To reach V2 1520 uL of deionised water was added to the pool. This resulted in an equimolar pool of primer pairs, each at a final concentration of 0.5 uM.

Extension primers were ordered at a resuspended concentration of 500 uM and pooled according to individual molecular weight as shown in Table 42 in Appendix B. A total initial volume of 832.57 uL of undiluted extension primers were made up. A test extension primer reaction was run on the Sequenom and concentrations were adjusted accordingly. An additional 4 parts and one half volume were added for rs2073815 and rs11748256 respectively. This was done to obtain an approximately equal distribution of each extension primer represented in the total pool. A total volume of 2000 uL was then made up using distilled water.

Amplification and Purification Steps

Before spotting samples onto the silica chips, three preparations were involved namely a PCR, SAP clean and extension PCR. The first PCR amplified each region of interest from genomic DNA, using forward and reverse primers. This reaction works in the same way as a standard PCR. After amplification the product was purified using shrimp alkaline phosphatase (SAP) to remove any excess dTNPs or impurities from the solution. Samples were incubated at 37 °C for 40 min followed by 85 °C for 5 min and held at 12 °C. Following the clean the final extension reaction was undertaken. During this step each extension primer amplifies the region of interest adjacent to each SNP assayed. The tables below outline the reaction conditions used for these steps.

Table 31: PCR Preparation

Reagent	Final Concentration in 5 uL Rxn	Volume of Reagent in 5 uL Rxn (uL)
Deionised water	-	1.8
10x PCR Buffer	2x	0.5
25mM $MgCl_2$	2 mM	0.4
25mM dNTP Mix	500 uM	0.1
1uM Primer Mix	0.5 uM	1.0
5U/uL Enzyme	1 U/rxn	0.2
DNA	10ng	1.0

Table 32: PCR Thermocycling Conditions

Temperature	Time	Cycles
95 °C	2 min	1
95 °C 56 °C 72 °C	30 sec 30 sec 60 sec	45
72 °C	5 min	1
12 °C	∞	Hold

Table 33: SAP Reaction

Reagent	Final Concentration in 7 uL Rxn	Volume of Reagent in 2 uL Mix
Deionised water	-	1.53
SAP Buffer	0.234 x	0.17
SAP Enzyme	0.5U	0.30

Table 34: Extension Primer Reaction

Reagent	Final Concentration	Volume of Reagent in 2 uL Mix
Deionised water	-	0.619
iPLEX Buffer	0.222 x	0.2
IPLEX Termination	1x	0.2
Extend Primer mis	14 uM	0.94
iPLEX Enzyme	1x	0.041

Table 35: Thermocycling conditions used for extension reaction

Temperature	Time	Cycles
94 °C	30 sec	1
94 °C 52 °C 80 °C	5 sec 5 sec 3 sec	40 (5)
72 °C	3 min	1
12 °C	∞	Hold

Desalting Samples

A final purification reaction was performed before dispensing samples to a chip. Resin was used to remove any remaining impurities from the solution. Clean resin was spread out across the accompanying 96 well plate mould so that exactly 15mg was added into each well. The resin was left to dry for 10 min and added after the addition of 41 uL of deionised water to each well. The plate was sealed and rotated for 15 min, followed by a 5 min spin in the centrifuge at 3200 x g.

Dispensing to Chip

The nanodispenser robot was used to distribute each sample onto a chip. Calibrant was added and used to determine the optimum dispense speed each day. On average a dispensing speed of 80 mm/sec was used. A 96 well to chip mapping method was used and all maintenance was adhered to including daily (ethanol) and weekly (NaOH) clean.

Acquiring Data with MALDI-TOF

Plain text files were created with sample information for each chip that was set-up and run. The Typer software was used to import the assay and sample information for each new plate

created. Once created, plates were linked with the corresponding chip using ChipLinker software. The iPLEX chemistry was selected with Genotype+Area data collection. All SpectroCHIPs were run on the MassARRAY Typer Workstation with settings for iPLEXGold.

Data Files

A PED file and MAP file were created using the genotype data generated by Sequenom. The PED file was drawn up in excel and included columns for Family ID, Individual ID, Paternal ID, Maternal ID, Sex, Phenotype and Genotype. Since this was a case-control study, individuals are not related and Family ID was assigned as a unique numerical value for each sample. The Individual ID corresponded to the information imported from the text files where samples had a unique identifier corresponding to our databases. Paternal and Maternal ID were set to 0 as all individuals were unrelated. Sex was assigned to each sample where 1=male and 2=female. The phenotype of interest in our study was migraine where 1=unaffected and 2=affected. A separate PED file was also created for migraine subclassifications so that MA and MO could be compared to controls to identify any associations with these sub-classifications.

The MAP file included headings for chromosome location, rs identifier or SNP ID, genetic distance between markers in cM and the base pair position of the marker. Genetic distance was set to 0 for the purposes of this analysis. All PED and MAP files were converted to text files and subsequently renamed to .ped and .map files respectively.

Analysis

Plink v1.07 was used for all analysis. A logistic regression model was used to test for association between the SNPs genotyped and migraine susceptibility. The model was adjusted for the covariate gender to avoid skewing of results due to sex bias. A regression model was utilised for overall migraine cases as well as the subtypes MA and MO. In total 1128 individuals comprising 544 cases and 584 controls were included in the analysis. Subtype analysis comprised 381 MA cases and 163 MO cases. Linkage disequilibrium was calculated across the 24 SNPs tested and haplotype blocks were constructed using this data. Association between haplotypes and migraine susceptibility was calculated in both Plink and Haploview. Hardy Weinberg Equilibrium (HWE)

was calculated and a threshold of $p<0.002$ was used taking into account multiple testing (0.05/24).

The command prompts used are outlined below:

Logistic Regression: (factoring in for sex)

--plink –-noweb –-ped Migraine_Overall.ped –-map NEMP_MAP1.map –-logistic ---sex

HWE:

--plink –-noweb –-ped Migraine_Overall.ped –-map NEMP_MAP1.map –-hardy

LD:

--plink –-noweb –-ped Migraine_Overall.ped –-map NEMP_MAP1.map –-r

Haplotype Block Construction:

--plink –-noweb –-ped Migraine_Overall.ped –-map NEMP_MAP1.map –-hap plink.blocks –-hap-freq

Haplotype Association:

--plink –-noweb –-ped Migraine_Overall.ped –-map NEMP_MAP1.map –-hap-assoc

For MA and MO analysis, the appropriate PED file was substituted each time. For MA: MA_Overall.ped and for MO: MO_Overall.ped

Results

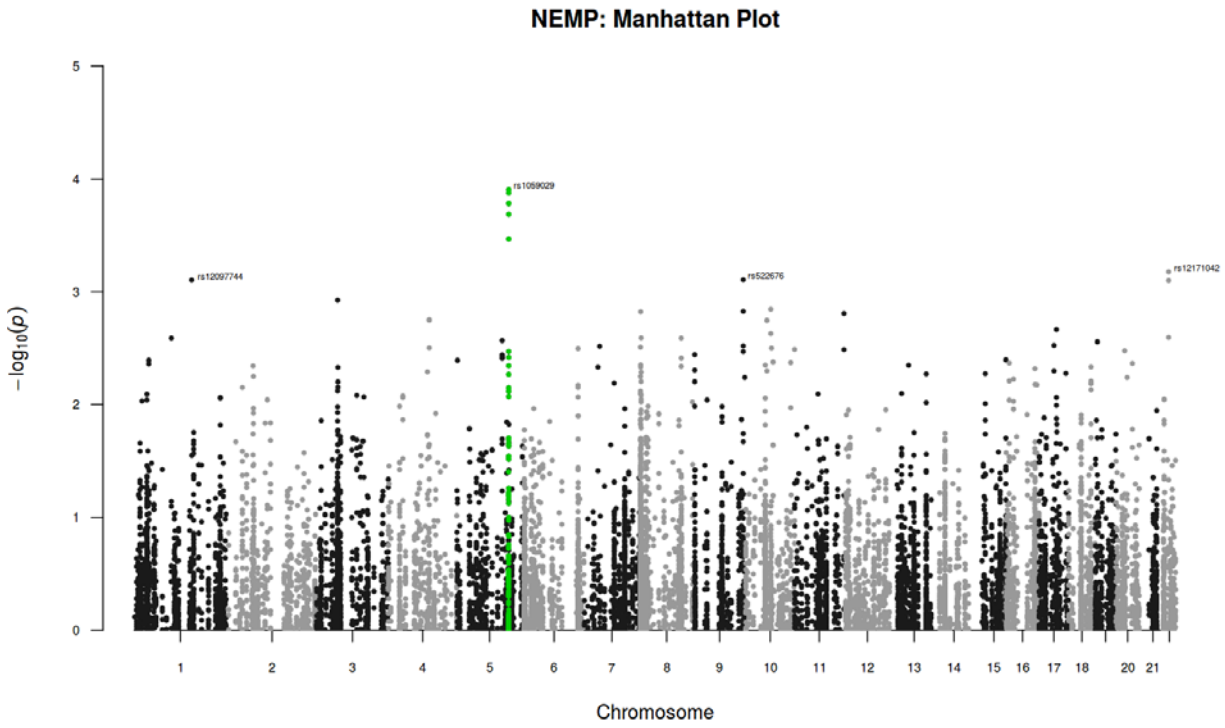
Norfolk Island

The initial logistic regression analysis using Norfolk Island samples showed that in total 667 NEMP SNPs are significantly associated with migraine. The results of the logistic regression are shown in the Manhattan plot below, Figure 33. A clear peak can be seen on chromosome 5, coloured in green, which shows the link between migraine and the *PCDHG* gene cluster. The logistic regression model factored in age, gender and relatedness and tested for association between 16280 previously genotyped NEMP SNPs and migraine susceptibility. It was found that 667 NEMP SNPs were significantly associated with migraine in the Norfolk Island population with the most significant associations showing $p<0.0001$. The genes which were found to be most significantly involved in migraine susceptibility in the Norfolk Island population are CUB and Sushi multiple domains (*CSMD*) 1 and 3, phosphatidylserine decarboxylase (*PISD*), fatty acid elongase 6 (*ELOVL6*), casein kinase 1 gamma 3 (*CSNK1G3*), sarcosine dehydrogenase (*SARDH*) and protocadherin gamma (*PCDHG*) C3 and B4.

In order to select SNPs for a Sequenom plex results were prioritised according to a combination criteria of smallest p value and genomic location. Multiple SNPs were selected

from genes which recurred as being significantly associated with migraine. In this way it was attempted to represent a spread of SNPs across the most significantly associated genes in Norfolk Island for replication. A list comprising 32 of our highest priority SNPs is shown in (Table 40) Appendix B. Due to primer interactions only a subset of the desired SNPs could be included in the final design and 21 NEMP SNPs were included for replication in an outbred case-control population.

Figure 33: Logistic regression analysis of genotyped NEMP SNPs in Norfolk Island



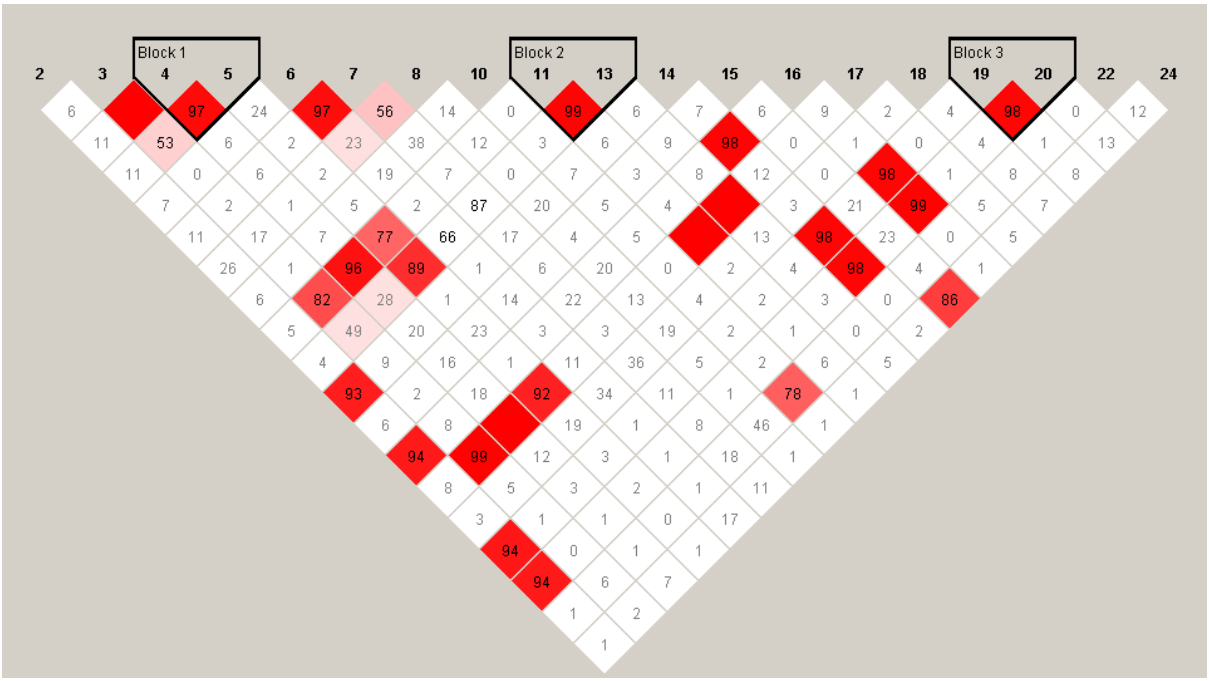
Each of the 16820 SNPs included in the logistic regression model are represented by a dot according to genomic location. Points are arranged according to chromosome number as shown on the x axis and the p value is reflected on the y axis. Shown in green is a significant peak on chromosome 5, in the PCDH gene region where multiple variants show a highly significant association with migraine susceptibility.

Replication Study

The logistic regression model showed a significant association with several of the tested SNPs and migraine association as shown in (Table 43) Appendix C. When investigated according to migraine subtype it is clear that these markers are significantly associated with MA and not at all with MO. It is possible that the results were biased in favour of MA association due to the low number of MO cases available in our migraine case cohort, however for the most part the evidence suggests an MA specific association. When correcting for multiple testing, only a single SNP was out of HWE (rs13361997) as shown in (Table 44) Appendix C. This SNP was not found to be significantly associated with migraine and should be disregarded from further analysis and removed before publication.

The haplotype association test showed 3 haplotype blocks amongst the tested SNPs with incomplete LD between SNP 4 and 5, 11 and 13 and 19 and 20 as shown in Figure 34 below. The composition of each haplotype block is illustrated in Table 45 Appendix C which shows the grouping of SNPs and the correlation between each haplotype block and migraine susceptibility. Haplotype block 1 and 2 show significant association with migraine susceptibility. Given the complex nature of migraine which is a multifactorial disease, it is important to consider interactions between SNPs and to examine them as groups and not just individual variants.

Figure 34: LD plot of tested SNPs as calculated and illustrated in Haploview

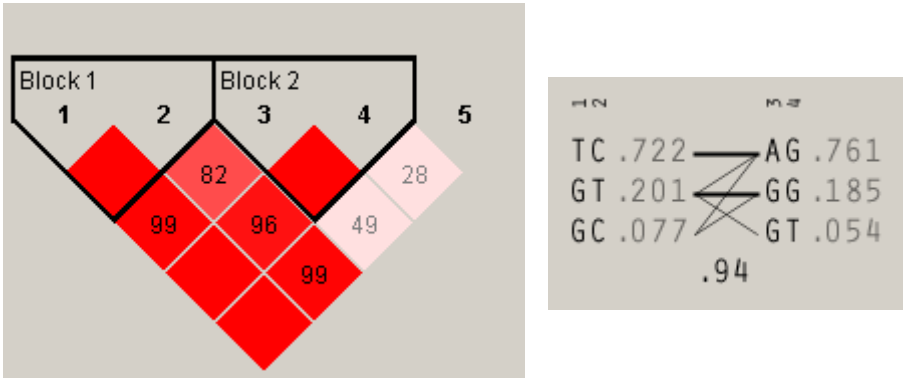


The haplotype association test showed 3 haplotype blocks amongst the tested SNPs with incomplete LD between SNP 4 and 5, 11 and 13 and 19 and 20. Linkage disequilibrium is illustrated in red.

The SNPs in *PCDHG* were further considered according to LD and haplotype analysis was undertaken in Haploview 4.2. Six SNPs located in the *PCDHG* gene cluster were included on the Sequenom assay and genotyped across our case-control population. One SNP, rs13361997 was not in HWE equilibrium and excluded from further analysis. A ped and map file was drawn up for the five SNPs located in *PCDHG* to assess LD between tested SNPs in this gene region and for further haplotype based association testing. As seen in Figure 35 below,

there is a high level of LD between the *PCDHG* SNPs with 2 haplotype blocks estimated. The haplotype association test showed that individuals with the G allele for both rs11748256 and rs1195229 are at an increased risk of developing migraine. The results suggest that multiple risk alleles in the *PCDHG* gene region increases the risk of developing migraine and this further strengthens the link between migraine susceptibility and this locus. Given the complex nature of migraine which is a multifactorial disease, it is important to consider interactions between SNPs and to examine them as groups and not just individual variants. Table 46 in Appendix C shows the composition of each haplotype block and the correlation between each haplotype block and migraine susceptibility.

Figure 35: Haplotype analysis of 5 tested PCDHG variants in outbred Australian population



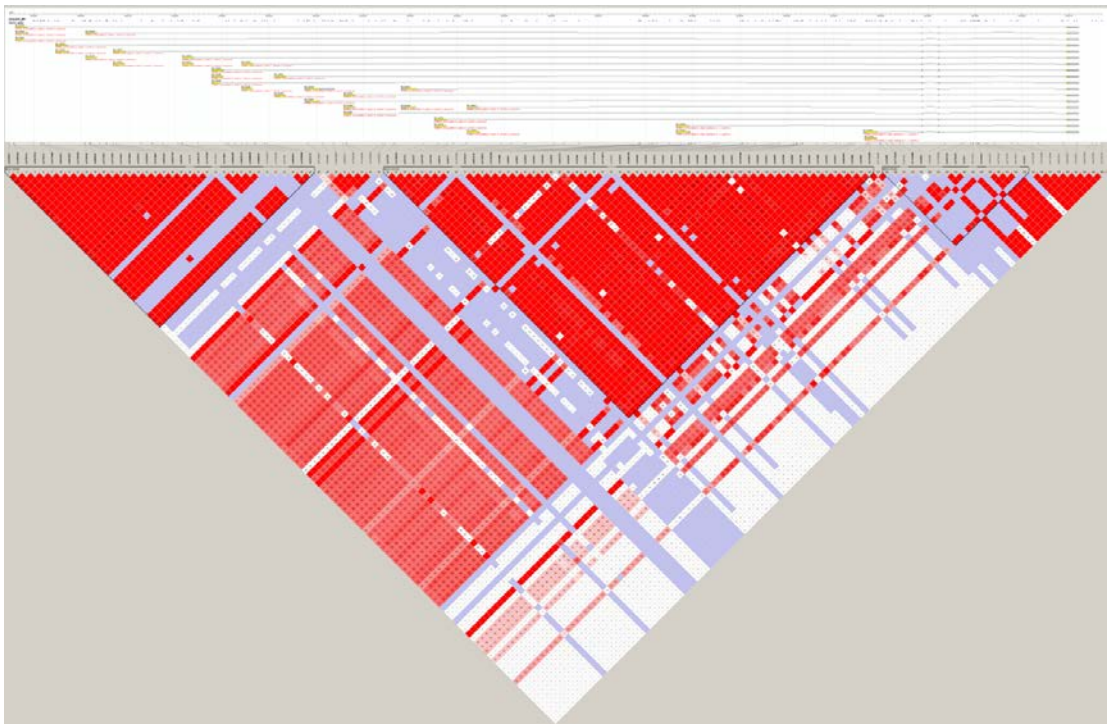
The 5 tested SNPs located within the *PCDH* gene region show significant LD as depicted by the red blocks. Individuals with multiple risk variants in this region have a further increased risk of developing migraine.

Linkage disequilibrium was considered across the entire *PCDHG* gene region using available data from HapMap and compared between different population groups. Norfolk Island is a unique admixture of paternal European ancestry in combination with Polynesian maternal origins. The Polynesian lineage descends from an Asian background and so the CEU and

CHB population groups were used for this comparison. As can be seen in Figure 36 and 37 below there are minimal differences in LD between these population groups for the *PCDHG* gene region.

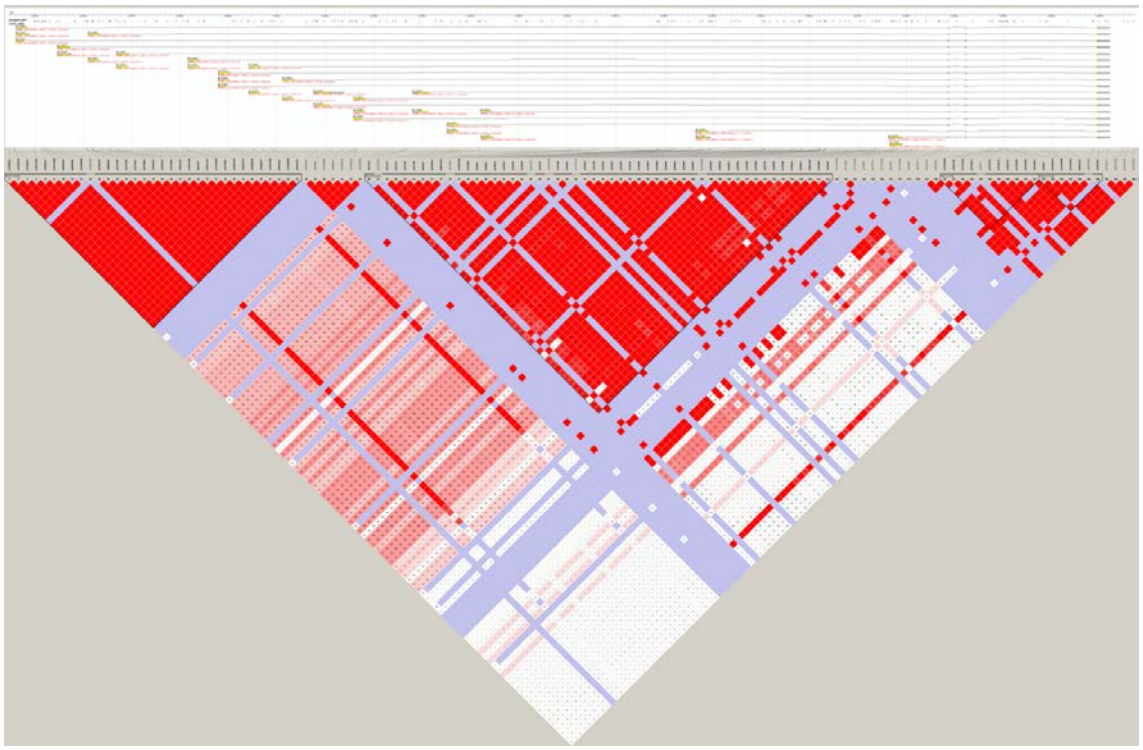
Haplotype blocks are slightly more broken up in the Asian compared to European population. To compare these populations with our unique Norfolk Island population, available genotype data was extracted for 165 SNPs across the *PCDHG* gene region and analysed in Haploview. Plink v1.07 was used to extract the relevant SNPs and generate new ped and map files to load into Haploview. As can be seen in Figure 38 below, LD in the *PCDHG* gene region in the Norfolk Island population more closely resembles Asian architecture than the European equivalent. The resolution is limited by the number of SNPs with available genotype data and more in-depth typing in this region may provide a clearer picture.

Figure 36: LD across the *PCDHG* gene region in the Ceu population according to HapMap



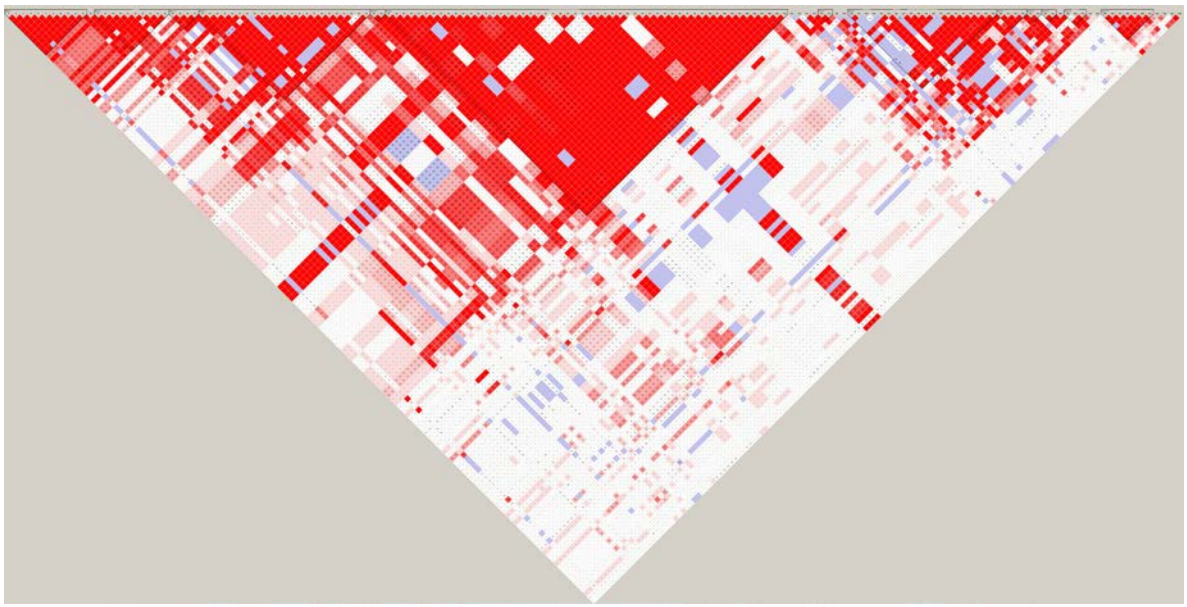
The large red blocks indicate a high level of LD across *PCDHG* in the Caucasian population. Large portions of this gene region are inherited across generations together. Variants within the same LD block are likely to be passed onto offspring together.

Figure 37: LD across the PCDHG gene region in the CHB population according to HapMap



Haplotype blocks are slightly more broken up in the Asian (CHB) compared to European populations.

Figure 38: LD across the *PCDHG* gene region in the Norfolk Island population according to available genotype information



165 SNPs across the *PCDHG* gene region were included from the Norfolk Island population illustrating that LD in the *PCDHG* gene region in the Norfolk Island population more closely resembles Asian architecture compared to the European equivalent.

Key Findings

The Norfolk Island pedigree was used as a discovery population to identify variants in mitochondrial related genes or NEMPs which are involved in migraine susceptibility. The majority of functional proteins and enzymes are imported into the mitochondrial matrix after being transcribed by nuclear encoded genes. In order to comprehensively investigate mitochondrial dysfunction in relation to migraine susceptibility, these 1000+ NEMP genes need to be investigated. Genotype data available from previous GWAS and genotyping studies was utilised to investigate all typed NEMP variants in the Norfolk Island population. The same 315 individuals which underwent full mitochondrial genome sequencing were

selected for statistical analysis. This provided a complete data set for these 315 individuals in terms of mitochondrial dysfunction caused by both the mitochondrial genome and mitochondrially related NEMP genes.

In total genotype data was available for 16280 previously genotyped NEMP SNPs in the selected individuals and a logistic regression model was used to compare variant frequencies between migraineurs and unaffected individuals. The model corrected for relatedness, gender and age in order to prevent bias or skewing of results. After correcting for all covariates, 667 NEMP SNPs were found to exceed the threshold $p<0.05$ for significant association. An adjusted threshold for multiple testing was not used as this process disregards valid findings [355]. The significantly associated SNPs were ranked according to p value, and it was found that variants in *PCDHGB5* are the most significantly associated autosomal NEMP variants involved in migraine susceptibility. Table 36 provides a brief summary of the most significant findings. A more comprehensive list of significantly associated SNPs can be found in full in Appendix A.

Table 36: Most significantly associated autosomal NEMP SNPs associated with migraine in the Norfolk Island population

Chr	Gene	Position	Allele	Test	Odds Ratio	Chi Square	P Value
5	PCDHGB5_rs1059029	140822723	G	ADD	2.182	3.838	0.0001242
5	PCDHGB5_rs3749767	140789933	T	ADD	2.17	3.821	0.0001328
5	PCDHGB5_rs13361997	140863674	C	ADD	2.37	3.768	0.0001648
5	PCDHGB5_rs11952292	140871249	T	ADD	3.061	3.712	0.0002058
5	PCDHGB5_rs6860615	140862840	G	ADD	1.906	3.582	0.0003411
22	PISD_rs12171042	32011225	C	ADD	1.83	3.403	0.0006661
9	SARDH_rs522676	136579589	C	ADD	0.485	-3.358	0.0007839

As shown in Table 36 above, the top 5 SNPs are all located in the *PCDH* gene region with rs1059029 (p=0.0001242), rs3749767 (p=0.0001328), rs13361997 (p=0.0001648), rs11952292 (p=0.0002058) and rs6860615 (p=0.0003411) all presenting with p<0.0004. The corresponding odds ratios are also high, where an individual carrying the G allele for rs1059029 is twice as likely to develop migraine compared to an individual who carries the ancestral allele. Similarly the other identified variants in the *PCDH* gene region confer significant disease risk and increase an individual's risk of developing migraine on average by two fold.

SNPs were prioritised according to genomic location and p value as shown in Appendix B and a selection of 21 SNPs was carried forward to a replication study. The three significantly associated mitochondrial SNPs were also included in the final plex. This part of the study is described in Chapter 5. While Norfolk Island is a valuable population for discovery of genetic variants involved in complex disease, any discoveries need to be replicated in an outbred population to assess the relevance of these variants in the population as a whole. A large

outbred Australian Caucasian migraine case control population was genotyped using the Sequenom platform in order to assess the role of the 21 prioritised NEMP SNPs in the general Australian population. A logistic regression model was used to assess differences in allele frequencies between cases and controls. The model was corrected for gender to avoid skewing of results and migraine was further assessed according to subtype.

As shown in Table 38, eight out of the 21 tested NEMP variants were found to be significantly associated with migraine susceptibility. This association was shown to occur in the MA migraine subtype and is of particular relevance to this patient group. These nine variants are significantly associated with migraine in both the genetically isolated Norfolk Island population and the outbred Australian Caucasian population, presenting strong evidence that NEMPs play a key role in migraine pathogenesis and that further focus on this area is warranted. Variants playing a key role in MA were identified in the genes *ELOVL6* ($p=2.5 \times 10^{-5}$), *SARDH* (p=0.000248), *CSNK1G3* (p<0.009) and the *PCDHG* (p<0.03) family. Four variants identified were found in to occur in the *PCDH* gene cluster, suggesting a particularly important role for this locus in migraine pathogenesis. Interestingly the most significant finding in *ELOVL6* rs7681294 ($p=2.5 \times 10^{-5}$) confers a protective effect indicated by an odds ratio of 0.6715. Table 37 illustrates the relative disease risk of each significant finding.

Given the finding that four out of the eight significantly associated variants (Table 38) in both the Norfolk Island population and the outbred Australian Caucasian population are found in the *PCDH* gene region, further haplotype analysis was conducted. It was found that individuals carrying multiple risk variants have a further increased risk ($p=3.318 \times 10^{-5}$) of

developing migraine compared to individuals only carrying a single risk variant. Specifically the haplotype association test showed that individuals with the G allele for both rs11748256 and rs1195229 are at the greatest increased risk of developing migraine as shown in Table 37 below.

Table 37: Haplotype analysis between tested *PCDH* gene variants in outbred population

Loc	Hap	F_A	F_U	CHISQ	DF	P value	SNPs
H1	GGT	0.04462	0.06176	3.036	1	0.08143	rs11748256|rs6860615|rs11952292
H1	GGG	0.2245	0.1537	17.23	1	3.318e-005	rs11748256|rs6860615|rs11952292
H1	AGG	0.04352	0.03937	0.2279	1	0.6331	rs11748256|rs6860615|rs11952292
H1	ATG	0.6873	0.7452	8.638	1	0.003292	rs11748256|rs6860615|rs11952292

LD across the *PCDH* gene region was assessed using available HapMap data in European and Asian populations and compared to LD Norfolk Island in the same region calculated on the available genotype data. As shown in the results, LD in the *PCDHG* gene region in the Norfolk Island population more closely resembles Asian architecture than the European equivalent.

Overall 667 NEMP SNPs were identified in the Norfolk Island population as being significantly associated with migraine susceptibility and 21 prioritised SNPs were carried forward to replication. In total 9 NEMP SNPs were found to be significantly associated with migraine susceptibility in both the Norfolk Island and Australian Caucasian population, presenting strong evidence that NEMP variants modulate mitochondrial function in a manner which corresponds to migraine pathophysiology. In particular the *PCDH* gene region was

identified as a key component of this process, with four out of the nine significantly associated variants found in this region. Further investigation into the role of NEMPs in migraine susceptibility would be useful.

Discussion

Mitochondria function primarily to produce a constant supply of energy to the cells of the body in the form of ATP. The most efficient conversion of calories from our fuel intake (food) into useable energy is through the oxidative phosphorylative chain under oxidative conditions where glucose is converted to ATP. The main metabolic pathways include glycolysis, the conversion of acetyl-CoA to GTP and other intermediates through the citric acid cycle, the pentose phosphate pathway, the urea cycle, fatty acid oxidation and gluconeogenesis. Molecules from our dietary intake are metabolised according to their properties and the end products are passed along the OXPHOS units in the mitochondria to produce energy [357].

Reactive Oxygen Species (ROS) are produced as bi-products of the energy conversion process and can have damaging effects on cells if they are allowed to accumulate [379]. Additional functions of the mitochondria include calcium homeostasis which is critical for neuronal function and initiation of apoptosis. Mitochondria occur in proportion to each tissue's energy requirements with muscle and nervous tissues containing several thousand mitochondria per cell [380]. It has been well established that mitochondrial dysfunction affects the tissues with the highest energy requirements and that the most severe mitochondrial disorders are neuromuscular diseases [381]. It has been hypothesised that the role of mitochondrial dysfunction in neurological conditions has been overlooked by the medical community and that further scientific investigations in this arena are warranted [357].

This is the first study to comprehensively investigate the role of NEMPs which are critically involved in mitochondrial function in relation to migraine susceptibility.

Our logistic regression model which factored in for the covariate gender showed a significant association between a number of SNPs and migraine susceptibility. Variants playing a key role in MA were identified in the genes *ELOVL6*, *SARDH*, *CSNK1G3* and the *PCDHG* family. Four variants were found to be in the same gene cluster, suggesting a particularly important role for this locus in migraine pathogenesis. Further investigation of the *PCDHG* gene family would be useful to pursue in migraine affected families and other outbred population groups. Information about each SNP is summarised in Table 38 below.

Table 38: Most significantly associated variants in relation to migraine susceptibility and corresponding locus information

Chr	SNP	Odds Ratio	P Value	Function	Gene
4	rs7681294	0.6715	2.55E-05	Intron Variant	ELOVL6
9	rs2073815	1.434	0.000248	Synonomous Codon	SARDH
5	rs9327298	0.146	0.007141	Intron Variant	CSNK1G3
5	rs6860615	1.313	0.008661	Intron Variant	PCDHGC3
5	rs4530754	1.285	0.009949	Intron Variant	CSNK1G3
5	rs3749767	1.296	0.02412	Intron Variant, missense	PCDHGB4
5	rs11748256	1.272	0.02846	Intron Variant	PCDHGC3
5	rs11952292	0.6147	0.03245	Intron Variant, Synonomous Codon	PCDHGC3

Fatty Acid Elongase 6 (*ELOVL6)*

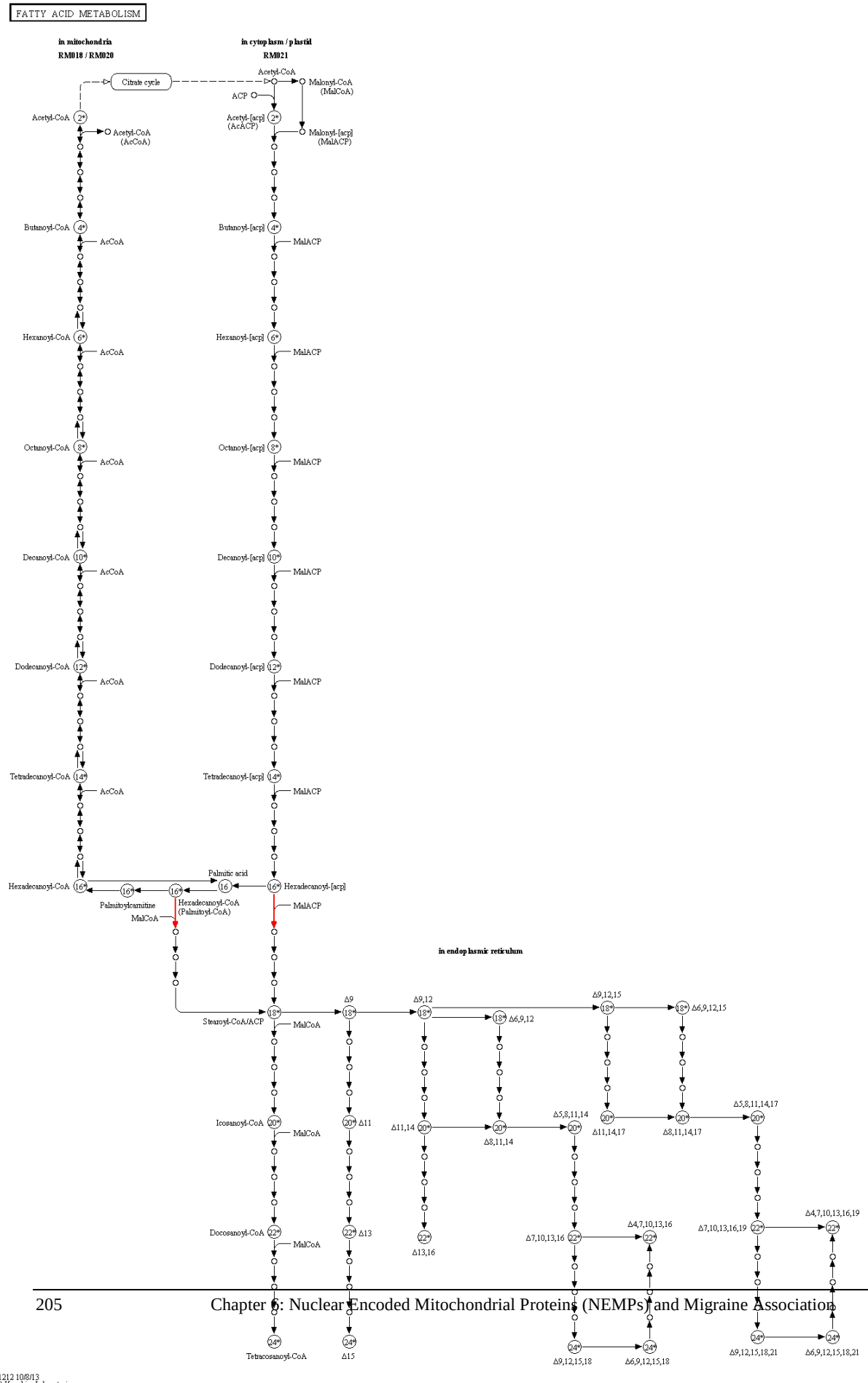
Fatty acid elongases such as ELOVL6 use malonyl-CoA as a 2-carbon donor in the first and rate-limiting step of fatty acid elongation [382]. This gene encodes for an enzyme in humans which catalyses the elongation of saturated and monounsaturated fatty acids with 12, 14 and 16 carbons. It has been found to be expressed in fatty tissues of the body and a recent study

has provided evidence for this to be a new candidate gene involved in energy deficiencies. Specifically variants in this gene have been associated with insulin sensitivity, suggesting an important role in metabolic processes [383]. Figure 39 taken from the KEGG database, below shows the major steps involved in fatty acid metabolism. Shown in red is the reaction governed by ELOVL6, which still takes place within the mitochondria itself [384, 385].

Within this key gene, we found rs7681294 to be highly associated with migraine susceptibility (p=2.55E-05). The odds ratio of 0.6715 suggests that individuals who carry the C>T change are protected by the T allele. A relevant hypothesis would be that this variant somehow increases the enzymatic function of ELOVL6, thereby increasing the efficiency of energy metabolism which has the knock on effect of neurons being provided with a continuous energy source from fat stores. This way even when food supply is scarce the body is easily able to metabolise fat reserves and individuals who carry the T allele are protected from migraine in this way. Extensive functional studies would be needed to examine this hypothesis and elucidate the mechanism in which this variant is functioning to prevent migraine attacks.

Even though this variant is situated within an intronic region and therefore traditionally considered to be less functionally relevant, new discussions and research are showing that non-coding regions of the genome are extremely important for regulation of body functions. It has been suggested that intronic regions are critical for gene expression and downstream pathway regulation.

Figure 39: Role of ELOVL6 in Fatty Acid Metabolism



Sarcosine Dehydrogenase (SARDH)

Sarcosine dehydrogenase situated on chromosome 9 encodes for an enzyme which is localised to the mitochondrial matrix and catalyzes the oxidative demethylation of sarcosine. Mutations in this gene have been associated with sarcosinemia, a mild inborn error of metabolism [386]. Some reports have suggested severe problems associated with this disease including mental retardation and neurological problems [387]. However the majority of cases are thought to be fairly benign. There has been a lot of controversy over this disease and the variable phenotype. Evidence suggests that this could be a complex disorder rather than a straight forward autosomal recessive model as previously thought. This complexity would explain the variable penetrance found across individuals. The final phenotype is dependent on the functional importance of the *SARDH* variant and also if there are additional contributing risk variants at other loci having a synergistic effect [388].

Sarcosine dehydrogenase removes the methyl group from sarcosine, a unique amino acid intermediate, to form glycine and an active 1 carbon unit. Disturbances in this pathway can be caused by severe folate deficiency, dysfunction of the electron transfer protein or dysfunction of sarcosine dehydrogenase. This process occurs in the mitochondria of liver and kidney cells and is an important part of 1-C metabolism. Given the neurological symptoms which have been associated with this gene it is feasible for it to play a role in migraine pathogenesis. It is difficult to hypothesize what the exact mechanism involved would be. According to our logistic regression model individuals carrying the C>T change are at a significantly increased risk of developing migraine ($p=0.000248$). The odds ratio (1.434) shows that individuals carrying the T allele are more susceptible. It is possible that this variant could be causing a mild enzymatic dysfunction and subsequently lowering the threshold for developing

migraines. Since migraine is a complex disease it is likely that this variant would act in conjunction with many other variants to bring about the final phenotype.

Casein Kinase I, Gamma 3 (CSNK1G3)

The casein kinase gene family encodes for serine and threonine kinases which are involved in phophorylating acidic substrates using ATP as a phosphate donor. No previous association have been made between this gene and human disease. There is limited information available about the function of this gene and variations within this region have not been studied in detail. Further investigations are required to draw conclusions.

Protocadherin Gamma Subfamily C (PCDHGC3 and PGDHGB4)

These genes are part of a large cluster of tandemly linked genes on chromosome 5. Their organisation has been described as similar to the immunoglobulin genes and it is thought that they are regulated in a unique way rather than the classic single promoter region per gene model. The gamma cluster to which both *PCDHG3* and *PCDHB4* belong consists of 22 genes subdivided into 3 families namely A, B and C. Subfamily A contains 12 genes, subfamily B contains 7 genes and 2 pseudogenes, and the more distantly related subfamily C contains 3 genes. *PCDHGC3* belongs to subfamily C, while *PCDHGB4* belongs to subfamily B. While these subfamilies are not exclusively linked, there is a shared gene element between them with 3 exons shared by all genes in this cluster. These genes encode for cadherin-like cell adhesion proteins and are thought to play a critical role in the establishment and function of specific cell-cell connections in the brain [389].

While the actual biological mechanism involved in a migraine attack is still debated, it is thought to be caused by activation of the trigeminal nerve causing pain sensation in the sensor

cortex of the brain and/or a dysfunction of the neuronal nuclei located within the brain stem [390].The trigeminal vascular theory states that activation of the trigeminal nerve system by a neural, vascular or neurovascular trigger leads to a migraine. The trigeminal nerves carry pain signals from the meninges and blood vessels infusing the meninges to the trigeminal nucleus in the brain stem which in turn sends signals to the sensor cortex via the thalamus. The sensor cortex processes pain signals and other senses, thus leading to the sensation of pain experienced during migraine attacks [391]. Dysfunction of neuronal nuclei can be explained by migraine pain and trigeminovascular activation being caused by a central mechanism which may not require a primary sensory input [22, 23]. The most recent theory explaining migraine pathogenesis describes migraine as a dysfunction of the subcortical brain structures including the brainstem and diencephalic nuclei which are involved in modulating sensory inputs. The theory suggests that aura is triggered by dysfunction of these nuclei and that the same mechanism is responsible for the pain and other symptoms experienced during migraine attacks [24].

Based on these theories, especially the idea of dysfunction within the neuronal nuclei, it is clear that signalling molecules/proteins involved in neuronal function and control are key targets for migraine pathogenesis. A disruption in cell-cell connections within the brain would lead to dysfunction within the neuronal nuclei as described in recent theories. Variants within the large *PDHG* gene region could alter the way in which connections are established and maintained in the brain, making an individual more susceptible towards developing migraines. We found 4 variants within this region to be very strongly associated with migraine susceptibility in a large case-control cohort. An association with multiple variants within this region strengthens the evidence that this gene cluster is a key component in migraine pathogenesis. Haplotype analysis showed some level of LD between these variants

and a strong association between haplotypes containing these variants and migraine susceptibility. Given the discovery of this gene cluster being involved in migraine susceptibility in the genetically isolated Norfolk Island population and the very clear replication in a large outbred population, this gene should be investigated further in great detail.

Conclusion

This is the first molecular genetic study to comprehensively investigate the role of NEMPs in migraine susceptibility we present empirical evidence for the first time to establish the link between mitochondrial dysfunction and migraine.

Chapter 7: Conclusions and Future Directions

It was hypothesised that development of migraine is influenced by mitochondrial dysfunction. The aim of this project was to conduct a complete mitochondrial genome scan to identify the full spectrum of mtDNA variation in the Norfolk Island pedigree samples and to determine whether the variants are associated with risk of migraine. It was further aimed to investigate whether these mtDNA variants or mitochondrial influencing variants from nuclear encoded genes modify migraine risk associated with key environmental factors.

The specific objectives were to:

- Obtain entire mitochondrial genome sequence information from selected samples using the Ion Torrent Platform.
- Align the mtDNA sequence information with independent worldwide samples to identify unique variants and determine frequencies of known variants as well as phylogenetic haplogroups.
- Collect genotype data from nuclear encoded mitochondrial protein (NEMP) genes.
- Statistically model the association of mtDNA variants and haplogroups with heritable migraine traits.
- Perform validation studies in independent cohorts.

All of the aims were achieved and new avenues of research to pursue in future studies were identified. Through full mitochondrial genome sequencing on the Ion Torrent platform 3 homoplasmic and 11 heteroplasmic variants were identified to be significantly associated with migraine susceptibility in the Norfolk Island population. Haplogroup K was also found to be associated with migraine in the Norfolk Island pedigree. The role of NEMPs was further

investigated as specified by the aims and it was found that the *PCDHG* gene region plays a particularly important role in migraine pathogenesis. It would be valuable to further investigate full mitochondrial genome information in an Australian outbred Caucasian population, specifically with regard to heteroplasmic variants. It would also be useful to further explore the role of NEMPs in migraine through deep sequencing, genotyping and gene expression studies in case-control populations, the Norfolk Island pedigree and migraine family samples.

Functional Studies

A recent study which aimed to develop an efficient method of assessing mitochondrial function in rat models for a range of neurological diseases made the finding that mitochondria are dysfunctional in a chronic migraine rat model [392]. The study identified a decreased spare respiratory capacity in the trigeminal nucleus caudalis (TNC) in a chronic migraine rat model. As stated throughout this thesis, it is hypothesised that the trigeminal nucleus plays a key role in migraine pathophysiology and so this finding of reduced oxidative phosphorylation in this key area of the brain is irrefutable evidence that mitochondrial function is linked to migraine. Many neurological disorders have already been associated with mitochondrial dysfunction including Alzheimer's disease, Parkinson's disease, amyotropic lateral sclerosis, Huntington's disease and Friedreich's ataxia showing that there is a clear link between bioenergetics and neurological disease [393, 394].

Given these links and the existing evidence of mitochondrial dysfunction and migraine it was logical to develop an accurate method capable of measuring mitochondrial function from brain sections. Nathan T Fried developed a technique to assess basal respiration, ATP turnover, proton leak, maximal respiration and non-mitochondrial respiration using thin brain sections. They found that sectioning at 37 °C produced more reproducible results as enzymes were kept at physiologically functional temperatures. It was demonstrated that nearly all of the oxygen consumed by the neurological cells of brain sections is due to mitochondrial respiration and that neurons are almost entirely dependent on the production of ATP via the oxidative phosphorylative chain [392].

These findings confirmed previously established knowledge that neuronal tissue is incredibly energy demanding, with oxygen consumption levels of a firing neuron thought to represent as

much as 80 % of a particular cell's maximum respiration and therefore even minimal changes to the spare capacity of a cell's mitochondria could have profound effects on neuronal function. This study showed for the first time that there is a decreased spare respiratory capacity within the TNC of a rat model with chronic migraine. The authors also demonstrated increases in glutamate in the TNC following exposure to glycerol trinitrate which is a classic molecule implicated in migraine pathogenesis. This functional work shows a clear relationship between mitochondrial dysfunction and migraine and strengthens the findings of this thesis, that molecular genetic variation which alters mitochondrial function is linked with the risk of migraine susceptibility.

Therapeutic Response

Riboflavin (B2)

Riboflavin, more commonly known as vitamin B2, acts as an essential precursor to the production of coenzymes flavin mononucleotide (FMN) and flavin-adenosine-dinucleotide (FAD). FMN and FAD are mitochondrial coenzymes which function to transport electrons across complex I and III of OXPHOS through the transport of hydrogen ions. Sufficient levels of riboflavin are essential for the production of ATP via OXPHOS in the mitochondria as well as the breakdown of amino acids, fatty acids and purines. A number of studies have provided direct clinical evidence of this link between riboflavin and mitochondrial function. It has been shown that riboflavin improves the biochemical and clinical abnormalities of patients with MELAS and other known mitochondrial diseases [395, 396]. These therapeutic responses show a clear link between mitochondrial function and riboflavin. Other research has investigated the effect of riboflavin supplementation on migraine frequency and severity and in all cases it has been shown that riboflavin reduces the burden of migraine for sufferers [367, 397] [369]. Given the clear link between riboflavin and mitochondrial function, this

finding clearly establishes the link between migraine and mitochondrial function demonstrating that migraine is caused at least in part by an impaired oxidative metabolism [398]. This is clear therapeutic evidence that mitochondrial dysfunction plays a key role in migraine and further supports the findings of this thesis that molecular genetic factors affecting mitochondrial function increase an individual's susceptibility towards developing migraine.

Coenzyme Q10

Further clinical evidence that migraine is a broad spectrum metabolic disorder is the clear therapeutic response to coenzyme Q10. Both riboflavin and coenzyme Q10 have been recommended as safe prophylactic migraine treatments [399]. The literature reflects that brain energy metabolism is abnormal in all types of migraine, with the most severe forms showing the greatest metabolic abnormalities [370]. Biochemical and phosphorus magnetic resonance studies have demonstrated marked metabolic defects in the brains of hemiplegic migraine and migraineous stroke sufferers. These metabolic abnormalities extend to the muscle cells, which also require a large supply of energy, in severely affected patients which shows the clear mitochondrial link. There are also strong similarities between migraine and some inborn errors of metabolisms, specifically the mitochondrial encephalomyopathies where there is an overlap of symptoms including lactic acidosis, stroke and migraineous headaches. Indeed there is strong biochemical and morphological evidence that migraine is a mitochondrial disorder. Until recently the molecular genetic evidence has been lacking, but we now show for the first time a clear molecular genetic link between genomic variants influencing mitochondrial function and migraine susceptibility [400].

Through next generation sequencing technologies a recent study showed a link for the first time between two mitochondrial variants and paediatric vomiting syndrome [339]. While this differs from adult onset migraine, it does demonstrate the power of next generation sequencing to unravel the molecular links with migraine. It has been suggested in many review papers that previous studies were negative because of low level heteroplasmic variants which could not be detected. This thesis validates these arguments and demonstrates a greater link between heteroplasmic (possibly accumulated) variants than homoplasmic germline variants [400].

Coenzyme Q10 acts as an electron carrier in the mitochondrial respiratory chain. Several clinical trials demonstrate a high response rate to treatment with Q10 with migraine sufferers experiencing a decrease in the frequency and severity of their attacks. Treatment with Q10 has been reported to be well tolerated with very few side effects and has been recommended as a safe prophylactic migraine treatment. Given coenzyme Q10's key role in mitochondrial respiration, the therapeutic response to this molecule is strong evidence that mitochondrial dysfunction attributes to the migraine phenotype in those individuals who respond positively to Q10 treatment [399, 401].

Mitochondria and Age Related Diseases

Age-related and lifestyle induced diseases are on the rise in developed countries. The dramatic increase in diabetes, cardiovascular disease, cancers and other disorders has been documented across all the developed regions of the world and more recently similar trends in developing countries are starting to appear. Scientists aim to understand what the link between our genetic makeup and these diseases are in an effect to better identify individuals at an increased risk for early intervention strategies. The rate of evolution is slow and our molecular clocks do not change over a period of one generation. Since our genes have not changed much over the last 70 years, it is obvious that environmental changes are causing this rapid rise of disease. While pollution, stress and other environmental factors could be playing a role, the biggest change to the average human lifestyle has been diet and the availability of high fat, sugar and calorie food in almost unlimited quantities [402-404].

The clear relationship between food and disease has led researchers to try and understand what the genetic links are between diet and disease. Thus far nuclear DNA studies have had limited success in explaining these tendencies and a Mendelian system of inheritance very clearly does not hold true for these complex lifestyle disorders. Given the increase of disease frequency and severity with an increase in age, it is logical that what we may be searching for is a gradual accumulation of mutations over time. It would also make sense for this to be occurring in a genetic system where there are multiple copies of a gene, instead of just two and for an accumulation of changes to eventually reach a threshold where disease develops. Research has also shown a clear difference between the same disease in different ethnic groups giving rise to the thought process that the genetic system is differing according to

regional origin. In summary we are looking for a genetic system which differs according to region, accumulates mutations over time and is linked to diet and metabolism. Mitochondrial genetics fits perfectly with all of these descriptions and is the next obvious place to look [357, 405].

The first evidence that mitochondria could play a key role in aging and age related degenerative diseases is the relationship between systematic disease caused by mtDNA mutations and the delayed onset of symptoms which then become progressive as time goes on. Pathogenic mutations including rearrangements mutations, polypeptide gene missense mutations and protein synthesis mutations (RNAs) affect the brain, heart, skeletal muscle and endocrine system in s similar way to what we observe in aging individuals. A further observation is the link between severity of disease and tissue distribution. It has been found that the distribution of a mutation has a larger effect on disease outcome than the actual mutation itself. The level at which the mutation occurs and the percentage of mitochondria it is found in greatly influences the severity of disease [406-408].

This trend ties in nicely with the findings outlined in this thesis of heteroplasmic variants and migraine susceptibility. It was found that there is a stronger link between heteroplasmic variants and migraine susceptibility than their homoplasmic counterparts. This further explains the variable expressivity and complex inheritance patterns observed in migraine sufferers. As has been found repeatedly, the tissues with the highest energy requirement are usually the most adversely affected and is makes sense that the more affected mitochondria in a tissue population, the worse the disease outcome. An excellent known example of this phenomenon is a mtDNA mutation in the ATPase6 gene where T>G at position 8993 causes

NARP when present at low frequencies, but is lethal when present in a higher percentage of cells, causing lethal childhood Leigh syndrome [357].

The Link between Region of Origin and Disease

Throughout human evolution our ancestors have continually been migrating across the globe to settle in locations with a variety of climates. Initially the first people left Africa around 70 000 years ago with a subsequent second migration into Asia approximately 30 000 years ago. These migrations represented drastic climatic changes for those leaving Africa. Just like any other mammal, humans have evolved to adapt to their changing environment. One form of adaptation has been the acquisition of mtDNA mutations that partially uncoupled OXPHOS allowing for an increase in mitochondrial heat production to compensate for cold European temperatures. This change in physiology is evidenced by the observations that individuals with different haplogroups exhibit different responses to mitochondrial disease [357].

The mitochondrial genome mutates at a faster rate than nuclear DNA and once a mutation has been acquired it is passed along the maternal lineage. All individuals descending from that maternal line with the new mutation also possess the change in their genetic code. In this way researchers have been able to track human migration through these acquired molecular markers in conjunction with paleo studies. Once the maternal origin of a particular marker has been matched with a geographic origin we know that all individuals carrying that marker descended from the same maternal line and a change in location is due to migration. Multiple mutations have taken place and been passed down through evolution and those occurring in

the hypervariable loop of the mitochondrial genome have been grouped and assigned letters/haplogroups according to geographic origin [409].

Research has found that haplogroup J variants have a functional effect and are not just markers of human evolution. Individuals possessing haplogroup J have partially uncoupled OXPHOS which reduces the efficacy of ATP output in favour of heat production. As a result of this pre-existing tendancy, individuals belonging to haplogroup J who have very mild complex I mtDNA missense mutations have exacerbated symptoms associated with LHON compared to individuals belonging to other haplogroups [284]. This theory has been further validated by the observation that individuals respond differentially to treatment with riboflavin according to haplogroup. This means that migraine sufferers treated with riboflavin will respond differently according to their haplogroup and provides a direct link between migraine and haplogroup [369]. This corresponds clearly with the finding that individuals who possess haplogroup K are at an increased risk of developing migraine. Further research into this area would be useful, especially in light of variable therapeutic response.

Other lines of evidence prove that there is a connection between mitochondrial haplogroups and disease and that these markers are adaptive, not just neutral markers of evolution. The haplogroups have also been positively linked with increased lifespan and may play a role in the way in which we age. Some of the altered physiological function caused by mitochondrial mutations may actually be beneficial when exposed to the correct environmental stimuli. Studies have found haplogroup J, U and WIX to be overrepresented in centenarians and these same mtDNA lineages have also been found to be protective against neurodegenerative diseases [410-413].

The argument which explains how a haplogroup can cause exacerbated disease in one instance, but longevity in another resides in the physiological change caused by the relevant mitochondrial mutations. Individuals with uncoupled mitochondria burn more calories more rapidly to generate the required ATP plus heat and as a result the electron transport chains of these individuals exist in a more oxidised state which ultimately reduces the production of ROS. Hence individuals with uncoupled mitochondria are more prone to clinical problems related to energy deficiencies such as LHON and bipolar mental illness, but are protected from the ageing process through lower levels of damaging ROS [357].

Co-morbid Spectrum Disorders

Recently the idea of a common underlying mechanism which causes a spectrum type of disorder has arisen. The theory is based on the idea that depression, migraine, fibromyalgia and chronic anxiety disorders which have a high degree of overlap in symptoms and therapeutic response are a group of related conditions or 'affective spectrum disorders". It has also been suggested that mitochondrial dysfunction in conjunction with inflammation may cause a spectrum of disease based on complex and variable interactions. The most commonly used drugs prescribed to treat depression have been shown to alter mitochondrial function and inflammation pathways, suggesting that these are key components involved in the pathophysiology of depressive like disorders. Many treatments for depression are also used to treat migraine given the known co-morbidity between these two disorders and Gardner suggest that mitochondrial targeted treatments have an increased efficacy in the treatment of affective spectrum disorders [414].

Studies have identified a definitive relationship between mitochondrial dysfunction and major depressive disorders where post mortem studies have identified alterations of translational products linked to mitochondrial function in the frontal, prefrontal and tertiary visual cortices [415]. Alerted gene expression of both mtDNA encoded and NEMP encoded transcripts have been reported in major depression in addition to decreases of respiratory chain enzyme ratios and ATP production. These findings collectively highlight the role of mitochondrial function in depression [416, 417]. A clear link has furthermore been defined between mitochondrial dysfunction and other affective spectrum disorders with mitochondrial abnormalities such as RFF, COX negative fibres and reduced ATP output identified in fibromyalgia and migraine [334, 418].

Gardner definse the role of inflammation in the affective spectrum disorders and further suggest that there is a combined role between inflammation and mitochondrial dysfunction in these spectrum disorders. Elevated levels of cytokines as well as alterations in inflammatory pathways have been reported in all the spectrum disorders, suggesting a complex interaction between mitochondrial dysfunction and inflammation [414, 419, 420]. Perusing this relatively novel avenue for treatment of affective spectrum disorders could be useful. Given the complex nature of these disorders, it is logical that multiple pathways are involved which can vary between individuals and according to environmental responses.

Summary and Concluding Comments

Migraine is a complex disorder, caused by an interplay between genetic and environmental factors. While a lot of progress has been made over the centuries with regards to understanding this disease, the complete disease aetiology remains to be fully elucidated. Many prophylactic and acute prescription medications are available for the treatment of migraine sufferers, however these treatments are still only effective for less than half of all patients. Migraine poses a significant burden to both the economy and also the quality of life of those affected. Many studies have shown that the cost of leaving this disease untreated is significantly higher due to lost productivity in the workplace, than the cost to effectively address disease burden through the development of new therapeutics.

Previous molecular genetic studies have made significant progress towards improving our understanding of migraine and ultimately developing more effective treatment strategies. Genetic studies have been particularly successful with the rare subtypes of migraine such as FHM which have clear patterns of inheritance through families and usually can be linked to a single causative variant. With the more common subtypes of migraine which are much more complex in nature, some progress has been made, but much more work is still required. With the significant improvements in technology and the rapid development of next generation sequencing technologies, our ability to investigate complex disease is improving at a rapid rate. This study harnessed the power of new technology and aimed to completely sequence full human mitochondria in a large subset of individuals belonging to a large genetically isolated pedigree.

This goal was achieved, initially by developing the methodology needed to carry out this experiment on a large scale in a small pilot project. Making use of available techniques to optimise this process in the most cost effective manner possible was a very important step and ultimately enabled this research to be undertaken on a larger scale. In total 315 unique individuals underwent full mitochondrial genome sequencing including 80 migraine sufferers and 235 healthy controls. These samples constituted 152 males and 163 females. In this study unprecedented levels of coverage were achieved, made possible by new technologies and the power of the optimised methodology. This allowed for the first time the investigation of the potential role of heteroplasmic variants in disease susceptibility.

The Torrent Suite was used extensively to analyse sequencing results and data further underwent a rigorous bioinformatics analysis process, using custom scripts and in-house developed pipelines. These techniques are on the cutting edge of development and are pushing science to the next frontier. This research represents a significant step forward in comparison to what was previously possible. Initial analysis identified three SNPs which are significantly associated with migraine in the Norfolk Island population namely mt 930 G>A, mt 6480 G>A and mt 11930 A>G. One SNP is novel and has never been documented before (mt 11930 A>G), 1 SNP was defined as a common variant with a MAF>0.01 (mt 930 G>A) and 2 SNPs were defined as rare variants with a MAF <0.05 (mt 6480 G>A and mt 11930 A>G). Unfortunately these results were not replicated in an outbred population suggesting that these variants are Norfolk specific and only play a role in migraine susceptibility in this particular pedigree.

Despite the lack of replication, this discovery still represents an important finding as it points us to the correct regions of the mitochondrial genome to investigate further in a large scale re-sequencing project. The mitochondrial regions they are found in represent a lot of biological plausibility and this study has provided the justification to further investigate these regions in other migraine family samples. Of great interest is the finding that 11 heteroplasmic variants are significantly associated with migraine as this is the first time a molecular genetic study has shown that there could be a relationship between heteroplasmic variants and complex disease. This idea opens the door for the possibility that mutations acquired during an individual's lifetime may play a significant role in disease progression once a defining threshold is met. This line of thinking aligns with new progressive ideas about mitochondrial disease kinetics, ageing and lifestyle issues.

Of further interest is the discovery that haplogroup K significantly increases an individual's risk of developing migraine. This finding supports the theory of adaptation according to geographical region and temperature zones where specific mitochondrial changes alter an individual's vulnerability to energy reliant processes. The theory is based on the idea that European individuals evolved to have uncoupled mitochondria, favouring heat production over ATP production in order to survive the cold Northern climates. Hence certain haplogroups may predispose, protect against or aggravate mitochondrial related disorders. Previous studies have found disease to be associated with an individual's haplogroup and our findings support this line of evidence.

Technology is now moving forward at an unprecedented rate and we should harness every advantage to pursue our understanding of complex disease and bring about a new era of

modern medicine. As the costs of sequencing decrease, more ambitious projects fall into our reach. Future studies which involve full mitochondrial genome sequencing in multiple large case-control cohorts, twin samples and migraine family samples are warranted and will shed light on our existing findings. It would be useful to investigate heteroplasmy on a much larger scale and also in a more quantitative manner, rather than using categories in order to be more certain of our findings. Bioinformatic techniques need to be developed to match the technologies available in the laboratory and to keep up with the current changing pace. Effective analysis techniques are needed to better understand our sequence data. Further analysis according to haplogroups would be interesting from a population genetics perspective.

As mentioned throughout this thesis, mitochondrial function is largely governed by nuclear encoded mitochondrial proteins which are imported into the mitochondrial matrix to perform their vital functions. The majority of proteins involved in mitochondrial function are produced in this way and therefore it is imperative to investigate these NEMP genes in a comprehensive manner in order to fully assess the relationship between mitochondrial dysfunction and migraine susceptibility. This study touched the tip of the iceberg in this arena, using available data in the Norfolk Island pedigree and a large Sequenom plex for validation of findings. The *PCDHG* gene cluster was identified to play a pivotal role in migraine susceptibility in both the Norfolk Island population and an outbred Australian Caucasian population. This finding is in line with a previous GWAS which identified this region to be important in common subtypes of migraine.

Future studies which investigate all 1000+ NEMP genes through comprehensive deep sequencing would be very useful. Gene expression studies investigating expression levels and pathways would also be useful. This research is original and has provided sound scientific evidence to support new ideas and hypothesis. As is typical of original research and pushing the known boundaries, it has also created new unanswered questions.

Bibliography

1. Olesen, J., *Preface to the Second Edition.* Cephalalgia, 2004. **24**(1 suppl): p. 9-10.
2. Olesen, J. and R.B. Lipton, *Migraine classification and diagnosis. International Headache Society criteria.* Neurology, 1994. **44**(6 Suppl 4): p. S6-10.
3. Eriksen, M.K., L.L. Thomsen, and J. Olesen, *New international classification of migraine with aura (ICHD-2) applied to 362 migraine patients.* Eur J Neurol, 2004. **11**(9): p. 583-91.
4. Leonardi, M., et al., *The global burden of migraine: measuring disability in headache disorders with WHO's Classification of Functioning, Disability and Health (ICF).* J Headache Pain, 2005. **6**(6): p. 429-40.
5. Holmes, W.F., E.A. MacGregor, and D. Dodick, *Migraine-related disability: impact and implications for sufferers' lives and clinical issues.* Neurology, 2001. **56**(6 Suppl 1): p. S13-9.
6. Leonardi, M., et al., *Health-related quality of life, disability and severity of disease in patients with migraine attending to a specialty headache center.* Headache, 2010. **50**(10): p. 1576-86.
7. Lipton, R.B., et al., *Migraine headache disability and health-related quality-of-life: a population-based case-control study from England.* Cephalalgia, 2003. **23**(6): p. 441-50.
8. Lipton, R.B., et al., *The family impact of migraine: population-based studies in the USA and UK.* Cephalalgia, 2003. **23**(6): p. 429-40.
9. Berg, J., *Economic evidence in migraine and other headaches: a review.* Eur J Health Econ, 2004. **5 Suppl 1**: p. S43-54.
10. Gerth, W.C., et al., *The multinational impact of migraine symptoms on healthcare utilisation and work loss.* Pharmacoeconomics, 2001. **19**(2): p. 197-206.
11. Gustavsson, A., et al., *Cost of disorders of the brain in Europe 2010.* Eur Neuropsychopharmacol, 2011. **21**(10): p. 718-79.
12. Schultz, A.B., C.Y. Chen, and D.W. Edington, *The cost and impact of health conditions on presenteeism to employers: a review of the literature.* Pharmacoeconomics, 2009. **27**(5): p. 365-78.
13. Lanteri-Minet, M., et al., *Anxiety and depression associated with migraine: influence on migraine subjects' disability and quality of life, and acute migraine management.* Pain, 2005. **118**(3): p. 319-26.
14. Tan, H.J., et al., *The coexistence of anxiety and depressive personality traits in migraine.* Singapore Med J, 2007. **48**(4): p. 307-10.
15. Kurth, T., et al., *Migraine, headache, and the risk of stroke in women: a prospective study.* Neurology, 2005. **64**(6): p. 1020-6.
16. Katsarava, Z., K. Rabe, and H.C. Diener, *From migraine to stroke.* Intern Emerg Med, 2008. **3 Suppl 1**: p. S9-16.
17. Narbone, M.C., S. Gangemi, and M. Abbate, *Migraine and stroke: from a questioned relationship to a supported comorbidity.* Neurol Sci, 2008. **29 Suppl 1**: p. S7-11.
18. Smitherman, T.A., et al., *The use of antidepressants for headache prophylaxis.* CNS Neurosci Ther, 2011. **17**(5): p. 462-9.
19. Breslau, N., et al., *Comorbidity of migraine and depression: investigating potential etiology and prognosis.* Neurology, 2003. **60**(8): p. 1308-12.

20. Headache Classification Subcommittee of the International Headache, S., *The International Classification of Headache Disorders: 2nd edition.* Cephalalgia, 2004. **24 Suppl 1**: p. 9-160.
21. de Almeida, R.F., et al., *Migraine with persistent visual aura: response to furosemide.* Clinics, 2009. **64**(4): p. 375-6.
22. Lambert, G.A., L. Truong, and A.S. Zagami, *Effect of cortical spreading depression on basal and evoked traffic in the trigeminovascular sensory system.* Cephalalgia : an international journal of headache, 2011. **31**(14): p. 1439-51.
23. Goadsby, P.J. and S. Akerman, *The trigeminovascular system does not require a peripheral sensory input to be activated - migraine is a central disorder Focus on 'Effect of cortical spreading depression on basal and evoked traffic in the trigeminovascular sensory system'.* Cephalalgia : an international journal of headache, 2012. **32**(1): p. 3-5.
24. Akerman, S., P.R. Holland, and P.J. Goadsby, *Diencephalic and brainstem mechanisms in migraine.* Nature reviews. Neuroscience, 2011. **12**(10): p. 570-84.
25. Sparaco, M., et al., *Mitochondrial dysfunction and migraine: evidence and hypotheses.* Cephalalgia : an international journal of headache, 2006. **26**(4): p. 361-372.
26. Stuart, S. and L.R. Griffiths, *A possible role for mitochondrial dysfunction in migraine.* Mol Genet Genomics, 2012. **287**(11-12): p. 837-44.
27. Attwell, D. and S.B. Laughlin, *An energy budget for signaling in the grey matter of the brain.* Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism, 2001. **21**(10): p. 1133-45.
28. Attwell, D., et al., *Glial and neuronal control of brain blood flow.* Nature, 2010. **468**(7321): p. 232-43.
29. Ko, K.R., A.C. Ngai, and H.R. Winn, *Role of adenosine in regulation of regional cerebral blood flow in sensory cortex.* The American journal of physiology, 1990. **259**(6 Pt 2): p. H1703-8.
30. Ido, Y., et al., *NADH: sensor of blood flow need in brain, muscle, and other tissues.* FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2001. **15**(8): p. 1419-21.
31. Tanji, K., et al., *Neuropathological features of mitochondrial disorders.* Seminars in Cell & Developmental Biology, 2001. **12**(6): p. 429-39.
32. Wieser, T., et al., *Absence of known familial hemiplegic migraine (FHM) mutations in the CACNA1A gene in patients with common migraine: implications for genetic testing.* Clinical chemistry and laboratory medicine : CCLM / FESCC, 2003. **41**(3): p. 272-5.
33. Chuquet, J., L. Hollender, and E.A. Nimchinsky, *High-resolution in vivo imaging of the neurovascular unit during spreading depression.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2007. **27**(15): p. 4036-44.
34. Metea, M.R. and E.A. Newman, *Glial cells dilate and constrict blood vessels: a mechanism of neurovascular coupling.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2006. **26**(11): p. 2862-70.
35. Sparaco, M., et al., *Mitochondrial dysfunction and migraine: evidence and hypotheses.* Cephalalgia : an international journal of headache, 2006. **26**(4): p. 361-72.
36. Montagna, P., et al., *Mitochondrial abnormalities in migraine. Preliminary findings.* Headache, 1988. **28**(7): p. 477-80.

37. Bresolin, N., et al., *Muscle mitochondrial DNA deletion and 31P-NMR spectroscopy alterations in a migraine patient.* Journal of the neurological sciences, 1991. **104**(2): p. 182-9.
38. Uncini, A., et al., *Abnormal Brain and Muscle Energy-Metabolism Shown by P-31-Mrs in Familial Hemiplegic Migraine.* Journal of the neurological sciences, 1995. **129**(2): p. 214-222.
39. Bonilla, E., et al., *New morphological approaches to the study of mitochondrial encephalomyopathies.* Brain Pathology, 1992. **2**(2): p. 113-9.
40. Tyni, T., et al., *Mitochondrial fatty acid beta-oxidation in the human eye and brain: implications for the retinopathy of long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency.* Pediatric research, 2004. **56**(5): p. 744-50.
41. Casas, F., et al., *New molecular aspects of regulation of mitochondrial activity by fenofibrate and fasting.* FEBS letters, 2000. **482**(1-2): p. 71-4.
42. Chinnery, P.F., et al., *Clinical progression of mitochondrial myopathy is associated with the random accumulation of cytochrome c oxidase negative skeletal muscle fibres.* Journal of the neurological sciences, 2003. **211**(1-2): p. 63-66.
43. Taylor, R.W., et al., *The determination of complete human mitochondrial DNA sequences in single cells: implications for the study of somatic mitochondrial DNA point mutations.* Nucleic Acids Research, 2001. **29**(15): p. art. no.-e74.
44. Uncini, A., et al., *Abnormal brain and muscle energy metabolism shown by 31P-MRS in familial hemiplegic migraine.* Journal of the neurological sciences, 1995. **129**(2): p. 214-22.
45. Majamaa, K., et al., *Mitochondrial DNA haplogroup U as a risk factor for occipital stroke in migraine.* Lancet, 1998. **352**(9126): p. 455-6.
46. Skinhoj, E., *Hemodynamic studies within the brain during migraine.* Archives of neurology, 1973. **29**(2): p. 95-8.
47. Okada, H., et al., *Plasma lactic acid and pyruvic acid levels in migraine and tension-type headache.* Headache, 1998. **38**(1): p. 39-42.
48. Dimauro, S., et al., *Mitochondrial Myopathies.* Annals of Neurology, 1985. **17**(6): p. 521-538.
49. Littlewood, J., et al., *Low Platelet Monoamine-Oxidase Activity in Headache - No Correlation with Phenolsulphotransferase, Succinate-Dehydrogenase, Platelet Preparation Method or Smoking.* Journal of Neurology Neurosurgery and Psychiatry, 1984. **47**(4): p. 338-343.
50. Sangiorgi, S., et al., *Abnormal Platelet Mitochondrial-Function in Patients Affected by Migraine with and without Aura.* Cephalalgia : an international journal of headache, 1994. **14**(1): p. 21-23.
51. Kabbouche, M.A., et al., *Carnitine palmityltransferase II (CPT2) deficiency and migraine headache: two case reports.* Headache, 2003. **43**(5): p. 490-5.
52. Wanic-Kossowska, M., *[Protective role of carnitine in acetate metabolism of patients with uremia treated by hemodialysis]*. Polskie Archiwum Medycyny Wewnetrznej, 1997. **97**(6): p. 534-40.
53. Chance, B., et al., *Phosphorus nuclear magnetic resonance spectroscopy in vivo.* Circulation, 1985. **72**(5 Pt 2): p. IV103-10.
54. Arnold, D.L., D.J. Taylor, and G.K. Radda, *Investigation of human mitochondrial myopathies by phosphorus magnetic resonance spectroscopy.* Annals of Neurology, 1985. **18**(2): p. 189-96.

55. Barbiroli, B., et al., *Defective brain energy metabolism shown by in vivo 31P MR spectroscopy in 28 patients with mitochondrial cytopathies.* Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism, 1993. **13**(3): p. 469-74.
56. Montagna, P., P. Cortelli, and B. Barbiroli, *Magnetic resonance spectroscopy studies in migraine.* Cephalalgia : an international journal of headache, 1994. **14**(3): p. 184-93.
57. Argov, Z., et al., *Bioenergetic heterogeneity of human mitochondrial myopathies: phosphorus magnetic resonance spectroscopy study.* Neurology, 1987. **37**(2): p. 257-62.
58. Welch, K.M., et al., *Preliminary observations on brain energy metabolism in migraine studied by in vivo phosphorus 31 NMR spectroscopy.* Neurology, 1989. **39**(4): p. 538-41.
59. Montagna, P., et al., *P-31-Magnetic Resonance Spectroscopy in Migraine without Aura.* Neurology, 1994. **44**(4): p. 666-669.
60. Barbiroli, B., et al., *Abnormal brain and muscle energy metabolism shown by 31P magnetic resonance spectroscopy in patients affected by migraine with aura.* Neurology, 1992. **42**(6): p. 1209-14.
61. Barbiroli, B., et al., *Complicated migraine studied by phosphorus magnetic resonance spectroscopy.* Cephalalgia : an international journal of headache, 1990. **10**(5): p. 263-72.
62. Sacquegna, T., et al., *Brain Energy-Metabolism Studied by P-31-Mr Spectroscopy in a Case of Migraine with Prolonged Aura.* Acta Neurologica Scandinavica, 1992. **86**(4): p. 376-380.
63. Montagna, P., P. Cortelli, and B. Barbiroli, *A case of cluster headache associated with mitochondrial DNA deletions.* Muscle & nerve, 1998. **21**(1): p. 127-9.
64. Macgregor, S., et al., *Bias, precision and heritability of self-reported and clinically measured height in Australian twins.* Human genetics, 2006. **120**(4): p. 571-80.
65. Visscher, P.M., et al., *Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings.* PLoS genetics, 2006. **2**(3): p. 316-325.
66. Ulrich, V., et al., *The inheritance of migraine with aura estimated by means of structural equation modelling.* Journal of Medical Genetics, 1999. **36**(3): p. 225-227.
67. Honkasalo, M.L., et al., *Migraine and Concomitant Symptoms among 8167 Adult Twin-Pairs.* Headache, 1995. **35**(2): p. 70-78.
68. Mulder, E.J., et al., *Genetic and environmental influences on migraine: a twin study across six countries.* Twin research : the official journal of the International Society for Twin Studies, 2003. **6**(5): p. 422-31.
69. Wessman, M., et al., *Migraine: a complex genetic disorder.* Lancet Neurol, 2007. **6**(6): p. 521-32.
70. Chasman, D.I., et al., *Genome-wide association study reveals three susceptibility loci for common migraine in the general population.* Nat Genet, 2011. **43**(7): p. 695-8.
71. Refshauge, W.F. and R.J. Walsh, *Pitcairn Island: fertility and population growth, 1790-1856.* Annals of human biology, 1981. **8**(4): p. 303-12.
72. Bellis, C., et al., *Phenotypical Characterisation of the Isolated Norfolk Island Population Focusing on Epidemiological Indicators of Cardiovascular Disease.* Human Heredity, 2005. **60**(4): p. 211-219.

73. Ober, C., M. Abney, and M.S. McPeek, *The genetic dissection of complex traits in a founder population.* American Journal of Human Genetics, 2001. **69**(5): p. 1068-79.
74. Hill, W.G. and T.F. Mackay, *D. S. Falconer and Introduction to quantitative genetics.* Genetics, 2004. **167**(4): p. 1529-36.
75. Maher, B.H., et al., *An X chromosome association scan of the Norfolk Island genetic isolate provides evidence for a novel migraine susceptibility locus at Xq12.* PLoS One, 2012. **7**(5): p. e37903.
76. Bellis, C., et al., *Linkage mapping of CVD risk traits in the isolated Norfolk Island population.* Human genetics, 2008. **124**(5): p. 543-52.
77. Benton, M., et al., *Complete mitochondrial genome sequencing reveals novel haplotypes in a Polynesian population.* PloS one, 2012. **7**(4): p. e35026.
78. Leonardi, M., et al., *The global burden of migraine: measuring disability in headache disorders with WHO's Classification of Functioning, Disability and Health (ICF).* The journal of headache and pain, 2005. **6**(6): p. 429-40.
79. Rose, F.C. and R. Capildeo, *Migraine: definition and classification.* Cephalalgia : an international journal of headache, 1983. **3**(4): p. 225-9.
80. Nappi, G., et al., *Migraine as a complex disease: Heterogeneity, comorbidity and genotype-phenotype interactions.* Functional Neurology, 2000. **15**(2): p. 87-93.
81. Maher, B.H. and L.R. Griffiths, *Identification of molecular genetic factors that influence migraine.* Molecular Genetics and Genomics, 2011. **285**(6): p. 433-446.
82. Colson, N.J., et al., *The role of vascular and hormonal genes in migraine susceptibility.* Molecular genetics and metabolism, 2006. **88**(2): p. 107-13.
83. Hans, M., et al., *Functional consequences of mutations in the human alpha(1A) calcium channel subunit linked to familial hemiplegic migraine.* Journal of Neuroscience, 1999. **19**(5): p. 1610-1619.
84. Dodick, D.W. and J.J. Gargus, *Why migraines strike.* Scientific American, 2008. **299**(2): p. 56-63.
85. Leao, A.A.P., *Spreading depression of activity in the cerebral cortex.* Journal of Neurophysiology, 1944. **7**(6): p. 359-390.
86. Levy, D., *Migraine pain, meningeal inflammation, and mast cells.* Current pain and headache reports, 2009. **13**(3): p. 237-40.
87. Lauritzen, M., *Cerebral blood flow in migraine and cortical spreading depression.* Acta neurologica Scandinavica. Supplementum, 1987. **113**: p. 1-40.
88. Boekema, E.J. and H.P. Braun, *Supramolecular structure of the mitochondrial oxidative phosphorylation system.* The Journal of biological chemistry, 2007. **282**(1): p. 1-4.
89. *The International Classification of Headache Disorders: 2nd edition.* Cephalalgia : an international journal of headache, 2004. **24 Suppl 1**: p. 9-160.
90. Terwindt, G.M., et al., *Variable clinical expression of mutations in the P/Q-type calcium channel gene in familial hemiplegic migraine. Dutch Migraine Genetics Research Group.* Neurology, 1998. **50**(4): p. 1105-10.
91. Lykke Thomsen, L., et al., *An epidemiological survey of hemiplegic migraine.* Cephalalgia : an international journal of headache, 2002. **22**(5): p. 361-75.
92. Rothner, A.D., *Complicated migraine and migraine variants.* Seminars in pediatric neurology, 2001. **8**(1): p. 7-12.
93. Curtain, R.P., et al., *Minor head trauma-induced sporadic hemiplegic migraine coma.* Pediatric neurology, 2006. **34**(4): p. 329-32.

94. Kors, E.E., et al., *Delayed cerebral edema and fatal coma after minor head trauma: role of the CACNA1A calcium channel subunit gene and relationship with familial hemiplegic migraine.* Annals of Neurology, 2001. **49**(6): p. 753-60.
95. Colson, N.J., et al., *The search for migraine genes: an overview of current knowledge.* Cellular and molecular life sciences : CMLS, 2007. **64**(3): p. 331-44.
96. Cuenca-Leon, E., et al., *Familial hemiplegic migraine: linkage to chromosome 14q32 in a Spanish kindred.* Neurogenetics, 2009. **10**(3): p. 191-8.
97. Ophoff, R.A., et al., *Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4.* Cell, 1996. **87**(3): p. 543-52.
98. de Vries, B., et al., *Molecular genetics of migraine.* Human genetics, 2009. **126**(1): p. 115-32.
99. Battistini S, S.S., Piatti M, Gelfi C, Righetti PG, Rocchi R, Giannini F, Battistini N, Guazzi GC, Ferrari M, Carrera P *A new CACNA1A gene mutation in acetazolamide-responsive familial hemiplegic migraine and ataxia.* Neurology, 1999. **53**(1): p. 38-43.
100. Ducros A, D.C.e.a., *The Clinical Spectrum Of Familial Hemiplegic Migraine Associated With Mutations In A Neuronal Calcium Channel.* The New England Journal of Medicine, 2001. **345**: p. 17-25.
101. Ophoff RA, T.G., Vergouwe MN, Van Eijk R, Oefner PJ, Hoffman SM, et al, *Familial hemiplegic migraine and episodic ataxia type 2 are caused by mutations in the Ca2+ channel gene CACNL1A4.* Cell, 1996. **87**: p. 543–552.
102. Ducros A, D.C., Joutel A, Vahedi K et al, *Recurrence of the T666M Calcium Channel CACNA1A Gene Mutation in amilial Hemiplegic Migraine with Progressive Cerebellar Ataxia.* American Journal of Human Genetics, 1999. **68**: p. 89–98.
103. Carrera P, P.M., Stenirri S, Grimaldi LM, Marchioni E, Curcio M, Righetti PG, Ferrari M, Gelfi C, *Genetic heterogeneity in Italian families with familial hemiplegic migraine.* Neurology, 1999. **53**(1): p. 26-33.
104. Cricchi F, D.L.C., Grieco GS, Rengo C, Cardinale A, Racaniello M, Santorelli FM, Nappi G, Pierelli F, Casali C, *Early-onset progressive ataxia associated with the first CACNA1A mutation identified within the I-II loop.* Journal of Neurological Sciences, 2007. **254**: p. 67-71.
105. D'Onofrio, M., et al., *The interplay of two single nucleotide polymorphisms in the CACNA1A gene may contribute to migraine susceptibility.* Neuroscience letters, 2009. **453**(1): p. 12-15.
106. Plomp, J.J., A.M. van den Maagdenberg, and S. Kaja, *The ataxic Cacna1a-mutant mouse rolling nagoya: an overview of neuromorphological and electrophysiological findings.* Cerebellum, 2009. **8**(3): p. 222-30.
107. Ferrari, M.D., et al., *Meta-analysis of rizatriptan efficacy in randomized controlled clinical trials.* Cephalalgia : an international journal of headache, 2001. **21**(2): p. 129-36.
108. Artto, V., et al., *Treatment of hemiplegic migraine with triptans.* European journal of neurology : the official journal of the European Federation of Neurological Societies, 2007. **14**(9): p. 1053-6.
109. Weinberger, J., *Interactions between migraine and stroke.* Current treatment options in neurology, 2006. **8**(6): p. 513-7.
110. Vanmolkot, K.R., et al., *Two de novo mutations in the Na,K-ATPase gene ATP1A2 associated with pure familial hemiplegic migraine.* European journal of human genetics : EJHG, 2006. **14**(5): p. 555-60.

111. Fernandez, F., N.J. Colson, and L.R. Griffiths, *Pharmacogenetics of migraine: genetic variants and their potential role in migraine therapy.* Pharmacogenomics, 2007. **8**(6): p. 609-622.
112. Maher, B.H. and L.R. Griffiths, *Identification of molecular genetic factors that influence migraine.* Molecular genetics and genomics : MGG, 2011. **285**(6): p. 433-46.
113. Colson, N.J., et al., *Investigation of hormone receptor genes in migraine.* Neurogenetics, 2005. **6**(1): p. 17-23.
114. Schurks, M., P.M. Rist, and T. Kurth, *5-HTTLPR polymorphism in the serotonin transporter gene and migraine: a systematic review and meta-analysis.* Cephalalgia : an international journal of headache, 2010. **30**(11): p. 1296-305.
115. Hamel, E., *Serotonin and migraine: biology and clinical implications.* Cephalalgia : an international journal of headache, 2007. **27**(11): p. 1293-300.
116. Juhasz, G., et al., *Association analysis of 5-HTTLPR variants, 5-HT2a receptor gene 102T/C polymorphism and migraine.* Journal of neurogenetics, 2003. **17**(2-3): p. 231-40.
117. Marziniak, M., et al., *A functional serotonin transporter gene polymorphism is associated with migraine with aura.* Neurology, 2005. **64**(1): p. 157-9.
118. Panconesi, A., *Serotonin and migraine: a reconsideration of the central theory.* The journal of headache and pain, 2008. **9**(5): p. 267-76.
119. Fernandez, F., et al., *Association between a 19 bp deletion polymorphism at the dopamine beta-hydroxylase (DBH) locus and migraine with aura.* Journal of the neurological sciences, 2006. **251**(1-2): p. 118-123.
120. Fernandez, F., et al., *Association between migraine and a functional polymorphism at the dopamine beta-hydroxylase locus.* Neurogenetics, 2009. **10**(3): p. 199-208.
121. Todt, U., et al., *New genetic evidence for involvement of the dopamine system in migraine with aura.* Human genetics, 2009. **125**(3): p. 265-279.
122. Anttila, V., et al., *Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1.* Nat Genet, 2010. **42**(10): p. 869-73.
123. Vikelis, M. and D.D. Mitsikostas, *The role of glutamate and its receptors in migraine.* CNS & neurological disorders drug targets, 2007. **6**(4): p. 251-7.
124. Formicola, D., et al., *Common variants in the regulative regions of GRIA1 and GRIA3 receptor genes are associated with migraine susceptibility.* BMC medical genetics, 2010. **11**: p. 103.
125. Lafreniere, R.G., et al., *A dominant-negative mutation in the TRESK potassium channel is linked to familial migraine with aura.* Nature medicine, 2010. **16**(10): p. 1157-60.
126. Andres-Enguix, I., et al., *Functional analysis of missense variants in the TRESK (KCNK18) K channel.* Scientific reports, 2012. **2**: p. 237.
127. Di Legge, S. and F. Sallustio, *Migraine relief after patent foramen ovale closure: should vascular risk factor control be achieved first?* Stroke; a journal of cerebral circulation, 2006. **37**(12): p. 2876; author reply 2877.
128. Goyette, P., et al., *Human methylenetetrahydrofolate reductase: isolation of cDNA mapping and mutation identification.* Nature genetics, 1994. **7**(4): p. 551.
129. Heux, S., et al., *The methylentetrahydrofolate reductase gene variant (C677T) as a risk factor for essential hypertension in Caucasians.* Hypertension research : official journal of the Japanese Society of Hypertension, 2004. **27**(9): p. 663-7.

130. Frosst, P., et al., *A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase.* Nature genetics, 1995. **10**(1): p. 111-113.
131. Graham, I.M. and P. O'Callaghan, *Vitamins, homocysteine and cardiovascular risk.* Cardiovascular drugs and therapy / sponsored by the International Society of Cardiovascular Pharmacotherapy, 2002. **16**(5): p. 383-9.
132. Taylor, L.M., Jr., et al., *The association of elevated plasma homocyst(e)ine with progression of symptomatic peripheral arterial disease.* Journal of vascular surgery : official publication, the Society for Vascular Surgery [and] International Society for Cardiovascular Surgery, North American Chapter, 1991. **13**(1): p. 128-36.
133. Mizrahi, E.H., et al., *Further evidence of interrelation between homocysteine and hypertension in stroke patients: a cross-sectional study.* The Israel Medical Association journal : IMAJ, 2003. **5**(11): p. 791-4.
134. Hering-Hanit, R., et al., *Evidence for activation of the coagulation system in migraine with aura.* Cephalalgia : an international journal of headache, 2001. **21**(2): p. 137-9.
135. Parsons, A.A. and P.J. Strijbos, *The neuronal versus vascular hypothesis of migraine and cortical spreading depression.* Current opinion in pharmacology, 2003. **3**(1): p. 73-7.
136. Das, U.N., *Folic acid says NO to vascular diseases.* Nutrition, 2003. **19**(7-8): p. 686-92.
137. van der Put, N.M., et al., *A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects?* American journal of human genetics, 1998. **62**(5): p. 1044-51.
138. Weisberg, I.S., et al., *The 1298A-->C polymorphism in methylenetetrahydrofolate reductase (MTHFR): in vitro expression and association with homocysteine.* Atherosclerosis, 2001. **156**(2): p. 409-15.
139. Kowa, H., et al., *The homozygous C677T mutation in the methylenetetrahydrofolate reductase gene is a genetic risk factor for migraine.* American Journal of Medical Genetics, 2000. **96**(6): p. 762-764.
140. Kara, I., et al., *Association of the C677T and A1298C polymorphisms in the 5,10 methylenetetrahydrofolate reductase gene in patients with migraine risk.* Brain research. Molecular brain research, 2003. **111**(1-2): p. 84-90.
141. Lea, R.A., et al., *The methylenetetrahydrofolate reductase gene variant C677T influences susceptibility to migraine with aura.* BMC medicine, 2004. **2**: p. 3.
142. Scher, A.I., et al., *Migraine and MTHFR C677T genotype in a population-based sample.* Annals of Neurology, 2006. **59**(2): p. 372-5.
143. Oterino, A., et al., *MTHFR T677 homozygosis influences the presence of aura in migraineurs.* Cephalalgia : an international journal of headache, 2004. **24**(6): p. 491-4.
144. Oterino, A., et al., *Thymidylate synthase promoter tandem repeat and MTHFD1 R653Q polymorphisms modulate the risk for migraine conferred by the MTHFR T677 allele.* Brain research. Molecular brain research, 2005. **139**(1): p. 163-8.
145. Todt, U., et al., *MTHFR C677T polymorphism and migraine with aura.* Annals of Neurology, 2006. **60**(5): p. 621-2; author reply 622-3.
146. Kaunisto, M.A., et al., *Testing of variants of the MTHFR and ESR1 genes in 1798 Finnish individuals fails to confirm the association with migraine with aura.* Cephalalgia : an international journal of headache, 2006. **26**(12): p. 1462-72.

147. Rubino, E., et al., *Association of the C677T polymorphism in the MTHFR gene with migraine: a meta-analysis.* Cephalalgia : an international journal of headache, 2009. **29**(8): p. 818-25.
148. Lea, R., et al., *The effects of vitamin supplementation and MTHFR (C677T) genotype on homocysteine-lowering and migraine disability.* Pharmacogenetics and Genomics, 2009. **19**(6): p. 422-428.
149. Wang, X., et al., *Efficacy of folic acid supplementation in stroke prevention: a meta-analysis.* Lancet, 2007. **369**(9576): p. 1876-82.
150. Sharma, P., *Meta-analysis of the ACE gene in ischaemic stroke.* Journal of neurology, neurosurgery, and psychiatry, 1998. **64**(2): p. 227-30.
151. Rigat, B., et al., *An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels.* The Journal of clinical investigation, 1990. **86**(4): p. 1343-6.
152. Paterna, S., et al., *Migraine without aura and ACE-gene deletion polymorphism: is there a correlation? Preliminary findings.* Cardiovascular drugs and therapy / sponsored by the International Society of Cardiovascular Pharmacotherapy, 1997. **11**(4): p. 603-4.
153. Paterna, S., et al., *Angiotensin-converting enzyme gene deletion polymorphism determines an increase in frequency of migraine attacks in patients suffering from migraine without aura.* European neurology, 2000. **43**(3): p. 133-6.
154. Kowa, H., et al., *Association of the insertion/deletion polymorphism of the angiotensin I-converting enzyme gene in patients of migraine with aura.* Neuroscience letters, 2005. **374**(2): p. 129-31.
155. Lea, R.A., et al., *Genetic variants of angiotensin converting enzyme and methylenetetrahydrofolate reductase may act in combination to increase migraine susceptibility.* Brain research. Molecular brain research, 2005. **136**(1-2): p. 112-7.
156. Menon, S., et al., *Association study of calcitonin gene-related polypeptide-alpha (CALCA) gene polymorphism with migraine.* Brain research, 2011. **1378**: p. 119-124.
157. Menon, S., et al., *Association of a Notch 3 gene polymorphism with migraine susceptibility.* Cephalalgia : an international journal of headache, 2011. **31**(3): p. 264-270.
158. MacGregor, E.A., *Headache and hormone replacement therapy in the postmenopausal woman.* Current treatment options in neurology, 2009. **11**(1): p. 10-7.
159. Maggioni, F., et al., *Headache during pregnancy.* Cephalalgia : an international journal of headache, 1997. **17**(7): p. 765-9.
160. Osterlund, M.K. and Y.L. Hurd, *Estrogen receptors in the human forebrain and the relation to neuropsychiatric disorders.* Progress in neurobiology, 2001. **64**(3): p. 251-67.
161. MacGregor, E.A., *Oestrogen and attacks of migraine with and without aura.* Lancet neurology, 2004. **3**(6): p. 354-361.
162. Colson, N.J., et al., *The estrogen receptor 1 G594A polymorphism is associated with migraine susceptibility in two independent case/control groups.* Neurogenetics, 2004. **5**(2): p. 129-133.
163. Oterino, A., et al., *Association of migraine and ESR1 G325C polymorphism.* Neuroreport, 2006. **17**(1): p. 61-64.
164. Oterino, A., et al., *Multilocus Analyses Reveal Involvement of the ESR1, ESR2, and FSHR Genes in Migraine.* Headache, 2008. **48**(10): p. 1438-1450.

165. Schurks, M., P.M. Rist, and T. Kurth, *Sex hormone receptor gene polymorphisms and migraine: A systematic review and meta-analysis.* Cephalalgia : an international journal of headache, 2010. **30**(11): p. 1306-1328.
166. Schurks, M., P.M. Rist, and T. Kurth, *5-HTTLPR polymorphism in the serotonin transporter gene and migraine: A systematic review and meta-analysis.* Cephalalgia : an international journal of headache, 2010. **30**(11): p. 1296-1305.
167. Corominas, R., et al., *Lack of association of hormone receptor polymorphisms with migraine.* European journal of neurology : the official journal of the European Federation of Neurological Societies, 2009. **16**(3): p. 413-5.
168. Joshi, G., S. Pradhan, and B. Mittal, *Role of the oestrogen receptor (ESR1 PvuII and ESR1 325 C->G) and progesterone receptor (PROGINS) polymorphisms in genetic susceptibility to migraine in a North Indian population.* Cephalalgia : an international journal of headache, 2010. **30**(3): p. 311-20.
169. Russell, M.B. and J. Olesen, *A nosographic analysis of the migraine aura in a general population.* Brain : a journal of neurology, 1996. **119 (Pt 2)**: p. 355-61.
170. Hadjikhani, N., et al., *Mechanisms of migraine aura revealed by functional MRI in human visual cortex.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(8): p. 4687-92.
171. Olesen, J., et al., *Timing and topography of cerebral blood flow, aura, and headache during migraine attacks.* Annals of Neurology, 1990. **28**(6): p. 791-8.
172. Geraud, G., et al., *[Positron emission tomographic studies of migraine]*. Revue Neurologique, 2005. **161**(6-7): p. 666-70.
173. Etminan, M., et al., *Risk of ischaemic stroke in people with migraine: systematic review and meta-analysis of observational studies.* BMJ, 2005. **330**(7482): p. 63.
174. Enyedi, P. and G. Czirjak, *Molecular background of leak K+ currents: two-pore domain potassium channels.* Physiological reviews, 2010. **90**(2): p. 559-605.
175. Ferrari, U., et al., *Calcineurin inhibitor-induced headache: clinical characteristics and possible mechanisms.* Headache, 2005. **45**(3): p. 211-4.
176. Anttila, V., et al., *Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1.* Nature genetics, 2010. **42**(10): p. 869-73.
177. Ligthart, L., et al., *Meta-analysis of genome-wide association for migraine in six population-based European cohorts.* European journal of human genetics : EJHG, 2011. **19**(8): p. 901-7.
178. Chasman, D.I., et al., *Genome-wide association study reveals three susceptibility loci for common migraine in the general population.* Nature genetics, 2011. **43**(7): p. 695-8.
179. Peier, A.M., et al., *A TRP channel that senses cold stimuli and menthol.* Cell, 2002. **108**(5): p. 705-715.
180. Dray, A., *Neuropathic pain: emerging treatments.* British Journal of Anaesthesia, 2008. **101**(1): p. 48-58.
181. Freilinger, T., et al., *Genome-wide association analysis identifies susceptibility loci for migraine without aura.* Nature genetics, 2012. **44**(7): p. 777-U205.
182. Flavell, S.W., et al., *Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number.* Science, 2006. **311**(5763): p. 1008-12.
183. Law, C., et al., *Clinical features in a family with an R460H mutation in transforming growth factor beta receptor 2 gene.* Journal of Medical Genetics, 2006. **43**(12): p. 908-16.

184. Cox, H.C., et al., *Heritability and genome-wide linkage analysis of migraine in the genetic isolate of Norfolk Island.* Gene, 2012. **494**(1): p. 119-23.
185. Jones, K.W., et al., *Migraine with aura susceptibility locus on chromosome 19p13 is distinct from the familial hemiplegic migraine locus.* Genomics, 2001. **78**(3): p. 150-154.
186. Nyholt, D.R., et al., *Familial typical migraine - Linkage to chromosome 19p13 and evidence for genetic heterogeneity.* Neurology, 1998. **50**(5): p. 1428-1432.
187. Russo, L., et al., *A new susceptibility locus for migraine with aura in the 15q11-q13 genomic region containing three GABA-A receptor genes.* American Journal of Human Genetics, 2005. **76**(2): p. 327-333.
188. Maher, B.H., et al., *Confirmation that Xq27 and Xq28 are susceptibility loci for migraine in independent pedigrees and a case-control cohort.* Neurogenetics, 2012. **13**(1): p. 97-101.
189. Nyholt, D.R., et al., *Evidence for an X-linked genetic component in familial typical migraine.* Human molecular genetics, 1998. **7**(3): p. 459-463.
190. Bjornsson, *Localization of a gene for migraine without aura to chromosome 4q21 (vol 73, pg 986, 2003).* American Journal of Human Genetics, 2005. **76**(4): p. 715-715.
191. Wessman, M., et al., *A susceptibility locus for migraine with aura, on chromosome 4q24.* American Journal of Human Genetics, 2002. **70**(3): p. 652-662.
192. Cox, H.C., et al., *A genome-wide analysis of 'Bounty' descendants implicates several novel variants in migraine susceptibility.* Neurogenetics, 2012. **13**(3): p. 261-6.
193. Maher, B.H., et al., *An X Chromosome Association Scan of the Norfolk Island Genetic Isolate Provides Evidence for a Novel Migraine Susceptibility Locus at Xq12.* PloS one, 2012. **7**(5).
194. Geetha, R.V.A.R.V.P.L.T., *Mitochondrial DNA and Inherited Diseases â€“ A Comprehensive Review.* Drug Invention Today, 2011. **3**(9).
195. Wong, L.J. and R.G. Boles, *Mitochondrial DNA analysis in clinical laboratory diagnostics.* Clinica chimica acta; international journal of clinical chemistry, 2005. **354**(1-2): p. 1-20.
196. Giles, R.E., et al., *Maternal inheritance of human mitochondrial DNA.* Proceedings of the National Academy of Sciences of the United States of America, 1980. **77**(11): p. 6715-9.
197. Alt, K.W., *The seven daughters of Eve - Why we are the descendants of seven women - Revolutionary findings of genetic research.* Homo-Journal of Comparative Human Biology, 2003. **54**(2): p. 180-180.
198. Barbour, V., *Genetic diversity studies support "out of Africa" theory.* Lancet, 2003. **362**(9396): p. 1635.
199. Taylor, S.W., et al., *Characterization of the human heart mitochondrial proteome.* Nature biotechnology, 2003. **21**(3): p. 281-6.
200. Cotter, D., et al., *MitoProteome: mitochondrial protein sequence database and annotation system.* Nucleic Acids Research, 2004. **32**(Database issue): p. D463-7.
201. Neupert, W., *PROTEIN IMPORT INTO MITOCHONDRIA.* Annual Review of Biochemistry, 1997. **66**(1): p. 863-917.
202. Bandiera, S., et al., *Nuclear outsourcing of RNA interference components to human mitochondria.* PloS one, 2011. **6**(6): p. e20746.
203. Elson, J.L., et al., *Associating mitochondrial DNA variation with complex traits.* American Journal of Human Genetics, 2007. **80**(2): p. 378-82; author reply 382-3.
204. Gibbons, A., *Calibrating the mitochondrial clock.* Science, 1998. **279**(5347): p. 28-9.

205. Mandavilli, B.S., J.H. Santos, and B. Van Houten, *Mitochondrial DNA repair and aging.* Mutation research, 2002. **509**(1-2): p. 127-51.
206. Jiang, Y., et al., *Mitochondrial DNA mutation detection by electrospray mass spectrometry.* Clinical chemistry, 2007. **53**(2): p. 195-203.
207. Zaragoza, M.V., et al., *Mitochondrial DNA Variant Discovery and Evaluation in Human Cardiomyopathies through Next-Generation Sequencing.* PloS one, 2010. **5**(8): p. e12295.
208. He, Y., et al., *Heteroplasmic mitochondrial DNA mutations in normal and tumour cells.* Nature, 2010. **464**(7288): p. 610-4.
209. Tang, S. and T.S. Huang, *Characterization of mitochondrial DNA heteroplasmy using a parallel sequencing system.* BioTechniques, 2010. **48**(4): p. 287-292.
210. Yang, J.L., et al., *Mitochondrial DNA damage and repair in neurodegenerative disorders.* DNA Repair, 2008. **7**(7): p. 1110-20.
211. Egensperger, R., et al., *Association of the mitochondrial tRNA(A4336G) mutation with Alzheimer's and Parkinson's diseases.* Neuropathology and Applied Neurobiology, 1997. **23**(4): p. 315-321.
212. Grazina, M., et al., *Genetic basis of Alzheimer's dementia: role of mtDNA mutations.* Genes Brain and Behavior, 2006. **5**: p. 92-107.
213. DiMauro, S., et al., *Mitochondrial myopathies.* Annals of Neurology, 1985. **17**(6): p. 521-38.
214. Lestienne, P. and G. Ponsot, *Kearns-Sayre syndrome with muscle mitochondrial DNA deletion.* Lancet, 1988. **1**(8590): p. 885.
215. Nishigaki, Y., et al., *Extensive screening system using suspension array technology to detect mitochondrial DNA point mutations.* Mitochondrion, 2010. **10**(3): p. 300-308.
216. Ino, H., et al., *Mitochondrial leucine tRNA mutation in a mitochondrial encephalomyopathy.* Lancet, 1991. **337**(8735): p. 234-5.
217. Tulinius, M., et al., *Leigh syndrome with cytochrome-c oxidase deficiency and a single T insertion nt 5537 in the mitochondrial tRNA(Trp) gene.* Neuropediatrics, 2003. **34**(2): p. 87-91.
218. Shtilbans, A., et al., *G8363A mutation in the mitochondrial DNA transfer ribonucleic acid(Lys) gene: Another cause of Leigh syndrome.* Journal of child neurology, 2000. **15**(11): p. 759-761.
219. Hanna, M.G., et al., *MELAS: a new disease associated mitochondrial DNA mutation and evidence for further genetic heterogeneity.* Journal of Neurology Neurosurgery and Psychiatry, 1998. **65**(4): p. 512-517.
220. Kobayashi, Y., et al., *The Mutant Mitochondrial Genes in Mitochondrial Myopathy, Encephalopathy, Lactic-Acidosis and Stroke-Like Episodes (Melas) Were Selectively Amplified through Generations.* Journal of Inherited Metabolic Disease, 1992. **15**(5): p. 803-808.
221. Hao, H.L. and C.T. Moraes, *Functional and molecular mitochondrial abnormalities associated with a C->T transition at position 3256 of the human mitochondrial genome - The effects of a pathogenic mitochondrial tRNA point mutation in organelle translation and RNA processing.* Journal of Biological Chemistry, 1996. **271**(4): p. 2347-2352.
222. Hayashi, J.I., et al., *Accumulation of Mtdna with a Mutation at Position-3271 in Transfer Rna(Leu)(Uur) Gene Introduced from a Melas Patient to Hela-Cells Lacking Mtdna Results in Progressive Inhibition of Mitochondrial Respiratory-Function.* Biochemical and Biophysical Research Communications, 1993. **197**(3): p. 1049-1055.

223. Uziel, G., et al., *Neuromuscular syndrome associated with the 3291T -> C mutation of mitochondrial DNA: a second case.* Neuromuscular Disorders, 2000. **10**(6): p. 415-418.
224. Sternberg, D., et al., *Mitochondrial DNA transfer RNA gene sequence variations in patients with mitochondrial disorders.* Brain : a journal of neurology, 2001. **124**: p. 984-994.
225. Melone, M.A.B., et al., *Revelation of a new mitochondrial DNA mutation (G12147A) in a MELAS/MERFF phenotype.* Archives of neurology, 2004. **61**(2): p. 269-272.
226. Hammans, S.R., et al., *The Mitochondrial-DNA Transfer Rna(Lys) a-]G(8344) Mutation and the Syndrome of Myoclonic Epilepsy with Ragged-Red Fibers (Merrf) - Relationship of Clinical Phenotype to Proportion of Mutant Mitochondrial-DNA.* Brain : a journal of neurology, 1993. **116**: p. 617-632.
227. Silvestri, G., et al., *A New Mtdna Mutation in the Trans Rnalys Gene Associated with Myoclonic Epilepsy and Ragged-Red Fibers (Merrf).* American Journal of Human Genetics, 1992. **51**(6): p. 1213-1217.
228. Ozawa, M., et al., *Myoclonus epilepsy associated with ragged-red fibers: A G-to-A mutation at nucleotide pair 8363 in mitochondrial tRNA(Lys) in two families.* Muscle & nerve, 1997. **20**(3): p. 271-278.
229. Rojo, A., et al., *NARP-MILS syndrome caused by 8993 T > G mitochondrial DNA mutation: a clinical, genetic and neuropathological study.* Acta neuropathologica, 2006. **111**(6): p. 610-616.
230. Ito, M., et al., *Screening for mitochondrial,DNA heteroplasmy in children at risk for mitochondrial disease.* Mitochondrion, 2001. **1**(3): p. 269-278.
231. Taylor, R.W., et al., *A novel mitochondrial DNA point mutation in the tRNA(Ile) gene: Studies in a patient presenting with chronic progressive external ophthalmoplegia and multiple sclerosis.* Biochemical and Biophysical Research Communications, 1998. **243**(1): p. 47-51.
232. Schaller, A., et al., *Impairment of mitochondrial tRNA(Ile) processing by a novel mutation associated with chronic progressive external ophthalmoplegia.* Mitochondrion, 2011. **11**(3): p. 488-496.
233. Vives-Bauza, C., et al., *Genotype-phenotype correlation in the 5703G > A mutation in the tRNA(Asn) gene of mitochondrial DNA.* Journal of Inherited Metabolic Disease, 2003. **26**(5): p. 507-508.
234. Brown, M.D., et al., *Phylogenetic analysis of Leber's hereditary optic neuropathy mitochondrial DNA's indicates multiple independent occurrences of the common mutations.* Human mutation, 1995. **6**(4): p. 311-25.
235. Zhao, H., et al., *Maternally inherited aminoglycoside-induced and nonsyndromic deafness is associated with the novel C1494T mutation in the mitochondrial 12S rRNA gene in a large chinese family.* American Journal of Human Genetics, 2004. **74**(1): p. 139-152.
236. Guan, M.X., *Molecular pathogenetic mechanism of maternally inherited deafness.* Mitochondrial Pathogenesis: From Genes and Apoptosis to Aging and Disease, 2004. **1011**: p. 259-271.
237. Santorelli, F.M., et al., *Maternally inherited cardiomyopathy and hearing loss associated with a novel mutation in the mitochondrial tRNA(Lys) gene (G8363A).* American Journal of Human Genetics, 1996. **58**(5): p. 933-939.
238. Chomyn, A., et al., *The mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episode syndrome-associated human mitochondrial tRNALeu(UUR) mutation causes aminoacylation deficiency and concomitant reduced association of*

mRNA with ribosomes. The Journal of biological chemistry, 2000. **275**(25): p. 19198-209.
239. Darin, N., et al., *The incidence of mitochondrial encephalomyopathies in childhood: Clinical features and morphological, biochemical, and DNA abnormalities.* Annals of Neurology, 2001. **49**(3): p. 377-383.
240. Sparaco, M., et al., *Neuropathology of Mitochondrial Encephalomyopathies Due to Mitochondrial-DNA Defects.* Journal of Neuropathology and Experimental Neurology, 1993. **52**(1): p. 1-10.
241. Shimomura, T., et al., *Point mutation in platelet mitochondrial tRNA(Leu(UUR)) in patient with cluster headache.* Lancet, 1994. **344**(8922): p. 625.
242. Cortelli, P., et al., *Lack of Association between Mitochondrial T(Rnaleu(Uur)) Point Mutation and Cluster Headache.* Lancet, 1995. **345**(8957): p. 1120-1121.
243. Klopstock, T., et al., *Mitochondrial DNA in migraine with aura.* Neurology, 1996. **46**(6): p. 1735-1738.
244. Seibel, P., et al., *Investigation on the mitochondrial transfer RNA(Leu(UUR)) in blood cells from patients with cluster headache.* Journal of neurology, 1996. **243**(4): p. 305-307.
245. Haan, J., et al., *Search for mitochondrial DNA mutations in migraine subgroups.* Cephalalgia : an international journal of headache, 1999. **19**(1): p. 20-22.
246. Buzzi, M.G., et al., *mtDNA A3243G MELAS mutation is not associated with multigenerational female migraine.* Neurology, 2000. **54**(4): p. 1005-7.
247. Di Gennaro, G., et al., *Assessing the relative incidence of mitochondrial DNA A3243G in migraine without aura with maternal inheritance.* Headache, 2000. **40**(7): p. 568-571.
248. Takeshima, T. and K. Nakashima, *[Current topics: molecular biology of migraine]*. Nihon Naika Gakkai zasshi. The Journal of the Japanese Society of Internal Medicine, 2001. **90**(4): p. 648-53.
249. Finnila, S., et al., *Increased risk of sensorineural hearing loss and migraine in patients with a rare mitochondrial DNA variant 4336A>G in tRNAGln.* Journal of Medical Genetics, 2001. **38**(6): p. 400-5.
250. Zaki, E.A., et al., *Two common mitochondrial DNA polymorphisms are highly associated with migraine headache and cyclic vomiting syndrome.* Cephalalgia : an international journal of headache, 2009. **29**(7): p. 719-28.
251. Boles, R.G., et al., *Are pediatric and adult-onset cyclic vomiting syndrome (CVS) biologically different conditions? Relationship of adult-onset CVS with the migraine and pediatric CVS-associated common mtDNA polymorphisms 16519T and 3010A.* Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society, 2009. **21**(9): p. 936-e72.
252. Finnila, S., I.E. Hassinen, and K. Majamaa, *Restriction fragment analysis as a source of error in detection of heteroplasmic mtDNA mutations.* Mutation research, 1999. **406**(2-4): p. 109-14.
253. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Research, 1988. **16**(3): p. 1215.
254. Almasy, L. and J. Blangero, *Multipoint quantitative-trait linkage analysis in general pedigrees.* American Journal of Human Genetics, 1998. **62**(5): p. 1198-211.
255. Visscher, P.M., W.G. Hill, and N.R. Wray, *Heritability in the genomics era--concepts and misconceptions.* Nature reviews. Genetics, 2008. **9**(4): p. 255-66.

256. Egger, G., et al., *Epigenetics in human disease and prospects for epigenetic therapy.* Nature, 2004. **429**(6990): p. 457-63.
257. Bird, A., *Perceptions of epigenetics.* Nature, 2007. **447**(7143): p. 396-8.
258. Klein, G., *Epigenetics: surveillance team against cancer.* Nature, 2005. **434**(7030): p. 150.
259. Eising, E., et al., *Epigenetic mechanisms in migraine: a promising avenue?* BMC Med, 2013. **11**: p. 26.
260. Takai, D. and P.A. Jones, *Comprehensive analysis of CpG islands in human chromosomes 21 and 22.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3740-5.
261. Riggs, A.D., *X inactivation, differentiation, and DNA methylation.* Cytogenet Cell Genet, 1975. **14**(1): p. 9-25.
262. Riggs, A.D., *X chromosome inactivation, differentiation, and DNA methylation revisited, with a tribute to Susumu Ohno.* Cytogenet Genome Res, 2002. **99**(1-4): p. 17-24.
263. Anamika, K., et al., *Lessons from genome-wide studies: an integrated definition of the coactivator function of histone acetyl transferases.* Epigenetics Chromatin, 2010. **3**(1): p. 18.
264. Schubeler, D., et al., *The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote.* Genes Dev, 2004. **18**(11): p. 1263-71.
265. Carrozza, M.J., et al., *The diverse functions of histone acetyltransferase complexes.* Trends Genet, 2003. **19**(6): p. 321-9.
266. Carninci, P., *Tagging mammalian transcription complexity.* Trends Genet, 2006. **22**(9): p. 501-10.
267. Willingham, A.T. and T.R. Gingeras, *TUF love for "junk" DNA.* Cell, 2006. **125**(7): p. 1215-20.
268. Sorm, F., et al., *5-Azacytidine, a new, highly effective cancerostatic.* Experientia, 1964. **20**(4): p. 202-3.
269. Jones, P.A. and S.M. Taylor, *Cellular differentiation, cytidine analogs and DNA methylation.* Cell, 1980. **20**(1): p. 85-93.
270. Issa, J.P., et al., *Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in hematopoietic malignancies.* Blood, 2004. **103**(5): p. 1635-40.
271. Nandakumar, V., M. Vaid, and S.K. Katiyar, *(-)-Epigallocatechin-3-gallate reactivates silenced tumor suppressor genes, Cip1/p21 and p16INK4a, by reducing DNA methylation and increasing histones acetylation in human skin cancer cells.* Carcinogenesis, 2011. **32**(4): p. 537-44.
272. Marks, P.A., T. Miller, and V.M. Richon, *Histone deacetylases.* Curr Opin Pharmacol, 2003. **3**(4): p. 344-51.
273. Laird, P.W., *The power and the promise of DNA methylation markers.* Nat Rev Cancer, 2003. **3**(4): p. 253-66.
274. Dehan, P., et al., *DNA methylation and cancer diagnosis: new methods and applications.* Expert Rev Mol Diagn, 2009. **9**(7): p. 651-7.
275. Infantino, V., et al., *Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome.* Mol Genet Metab, 2011. **102**(3): p. 378-82.
276. Papageorgiou, E.A., et al., *Fetal-specific DNA methylation ratio permits noninvasive prenatal diagnosis of trisomy 21.* Nat Med, 2011. **17**(4): p. 510-3.

277. Iacobazzi, V., et al., *Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool.* Mol Genet Metab, 2013. **110**(1-2): p. 25-34.
278. Chestnut, B.A., et al., *Epigenetic regulation of motor neuron cell death through DNA methylation.* J Neurosci, 2011. **31**(46): p. 16619-36.
279. Coskun, P., et al., *A mitochondrial etiology of Alzheimer and Parkinson disease.* Biochim Biophys Acta, 2012. **1820**(5): p. 553-64.
280. Finnila, S., et al., *Increased risk of sensorineural hearing loss and migraine in patients with a rare mitochondrial DNA variant 4336A>G in tRNAGln.* J Med Genet, 2001. **38**(6): p. 400-5.
281. Shoffner, J.M., et al., *Mitochondrial DNA variants observed in Alzheimer disease and Parkinson disease patients.* Genomics, 1993. **17**(1): p. 171-84.
282. Wragg, M.A., et al., *No association found between Alzheimer's disease and a mitochondrial tRNA glutamine gene variant.* Neuroscience letters, 1995. **201**(2): p. 107-110.
283. Hutchin, T. and G. Cortopassi, *A mitochondrial DNA clone is associated with increased risk for Alzheimer disease.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(15): p. 6892-5.
284. Brown, M.D., et al., *The role of mtDNA background in disease expression: a new primary LHON mutation associated with Western Eurasian haplogroup J.* Hum Genet, 2002. **110**(2): p. 130-8.
285. Lott, M.T., et al., *mtDNA Variation and Analysis Using MITOMAP and MITOMASTER.* Curr Protoc Bioinformatics, 2013. **1**(123): p. 1.23.1-1.23.26.
286. Sherratt, H.S., *Mitochondria: structure and function.* Revue Neurologique, 1991. **147**(6-7): p. 417-30.
287. Wallace, D.C., *Mitochondrial defects in neurodegenerative disease.* Mental retardation and developmental disabilities research reviews, 2001. **7**(3): p. 158-66.
288. Oldfors, A. and M. Tulinius, *Mitochondrial encephalomyopathies.* Journal of Neuropathology and Experimental Neurology, 2003. **62**(3): p. 217-27.
289. Jenuth, J.P., A.C. Peterson, and E.A. Shoubridge, *Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice.* Nature genetics, 1997. **16**(1): p. 93-95.
290. Silvestri, G., et al., *Single-fiber PCR in MELAS(3243) patients: Correlations between intratissue distribution and phenotypic expression of the mtDNAA(3243G) genotype.* American journal of medical genetics, 2000. **94**(3): p. 201-206.
291. Isobe, K., et al., *Nuclear-recessive mutations of factors involved in mitochondrial translation are responsible for age-related respiration deficiency of human skin fibroblasts.* Journal of Biological Chemistry, 1998. **273**(8): p. 4601-4606.
292. Dunbar, D.R., et al., *Different Cellular Backgrounds Confer a Marked Advantage to Either Mutant or Wild-Type Mitochondrial Genomes.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(14): p. 6562-6566.
293. *mtSNP: a database of human mitochondrial genome polymorphisms.* The New York Academy of Sciences, 2004. **1011**: p. 7-20.
294. Egensperger, R., et al., *Association of the mitochondrial tRNA(A4336G) mutation with Alzheimer's and Parkinson's diseases.* Neuropathol Appl Neurobiol, 1997. **23**(4): p. 315-21.
295. Grazina, M., et al., *Genetic basis of Alzheimer's dementia: role of mtDNA mutations.* Genes Brain Behav, 2006. **5 Suppl 2**: p. 92-107.

296. Tanaka, N., et al., *Mitochondrial DNA variants in a Japanese population of patients with Alzheimer's disease.* Mitochondrion, 2010. **10**(1): p. 32-37.
297. Button, K.S., et al., *Power failure: why small sample size undermines the reliability of neuroscience.* Nat Rev Neurosci, 2013. **14**(5): p. 365-376.
298. Ojaimi, J., et al., *Mitochondrial DNA in stroke and migraine with aura.* Cerebrovascular diseases, 1998. **8**(2): p. 102-6.
299. Chagnon, P., et al., *Phylogenetic analysis of the mitochondrial genome indicates significant differences between patients with Alzheimer disease and controls in a French-Canadian founder population.* American journal of medical genetics, 1999. **85**(1): p. 20-30.
300. Hutchin, T.P., et al., *Mitochondrial DNA mutations in Alzheimer's disease.* Biochemical and biophysical research communications, 1997. **241**(2): p. 221-5.
301. Bandmann, O., et al., *Mitochondrial DNA polymorphisms in pathologically proven Parkinson's disease.* Journal of neurology, 1997. **244**(4): p. 262-265.
302. Mayr-Wohlfart, U., G. Rodel, and A. Henneberg, *Mitochondrial tRNA(Gln) and tRNA(Thr) gene variants in Parkinson's disease.* European journal of medical research, 1997. **2**(3): p. 111-3.
303. Drmanac, R., et al., *Human Genome Sequencing Using Unchained Base Reads on Self-Assembling DNA Nanoarrays.* Science, 2010. **327**(5961): p. 78-81.
304. Thomas, R.K., et al., *Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing (vol 12, pg 852, 2006).* Nature medicine, 2006. **12**(10): p. 1220-1220.
305. Ley, T.J., et al., *DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome.* Nature, 2008. **456**(7218): p. 66-72.
306. Ng, S.B., et al., *Exome sequencing identifies the cause of a mendelian disorder.* Nature genetics, 2010. **42**(1): p. 30-U41.
307. Venter, J.C., et al., *Environmental genome shotgun sequencing of the Sargasso Sea.* Science, 2004. **304**(5667): p. 66-74.
308. Sogin, M.L., et al., *Microbial diversity in the deep sea and the underexplored "rare biosphere".* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(32): p. 12115-12120.
309. Green, R.E., et al., *Analysis of one million base pairs of Neanderthal DNA.* Nature, 2006. **444**(7117): p. 330-336.
310. Arnaud, C.H., *The $1,000 genome.* Chemical & Engineering News, 2005. **83**(51): p. 60-61.
311. Sanger, F., *Citation Classic - DNA Sequencing with Chain-Terminating Inhibitors.* Current Contents/Life Sciences, 1988(50): p. 23-23.
312. Smith, L.M., et al., *Fluorescence Detection in Automated DNA-Sequence Analysis.* Nature, 1986. **321**(6071): p. 674-679.
313. Metzker, M.L., *Applications of Next-Generation Sequencing Sequencing Technologies - the Next Generation.* Nature Reviews Genetics, 2010. **11**(1): p. 31-46.
314. Themissen, A.J.P., *CMOS image sensors: State-of-the-art.* Solid-State Electronics, 2008. **52**(9): p. 1401-1406.
315. Rothberg, J.M., et al., *An integrated semiconductor device enabling non-optical genome sequencing.* Nature, 2011. **475**(7356): p. 348-352.
316. Sakurai, T. and Y. Husimi, *Real-Time Monitoring of DNA-Polymerase Reactions by a Micro Isfet Ph Sensor.* Analytical chemistry, 1992. **64**(17): p. 1996-1997.

317. Broekhuizen, L.N., et al., *Physical activity, metabolic syndrome, and coronary risk: the EPIC-Norfolk prospective population study.* Eur J Cardiovasc Prev Rehabil, 2011. **18**(2): p. 209-17.
318. Goadsby, P.J., *Migraine: diagnosis and management.* Intern Med J, 2003. **33**(9-10): p. 436-42.
319. Sun-Edelstein, C., M.E. Bigal, and A.M. Rapoport, *Chronic migraine and medication overuse headache: clarifying the current International Headache Society classification criteria.* Cephalalgia, 2009. **29**(4): p. 445-52.
320. Mulder, E.J., et al., *Genetic and environmental influences on migraine: a twin study across six countries.* Twin Res, 2003. **6**(5): p. 422-31.
321. Cha, Y.H., M.J. Kane, and R.W. Baloh, *Familial clustering of migraine, episodic vertigo, and Meniere's disease.* Otol Neurotol, 2008. **29**(1): p. 93-6.
322. Karwautz, A.F., et al., *Family-based analysis of serotonin transporter gene polymorphisms in migraine with and without aura.* Cephalalgia, 2007. **27**(7): p. 773-80.
323. Nyholt, D.R., et al., *Genomewide significant linkage to migrainous headache on chromosome 5q21.* Am J Hum Genet, 2005. **77**(3): p. 500-12.
324. Anttila, V., et al., *Consistently replicating locus linked to migraine on 10q22-q23.* Am J Hum Genet, 2008. **82**(5): p. 1051-63.
325. Wessman, M., et al., *A susceptibility locus for migraine with aura, on chromosome 4q24.* Am J Hum Genet, 2002. **70**(3): p. 652-62.
326. Soragna, D., et al., *A locus for migraine without aura maps on chromosome 14q21.2-q22.3.* Am J Hum Genet, 2003. **72**(1): p. 161-7.
327. Russo, L., et al., *A new susceptibility locus for migraine with aura in the 15q11-q13 genomic region containing three GABA-A receptor genes.* Am J Hum Genet, 2005. **76**(2): p. 327-33.
328. Noseda, R. and R. Burstein, *Migraine pathophysiology: anatomy of the trigeminovascular pathway and associated neurological symptoms, CSD, sensitization and modulation of pain.* Pain, 2013. **154 Suppl 1**.
329. Leao, A.A., *Further observations on the spreading depression of activity in the cerebral cortex.* J Neurophysiol, 1947. **10**(6): p. 409-14.
330. Piilgaard, H. and M. Lauritzen, *Persistent increase in oxygen consumption and impaired neurovascular coupling after spreading depression in rat neocortex.* J Cereb Blood Flow Metab, 2009. **29**(9): p. 1517-27.
331. Akerman, S., P.R. Holland, and P.J. Goadsby, *Diencephalic and brainstem mechanisms in migraine.* Nat Rev Neurosci, 2011. **12**(10): p. 570-84.
332. Maher, B.H. and L.R. Griffiths, *Identification of molecular genetic factors that influence migraine.* Mol Genet Genomics, 2011. **285**(6): p. 433-46.
333. Yorns, W.R., Jr. and H.H. Hardison, *Mitochondrial dysfunction in migraine.* Semin Pediatr Neurol, 2013. **20**(3): p. 188-93.
334. Sparaco, M., et al., *Mitochondrial dysfunction and migraine: evidence and hypotheses.* Cephalalgia, 2006. **26**(4): p. 361-72.
335. Sangiorgi, S., et al., *Abnormal platelet mitochondrial function in patients affected by migraine with and without aura.* Cephalalgia, 1994. **14**(1): p. 21-3.
336. D'Andrea, G., et al., *Abnormal platelet trace amine profiles in migraine with and without aura.* Cephalalgia, 2006. **26**(8): p. 968-72.
337. Montagna, P., P. Cortelli, and B. Barbiroli, *Magnetic resonance spectroscopy studies in migraine.* Cephalalgia, 1994. **14**(3): p. 184-93.

338. Brenner, S.R., *Mitochondrial DNA haplogroups influence the therapeutic response to riboflavin in migraineurs.* Neurology, 2010. **74**(2): p. 182-3; author reply 183.
339. Zaki, E.A., et al., *Two common mitochondrial DNA polymorphisms are highly associated with migraine headache and cyclic vomiting syndrome.* Cephalalgia, 2009. **29**(7): p. 719-28.
340. Blakely, E.L., et al., *A mitochondrial cytochrome b mutation causing severe respiratory chain enzyme deficiency in humans and yeast.* FEBS J, 2005. **272**(14): p. 3583-92.
341. Macgregor, S., et al., *Legacy of mutiny on the Bounty: founder effect and admixture on Norfolk Island.* Eur J Hum Genet, 2010. **18**(1): p. 67-72.
342. Bellis, C., et al., *Phenotypical characterisation of the isolated norfolk island population focusing on epidemiological indicators of cardiovascular disease.* Hum Hered, 2005. **60**(4): p. 211-9.
343. Gioio, A.E., et al., *Local synthesis of nuclear-encoded mitochondrial proteins in the presynaptic nerve terminal.* J Neurosci Res, 2001. **64**(5): p. 447-53.
344. Peltonen, L., A. Jalanko, and T. Varilo, *Molecular genetics of the Finnish disease heritage.* Hum Mol Genet, 1999. **8**(10): p. 1913-23.
345. Angius, A., et al., *Archival, demographic and genetic studies define a Sardinian sub-isolate as a suitable model for mapping complex traits.* Hum Genet, 2001. **109**(2): p. 198-209.
346. Refshauge, W.F. and R.J. Walsh, *Pitcairn Island: fertility and population growth, 1790-1856.* Ann Hum Biol, 1981. **8**(4): p. 303-12.
347. Benton, M.C., et al., *Mapping eQTLs in the Norfolk Island genetic isolate identifies candidate genes for CVD risk traits.* Am J Hum Genet, 2013. **93**(6): p. 1087-99.
348. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* Am J Hum Genet, 2007. **81**(3): p. 559-75.
349. Ratan, A., et al., *Comparison of sequencing platforms for single nucleotide variant calls in a human sample.* PLoS One, 2013. **8**(2): p. e55089.
350. Ewing, B. and P. Green, *Base-calling of automated sequencer traces using phred. II. Error probabilities.* Genome Res, 1998. **8**(3): p. 186-94.
351. Ruiz-Pesini, E., et al., *An enhanced MITOMAP with a global mtDNA mutational phylogeny.* Nucleic Acids Res, 2007. **35**(Database issue): p. D823-8.
352. Lorenz, R., et al., *ViennaRNA Package 2.0.* Algorithms Mol Biol, 2011. **6**: p. 26.
353. Ingman, M. and U. Gyllensten, *mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences.* Nucleic Acids Res, 2006. **34**(Database issue): p. D749-51.
354. Petros, J.A., et al., *mtDNA mutations increase tumorigenicity in prostate cancer.* Proc Natl Acad Sci U S A, 2005. **102**(3): p. 719-24.
355. Browning, B.L., *PRESTO: rapid calculation of order statistic distributions and multiple-testing adjusted P-values via permutation for one and two-stage genetic association studies.* BMC Bioinformatics, 2008. **9**: p. 309.
356. Warner, P., *Testing association with Fisher's Exact test.* J Fam Plann Reprod Health Care, 2013. **39**(4): p. 281-4.
357. Wallace, D.C., *A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine.* Annu Rev Genet, 2005. **39**: p. 359-407.
358. Del Zotto, E., et al., *Migraine and ischemic stroke: a debated question.* J Cereb Blood Flow Metab, 2008. **28**(8): p. 1399-421.

359. *Oral contraceptives and stroke in young women. Associated risk factors.* JAMA, 1975. **231**(7): p. 718-22.
360. Lipton, R.B. and S.D. Silberstein, *Why study the comorbidity of migraine?* Neurology, 1994. **44**(10 Suppl 7): p. S4-5.
361. Welch, K.M., *Stroke and migraine--the spectrum of cause and effect.* Funct Neurol, 2003. **18**(3): p. 121-6.
362. Chabriat, H., et al., *Autosomal dominant migraine with MRI white-matter abnormalities mapping to the CADASIL locus.* Neurology, 1995. **45**(6): p. 1086-91.
363. Verin, M., et al., *New phenotype of the cerebral autosomal dominant arteriopathy mapped to chromosome 19: migraine as the prominent clinical feature.* J Neurol Neurosurg Psychiatry, 1995. **59**(6): p. 579-85.
364. Shimomura, T., et al., *Platelet superoxide dismutase in migraine and tension-type headache.* Cephalalgia, 1994. **14**(3): p. 215-8; discussion 181.
365. Torres, J., V. Darley-Usmar, and M.T. Wilson, *Inhibition of cytochrome c oxidase in turnover by nitric oxide: mechanism and implications for control of respiration.* Biochem J, 1995. **312 (Pt 1)**: p. 169-73.
366. Balbi, T., et al., *The presence of trace amines in postmortem cerebrospinal fluid in humans.* J Forensic Sci, 2005. **50**(3): p. 630-2.
367. Boehnke, C., et al., *High-dose riboflavin treatment is efficacious in migraine prophylaxis: an open study in a tertiary care centre.* Eur J Neurol, 2004. **11**(7): p. 475-7.
368. Arts, W.F., et al., *NADH-CoQ reductase deficient myopathy: successful treatment with riboflavin.* Lancet, 1983. **2**(8349): p. 581-2.
369. Di Lorenzo, C., et al., *Mitochondrial DNA haplogroups influence the therapeutic response to riboflavin in migraineurs.* Neurology, 2009. **72**(18): p. 1588-94.
370. Uncini, A., et al., *Abnormal brain and muscle energy metabolism shown by 31P-MRS in familial hemiplegic migraine.* J Neurol Sci, 1995. **129**(2): p. 214-22.
371. Mauskop, A. and J. Varughese, *Why all migraine patients should be treated with magnesium.* J Neural Transm, 2012. **119**(5): p. 575-9.
372. Welch, K.M. and N.M. Ramadan, *Mitochondria, magnesium and migraine.* J Neurol Sci, 1995. **134**(1-2): p. 9-14.
373. Roos-Araujo, D., et al., *Epigenetics and migraine; complex mitochondrial interactions contributing to disease susceptibility.* Gene, 2014. **543**(1): p. 1-7.
374. Xing, G., Z. Chen, and X. Cao, *Mitochondrial rRNA and tRNA and hearing function.* Cell Res, 2007. **17**(3): p. 227-39.
375. Pagliarini, D.J., et al., *A mitochondrial protein compendium elucidates complex I disease biology.* Cell, 2008. **134**(1): p. 112-23.
376. Hendrickson, S.L., et al., *Genetic variants in nuclear-encoded mitochondrial genes influence AIDS progression.* PLoS One, 2010. **5**(9): p. e12862.
377. Colson, N.J., et al., *The estrogen receptor 1 G594A polymorphism is associated with migraine susceptibility in two independent case/control groups.* Neurogenetics, 2004. **5**(2): p. 129-33.
378. Shchepinov, M.S., et al., *Matrix-induced fragmentation of P3'-N5' phosphoramidate-containing DNA: high-throughput MALDI-TOF analysis of genomic sequence polymorphisms.* Nucleic acids research, 2001. **29**(18): p. 3864-72.
379. McCord, J.M., *The evolution of free radicals and oxidative stress.* Am J Med, 2000. **108**(8): p. 652-9.

380. Lightowlers, R.N., et al., *Mammalian mitochondrial genetics: heredity, heteroplasmy and disease.* Trends Genet, 1997. **13**(11): p. 450-5.
381. Cordeiro, M., et al., *The brain-heart connection in mitochondrial respiratory chain diseases.* Neuroradiol J, 2009. **22**(5): p. 558-63.
382. Moon, Y.S., et al., *Suppression of fatty acid synthase promoter by polyunsaturated fatty acids.* J Lipid Res, 2002. **43**(5): p. 691-8.
383. Morcillo, S., et al., *ELOVL6 genetic variation is related to insulin sensitivity: a new candidate gene in energy metabolism.* PLoS One, 2011. **6**(6): p. e21198.
384. Kanehisa, M., et al., *Data, information, knowledge and principle: back to metabolism in KEGG.* Nucleic Acids Res, 2014. **42**(Database issue): p. D199-205.
385. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.
386. Scott, C.R., et al., *Clinical and cellular studies of sarcosinemia.* J Pediatr, 1970. **77**(5): p. 805-11.
387. Meissner, T. and E. Mayatepek, *Sarcosinaemia in a patient with severe progressive neurological damage and hypertrophic cardiomyopathy.* J Inherit Metab Dis, 1997. **20**(5): p. 717-8.
388. Bar-joseph, I., et al., *Mutations in the sarcosine dehydrogenase gene in patients with sarcosinemia.* Hum Genet, 2012. **131**(11): p. 1805-10.
389. Pruitt, K.D., et al., *RefSeq: an update on mammalian reference sequences.* Nucleic Acids Res, 2014. **42**(Database issue): p. D756-63.
390. Ho, T.W., L. Edvinsson, and P.J. Goadsby, *CGRP and its receptors provide new insights into migraine pathophysiology.* Nat Rev Neurol, 2010. **6**(10): p. 573-82.
391. Oshinsky, M.L. and J. Luo, *Neurochemistry of trigeminal activation in an animal model of migraine.* Headache, 2006. **46 Suppl 1**: p. S39-44.
392. Fried, N.T., et al., *Functional Mitochondrial Analysis in Acute Brain Sections from Adult Rats Reveals Mitochondrial Dysfunction in a Rat Model of Migraine.* Am J Physiol Cell Physiol, 2014.
393. Dave, K.R., et al., *Hyperbaric oxygen therapy protects against mitochondrial dysfunction and delays onset of motor neuron disease in Wobbler mice.* Neuroscience, 2003. **120**(1): p. 113-20.
394. Sudakov, N.P., et al., *[Mitochondrial dysfunction in neurodegenerative diseases]*. Zh Nevrol Psikhiatr Im S S Korsakova, 2010. **110**(9): p. 87-91.
395. Antozzi, C., et al., *Late-onset riboflavin-responsive myopathy with combined multiple acyl coenzyme A dehydrogenase and respiratory chain deficiency.* Neurology, 1994. **44**(11): p. 2153-8.
396. Scholte, H.R., et al., *Riboflavin-responsive complex I deficiency.* Biochim Biophys Acta, 1995. **1271**(1): p. 75-83.
397. Schoenen, J., M. Lenaerts, and E. Bastings, *High-dose riboflavin as a prophylactic treatment of migraine: results of an open pilot study.* Cephalalgia, 1994. **14**(5): p. 328-9.
398. Colombo, B., L. Saraceno, and G. Comi, *Riboflavin and migraine: the bridge over troubled mitochondria.* Neurol Sci, 2014. **35 Suppl 1**: p. 141-4.
399. Sandor, P.S., et al., *Efficacy of coenzyme Q10 in migraine prophylaxis: a randomized controlled trial.* Neurology, 2005. **64**(4): p. 713-5.
400. Markley, H.G., *CoEnzyme Q10 and riboflavin: the mitochondrial connection.* Headache, 2012. **52 Suppl 2**: p. 81-7.

401. Rozen, T.D., et al., *Open label trial of coenzyme Q10 as a migraine preventive.* Cephalalgia, 2002. **22**(2): p. 137-41.
402. Harrison, D.E. and J.R. Archer, *Genetic differences in effects of food restriction on aging in mice.* J Nutr, 1987. **117**(2): p. 376-82.
403. Masoro, E.J., *Dietary restriction and aging.* J Am Geriatr Soc, 1993. **41**(9): p. 994-9.
404. McCarter, R.J. and J. Palmer, *Energy metabolism and aging: a lifelong study of Fischer 344 rats.* Am J Physiol, 1992. **263**(3 Pt 1): p. E448-52.
405. Wallace, D.C., *Mitochondrial diseases in man and mouse.* Science, 1999. **283**(5407): p. 1482-8.
406. Shoffner, J.M., et al., *Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation.* Cell, 1990. **61**(6): p. 931-7.
407. Holt, I.J., et al., *A new mitochondrial disease associated with mitochondrial DNA heteroplasmy.* Am J Hum Genet, 1990. **46**(3): p. 428-33.
408. Tatuch, Y., et al., *Heteroplasmic mtDNA mutation (T----G) at 8993 can cause Leigh disease when the percentage of abnormal mtDNA is high.* Am J Hum Genet, 1992. **50**(4): p. 852-8.
409. Hunter, P., *The genetics of human migrations: Our ancestors migration out of Africa has left traces in our genomes that explain how they adapted to new environments.* EMBO Rep, 2014. **15**(10): p. 1019-22.
410. Ross, O.A., et al., *Mitochondrial DNA polymorphism: its role in longevity of the Irish population.* Exp Gerontol, 2001. **36**(7): p. 1161-78.
411. Niemi, A.K., et al., *Mitochondrial DNA polymorphisms associated with longevity in a Finnish population.* Hum Genet, 2003. **112**(1): p. 29-33.
412. Tanaka, M., et al., *Mitochondrial genotype associated with longevity and its inhibitory effect on mutagenesis.* Mech Ageing Dev, 2000. **116**(2-3): p. 65-76.
413. van der Walt, J.M., et al., *Mitochondrial polymorphisms significantly reduce the risk of Parkinson disease.* Am J Hum Genet, 2003. **72**(4): p. 804-11.
414. Gardner, A. and R.G. Boles, *Beyond the serotonin hypothesis: mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders.* Prog Neuropsychopharmacol Biol Psychiatry, 2011. **35**(3): p. 730-43.
415. Karry, R., E. Klein, and D. Ben Shachar, *Mitochondrial complex I subunits expression is altered in schizophrenia: a postmortem study.* Biol Psychiatry, 2004. **55**(7): p. 676-84.
416. Gardner, A. and R.G. Boles, *Mitochondrial energy depletion in depression with somatization.* Psychother Psychosom, 2008. **77**(2): p. 127-9.
417. Gardner, A. and R.G. Boles, *Symptoms of somatization as a rapid screening tool for mitochondrial dysfunction in depression.* Biopsychosoc Med, 2008. **2**: p. 7.
418. Le Goff, P., *Is fibromyalgia a muscle disorder?* Joint Bone Spine, 2006. **73**(3): p. 239-42.
419. Zhang, Z., et al., *High plasma levels of MCP-1 and eotaxin provide evidence for an immunological basis of fibromyalgia.* Exp Biol Med (Maywood), 2008. **233**(9): p. 1171-80.
420. Bo, S.H., et al., *Cerebrospinal fluid cytokine levels in migraine, tension-type headache and cervicogenic headache.* Cephalalgia, 2009. **29**(3): p. 365-72.

Appendices

Appendix A

Table 39: NEMP SNPs significantly associated with migraine in the Norfolk Island population

Chr	Gene	Position	Allele	Test	Odds Ratio	Chi Square	P Value
5	PCDHGB5_rs1059029	140822723	G	ADD	2.182	3.838	0.0001242
5	PCDHGB5_rs3749767	140789933	T	ADD	2.17	3.821	0.0001328
5	PCDHGB5_rs13361997	140863674	C	ADD	2.37	3.768	0.0001648
5	PCDHGB5_rs11952292	140871249	T	ADD	3.061	3.712	0.0002058
5	PCDHGB5_rs6860615	140862840	G	ADD	1.906	3.582	0.0003411
22	PISD_rs12171042	32011225	C	ADD	1.83	3.403	0.0006661
9	SARDH_rs522676	136579589	C	ADD	0.485	-3.358	0.0007839
1	DAP3_rs12097744	155700042	G	ADD	9.136	3.357	0.0007882
22	PISD_rs9956	32015450	C	ADD	1.817	3.354	0.0007959
3	FHIT_rs687342	59853580	G	ADD	2.241	3.242	0.001185
10	KIAA1279_rs2487710	70782126	A	ADD	2.381	3.189	0.001427
9	SARDH_rs493901	136600201	C	ADD	0.5578	-3.176	0.001491
8	CSMD1_rs17066503	3452058	C	ADD	3.349	3.174	0.001503
11	ACAD8_rs1048761	134124124	T	ADD	0.4674	-3.164	0.001557
4	ELOVL6_rs11733718	111073867	G	ADD	0.5673	-3.127	0.001769
4	ELOVL6_rs900328	111074900	C	ADD	0.5687	-3.125	0.001778
10	TFAM_rs11006130	60150445	G	ADD	1.964	3.123	0.00179
17	PHB_rs6917	47481543	T	ADD	2.135	3.067	0.002163

10	KIAA1279_rs12250707	70779436	T	ADD	2.345	3.042	0.002348
22	PISD_rs5994415	32004588	A	ADD	1.726	3.018	0.002542
8	CSMD1_rs7833750	4838089	A	ADD	2.113	3.015	0.002568
1	DBT_rs3806237	100715782	C	ADD	6.973	3.014	0.00258
8	CSMD3_rs16883388	113316520	C	ADD	6.973	3.014	0.00258
5	CSNK1G3_rs4530754	122855416	G	ADD	0.5767	-2.999	0.002708
19	CLPP_rs7253024	6362724	T	ADD	2.109	2.992	0.002775
17	PTRF_rs1905339	40582296	C	ADD	1.656	2.969	0.002985
9	SARDH_rs916620	136596750	A	ADD	0.5372	-2.966	0.003012
7	MRPL32_rs10486743	42979942	C	ADD	2.57	2.964	0.003034
8	CSMD1_rs6993396	2924014	T	ADD	0.5707	-2.958	0.003096
4	ELOVL6_rs7681294	111082310	T	ADD	0.5882	-2.954	0.003141
10	PCBD1_rs877034	72639621	T	ADD	1.804	2.952	0.003157
6	MTHFD1L_rs563440	151288991	G	ADD	0.4867	-2.949	0.003186
10	MTG1_rs4838680	135217956	G	ADD	1.687	2.944	0.003241
11	ACAD8_rs11223738	134128313	T	ADD	0.4918	-2.942	0.003256
20	CST2_rs6076132	23801747	C	ADD	3.807	2.935	0.003336
5	PCDHGB5_rs4151698	140753245	G	ADD	2.158	2.931	0.003384
9	SARDH_rs2073815	136573412	C	ADD	1.683	2.93	0.003393
9	AK3_rs3847258	4717635	C	ADD	1.874	2.909	0.003621
5	CSNK1G3_rs9327298	122850321	A	ADD	1.697	2.907	0.003648
5	CSNK1G3_rs7737667	122875622	G	ADD	1.697	2.907	0.003648
5	PCDHGB5_rs11748256	140861801	G	ADD	1.714	2.893	0.003817
5	CSNK1G3_rs2052485	122882219	A	ADD	0.5863	-2.892	0.003828
5	CSNK1G3_rs6595459	122908361	A	ADD	0.5863	-2.892	0.003828
5	CSNK1G3_rs7705070	122862876	T	ADD	1.696	2.891	0.003838

8	CSMD3_rs16883751	113500330	A	ADD	2.794	2.888	0.003875
5	CSNK1G3_rs10037048	122961813	C	ADD	0.5848	-2.888	0.003881
15	PCSK6_rs3784457	101965792	T	ADD	2.54	2.877	0.00401
1	NDUFS5_rs10888650	39507161	G	ADD	0.5823	-2.874	0.004047
5	SDHAP3_rs7734561	1594096	G	ADD	0.4644	-2.874	0.00405
10	VDAC2_DUP_01_rs7894555	76998231	G	ADD	0.3733	-2.863	0.004193

Appendix B

Table 40: Prioritised SNPs for replication in MAP1 and MAP3

Chr	Gene_rsID	Bp Position	Allele	Test	Odds Ratio	p Value
8	CSMD1_rs6993396	2924014	T	ADD	0.5707	0.003096
8	CSMD1_rs7828513	3432054	T	ADD	0.5825	0.004645
8	CSMD1_rs17066503	3452058	C	ADD	3.349	0.001503
8	CSMD1_rs7815959	3460063	C	ADD	1.73	0.004444
22	PISD_rs5994415	32004588	A	ADD	1.726	0.002542
22	PISD_rs12171042	32011225	C	ADD	1.83	0.0006661
22	PISD_rs9956	32015450	C	ADD	1.817	0.0007959
4	ELOVL6_rs11733718	111073867	G	ADD	0.5673	0.001769
4	ELOVL6_rs900328	111074900	C	ADD	0.5687	0.001778
4	ELOVL6_rs7681294	111082310	T	ADD	0.5882	0.003141
8	CSMD3_rs16883344	113282795	A	ADD	3.833	0.004581
8	CSMD3_rs16883388	113316520	C	ADD	6.973	0.00258
8	CSMD3_rs16883751	113500330	A	ADD	2.794	0.003875
5	CSNK1G3_rs9327298	122850321	A	ADD	1.697	0.003648
5	CSNK1G3_rs4530754	122855416	G	ADD	0.5767	0.002708
5	CSNK1G3_rs7705070	122862876	T	ADD	1.696	0.003838
5	CSNK1G3_rs7737667	122875622	G	ADD	1.697	0.003648
5	CSNK1G3_rs2052485	122882219	A	ADD	0.5863	0.003828
5	CSNK1G3_rs6595459	122908361	A	ADD	0.5863	0.003828
5	CSNK1G3_rs10037048	122961813	C	ADD	0.5848	0.003881
9	SARDH_rs2073815	136573412	C	ADD	1.683	0.003393
9	SARDH_rs522676	136579589	C	ADD	0.485	0.0007839
9	SARDH_rs916620	136596750	A	ADD	0.5372	0.003012
9	SARDH_rs493901	136600201	C	ADD	0.5578	0.001491
5	PCDHGB5_rs4151697	140743661	G	ADD	2.112	0.004523
5	PCDHGB5_rs4151698	140753245	G	ADD	2.158	0.003384
5	PCDHGB5_rs3749767	140789933	T	ADD	2.17	0.0001328
5	PCDHGB5_rs1059029	140822723	G	ADD	2.182	0.0001242
5	PCDHGB5_rs11748256	140861801	G	ADD	1.714	0.003817
5	PCDHGB5_rs6860615	140862840	G	ADD	1.906	0.0003411
5	PCDHGB5_rs13361997	140863674	C	ADD	2.37	0.0001648
5	PCDHGB5_rs11952292	140871249	T	ADD	3.061	0.0002058

Name convention used for novel SNP in Sequenom plex design:

SNP_ID SEQUENCE

11930_A-G

ATGCTAAAACTAATCGTCCCAACAATTATATTACTACCACTGACATGACTT
TCCAAAAAACACATAATTTGAATCAACACAACCACCCACAGCCTAATTATTAGCA
TCATCCCTCTACTATTTTTTAACCAAATCAACAACAACCTATTTAGCTGTTCCCCA
ACCTTTTCCTCCGACCCCCTAACAACCCCCCTCCTAATACTAACTACCTGACTCCT
ACCCCTCACAATCATGGCAAGCCAACGCCACTTATCCAGTGAACCACTATCACGA
AAAAAACTCTACCTCTCTATACTAATCTCCCTACAAATCTCCTTAATTATAACATT
CACAGCCACAGAACTAATCATATTTTATATCTTCTTCGAAACCACACTTATCCCC
ACCTTGGCTATCATCACCCGATGAGGCAACCAGCCAGAACGCCTGAACGCAGGC
ACATACTTCCTATTCTACACCCTAGTAGGCTCCCTTCCCCTACTCATCGCACTAAT
TTACACTCACAACACCCTAGGCTCACTAAACATTCTACTACTCACTCTCACTGCC
CAAGAACTATCAAACTCCTGAGCCAACAACTTAATATGACTAGCTTACACAATAG
CTTTTATAGTAAAGATACCTCTTTACGGACTCCACTTATGACTCCCTAAAGCCCAT
GTCGAAGCCCCCATCGCTGGGTCAATAGTACTTGCCGCAGTACTCTTAAAACTAG
GCGGCTATGGTATAATACGCCTCACACTCATTCTCAACCCCCTGACAAAACACAT
AGCCTACCCCTTCCTTGTACTATCCCTATGAGGCATAATTATAACAAGCTCCATCT
GCCTACGACAAACAGACCTAAAATCGCTCATTGCATACTCTTCAATCAGCCACAT
AGCCCTCGTAGTAACAGCCATTCTCATCCAAACCCCCTGAAGCTTCACCGGCGCA
GTCATTCTCATAATCGCCCACGGGCTTACATCCTCATTACTATTCTGCCTAGCAAA
CTCAAACTACGAACGCACTCACAGTCGCATCATAATCCTCTCTCAAGGACTTCAA
ACTCTACTCCCACTAATAGCTTTTTGATGACTTCTAGCAAGCCTCGCTAACCTCGC
CTTACCCCCCACTATTAACCTACTGGGAGAACTCTCTGTGCTAGTAACCACGTTCT
CCTGATCAAAT[A/G]TCACTCTCCTACTTACAGGACTCAACATACTAGTCACAGCC
CTATACTCCCTCTACATATTTACCACAACACAATGGGGCTCACTCACCCACCACA
TTAACAACATAAAACCCTCATTCACACGAGAAAACACCCTCATGTTCATACACCT
ATCCCCCATTCTCCTCCTATCCCTCAACCCCGACATCATTACCGGGTTTTCCTCTT
GTAAATATAGTTTAACCAAAACATCAGATTGTGAATCTGACAACAGAGGCTTACG
ACCCCTTATTTACCGAGAAAGCTCACAAGAACTGCTAACTCATGCCCCCATGTCT
AACAACATGGCTTTCTCAACTTTTAAAGGATAACAGCTATCCATTGGTCTTAGGC
CCCAAAAATTTTGGTGCAACTCCAAATAAAAGTAATAACCATGCACACTACTATA
ACCACCCTAACCCTGACTTCCCTAATTCCCCCCATCCTTACCA

Table 41: Final Plex Design

SNP ID	Forward Primer ID	Forward Primer Sequence	Reverse Primer ID	Reverse Primer Sequence	Extended Pr ID
rs11952292_W1	rs11952292_W1_F	ACGTTGGATGTAAGACCCCTCAGCGTTCAG	rs11952292_W1_R	ACGTTGGATGCCCTCACCTGGCTCCGCTC	rs11952292_
rs199476128_W1	rs199476128_W1_F	ACGTTGGATGAATACCAAACGCCCCTCTTC	rs199476128_W1_R	ACGTTGGATGTAGTATAGTGATGCCAGCAG	rs199476128 E
rs3749767_W1	rs3749767_W1_F	ACGTTGGATGTTGACACAGAGACCAGGATG	rs3749767_W1_R	ACGTTGGATGATTCTGGCCATTGCCTTGCG	rs3749767_W
rs4530754_W1	rs4530754_W1_F	ACGTTGGATGGAATTCCAGTGAGATTGGAG	rs4530754_W1_R	ACGTTGGATGATCCTATGGTCCTTCCCTTC	rs4530754_W
rs6595459_W1	rs6595459_W1_F	ACGTTGGATGTAGAACTGTGTGGCCCTTAC	rs6595459_W1_R	ACGTTGGATGAGTAGCAAACCACCATGCAG	rs6595459_W
rs16883751_W1	rs16883751_W1_F	ACGTTGGATGGTTTCTCATCACTGAACAGG	rs16883751_W1_R	ACGTTGGATGTTGGACACTGGGAGATTATA	rs16883751_
rs7828513_W1	rs7828513_W1_F	ACGTTGGATGTGATCCACCCACTTTGGCCT	rs7828513_W1_R	ACGTTGGATGCTCCTACTATACTCAGACAC	rs7828513_W
rs9327298_W1	rs9327298_W1_F	ACGTTGGATGAGGTGTATTGGAAGCATGGG	rs9327298_W1_R	ACGTTGGATGTGGAGTGTTACCAGCTCTAC	rs9327298_W
rs2073815_W1	rs2073815_W1_F	ACGTTGGATGAACTACTCCGTCGTCTTCCC	rs2073815_W1_R	ACGTTGGATGTCCCCAGCAGGAGCTGTAG	rs2073815_W
11930_A-G_W1	11930_A-G_W1_F	ACGTTGGATGTCTCTGTGCTAGTAACCACG	11930_A-G_W1_R	ACGTTGGATGGGGCTGTGACTAGTATGTTG	11930_A-G_
rs13361997_W1	rs13361997_W1_F	ACGTTGGATGGCTAAGTGAAAAGCTTGCTC	rs13361997_W1_R	ACGTTGGATGTCCCTTTATCTGACACCAAG	rs13361997_
rs916620_W1	rs916620_W1_F	ACGTTGGATGTAACTCTTATGCTGAAGCGG	rs916620_W1_R	ACGTTGGATGCTTTCTCCCAGAAGCCTTAG	rs916620_W
rs16883388_W1	rs16883388_W1_F	ACGTTGGATGGGAATCAGAGAGAATAAAC	rs16883388_W1_R	ACGTTGGATGTGATATTTGACCAAGTAGCC	rs16883388_
rs41352944_W1	rs41352944_W1_F	ACGTTGGATGGGTCACACGATTAACCCAAG	rs41352944_W1_R	ACGTTGGATGTTTAGCTTTATTGGGGAGGG	rs41352944_
rs7737667_W1	rs7737667_W1_F	ACGTTGGATGACCTTGTATCCCTGTGTTAC	rs7737667_W1_R	ACGTTGGATGATACTAAAGGACAAAGGAC	rs7737667_W
rs9956_W1	rs9956_W1_F	ACGTTGGATGGACCAGGTAGGACTTGAATG	rs9956_W1_R	ACGTTGGATGAGGAACGGGATAGGTTGAGG	rs9956_W1_
rs10037048_W1	rs10037048_W1_F	ACGTTGGATGGAATGTGCTTCCTGACAAAG	rs10037048_W1_R	ACGTTGGATGCTGATCACTTATGGTCTTC	rs10037048_
rs5994415_W1	rs5994415_W1_F	ACGTTGGATGATCTATTCCCAACTGACTGC	rs5994415_W1_R	ACGTTGGATGGTAGATCTCAAGCTCTAGCC	rs5994415_W
rs11748256_W1	rs11748256_W1_F	ACGTTGGATGACCAGGTACTTGTTTTGGTG	rs11748256_W1_R	ACGTTGGATGGACTTACCTAAGCTAAACAAC	rs11748256_
rs16883344_W1	rs16883344_W1_F	ACGTTGGATGGCTTGGAATTCCAAAGTTAC	rs16883344_W1_R	ACGTTGGATGTTGCTAGGAACTGAGAAGAC	rs16883344_
rs7681294_W1	rs7681294_W1_F	ACGTTGGATGCCAGTGCAATGTAATCCAAG	rs7681294_W1_R	ACGTTGGATGGTTCAGGCTGTATATAATCC	rs7681294_W

rs6860615_W1	rs6860615_W1_F	ACGTTGGATGAGATGACAGTCACTGCAGAA	rs6860615_W1_R	ACGTTGGATGTCCCTTCACTTATGACTGAC	rs6860615_W
rs4151698_W1	rs4151698_W1_F	ACGTTGGATGCCCCAAGGTCCTAAGAGATA	rs4151698_W1_R	ACGTTGGATGGCATTTAGAAGAGTGTAAGAG	rs4151698_W
rs7705070_W1	rs7705070_W1_F	ACGTTGGATGTTCATTGCTATAGCTCTGTC	rs7705070_W1_R	ACGTTGGATGCCAGAACACAATGAAGATTGC	rs7705070_W

Table 42: Extension Primer Pooling

Assay	Original Stock Concentration	UEP Mass	Target Reaction Concentration (uM)	Target Working Concentration (uM)	Volume from Stock (uL)
rs11952292	500	4489.9	0.590	5.64	22.58
rs199476128	500	4810.2	0.658	6.30	25.22
rs3749767	500	4899.2	0.677	6.48	25.92
rs4530754	500	5043.3	0.706	6.76	27.03
rs6595459	500	5337.5	0.763	7.30	29.20
rs16883751	500	5492.6	0.791	7.57	30.3
rs7828513	500	5554.6	0.802	7.68	30.73
rs9327298	500	5706.7	0.829	7.94	31.76
rs2073815	500	5834.8	0.852	8.15	32.61
11930a_g	500	5916.9	0.866	8.29	33.15
rs13361997	50	6067	0.891	8.53	34.11
rs916620	500	6207	0.913	8.75	34.98
rs16883388	500	6320.1	0.931	8.92	35.67
rs41352944	500	6431.2	0.949	9.09	36.34
rs7737667	500	6536.3	0.965	9.24	36.96
rs9956	500	6552.3	0.968	9.26	37.06
rs10037048	500	6647.3	0.982	9.40	37.61
rs5994415	500	7006.6	1.035	9.91	39.62
rs11748256	500	7149.7	1.055	10.10	40.4
rs16883344	500	7217.7	1.064	10.19	40.76
rs7681294	500	7286.8	1.074	10.28	41.13
rs6860615	500	7384.8	1.087	10.41	41.64
rs4151698	500	7694	1.128	10.80	43.21
rs7705070	500	7974.2	1.164	11.14	44.58

Appendix C

Table 43: Logistic Regression Analysis between 24 selected SNPs and Migraine Association

CHR	SNP	BP	A1	TEST	OR	STAT	P Migraine Overall	P MA	P MO
4	rs7681294	1.11E+08	C	ADD	0.7296	-3.576	0.0003492	2.55E-05	0.2771
5	rs9327298	1.23E+08	A	ADD	0.07592	-3.558	0.0003733	0.007141	0.9956
9	rs2073815	1.37E+08	C	ADD	1.371	3.35	0.0008086	0.000248	0.9643
5	rs6860615	1.41E+08	G	ADD	1.388	3.279	0.00104	0.008661	0.4156
5	rs3749767	1.41E+08	T	ADD	1.367	2.81	0.004956	0.02412	0.5187
5	rs11748256	1.41E+08	G	ADD	1.297	2.458	0.01396	0.02846	0.7193
5	rs7705070	1.23E+08	T	ADD	0.7955	2.234	0.0255	0.0935	0.6719
5	rs4530754	1.23E+08	G	ADD	1.199	1.972	0.04864	0.009949	0.2127
5	rs11952292	1.41E+08	T	ADD	0.6783	-1.908	0.05638	0.03245	0.954
5	rs7737667	1.23E+08	G	ADD	0.8347	-1.794	0.07286	0.1558	0.8067
8	rs7828513	3432054	T	ADD	1.191	1.684	0.09213	0.1393	0.8285
8	rs16883344	1.13E+08	A	ADD	1.392	1.489	0.1364	0.3709	0.5556
5	rs10037048	1.23E+08	C	ADD	1.135	1.391	0.1643	0.05318	0.2155
2	rs9956	32015450	C	ADD	0.8675	-1.361	0.1736	0.1037	0.5734
5	rs13361997	1.41E+08	C	ADD	0.8705	-1.181	0.2377	0.1859	0.7222
5	rs6595459	1.23E+08	A	ADD	1.111	1.147	0.2516	0.05761	0.1469
8	rs16883751	1.14E+08	A	ADD	0.7813	-1.115	0.265	0.5374	0.3743
5	rs4151698	1.41E+08	G	ADD	1.159	0.86	0.3898	0.6146	0.5111
26	rs41352944	930	A	ADD	1.089	0.3023	0.7625	0.4746	0.1641
22	rs5994415	32004588	A	ADD	1.025	0.2258	0.8213	0.9339	0.6168
9	rs916620	1.37E+08	A	ADD	0.99	-0.11	0.9124	0.4356	0.3605
8	rs16883388	1.13E+08	C	ADD	1.011	0.03409	0.9728	0.6235	0.3194

26	rs199476128	6480	A	ADD	2.10E+09	0.00088	0.9993	0.9993	0.9994
26	11930A>G	11930	0	ADD	NA	NA	NA	NA	NA

CHR	SNP	TEST	A1	A2	GENO	O(HET)	E(HET)	P
4	rs7681294	UNAFF	C	T	137/254/156	0.4644	0.4994	0.104
5	rs9327298	UNAFF	A	G	6/29/209	0.1189	0.1539	0.003048
5	rs4530754	UNAFF	G	A	92/271/178	0.5009	0.4874	0.5381
5	rs7705070	UNAFF	T	C	34/217/295	0.3974	0.3857	0.5788
5	rs7737667	UNAFF	G	T	34/229/290	0.4141	0.3928	0.234
5	rs10037048	UNAFF	C	A	91/265/180	0.4944	0.4862	0.7231
5	rs6595459	UNAFF	A	G	94/269/172	0.5028	0.4894	0.5959
5	rs4151698	UNAFF	G	A	2/65/487	0.1173	0.1168	1
5	rs3749767	UNAFF	T	C	19/157/379	0.2829	0.2896	0.5591
5	rs11748256	UNAFF	G	A	27/182/330	0.3377	0.342	0.8009
5	rs6860615	UNAFF	G	T	37/202/313	0.3659	0.375	0.5709
5	rs13361997	UNAFF	C	A	25/113/426	0.2004	0.2472	4.78E-05
5	rs11952292	UNAFF	T	G	1/65/481	0.1188	0.115	0.7121
8	rs7828513	UNAFF	T	C	48/140/196	0.3646	0.4257	0.005701
8	rs16883344	UNAFF	A	G	1/37/513	0.06715	0.06828	0.5019
8	rs16883388	UNAFF	C	T	1/19/547	0.03351	0.03635	0.1719
8	rs16883751	UNAFF	A	C	1/49/503	0.08861	0.08797	1
9	rs2073815	UNAFF	C	T	65/228/178	0.4841	0.4712	0.6248
9	rs916620	UNAFF	A	G	80/256/223	0.458	0.4673	0.6511
22	rs5994415	UNAFF	A	G	29/167/327	0.3193	0.3377	0.2427
22	rs9956	UNAFF	C	A	39/165/318	0.3161	0.3572	0.009927
26	rs41352944	UNAFF	A	G	0/0/0	nan	nan	NA
26	rs199476128	UNAFF	A	G	0/0/0	nan	nan	NA
26	11930A>G	UNAFF		A	0/0/0	nan	nan	NA

Table 44: Hardy Weinbery Equillibrium

Table 45: Haplotype Structure of tested variants and Association with Migraine Susceptibility

LOC	HAPLOTYPE	F_A	F_U	CHISQ	DF	P	SNPs
H1	OMNIBUS	NA	NA	12.75	4	0.01258	rs9327298|rs4530754|rs7705070|rs7737667|rs10037048|rs6595459
H1	GGCTCA	0.4294	0.4007	1.664	1	0.197	rs9327298|rs4530754|rs7705070|rs7737667|rs10037048|rs6595459
H1	GGCTAA	0.03052	0.03317	0.111	1	0.739	rs9327298|rs4530754|rs7705070|rs7737667|rs10037048|rs6595459
H1	AATGAG	0.05684	0.09985	12.3	1	0.000452	rs9327298|rs4530754|rs7705070|rs7737667|rs10037048|rs6595459
H1	GATGAG	0.1783	0.1689	0.3038	1	0.5815	rs9327298|rs4530754|rs7705070|rs7737667|rs10037048|rs6595459
H1	GACTAG	0.3049	0.2975	0.1287	1	0.7198	rs9327298|rs4530754|rs7705070|rs7737667|rs10037048|rs6595459
H2	OMNIBUS	NA	NA	7.241	2	0.02677	rs11748256|rs6860615
H2	GG	0.2642	0.2143	6.899	1	0.008623	rs11748256|rs6860615
H2	AG	0.04193	0.03947	0.07777	1	0.7803	rs11748256|rs6860615
H2	AT	0.6939	0.7462	6.853	1	0.008851	rs11748256|rs6860615
H3	OMNIBUS	NA	NA	2.736	2	0.2546	rs13361997|rs11952292
H3	CT	0.044	0.05985	2.642	1	0.1041	rs13361997|rs11952292
H3	CG	0.067	0.06262	0.1654	1	0.6843	rs13361997|rs11952292
H3	AG	0.889	0.8775	0.6629	1	0.4156	rs13361997|rs11952292

Table 46: Haplotype Structure of tested *PCDH* variants and Association with Migraine Susceptibility using Haploview

LOCUS	Haplotype	F_A	F_U	ChiSQ	DF	P	SNPs
H1	Omnibus	NA	NA	19.53	3	0.000213	rs11748256/rs686015/rs119522
H1	GGT	0.04462	0.06176	3.036	1	0.08143	rs11748256/rs686015/rs119522
H1	GGG	0.2245	0.1537	17.23	3	3.32E-05	rs11748256/rs686015/rs119522
H1	AGG	0.04352	0.03937	0.2279	1	0.6331	rs11748256/rs686015/rs119522
H1	ATG	0.6873	0.7452	8.638	1	0.003292	rs11748256/rs686015/rs119522

Appendix D

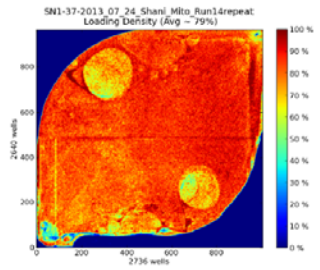
Run Report for Auto_user_SN1-37-2013_07_24_Shani_Mito_Run14repeat_61

Run Summary

294 M	86
Total Bases	Key Signal

79%
ISP Loading

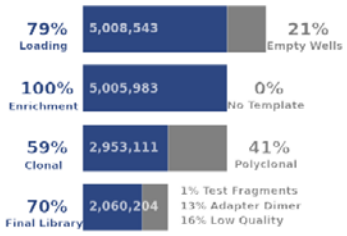
ISP Density



2,060,204
Total Reads

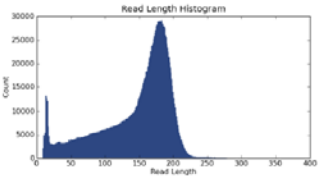
41%
Usable Reads

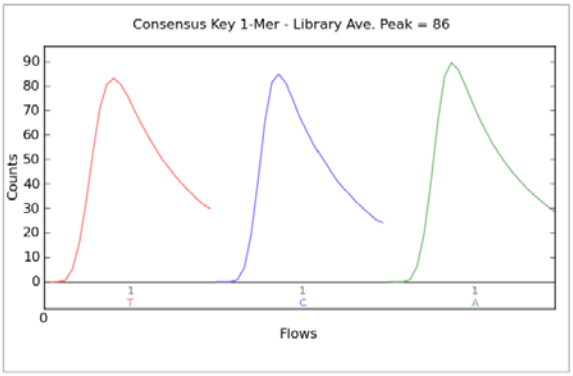
ISP Summary



142 bp	162 bp	183 bp
Mean	Median	Mode

Read Length





Addressable Wells	6,348,235	
With ISPs	5,008,543	78.9%
Live	5,005,983	99.9%
Test Fragment	21,477	00.4%
Library	4,984,506	99.6%

Library ISPs	4,984,506	
Filtered: Polyclonal	2,052,872	41.2%
Filtered: Low Quality	480,179	09.6%
Filtered: Primer Dimer	391,251	07.8%
Final Library ISPs	**2,060,204**	**41.3%**

Barcode Name	Sample	Bases	$\geq Q20$	Reads	Mean Read Length
No barcode	3611	2,630,599	1,885,453	59,751	44 bp
IonXpress_001	2560	6,310,553	5,620,341	45,457	138 bp
IonXpress_002	3601	5,474,537	4,846,400	36,137	151 bp
IonXpress_003	1155	7,177,342	6,358,094	46,318	154 bp
IonXpress_004	2568	6,471,332	5,731,910	42,763	151 bp
IonXpress_005	2589	7,179,829	6,326,698	48,326	148 bp
IonXpress_006	1159	6,526,000	5,736,277	42,687	152 bp
IonXpress_007	2569	5,408,341	4,796,437	37,375	144 bp
IonXpress_008	2591	6,250,843	5,563,696	46,015	135 bp
IonXpress_009	2587	5,934,983	5,288,132	43,459	136 bp
IonXpress_010	3603	6,213,771	5,497,466	41,987	147 bp
IonXpress_011	1151	6,611,446	5,878,180	46,505	142 bp

ion torrent by life technologies

1

IonXpress_012	6891	6,001,672	5,346,440	42,017	142 bp
IonXpress_013	2549	6,443,426	5,727,505	44,097	146 bp
IonXpress_014	2600	7,134,890	6,322,805	49,439	144 bp
IonXpress_015	6848	6,328,796	5,597,816	43,653	144 bp
IonXpress_016	2555	7,265,711	6,454,993	55,464	130 bp
IonXpress_017	2583	5,314,884	4,705,111	36,568	145 bp
IonXpress_018	3607	4,571,905	4,020,990	28,412	160 bp
IonXpress_019	1156	6,264,802	5,538,571	40,479	154 bp
IonXpress_020	3611	6,800,530	5,980,774	44,196	153 bp
IonXpress_021	6807	5,680,781	5,049,787	38,598	147 bp
IonXpress_022	6922	6,340,396	5,600,834	43,470	145 bp
IonXpress_023	6835	5,724,678	5,060,341	37,305	153 bp
IonXpress_024	6802	6,949,897	6,197,170	47,582	146 bp
IonXpress_025	2576	6,953,041	6,190,163	49,917	139 bp
IonXpress_026	3608	7,772,791	6,902,840	52,431	148 bp
IonXpress_027	1169	5,370,961	4,775,662	35,682	150 bp
IonXpress_028	2547	6,236,779	5,572,308	44,998	138 bp
IonXpress_029	6818	5,743,553	5,107,574	41,864	137 bp
IonXpress_030	1163	6,143,027	5,458,822	45,001	136 bp
IonXpress_031	2543	6,280,505	5,616,214	48,089	130 bp
IonXpress_032	6836	4,802,584	4,284,442	39,062	122 bp
IonXpress_033	2536	6,391,451	5,689,130	47,588	134 bp
IonXpress_034	3609	4,908,997	4,356,310	32,472	151 bp
IonXpress_035	1174	2,627,051	2,332,625	17,708	148 bp
IonXpress_036	1177	6,331,306	5,640,552	42,179	150 bp
IonXpress_037	2409	5,975,147	5,258,472	39,355	151 bp
IonXpress_038	1194	6,430,342	5,647,429	41,533	154 bp
IonXpress_039	1195	5,487,737	4,822,415	34,614	158 bp
IonXpress_040	2401	5,010,246	4,405,439	31,986	156 bp
IonXpress_041	2425	6,510,838	5,793,076	44,861	145 bp
IonXpress_042	2413	5,097,585	4,484,557	31,480	161 bp
IonXpress_043	2412	7,510,001	6,591,411	47,666	157 bp
IonXpress_044	2426	6,215,435	5,480,770	40,226	154 bp

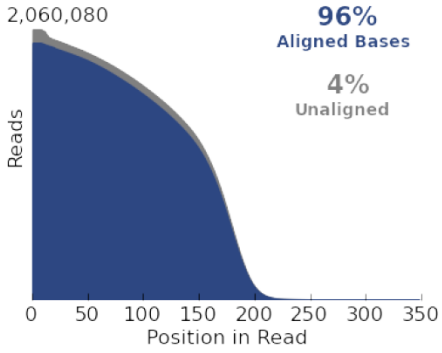
IonXpress_045	2466	5,687,150	5,031,238	37,270	152 bp
IonXpress_046	2467	5,169,117	4,568,131	34,347	150 bp
IonXpress_047	1167	6,166,744	5,505,130	41,941	147 bp
IonXpress_048	1168	6,540,459	5,846,583	49,750	131 bp

Test Fragment	Reads	Percent 50AQ17	Read Length Histogram
TF_D	**9,602**	**93%**	
TF_A	**11,429**	**92%**	

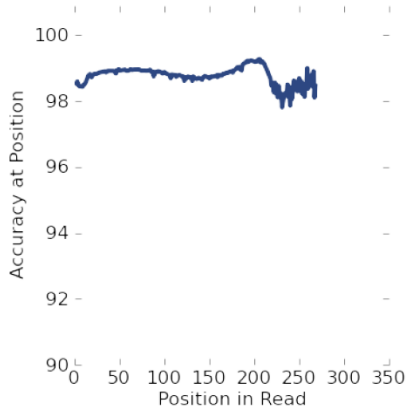
Alignment Summary *(aligned to Human)*

281 M
Total Alignment Bases

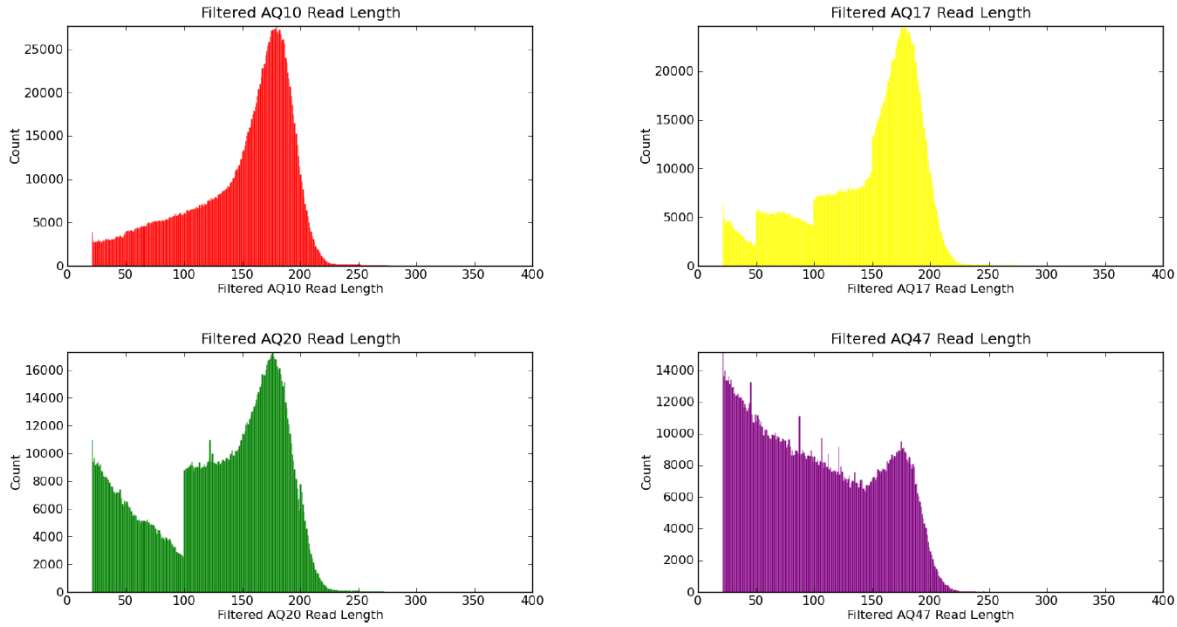
17012.0X
Average Coverage Depth of Reference



98.8%
Mean Raw Accuracy 1x



	AQ17	AQ20	Perfect
Total Number of Bases [Mbp]	255 M	216 M	161 M
Mean Length [bp]	143	128	100
Longest Alignment [bp]	327	315	303
Mean Coverage Depth	15438.8	13040.4	9763.2



variantCaller

Library type: Whole Genome
Targeted regions: None
Hotspot regions: None
Configuration: Germ Line - High Stringency

Barcode Name	Sample Name	Variants	Download Links
IonXpress_001	2560	45	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_002	3601	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_003	1155	56	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_004	2568	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_005	2589	55	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_006	1159	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_007	2569	63	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_008	2591	50	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_009	2587	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_010	3603	42	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_011	1151	46	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_012	6891	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_013	2549	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_014	2600	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_015	6848	50	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_016	2555	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_017	2583	50	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_018	3607	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_019	1156	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_020	3611	51	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_021	6807	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_022	6922	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_023	6835	46	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_024	6802	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_025	2576	51	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_026	3608	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_027	1169	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_028	2547	42	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_029	6818	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_030	1163	53	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_031	2543	51	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_032	6836	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_033	2536	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_034	3609	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_035	1174	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_036	1177	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_037	2409	54	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_038	1194	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_039	1195	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_040	2401	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_041	2425	45	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_042	2413	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_043	2412	51	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_044	2426	46	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_045	2466	49	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_046	2467	48	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_047	1167	47	[VCF.GZ] [VCF.GZ.TBI]
IonXpress_048	1168	50	[VCF.GZ] [VCF.GZ.TBI]

Analysis Details

Run Name	R_2013_07_24_12_22_10_user_SN1-37-2013_07_24_Shani_Mito_Run14repeat
Run Date	July 24, 2013, 12:22 p.m.
Run Flows	500
Projects	mito
Sample	1169, 6891, 1159, 1174, 2589, 6848, 2426, 2555, 2576, 2536, 2413, 2412, 2600, 6802, 3601, 6807, 3609,
Reference	
PGM	sn11c110101
Flow Order	TACGTACGTCTGAGCATCGATCGATGTACAGC
Library Key	TCAG
TF Key	ATCG
Chip Check	Passed
Chip Type	316D
Chip Data	single
Barcode Set	IonXpress
Analysis Name	Auto_user_SN1-37-2013_07_24_Shani_Mito_Run14repeat_61
Analysis Date	July 24, 2013, 4:48 p.m.
Analysis Flows	500
runID	HPY7K

Software Version

Torrent_Suite	3.6
host	FPYM0R1
ion-alignment	3.6.3-1
ion-analysis	3.6.30-1
ion-dbreports	3.6.37-1
ion-gpu	3.6.5-1
ion-pipeline	3.6.21-1
ion-plugins	3.6.32-1
ion-torrentr	3.6.9-1
Script	21.1.4
LiveView	515
DataCollect	445
OS	20
Graphics	32