Heritability and Genome-wide Linkage of Complex Diseases in the Norfolk Island Population Isolate

Elizabeth Matovinovic
Bachelor of Arts
Master of Science

School of Medical Science
Griffith Health
Genomics Research Centre
Griffith University

Submitted in fulfillment of the requirements of the degree of Doctor of Philosophy (PhD)

September 2011
Abstract

Complex diseases such as cardiovascular disease, pterygia, glaucoma, and myopia are caused by polygenic, environmental and lifestyle factors. Studying genetics of complex diseases within population isolates has multiple benefits over studies conducted in single families and unrelated populations. Generally, there is increased power to detect heritable effects. Also, the presence of multiple households within an extended pedigree disentangles confounding environmental variables from true genetic effects. Isolated populations also contain greater potential for identifying underlying causal quantitative trait loci. This thesis work focuses on heritability and genome-wide linkage analyses of cardiovascular disease and three eye diseases: pterygia, glaucoma, and myopia in the Norfolk Island population isolate.

The Norfolk Island pedigree was constructed as part of the Norfolk Island Health Study conducted by Professor Lyn Griffiths at the Griffith University Genomics Research Centre. The complete Norfolk Island pedigree includes 6537 individuals and 11 meiotic generations dating back to the original Bounty ship mutineer and Tahitian founders. Three hundred and seventy-two individuals within the pedigree were genotyped with 400 microsatellite markers spanning chromosomes 1-22. Ocular phenotypic data was obtained through collaboration with the Norfolk Island Eye Study headed by Dr David Mackey (Centre for Eye Research), which characterised epidemiology of eye disease in this population.
This study investigates gene-environment interactions of four complex diseases: cardiovascular disease, pterygia, glaucoma, and myopia. Specific aims for each disease-based set of analyses consist of the following:

1) Create quantitative phenotypes related to the disease. Disease classifications of complex diseases such as cardiovascular disease and glaucoma are often clinical generalisations of multiple distinct disease processes, with similar symptoms but different causes. Hence, in addition to dichotomous yes/no clinically classified disease traits, this study explores the value of using composite, endophenotypic, and environmental based quantitative traits for gene discovery in complex diseases.

2) Calculate the extent of heritability for each phenotypic trait within the Norfolk Island extended pedigree. Heritability estimates calculate the extent of inherited genetic contribution to a disease. However, environmental and lifestyle factors passed down through generations may also influence the estimate.

3) Conduct genome wide linkage scanning to determine whether phenotypes map to particular genetic loci. This approach has been used to identify disease related genes in extended pedigrees.

Previous linkage mapping of individual cardiovascular disease risk factors has been conducted with limited success in identifying causative disease genes. In this study, we applied the Framingham risk score used to predict cardiovascular disease risk in a clinical setting, to heritability and linkage based gene mapping. This composite risk score consists of eight weighted risk factors: age, sex, systolic/diastolic blood pressure, total/high density lipid cholesterol, presence/absence of cigarette smoking, and diabetes status. Since a composite score based on several risk factors is more
effective at predicting cardiovascular disease events than single risk factors, it may also be useful for identifying chromosome regions involved in regulating this polygenic disease. We discovered heritability of the Framingham risk score was negligible and insignificant ($h^2=0.02$, $p=0.42$) whereas individual component risk factors were previously shown to be heritable. Similarly, subsequent linkage analyses were negative (highest LOD 0.01 on chromosome 20).

Pterygia are common ocular surface lesions originating in the limbal conjunctiva with progressive involvement of the cornea. Lesions occur frequently at the nasal limbus with a characteristic wing-like appearance. Pterygia aetiology involves a strong environmental component that includes exposure to ultraviolet radiation, proximity to the equator, dry climates, and outdoor lifestyle. To determine the extent of genetic contribution to this disease, heritability was estimated for pterygia using a dichotomous trait based on presence/absence of the condition. In this study pterygia was not heritable and results were insignificant ($h^2=0.02$, $p=0.44$). However, covariate ultraviolet autofluorescence, a marker for sun exposure, did contribute significantly to explaining pterygia variability ($p=0.01$) indicating an environmental contribution to this condition.

Primary open-angle glaucoma (POAG) is a chronic, progressive multifactorial optic neuropathy with a characteristic acquired loss of optic nerve fibers. Such loss develops in the presence of open anterior chamber angles, characteristic visual field abnormalities, and involves ocular hypertension that gradually damages the eye. One approach to studying genetics of this complex disease involves investigating individual quantitative endophenotypes underlying the disease. Quantitative endophenotypes place all individuals along a continuum of risk, which offers
substantially more information than dichotomous measures of disease status. Intraocular pressure, optic nerve cupping measured by cup-to-disc ratio, central corneal thickness, and anterior chamber depth are putative POAG endophenotypes. In this study, adjusted heritability for dichotomous POAG status was low ($h^2=0.20$) and insignificant ($p=0.14$). Vertical cup-to-disc ratio was the only significant covariate for glaucoma ($p=1\times10^{-6}$) accounting for 0.13 of variation. Central corneal thickness was the most heritable measure with unadjusted $h^2=0.85$ ($p=1.5\times10^{-6}$). Adjusted heritability for central corneal thickness decreased to $h^2=0.77$ ($p=5.7\times10^{-6}$), with intraocular pressure, pterygium, and sphere as significant covariates. Adjusted intraocular pressure heritability was 0.39 ($p=0.008$) with significant covariates anterior chamber depth and ultraviolet autofluorescence, a biomarker for sun exposure accounting for 0.01 of trait variation. Anterior chamber depth adjusted ($h^2=0.37$, $p=0.02$) and unadjusted ($h^2=0.35$, $p=0.02$) heritabilities were similar with significant covariates: age, sex, intraocular pressure, and pterygium. Unadjusted cup-to-disc ratio heritability estimate ($h^2=0.47$, $p=0.001$) was considerably lower than adjusted value ($h^2=0.69$, $p=1\times10^{-5}$). Significant covariates for this trait included age, sex, glaucoma, visual acuity, and pterygium. Hence, these ocular endophenotypes were highly heritable in the Norfolk Island population and estimates are consistent with results from other studies.

Genome-wide linkage scans of these glaucoma related ocular endophenotypes mapped to several loci. Highest peak was LOD 2.5 for cup-to-disc ratio on chromosome 2. Other cup-to-disc ratio peaks >1 were found on chromosomes 1 (LOD 1.8), 9 (LOD 1.2) and 17 (LOD 1.1). Intraocular pressure maximum peak was on chromosome 15 (LOD 1.3), followed by peaks on chromosomes 22 (LOD 1.1) and 8 (LOD 1.2). Chromosomes 20 (LOD 1.9), 11 (LOD 1.7), 14 (LOD 1.2), 4 (LOD
1.3), and 3 (LOD 1.4) were mapped with central corneal thickness. Anterior chamber depth resulted in three peaks on chromosome 20 (LOD 1.8), 17 (LOD 1.3), and 14 (LOD 1.2), 3 (LOD1.5) and 4 (LOD 1.5). In summary, linkage mapping of POAG and corresponding endophenotypes resulted in several peaks, some replicated from other studies, and one providing suggestive evidence of linkage for cup-to-disc ratio on chromosome 12.

The fourth objective of this thesis was to map loci using individual levels of ultraviolet (UV) radiation exposure, a risk factor for myopia. We created a quantitative trait for ultraviolet light exposure by measuring conjunctival UV autofluorescence, a method used in clinical settings to detect sun damage in skin and eyes. We then conducted heritability analyses and genome wide linkage mapping in the Norfolk Island population isolate. We discovered that conjunctival UV autofluorescence is a highly heritable trait ($h^2=0.73$, $p=0.003$). Significant covariates were age, height, glaucoma status, pterygium status, visual acuity, cylinder, axis, and anterior chamber depth, which accounted for 0.27 of the total phenotypic variance of UVAF. Sex was a significant covariate until height was adjusted. Also, eye color was significant prior to adjusting for glaucoma status.

We also identified regions of significant linkage at high myopia locus 7q35 (LOD 4.06) and suggestive evidence at ocular refraction locus 12q24 (LOD 2.5). These results may suggest a common mechanism involving ultraviolet radiation between these two loci. Another region of significant linkage occurred at locus 7q22-q36 (LOD 3.1). Experimental studies show ultraviolet light causes collagen crosslinking, which prevents eye elongation and blocks myopia progression. The ultraviolet autofluorescence measurements comprising the quantitative trait in our analysis are
thought to quantify collagen crosslinking, which supports the experimental model. Hence, these results provide genetic evidence for interactions between UV light and myopia progression, reinforcing experimental and observational studies, and warranting further research into a causative biological mechanism of action. To conclude, this approach was notably effective at mapping myopia related loci previously identified from family studies that used disease status as the phenotype.

In summary, this thesis embodies various phenotypic approaches to designing studies for heritability and linkage analyses of complex cardiovascular and eye disease traits. Phenotypes were created from dichotomous clinical disease classifications, composite quantitative risk factor traits, endophenotypes, and environmental exposure. The most significant heritability estimates and LOD scores were obtained using quantitative traits with abundant covariates. Future studies should fine map significant peak regions identified in this research to narrow down putative genes involved in the complex diseases studied.
Acknowledgements

This research was completed under guidance and supervision of Professor Lyn R Griffiths, at the Genomics Research Centre, from 2007 to 2011 as fulfillment of the Doctor of Philosophy Program in the School of Medical Science and the Griffith Health Institute, Griffith University, Gold Coast, Australia.

I am immeasurably grateful to Lyn for the extraordinary opportunity to study gene-environment interactions in a special population consisting of HMS Bounty mutineer descendants, and to experience living in such a beautiful country as Australia. I benefitted greatly from her wonderful can-do attitude, ability to minimize potentially great obstacles, and to keep moving forward regardless. These are invaluable qualities to surround a PhD student, especially when faced with the daunting task of completing a thesis. I also thank Lyn for insisting I attend the 2011 GeneMappers conference, one of the most intimate and scientifically engaging meetings I have experienced. I hope to be involved with the Australian biomedical genetic research community in the future.

I thank Dr. Rod Lea for co-supervising my work, being approachable and accessible. I appreciate Rod's ability to simplify complicated statistical formulas into easy to understand concepts directly applicable to the task at hand. I also enjoyed exposure to his students' research projects involving gene-environment interactions.

I thank Dr. Larisa Haupt for bringing Lyn's research to my attention, encouraging me to pursue this PhD, and simplifying the whole "move to Australia for a few years"
process. I will always remember our conversation at Park Chow in San Francisco when you first mentioned Lyn and genetic epidemiology related PhD opportunities at Griffith University.

I thank my collaborators Drs David Mackey and Alex Hewitt for initiating the Norfolk Island Eye Study and providing valuable input for manuscripts. I also thank Justin Sherwin for collecting the eye measurements and further supporting the research collaboration.

Thanks to Jeffrey Olive, Gail Davies, and Fred So for providing consistent support to the GRC and Norfolk Island Health Study. Thanks to Hannah Cox for orienting me to previous research conducted with the study and emailing many files needed to complete the work. Also, I will always fondly think back to the time in Napa, CA when Lyn sent Hannah over to help get me started with SOLAR. I also thank Bridget Maher and Emma Cowie for inviting me, a foreigner, out to do fun things. Thanks to Dr. Louise Aldridge for many fun times. Thanks to Nzie Okpokam for help making figures and for some good existential conversation.

Thanks to the Texas Biomedical Research Institute for providing SOLAR support and hosting my confirmation seminar. Thanks to Drs Claire Bellis, Jac Charlesworth, and Richard Polich for helping with SOLAR glitches. Thanks to Dr. Sue Rutherford-Siegel for excellent feedback during the confirmation. Thanks very much to Dr. John Blangero for creating SOLAR and making heritability and linkage analyses feasible for geneticists.
Thanks to the Norfolk Island resident participants. Special thanks to Charisse Clark for the beautiful book and poignant details regarding the author. Thanks to Les Quintal for the lovely glass of wine conversation on what it's like to be a Bounty descendant.

Thanks to Andrew McGrath for inviting me to stay with him while finishing up the thesis, and providing amazing views of Brisbane and Alexandra Headland on the Sunshine Coast. Thanks to Wes Pierson for taking me in during the country-to-country transition. Thanks to Joe Maisel for guitar, support, and Petaluma. Thanks to Blair Fuller for gifting me with an enormous epiphany during this time. Thanks to the kookaburras, lorikeets, whip birds, galahs, and cockatoos for providing the most marvelous, entertaining backdrop to a fascinating PhD experience.
Statement of Originality

The material presented in this thesis has not previously been submitted for a degree or diploma in any university and to the best of my knowledge contains no material previously published or written by another person except where due acknowledgement is made in the thesis itself.

__________________________
Elizabeth Matovinovic
# Table of Contents

Abstract .................................................................................................................. I
Acknowledgements............................................................................................... VII
Statement of Originality......................................................................................... X
Table of Contents .................................................................................................. XI
List of Figures ......................................................................................................... XV
List of Tables ......................................................................................................... XVI
List of Abbreviations ............................................................................................ XVII
Publication List ...................................................................................................... XIX
General Introduction ............................................................................................. 1
Research Aims ........................................................................................................ 6
Research Significance ............................................................................................ 8

Chapter 1: Isolated Population Genetics .............................................................. 11
1.1 Overview .......................................................................................................... 11
   1.1.1 Founder Effect ........................................................................................... 11
   1.1.2 Population Isolates .................................................................................... 12
   1.1.3 Population Bottlenecks ............................................................................. 14
1.3 Complex Disease Gene Mapping ...................................................................... 19
   1.3.1 Family Based Linkage Analysis ................................................................. 19
   1.3.3 Association Analysis ............................................................................... 22
   1.3.4 Genetic Markers Used in Linkage and Association Analyses .............. 23

Chapter 2: Disease Background .......................................................................... 26
   2.2.1 Outcomes Predicted By Risk Score Methods ........................................... 29
   2.2.2 Risk Factors and Scoring Methods ............................................................ 30
2.3 Pterygia ............................................................................................................ 37
   2.3.1 Ultraviolet Radiation and Pterygia ........................................................... 38
2.4 Myopia ............................................................................................................. 39
2.5 Glaucoma ........................................................................................................ 42

Chapter 3: Materials and Methods ..................................................................... 45
3.1 Overview .......................................................................................................... 45
3.2 Ethics ................................................................................................................. 45
   3.2.1 NIHS ....................................................................................................... 45
   3.2.2 NIES ..................................................................................................... 46
3.3 Subject Recruitment ......................................................................................... 46
   3.3.1 NIHS Subject Recruitment .................................................................... 46
Chapter 6: Heritability and Linkage Scan of Glaucoma Endophenotypes in the Norfolk Island Isolate .........................103
6.1 Overview ..............................................................................103
6.2 Background ........................................................................103
   6.2.1 Genetics of Ocular Metric Endophenotypes ..................105
6.3 Methods .............................................................................109
   6.3.1 Ethics ...........................................................................109
   6.3.2 Participant recruitment ..................................................109
   6.3.3. Ophthalmic Evaluation ................................................110
   6.3.4 Pedigree Construction and Genotyping .........................111
   6.3.5 Statistical Analyses .......................................................113
6.4 Results ...............................................................................114
6.5 Discussion .........................................................................116
6.6 Conclusions .......................................................................118

Chapter 7: Genome-wide Linkage Scan of Ultraviolet Radiation Exposure Maps to Myopia Related Loci in the Norfolk Island Isolate .................................................................128
7.1 Background .......................................................................128
7.2 Methods ............................................................................129
   7.2.1 Study Population ..........................................................129
   7.2.2 Ethics ...........................................................................129
   7.2.3 Participant Recruitment ...............................................130
   7.2.4 Conjunctival Ultraviolet Autofluorescence ....................130
   7.2.5 Pedigree Construction ..................................................131
   7.2.6 DNA Extraction and Genotyping .................................132
   7.2.7 Statistical Analysis: Heritability and Linkage ................133
7.3 Results ..............................................................................134
   7.3.1 Sample Population .......................................................134
   7.3.2 Heritability Analyses .....................................................135
   7.3.3 Linkage ......................................................................135
7.4 Discussion .........................................................................136
7.5 Conclusion ........................................................................139

Chapter 8: Conclusions ..............................................................144
8.1 Overview ...........................................................................144
8.2 Framingham Risk Score ......................................................145
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Glaucoma Pathology</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>CVD risk prediction to observed risk ratios by observed population</td>
<td>64</td>
</tr>
<tr>
<td>4.2</td>
<td>Framingham-Wilson based risk score formula</td>
<td>69</td>
</tr>
<tr>
<td>5.1</td>
<td>Power simulation</td>
<td>78</td>
</tr>
<tr>
<td>6.1</td>
<td>Representation of glaucoma ancestry</td>
<td>89</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Intraocular pressure LOD graph for chromosomes 1-22</td>
<td>95</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Anterior chamber depth LOD graph for chromosomes 1-22</td>
<td>95</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Vertical cup-to-disc ratio LOD graph for chromosomes 1-22</td>
<td>96</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Central corneal thickness LOD graph for chromosomes 1-22</td>
<td>96</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Ultraviolet exposure LOD scores across chromosomes 1-22</td>
<td>108</td>
</tr>
<tr>
<td>7.1.2</td>
<td>LOD scores for ultra violet fluorescence on chromosome 7</td>
<td>108</td>
</tr>
<tr>
<td>7.1.3</td>
<td>LOD scores for ultra violet fluorescence on chromosome 12</td>
<td>109</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1 Genetic studies of complex diseases in isolated populations........... 10
Table 1.2 Linkage studies in families with cardiovascular disease .............. 11
Table 2.1 Summary of cardiovascular disease risk factors ....................... 21
Table 2.2 Risk scoring methods............................................................... 25
Table 2.3 Studies of myopia phenotypes with positive association results..... 33
Table 3.1 Number of relative pairs within the Norfolk Island linkage pedigree................................................................. 50
Table 4.1 CVD risk factors comprising the Framingham-Wilson Risk Score.. 71
Table 4.2 Proportion of covariate contribution to explaining FRS variation in the Norfolk Island isolate......................................................... 71
Table 4.3 Bivariate analyses of continuous traits comprising the FRS in the Norfolk Island isolate........................................................................... 72
Table 5.1 Baseline characteristics between two groups................................ 79
Table 5.2 Prevalence of unilateral pterygium by Pitcairn ancestry............... 79
Table 5.3 Prevalence of bilateral pterygium by Pitcairn ancestry............... 80
Table 5.4 Comparison of pterygium severity and UV autofluorescence between Pitcairn and non-Pitcairn ancestry...................................................... 81
Table 5.5 Covariates used in pterygium heritability........................................ 82
Table 6.1 Study population summary............................................................. 90
Table 6.2 Correlations between endophenotypic traits.................................. 91
Table 6.3 Heritability of endophenotypes...................................................... 91
Table 6.4 Bivariate analysis of endophenotypes............................................ 92
Table 6.5 Summary of LOD scores for linkage mapping of endophenotypes.. 92
Table 7.1 Demographic comparison of Bounty and non-Bounty descended Norfolk Island residents............................................................... 98
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI</td>
<td>Applied Biosystems</td>
</tr>
<tr>
<td>ACD</td>
<td>Anterior Chamber Depth</td>
</tr>
<tr>
<td>AGRF</td>
<td>Australian Genome Research Facility</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CCT</td>
<td>Central Corneal Thickness</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary Heart Disease</td>
</tr>
<tr>
<td>Cl</td>
<td>Chloride</td>
</tr>
<tr>
<td>cM</td>
<td>centi Morgans</td>
</tr>
<tr>
<td>C:D</td>
<td>Cup-to-disc ratio</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribo Nucleoside Triphosphate</td>
</tr>
<tr>
<td>DZ</td>
<td>Dizygotic</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FHS</td>
<td>Framingham Heart Study</td>
</tr>
<tr>
<td>FRS</td>
<td>Framingham Risk Score</td>
</tr>
<tr>
<td>GWS</td>
<td>Genome Wide Scan</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide Association Scan</td>
</tr>
<tr>
<td>HDLc</td>
<td>High Density Lipoprotein Cholesterol</td>
</tr>
<tr>
<td>HMAS</td>
<td>His Majesty’s Armed Ship</td>
</tr>
<tr>
<td>IBD</td>
<td>Identity By Descent</td>
</tr>
<tr>
<td>IOP</td>
<td>Intraocular Pressure</td>
</tr>
<tr>
<td>K</td>
<td>Potassium</td>
</tr>
<tr>
<td>Lp(a)</td>
<td>Lipoprotein A</td>
</tr>
<tr>
<td>LDLC</td>
<td>Low Density Lipoprotein Cholesterol</td>
</tr>
<tr>
<td>LOD</td>
<td>Logarithm of the odds</td>
</tr>
<tr>
<td>LUM</td>
<td>Lumican</td>
</tr>
<tr>
<td>MgCl2</td>
<td>Magnesium Chloride</td>
</tr>
<tr>
<td>MIBD</td>
<td>Multipoint Identity By Descent</td>
</tr>
<tr>
<td>MIM</td>
<td>Mendelian Inheritance in Man</td>
</tr>
<tr>
<td>min</td>
<td>Minute</td>
</tr>
<tr>
<td>µL</td>
<td>Micro Litres</td>
</tr>
<tr>
<td>mL</td>
<td>Milli Litre</td>
</tr>
<tr>
<td>mM</td>
<td>Milli Molar</td>
</tr>
<tr>
<td>MMP</td>
<td>Matrix Metalloproteinase</td>
</tr>
<tr>
<td>MZ</td>
<td>Monozygotic</td>
</tr>
<tr>
<td>Na</td>
<td>Sodium</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NIES</td>
<td>Norfolk Island Eye Study</td>
</tr>
<tr>
<td>NIHS</td>
<td>Norfolk Island Health Study</td>
</tr>
<tr>
<td>NKM</td>
<td>Sodium-Potassium-Magnesium</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ng</td>
<td>Nanogram</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometre</td>
</tr>
<tr>
<td>NP</td>
<td>Non-Parametric</td>
</tr>
<tr>
<td>NPL</td>
<td>Non-Parametric LOD</td>
</tr>
<tr>
<td>OD</td>
<td>Optical Density</td>
</tr>
<tr>
<td>pmol</td>
<td>pico mol</td>
</tr>
<tr>
<td>POAG</td>
<td>Primary Open Angle Glaucoma</td>
</tr>
<tr>
<td>PREST</td>
<td>Pedigree RELationship Statistical Test</td>
</tr>
<tr>
<td>QTL</td>
<td>Quantitative Trait Loci</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction Fragment Length Polymorphism</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions Per Minute</td>
</tr>
<tr>
<td>RR</td>
<td>Relative Risk</td>
</tr>
<tr>
<td>sibpair</td>
<td>sibling pair</td>
</tr>
<tr>
<td>SOLAR</td>
<td>Sequential Oligogenic Linkage Analysis Routines</td>
</tr>
<tr>
<td>STR</td>
<td>Short Tandem Repeat</td>
</tr>
<tr>
<td>TAC</td>
<td>Trigeminal Autonomic Cephalalgias TBRI Texas Biomedical Research Institute</td>
</tr>
<tr>
<td>TC</td>
<td>Total Cholesterol</td>
</tr>
<tr>
<td>TCA</td>
<td>Trait Component Analysis</td>
</tr>
<tr>
<td>TDT</td>
<td>Transmission Disequilibrium Test</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>UVAF</td>
<td>Ultraviolet Autofluorescence</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>VC</td>
<td>Variance Component</td>
</tr>
</tbody>
</table>
Publication List

Publications Arising From This Research


Matovinovic E, Sherwin J, Hewitt AW, Mackey DA, Cox, H, Lea RA, Griffiths LR. 

**Genome-wide Linkage Scan of Ultraviolet Radiation Exposure Maps to Myopia Related Loci in the Norfolk Island Isolate.** The Lancet (Under Review)

Matovinovic E, Lea RA, Okpokam N, Hewitt AW, Mackey DA, Griffiths LR. 

**Heritability and Linkage Scan of Glaucoma Endophenotypes in the Norfolk Island Isolate.** (In Preparation)

Matovinovic E, Lea RA, Griffiths LR. **Heritability of the Framingham Risk Score in the Norfolk Island Population Isolate.** (In Preparation)

**Conference Presentations**

Matovinovic E, Lea R, Cox H, Sherwin JC, Hewitt AW, Mackey L, and Griffiths L. 

**Heritability of Refractive Error in the Genetic Isolate of Norfolk Island.** GeneMappers, Hobart, Tasmania, April 2011.


Griffiths LR, Cox HC, Lea RA, Bellis C, Carless M, Dyer TD, Charlesworth J, Matovinovic E, MacGregor S, and Blangero J. **Genome-wide association analysis**
of the Norfolk Island isolate implicates novel variants in migraine susceptibility.
ASHG, Washington DC, USA, November 2010.

TOWARDS the end of 1789, nine mutineers, between ten and twelve native women and six or seven native men left Tahiti in the Bounty. Their fate was unknown to the rest of the world until 1808, when the captain of the American sealer Topaze found Alexander Smith, otherwise John Adams, the only surviving mutineer, eight or nine Polynesian women and several children living in happy isolation on Pitcairn, a volcanic island some four and a half miles in circumference in the remote South Pacific. Dr. Shapiro's lively book embodies the preliminary results of an incredibly energetic ten days' visit to Pitcairn, where he landed exactly a hundred and forty-seven years after the Bounty had sailed from Deptford for the Society archipelago to collect and transport breadfruit plants to the West Indies. Where so much is good, it is difficult to single out particular chapters for consideration, but two of these entitled "Anglo-Polynesian" and "Breeding and Inbreeding" are likely to be of special interest to readers of the present notice and will form its substance. The crossing of distinctive racial elements on Pitcairn forms perhaps the most perfect laboratory situation that the student of miscegenation and its biological effects has yet been able to find. The composition of the parental groups, particularly as regards the European side, is fairly well known. Of the Bounty's company, "Christian came from an ancient, respected, and influential Manx family. Edward Young was the nephew of a baronet. Quintal, McCoy, Adams, and Mills, the other four mutineers who founded families, were ordinary seamen of the working-class. "When an influx of fresh white blood took place, the fact was invariably noted. Details of the Polynesian women are less complete. "One or two had excellent connections, and all of them belonged to a stock famous for its physical beauty." Most important, "the Pitcairn Islanders, unlike other half-caste populations, have never had to eat the bitter bread of social and economic prejudice. Isolated on their fertile island, they are free of the usual disabilities under which half-castes ordinarily labour. They have an abundance of food, nutritious and varied. They are inferior to no one. In other words, they have been allowed to develop in an environment that is wholesome and in a manner dictated by their innate capabilities."
The Norfolk Island study population possesses a unique set of characteristics including founder effect, geographic isolation, limited variation in environmental variation, exhaustive genealogy with well-documented historical records, and higher levels of linkage disequilibrium (LD) extending up to 9.5–11 Mb (Bellis, Cox et al. 2008).

Norfolk Island (34.6km²) is situated 1600km northeast of Sydney, on the Norfolk Ridge that runs from New Zealand to New Caledonia. It was originally a penal colony outpost of Sydney until the last convict settlers were transported to Tasmania in the 1850s (Hoare 1999). In 1789, the Bounty was originally bound for the West Indies until mutineers took control of the ship casting the captain and 18 men adrift in the open sea. Reports indicate the mutineers sailed around the South Pacific for several months before anchoring at Tahiti where several mutineers left the crew and settled. In the same year, nine remaining crewmen, including acting lieutenant Fletcher Christian, six Tahitian men, twelve Tahitian women and a baby girl sailed the HMS Bounty to Tonga and Fiji and finally settled on Pitcairn Island (Hoare 1999) (Eastman 1999).

In June 1856, 194 Pitcairn Islanders; 40 men, 47 women, 53 boys, 53 girls and a baby boy born on the relocation voyage settled uninhabited Norfolk Island. In 2001, the Island’s permanent population totalled 1574 individuals, from which 756 claimed to be of Pitcairn descent (Mathews 2001). The population supports itself from local produce, however as a result of both isolation and small land mass the population is
highly dependent on imports of primary produce and manufactured goods (Mathews and Norfolk 2001). The islanders have a relatively homogeneous lifestyle and diet due to their isolation, strict quarantine and immigration laws and community centred culture. It is expected that the effects of non-genetic variables should be minimised in this population. Also, genetic variables should be reduced due to geographical isolation, known founder effect, early population bottlenecks and inbreeding, as well as strict immigration laws preventing new founders from settling the island (Bellis, Hughes et al. 2005). Genealogical records confirm Pitcairn Island was settled in 1790 by nine Isle of Man Bounty ship mutineers, twelve Tahitian women, and six Tahitian men (Hoare 1999). Pedigree reconstruction confirmed current descendents possess lineages to all nine mutineers, six Tahitian women and three Caucasian sailors who joined the colony in the early 19th century (Bellis, Hughes et al. 2005; McEvoy, Zhao et al. 2009; Macgregor, Hewitt et al. 2010). Census data in 2001 suggested that there were 1,574 permanent residents, of whom 309 were under age 15 years (Mathews 2001). The Jan 2008 census indicated 1619 residents lived on the island.

The Norfolk Island Health Study

In 2000, Professor Lyn Griffiths along with Genomics Research Centre at Griffith University initiated the Norfolk Island Health Study (NIHS) to study cardiovascular disease on Norfolk Island (Bellis, Hughes et al. 2005). The complete Norfolk pedigree includes 11 meiosis and 6537 individuals dating back to the original Bounty mutineer and Tahitian founders. The complete pedigree was simplified to remove inbreeding loops from early generations by blanking the parents of nonfounders in the first four generations of the trimmed pedigree. The trimmed pedigree structure comprises 1064 individuals, 372 of which have both cardiovascular disease related phenotypic and genotypic data available. These individuals were genotyped using
400 microsatellite markers with an average heterozygosity level of 76% spaced at 10cM. Linkage mapping of 600 individuals from Norfolk Island identified a LOD score of 2.01 on chromosome 1p36 for systolic blood pressure (Bellis, Cox et al. 2008). Principal component and linkage analysis support the clustering of cardiovascular disease risk traits and provide interesting evidence of a region on chromosome 5q35 segregating with weight, waist circumference, HDL and total triglyceride levels (Cox, Bellis et al. 2009). Recently, linkage mapping of migraine related loci have been conducted in the Norfolk Island population (Cox et al, under review).

The Norfolk Island Eye Study

The purpose of the Norfolk Island Eye Study (NIES) was to conduct an epidemiological survey of eye diseases and endophenotypes on Norfolk Island (Mackey, Sherwin et al. 2011; Sherwin, Hewitt et al. 2011; Sherwin, Kearns et al. 2011; Sherwin, Kelly et al. 2011).

All 1,275 permanent residents of Norfolk Island aged over 15 years were invited to participate, including 602 individuals involved in the 2001 NIHS study. Participants completed a detailed questionnaire and underwent a comprehensive eye assessment including stereo disc and retinal photography, ocular coherence topography and conjunctival autofluorescence assessment. Additionally, blood or saliva was taken for DNA testing. Seven hundred eighty-one participants aged over 15 years were seen (54% female), comprising 61% of the permanent island population. Three hundred forty-three people (43.9%) could trace their family history to the Pitcairn Islanders. The Norfolk Island Eye Study participants with phenotypic data were located within the complete NIHS pedigree using first name, surname, and date of
birth. Of the 800 individuals sampled from NIES, 495 were located in the complete NIHS pedigree, and approximately 200 (depending on analysis) were used in subsequent SOLAR analyses.
Research Aims

This study aimed to investigate genetics of four complex diseases in the Norfolk Island population isolate: cardiovascular disease, pterygia, glaucoma, and myopia. Complex diseases are caused by a combination of genetic, environmental, and lifestyle factors. Studying complex diseases in extended pedigrees from population isolates have multiple benefits over smaller family studies and heterogeneous populations. Generally, there is increased power to detect heritable effects. Also, environmental effects minimally confound genetic effects because multiple households are included within pedigrees. Isolated populations also contain greater potential for identifying causal quantitative trait loci.

The specific aims are categorized by study disease.

For cardiovascular disease:
1) Create a quantitative phenotypic trait based on the Framingham Risk Score
2) Estimate the heritable component of the Framingham Risk Score
3) Conduct a genome-wide linkage scan to detect quantitative trait loci for cardiovascular disease

For pterygia:
1) Use the binary clinical classification of pterygia as a dichotomous trait
2) Estimate heritable component of pterygia
3) Conduct a genome-wide linkage scan to detect quantitative trait loci for pterygia
For glaucoma:

1) Use several endophenotypes for glaucoma as quantitative traits

2) Estimate heritable component of glaucoma endophenotypes using an abundant selection of genetic and environmental covariates

3) Conduct a genome-wide linkage scan to identify quantitative trait loci for glaucoma

For myopia:

1) Create an environmental based phenotype that quantifies ultraviolet radiation exposure

2) Estimate the heritable component of this trait, adjusting for several covariates

3) Conduct a genome-wide linkage scan to identify quantitative trait loci for myopia
Research Significance

Cardiovascular Disease
Fifty years ago, cardiovascular genetics did not exist (Milewicz and Seidman 2000). Genetics was an emerging field of basic research not yet integrated with clinical science. Today, cardiovascular genetics is a large interdisciplinary field integrating high throughput technology with clinical medicine. Insights from these discoveries have uprooted traditional clinical disease classifications, and molecular genetics have been incorporated into the study of pathology.

The significance of these contributions can be observed at the Online Mendelian Inheritance in Man® website (http://www.ncbi.nlm.nih.gov/omim) that documents over 12,000 human genes and genetic phenotypes. Here, one third of the 800 inherited human phenotypes listed in the database have a cardiovascular component.

Glaucoma
Glaucoma is a substantial health problem. Glaucoma is the leading cause of irreversible blindness worldwide. Chronic glaucoma prevalence in the US population 40 years and older is 1.86% (Quigley 1996). With an aging population, this prevalence is projected to increase by 50% in 2020 (Quigley 1996). Australian prevalence is similar at 3.0% (Mutti, Mitchell et al. 2002). Hence, any advance in understanding this disease will affect a great number of people. Second, identifying glaucoma related genes contribute to further understanding pathogenesis at the cellular and molecular level. Model systems, such as transgenic mice, can be created to study disease progression, gene-environment interactions, and test new therapies.
In addition, genetic information may improve clinician counseling, monitoring and treatment of patients. A patient with a mutation known to produce optic nerve cupping and visual field loss rapidly, for example, might be treated in a preventative manner prior to disease manifestation. Also, addressing the underlying genetic cause may shed light on other risk factors or systemic diseases related to glaucoma. Finally, identifying genes involved in disease pathogenesis may provide novel targets for therapy.

**Pterygia**

Studying pterygia genetics has several public health benefits. Pterygia may be used as a biomarker for other diseases correlated with ultraviolet radiation. Ultraviolet radiation exposure is strongly associated with eye disease, both acute (acute photokeratitis or flash burns) and chronic (pterygium, climatic droplet keratopathy, cataract, ocular surface squamous neoplasia, eyelid tumours, basal and squamous cell carcinoma and melanoma) (Oliva and Taylor 2005). Therefore, documenting increased risk of developing ultraviolet radiation related disease, such as pterygia, whether due to genetics or lifestyle, has important public health consequences. At present, public health experts advocate that sun exposure should be guided by level of ambient ultraviolet radiation and skin type (Lucas, McMichael et al. 2008). Knowledge of an individual’s ultraviolet radiation response through pterygia detection may lead to targeted prevention and therapeutic strategies.

**Myopia**

Currently, over 80% of young adults in East Asian cities have myopia (nearsightedness) (Wu, Seet et al. 2001; He, Zeng et al. 2004; Lin, Shih et al. 2004), in the US prevalence has doubled to 40% in 30 years (Vitale, Sperduto et al. 2009)
suggested strong environmental influences. Myopic trait heritability varies, ranging from high (Chen, Scurrah et al. 2007; Klein, Suktitipat et al. 2009) in urbanized settings to low (Biino, Palmas et al. 2005; Vitart, Bencic et al. 2010) in island populations further reflecting environmental contributions (Young, Leary et al. 1969; Morgan, Speakman et al. 1975; Lin, Hung et al. 1988; Visscher, Hill et al. 2008). Recently, more time spent outdoors has emerged as a protective factor against myopia (Mutti 2010) while risk factors such as education level (Tay, Au Eong et al. 1992) and near work (Saw, Chan et al. 2008), coincide with more time indoors. Hence, variability in UV light exposure among these environmental factors is an underlying trait that may also play a role in myopia (Prepas 2008). The significance of studying ultraviolet radiation and myopia is enormous due to the ubiquitous environmental human exposure to the sun, which plays a dose dependent role in preventing and causing several diseases. Moreover, since the extent of ultraviolet radiation exposure is determined by personal lifestyle choices, this topic is practically relevant and significant to everyone.
Chapter 1: Isolated Population Genetics

1.1 Overview

1.1.1 Founder Effect

A founder effect occurs when a new colony is established from a few members (<200) of the large original population. This small population size means that the colony may have reduced genetic variation from the original population and a non-random sample of the genes in the original population. Moreover, social and cultural isolation often accompanies these populations.

Chromosome and mitochondrial DNA analysis confirm that these populations are less genetically diverse than outbred populations. This is also evident by the high frequency of rare mendelian disorders, particularly in the Finnish, Amish, Sardinian and Bedouin isolates (Arcos-Burgos and Muenke 2002). Many disorders in these populations can be traced back to the population founders. A trait allele that is introduced into a population by one of the original founders is referred to as founder effect. Founder mutations are particularly prevalent in the Dutch population (Zeegers, van Poppel et al. 2004). This genetically and environmentally isolated type of population has advantages for identifying genes involved in familial-based diseases.
1.1.2 Population Isolates

Many population isolates are known to exist worldwide including Saguenay-Lac St-Jean (Canada) (De Braekeleer 1991), Kosraean (Shmulewitz, Heath et al. 2006), Icelandic (Moises, Yang et al. 1995), Croatian (Safra, Schaible et al. 2006; Navarro, Vitart et al. 2010), Sardinian (Biino, Palmas et al. 2005) old order Amish (Ginns, Ott et al. 1996), Hutterites (Ober, Abney et al. 2001), Bedouins (Guilford, Ben Arab et al. 1994), Pasia Community (Arcos-Burgos, Castellanos et al. 2004), Pima Indians (Hanson, Ehm et al. 1998) and Finnish (Pajukanta, Terwilliger et al. 1999) populations. This study focuses on the Norfolk Island population isolate (Bellis, Hughes et al. 2005; Bellis, Cox et al. 2008). Table 1.1 below summarizes genetic studies of complex diseases in isolated populations.

Table 1.1 Genetic studies of complex diseases in isolated populations adapted from (Kristiansson, Naukkarinen et al. 2008)

<table>
<thead>
<tr>
<th>Complex disease/trait</th>
<th>Gene/locus</th>
<th>Population/isolate</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective disorders</td>
<td>Several</td>
<td>N. Sweden</td>
<td>(Venken, Claes et al. 2005)</td>
</tr>
<tr>
<td>Asthma</td>
<td>IRAK-M (interleukin-1 receptor-associated kinase M)</td>
<td>Sardinia</td>
<td>(Terwilliger, Zollner et al. 1998)</td>
</tr>
<tr>
<td>Asthma</td>
<td>CHI3L1 (chitinase 3-like 1)</td>
<td>Hutterites</td>
<td>(Ober, Tan et al. 2008)</td>
</tr>
<tr>
<td>Asthma</td>
<td>NPSR1 (neuropeptide S receptor 1)</td>
<td>Finland</td>
<td>(Laitinen, Polvi et al. 2004)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>4q25</td>
<td>Iceland</td>
<td>(Gudbjartsson, Arnar et al. 2007)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>5q31-34</td>
<td>Antioquia (Colombia), Central Valley of Costa Rica</td>
<td>(Herzberg, Jasinska et al. 2006)</td>
</tr>
<tr>
<td>Blood pressure (systolic)</td>
<td>1p36</td>
<td>Norfolk Island</td>
<td>(Bellis, Cox et al. 2008)</td>
</tr>
<tr>
<td>Disorder</td>
<td>Chromosomal Location</td>
<td>Country</td>
<td>Reference</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Bone mineral density</td>
<td>Several</td>
<td>Iceland</td>
<td>(Styrkarsdottir, Halldorsson et al. 2008)</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>5p12, 2q35, 16q12</td>
<td>Iceland</td>
<td>(Stacey, Manolescu et al. 2008)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>9p21</td>
<td>Iceland</td>
<td>(Helgadottir, Thorleifsson et al. 2008)</td>
</tr>
<tr>
<td>Coronary heart disease</td>
<td>8p22</td>
<td>French Canadian</td>
<td>(Engert, Lemire et al. 2008)</td>
</tr>
<tr>
<td>Crohn’s disease</td>
<td>Several</td>
<td>French Canadian</td>
<td>(Raelson, Little et al. 2007)</td>
</tr>
<tr>
<td>Exfoliation glaucoma</td>
<td>LOXL1 (lysyl oxidase-like 1)</td>
<td>Iceland</td>
<td>(Thorleifsson, Magnusson et al. 2007)</td>
</tr>
<tr>
<td>Fasting glucose levels</td>
<td>G6PC2 (glucose-6-phosphatase, catalytic, ABCB11 (ATP-binding cassette, sub-family B(MDR/TAP), member 11) region2</td>
<td>Finland, Sardinia</td>
<td>(Chen, Saxena et al. 2008)</td>
</tr>
<tr>
<td>Height</td>
<td>Several loci</td>
<td>Iceland</td>
<td>(Gudbjartsson, Walters et al. 2008)</td>
</tr>
<tr>
<td>Height</td>
<td>GDF5 (growth differentiation factor 5) UQCC (ubiquinol-cytochrome c reductase complex chaperone) locus</td>
<td>Finland, Sardinia, Amish</td>
<td>(Sanna, Jackson et al. 2008)</td>
</tr>
<tr>
<td>Migraine</td>
<td>ADABR2</td>
<td>Norfolk Island</td>
<td>unpublished</td>
</tr>
<tr>
<td>Nicotine dependence and smoking-related diseases</td>
<td>15q24</td>
<td>Iceland</td>
<td>(Thorgeirsson, Geller et al. 2008)</td>
</tr>
<tr>
<td>Obesity</td>
<td>FTO (fat mass and obesity associated), PFKP (phosphofructokinase, platelet)</td>
<td>Sardinia</td>
<td>(Scuteri, Sanna et al. 2007)</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>GBA (β-glucocerebrosidase)</td>
<td>Ashkenazi</td>
<td>(Gan-Or, Giladi et al. 2008)</td>
</tr>
<tr>
<td>Pigmentation</td>
<td>Several</td>
<td>Iceland</td>
<td>(Sulem, Gudbjartsson et al. 2007)</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>Xp11.22, 2p15, 17q</td>
<td>Iceland</td>
<td>(Gudmundsson, Sulem et al. 2008)</td>
</tr>
<tr>
<td>Psychotic and bipolar spectrum</td>
<td>TSNAX (translin-associated factor X) DISC1(disrupted in</td>
<td>Finland</td>
<td>(Palo, Antila et al. 2007)</td>
</tr>
</tbody>
</table>
1.1.3 Population Bottlenecks

A novel haplotype surrounds the founder trait allele. This haplotype degrades with each meiosis due to recombination related events until ultimately the trait is lost. Haplotype signatures in population isolates depend on population demographics (Peltonen, Palotie et al. 2000). A limited number of founders combined with isolation produce high rates of endogenous and consanguineous marriages, which conserve trait alleles and haplotypes in such populations. Significant haplotype reductions can arise due to natural disasters such as disease, war or large storms. This phenomenon is referred to as a population bottleneck. This event is generally followed by rapid expansion of a small random number of original population members. Recovery from a severe bottleneck in isolated populations leads to inbreeding and genetic drift. In the case of genetic disease, this may cause alleles to be lost or pushed to a much higher frequency (Maruyama and Fuerst 1984; Maruyama and Fuerst 1985; Maruyama and Fuerst 1985; Luikart, Allendorf et al. 1998; Hey and Harris 1999) compared to outbred populations as has been observed in Polynesians (Flint, Boyce et al. 1989) and Hutterites (Newman, Hoffjan et al. 2004). Reduced genetic variation means that the population may not be able to adapt to new selection pressures, such as climatic change or a shift in available resources, because
the genetic variation that selection would act on may have already drifted out of the population.

1.1.4 Linkage Disequilibrium

The older a trait allele, the shorter the surrounding haplotype sequence as observed in older outbred populations (Kristiansson, Naukkarinen et al. 2008). Comparably, a small number of isolated founders generally result in extended haplotype blocks surrounding trait alleles. Consequently, isolated founder populations have extended regions of linkage disequilibrium compared to older outbred populations (Katoh, Mano et al. 2002; Rahman, Jones et al. 2003). In terms of positional cloning, a lower density marker map is typically more informative in young founder isolates compared to old outbred populations. The co-inheritance of a trait allele with adjacent DNA markers at the population level is known as linkage disequilibrium (LD). The extent of LD around a particular allele is the result of selection and population history (population size, genetic drift, admixture, mutation and recombination events), which act to degrade founder chromosome containing trait alleles (Palmer and Cardon 2005). Linkage disequilibrium is proportional to the number of meioses that took place after the bottleneck. Due to the heterogeneous nature, LD patterns across the genome are unpredictable and can only be assessed through sample analysis. However, a young isolated founder population should exhibit a greater extent of LD due to conservation of the founder chromosomes. Founder effect combined with isolation should result in a select number of susceptibility genes over time, with unique patterns of haplotype sharing and LD in the vicinity of disease.
The combination of both geographical and cultural isolation leads to individuals sharing a common environment. Differences in lifestyle factors such as diet, exercise and sanitation are minimized compared to outbred populations. This has been observed particularly in Amish and Hutterite populations (Ginns, Ott et al. 1996; Hsueh, Mitchell et al. 2000; Ober, Tsalenko et al. 2000). The homogeneous environment that is shared by individuals in such populations is beneficial for complex disease studies such as cardiovascular and eye diseases where a threshold effect for the disease may exist that is heavily influenced by environmental factors such as diet (Shmulewitz, Heath et al. 2006) and sun exposure. The reduction of such variables makes disease gene mapping more robust to detect true effects.

1.2 Demographic History of the Norfolk Island Isolate

Norfolk Island (34.6km$^2$) is situated 1600km northeast of Sydney, on the Norfolk Ridge that runs from New Zealand to New Caledonia. Two smaller volcanic islands, Nepean and Phillip are flora and fauna reserves located south of Norfolk Island. Captain James Cook discovered the uninhabited Norfolk Island on October 10 1774 and claimed the island under the British Crown. The first penal settlement was established in 1788. The initial population comprised 22 individuals of whom 15 were convicts. Due to convict uprisings, the first penal settlement closed in 1814 and the island was abandoned. Norfolk Island was re-occupied in 1827 as a convict station under the British Empire, however, more violent uprisings forced the station to be closed. All inhabitants were relocated to Hobart, Tasmania in 1856. Norfolk was to remain uninhabited until June 1856 when it was suggested as a new home for Pitcairn Islanders, founders of the modern Norfolk Island population (Hoare 1999).
Pitcairn Island settlement originates from the infamous mutiny aboard the HMS Bounty ship. In 1789, the Bounty was originally bound for the West Indies until mutineers took control of the ship casting the captain and 18 men adrift in the open sea. Reports indicate the mutineers sailed around the South Pacific for several months before anchoring at Tahiti where several mutineers left the crew and settled. In the same year, nine remaining crewmen, including acting lieutenant Fletcher Christian, six Tahitian men, twelve Tahitian women and a baby girl sailed the HMS Bounty to Tonga and Fiji and finally settled on Pitcairn Island (Hoare 1999) (Eastman 1999).

In 1767, Pitcairn Island was discovered as a small volcanic island located in the Pacific Ocean approximately 4,830 km northeast of Norfolk Island. The island was uninhabited until the 28 individuals comprising the Bounty landed on its shores in January 1790. The men destroyed the ship to prevent discovery and ensure isolation, as they were fugitives of the British Crown.

The initial years of settlement were turbulent due to rampant violence. All the Tahitian men and seven of the mutineers met brutal deaths. Of the two remaining mutineers, only John Adams survived. Records indicate that the population descended from twelve maternal Tahitian lineages (Polynesian ancestry) and nine paternal European lineages (Isle of Man and British ancestry) (Hoare 1999). Due to the limited number of founders and the geographical isolation of Pitcairn, the early stages of population expansion included complex relationships with high levels of inbreeding (Bellis, Hughes et al. 2005). The small population suffered epidemics,
drought, food shortages and consequently malnutrition. Finally the community was relocated to Norfolk Island in 1856.

In June 1856, 194 Pitcairn Islanders; 40 men, 47 women, 53 boys, 53 girls and a baby boy born on the relocation voyage settled the then uninhabited island of Norfolk. Strict immigration laws allowing only those of Pitcairn descent to occupy Norfolk Island were established. In 1863, 27 islanders left Norfolk to resettle Pitcairn Island (Hoare 1999).

In 2001, the Island’s permanent population totalled 1574 individuals, from which 756 claimed to be of Pitcairn descent (Mathews 2001). The population supports itself from local produce, however as a result of both isolation and small land mass the population is highly dependent on imports of primary produce and manufactured goods (Mathews and Norfolk 2001). The islanders live a relatively homogeneous lifestyle and diet due to their isolation, strict quarantine and immigration laws and community centred culture. It is expected that the effects of non-genetic variables should be minimised in this population. Also, genetic variables should be reduced due to geographical isolation, known founder effect, early population bottlenecks and inbreeding, as well as strict immigration laws preventing new founders from settling the island (Bellis, Hughes et al. 2005).

Anthropologists from the Island have maintained genealogical history of the population extending back to the initial founding of Pitcairn Island by the Bounty mutineers and Tahitian women. The Norfolk pedigree contains 6537 individuals and eleven meiosis spanning 240 years (Bellis, Hughes et al. 2005). DNA and phenotypic
information is available for individuals comprising the lower four generations of the pedigree structure (Bellis, Hughes et al. 2005).

The first “genetic” study on Norfolk Island was published in 1927 (Shapiro, 1927). Dr. Shapiro “found these islanders to be of sound physique, taller than the average English and Tahitians, and of good mentality.” “The Norfolk Islanders prove that, when the stock is sound to begin with, intensive in-breeding makes for no decrease in stamina. Likewise, race mixture, in his opinion, brings no deterioration.” Notably 75 years later MacGregor et al. found that the more inbred individuals of Norfolk Island had reduced height (Macgregor, Bellis et al. 2010).

1.3 Complex Disease Gene Mapping

Family based genome-wide linkage studies and population based gene association studies using candidate genes have been used to discover or characterise causes of complex disorders such as CVD (Keating 1992; Seda, Tremblay et al. 2008), glaucoma (Crooks, Allingham et al. 2011) and myopia (Paget, Julia et al. 2008).

1.3.1 Family Based Linkage Analysis

Family based linkage analysis is an approach to identify genes without a priori assumptions regarding the genetic basis of the disease (Dyer, Blangero et al. 2001). Linkage analysis is a hypothesis generating method intended to localise genomic
regions that might contain genes influencing a trait. It relies on co-segregation of chromosomal regions that strays from Mendel’s second law of independent assortment. Nonrandom assortment or linkage of alleles during meiotic recombination is more likely to occur when alleles are closer to each other. Predefined markers are used to track genome regions. Table 1.2 below lists cardiovascular disease gene regions identified using linkage analysis.

**Table 1.2 Linkage studies in families with cardiovascular disease**

<table>
<thead>
<tr>
<th>Study</th>
<th>Popn</th>
<th>#families/ #individuals</th>
<th>Clinical event</th>
<th>Chrom. region</th>
<th>LOD score</th>
<th>Gene family</th>
<th>Analytic method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pajukanta, 2000</td>
<td>Finland</td>
<td>156/364</td>
<td>Premature CAD</td>
<td>2q21-22 Xq23-26</td>
<td>3.7</td>
<td>no</td>
<td>Allele sharing (MAPMAKE R SIBS)</td>
</tr>
<tr>
<td>Francke, 2001</td>
<td>Mauritius</td>
<td>99/535</td>
<td>CAD</td>
<td>16p13</td>
<td>3.06</td>
<td>no</td>
<td>Allele sharing (MAPMAKE R SIBS)</td>
</tr>
<tr>
<td>Broeckel, 2002</td>
<td>Germany</td>
<td>513/1406</td>
<td>MI</td>
<td>14</td>
<td>3.9</td>
<td>no</td>
<td>Variance component (SOLAR)</td>
</tr>
<tr>
<td>Harrap, 2002</td>
<td>Australia</td>
<td>61/161</td>
<td>CAD</td>
<td>2q36</td>
<td>2.6</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Wang, 2003</td>
<td>USA</td>
<td>1/25</td>
<td>CAD or MI</td>
<td>15q26</td>
<td>4.19</td>
<td>MEF2</td>
<td>Allele sharing (SAGE)</td>
</tr>
<tr>
<td>Helgadottir, 2004</td>
<td>Iceland</td>
<td>296/2454</td>
<td>MI</td>
<td>13q12-13</td>
<td>2.86</td>
<td>ALOX5 AP</td>
<td>Allele sharing</td>
</tr>
<tr>
<td>Hauser, 2004</td>
<td>Euro-American</td>
<td>438/1168</td>
<td>Premature CAD</td>
<td>3q13</td>
<td>3.3</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Samani, 2005</td>
<td>UK</td>
<td>1933 /4175</td>
<td>CAD or MI</td>
<td>2q14-21</td>
<td>1.15</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Helgadottir, 2006</td>
<td>Iceland</td>
<td>296/2454</td>
<td>MI</td>
<td>12q22</td>
<td>NA</td>
<td>LTA4H</td>
<td></td>
</tr>
<tr>
<td>Farrall, 2006</td>
<td>Europe</td>
<td>2036/2658</td>
<td>CAD or MI</td>
<td>17q21</td>
<td>2.68</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

Several steps are involved in conducting linkage analysis. First, polymorphic markers are genotyped at known locations across the genome. These marker data, along with phenotype and pedigree data are entered into biostatistical algorithms to calculate the extent the marker information is identical by descent (IBD) among family members in the pedigree and how this degree of genetic similarity for a marker correlates with the phenotypic resemblance among family members. The
logarithm of the odds (LOD) score is calculated for the probability that a marker and putative causal locus are linked in pedigrees and contrasted to the null hypothesis of the absence of linkage in the pedigrees. A LOD score of 3 suggests significant linkage for autosomal markers while 2 is the cutoff for x-linked genes. A LOD score less than -2 excludes linkage.

1.3.2 Parametric vs. Non-Parametric Analyses

There are two statistical models of linkage analysis, parametric and non-parametric: Parametric analysis evaluates the transmission of marker alleles under a defined disease inheritance model to produce an LOD score. LOD scores are calculated to measure the likelihood that inheritance of an allele at a specific marker in affected individuals is the result of linkage opposed to chance. LOD scores are calculated using logs to the base 10 and a parameter θ, known as the recombination fraction in 2-point analysis or map position in multipoint analysis (Dawn Teare and Barrett 2005). For X-linked markers, evidence of linkage is indicated by an LOD score equal or greater than 2. For autosomal markers a score > 3 indicates linkage. A LOD score less than -2 excludes linkage.

Non-parametric analysis reports allele-sharing frequencies in affected relatives without any assumed parameters; it is model free. This analysis is ideal for complex traits with an undefined genetic model such as CVD, pterygia, glaucoma, and myopia. Nonparametric analysis uses all the genotype data for markers simultaneously to estimate the proportion of alleles shared at any point along the
chromosome. Excess allele sharing between affected relatives is indicative of linkage to the trait. This method assumes that in the presence of a disease locus, affected relatives will exhibit excess sharing of haplotypes that are IBD regardless of whether a disease model is specified. Hence this form of analysis is also referred to as allele sharing methods. Linkage is reported as a likelihood ratio using natural logarithms. This score can be converted to a traditional LOD score by dividing by 4.6 (ie, $2x\log_{10}$) (Dawn Teare and Barrett 2005).

1.3.3 Association Analysis

Association analysis compares differences in disease allele frequencies between affected (cases) and unaffected (controls) individuals matched for age, sex, and other variables thought to confound the analysis. Allele distribution is determined using non-parametric Chi-Square analysis, logistic regression and odds ratios. More recently family-based association testing and DNA pooling approaches have gained interest. Marker and trait associations can be causal, spurious, indirectly associated through proximity to the true causal variant (linkage disequilibrium), the result of admixture by a population substructure.

Association studies have more power to detect genes of small effect but require the need for greater marker density and large cohort numbers. Other limitations include the presence of population substructure (stratification), which can generate false positive associations (Hoggart, Parra et al. 2003). Population stratification can arise if the total population is derived from a combination of subpopulations leading to racial admixture, as is the case with Norfolk Island (Macgregor, Bellis et al. 2010),
African American, African Caribbean, Hispanic American populations (Hoggart, Parra et al. 2003). Studies have shown substantial differences in the frequencies of common polymorphisms and LD in different ethnic populations (Carlson and Monti 2003). Such admixture can be controlled for if specific allele frequencies for each subpopulation are known for a set of markers. Alternatively a family-based approach eliminates confounding admixture and is suitable for linkage analyses with the Norfolk Island population (Macgregor, Bellis et al. 2010).

1.3.4 Genetic Markers Used in Linkage and Association Analyses

Genetic markers are variable DNA sequences that localise to a single locus. Markers are generally heterogeneous, which facilitates inter-individual comparisons. There are many genetic markers available to study traits. This research is particularly concerned with short tandem repeats (STRs) and single nucleotide polymorphisms (SNPs).

Single nucleotide polymorphisms are genetic variants that arise from a single base mutation. They are abundant and have a low mutation rate in the human genome (Carlson and Monti 2003). Some are functional and cause disease, while others comprise intron regions of the genome that do not encode genes. SNP alleles are often in linkage disequilibrium with other alleles adjacent or at nearby sites.

Microsatellites, or short tandem repeats (STRs) are repetitive DNA sequences dispersed evenly at fixed chromosomal locations throughout the genome and frequently occur within genes (Bennett 2000). The repeat sequence may consist of 1-5 base pairs designated (X)n , where X is the repeat unit and n is the total number of
repeats. Hence (CA)15 denotes an allele with 15 CA repeats. Microsatellites are highly polymorphic as the repeat expansions often display a high degree of variability in numbers.

Microsatellites may have originated through slipped strand mispairing (SSM) during DNA replication (Bennett 2000). In prokaryotes, they act as contingency genes while in eukaryotes the function is unknown. Until recently microsatellites were considered to be extra non-coding DNA. Due to their highly polymorphic nature, fast mutational rate and high density in the genome, microsatellites have more information content than SNPs, making them ideal for linkage studies (Broman, Murray et al. 1998; Weiss and Terwilliger 2000).

Genotyping of STRs is achieved using polymerase chain reaction (PCR) and capillary electrophoresis. PCR is a technique used to amplify a target DNA sequence in a genome. It can amplify as little as 0.2 μL of DNA a million fold (Baumforth, Nelson et al. 1999) and increased output is maximized by multiplexing (pooling) PCR products. PCR fragments may be analysed by agarose gel electrophoresis (AGtE). Though this technique is simple and effective, it requires manual handling and has a low resolution and sensitivity. It has become a standard procedure to genotype microsatellites using ABI Genetic Analysers. These automated machines use capillary electrophoresis, electrokinetic injection and fluorescence to accurately genotype PCR products to a resolution of two nucleotides (Butler, Buel et al. 2004).

1.3.5 SOLAR
This research uses the statistical program Sequential Oligogenic Linkage Analysis Routines (v.4.2.0), (SOLAR) (Blangero and Almasy 1997; Almasy and Blangero 1998) for heritability and linkage analyses. This program can perform many types of analyses, including genetic variance components analysis, linkage analysis, quantitative genetic analysis, and covariate screening. The program is specifically designed for linkage analysis of quantitative traits that may involve multiple loci (oligogenic), dominance effects, and epistasis. Although SOLAR is designed for quantitative traits, qualitative traits can also be analysed by appropriately coding trait data. SOLAR calculates multi-point non-parametric LOD scores using IBD sharing probabilities in relatives, negating the need for assumptions regarding the genetic model and is capable of analysing SNP and STR data.

The computational demands of linkage calculations limits pedigree size, however trimming or splitting the pedigree using a subset of the genealogy can overcome these limitations. Cutting the pedigree at different generations breaks original loops and reduces power to detect linkage (Falchi, Forabosco et al. 2004). It has been shown that analysing extended pedigrees yield greater statistical power than only analysing select groups of relatives (Dyer, Blangero et al. 2001). SOLAR has been used to identify cardiovascular disease genes (Broekel, Hengstenberg et al. 2002) and also used in large complex pedigrees from Sardinia (Biino, Palmas et al. 2005) to identify eye disease genes.
Chapter 2: Disease Background

2.1 Overview

This chapter summarises literature for cardiovascular disease risk scores, pterygia, myopia, and glaucoma diseases.

2.2 Defining Cardiovascular Disease

Cardiovascular disease (CVD) refers to a range of diseases impacting the heart or blood vessels such as coronary artery disease (CAD), coronary heart disease (CHD), heart attack, heart failure, high blood pressure and stroke. There are several etiologies of CVD ranging from rare single gene mendelian disorders to “common disease, common variants” (Hirschhorn and Daly 2005) causes where multiple biological processes combine with environmental triggers to result in various clinical outcomes. This multiplicity of genetic and environmental risk factors provides a challenging environment for elucidating mechanisms of action that can be extrapolated or replicated across various study populations.

Although CVD is a major cause of mortality worldwide, trends in industrialized countries have been stabilizing in certain populations. In the US, from 1980-2000, the age-adjusted CVD death rate declined substantially from 543 to 267 deaths per 100,000 among men and from 263 to 134 deaths per 100,000 among women (Ford, Ajani et al. 2007; Ford and Capewell 2007). Half of this decrease was attributed to advances in secondary preventive therapies following myocardial infarction (MI) or
vascularization, initial treatments for acute MI or unstable angina, treatments for heart failure, revascularization for chronic angina, and other therapies (Ford and Capewell 2007). The other half of this decrease was due to reductions in risk factors, including total cholesterol, systolic blood pressure, smoking prevalence, and physical inactivity. However, these reductions were partially offset by increases in body-mass index (BMI) and diabetes, which accounted for an increased number of deaths (Ford, Ajani et al. 2007; Ford and Capewell 2007).

Similar trends have been observed for England and Wales where the age-adjusted mortality rate decreased by 54.7% in men and by 48.3% in women (O'Flaherty, Ford et al. 2008). However, among men aged 35-44 years, CHD mortality rates in 2002 increased after 20 years of steady decline. Recent declines in CHD mortality rates are tapering in both men and women aged 45-54. Among older adults, however, mortality rates continued to decrease steadily throughout the period. Analogous to the US study, leveling CHD mortality rates in these countries were attributed to increases in BMI and diabetes.

The primary cause of CHD is atherosclerosis, a progressive chronic disease process characterised by accumulated lipids and fibrous elements in medium and large arteries (Lusis, Mar et al. 2004). Both genetic and environmental risk factors contribute to its development (Anderson, Odell et al. 1991; Wilson, D'Agostino et al. 1998; Assmann, Cullen et al. 2002; Conroy, Pyorala et al. 2003).

Family and twin studies have consistently found that genetic factors contribute importantly to premature CHD (Lusis, Mar et al. 2004). Identical twin studies indicate that premature cardiac death in one sibling is associated with an 8-fold
greater risk to the surviving male sibling and a 15-fold increase in risk to female siblings (Marenberg, Risch et al. 1994). Premature cardiac death risk is significantly lower in fraternal twins (2.6 for women and 3.8 for men). Other studies also indicate that family history in a parent or a sibling is associated with atherosclerotic CVD in the form of stroke or peripheral arterial disease (Williams, Hunt et al. 2001; Lloyd-Jones, Nam et al. 2004).

Several rare Mendelian disorders contribute to CVD, for example, hypercholesterolemia (Jensen 2002) involves mutations in the low-density lipoprotein receptor gene or the apolipoprotein (apo) B gene. Hyperhomocystinuria is associated with mutations in the 5,10-methylenetetrahydrofolate reductase gene (Rees and Rodgers 1993; Cho, Hong et al. 2006). Hutchinson-Gilford progeria syndrome (DeBusk 1972) is caused by mutations in the lamin A/C gene. Tangier disease is related to mutations in the ATP binding cassette, subfamily A, member 1 gene (Brooks-Wilson, Marcil et al. 1999). An autosomal-dominant form of CAD is associated with the MADS box transcription enhancer factor 2, polypeptide A gene (Wang, Fan et al. 2003). Some variants of voltage gated sodium and potassium channels have also been implicated in CAD (Loussouarn, Baro et al. 2006; Priori and Napolitano 2006).

The most common forms of CVD are multifactorial resulting from many genes, each with a relatively small effect, working alone or in combination with modifier genes and environmental factors. The common disease, common variants hypothesis proposes that genetic variants present in many normal individuals contribute to overall CVD risk (Hirschhorn and Daly 2005).
Family history is a significant determinant for atherosclerosis, and genetic factors contribute to diabetes, hypertension, and low and high-density lipoprotein cholesterol levels. However, the contribution of family history is not fully explained by these known cardiac risk factors, suggesting that other variables also contribute to cardiovascular risk (Snowden, McNamara et al. 1982; Marenberg, Risch et al. 1994; Nasir, Michos et al. 2004).

Migration studies show that the environment can trigger variation in disease incidence between genetically similar populations. For example, Japanese have lower incidence of CAD than Americans, yet Japanese Americans have the same incidence as other Americans (Nichaman, Hamilton et al. 1975). Similarly, the Seven Countries study discovered varying plasma lipid levels and CAD in seven disparate populations unexplained by genetic differences (Keys, Menotti et al. 1986). Moreover, The Arizona Pima indigenous people had modest incidence of obesity and diabetes prior to adopting a Western lifestyle, now obesity and diabetes prevalence exceeds the rest of the country (Norman, Bogardus et al. 1995). Interestingly, the Pima Natives have a low incidence of CAD despite the high frequency of diabetes, demonstrating the interplay between genes and the environment.

## 2.2.1 Outcomes Predicted By Risk Score Methods

In the clinical setting, risk for developing various CVD outcomes is estimated using composite risk factor scores. In the Framingham study, risk formulas were developed for CHD, CHD death, CVD, CVD death, myocardial infarction, and stroke (Anderson, Odell et al. 1991). Cardiovascular death is the outcome predicted in the Systematic Coronary Risk Evaluation (SCORE) method. Consensus is required on
the ideal outcome for a risk score for making genetic studies comparable to research results. For example, raised cholesterol treatments and strategies to stop smoking are aimed at reducing the likelihood of a CHD event, whereas the aim of antihypertensive treatment is to prevent stroke and CHD. The outcome "any cardiovascular event", as recommended in the Joint British Societies guidelines (2005), encompasses all sequelae of cardiovascular disease: death, morbidity, disability, poor quality of life, loss of independence, and use of health services. This outcome may be favoured by healthcare professionals, but is difficult to define and use in comparing clinical studies.

2.2.2 Risk Factors and Scoring Methods

Interestingly, the past 20 years worth of genetic research on CVD has not identified predictive factors that can be used to estimate risk in the clinical setting. Similarly, although the environment is thought to play a significant role in the onset of CVD, smoking is the only environmental variable that has predictive value for this disease.

<table>
<thead>
<tr>
<th>Genetic or Heritable Risk Factors</th>
<th>% Heritability Estimate</th>
<th>Use in CVD Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial infarction</td>
<td>25-60</td>
<td>Yes</td>
</tr>
<tr>
<td>Stroke</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>40-60</td>
<td>Yes</td>
</tr>
<tr>
<td>HDL cholesterol</td>
<td>45-75</td>
<td>Yes</td>
</tr>
<tr>
<td>Total triglycerides</td>
<td>40-80</td>
<td>Yes</td>
</tr>
<tr>
<td>Body mass index</td>
<td>25-60</td>
<td>No</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>50-70</td>
<td>Yes</td>
</tr>
<tr>
<td>diastolic blood pressure</td>
<td>50-65</td>
<td>Yes</td>
</tr>
<tr>
<td>Lp(A) levels</td>
<td>90</td>
<td>No</td>
</tr>
<tr>
<td>Homocysteine levels</td>
<td>45</td>
<td>No</td>
</tr>
<tr>
<td>Type II diabetes</td>
<td>40-80</td>
<td>Yes</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>20-50</td>
<td>No</td>
</tr>
<tr>
<td>c-reactive protein</td>
<td>n/a</td>
<td>No</td>
</tr>
</tbody>
</table>
Cardiovascular disease risk score formulas have been developed for clinical use with various combinations of factors listed in Table 2.1. Sex is one of the simplest variables to measure and has the most predictive value for CVD. Consequently, separate formulas are often provided for men and women (Silberberg 1999). Age is also among the most predictive factors, however, risk can be underrepresented in younger populations due to the presence of rare monogenetic forms of CVD resulting in a bimodal distribution of disease (Ridker and Cook 2005). Hence, age may be a surrogate for atherosclerotic plaque formation (Grundy 2001).

Lipid measures have high heritability estimates (Table 2.1) and are often used in various CVD risk scores (Table 2.2). Risk scores that include HDL cholesterol improve sensitivity for identifying individuals at high risk for CHD (Jones, Walker et al. 2001). Serum triglyceride measurements are included in the Prospective Cardiovascular Munster study (PROCAM) (Assmann, Cullen et al. 2002) and the Framingham–D’Agostino methods (D’Agostino, Russell et al. 2000). Elevated triglycerides may be significant as an independent risk factor (Hokanson and Austin 1996; Patel, Barzi et al. 2004) or as comprising part of metabolic syndrome (Alberti, Zimmet et al. 2005; Shaw, Zimmet et al. 2005; Zimmet, Magliano et al. 2005; Zimmet, Alberti et al. 2005). In the Framingham–Wilson method, total cholesterol may be replaced by LDL cholesterol without impacting predictive value (Wilson,
D’Agostino et al. (1998). Despite its high estimated heritability (Table 2.1), lipoprotein(a) and association with CVD, this variable is not used in CVD risk estimations due to large racial variation and unstandardised measurement techniques.

Cigarette smoking is a powerful risk factor and doubles CHD mortality risk (Ball and Turner 1974), but the mechanism of action is unknown. Smoking may effect atheroma, thrombosis, or the heart by altering arrhythmias or pump failure. Increased coronary atheroma has been reported in smokers (Auerbach, Hammond et al. 1965), as well as increases in pro-thrombotic factors including platelet survival (Mustard and Murphy 1963) and aggregation, raised plasma concentrations of fibrinogen and coagulation factors (Meade, Chakrabarti et al. 1979).

Although stroke is more closely linked with hypertension than with smoking, a differential association with cigarette smoking occurred in the Framingham study, where male smokers had a threefold increase in the risk of developing a stroke, whereas no relation was found between smoking and the risk of stroke in females (Kannel 1976). In the Multiple Risk Factor Intervention Trial (Ockene and Shaten 1991; Neaton and Wentworth 1992), serum thiocyanate improved prediction of CHD events beyond other measures of cigarette smoking.

Diabetes more than doubles CVD risk (Watkins 2003). Consequently, most risk scoring methods incorporate a diabetes factor (Wilson, D’Agostino et al. 1998) (Conroy, Pyorala et al. 2003). Glycosylated haemoglobin is included in the UK Prospective Diabetes Study (UKPDS) risk score (Stevens, Kothari et al. 2001) and the Royal College of Physicians of Edinburgh method (Lewis, MacLeod et al. 2005), whereas in the Cardiff score (McEwan, Peters et al. 2006) duration of diabetes is
used. Both Framingham-Wilson and The New Zealand CVD risk-benefit prediction guide defines diabetes as someone taking insulin, oral hypoglycemics or a fasting blood glucose concentration >8.0 mmol/L (Wilson, D'Agostino et al. 1998; Jackson 2000).

Scoring methods including Framingham, SCORE, and DECODE tend to underestimate risk in diabetic populations (Coleman, Stevens et al. 2007). Guzder et al. observed better estimation of CHD risk using the UKPDS method compared with Framingham (Guzder, Gatling et al. 2005). In the Asia Pacific Cohort Studies Collaboration (Woodward, Barzi et al. 2006), authors improve risk prediction by incorporating diabetes as an independent risk factor compared with use of separate specific charts or classifying all patients with diabetes as high risk.

All these factors involve multiple genes except lipoprotein-a and gender. Blood pressure, HDL levels, and total cholesterol exhibit smooth population distributions characteristic of polygenic traits. Large genome scans for risk factors such as diabetes, obesity, and hypertension failed to reveal loci accounting for substantial population variation, suggesting the involvement of many genes. Thus, at least hundreds of genes are thought to be involved in CVD susceptibility, which may have additive, multiplicative or synergistic effects (Law and Wald 2002). For example, high cholesterol levels amplify the effects of hypertension. Also, smoking has a greater effect on CVD in males than females.

Multivariable risk scoring methods are incorporated in national guidelines (Jackson 2000; Pearson, Blair et al. 2002; 2005; van Wyk, van Wijk et al. 2006; Crooke 2007;
Graham, Atar et al. 2008) for clinicians to calculate an individual’s risk of a cardiovascular event within a specified period. Several multivariable formulas exist to predict CVD or CAD risk over a certain period of time including: New Zealand guidelines group, Framingham, Sheffield, UKPDS (Jackson, Lawes et al. 2005). Cardiovascular risk scoring methods are summarized below in Table 2.2.

**Table 2.2** Risk scoring methods from Beswick et al, 2008

<table>
<thead>
<tr>
<th>Risk scoring method</th>
<th>Recent guidelines</th>
<th>Outcome and prediction period</th>
<th>Risk Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dundee (Tunstall-Pedoe and Woodward 2006), Multiple logistic function; Equation, rank, risk disk (men)</td>
<td>CHD 5 years</td>
<td>Age SBP, Total cholesterol Smoking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 years MI or CHD death</td>
<td>Age Sex SBP, Total cholesterol, HDL cholesterol Smoking, Diabetes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI, CHD death, CHD, stroke, CVD, CVD death 4–12 years</td>
<td>Age Sex SBP, DBP, Total cholesterol, HDL cholesterol, Smoking, Diabetes, LVH</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stroke 10 years</td>
<td>Age Sex SBP, Antihypertensive medication Smoking, Diabetes, LVH, AF, CVD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHD, CHD 10 years</td>
<td>Age Sex SBP, DBP, Total cholesterol or LDL cholesterol, HDL cholesterol Smoking, Diabetes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHD 1–4 years</td>
<td>Age Sex SBP, Antihypertensive medication, Total cholesterol, HDL cholesterol Triglycerides Smoking, Diabetes, BMI, Menopausal status, Alcohol intake</td>
<td></td>
</tr>
<tr>
<td>FINRISK (Bhopal, Fischbacher et al. 2005), Logistic regression, Equation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Model/Method</td>
<td>Equation</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Framingham, (D'Agostino, Russell et al. 2000), Weibull accelerated failure time model, Regression equation, Point scoring Global coronary risk score (Kornitzer and Koyunco 2000), Not described, Point scoring</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHD</td>
<td>Age SBP, Total cholesterol, HDL cholesterol, Smoking</td>
<td></td>
</tr>
<tr>
<td>PRECARD (Thomsen, Davidsen et al. 2001)</td>
<td>Cox regression model, Computer program, Equation</td>
<td>Fatal/nonfatal MI 10 years</td>
<td>Sex SBP, Total cholesterol, HDL cholesterol, Smoking, Diabetes, BMI, Previous MI, Family history</td>
</tr>
<tr>
<td></td>
<td>PROCAM (Assmann, Cullen et al. 2002) Cox proportional hazards model, Point scoring</td>
<td>CHD 10 years</td>
<td>Age Men SBP, Total cholesterol, HDL cholesterol, Triglycerides Smoking, Diabetes Family history, Region</td>
</tr>
<tr>
<td></td>
<td>REGICOR Recalibration of Framingham Wilson for Spanish Population (Marrugat, D'Agostino et al. 2003), Equation</td>
<td>CHD death or nonfatal MI 10 years</td>
<td>Age Sex Hypertension, Total cholesterol, HDL cholesterol Smoking, Diabetes</td>
</tr>
<tr>
<td></td>
<td>RISCARD (Menotti, Lanti et al. 2002), Log-linear model incorporating Weibull hazard distribution, accelerated failure time model, Computer program, chart</td>
<td>CHD, CVD, cerebro-vascular 5, 10, 15 years</td>
<td>Age Sex Mean blood pressure, Non-HDL cholesterol (total cholesterol minus HDL cholesterol), HDL cholesterol Smoking, Diabetes, BMI, Heart rate</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Variables</td>
<td>Notes</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>SCORE</td>
<td>(Conroy, Pyorala et al. 2003), Weibull proportional hazards model, Charts, Internet resource</td>
<td>CVD death 10 years</td>
<td>Age Sex SBP, Total cholesterol, HDL cholesterol Smoking, High or low risk population</td>
</tr>
<tr>
<td>UKPDS</td>
<td>(Stevens, Kothari et al. 2001), Maximal likelihood estimation, Fully parametric model, Equation</td>
<td>Fatal MI/sudden death 20 years</td>
<td>Age Sex SBP, Total cholesterol, HDL cholesterol Smoking, HbA1c, Ethnicity</td>
</tr>
<tr>
<td>Westlund score</td>
<td>(Westlund and Nicolaysen 1972), Product of relative risks, Point scoring</td>
<td>MI 10 years</td>
<td>Sex SBP, Total cholesterol Smoking, Family history New methods</td>
</tr>
<tr>
<td>Cuore</td>
<td>(Ferrario, Chiodini et al. 2005), Cox proportional hazards model</td>
<td>CHD 10 years</td>
<td>Age SBP, Antihypertensive medication, Total cholesterol, HDL cholesterol, Smoking, Diabetes Family history</td>
</tr>
<tr>
<td>DECODE</td>
<td>(Balkau, Hu et al. 2004) Cox proportional hazards model</td>
<td>CVD death 5 years</td>
<td>Age Sex SBP, Total cholesterol Smoking, Diabetes, Fasting plasma glucose, Region</td>
</tr>
<tr>
<td>RISKARD</td>
<td>(Menotti, Lanti et al. 2005)</td>
<td>CVD (major coronary, cerebrovascular and peripheral artery disease events)</td>
<td>Age Sex SBP, Total cholesterol Smoking, Diabetes</td>
</tr>
<tr>
<td>Zhang</td>
<td>(Zhang, Attia et al. 2005) Cox proportional hazards model</td>
<td>CHD, stroke</td>
<td>Age SBP, DBP, Total cholesterol Smoking, BMI</td>
</tr>
<tr>
<td>Royal College of Physicians of Edinburgh Diabetes Research Group</td>
<td>2005</td>
<td>Macrovascular disease</td>
<td>Age SBP, DBP, Total cholesterol Smoking, HbA1c Urinary albumin</td>
</tr>
</tbody>
</table>

CHD, coronary heart disease; SBP, systolic blood pressure; MI, myocardial infarction; CVD, cardiovascular disease; DBP, diastolic blood pressure; LVH, left ventricular hypertrophy; WHO, World Health Organization; ISH, International Society of Hypertension; JBS, Joint British Societies; AF, atrial fibrillation; NCEP ATPIII, National Cholesterol Education Program, Adult Treatment Guidelines; LDL, low-density lipoprotein; BMI, body mass index; PROCAM, Prospective Cardiovascular Munster study; RCV-ap, rel Riesgo Cardiovascular en Atencion Primaria; SCORE, Systematic Coronary Risk Evaluation; UKPDS, UK Prospective Diabetes Study; HbA1c, glycosylated hemoglobin; DECODE, Diabetes Epidemiology Collaborative Analysis of Diagnostic Criteria in Europe study. 

2.3 Pterygia

The aetiology of pterygia has been of interest to researchers for many years (Enroth 1951; Hilgers 1960); however, its pathogenesis has yet to be determined. Solar ultraviolet radiation (UVR) appears to influence the pterygia phenotype (Saw and Tan 1999). Recent reviews suggest that UVR induces pterygia in a complex and multifactorial manner, involving immunological, viral and genetic factors including cytokines, growth factors, extracellular matrix modulators and matrix metalloproteinases (Di Girolamo, Chui et al. 2004; Chui, Di Girolamo et al. 2008; Bradley, Yang et al. 2009).

Hypotheses suggesting an inherited contribution to pterygia aetiology originated more than half a century ago (Enroth 1951; Hilgers 1960). Supporting data are derived from studies implicating familial aggregation of cases (Zhang 1987); ethnic (Luthra, Nemesure et al. 2001; Paula, Thorn et al. 2006) and gender (Tan, Lim et al. 2006) differences in pterygia prevalence as well as in recurrence rates (Kandavel, Kang et al. 2010); and clustering of pterygia in monozygotic twins (Faraldi and Gracis 1976). Additionally, people with pterygia are more likely to have family members with pterygia (Booth 1985). Genetic associations with pterygia have included loss of heterozygosity, and mutations in proto-oncogenes and tumour suppressor genes (Detorakis and Spandidos 2009). To date, no single gene has been found to account for the majority of disease.
2.3.1 Ultraviolet Radiation and Pterygia

At present, public health experts advocate that sun exposure should be guided by level of ambient UVR and skin type (Lucas, McMichael et al. 2008). Knowledge of an individual’s response to UVR has implications for many different ocular and non-ocular diseases and may lead to the development of targeted prevention and therapeutic strategies. Further work on linkage analysis and genome-wide association studies may be useful in further determining the genetics of this trait. Exposure to ultraviolet B wavelength (UVB) radiation (Saw and Tan 1999), as well as increasing age, male gender (Wong, Foster et al. 2001), outdoor work (Khoo, Saw et al. 1998) and rural residence (McCarty, Fu et al. 2000) are recognized risk factors for pterygia development. A combined increase in pterygium prevalence and in median UVAF provide further epidemiological evidence that UVR contributes to the aetiology of this disease. UVR contributes to the pathogenesis either directly or through the formation of radical oxygen species (Tsai, Cheng et al. 2005). Epidemiological associations with UVB and pterygium are also strengthened by findings at a molecular level demonstrating that interleukin 6 and 8 (Di Girolamo, Kumar et al. 2002), matrix metalloproteinase-1 (Di Girolamo, Coroneo et al. 2003), HB endothelial derived growth factor (HB-EGF) (Nolan, DiGirolamo et al. 2003) and vascular EGF(Di Girolamo, Wakefield et al. 2006) are significantly expressed in pterygia and induced by UVB. Other postulated risk factors for pterygia include G6PD deficiency (Peiretti, Mandas et al. 2010), COX expression (Karahan, Baspinar et al. 2008) and viruses. For example, human papilloma virus is another potential risk factor for pterygium (Gallagher, Giannoudis et al. 2001) and may trigger pterygium pathogenesis through the inactivation of p53 (Tsai, Chang et al. 2009).
Pterygia have been suggested as a biomarker for other ophthalmic and non-ophthalmic diseases strongly correlated with UVR. Exposure to UVR is strongly associated with eye disease, both acute (acute photokeratitis or flash burns) and chronic (pterygium, climatic droplet keratopathy, cataract, ocular surface squamous neoplasia, eyelid tumours, basal and squamous cell carcinoma and melanoma) (Oliva and Taylor 2005). Therefore, the importance of documenting increased risk of developing UVR-related disease, whether due to genetic or lifestyle reasons, has important public health ramifications.

2.4 Myopia

Nearsightedness, or myopia, is a vision condition in which close objects are seen clearly, but objects farther away appear blurred. Nearsightedness occurs if the eyeball is too long or the cornea, the clear front cover of the eye, has too much curvature. As a result, the light entering the eye isn’t focused correctly and distant objects look blurred. Currently, over 80% of young adults in East Asian cities have myopia (Wu, Seet et al. 2001; He, Zeng et al. 2004; Lin, Shih et al. 2004), in the US prevalence has doubled to 40% in 30 years (Vitale, Sperduto et al. 2009) suggesting strong environmental influences. Myopic trait heritability varies, ranging from high (Chen, Scurrah et al. 2007; Klein, Suktitipat et al. 2009) in urbanized settings to low (Biino, Palmas et al. 2005; Vitart, Bencic et al. 2010) in island populations further reflecting environmental contributions (Young, Leary et al. 1969; Morgan, Speakman et al. 1975; Lin, Hung et al. 1988; Visscher, Hill et al. 2008). Recently, more time spent outdoors has emerged as a protective factor against myopia (Mutti 2010) while
risk factors such as education level (Tay, Au Eong et al. 1992) and near work (Saw, Chan et al. 2008), coincide with more time indoors. Hence, variability in UV light exposure among these environmental factors is an underlying trait that may also play a role in myopia (Prepas 2008).

Myopia genetics reflect complex interactions and multifactorial pathways. Linkage studies have mapped almost 20 loci for high myopia, moderate myopia, and refraction as quantitative traits (Table 2.3) (Wojciechowski 2011). Association studies have identified variants in at least 25 genes putatively involved in ocular refraction (Table 2.3). However, replicating these results has remained a challenge and refractive error susceptibility alleles generally have low or modest effect sizes. Hence, it is likely that more undiscovered variants exist and that variants in several genes interact together, and with environmental factors, to mediate refraction and myopia (Wojciechowski 2011).

To date, most genes found to be associated with myopia are involved in connective tissue remodelling mechanisms. This group includes matrix metalloproteinases, growth factors, collagens, and proteoglycans (Wojciechowski 2011). Recently, two studies have implicated mitochondrial-mediated apoptosis as a novel mechanism for mediating refractive error (Andrew, Maniatis et al. 2008; Nakanishi, Yamada et al. 2009). Other pathways which may regulate refractive variation include Ras specific guanine nucleotide releasing factor 1 (Hysi, Young et al. 2010) and rod mediated visual signal transduction proteins (Lin, Wan et al. 2009; Solouki, Verhoeven et al. 2010) were identified in recent genome wide association studies. Muscarinic acetylcholine receptor genes have also been identified.
Table 2.3 Studies of myopia phenotypes with positive association results

<table>
<thead>
<tr>
<th>Study</th>
<th>Region</th>
<th>Myopia Phenotype</th>
<th>Gene</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Lam, Tam et al. 2003)</td>
<td>Hong Kong (Chinese)</td>
<td>high</td>
<td>TGF1</td>
<td>18p11.3</td>
</tr>
<tr>
<td>(Han, Leung et al. 2009)</td>
<td>Taiwan (Han)</td>
<td>high</td>
<td>HGF</td>
<td>7q21.1</td>
</tr>
<tr>
<td>(Lin, Wan et al. 2006)</td>
<td>Taiwan (Chinese)</td>
<td>high</td>
<td>TGFBI</td>
<td>19q13.1</td>
</tr>
<tr>
<td>(Wang, Chiang et al. 2006)</td>
<td>Taiwan (Chinese)</td>
<td>high</td>
<td>LUM</td>
<td>12q21.3-q22</td>
</tr>
<tr>
<td>(Mutti, Cooper et al. 2007)</td>
<td>US (62% Cauc.)</td>
<td>high</td>
<td>COL2A1</td>
<td>12q13.11</td>
</tr>
<tr>
<td>(Tang, Yip et al. 2007)</td>
<td>Hong Kong (Chinese)</td>
<td>high</td>
<td>MYOC</td>
<td>1q23-q24</td>
</tr>
<tr>
<td>(Inamori, Ota et al. 2007)</td>
<td>Japan (Japanese)</td>
<td>refraction</td>
<td>COL1A1</td>
<td>17q21.33</td>
</tr>
<tr>
<td>(Andrew, Maniatis et al. 2008)</td>
<td>UK (Cauc.)</td>
<td>high</td>
<td>MFN1/PSARL/CHRM1</td>
<td>3q26.33/3q27.1/3q26.3-q27</td>
</tr>
<tr>
<td>(Lin, Wan et al. 2009)</td>
<td>Taiwan (Han)</td>
<td>high</td>
<td>TGFBI</td>
<td>1q41</td>
</tr>
<tr>
<td>(Lin, Wan et al. 2009)</td>
<td>Taiwan (Han)</td>
<td>high</td>
<td>TGFBI</td>
<td>1q41</td>
</tr>
<tr>
<td>(Ng, Lam et al. 2009)</td>
<td>Hong Kong (Han)</td>
<td>mild</td>
<td>PAX6</td>
<td>11p13</td>
</tr>
<tr>
<td>(Yanovitch, Li et al. 2009)</td>
<td>US (Cauc.)</td>
<td>high</td>
<td>HGF</td>
<td>7q21.1</td>
</tr>
<tr>
<td>(Han, Yap et al. 2006)</td>
<td>China (Han)</td>
<td>high</td>
<td>PAX6</td>
<td>11p13</td>
</tr>
<tr>
<td>(Vatavuk, Skunca et al. 2009)</td>
<td>Croatia (Cauc.)</td>
<td>high</td>
<td>MYOC</td>
<td>1q23-q24</td>
</tr>
<tr>
<td>(Zha, Leung et al. 2009)</td>
<td>Hong Kong (Chinese)</td>
<td>high</td>
<td>TGFBI</td>
<td>19q13.1</td>
</tr>
<tr>
<td>(Metlapally, Li et al. 2009)</td>
<td>US/Wales (Cauc.)</td>
<td>high</td>
<td>COL2A1</td>
<td>12q13.11</td>
</tr>
<tr>
<td>(Khor, Grignani et al. 2009)</td>
<td>Singapore (Chinese)</td>
<td>progression</td>
<td>CMET</td>
<td>7q31</td>
</tr>
<tr>
<td>(Chen, Wang et al. 2009)</td>
<td>Taiwan (Han)</td>
<td>high</td>
<td>LUM</td>
<td>12q21.3-q22</td>
</tr>
<tr>
<td>(Hall, Gale et al. 2009)</td>
<td>UK (Cauc.)</td>
<td>all</td>
<td>MMP3/9</td>
<td>11q22.3/20q11.2 -q13.1</td>
</tr>
<tr>
<td>(Zayats, Yanovitch et al. 2009)</td>
<td>US/Wales (Cauc.)</td>
<td>high</td>
<td>MYOC</td>
<td>1q23-q24</td>
</tr>
<tr>
<td>(Liu, Lin et al. 2009)</td>
<td>Taiwan (N/A)</td>
<td>high</td>
<td>BMP2K</td>
<td>4q21.21</td>
</tr>
<tr>
<td>(Nishizaki, Ota et al. 2009)</td>
<td>Japan (Japanese)</td>
<td>high</td>
<td>UMODL1</td>
<td>21q22.3</td>
</tr>
<tr>
<td>(Veerappan, Pertile et al. 2010)</td>
<td>Australia (N/A)</td>
<td>low/mod.</td>
<td>HGF</td>
<td>7q21.1</td>
</tr>
</tbody>
</table>

Adapted from (Wojciechowski 2011)
2.5 Glaucoma

Glaucoma is a leading cause of visual impairment and irreversible blindness affecting 70 million people worldwide (Quigley 1996). Chronic glaucoma prevalence in the US population 40 years and older is 1.86% (Friedman, Jampel et al. 2006). With an aging population, this prevalence is projected to increase by 50% in 2020 (Friedman, Jampel et al. 2006). Australian prevalence is higher at 3.0% (Mutti, Mitchell et al. 2002). Glaucoma refers to a group of heterogeneous diseases causing optic nerve fiber damage (Roberts 1905). Aqueous is produced by the ciliary body, which enters the posterior chamber between the iris and the lens and passes through the pupil into the anterior chamber then to the trabecular meshwork to the anterior chamber angle (Figure 2.1).

Main resistance to aqueous outflow from the anterior chamber occurs at the trabecular meshwork and Schlem’s canal (Figure 2.1). Normal intraocular pressure (IOP) is 16 +/- 5 mm Hg. It varies during the day with highest readings in the early morning. Thus a single normal reading does not exclude glaucoma. Glaucomatous field loss mainly involves the central 30 degrees of vision. The earliest changes are enlargement of the blind spot, nasal scotomas (nasal step) followed by peripheral arcuate defects. Further loss increase constriction of the visual field that may result in tunnel vision.
Primary open angle glaucoma (POAG) is the main type of glaucoma in most populations. Since eye pressure plays a role in damaging optic nerve fibers, IOP $>22\text{mmHg}$ is a defining criterion for the POAG form of the disease (Quigley 1993). However, normal tension glaucoma (NTG) can occur with normal eye pressure (Werner 1996) and prevalence of this form is significantly higher in Asian than Caucasian populations (Shiose, Kitazawa et al. 1991; Fuse 2010). Also, glaucoma can manifest as juvenile or adult onset, yet further evidence of the disease complexity and multiple etiologies.

At least 14 genetic loci have been associated with glaucoma susceptibility; of these myocilin (GLC1A) located on chromosome 1q23, optineurin (GLC1E) on 10p15-p14, and WDR36 (GLC1G) on 5q22 play a causative role in the onset of this disease (Fuse 2010). Myocilin (MYOC) on 1q24.3 has been shown to account for 3% of adult
onset POAG (Stone, Fingert et al. 1997; Fingert, Heon et al. 1999; Alward 2000; Alward, Kwon et al. 2002). Optineurin mutations may play a role in both NTG and adult-onset primary open angle glaucoma by inducing death in retinal ganglion cells (Quigley 1993). At least 27 other loci related to glaucoma have been identified through association studies; interestingly, none correspond to eye phenotypes but instead play systemic roles and are implicated in other diseases (Fuse 2010).

In addition to IOP, other POAG risk factors include age, race, central corneal thickness (CCT), high myopia, diabetes (Bonovas, Peponis et al. 2004), cigarette smoking, (Bonovas, Filioussi et al. 2004) and positive family history (Rosenthal and Perkins 1985; Tielsch, Katz et al. 1994; Wolfs, Klaver et al. 1998; Weih, Nanjan et al. 2001) are potential risk factors with significant associations. The lifetime risk of glaucoma in relatives of patients with glaucoma is 22% while it is 2.3% in controls resulting in a risk ratio of 9.2 (Wolfs et al, 1998 see Fuse). Despite this increased risk, heritability studies of POAG conducted in twins report low estimates 0.13 (Teikari 1987). In addition to IOP, and CCT, other ocular metrics such as vertical cup-to-disc ratio (C:D) and anterior chamber depth (ACD) have also been studied in glaucoma (van Koolwijk, Despriet et al. 2007; Charlesworth, Kramer et al. 2010).
3.1 Overview

This section summarises methodology for pedigree/phenotypic/genotypic data necessary to conduct the heritability and linkage analyses described in this thesis. Phenotypic data was collected from the Norfolk Island Health Study (NIHS) in 2000 and the Norfolk Island Eye Study (NIES) in 2007. Genetic data used in this study was collected in 2000 for the NIHS by the Genomics Research Centre, Griffith University Gold Coast, Australia. Most of these methods have been published elsewhere (Bellis, Hughes et al. 2005; Bellis, Cox et al. 2008; Bellis, Cox et al. 2008; Mackey, Sherwin et al. 2011).

3.2 Ethics

3.2.1 NIHS

Griffith University, Human Research and Ethics Committee approved all NIHS research protocols prior to beginning the study. All patients gave informed written consent prior to participation.
3.2.2 NIES

Griffith University, Human Research and Ethics Committee in Queensland and the Human Research and Ethics Committee at the Royal Victorian Eye and Ear Hospital in Melbourne, approved the NIES. Consent was obtained to conduct the ophthalmic examination and link data with earlier cardiovascular and genetic research as well as ongoing genetic eye research. In addition, local community consultation with Norfolk Island hospital administration, local doctors, local optometrist and visiting ophthalmologists ensured that all concerns were met regarding possible long term impact of the study.

3.3 Subject Recruitment

3.3.1 NIHS Subject Recruitment

Subject recruitment took place in 2000 with permanent resident status as inclusion criteria. Subjects were recruited through a range of local media announcements including radio and newspaper. Phenotypic data and non-fasting venous blood was collected from 600 individuals. Phenotypic data was obtained through a comprehensive medical questionnaire that included family history. Blood specimens and questionnaires for each subject were uniquely barcoded. Blood samples and questionnaires were shipped to Australia. Queensland Medical Laboratories, Brisbane, Australia, conducted a full blood biochemistry analysis. Remaining blood was stored at -80°C until DNA extraction at the Genomics Research Centre, Griffith
University (Gold Coast), Australia. All phenotypic data, including blood biochemistry was entered into a Microsoft Excel spreadsheet.

3.3.2 NIES Subject Recruitment

This methods section was summarised from Mackey et al (Mackey, Sherwin et al. 2011). A pilot research team visited Norfolk Island to discuss the study with local community. Original NIHS participants were mailed invitations to participate in the NIES. Articles were presented in the local newspaper and radio. Word of mouth, particularly within families, resulted in many other participants. Off-island relatives, mainly students at Australian universities, were contacted by family or the Islander Facebook page and were seen in Brisbane, Melbourne or Sydney with TEST fieldwork.

Participants were invited to contact the eye research clinic at the Norfolk Island Hospital personally or by phone to make an appointment. Prior to an eye examination, a member of the research team would ensure that the administrative components of the protocol were appropriately addressed. An appointment list of attending participants was reviewed and personal information and study family numbers (if assigned) checked for accuracy. Information about special requests and tests (eg DNA collection) was highlighted on the appointment list, particularly if the subject had not previously attended the cardiovascular study. Various clinic rooms were set up at the Norfolk Island district hospital. Health care services for the
geographically isolated community of 2,000 are based at the 20-bed hospital enterprise, staffed by 2.5 full-time equivalent medical practitioners. In addition eye care is available at a private optometry practice. Australian medical students on remote attachments through the John Flynn Placement Program spend a term each year of their training on the island (Sherwin, Mitchell et al. 2008). The eye examination for the Norfolk Island Eye Study was conducted by a visiting team consisting of ophthalmologists, an ophthalmology trainee, a medical student, orthoptists, optometrists and a medical photographer.

3.4 Demographics and Phenotyping

3.4.1 NIHS

In the year 2000, approximately 1200 permanent residents resided on Norfolk Island. Of these individuals, approximately 900 were aged 18 years old or older at the time of recruitment. The NIHS clinic successfully recruited 600 (261 males, 339 females) permanent residents aged 18 years or older (2/3 of the permanent adult population).

Obesity markers consisted of arthropometric measures that included basal metabolic rate (calculated as kg/m2), hip and waist circumference, percentage body fat and weight. A trained member of the NIHS clinic recorded all measurements. Systolic and diastolic blood pressure measurements, markers of hypertension were recorded using an automatic electronic blood pressure monitor. Lastly, a standard blood screen was performed by Queensland Medical Laboratories, Brisbane and the following
measures of serum cholesterol (dislipidemia indicators) were obtained from this screen: HDLc (mmol/L), LDLc (mmol/L), TC (mmol/L), and TG (mmol/L).

3.4.2 NIES

Most participants appreciated the free comprehensive eye check with equipment that is not normally available on the island. Following completion of the examination session, each participant was given a detailed debriefing of the results of the various tests conducted, a summary of his/her ocular health status and provided with an opportunity to ask any questions. Any patients seen by ophthalmologists or optometrists previously had report letters sent to their eye care providers in Australia, New Zealand or on the island. Any newly diagnosed pathology was referred to the local optometrist (e.g. suspicious discs or pressure) for visual field testing (Carl Zeiss, Dublin, CA, U.S.A.). Baseline photographs were provided to the local hospital and optometry clinic for all participants’ future use. Receiving feedback on the status of their eye health, and undergoing a normal baseline examination and ocular imaging are important benefits that participants received.

3.4.2.1 Ophthalmic Examination

With the exception of a slit lamp and automated perimeter at the local optometrist’s office, most of the equipment required for the study was borrowed from the Center for Eye Research, Australia. The following equipment was brought to Norfolk Island: Nidek 3DX, a Kowa portable fundus camera, UV, IOLMaster, tonopen, pachymeter, UV auto fluorescence and an OCT (type on loan from Zeiss).
The examination protocol consisted of:

1) Visual acuity Log MAR at 3m (Precision Vision, LaSalle, IL, U.S.A)

2) Binocular vision function tests
   a) Cover test
   b) Four dioptre base-out prism test
   c) Ocular motility
   d) Stereopsis, Lang stereo card (Lang, Forch, Switzerland)

3) Eye dominance

4) Anterior chamber depth, keratometry, axial length. IOLMaster (Carl Zeiss Meditek, Dublin, CA, U.S.A)

5) Conjunctival UV auto-fluorescence photos

6) Autorefraction, pre and post-cycloplegia (Nidek, Gamagori, Japan).

7) Intraocular pressure- Tonopen handheld tonometer (Reichert, Tustin, CA. U.S.A).

8) Central Corneal thickness Pachmate ultrasonic pachymeter (DGH, Exton, USA)

9) Dilated stereoscopic optic disc photographs Nidek 3-DX (Nidek, Gamagori, Japan)

10) Disc-centred and macular-centred retinal photographs Genesis D hand held retinal camera (Kowa, Tokyo, Japan)

11) Ocular Coherence Tomography (OCT) Cirrus (Carl Zeiss Meditek, Dublin, CA, U.S.A) Macular thickness, RNFL thickness and optic head components were utilised on each patient using the (Three components were assessed; fast RNFL, fast macula and fast optic nerve head.

12) DNA collection, where participants in the original study had inadequate DNA available or for new participants.
The Nidek 3-DX stereo images were digitised and analysed with a stereo Z-screen (StereoGraphics, Beverly Hills, CA, USA) using high-resolution, 16-bit colour images, high-speed modulating stereoscopic panel and passive polarised eyewear. The cup and disc were delineated with custom software (developed by James Morgan; Cardiff, UK) and adjusted for magnification using refraction and Keratometry readings. Further analysis of optic disc stereo photos, retinal vasculature and conjunctival auto-fluorescence are in process. UV areas of fluorescence were measured using area calculation functions in image processing software (Adobe Photoshop). Method details are published in the parallel twin study (Mackey, Mackinnon et al. 2009). Main differences between the studies are the Nidek autorefractor and Cirrus OCT.

Chi-square tests were used to assess differences in categorical data and the t-test was used for continuous data. Mann-Whitney Test was used to assess differences in non-parametric data. All analyses were performed in using GraphPad Prism version 5.00 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com, and the results were reported significant for P< 0.05.
3.5 Laboratory Work

3.5.1 Blood Samples

Autoclaved or filter sterilised, deionised water from a Milli-Q Water Purification System was used in all aqueous solutions, dilutions and reactions. DNA was extracted from venous blood using analytical grade reagents supplied by Sigma-Aldrich (St Louis, CA), Astral Scientific (Gymea, Australia), and Gibco (Rockville, MD) unless otherwise stated. Proteinase K was obtained from Astral Scientific (Gymea, Australia). Quantum Scientific Pty Ltd (Milton, QLD) supplied all disposable laboratory equipment. Chemicals required for DNA extraction were obtained from Sigma Chemical Company (St. Louis, MO), Astral Scientific (Gymea, Australia), and Gibco (Rockville, MD). A NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Inc.) was used to quantify all genomic DNA and oligonucleotides.

3.5.2 DNA Extraction

While the majority of samples were collected and prepared in 2000 and 2001, DNA stocks were periodically replenished from original blood samples stored at the Genomics Research Centre. DNA was extracted from venous blood using a modified version of a standard salting-out procedure (Miller et al., 1988) involving DNA isolation then precipitation reaction.
3.5.3 DNA Isolation

The DNA isolation stage required red blood cell removal, followed by overnight incubation to ensure complete lymphocyte lysis and the release of DNA into solution. Blood samples were removed from -80°C storage, thawed and transferred to a 50mL centrifuge tube. The original storage vile was rinsed with sodium-potassium-magnesium (NKM) buffer to ensure the entire blood specimen was removed and transferred to the centrifuge tube. Each sample was brought to a final volume of 25mL with NKM buffer and shaken vigorously. Samples were centrifuged at 4°C, 25min at 4,800 rpm. The supernatant was discarded and each sample was brought to a final volume of 25mL with RSB buffer and centrifuged at 4°C for 15 minutes at 4000 rpm. The supernatant was discarded. The DNA pellet was resuspended in 1mL of RSB buffer, prior to the addition of 4mL of lympholysis solution and 250µL of Proteinase K. Samples were placed in a 37°C shaking water bath overnight. Once complete lymphocyte lysis was ensured, a salting out procedure was employed to precipitate DNA from the solution. Samples were removed from the water bath. Two millilitres of saturated sodium chloride (NaCl) solution was added to each specimen. Samples were mixed for 15 seconds by inversion, prior to centrifugation at 4°C for 15 minutes at 2,500rpm. This step ensures the removal of proteins from the lymphocytes. The DNA-containing supernatant was collected and transferred to a 15mL centrifuge tube. This centrifugation step was repeated to ensure maximum removal of proteins from solution. The supernatant was collected and transferred to a new 50mL centrifuge tube.
The volume of each sample was approximated and 2 volumes of room temperature absolute ethanol were added to each sample. Tubes were gently swirled to precipitate the DNA strands. DNA was removed using an inoculation loop and transferred into a new tube containing 2mL of Tris-EDTA (TE) buffer at pH 8. DNA was dissolved in TE by incubation at 37°C for 2 hours, mixing at regular intervals. DNA can be stored for infinite amounts of time at 4°C in TE buffer.

3.5.4 Ethanol Precipitation of DNA Stocks

Suspension of DNA in TE buffer permits long term storage of DNA stocks. Unfortunately, the presence of ethylenediaminetetraacetic acid (EDTA) is known to inhibit PCR reactions. Therefore, prior to experimentation, especially those involving PCR the dissolved DNA requires precipitation with ethanol to remove the EDTA and/or any residual tri-phosphates and suspension in sterile water.

To perform an ethanol precipitation a 100µL aliquot of DNA suspended in TE buffer is added to a 1.5mL centrifuge tube containing 200µL of chilled absolute ethanol and 10µL of sodium acetate solution (pH 5.5). The solution is mixed, frozen with liquid nitrogen and centrifuged at 4°C for 15 minutes at 10,000rpm. The supernatant was discarded and the DNA pellet dried using a savant speed vacuum. The pellet was resuspended in 100µL of sterile Milli-Q water and incubated in a 37°C oven for 24 hours.

Double-stranded DNA was quantified using a NanoDrop spectrophotometer (NanoDrop Technologies, Inc.), which uses fiber optics and surface tension to
accurately measure DNA concentration and quality (purity). Only 0.5-2.0 µl of sample was required for quantitation.

All DNA was diluted to a working concentration of µg/mL. DNA purity was measured by the ratio of absorbance at 260nm and 280nm. Pure double-stranded DNA has a purity of 1.8; other values indicate contamination by proteins, phenols or other nucleic acids.

3.6 Genome-wide Linkage Scan

A genome-wide linkage scan of microsatellite (STR) markers were used to identify genomic regions involved in the predisposition of four common disorders: cardiovascular disease, pterygia, glaucoma, and myopia in the Norfolk Island extended pedigree. DNA from all NIHS subjects (N=600) was genotyped at the Australian Genome Research Facility (AGRF), Melbourne, Australia using the Applied Biosystems PRISM Human Linkage Mapping Set version 2.5 (medium density) [MD10-LMSV2.5]. Markers included in this mapping set were selected according to chromosomal location and heterozygosity from the Généthon human linkage map (Dib et al., 1996; Gyapay et al., 1994; Weissenbach et al., 1992). In total, the linkage map consisted of 400 highly polymorphic STR markers with an expected average heterozygosity of 0.79 and an average distance between adjacent markers of 9.2 centi Morgans (cM) (maximum gap 26.11cM).
Markers were individually amplified by PCR. All primers were labelled with the fluorescent dyes FAM, HEX and NED (Applied Biosystems LMSV2.5). PCR reactions were performed in 386-well plates in a final volume of 6uL. Each reaction contained 30ng of genomic DNA, 1 x PCR buffer, 0.5pmol of each primer, 0.5 units (U) of AmpliTaq Gold, 2.5mM MgCl2, and 0.25mM of dATP, dGTP, dCTP and dTTP. PCR amplification was performed in a PTC-225 DNA Engine Tetra (MJ Research inc, Waltham, MA, USA) using the following cycling parameters: initial denaturation at 94°C for 10 minutes; 30 cycles of denaturation, annealing and extension at 94°C/15 seconds (s), 55°C/15s and 72°C for 30s, respectively; and a final extension at 72°C for 5 minutes.

After PCR, amplicons were denatured using a formamide loading buffer. Where possible products from the same DNA sample were multiplexed (10-20 products per lane) with an internal size standard. Samples were resolved by capillary electrophoresis at 80-100V on a 6% polyacridamide gel using an Applied Biosystems 3730 Genetic Analyser. In total, 28 multiplexed markers were used to genotype markers. Quality control within and between plates was ensured by the inclusion of internal positive controls. Raw electrophoresis data was transferred to an off-line computer. Tracking for each sample lane was assigned using Genescan version 3.1 (Applied Biosystems) software. Each gel batch was manually screened prior to analysis. Minor variations were corrected using a standard curve generated using the Local Southern Method (Genescan version 3.1 software) and size standard patterns for each sample lane of any one gel were manually corrected.
Raw electrophoretogram data was imported into Genotyper version 2.1 software (Applied Biosystems) for interpretation and genotype assignment. For each STR marker, peaks in the electrophoretogram were sized against the standard curve and designated as alleles. All peaks were greater than 10 fluorescent units in height. Stutter peaks, which can arise as a result of phenomena such as Taq polymerase slippage and low signals in relation to main peaks, were filtered out. Data was initially screened for typing errors using PedManager. (http://www-genome.wi.mit.edu/)

3.7 The Norfolk Genealogy

Genealogists from Norfolk Island have maintained records of individuals who contributed to the present population. This information was entered into the genealogy program Brother’s Keeper 6.0 (Rockford, MI USA). At the time of this study the complete Norfolk genealogy encompassed 6,379 individuals and is highly complex, containing multiple inbreeding loops particularly in the early generations (Bellis, Hughes et al. 2005). This initial version of the pedigree spanned 11 generations and comprised 2,185 families. 62.8% (N=377) of subjects possessed ancestral lineages either directly or through marriage to the Norfolk genealogy and fall within the most recent five pedigree generations. The remaining 37.2% (N=223) of subjects likely represent new immigrants settling Norfolk Island as they hold permanent resident status. These individuals did not possess ancestral heritage to the founders or links via marriage at the time of sampling. Given their permanent resident status they may be considered potential new founders for future investigations, particularly longitudinal studies.
The complete 6,379 member pedigree compiled from genealogical and participant questionnaires was validated using the Pedigree Relationship Statistical Test (PREST) (McPeek and Sun 2000) prior to all genetic analyses (Bellis, Cox et al. 2008). The program PREST was used to detect mis-specified relationships in a pedigree by determining whether the observed pattern of identity-by-descent (IBD) allele sharing fits the expected pattern of IBD allele sharing given the degree of relationship between relative pairs. Pedigree errors such as mis-specified paternity were resolved by reconfiguring the pedigree structure. Re-assignments were confirmed by re-running the pedigree through PREST. This produced a final inferred pedigree structure of 6,537 (Bellis, Cox et al. 2008). The number of pedigree members is inflated in comparison to the original structure. The additional individuals were imputed into the pedigree in cases where only one parent was specified to ensure the pedigree conformed to standard linkage format.

The complete pedigree structure is computationally demanding for IBD and quantitative trait locus (QTL) estimation, therefore pedigree trimming was required. The program PEDTRIM, included in the PEDSYS database suite (Dyke 1996) was employed to partition the genealogy and remove uninformative individuals. This program removes individuals if their phenotypic or genotypic data is missing or incomplete, or if they have no offspring or no parents and a single offspring (Dyke 1996). PEDTRIM produced a pedigree structure of 1,078 individuals, which included 377 ascertained subjects and one inbreeding loop. The 1,078-member pedigree structure was employed in all linkage investigations described in this thesis.
The number and type of pairwise relationships in the 1,078 member pedigree were calculated using SOLAR (Sequential Oligogenic Linkage Analysis Routines) version 4.0.1 and are detailed in Table 3.1 (Bellis, Cox et al. 2008). Using the program SOLAR version 4.0.1, the mean inbreeding coefficient was determined to be 0.0044, with a maximum observed inbreeding coefficient of 0.0684, reflecting the presence of consanguineous unions in the pedigree (Bellis et al., 2007). Interestingly, the level of kinship (the pairwise coefficient of relationship) revealed the majority of sampled individuals were related by less than third-degree on average (=0.125) (Bellis et al., 2007). This number is likely a reflection of the number of newly married-in individuals (founders) (N=124) amongst the ascertained subjects and also the sampling of individuals in the genealogy.

Statistical analyses conducted at Griffith University were performed on a Sun Solaris High Performance Computing (HPC) cluster. Griffith's Sun Solaris HPC cluster consisted of 1 head node and 8 computational nodes. Each computational node was a SunFire v880 with 8x Ultra SPARC-III 900MHz CPUs, 8GB of memory and 144GB of internal disk.

Table 3.1 Number of relative pairs within the Norfolk Island linkage pedigree (Bellis, Cox et al. 2008)

<table>
<thead>
<tr>
<th>Pairwise Relationship(s)</th>
<th>Trimmed Pedigree (N)</th>
<th>Ascertained Subjects (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedigree members</td>
<td>1 078</td>
<td>377</td>
</tr>
<tr>
<td>Founders</td>
<td>587</td>
<td>124</td>
</tr>
<tr>
<td>Relative pairs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------</td>
<td>---</td>
</tr>
<tr>
<td>Non-founders</td>
<td>491</td>
<td>253</td>
</tr>
<tr>
<td>Parent-offspring</td>
<td>982</td>
<td>142</td>
</tr>
<tr>
<td>Siblings</td>
<td>270</td>
<td>80</td>
</tr>
<tr>
<td>Grandparent-grandchild</td>
<td>908</td>
<td>19</td>
</tr>
<tr>
<td>Avuncular</td>
<td>612</td>
<td>120</td>
</tr>
<tr>
<td>Half Siblings</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>3rd degree</td>
<td>1677</td>
<td>283</td>
</tr>
<tr>
<td>4th degree</td>
<td>1317</td>
<td>315</td>
</tr>
<tr>
<td>5th degree</td>
<td>952</td>
<td>404</td>
</tr>
<tr>
<td>6th degree</td>
<td>648</td>
<td>372</td>
</tr>
<tr>
<td>7th degree</td>
<td>286</td>
<td>239</td>
</tr>
<tr>
<td>8th degree</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Unrelated</td>
<td>225 349</td>
<td>27 670</td>
</tr>
</tbody>
</table>

3.7.1 Pedigree Cleaning

The confirmation of relationships in a pedigree and the elimination of typing errors is a vital step to obtain accurate estimates of association or linkage. Although the AGRF employs pedigree-checking procedures to detect typing errors, the Norfolk pedigree required more comprehensive data cleaning strategies due to its size and complexity and as the structure had not been validated at the time of the genome wide scan (GWS). While pedigree verification and trimming was specific to the GWS, the
methods used to eliminate typing errors and construct of IBD matrices are applicable to both genome-wide and fine mapping of STR markers.

Typing errors were resolved using the program Simwalk2 (Sobel, Papp et al. 2002) and the PEDSYS (Dyke 1996) program INFER. The program PREST (McPeek and Sun 2000) was used to detect wrongly specified relationships and confirm relationship re-assignments according to the pattern of IBD allele sharing in the complete pedigree structure. The program Loki (Heath, Snow et al. 1997) was used to compute IBD and multipoint identity-by-descent (MIBD) matrices for each relative pair, which are required for two-point and multipoint linkage analysis, respectively. The programs PREST, Simwalk2 and Loki require the order and position of each marker along a chromosome to be specified. Positional information for all autosomal markers was obtained online from Marshfield Centre for Medical Genetics (http://research.marshfieldclinic.org/genetics/MarkerSearch/markersearch_index.asp).

All analyses were conducted using sex-averaged marker maps in cM.

The validated 6,537-member pedigree was computationally demanding due to size, pedigree sampling and the presence of multiple inbreeding loops, hence pedigree partitioning was required. The program PEDTRIM was used to produce a trimmed 1,078-member pedigree, which included 377 ascertained subjects and one inbreeding loop. Calculations of the inherent of this trimmed pedigree indicate that a QTL accounting for approximately 30% of the variation would be detected with 50% power (Bellis, Cox et al. 2008).
3.8 Data Screening

To provide an overview of the population demographics, descriptive statistics for the entire sample (N=600) and pedigree subset (N=377) were generated using SPSS (versions 14.0 and 17.0). For continuous variables mean values were calculated and the results presented followed by standard deviation and number of individuals included in the analysis. Continuous measurements were screened for the presence of outliers and conformance to normality. Measurements greater than or equal to 4 standard deviations from the mean were considered extreme outliers and were either blanked or in the case of data entry errors, corrected prior to analysis. Where possible, traits whose distribution deviated from normality were transformed prior to analysis. Parametric and non-parametric correlations, chi-square analysis, student’s t-tests, logistic regression and one-way analysis of variance (ANOVA) were all undertaken in SPSS (versions 14.0 and 17.0).

3.9 Quantitative Genetic Analysis

Quantitative genetics is concerned with complex (multifactorial) traits that display continuous variation. The variation in the trait is the result of multiple genetic and non-genetic (environmental) factors and their interactions. Although quantitative genetics is primarily concerned with continuous or metric traits (also known as
quantitative traits), the analysis of categorical and dichotomous traits is also possible by this method. Categorical traits must be either re-coded as a dichotomous variable or possess enough categories for the trait to be considered quantitative. Dichotomous traits can be analysed by assuming an underlying genetic liability (Duggirala, Williams et al. 1997) (Hartl and Clark, 2007). Assuming this model, the presence of a dichotomous trait is determined by an underlying risk (liability), which follows a continuous distribution. An individual expresses the trait if the liability exceeds a certain threshold.

Quantitative genetic theory is concerned with partitioning the total phenotypic variance of a trait of interest into components (Amos 1994). In the simplest variance component (VC) model, the total phenotypic variance is $\sigma^2_p$ the sum of the genotypic variance ($\sigma^2_G$) and the environmental variance ($\sigma^2_E$) (Hartl and Clark 2007):

$$\sigma^2_p = \sigma^2_G + \sigma^2_E$$

These variance components may be further partitioned, depending on the model. For instance, the genetic variance can be partitioned into components surmising the additive genetic variance attributable to individual alleles and dominance variance, the combination of alleles into genotypes (Hartl and Clark 2007). It is theorised that partitioning environmental components should facilitate the characterisation of the genetic component influencing the trait.

Variance component methods have been used extensively in the analysis of complex polygenic phenotypes and have successfully localised QTLs influencing a diverse
range of phenotypes. VC based programs such as SOLAR, enable powerful penetrance model-free linkage analysis of quantitative (and qualitative) traits in pedigrees of arbitrary size and complexity (Almasy and Blangero, 1998; Blangero et al., 2000; Williams-Blangero et al., 1999). SOLAR is particularly versatile, as it enables modelling of genotype by environment interactions, epistasis and pleiotropy, as well as single point and multipoint linkage, association, multivariate, mitochondrial and oligogenic analyses (Almasy et al., 1997; Blangero et al., 2001; Comuzzie et al., 1997; Duggirala et al., 1997; Williams et al., 1997). Due to the size and complexity of the Norfolk genealogy, VC linkage methods are well suited to approximating the genomic location of susceptibility loci influencing selected phenotypes. The following sections describe the application of VC methods in heritability estimation and linkage analysis with particular reference to the SOLAR software package.

3.9.1 Heritability Estimation

Heritability is a measure of the genetic component influencing a trait of interest. Heritability is measured on a scale ranging from 0 to 1. A value of 0 indicates the phenotype is completely controlled by non-genetic or environmental factors. A value of 1 indicates the trait is completely controlled by genetic factors. Generally, heritability estimates fall within these 2 extremes, indicating the trait is influenced by a combination of genetic and non-genetic factors. For example, a heritability value of 0.62 indicates 62% of the phenotypic variation is attributable to genetic factors and
38% to non-genetic factors. Heritability estimates provide statistical evidence of the involvement of a genetic component to the phenotype, and thus support deeper investigations by linkage analysis. Gene mapping of a trait purely controlled by non-genetic factors would be unproductive.

There are various methods of estimating inheritance. Quantitative genetics distinguishes two measures of the heritable component of a phenotype: the narrow-sense heritability ($h^2$) and the broad-sense heritability ($H^2$). Both measures are expressed as ratios of the total phenotypic variance. While both measures incorporate the additive genetic variance, the broad sense heritability incorporates two additional variance components, dominance variance and epistatic variance (Hartl and Clark, 2007). This section will specifically focus on modelling inheritance using generalised VC methods in the absence of dominance and/or epistatic effects.

In its simplest form, the narrow sense heritability for a quantitative trait is defined as the ratio of the additive genetic variance ($\sigma^2_G$) to the total phenotypic variance ($\sigma^2_P$) (Hartl and Clark 2007):

$$h^2 = \frac{\sigma^2_G}{\sigma^2_P}$$

The heritability model therefore reflects the proportion of the phenotypic variance explained by genetic factors (Göring, Terwilliger et al. 2001).
For a polygenic model, the estimated additive heritability is expressed as the sum of the genetic effects attributable to a QTL ($\sigma^2_q$) and residual (additive) polygenic effects ($\sigma^2_p$) to the total phenotypic variance ($\sigma^2_P$). The additive heritability ($h^2$) of a polygenic model expressed as a function of these variance components given by (Göring, Terwilliger et al. 2001):

$$h^2 = \frac{\sigma^2_q + \sigma^2_p}{\sigma^2_P}$$

This model assumes a major trait locus (QTL) and an infinite number of genetic factors with small additive effects contributing to the overall phenotypic variance. The residual heritability is attributable to the effects of non-genetic factors.

The genetic component of the phenotype of interest is estimated using samples of related individuals, which may include sibling-pairs (sib pairs), monozygotic (MZ) and dizygotic (DZ) twins, nuclear families or extended pedigrees (Williams-Blangero and Blangero 2006). Heritability is estimated by comparing how the phenotype co-varies in relative pairs compared to that which is expected given the degree of relationship between the two individuals. This is the genetic covariance (or correlation). The degree to which the observed covariance fits that which is expected assuming the phenotype is purely attributable to genetic factors provides an overall estimate of the genetic component (Hartl and Clark 2007). In addition to estimating the additive heritability, programs such as SOLAR provide p-values calculated by
likelihood ratio tests, where the likelihood of the model is estimated and compared to the likelihood of the model where the heritability is constrained to zero.

Valid estimates of trait heritability are dependent on accurate phenotyping, adequate sampling of individuals and for families, accurate knowledge of the pedigree structure. In the case of continuous phenotypes, the trait is assumed to have an underlying multivariate normal distribution. Deviations from normality require correction by transformation and/or the removal of outliers. Where traits cannot be normalised, some statistical programs offer alternative functions (eg tdist and lod adjustment functions in SOLAR) to correct the test statistic for the effects of the statistical violation.

In addition to the variance components, the covariate effects of dichotomous and/or continuous variables of interest (and their interactions) can be included in the polygenic model (Duggirala, Williams et al. 1997). This allows for the effects of the variance components and covariates on the additive heritability to be simultaneously estimated by maximum likelihood techniques. This is a mixed polygenic model. The most commonly screened covariates in genetic models are age, age-squared, sex and age-by-sex interaction terms and/or environmental factors. The inclusion or exclusion of covariates in the heritability model is determined by a statistical threshold (p-value) for significance, typically p<0.05.

Heritability screening of dichotomous and continuous phenotypes in the Norfolk Island pedigree utilised maximum likelihood VC methods implemented by the statistical program SOLAR version 4.0.1.
3.9.2 Variance Component Linkage Methods

Once the additive heritability is estimated and statistical evidence supporting a genetic component for the trait of interest is provided, linkage may be tested at defined chromosomal locations throughout the genome to approximate a QTL. In the simplest VC linkage model, the total phenotypic variance of the quantitative trait of interest is expressed as the sum of the QTL specific effects, residual additive polygenic effects and individual-specific random environmental effects (Williams and Blangero, 1999). This model is extended to related individuals by observing how the phenotypes of relative pairs co-vary given the degree of relationship between the individuals and the proportions of genes they share IBD at a specific marker locus (Amos, 1994). This is the genetic covariance between the relative pairs, or in the case of a pedigree, the covariance matrix. By observing the trait covariance between different classes of relatives at specific locations throughout the genome, evidence for a QTL can be obtained.

The covariance matrix for a pedigree ($\Omega$) is estimated as:

$$\Omega = \Pi \sigma_q^2 + 2\Phi \sigma_p^2 + \sigma_e^2$$
Where $\Pi$ is a matrix with elements $(\pi_{ij})$ providing the estimated proportion of genes individuals $i$ and $j$ share IBD at a given chromosomal position ($q$); $(\sigma^2_q)$ is the variance component corresponding to the additive genetic effects due to the locus (the QTL effect size); $\Phi$ is the kinship coefficient matrix; $(\sigma^2_p)$ is the variance due to additive effects of genes elsewhere in the genome; $I$ is the identity matrix; and $(\sigma^2_e)$ is the variance due to environmental effects specific to the individual (Almasy, Williams et al. 1999). The IBD matrix is used to conduct two-point (single-point) linkage analysis, where linkage is tested between each marker and the phenotype independently. For multipoint linkage analysis the IBD matrix is a multipoint IBD (MIBD) matrix, which utilises marker information across a single chromosome, generally at 1cM increments.

Therefore, the VC method can be used in pedigrees of arbitrary size and complexity by specifying the covariance matrix for relative pairs. The covariance matrix can be surmised as a function of the relationship between the genes shared IBD at a defined chromosomal position and the QTL effect size, kinship and the additive genetic variance, and the identity matrix and the environmental variance. If the variance component is significantly greater than zero, there is evidence for a QTL at a given chromosomal location influencing the phenotype.

### 3.13.3 The LOD Score

To estimate linkage, the VC method measures how the phenotype co-varies in relative pairs, given the degree of relationship between the two individuals and
observed IBD sharing at a defined chromosomal location (Göring, Terwilliger et al. 2001). The presence of a QTL at a particular chromosomal location (linkage) is tested by comparing the likelihood of a linkage model to a basic polygenic model with no linkage at each marker.

This maximum likelihood-ratio test is given in the form of a logarithm of the odds (LOD) score \((Z)\), which is given by (Göring, Terwilliger et al. 2001);

\[
Z = \log_{10} \frac{\max L (\sigma_q^2, \sigma_p^2, \sigma_e^2)}{\max L (\sigma_q^2 = 0, \sigma_p^2, \sigma_e^2)}
\]

Where, \( L(\cdot) \) represents the likelihood. Under the null hypothesis of no linkage the QTL effect size is constrained to zero \( (\sigma_q^2 = 0) \). Under the alternate hypothesis of linkage the difference of the \( \log_{10} \) likelihoods produces a LOD score. If the variance component is significantly greater than zero at a tested chromosomal location, there is evidence for the presence of a QTL.

This LOD score is equivalent to the LOD scores of parametric linkage analysis (Blangero et al., 2001). Like parametric LOD scores, under the null hypothesis of no linkage, the VC LOD score is distributed as an equal mixture of a chi-square \( (\chi^2) \) random variable a point mass of zero and a degree of freedom of one (Göring et al., 2001). As a result, point-wise p-values can be estimated for each LOD score value (Nyholt, 2000). Alternatively, p-values can be estimated for each LOD score by performing simulations in SOLAR for continuous phenotypes. Linkage analysis of
dichotomous and continuous quantitative (continuous) phenotypes in the Norfolk Island pedigree utilised maximum likelihood VC methods implemented by the statistical program SOLAR version 4.0.1.
Chapter 4: Heritability of the Framingham Risk Score in the Norfolk Island Isolate

4.1 Overview

A main objective of the Framingham Heart Study was to identify a single cause for coronary heart disease. It became apparent that the disease process was more complex and multivariable models were necessary to quantify the contributions of multiple risk factors. The Framingham Study produced several multivariable models to determine the function that best predicted the likelihood of an event based on readily available and measurable risk factors. Framingham risk score (FRS) models are able to predict 10-year risk of a cardiac event with respectable accuracy (Brindle, Emberson et al. 2003) (Havranek 2007). The purpose of this study was to determine whether the FRS could be used for gene mapping in linkage analyses as a new approach to CVD gene discovery. The rationale was that since this composite score is reliable at predicting CVD outcomes, it might also be useful for identifying genetic contributions to this multigene disease. The research plan was to calculate the FRS for individuals within the Norfolk Island pedigree using cardiovascular disease biomarkers collected during the Norfolk Island Health Study (NIHS), estimate FRS heritability, and then conduct linkage mapping of this trait to identify putative cardiovascular disease loci.
4.2 Background

4.2.1 Framingham Heart Study

The Framingham Heart Study began in 1948 with the recruitment of 5209 residents of Framingham, Massachusetts, USA. Subsequently in 1971–75 a second generation of 5124 subjects (original cohort offspring and spouses) were recruited and in 2002–2005 a third generation of 4095 subjects (all offspring of the second generation) were recruited. First-generation subjects have been examined every two years for 60 years and second-generation subjects have been examined every four years. At each examination, subjects complete questionnaires, undergo a medical examination, an electrocardiogram, and have blood drawn. Methods for recruitment and examinations have been previously published (Anderson, Odell et al. 1991; Wilson, D'Agostino et al. 1998).

4.2.2 Framingham Risk Formula

The original Framingham formula is mathematically complex and has been converted into several risk tables used in clinical settings across the UK, Europe, and New Zealand. This formula uses eight weighed risk factors (age, sex, systolic or diastolic blood pressure, total and HDL cholesterol, and the presence or absence of cigarette smoking, and diabetes mellitus to calculate risk. Left ventricular hypertrophy was also used in the original Framingham formula but was dropped from more recent
versions such as the Framingham-Wilson method (Wilson, D'Agostino et al. 1998; Watkins 2003).

The first models were based on logistic regression and discriminant function analysis (Truett, Cornfield et al. 1967). As more data were accumulated due to longer follow up, survival analysis techniques were used to update models (Kannel 1976; Anderson, Odell et al. 1991; Wilson, D'Agostino et al. 1998). Multivariable models have been produced for specific events including stroke, peripheral vascular disease and congestive heart failure as well as subsequent events by quantifying the effects of risk factors on repeat events in persons who have a history of coronary disease (D'Agostino, Russell et al. 2000).

4.2.3 Framingham Model External Validity

Framingham models were developed in primarily Caucasian populations without external validity in populations that differ ethnically, racially, according to risk factor prevalence or event incidence. In 1999, the National Heart Lung and Blood Institute convened a workshop to assess the validity of the Framingham coronary heart disease functions in ethnically and racially diverse populations including whites, blacks, Native Americans, Japanese American men and Hispanic men. The Framingham models were compared with the performance of models developed on each population. The workshop concluded that both groups of models assigned similar weights to each risk factor, identified CVD, and calibrated similarly with predicted/actual event rates in whites and blacks. The workshop also concluded that
Framingham models could be applied to other ethnic groups after calibration (D'Agostino, Grundy et al. 2001).

Recently, Beswick and Brindle reviewed the literature to determine the predictive value of Framingham scores across populations. Figure 4.1 below summarizes differences in predicted CVD relative risk scores across various populations.

**Figure 4.1** CVD risk prediction to observed risk ratios by population (Taken from Beswick, 2006 (Beswick and Brindle 2006))

Overall, Framingham score predictions were relatively accurate for the New Zealand cohort (10% Maori, 5% Pacific Islanders, 85% European) in both men and women (Milne, Gamble et al. 2003). However, in the Australian aboriginal study, a
population with a high CVD rate, risk was underestimated in all age groups (Chang, Hawes et al. 2005). Similarly, underestimation was observed in UK diabetics (McEwan, Williams et al. 2004). Conversely, in a low-risk Chinese cohort, Zhang et al. found overestimation at all levels of cardiovascular risk (Zhang, Attia et al. 2005).

In the British Regional Heart Study prediction accuracy reflected racial variation in CHD risk where, overprediction by the Framingham–Anderson method was apparent in all regions. This overprediction was greater in the South, Midlands, and Wales, where CHD rates are lower than in Scotland and northern England (Brindle, Emberson et al. 2003). Hence, there is a tendency to overestimate risk in low-risk populations and underestimate in high-risk groups. A confounding aspect to this cross-cultural comparison is that different CVD related outcomes are being studied.

4.2.4 Calibrating Risk Scoring Methods

Due to population related differences in predictive power of these formulas, risk scoring methods require calibration if they are to be transferable between ethnic groups (D'Agostino, Grundy et al. 2001). One strategy for calibrating risk-scoring methods is to generate additional risk equations for each population (Ji, Chang et al. 2005). Another approach is to adjust for background risk based on regional health and mortality data. For example, the WHO is currently developing CVD risk scoring charts applicable to each of the global regions using data from the global burden of disease study (Beswick and Brindle 2006). Since age is a major CVD risk factor, a simplified approach may be to modify age based on population risk; Aarabi and Jackson suggest adding 10 years to the age of South Asian patients when calculating
Framingham risk based on an overview of previous prospective studies (Aarabi and Jackson 2005).

Another factor impacting risk score ability to predict CVD is antihypertensive treatment, which obscures the hereditary contribution to blood pressure variability (Cui and Sheffield 2003). In the US National Cholesterol Education Program (NCEP) Adult Treatment Panel III (ATPIII) modification of the Framingham–Wilson score and the Framingham–D’Agostino method, antihypertensive treatment is included. The role of antihypertensive treatment in CVD risk reduction is unclear since high blood pressure associated risk is not entirely reversed by antihypertensive drugs (Blacher, Evans et al. 2004). Development of a risk score in patients treated for hypertension such as that derived in the INdividual DataANalysis of Antihypertensive intervention (INDANA) pooling project (Pocock, McCormack et al. 2001) represents an alternative approach.

Differences between populations may be explained by additional risk factors. Candidate risk factors include thrombotic and inflammatory activity. In the WHOMONICA project, fibrinogen, von Willebrand factor, and D-dimer were associated with CHD differences between European countries (Davies, Hobbs et al. 2001). Fibrinogen may also be a marker of differences between social classes (Markowe, Marmot et al. 1985). In a study of ethnic differences of CHD risk (Matthews, Sowers et al. 2005), predictive factors additional to those in Framingham-based assessment were educational attainment, geographic location, waist circumference, lipoprotein(a), and c-reactive protein (CRP). The INTERHEART study (Yusuf, Hawken et al. 2004), however, has demonstrated very similar relative risks associated with a range of risk factors among very diverse study populations.
4.3 Methods

4.3.1 Study Population

The NIHS 2000 cohort consisted of 602 individuals. Information on these subjects was obtained from a questionnaire detailing ancestry, lifestyle choices, and medical history. Individual medical histories for complex disorders were provided. Phenotypic information from questionnaires and biochemical testing has been entered into a database for comparative analysis with experimental results. Blood samples were collected (20ml) and sent to the Queensland Medical Laboratory (Southport, Australia) for plasma biochemical analysis.

The complete 6,379 member pedigree compiled from genealogical and participant questionnaires was validated using the Pedigree RELationship Statistical Test (PREST) (McPeek and Sun 2000) prior to all genetic analyses (Bellis, Cox et al. 2008). Using marker data, the program PREST detects mis-specified relationships in a pedigree by determining whether the observed pattern of identity-by-descent (IBD) allele sharing fits the expected IBD pattern given the degree of relationship between relative pairs. Pedigree errors such as mis-specified paternity were resolved by reconfiguring the pedigree structure. Re-assignments were confirmed by re-running the pedigree through PREST. This produced a final inferred pedigree structure of 6,537 (Bellis, Cox et al. 2008). The number of pedigree members is inflated when compared to the original structure. The additional individuals were imputed into the pedigree in cases where only one parent was specified to ensure the pedigree
conformed to standard linkage format. Genotyping was conducted for 372 individuals with key positions in the pedigree.

### 4.3.2 Cardiovascular Disease Risk Score

Cardiovascular disease risk scores were determined using the Framingham–Wilson method outlined in Figure 4.2 below:

![Figure 4.2 Framingham-Wilson based risk score formula (Wilson, D'Agostino et al. 1998)](image)

Figure 4.2 Framingham-Wilson based risk score formula (Wilson, D'Agostino et al. 1998) CHD score sheet for men using TC or LDL-C categories. Uses age, TC (or LDL-C), HDL-C, blood pressure, diabetes, and smoking. Estimates risk for CHD over a period of 10 years based on Framingham experience in men 30 to 74 years old at baseline. Average risk estimates are based on typical Framingham subjects, and estimates of idealized risk are based on optimal blood pressure, TC 160 to 199 mg/dL (or LDL 100 to 129 mg/dL), HDL-C of 45 mg/dL in men, no diabetes, and no smoking. Use of the LDL-C categories is appropriate when fasting LDL-C measurements are available. Pts indicates points.
4.3.3 Ethics

Ethical clearance was granted by the Griffith University Human Research Ethics Committee prior to commencing this study. In addition, signed informed consent was obtained from each individual aged over 18 years prior to participation.

4.3.4 Heritability and Bivariate Analysis

Heritability analyses were conducted using variance components-based methodology implemented in the Sequential Oligonucleotide Linkage Analysis Routines (SOLAR) version 4.0.6 software package (Almasy and Blangero 1998). Heritability estimates were calculated as the ratio of the trait variance that is explained by additive polygenic effects to total phenotypic variance of the trait (Göring, Terwilliger et al. 2001). The applied polygenic model assumes an infinite number of genetic factors, each with a small additive effect contributing to the trait variance. Estimates were screened for the covariate effects using the automodel prompt in SOLAR. Bivariate analyses conducted with SOLAR were used to calculate genotypic, environmental, and phenotypic correlations between metrics.

4.4 Results

Framingham Wilson scores were calculated for N=372 population using variables described in table 4.1:
### Table 4.1 CVD risk factors comprising the Framingham-Wilson Risk Score

<table>
<thead>
<tr>
<th>Fram-Wilson Risk Factor</th>
<th>Fram-Wilson Definition</th>
<th>Norfolk Data Used in Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>30-74</td>
<td>Any</td>
</tr>
<tr>
<td>LDL-C (mg/dl or mmol/L)</td>
<td>All</td>
<td>Any</td>
</tr>
<tr>
<td>HDL-C (mg/dl or mmol/L)</td>
<td>All</td>
<td>Any</td>
</tr>
<tr>
<td>Diastolic/systolicBP(mm/hg)</td>
<td>All</td>
<td>Any</td>
</tr>
<tr>
<td>Diabetes</td>
<td>a) Insulin/ hypoglycemic tx or b) Glucose: blood &gt;150 or fasting &gt;140 dichotomous yes/no</td>
<td>dichotomous yes/no</td>
</tr>
<tr>
<td>Smoker</td>
<td>Regularly smoked ≥1 cigarette/day for 12 months dichotomous yes/no</td>
<td>Have you ever smoked cigarettes, cigars or a pipe? Never, yes current, yes ex If yes at what age? Do you smoke at all now? dichotomous yes/no</td>
</tr>
<tr>
<td>TC (mg/dl or mmol/L)</td>
<td>All</td>
<td>Any</td>
</tr>
</tbody>
</table>

### Table 4.2 Covariates that explain FRS variability in the Norfolk Island isolate.

Significant covariates include age, hdl, dbp, total cholesterol, and diabetes status. Sex, ldl, sbp, and smoking status were insignificant in this analysis.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>2.6x10-9</td>
</tr>
<tr>
<td>Sex</td>
<td>0.3</td>
</tr>
<tr>
<td>Ldl</td>
<td>0.1</td>
</tr>
<tr>
<td>Hdl</td>
<td>5.7x10-7</td>
</tr>
<tr>
<td>Sbp</td>
<td>0.6</td>
</tr>
<tr>
<td>Dbp</td>
<td>2.9x10-5</td>
</tr>
<tr>
<td>Smoker</td>
<td>0.3</td>
</tr>
<tr>
<td>Diabetic</td>
<td>0.02</td>
</tr>
<tr>
<td>Totchole</td>
<td>0.0008</td>
</tr>
</tbody>
</table>
Table 4.3 Bivariate analyses of FRS continuous traits in the Norfolk Island isolate.

<table>
<thead>
<tr>
<th>Trait 1 $h^2$ (SE)</th>
<th>Trait 2 $h^2$ (SE)</th>
<th>$\rho_e$ (SE)</th>
<th>$\rho_g$ (SE)</th>
<th>$\rho_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ldl 0.45 (0.13)</td>
<td>hdl 0.42 (0.14)</td>
<td>-0.03 (0.18) p=0.18</td>
<td>-0.25 (0.25) p=0.34</td>
<td>-0.12 p=0.02</td>
</tr>
<tr>
<td>bp 0.24 (0.12)</td>
<td>hdl 0.42 (0.14)</td>
<td>-0.21 (0.16) p=0.16</td>
<td>0.13 (0.50) p=0.78</td>
<td>-0.12 p=0.03</td>
</tr>
<tr>
<td>ldl 0.45 (0.13)</td>
<td>bp 0.24 (0.12)</td>
<td>0.31 (0.14) p=0.05</td>
<td>0.35 (0.26) p=0.30</td>
<td>0.31 p=1.5x10^{-9}</td>
</tr>
</tbody>
</table>

Framingham risk score values calculated for the study population ranged from 17 to -16 units. Framingham risk score heritability was low at $h^2=0.02$ and insignificant (p=0.4). Table 4.2 shows results of a covariate screen for FRS heritability indicating proportion of FRS variation explained by covariates. Age, hdl, dbp, diabetes status, and total cholesterol were significant covariates and accounted for 50% of FRS variation. Table 4.3 summarizes bivariate analyses of FRS continuous traits. Environmental correlation ($\rho_e=0.31$) was significant between ldl and bp (p=0.05). Phenotypic correlations were significant among all three trait combinations, with negative relationship between ldl/hdl ($\rho_p=-0.12$, p=0.02), bp/hdl ($\rho_p=-0.12$, p=0.03), and positive correlation between ldl/bp ($\rho_p=0.31$, p=1.5x10^{-9})

4.5 Discussion

Although a composite risk score based on several risk factors is more effective at predicting CVD outcomes than single risk factors, the FRS wasn't heritable in the Norfolk Island isolate and wasn't effective in identifying chromosome regions involved in regulating the complex mechanisms contributing to this multigenic disease. A main reason for this negligible and insignificant heritability estimate may be that several factors included in the FRS are individually not heritable. For
example, age was the most significant trait contributing to explain FRS variability, yet is not a heritable factor but rather a surrogate for many contributors to heart disease that converge with time. Similarly, sex and smoking status aren't heritable traits and diabetes status has a strong environmental component, which may not be passed down through generations. Individually heritable traits comprising the FRS include ldl, hdl, bp, and TC; these results along with linkage analyses have been previously published for this population (Bellis, Cox et al. 2008). In the previous study, CVD-risk traits SBP, DBP, total cholesterol, triglycerides, HDLc, LDLc, and BMI were analysed in the initial autosomal genome-wide scan (Bellis et al., 2008). A maximum LOD score of 2.01 on chromosome 1p36 was detected for SBP. Additional QTLs were detected on chromosome 2q37.1 for total triglycerides, 18q22.3 and 20p12.3 for HDLc and chromosomes 8p12 for SBP. As CVD-risk phenotypes have tendency to aggregate, multivariate analysis approaches have the potential locate QTLs undetectable by univariate analysis (Cai et al., 2004; He et al., 2008; Shmulewitz et al., 2001; Tang et al., 2003).

In another NIHS study, multivariate analyses were used to categorize cardiovascular disease traits into 4 principal components, which accounted for 83% of the variation in the original 11 variables (Cox, Bellis et al. 2009). One principal component explained 44.35% of the overall variance, was loaded with body size indicators and interpreted as obesity risk. The remaining principal components explained 9.92-16.87% of the overall variance and were interpreted as obesity and Syndrome X risk, essential hypertension risk and stroke and heart attack risk. A significant genetic component was detected for 3 principal components; the highest LOD score (LOD 1.85) occurred on chromosome 5q35. The 5q34-ter chromosomal region is often
impllicated for lipid level regulation, body mass determination and even type II diabetes mellitus (Almasy et al., 1999; Elbein and Hasstedt, 2002; Feitosa et al., 2002; Hager et al., 1998; He et al., 2008; Pe´russe et al., 2001; Platte et al., 2003; Reynisdottir et al., 2003; Shmulewitz et al., 2006; Zhao et al., 2007). These results support the clustering of CVD risk traits and highlight a region segregating with weight, waist circumference, HDLc and total triglyceride levels.

Hence, perhaps a composite score combining weighted contributions of these traits would yield higher heritability estimates. Although insignificant, moderate genetic correlations between ldl/hdl ($\rho_g = -0.25$, $p=0.34$) and ldl/bp ($\rho_g = 0.35$, $p=0.30$) were found in bivariate analyses. These results may suggest that composite score consisting of these traits may strengthen linkage analyses and help to identify common loci.

### 4.6 Conclusions

The FRS is not heritable in the Norfolk Island population isolate. Bivariate analyses of genetic associations between individual traits were insignificant but suggest relationships between ldl/hdl ($\rho_g = -0.25$), ldl/bp ($\rho_g = 0.35$), and bp/hdl ($\rho_g = 0.13$). More sampled individuals are needed to confirm significance of these relationships and explore bivariate and trivariate linkage mapping with these traits. Alternatively, these traits may be combined to form a composite score, which can be applied to heritability and linkage analyses.
Chapter 5: Pterygia Prevalence Differs Among Norfolk Island Residents in the Norfolk Island Eye Study

5.1 Overview

Pterygia are a common disorder of the ocular surface known to be strongly associated with sun exposure. One aim of this study was to examine differences in pterygia prevalence among Norfolk Island residents. Another aim was to estimate pterygia heritability in the Norfolk Island population isolate.

5.2 Background

Pterygia have a complex etiology thought to involve gene-environment interplay. It was noted half a century ago that pterygia is heritable (Enroth 1951; Hilgers 1960). Studies have shown familial aggregation of cases (Zhang 1987); ethnic (Luthra, Nemesure et al. 2001; Paula, Thorn et al. 2006) and gender (Tan, Lim et al. 2006) (Wong, Foster et al. 2001) differences in pterygia prevalence and recurrence rate (Kandavel, Kang et al. 2010); and clustering of pterygia in monozygotic twins (Faraldi and Gracis 1976). Pterygium prevalence is higher in individuals with a family history of the condition, suggesting either genetic components (Booth 1985) (Bradley, Yang et al. 2009) or lifestyle behaviours are passed down through generations. Solar ultraviolet radiation (UVR), specifically ultraviolet B wavelength (UVB) (Saw and Tan 1999), is an important risk factor for pterygia. Recent reviews
suggest that UVR can induce pterygia through immunological, viral, cell-related gene expression and extracellular remodelling (Di Girolamo, Chui et al. 2004; Chui, Di Girolamo et al. 2008; Bradley, Yang et al. 2009). Experimental studies demonstrate that growth factors (Nolan, DiGirolamo et al. 2003) (Di Girolamo, Wakefield et al. 2006), cytokines (Di Girolamo, Kumar et al. 2002) and extracellular remodelling proteins (Di Girolamo, Coroneo et al. 2003) are expressed in pterygia and can be induced by UVB. Ultraviolet radiation may contribute to formation of radical oxygen species (Tsai, Cheng et al. 2005). Other pterygia risk factors include increasing age, outdoor work (Khoo, Saw et al. 1998), and rural residence (McCarty, Fu et al. 2000).

Population isolates have been studied to elucidate the genetic basis of ophthalmologic diseases (Sherwin, Hewitt et al. 2008). Founder populations that remained in isolation throughout generations have a smaller gene pool and reduced confounding environmental factors due to a homogenous diet, lifestyle, and cultural and physical environment. Increased genetic and environmental homogeneity can simplify associations with disease risk (Shifman and Darvasi 2001). Genetic studies conducted in population isolates also have increased power to detect heritable effects and greater potential for identifying quantitative trait loci.

Norfolk is a small island (34.6km²) thus residents have a similar environment. The Norfolk Island population isolate originally settled Pitcairn Island in 1790, and were founded by mutineers from the HMS Bounty. In 1856, 193 people living on Pitcairn Island relocated to Norfolk Island. Approximately 43% of the permanent Norfolk Island population can trace their ancestry to Pitcairn Island. Norfolk Island residents with Pitcairn Island ancestry have approximately 88% European and 12% Polynesian
genetic admixture, though substantial variation exists between individuals (McEvoy, Zhao et al. 2009). Interestingly, founder effect and admixture have been shown to influence a variety of cardiovascular traits in the Norfolk Island population (Macgregor, Bellis et al. 2010). The Norfolk Island Eye Study (Mackey, Sherwin et al. 2011) researchers conducted various projects including prevalence of ocular diseases (Sherwin, Kearns et al. 2011). Distribution of UVAF measurements in the Norfolk Island population has been described elsewhere (Sherwin, Hewitt et al. 2011).

We hypothesised that genetic and environmental factors could interact and contribute to the development of pterygium. In this study, we examined differences in pterygia prevalence among Norfolk Island residents and sought to estimate heritability in the genetic isolate.

5.3 Methods

5.3.1 Study Population and Recruitment

We conducted a population-based study on Norfolk Island, an external territory of Australia in 2007-8. Permanent residents aged ≥15 years were recruited by radio and newspaper advertisements, referral from healthcare providers, and word of mouth. In addition, subjects enrolled in a previous study of cardiovascular disease conducted on the Island by Griffith University were sent invitations (Bellis, Hughes et al. 2005). There were no exclusion criteria.
Conjunctival ultraviolet autofluorescence (UVAF) photography was performed on 641/781 (82.1%) participants in the Norfolk Island Eye Study (NIES), accounting for 61% of permanent residents aged ≥15, and nearly 75% of individuals ≥50 years.

5.3.2 Ethical Approval

The NIES received approval from the Human Research and Ethics Committee at both Griffith University and the Royal Victorian Eye and Ear Hospital in Melbourne. Consent was obtained to conduct the ophthalmic examination and link with earlier cardiovascular disease genetic research as well as ongoing genetic eye research. In addition, there was local community consultation with the hospital administration, local doctors, local optometrist and visiting ophthalmologists to address concerns regarding the possible long-term impact of the study.

5.3.3 Pterygium Analysis

Clinical examination of the anterior segment was performed with a slitlamp biomicroscope (Topcon, Newington, Australia). Diagnosis of pterygium was confirmed by anterior segment control photography. Pterygia were graded by one individual according to criteria outlined in Tan et al (Tan, Chee et al. 1997). There are three grades based on the visibility of underlying episcleral vessels: Grade 1 (vessels clearly visible), Grade 2 (vessels partially visible) and Grade 3 (opaque vessels). Patterns of autofluorescence in pterygia were assessed according to criteria from Ooi et al. (Ooi, Sharma et al. 2007) (leading edge; limbus; both leading edge and limbus; none).
5.3.4 Conjunctival UV Auto-Fluorescence

UV photographic methods have been published elsewhere (Ooi, Sharma et al. 2006; Ooi, Sharma et al. 2007). In brief, photographs were taken using both reflected visible light (control) and UV-induced fluorescence with the aid of two portable photographic systems. Each system consisted of a height-adjustable table equipped with subject head-rest, camera positioning assembly, digital single-lens reflex camera, macro lens, and filtered electronic flash. Each eye was photographed at 0.94 magnification, with separate views of nasal and temporal regions from both eyes. Colored low-voltage light emitting diodes (LED) were positioned on stands in the subject’s visual field, 35° to the camera–subject axis to aid fixation.

The UV-induced fluorescence photography was based on standard principles, using a specially adapted electronic flash system fitted with UV-transmission filters (300-400 nm, peak 365 nm) as the excitation source. Subject fluorescence was recorded with a Nikon D100 (Nikon, Melville, New York, USA) digital camera and 105 mm f/2.8 Micro Nikon (Nikon, Melville, New York, USA) lens fitted with infrared and UV barrier filters. Thus, the camera recorded only fluorescence. Images were saved in RGB format at the D100 settings of JPEG fine (1:4 compression) and large resolution. Each photograph was verified immediately and reshot, if necessary, to obtain a clearer image.

Adobe Photoshop (Adobe Systems Inc, San Jose, California, USA) was used to perform the quantitative analysis. Four photos were analysed per person (right/left,
nasal/temporal). Setting lengths required for UV analysis were pixel=3008 and logical=2.4, with resulting area expressed in mm$^2$.

5.3.5 Genealogical Structure

Genealogical records confirm nine English HMS Bounty ship mutineers, twelve Tahitian women, and six Tahitian men (Hoare 1999) settled Pitcairn Island in 1790. Pedigree reconstruction confirmed current descendants possess lineages to all nine mutineers, six Tahitian women and three Caucasian sailors who joined the colony in the early 19th century (Macgregor, Bellis et al. 2010; Bellis, Hughes et al. 2005; McEvoy, Zhao et al. 2009). To facilitate heritability analyses, the large, complex pedigree (N=6537) was trimmed (N=1,078) using the pedigree database management system PEDSYS (Texas Biomedical Research Institute, San Antonio, Texas, USA) (Dyke 1996).

5.3.6 Power Calculation and Heritability Analyses

Power calculation for the pedigree and heritability analysis of pterygium were conducted using variance components-based methodology implemented in the Sequential Oligonucleotide Linkage Analysis Routines (SOLAR) version 4.0.6 software package (Almasy and Blangero 1998). The power calculation was conducted using the SOLAR h2power simulation to estimate frequency of significantly non-zero heritability estimates assuming a portion of the phenotypic variance is due to genetics (Williams and Blangero 2004). SOLAR asymptotically distributes twice the difference in the loglikelihoods of the polygenic and sporadic
models (1/2:1/2) of a chi-square random variate with one degree of freedom and a point mass at zero. Significance means the probability of obtaining the observed chi-square value, in the absence of a genetic effect, is \( \leq 0.05 \).

The model was analysed under a classical liability threshold (Falconer 1989) that assumes discrete traits are actually unobserved continuous traits where threshold values partition the distribution into two intervals that become distinct phenotypes. Heritability estimates were calculated as the ratio of the trait variance explained by additive polygenic effects to total phenotypic variance of the trait (Dyer, Blangero et al. 2001). The applied polygenic model assumes an infinite number of genetic factors, each with a small additive effect contributing to the trait variance. Estimates were screened for covariate effects of age, sex, and several eye metrics. Covariates with p-values less than or equal to 0.10 were retained in the final model. Assuming a liability threshold model also enabled dichotomous trait heritability analysis (Duggirala, Williams et al. 1997). A Kullback-Leibler \( r^2 \) was calculated to determine covariate contribution to the analysis (Kullback 1968).

### 5.3.7 Relative Risk Analysis

Total UVAF was divided into quartiles. Time outdoors categories were combined: \(~1/4, ~half, ~3/4\) of day. Differences in categorical variables were assessed with the Chi-square test. Differences in two continuous variables were assessed with Mann Whitney U test. Trends across categories were assessed using Cuzick’s nonparametric test for trend. All p-values presented are two-tailed. The odds ratio (OR) of pterygium in those with and without Pitcairn ancestry was stratified by
variables of interest. The Mantel-Haenszel test for heterogeneity across strata was used to assess OR differences related to these variables. Statistical significance was set at <0.05. All analyses were conducted in Intercooled Stata 8.2 for Windows (StataCorp, College Station, 2007).

5.4 Results

5.4.1 Sample Population

Approximately half the study population could trace their ancestry to Pitcairn Island (Table 5.1). Mean age of all study participants was 54.1 years (SD 16.2), with age ranging from 15 to 89 years. Mean age was lower in the Pitcairn ancestry group (51.7 vs. 57.4 years, p=0.0002) compared to other residents. There were no statistically significant differences in the prevalence of hypertension (17.9% vs. 15.2%, p=0.355), cigarette smoking (59.1% vs. 52.6%, p=0.098), diabetes (Type 1 or 2) (3.8% vs. 4.9%, p=0.991) or cataract (25.2% vs. 19.8%, p=0.105) between Pitcairn descendants and other residents respectively. Pitcairn Island descendants spent a greater proportion of their lifespan on Norfolk Island (71.9% vs. 52.9%, p=0.001).

5.4.2 Prevalence Estimates by Ancestry

The prevalence of unilateral pterygium was higher in Pitcairn descendants, 8.5% (95% CI 5.4-11.6% vs. 4.3% (95% CI 2.1-6.5%), Chi-square=4.62, P=0.032. For bilateral pterygium, prevalence was also higher in people with Pitcairn ancestry, 6.0%
(95% CI 3.4-8.6%) vs. 3.1% (95% CI 1.2-5.0%), Chi-square=3.07, P=0.080. Prevalence estimates were calculated for unilateral (Table 5.2) and bilateral (Table 5.3) pterygium. In both cases, pterygium prevalence was two-fold higher in Pitcairn descendants, with males having a higher risk than females.

After controlling for age and sex, the OR of Pitcairn ancestry for pterygium in any eye was 2.30 (95% CI 1.35-3.92, p=0.002). The association remained significant after further controlling for time spent on Norfolk Island, and conjunctival UVAF 1.96 (1.14-3.39, p=0.015).

Ninety-nine pterygia-positive eyes (7.7% of total) were graded for severity and degree of autofluorescence (Table 5.4). All pterygia were found on the nasal surface. More than half (66.7%) of pterygia detected in either or both eyes were found in Pitcairn Island descendants.

Prevalence estimates were calculated for unilateral (Table 5.2) and bilateral (Table 5.3) pterygium. In both cases, pterygium prevalence was approximately two-fold (OR 2.05 (CI 1.05-3.98) unilateral; 6.0 vs 3.1 OR1.99 (CI 0.91-4.35) bilateral) higher in Pitcairn descendants, with odds significantly higher in males (OR 3.87 (CI 1.23 – 12.14) bilateral).

Ninety-nine pterygia positive eyes (7.7% of total) were graded for severity and degree of autofluorescence (Table 5.4). All pterygia were found on the nasal surface. Approximately two thirds (66.7%) of all unilateral and bilateral cases were of Pitcairn descent.
5.4.3 Power and Heritability Analyses

Power for discrete traits was modelled with a 20% prevalence across the N=6537 pedigree. The model assumes complete phenotypic information for each individual. Results are summarised in Figure 5.1 and show that with complete phenotypic information, 80% of replicated simulations were able to detect heritability estimates >0.40.

Two hundred and seventeen individuals were detected in the N=6537 extended pedigree and included in the pterygium heritability analysis. Of these, 46 (21%) pterygia-positive participants were identified. Pterygia heritability did not differ significantly from a model of h^2=0.024 (SE 0.16, p=0.44). Table 5.5 lists covariates used in the heritability calculation; visual acuity with pinhole (p=0.07), UVAF (p=0.01) and axis (p=0.03) were significant in the model while age and sex were insignificant. The proportion of pterygium variability attributable to covariates was estimated as 9.0% (Kullback-Leibler r^2).

5.5 Discussion

We performed a cross-sectional survey of ocular metrics on Norfolk Island to elucidate pterygium prevalence and UVR exposure in this population. We found that the Pitcairn Island descendants had a doubled risk of developing pterygium compared to other residents. Moreover, Pitcairn descendants had larger UVAF measurements and were twice as likely to spend more than 3/4 of their day outside compared to
A combined increase in both pterygium prevalence and median UVAF further supports the role for UVR in pterygia etiology.

Although pterygium heritability was negligible and insignificant ($h^2 = 0.024$, $p=0.44$), significant covariates were UVAF, visual acuity through pinhole, and cylinder axis. The association with cylinder axis can be explained since pterygium occurs at the 180° (or 0°) end of the axis and will distort the cornea. Although male gender (Wong, Foster et al. 2001) has previously been found to be a risk factor for pterygium development, sex was not a significant covariate in the heritability analysis. An unequal variance Welch's test indicated that UVAF measurements were significantly lower in females ($M=0.66$, $SD=0.83$) than in males ($M=0.80$, $SD=1.00$), $F(1,430)=9.35$, $p=0.0024$, which may explain the gender differences in pterygium risk.

Pitcairn descendants also spent significantly more time on Norfolk Island, which is located in the southern hemisphere south of New Zealand. An inverse relationship between latitude and pterygium prevalence has been observed (Mackenzie, Hirst et al. 1992). In an Australian case-control study, participants who spent their first five years at latitudes <30° had 40 times greater pterygium risk than those living at latitudes more than 40° (Mackenzie, Hirst et al. 1992). Both Pitcairn and non-Pitcairn descendents reported similar use of protective behaviours such as wearing sunglasses, hats, and staying in the shade. Although smoking history has been found to be protective for developing pterygium (West and Munoz 2009), this relationship wasn't observed in our study.

Pterygia may be used as a biomarker for other diseases that have a strong correlation with UVR. UVR exposure is strongly associated with eye disease, both acute
photokeratitis or flash burns) and chronic (pterygium, climatic droplet keratopathy, cataract, ocular surface squamous neoplasia, and eyelid tumours; basal and squamous cell carcinoma and melanoma) (Oliva and Taylor 2005). Therefore, documenting increased risk of developing UVR-related disease, whether due to genetics or lifestyle, has important public health ramifications. At present, public health experts advocate that sun exposure should be guided by level of ambient UVR and skin type (Lucas, McMichael et al. 2008). Hence, knowledge of an individual’s UVR response may lead to targeted prevention and therapeutic strategies.

In conclusion, we discovered population differences in pterygium prevalence on Norfolk Island. We also found that although pterygium was not significantly heritable given the limited variability of discrete traits, visual acuity, UVAF, and axis measures were significant covariates in the analysis. Hence, further genetic work involving discrete traits and the Norfolk Island extended pedigree will require more phenotyped individuals or should focus on continuous variable environmental factors.
Table 5.1 Baseline characteristics between two groups

<table>
<thead>
<tr>
<th>Category</th>
<th>Pitcairn</th>
<th></th>
<th>Non-Pitcairn</th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (% of total)</td>
<td>318</td>
<td>49.61</td>
<td>323</td>
<td>50.39</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>175</td>
<td>55.03</td>
<td>187</td>
<td>57.89</td>
<td>0.465</td>
</tr>
<tr>
<td>Male</td>
<td>143</td>
<td>44.97</td>
<td>136</td>
<td>42.11</td>
<td></td>
</tr>
<tr>
<td>Age group, N (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>15-29</td>
<td>37</td>
<td>11.64</td>
<td>15</td>
<td>4.64</td>
<td></td>
</tr>
<tr>
<td>30-39</td>
<td>47</td>
<td>14.78</td>
<td>31</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>59</td>
<td>18.55</td>
<td>57</td>
<td>17.65</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>74</td>
<td>23.27</td>
<td>76</td>
<td>23.53</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>49</td>
<td>15.41</td>
<td>82</td>
<td>25.39</td>
<td></td>
</tr>
<tr>
<td>70 +</td>
<td>52</td>
<td>16.35</td>
<td>62</td>
<td>19.2</td>
<td></td>
</tr>
<tr>
<td>Lifetime on Norfolk N (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.969</td>
</tr>
<tr>
<td>No (left island at some time)</td>
<td>276</td>
<td>86.79</td>
<td>280</td>
<td>86.69</td>
<td></td>
</tr>
<tr>
<td>Yes (all of life)</td>
<td>42</td>
<td>13.21</td>
<td>43</td>
<td>13.31</td>
<td></td>
</tr>
<tr>
<td>UVAF mm² (% of total in quartile)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>1st quartile (&lt;=14.4)</td>
<td>57</td>
<td>17.92</td>
<td>102</td>
<td>31.58</td>
<td></td>
</tr>
<tr>
<td>2nd quartile (14.5 – 28.0)</td>
<td>72</td>
<td>22.64</td>
<td>89</td>
<td>27.55</td>
<td></td>
</tr>
<tr>
<td>3rd quartile (28.1 – 47.7)</td>
<td>84</td>
<td>26.42</td>
<td>76</td>
<td>23.53</td>
<td></td>
</tr>
<tr>
<td>4th quartile (&gt;=47.8)</td>
<td>105</td>
<td>33.02</td>
<td>56</td>
<td>17.34</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.2 Prevalence of unilateral pterygium by Pitcairn ancestry

<table>
<thead>
<tr>
<th>Category</th>
<th>Pitcairn Ancestry</th>
<th>Non-Pitcairn</th>
<th>Stratum-specific</th>
<th>P</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>N pterygium (%)</td>
<td>N pterygium (%)</td>
<td>OR (95%CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>11 (6.3)</td>
<td>7 (3.7)</td>
<td>1.72 (0.65 – 4.57)</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16 (11.2)</td>
<td>7 (5.2)</td>
<td>2.32 (0.92 – 5.88)</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Age group, N (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-29</td>
<td>1 (2.7)</td>
<td></td>
<td></td>
<td></td>
<td>0.52</td>
</tr>
<tr>
<td>30-39</td>
<td>5 (10.6)</td>
<td>1 (3.2)</td>
<td>3.58 (0.38 – 33.30)</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>7 (11.9)</td>
<td>4 (7.0)</td>
<td>1.78 (0.49 – 6.52)</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>5 (6.8)</td>
<td>1 (1.3)</td>
<td>5.43 (0.60 – 49.07)</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>6 (12.2)</td>
<td>7 (8.5)</td>
<td>1.49 (0.47 – 4.76)</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>70+</td>
<td>3 (5.8)</td>
<td>1 (1.6)</td>
<td>3.73 (0.37 – 38.00)</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Lived on Norfolk all of life, N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (left island at some time)</td>
<td>22 (8.0)</td>
<td>13 (4.6)</td>
<td>1.77 (0.87-3.61)</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Yes (all of life)</td>
<td>5 (11.9)</td>
<td>1 (2.3)</td>
<td>5.68 (0.60-53.56)</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>UVAF mm2 (% of total in)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile (&lt;=14.4)</td>
<td>2 (3.5)</td>
<td>3 (2.9)</td>
<td>1.20 (0.19-7.44)</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>2nd quartile (14.5 – 28.0)</td>
<td>6 (8.3)</td>
<td>3 (3.4)</td>
<td>2.61 (0.62-10.95)</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>3rd quartile (28.1 – 47.7)</td>
<td>6 (7.1)</td>
<td>5 (6.6)</td>
<td>1.09 (0.32-3.75)</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>4th quartile (&gt;=47.8)</td>
<td>13 (12.4)</td>
<td>3 (5.4)</td>
<td>2.49 (0.67-0.27)</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>P Trend</td>
<td>P=0.857</td>
<td>P=0.819</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR- Odds ratio of unilateral pterygium in those with Pitcairn ancestry compared to those without Pitcairn ancestry

*p values relate to Mantel-Haenszel test for heterogeneity across strata
Table 5.3 Prevalence of bilateral pterygium by Pitcairn ancestry

<table>
<thead>
<tr>
<th>Category</th>
<th>Pitcairn Ancestry</th>
<th>Non-Pitcairn Ancestry</th>
<th>Stratum-specific OR (95%CI)</th>
<th>P</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>19 (6.0)</td>
<td>10 (3.1)</td>
<td>1.99 (0.91-4.35)</td>
<td>0.085</td>
<td>-</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.055</td>
</tr>
<tr>
<td>Female</td>
<td>4 (2.3)</td>
<td>6 (3.2)</td>
<td>0.71 (0.20 – 2.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>15 (10.5)</td>
<td>4 (2.9)</td>
<td>3.87 (1.23 – 12.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age group, N (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.925</td>
<td></td>
</tr>
<tr>
<td>15-29</td>
<td>0</td>
<td>0</td>
<td>0.99 (0.15 – 6.36)</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>30-39</td>
<td>3 (6.4)</td>
<td>2 (6.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49</td>
<td>3 (5.1)</td>
<td>1 (1.8)</td>
<td>2.90 (0.30 – 30.31)</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>5 (6.8)</td>
<td>2 (2.6)</td>
<td>2.57 (0.50 – 13.47)</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>3 (6.1)</td>
<td>2 (2.4)</td>
<td>2.51 (0.41 – 16.43)</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>70+</td>
<td>5 (9.6)</td>
<td>3 (4.8)</td>
<td>1.99 (0.47 – 9.33)</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>P Trend</td>
<td></td>
<td></td>
<td></td>
<td>0.113</td>
<td>0.699</td>
</tr>
<tr>
<td>Lifetime on Norfolk N (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.241</td>
<td></td>
</tr>
<tr>
<td>No (left island at some time)</td>
<td>18 (6.5)</td>
<td>8 (2.9)</td>
<td>2.37 (1.01 – 5.57)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Yes (all of life)</td>
<td>1 (2.4)</td>
<td>2 (4.7)</td>
<td>0.50 (0.04 – 5.84)</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>UVAF mm² (% of total in quartile)</td>
<td></td>
<td></td>
<td></td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td>1st quartile (&lt;=14.4)</td>
<td>1 (1.8)</td>
<td>1 (1.0)</td>
<td>1.80 (0.11 – 29.70)</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>2nd quartile (14.5 – 28.0)</td>
<td>4 (5.6)</td>
<td></td>
<td>0 -</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>3rd quartile (28.1 – 47.7)</td>
<td>4 (4.8)</td>
<td>6 (7.9)</td>
<td>0.58 (0.15 – 2.16)</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>4th quartile (&gt;=47.8)</td>
<td>10 (9.5)</td>
<td>3 (5.4)</td>
<td>1.86 (0.49 – 7.11)</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>P Trend</td>
<td>P=0.058</td>
<td>P=0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR- Odds ratio of bilateral pterygium in those with Pitcairn ancestry compared to those without Pitcairn ancestry

* p values relate to Mantel-Haenszel test for heterogeneity across strata
### Table 5.4 Comparison of pterygium severity and UV autofluorescence between Pitcairn and Non-Pitcairn ancestry

<table>
<thead>
<tr>
<th>Category</th>
<th>Pitcairn (N=65)</th>
<th>Non-Pitcairn (N=34)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><strong>Grade of Pterygia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1- vessels clearly visible</td>
<td>7</td>
<td>10.77</td>
<td>10</td>
</tr>
<tr>
<td>2- vessels partly visible</td>
<td>40</td>
<td>61.54</td>
<td>18</td>
</tr>
<tr>
<td>3- vessels opaque</td>
<td>18</td>
<td>27.69</td>
<td>6</td>
</tr>
<tr>
<td><strong>Pattern of Autofluorescence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1- leading edge</td>
<td>10</td>
<td>15.38</td>
<td>2</td>
</tr>
<tr>
<td>2- limbus</td>
<td>18</td>
<td>27.69</td>
<td>12</td>
</tr>
<tr>
<td>3- both leading edge and limbus</td>
<td>33</td>
<td>50.77</td>
<td>16</td>
</tr>
<tr>
<td>4- none</td>
<td>4</td>
<td>6.15</td>
<td>4</td>
</tr>
<tr>
<td><strong>Degree Autofluorescence (median; mm²)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(total of eye [sum of four measurements])</td>
<td>25.8</td>
<td>19.8</td>
<td>0.127+</td>
</tr>
<tr>
<td>(nasal [sum of two measurements])</td>
<td>12.1</td>
<td>9.0</td>
<td>0.074+</td>
</tr>
<tr>
<td>(temporal [sum of two measurements])</td>
<td>14.0</td>
<td>9.8</td>
<td>0.093+</td>
</tr>
</tbody>
</table>

* P value refers to Chi-square test
+ P value refers to Mann-Whitney Test
Table 5.5 Covariates used in pterygium heritability analysis

<table>
<thead>
<tr>
<th>Covariate</th>
<th>p-value</th>
<th>Chi (deg=1)</th>
<th>loglikelihood without covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>0.16</td>
<td>1.96</td>
<td>-62.74</td>
</tr>
<tr>
<td>Age</td>
<td>0.56</td>
<td>0.34</td>
<td>-61.92</td>
</tr>
<tr>
<td>Age*sex</td>
<td>0.23</td>
<td>1.43</td>
<td>-62.47</td>
</tr>
<tr>
<td>Age^2</td>
<td>0.89</td>
<td>0.02</td>
<td>-61.76</td>
</tr>
<tr>
<td>Age^2*sex</td>
<td>0.56</td>
<td>0.33</td>
<td>-61.92</td>
</tr>
<tr>
<td>Height</td>
<td>0.14</td>
<td>2.13</td>
<td>-62.82</td>
</tr>
<tr>
<td>UVAF</td>
<td>0.001</td>
<td>11.29</td>
<td>-67.40</td>
</tr>
<tr>
<td>Glaucoma status</td>
<td>0.25</td>
<td>1.28</td>
<td>-62.40</td>
</tr>
<tr>
<td>Visual acuity</td>
<td>0.11</td>
<td>2.49</td>
<td>-63.00</td>
</tr>
<tr>
<td>Visual acuity through pinhole</td>
<td>0.03</td>
<td>4.77</td>
<td>-64.14</td>
</tr>
<tr>
<td>Cylinder</td>
<td>0.23</td>
<td>1.23</td>
<td>-62.37</td>
</tr>
<tr>
<td>Cylinder axis</td>
<td>0.06</td>
<td>3.41</td>
<td>-63.46</td>
</tr>
<tr>
<td>Anterior chamber depth</td>
<td>0.13</td>
<td>2.2</td>
<td>-62.85</td>
</tr>
<tr>
<td>Horizontal corneal curvature (k-value)</td>
<td>0.46</td>
<td>0.5</td>
<td>-62.02</td>
</tr>
<tr>
<td>Eye color</td>
<td>0.15</td>
<td>0.20</td>
<td>-62.80</td>
</tr>
<tr>
<td>Time spent outdoors</td>
<td>0.61</td>
<td>0.27</td>
<td>-61.89</td>
</tr>
</tbody>
</table>
Figure 5.1 Power simulation of the proportion of replicates that detected significant genetic effect given 20% prevalence of a discrete trait by heritability estimate assuming phenotypic information for all individuals in the Norfolk Island pedigree N=6537.

5.6 Conclusion

Differences were observed in pterygium prevalence between Norfolk Island residents with and without Pitcairn Island ancestry. Heritability estimates for pterygium were negligible and not statistically significant ($h^2=0.024$, $p=0.44$), reflecting low trait heritability and limited analysable information of dichotomous traits for the given trait prevalence and number of phenotyped individuals.
Chapter 6: Heritability and Linkage Scan of Glaucoma Endophenotypes in the Norfolk Island Isolate

6.1 Overview

The goal of this study was to investigate heritability and linkage mapping of quantitative ocular traits intraocular pressure (IOP), central corneal thickness (CCT), vertical cup-to-disc ratio (C:D), and anterior chamber depth (ACD) in the Norfolk Island population isolate. Hence, heritability estimates were determined and subsequent linkage mapping was performed. This work combines pedigree and genetic resources from the Norfolk Island Health Study (NIHS) (Macgregor, Bellis et al. 2010; Bellis, Hughes et al. 2005; Bellis, Cox et al. 2008; Bellis, Cox et al. 2008; Cox, Bellis et al. 2009) with phenotypic data from the Norfolk Island Eye Study (NIES), which characterised epidemiology and environmental contributions to eye disease in this population (Mackey, Sherwin et al. 2011) (Sherwin, Kearns et al. 2011).

6.2 Background

Glaucoma is a leading cause of visual impairment and irreversible blindness affecting 70 million people worldwide (Quigley 1996). Chronic glaucoma prevalence in the
US population 40 years and older is 1.86% (Friedman, Jampel et al. 2006). With an aging population, this prevalence is projected to increase by 50% in 2020 (Friedman, Wolfs et al. 2004). Glaucoma prevalence in Australia is higher than the US at 3.0% (Mutti, Mitchell et al. 2002).

Glaucoma refers to a group of heterogeneous diseases causing optic nerve fiber damage. Primary open angle glaucoma (POAG) is the main type of glaucoma in most populations. Since eye pressure plays a role in damaging optic nerve fibers, intraocular pressure (IOP) >22mmHg is a defining criterion for the POAG form of the disease (Quigley 1993). However, normal tension glaucoma (NTG) can occur with normal eye pressure (Werner 1996) and prevalence of this form is significantly higher in Asian than Caucasian populations (Shiose, Kitazawa et al. 1991; Fuse 2010). Also, glaucoma can manifest as juvenile or adult onset, yet further evidence of the disease complexity and multiple etiologies comprising this condition.

At least 14 genetic loci have been associated with glaucoma susceptibility; of these myocilin (GLC1A) located on chromosome 1q23, optineurin (GLC1E) on 10p15-p14, and WDR36 (GLC1G) on 5q22 play a causative role in the onset of this disease (Fuse 2010). Myocilin (MYOC) on 1q24.3 has been shown to account for 3% of adult onset POAG (Stone, Fingert et al. 1997; Fingert, Heon et al. 1999; Alward 2000; Alward, Kwon et al. 2002). Optineurin mutations may play a role in both NTG and adult-onset POAG by inducing death in retinal ganglion cells (Quigley 1993). At least 27 other loci related to glaucoma have been identified through association studies; interestingly, none correspond to specific eye functions but instead are involved in systemic processes implicated in other diseases (Fuse 2010).
In addition to IOP, other POAG risk factors include age, race, central corneal thickness (CCT), high myopia, diabetes (Bonovas, Peponis et al. 2004), cigarette smoking, (Bonovas, Filioussi et al. 2004) and positive family history (Rosenthal and Perkins 1985; Tielsch, Katz et al. 1994; Wolfs, Klaver et al. 1998; Weih, Nanjan et al. 2001). The lifetime risk of glaucoma in relatives of patients with glaucoma is 22% while it is 2.3% in controls resulting in a risk ratio of 9.2 (Wolfs, Klaver et al. 1998). Despite this increased risk, heritability studies of POAG conducted in twin studies report low estimates ($h^2=0.13$) (Teikari 1987). In addition to IOP, and CCT, other ocular metrics such as vertical cup-to-disc ratio (C:D) and anterior chamber depth (ACD) have been studied in glaucoma (van Koolwijk, Despriet et al. 2007; Charlesworth, Kramer et al. 2010).

6.2.1 Genetics of Ocular Metric Endophenotypes

This section summarises heritability studies and linkage analyses conducted for IOP, CCT, C:D, and ACD.

6.2.1.1 Intraocular Pressure

Elevated IOP is a principal risk factor for POAG and can cause optic nerve damage (Quigley, West et al. 1993). In the US, at least three million people have elevated IOP (ocular hypertension) and are consequently at an increased risk for developing
glaucoma (Friedman, Wolfs et al. 2004). This somewhat broad range of heritabilities may reflect the presence of non-genetic components regulating this trait. For example, associations between IOP and systemic blood pressure - a trait known to be influenced by lifestyle, have been documented (McLeod, West et al. 1990; Dielemans, Vingerling et al. 1995; Tielsch, Katz et al. 1995).

Several twin studies have estimated heritabilities for IOP. Two twin studies calculated lower heritabilities at 0.35 (Schwartz, Reuling et al. 1973) and 0.36 (Klein, Klein et al. 2004), while five other studies derived higher heritabilities ranging from 0.62-0.67 (Kalenak and Paydar 1995; Carbonaro, Andrew et al. 2008; Carbonaro, Andrew et al. 2009) in older women (Parssinen, Era et al. 2007) and Chinese children (Zheng, Xiang et al. 2009) populations.

The published range of heritability estimates for IOP in family studies is $h^2=0.29-0.48$ (Levene, Workman et al. 1970; Chang, Congdon et al. 2005; Duggal, Klein et al. 2005; van Koolwijk, Despriet et al. 2007; Charlesworth, Kramer et al. 2010; Lee, Woo et al. 2010), which is lower than the range for twin studies. Heritability in Dutch extended pedigrees (N=2620) was $h^2=0.35$ (95% CI, 0.27-0.43) (van Koolwijk, Despriet et al. 2007), $h^2=0.42$ in a combined family pedigree study of Australian and US populations (Charlesworth, Kramer et al. 2010), and $h^2=0.48$ in a Mongolian population (Lee, Woo et al. 2010).

In one study, linkage mapping of IOP identified seven regions of interest on chromosomes 2, 5, 6, 7, 12, 15, and 19 (Duggal, Klein et al. 2007). Chromosomes 6 and 13 have been implicated in another IOP linkage study (Duggal, Klein et al. 2005).
Evidence for linkage of maximum IOP has been mapped to chromosome 10q22 (Charlesworth, Dyer et al. 2005) near region 10p15-p14, which is a locus shown to be involved in up to 17% of low tension glaucoma pedigrees (Rezaie, Child et al. 2002). Chromosome 2p16.3-p15 is implicated in IOP onset over 50 years, and middle to high IOP (Suriyapperuma, Child et al. 2007) while 2cen-q13 is found in normal tension glaucoma, with onset over 50 years (Stoilova, Child et al. 1996). Chromosome 3q21-q24 is implicated in IOP onset over 50 years (Wirtz, Samples et al. 1997).

6.2.1.2 Central Corneal Thickness

Central corneal thickness is a potential confounder for IOP. Corneal thickness <588 μ is considered a risk factor for POAG (Hansen and Ehlers 1971; Stodtmeister 1998; Gordon, Beiser et al. 2002; Kohlhaas, Boehm et al. 2006). Heritability of CCT is consistently high in both twin and family studies ranging from 0.65-0.95, which may indicate minimal environmental contribution influencing this trait. Twin studies derived similar heritability estimates in females (h²=0.91) and males (h²=0.88) in a Chinese population (Zheng, Ge et al. 2008). In another twin study, Toh et al. estimated heritability to be h²=0.95 (Toh, Liew et al. 2005). Family studies derived similar heritabilities (h²=0.65-0.72) (Alsbirk 1978; Landers, Hewitt et al. 2009; Charlesworth, Kramer et al. 2010). Linkage mapping of CCT identified loci on chromosomes 13 and 16 (Lu, Dimasi et al. 2010).
6.2.1.3 Vertical Cup-to-disc Ratio

Cup-to-disc ratio heritability estimates in twins or nuclear family studies range from \( h^2 = 0.48-0.70 \). Although no twin studies examined vertical C:D, heritability in horizontal C:D was estimated to be \( h^2 = 0.70 \) (Schwartz, Reuling et al. 1975). Family study estimates of vertical C:D ranged from \( h^2 = 0.48-0.66 \) (Klein, Klein et al. 2004; Chang, Congdon et al. 2005; Charlesworth, Dyer et al. 2005; van Koolwijk, Despriet et al. 2007). Maximum C:D has been mapped with significant evidence of linkage to glaucoma causative gene MYOC on 1q24.3 (Charlesworth, Dyer et al. 2005).

6.2.1.4 Anterior Chamber Depth

Generally, shallow ACD indicates various eye problems including glaucoma. Heritability of ACD ranges from 51-94%. Anterior chamber depth heritability estimates in the Chinese Guangzhou twin eye study were high at 0.90 (He, Wang et al. 2008) and also high (\( h^2 > 0.90 \)) in two other twin studies (Tu, Yin et al. 2008) (Lyhne, Sjolie et al. 2001). Sex based heritability differences were calculated in another twin study and were considerably lower in males (\( h^2 = 0.51 \)) than females (\( h^2 = 0.78 \)) (Dirani, Chamberlain et al. 2006). In a family study, heritability was similar between sexes at (\( h^2 = 0.44 \) males, \( h^2 = 0.47 \) females) (Biino, Palmas et al. 2005). In other family studies, ACD heritability was consistently between 64-78% (Chen, Scurrah et al. 2007; Klein, Suktitipat et al. 2009), (Sorsby and Fraser 1964),

6.3 Methods

6.3.1 Ethics

The original NIHS received ethics approval from the Griffith University, Human Research and Ethics Committee in Queensland. This same committee, in addition to the Human Research and Ethics Committee at the Royal Victorian Eye and Ear Hospital in Melbourne, approved the NIES. Consent was obtained to conduct the ophthalmic examination then link with the earlier cardiovascular and genetic research as well as ongoing genetic eye research. In addition, there was community consultation with hospital administration, local doctors, local optometrist and visiting ophthalmologists to meet all concerns regarding possible long-term impacts of the study.

6.3.2 Participant recruitment

The NIES consisted of conducting a population-based study on Norfolk Island, which was declared an external territory of Australia in 2007. Permanent residents aged ≥15 years were recruited by radio and newspaper advertisements, referral from healthcare providers, and word of mouth. In addition, subjects enrolled in the earlier NIHS
(2000) were sent invitations to participate in the NIES (Bellis, Hughes et al. 2005). There were no exclusion criteria other than non-permanent resident status.

6.3.3. Ophthalmic Evaluation

The examination protocol consisted of the following:

1) Visual acuity Log MAR at 3m (Precision Vision, LaSalle, IL, U.S.A).
2) Binocular vision function tests
   a) Cover test
   b) Four Dioptre base-out prism test
   c) Ocular motility
   d) Stereopsis, Lang stereo card (Lang, Forch, Switzerland)
3) Eye dominance
4) Anterior chamber depth, keratometry, axial length. IOLMaster (Carl Zeiss Meditek, Dublin, CA, U.S.A)
5) Conjunctival UV auto-fluorescence photos
6) Autorefraction, pre and post-cycloplegia (Nidek, Gamagori, Japan).
7) Intraocular pressure- Tonopen handheld tonometer (Reichert, Tustin, CA. U.S.A)
8) Central Corneal thickness Pachmate ultrasonic pachymeter (DGH, Exton, USA)
9) Dilated stereoscopic optic disc photographs Nidek 3-DX (Nidek, Gamagori, Japan)
10) Disc-centred and macular-centred retinal photographs Genesis D hand held retinal camera (Kowa, Tokyo, Japan)

11) Ocular Coherence Tomography (OCT) Cirrus (Carl Zeiss Meditek, Dublin, CA, U.S.A) Macular thickness, RNFL thickness and optic head components were used to assess three measures: fast RNFL, fast macula and fast optic nerve head

12) DNA collection for NIHS participants requiring replenished DNA stocks, and for new participants

After completing the examination, each participant was debriefed with examination results, and was provided with an opportunity to ask questions. Reports were mailed to eye care providers for patients previously seen by ophthalmologists or optometrists in Australia, New Zealand or on the island. Any newly diagnosed pathology such as suspicious discs or pressure for FDT visual field testing (Carl Zeiss, Dublin, CA, U.S.A.) was referred to the local optometrist. Baseline photographs were given to the local hospital and optometry clinic for participants' future use.

6.3.4 Pedigree Construction and Genotyping

Genealogical records confirm nine Isle of Man Bounty ship mutineers, twelve Tahitian women, and six Tahitian men (Hoare 1999) settled Pitcairn Island in 1790. Pedigree reconstruction confirmed current descendents possess lineages to all nine mutineers, six Tahitian women and three Caucasian sailors who joined the colony in the early 19th century (Bellis, Hughes et al. 2005; McEvoy, Zhao et al. 2009; Macgregor, Hewitt et al. 2010). Three hundred and seventy-seven individuals
indicate familial links to the 17 founders and were genotyped using a 400 STR panel of markers across the genome.

The complete Norfolk pedigree includes 11 meiosis and 6537 individuals dating back to the original Bounty mutineer and Tahitian founders. To facilitate heritability analyses, the large, complex pedigree (N=6,537) was trimmed (N=1,078) using the pedigree database management system PEDSYS (Texas Institute for Biomedical Research, San Antonio, Texas, USA) (Dyke 1996).

The complete pedigree was simplified to remove inbreeding loops from early generations by blanking the parents of all nonfounders in the first four generations of the trimmed pedigree. The Pedigree RElationship Statistical Test (PREST) was used to verify the pedigree structure and detect relationship misspecification (McPeek and Sun 2000). The trimmed or simplified pedigree structure comprises 1064 individuals, 372 of which have both phenotypic and genotypic data available. Genotyping was performed at AGRF high-throughput typing facility using the ABI Prism Linkage Mapping Set Version 2, a set of fluorescently labelled primer pairs from Perkin Elmer. The medium density set contains 400 microsatellite (STR) markers with an average heterozygosity level of 76% spaced at 10cM. Prior to statistical analysis, STR genotyped data was screened for errors and IBD sharing probabilities were generated using PEDSYS, MENDEL and SOLAR (v. 4.0.7). This information was used to validate the pedigree structure derived from questionnaires and genealogical records. Mendelian inconsistencies were resolved using PEDSYS–INFER (Dyke 1996) and Simwalk2 (Sobel, Papp et al. 2002). Chromosomal marker maps were obtained from the Marshfield Centre for Medical Genetics.
Cyto bands for markers were obtained from the University of California Santa Cruz (UCSC) Genome Browser (http://www.genome.ucsc.edu/cgi-bin/hgGateway). Norfolk Island Eye Study participants with phenotypic data were located within the complete NIHS pedigree using first name, surname, and date of birth. The glaucoma pedigree figure was created using R-Graphviz (Zhao 2006).

6.3.5 Statistical Analyses

Heritability analyses were conducted using variance components-based methodology implemented in the Sequential Oligonucleotide Linkage Analysis Routines (SOLAR) version 4.0.6 software package (Almasy and Blangero 1998). Heritability estimates were calculated as the ratio of the trait variance explained by additive polygenic effects to total phenotypic variance of the trait (Dyer, Blangero et al. 2001). The applied polygenic model assumes an infinite number of genetic factors, each with a small additive effect contributing to the trait variance. Estimates were screened for covariate effects of sex, age, age*sex, age^2, age^2*sex, height, ultraviolet autofluorescence, glaucoma status, visual acuity pinhole, sphere, cylinder, axis, kvalue-v, kvalue axis, central corneal thickness, anterior chamber, intraocular pressure, central corneal thickness, pterygium, eye color, visual acuity, axial length, outdoors, and kvalue-h. Glaucoma was treated as a continuous variable using "EnableDiscrete 0" to circumvent convergence errors due to discrete trait modelling in SOLAR. Covariates with p-values less than or equal to 0.10 were retained in the final model. SOLAR requires kurtosis to be <0.8 and standard deviation <0.5 in
order to proceed with analysis. Vertical cup:disc ratio values were multiplied by a factor of 5.2. Anterior chamber depth was normalised by log transforming and multiplying by three. SOLAR bivariate analysis was used to calculate genotypic, environmental, and phenotypic correlations between metrics. Genome-wide linkage mapping of adjusted eye metric traits was conducted with SOLAR using multipoint analysis.

6.4 Results

Eight hundred Norfolk Island residents participated in the study and received exams consisting of several ocular measurements. Of these, 495 participants were identified within the complete NIHS pedigree. Heritability estimates for ocular traits were calculated in this NIES subset. Characteristics of the NIES study population have been described (Mackey, Sherwin et al. 2011). Table 6.1 summarises trait distributions in the NIES population and NIHS subset. Trait means are similar in both groups. Anterior chamber depth was log transformed and multiplied by three to create a normal distribution. Multiplying by 5.2 increased C:D standard deviation. Intraocular pressure was log transformed and multiplied by 100. Figure 6.1 traces the lineage of a glaucoma participant back to a member of the founder population. Sixty individuals were diagnosed with glaucoma in the N=800 NIES study population. Of these, 22 individuals were identified in the N=495 NIES subset used in the analyses.
Phenotypic trait correlations are presented in Table 6.2. Highest correlation is between age and C:D ($r^2=0.27$) followed by IOP and CCT ($r^2=0.22$). A negative correlation resulted between ACD and age ($r^2=-0.19$).

Table 6.3 lists ocular trait heritability estimates and significant covariates. Glaucoma adjusted heritability was $h^2=0.20$ and insignificant ($p=0.14$). The only significant covariate was C:D ($p=1\times10^{-6}$) accounting for 0.13 of variation. Unadjusted estimate was considerably lower ($h^2=0.01$, $p=0.48$). Central corneal thickness was the most heritable measure with unadjusted $h^2=0.85$ ($p=1.5\times10^{-6}$). Adjusted heritability for CCT decreased to $h^2=0.77$ ($p=5.7\times10^{-6}$). Significant covariates were IOP, pterygium, and sphere, which adjusted the estimate to $h^2=0.85$ ($p=1.5\times10^{-6}$) (Table 6.3). Covariates explained 0.08 of CCT variability. Adjusted IOP heritability was 0.39 ($p=0.008$) with significant covariates ACD and ultraviolet autofluorescence (UVAF). Unadjusted IOP heritability was slightly lower ($h^2=0.38$, $p=0.007$) as covariates accounted for 0.01 of trait variation. Anterior chamber depth adjusted ($h^2=0.37$, $p=0.02$) and unadjusted ($h^2=0.35$, $p=0.02$) heritabilities were similar. Covariates age*sex, age^2, age^2*sex, IOP, pterygium accounted for 0.08 of trait variation. Unadjusted C:D heritability estimate $h^2=0.47$ ($p=0.001$) was considerably lower than adjusted value $h^2=0.69$ ($p=1\times10^{-5}$). Significant covariates were sex, age*sex, glaucoma, visual acuity pinhole, k-value, and pterygium. Covariates accounted for 0.26 of C:D trait variation.

Table 6.4 summarises results from bivariate analyses of traits. All genetic correlations were insignificant. Highest genetic association was between IOP and CCT ($\rho_e=0.43$, $p=0.13$). Significant environmental analyses were between glaucoma
and C:D ($\rho_\rho=0.66$, $p=1.6\times10^{-11}$), and CCT and C:D ($\rho_\rho=-0.69$, $p=0.04$). Significant phenotypic analyses were between IOP and CCT ($\rho_\rho=0.16$, $p=0.02$) and glaucoma and C:D ($\rho_\rho=0.36$, $p=4.7\times10^{-9}$).

Table 6.5 summarises LOD scores for ocular traits adjusted for covariates described in Table 6.3. Highest peak was LOD 2.5 for C:D on chromosome 2. Other C:D peaks >1 were found on chromosomes 1 (LOD 1.8), 9 (LOD 1.2) and 17 (LOD 1.1). Intraocular pressure maximum peak was on chromosome 15 (LOD 1.3), followed by peaks on chromosomes 22 (LOD 1.1) and 8 (LOD 1.2). Chromosomes 20 (LOD 1.9), 11 (LOD 1.7), 14 (LOD 1.2), 4 (LOD 1.3), and 3 (LOD 1.4) were mapped with CCT. Anterior chamber depth resulted in three peaks on chromosome 20 (LOD 1.8), 17 (LOD 1.3), and 14 (LOD 1.2), 3 (LOD 1.5) and 4 (LOD 1.5). Figures 6.2.1-4 summarise genome-wide scans by illustrating LOD peaks for traits on chromosomes 1-22. Figure 6.2.3 illustrates that C:D maximum peak on chromosome 2 (LOD 2.5) reached suggestive evidence for linkage.

6.5 Discussion

Overall, glaucoma endophenotype heritabilities are consistent with the literature; adjustment increased heritability for glaucoma, IOP, and C:D while slightly decreasing CCT. Adjusted IOP heritability ($h^2=0.38$) was within range ($h^2=0.29-0.42$) of estimates found in family studies (van Koolwijk, Despriet et al. 2007) (Chang and Stulting 2005) (Duggal, Klein et al. 2005) (Charlesworth, Kramer et al. 2010) (Levene, Workman et al. 1970) (Lee, Woo et al. 2010). Ultraviolet
autofluorescence was a significant covariate and may be a surrogate for exercise instead of signifying a true biological relationship. For CCT, adjusted heritability (h²=0.77) was higher than the published range for family studies (h²=0.65-0.72 (Alsbirk 1978) (Landers, Hewitt et al. 2009) (Charlesworth, Kramer et al. 2010). This difference may be due to the comprehensive list of covariates available in this study to adjust for confounders. Unadjusted CCT heritability (h²=0.85) is nearer to the twin range (h²=0.88-0.95) (Zheng, Ge et al. 2008) (Toh, Liew et al. 2005). Covariate contribution to explaining CCT was 0.08. As expected, IOP was a significant covariate for CCT. Controlling for CCT did not strengthen heritability of IOP as proposed by Charlesworth (Charlesworth, Dyer et al. 2005). Bivariate analysis further supported a genetic correlation between these traits (ρg=0.43 p=0.13) but results were not significant. However, the reverse relationship did not occur for IOP heritability. Vertical cup:disc ratio adjusted heritability (h²=0.65) was within the published range (h²=0.48-0.70) (van Koolwijk, Despriet et al. 2007) (Klein, Klein et al. 2004) (Chang, Congdon et al. 2005) (Charlesworth, Dyer et al. 2005).

For IOP, the highest LOD score occurred on chromosome 15 (LOD 1.3), where a peak was also found in the Beaver Dam Eye Study (Duggal, Klein et al. 2005). Linkage results for CCT did not replicate any previously published peaks. Although a previous study indicated linkage to chromosome 3 for IOP (Wirtz, Samples et al. 1997), in this study chromosome 3 was identified with CCT (LOD 1.4) and ACD (LOD 1.5). Linkage results for C:D show highest peaks at chromosomes 2 (2.5) and 1 (1.8), corresponding to peaks obtained for IOP in another study (Duggal, Klein et al. 2007). Linkage peaks on chromosomes 3, 4, and 14 occurred for both CCT and ACD. Adjusted ACD heritability (h²=0.37) was lower than the published range
(h^2=0.51-0.64) (Sorsby and Fraser 1964; Alsbirk 1975; Alsbirk 1977; Lyhne, Sjolie et al. 2001; Biino, Palmas et al. 2005; Dirani, Chamberlain et al. 2006; Chen, Scurrah et al. 2007; He, Huang et al. 2008; Tu, Yin et al. 2008; Klein, Suktitipat et al. 2009).

However, adjusting for covariates not included in other studies such as pterygium explained 0.08 of the variation.

6.6 Conclusions

Heritability of glaucoma endophenotypes in NI population was consistent with other populations. Bivariate analyses did not reveal any significant genetic relationships between covariates. Linkage scan indicates suggestive linkage of C:D on chromosome 2. In this large pedigree based study, we examined genetic determinants of eye diseases. In secondary analyses that used alternative residuals without adjustments for trait specific covariates, the heritability estimates were generally higher.
Figure 6.1 Representation of glaucoma ancestry

Individual (577) diagnosed with glaucoma; ancestry spanning nine generations to the Mutiny on the Bounty founders within the Norfolk Island pedigree.
### Table 6.1 Study population summary

<table>
<thead>
<tr>
<th>Glaucoma Trait</th>
<th>N NIES</th>
<th>Mean NIES</th>
<th>N NIHS</th>
<th>Mean NIHS</th>
<th>SD</th>
<th>Kurtosis</th>
<th>skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>802</td>
<td>52.1</td>
<td>496</td>
<td>52.1</td>
<td>17.5</td>
<td>-0.32</td>
<td>-0.33</td>
</tr>
<tr>
<td>Height</td>
<td>665</td>
<td>169.7</td>
<td>390</td>
<td>169.2</td>
<td>-0.15</td>
<td>0.06</td>
<td>-0.16</td>
</tr>
<tr>
<td>CCT</td>
<td>788</td>
<td>546.4</td>
<td>489</td>
<td>543.6</td>
<td>33.9</td>
<td>0.08</td>
<td>-0.15</td>
</tr>
<tr>
<td>ACD</td>
<td>789</td>
<td>3.3</td>
<td>489</td>
<td>3.3</td>
<td>0.4</td>
<td>1.02</td>
<td>0.47</td>
</tr>
<tr>
<td>IOP</td>
<td>799</td>
<td>16.1</td>
<td>15.9</td>
<td>15.9</td>
<td>3.25</td>
<td>0.93</td>
<td>0.46</td>
</tr>
<tr>
<td>C:D</td>
<td>782</td>
<td>0.4</td>
<td>0.41</td>
<td>0.41</td>
<td>0.20</td>
<td>-0.24</td>
<td>0.54</td>
</tr>
<tr>
<td>% female</td>
<td>54 (n=441)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 6.2 Correlations between endophenotypic traits

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>CCT</th>
<th>ACD</th>
<th>IOP</th>
<th>C:D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.00</td>
<td>-0.06</td>
<td>-0.20</td>
<td>-0.10</td>
<td>0.27</td>
</tr>
<tr>
<td>CCT</td>
<td>-0.06</td>
<td>1.00</td>
<td>-0.06</td>
<td>0.22</td>
<td>-0.10</td>
</tr>
<tr>
<td>ACD</td>
<td>-0.20</td>
<td>-0.06</td>
<td>1.00</td>
<td>-0.06</td>
<td>-0.07</td>
</tr>
<tr>
<td>IOP</td>
<td>-0.10</td>
<td>0.22</td>
<td>-0.06</td>
<td>1.00</td>
<td>0.05</td>
</tr>
<tr>
<td>C:D</td>
<td>0.27</td>
<td>-0.10</td>
<td>-0.07</td>
<td>0.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>
### Table 6.3 Heritabilities of endophenotypes

<table>
<thead>
<tr>
<th>Trait</th>
<th>Covariates p-value &lt;0.11</th>
<th>Proportion due to Covariates</th>
<th>$H^2_{\text{adj.}}$ (SE) p-value</th>
<th>$H^2_{\text{unadj.}}$ (SE) p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glaucoma</td>
<td>C:D p=1x10^{-6}</td>
<td>0.13</td>
<td>0.20 (0.20) p=0.14</td>
<td>0.007 (0.15) p=0.48</td>
</tr>
<tr>
<td>IOP</td>
<td>UVAF p=0.07</td>
<td>0.01</td>
<td>0.39 (0.17) p=0.008</td>
<td>0.38 (0.17) p=0.008</td>
</tr>
<tr>
<td></td>
<td>ACD p=0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCT</td>
<td>sphere p=0.0003</td>
<td>0.08</td>
<td>0.77 (0.15) p=5.7x10^{-6}</td>
<td>0.85 (0.14) p=1.5x10^{-6}</td>
</tr>
<tr>
<td></td>
<td>IOP p=0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pterygium p=0.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:D</td>
<td>sex p=0.05</td>
<td>0.26</td>
<td>0.69(0.14) p=6x10^{-5}</td>
<td>0.47(0.15) p=0.001</td>
</tr>
<tr>
<td></td>
<td>age*sex p=0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>glaucoma p=7x10^{-6}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>visual acuity pinhole p=0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>k-value p=0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pterygium p=0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACD</td>
<td>age*sex p=0.07</td>
<td>0.08</td>
<td>0.37(0.18) p=0.02</td>
<td>0.35(0.18) p=0.02</td>
</tr>
<tr>
<td></td>
<td>age^2 p=8.2x10^{-5}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>age^2*sex p=0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IOP p=0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pterygium p=0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6.4 Bivariate analysis of endophenotypes

<table>
<thead>
<tr>
<th>Trait 1</th>
<th>Trait 2</th>
<th>$\rho_t$</th>
<th>P</th>
<th>$\rho_e$</th>
<th>P</th>
<th>$\rho_p$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glaucoma</td>
<td>IOP</td>
<td>0.19 (0.53)</td>
<td>0.74</td>
<td>0.12 (0.17)</td>
<td>0.51</td>
<td>0.13</td>
<td><strong>0.04</strong></td>
</tr>
<tr>
<td>Glaucoma</td>
<td>CCT</td>
<td>-0.06 (0.35)</td>
<td>0.87</td>
<td>-0.03 (0.30)</td>
<td>0.91</td>
<td>-0.03</td>
<td>0.58</td>
</tr>
<tr>
<td>IOP</td>
<td>C:D</td>
<td>-0.05 (0.30)</td>
<td>0.86</td>
<td>0.24 (0.29)</td>
<td>0.40</td>
<td>0.12</td>
<td>0.07</td>
</tr>
<tr>
<td>IOP</td>
<td>CCT</td>
<td>0.43 (0.25)</td>
<td>0.13</td>
<td>-0.14 (0.34)</td>
<td>0.66</td>
<td>0.16</td>
<td><strong>0.02</strong></td>
</tr>
<tr>
<td>CCT</td>
<td>C:D</td>
<td>0.23 (0.20)</td>
<td>0.23</td>
<td>-0.69 (0.48)</td>
<td><strong>0.04</strong></td>
<td>-0.05</td>
<td>0.48</td>
</tr>
<tr>
<td>Glaucoma</td>
<td>C:D</td>
<td>-1.00 (na)</td>
<td>0.34</td>
<td>0.66 (0.12)</td>
<td><strong>1.6 \times 10^{-11}</strong></td>
<td>0.36</td>
<td><strong>4.7 \times 10^{-9}</strong></td>
</tr>
<tr>
<td>ACD</td>
<td>IOP</td>
<td>0.35 (0.5)</td>
<td>0.46</td>
<td>-0.26 (0.16)</td>
<td>0.13</td>
<td>-0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>ACD</td>
<td>CCT</td>
<td>-0.14 (0.26)</td>
<td>0.59</td>
<td>-0.03 (0.36)</td>
<td>0.94</td>
<td>-0.08</td>
<td>0.21</td>
</tr>
<tr>
<td>ACD</td>
<td>C:D</td>
<td>-0.15 (0.30)</td>
<td>0.30</td>
<td>-0.03 (0.2)</td>
<td>0.87</td>
<td>-0.08</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Table 6.5: Summary of LOD scores for linkage mapping of endophenotypes

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Map Distance (cM)</th>
<th>Multivariable Adjusted LOD Score &gt;1*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glaucoma</td>
<td>Na</td>
<td>Na</td>
</tr>
<tr>
<td>IOP</td>
<td>15(116)</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>22(51)</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>8(60)</td>
<td>1.2</td>
</tr>
<tr>
<td>CCT</td>
<td>20(15)</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>11(40-58)</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>14(59)</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>4(102)</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>3(55)</td>
<td>1.4</td>
</tr>
<tr>
<td>C:D</td>
<td>2(244)</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>1(172)</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>9(82)</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>17(95)</td>
<td>1.1</td>
</tr>
<tr>
<td>ACD</td>
<td>20(61)</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>17(66)</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>14(162)</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>3(100)</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>4(32)</td>
<td>1.5</td>
</tr>
</tbody>
</table>

*LOD scores were calculated from residuals from multivariable models. For covariates included in the models, see Table 6.3.
Figure 6.2.1 Intraocular pressure LOD graph for chromosomes 1-22
Figure 6.2.2 Anterior chamber depth LOD graph for chromosomes 1-22
Figure 6.2.3 Vertical cup-to-disc ratio LOD graph for chromosomes 1-22
Figure 6.2.4 Central corneal thickness LOD graph for chromosomes 1-22
Chapter 7: Genome-wide Linkage Scan of Ultraviolet Radiation Exposure Exposure Maps to Myopia Related Loci in the Norfolk Island Isolate

7.1 Background

Currently, over 80% of young adults in East Asian cities have myopia (nearsightedness) (Wu, Seet et al. 2001; He, Zeng et al. 2004; Lin, Shih et al. 2004), while in the US prevalence has doubled to 40% in 30 years (Vitale, Sperduto et al. 2009) suggesting strong environmental influences. Myopic trait heritability varies, ranging from high (Chen, Scurrah et al. 2007; Klein, Suktitipat et al. 2009) in urbanised settings to low (Biino, Palmas et al. 2005; Vitart, Bencic et al. 2010) in island populations, further reflecting environmental contributions (Young, Leary et al. 1969; Morgan, Speakman et al. 1975; Lin, Hung et al. 1988; Visscher, Hill et al. 2008). Recently, more time spent outdoors has emerged as a protective factor against myopia (Mutti 2010) while risk factors such as education level (Tay, Au Eong et al. 1992) and near work (Saw, Chan et al. 2008), coincide with more time indoors. Hence, variability in UV light exposure among these environmental factors is an underlying factor that may also play a role in myopia (Prepas 2008). We created a quantitative trait for UV light exposure by measuring conjunctival UV autofluorescence, a method used in clinical settings to detect sun damage in skin (Asawanonda and Taylor 1999; Sandby-Moller, Thieden et al. 2004) and eyes (Ooi,
Sharma et al. 2006; Ooi, Sharma et al. 2007). We then conducted heritability analyses and genome wide linkage mapping in the Norfolk Island population isolate.

### 7.2 Methods

#### 7.2.1 Study Population

Genetic studies conducted in population isolates have increased power to detect heritable effects bestowing greater potential to identify underlying causal quantitative trait loci with less environmental confounding of genetic effects because multiple households are included within pedigrees (Blangero, Williams et al. 2003).

#### 7.2.2 Ethics

The original Norfolk Island Health Study (NIHS) received ethics approval from the Human Research and Ethics Committee at Griffith University. This same committee, in addition to the Human Research and Ethics Committee at the Royal Victorian Eye and Ear Hospital in Melbourne, approved the Norfolk Island Eye Study (NIES). Consent was obtained to conduct the ophthalmic examination and link this with the earlier cardiovascular and genetic research as well as ongoing genetic eye research. In addition, community consultation took place with hospital administration, local doctors, local optometrist and visiting ophthalmologists to meet all concerns regarding possible long-term impact of the study.
7.2.3 Participant Recruitment

The NIES research consisted of conducting a population-based study on Norfolk Island, which was declared an external territory of Australia in 2007. Permanent residents aged ≥15 years were recruited by radio and newspaper advertisements, referral from healthcare providers, and word of mouth. In addition, subjects enrolled in the NIHS were sent invitations to participate in the NIES (Bellis, Hughes et al. 2005). There were no exclusion criteria other than non-permanent resident status.

7.2.4 Conjunctival Ultraviolet Autofluorescence

Conjunctival ultraviolet autofluorescence (UVAF) photographs were taken for 641/781 (82.1%) of participants in the NIES comprising 61% of residents aged >15 years, and nearly 75% of those were >50 years. Ultraviolet photographic methods have been published elsewhere (Ooi, Sharma et al. 2006; Ooi, Sharma et al. 2007). In brief, photographs were taken using both reflected visible light (control) and UV-induced fluorescence. Each system consisted of height-adjustable table equipped with subject head-rest, camera positioning assembly, digital single-lens reflex camera, macro lens, and filtered electronic flash. Nasal and temporal regions of each eye were photographed at 0.94 magnification. Coloured low-voltage light emitting diodes were positioned on stands in the subject’s visual field 35° to the camera–subject axis to aid fixation. The UV-induced fluorescence images were photographed using an electronic flash system with UV-transmission filters (300-400nm) as the
excitation source. Subject fluorescence was recorded with a Nikon D100 (Nikon, Melville, New York, USA) digital camera and 105 mm f/2.8 Micro Nikon lens fitted with infrared and UV barrier filters. Thus, the camera recorded only fluorescence. Images were saved in RGB format at the D100 settings of JPEG fine (1:4 compression) and large resolution. Adobe Photoshop (Adobe Systems Inc, San Jose, California, USA) was used to perform the quantitative analysis. Four photos were analysed per person (right and left nasal/temporal). Settings required for the UV analysis were pixel length = 3008 and logical length = 2.4. Results were expressed in mm².

**7.2.5 Pedigree Construction**

Genealogical records confirm nine Isle of Man Bounty ship mutineers, twelve Tahitian women, and six Tahitian men (Hoare 1999) settled Pitcairn Island in 1790. Pedigree reconstruction confirmed current descendent possess lineages to all nine mutineers, six Tahitian women and three Caucasian sailors who joined the colony in the early 19th century (Bellis, Hughes et al. 2005; McEvoy, Zhao et al. 2009; Macgregor, Hewitt et al. 2010). The complete Norfolk pedigree includes 11 meiosis and 6537 individuals. The complete pedigree was simplified by removing inbreeding loops from early generations by blanking the parents of all nonfounders in the first four generations of the trimmed pedigree. The large complex pedigree (N=6,537) was then trimmed (N=1,078) using the pedigree database management system PEDSYS (Texas Institute for Biomedical Research, San Antonio, Texas, USA) (Dyke 1996). The Pedigree RELationship Statistical Test (PREST) was used to verify the
pedigree structure and detect relationship misspecification (McPeek and Sun 2000). The trimmed pedigree structure comprises 1064 individuals, 372 of which were genotyped for genetic analyses. Blood samples were collected (20ml) and sent to the Queensland Medical Laboratory (Southport, Australia) for plasma biochemical analysis, then stored at the Genomics Research Centre at -80°C.

7.2.6 DNA Extraction and Genotyping

DNA was extracted following a standard salting out protocol (Miller et al., 1988). After extraction, DNA was dissolved in 1.8mL of TE buffer (1M Tris-HCL, pH8.0; 0.5M EDTA, pH8.0) and stored at 4°C. Prior to PCR, DNA was precipitated with ethanol to remove residual EDTA. Re-precipitation of stock DNA samples required 100μL of DNA dissolved in TE buffer. One hundred microlitres of DNA was placed in 2mL tubes containing 200μL of chilled absolute ethanol and 10μL of 3M sodium acetate. Samples were inverted and frozen using liquid nitrogen. Samples were centrifuged for 15 minutes at 10,000rpm at 4°C. The supernatant was removed and the DNA pellet was dried using a savant speed vacuum for 2-minute intervals. The pellet was resuspended in 100μL of sterile milli Q water and incubated overnight at 37°C. All DNA samples were quantified using a Nanodrop™ spectrophotometer and re-eluted to a standard concentration of 20ng/μL.

Genotyping was performed at AGRF high-throughput typing facility using the ABI Prism Linkage Mapping Set Version 2, and a set of fluorescently labelled primer pairs from Perkin Elmer. The medium density set contains 400 microsatellite (STR)
markers with an average heterozygosity level of 76% spaced at 10cM. Prior to statistical analysis STR genotyped data was screened for errors and identity-by-descent sharing probabilities were generated using PEDSYS, MENDEL and SOLAR (v. 4.0.7). This information was used to validate the pedigree structure derived from questionnaires and genealogical records. Mendelian inconsistencies were resolved using PEDSYS -INFER (Dyke 1996) and Simwalk2 (Sobel, Papp et al. 2002). Chromosomal marker maps were obtained from the Marshfield Centre for Medical Genetics (http://research.marshfieldclinic.org/genetics). Cytobands for markers were obtained from the University of California Santa Cruz (UCSC) Genome Browser (http://www.genome.ucsc.edu/cgi-bin/hgGateway). Phenotyped NIES participants were located within the complete NIHS pedigree using first name, surname, and date of birth then assigned corresponding NIHS unique identifiers.

7.2.7 Statistical Analysis: Heritability and Linkage

UVAF values were log transformed using a Box-Cox transformation prior to analyses. Heritability analyses were conducted using variance components-based methodology implemented in the Sequential Oligonucleotide Linkage Analysis Routines (SOLAR) version 4.0.6 software package (Almasy and Blangero 1998). Heritability estimates were calculated as the ratio of the trait variance explained by additive polygenic effects to total phenotypic variance of the trait (Dyer, Blangero et al. 2001). The applied polygenic model assumes an infinite number of genetic factors, each with a small additive effect contributing to the trait variance. Heritability analyses included screening for covariate effects of sex, age, age*sex, age^2,
age^2*sex, height, glaucoma status, visual acuity pinhole, sphere, cylinder, axis, kvalue-v, kvalue axis, central corneal thickness, anterior chamber, intraocular pressure, central corneal thickness, pterygium, eye color, visual acuity, axial length, outdoors, and kvalue-h. Covariates with p-values less than or equal to 0.10 were retained in the final model. SOLAR requires kurtosis to be <0.8 and standard deviation <0.5 in order to proceed with the analysis. Genomewide linkage mapping of UVAF was conducted with SOLAR using multipoint analysis.

7.3 Results

7.3.1 Sample Population

Approximately half the NIES study population could trace their ancestry to the HMS Bounty founding population (Table 7.1). Mean age of all study participants was 54.1 years, ranging from 15-89 years. Mean age was lower in the Pitcairn ancestry group (51.7 vs 57.4 years, P=0.0002). There were no statistically significant differences in hypertension rates (17.9% vs. 15.2%, P=0.355), smoking (59.1% vs. 52.6%, P=0.098), diabetes (Type 1 or 2) (3.8% vs. 4.9%, P=0.991) or history of cataract or cataract surgery in either eye (25.2% vs. 19.8%, P=0.105). Pitcairn Island ancestors spent a greater proportion of their lifespan on Norfolk Island (71.9% vs. 52.9%, P<0.001).
7.3.2 Heritability Analyses

Four hundred and ninety-five individuals from the NIES with UVAF phenotypic information were located in the NIHS complete pedigree. Of these, SOLAR used phenotypic information from 166 individuals to calculate UVAF heritability estimates. Prior to analysis, UVAF was log transformed to normalise the distribution and remove kurtosis.

The heritability estimate for UVAF, calculated as the contribution of genetic factors to the total phenotypic variance was $h^2=0.73$ (p=0.003). Significant covariates were age$^2$, height, glaucoma status, pterygium status, visual acuity, visual acuity through pinhole, cylinder, axis, anterior chamber depth, which accounted for 0.27 of the total phenotypic variance of UVAF. Sex was a significant covariate until height was adjusted. Also, eye color was significant prior to adding glaucoma status to the model.

7.3.3 Linkage

Genome-wide linkage results are shown in Figure 7.1.1 with the solid line indicating the suggestive (LOD 2.0) evidence of linkage threshold. Figure 7.1.2 shows chromosome 7 with the largest LOD score of 4.06 occurring at marker D7S661. A second significant peak (LOD 3.1) occurs near marker D7S530. The entire coupled peak region spans 105-180 cM corresponding to chromosome 7q22-36. Figure 7.1.3 represents the third highest genome-wide peak on chromosome 12q24 (LOD 2.5)
between D12S324-DS12S1659 (145-160cM). The entire peak region spans markers D12S351-D12S1723 (105-180cM).

7.4 Discussion

This study concluded that conjunctival UV autofluorescence is a highly heritable trait ($h^2=0.73$, $p=0.003$). Regions of significant linkage at high myopia locus 7q35 (Naiglin, Gazagne et al. 2002; Klein, Duggal et al. 2011) (LOD 4.06) were discovered and suggestive evidence at ocular refraction locus 12q24 (Wojciechowski, Stambolian et al. 2009) (LOD 2.5) was also indicated.

Chromosome 7q36 has been identified in two high myopia family studies (Naiglin, Gazagne et al. 2002; Klein, Duggal et al. 2011). The highest peak (D7S661 LOD 4.06) in our study located at 7q35 is approximately 30cM nearer to the centromere than the highest peak (LOD 2.81) reported in the genome-wide scan for familial high myopia conducted in 21 French and two Algerian families (Naiglin, Gazagne et al. 2002). In the Beaver Dam Eye Study, linkage with the refraction phenotype overlapped the Naiglin et al peak and the UVAF region with the highest peak occurring within the MYP4 locus at rs2536077 (164.1 cM) (Klein, Duggal et al. 2011). Hence, peaks from all three studies occur within the 75cM wide peak range (Figure 7.1.2) flanked between markers D7S515-D7S636 (7q35-36.1).
Chromosome 12q has been implicated in several familial myopia studies (Young, Ronan et al. 1998; Nurnberg, Jacobi et al. 2008; Li, Guggenheim et al. 2009; Wojciechowski, Stambolian et al. 2009). In an international collaboration involving 254 families (Li, Guggenheim et al. 2009), sphere measurements mapped to 12q21.2-24.12 (89.57-126.16 cM), a 36.59 cM spread within the MYP3 locus, overlapping with results from this study (Figure 7.1.3). This study shows slight shifts in peaks between Asians and Caucasians.

Another relevant study consisted of a linkage meta-analysis of refraction from Caucasians and Amish within the Myopia Family Study (Wojciechowski, Stambolian et al. 2009). In the Caucasian population, the maximum LOD score was 4.58 (Ppw=0.00037 Pgw=0.13) at 142cM, only 20cM from the maximum peak on chromosome 12 in this study (Figure 7.1.3). Interestingly, a smaller peak was also detected at 44cM (LOD=2.21, Ppw=0.0053), which corresponds to the 3rd highest peak on chromosome 12 found in the UVAF analysis (30-55cM peak at LOD= 0.75) (Figure 7.1.3).

A large Italian-German family study identified the MYP3 locus for autosomal dominant high myopia within a 30cM critical region (LOD 3.85) between D12S1684-D12S1605 (Marshfield 86.4-116.66cM) (Young, Ronan et al. 1998), 38cM away from the Caucasian study (Wojciechowski, Congdon et al. 2005). Results from Mutti et al (Mutti, Cooper et al. 2007) reinforced linkage to this locus in a study of mild myopia in an ethnically diverse sample of American school children.
The MYP3 region contains genes coding for the proteoglycans lumican (LUM), decorin, the dermatan sulfate proteoglycan 3 precursor (DSPG3), and fibromodulin (FMOD) (Wojciechowski 2011). These proteoglycans regulate collagen fibril assembly and are expressed in ocular tissues. A case–control study of Taiwanese individuals (Zhang, Zhu et al. 2009) found that a single nucleotide polymorphism in the 5' promoter region of the lumican gene was significantly associated with high myopia. Paluru et al (Paluru, Scavello et al. 2004) screened the same family used to identify the MYP3 locus for sequence alterations in the LUM and FMOD genes and found no relationship between affection status and polymorphisms within these genes. However, their analysis did not include alterations in the 5' promoter region of the lumican gene, which may have been informative (Wojciechowski, Congdon et al. 2005).

Experimental studies show UV light causes collagen crosslinking, which prevents eye elongation and blocks myopia progression (Friedlaender 2005; Wollensak, Iomdina et al. 2005; Wollensak and Iomdina 2009). Eye elongation is a factor that determines ocular refraction. The UVAF measurements are thought to quantify collagen crosslinking (Asawanonda and Taylor 1999; Sandby-Moller, Thieden et al. 2004), which supports the experimental model. Our linkage signal may correspond to the MYP3 locus and/or to the chromosome 12 proteoglycan gene complex regulating collagen fibril assembly.

Studies indicate that chromosome 12 is one of 2 chromosomes with ethnic differences in inversion rates between Africans and Caucasians, Asians, Hispanics (Jorgenson, Tang et al. 2005). Although no chromosome comparisons have been made with
Polynesians, this slight shift in MYP3 peak across studies may reflect admixture differences between study populations.

7.5 Conclusion

In this study, we discovered that conjunctival UV autofluorescence is a highly heritable trait ($h^2=0.73$, $p=0.003$). We also identified regions of significant linkage at high myopia locus 7q35 (Naiglin, Gazagne et al. 2002; Klein, Duggal et al. 2011) (LOD 4.06) and suggestive evidence at ocular refraction locus 12q24 (Wojciechowski, Stambolian et al. 2009) (LOD 2.5). Ultraviolet autofluorescence is the first phenotype to identify both MYP3 and MYP4 loci in the same genome-wide linkage scan suggesting UV radiation may play a role across heterogenic forms of myopia.

Hence, these results provide genetic evidence for gene-environment interactions between UV light and myopia progression, reinforcing experimental and observational studies, and warranting further research into a causative biological mechanism of action. Given previous work and the biological plausibility of a role for UVAF in myopia, these two chromosomal regions should be investigated further using fine-mapping techniques.
### Table 7.1 Demographic comparison of Bounty and non-Bounty descended Norfolk Island residents

<table>
<thead>
<tr>
<th>Category</th>
<th>Bounty</th>
<th>Non-Bounty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Overall</td>
<td>318</td>
<td>49.6</td>
</tr>
<tr>
<td>Female</td>
<td>175</td>
<td>55</td>
</tr>
<tr>
<td>Male</td>
<td>143</td>
<td>45</td>
</tr>
<tr>
<td>Age 15-49</td>
<td>143</td>
<td>45</td>
</tr>
<tr>
<td>&gt;=50</td>
<td>175</td>
<td>55</td>
</tr>
<tr>
<td>Lifetime on Norfolk Island</td>
<td>42</td>
<td>13.2</td>
</tr>
<tr>
<td>UVAF mm²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;=14.4</td>
<td>57</td>
<td>17.9</td>
</tr>
<tr>
<td>14.5 – 28.0</td>
<td>72</td>
<td>22.6</td>
</tr>
<tr>
<td>28.1 – 47.7</td>
<td>84</td>
<td>26.4</td>
</tr>
<tr>
<td>&gt;=47.8</td>
<td>105</td>
<td>33</td>
</tr>
</tbody>
</table>
Figure 7.1.1 Ultraviolet exposure LOD scores across chromosomes 1-22
Figure 7.1.2 LOD scores for ultraviolet fluorescence on chromosome 7
Figure 7.1.3 LOD scores for ultraviolet fluorescence on chromosome 12
Chapter 8: Conclusions

8.1 Overview

Norfolk Island (34.6km$^2$) is situated 1600km northeast of Sydney, on the Norfolk Ridge that runs from New Zealand to New Caledonia. Norfolk Island contains a founder population descended from 11 British Bounty Mutineers and 6 Tahitian women, who colonised nearby Pitcairn Island in 1790 (Macgregor, Bellis et al. 2010). In 1856, the small community of 193 relocated to Norfolk Island (then uninhabited) when population growth became unsustainable on Pitcairn (Hoare 1999). The present day Norfolk Islanders have maintained a relatively homogeneous lifestyle due to geographic isolation, strict quarantine and immigration laws, and community centred culture. Given the unique population history and presence of an extended genealogy with unbroken lineage to the original Bounty mutineer founders, this population is of particular interest for studies of complex disease. In an effort to localise susceptibility loci the research described in this thesis involved genome wide molecular genetic analysis of two complex phenotypes, CVD risk and migraine in the Norfolk Island Bounty mutineer pedigree.
8.2 Framingham Risk Score

8.2.1 Conclusions

Framingham risk score (FRS) models are able to predict 10-year risk of a cardiac event with respectable accuracy (Brindle, Emberson et al. 2003; Havranek 2007). The purpose of this study was to determine whether the FRS could be used for gene mapping in linkage analyses as a new approach to cardiovascular disease (CVD) gene discovery. The rationale was since this composite score is reliable at predicting CVD outcomes it may also be useful in identifying genetic contributions to this multigenic disease. The research plan was to calculate the FRS for individuals within the Norfolk Island pedigree, estimate FRS heritability, and then conduct linkage mapping of this trait to identify putative cardiovascular disease loci.

Although a composite risk score based on several risk factors is more effective at predicting CVD outcomes than single risk factors, the FRS was not heritable in the Norfolk Island isolate and did not identify any chromosome regions involved in regulating the complex mechanisms contributing to this multigenic disease. A possible explanation may be that age, sex, and smoking are major contributing factors to the FRS formula yet are not genetic based or heritable, diluting the contribution of heritable factors comprising the score.

Bivariate analyses of genetic associations between individual traits were insignificant but suggest relationships between low and high density lipids and blood pressure.
More sampled individuals are needed to confirm significance of these relationships and explore bivariate linkage mapping with these traits.

8.2.2 Future Directions

Future work on this topic may include estimating heritabilities for other composite cardiovascular risk scores comprised of only individually heritable, continuous traits. In 2009-10 the Genomics Research Centre initiated a second wave of Norfolk Island Health Study participant recruitment on Norfolk Island. As with the 2000 study, signed informed consent, medical and genealogical questionnaire responses, EDTA venous blood specimens for biochemical measures and DNA extraction were obtained from volunteers. Calculating a second FRS estimate for the same participants using this follow-up data would enable an intra-individual comparison between two sets of scores and provide an interesting comparison of Framingham risk scores.

8.3 Pterygia

8.3.1 Conclusions

Pterygia is a common disorder of the ocular surface known to be strongly associated with sun exposure. One aim of this study was to examine differences in pterygia prevalence among Norfolk Island residents. Another aim was to estimate pterygia
heritability in the Norfolk Island population isolate.

Pterygia heritability did not differ significantly from a model of $h^2 = 0.024$ (SE 0.16, p=0.44). Significant covariates were visual acuity with pinhole (p=0.07), UVAF (p=0.01) and axis (p=0.03) while age and sex were insignificant. The proportion of pterygia variability attributable to covariates was estimated as 9.0% (Kullback-Leibler $r^2$). The association with cylinder axis can be explained since pterygium occurs at the $180^\circ$ (or $0^\circ$) end of the axis and will distort the cornea. Although male gender (Wong, Foster et al. 2001) has previously been found to be a risk factor for pterygium development, sex was not a significant covariate in the heritability analysis. An unequal variance Welch's test indicated that UVAF measurements were significantly lower in females (M=0.66, SD=0.83) than in males (M=0.80, SD=1.00), F (1,430)=9.35, p=0.0024, which may explain the gender differences in pterygium risk. Hence, the UVAF covariate may account for the gender differences.

We also performed a cross-sectional survey of ocular metrics on Norfolk Island to elucidate pterygium prevalence and UVR exposure in this population. We found that the Pitcairn Island descendants had a doubled risk of developing pterygium compared to other residents. Moreover, Pitcairn descendants had larger UVAF measurements and were twice as likely to spend more than 3/4 of their day outside compared to other residents. A combined increase in both pterygium prevalence and median UVAF further supports the role for UVR in pterygium etiology.
8.3.2 Future Directions

Future work with pterygia may include increasing the number key individuals sampled within the pedigree to increase significance of the heritability estimate. Other work may involve creating a continuous phenotype based on grading pterygia levels to determine whether more variability in the trait values would improve heritability estimates. Further work elucidating pterygia genetics would also benefit from cell and animal models involving UV radiation and gene expression.

Pterygia may be used as a biomarker for other diseases that have a strong correlation with UVR. UVR exposure is strongly associated with eye disease, both acute (acute photokeratitis or flash burns) and chronic (pterygium, climatic droplet keratopathy, cataract, ocular surface squamous neoplasia, and eyelid tumours; basal and squamous cell carcinoma and melanoma) (Oliva and Taylor 2005). Therefore, documenting increased risk of developing UVR-related disease, whether due to genetics or lifestyle, has important public health ramifications. At present, public health experts advocate that sun exposure should be guided by level of ambient UVR and skin type (Lucas, McMichael et al. 2008). Hence, knowledge of an individual’s pterygia status or UV response may lead to targeted prevention and therapeutic strategies.
8.4 Glaucoma

8.4.1 Conclusions

The goal of this study was to investigate heritability and linkage mapping of quantitative ocular traits intraocular pressure (IOP), central corneal thickness (CCT), vertical cup-to-disc ratio (C:D), and anterior chamber depth (ACD) in the Norfolk Island population isolate. Hence, heritability estimates were determined and subsequent linkage mapping was performed. This work combines pedigree and genetic linkage resources from the Norfolk Island Health Study (NIHS) (Macgregor, Bellis et al. 2010; Bellis, Hughes et al. 2005; Bellis, Cox et al. 2008; Bellis, Cox et al. 2008; Cox, Bellis et al. 2009) with phenotypic data from the Norfolk Island Eye Study (NIES), which was designed to characterise epidemiology, genetics, and environmental contributions to eye disease in this population (Mackey, Sherwin et al. 2011) (Sherwin, Kearns et al. 2011).

Overall, glaucoma endophenotype heritabilities are consistent with the literature. Covariate adjustment increased heritability estimates for glaucoma and ACD, decreased CCT estimates and did not effect IOP calculations. Intraocular pressure heritability ($h^2=0.38$) was within the range of estimates found in family studies ($h^2=0.29-0.42$) (van Koolwijk, Despriet et al. 2007) (Chang and Stulting 2005).
Ultrap violet autofluorescence was a significant covariate and may be a surrogate for exercise. For CCT, adjusted heritability ($h^2 = 0.77$) was higher than the family studies range ($h^2 = 0.65\text{-}0.72$) (Alsbirk 1978) (Landers, Hewitt et al. 2009) (Charlesworth, Kramer et al. 2010). This difference may be due to the comprehensive list of covariates used for adjustment. Unadjusted CCT ($h^2 = 0.85$) is nearer the twin range ($h^2 = 0.88\text{-}0.95$) (Zheng, Ge et al. 2008) (Toh, Liew et al. 2005). Covariate contribution to explaining CCT was 0.08. As expected, IOP was a significant covariate for CCT. Controlling for CCT did not strengthen heritability or linkage of IOP to chromosome 10 as proposed by Charlesworth (Charlesworth, Dyer et al. 2005). Bivariate analysis further supported a genetic correlation between these traits ($\rho_g = 0.43$, $p = 0.13$) but results were not significant. However, the reverse relationship was not found for IOP heritability. Vertical C:D ratio adjusted heritability ($h^2 = 0.65$) was within the published range ($h^2 = 0.48\text{-}0.70$) (van Koolwijk, Despriet et al. 2007) (Klein, Klein et al. 2004) (Chang, Congdon et al. 2005) (Charlesworth, Dyer et al. 2005).

For IOP, the highest LOD score on chromosome 15 (LOD 1.3) was also found in the Beaver Dam Eye Study (Duggal, Klein et al. 2005). Linkage results for CCT did not replicate any previously published peaks. Although previous study has shown linkage to chromosome 3 for IOP (Wirtz, Samples et al. 1997), in this study it has been identified with CCT (LOD1.4) and ACD (LOD1.5). Linkage results for C:D show highest peaks at chromosomes 2 (LOD 1.9) and 12 (LOD 2.6), corresponding to peaks obtained for IOP in another study (Duggal, Klein et al. 2007). Linkage peaks on chromosomes 3, 4, and 14 were found for both CCT and ACD. Adjusted ACD
heritability $h^2=0.37$ was lower than the published range ($h^2=0.51-0.64$) (Sorsby and Fraser 1964; Alsbirk 1975; Alsbirk 1977; Lyhne, Sjolie et al. 2001; Biino, Palmas et al. 2005; Dirani, Chamberlain et al. 2006; Chen, Scurrah et al. 2007; He, Huang et al. 2008; Tu, Yin et al. 2008; Klein, Suktitipat et al. 2009). However, adjusting for covariates including pterygium explained 0.08 of the variation, which may account for this lower heritability.

Heritability of glaucoma endophenotypes in NI population was consistent with other populations. Bivariate analyses did not reveal any significant genetic relationships between covariates. Linkage scan indicates suggestive linkage of C:D on chromosome 2. In secondary analyses that used alternative residuals without adjustments for trait specific covariates, the heritability estimates were generally higher and the linkage peaks from the primary analyses were reproduced.

**8.4.2 Future Directions**

More ocular endophenotypes will be created based on residuals to see whether overlapping peaks may be strengthened. Also, it will be useful to modify existing endophenotypic traits to reflect more of a glaucoma phenotype, values such as maximum IOP, maximum C:D, lowest ACD, and lowest CCT may reinforce existing peaks and identify potential new loci involved in glaucoma. Also, peak regions with LOD $>3$ should be fine mapped to further localise putative glaucoma genes.
Moreover, cardiovascular disease risk factors and migraine, which were studied as part of the NIHS, are known to be associated with glaucoma and data could be crosslinked to allow broader analyses (Mackey, Sherwin et al. 2011).

8.5 Myopia

8.5.1 Conclusions

We created a quantitative trait for ultraviolet (UV) radiation exposure by measuring conjunctival UV autofluorescence (UVAF), a method used in clinical settings to detect sun damage in skin (Asawanonda and Taylor 1999; Sandby-Moller, Thieden et al. 2004) and eyes (Ooi, Sharma et al. 2006; Ooi, Sharma et al. 2007). We then conducted heritability analyses and genome wide linkage mapping in the Norfolk Island population isolate.

In this study we discovered that conjunctival UVAF is a highly heritable trait ($h^2=0.73$ $p=0.003$). We also identified regions of significant linkage at high myopia locus 7q35(Naiglin, Gazagne et al. 2002; Klein, Duggal et al. 2011) (LOD 4.06) and suggestive evidence at ocular refraction locus 12q24(Wojciechowski, Stambolian et al. 2009)(LOD 2.5).
Chromosome 7q36 has been identified in two high myopia family studies (Naiglin, Gazagne et al. 2002; Klein, Duggal et al. 2011). The highest peak (LOD 4.06) in our study located at 7q35 is approximately 30cM nearer to the centromere than the highest peak (LOD 2.81) reported in the genome wide scan for familial high myopia conducted in 21 French and two Algerian families (Naiglin, Gazagne et al. 2002). In the Beaver Dam Eye Study, linkage with the refraction phenotype overlapped Naiglin and the UVAF region with the highest peak occurring within the MYP locus (Klein, Duggal et al. 2011). Hence, peaks from all three studies occur within the 75cM wide peak range.

Chromosome 12q has been implicated in several familial myopia studies (Young, Ronan et al. 1998; Nurnberg, Jacobi et al. 2008; Li, Guggenheim et al. 2009; Wojciechowski, Stambolian et al. 2009). An international collaboration involving 254 families (Li, Guggenheim et al. 2009), sphere measurements mapped within a 36.6cM spread of the MYP3 locus, overlapping with results from the UVAF study. This study shows slight shifts in peaks between Asians and Caucasians.

Another relevant study conducted a linkage meta-analysis of refraction from Caucasians and Amish within the Myopia Family Study (Wojciechowski, Stambolian et al. 2009). In the Caucasian population, the maximum LOD score was only 20cM from the maximum peak on chromosome 12 in this study. Interestingly, a smaller peak was also detected at 44cM, which corresponds to the third highest peak on chromosome 12 found in the UVAF analysis.
A large Italian-German family study identified the MYP3 locus for autosomal dominant high myopia (Young, Ronan et al. 1998), 38cM away from the Caucasian study (Wojciechowski, Congdon et al. 2005). Results from Mutti et al (Mutti, Cooper et al. 2007) reinforced linkage to this locus in a study of mild myopia in an ethnically diverse sample of American school children.

The MYP3 region contains genes coding for the proteoglycans lumican (LUM), decorin, the dermatan sulfate proteoglycan 3 precursor (DSPG3), and fibromodulin (FMOD)(Wojciechowski 2011). These proteoglycans regulate collagen fibril assembly and interaction and are expressed in ocular tissues. Experimental studies show UV light causes collagen crosslinking, which prevents eye elongation and blocks myopia progression (Friedlaender 2005; Wollensak, Iomdina et al. 2005; Wollensak and Iomdina 2009). Eye elongation is a factor that determines ocular refraction. The UVAF measurements are thought to quantify collagen crosslinking (Asawanonda and Taylor 1999; Sandby-Moller, Thieden et al. 2004), which supports the experimental model. Our linkage signal may correspond to the MYP3 locus and/or to the chromosome 12 proteoglycan gene complex regulating collagen fibril assembly.

Ultraviolet autofluorescence is the first phenotype to identify both MYP3 and MYP4 loci in the same genome-wide linkage scan suggesting UV radiation may play a role across heterogenic forms of myopia.

Hence, these results provide genetic evidence for gene-environment interactions between UV light and myopia progression, reinforcing experimental and
observational studies, and warranting further research into a causative biological mechanism of action.

8.5.2 Future Directions

Given previous work and the biological plausibility of a role for UVAF in myopia, these two chromosomal regions should be investigated further using fine-mapping techniques. These markers should have an increased density compared to the 10cM map in order to refine linkage signals with approximately 10 additional markers at a 1cM density should be typed. These additional microsatellite short tandem repeat markers can be selected from the ABI Linkage Mapping sets, as well as several human genome databases and a set of fluorescently labelled primer pairs obtained from Perkin Elmer and genotyping will be performed at AGRF high-throughput typing facility. Also, functional SNPs in high myopia and axial length loci could be tested as covariates to determine the impact on linkage peaks. UVAF should also be conducted in other populations, perhaps sib pairs to determine the relationship to these genes and to confirm and refine implicated regions. Lastly, some experiments should be conducted to specify the conjunctiva component causing the autofluorescence in this UV photographic method.
8.6 NIHS

Linkage mapping described in this thesis was conducted with 10cM density analysed for quantitative and dichotomous phenotypes present in the Norfolk genealogy. These studies detected moderate linkage signals requiring further verification by increasing sample size and/or marker density. Linkage results described in this thesis should also be reanalysed using a high density SNP panel.

8.6.1 Exome sequencing

Whole genome sequencing will be feasible in several years when costs are reduced. Until then, exome sequencing is an efficient strategy to selectively sequence coding regions (or exons) of the human genome requiring sequencing of merely 1/20th of the whole genome to identify novel genes associated with rare and common disorders (Teer and Mullikin 2010). Approximately 180,000 exons exist constituting 1% of the human genome or 30Mb (Tarpey, Smith et al. 2009). It is estimated that the protein coding regions of the human genome constitute 85% of disease-causing mutations (Garber 2008).

Exome sequencing has the potential to be clinically relevant in genetic diagnosis due to current understanding of functional consequences in sequence variation (Garber 2008). The goal of this approach is to identify the functional variation responsible for
common diseases such as cardiovascular disease and glaucoma without the high costs associated with whole-genome sequencing whilst maintaining high coverage in sequence depth (Garber 2008).

8.6.2 Future Collaborations

Since Norfolk Island residents have a limited health care system and receive minimal medical attention, genetic studies involving clinical examinations are of mutual benefit to participants and investigators. Hence, more partnerships like the NIHS-NIES collaboration are encouraged. However, of the 800 individuals sampled from NIES, only 500 were matched to individuals in the NIHS pedigree, and from that <200 were used in subsequent SOLAR analyses. Future collaborations must begin with a list of NIHS subjects from which to sample to optimize pedigree based genetic studies in this population.

Many complex diseases are studied as simple affected-versus-normal dichotomies at the clinical level. Rigorous quantification of intermediate disease states or disease liability is desirable because it increases the information available for mapping studies.
8.7 Final Conclusions

This thesis work investigated genetics of four complex diseases in the Norfolk Island population isolate. Where studies with the Framingham risk score built on previous findings on cardiovascular disease from the Norfolk Island Health Study, the ocular endophenotype and ultraviolet autofluorescence trait research was the first attempt to elucidate eye genetics in this population. Although the Framingham risk score approach had limited success with heritability estimates and linkage analysis, significant LOD scores were obtained for myopia related loci using a novel environmental exposure based phenotype as well as suggestive evidence of linkage for glaucoma related traits. In summary, this population has been very useful in identifying loci related to highly heritable quantitative traits. This success will lead to future studies involving next generation exome sequencing that will identify more genes related to these complex disease traits. Large pedigrees such as the Norfolk Island population isolate are emerging as an important resource for gene discovery of complex diseases.

Computational limitations that led to trimming the large pedigree will be overcome in the future with increasing computer power, which may further strengthen the analytical potential of conducting genetic studies in this population.
References

Arcos-Burgos, M., F. X. Castellanos, et al. (2004). "Pedigree disequilibrium test (PDT) replicates association and linkage between DRD4 and ADHD in


Dyke, B. (1996). PEDSYS: A pedigree data management system, 2.0 edn., Population Genetics Laboratory, Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio.


New York, Longman

Wiley.


