Legacy of Mutiny on the Bounty: Founder Effect and Admixture on Norfolk Island

Stuart Macgregor¹, Claire Bellis²,³, Rod A. Lea²,⁴, Hannah Cox²,
Tom Dyer³, John Blangero³, Peter M. Visscher¹, Lyn R. Griffiths²*

¹Queensland Institute of Medical Research, Brisbane, QLD 4006, Australia
²Genomics Research Centre, Griffith Institute for Health and Medical Research, Griffith University, Southport, QLD 4215, Australia
³Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas, 78227-5301, USA
⁴Kenepuru Science Centre, Institute of Environmental Science and Research, Wellington, New Zealand

Keywords: isolated population, extended pedigree, cardiovascular disease, inbreeding

Running Title: Founder Effect and Admixture on Norfolk Island

*Correspondence to Professor Lyn Griffiths
l.griffiths@griffith.edu.au

Genomics Research Centre,
Griffith Institute for Health and Medical Research
Griffith University,
Gold Coast Campus
Parklands Drive
Southport,

Australia,

4215

tel. +61 7 5552 8664

fax. +61 7 5552 9081
Abstract

The population of Norfolk Island, located off the eastern coast of Australia, possesses an unusual and fascinating history. Most present day islanders are related to a small number of ‘Bounty’ mutineer founders. These founders consisted of Caucasian males and Polynesian females and led to an admixed present day population. By examining a single large pedigree of 5742 individuals, spanning >200 years, we analyzed the influence of admixture and founder effect on various cardiovascular disease (CVD) related traits. Because of the relative isolation of the population, on average one third of the genome of present day islanders (single large pedigree individuals) is derived from 17 initial founders. The proportion of Polynesian ancestry in present day individuals was found to significantly influence total triglycerides, body mass index, systolic blood pressure and diastolic blood pressure. For various cholesterol traits the influence of ancestry was less marked but overall the direction of effect for all CVD related traits was consistent with Polynesian ancestry conferring greater CVD risk. Marker derived homozygosity was computed and agreed with measures of inbreeding derived from pedigree information. Founder effect (inbreeding and marker derived homozygosity) significantly influenced height. In conclusion, both founder effect and extreme admixture have substantially influenced the genetic architecture of a variety of CVD related traits in this population.
Introduction

Many population isolates have been used to help dissect the basis of complex traits. Isolates have one or more of the following properties; a small number of initial founding individuals\textsuperscript{1,2}, long isolation with limited immigration/inter-marriage\textsuperscript{3,4,5}, unusually high levels of inbreeding\textsuperscript{6,7}, extensive sampling of the target population\textsuperscript{1,2} and/or relatively homogeneous environment\textsuperscript{8}. Isolated populations have proven particularly valuable for the purposes of mapping genes involved in rare Mendelian monogenic disorders\textsuperscript{9}. Speculation surrounding the potential advantages of utilizing isolated populations in the study of complex multifactorial diseases has led to the collection of many contrasting population isolates. Several studies have revealed differences between populations both compared to outbred cohorts and within the isolates themselves\textsuperscript{4,10,11,12,13}. Another approach in gene mapping studies is admixture mapping. A number of studies have focused on admixed populations, where individuals with differing ancestries marry\textsuperscript{14,15}. Perhaps unsurprisingly, there have been few studies to date focusing on populations with both admixture and a history of isolation.

Cardiovascular disease (CVD) and related phenotypes such as blood pressure and obesity are known to be influenced by both genetic and environmental factors. Gene mapping studies have identified some loci underlying the heritable component of these phenotypes\textsuperscript{16,17} but the majority of the genetic component of these traits remains to be discovered.

This study involved the participation of an isolated population from Norfolk Island, located off the eastern coast of Australia. Norfolk Island is a small volcanic island 1,600 kilometers east, northeast of Sydney, Australia. The initial founders of Norfolk were relocated from Pitcairn Island and possess a fascinating history - their population was started in 1790 on Pitcairn by the original ‘Bounty’ mutineers (Isle of Mann and British ancestry) and Tahitian women (Polynesian ancestry). Fletcher Christian, the Bounty’s acting Lieutenant, led the mutiny which resulted in a small number of Caucasian men and Tahitian women leaving descendants on the island (a small number of
Tahitian men also traveled to the island but left no descendants. The island community remained isolated from immigration until 1856. In 1856, limited by Pitcairn’s meager resources, most of the islanders relocated to uninhabited Norfolk Island. Present-day Norfolk Islanders have maintained a detailed family genealogy, with family histories traceable back to original founders. The combination of a small number of original founders with diverse ancestries, together with an extremely strict immigration policy and Norfolk’s obvious geographical isolation presents quite a unique population for the investigation of complex multifactorial diseases. Consistent with their isolated history, we recently showed that a set of individuals drawn from Norfolk Island show long stretches of linkage disequilibrium. Further background information on the Norfolk Island data and on Polynesian settlement more generally was published previously.

Materials and Methods

The main focus in this study was on the legacy of a small number of “original founders” and the set of individuals used differs slightly from that used in Bellis et al. There were 9 male founders (Bounty mutineers of Caucasian ancestry) and 12 female founders (Polynesian ancestry). This set of individuals was supplemented by two further Caucasian males in the mid 19th century. In total, 11 male founders and 6 female founders contribute genes to present day individuals - this set of “original founder” individuals appear as filled symbols in the pedigree of 5742 connected individuals in Figure 1. Of the 5742 individuals, 337 had phenotype/genotype information available. Of the 337, 295 were non-founders in the pedigree, with the remainder being married-in individuals.

A present day phenotyped individual is known to have all ancestors in the set of “original founders” (Figure 2). On Figure 2 the lineages can be followed to the initial founders, with the phenotyped individual being related to all six Polynesian “original founders”. There are ~6 generations from most “original founders” to the individual in Figure 2 and since most of the
phenotyped individuals are younger than this there are generally 6-7 generations between “original founders” and phenotyped individuals. If there were a single lineage from founder to present day, phenotyped individuals would be expected to share roughly 0.008-0.015 (1/128 to 1/64) of their genome with any given “original founder”.

In addition to the phenotyped individuals from the main pedigree, a further 256 phenotyped/genotyped individuals were available for study. In total 593 individuals had phenotypes or genotypes available (in virtually all cases both).

Nine quantitative traits were considered (Table 1). Full details for the collection of these data were given previously²⁰. Informed consent was obtained from all participants. Prior to collection of samples, the Griffith University Human Research Ethics Committee granted ethical clearance. A panel of 382 microsatellite markers were typed across all individuals with blood available - full details are in Bellis et al²². Marker information was used to validate the specified pedigree relationships and to assess if “unrelateds” could be connected to the main pedigree using the program GRR²³. A number of “unrelateds” were joined to the initially specified main pedigree, with several mis-specified relationships within the pedigree corrected in the light of the marker data. The marker data suggested that some of the “unrelateds” were distantly related to the main pedigree but was not sufficient to explicitly join such individuals to the pedigree.

Ancestry was assessed by calculating the expected contribution of each set of parents to their offspring. For example, an individual with a 100% Polynesian mother and 100% Caucasian father would have 50% Polynesian and 50% Caucasian ancestry. For the early part of the islanders’ history, all individuals’ ancestry was derived from the “original founders”. The vast majority of individuals who came to the island in more recent years were Caucasian and all of the recent founder individuals outside the set of “original founders” are assumed here to have Caucasian ancestry. Ancestry values for the main pedigree were computed using PEDIG²⁴. Inbreeding coefficients (F) based on the pedigree information were computed in PEDIG using the method of Meuwissen and Luo²⁵. F is the probability that the two alleles at a locus in an individual are
identical by descent.

Genome-wide homozygosity was calculated by computing the proportion of autosomal markers that were homozygous. Inbreeding will increase homozygosity, with inbred individuals displaying increased homozygosity as a result of them inheriting the same alleles identical by descent down both sides of an inbreeding loop. The expected amount of homozygosity in an inbred individual is

$$H_0 + (1 - H_0) F$$

where $H_0$ is the homozygosity in the outbred population. This predicts that a graph of homozygosity versus $F$ will have gradient $1 - H_0$.

The relationship between each trait and ancestry, marker homozygosity and $F$ was assessed using a mixed model approach in SOLAR\textsuperscript{26} - this models the relationship between each individual in the pedigree by fitting a variance component for the additive genetic effect transmitted to each offspring. This allows estimation of the additive genetic variance and hence the heritability of each trait. For each of ancestry, marker homozygosity and $F$ in turn, a fixed effect was fitted to test for an effect of each on the traits of interest. Also fitted in each case were fixed effects for age, sex, age$^2$ and an age by sex interaction. Strictly speaking, the estimates of the variance components from the maximum likelihood approach implemented in SOLAR are biased but in practice the bias is small - this was confirmed by running a subset of the data through the program ASREML\textsuperscript{27} which implements restricted maximum likelihood (data not shown).

Results

Of the 5742 individuals in the main pedigree, 1503 were inbred. The average $F$ of the inbred individuals was 0.044, with a maximum of 0.16 (0.16 indicates slightly more inbreeding than would result from an avuncular marriage). Of the 337 individuals with phenotypes/genotypes, 60 are inbred - the average inbreeding in these individuals was 0.026, with a maximum of 0.081.

Figure 3 shows the relationship between marker derived homozygosity and $F$ for the 60 inbred individuals. The gradient is 1.05 (standard error 0.15); this is not significantly different from the
prediction of 1-0.21=0.79. For individuals who are (based on the available pedigree information) not inbred (F=0) there is some variability in their marker derived homozygosity but the mean value in the non-inbred individuals is 0.21 - this is the same as the intercept derived from the inbred individuals. A small number of individuals from outside the main pedigree had marker homozygosity values much higher than expected if they were not inbred - it is likely that some of these individuals are in fact related to the main pedigree but the available marker data was not sufficient to allow these individuals to be explicitly connected to the main pedigree (using the program GRR).

The average proportion of the genome from initial founders obviously decreases over time (as more new founders are "married-in") but the contribution of the initial founders remains reasonably high. There are 295 non-founder individuals in the main pedigree with phenotypes. On average these individuals carry 18% of their genome from the male original founders and 14% of their genome from the female initial founders - histograms are in Figure 4. The present day individual from figure 2 is on the right hand side of the histograms - she is expected to have inherited 61% of her genome from founder males and 39% from founder females and has F=0.029. The female specific histogram in Figure 2 can also be interpreted as showing the degree of Polynesian/Caucasian admixture in the Norfolk island population (assuming that all new founders were Caucasian). Considering each initial founder separately, some of the original founders contribute much more to present day phenotyped individuals than others (range 0.2% to 3.1%).

There is a strong relationship between inbreeding and ancestry - individuals with ancestors who intermarried with others from small group of founders will be inbred and also carry a large proportion of the original founder individuals’ DNA. The correlation between F and Polynesian ancestry was 0.78. Graphs showing the relationship between proportion of genome from initial founders (i.e. ancestry) and inbreeding coefficient are in Figure 5. The correlation between marker derived homozygosity and Polynesian ancestry was much lower (0.19).

Heritabilities (h²) for each trait were estimated using the mixed model (Table 1). The mixed
model was also used to estimate the effect of Polynesian ancestry, taking into account the known relationships between individuals. Polynesian founder ancestry significantly increased total triglycerides, BMI, SBP and DBP for individuals with more Polynesian ancestry (table 1). There was also an indication that Polynesian ancestry increased total cholesterol. Although the effects of ancestry on the other traits were non-significant, the effect direction was always consistent with Polynesian ancestry leading toward less favorable outcomes with respect to CVD risk (increases in all traits with the exception of HDL and height). It appears likely that having greater Polynesian ancestry generally increased individuals susceptibility to traits associated with CVD when those individuals were ‘exposed’ to a western diet.

Table 2 shows the effect of genome-wide homozygosity (GWH) and inbreeding (F) on each trait. For GWH, N=593; for F, N is only 60 because only individuals from the main pedigree can be assessed for their F value, and only a subset of individuals in the main pedigree are inbred. The only trait significantly affected by marker derived homozygosity was height and this result was also confirmed in the small set of inbred individuals. Total cholesterol and SBP were significantly related to F, although interpretation of this is made difficult by the strong correlation between F and ancestry. The individuals who have high levels of Polynesian ancestry tend to have maintained that high level as a result of them being inbred descendants of the small number of Polynesian founders. Due to the strong correlation between F and ancestry in the pedigreed individuals in the main pedigree, it is difficult to fully disentangle their effects. Fitting both F in the model after ancestry resulted in F becoming non-significant (and vice-versa). For genome-wide homozygosity and ancestry, the correlation was substantially lower (0.19).

The mixed model results shown are for the main pedigree together with the additional individuals (N=593). The results were similar if the analysis was restricted to just the N=295 non-founder individuals in the main pedigree.

We attempted to fit dominance components of variance in the mixed model but with little success. For some of the traits, including a dominance component in the model led to convergence
problems. For other traits, estimates of the dominance variance were estimated but 1) our confidence that these results were correct and not convergence artefacts was low and 2) the estimated standard error of the dominance components were very large. Such problems were not entirely surprising because our pedigree has somewhat limited information to estimate dominance variance because dominance variance can only be estimated from a limited number of (phenotyped) relative pairs such as full sibs. Convergence problems are common when there is only limited information to estimate a parameter in the mixed model.

Discussion

The population of Norfolk Island possesses an unusual and fascinating history. The data presented here offer a unique opportunity to the effects of both ancestry and founder effect on CVD related traits. To allow analysis of ancestry, an accurate pedigree structure defining the “original founders” of interest is vital. For the Norfolk Island population we took advantage of well documented historical records and also confirmed the pedigree structure using marker data - individuals with high pedigree derived inbreeding (F) values also had higher levels of genome-wide homozygosity. Admixed ancestry remains a major factor in the genetic composition of this population because immigration to the island in recent years has been strictly regulated. The vast majority of recent immigrants are thought to be Caucasian - to allow analysis here we assumed all married-in individuals were Caucasian.

Due to the population history of the islanders, there was a strong correlation between male founder ancestry and female founder ancestry. Here, because the female founder ancestry is in the minority (with all new immigration resulting from Caucasians coming to the island), it was assumed that it was admixture of these Polynesians into an otherwise Caucasian sample was the ‘event’ of primary interest. It is possible to interpret the results as genes from English seaman being different (increasing cholesterol, triglycerides, SBP, DBP) from Caucasians more generally, with Polynesian similar to general Caucasians.
Our focus was on the “genome-wide” effect of ancestry on each trait of interest. This approach was taken because the power to detect ancestry effects was high and indeed an effect of ancestry was shown for a number of the traits. A possible next step would be to use the marker data to further identify specific chromosomal regions (admixture mapping) underlying the observed effect of ancestry. However, the limited sample size available here would mean that power to detect any particular chromosomal region would be low and we hence did not attempt this.

Previous studies of founder populations have considered CVD risk related phenotypes. Abney et al\textsuperscript{28} and Ober et al\textsuperscript{8}, examine a Hutterite data set and report substantial dominance components of variance for the traits SBP and LDL but no dominance variance for triglycerides, HDL, height, BMI, and DBP. Campbell et al\textsuperscript{29} examine the effect of inbreeding on these traits in another founder population from the Dalmatian Islands in Croatia. Campbell et al find that inbreeding affects the traits (SBP and LDL) that were found to have positive dominance components of variance but not the other traits. Here we find a slight effect of pedigree derived inbreeding (inbreeding coefficient, F) on SBP but this appears more readily explicable as an effect of Polynesian ancestry than inbreeding. There was no effect of marker derived homozygosity (a proxy for inbreeding) on SBP in the larger set of 593 individuals with available data. We found no effect of inbreeding on LDL in the Norfolk population. Dominance effects have not been reported for any of these traits in large twin studies\textsuperscript{30,31,32}, although this is not necessarily unexpected as twin samples are typically drawn from the general population - the genetic composition of these traits in likely to differ in isolated populations compared with general population samples.

The main finding apropos of homozygosity/inbreeding was that height appeared to be slightly decreased in inbred individuals - there was no significant effect of Polynesian ancestry on height. A number of traits appeared to be influenced by the admixture of Polynesian and Caucasian genetic backgrounds, with Polynesian ancestry significantly increasing trait values for total triglycerides, BMI, SBP and DBP. For the other traits, the effect of ancestry was not significant but in all cases, the direction of effect was consistent with Polynesian ancestry increasing CVD risk. Whilst
inbreeding cannot be ruled out as causing these increases, the lack of association with marker derived homozygosity suggests that ancestry is the main factor here.

Acknowledgments

This work was supported by grants from the Australian National Health and Medical Research Council (Medical Genomics Program grant) and the Griffith Medical Research College (a joint program of Griffith University and the Queensland Institute of Medical Research, QIMR, Herston, QLD, Australia, 4006). SM and PMV are supported by the NHMRC fellowship scheme.

We would particularly like to thank Norfolk Island volunteers for their participation in this study.
References


isolated norfolk island population. *Heredity* 2008; **100**: 366—373.


30. Beekman M, Heijmans BT, Martin NG, et al.. Heritabilities of apolipoprotein and lipid levels


**Table 1: Effect of ancestry on CVD related traits**

<table>
<thead>
<tr>
<th>Trait</th>
<th>$h^2$ (s.e.)</th>
<th>effect of Polynesian ancestry</th>
<th>p-value for Polynesian ancestry</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCholesterol</td>
<td>0.40 (0.12)</td>
<td>1.04</td>
<td>0.08896</td>
</tr>
<tr>
<td>TTriglycerides</td>
<td>0.14 (0.16)</td>
<td>1.55</td>
<td>0.00424</td>
</tr>
<tr>
<td>HDL</td>
<td>0.22 (0.16)</td>
<td>-0.74</td>
<td>0.18456</td>
</tr>
<tr>
<td>LDL</td>
<td>0.41 (0.11)</td>
<td>0.70</td>
<td>0.26285</td>
</tr>
<tr>
<td>Height</td>
<td>0.76 (0.10)</td>
<td>-0.37</td>
<td>0.53888</td>
</tr>
<tr>
<td>Weight</td>
<td>0.30 (0.13)</td>
<td>0.80</td>
<td>0.14003</td>
</tr>
<tr>
<td>BMI</td>
<td>0.17 (0.13)</td>
<td>1.15</td>
<td>0.03498</td>
</tr>
<tr>
<td>SBP</td>
<td>0.24 (0.12)</td>
<td>1.75</td>
<td>0.00061</td>
</tr>
</tbody>
</table>
| DBP            | 0.06 (0.12)  | 1.15                          | 0.01521                         

Table 1 caption

$h^2$ is the heritability of each trait. s.e. denotes the standard error. N=593. The scale for effect size is for each trait transformed to N(0,1), with Polynesian ancestry as a proportion between 0 and 1. For example, an individual with 0.3 Polynesian ancestry would on average have ~0.11 standard deviations lower height than an individual with no Polynesian ancestry. Having all traits on the same scale allows all traits to be easily compared in terms of effect sizes.
Table 2: Effect of homozygosity and inbreeding on CVD related traits

<table>
<thead>
<tr>
<th>Trait</th>
<th>GWH effect size</th>
<th>p-value for GWH</th>
<th>F effect size</th>
<th>p-value for F</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCholesterol</td>
<td>0.66</td>
<td>0.66048</td>
<td>17.17</td>
<td>0.01898</td>
</tr>
<tr>
<td>TTriglycerides</td>
<td>1.28</td>
<td>0.41355</td>
<td>15.27</td>
<td>0.05299</td>
</tr>
<tr>
<td>HDL</td>
<td>-0.21</td>
<td>0.89381</td>
<td>-1.87</td>
<td>0.78042</td>
</tr>
<tr>
<td>LDL</td>
<td>0.60</td>
<td>0.69520</td>
<td>9.29</td>
<td>0.22282</td>
</tr>
<tr>
<td>Height</td>
<td>-2.72</td>
<td>0.01782</td>
<td>-13.33</td>
<td>0.02610</td>
</tr>
<tr>
<td>Weight</td>
<td>-2.18</td>
<td>0.12699</td>
<td>3.37</td>
<td>0.67334</td>
</tr>
<tr>
<td>BMI</td>
<td>-1.32</td>
<td>0.40732</td>
<td>11.22</td>
<td>0.19824</td>
</tr>
<tr>
<td>SBP</td>
<td>0.35</td>
<td>0.80152</td>
<td>16.38</td>
<td>0.00812</td>
</tr>
<tr>
<td>DBP</td>
<td>1.31</td>
<td>0.40785</td>
<td>3.02</td>
<td>0.69493</td>
</tr>
</tbody>
</table>

Table 2 caption

Genome-wide homozygosity (GWH) takes the range of values as shown in figure 3, with most values in the range 0.2-0.3. F takes values between 0 and 0.081 for these individuals. N=593 for GWH and N=60 for F.
Figure Legends

Figure 1: Full Pedigree

Figure 2: One present day individual

Figure 3: Marker derived homozygosity and inbreeding coefficient

Figure 4: (4A males, 4B females) Contribution of “original” founders

Figure 5: (5A males, 5B females) Relationship between ancestry and inbreeding
A present day individual with all relatives in the set of "original founders"
Validation of increased marker homozygosity in the presence of inbreeding.

Proportion of autosomal markers that are homozygous:

\[ y = 0.21 + 1.05x \]

Pedigree inbreeding coefficient, \( F \)

Correlation: 0.68
Average proportion of genome in phenotyped individuals from founder males

N=295 mean=0.18

Histogram: Male Founders
Histogram: Female Founders

Average proportion of genome in phenotyped individuals from founder females

N=295 mean=0.14

Histogram: Female Founders
Male founder ancestry versus inbreeding for N=60 inbred individuals. The proportion initial male founder ancestry is given by the equation $y = 0.240 + 4.364x$, with a correlation of $0.65$. The plot shows a linear relationship between the pedigree inbreeding coefficient, $F$, and the proportion initial male founder ancestry.
Female founder ancestry versus inbreeding for N=60 inbred individuals.

Pedigree inbreeding coefficient, $F$

Proportion initial female founder ancestry

$y = 0.175 + 3.990x$, correlation = 0.78

For N=60 inbred individuals.

Female founder ancestry versus inbreeding.