Leprosy epidemics during history increased protective allele frequency of \textit{PARK2}/\textit{PACRG} genes in the population of the Mljet Island, Croatia

A. Bakija-Konsuo\textsuperscript{a,*}, R. Muli\textsuperscript{b}, V. Boraska\textsuperscript{c}, M. Pehlic\textsuperscript{c}, J.E. Huffman\textsuperscript{d}, C. Hayward\textsuperscript{d}, M. Marlais\textsuperscript{e}, T. Zemunik\textsuperscript{c}, I. Rudan\textsuperscript{f}

\textsuperscript{a} Clinic for Dermatovenerology Cutis, Vukovarska 22, Dubrovnik, Croatia
\textsuperscript{b} Department of Public Health, Medical School, University of Split, Split, Croatia
\textsuperscript{c} Department of Medical Biology, Medical School, University of Split, Split, Croatia
\textsuperscript{d} MRC Human Genetics Unit; Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, Scotland EH4 2XU, UK
\textsuperscript{e} Imperial College School of Medicine, Imperial College London, London, UK
\textsuperscript{f} The Croatian Centre for Global Health Medical School, University of Split, Split, Croatia

\begin{abstract}
\textbf{Introduction:} Two regulatory polymorphisms (rs1040079 and rs9356058) shared by \textit{PARK2} and \textit{PACRG} genes were identified as major risk variants for leprosy susceptibility. The aim of this study was to investigate if allele frequencies of these polymorphisms in the isolated population of the island of Mljet, which served as a quarantine for leprosy patients during past centuries, were different to allele frequencies in two control populations with no history of leprosy.

\textbf{Subjects and methods:} This study included 88 unrelated Caucasian individuals from the island of Mljet while two control groups included 93 individuals from the island of Rab and 160 individuals from the region of Split. Genotyping for rs1040079 and rs9356058 was performed by \textit{real-time} PCR analysis. We also compared the allele frequency of the rs9356058 polymorphism from the population of Mljet with allele frequencies derived from the existing genome wide association scans in two additional island populations, Vis (924 subjects) and Korcula (909 subjects).

\textbf{Results:} We found a significant increase in the frequency of rs9356058 allele C in the population of Mljet when compared to both control groups. We also observed a significant increase in the frequency of rs1040079 allele A in the population of Mljet when compared with the population of Rab, however this increase was not significant when compared with the population of Split. Allele frequencies of both examined polymorphisms did not differ between the two control populations. Protective haplotype rs9356058-rs1040079 CA was also more frequent in the population of Mljet compared with the Rab and Split populations. In addition, an increase of frequency of rs9356058 allele C was also observed in the population of Mljet when compared with the frequency in the Korcula population.

\textbf{Conclusion:} The results of our study show the association of polymorphisms rs9356058 and rs1040079 in gene \textit{PARK2}/\textit{PACRG} with leprosy. The results of our study indicate that exposure to leprosy and mortality in the population caused by leprosy on Mljet resulted in the selection of rs9356058 "protective" C allele in the \textit{PARK2} gene, while this was not observed in the two control groups. This is the first study to assess the genetic susceptibility to leprosy in a European population.

\end{abstract}

\section{Introduction}
The prevalence of leprosy in the world is in rapid decline, primarily thanks to improved diagnostics and efficient multi-drug therapy. However, there are still 250,000 new patients each year, mostly in Africa, Asia, India and South America where leprosy is still a public health issue [1–3]. There is no indigenous leprosy nowadays in Croatia or anywhere in Europe [3–5].

Leprosy or Hansen’s disease ravaged across Europe, including Croatia, for centuries. During the Crusades (in the period between the 11th and 13th century) there were many leprosy patients across Europe, and the disease itself assumed the features of a pandemic. The largest number of leprosy patients at the time resided in Dalmatia, along one of the routes used by European crusaders to return.
home from the Holy Land, which was of particular importance for the maritime city of Dubrovnik in the sense of early prevention of the disease’s expansion. Dated 1272, the Statute of the City of Dubrovnik provides the first mention of the isolation of leprosy patients. It is a known fact that the concept of quarantine was “invented” by the Dubrovnik Republic, and the Dubrovnik administration designed the concept of quarantine as a result of its experiences with leprosy patients [6–9]. In order to improve the lives of quarantined persons, as well as to tighten epidemiological measures, a quarantine was created in 1397 in the Benedictine Monastery of St Mary on the island of Mljet and this functioned with minor interruptions for more than 130 years [6–11]. Mljet population has been isolated during centuries because of its geographical position and there have been no considerable migrations and no evidence of any major epidemics of infectious diseases in the last few centuries, which is methodologically important for this study [12].

Ongoing efforts to identify genetic risk factors for leprosy include candidate gene association studies and genome-wide linkage analysis in samples from endemic countries. Numerous candidate gene studies have identified variants in VDR, HLA-DRB2 specificities, TAP1 and TAP2, CIITA, COL3A, SEC11A1 (also called IRAMP1), IL-10 and TNF-α genes to be associated with leprosy and its subtypes. These associations reflect the differential susceptibility to this polygenic disease in different populations [1,2,13–18]. The replication of these findings in different populations is a crucial step towards the validation of these studies.

The first leprosy study performing genome-wide linkage analyses reported linkage of chromosome region 10q13 [19]. Tosh et al. showed marginal evidence for linkage to chromosome region 20p12 while a genome-wide scan analysis provided by Mira et al. detected linkage to chromosome region 6q25–q26 with the disease [20,21]. Their evidence suggested that the 6q25 locus is involved in leprosy, both the paucibacillary and multibacillary types. Mira et al. showed that only two “tag” single nucleotide polymorphisms (SNPs) located within the first intron of PACRG, rs9356058 and rs1040079 were needed to capture all association information between the association region and leprosy [22].

The PARK2, also named Parkin gene maps to 6q25.2-q27, the region to which autosomal recessive juvenile parkinsonism maps [23]. The precise function of this gene is unknown; however, the encoded protein is a component of a multiprotein E3 ubiquitin ligase complex that mediates the targeting of substrate proteins for proteasomal degradation. Mutations in this gene are known to cause Parkinson’s disease and autosomal recessive juvenile Parkinson’s disease. In addition, PARK2, and to a lesser extent, PACRG, were found to be expressed in Schwann cells and macrophages, the primary host cells of Mycobacterium leprae, the causative agent of leprosy. Mira et al. noted that both genes are linked to the ubiquitin-mediated proteolysis system, which has heretofore received little attention in the study of leprosy pathogenesis and the control of M. leprae in the human host [22].

The aim of this study was to investigate rs1040079 and rs9356058 allele frequencies in the isolated population of the island of Mljet, which served as a quarantine for leprosy patients during the past centuries, and to compare them with allele frequencies in two control populations with no history of leprosy, one as an isolated island population and another as a mixed urban population.

2. Subjects and methods

2.1. Study design

The study compares the frequencies of the two abovementioned SNPs among European sub-populations of the same genetic origin, with one of them being affected and decimated by leprosy (Mljet) and two other control populations that were spared (Split and Rab).

Samples involved in this study were very carefully selected. The group of samples from Mljet were selected because Mljet had a very well documented history of leprosy patients and leprosarium in contrast to control populations from Split and Rab. The village of Barbat on the island of Rab was formed by the settlers from southern Dalmatia in 18th century and therefore it was chosen as a negative control population. Samples from Split were selected by random choice and represent general population from Split area [24]. Historic evidence suggests that this urban region had no history of exposure to leprosy, so Split was chosen as another unaffected control group. The islands of Mljet and Rab were of particular interest for this investigation because of their geographic isolation and very low immigration rates during past centuries. Indeed, because of their isolation any differences in the frequency of SNPs between examined and controls populations would be likely to remain detectable to date, rather than levelled out by gene flow from the large outbred population.

2.2. Subjects

The study included 88 unrelated Caucasian individuals (42 male) from the island of Mljet. The two control groups included 93 individuals from the island of Rab (47 male) and 160 individuals from the region of Split (74 male). All investigated populations (Mljet, Split and Rab) originate form the same ancestral population (Mediterranean part of Croatian population). In order to confirm the value of our results, we compared the allele frequency of rs9356058 polymorphism from the population of Mljet with allele frequencies in two additional island populations, Vis (924 subjects) and Korcula (909 subjects), that also originate from the same ancestral population, rs9356058 genotype data for Vis and Korcula were derived from the existing genome-wide scans association. Historical data shows that only sporadic cases of leprosy existed on the islands of Vis and Korcula [7,11]. This study was approved by the ethics committee of Medical School, University of Split, Croatia, and informed consent from all individuals was obtained prior to inclusion in the study. According to the population census in 2001. There are 1111 permanent inhabitants on the Island of Mljet, on the Island of Rab (town Barbat) there are 1205 permanent inhabitants, while in the region of Split there are 463 676 permanent inhabitants [25].

2.3. Gene polymorphism

Genomic DNA was extracted from peripheral blood leukocytes using the Nucleon DNA Blood Kit (Tepnel, UK). Two SNPs (rs1040079 and rs9356058) were selected across the PARK2 gene according to literature data [15,22]. “Real-time” polymerase chain reaction (PCR) genotyping for these two PARK2 polymorphisms was performed using an ABIPRISM 7500 Sequence Detection System (Applied Biosystems, Foster City, USA) and pre-developed TaqMan assay reagents, C-1575898-10 for rs1040079 SNP and C-179028-10 (Applied Biosystems, Foster City, USA) and pre-developed TaqMan assay reagents, C-1575898-10 for rs9356058 SNP. PCR reaction was carried out according to the manufacturer’s protocol.

2.4. Statistical analysis

Prior to association analysis, we performed quality control (QC) of the obtained genotypes. We tested Hardy–Weinberg equilibrium (HWE) and minor allele frequencies (MAF) for all samples using Plink (version 1.00, http://pngu.mgh.harvard.edu/purcell/plink/). MAFs of both SNPs were compared with the National Centre for Biotechnology Information SNP database (NCBI dbSNP) MAFs for the central European (CEU) population (www.ncbi.nlm.nih.gov/projects/SNP/). Comparisons of allele and haplotype frequencies
Association analysis for rs9356058 and rs1040079 between population of Mljet and two control groups, Split and Rab.

Table 1

<table>
<thead>
<tr>
<th>Population</th>
<th>Alleles (major:minor)</th>
<th>%Geno</th>
<th>HWpval</th>
<th>MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9356058</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mljet</td>
<td>C:T</td>
<td>100.0</td>
<td>0.0786</td>
<td>0.483 (T)</td>
</tr>
<tr>
<td>Split</td>
<td>T:C</td>
<td>100.0</td>
<td>0.795</td>
<td>0.366</td>
</tr>
<tr>
<td>Rab</td>
<td>T:C</td>
<td>98.9</td>
<td>0.0164</td>
<td>0.337</td>
</tr>
<tr>
<td>rs1040079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mljet</td>
<td>A:G</td>
<td>100.0</td>
<td>0.0652</td>
<td>0.182</td>
</tr>
<tr>
<td>Split</td>
<td>A:G</td>
<td>100.0</td>
<td>0.1272</td>
<td>0.266</td>
</tr>
<tr>
<td>Rab</td>
<td>A:G</td>
<td>100.0</td>
<td>0.3856</td>
<td>0.355</td>
</tr>
</tbody>
</table>

a Alleles — the major and minor alleles for this marker.
b %Geno — the percentage of non-missing genotypes for this marker.
c HWpval — the Hardy–Weinberg equilibrium p value.
d MAF — the minor allele frequency for this marker.

between all three studied populations were done using Haploview. Linkage disequilibrium (LD) for two investigated SNPs was calculated using Haploview [26]. P-values less than 0.05 were considered nominally significant.

3. Results

SNP rs1040079 was found to be in HWE in all examined populations, whereas rs9356058 slightly deviated from HWE (p = 0.0164) in the population of Rab. The percentage of missing genotypes for SNPs rs9356058 and rs1040079 was less than 1%. QC results are shown in Table 1. SNPs rs9356058 and rs1040079 are not in tight LD with each other (Mljet $r^2 = 0.17$, Split $r^2 = 0.18$, Rab $r^2 = 0.28$).

The results of association analysis are shown in Table 2. We found a significant difference in the frequency of rs9356058 allele C between the population of Mljet and Rab, before and after Bonferroni correction. We observed a significant difference in the same allele frequency between the populations of Mljet and the region of Split, while the control populations of Rab and Split did not show a significant difference in the frequency of allele C.

Association analysis showed a significant difference in the frequency of rs1040079 allele A between the populations of Mljet and Rab, before and after Bonferroni correction. There was no significant difference in the frequency of this allele between the populations of Mljet and Split after Bonferroni correction, there was also no significant difference in the frequency of this allele between two control populations of Rab and Split.

rs9356058-rs1040079 CA haplotype frequency was higher in the population of Mljet in relation to Rab (p = 0.0013) and Split (p = 0.0017). In the same time, rs9356058-rs1040079 TG haplotype had higher frequency in the population of Rab in relation to Mljet (p = 7.5 × 10^{-5}). Results of rs9356058-rs1040079 haplotype association analyses between population of Mljet and two control groups, Split and Rab are presented in Table 3.

Comparison of rs9356058 allele frequency between populations of Mljet and two additional populations, Vis and Korcula, showed a significant difference in frequency between Mljet and Korcula (p = 0.0468), with a higher frequency of C allele in the population of Mljet. However, the difference was not present when we compared the allele frequencies of the same polymorphism in populations of Mljet and Vis (0.6832).

4. Discussion

In our study we analyzed differences in the frequencies of two polymorphisms of the PARK2/PACRG gene, rs9356058 and rs1040079 in populations from Mljet, Rab and Split. There is no documented history of leprosy patients in both control regions according to literature data. Rab is isolated homogenous population while Split is mixed urban population. We found a significant increase in the frequency of rs9356058 allele C in the population of Mljet when compared both with the population of Rab and the population of Split. Also, rs9356058 allele C had a significantly higher frequency in Mljet than in the additionally investigated island population of Korcula, thus confirming our main result. We also observed a significant increase in the frequency of rs1040079 allele A in the population of Mljet when compared with the population of Rab, however this increase was not significant when compared with the population of Split. We would suggest that one explanation for the lack of significant difference in rs1040079 allele A between Mljet and Split is due to the mixed urban population in Split, and the difference in rs1040079 allele A is retained when comparing Mljet and Rab as both are isolated populations. The allele frequencies of both examined SNPs did not differ between the two control populations, Split and Rab. An interesting observation which further corroborates our hypothesis of the protective role of PARK2 against leprosy comes from haplotype analysis. We observed a significantly higher frequency of the rs9356058-rs1040079 CA haplotype in the population of Mljet than in Rab and Split.

Mira et al confirmed the strong association of the PARK2/PACRG gene, rs9356058 and rs1040079 to leprosy. Common allele T of PARK2 rs9356058 and rare allele C of rs1040079 were independently associated with an increased risk of leprosy [21,22]. Malhorta et al studied an ethnically homogeneous population of Indian leprosy patients and controls for associations with six SNPs in the common regulatory region of PARK2 and PACRG. After Bonferroni correction, they found no significant associations, in contrast with the findings in Vietnamese and Brazilian populations reported by Mira et al. They concluded that the risks associated with these SNPs vary in different populations [15].

Table 2

<table>
<thead>
<tr>
<th>SNP</th>
<th>Assoc allele</th>
<th>1./2. population ratio counts</th>
<th>1./2. population frequencies</th>
<th>Chi square ($\chi^2$) test</th>
<th>P value</th>
<th>Bonferroni correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mljet-Rab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9356058</td>
<td>C</td>
<td>91.85, 62:122</td>
<td>0.517, 0.337</td>
<td>11.938</td>
<td>5.00E-04</td>
<td>1.00E-03</td>
</tr>
<tr>
<td>rs1040079</td>
<td>A</td>
<td>144.32, 120:66</td>
<td>0.818, 0.645</td>
<td>13.712</td>
<td>2.00E-04</td>
<td>4.00E-04</td>
</tr>
<tr>
<td>Mljet-Split</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9356058</td>
<td>C</td>
<td>91.85, 117:203</td>
<td>0.517, 0.366</td>
<td>10.692</td>
<td>0.0011</td>
<td>0.0022</td>
</tr>
<tr>
<td>rs1040079</td>
<td>A</td>
<td>144.32, 235:85</td>
<td>0.818, 0.734</td>
<td>4.425</td>
<td>0.0354</td>
<td>0.0708</td>
</tr>
<tr>
<td>Rab-Split</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9356058</td>
<td>C</td>
<td>117.203, 62:122</td>
<td>0.366, 0.337</td>
<td>0.419</td>
<td>0.5173</td>
<td>/</td>
</tr>
<tr>
<td>rs1040079</td>
<td>A</td>
<td>235.85, 120:66</td>
<td>0.734, 0.645</td>
<td>4.472</td>
<td>0.0345</td>
<td>0.069</td>
</tr>
</tbody>
</table>

Chi square ($\chi^2$) test $p$ value and Bonferroni correction.
Considering the fact that Mljet served the Dubrovnik Republic as a quarantine (for leprosy, as well as some other diseases undefined at the time, such as plague, lupus etc.) and taking into account that only several European populations have remained geographically isolated (as is the case with island of Mljet), we are provided with a unique opportunity to join in the current scientific discussions investigating the populations of the Adriatic islands [7–10]. In Croatia, as well as in Europe there is no indigenous leprosy nowadays. However, using the previously mentioned historical data, we were able to study the genetic susceptibility to leprosy for the first time in a European population. The results of our study indicate that the exposure to leprosy on the island of Mljet resulted in the selection of “protective” C alleles in the gene PARK2 for the SNP rs9356058, while in two control groups where no leprosarium are recorded in historical data, this selection was not observed.

During the past decades, leprosy has been studied from a somewhat unusual perspective for an infectious disease; modern methods for experimental analysis have been applied to demonstrate the existence of an important genetic effect controlling host susceptibility to leprosy and its phenotypes. Complete understanding of the genetic mechanisms of leprosy susceptibility may ultimately lead to exciting new perspectives for treatment and prevention of leprosy, as well as for other infectious diseases [30, 31].

Our study shows that the exposure to different environmental factors (in this case dying of leprosy) may influence the genetic structure of a population. We observe borderline non-significant deviation from HW for both investigated SNPs in Mljet population (rs9356058, \( p = 0.0786 \) and rs1040079, \( p = 0.0652 \)) (Table 1). This observation reinforces and supports the fact that negative selection caused by leprosy epidemics affected genotype distribution by loss of susceptible alleles and is visible through slight HW disequilibrium. Also this gives further evidence that Mljet was an isolated population during last several centuries. It is known that changes in allele frequencies induced by negative/positive selection can be detectable after long periods of time but can also point towards a population bottlenecks effect [32]. This study also shows that isolated populations have the advantage of easier gene identification, compared to more outbred populations. Although we cannot rule out the possibility that the enrichment of this allele is due to a simple overrepresentation amongst cases from the start, our results give further evidence that the increase in the protective allele in Mljet is due to positive selection as a result of leprosy.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>1/z. population frequencies</th>
<th>Chi square (y2) test</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mljet-Rab</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>0.505: 0.338</td>
<td>10.364</td>
<td>0.0013</td>
</tr>
<tr>
<td>TA</td>
<td>0.313: 0.307</td>
<td>0.015</td>
<td>0.9021</td>
</tr>
<tr>
<td>TG</td>
<td>0.170: 0.353</td>
<td>15.679</td>
<td>7.5 \times 10^{-5}</td>
</tr>
<tr>
<td>Mljet-Split</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>0.504: 0.359</td>
<td>9.867</td>
<td>0.0017</td>
</tr>
<tr>
<td>TA</td>
<td>0.314: 0.375</td>
<td>1.862</td>
<td>0.1724</td>
</tr>
<tr>
<td>TG</td>
<td>0.169: 0.259</td>
<td>5.263</td>
<td>0.0218</td>
</tr>
<tr>
<td>Rab-Split</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>0.338: 0.361</td>
<td>0.288</td>
<td>0.5915</td>
</tr>
<tr>
<td>TA</td>
<td>0.307: 0.373</td>
<td>2.227</td>
<td>0.1357</td>
</tr>
<tr>
<td>TG</td>
<td>0.354: 0.261</td>
<td>4.83</td>
<td>0.028</td>
</tr>
</tbody>
</table>

The results of our study show a significant increase in the frequencies of rs9356058 allele C, rs1040079 allele A and rs9356058-rs1040079 CA haplotype in the population of Mljjet which has a history of leprosy, whereas there was no change in allele/haplotype frequencies in two control groups with no history of leprosy. This is the first study to assess genetic susceptibility to leprosy in a European population and the first study to use historical populations to do so. This study also provides further evidence for the importance of the PARK2/PARCGR region in human genetic susceptibility to leprosy.

References