Gene 532 (2013) 132–139

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short Communication

Genetic risk assessment for cardiovascular disease in Azoreans (Portugal): A general population-based study

Manuela Correia [a,1], Claudia C. Branco [a,b,c,*,1], Nicolas Bruffaerts [a], Ainhoa Balagué [a], Luisa Mota-Vieira [a,b,c]

[a] Molecular Genetics and Pathology Unit, Hospital of Divino Espirito Santo of Ponta Delgada, EPE, São Miguel Island, Azores, Portugal
[b] Instituto Gulbenkian de Ciência, Oeiras, Portugal
[c] Center for Biodiversity, Functional & Integrative Genomics (BioFig), Lisbon, Portugal

ARTICLE INFO

Article history:
Accepted 29 August 2013
Available online 13 September 2013

Keywords:
9p21
USF1
LDLR
Cardiovascular disease
Genetic risk
Azores

ABSTRACT

The identification of clinically validated genetic variants contributing to complex disorders raise the possibility to investigate individuals' risk. In this line of research, the present work aimed to assess the genetic risk for cardiovascular disease (CVD) in Azoreans. Genotyping of 19 SNPs – 9 on 9p21, 5 on *LDLR* and 5 on *USF1* – was performed by TaqMan assays on 170 healthy Azorean individuals. Results demonstrate that the most frequent haplotype in 9p21, with a frequency of 41.4%, is TGGGCGCGC, which harbors all risk alleles. Considering haplotype homozygosity data show that females present higher value of homozygosity for both *LDLR* (13.5%) and *USF1* (15.3%), whereas males present higher value for the 9p21 region (8.2%). Interestingly, genetic profile analysis revealed differences in terms of geographic and gender distribution. The Azorean Central group presented a higher risk for atherosclerosis, 2.7 times higher when compared to the Eastern group, while the Eastern group shows 1.5 times higher risk for dyslipidemias. Moreover, Azorean females demonstrated a 4 times higher risk for dyslipidemias when compared to males, whereas males have an increased risk for atherosclerosis. Although allele frequencies in Azoreans were similar to those reported for the HapMap CEU population, the differences in terms of haplotype and genetic profile distribution must be taken in consideration when assessing genetic risk. Taken together, the data here presented evidence for the need to perform biomedical research and epidemiologic analysis in Azoreans with the aim of developing strategies to CVD prevention, health promotion and population education.

1. Introduction

Cardiovascular disease (CVD) comprises a group of complex disorders that are highly prevalent and the leading cause of deaths worldwide. Research on European CVD mortality rates reported regional variations, both between and within countries, mainly due to differences in risk factors (Müller-Nordhorn et al., 2008). Based on the Euro Heart Survey 2006 (Scholte op Reimer et al., 2006), the number of deaths per 100,000 inhabitants in 2004, due to CVD, places Portugal with higher values compared with south western European countries, like Spain and France. A similar analysis comparing Portuguese regions reveals that Azoreans, a young admixed islands population with the same lifestyle habits, presents the highest value of CVD mortality (Healing the Heart, 2007). These and other biomedical data address key questions to be investigated in Azoreans.

CVD and their underlying pathologic processes, atherosclerosis and thrombosis, result from the interplay of both traditional (potentially modifiable, such as smoking, hypertension, hyperlipidemia, diabetes, overweight) and genetic (non modifiable) factors (Ho et al., 2011; Peden and Farrall, 2011). Presently, more than 30 genetic variations in distinct *loci* have been associated with CVD, which explain only a small proportion of heritability and demonstrate the high complexity of these diseases (Peden and Farrall, 2011; Roberts and Stewart, 2012). Of considerable relevance is the 9p21 *locus*, the first genetic risk whose effect is independent from traditional CVD risk factors. This *locus* co-localizes with exons 13-19 of *CDKN2B-AS*, a large non-coding antisense RNA gene, formerly known as ANRIL (antisense non-coding RNA gene; Pasmant et al., 2011). Carriers of the risk allele have ~2 fold increased probability of developing coronary artery disease (CAD) and/or myocardial infarction (MI; Roberts and Stewart, 2012). The biological pathway through which 9p21 acts in development of CVD is not well established; however, until now, the majority of studies reveal a role in atherosclerosis. For example, Harismendy et al. (2011) demonstrate that risk variants in 9p21

Abbreviations: CAD, Coronary artery disease; CHD, Coronary heart disease; CVD, Cardiovascular disease; GP, Genetic profiles; HWE, Hardy-Weinberg equilibrium; INE, National Institute of Statistics; LD, Linkage disequilibrium; LDLR, Low density lipoprotein receptor; MAF, Minor allele frequency; MI, Myocardial infarction; ROH, Run of homozygosity; SNP, Single nucleotide polymorphism; USF1, Upstream stimulatory factor 1.

* Corresponding author at: Molecular Genetics and Pathology Unit, Hospital of Divino Espirito Santo of Ponta Delgada, EPE, 9500-370 Ponta Delgada, São Miguel Island, Azores, Portugal. Tel.: +351 296 203 631; fax: +351 296 203 090.
E-mail address: claudiacbranco@hdes.pt (C.C. Branco).

[1] Both authors contributed equally to this work.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.08.099

disrupt the binding site of STAT1, a transcription factor involved in inflammatory response. Moreover, these authors also demonstrate that 9p21 is involved in long-range enhancer interactions and alters expression of neighboring genes.

Less relevant in terms of risk increase, but notable among the traditional risk factors for CVD, are genes implicated in dyslipidemia. The most important is *LDLR* (low density lipoprotein receptor), which encodes a cell surface protein that mediates specific uptake and degradation of LDL. Located on 19p13.1–13.3 region, the *LDLR* gene spans 45 kb and comprises 18 exons and 17 introns (Setia et al., 2012). According to the British Heart Foundation, 1741 allelic variants have been described, most of them associated with familial hypercholesterolemia (http://www.ucl.ac.uk/ldlr/Current/index.php?select_db=LDLR; access date June 14th, 2013). *LDLR* mutations result in ineffective clearance of serum LDL cholesterol, and contribute to early-onset atherosclerosis and CVD (Franceschini et al., 2009). Moreover, Martinelli et al. (2010) demonstrated that a *LDLR* polymorphism (rs688), associated with CVD, is independent of plasma lipids, suggesting a higher complexity in the role played by *LDLR* in CVD development.

Another gene implicated in genetic dyslipidemia is the *USF1* (upstream stimulatory factor 1), located on chromosome 1q22-23, which encodes for an ubiquitously expressed cellular transcription factor. It is a member of the basic helix-loop-helix-leucine zipper superfamily of transcription factors, and regulates other genes involved in lipid and glucose metabolism (Naukkarinen et al., 2009). Research conducted in different European populations and US whites (Collings et al., 2008; Lee et al., 2007; Reiner et al., 2007) suggest that *USF1* is associated with familial combined hyperlipidemia, a common genetic dyslipidemia, as well as with other dyslipidemic disorders. Two independent studies identified an association between *USF1* variants and coronary atherosclerosis, reporting both protective (rs3737787, OR = 0.79, 95% CI 0.63-0.98; Reiner et al., 2007) and risk effects (rs2516839, OR = 2.39, 95% CI 1.30–4.40; Kristiansson et al., 2008), which, in turn, make this gene an interesting candidate in the understanding of the genetic architecture of CVD.

The identification of clinically validated genetic variants, common or rare, contributing to complex disorders raise the possibility to investigate individuals who have an early age onset disease and/or a positive family history. Nowadays, these variants, known as biomarkers, are also being used to target diagnostic, preventive and/or therapeutic interventions. In this line of research, the present work aims to assess the genetic risk for CVD in Azoreans, based on the analysis of 9p21 region and two genes – *LDLR* and *USF1* – previously reported as disease associated. The knowledge here obtained constitutes a valuable contribution to understanding the CVD burden in Azoreans and is crucial for developing prevention strategies and better healthcare measures.

2. Material and methods

2.1. Ethics statement

The population sample was composed of unrelated healthy blood donors from the anonymized Azorean DNA bank located at the Hospital of Divino Espírito Santo of Ponta Delgada, EPE. This DNA bank was established after approval by the local Health Ethics Committee and follows the international ethical guidelines, which include written informed consent, confidentiality, anonymity of personal data and abandonment option in case of expressed will.

2.2. Bio-demographic data of Azores population and study group

The Azores archipelago is composed of 9 small islands dispersed among three geographic groups: Eastern — Santa Maria, São Miguel; Central — Terceira, Graciosa, São Jorge, Pico, Faial; and Western — Corvo, Flores. The total population consists of 246,772 inhabitants, unevenly distributed by the islands, with the Eastern group comprising around 58.1% of the islanders (143,408), the Central having 99,141 inhabitants (40.2%), and the remaining 4223 (1.7%) in the Western group (Portugal Census, 2011). The majority of the population lives in small rural localities, which are characterized by agriculture and cattle breeding economy, and its inhabitants show similar life styles as well as eating habits.

The study group was composed of 170 DNA samples of healthy unrelated blood donors from the anonymized Azorean DNA bank located at the Hospital of Divino Espírito Santo of Ponta Delgada, EPE, on São Miguel Island (Mota-Vieira et al., 2006). The sample was geographically representative of the 9 islands: São Miguel (n = 60), Santa Maria (n = 10), Faial (n = 10), Graciosa (n = 10), Pico (n = 10), São Jorge (n = 10), Terceira (n = 30), Flores (n = 20) and Corvo (n = 10). Gender representativeness was also considered, in a total of 85 males and 85 females, which were also half and half distributed by each island sample. This sample size is adequate to grasp the genetic diversity observed in the Azoreans, since the minor allele frequency (MAF) of each selected single nucleotide polymorphism (SNP) here analyzed is higher than 10%.

2.3. SNP selection and genotyping

Three genomic regions associated with CVD risk, namely 9p21, *LDLR* gene (19p13.1–13.3) and *USF1* gene (1q22-q23), were assessed (Table 1). SNPs were selected based on known associations established from GWAS and candidate gene studies. A total of 19 SNPs were selected, 9 at 9p21 (rs10116277, rs10757274, rs4977574, rs2383206, rs944797, rs2383207, rs1537375, rs107557278, rs1333049), 5 at *LDLR* (rs6511720, rs2228671, rs688, rs1433099, rs2738466), and 5 at *USF1* (rs10908821, rs3737787, rs2516839, rs1556259, rs2774279), covering 44.1, 40.4 and 9.0 kbp, respectively.

Excepting rs944797, all SNPs were genotyped by TaqMan® Pre-Designed SNP Genotyping Assays (Applied Biosystems) on a 7500 Fast Real-Time PCR System. For rs944797, an in-house TaqMan® assay consisting of 17 μL reaction volume of TaqMan® Universal PCR Master Mix, 31.3 ng of genomic DNA, 590 nM of each primers 5′-TTGGTGGCTTAAAGTTAGGCTGA-3′ (forward), 5′-CCTGGCATAAG TGTTAGTACCCTGT-3′ (reverse), and 290 nM of each probes (wt-5′-FAM-CCCCAGTCTTCCCTCCTTGA-BBQ-3′, mut-5′-YAK-CCCGGTCTTC CCTCCTTGA-BBQ-3′). The DNA amplification started with pre-PCR holding at 60 °C for 1 min, followed by enzyme activation at 95 °C for 10 min, then two-stage polymerase run of 40 cycles at 95 °C for 15 s and 60 °C for 1 min, and a final post-PCR holding at 60 °C for 1 min.

2.4. Statistical analysis

Allelic and genotypic frequencies were estimated by direct count. Hardy-Weinberg equilibrium was determined by Arlequin software v.3.5.1.2 (Schneider et al., 2000), and results were considered statistically significant at $p < 0.05$. Haplotypes were indirectly inferred by Arlequin, using maximum likelihood method, as well as directly by homozygosity analysis. Linkage disequilibrium (LD) analysis was also calculated by Haploview v.4.2 (Barrett et al., 2005). The solid spine of LD method was used to search for strong LD from the first SNP to the last along the legs of the triangle in the LD plot using $D' > 0.8$ to define the LD blocks.

An analysis of genetic profiles for the 3 regions (9p21 + *LDLR* + *USF1*) was carried out based on the reported phenotypic associations and the OR/HR values for populations of European ancestry (Table 1). Two SNPs, one in *LDLR* (rs2738466) and one in *USF1* (rs10908821), were excluded from the analysis, since they are delimiting haplotypes and do not present information for risk alleles. Considering current scientific knowledge about biological pathways, individuals were categorized into three risk groups: 1. high risk for atherosclerosis, 2. low risk for atherosclerosis, and 3. high risk for dyslipidemias.