Disease Markers 32 (2012) 231–239
DOI 10.3233/DMA-2011-0879
IOS Press

Replication of relevant SNPs associated with cardiovascular disease susceptibility obtained from GWAs in a case-control study in a Canarian population

Francisco Rodríguez Esparragón[a,b,1,*], Osmel Companioni[a,1], Miguel García Bello[a], Nisa Buset Ríos[a] and José Carlos Rodríguez Pérez[a,b,c]
[a]*Unidad de Investigación, HIRICARE Study Group, Hospital Universitario de Gran Canaria, Las Palmas de Gran Canaria, Gran Canaria, Spain*
[b]*Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Gran Canaria, Spain*
[c]*Servicio de Nefrología, Hospital Universitario de Gran Canaria, Las Palmas de Gran Canaria, Gran Canaria, Spain*

Abstract. Recent genome-wide single nucleotide polymorphism (SNP) association studies (GWAS) have identified a number of SNPs that were significantly associated with coronary artery disease (CAD) and myocardial infarction (MI). We tested for replication of the previously described association with CAD in our case-control datasets of SNPs variants located at 1p13.1, 2q33.1, 10q11.1, 9p21, and 21q22. We observed a small significant risk associated of the SNP rs10757274 with CAD in the PROCAGENE study. Besides, the multilocus combination rs10757274 and rs1333048 gave a near significant result. We confirmed that the SNP rs10757274 showed association with CAD in the PROCAGENE study, although after applying the Bonferroni correction was not longer significant. Independent replication studies in other populations are needed to unequivocally confirm the association.

Keywords: Genome-wide association studies, polymorphisms, coronary artery disease, myocardial infarction

1. Background

Recent genome wide association studies (GWAS) have shown a consistent association of single nucleotide polymorphisms (SNPs) with coronary artery disease (CAD) and myocardial infarction (MI) in new chromosomal regions mapped at 1p13.1, 2q36.3, 9p21 and 10q11.21 [21,22,27,34]. The SNPs located at the 9p21 region have been successfully replicated in Europeans [31] and also have been associated with abdominal aortic and intracranial aneurysms, arterial stiffness and coronary calcium [8]. The signals in 9p21 reside in a region near CDKN2A/B genes encoding p16INK4a and p15INK4b, both are cyclins which are involved in cell proliferation, aging and tumor suppression [34]. Other GWAS found SNPs associated which are located in other relevant genes to the atherosclerotic process that underlie CAD. However, many independent replication studies in other populations are needed to unequivocally confirm the GWAS association. Here, we carried out a genetic case-control study in subjects from Gran Canaria, Spain, with the aim to replicate in this population SNPs previously associated with CAD and MI in GWAS studies.

2. Methods

2.1. Study population

During 1996–1998, coronary artery disease (CAD) patients and matched controls from Gran Canaria pop-

[1]These authors equally contributed to this work.
[*]Corresponding author. E-mail: frodesp@gobiernodecanarias.org.

ISSN 0278-0240/12/$27.50

ulation were collected for a genetic study named PROCAGEN [26]. A case was defined as an individual with a diagnosis of MI or unstable angina pectoris, as well as evidence of CAD by coronary angiography. All incident cases were recruited during study period. A control was defined as an individual who had not suffered cardiovascular disease and was matched to cases by age and gender. All individuals were of Canary/Spanish ancestry, defined by being born in the Canary Islands or to have Spanish parents who have resided in the islands for over 10 years. Clinical variables measured were: body mass index (BMI), blood pressure, glucose, serum creatinine, total cholesterol and its fractions, triglycerides, homocysteine and urinary albumin. The study was approved by our institutional research ethics committee, and informed consent was obtained from all participants. All study participants were written informed about the future use of DNA samples for further genetic determinations. Researches were carried out in compliance with the Helsinki Declaration (http://www.wma.net/es/30publications/10policies/b3/). The detailed characteristics of the participants of PROCAGENE study have described previously [26].

2.2. SNP selection and genotyping

Single nucleotide polymorphisms were selected from GWAS (http://www.genome.gov/gwastudies/) studies reported to be associated with CAD or MI. Studied variants had minor allele frequency $> 10\%$ according to data from HapMap European population (CEU) and specific articles. All subjects from whom DNA samples were available were expected to be genotyped for the rs1333049, rs10757274, rs599839, rs9982601, rs6725887 and rs501120 genetic variants. For each studied SNP we designed a PCR-RFLP that generates a recognition pattern distinguishable in agarose gels. This task was carried out using DNA sequences flanking the SNP from dbSNP, and the software Primer3 version 0.4.0 (http://frodo.wi.mit.edu/primer3/) to design specific primers, and the software GeneRunner (http://www.generunner.net/) and Webcutter 2.0 (http://bio.lundberg.gu.se/cutter2/) for selection of restriction endonucleases. Because of three SNPs had no recognition sites for restriction enzymes, we looked for tagSNPs in linkage disequilibrium with them. We selected those at the minimum physical distance from the SNP of interest, exploring up to 14 kilobases and ensuring that the chromosomal region of tagSNP belongs to the same gene. This task was performed with the HapMap Genome Browser (http://snp.cshl.org/), Release 24 Phase II, tag SNP data, and the configuration restrictions: $r^2 > 0.9$, minor allele frequency (MAF) > 0.1 in CEU, algorithm Tagger pairwise. After that, we designed a PCR-RFLP as described above. The characteristics of SNPs genotyped, their primers and restriction enzymes are detailed in Table 1.

The PCR reactions were prepared in a final volume of 15μl with the following conditions: 200μM dNTP, 10% DMSO, 1.5 mM MgCl2, 0.5μM primers, BSA 0.8μg/μl, 0.5U Taq DNA Polymerase from Bioline and 100ng of genomic DNA. Amplification steps were an initial denaturalization step of 95°C for 4 minutes followed by 35 cycles of 92°C for 30 seconds, primers-dependent specific Tm annealing for 30 seconds, and 72°C for 15 or 30 seconds; a final extension step was done at 72°C for 7 minutes. Digestion with specific restriction enzymes (NEB, USA) was carried out overnight following conditions from the supplier. All products were visualized by electrophoresis in agarose gels.

2.3. Statistical analysis

Allele and genotype frequencies, Hardy-Weinberg equilibrium and linkage disequilibrium in cases and controls separately were estimated via Pypop software version 0.7.0 [14]. Comparison between cases and controls for quantitative variables was performed by Student-t test or Wilcoxon multiple range test depending on statistical distribution. Qualitative variables were compared by chi-square test. Statistical association of the SNP with susceptibility to CAD was tested by odds ratio with 95% CI. A p value < 0.05 was considered to be statistically significant. Bonferroni correction reduced this value to $p = 0.008$ (0.05/6). Most of the analyses were performed using the statistical package SPSS version 15.0 unless otherwise specified.

In order to detect gene-gene interactions that predispose to disease we used MDR (Multifactor Dimensionality Reduction) software version 2.0 beta 8.2, freely available at http://www.epistasis.org/ [9]. MDR is a non-parametric and model free method that labels multilocus genotypes to high or low risk categories if the ratio of cases/controls in each contingency table cell exceeds some threshold or not. The data are subjected to cross validation by random division into a training and testing set, the model is developed for each 9/10 of individuals and used to predict disease status in the remaining 1/10, then, the selected combination of loci is the one that maximizes balanced testing accuracy

Table 1
Characteristics of the SNPs genotyped in the Procagene population

SNP	GWAS Study	Risk allele	OR (CI 95%) initial study	Chromosome/ Gene	Primers	Enzyme fragments	Restriction
rs1333049 (tagSNP: rs1333048)	(1)	C	1.47 (1.27–1.70)	9p21	ACCCGAAGTAGAGCTGCAAA CACAAGTTGGAATATGAAGCAGA	Dra I	AA: 84 and 68 pb CC: 152 bp
rs10757274 (A/G)	(2)	G	1.38 (1.19–1.60)	9p21	GTTTCTGCACATGGTGATGG CTGCCTCACTCTCCAGTTCC	BsmA I	AA:172 and 78bp GG: 250 bp
rs9982601 C/T	(3)	T	1.2 (1.14–1.27)	21q22.11 SLC5A3/ MRPS6/KCNE2	TATGGAGCCAGGAAGACAGG GCCCATCTGAGTTCCACTGT	Nco I	CC: 209 bp TT: 246 bp
rs6725887 (tagSNP: rs6722332)	(3)	C	1.17 (1.11–1.23)	2q33 WDR12	AAGCAAATCTGAATTCAAACCAA GATGCAGATGCAGAAAGCAG	Dra I	TT: 182 and 30 bp CC:212 bp
rs599839	(4)	A	1.29 (1.18–1.40)	1p13.3 SORT1	CCAGAAGGTGGAAGTTGCAG AGGCACCTGGTCTTCAGAAA	Taq I	AA: 246 bp GG:141 and 105 bp
rs501120 (tagSNP: rs671765) (tagSNP de)	(4)	T	1.33 (1.20-1.48)	10q11.21 CXCL12	TCAAGACTGCCCATTCCTTC CAGACCAGAGAGCCACATGA	Hha I	CC: 142 and 48bp TT: 190 bp

and cross validation consistency. Permutation testing is performed to calculate the statistical significance associated to testing accuracy. We used the default settings and repeated the analysis 10 times with different random seeds to avoid biased results. Multiple comparisons correction is not necessary for MDR because use of permutation testing to estimate p-value [25].

Statistical power was estimated using Quanto software version 1.2.4 downloaded from http://hydra.usc.edu/gxe specifying the following parameters: frequency of risk alleles in controls, unmatched case-control study, and hypothesis gene only, $N = 281$ pairs, population risk ($\mathrm{Kp} = 0.0021$), significance 0.008 and one sided test, genetic models (dominant, recessive and additive). $\mathrm{Kp} = 0.0021$ is the global incidence rate of coronary events reported in the MONICA study (210 per 100 000 individuals). Based on the power analyses we had on average $\sim$63% power to detect a genotype relative risk of 1.5 at an alpha level of 0.008 for variants with a minor allele frequency $\geqslant 10\%$. If all variants were considered the study was unpowered. Multiple comparisons were also corrected by the false discovery rate (FDR), based on the method of Benjamini and Hochberg [2]. To this end q-values were generated from p-values obtained to evaluate the hypotheses of SNP association with phenotypes of interest by using the program QVALUE [32] downloaded from http://www.bioconductor.org and assuming 10% of FDR.

To assess the independent variable predictor ability for CAD, we performed multiple logistic regression analysis. First we made exploratory analysis by backward stepwise regression and tested which coefficients were significant for inclusion or elimination from the model by applying the Hosmer-Lemeshow Goodness-of-Fit Test. As discussed above, some quantitative variables were seen significantly altered by drug treatment or provided redundant information. Those variables were excluded from the constructed model. The model included the classical risk factors whether they differed in univariate analysis with significance as well as those that attained statistical significance in the univariate analysis as well as genotypes. Models were performed coding genotypes according dominant, recessive or additive effects. The contribution of analyzed genotypes and reported CAD risk factors in predicting the dependent variable (coronary event) when all other independent variables were allowed for was expressed as an OR with 95% CIs. All the categorical variables were transformed in dummy variables, taking the value 1 for the presence of the characteristic and 0 for its absence. Taking a cut-off point of 50%, the model with the lower percentage classified correctly 83.3% of individuals. A second group of models are built using only the genotypes coded for dominant inheritance models, recessive or additive and including interactions. In these variables the phenotypic inclusion meant a loss of statistical power and the percentage of overall explanation.

Table 2
Statistical comparison of main clinical variables

Clinical variable	Cases	Controls	p-value
Age (years)	55.90 ± 10.26	54.59 ± 10.99	0.126
BMI (Kg/m2)	27.33 ± 3.70	27.28 ± 3.79	0.809
SBP (mmHg)	135.52 ± 24.21	135.92 ± 25.97	0.944
DBP (mmHg)	76.26 ± 13.63	84.48 ± 12.07	< 0.001
Total cholesterol (mg/dL) (mg/dL)	197.24 ± 42.78	232.09 ± 40.47	< 0.001
Triglycerides (mg/dL)	124.13 ± 67.77	149.71 ± 73.54	< 0.001
HDL-cholesterol (mg/dL)	35.32 ± 9.29	50.10 ± 12.33	< 0.001
LDL-cholesterol (mg/dL)	137.11 ± 37.02	152.12 ± 35.91	< 0.001
Fibrinogen (mg/dL)	327.52 ± 81.08	481.73 ± 302.85	0.030
Lpa (mg/dL)	53.89 ± 68.53	36.51 ± 43.02	< 0.001
Homocysteine (μM)	14.64 ± 5.31	15.80 ± 7.31	0.066
Drinkers	53.38%	47.23%	0.134
Smokers	49.11%	26.71%	< 0.001
Hipertensives	49.11 %	33.87 %	< 0.001
Sex (Men)	77.93 %	73.28 %	0.225
Diabetes mellitus type 2	33.09 %	12.05 %	< 0.001

3. Results

3.1. Study patients, lifestyle and cardiovascular risk factors

As can be observed in Table 2, mean age values, body mass index (BMI), systolic blood pressure (SBP), total homocysteine concentration (tHcy) were homogeneous in our sample. Similarly there were not differences in the frequency of heavy drinkers and the percentage of sex composition between groups. Diastolic blood pressure (DBP), total cholesterol, triglycerides, HDL-cholesterol, LDL-cholesterol and fibrinogen had significant minor levels in coronary artery disease (CAD) patients with respect to healthy control subjects. This pattern it's thought to be because of the efficacy of antihypertensive and antilipidemic medication in our sample [26]. However, Lpa levels and the number of smokers, hypertensive and diabetics were significant higher in cases versus controls.

3.2. Genotypes and association with CAD

All evaluated SNPs except rs1333048 were found to be in Hardy-Weinberg equilibrium in controls. The relative frequencies and association status for different genetic models of the SNPs genotyped are shown in additional Table 3. The unique SNP statistically associated with CAD was rs10757274 with an unadjusted OR of 1.59 (1.01–2.49) for the dominant model. However, this association did not resist either Bonferroni or FDR correction for multiple testing. No significant differences between genotypes were observed among cases and controls for the other polymorphisms. Pypop software showed that SNPs rs10757274 and rs1333048 were in linkage disequilibrium both in controls (D' $=$ 0.807, $W_n = 0.647$, $p < 0.001$) (r statistic $= 0.6474$; $p < 0.001$) and cases (D' $= 0.967$, $W_n = 0.812$, $p <$ 0.001) (r statistic $= 0.8126$; $p < 0.001$), as could be expected due to their proximity in 9p21 (22,086,045 and 22,115,337 base pairs respectively).

As distinctive of our sample with respect to European HapMap population (CEU), we did not observe the genotype AA of the rs1333048 and we did observe the genotype CC of rs671765, both polymorphisms are tagSNPs of rs1333049 and rs501120 respectively. In opposition to the absence of genotype CC (rs671765) in CEU population, its presence in our sample might be explained by genetic flow from sub-Saharan Africa based on genotype frequency (0.22) of African HapMap (YRI) population.

By means of Multifactor Dimensionality Reduction (MDR) software the best combination obtained was rs10757274 and rs1333048 (Fig. 1) with both maximum testing balance accuracy (0.567) and cross validation consistency of 10 which correspond to a p-value in the range of (0.049–0.05) from the permutation test. Other combinations were found not significant. We based this finding primarily on a testing balanced accuracy > 0.55, secondly on a cross validation consistency of 10 and a p-value near significance. In addition, its graphical model shows that distribution of high and low risk genotypes vary across cells and the dendrogram placed both SNPs in the same branch with red color indicating the informativeness of the analysis done with the MRD method to detect such interaction even considering that the inclusion of SNPs that are highly correlated may lead to unstable results. This is

Table 3
Association analysis of SNPs with coronary artery disease

	rs1333048	rs10757274	rs9982601	rs6722332	rs599839	rs671765
Risk allele	C	G	T	C	A	T
Reference allele	A	A	C	T	G	C
Cases ($N = 281$)						
HWE (p-value)	0	0.268	0.975	0.002	0.995	0.215
Risk allele frequency	0,699	0,621	0,121	0,144	0,733	0,840
Reference allele frequency	0,301	0,379	0,879	0,856	0,267	0,160
Controls ($N = 307$)						
HWE (p-value)	0	0.429	0.319	0.99	0.25	0.53
Risk allele frequency	0,681	0,578	0,104	0,127	0,738	0,805
Unadjusted OR (CI 95%)						
Additive model	1.17 (0.83–1.63)	1.19 (0.94–1.51)	1.17 (0.82–1.66)	1.14 (0.83–1.57)	0.97 (0.75–1.26)	1.27 (0.94–1.71)
Dominant model	–	1,59 (1.01–2.49)*	1.29 (0.87–1.93)	0.96 (0.66–1.4)	0.86 (0.47–1.59)	0.91 (0.37–2.23)
Recessive model	–	1,11 (0.79–1.56)/	0.62 (0.18–2.14)/	0.37 (0.13–1.07)/	1.08 (0.78–1.5)/	1.4 (0.99–1.99)/

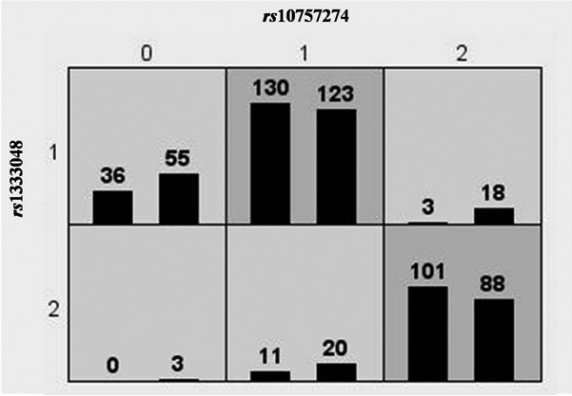

Fig. 1. Multifactor dimensionality reduction reveals gene-gene interactions associated with coronary artery disease susceptibility in Canarian population. Best two-locus multifactor dimensionality reduction (MDR) model. Light gray cells are low risk, while dark gray cells are high risk. The number of cases is shown in the histogram on the left in each cell, while controls are shown by the histogram on the right for each genotype combination.

because highly correlated SNPs may compete for the best predictor.

By using logistic regression we found that models provided a good overall fit of about 83% and identified an associated risk for variant rs10757274 in a dominant mode of inheritance of OR: 2.074 (1.059–4.067). To note that for the remaining evaluated SNPs, we found a modest risk on the threshold of significance. Age, gender, tobacco consumption, alcohol intake, Lpa levels, diabetes and the total cholesterol to HDL-cholesterol levels were all statistical significant predictable variables in analyzed models. In an attempt to approach the evaluation of two causal loci, logistic model were evaluated for interactions. We found that the interaction was significant when genotypes were codes in a recessive form for the SNPs rs10757271 and rs1333048 but showed an imbalance between the certainty of the estimate and the width of a confidence interval 13.428 (3.113 to 57.919).

4. Discussion

We tested for replication of the previously described SNPs to be associated with coronary artery disease (CAD) in our case-control datasets for SNPs variants located at 1p13.1, 2q33.1, 10q11.1, 9p21, and 21q22. We have shown that only a small significant effect was found for the rs10757274 variant under dominant inheritance model. The associated risk with the rs10757274 variant was determined using new bioinformatic approaches and through logistic regression. We also observed by using Multifactor Dimensionality Reduction (MDR) software that the best multilocus model fit with variants located at the 9p21 chromosome region. Interestingly interaction was also detected by logistic regression analysis.

The Spanish rate of CAD is located among the lowest in Europe [3]. However there are important differences between territories being insular regions which possess the higher rates. Among the Canary Islands, CAD prevalence is particularly high Gran Canaria. The present Canarian population is the result of successive European settlements formed through successive waves with a drastic effect on the original Aboriginal population of Berber origin [17]. Although previous studies have demonstrated the presence at low frequencies of Berbers alleles in this population [24], a general consensus today is to consider that it is a European population. We and others have also found the absence of stratification so that the Canarian population is also considered genetically homogeneous [18].

Genetic association studies in this population have attempted to assess the relevance of the association of

certain genetic variants in genes involved in the maintenance of vascular homoeostasis, lipid metabolism or oxidative stress and vascular disease [6,10]. However, for the most of the studied variants risk allele frequencies were found not to be significantly different to those obtained in other studied populations were association was found not to be replicated. This is thought to occur mainly because the lack of statistical power in small case-control studies. The development of high-density genotyping arrays now provides improved resolution for an unbiased genome-wide assessment of common variants associated with common diseases. Several GWAS have found a consistent association of several chromosomal regions with CAD and MI. However, given the polygenic nature of complex traits, selection may exert its influence on them by altering allele frequencies at many associated loci, a possibility which has been recently explored [7]. The most replicated signal obtained from GWAS is located at 9p21.3, near to tumor suppressor genes CDKN2A and CDKN2B [4,5,13,27,29,31]. Interestingly associations are independent of other cardiovascular risk factors. Based on a meta-analysis, 25% of Europeans have 2 copies of the rs1333049 risk allele, which results in an increase of 1.6 of the risk of CAD [29]. The SNPs rs10757274 belongs to the same chromosomal region (9p21.3) of rs1333049. The risk allele (A) of rs599839 is associated with low expression levels of SORT1, high LDL-cholesterol and increased risk of CAD. SORT1 is a multi-ligand receptor that binds RAP (receptor associated protein LDL), lipoprotein lipase and apolipoprotein AV [15]. We evaluated without finding a possible association with CAD of SNPs located on other chromosomes. The rs501120 is located upstream of the CXCL12 gene (factor-1-derived stromal cells) at 10q11.1 and encodes a chemokine with a central role in tissue regeneration and angiogenesis in ischemic heart disease through endothelial progenitor cell recruitment [27]. The signal from the SNP rs9982601 is located in SLC5A3 in the 21q22 region, a gene whose function is to transport Na^+ and myo-inositol in response to hypertonic stress [22]. The rs6725887 is placed at the WDR12, a component of a nucleolar protein complex that affects maturation of the large ribosomal subunit [22].

Here we observed that according to our statistical power study by means of Quanto, the additive genetic model will allows a stable detection of genetic effect for all SNPs at odds ratio (OR) values in the interval 1.34–1.61. The dominant and recessive models are affected for high and low risk allele frequencies respectively. However, with the exception of rs10757274 (dominant model) in single gene analysis we were unable to replicate the previous association results. Moreover when we applied the Bonferroni correction for multiple comparisons even this small significant association disappeared.

We intend to detect multilocus effects in our sample using the MDR, this sofware has identified gene-gene interactions in cardiovascular diseases as hypertension [28,36], atrial fibrillation [35], coronary artery calcification [1], myocardial infarction [19] and thrombotic stroke [30]. MDR was run with different random seeds to detect a true signal independent of how divided the data, although near significance (p-value: 0.049–0.05) the best multilocus combination we found was composed by rs10757274 and rs1333048, these SNPs are in linkage disequilibrium in our sample, and as discussed rs10757274 shows statistical signification before applying multiple comparisons correction.

However, given the weak association obtained the described interaction as well as is biological plausibility our study trend to reaffirm the role of 9p21 zone in the predisposition to CAD. The biological pathway that connect this region with CAD has not been established, but presumably could be due to the participation of the tumor suppressor genes CDKN2A/B that control cell cycle and/or through the ANRIL, a non coding antisense RNA located near from the above genes [23]. Several studies using real time-PCR have shown that expression levels of these three genes are co-regulated and several transcripts of ANRIL are increased in risk allele homozygotes versus the homozygotes for the reference allele, and associated with atherosclerosis severity in blood and carotid plaques [11,12,16]. Also, it was demonstrated that senescence of smooth muscle from coronary endarterectomy is associated with high protein levels of CDKN2A (p16) measured by Western-Blots [20]. The implication of ANRIL in cellular aging has provided an attractive unifying hypothesis to explain its association with various susceptibility risk factors. Other underlying cause of association could be that CAD risk SNPs at 9p21.3 are in linkage disequilibrium with a causal variant in another chromosomal location.

Our study was initially designed for studying CAD in native Gran Canaria Island habitants. The original sample contained 619 subjects of which 304 were cases and 315 were controls. Due to polygenic nature of CAD and even with such a small size, our replication study sample may still have a reasonable power to detect statistical significant association since all selected

SNPs had an a priori minor allele frequency $> 10\%$. Nevertheless, our study has limitations. First, the small size is still an important limitation, which may have limited the statistical power to detect association. Second, we only selected six SNPs in five loci associated with CAD and MI when at least 12 SNPs have been replicated to be associated with CAD and MI and third, in a recent large-scale study [33], authors have identified 13 novel loci harboring one or more SNPs that were associated with CAD at $P < 5 \times 10$–18 and confirmed the association of 10 of 12 previously reported CAD loci.

5. Conclusions

Our main conclusion was that SNP rs10757274 showed association with coronary artery disease in the PROCAGENE study, although after applying the Bonferroni correction was not longer significant. Besides, the multilocus combination rs10757274 and rs1333048 gave a near significant result. Taking into account our small sample size we consider these results as exploratory and preliminary, and follow up studies are required. Replication studies are important because they confirm a modest risk in genetic poblational series even with small sample sizes ensuring the certainty of the associations in different populations and the future reliability in the setting of cardiovascular risk equations.

Acknowledgements

We thank all patients and their families for their help and generosity. Lidia Estupiñán and Erika Hernández kindly helped us with the genetic studies. This work was supported in part by a grant from Caja Rural de Canarias and Fundación Mapfre-Guanarteme.

Author disclosure statement

No competing financial interests exist.

References

[1] L. Bastone, M. Reilly, D.J. Rader and A.S. Foulkes, MDR and PRP: a comparison of methods for high-order genotype-phenotype associations, Hum Hered, **58** (2004), 82-92.

[2] Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing, Journal of the Royal Statistical Society, Series B, **57** (1995), 289-300.

[3] M.R. Boix, S.N. Aragones and M.J. Medrano Albero, Trends in mortality from ischemic heart disease in 50 Spanish provinces, Rev Esp Cardiol, **56** (2003), 850-856.

[4] S.N. Chen, C.M. Ballantyne, A.M. Jr. Gotto and A.J. Marian, The 9p21 susceptibility locus for coronary artery disease and the severity of coronary atherosclerosis, BMC Cardiovasc Disord **9** (2009), 3.

[5] X. Cheng, L. Shi, S. Nie. F. Wang, X. Li, C. Xu, P. Wang, B. Yang, Q. Li, Z. Pan, Y. Li, H. Xia, C. Zheng, Y. Ke, Y. Wu, T. Tang, X. Yan, Y. Yang, N. Xia, R. Yao, B. Wang, X. Ma, Q. Zeng, X. Tu, Y. Liao and Q.K. Wang, The same chromosome 9p21.3 locus is associated with type 2 diabetes and coronary artery disease in a Chinese Han population; Diabetes **60** (2011), 680-684.

[6] H.M. Colhoun, P.M. McKeigue and S.G. Davey, Problems of reporting genetic associations with complex outcomes, Lancet **361** (2003),865-872.

[7] K. Ding and I.J. Kullo, Geographic differences in allele frequencies of susceptibility SNPs for cardiovascular disease, BMC Med Genet **12** (2011), 55.

[8] D. Girelli, N. Martinelli, F. Peyvandi and O.Olivieri, Genetic architecture of coronary artery disease in the genome-wide era: implications for the emerging "golden dozen" loci, Semin Thromb Hemost **35** (2009), 671-682.

[9] L.W. Hahn, M.D. Ritchie and J.H. Moore, Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions, Bioinformatics **19** (2003), 376-382.

[10] A.T. Hattersley and M.I. McCarthy, What makes a good genetic association study?, Lancet **366** (2005), 1315-1323.

[11] L.M. Holdt, F. Beutner, M. Scholz, S. Gielen, G. Gäbel, H. Bergert, G. Schuler, J. Thiery and D. Teupser, ANRIL expression is associated with atherosclerosis risk at chromosome 9p21, Arterioscler Thromb Vasc Biol **30** (2010), 620-627.

[12] O. Jarinova, A.F. Stewart, R. Roberts, G. Wells, P. Lau, T. Naing, C. Buerki, B.W. McLean, R.C. Cook, J.S. Parker and R. McPherson, Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus, Arterioscler Thromb Vasc Biol **29** (2009), 1671-1677.

[13] J. Kumar, S. Yumnam, T. Basu, A. Ghosh, G. Garg, G. Karthikeyan and S. Sengupta, Association of polymorphisms in 9p21 region with CAD in North Indian population: replication of SNPs identified through GWAS, Clin Genet **79** (2011), 588-593.

[14] A.K. Lancaster, R.M. Single, O.D. Solberg, M.P. Nelson and G. Thomson, PyPop update–a software pipeline for large-scale multilocus population genomics, Tissue Antigens **69** Suppl 1 (2007),192-197.

[15] P. Linsel-Nitschke, J. Heeren, Z. Aherrahrou, P. Bruse, C. Gieger, T. Illig, H. Prokisch, K. Heim, A. Doering, A. Peters, T. Meitinger, H.E. Wichmann, A. Hinney, T. Reinehr, C. Roth, J.R. Ortlepp, M. Soufi, A.M. Sattler, J. Schaefer, K. Stark, C. Hengstenberg, A. Schaefer, S. Schreiber, F. Kronenberg, N.J. Samani, H. Schunkert and J. Erdmann, Genetic variation at chromosome 1p13.3 affects sortilin mRNA expression, cellular LDL-uptake and serum LDL levels which translates to the risk of coronary artery disease, Atherosclerosis **208** (2010), 183-189.

[16] Y. Liu, H.K. Sanoff, H. Cho, C.E. Burd, C. Torrice, K.L. Mohlke, J.G. Ibrahim and N.E. Thomas, INK4/ARF tran-

script expression is associated with chromosome 9p21 variants linked to atherosclerosis, PLoS One **4** (2009), e5027.

[17] N. Maca-Meyer, J. Villar, L. Perez-Mendez, A. Cabrera de León and C. Flores, A tale of aborigines, conquerors and slaves: Alu insertion polymorphisms and the peopling of Canary Islands. Ann Hum Genet **68** (2004), 600-605.

[18] A. Macias-Reyes, F. Rodriguez-Esparragon, A. Caballero-Hidalgo, Y. Hernández-Trujillo, A. Medina and J.C. Rodríguez-Pérez, Insight into the role of CYBA A640G and C242T gene variants and coronary heart disease risk. A case-control study, Free Radic Res **42** (2008), 82-92.

[19] M.N. Mannila, R.S. Lovely, S.C. Kazmierczak, P. Eriksson, A. Samnegrd, D.H. Farrell and A. Hamsten, Elevated plasma fibrinogen gamma' concentration is associated with myocardial infarction: effects of variation in fibrinogen genes and environmental factors, J Thromb Haemost **5** (2007), 766-773.

[20] C. Matthews, I. Gorenne, S. Scott, N. Figg, P. Kirkpatrick, A. Ritchie, M. Goddard and M. Bennett, Vascular smooth muscle cells undergo telomere-based senescence in human atherosclerosis: effects of telomerase and oxidative stress, Circ Res **99** (2006), 156-164.

[21] R. McPherson, A. Pertsemlidis, N. Kavaslar, A. Stewart, R. Roberts, D.R. Cox, D.A. Hinds, L.A. Pennacchio, A. Tybjaerg-Hansen, A.R. Folsom, E. Boerwinkle, H.H. Hobbs and J.C. Cohen, A common allele on chromosome 9 associated with coronary heart disease, Science **316** (2007), 1488-1491.

[22] Myocardial Infarction Genetics Consortium, S. Kathiresan, B.F. Voight, S. Purcell, K. Musunuru, D. Ardissino, P.M. Mannucci, S. Anand, J.C. Engert, N.J. Samani, H. Schunkert, J. Erdmann, M.P. Reilly, D.J. Rader, T. Morgan, J.A. Spertus, M. Stoll, D. Girelli, P.P. McKeown, C.C. Patterson, D.S. Siscovick, C.J. O'Donnell, R. Elosua, L. Peltonen, V. Salomaa, S.M. Schwartz, O. Melander, D. Altshuler, D. Ardissino, P.A. Merlini, C. Berzuini, L. Bernardinelli, F. Peyvandi, M. Tubaro, P. Celli, M. Ferrario, R. Fetiveau, N. Marziliano, G. Casari, M. Galli, F. Ribichini, M. Rossi, F. Bernardi, P. Zonzin, A. Piazza, P.M. Mannucci, S.M. Schwartz, D.S. Siscovick, J. Yee, Y. Friedlander, R. Elosua, J. Marrugat, G. Lucas, I. Subirana, J. Sala, R. Ramos, S. Kathiresan, J.B. Meigs, G. Williams, D.M. Nathan, C.A. MacRae, C.J. O'Donnell, V. Salomaa, A.S. Havulinna, L. Peltonen, O. Melander, G. Berglund, B.F. Voight, S. Kathiresan, J.N. Hirschhorn, R. Asselta, S. Duga, M. Spreafico, K. Musunuru, M.J. Daly, S. Purcell, B.F. Voight, S. Purcell, J. Nemesh, J.M. Korn, S.A. McCarroll, S.M. Schwartz, J. Yee, S. Kathiresan, G. Lucas, I. Subirana, R. Elosua, A. Surti, C. Guiducci, L. Gianniny, D. Mirel, M. Parkin, N. Burtt, S.B. Gabriel, N.J. Samani, J.R. Thompson, P.S. Braund, B.J. Wright, A.J. Balmforth, S.G. Ball and A.S. Hall, Wellcome Trust Case Control Consortium, H. Schunkert, J. Erdmann, P. Linsel-Nitschke, W. Lieb, A. Ziegler, I. König, C. Hengstenberg, M. Fischer, K. Stark, A. Grosshennig, M. Preuss, H.E. Wichmann, S. Schreiber, H. Schunkert, N.J. Samani, J. Erdmann, W. Ouwehand, C. Hengstenberg, P. Deloukas, M. Scholz, F. Cambien, M.P. Reilly, M. Li, Z.Chen, R. Wilensky, W. Matthai, A. Qasim, H.H. Hakonarson, J. Devaney, M.S. Burnett, A.D. Pichard, K.M. Kent, L. Satler, J.M. Lindsay, R. Waksman, C.W. Knouff, D.M. Waterworth, M.C. Walker, V. Mooser, S.E. Epstein, D.J. Rader, T. Scheffold, K. Berger, M. Stoll, A. Huge, D. Girelli, N. Martinelli, O. Olivieri, R. Corrocher, T. Morgan, J.A. Spertus, P. McKeown, C.C. Patterson, H. Schunkert, E. Erdmann, P. Linsel-Nitschke, W. Lieb, A. Ziegler, I.R. König, C. Hengstenberg, M. Fischer, K. Stark, A. Grosshennig, M. Preuss, H.E. Wichmann, S. Schreiber, H. Hólm, G. Thorleifsson, U. Thorsteinsdottir, K. Stefansson, J.C. Engert, R. Do, C. Xie, S. Anand, S. Kathiresan, D. Ardissino, P.M. Mannucci, D. Siscovick, C.J. O'Donnell, N.J. Samani, O. Melander, R. Elosua, L. Peltonen, V. Salomaa, S.M. Schwartz and D. Altshuler, Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet **41** (2009), 334-341.

[23] E. Pasmant, A. Sabbagh, M. Vidaud and I.Bièche, ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS, FASEB J **25** (2011), 444-448.

[24] M. Pino-Yanes, A. Corrales, S. Basaldua, A. Hernández, L. Guerra, J. Villar and C. Flores, North African influences and potential bias in case-control association studies in the Spanish population, PLoS One **6** (2011), e18389.

[25] M.D. Ritchie, L.W. Hahn, N. Roodi, L.R. Bailey, W.D. Dupont, F.F. Parl and J.H. Moore, Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer, Am J Hum Genet **69** (2001), 138-147.

[26] J.C. Rodriguez-Perez, F. Rodriguez-Esparragon, O. Hernandez-Perera, A. Anabitarte, A. Losada, A. Medina, E. Hernández, D. Fiuza, O. Avalos, C. Yunis and C.M. Ferrario, Association of angiotensinogen M235T and A(-6)G gene polymorphisms with coronary heart disease with independence of essential hypertension: the PROCAGENE study. Prospective Cardiac Gene, J Am Coll Cardiol **37** (2001), 1536-1542.

[27] N.J. Samani, J. Erdmann, A.S. Hall, C.Hengstenberg, M. Mangino, B. Mayer, R.J. Dixon, T. Meitinger, P. Braund, H.E. Wichmann, J.H. Barrett, I.R. König, S.E. Stevens, S. Szymczak, D.A. Tregouet, M.M. Iles, F. Pahlke, H. Pollard, W. Lieb, F. Cambien, M. Fischer, W. Ouwehand, S. Blankenberg, A.J. Balmforth, A. Baessler, S.G. Ball, T.M. Strom, I. Braenne, C. Gieger, P. Deloukas, M.D. Tobin, A. Ziegler, J.R. Thompson and H. Schunkert, WTCCC and the Cardiogenics Consortium, Genomewide association analysis of coronary artery disease, N Engl J Med **357** (2007), 443-453.

[28] H. Sanada, J. Yatabe, S. Midorikawa, S. Hashimoto, T. Watanabe, J.H. Moore, M.D. Ritchie, S.M. Williams, J.C. Pezzullo, M. Sasaki, G.M. Eisner, P.A. Jose and R.A. Felder, Single-nucleotide polymorphisms for diagnosis of salt-sensitive hypertension, Clin Chem **52** (2006), 352-360.

[29] H. Schunkert, A. Gotz, P. Braund, R. McGinnis, D.A. Tregouet, M. Mangino, P. Linsel-Nitschke, F. Cambien, C. Hengstenberg, K. Stark, S. Blankenberg, L. Tiret, P. Ducimetiere, A. Keniry, M.J. Ghori, S. Schreiber, N.E. El Mokhtari, A.S. Hall, R.J. Dixon, A.H. Goodall, H. Liptau, H. Pollard, D.F. Schwarz, L.A. Hothorn, H.E. Wichmann, I.R. König, M. Fischer, C. Meisinger, W. Ouwehand, P. Deloukas, J.R. Thompson, J. Erdmann, A. Ziegler and N.J. Samani, Cardiogenics Consortium, Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease, Circulation **117** (2008), 1675-1684.

[30] C.D. Shen, W.L. Zhang, K. Sun, Y.B. Wang, Y.S. Zhen and R.T. Hui, Interaction of genetic risk factors confers higher risk for thrombotic stroke in male Chinese: a multicenter case-control study, Ann Hum Genet **71** (2007), 620-629.

[31] G.Q. Shen, S. Rao, N. Martinelli, L.Li, O. Olivieri, R. Corrocher, K.G. Abdullah, S.L. Hazen, J. Smith, J. Barnard, E.F. Plow, D. Girelli and Q.K. Wang, Association between four SNPs on chromosome 9p21 and myocardial infarction is replicated in an Italian population, J Hum Genet **53** (2008), 144-150.

[32] J.D. Storey, A direct approach to false discovery rates, Journal of the Royal Statistical Society, Series B **64** (2002), 479-498.

[33] H. Schunkert, I.R. Konig, S. Kathiresan, M.P. Reilly, T.L. Assimes, H. Holm, M. Preuss, A.F. Stewart, M. Barbalic, C. Gieger, D. Absher, Z. Aherrahrou, H. Allayee, D. Altshuler, S.S. Anand, K. Andersen, J.L. Anderson, D. Ardissino, S.G. Ball, A.J. Balmforth, T.A. Barnes, D.M. Becker, L.C. Becker, K. Berger, J.C. Bis, S.M. Boekholdt, E. Boerwinkle, P.S. Braund, M.J. Brown, M.S. Burnett, I. Buysschaert, J.F. Carlquist, L. Chen, S. Cichon, V. Codd, R.W. Davies, G. Dedoussis, A. Dehghan, S. Demissie, J.M. Devaney, P. Diemert, R. Do, A. Doering, S. Eifert, N.E. Mokhtari, S.G. Ellis, R. Elosua, J.C. Engert, S.E. Epstein, U. de Faire, M. Fischer, A.R. Folsom, J. Freyer, B. Gigante, D. Girelli, S. Gretarsdottir, V. Gudnason, J.R. Gulcher, E. Halperin, N. Hammond, S.L. Hazen, A. Hofman, B.D. Horne, T. Illig, C. Iribarren, G.T. Jones, J.W. Jukema, M.A. Kaiser, L.M. Kaplan, J.J. Kastelein, K.T. Khaw, J.W. Knowles, G. Kolovou, A. Kong, R. Laaksonen, D. Lambrechts, K. Leander, G. Lettre, M. Li, W. Lieb, C. Loley, A.J. Lotery, P.M. Mannucci, S. Maouche, N. Martinelli, P.P McKeown, C. Meisinger, T. Meitinger, O. Melander, P.A. Merlini, V. Mooser, T. Morgan, T.W. Muhleisen, J.B. Muhlestein, T. Munzel, K. Musunuru, J. Nahrstaedt, C.P. Nelson, M.M. Nothen, O. Olivieri, R.S. Patel, C.C. Patterson, A. Peters, F. Peyvandi, L. Qu, A.A. Quyyumi, D.J. Rader, L.S. Rallidis, C. Rice, F.R: Rosendaal, D. Rubin, V. Salomaa, M.L. Sampietro, M.S. Sandhu, E. Schadt, A. Schafer, A. Schillert, S. Schreiber, J. Schrezenmeir, S.M. Schwartz, D.S. Siscovick, M. Sivananthan, S. Sivapalaratnam, A. Smith, T.B. Smith, J.D. Snoep, N. Soranzo, J.A. Spertus, K. Stark, K. Stirrups, M. Stoll, W.H. Tang, S. Tennstedt, G. Thorgeirsson, G. Thorleifsson, M. Tomaszewski, A.G. Uitterlinden, A.M. van Rij, B.F. Voight, N.J. Wareham, G.A. Wells, H.E. Wichmann, P.S. Wild, C. Willenborg, J.C. Witteman, B.J. Wright, S. Ye, T. Zeller, A. Ziegler, F. Cambien, A.H. Goodall, L.A. Cupples, T. Quertermous, W. Marz, C. Hengstenberg, S. Blankenberg, W.H. Ouwehand, A.S. Hall, P. Deloukas, J.R. Thompson, K. Stefansson, R. Roberts, U. Thorsteinsdottir, C.J. O'Donnell, R. McPherson, J. Erdmann and N.J. Samani, Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease, Nat Genet **43** (2011), 333-338.

[34] Wellcome Trust Case Control Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls, Nature, **447** (2007), 661-678.

[35] C.T. Tsai, L.P. Lai, J.L. Lin, F.T. Chiang, J.J. Hwang, M.D. Ritchie, J.H. Moore, K.L. Hsu, C.D. Tseng, C.S. Liau and Y.Z. Tseng, Renin-angiotensin system gene polymorphisms and atrial fibrillation, Circulation **109** (2004), 1640-1646.

[36] S.M. Williams, M.D. Ritchie, J.A. Phillips, III E. Dawson, M. Prince, E. Dzhura, A. Willis, A. Semenya, M. Summar, B.C. White, J.H. Addy, J. Kpodonu, L.J. Wong, R.A. Felder, P.A. Jose, J.H. Moore, Multilocus analysis of hypertension: a hierarchical approach, Hum Hered **57** (2004), 28-38..