ORIGINAL ARTICLE

Transcription factor 21 gene and prognosis in a coronary population

Marina Raquel Santos, Maria Isabel Mendonça, Margarida Temtem, Débora Sá, Ana Célia Sousa, Sónia Freitas, Mariana Rodrigues, Sofia Borges, Graça Guerra, Ilídio Ornelas, António Drumond, Roberto Palma dos Reis

a Centro de Investigação Dra. Maria Isabel Mendonça, Hospital Dr. Nélio Mendonça, SESARAM EPERAM, Funchal, Portugal
b Serviço de Cardiologia, Hospital Dr. Nélio Mendonça, SESARAM EPERAM, Funchal, Portugal
c NOVA Medical School, Faculdade de Ciências Médicas, Lisboa, Portugal

Received 16 November 2022; accepted 27 February 2023
Available online 29 June 2023

KEYWORDS
Transcription factor 21; Major adverse cardiovascular events; Coronary artery disease; Genetic susceptibility; Risk factors

Abstract

Introduction and Objectives: Transcription factor 21 (TCF21) is a member of the basic helix-loop-helix (bHLH) transcription factor family, and is critical for embryogenesis of the heart. It regulates differentiation of epicardium-derived cells into smooth muscle cell (SMC) and fibroblast lineages. The biological role of TCF21 in the progression of atherosclerosis is the subject of debate. The aim of this study was to investigate the impact of the TCF21 rs12190287 gene variant on the prognosis of coronary artery disease (CAD) in a Portuguese population from Madeira island.

Methods: We analyzed major adverse cardiovascular events (MACE) in 1713 CAD patients, mean age 53.3±7.8, 78.7% male, for 5.0±4.3 years. Genotype and allele distribution between groups with and without MACE was determined. The dominant genetic model (heterozygous GC plus homozygous CC) was used and compared with the wild GG to assess survival probability. Cox regression with risk factors and genetic models assessed variables associated with MACE. Kaplan-Meier analysis was used to estimate survival.

Results: The wild homozygous GG, heterozygous GC and risk CC genotypes were found in 9.5%, 43.2% and 47.3% of the population, respectively. The dominant genetic model remained in the equation as an independent risk factor for MACE (HR 1.41; p=0.033), together with multivessel disease, chronic kidney disease, low physical activity and type 2 diabetes. The C allele in the dominant genetic model showed worse survival (22.5% vs. 44.3%) at 15 years of follow-up.
Introduction

Coronary artery disease (CAD) is a chronic multifactorial disorder of the coronary arteries that develops silently and generally presents at an advanced stage by the time symptoms appear.\(^1\)\(^2\) It is thus crucial to promptly identify individuals with a high risk of developing this condition. Both genetic and environmental factors contribute to disease progression. Extensive epidemiological research has established the importance of traditional risk factors such as diabetes, hyperlipidemia and hypertension in the development of CAD.\(^3\)\(^4\) However, some patients can develop vascular disease despite these conventional risk factors being controlled or absent.

Genome-wide association studies (GWAS) have found several genes strongly associated with CAD,\(^5\)\(^6\) 33 of which were formally investigated in the GENEMACOR study designed to investigate the genetic profile of a Portuguese population.\(^7\) One of the most significant genes in our population was TCF21, which encodes transcription factor 21 (TCF21), a transcription factor belonging to the cell-type-specific class II basic helix-loop-helix (bHLH) family at chromosome 6q23.2. Recent research shows that vascular smooth muscle cells (VSMCs) can de-differentiate, multiply and migrate in a process known as phenotypic modulation in response to various stimuli. This mechanism influences the evolution of human atherosclerosis. Based on experimental studies, Wirka et al.,\(^8\) demonstrated a specific knockout of TCF21 in mice with loss of TCF21 expression that significantly inhibited SMC phenotypic modulation, leading to fewer fibroblasts within lesions and plaque instability.\(^9\) The G>C TCF21 rs12190287 single-nucleotide polymorphism (SNP) has allelic specificity, with the risk allele C inducing reduced transcriptional activity, in contrast to the protective G allele, which increases this activity. This SNP could therefore affect regulation of gene expression.\(^9\)\(^10\) Changes in the expression and functionality of target RNAs could play a role in the occurrence of adverse cardiovascular events after a first acute coronary syndrome and potentially be a marker of genetic susceptibility to CAD. Loss of TCF21 function has also been associated with epigenetic alterations in different types of human cancers.\(^11\)
However, the biological role of TCF21 in determining the fate of epicardial-derived cells and the progression of atherosclerosis remains controversial.

**Objective**

We aimed to investigate the impact of the TCF21 rs12190287 gene variant on the prognosis and progression of atherosclerosis in a coronary disease cohort in the Madeira archipelago.

**Methods**

**Study population**

The study population consisted of 1713 CAD patients from the coronary arm of the GENEMACOR study (mean age 53.3±7.8 years, 78.7% male) prospectively followed for 5.0±4.3 years. Our cohort included individuals with at least 70% stenosis in one or more main coronary arteries (or their primary branches) on coronary angiography. Patient post-acute coronary syndrome (ACS) was considered eligible to enter the study after stabilization and hospital discharge in the chronic phase (≥6 months after the acute event). The study protocol was approved by the hospitals’ ethics committees under protocol 50/2012. Written informed consent was obtained from all subjects.

**Patient outcomes**

A composite of recurrent major adverse cardiovascular events (MACE) was defined as recurrent ACS (myocardial infarction [MI] or unstable angina), coronary revascularization (percutaneous or surgical), hospitalization (readmission due to heart failure or ischemic stroke) and cardiovascular mortality.

MI was defined as typical chest discomfort or pain with elevated creatinine kinase-myocardial band (CK-MB) (>1.5 times the upper limit of normal) and cardiac troponin levels above the average upper limit. Unstable angina was defined as an episode of typical discomfort or pain at rest for more than 10 min or two episodes persisting for more than 5 min, with negative cardiac biomarkers. Changes in the electrocardiogram may improve the specificity of this definition.

Coronary revascularization was defined as any percutaneous coronary intervention or coronary artery bypass graft procedure performed even in the absence of MI.

Renal function was assessed using the Cockcroft-Gault formula to estimate creatinine clearance and hence glomerular filtration rate (GFR), which was considered low if <60 ml/min/1.73 m². Multivessel disease was defined as the presence of ≥70% luminal stenosis in ≥2 major epicardial arteries. Left main disease was considered significant with ≥50% stenosis. In patients suffering multiple events, only the time of the first event was used for further analyses. Two independent cardiologists validated all potential cardiovascular events based on discharge reports or death with a detailed summary of the circumstances.

The criteria used for cardiovascular mortality were those of the International Classification of Diseases 10th Revision (ICD-10) codes I00-I25, I27, I30-I52, and I60-I72.

**Genetic analysis**

Of the 33 SNPs significantly associated with CAD in the GENEMACOR study, we only considered the TCF21 rs12190287 gene variant for the present study.

Genotyping was performed in peripheral leukocyte DNA using the Taqman genotyping assay (Applied Biosystems, USA) with allelic discrimination. Labeled probes and primers pre-established by the supplier (7300 Real-Time PCR System, Applied Biosystems, USA) were used without prior knowledge of the individual’s clinical data. The quality of the genotyping technique was checked by including one non-template control in each plate of 96 wells and a blind duplicate, accounting for 20% of all samples.

**Statistical analysis**

Quantitative data were compared between subjects with and without MACE by the unpaired Student’s t test. A one-way ANOVA test was applied to analyze the differences in the TCF21 variant between different genotypes. Categorical data were compared using the chi-square test, and the gene counting method was used to estimate allelic frequencies. Hardy-Weinberg equilibrium was tested under the null hypothesis of the predicted segregation ratio of specific matching genotypes (>0.05) using the chi-square goodness-of-fit test. TCF21 variants were assessed according to all genetic models: dominant (0, wild-type homozygous vs. 1, heterozygous wild-type plus risk homozygous), recessive (0, wild-type homozygous plus risk heterozygous vs. 1, risk homozygous) and additive (0, heterozygous vs. 1, wild-type homozygous). After consideration of all the genetic models, the dominant model was selected and used in the subsequent statistical analysis. Cox regression analysis adjusted for confounders was performed to assess independent variables associated with prognosis. All the risk factors for MACE were entered in the model: age, gender, body mass index, smoking, dyslipidemia, physical inactivity, kidney failure, heart failure, multivessel disease and the TCF21 dominant model. MACE were the dependent variable in this analysis. Hazard ratios (HR) and corresponding 95% confidence intervals (CI) were estimated. The Kaplan-Meier method was used to estimate survival. The statistical analysis was carried out using IBM SPSS version 25.0 (IBM, Armonk, NY, USA). All p-values were two-sided and considered statistically significant for p<0.05.

**Results**

**Patient characteristics**

Demographic, behavioral and clinical characteristics of the 1713 coronary patients with and without MACE are described in Table 1. We noticed a highly statistically significant difference (p<0.0001) between the two groups regarding physical inactivity, diabetes, renal impairment and multivessel dis-
ease, which were more prevalent in subjects with MACE. This group was also older (p=0.014) and more likely to be hypertensive (p=0.003) and dyslipidemic (p=0.025). By contrast, smoking was more predominant in subjects without events (p=0.006).

Genetic analysis of the TCF21 rs12190287 G>C variant

No significant deviation from Hardy-Weinberg equilibrium was found, according to the chi-square goodness-of-fit chi-square test (chi-square=0.103; p=0.05).

Comparison of the distribution of the TCF21 rs12190287 G>C genotype between the two groups studied (with and without MACE) revealed a significant difference, with a higher percentage of the CC risk genotype in the MACE group (50.3% vs. 45.7%). By contrast, in the non-MACE group, the GG genotype was predominant, with 10.7% vs. 7.2% (p=0.032).

As expected, the same trend persisted regarding allele distributions, with 71.6% of the C risk allele in the MACE group vs. 67.5% in the group without events. Similarly, the prevalence of the G allele was 28.4% in the MACE group and 32.5% in the non-MACE group (p=0.014) (Table 2).

**Table 1** Demographic and clinical characteristics of the study population.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total (n=1713)</th>
<th>MACE (n=584)</th>
<th>No MACE (n=1129)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>53.3±7.8</td>
<td>54.0±7.5</td>
<td>53.0±8.0</td>
<td>0.014</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>1348 (78.8%)</td>
<td>459 (78.6%)</td>
<td>889 (78.7%)</td>
<td>0.944</td>
</tr>
<tr>
<td>Physical inactivity, n (%)</td>
<td>1073 (62.6%)</td>
<td>417 (71.4%)</td>
<td>656 (58.1%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Smoking status, n (%)</td>
<td>807 (47.1%)</td>
<td>248 (42.5%)</td>
<td>559 (49.5%)</td>
<td>0.006</td>
</tr>
<tr>
<td>Hypertension, n (%)</td>
<td>1216 (71.0%)</td>
<td>441 (75.5%)</td>
<td>775 (68.6%)</td>
<td>0.003</td>
</tr>
<tr>
<td>Diabetes, n (%)</td>
<td>541 (34.0%)</td>
<td>238 (42.3%)</td>
<td>303 (29.4%)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dyslipidemia, n (%)</td>
<td>1526 (89.1%)</td>
<td>534 (91.4%)</td>
<td>992 (87.9%)</td>
<td>0.025</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>28.7±4.4</td>
<td>28.7±4.2</td>
<td>28.7±4.5</td>
<td>0.939</td>
</tr>
<tr>
<td>CAD family history, n (%)</td>
<td>411 (24.0%)</td>
<td>148 (25.3%)</td>
<td>263 (23.3%)</td>
<td>0.347</td>
</tr>
<tr>
<td>Creatinine clearance &lt;60 ml/min, n (%)</td>
<td>125 (7.3)</td>
<td>65 (11.1)</td>
<td>60 (5.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Multivessel disease, n (%)</td>
<td>820 (49.2%)</td>
<td>365 (63.3%)</td>
<td>455 (41.8%)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

BMI: body mass index; CAD: coronary artery disease; MACE: major adverse cardiovascular events. Statistically significant for p<0.05.

**Table 2** TCF21 genotypes and allele distribution in groups with and without major adverse cardiovascular events.

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Total (n=1713)</th>
<th>MACE (n=584)</th>
<th>No MACE (n=1129)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>G/G</td>
<td>163 (9.5%)</td>
<td>42 (7.2%)</td>
<td>121 (10.7%)</td>
<td>0.032</td>
</tr>
<tr>
<td>G/C</td>
<td>740 (43.2%)</td>
<td>248 (42.5%)</td>
<td>492 (43.6%)</td>
<td>0.003</td>
</tr>
<tr>
<td>C/C</td>
<td>810 (47.3%)</td>
<td>294 (50.3%)</td>
<td>516 (45.7%)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alleles</th>
<th>Total (n=1713)</th>
<th>MACE (n=584)</th>
<th>No MACE (n=1129)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>1066 (31.13%)</td>
<td>332 (28.4%)</td>
<td>734 (32.5%)</td>
<td>0.014</td>
</tr>
<tr>
<td>C</td>
<td>2360 (68.93%)</td>
<td>836 (71.6%)</td>
<td>1524 (67.5%)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

MACE: major adverse cardiovascular events. Statistically significant for p<0.05.

Independent risk factors for major adverse cardiovascular events

On multivariate analysis to predict risk factors for developing MACE, the most significant were multivessel disease (HR 1.81; p<0.0001), the TCF21 dominant genetic model (HR 1.41; p=0.033), physical inactivity (HR 1.40; p<0.0001), kidney failure (HR 1.36; p=0.021) and diabetes (HR 1.24; p=0.013) (Table 3).

Survival analysis: Kaplan-Meier estimator

We used the Kaplan-Meier method to compare the dominant genetic model (GC+CC) with the wild-type GG genotype to assess patients’ probability of survival.

C allele carriers (GC+CC) showed worse survival throughout the follow-up period (p=0.015). Specifically, at 15-year follow-up, GC+CC showed shorter survival (22.5%) compared to the wild-type GG (44.3%) (Figure 1).

Discussion

The interaction between environmental and genetic factors contributes to the pathogenesis of coronary heart disease, and the discovery of new loci and mechanisms will improve the ability to predict cardiovascular risk. Identifying population subgroups that would benefit from lifestyle modification...
Table 3 Independent risk factors for major adverse cardiovascular events.

<table>
<thead>
<tr>
<th>Variables</th>
<th>HR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivessel disease</td>
<td>1.809 (1.525−2.146)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TCF21 (GC+CC)</td>
<td>1.408 (1.027−1.930)</td>
<td>0.033</td>
</tr>
<tr>
<td>Physical inactivity</td>
<td>1.402 (1.168−1.682)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Creatinine clearance &lt;60 ml/min</td>
<td>1.357 (1.047−1.760)</td>
<td>0.021</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1.237 (1.046−1.464)</td>
<td>0.013</td>
</tr>
</tbody>
</table>

CI: confidence interval; HR: hazard ratio.

Cox regression analysis, forward Wald method (IBM SPSS v. 25.0). Hypertension, smoking status and dyslipidemia did not remain in the equation.

Statistically significant for p<0.05.

Figure 1 Kaplan-Meier survival time plot for major adverse cardiovascular events (MACE) according to TCF21 gene variant (GC+CC vs. GG).

In our study, estimated GFR in the MACE group was low (<60 ml/min), and it is known that GFR is an important marker of renal impairment and a risk factor for adverse cardiac events. Inrig et al. 21 stated that ~40% of patients who present with CAD have some degree of renal impairment. 21, 22

As expected, hypertension was more frequent in MACE subjects, as it accelerates the atherosclerotic process, promoting premature CAD and its common sequelae. It is in agreement with Kannel’s findings, indicating that preventive management of hypertension should be broadened and requires consideration of the individual’s complete cardiovascular risk profile. 23

However, the most statistically significant risk factors for MACE development were multivessel disease, physical inactivity, renal impairment, diabetes, and an unfavorable genetic profile with a risk C allele. GWAS and post-GWAS studies have led to substantial progress in understanding the genetic architecture of atherosclerosis progression. Integrating additional ‘omic’ data (genetic and epigenetic), individual clinical data, and environmental influences could improve disease risk stratification and preventive measures. 24, 25

Our results showed a statistically significant difference between the two study groups (with and without MACE) in the frequency of the TCF21 rs12190287 variant, with a higher frequency of the GG native genotype in the non-MACE group. By contrast, the CC genotype presented a higher frequency in the MACE group. Our results are in line with those of Miller et al., 26 who, almost a decade ago, investigated this variant, located within the 3′-UTR region of the TCF21 gene. They deduced that GC alleles of the TCF21 polymorphism altered the microRNA (miRNA)-224 binding sequence, changing RNA secondary structure and dysregulating TCF21 gene expression. The C risk allele combined with disruption of miRNA-224 and downregulated TCF21 expression increases CAD risk in response to disease-relevant stimuli. 27

Abdou et al. 28 also studied this SNP in an Egyptian population selected from the Cardiology Department of Menoufia University. They concluded that the CC genotype and C allele of TCF21 (rs12190287 G>C) were genetic risk factors for CAD development.

Although for a long time the focus was on protein-coding variants, recent research has expanded the spectrum of CAD-associated variants and highlighted the importance of regulatory variants in miRNA sequences and binding sites. Following the contributions of the ENCODE and FANTOM consortia to knowledge of the human genome, significant
attention has been paid to so-called genomic noise – miRNAs and long noncoding RNAs (lncRNA) involved in modulating gene expression. Several reports showed that polymorphisms in lncRNA or their target genes play a role in the development of adverse cardiovascular phenotypes. TCF21 protein and miR-224 were expressed in the human diseased vessel wall in vivo with an inversely regulated expression in atherosclerotic lesions.  

Using laser microdissection in endarterectomy samples from vessels with stable or ruptured plaque, Nurnberg et al. employed modern genomic methods to demonstrate that the expression of TCF21 was significantly lower in the ruptured plaque. By decreasing TCF21 expression, the mutated C risk allele may interfere in fibrous cap formation, increasing coronary risk. Nevertheless, further studies are needed to assess how changes in TCF21 expression affect the fibrous cap and consequently plaque vulnerability and rupture.

**Study limitations**

The patients’ medications were not standardized before randomization. Patients were on various types and doses of statins, antihypertensive and antiplatelet drugs and beta-blockers, which could have affected the clinical and laboratory variables. Second, in our genetic analysis, the sample size needed to be larger to increase the frequency of rare alleles and improve test power. Finally, in our genetic laboratory it was not possible to assess miRNA function, which is known to be involved in post-transcriptional regulation of gene expression and may affect both the stability and translation of mRNAs.

**Conclusions**

Multivessel disease, diabetes, renal impairment, sedentary lifestyle, and an unfavorable genetic profile were the main significant risk factors for atherosclerosis progression and MACE occurrence in a Portuguese coronary population from the Madeira archipelago. In this population, the C allele in the dominant model (CC>CG) of the TCF21 rs12190287 G>C polymorphism was a significant genetic risk factor for MACE occurrence and a marker of severe prognosis and death. This gene variant may influence fundamental smooth muscle processes in response to vascular stress, accelerating atherosclerosis progression, and may represent a target for future therapies. These findings could impact future preventive and management protocols in this group of patients.

**Conflicts of interest**

The authors have no conflicts of interest to declare.

**References**


