Background—After age, sex is the most important risk factor for coronary artery disease (CAD). The mechanism through which women are protected from CAD is still largely unknown, but the observed sex difference suggests the involvement of the reproductive steroid hormone signaling system. Genetic association studies of the gene-encoding Estrogen Receptor α (ESR1) have shown conflicting results, although only a limited range of variation in the gene has been investigated.

Methods and Results—We exploited information made available by advanced new methods and resources in complex disease genetics to revisit the question of ESR1’s role in risk of CAD. We performed a meta-analysis of 14 genome-wide association studies (CARDioGRAM discovery analysis, N = 87,000) to search for population-wide and sex-specific associations between CAD risk and common genetic variants throughout the coding, noncoding, and flanking regions of ESR1. In addition to samples from the MiGen (N = 6,000), WTCCC (N = 7,400), and Framingham (N = 3,700) studies, we extended this search to a larger number of common and uncommon variants by imputation into a panel of haplotypes constructed using data from the 1000 Genomes Project. Despite the widespread expression of ERα in vascular tissues, we found no evidence for involvement of common or low-frequency genetic variation throughout the ESR1 gene in modifying risk of CAD, either in the general population or as a function of sex.

Conclusions—We suggest that future research on the genetic basis of sex-related differences in CAD risk should initially prioritize other genes in the reproductive steroid hormone biosynthesis system. (Circ Cardiovasc Genet. 2011;4:647-654.)

Key Words: coronary artery disease ■ estrogen receptor alpha ■ menopause ■ polymorphism ■ single nucleotide ■ genetic association studies ■ meta-analysis

After age, sex is the most important risk factor for coronary artery disease (CAD), with women aged 35 to 74 years having 2 to 3 times lower myocardial infarction (MI) incidence than age-matched men.1 The mechanism through which women are protected from MI/CAD is still largely unknown, but the observed sex difference and the fact that CAD risk in postmenopausal women approaches that of males suggests the involvement of the sex steroid hormone system. This hypothesis was initially supported by the results of observational studies that showed lower CAD risk among postmenopausal women undergoing hormone replacement therapy2,3; however, initial clinical trials of hormone replacement therapy have shown unexpected negative results,4,5 even unanticipated harm, al-

1

2

3

4

5
through the timing of initiation of therapy may explain these conflicting results.6–8

Clinical Perspective on p 654

The fact that CAD clusters in families9 (estimated heritability 38% to 57%10), coupled with the observation of sex- and menopause-related differences in risk, suggests that inter-individual variation in CAD risk may be partly mediated by population-level genetic variation in the genes that encode elements of the sex steroid hormone system. ERα is an important signaling gateway within this system and is expressed in multiple cardiovascular tissues in both males and females.11 The gene encoding ERα, ESR1, has been the subject of several candidate gene association studies in relation to CAD over the past decade, with generally inconsistent results12–14; however, only a very limited range of the genetic variation in ESR1 has been investigated, and the role of this gene in CAD risk remains to be clarified.

The last 5 to 7 years have seen a paradigm shift in our approach to investigating the genetic basis of complex diseases. Advanced new methods, including high-throughput genotyping,15 genome-wide association studies (GWAS),16 genotype imputation,17 second-generation sequencing,18 along with the availability of resources describing natural human genetic variation (eg, HapMap19, 1000 Genomes Project20), allow us to explore the effect of genetic variation on phenotype more thoroughly. Also important is the manner and volume in which raw genetic data are now generated and disseminated under a model of cross-study cooperation and public data deposition, which has been key to overcoming many of the problems that limited the success of candidate gene association studies for complex diseases.

While no genome-wide significant evidence for the involvement of ESR1 variation in CAD risk has been reported in recent GWAS, data from these studies may still provide important information either to support or refute this hypothesis. The fact that many robust new GWAS loci for complex diseases had previously been investigated as candidate genes (eg, LDLR in CAD21 and several recently confirmed loci for low-density lipoprotein, high-density lipoprotein, and triglycerides22) highlights the importance of revisiting the role of candidate genes in complex diseases.23

Therefore, in this article, we bring these powerful post-genomic methods and resources to bear on a classical CAD candidate gene to resolve a long-running unanswered question in cardiovascular genetics. For common variation in a genomic region centered on ESR1, we report the results of a large meta-analysis of GWAS of MI and CAD and explore possible sex-specific differences. We also investigate the effect on CAD risk of low-frequency variation in this region.

Materials and Methods

Coronary Artery Disease GWAS Meta-Analysis

The Coronary ARtery Disease Genome-wide Replication And Meta-analysis (CARDIoGRAM) Consortium was formed with the purpose of identifying novel susceptibility loci for CAD. Briefly, the CARDIoGRAM discovery analysis combined data from 14 published and unpublished primary GWAS in individuals of European ancestry, including 22 233 cases with CAD (stable or unstable coronary events) (30.9% of which were females) and 64 762 controls (58.1% of which were females).

Each primary GWAS performed a logistic regression analysis to test for association between genotyped and imputed (using the HapMap Phase II reference panel24) single-nucleotide polymorphisms (SNPs) and risk of CAD under an additive disease model adjusted for age and sex. (See the online-only Supplemental Methods for a more detailed summary of the genotyping and quality control methods used.)

In this study, we meta-analyzed these study-level results using inverse-variance weighting under a fixed-effects model. We performed a random-effects meta-analysis for SNPs with significant between-study heterogeneity (P-heterogeneity <0.1), on the basis of Cochran’s Q statistic. These analyses were carried out for each of 535 SNPs in a genomic region containing the entire coding and noncoding region of ESR1 (see the online-only Supplemental Table 1 and a 50kb region upstream and downstream of the gene (5'547kb; Chr6, 151927808 to 152474406, GRCh37.p1).

Sex-Stratified Analysis

An equivalent analysis to that described above was performed separately for females and males in 13 of the 14 contributing studies (data for the Cohorts for Heart and Aging Research in Genomic Epidemiology [CHARGE] Consortium not available), and the results were meta-analyzed in a similar way. We also formally tested for interaction between each SNP and sex by using the sex-specific effects and variances within each study to estimate those of the SNP×sex interaction term (online-only Supplemental Methods). We then meta-analyzed the results as described for the unstratified analysis.

Fine-Mapping Analysis

Public GWAS Data Sources

To perform fine-mapping studies in the region of interest, we used publicly available genotype and phenotype data from 3 large published GWAS: (1) The Myocardial Infarction Genetics Consortium (MiGen25) is a case-control GWAS consisting of 2967 cases of early-onset MI and 3075 age- and sex-matched controls from 6 international sites in the United States and Europe; (2) The Welcome Trust Case-Control Consortium (WTCCC23) is a case-control GWAS of CAD consisting of 1988 cases and 5380 controls from the United Kingdom; (3) The Framingham Share Initiative dataset includes genetic data and longitudinal phenotype data, such as incidence of major cardiovascular events, for ∼9000 individuals from the Framingham Heart Study (http://www.framinghamheartstudy.org), of which we have included 3717 in the present study (selected to maximize the number of subjects free from cardiovascular disease at baseline who had genetic data and complete follow-up data; 464 events [see below for phenotype definition]; mean follow-up, 13.5 years; online-only Supplementary Appendix 1; Luís-Ganella et al, unpublished data, 2011).

The phenotypic characteristics of these studies were as follows: MiGen cases were males aged <50 years or females aged <60 years who were diagnosed with MI on the basis of autopsy evidence, a combination of chest pain and electrocardiographic evidence, or elevation of cardiac biomarkers; WTCCC cases had a validated history of either MI or coronary revascularization (coronary artery bypass surgery or percutaneous coronary angioplasty) before their 66th birthday; in the Framingham sample, events included incident cases with MI, angina, coronary revascularization, and death because of CAD.

Of the 6042 individuals in the MiGen sample, 2681 were previously included in the CARDIoGRAM discovery meta-analysis. All of the WTCCC cases (N = ∼1988) and approximately half of the controls (N = ∼2938) were also included in the CARDIoGRAM meta-analysis, as were many of the individuals in the Framingham sample, as part of the CHARGE Consortium.26

Genome-wide genotype data and associated phenotype data for the MiGen and Framingham samples were obtained via the database of Genotypes And Phenotypes (dbgap.ncbi.nlm.nih.gov; project num-
ber 2392). Data for the WTCCC sample were obtained from the European Genotype Archive (www.ebi.ac.uk/ega) with permission from the WTCCC Data Access Committee (www.wtccc.org.uk).

1000 Genomes Imputation in MIGen, WTCCC, and Framingham

See the online-only Supplemental Figure 1 for a summary of the quality control steps, imputation process, association analyses, and meta-analyses performed for this analysis.

Imputation of untyped genetic variants in individuals from the MIGen, WTCCC, and Framingham samples was performed using IMPUTE2. Imputation was performed for SNPs in the region of interest using a reference panel of phased haplotypes (available from http://www.1000genomes.org/sequencedata/1000GenomesPhasingPanel). Based on the August 2010 data release from the 1000 Genomes Project (1kG; 566 haplotypes from populations of European ancestry, EUR: CEU, TSI, GRB, FIN, and IBS). As input for this process, we included only directly genotyped SNPs with high call rate (>95%) and whose genotype frequencies were in Hardy-Weinberg equilibrium (P<10^-6). We carried forward to the analysis stage only those SNPs imputed with high quality (IMPUTE2 INFO metric ≥0.5).

Association and Meta-Analysis of Genotyped and Imputed SNPs in MIGen, WTCCC, and Framingham

A logistic regression analysis of association between allele dosage of imputed and genotyped SNPs and MI/CAD was performed separately in the MIGen, WTCCC, and Framingham studies, with adjustment for sex. Adjustment for age or other clinical covariates was not possible because no further phenotype data were available in all studies; however, the association results in the Framingham and MIGen samples were very similar after additional adjustment for age at event (data not shown), and both the MIGen and WTCCC studies were age- and sex-matched by design. To account for interrelatedness, the analysis of the Framingham sample was also adjusted for the first 2 genetic principal components. The results from these 3 studies were meta-analyzed as described above for the CARDIoGRAM analysis.

Statistical Analysis

Apart from imputation, all analyses were performed using R version 2.11 (packages and functions indicated below by <package>:::<function>). Fixed- and random-effects meta-analyses were performed using metan::meta.DSL. Association testing was performed using stats::glm for the case-control studies and rmeta::meta.DSL for the unstratified meta-analysis (p<0.05, and 0.01, respectively (online-only Supplemental Figure 3)). We used the same criterion for declaring statistical significance as for the unstratified meta-analysis (p<10^-6). In females, we had high power (~80%) to detect ORs of ≥1.18, ≥1.47, and ≥1.58 and moderate power (~50%) to detect ORs of ≥1.15, ≥1.37, and ≥1.45 for SNPs with MAF ≥0.15, ≥0.05, and ≥0.01, respectively (online-only Supplemental Table 2). In males, we had high power (~80%) to detect ORs of ≥1.15, ≥1.23, and ≥1.39 and moderate power (~50%) to detect ORs of ≥1.12, ≥1.18, and ≥1.39 for SNPs with MAF ≥0.15, ≥0.05, and ≥0.01, respectively (online-only Supplemental Table 2).

One SNP, lying ~35 kb upstream of the most distal noncoding exon (Figure B), exceeded the threshold for regional significance in the test for interaction between sex and genotype as a predictor of CAD risk (rs9479087, MAF: 0.183 in CARDIoGRAM, P=0.0026; online-only Supplemental Figure 3); however, this variant was not significantly associated with risk in either males (P=0.0026; pooled OR [95% CI]: 1.07 [1.03, 1.13]) or females (P=0.057; pooled OR [95% CI]: 0.94 [0.89, 1.00]) at the regional significance level; no other regionally significant evidence for association was observed either among females (top result: rs6927702 in intron 3, P=0.0081, Figure B) or males (top result: rs9479087, P=0.0026; Figure B).

Fine-Mapping Analysis

While the density of SNP data in the HapMap II panel (CARDIoGRAM results) for this region is quite high (mean=1.15 SNPs/kb), it is possible that some stronger true
Figure. Regional association results for the ESR1 gene region. Results of CARDioGRAM global and sex-stratified meta-analysis and fine mapping analysis for a region of chromosome 6 containing the coding and noncoding exons of ESR1 and 50 kb of the upstream and downstream flanking regions (−log10(P-value) shown as black points). Results shown are for a fixed- or random-effects meta-analysis in the absence or presence of between-study heterogeneity, respectively (see Methods). The position of the ESR1 gene is shown at the bottom of the plot (dotted line), with coding and noncoding exons shown as long and short vertical bars, respectively. Regional recombination rate (HapMap II) is shown as a gray line in plot A. A, Regional association plot of global meta-analysis results from the CARDioGRAM study. B, Regional association plot of sex-stratified meta-analysis results from the CARDioGRAM study. Results for the test for single-nucleotide polymorphisms (SNP)-sex interaction are shown in the top panel; results for the association test in females and males are shown in the middle and bottom panels, respectively. C, Regional association plot of results from fine-mapping meta-analysis (MiGen, WTCCC, and Framingham studies). Results for SNPs that were previously analyzed in the CARDioGRAM study are shown as black points (ie, SNPs that were directly genotyped or imputed using a reference panel of haplotypes generated from the Phase II HapMap CEU genotypes). Results for additional SNPs that were imputed using a reference panel of haplotypes generated using data from the 1000 Genomes project (August 2010 release) are shown as gray points.
association signals are not captured by these common geno-
typed and imputed variants. Such signals might be detected
by analyzing a higher density map of common and low-
frequency SNPs in this region. To explore this possibility, we
imputed ~2500 additional variants from the 1 kG reference
panel (~4.52 SNPs/kb), 1451 of which were imputed with
high quality in all 3 samples (~2.7 SNPs/kb; see the
online-only Supplemental Figure 1). Imputation in the 1 kG
panel allowed us to test ~800 additional SNPs within the
region of interest that were not included in the CARDio-
GRAM meta-analysis. Newly imputed SNPs had a wide
range of MAF, although a large proportion had MAF in the
range 0.0 to 0.05 (online-only Supplemental Figure 4).

After testing for association between SNPs in the 1 kG
panel and CAD in the MIGen, WTCCC, and Framingham
samples, meta-analyzing the results and correcting for mul-
tiple testing (neff =3.8×10^{-5}), we observed no
global significant evidence for association in this region
(Figure C). This analysis had high power (~80%) to detect
ORs of ~1.21, ~1.44, ~2.09, and ~3.14 and moderate
power (~50%) to detect ORs of ~1.18, ~1.35, ~1.85, and
~2.59 for SNPs with MAF ~0.15, ~0.05, ~0.01, and
<0.01, respectively (online-only Supplemental Table 2). The
strongest association was observed for variant 6 to
rs9340799, previously known as the
PvuII
variants,
which is inherited on the Y chromosome in males. Since, as far as we are aware, no evidence of association between CAD and SRY has been
reported, it is not appropriate to consider sex as being
causally associated with CAD risk. Rather, sex is a trait that
is strongly associated with CAD risk via unknown and
potentially modifiable factors (eg, physiological, environ-
mental, behavioral factors, etc), whose effects we can partly
capture by using sex as a proxy variable. It is important to
identify and understand these factors because the ability to
modify even a fraction of sex-associated CAD risk might
have a marked impact on prevention, possibly more so than
by modifying other CVRFs.

All of the loci identified by GWAS to date as being
associated with CAD risk are located on autosomes, and it
seems likely that most or all of the loci that explain the
remaining heritability of CAD risk will also be autosomal.
Consequently, these loci are in linkage equilibrium with
SRY and have equal genotype frequencies in males and females.
This leads us to the simple but important conclusion that
differences in CAD risk between sex cannot be directly
caused by genetic factors but can only arise because of an
interaction between sex and other processes associated with
risk. Consequently, the present study, like all association
studies of primary autosomal genetic variation, does not
attempt to explain differences in risk between sex. Instead,
we search for population-level differences in CAD risk that
are driven by ESR1 variation and whose effects may or may
not be different among females compared with males (ie, that
interact with sex).

Over the past decade, candidate-gene association stud-
ies have reported generally inconsistent results regarding
the role of ESR1 genetic variation in CAD risk. An initial
meta-analysis including ~7000 individuals supported associ-
ation, but this result was not upheld by 2 subsequent
meta-analyses representing ~16 000 individuals and ~32 000
individuals; however, these studies have been restricted to a very
limited number of SNPs (especially rs2234693 and
rs9340799, previously known as the
PvuII
and
XbaI
variants, which lie in Intron 1) out of the thousands now known to
lie within the gene region. We estimate that the 4 most widely
studied variants collectively capture (with \(\rho^2\geq0.8\)) only \(~2%\)
of the 1450 SNPs tested in our study (data not shown).
Therefore, although recent reports have found no evidence of
association between ESR1 variation and CAD risk, this
question remains unanswered until a more complete survey of
the gene is carried out. The potential gain to be made from
this is illustrated by recent advances in understanding ESR1’s
role in modulating bone mineral density and fracture risk,
phenotypes that show intriguingly similar patterns of sex-
specific and menopause-related risk to those observed for
cardiovascular risk. While candidate gene studies of the role
of ESR1 variation in bone mineral density and fracture risk
also examined a limited range of genetic variation and
obtained similarly inconsistent results, a large meta-
analysis of several GWAS subsequently confirmed the involve-
ment of ESR1 variation in modulating these phenotypes,
with highly significant evidence for association in the up-
stream noncoding regulatory region of the gene, in stark

Discussion

In this study, we exploited post-genomic tools and resources
to expand on previous candidate association studies of ESR1
in 2 main ways: (1) we analyzed a large number of common
and uncommon genetic variants in the coding, noncoding,
and flanking regions of the gene, capturing a large proportion
of the genetic variation throughout the gene and its regulatory
regions; (2) we performed these analyses in large samples of
up to ~85 000 individuals representing multiple populations
of European descent, which increases our power to detect
subtle risk effects.

Despite this study’s power to detect case-control differ-
ences in CAD risk of as low as 10% for a broad range of
genetic variation throughout this region, we found no evi-
dence of involvement of ESR1 in modifying CAD risk either
at the population level or as a function of sex. We consider
these results surprising, given ERα’s central role in estrogen
and androgen signaling, its widespread expression in vascular
tissues, and the importance of sex for CAD risk.

After age, male sex remains the most important indepen-
dent cardiovascular risk factor (CVRF), and has a far greater
impact on total risk than other important risk factors, such as
smoking, lipid profile, and diabetes. The physiological basis
of this sex difference remains unclear, and limited research
into this question has been carried out, compared with that for
other risk factors, mainly because sex is nonmodifiable; how-
over, rather than considering male sex as a nonmodifiable
cause of increased CAD risk, it is important to remember that
sex is a simple mendelian trait determined by the presence or
absence of a single gene, SRY, which is inherited on the Y
chromosome in males. Since, as far as we are aware, no evidence of association between CAD and SRY has been
reported, it is not appropriate to consider sex as being
causally associated with CAD risk. Rather, sex is a trait that
is strongly associated with CAD risk via unknown and
potentially modifiable factors (eg, physiological, environ-
mental, behavioral factors, etc), whose effects we can partly
capture by using sex as a proxy variable. It is important to
identify and understand these factors because the ability to
modify even a fraction of sex-associated CAD risk might
have a marked impact on prevention, possibly more so than
by modifying other CVRFs.
contrast with the lack of association we have observed for CAD.

In the discovery stage of the CARDIoGRAM study, the direction of effect of the lead SNP was largely consistent across the contributing studies (online-only Supplemental Figure 1) but fell well short of the threshold for regional statistical significance. The region of high LD containing this SNP was located within the 5′ regulatory region but did not coincide with the previously reported signal for bone mineral density and fracture risk.34

We found no broadly convincing evidence of association between ESR1 variation and CAD risk as a function of sex. Although the P-value of the sex interaction test for 1 SNP exceeded the significance threshold set, with opposing effects observed among males and females, this variant was not significantly associated with CAD risk in either sex considered separately (online-only Supplemental Figure 3). Considering the additional fact that this variant lies at a considerable distance from the regulatory (≈35 kb) and coding (≈186 kb) regions of the gene, we feel that these results do not provide strong evidence of a robust sex-specific association at this locus. In addition to sex, another potential modifier of the putative association between ESR1 variation and CAD risk is menopausal status among women. Although we were unable to investigate this issue directly, we provide some initial data on this question based on age data from the MIGen study, and we find no evidence of significantly different effects of ESR1 variation on cardiovascular risk as a function of menopausal status (see the online-only Supplemental Note).

In the fine-mapping analysis, imputation using data from the 1000 Genomes Project allowed us to analyze a much denser map of common variants in the region (online-only Supplemental Figure 4) and especially to explore the role of variants with frequencies below 0.05, which are under-represented in haplotype panels based on data from the HapMap project but which are a potentially important source of risk variance in complex diseases35,36; however, we found no additional evidence of association with CAD risk for any of these additional variants.

We highlight the fact that this study is well-powered to detect genetic risk effects with sizes and frequencies that are generally plausible for common complex diseases. For example, in the CARDIoGRAM discovery analysis, we have high power (≈80%) to detect common variants with MAF ≥0.15 that carry risk effects as low as OR ≈1.1 and low-frequency variants (0.01 ≤ MAF ≤ 0.05) that carry risk effects of OR ≈1.3. In addition, the fine mapping analysis was also powered to detect associations for rare imputed variants with MAF ≤0.01 and effect sizes of approximately OR≈3. Weaker and/or rarer risk effects than these are likely to have limited clinical relevance at the population level. In these power computations, we used stringent statistical significance thresholds that account for multiple testing (see the online-only Supplemental Methods).

The most likely explanation for lack of observable association in this analysis is that no true association exists in this gene, although we note the following limitations in this study’s ability to draw this conclusion: First, this study does not address this question in populations with non-European ancestry. Second, some truly associated variants in this gene may not have been detected by this study, although these are unlikely to be simple primary sequence variants with low allelic diversity, such as common or uncommon SNPs, low-copy number polymorphisms, or insertions/deletions. This analysis was also unable to detect very weak or very rare effects (online-only Supplemental Table 2). Third, this study cannot address the role of other potentially relevant forms of variation related to ERα, such as epistasis or epigenetic (eg, promoter methylation), post-transcriptional, or post-translational variation; however, if such variation exists, it is likely to be largely independent of primary sequence variation. Fourth, this study suggests that menopausal status does not modify the effect of ESR1 variation on female CAD risk but cannot discount this possibility because of the size and imprecise design of that analysis. Fifth, these analyses were not adjusted for classical CVRFs, although a true SNP-CAD association would only be masked by confounding if the SNP had opposing effects on CAD risk and CVRF profile, which seems unlikely. Sixth, most of the studies included in these meta-analyses had a case-control design, which could lead to a bias against the discovery of variants that reduce survival.

Finally, it is important to note that we have analyzed the genetic variation in only 1 of the genes that encode components of the steroid sex hormone system. A more thorough exploration of this system may help to clarify the role of this system in the pathophysiology of coronary risk.

Conclusions

In conclusion, on the basis of data from a large number of subjects representing multiple samples from several populations, we find no evidence for involvement of common or uncommon genetic variation in the coding, noncoding, or flanking regions of the ESR1 gene in modifying risk of CAD, irrespective of sex; however, data from observational studies and subanalysis of clinical trials continue to support the involvement of the steroid hormone system in modulating CAD risk. Therefore, we consider that the next step in exploring the role of the sex hormone biosynthesis system in modulating CAD risk should initially be to prioritize the investigation of other genes within this system.

Acknowledgments

See the online-only Supplemental Appendix for full list of contributors from the CARDIoGRAM Consortium. We thank all contributing members of the CARDIoGRAM Consortium for the use of cohort-level summary association results. We thank the authors of the MiGen, WTCCC, and Framingham GWAS and 1000 Genomes project for making their data publicly available and the authors of IMPUTE2 for making 1-kb-based phased haplotypes available for public use. A full list of the investigators who contributed to the generation of the WTCCC data are available from http://www.wtccc.org.uk. This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or the National Heart, Lung, and Blood Institute. We also thank Ana Paula Dantas and Jana Selent for interesting discussions at the design stage.

Sources of Funding

Information on funding sources for the studies that contributed to The CARDIoGRAM Consortium were provided previously.21 The
Myocardial Infarction Genetics (MiGen) Consortium\(^1\) was funded by grant R01 HL087676 from the National Institutes of Health. The Wellcome Trust Case-Control Consortium 2 was funded by the Wellcome Trust under award 085475. The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood institute in collaboration with Boston University (Contract No. N01-HC-25195). This work was supported by a grant from ACCIÓ (RD08-10024), the European Regional Development Fund, the Spanish Ministry of Science and Innovation through the Carlos III Health Institute [CIBER Epidemiología y Salud Pública, Red HERACLES RD06/009, P1061254, P109/090506], and by the Catalan Research and Technology Innovation Interdepartmental Commission [SGR 1195]. Dr Lucas was funded by the Juan de la Cierva Program, Ministerio de Educación (JCI-2009-04068).

References


CLINICAL PERSPECTIVE

After age, male sex is the most important risk factor for cardiovascular disease; however, little research has been carried out to understand the underlying causes of sex-related risk, in comparison with work on modifiable risk factors. Although the physiological differences between the sexes are evident, the genetic differences are minimal because sex is determined by the presence (in males) of a single gene, SRY, on chromosome Y. This leads to the important conclusion that male sex itself is not a cardiovascular risk factor but a proxy variable that captures a large fraction of risk via other unknown, possibly modifiable, metabolic factors that differ between males and females. Sex hormone metabolism is a prime candidate system for explaining sex differences in risk. In this highly powered study, the authors perform an extensive survey of a broad range of genetic variation in the ESR1 gene, which encodes the principal candidate for explaining sex-related differences in coronary artery disease (CAD) risk, Estrogen Receptor α, ERα. Despite ERα’s central role in sex hormone signaling, its widespread expression in vascular tissues, and the importance of sex for CAD risk, the authors find no evidence for the involvement of genetic variation in ESR1 in modifying CAD risk, either in the general population or separately in males and females. Against the context of a history of inconsistent results regarding this question, this study provides a reasonably conclusive answer and may stimulate a renewed effort to explore other elements of the sex hormone system to explain sex differences in CAD risk.
Post-Genomic Update on a Classical Candidate Gene for Coronary Artery Disease: \textit{ESRI}

Gavin Lucas, Carla Llufs-Ganella, Isaac Subirana, Mariano Sentí, Christina Willenborg, Muntaser D. Musameh, Stephen M. Schwartz, Christopher J. O'Donnell, Olle Melander, Veikko Salomaa and Roberto Elosua

\textit{Circ Cardiovasc Genet}. 2011;4:647-654; originally published online October 9, 2011; doi: 10.1161/CIRCGENETICS.111.960583

\textit{Circulation: Cardiovascular Genetics} is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/4/6/647

Data Supplement (unedited) at:
http://circgenetics.ahajournals.org/content/suppl/2011/10/09/CIRCGENETICS.111.960583.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in \textit{Circulation: Cardiovascular Genetics} can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to \textit{Circulation: Cardiovascular Genetics} is online at:
http://circgenetics.ahajournals.org//subscriptions/
Post-genomic update on a classical candidate gene for coronary artery disease: \textit{ESR1}

Gavin Lucas MSc, PhD*1; Carla Lluís-Ganella, MSc*1; Isaac Subirana, MSc2,1; Mariano Sentí, MD, PhD1,3; Christina Willenborg4,5; Muntaser Musameh MD, PhD6; CARDIoGRAM Consortium†; Stephen M Schwartz MD, PhD7,8; Christopher J O’Donnell MD MPH9,10; Olle Melander MD, PhD11; Veikko Salomaa MD, PhD12; Roberto Elosua, MD, PhD1,2.

* These authors contributed equally to this work

1 Cardiovascular Epidemiology and Genetics, IMIM. Barcelona, Spain.
2 Epidemiology and Public Health Network (CIBERESP), Barcelona, Spain.
3 Pompeu Fabra University. Barcelona, Spain.
4 Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany.
5 Medizinische Klinik II, Universität zu Lübeck, Lübeck, Germany.
6 Department of Cardiovascular Sciences, University of Leicester, United Kingdom.
7 Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, Washington, USA.
8 Department of Epidemiology, University of Washington, Seattle, Washington, USA.
9 National, Heart, Lung, and Blood Institute and Framingham Heart Study, Framingham, Massachusetts, USA.
10 Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA.
11 Department of Clinical Sciences, Hypertension and Cardiovascular Diseases, University Hospital Malmö, Lund University, Malmö, Sweden.
12 National Institute for Health and Welfare, Helsinki, Finland.

† See Supplementary Appendix for a full list of contributors
### Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Methods</td>
<td>S2</td>
</tr>
<tr>
<td>a. CARDIoGRAM discovery analysis methods summary</td>
<td></td>
</tr>
<tr>
<td>b. Test for interaction between SNP and gender</td>
<td></td>
</tr>
<tr>
<td>c. Power calculations</td>
<td></td>
</tr>
<tr>
<td>Supplementary Note</td>
<td>S4</td>
</tr>
<tr>
<td>Preliminary age-stratified analysis to explore possible menopause-related ESR1 effects among women</td>
<td></td>
</tr>
<tr>
<td>Supplementary Table 1</td>
<td>S5</td>
</tr>
<tr>
<td>Chromosomal positions of coding and non-coding regions of ESR1</td>
<td></td>
</tr>
<tr>
<td>Supplementary Table 2</td>
<td>S6</td>
</tr>
<tr>
<td>Power computation</td>
<td></td>
</tr>
<tr>
<td>Supplementary Figure 1</td>
<td>S9</td>
</tr>
<tr>
<td>Summary of quality control and analysis procedures in the fine mapping.</td>
<td></td>
</tr>
<tr>
<td>Supplementary Figure 2</td>
<td>S10</td>
</tr>
<tr>
<td>CARDIoGRAM meta-analysis results for the top SNP in the region of interest.</td>
<td></td>
</tr>
<tr>
<td>Supplementary Figure 3</td>
<td>S11</td>
</tr>
<tr>
<td>CARDIoGRAM meta-analysis results for the SNP with the greatest difference in association between males and females (strongest interaction with gender).</td>
<td></td>
</tr>
<tr>
<td>Supplementary Figure 4</td>
<td>S12</td>
</tr>
<tr>
<td>Distribution of minor allele frequencies (MAF) for SNPs analyzed in this study.</td>
<td></td>
</tr>
<tr>
<td>Supplementary References</td>
<td>S13</td>
</tr>
<tr>
<td>Supplementary Appendix 1</td>
<td>S14</td>
</tr>
<tr>
<td>Process of selection of participants from the Framingham study</td>
<td></td>
</tr>
<tr>
<td>Supplementary Appendix 2</td>
<td>S15</td>
</tr>
<tr>
<td>CARDIoGRAM Investigators</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Methods

a. **CARDIoGRAM discovery analysis methods summary**

Genotyping in individual discovery GWA studies was carried out on Affymetrix or Illumina platforms. Approximately 2.3 million imputed genotypes were generated using the MACH, IMPUTE, or BIMBAM imputation algorithms and the HapMap Phase II reference panel\(^1\)). Each primary discovery GWAS performed a logistic regression analysis to test for association between genotyped and imputed SNPs and risk of MI/CAD under an additive disease model adjusted for age and sex and taking into account the uncertainty of imputed genotypes. In every study, the variance inflation factor \(\lambda\) was estimated from genotyped SNPs and also used for adjustment.

Quality control of these data was performed centrally according to established criteria including a check of consistency of the given alleles across all studies, quality of the imputation, deviation from Hardy-Weinberg equilibrium in the controls, minor allele frequency, and call rate.

In the present study, a meta-analysis was performed separately for every SNP from every CARDIoGRAM study that passed the quality criteria. The default meta-analysis was a fixed effect model with inverse variance weighting and calculation of two homogeneity statistics: Cochran’s Q- and \(I^2\) statistic. When there was no indication of heterogeneity for a SNP (P for Q > 0.01), the fixed effect model was maintained. When heterogeneity was present (P for Q < 0.01), a random effects model (DerSimonian-Laird) was used for that SNP.

b. **Test for interaction between SNP and gender**

To formally test for interaction between each SNP and gender in the CARDIoGRAM and fine mapping analyses (data not shown for the latter), we performed the following steps:

i. Within each CARDIoGRAM study, we computed the beta for the SNP-gender interaction term as the absolute difference between the betas for females and males.

ii. Within each CARDIoGRAM study, we computed the standard error of the SNP-gender interaction term as square root of the sum of the squares of the standard errors of the \(\beta\) from the female and male analysis.

iii. We then used these betas and standard errors to perform fixed or random effects meta-analyses according to the same protocol as that used for the un-stratified analysis.

c. **Power Calculations**

We performed a post-hoc calculation of our analyses’ power to detect significant associations. We allowed that power is determined by sample size, the proportion of cases and controls for the case-control studies or the number of events for the cohort study, the effect of a variant on risk
(e.g. OR), and the frequency of the minor allele (MAF) of the associated variant, the p-value threshold required to declare statistical significance, LD between correlated and causal variants, genotyping error, the quality of imputation for imputed variants, between-study heterogeneity in the meta-analysis, and possibly other factors. Of these, MAF is the most important non-modifiable determinant of power, and so we estimated power for a series of representative sub-ranges of MAF. Rather than attempting to parameterize all of the other factors, we captured their effects by using the standard error (SE) from the meta-analysis of all three studies, which is inversely correlated with power. In these power calculations, the variant’s effect on disease risk was taken as the beta from the meta-analysis of all studies, and thus represented the odds ratios for the case-control studies and hazard ratio for cohort studies, where applicable; ORs and HRs are considered to be comparable because the prevalence of the phenotype in the cohort studies is relatively low. All power computations were based on an alpha value (Type I error rate) equivalent to the threshold required to declare a statistically significant association after adjustment for multiple testing (see main text). Within each analysis we performed the following steps:

i. For each SNP in the analysis, MAF was taken to be the mean MAF across all studies.

ii. SNPs were binned according to the following sub-ranges of MAF: (0,0.01], (0.01,0.02], (0.02,0.03], (0.03,0.04], (0.04,0.05], (0.05,0.06], (0.06,0.07], (0.07,0.08], (0.08,0.09], (0.09,0.1], (0.1,0.15], (0.15,0.2], (0.2,0.3], (0.3,0.4] and (0.4,0.5].

iii. For each sub-range of MAF the mean of the SE of all SNPs within that sub-range was computed, and used to compute and express the power of the analysis in the following two ways.

iv. The minimum effect size (beta) the analysis had high (~80%) or moderate (~50%) power to detect. The definitions of high and moderate power were selected arbitrarily to indicate where our analysis was well powered to detect risk effects (high power), but also to allow for the fact that, if multiple independent but more subtle effects were present, at least some proportion of these could also be detected (e.g. 50%, moderate power).

v. The power of the analysis to detect each of a series of effect sizes (betas, corresponding to the following odds ratios: 1.05, 1.1, 1.2, 1.3, 1.5, 1.7, 2, 2.5 and 3). These data were computed to help indicate the circumstances under which our study was unable to provide conclusive information, e.g. for rarer variants or for more subtle effect sizes.

The results of these power calculations are shown in Supplementary Table 2
Supplementary Note

Preliminary age-stratified analysis to explore possible menopause-related ESR1 effects among women

After age and gender, menopausal status among women appears to be one of the strongest determinants of CAD risk. We explored the possibility that the effect of genetic variation in ESR1 on CAD risk may vary among women according to menopausal status. Although this variable was not available for any of the CARDIoGRAM studies or for the three studies included in the fine mapping analysis, we attempted to capture most of its variation using a proxy variable based on age (<50 years or ≥50 years), and then tested for interaction between this proxy variable and genotype. This analysis was performed only for MIGen owing to the lack of age data for the WTCCC sample, and the low number of events in the Framingham study.

We observed no regionally significant interaction between this proxy variable and genotype for any variant in the region of interest (minimum p-value=0.0012 for rs11968025), although we note the limited sample size of this analysis (number of females aged <50 yrs and ≥50 yrs was 832 (of which 389 were cases) and 582 (of which 274 were cases), respectively).
## Supplementary Tables

### Supplementary Table 1. Chromosomal locations of coding and non-coding exons in \textit{ESR1}.

<table>
<thead>
<tr>
<th>Exon Name</th>
<th>Coding</th>
<th>Start</th>
<th>Stop</th>
<th>Length (bp)</th>
<th>Position with respect to translation start site</th>
<th>AA length</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>-</td>
<td>151977808</td>
<td>151977899</td>
<td>91</td>
<td>-151240</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>-</td>
<td>152011675</td>
<td>152011800</td>
<td>125</td>
<td>-117373</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>-</td>
<td>152023011</td>
<td>152023141</td>
<td>130</td>
<td>-106037</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>-</td>
<td>152112508</td>
<td>152112595</td>
<td>87</td>
<td>-16540</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>-</td>
<td>152112697</td>
<td>152112848</td>
<td>151</td>
<td>-16351</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>-</td>
<td>152125065</td>
<td>152125160</td>
<td>95</td>
<td>-3983</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-</td>
<td>152125748</td>
<td>152126956</td>
<td>1208</td>
<td>-3300</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>-</td>
<td>152128494</td>
<td>152128645</td>
<td>151</td>
<td>-554</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-</td>
<td>152128816</td>
<td>152128978</td>
<td>162</td>
<td>-232</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>+</td>
<td>152128979</td>
<td>152129499</td>
<td>521</td>
<td>-70</td>
<td>151</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>152163732</td>
<td>152163922</td>
<td>190</td>
<td>34684</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>152201790</td>
<td>152201906</td>
<td>116</td>
<td>72742</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>152265308</td>
<td>152265643</td>
<td>335</td>
<td>136260</td>
<td>112</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>152332791</td>
<td>152332929</td>
<td>138</td>
<td>203743</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>152382126</td>
<td>152382259</td>
<td>133</td>
<td>253078</td>
<td>45</td>
</tr>
<tr>
<td>7</td>
<td>+</td>
<td>152415520</td>
<td>152415703</td>
<td>183</td>
<td>286472</td>
<td>61</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>152419867</td>
<td>152420102</td>
<td>235</td>
<td>290819</td>
<td>77</td>
</tr>
<tr>
<td>3' UTR</td>
<td>-</td>
<td>152420103</td>
<td>152424406</td>
<td>4303</td>
<td>291055</td>
<td></td>
</tr>
</tbody>
</table>

AA: Amino Acid; bp: base pairs. 152177055

* Name assigned by Koš et al.\textsuperscript{3} to non-coding exons, or sequentially for coding exons

\textsuperscript{3} Position in GRCh37.p1 determined using information provided by Koš et al. for non-coding exons and the Exon 1 start point, and from the Ensembl exon report for the coding exons (ENSG00000091831; www.ensembl.org) otherwise.

\textsuperscript{5} Translation start codon begins at 152129048, 70bp downstream of the beginning of Exon 1

\textsuperscript{11} Common splice acceptor site reported by Koš et al.
### Supplementary Table 2. Power computation – see Supplementary Methods for details

<table>
<thead>
<tr>
<th>Minor Allele Frequency range</th>
<th>Number of SNPs</th>
<th>high power (0.8)</th>
<th>detectable (0.5)</th>
<th>Power (to detect OR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>4</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CARDioGRAM (n=86,995)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.01, 0.02)</td>
</tr>
<tr>
<td>(0.03, 0.04)</td>
</tr>
<tr>
<td>(0.05, 0.06)</td>
</tr>
<tr>
<td>(0.07, 0.08)</td>
</tr>
<tr>
<td>(0.09, 0.1)</td>
</tr>
<tr>
<td>(0.15, 0.2)</td>
</tr>
<tr>
<td>(0.3, 0.4)</td>
</tr>
<tr>
<td>(0.4, 0.5)</td>
</tr>
</tbody>
</table>

| Number of SNPs | 0   | 4    | 1    | 8    | 1    | 10   | 8    | 16   | 10   | 19   | 65   | 71   | 102  | 89   | 96   |

<table>
<thead>
<tr>
<th>CARDioGRAM Females (n=30,915)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.01, 0.02)</td>
</tr>
<tr>
<td>(0.03, 0.04)</td>
</tr>
<tr>
<td>(0.05, 0.06)</td>
</tr>
<tr>
<td>(0.07, 0.08)</td>
</tr>
<tr>
<td>(0.09, 0.1)</td>
</tr>
<tr>
<td>(0.15, 0.2)</td>
</tr>
<tr>
<td>(0.3, 0.4)</td>
</tr>
<tr>
<td>(0.4, 0.5)</td>
</tr>
<tr>
<td>(0.6, 0.7)</td>
</tr>
<tr>
<td>(0.8, 0.9)</td>
</tr>
<tr>
<td>(1, 1.1)</td>
</tr>
</tbody>
</table>
### CARDioGRAM Males

<table>
<thead>
<tr>
<th>Number of SNPs</th>
<th>0</th>
<th>1</th>
<th>7</th>
<th>6</th>
<th>3</th>
<th>2</th>
<th>3</th>
<th>14</th>
<th>14</th>
<th>23</th>
<th>68</th>
<th>75</th>
<th>100</th>
<th>94</th>
<th>97</th>
</tr>
</thead>
<tbody>
<tr>
<td>high power (0.8)</td>
<td>--</td>
<td>1.488</td>
<td>1.382</td>
<td>1.304</td>
<td>1.261</td>
<td>1.226</td>
<td>1.336</td>
<td>1.352</td>
<td>1.344</td>
<td>1.202</td>
<td>1.211</td>
<td>1.152</td>
<td>1.123</td>
<td>1.101</td>
<td>1.096</td>
</tr>
<tr>
<td>moderate power (0.5)</td>
<td>--</td>
<td>1.387</td>
<td>1.305</td>
<td>1.244</td>
<td>1.21</td>
<td>1.182</td>
<td>1.269</td>
<td>1.281</td>
<td>1.275</td>
<td>1.164</td>
<td>1.171</td>
<td>1.123</td>
<td>1.1</td>
<td>1.082</td>
<td>1.078</td>
</tr>
<tr>
<td>Power (to detect OR)</td>
<td>1.05</td>
<td>--</td>
<td>0.00048</td>
<td>0.0008</td>
<td>0.0014</td>
<td>0.0021</td>
<td>0.0029</td>
<td>0.0015</td>
<td>0.0016</td>
<td>0.0021</td>
<td>0.0068</td>
<td>0.0099</td>
<td>0.017</td>
<td>0.045</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>--</td>
<td>0.003</td>
<td>0.0072</td>
<td>0.016</td>
<td>0.029</td>
<td>0.047</td>
<td>0.021</td>
<td>0.024</td>
<td>0.034</td>
<td>0.13</td>
<td>0.19</td>
<td>0.32</td>
<td>0.62</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>1.2</td>
<td>--</td>
<td>0.043</td>
<td>0.13</td>
<td>0.28</td>
<td>0.45</td>
<td>0.63</td>
<td>0.29</td>
<td>0.23</td>
<td>0.29</td>
<td>0.76</td>
<td>0.75</td>
<td>0.95</td>
<td>0.96</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td>--</td>
<td>0.22</td>
<td>0.49</td>
<td>0.77</td>
<td>0.91</td>
<td>0.99</td>
<td>0.62</td>
<td>0.56</td>
<td>0.58</td>
<td>0.94</td>
<td>0.86</td>
<td>0.97</td>
<td>0.98</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>--</td>
<td>0.83</td>
<td>0.97</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.98</td>
<td>0.97</td>
<td>0.96</td>
<td>1</td>
<td>0.98</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.7</td>
<td>--</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

### CARDioGRAM Gender*Genotype Interaction

<table>
<thead>
<tr>
<th>Number of SNPs</th>
<th>0</th>
<th>0</th>
<th>4</th>
<th>5</th>
<th>3</th>
<th>10</th>
<th>12</th>
<th>11</th>
<th>13</th>
<th>21</th>
<th>60</th>
<th>75</th>
<th>103</th>
<th>85</th>
<th>97</th>
</tr>
</thead>
<tbody>
<tr>
<td>high power (0.8)</td>
<td>--</td>
<td>--</td>
<td>1.646</td>
<td>1.535</td>
<td>1.525</td>
<td>1.753</td>
<td>1.523</td>
<td>1.345</td>
<td>1.382</td>
<td>1.3</td>
<td>1.265</td>
<td>1.237</td>
<td>1.185</td>
<td>1.17</td>
<td>1.161</td>
</tr>
<tr>
<td>moderate power (0.5)</td>
<td>--</td>
<td>--</td>
<td>1.506</td>
<td>1.422</td>
<td>1.415</td>
<td>1.586</td>
<td>1.413</td>
<td>1.276</td>
<td>1.305</td>
<td>1.24</td>
<td>1.213</td>
<td>1.191</td>
<td>1.15</td>
<td>1.138</td>
<td>1.131</td>
</tr>
<tr>
<td>Power (to detect OR)</td>
<td>1.05</td>
<td>--</td>
<td>0.00032</td>
<td>0.00042</td>
<td>0.00043</td>
<td>0.00033</td>
<td>0.00053</td>
<td>0.001</td>
<td>0.00085</td>
<td>0.0014</td>
<td>0.002</td>
<td>0.026</td>
<td>0.0059</td>
<td>0.0081</td>
<td>0.0099</td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>--</td>
<td>0.0015</td>
<td>0.0024</td>
<td>0.0025</td>
<td>0.0018</td>
<td>0.0039</td>
<td>0.011</td>
<td>0.008</td>
<td>0.017</td>
<td>0.028</td>
<td>0.04</td>
<td>0.11</td>
<td>0.16</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>1.2</td>
<td>--</td>
<td>0.017</td>
<td>0.033</td>
<td>0.034</td>
<td>0.025</td>
<td>0.063</td>
<td>0.2</td>
<td>0.14</td>
<td>0.3</td>
<td>0.44</td>
<td>0.57</td>
<td>0.88</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td>--</td>
<td>0.089</td>
<td>0.17</td>
<td>0.18</td>
<td>0.11</td>
<td>0.26</td>
<td>0.65</td>
<td>0.51</td>
<td>0.79</td>
<td>0.9</td>
<td>0.97</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>--</td>
<td>0.49</td>
<td>0.72</td>
<td>0.74</td>
<td>0.37</td>
<td>0.7</td>
<td>0.97</td>
<td>0.95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.7</td>
<td>--</td>
<td>0.86</td>
<td>0.97</td>
<td>0.98</td>
<td>0.66</td>
<td>0.94</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>--</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>0.93</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
### Post-genomic update on a classical candidate gene for coronary artery disease: ESR1

<table>
<thead>
<tr>
<th>Number of SNPs</th>
<th>22</th>
<th>139</th>
<th>187</th>
<th>12</th>
<th>28</th>
<th>28</th>
<th>18</th>
<th>54</th>
<th>70</th>
<th>20</th>
<th>206</th>
<th>80</th>
<th>215</th>
<th>219</th>
<th>153</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OR detectable</strong>&lt;br&gt;high power (0.8)</td>
<td>3.139</td>
<td>2.094</td>
<td>1.644</td>
<td>1.667</td>
<td>1.409</td>
<td>1.436</td>
<td>1.389</td>
<td>1.333</td>
<td>1.326</td>
<td>1.334</td>
<td>1.247</td>
<td>1.214</td>
<td>1.178</td>
<td>1.17</td>
<td>1.193</td>
</tr>
<tr>
<td>moderate power (0.5)</td>
<td>2.586</td>
<td>1.847</td>
<td>1.511</td>
<td>1.528</td>
<td>1.33</td>
<td>1.35</td>
<td>1.314</td>
<td>1.27</td>
<td>1.264</td>
<td>1.27</td>
<td>1.201</td>
<td>1.175</td>
<td>1.146</td>
<td>1.139</td>
<td>1.158</td>
</tr>
<tr>
<td>1.05</td>
<td>0.000057</td>
<td>0.00009</td>
<td>0.00017</td>
<td>0.00019</td>
<td>0.00037</td>
<td>0.00037</td>
<td>0.00046</td>
<td>0.00069</td>
<td>0.00087</td>
<td>0.00091</td>
<td>0.0018</td>
<td>0.0025</td>
<td>0.0054</td>
<td>0.0066</td>
<td>0.0067</td>
</tr>
<tr>
<td>1.1</td>
<td>0.0013</td>
<td>0.00035</td>
<td>0.0011</td>
<td>0.0013</td>
<td>0.004</td>
<td>0.0042</td>
<td>0.0058</td>
<td>0.011</td>
<td>0.015</td>
<td>0.017</td>
<td>0.041</td>
<td>0.062</td>
<td>0.15</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>1.2</td>
<td>0.0066</td>
<td>0.0035</td>
<td>0.018</td>
<td>0.025</td>
<td>0.094</td>
<td>0.1</td>
<td>0.14</td>
<td>0.25</td>
<td>0.33</td>
<td>0.34</td>
<td>0.63</td>
<td>0.72</td>
<td>0.88</td>
<td>0.9</td>
<td>0.76</td>
</tr>
<tr>
<td>1.3</td>
<td>0.026</td>
<td>0.02</td>
<td>0.11</td>
<td>0.15</td>
<td>0.44</td>
<td>0.43</td>
<td>0.53</td>
<td>0.73</td>
<td>0.76</td>
<td>0.75</td>
<td>0.92</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.95</td>
</tr>
<tr>
<td>1.5</td>
<td>0.021</td>
<td>0.16</td>
<td>0.57</td>
<td>0.62</td>
<td>0.92</td>
<td>0.84</td>
<td>0.91</td>
<td>0.93</td>
<td>0.91</td>
<td>0.89</td>
<td>0.96</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1.7</td>
<td>0.078</td>
<td>0.43</td>
<td>0.84</td>
<td>0.82</td>
<td>0.97</td>
<td>0.95</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.96</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1.7</td>
<td>0.24</td>
<td>0.72</td>
<td>0.95</td>
<td>0.89</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>0.53</td>
<td>0.9</td>
<td>0.99</td>
<td>0.95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.72</td>
<td>0.95</td>
<td>0.99</td>
<td>0.98</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**

1. Within each analysis, the number of SNPs whose mean SE was used to compute power is shown for each sub-range of MAF.
2. 'OR detectable' indicates the minimum risk effect detectable (expressed as the exponent of the beta from the meta-analysis) with high or moderate power. 'Power' indicates the study's power to detect the effect sizes (odds ratios) shown.
3. Effect sizes detectable or for which power is shown are expressed as the exponent of the absolute beta from the meta-analysis (i.e. the odds ratio computed with the lower risk group set as the reference group). Thus, in the CARDioGRAM, Females, Males, and Fine Mapping analyses, these are the odds ratios associated for each additional copy of the risk allele; in the Gender*Genotype Interaction analysis these are the odds ratios for difference in risk between sexes.
4. Power does not increase linearly with increase in MAF because these data are based on empirical SE values, which may be affected by other factors (e.g. imputation quality, between-study heterogeneity in the meta-analysis, etc.) for SNPs in some sub-ranges of MAF compared to others.
5. In the computation of power for given effect size, scenarios with high power (≥80%) are shaded dark grey, those with moderate power (≥50% and <80%) are shaded light grey, and those with power lower than 50% are unshaded.
Supplementary Figures

Supplementary Figure 1. Summary of quality control and analysis procedures in the fine mapping analysis.

<table>
<thead>
<tr>
<th>&lt;QC/analysis step&gt;</th>
<th>MIGen</th>
<th>WTCCC</th>
<th>FHS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study design</strong></td>
<td>case/control</td>
<td>case/control</td>
<td>cohort</td>
</tr>
<tr>
<td>prior subject-level QC (sample call rate)</td>
<td>≥0.95</td>
<td>≥0.95</td>
<td>≥0.95</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N (cases/controls)</td>
<td>6042</td>
<td>7368</td>
<td>3717*</td>
</tr>
<tr>
<td>Number of genotyped SNPs in the region of interest</td>
<td>149</td>
<td>81</td>
<td>115</td>
</tr>
<tr>
<td>SNP level QC (SNPs removed: ≤95% call rate; HWE p≤10^-6)</td>
<td>0.0</td>
<td>0.5</td>
<td>9.2</td>
</tr>
<tr>
<td><strong>Number of SNPs passing QC</strong></td>
<td>149</td>
<td>76</td>
<td>105</td>
</tr>
<tr>
<td><strong>Imputation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of genotyped and imputed SNPs</td>
<td>2473†</td>
<td>2472†</td>
<td>2477†</td>
</tr>
<tr>
<td>post-imputation QC (SNPs removed: IMPUTE2 INFO&lt;0.5)</td>
<td>891</td>
<td>970</td>
<td>814</td>
</tr>
<tr>
<td>Number of SNPs genotyped or imputed with high quality</td>
<td>1582†</td>
<td>1502†</td>
<td>1663†</td>
</tr>
<tr>
<td>remove SNPs not present in all studies (number of SNPs removed)</td>
<td>131</td>
<td>51</td>
<td>212</td>
</tr>
<tr>
<td><strong>Number of SNPs common to all studies</strong></td>
<td>1451</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The publicly available dataset for the Framingham study contained genotype and phenotype data for 9,270 individuals. In the current analysis, we included 3,717 individuals from the original and offspring cohorts for whom survival data were available for the follow-up periods beginning at visits 15 and 5 respectively. The sample selection and process used to filter these individuals is described in more detail in Lluís-Ganella et al. 2011 (submitted).

† For all three studies, all genotyped SNPs were also present in the 1kG reference panel. These values show the total number of SNPs, including genotyped SNPs, after imputation, post-imputation QC, and filtering to include SNPs common to all studies.
Supplementary Figure 2. CARDIoGRAM global meta-analysis results for the top SNP in the region of interest. Total sample size, number of cases, OR and 95% CI are shown for each contributing study, in addition to global sample sizes, OR, 95% CI, and p-values for association and heterogeneity. Note that only 12 of the 14 CARDIoGRAM studies are represented, as data for this variant was not available in the LURIC 1 and LURIC 2 samples.
Supplementary Figure 3. Meta-analysis results for the CARDIoGRAM SNP with the greatest difference in association between males and females (strongest interaction with gender). Total sample size, number of cases, OR and 95% CI are shown for each contributing study, in addition to global sample sizes, OR, 95%CI, and p-values for association and heterogeneity. Note that only 11 of the 14 CARDIoGRAM studies are represented, as data for this variant was not available in the LURIC 1, LURIC 2 and CHARGE samples.
Supplementary Figure 4. Distribution of minor allele frequencies (MAF) for SNPs analyzed in this study.

Data are shown as vertical bars whose widths are proportional to the ranges of MAF indicated on the x-axis, and whose heights correspond to the absolute number (left y-axis) of SNPs whose MAF falls within those ranges (MAF computed as the weighted mean in the MiGen, WTCCC and Framingham samples).

The number of SNPs within the region of interest that were genotyped or imputed in (a) the CARDIoGRAM meta-analysis (corresponding to the high-quality SNPs from the HapMapII reference panel) are indicated as white bars; (b) the number of additional SNPs imputed in the fine mapping analysis in this study are shown as light grey bars; the total the number of SNPs analyzed in the fine mapping analysis in this study (a plus b, corresponding to high-quality SNPs from the 1kG reference panel) are shown as dark grey bars.

Within each sub-range of MAF, the vertical black lines and diamonds at the top of the graph represent the proportions of SNPs analyzed in the fine mapping analysis (1kG panel, see (c) above). The portion of the line above the diamond represents the percentage (right y-axis) of these SNPs that were included in the HapMapII panel, and the portion below the diamond represents those additional SNPs that were imputed in the present study. This graph shows that many additional SNPs with a broad range of MAFs were imputed in this study, but that the greatest gain of information was obtained for rarer SNPs.

![Graph showing distribution of minor allele frequencies (MAF) for SNPs analyzed in this study.](image-url)
Supplementary References


Supplementary Appendix 1. Process of selection of participants from the Framingham study (from Lluis-Ganella et al., submitted).

**Mean follow-up: 13.32 years**
Supplementary Appendix 2. CARDIoGRAM Investigators

Executive Committee: Sekar Kathiresan1,2,3, Muredach P. Reilly4, Niles J. Samani5,6, Heribert Schunkert7

Executive Secretary: Jeanette Erdmann7


ADVANCE: Devin Absher18, Themistocles L. Assimesea, Stephen Fortmanns, Alan Go27, Mark Hlatkj y, Carlos Iribarren27, Joshua Knowless, Richard Myers18, Thomas Quertermous8, Steven Sidney27, Neil Risch28, Hua Tang29

CADomics: Stefan Blankenberg30, Tanja Zeller30, Arne Schillert12, Philipp Wild30, Andreas Ziegler12, Renate Schnabel30, Christoph Sinning30, Karl Lackner31, Laurence Tiret32, Viviane Nicaud32, Francois Cambien32, Christoph Bickel30, Hans J. Rupprecht30, Claire Perret32, Carole Proust32, Thomas Munzel30


deCODE: Solveig Gretarsdottir16, Jeffrey R. Gulcher16, Hilma Holm16, Augustine Kong16, Kari Stefansson16,17, Gudmundur Thorgeirsson53,17, Karl Andersen53,17, Gudmar Thorleifsson16, Unnur Thorsteinsdottir16,17

GERMIFS I and II: Jeanette Erdmann7, Marcus Fischer11, Anika Grosshennig12,7, Christian Hengstenberg11, Inke R. König12, Wolfgang Liebs4, Patrick Linsel-Nitschke7, Michael Preuss12,7, Klaus Stark11, Stefan Schreiber55, H.-Erich Wichmann56,58,59, Andreas Ziegler12, Heribert Schunkert7


MiGen: Benjamin F. Voight2,3,24, Kiran Musunurui,2,3, Candace Guiducci3, Noel Burtt3, Stacey B. Gabriels, David S. Siscovick50, Christopher J. O’Donnell47,
Post-genomic update on a classical candidate gene for coronary artery disease: ESR1

Roberto Elosua, Leena Peltonen, Veikko Salomaa, Stephen M. Schwartz, Olle Melander, David Altshuler, Sekar Kathiresan, Roberto Robert

**OHGS:** Alexandre F. R. Stewart, Li Chen, Sonny Dandona, George A. Wells, Olga Jarinova, Ruth McPherson, Robert Roberts

**PennCATH/MedStar:** Muredach P. Reilly, Mingyao Li, Liming Qu, Robert Wilensky, William Matthai, Hakon H. Hakonarson, Joe Devaney, Mary Susan Burnett, Augusto D. Pichard, Kenneth M. Kent, Lowell Satler, Joseph M. Lindsay, Ron Waksman, Christopher W. Knouff, Dawn M. Waterworth, Max C. Walker, Vincent Mosser, Stephen E. Epstein, Daniel J. Rader

**WTCCC:** Nilesh J. Samani, John R. Thompson, Peter S. Braund, Christopher P. Nelson, Benjamin J. Wright, Anthony J. Balmforth, Stephen G. Ball, Alistair S. Hall, Wellcome Trust Case Control Consortium

**Affiliations**

1 Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, MA, USA.
2 Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA.
3 Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, MA, USA.
4 The Cardiovascular Institute, University of Pennsylvania, Philadelphia, PA, USA.
5 Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK.
6 Leicester National Institute for Health Research Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, LE3 9QP, UK.
7 Medizinische Klinik II, Universitat zu Lubeck, Lubeck, Germany.
8 Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA.
9 University of Texas Health Science Center, Human Genetics Center and Institute of Molecular Medicine, Houston, TX, USA.
10 Division of Cardiovascular and Neuronal Remodelling, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, UK.
11 Klinik und Poliklinik fur Innere Medizin II, Universitat Regensburg, Regensburg, Germany.
12 Institut fur Medizinische Biometrie und Statistik, Universitat zu Lubeck, Lubeck, Germany.
13 Science Center, Tampere University Hospital, Tampere, Finland.
14 The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada.
15 Department of Health Sciences, University of Leicester, Leicester, UK.
16 deCODE Genetics, 101 Reykjavik, Iceland.
17 University of Iceland, Faculty of Medicine, 101 Reykjavik, Iceland.
18 Hudson Alpha Institute, Huntsville, Alabama, USA.
19 Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, 40 Ruskin Street, Ottawa, Ontario, Canada, K1Y 4W7.
20 Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.
21 National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA.
22 The Blavatnik School of Computer Science and the Department of Molecular Microbiology and Biotechnology, Tel-Aviv University, Tel-Aviv, Israel, and the International Computer Science Institute, Berkeley, CA, USA.
23 Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA.
Post-genomic update on a classical candidate gene for coronary artery disease: ESR1

57 Institut fur Humangenetik, Helmholtz Zentrum Munchen, Deutsches Forschungszentrum fur Umwelt und Gesundheit, Neuherberg, Germany.
58 Institute of Medical Information Science, Biometry and Epidemiology, Ludwig-Maximilians-Universitat Munchen, Germany.
59 Klinikum Grosshadern, Munich, Germany.
60 Institut fur Humangenetik, Technische Universitat Munchen, Germany.
61 Division of Endocrinology and Diabetes, Graduate School of Molecular Endocrinology and Diabetes, University of Ulm, Ulm, Germany.
62 Division of Endocrinology, Department of Medicine, Medical University of Graz, Austria.
63 Synlab Center of Laboratory Diagnostics Heidelberg, Heidelberg, Germany.
64 Division of Clinical Chemistry, Department of Medicine, Albert Ludwigs University, Freiburg, Germany.
65 LURIC non profit LLC, Freiburg, Germany.
66 Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University Graz, Austria.
67 Institute of Public Health, Social and Preventive Medicine, Medical Faculty Manneim, University of Heidelberg, Germany.
68 Cardiology Group Frankfurt-Sachsenhausen, Frankfurt, Germany.
69 Cardiovascular Epidemiology and Genetics Group, Institut Municipal d’Investigacio Medica, Barcelona. Ciber Epidemiologia y Salud Publica (CIBERSP), Spain.
70 Chronic Disease Epidemiology and Prevention Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland.
71 Department of Molecular Biology and Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, USA.
72 The Center for Applied Genomics, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania, USA.
73 Cardiovascular Research Institute, Medstar Health Research Institute, Washington Hospital Center, Washington, DC 20010, USA.
74 Genetics Division and Drug Discovery, GlaxoSmithKline, King of Prussia, Pennsylvania 19406, USA.
75 The Institute for Translational Medicine and Therapeutics, School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
76 Department of Cardiovascular Surgery, University of Leicester, Leicester, UK.
77 Division of Cardiovascular and Diabetes Research, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, LS2 9JT, UK.
78 LIGHT Research Institute, Faculty of Medicine and Health, University of Leeds, Leeds, UK.