

Disclaimer: The manuscript and its contents are confidential,
intended for journal review purposes only, and not to be further
disclosed.

URL: <http://circ-submit.aha-journals.org>

Title: Loss of cardio-protective effects at the ADAMTS7 locus due to
gene-smoking interactions

Manuscript number: CIRCULATIONAHA/2016/022069R2

Author(s): Danish Saleheen, University of Pennsylvania

Wei Zhao, Translational Medicine and Human Genetics

Robin Young, Department of Public Health and Primary Care,

University of Cambridge, UK

Christopher Nelson, University of Leicester

Weang Kee Ho, University of Cambridge

Jane Ferguson, Vanderbilt University Medical Center

Asif Rasheed, Center for Non-Communicable Diseases, Karachi, Pakistan

Kristy Ou, University of Pennsylvania

Sylvia Nurnberg, University of Pennsylvania

Robert Bauer, Columbia University

Anuj Goel, University of Oxford

Ron Do, The Charles Bronfman Institute of Personalized Medicine,
Icahn School of Medicine at Mount Sinai, New York, USA; Department
of Genetics and Genomic Sciences, Icahn School of Medicine, NY, USA

Alexandre Stewart, University of Ottawa Heart Institute

Jaana Hartiala, Institute for Genetic Medicine, University of
Southern California Keck School of Medicine

Weihua Zhang, Department of Epidemiology and Biostatistics, Imperial
College London

Gudmar Thorleifsson, deCODE Genetics

Rona Strawbridge, Karolinska Institutet

Juha Sinisalo, Helsinki University Central Hospital

Stavroula Kanoni, William Harvey Research Institute, Barts and The
London School of Medicine and Dentistry, Queen Mary University of London

Sanaz Sedaghat, Erasmus Medical Center

Eirini Marouli, William Harvey Research Institute, Barts and The
London School of Medicine and Dentistry, Queen Mary University of London

Kati Kristiansson, National Institute for Health and Welfare

Jing Hua Zhao, University of Cambridge School of Clinical Medicine

Robert Scott, University of Cambridge

Dominique Gauguier, INSERM UMRS_1138, Paris, France

Svati Shah, Duke University

Albert Smith, Icelandic Heart Association

Natalie Van Zuydam, Medical Research Institute

Amanda Cox, Center for Genomics and Personalized Medicine Research,

Wake Forest School of Medicine, Winston-Salem, NC, USA

Christina Willenborg, Medizinische Klinik II - AG Kardiovaskuläre

Genomik, Universität zu Lübeck

Thorsten Kessler, Deutsches Herzzentrum München

Lingyao Zeng, Technische Universität München

Michael Province, Washington University School of Medicine

Andrea Ganna, Karolinska Institutet

Lars Lind, Uppsala University

Nancy Pedersen, Karolinska Institutet

Charles White, Boston University School of Public Health

Anni Joensuu, National Institute for Health and Welfare

Marcus Kleber, Medical Faculty of Mannheim, University of Heidelberg

Alistair Hall, Leeds University

Winfried März, Synlab Academy, Synlab Services GmbH, Mannheim,

Germany, Clinical Institute of Medical and Chemical Laboratory

Diagnostics, Medical University of Graz, Graz, Austria, Medical

Clinic V, UMM

Veikko Salomaa, National Institute for Health and Welfare

Christopher O'Donnell, Boston Veterans Affairs Healthcare System

Erik Ingelsson, Uppsala University

Mary Feitosa, Washington University School of Medicine

Jeanette Erdmann, University of Lübeck

Donald Bowden, Wake Forest University School of Medicine

Colin Palmer, University of Dundee

Vilmundur Gudnason, Icelandic Heart Association

Ulf de Faire, Karolinska Institute

Pierre Zalloua, Lebanese American University

Nicholas Wareham, University of Cambridge

John Thompson, University of Leicester

Kari Kuulasmaa, National Institute for Health and Welfare (THL)

George Dedoussis, Department of Dietetics-Nutrition, Harokopio

University, 70 El. Venizelou Str

Markus Perola, National Public Health Institute

Abbas Dehghan, Erasmus University Medical Center

John Chambers, Imperial College London

Jaspal Kooner, Imperial College London

Hooman Allayee, Institute for Genetic Medicine, University of

Southern California Keck School of Medicine

Panos Deloukas, Queen Mary University of London

Ruth McPherson, Atherogenomics Laboratory, Ruddy Canadian

Cardiovascular Genetics Centre

Kari Stefansson, deCODE Genetics

Heribert Schunkert, Deutsches Herzzentrum München

Sekar Kathiresan, Massachusetts General Hospital

Martin Farrall, University of Oxford

Philippe Frossard, Center for Non-Communicable Diseases, Karachi, Pakistan

Daniel Rader, University of Pennsylvania School of Medicine

Nilesh Samani, Department of Cardiovascular Sciences, University of

Leicester, Leicester, LE3 9QP, UK

Muredach Reilly, Columbia University

1 Loss of cardio-protective effects at the *ADAMTS7* locus due to gene-smoking interactions

2 Danish Saleheen PhD,^{1,2} Wei Zhao MSc,¹ Robin Young PhD,³ Christopher P Nelson PhD,⁴
3 WeangKee Ho PhD,³ Jane Ferguson PhD,⁵ Asif Rasheed MBBS,² Kristy Ou BS,⁵ Sylvia T Nurnberg
4 PhD,^{5,7} Robert C. Bauer PhD,^{7,6} Anuj Goel MSc,⁶ Ron Do PhD,^{7,8} Alexandre FR Stewart PhD,⁹ Jaana
5 Hartiala PhD,¹⁰ Weihua Zhang PhD,^{11,12} Gudmar Thorleifsson PhD,^{50,51} Rona J Strawbridge PhD,¹³
6 Juha Sinisalo PhD,⁶¹ Stavroula Kanoni PhD,¹⁵ Sanaz Sedaghat PhD,¹⁶ Eirini Marouli PhD,^{15,17} Kati
7 Kristiansson PhD,¹⁸ Jing Hua PhD,¹⁹ Robert Scott PhD,¹⁹ Dominique Gauguier PhD,²⁰ Svati Shah
8 MD,²¹ Albert Vernon Smith PhD,^{22,36} Natalie R van Zuydam PhD,²³ Amanda J Cox PhD,²⁴ Christina
9 Willenborg PhD,^{25,26} Thorsten Kessler MD,^{27,28} Lingyao Zeng PhD,^{27,29} Michael A Province PhD,³⁰
10 Andrea Ganna PhD,^{31,32} Lars Lind PhD,³³ Nancy L. Pedersen PhD,³⁴ Charles C White PhD,³⁵ Anni
11 Joensuu MSc,^{18,37} Marcus E Kleber PhD,³⁸ Alistair S Hall PhD,³⁹ Winfried März PhD,⁴⁰ Veikko
12 Salomaa PhD,¹⁸ Christopher O'Donnell MD,⁴¹ Erik Ingelsson PhD,^{42,43} Mary F Feitosa PhD,³⁰
13 Jeanette Erdmann PhD,^{25,26} Donald W. Bowden PhD,²⁴ Colin N Palmer PhD,²³ Vilmundur Gudnason
14 PhD,^{22,36} Ulf De Faire PhD,⁴⁴ Pierre Zalloua PhD,⁴⁵ Nick Wareham PhD,²⁰ John R Thompson PhD,⁴⁶
15 Kari Kuulasmaa PhD,¹⁸ George Dedoussis PhD,¹⁷ Markus Perola PhD,^{18,37} Abbas Dehghan PhD,¹⁶
16 John C Chambers PhD,^{11,12,47} Jaspal Kooner MD,^{11,47,48} Hooman Allayee PhD,¹⁰ Panos Deloukas
17 PhD,^{15,49} Ruth McPherson PhD,⁹ Kari Stefansson PhD,^{50,51} Heribert Schunkert MD,^{27,29} Sekar
18 Kathiresan MD,⁵²⁻⁵⁵ Martin Farrall PhD,⁶ EPIC-CVD,³ Philippe Frossard DSC,² Daniel J Rader
19 MD,^{56,57} Nilesh Samani MD,⁴ PROMIS,² CARDIoGRAMplusC4D, Muredach P. Reilly MD.⁷⁶

20 Collaborators:

21 Stanley L. Hazen MD,⁵⁸ W.H. Wilson Tang MD,⁵⁸ Perttu P Salo PhD,^{18,37} Marja-Liisa Lokki PhD,⁶⁰
22 Markku S Nieminen PhD,⁶¹ Antti-Pekka Sarin MSc,^{18,37} Alun Evans MSc,⁶³ Jean Ferrières MD,⁶⁴
23 Jarmo Virtamo PhD,¹⁸ Frank Kee PhD,⁶⁶ David-Alexandre Trégouët PhD,⁶⁷ Dominique Arveiler
24 PhD,⁶⁸ Philippe Amouyel PhD,⁶⁹ Paolo Brambilla PhD,⁷⁰ Annette Peters PhD,⁷¹ Melanie
25 Waldenberger PhD,^{71,75} Giovanni Veronesi PhD,⁷² Giancarlo Cesana PhD,⁷³ Satu Männistö PhD,¹⁸
26 Pekka Jousilahti PhD,¹⁸ Antti M Jula PhD,¹⁸ Kennet Harald PhD,¹⁸ Albert Hofman PhD,¹⁶ Oscar H.
27 Franco PhD,¹⁶ Andre G. Uitterlinden PhD.⁷⁴

28
29 1. Department of Biostatistics and Epidemiology, University of Pennsylvania, USA
30 2. Center for Non-Communicable Diseases, Karachi, Pakistan.
31 3. Department of Public Health and Primary Care, University of Cambridge, UK
32 4. Department of Cardiovascular Sciences, University of Leicester, Leicester, LE3 9QP, UK
33 5. Cardiology Division, Department of Medicine, Vanderbilt University, Nashville, USA
34 6. Division of Cardiovascular Medicine, Radcliffe Department of Medicine & Wellcome Trust Centre for Human
35 Genetics, University of Oxford, UK
36
37 7. The Charles Bronfman Institute of Personalized Medicine, Icahn School of Medicine at Mount Sinai, New
38 York, NY, USA
39
40 8. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY,
41 USA
42
43 9. Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada
44
45 10. Institute for Genetic Medicine and Department of Preventive Medicine, Keck School of Medicine, University
46 of Southern California, Los Angeles, CA 90033

47 11. Department of Epidemiology and Biostatistics, Imperial College London, London, UK
48 12. Department of Cardiology, Ealing Hospital NHS Trust, Middlesex, UK
49 13. Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
50 14. Department of Medicine, Department of Cardiology, Helsinki University Central Hospital, Helsinki, Finland
51 15. William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary
52 University of London, London, UK
53 16. Department of Epidemiology, Erasmus University Medical center, Rotterdam, The Netherlands
54 17. Department of Dietetics-Nutrition, Harokopio University, 70 El. VenizelouStr, Athens, Greece
55 18. National Institute for Health and Welfare, Helsinki, Finland
56 19. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge School of Clinical
57 Medicine, Box 285, Cambridge Biomedical Campus, Cambridge, CB2 0QQ
58 20. INSERM, UMRS1138, Centre de Recherche des Cordeliers, Paris, France
59 21. Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, NC 27710
60 22. Icelandic Heart Association, Kopavogur, Iceland
61 23. Medical Research Institute, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK
62 24. Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-
63 Salem, NC, USA
64 25. Institut für Integrative und Experimentelle Genomik, Universität zu Lübeck, Lübeck, Germany
65 26. DZHK (German Research Center for Cardiovascular Research) partner site Hamburg–Lübeck–Kiel,
66 Lübeck, Germany
67 27. Deutsches Herzzentrum München, Technische Universität München, München, Germany
68 28. Klinikum rechts der Isar, München, Germany
69 29. DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, München,
70 Germany
71 30. Department of Genetics, Washington University School of Medicine, St. Louis, MO
72 31. Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and
73 Harvard Medical School, Boston, Massachusetts 02114, USA
74 32. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center,
75 Cambridge, Massachusetts 02142, USA
76 33. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden
77 34. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
78 35. Department of Biostatistics Boston University School of Public Health Framingham Heart Study 801
79 Massachusetts Avenue, CT3 Boston, MA 02118
80 36. Faculty of Medicine, University of Iceland, Reykjavik, Iceland
81 37. University of Helsinki, Institute for Molecular Medicine, Finland (FIMM)
82 38. Department of Medicine, Mannheim Medical Faculty, Heidelberg University
83 39. Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, UK
84 40. Synlab Academy, Synlab Services GmbH, Mannheim, Germany and Clinical Institute of Medical and
85 Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria
86 41. National Heart, Lung, and Blood Institute and the Framingham Heart Study, National Institutes of Health,
87 Bethesda, MD
88 42. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala
89 University, Uppsala, Sweden
90 43. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine,
91 Stanford, California, USA
92 44. Division of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet,
93 Stockholm, Sweden
94 45. Lebanese American University, School of Medicine, Beirut 1102 2801, Lebanon
95 46. Department of Health Sciences, University of Leicester, Leicester, UK
96 47. Imperial College Healthcare NHS Trust, London, UK
97 48. Cardiovascular Science, National Heart and Lung Institute, Imperial College London, London, UK
98 49. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD),
99 King Abdulaziz University, Jeddah, Saudi Arabia
100 50. deCODE Genetics, Sturlugata 8, IS-101 Reykjavik, Iceland
101 51. University of Iceland, School of Medicine, 101 Reykjavik, Iceland

102 52. Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge,
103 Massachusetts, USA
104 53. Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA
105 54. Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA
106 55. Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA
107 56. Department of Genetics, University of Pennsylvania, USA
108 57. Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of
109 Medicine, University of Pennsylvania, USA
110 58. Departments of Cardiovascular Medicine and Cellular and Molecular Medicine, Cleveland Clinic,
111 Cleveland, OH 44195
112 60. Transplantation Laboratory, Haartman Institute, Helsinki, Finland
113 61. Heart and Lung Center HUH, Helsinki University Hospital and Helsinki University, Finland
114 63. Centre for Public Health, The Queen's University of Belfast, Belfast, Northern Ireland
115 64. INSERM UMR 1027, Department of Cardiology, Toulouse University School of Medicine, Rangueil
116 Hospital, 31059 Toulouse, France
117 66. UKCRC Centre of Excellence for Public Health (NI), Queens University of Belfast, Northern Ireland
118 67. Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé
119 (UMR_S) 1166, F-75013, Paris, France AND Sorbonne Universités, Université Pierre et Marie Curie, Paris,
120 France; and Institute for Cardiometabolism and Nutrition, Paris, France
121 68. Department of Epidemiology and Public Health, EA3430, University of Strasbourg, Faculty of Medicine,
122 Strasbourg, France
123 69. Institut Pasteur de Lille, INSERM U744, Université Lille Nord de France, F-59000 Lille, France
124 70. Laboratory Medicine, Hospital of Desio, Department of Health Sciences, University of Milano, Bicocca, Italy
125 71. Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany AND German Center for
126 Cardiovascular Disease Research (DZHK e.V.), Munich, Germany Melanie Waldenberger
127 72. EPIMED Research Center, Department of Clinical and Experimental Medicine, University of Insubria,
128 Varese, Italy
129 73. University of Milano, Bicocca, Italy
130 74. Department of Internal Medicine, Erasmus University Medical center, Rotterdam, The Netherlands
131 75. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
132 76. Cardiology Division, Department of Medicine and the Irving Institute for Clinical and Translational
133 Research, Columbia University Medical Center, New York, USA
134

135 **Corresponding authors:**

136 Danish Saleheen, MBBS, PhD
137 Assistant Professor of Biostatistics and Epidemiology
138 University of Pennsylvania
139 11-134 Translational Research Center
140 3400 Civic Center Boulevard
141 Philadelphia, PA 19104
142 Tel: 215-573-6323
143 Fax: 215-573-2094
144 Email: saleheen@mail.med.upenn.edu
145

146 Muredach P. Reilly, MBBCH, MSCE
147 Herbert and Florence Irving Professor of Medicine
148 Director of the Irving Institute for Clinical and Translational Research
149 Columbia University
150 622 West 168th Street, PH10-305
151 New York, NY 10032
152 Tel: 212-305-9453
153 Fax: 212-305-3213
154 Email: mpr2144@cumc.columbia.edu

155 **ABSTRACT**

156 **Background** Common diseases such as coronary heart disease (CHD) are complex in etiology. The
157 interaction of genetic susceptibility with lifestyle factors may play a prominent role. However, gene-
158 environment interactions for CHD have been difficult to identify. Here, we investigate interaction of
159 smoking behavior, a potent lifestyle factor, with genotypes that have been shown to associate with
160 CHD risk.

161 **Methods** We analyzed data on 60,919 CHD cases and 80,243 controls from 29 studies for gene-
162 smoking interactions for genetic variants at 45 loci previously reported to associate with CHD risk.
163 We also studied 5 loci associated with smoking behavior. Study specific gene-smoking interaction
164 effects were calculated and pooled using fixed-effects meta-analyses. Interaction analyses were
165 declared to be significant at a *P-value* < 1.0×10^{-3} (Bonferroni correction for 50 tests).

166 **Results** We identified novel gene-smoking interaction for a variant upstream of the *ADAMTS7* gene.
167 Every T allele of rs7178051 was associated with lower CHD risk by 12% in never-smokers (*P-value*:
168 1.3×10^{-16}) compared to 5% in ever-smokers (*P-value*: 2.5×10^{-4}) translating to a 60% loss of CHD
169 protection conferred by this allelic variation in people who smoked tobacco (*Interaction P-value*:
170 8.7×10^{-5}). The protective T allele at rs7178051 was also associated with reduced *ADAMTS7*
171 expression in human aortic endothelial cells and lymphoblastoid cell lines. Exposure of human
172 coronary artery smooth muscle cells to cigarette smoke extract led to induction of *ADAMTS7*.

173 **Conclusion** Allelic variation at rs7178051 that associates with reduced *ADAMTS7* expression
174 confers stronger CHD protection in “never-smokers” compared to “ever-smokers”. Increased
175 vascular *ADAMTS7* expression may contribute to the loss of CHD protection in smokers.

176 **Key words:** Gene-smoking interaction, gene-environment interaction, coronary heart disease,
177 *ADAMTS7*, smoking.

178 **Word count: 269**

179

180

181 **Clinical Perspective**

182 1) What is new?

- 183
- 184 • Using data on 60,919 CHD cases and 80,243 controls, this study conducted gene-
185 environment interaction analyses to investigate effect modification by smoking behavior at
186 established CHD and smoking related loci.
 - 187 • Cardio-protective effects conferred by allelic variation at the *ADAMTS7* locus attenuated by
188 60% in people who smoked tobacco compared to those who did not smoke.
 - 189 • Allelic variation at *ADAMTS7* associated with reduced CHD risk was associated with reduced
190 *ADAMTS7* expression in human aortic endothelial cells and lymphoblastoid cell lines.
 - 191 • Exposure of human coronary artery smooth muscle cells to cigarette smoke extract led to
192 induction of *ADAMTS7*.

193

194 2) What are the clinical implications?

- 195
- 196 • These human genomic data provide new insights into potential mechanisms of CHD in
197 cigarette smokers.
 - 198 • Findings from this study also point towards the directional impact of the *ADAMTS7* locus on
199 CHD.
 - 200 • *ADAMTS7* and its substrates have a specific role in cigarette smoking related CHD.
 - 201 • Inhibition of *ADAMTS7* is a novel potential therapeutic strategy for CHD that may have
202 particular benefits in individuals who smoke cigarettes.
- 203
- 204
- 205
- 206
- 207
- 208
- 209
- 210

211 **INTRODUCTION**

212 Coronary heart disease (CHD) is a complex disorder resulting from the interplay of lifestyle
213 and genetic factors.^{1, 2} Yet, gene-environment interactions for CHD have been difficult to identify.
214 Cigarette smoking is one of the strongest lifestyle risk factors for CHD but the underlying molecular
215 mechanisms of CHD in humans who smoke remain uncertain.³⁻⁵ Cigarette smoking accounts for
216 one-fifth of all CHD events globally and is responsible for ~1.6 million deaths attributable to CHD
217 annually.⁶ Genome-wide association studies (GWAS) have improved our understanding on the
218 genetic predisposition to both CHD and smoking behavior.⁷⁻¹⁰ Joint or interactive effects of genetic
219 variation and smoking behavior in the etiology of CHD, however, remain poorly understood. GWAS
220 can provide new opportunities to investigate gene-smoking interactions.

221 We hypothesized that genetic predisposition to CHD is modified by cigarette smoking at
222 loci discovered by GWAS to be associated with either CHD or smoking behavior. We conducted a
223 focused experiment at 50 main-effect loci (45 for CHD and 5 for smoking behavior) using genetic
224 data and information on smoking behavior in 60,919 CHD cases and 80,243 controls from 29
225 different studies. We report novel findings on gene-smoking interactions in CHD. Allelic variation on
226 chr.15q25.1 at *ADAMTS7* is associated with protection from CHD in “never-smokers” with
227 attenuation of this protective effect in people who smoked. Expression studies in relevant vascular
228 cells support a role for *ADAMTS7* in smoking induced CHD. These data provide the first insights on
229 the etiology of CHD in cigarette smokers and may present opportunities for targeted therapeutic
230 strategies to lower CHD risk in individuals who smoke cigarettes.

231

232

233

234

235

236

237

238

239 **METHODS**

240 **Summary of study Design**

241 All studies participating in the CARDIoGRAMplusC4D consortium⁷⁻⁹ that had information
242 available on smoking status, CHD risk and genotypes at the 50 CHD and smoking behavior-
243 associated loci were invited to participate. The current study had five inter-related components
244 (**eFigure-1**). First, as part of the quality control, we investigated the association of smoking status
245 with CHD risk within each study. Second, we performed an updated analysis of all the SNPs (\pm 50
246 KB) at the 45 established CHD loci to identify the variant with the strongest CHD association in our
247 study population at each established CHD locus. Effect estimates from each study in association
248 with CHD risk were obtained and pooled to identify the strongest CHD associated variant (“lead
249 variant”). Third, we investigated gene-smoking interactions at these 45 CHD loci and at 5 loci related
250 to smoking behavior. Fourth, for loci demonstrating differential CHD associations by smoking status,
251 we mapped the interaction region, examined linkage disequilibrium (LD) across the region and
252 performed conditional analyses to identify independent genetic signals. Finally, for loci exhibiting
253 gene-smoking interaction in CHD, we assessed eQTL data for association of variants with
254 expression of local genes in available datasets and examined expression of these genes in multiple
255 cell types that play prominent roles in smoking-CHD pathobiology.

256 **Harmonization of phenotypes and genotypes**

257 Summary level estimates for each study were shared via a secure FTP site. We used
258 “ever-smoking” as a primary exposure and data were harmonized by uniformly characterizing
259 participants in each study into two categories, “ever-smokers” and “never-smokers”. “Ever-smokers”
260 were defined as those who had smoked more than 100 cigarettes in a lifetime. For case-control
261 studies, information on “ever smoking” status collected at the time of enrollment was used for the
262 current analyses; whereas for prospective cohort studies, information on smoking status obtained at
263 the baseline visit was used for the current investigation. CHD was defined based on evidence from
264 angiography or history of verified myocardial infarction (MI), percutaneous coronary interventions
265 (PCI) or coronary artery bypass grafting (CABG) as published in CARDIoGRAMplusC4D projects.⁷⁻⁹
266 Genotype data generated through GWAS (directly genotyped or imputed) or cardio-metabochip
267 (directly genotyped only) arrays were obtained from each study and all genetic data were aligned
268 using the build-37 reference panel. Imputed SNPs were removed if they had any of the following: (i)
269 a minor allele frequency of <1%; (ii) info score of <0.90; or (iii) confidence score <0.90. For each
270 study, GWAS data were imputed using the Phase II CEU HapMap reference population.¹¹ Standard

271 quality control criteria were applied by each participating study, as described previously.⁷ All
272 participating studies in the CARDIoGRAMplusC4D consortium were approved by their locally
273 relevant institutional review boards and all participants gave written informed consent before their
274 enrollment in each study.⁷⁻⁹

275 **STATISTICAL ANALYSIS**

276 **Gene-smoking interaction analyses**

277 Initial quality control and association of established CHD loci with CHD risk: As part of an initial
278 quality control, effect estimates from each study were obtained for “ever-smoking” status and CHD
279 risk using a case-control logistic regression model adjusted for age and sex. Each participating study
280 also assessed and, if needed, controlled for population stratification by including principal
281 components as covariates in the regression model as described earlier.⁷⁻⁹ To identify variant(s) with
282 the most significant association with CHD risk at established CHD loci in our study population,
283 logistic regression analyses were conducted by each participating study for all the SNPs flanking
284 (± 50 kb) the lead variant previously reported at each CHD locus. Effect estimates and standard
285 errors were obtained and meta-analyzed using a fixed-effects inverse variance approach. All lead
286 variants identified through these analyses were further investigated for gene-smoking interactions in
287 CHD. One lead variant per locus was selected for primary gene-smoking interaction analyses.

288 Investigation of the APOE locus: Although APOE has been recently established as a GWAS locus,⁷
289 previous studies prior to GWAS have suggested that CHD risk is higher among carriers of the $\epsilon 4$
290 allele at the APOE locus in smokers than in non-smokers.¹²⁻¹⁴ Because the $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$ alleles at
291 the APOE locus are not captured by the GWAS platform, we specifically conducted genotyping for
292 rs429358 and rs7412 variants to capture the three epsilon (ϵ) alleles in 13,822 participants (including
293 7,286 first-onset myocardial infarction cases) in the PROMIS study.¹⁵

294 Gene-smoking interaction analyses at CHD and smoking loci: To assess gene-smoking interactions,
295 analyses were conducted within each study, adjusted for age, sex and other study specific
296 covariates (e.g., principal components), and variants were analyzed in association with CHD
297 separately in “ever-smokers” and “never-smokers”. Results from the two groups were then used to
298 test for interaction within each study. For the 50 variants, an interaction test statistic was calculated
299 within each study using the following equation as adapted from Teslovich *TM* et.al.¹⁶

$$\frac{(\beta_n - \beta_e)}{\sqrt{SE_n^2 + SE_e^2}}$$

300 where β_n and β_e are the beta coefficients for the SNP in never-smokers and ever-smokers
301 respectively, SE_n and SE_e are the standard errors for the log-ORs estimated for never-smokers and
302 ever-smokers, respectively. Study specific interaction beta(s) and se(s) were calculated within each
303 study and were pooled across studies using a fixed-effects meta-analysis. Interaction analyses were
304 declared to be significant at a P-value of $<1.0 \times 10^{-3}$ (Bonferroni correction for 50 tests).

305 Conditional analyses on chr.15q25.1: At chr.15q25.1, we observed two variants exhibiting gene-
306 smoking interactions for CHD. The proximity of these two signals raised the possibility that the
307 observed interactions may represent a single interaction locus across the entire region. To
308 investigate this possibility we undertook conditional analyses using an approximate conditional and
309 joint analyses approach, also known as GCTA (Genome-wide Complex Trait Analysis), as described
310 previously.¹⁷⁻²² Briefly, this method leverages summary-level statistics from a meta-analysis and uses
311 LD corrections between SNPs estimated from a reference sample. Such an approach has been
312 shown to yield similar results to that obtained from conditional analyses conducted on individual
313 participant data and has been successfully implemented in several other studies that have fine-
314 mapped loci for other complex traits.¹⁷⁻²² Using this approach, we first conducted separate
315 conditional analyses at the chr.15q25.1 locus to identify main-effect variant(s) independently
316 associated with CHD and smoking behavior, respectively. We used the meta-analyzed data for CHD
317 main effects in the CARDIoGRAMplus4D consortium to identify SNPs independently associated with
318 CHD risk and we used the genetic meta-analysis data from the Tobacco and Genetics Consortium
319 (TGC) in 140,000 participants to identify variants independently associated with smoking behavior.
320 We then estimated the effects of these independent variants on CHD risk stratified by smoking
321 status and mutually adjusted the effects of these variants for each other.

322

323 **Analysis of eQTLs and regulatory features at the chr15q25.1 gene-smoking interaction locus**

324 eQTL analyses: We mined publicly available databases to identify genotype-related expression
325 differences (eQTLs) in *ADAMTS7* and the *CHRNA4-A3-A5* gene cluster in order to understand the
326 directionality of the association of expression of these genes with CHD and smoking behavior.
327 Specifically, we investigated data available from the GTEx consortium,²³ the HapMap consortium
328 (restricted to European populations), and the Multiple Tissue Human Expression Resource
329 (MuTHER).²⁴ We also analyzed expression data in 147 donor HAOEC lines.²⁵ We used a nominal P-
330 value of 0.002 to account for multiple testing involved in the eQTL analyses.

331 Regulatory features of the chr. 15q25.1 region: Data from ENCODE²⁶ were explored as described in
332 eMethods. CHIP-seq experiments were performed on confluent HCASMC (Cell Applications 350-05a
333 & Lonza CC-2583; cultured in SmGM-2 BulletKit media; Lonza) as described.²⁷ TCF21 (Abcam
334 ab49475), Jun (Santa Cruz Biotechnology sc-1694), JunD (Santa Cruz Biotechnology sc-74), and
335 CEBP (Santa Cruz Biotechnology sc-150) transcription factor binding was interrogated and H3K27ac
336 data were acquired using the same CHIP protocol with an anti-H3K27ac antibody (Abcam; ab4729).
337 Reads were aligned to the human genome (GRCh37p13) using STAR.²⁸

338

339 **Analyses of ADAMTS7 and CHRN4-A3-A5 gene expression in vascular cells and tissues**

340 ADAMTS7 and CHRN4-A3-A5 gene expression in vascular cells: ADAMTS7 and CHRN4-A3-A5
341 mRNA levels were measured in cultured human coronary artery smooth muscle cells (HCASMC;
342 Lonza CC-2583, Lonza Walkersville, MD), human coronary artery endothelial cells (HCAEC, Lonza
343 CC-2585), human aortic smooth muscle cells (HAoSMC, Lonza CC-2571), human aortic endothelial
344 cells (HAoEC, Lonza CC-2535), human aortic adventitial fibroblasts (HAoAF, Lonza CC-7014), and
345 human acute monocytic leukemia cell line (THP-1, ATCC TIB-202). Further details are provided in
346 eMethods.

347 ADAMTS7 and CHRN4-A3-A5 gene expression in response to cigarette smoke extract: HCASMC
348 were grown to confluence and cigarette smoke extract experiments performed at passage-7.
349 Cigarette smoke extract was custom-prepared by Arista Laboratories (Richmond, VA). Briefly, the
350 condensate was generated by smoking Marlboro Red King Size Hard Pack cigarettes on an
351 analytical smoke machine under International Organization for Standardization smoking conditions.
352 The smoke condensate was collected on 92 mm filter pads and extracted from each pad in DMSO
353 by shaking to obtain a solution of ~20 mg/mL final concentration of the total particulate matter.
354 Serum starved (24 hrs) HCASMC were treated with 0.5% or 1.0% cigarette smoke extract (v/v) for 4,
355 12, and 24 hrs in serum reduced conditions (0.5% FBS in DMEM). Details on RNA preparation and
356 q-PCR are provided in eMethods.

357

358

359

360

361 RESULTS

362 Description of the participating studies

363 Of the 37 studies participating in the CARDIoGRAMplusC4D consortium, information on
364 “ever-smoking” was available in 30 studies, yielding a total sample size of 60,919 CHD cases and
365 80,243 controls. All studies recruited participants of European ancestry, except PROMIS (South
366 Asian),¹⁵ LOLIPOP (South Asian)²⁹ and FGENTCARD (Lebanese).³⁰ Number of CHD cases and
367 controls and percentages that were “ever-smokers” are provided in **eTable 1**. As expected, in all the
368 participating studies, association of “ever-smoking” status with CHD risk was directionally consistent
369 with an increased risk of CHD (**eFigure 2**).

370 New variants associated with CHD at established loci

371 **eFigure 3** and **eTable 2** present effect estimates for the association with CHD for (i) the
372 most significant variant that we identified at known CHD loci in the current CARDIoGRAMplusC4D
373 consortium analysis as well as for (ii) the top SNP previously reported at each of these established
374 CHD loci. Of the 45 established CHD loci, we identified 32 for which we discovered a more
375 statistically significant SNP in association with CHD risk in our dataset than the prior reported top
376 variant. All of these 32 SNPs were in moderate to high LD ($r^2 > 0.6$) with the previously published
377 variants.⁷⁻⁹ In our primary gene-smoking interaction analyses, at each of the CHD loci, we, therefore,
378 used the SNP with the most significant CHD association (**eFigure 3** and **eTable 2**). Because the
379 smoking behavior phenotype (captured as cigarettes per day [CPD]) was not available in all
380 CARDIoGRAMplusC4D studies, we used the top variant previously reported for CPD¹⁰ at each locus
381 (**eFigure 4**).

382 Analyses of the APOE locus.

383 The effect of rs6857, the lead CHD variant at the *APOE* locus, was similar in “ever-
384 smokers” compared to “never-smokers” (**eTable 3**). Specifically, the CHD OR for the T allele at
385 rs6857 was found to be 1.10 (P-value 7.93×10^{-4}) in “never-smokers” (12,159 CHD cases and 22,932
386 controls) which was quantitatively similar to the CHD OR of 1.09 (P-value: 8.68×10^{-5}) observed in
387 “ever-smokers” (23,753 CHD cases and 24,019 controls) (interaction P-value: 0.76) (**eFigure 5a**).
388 Investigation in the PROMIS study of the *APOE* epsilon genotypes yielded consistent findings; the
389 OR for CHD among $\epsilon 4$ carriers in “never-smokers” was 1.13 compared to the CHD OR of 1.07
390 observed in “ever-smokers” (interaction P-value: 0.82) (**eFigure 5a**).

391

392
393 Novel gene-smoking interaction effects on CHD at chromosome 15q25.1
394 Of the 50 loci, we identified effect-modification by “ever-smoking” status on CHD risk for the
395 lead variants at two distinct loci, rs7178051, in proximity of *ADAMTS7* (an established CHD locus),
396 and rs1051730, in proximity of *CHRNA4-A3-A5* (an established smoking behavior locus) (**eTable 3**).
397 Although associated with different traits and located in distinct LD blocks, these two variants reside
398 ~224 KBs apart on chr.15q25.1 and are in weak linkage disequilibrium (LD) ($r^2 = 0.22$), raising the
399 question of whether these two chr.15q25.1 gene-smoking interactions on CHD are independent of
400 each other.

401 At the *ADAMTS7* CHD locus, the T allele at the rs7178051 variant was found to be more
402 strongly and inversely associated with CHD risk in never-smokers (OR: 0.88; P-value: 7.02×10^{-16})
403 compared to a much weaker effect in ever-smokers (OR: 0.95; P-value: 8.64×10^{-4}) (P-value of
404 interaction: 8.57×10^{-5}) (**Table 1**). Thus, the protective impact of the rs7178051 T allele observed in
405 never-smokers was halved in people who smoked (**Figure-1**). This difference is not related to power
406 differences within strata because for this variant, there were less data available in the never-smoking
407 group (21,232 CHD cases and 38,713 controls) compared to the ever-smoking group (39,585 CHD
408 cases and 40,749 controls). There was no substantial evidence of heterogeneity for the interaction
409 beta across the participating studies (Heterogeneity chi-squared = 36.23 (d.f. = 25); P-value for the
410 χ^2 test of heterogeneity = 0.06; $I^2 = 31.0\%$; tau-squared ($\tau^2 = 0$). We further conducted sensitivity
411 analyses using a random effect model; the results remained unchanged and the interaction beta
412 remained significant (**eFigure 5b**). Although the frequency of rs7178051 was 39% in Europeans
413 compared to 28% in South Asians, further analyses stratified by ancestry (i.e., European versus non-
414 Europeans) showed similar results (**eFigure 5c**). Other variants discovered through prior CHD
415 GWAS at the *ADAMTS7* locus (e.g., rs7173743, rs4380028, rs3825807) were in moderate to high
416 LD ($r^2 > 0.50$) with rs7178051 and were also found to display a similar pattern of gene-smoking
417 interaction effects (**Table 1**).

418 At the *CHRNA4-A3-A5* smoking locus, the A allele at the rs1051730 variant had an inverse
419 trend (not significant after adjustment) of association with CHD in never-smokers (OR: 0.96; P-value:
420 1.56×10^{-2}) and a positive trend (not significant after adjustment) in ever-smokers (OR: 1.03; P-value:
421 1.53×10^{-2}) (P-value of interaction: 2.37×10^{-4}) (**Table 1 and eTable 3**). For this variant, data on
422 20,559 CHD cases and 38,198 controls were available in the never-smoking group whereas 38,923
423 CHD cases and 40,371 controls were available in the ever-smoking group. Similar gene-smoking

424 interaction patterns were observed for other variants (e.g., rs2036527, rs8034191) that have been
425 previously reported for CPD behavior at the *CHRNA4-A3-A5* gene cluster (**Table 1**).

426 Further interrogation of the chr15q21.1 region encompassing rs7178051 and rs1051730
427 across three distinct LD blocks (**Figure 1**) revealed multiple additional variants for which we
428 observed gene-smoking interactions in CHD (**Table 1** and **Figure 1**). Indeed, several SNPs (e.g.,
429 rs7178051, rs10083696, rs7176187, rs6495335, rs4887077) had genome-wide significant
430 associations with CHD in “never-smokers” but had weaker and less significant associations with
431 CHD in “ever-smokers” (**Figure 1**). Alleles clustered specifically around *ADAMTS7* rather than at the
432 *CHRNA4-A3-A5* genes appear to be protective of CHD in “never-smokers” but have attenuated
433 protective effects in “ever-smokers” (**Figure 2**).

434 Conditional analyses

435 To investigate the possibility that the two chr.15q25.1 gene-smoking interactions might
436 represent a single interaction locus across the entire region we undertook an approximate
437 conditional and joint analyses¹⁷⁻²² using summary data derived from CARDIoGRAMplus4D for CHD
438 and from the TGC for smoking behavior. In-addition to rs7178051, we identified one other variant,
439 rs11072794 in low LD with rs7178051 ($r^2=0.20$) that was associated independently with CHD
440 (**Figure 3a**; red triangles) (**Figure 3b & eFigure 6b**; red triangles). We also confirmed two variants
441 (rs1051730 and rs684513) located in two different LD blocks that were independently associated
442 with smoking behavior in the TGC data¹⁰ (**Figure 3d & eFigure 6b**; grey circles).

443 In analyses of the CHD variants, both rs7178051 and rs11072794 remained strongly
444 associated with CHD after adjusting for the top CPD variants (rs1051730 and rs684513) (**Figure 3d**,
445 red triangles) whereas their weak association with CPD was lost after adjusting for the top CPD
446 variants (**Figure 3d**; grey circles); e.g., the P-value for rs7178051 association with CPD was 1×10^{-5}
447 in unadjusted analyses but attenuated to 0.55 after adjusting for rs1051730 and rs684513. In
448 analyses of the CPD variants, both rs1051730 and rs684513 remained strongly associated with CPD
449 after adjusting for the top CHD variants (rs7178051 and rs11072794) (**Figure 3b**, grey circles)
450 whereas their weak association with CHD was lost after adjusting for the top CHD variants (**Figure**
451 **3b**, red triangles). As expected, conditional analyses that included all four of these variants resulted
452 in a null association of the region with both CHD and CPD (**eFigure 6b**). To underscore the validity
453 of the conditional approach using summary data, we used individual participant data from an
454 expanded PROMIS sample involving 9,025 MI cases and 8,506 controls. We found that the OR
455 conferred by allelic variation at rs7178051 remained associated with MI risk independent of the two

456 CPD variants (rs1051730 and rs684513) and rs11072794 (the second CHD SNP) (**eFigure 6c**).
457 Conversely, the apparent effect of allelic variation at rs1051730 (the top CPD variant) on CHD risk
458 was lost when we adjusted for the other three variants, rs7178051, rs11072794 and rs684513
459 (**eFigure 6c**).

460 Next, using summary level data we examined the association of each of these four variants
461 with CHD risk separately in “ever-smokers” and “never-smokers” while mutually adjusting for the
462 other three variants (**Figure 4 & eFigure 7**). In these analyses, only allelic variation at rs7178051
463 was found to have independent genome-wide significant effects on CHD in never-smokers.
464 rs7178051 was also the only one of these four variants with significant differences in the effect
465 estimate for gene-CHD associations between the two smoking groups (P-value for the χ^2 test of
466 heterogeneity: 5.4×10^{-5}) after adjusting for the effects of other variants (rs11072794, rs1051730 and
467 rs684513). These conditional analyses suggest that (a) variants located near the *ADAMTS7* gene
468 but not *CHRNA4-A3-A5* genes have independent effects on CHD, (b) a single independent gene-
469 smoking interaction signal for CHD exists on chr.15q.25.1 which is localized at the *ADAMTS7* CHD
470 locus (marked by rs7178051) and (c) an apparent interaction signal observed at the nearby
471 *CHRNA4-A3-A5* CPD locus (marked by rs1051730) is not independent of the *ADAMTS7*
472 (rs7178051) interaction signal.

473 To assess the robustness of conditional analyses methodology that uses summary level data
474 (i.e., GCTA)¹⁷⁻²², we conducted sensitivity analyses in the PROMIS dataset (9,025 MI cases and
475 8,506 controls). We assessed the association of rs7178051 (top CHD SNP) and rs1051730 (top
476 CPD SNP) after mutually adjusting for each other by conducting (i) standard logistic regression using
477 individual participant data and (ii) summary level data in PROMIS using the GCTA method (**eTable**
478 **4**). The top CHD SNP was found associated with CHD risk in PROMIS independent of the top CPD
479 variant using both the methods, in-contrast the effect on CHD of the top CPD SNP attenuated
480 sharply when adjusted for the top CHD SNP – the effect estimates obtained using the two methods
481 were very similar (**eTable 4**).

482 Finally, to further demonstrate that the gene-smoking interaction effect in CHD at rs7178051 is
483 independent of the *CHRNA4-A3-A5* CPD locus, we conducted sensitivity analyses in the PROMIS
484 study by restricting our gene-environment interaction analysis to subjects who do not carry the minor
485 alleles of rs1051730 and rs684513 (the two SNPs associated with CPD) at the *CHRNA4-A3-A5*
486 locus. The T allele at the rs7178051 variant was associated with CHD only in never-smokers (OR:
487 0.88; P-value: 0.01) compared to a weaker and non-significant association in ever-smokers (OR:

488 0.94; P-value: 0.21) (**eTable 5**). The effect estimates obtained in each of the categories defined by
489 smoking status in PROMIS were similar to the effect estimates obtained in our overall meta-analyses
490 that utilized data in all participants (**eTable 5**).

491

492 Analysis of eQTLs and regulatory features at the chr15q25.1 gene-smoking interaction locus.

493 We mined publicly available eQTL data from the HapMap consortium,¹¹ GTEx consortium²³
494 and the MuTHER consortium²⁴ as well as data from 147 HAoEC lines²⁵ to examine the association
495 between mRNA expression of *ADAMTS7* and *CHRN* genes with CHD, CPD and gene-smoking
496 interaction SNPs at the chr15q25.1 locus. SNP-mRNA associations with p-values <0.002 (correction
497 for 20 tests) are presented (**Figure 5**). The top two CHD variants (rs7178051, rs11072794) are
498 associated with reduced *ADAMTS7* expression (e.g., rs11072794 $p=6.01 \times 10^{-21}$ in MuTHER LCL,
499 $n=850$; and rs7178051 $p=0.0029$ in HAoEC, $n=147$) but have no association with expression of
500 *CHRN* genes in any cell or tissue examined. In contrast, the top two CPD variants (rs1051730 and
501 rs684513) were associated with *CHRN* gene expression (e.g., rs1051730 $p=6.9 \times 10^{-7}$ for *CHRNA5* in
502 GTEx skeletal muscle and nerve tissue) but have no association with *ADAMTS7* in these cells or
503 tissues. These findings complement conditional analyses suggesting that gene-CHD and gene-
504 smoking interaction effects on CHD are likely mediated by *ADAMTS7* whereas the smoking behavior
505 effect appears to be mediated through the *CHRNA3-5* gene cluster.

506 In analysis of data from the ENCODE project²⁶ and for human aortic tissue in NIH
507 Roadmap Epigenomics project, *ADAMTS7* was associated with RNAseq reads and an active
508 transcription mark, H3K36me3, whereas *CHRN* genes had low/absent RNAseq reads and were
509 positive for repressive marks, H3K27me3 and H3K9me3 (**eFigure 8**). In HCASMC ChIPseq data,
510 rs7178051 the top CHD and gene-smoking CHD interacting SNP, is located in a region with active
511 regulatory marks H3K4me1 and H3K4me3 as well as transcription factor binding site for TCF21, an
512 important HCASMC transcription factor also associated with CAD. This ChIPseq pattern was
513 observed also in human aortic tissue (**Figure 6**). These regulatory data suggest active transcription
514 of *ADAMTS7*, but not *CHRN* genes, in vascular cells and aortic tissue and reveal that rs7178051,
515 the lead gene-smoking CHD interaction SNP, overlaps active transcription marks and transcription
516 factor binding regions in HCASMC.

517 *ADAMTS7* and *CHRNA3-5* expression in vascular cells and their response to cigarette smoke
518 extract

519 In order to explore which genes at the chr15q25.1 locus are expressed in CHD-relevant
520 vascular cells, we performed q-PCR of *ADAMTS7* and the *CHRNA4-A3-A5* genes in primary human
521 vascular cells and in the THP1 human monocyte cell line (**eFigure 9 & Figure 5**). Whilst *ADAMTS7*
522 mRNA was expressed abundantly in all vascular cell types, mRNA was below detection or
523 expressed at a very low level for each of the genes in the *CHRNA4-A3-A5* cluster in any of these cell
524 types (**eFigure 9**). Next, we explored the effect of cigarette smoke extract on gene expression in
525 HCASMC, a cell type of particular relevance to vascular responses to cigarette smoke products^{31, 32}
526 as well as to *ADAMTS7* vascular functions in atherosclerosis and CHD.³³ In primary HCASMC,
527 cigarette smoke extract exposure increased *ADAMTS7* mRNA levels by over 2-fold (**Figure 5**) but
528 did not affect expression of the *CHRN* genes (not shown). Thus, in contrast to *CHRN* genes,
529 *ADAMTS7* is both expressed and modulated by cigarette smoke extract in CHD-relevant vascular
530 cells providing biological support for *ADAMTS7*, but not *CHRN* genes, in the gene-smoking
531 interaction at chr15q25.1.

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546 DISCUSSION

547 We conducted a gene-environment interaction study at fifty loci associated with either CHD
548 or smoking behavior and found evidence of effect-modification of genotype-related CHD risk by
549 smoking-behavior at the chr.15q21.1 CHD locus. Specifically, we observed highly significant
550 attenuation of the cardio-protective effects conferred by alleles at this locus in people who smoked
551 cigarettes. Conditional analyses identified an LD block located at the *ADAMTS7* gene that
552 accounted for both the main effect on CHD as well as the gene-smoking interactions in CHD. Data
553 from expression and cell studies support our genetic analysis, suggesting that the underlying
554 mechanism relates to genotype differences in the effect of smoking on expression of *ADAMTS7* in
555 vascular tissue.

556 Our findings have novel mechanistic and clinical implications. These human genomic data
557 provide new insights into the mechanism of CHD in cigarette smokers. Identification of gene-
558 smoking interaction at the chr15q21.1 locus suggests a specific role in smoking-related CHD for
559 *ADAMTS7* and its substrates, vascular matrix and vascular smooth muscle cell biology more
560 broadly. Such insights can help to prioritize translational strategies for smoking-related CHD and
561 present opportunities to study lifestyle interventions and pharmacological strategies to lower CHD in
562 individuals who smoke cigarettes. Thus, inhibition of *ADAMTS7* represents a novel potential
563 therapeutic strategy for CHD that may have particular benefits in individuals who smoke cigarettes.
564 All smokers should receive counseling for smoking cessation yet such broad public health strategies
565 have failed to reach or impact smoking behavior in a large portion of nicotine-addicted individuals.
566 Our data provides a human genomic context for consideration of targeting specific genetically at-risk
567 individuals via intensified preventive strategies and development of novel pharmacological
568 treatments.

569 Our study also represents a realistic strategy for performing gene-environment interaction
570 studies using contemporary genetic data. We illustrate that identifying joint effects of genetic and
571 lifestyle factors in CHD requires very large sample sizes, yet such analyses are biologically
572 informative when studies are adequately powered. In this context, an important observation in our
573 large sample is the lack of effect modification by smoking behavior on CHD at the *APOE* locus, a
574 previously reported smoking interaction locus.¹²⁻¹⁴ This finding is consistent with a recent meta-
575 analysis that found no evidence of effect modification by smoking for *APOE* genotypes and CHD
576 risk.³⁴ These studies raise concerns that most prior gene-environment interactions studies in CHD
577 have been prone to the same biases (i.e., limited statistical power and false positive associations) as

578 candidate gene studies investigating main effects in the pre-GWAS era. The present study differs
579 from previous studies by being much larger and, importantly, it includes genomic and functional
580 follow-up data supporting the plausibility of the observed gene-environment interaction.

581 *ADAMTS7* (or the A disintegrin and metalloproteinase with thrombospondin motifs-7) is a
582 member of the ADAMTS family of secreted zinc metalloproteases.^{35, 36} We previously discovered
583 and replicated genetic variation at the *ADAMTS7* locus in association with coronary atherosclerosis
584 and MI.⁷⁻⁹ Both *in vivo* and *in vitro* studies suggest that *ADAMTS7* modulates VSMC phenotype
585 switching and migration and that this may be mediated via cartilage oligomeric matrix protein
586 (COMP) or thrombospondin-1 (TSP-1),^{32,33} i.e. putative *ADAMTS7* substrates expressed in vascular
587 tissue. Genetic variation at *ADAMTS7*, however, has no relationship with traditional risk factors or
588 mechanistic biomarkers; hence the directional impact of *ADAMTS7* expression on CHD risk and the
589 underlying biological mechanisms have been unclear.³²

590 Our gene-smoking interaction analyses provide novel insights into the directional impact of
591 the *ADAMTS7* locus on CHD, the underlying mechanisms of CHD in smokers, and how such
592 findings ultimately might translate towards achieving health benefits in society. Our human eQTL
593 interrogations reveal that common alleles that relate to lower CHD risk at the *ADAMTS7* locus are
594 also associated with reduced *ADAMTS7* expression, implying an atherogenic role of the gene. This
595 is supported by our recent *in vivo* experimental studies; *Adamts7* deficiency protected against diet-
596 induced atherosclerosis in both the *Ldlr*^{-/-} and *ApoE*^{-/-} mouse models, reduced neointima formation
597 following arterial injury, and decreased VSMC migration *in vitro*.³³ In our smoking-stratified analyses,
598 we observed CHD protective effect which was attenuated in smokers. Thus, smoking exposure may
599 overcome the genetic effect of protective alleles that act by reducing *ADAMTS7* expression. Such a
600 possibility is supported by our HCASMC data that reveals increased *ADAMTS7* expression in
601 HCASMC exposed to cigarette smoke extract. These human genome-smoking studies are the first to
602 implicate a specific locus as causal in mediating increased risk of CHD in smokers. Additional
603 translational studies are needed to establish the precise mechanisms of atheroprotection for alleles
604 at the *ADAMTS7* locus, how cigarette smoking impacts these genetic effects, and whether deletion
605 or inhibition of *ADAMTS7* *in vivo* attenuates the specific acceleration of atherosclerosis conferred by
606 cigarette smoking.

607 Strengths and limitations of our study merit consideration. This is a large study that
608 conducted gene-smoking interaction analyses in CHD by using GWAS data. We observed
609 substantial heterogeneity across study samples in our initial quality control analyses of “ever-

610 smoking” status with CHD risk. This is similar, however, to the heterogeneity reported in a recent
611 meta-analysis that pooled risk ratios from all the past prospective studies with information on
612 association of “ever-smoking” with incident CHD events.⁵ We recognize that other smoking related
613 phenotypes are important e.g., “current smoking” may have a more pronounced role than “ever-
614 smoking” in plaque rupture and thrombosis in patients with MI. “Current smoking” status and MI
615 phenotypes were available only in a subset of our studies limiting statistical power. Given the use of
616 multiple studies and meta-analyses of data, we used only one analytical approach to investigate
617 gene-smoking interactions. This approach, however, was feasible and powerful in this large-scale
618 consortium setting. While we used a fixed effects approach in our meta-analyses, a random effects
619 meta-analysis yielded qualitatively similar results (data not shown). The lack of replication is partially
620 offset by a large sample size, consistency across study cohorts and racial groups and supplemental
621 genomic and experimental evidence supporting biological plausibility. This approach is also
622 consistent with recent recommendations³⁷ which favor use of a powerful discovery experiment using
623 all data rather than reducing power by splitting available dataset for discovery and validation. While
624 our *in vitro* studies support a role for ADAMTS7 in the gene-smoking interaction, it will be important
625 to confirm that *Adamts7* deficiency protect against cigarette-smoke acceleration of atherosclerosis in
626 rodent models.

627 Our interaction analyses, conditional analyses, eQTL interrogations and cell studies
628 suggest that *ADAMTS7*, but not the *CHRNA4-A3-A5* gene cluster, is likely causal at 15q21.1 for
629 gene-smoking interaction effects in CHD. Yet, analyses are not definitive. Although top interacting
630 SNPs and CHD SNPs (e.g., rs7178051) were associated with *ADAMTS7*, but not *CHRNA4-A3-A5*,
631 expression in LCLs, large-scale eQTL data and allele specific expression data (e.g., via RNA
632 sequencing) are not available for vascular tissues limiting causal inference. In our small HCAEC
633 datasets, we did however find that alleles at rs7178051 associate with *ADAMTS7* expression but not
634 with any *CHRNA4-A3-A5* genes suggesting, at least in one vascular cell type, that the gene-smoking
635 interaction is mediated via *ADAMTS7*.

636

637

638

639

640

641 **Conclusions**

642 We provide novel evidence for allelic variation exhibiting gene-smoking interaction in CHD
643 at the chr.15q21.1 locus. The protective effect conferred by variation at this locus in never-smokers
644 is markedly attenuated in people who are ever-smokers. Stepwise conditional analyses, gene
645 expression data in vascular cells, eQTL interrogation, and cigarette smoke extract exposure in
646 HCASMC suggest that *ADAMTS7* accounts for both the gene-smoking interaction in CHD and the
647 CHD main effect on chr.15q21.1. Our findings reveal interactions of genetic variants and key lifestyle
648 determinants in the etiology of CHD, provide new insights into the potential mechanisms of CHD in
649 cigarette smokers, and facilitate precision medicine advances in cigarette-smoking related CHD. Our
650 work motivates future large-scale studies investigating joint effects of genes and environment in
651 CHD using existing complex-disease consortia datasets and genome-wide discovery approaches.
652 This will provide opportunities to detect additional and novel loci displaying gene-environment
653 interactions revealing genetic contexts for targeting intensive lifestyle interventions and novel
654 therapeutics.

655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673

Disclaimer: The manuscript and its content are confidential, intended for journal review only, and not to be further distributed.

674

675 **Sources of Funding**

676 Dr. Saleheen has received funding from the National Institutes of Health, the Fogarty International,
677 the Wellcome Trust, the British Heart Foundation, Pfizer, Genentech, Regeneron and Eli Lilly
678 pharmaceuticals. This work was supported in part by R01-HL-111694 and K24-HL-107643 from the
679 National Institutes of Health to Dr. Reilly.

680

681 **PROMIS.** Genotyping in PROMIS was funded by the Wellcome Trust, UK and Pfizer. Fieldwork in
682 the PROMIS study has been supported through funds available to investigators at the Center for
683 Non-Communicable Diseases, Pakistan and the University of Cambridge, UK.

684

685 **EPIC-CVD Consortium.** CHD case ascertainment and validation, genotyping, and clinical chemistry
686 assays in EPIC-CVD were principally supported by grants awarded to the University of Cambridge
687 from the EU Framework Programme 7 (HEALTH-F2-2012-279233), the UK Medical Research
688 Council (G0800270) and British Heart Foundation (SP/09/002), the UK National Institute for Health
689 Research Cambridge Biomedical Research Centre, and the European Research Council (268834).
690 Scientists at the EPIC-CVD Coordinating Centre have also been supported by grants from the US
691 National Institutes of Health, Merck, Novartis, GlaxoSmithKline, and Pfizer.

692

693 **WTCCC.** Recruitment of the WTCCC CAD cases was funded by the British Heart Foundation.
694 Collection of controls and genotyping was funded by the Wellcome Trust. Chris Nelson and Nilesh
695 Samani are funded by the British Heart Foundation. Nilesh Samani is a NIHR Senior Investigator.

696

697 **SCARF-SHEEP.** The investigators would like to acknowledge the Swedish Heart-Lung Foundation,
698 the Swedish Research Council, the Strategic Cardiovascular Programme of Karolinska Institutet and
699 the Stockholm County Council, the Strategic support for epidemiological research at Karolinska
700 Institutet and the Stockholm County Council. Rona J Strawbridge is supported by SRP Diabetes
701 Program at Karolinska Institutet.

702 **CARDIOGENICS and THESIAS.** Professor Deloukas' work forms part of the research themes
703 contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit
704 which is supported and funded by the National Institute for Health Research. Analysis was supported
705 by BHF grant (Deloukas) RG/14/5/30893. Professor Schunkert was supported by grants from the
706 Fondation Leducq (*CADgenomics: Understanding CAD Genes, 12CVD02*), the German Federal

707 Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding
708 concept (*e:AtheroSysMed*, grant 01ZX1313A-2014), and the European Union Seventh Framework
709 Programme FP7/2007-2013 under grant agreement n° HEALTH-F2-2013-601456 (*CVgenes-at-*
710 *target*). Further grants were received from the DFG as part of the Sonderforschungsbereich CRC
711 1123 (B2). Thorsten Kessler MD was supported by a DZHK Rotation Grant.

712 **Rotterdam Study.** The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus
713 University Rotterdam; the Netherlands Organization for Scientific Research; the Netherlands
714 Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in
715 the Elderly; the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and
716 Sports; the European Commission; and the Municipality of Rotterdam. Support for genotyping was
717 provided by the Netherlands Organisation of Scientific Research (NOW) Investments (No.
718 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015;
719 RIDE2), the Netherlands Genomics Initiative/Netherlands Consortium for Healthy Aging project No.
720 050-060-810. Abbas Dehghan is supported by an NWO grant (veni, 916.12.154) and the EUR
721 Fellowship. Oscar H. Franco works in ErasmusAGE, a center for aging research across the life
722 course funded by Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA. Nestlé Nutrition (Nestec
723 Ltd.), Metagenics Inc., and AXA had no role in design and conduct of the study; collection,
724 management, analysis, and interpretation of the data; and preparation, review, or approval of the
725 manuscript.

726
727 **Expression studies.** These studies were supported in part by a Transatlantic Network of Excellence
728 grant 10CVD03 from the Fondation Leducq. Analysis of expression quantitative trait loci (eQTL) in
729 endothelial cells was supported by a Transatlantic Networks of Excellence Award (12CVD02) from
730 Fondation Leducq (to Dr. Jake Lusis and team)

731 **Diabetes Heart Study.** This study was supported in part by R01 HL67348, R01 HL092301, and R01
732 NS058700 to Donald W. Bowden.

733 **Cleveland Clinic Study.** This study was supported in part by NIH grants R01ES021801,
734 3R01ES021801-03S1, and R01ES025786.

735 **Family Heart Study (FamHS).** The FamHS is funded by NHLBI grant R01HL117078 grant.

736 **MORGAM.** This work has been sustained by the MORGAM Project's funding: European Union FP 7
737 projects ENGAGE (HEALTH-F4-2007-201413), CHANCES (HEALTH-F3-2010-242244) and
738 BiomarcCaRE (278913). This funding has supported central coordination, workshops and part of the

739 activities of the MORGAM Data Centre, at THL in Helsinki, Finland. MORGAM Participating Centres
740 are funded by regional and national governments, research councils, charities, and other local
741 sources. The PRIME Study was supported by grants from Inserm, Merck Sharp and Dohme-Chibret
742 Laboratory, the French Research Agency and the Foundation Heart and Arteries. We also thank the
743 following organisations that allowed the recruitment of participants for the PRIME: the health
744 screening centres organised by the SocialSecurity of Lille (Institut Pasteur), Strasbourg, Toulouse,
745 and Tourcoing; the occupational medicine services of Haute-Garonne and of the Urban Community
746 of Strasbourg; the Association Inter-entreprises des Services Médicaux du Travail de Lille et
747 environs; the Comité pour le Développement de la Médecine du Travail; the Mutuelle Générale des
748 Postes, Télégraphes et Téléphones du Bas-Rhin; the Laboratoire d'Analyses de l'Institut de Chimie
749 Biologique de la Faculté de Médecine de Strasbourg. We also gratefully acknowledge the teams of
750 the Lille, Strasbourg and Toulouse centres for their dedicate work and relentness energy in following
751 up their cohorts; the contribution of the members of the event validation committees : L Guize; C
752 Morrison; M-T Guillanneuf; and M Giroud and the Alliance Partnership Programme for its financial
753 support. The KORA study was initiated and financed by the Helmholtz Zentrum München – German
754 Research Center for Environmental Health, which is funded by the German Federal Ministry of
755 Education and Research (BMBF) and by the State of Bavaria. KK was supported by the Orion-
756 Farnos Research Foundation and Academy of Finland (grant number 250207).

757 **DILGOM.** This work was enabled through a grant #139635 from the Academy of Finland and a grant
758 from the Finnish Foundation for Cardiovascular Research. The DILGOM project is also supported by
759 the Academy of Finland (grant numbers 136895 and 263836). We are grateful for the THL DNA
760 laboratory for its skillful work to produce the DNA samples used in this study. We also acknowledge
761 the Academy of Finland (136895 and 263836), Funding from Academy of Finland, grant number:
762 118065, and Juho Vainio Foundation.

763

764

765

766

767

768

769

770

771 **Acknowledgements**

772 We would like to thank the CARDIoGRAMplusC4D consortium, the EPIC-CVD and the PROMIS
773 study for contributing data.

774 **PROMIS.** We also acknowledge the contributions made by the following: Zeeshan Ozair, Usman
775 Ahmed, Abdul Hakeem, Hamza Khalid, Naeem Khan, Sadiq Khan, Ayaz Ali, Madad Ali, Saeed
776 Ahmed, Muhammad Waqar Khan, Muhammad Razaq Khan, Abdul Ghafoor, Mir Alam, Riazuddin,
777 Muhammad Irshad Javed, Abdul Ghaffar, Tanveer Baig Mirza, Muhammad Shahid, Jabir Furqan,
778 Muhammad Iqbal Abbasi, Tanveer Abbas, Rana Zulfiqar, Muhammad Wajid, Irfan Ali, Muhammad
779 Ikhtlaq, Danish Sheikh, Muhammad Imran, Nadeem Sarwar, Adam Butterworth, Matthew Walker and
780 Hannah Lombardi, Shahid Abbas, Faisal Majeed, Saba Akhtar, Nadeem Qamar, Khan Shah Zaman,
781 Zia Yaqoob, Tahir Saghir, Syed Nadeem Hasan Rizvi, Anis Memon, Nadeem Hayyat Mallick,
782 Mohammad Ishaq, Syed Zahed Rasheed, Fazal-ur-Rehman Memon, Khalid Mahmood, and
783 Naveeduddin Ahmed.

784 **EPIC-CVD.** We thank all EPIC participants and staff for their contribution to the study, the laboratory
785 teams at the Medical Research Council Epidemiology Unit for sample management and Cambridge
786 Genomic Services for genotyping, Sarah Spackman for data management, and the team at the
787 EPIC-CVD Coordinating Centre for study coordination and administration.

788 **Disclosures**

789 Dr. Saleheen has received funding from Pfizer, Genentech, Regeneron and Eli Lilly pharmaceuticals.

790

791

792

793

794

795

796

797

798

799

800

801

802 References

- 803 (1) Mangino M, Spector T. Understanding coronary artery disease using twin studies. *Heart*.
804 2013;99:373-5.
- 805 (2) Lüscher TF, Libby P. Inflammation in atherosclerosis. *Nature*. 2011;479:53-61.
806 2004;5:189-218.
- 807 (3) Teo KK, Ounpuu S, Hawken S, Pandey MR, Valentin V, Hunt D, Diaz R, Rashed W, Freeman R,
808 Jiang L, Zhang X, Yusuf S; INTERHEART Study Investigators. Tobacco use and risk of myocardial
809 infarction in 52 countries in the INTERHEART study: a case-control study. *Lancet*. 2006;368:647-58.
- 810 (4) Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations
811 on male British doctors. *BMJ*. 2004;328:1519.
- 812 (5) Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women
813 compared with men: a systematic review and meta-analysis of prospective cohort studies. *Lancet*.
814 2011;378:1297-305.
- 815 (6) Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet*.
816 2003;362:847-52.
- 817 (7) CARDIoGRAMplusC4D Consortium., Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes
818 TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups K, König IR, Cazier
819 JB, Johansson A, Hall AS, Lee JY, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A,
820 Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K,
821 Lundmark P, Lyytikäinen LP, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van
822 Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I,
823 Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R; DIAGRAM Consortium.;
824 CARDIOGENICS Consortium., Doney AS, El Mokhtari N, Eriksson P, Fischer K, Fontanillas P,
825 Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han BG, Hunt SE,
826 Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C,
827 Leander K, Lokki ML, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S,
828 Morris AD, Müller-Nurasyid M; MuTHER Consortium., Nikus K, Peden JF, Rayner NW, Rasheed A,
829 Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AF, Tan ST,
830 Thorgeirsson G, van der Schoot CE, Wagner PJ; Wellcome Trust Case Control Consortium., Wells
831 GA, Wild PS, Yang TP, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P,
832 Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM,
833 Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason V, Hazen SL, Holm H, Iribarren C, Jang Y,
834 Kähönen M, Kee F, Kim HS, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R,
835 Lee JY, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader
836 DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Trégouët DA, Virtamo J,
837 Wallentin L, Wareham N, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen
838 T, Syvänen AC, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA,
839 Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell C,
840 Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J,
841 Palmer CN, Roberts R, Watkins H, Schunkert H, Samani NJ. Large-scale association analysis
842 identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45:25-33.

843 (8) Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in
844 Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.*
845 2011;43:339-44.

846 (9) Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF,
847 Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K,
848 Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis
849 JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I; Cardiogenics.,
850 Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S,
851 Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC,
852 Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S,
853 Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T,
854 Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW,
855 Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery
856 AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O,
857 Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt
858 J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi
859 AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS,
860 Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS,
861 Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K,
862 Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M,
863 Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg
864 C, Wittteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA,
865 Quertermous T, März W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P,
866 Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann
867 J; CARDIoGRAM Consortium., Samani NJ. Large-scale association analysis identifies 13 new
868 susceptibility loci for coronary artery disease. *Nat Genet* 2011;43:333-8.

869 (10) Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci
870 associated with smoking behavior. *Nat Genet.* 2010;42:441-7.

871 (11) International HapMap 3 Consortium., Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs
872 RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE,
873 Altshuler DM, Gibbs RA, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X,
874 Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA,
875 Muzny DM, Barnes C, Darvishi K, Hurler M, Korn JM, Kristiansson K, Lee C, McCarroll SA, Nemesh
876 J, Dermitzakis E, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Bonnen PE, Gibbs
877 RA, Gonzaga-Jauregui C, Keinan A, Price AL, Yu F, Anttila V, Brodeur W, Daly MJ, Leslie S,
878 McVean G, Moutsianas L, Nguyen H, Schaffner SF, Zhang Q, Ghorji MJ, McGinnis R, McLaren W,
879 Pollack S, Price AL, Schaffner SF, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti
880 PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare
881 D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE.
882 Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010;467:52-8.

883 (12) Humphries SE, Talmud PJ, Hawe E, Bolla M, Day IN, Miller GJ. Apolipoprotein E4 and coronary
884 heart disease in middle-aged men who smoke: a prospective study. *Lancet.* 2001;358:115-9.

885 (13) Grammer TB, Hoffmann MM, Scharnagl H, Kleber ME, Silbernagel G, Pilz S, Tomaschitz A,
886 Lerchbaum E, Siekmeier R, März W. Smoking, apolipoprotein E genotypes, and mortality (the
887 Ludwigshafen Risk and Cardiovascular Health study). *Eur Heart J.* 2013;34:1298-305.

888 (14) Gustavsson J, Mehlig K, Leander K, Strandhagen E, Björck L, Thelle DS, Lissner L, Blennow K,
889 Zetterberg H, Nyberg F. Interaction of apolipoprotein E genotype with smoking and physical inactivity
890 on coronary heart disease risk in men and women. *Atherosclerosis*. 2012;220:486-92.

891 (15) Saleheen D, Zaidi M, Rasheed A, Ahmad U, Hakeem A, Murtaza M, Kayani W, Faruqi A,
892 Kundi A, Zaman KS, Yaqoob Z, Cheema LA, Samad A, Rasheed SZ, Mallick NH, Azhar M, Jooma
893 R, Gardezi AR, Memon N, Ghaffar A, Fazal-ur-Rehman, Khan N, Shah N, Ali Shah A, Samuel M,
894 Hanif F, Yameen M, Naz S, Sultana A, Nazir A, Raza S, Shazad M, Nasim S, Javed MA, Ali SS,
895 Jafree M, Nisar MI, Daood MS, Hussain A, Sarwar N, Kamal A, Deloukas P, Ishaq M, Frossard P,
896 Danesh J. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic,
897 lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol*.
898 2009;24:329-38.

899 (16) Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP,
900 Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M,
901 Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M,
902 Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee JY, Park T, Kim K,
903 Sim X, Twee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X,
904 Luan J, Lamina C, Ziegler A, Zhang W, Zee RY, Wright AF, Witteman JC, Wilson JF, Willemsen G,
905 Wichmann HE, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF,
906 Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM,
907 Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJ, Scuteri A, Scott J,
908 Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruukonen A, Rudan I, Rose LM,
909 Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BW, Pedersen NL,
910 Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery
911 GW, Meitinger T, McPherson R, McCarthy MI, McArdle W, Masson D, Martin NG, Marroni F,
912 Mangino M, Magnusson PK, Lucas G, Luben R, Loos RJ, Lokki ML, Lettre G, Langenberg C, Launer
913 LJ, Lakatta EG, Laaksonen R, Kyvik KO, Kronenberg F, König IR, Khaw KT, Kaprio J, Kaplan LM,
914 Johansson A, Jarvelin MR, Janssens AC, Ingelsson E, Igl W, Kees Hovingh G, Hottenga JJ, Hofman
915 A, Hicks AA, Hengstenberg C, Heid IM, Hayward C, Havulinna AS, Hastie ND, Harris TB,
916 Haritunians T, Hall AS, Gyllenstein U, Guiducci C, Groop LC, Gonzalez E, Gieger C, Freimer NB,
917 Ferrucci L, Erdmann J, Elliott P, Ejebe KG, Döring A, Dominiczak AF, Demissie S, Deloukas P, de
918 Geus EJ, de Faire U, Crawford G, Collins FS, Chen YD, Caulfield MJ, Campbell H, Burt NP,
919 Bonnycastle LL, Boomsma DI, Boekholdt SM, Bergman RN, Barroso I, Bandinelli S, Ballantyne CM,
920 Assimes TL, Quertermous T, Altshuler D, Seielstad M, Wong TY, Tai ES, Feranil AB, Kuzawa CW,
921 Adair LS, Taylor HA Jr, Borecki IB, Gabriel SB, Wilson JG, Holm H, Thorsteinsdottir U, Gudnason V,
922 Krauss RM, Mohlke KL, Ordovas JM, Munroe PB, Kooner JS, Tall AR, Hegele RA, Kastelein JJ,
923 Schadt EE, Rotter JI, Boerwinkle E, Strachan DP, Mooser V, Stefansson K, Reilly MP, Samani NJ,
924 Schunkert H, Cupples LA, Sandhu MS, Ridker PM, Rader DJ, van Duijn CM, Peltonen L, Abecasis
925 GR, Boehnke M, Kathiresan S. Biological, clinical and population relevance of 95 loci for blood lipids.
926 *Nature* 2010;466:707-13..

927 (17) Yang J, Ferreira T, Morris AP, Medland SE; Genetic Investigation of ANthropometric Traits
928 (GIANT) Consortium.; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium.,
929 Madden PA, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy
930 MI, Hirschhorn JN, Goddard ME, Visscher PM. Conditional and joint multiple-SNP analysis of GWAS
931 summary statistics identifies additional variants influencing complex traits. *Nat Genet*. 2012;44:369-
932 3.

933 (18) Yang J, Lee SH, Goddard ME, Visscher PM. Genome-wide complex trait analysis (GCTA): methods,
934 data analyses, and interpretations. *Methods Mol Biol*. 2013; 1019:215-36.

- 935 (19) Keller MF, Ferrucci L, Singleton AB. Genome-Wide Analysis of the Heritability of Amyotrophic Lateral
936 Sclerosis. *JAMA Neurol.* 2014;71:1123-34.
- 937 (20) van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, van der Spek RA,
938 Vösa U, de Jong S, Robinson MR, Yang J, Fogh I, van Doormaal PT, Tazelaar GH, Koppers M, Blokhuis
939 AM, Sproviero W, Jones AR, Kenna KP, van Eijk KR, Harschnitz O, Schellevis RD, Brands WJ, Medic J,
940 Menelaou A, Vajda A, Ticozzi N, Lin K, Rogelj B, Vrabc K, Ravnik-Glavač M, Koritnik B, Zidar J,
941 Leonardis L, Grošelj LD, Millecamps S, Salachas F, Meininger V, de Carvalho M, Pinto S, Mora JS,
942 Rojas-García R, Polak M, Chandran S, Colville S, Swingler R, Morrison KE, Shaw PJ, Hardy J, Orrell
943 RW, Pittman A, Sidle K, Fratta P, Malaspina A, Topp S, Petri S, Abdulla S, Drepper C, Sendtner M,
944 Meyer T, Ophoff RA, Staats KA, Wiedau-Pazos M, Lomen-Hoerth C, Van Deerlin VM, Trojanowski JQ,
945 Elman L, McCluskey L, Basak AN, Tunca C, Hamzeiy H, Parman Y, Meitinger T, Lichtner P, Radivojkov-
946 Blagojevic M, Andres CR, Maurel C, Bensimon G, Landwehrmeyer B, Brice A, Payan CA, Saker-Delye S,
947 Dürr A, Wood NW, Tittmann L, Lieb W, Franke A, Rietschel M, Cichon S, Nöthen MM, Amouyel P,
948 Tzourio C, Dartigues JF, Uitterlinden AG, Rivadeneira F, Estrada K, Hofman A, Curtis C, Blauw HM, van
949 der Kooij AJ, de Visser M, Goris A, Weber M, Shaw CE, Smith BN, Pansarasa O, Cereda C, Del Bo R,
950 Comi GP, D'Alfonso S, Bertolin C, Sorarù G, Mazzini L, Pensato V, Gellera C, Tiloca C, Ratti A, Calvo A,
951 Moglia C, Brunetti M, Arcuti S, Capozzo R, Zecca C, Lunetta C, Penco S, Riva N, Padovani A, Filosto M,
952 Muller B, Stuit RJ; PARALS Registry.; SLALOM Group.; SLAP Registry.; FALS Sequencing Consortium.;
953 SLAGEN Consortium.; NNIPPS Study Group., Blair I, Zhang K, McCann EP, Fifita JA, Nicholson GA,
954 Rowe DB, Pamphlett R, Kiernan MC, Grosskreutz J, Witte OW, Ringer T, Prell T, Stubendorff B, Kurth I,
955 Hübner CA, Leigh PN, Casale F, Chio A, Beghi E, Pupillo E, Tortelli R, Logroscino G, Powell J, Ludolph
956 AC, Weishaupt JH, Robberecht W, Van Damme P, Franke L, Pers TH, Brown RH, Glass JD, Landers JE,
957 Hardiman O, Andersen PM, Corcia P, Vourc'h P, Silani V, Wray NR, Visscher PM, de Bakker PI, van Es
958 MA, Pasterkamp RJ, Lewis CM, Breen G, Al-Chalabi A, van den Berg LH, Veldink JH. Genome-wide
959 association analyses identify new risk variants and the genetic architecture of amyotrophic lateral
960 sclerosis. *Nat Genet.* 2016;48:1043-8.
- 961 (21) Torres JM, Gamazon ER, Parra EJ, Below JE, Valladares-Salgado A, Wachter N, Cruz M, Hanis CL,
962 Cox NJ. Cross-tissue and tissue-specific eQTLs: partitioning the heritability of a complex trait. *Am J Hum*
963 *Genet.* 2014;95:521-34.
- 964 (22) Bailey JN, Loomis SJ, Kang JH, Allingham RR, Gharahkhani P, Khor CC, Burdon KP, Aschard H,
965 Chasman DI, Igo RP Jr, Hysi PG, Glastonbury CA, Ashley-Koch A, Brilliant M, Brown AA, Budenz DL,
966 Buil A, Cheng CY, Choi H, Christen WG, Curhan G, De Vivo I, Fingert JH, Foster PJ, Fuchs C,
967 Gaasterland D, Gaasterland T, Hewitt AW, Hu F, Hunter DJ, Khawaja AP, Lee RK, Li Z, Lichter PR,
968 Mackey DA, McGuffin P, Mitchell P, Moroi SE, Perera SA, Pepper KW, Qi Q, Realini T, Richards JE,
969 Ridker PM, Rimm E, Ritch R, Ritchie M, Schuman JS, Scott WK, Singh K, Sit AJ, Song YE, Tamimi RM,
970 Topouzis F, Viswanathan AC, Verma SS, Vollrath D, Wang JJ, Weisschuh N, Wissinger B, Wollstein G,
971 Wong TY, Yaspan BL, Zack DJ, Zhang K, Study EN; ANZRAG Consortium., Weinreb RN, Pericak-Vance
972 MA, Small K, Hammond CJ, Aung T, Liu Y, Vithana EN, MacGregor S, Craig JE, Kraft P, Howell G,
973 Hauser MA, Pasquale LR, Haines JL, Wiggs JL. Genome-wide association analysis identifies TXNRD2,
974 ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. *Nat Genet.* 2016;48:189-94
- 975 (23) GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013;45:580-5.
- 976 (24) Grundberg E, Small KS, Hedman ÅK, Nica AC, Buil A, Keildson S, Bell JT, Yang TP, Meduri E,
977 Barrett A, Nisbett J, Sekowska M, Wilk A, Shin SY, Glass D, Travers M, Min JL, Ring S, Ho K,
978 Thorleifsson G, Kong A, Thorsteindottir U, Ainali C, Dimas AS, Hassanali N, Ingle C, Knowles D,
979 Krestyaninova M, Lowe CE, Di Meglio P, Montgomery SB, Parts L, Potter S, Surdulescu G,
980 Tzaprouni L, Tsoka S, Bataille V, Durbin R, Nestle FO, O'Rahilly S, Soranzo N, Lindgren CM,
981 Zondervan KT, Ahmadi KR, Schadt EE, Stefansson K, Smith GD, McCarthy MI, Deloukas P,
982 Dermitzakis ET, Spector TD; Multiple Tissue Human Expression Resource (MuTHER) Consortium..

- 983 Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet.* 2012;44:1084-
984 9.
- 985 (25) Erbilgin A, Civelek M, Romanoski CE, Pan C, Hagopian R, Berliner JA, Lusis AJ. Identification
986 of CAD candidate genes in GWAS loci and their expression in vascular cells. *J Lipid Res.*
987 2013;54:1894-905.
- 988 (26) ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE).
989 *PLoS Biol.* 2011;9:e1001046.
- 990 (27) Miller CL, Anderson DR, Kundu RK, Raiesdana A, Nürnberg ST, Diaz R, Cheng K, Leeper NJ,
991 Chen CH, Chang IS, Schadt EE, Hsiung CA, Assimes TL, Quertermous T. Disease-related growth
992 factor and embryonic signaling pathways modulate an enhancer of TCF21 expression at the 6q23.2
993 coronary heart disease locus. *PLoS Genet.* 2013;9:e1003652.
- 994 (28) Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras
995 TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29:15-21.
- 996 (29) Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, Frossard P, Been LF, Chia KS, Dimas AS,
997 Hassanali N, Jafar T, Jowett JB, Li X, Radha V, Rees SD, Takeuchi F, Young R, Aung T, Basit A,
998 Chidambaram M, Das D, Grundberg E, Hedman AK, Hydrie ZI, Islam M, Khor CC, Kowlessur S,
999 Kristensen MM, Liju S, Lim WY, Matthews DR, Liu J, Morris AP, Nica AC, Pinidiyapathirage JM,
1000 Prokopenko I, Rasheed A, Samuel M, Shah N, Shera AS, Small KS, Suo C, Wickremasinghe AR,
1001 Wong TY, Yang M, Zhang F; DIAGRAM; MuTHER, Abecasis GR, Barnett AH, Caulfield M, Deloukas
1002 P, Frayling TM, Froguel P, Kato N, Katulanda P, Kelly MA, Liang J, Mohan V, Sanghera DK, Scott J,
1003 Seielstad M, Zimmet PZ, Elliott P, Teo YY, McCarthy MI, Danesh J, Tai ES, Chambers JC. Genome-
1004 wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes
1005 susceptibility loci. *Nat Genet.* 2011;43:984-9.
- 1006 (30) Saade S, Cazier JB, Ghassibe-Sabbagh M, Youhanna S, Badro DA, Kamatani Y, Hager J,
1007 Yeretzian JS, El-Khazen G, Haber M, Salloum AK, Douaihy B, Othman R, Shasha N, Kabbani S,
1008 Bayeh HE, Chammas E, Farrall M, Gauguier D, Platt DE, Zalloua PA; FGENTCARD consortium.
1009 Large scale association analysis identifies three susceptibility loci for coronary artery disease. *PLoS*
1010 *One.* 2011;6:e29427.
- 1011 (31) Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, Viñuela A, Grundberg E, Nelson
1012 CP, Meduri E, Buil A, Cambien F, Hengstenberg C, Erdmann J, Schunkert H, Goodall AH,
1013 Ouwehand WH, Dermitzakis E, Spector TD, Samani NJ, Deloukas P. Cigarette smoking reduces
1014 DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation.
1015 *Epigenetics.* 2014;9:1382-96.
- 1016 (32) Starke RM, Ali MS, Jabbour PM, Tjoumakaris SI, Gonzalez F, Hasan DM, Rosenwasser RH,
1017 Owens GK, Koch WJ, Dumont AS. Cigarette smoke modulates vascular smooth muscle phenotype:
1018 implications for carotid and cerebrovascular disease. *PLoS One.* 2013;8:e71954.
- 1019 (33) Bauer RC, Tohyama J, Cui J, Cheng L, Yang J, Zhang X, Ou K, Paschos GK, Zheng XL,
1020 Parmacek MS, Rader DJ, Reilly MP. Knockout of *Adamts7*, a novel coronary artery disease locus in
1021 humans, reduces atherosclerosis in mice. *Circulation.* 2015;131:1202-13.
- 1022 (34) Holmes MV, Frikke-Schmidt R, Melis D, Luben R, Asselbergs FW, Boer JM, Cooper J, Palmén
1023 J, Horvat P, Engmann J, Li KW, Onland-Moret NC, Hofker MH, Kumari M, Keating BJ, Hubacek JA,

1024 Adamkova V, Kubinova R, Bobak M, Khaw KT, Nordestgaard BG, Wareham N, Humphries SE,
1025 Langenberg C, Tybjaerg-Hansen A, Talmud PJ. A systematic review and meta-analysis of 130,000
1026 individuals shows smoking does not modify the association of APOE genotype on risk of coronary
1027 heart disease. *Atherosclerosis*. 2014;237:5-12.

1028 (35) Pu X, Xiao Q, Kiechl S, Chan K, Ng FL, Gor S, Poston RN, Fang C, Patel A, Senver EC, Shaw-
1029 Hawkins S, Willeit J, Liu C, Zhu J, Tucker AT, Xu Q, Caulfield MJ, Ye S. ADAMTS7 cleavage and
1030 vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant.
1031 *Am J Hum Genet*. 2013;92:366-74.

1032 (36) Wang L, Zheng J, Bai X, Liu B, Liu CJ, Xu Q, Zhu Y, Wang N, Kong W, Wang X. ADAMTS-7
1033 mediates vascular smooth muscle cell migration and neointima formation in balloon-injured rat
1034 arteries. *Circ Res*. 2009;104:688-98.

1035
1036 (37) Hutter CM, Mechanic LE, Chatterjee N, Kraft P, Gillanders EM; NCI Gene-Environment Think
1037 Tank. Gene-environment interactions in cancer epidemiology: a National Cancer Institute Think Tank
1038 report. *Genet Epidemiol*. 2013;37:643-57.

1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072

Disclaimer: The manuscript and its contents are confidential, intended for journal review purposes only, and not to be further disclosed.

1073
1074
1075 **Figure Legends**
1076

1077 **Figure 1.** (a) Regional association analyses at the chromosome 15q25.1 locus in association with
1078 CHD risk stratified by smoking status. Association P-values for genetic variants with CHD risk in
1079 “never-smokers” (green squares) and “ever-smokers” (red triangles). (b) Longitudinal bars represent
1080 gene-smoking CHD interaction P-values at the chromosome 15q25.1 locus; bars in blue are P-
1081 values for variants listed in Table-1 and each variant has been assigned a unique identification
1082 number based on its physical location; (c) LD-blocks at the 15q25.1 locus visualized through
1083 HAPLOVIEW using LD estimates in the HapMAP-2 CEU reference population.
1084

1085 **Figure 2.** Several variants at chromosome 15q21.1 have stronger effects on CHD risk in “never-
1086 smokers” compared to “ever-smokers”. Variants with the strongest interaction P-value are displayed.
1087

1088 **Figure 3.** Step-wise conditional analysis of genetic variation at the chromosome 15q21.1 locus with
1089 CHD (red triangles) and smoking behavior (cigarettes per day, CPD; grey circles). At the
1090 chromosome 15q21.1 locus, analyses adjusted for rs7178051 and rs11072794 completely
1091 attenuated the gene-CHD associations whereas gene-smoking remained unchanged. Analyses
1092 adjusted for rs1051730 and rs684513 completely attenuated the gene-smoking associations
1093 whereas gene-CHD effect remained unchanged.

1094 **Figure 4.** Analyses mutually adjusted for rs7178051, rs11072794, rs1051730 and rs684513 at
1095 15q21.1 on CHD and smoking behavior; gene-CHD interaction analyses were only found significant
1096 for rs7178051.

1097 **Figure 5.** Genome browser view of regulatory features at rs7178051 on Chr15q21.1. ChIP-seq
1098 experiments were performed on confluent HCASMC for TCF21, Jun, JunD, CEBP and H3K4me1,
1099 H3K27me3, H3K27ac. DNaseI hypersensitivity data for human AoSMC were acquired from the
1100 ENCODE project. Human aortic tissue H3K4me1, H3K9me3, H3K27me3, and H3K36me3 ChIP-seq
1101 data were acquired from the NIH Roadmap Epigenomics Project. HCASMC = human coronary
1102 artery smooth muscle cells; AoSMC = human aortic smooth muscle cells.

1103 **Figure 6.** (a) *ADAMTS7* and *CHRNA3-3-A5* mRNA levels were measured in HCASMC. Cells were
1104 cultured to confluence, total RNA was extracted and cDNA generated. q-PCR was performed for
1105 *ACTB*, *GAPDH*, *TBP*, *ADAMTS7*, *CHRNA3*, *CHRNA5* (95°C 15s, 60°C 1min). Delta Cts

1106 were calculated as follows: $(Ct_{ACTB} + Ct_{GAPDH} + Ct_{TBP})/3 - Ct_{TARGET\ GENE}$. Fold changes are derived
1107 from delta delta Cts based on formula $FC = 2^{-\Delta\Delta Ct}$. (b) Confluent HCASMC were exposed to cigarette
1108 smoke extract. Serum starved (x24 hrs.) confluent HCASMC were treated with 0.5% or 1.0%
1109 cigarette smoke extract (v/v) for 4, 12, and 24 hrs. in serum reduced conditions (0.5% FBS in
1110 DMEM). Total RNA was extracted, cDNA generated preparation and q-PCR performed for
1111 *ADAMTS7* by Taqman and normalized to *GAPDH*. The Average Ct for *ADAMTS7* at baseline was
1112 28.25. Results were presented as means \pm SEM, and data were analyzed using Student's t-Test. (c)
1113 expression and eQTL Data from the GTEx consortium, the HapMap consortium (restricted to
1114 European populations), the Multiple Tissue Human Expression Resource (MuTHER) and in 147
1115 donor HAoEC lines. Association of the independent lead variants identified in our conditional
1116 analyses with expression of *ADAMTS7* and genes in the *CHRNA4-A3-A5* cluster. A P-value
1117 threshold of 0.002 was set to account for multiple testing involved in the eQTL analyses.

1118

1119

1120

1121

Disclaimer: The manuscript and its content are confidential, intended for journal review only, and not to be further disclosed.

Table-1. Novel genotype-smoking interaction findings in coronary heart disease at the chromosome 15q25.1 locus

Variant	Association	allele	LD with rs7178051*	LD with rs1051730^	Never Smokers					Ever Smokers					
					N cases	N controls	N Total	Beta (SE)	P-value	N cases	N controls	N Total	Beta (SE)	P-value	P-value interaction
*rs7178051 ⁴	CHD (NPR)	T/C	-	0.22	21232	38713	59945	-0.13 (0.01)	1.30E-16	39585	40749	80334	-0.05 (.01)	2.49E-04	8.57E-05
^rs1051730 ¹⁵	SB (known)	A/G	0.22	-	20559	38198	58757	-0.04 (0.02)	0.02	38923	40371	79294	0.03 (0.01)	0.02	2.37E-04
Other variants on chr.15q25.1 with significant gene-smoking interactions on CHD															
rs7173743 ¹	CHD (Known)	C/T	0.61	0.18	21050	37955	59005	-0.11 (0.01)	2.73E-13	39044	39559	78603	-0.04 (0.01)	8.60E-04	9.29E-05
rs10083696 ²	CHD (Novel)	A/G	1.0	0.22	19721	36206	55927	-0.11 (0.02)	1.60E-12	38807	40018	78825	-0.05 (0.01)	2.72E-04	5.15E-05
rs7176187 ³	CHD (Novel)	T/C	1.0	0.24	21232	38713	59945	-0.12 (0.01)	7.02E-16	39585	40749	80334	-0.04 (0.01)	8.64E-04	6.93E-05
rs6495335 ⁵	CHD (Novel)	G/T	1.0	0.22	20144	37217	57361	-0.13 (0.02)	2.39E-15	36448	38203	74651	-0.04 (0.01)	1.69E-03	9.51E-04
rs4380028 ⁶	CHD (Known)	T/C	1	0.22	21232	38713	59945	-0.12 (0.01)	2.20E-15	39585	40749	80334	-0.04 (.01)	1.03E-03	5.44E-04
rs3825807 ⁷	CHD (Known)	G/A	0.52	0.43	17137	28633	45771	-0.09 (0.02)	2.82E-08	30071	29014	59086	-0.03 (0.01)	0.04	2.6E-03
rs3813565 ⁸	CHD (NPR)	T/G	0.43	0.56	19466	35830	55296	-0.08 (0.02)	5.08E-07	36642	37759	74401	-0.01 (0.01)	0.42	3.05E-04
rs11638490 ⁹	CHD (NPR)	T/C	0.44	0.51	20465	37897	58362	-0.08 (0.01)	6.90E-08	38533	39690	78223	-0.01 (0.01)	0.28	2.25E-04
rs11072791 ¹¹	CHD (NPR)	A/C	0.44	0.51	19289	35944	55233	-0.08 (0.02)	2.83E-07	35245	36635	71880	-.005 (0.01)	0.68	1.06E-04
rs922692 ¹²	CHD (NPR)	A/C	0.44	0.50	20559	38198	58757	-0.08 (0.01)	2.81E-07	38923	40371	79294	-0.01 (0.01)	0.29	2.75E-04
rs11638372 ¹³	CHD (NPR)	T/C	0.44	0.50	21232	38713	59945	-0.08 (0.01)	6.92E-08	39585	40749	80334	-0.01 (0.01)	0.23	3.16E-04
rs4887077 ¹⁴	CHD (NPR)	T/C	0.44	0.50	21232	38713	59945	-0.08 (0.01)	4.71E-08	39585	40749	80334	-0.02 (0.01)	0.20	3.92E-05
rs12899135 ¹⁵	CHD (NPR)	G/A	0.39	0.56	20377	37440	57817	-0.07 (0.02)	3.97E-06	38382	39181	77563	-0.01 (0.01)	0.58	4.54E-04
rs684513 ¹⁸	SB (Known)	C/G	0.01	0.10	12517	21054	33572	-0.01 (0.02)	0.67	24641	24487	49129	0.03 (0.02)	0.18	0.08
rs2036527 ¹⁹	SB (Known)	A/G	0.17	0.90	20559	38198	58757	-0.04 (0.02)	0.02	38923	40371	79294	0.03 (0.01)	0.02	2.14E-04
rs10519203 ²⁰	CHD (NPR)	G/A	0.19	0.93	21232	38713	59945	-0.04 (0.01)	5.93E-03	39585	40749	80334	0.03 (0.01)	0.03	1.27E-04
rs8034191 ²¹	SB (Known)	C/T	0.19	1.0	19251	32131	51382	-0.05 (0.02)	2.62E-03	34925	34047	68972	0.02 (0.01)	0.06	3.91E-05

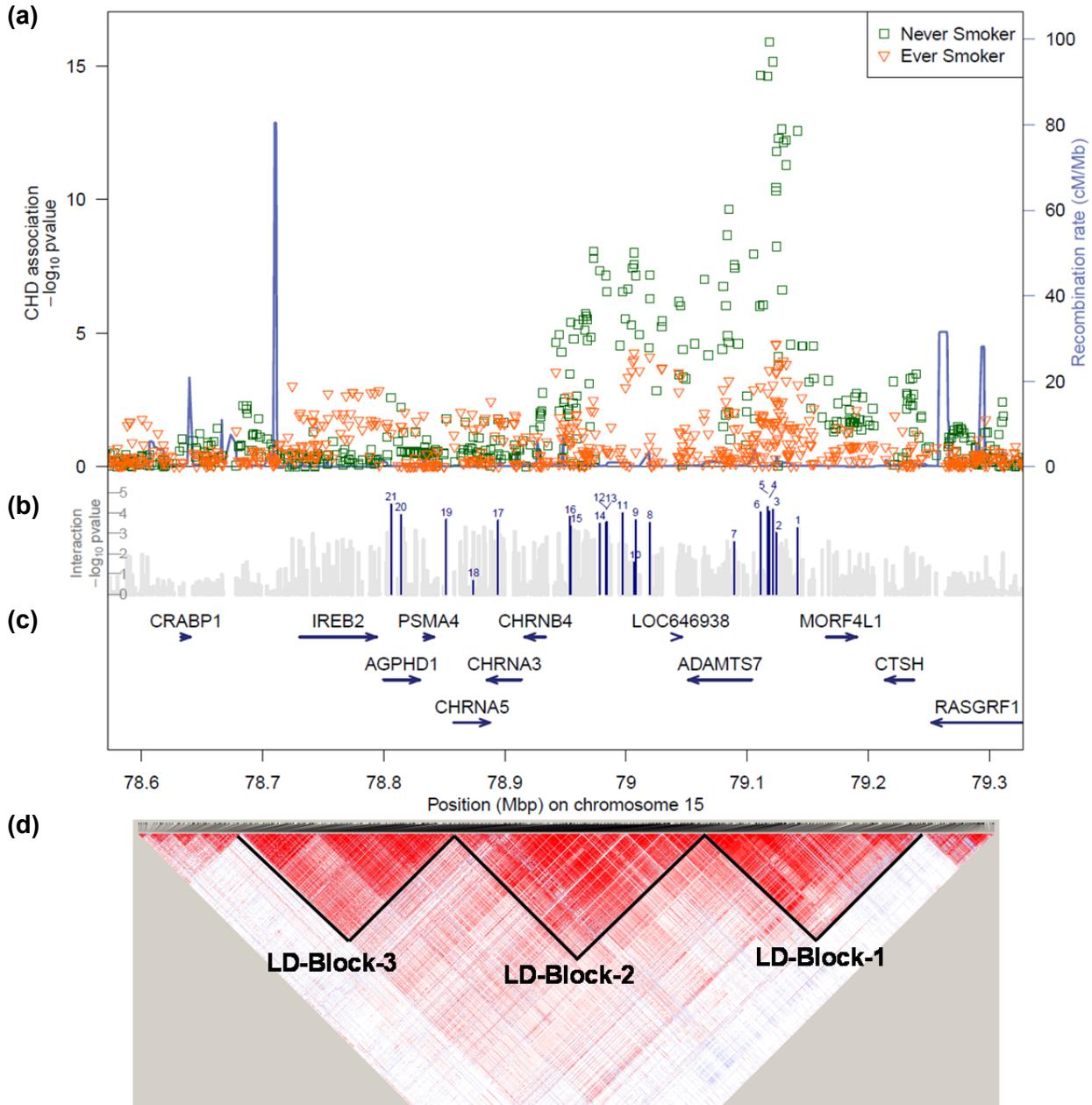
CHD = coronary heart disease; SB = smoking behavior; NPR: Not a previously reported variant with disease risk

*lead variant in association with CHD in our dataset; ^ lead variant in association with SB

¹⁻²¹each number refers to the physical location of the variant in figure-

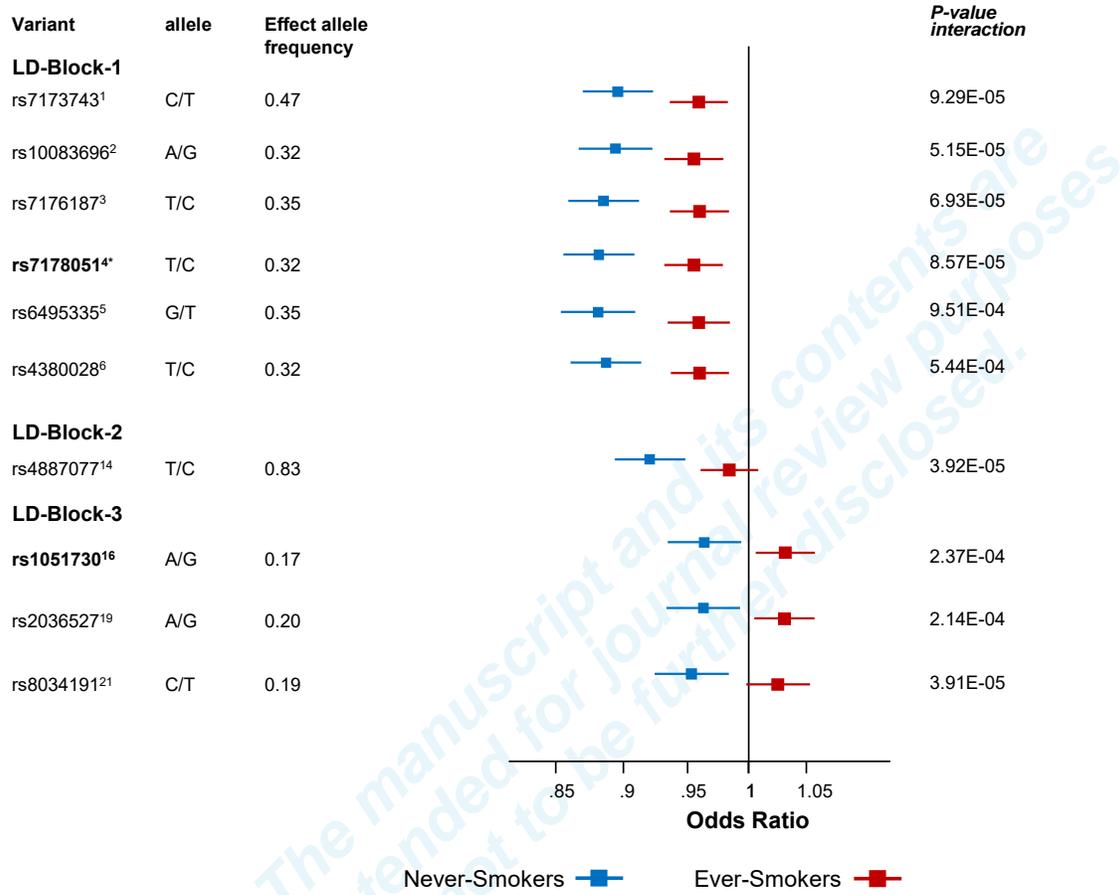
Disclaimer: This preprint is intended for journal publication only. It is not certified by peer review and is subject to change. All rights reserved. No reuse allowed without permission.

Figure-1. Analyses of the chromosome 15q25.1 locus association with CHD stratified by smoking status and gene-smoking CHD interaction analyses



1-rs7173743; 2-rs10083696; 3-rs7176187; 4-rs7178051; 5-rs6495335; 6-rs4380028; 7-rs3825807; 8-rs3813565; 9-rs11638490; 10-rs11072794; 11-rs11072791; 12-rs922692; 13-rs11638372; 14-rs4887077; 15-rs12899135; 16-rs17487514; 17-rs1051730; 18-rs637137; 19-rs2036527; 20-rs10519203; 21-rs8034191. LD 1-3 indicate three separate linkage disequilibrium blocks in European ancestry at the chromosome 15q25.1 locus.

Figure-2. Multiple variants at chromosome 15q21.1 have stronger effects on CHD risk in “never-smokers” compared to “ever-smokers”

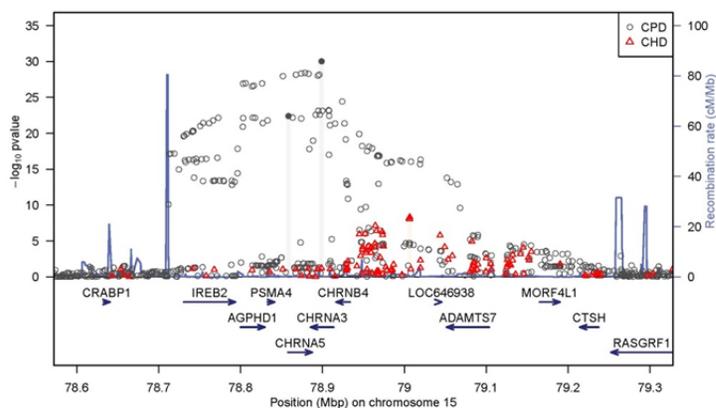


rs7178051 is the lead variant identified in association with CHD in our study population; whereas rs1051730 is the lead variant previously identified in association with smoking behavior . Variants are ordered based on their base pair position in **Figure-1**.

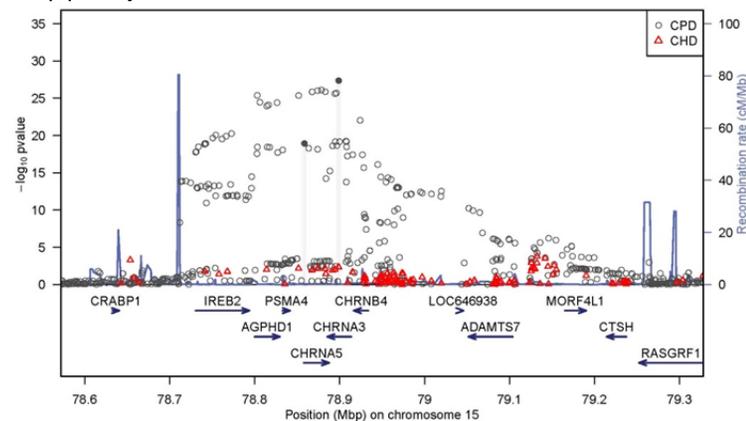
Figure 3. Conditional analysis of genetic variation at the chromosome 15q21.1 locus with coronary heart disease (CHD; red triangles) and smoking behavior (cigarettes per day, CPD; grey circles)

Stepwise conditional analyses for CHD risk and CPD behavior adjusting for top CHD variants at chromosome 15q21.1

(a) analyses conditioned on rs7178051

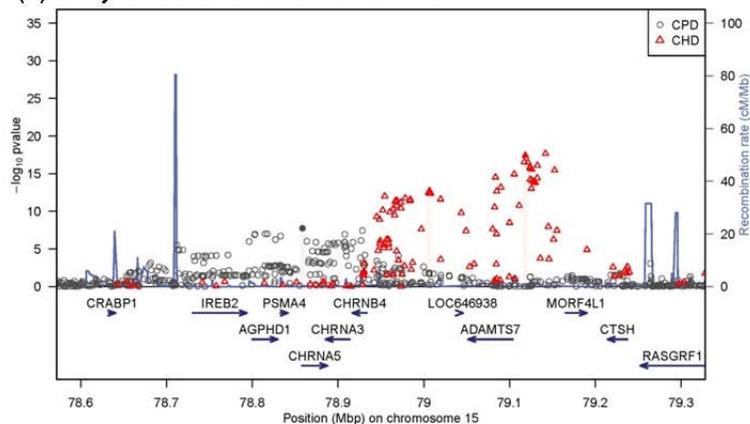


(b) analyses conditioned on rs7178051 and rs11072794



Stepwise conditional analyses for CHD risk and CPD behavior adjusting for top CPD variants at chromosome 15q21.1

(c) analyses conditioned on rs1051730



(d) Analyses conditioned on rs1051730 and rs684513

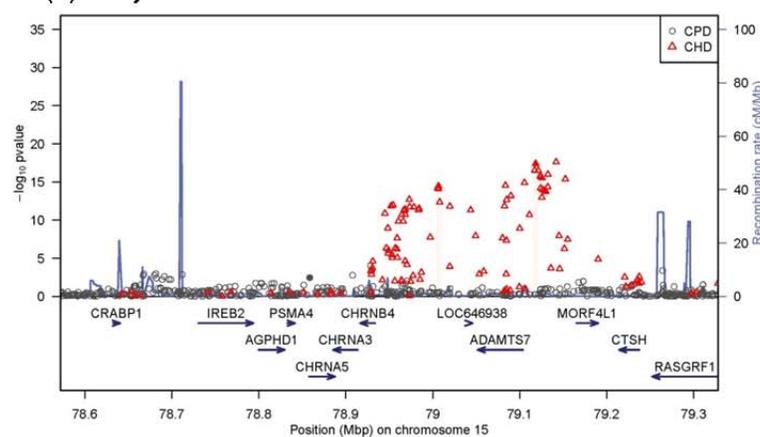
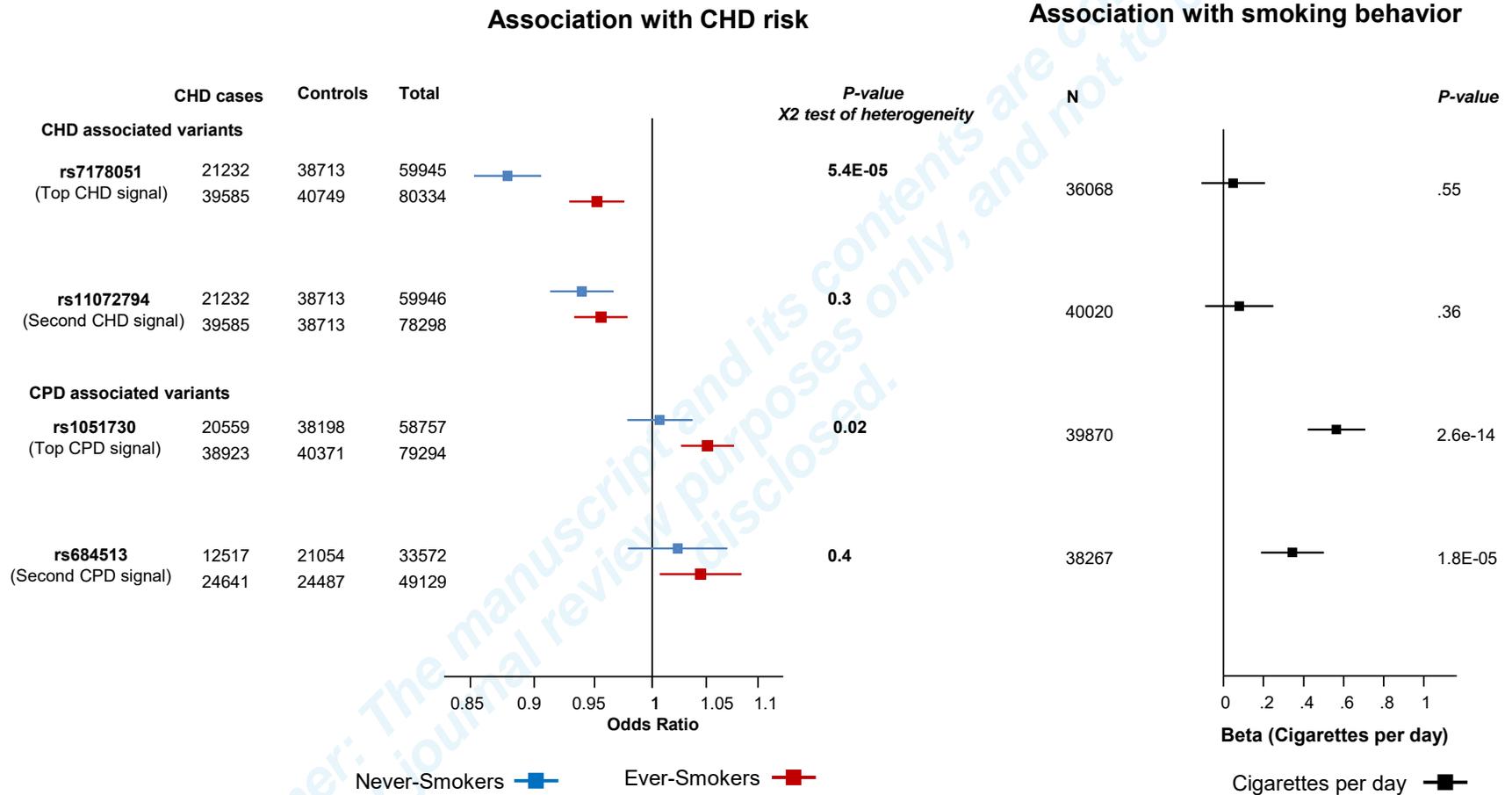


Figure 4. Mutually adjusted effects of 15q21.1 lead variants on coronary heart disease and smoking behavior



Gene-CHD and gene-smoking analyses for rs7178051 were adjusted for rs11072794, rs1051730 and rs684513;
 Gene-CHD and gene-smoking analyses for rs11072794 were adjusted for rs7178051, rs1051730 and rs684513;
 Gene-CHD and gene-smoking analyses for rs1051730 were adjusted for rs7178051, rs11072794 and rs684513;
 Gene-CHD and gene-smoking analyses for rs684513 were adjusted for rs7178051, rs11072794 and rs1051730.

Figure-5. (a) Expression of *ADAMTS7* and *CHRNA4-A3-A5* mRNAs and (b) cigarette smoke extract (CSE) induction of *ADAMTS7* mRNA in primary human coronary artery smooth muscle cells

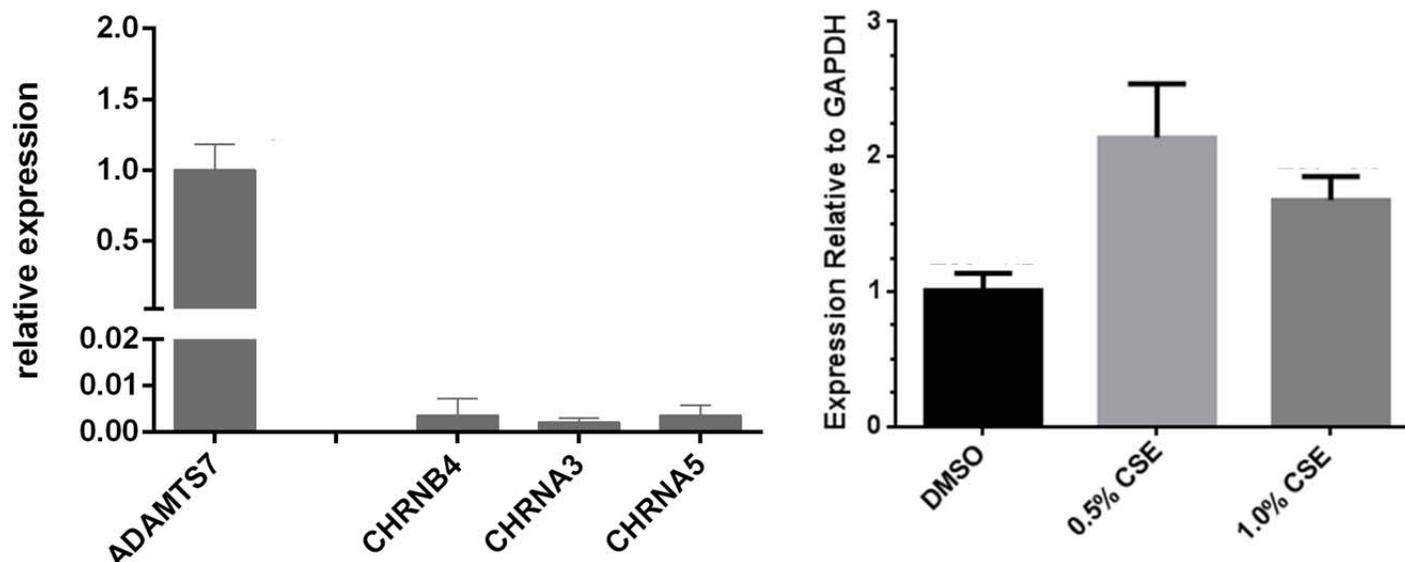
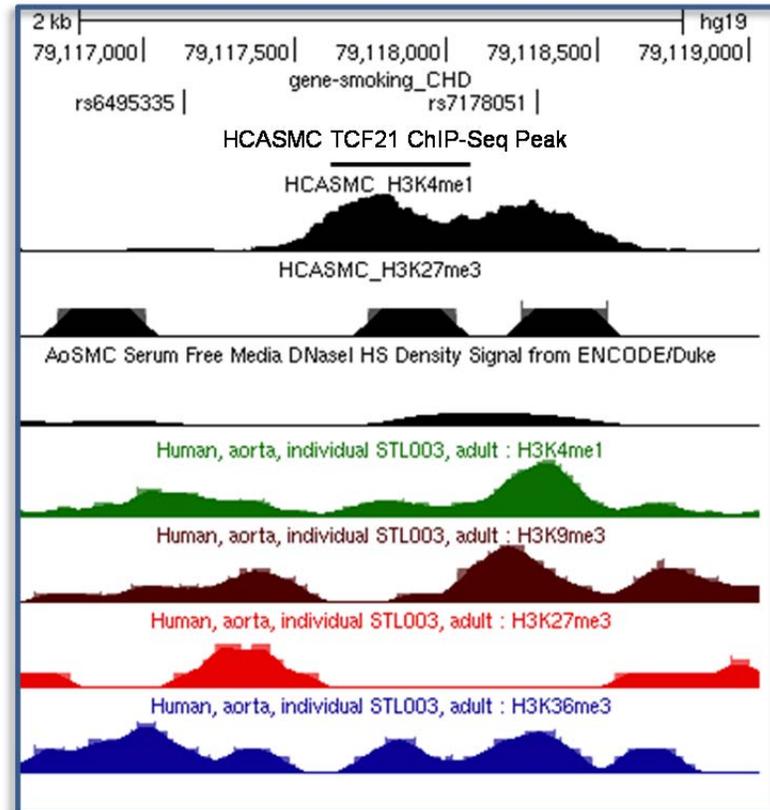


Figure-5c. Association of lead CHD and smoking behavior variants with candidate gene expression in available cells and tissues

Variant	Type	CHD direct.#	LCL in the MuTHER consortium (n=850)		HAEC (n=147)		HapMap CEU LCL (n=109)		GTEx Skeletal muscle (n= 142)		GTEx Nerve Tibial (n=101)	
			<i>ADAMTS7</i>	<i>CHRNA3-5</i>	<i>ADAMTS7</i>	<i>CHRNA3-5</i>	<i>ADAMTS7</i>	<i>CHRNA3-5</i>	<i>ADAMTS7</i>	<i>CHRNA3-5</i>	<i>ADAMTS7</i>	<i>CHRNA3-5</i>
rs7178051	Top CHD signal	-	4.1e-4 (-)	NS	0.0029 (-)	NS	NS	NS	NS	NS	NS	NS
rs11072794	Second CHD signal	-	6.0E-21 (-)	NS	NA	NS	0.0013 (-)	NS	NS	NS	NS	NS
rs1051730	Top CPD signal	-	NS	NS	NS	NS	NS	NS	NS	6.9E-7 (-) ¹	NS	6.9E-7 ¹
rs684513	Second CPD signal	-	NS	NS	NS	NS	NS	NS	NS	2.4E-7 (-) ¹	NS	NS

#Direction of association for the effect allele on CHD; NS: Not significant (P-value < 0.002; Bonferroni correction for 20 tests); HAEC: Human Aortic Endothelial Cells; LCL: lymphoblastoid cell lines; 1 Association with *CHRNA5* expression;

Figure-6. Genome browser view of regulatory features at rs7178051 on Chr15q21.1



Disclaimer: This content is intended for your use only.

Confidential,
not to be further