Pathway-specific Polygenic Risk Scores (PRS) identify OSA-related pathways

differentially moderating genetic susceptibility to CAD

Goodman: Pathway-specific effects of OSA on genetic CAD risk

Matthew O Goodman,^{1,2,3} Brian E Cade,^{1,2,3} Neomi Shah,⁴ Tianyi Huang,^{2,5} Hassan S Dashti,^{3,6,7}

Richa Saxena,^{1,2,3,6,7} Martin K Rutter,^{8,9} Peter Libby,¹⁰ Tamar Sofer,^{1,2} Susan Redline^{1,2}

Affiliations:

¹Division of Sleep and Circadian Disorders, Brigham and Women's Hospital and Harvard

Medical School, Boston, MA

²Division of Sleep Medicine, Harvard Medical School, Boston, MA

³Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

⁴Icahn School of Medicine at Mount Sinai, New York, NY

⁵Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical

School, Boston, MA

⁶Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

⁷Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital

and Harvard Medical School, Boston, MA, USA

⁸Division of Diabetes, Endocrinology and Gastroenterology, School of Medical Sciences,

Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic

Health Science Centre, Manchester, United Kingdom

⁹Diabetes, Endocrinology and Metabolism Centre, Manchester University NHS Foundation

Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom

¹⁰Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's

Hospital, Harvard Medical School, Boston, MA, USA.

Corresponding Author:

Matthew Goodman

Brigham and Women's Hospital

221 Longwood Avenue

Boston, MA 02115

(617) 525-9767

mogoodman@bwh.harvard.edu

Total word count: 8867 (including the title page, abstract, text, references, tables, and figures legends)

ABSTRACT

Background—

Obstructive sleep apnea (OSA) and its features, such as chronic intermittent hypoxia (IH), may differentially affect specific molecular pathways and processes in the pathogenesis of coronary artery disease (CAD) and influence the subsequent risk and severity of CAD events. In particular, competing adverse (e.g. inflammatory) and protective (e.g. increased coronary collateral blood flow) mechanisms may operate, but remain poorly understood. We hypothesize that common genetic variation in selected molecular pathways influences the likelihood of CAD events differently in individuals with and without OSA, in a pathway-dependent manner.

Methods—

We selected a cross-sectional sample of 471,877 participants from the UK Biobank, among whom we ascertained 4,974 to have OSA, 25,988 to have CAD, and 711 to have both. We calculated pathway-specific polygenic risk scores (PS-PRS) for CAD, based on 6.6 million common variants evaluated in the CARDIoGRAMplusC4D genome-wide association study (GWAS), annotated to specific genes and pathways using functional genomics databases. Based on prior evidence of involvement with IH and CAD, we tested PS-PRS for the HIF-1, VEGF, NFκB and TNF signaling pathways.

Results—

In a multivariable-adjusted logistic generalized additive model, elevated PS-PRSs for the KEGG VEGF pathway (39 genes) associated with protection for CAD in OSA (interaction odds ratio 0.86, $p = 6E-04$). By contrast, the genome-wide CAD PRS did not show evidence of statistical interaction with OSA.

Conclusions—

1
2 We find evidence that pathway-specific genetic risk of CAD differs between individuals with
3 and without OSA in a qualitatively pathway-dependent manner, consistent with the previously
4 studied phenomena whereby features of OSA may have both positive and negative effects on
5 CAD. These results provide evidence that gene-by-environment interaction influences CAD risk
6 in certain pathways among people with OSA, an effect that is not well-captured by the genome-
7 wide PRS. These results can be followed up to study how OSA interacts with genetic risk at the
8 molecular level, and potentially to personalize OSA treatment and reduce CAD risk according to
9 individual pathway-specific genetic risk profiles.

10

INTRODUCTION

Obstructive sleep apnea (OSA) is a common disorder characterized by recurrent episodes of hypoxemia during sleep, sleep fragmentation, and activation of the sympathetic nervous system, potentially stimulating pro-inflammatory and pro-thrombotic pathways that drive atherosclerosis [1]. Epidemiological studies implicate OSA as a potentially modifiable risk factor for cardiovascular disease (CVD), stimulating research aimed at quantifying the impact of OSA and its treatment on both coronary artery disease (CAD) and cerebrovascular disease [2]. While OSA increases the incidence of cerebrovascular disease, weaker and less consistent associations support associations between OSA and incident CAD [3]. Moreover, large clinical trials of the effect of positive airway pressure (PAP) treatment of OSA on composite CVD outcomes have provided equivocal results, in some cases with point estimates in the direction of increased risk in the treatment group [4-7].

Potential explanations for disappointing results of PAP RCTs include treatment non-adherence, non-optimal trial participant selection (e.g. advanced CVD at baseline) and limited statistical power [8, 9]. Another possibility is that putative CVD benefits of OSA treatment may vary across individuals, influenced by factors related to characteristics of OSA-related stress not captured by the traditional metric of OSA severity (apnea hypopnea index, AHI), as well as by underlying host characteristics [10, 11], including genetic predisposition to CVD and genetic variation within known OSA pathways [12, 13]. Moreover, some features of OSA might even prove protective for CVD in some individuals, and therefore treatment of OSA with PAP could increase CVD risk among those individuals [14]. This hypothesis derives support from human and animal studies showing that exposure to intermittent hypoxemia - a cardinal feature of OSA

- can promote collateral blood flow, reflective of well-known effects of ischemic preconditioning [15].

The need to develop a precision medicine-informed approach to CVD risk stratification in OSA agrees with data indicating that OSA-related CVD risk varies across subgroups defined by clusters of phenotypic traits [16]. A central question, and the one that we address here, is whether genetic variation in CAD risk interacts with OSA features to modulate CAD risk in a pathway-dependent manner. We conceptualize repeated intermittent hypoxia (IH) and reoxygenation as a modifying environment for preexisting genetic risk for CAD, locating our analysis within a gene-by-environment (GxE) interaction paradigm. Our pathway-level GxE approach provides a new scale on which to study the genetic architecture of CAD susceptibility and its modification by OSA.

We build on prior work that (1) identified genetic markers for CAD risk involved in atherosclerosis, small vessel disease, inflammatory pathways, and angiogenesis, with evidence from genome-wide association studies (GWAS), expression studies, animal experiments and clinical trials [17-19]; and (2) showed the influence of IH on a variety of pathways (in positive and negative directions) relevant to CAD, including redox, inflammatory and angiogenic pathways [20].

Our approach considers effect modification of genetic risk, aggregated to the pathway level, and investigates pathway-specific polygenic risk scores (PS-PRS) for CAD. PRSs have demonstrated success in identifying high risk populations, but may perform poorly due to overly simplistic dimension reduction when compared with non-linear methods to summarize risk across many alleles [21, 22]. Here we consider a novel PRS approach, where we group markers

by pathways to create multiple PS-PRSs aligned logically with genetic architecture, to facilitate investigation of GxE interaction at the pathway level.

We hypothesized that common genetic variation, in specific molecular pathways for CAD that are linked to the pathophysiology of OSA, influences the likelihood of CAD events differently in individuals with and without OSA, in a pathway-dependent manner. Our work was motivated by the potential for differential effects of OSA on multiple processes that impact CAD, such as inflammation and atherosclerosis, thrombosis, collateral vessel formation (via angiogenesis), control of coronary artery blood flow, and plaque disruption. Of central interest is the idea that OSA could be a “double-edged sword” with both positive or negative effects on CAD. In particular, intermittent hypoxia – a physiological stress characteristic of OSA – may exert either positive effects through ischemic preconditioning or negative effects by boosting atherosclerotic processes including inflammation [14, 15]. We pre-selected four regulatory pathways of interest based on our hypotheses of the impacts of chronic IH on angiogenic and inflammatory processes in CAD, the HIF-1, VEGF, NFκB and TNF signaling pathways. We focused on these four representative pathways for their recognized roles in these processes, which have been previously shown to involve differential expression of constituent genes in OSA [23-32].

By investigating whether OSA enhances, does not change, or ameliorates genetic risk of developing CAD among individuals with specific risk allele profiles within these pathways, we aim to provide new mechanistic insight into the biological effects of OSA on CAD. In addition, we explore the ability to identify patients with OSA who may have either protection against or increased risk of CAD events from their OSA status according to their specific genetic profile. Ultimately, we propose that this approach provides a novel step on the path toward

understanding the complex role of OSA in CAD and formulating a personalized approach to treatment.

METHODS

UK Biobank

The UK Biobank (UKBB) cohort consists of 502,620 participants, recruited between 2006 and 2010 from across the United Kingdom, and is described elsewhere in detail [33]. Participants were assessed at baseline using questionnaires that included components for ascertaining medical conditions, lifestyle, and demographic information. Additional interviews were conducted by trained medical staff to assess medical history, health status and medication intake. We used available self-reported doctor-diagnosed medical conditions and ICD10 codes derived from the UKBB hospital episode inpatient data set (retrieved 03/21/2019) to define medical conditions of interest. Participants were genotyped on two closely related genotyping arrays: UK BiLEVE: N = 49,939, and UK Biobank Axiom: 438,343. In total, 488,282 participants were genotyped. All genotypes were then imputed using reference sequence data as previously described [34]. The National Health Service National Research Ethics Service (ref. 11/NW/0382) gave approval for the UK Biobank study, with each participant providing written informed consent.

OSA status assessment

We defined cases and controls as follows: Cases of OSA must have had self-reported snoring AND at least one of the following sleep apnea diagnoses:

1) an ICD10¹ code G47.3 for sleep apnea (N = 6643)

OR

2) a self-report of doctor diagnosed sleep apnea documented during an interview with a trained medical professional (UKBB field 20002, coding 1123: sleep apnoea) (N = 1757).

Controls were non-cases. Individuals who could not be confidently defined as either cases or controls were excluded. To reduce the possibility of including cases of central sleep apnea ascertained using sleep apnea diagnostic codes, we required snoring (a cardinal symptom of OSA), documented by self-report in the UKBB sleep questions interview. Among participants without OSA we excluded those who self-reported both snoring and daytime sleepiness, two common symptoms of OSA. A total of 471,877 participants (4,974 OSA cases) with complete covariate data were available after making these exclusions.

CAD status assessment

Coronary artery disease (CAD) cases were assessed based on a composite definition, including UKBB participants with self-reported heart attack/myocardial infarction, phecode 411.2 or UKBB algorithmic ascertainment for myocardial infarction, including self-report, hospital admission, or death certificate.

Covariate Assessment

Diagnostic information for medical conditions were obtained from the UKBB data sets for self-reported doctor-diagnosed non-cancer illness and hospital episode inpatient diagnostic codes. ICD10 codes were collapsed to medically interpretable groupings using the phecode system [35], using Phecode Map version 1.2 for WHO ICD10's. Hypertension, type 1 (T1D) and

¹ ICD9 records were not analyzed as A) ICD9 codes come from a unique subset of 20k Scottish participants with available hospital inpatient episodes from 1980 to 1996, whereas the ICD10 data was collected almost exclusively after 1996 B) the fact that these ICD9s contain no instances of sleep apnea codes 327.23 or 327.2, nor any sleep codes in the 327 range.

type 2 diabetes (T2D), COPD, asthma, and pulmonary fibrosis were ascertained as composite phenotypes, as described in Supplementary Table S1. Additional covariates were self-reported race, genotype platform (BiLEVE vs Axiom) and UK Biobank-provided genetic PCs.

Polygenic Risk Score Construction

We based our analysis of genetic CAD risk on a published PRS described by Khera *et al* [36], with single nucleotide variant (SNV) effects estimated for each of 6,630,150 SNVs, using the linkage disequilibrium (LD) aware PRS algorithm LDpred [37] and summary statistics from the CARDIoGRAMplusC4D Consortium CAD GWAS of 60,801 cases and 123,504 controls of mainly European ancestry [38].

Pathway specific polygenic risk scores (PS-PRSs) were calculated for each individual as the effect-weighted sum of the count of risk alleles at each locus across all pathway SNVs:

$$PS - PRS_k = \sum_{i: a_i \in \text{pathway } k} w_i a_i$$

where w_i is the LD-adjusted additive genetic effect and a_i is the count (or imputed dosage) of the risk alleles at locus i of pathway k . PS-PRSs for select biological pathways were derived based on subsets of SNVs assigned to various pathways. Pathways are defined as sets of genes, and SNVs are assigned to genes via annotations derived from external reference data sets. We annotated SNVs in sequential rounds, prioritizing annotations with high specificity. SNVs assigned in prior rounds were set aside and were not annotated with less-specific data (Figure 1). We ordered the resources as follows: GenCode v34 exons [39], FANTOM5 [40], Promoter-Capture HiC from white blood cells [41], GTEx v7 eQTLs from artery, heart and whole blood tissues [42], GeneHancer v4.4 [43], and GenCode protein-coding regions (20kbp upstream to 10kbp downstream) and noncoding gene sequence regions. Further detail is given in the Methods

Supplement. We also calculated gene-specific risk scores, GSR_j , i.e. PS-PRSs for a single gene, where the above sum is taken over loci i : $a_i \in \text{gene } j$ for a single gene of interest.

Defining pathways of interest

Based on our *a priori* hypotheses on the role hypoxemia plays in CAD, inducing angiogenesis and inflammation, we selected four pathways from the literature, for which the nominal genes, namely *HIF1A*, *VEGFA*, *NFKB1/NFKB2* and *TNF*, have been shown to be differentially expressed in OSA. For each, we created gene-pathways using two approaches. First, we selected the core-gene pathways, defined as the nominal gene(s) and their receptors (for HIF1, core genes selected were: *HIF1A* and *ARNT*, for VEGF: *VEGFA*, *FLT*, and *KDR*, for NFκB: *NFKB1A*, *NFKB1*, *RELA*, *NFKB2*, *RELB*, and for TNF: *TNF*, *TNFRSF1A*, and *TNFRSF1B*). Second, we selected more comprehensive gene pathways based the Kyoto Encyclopedia of Genes and Genomes (KEGG) database: HIF-1 signaling (KEGG pathway code: hsa04066), VEGF signaling (hsa04370), NFκB signaling (hsa04064), and TNF signaling (hsa04668). We also included, as comparators, the full PRS on 6.6M SNVs, as well as a curated CAD/OSA pathway consisting of 225 genes involved in both OSA and CAD based on literature review. Because KEGG pathway boundaries are somewhat arbitrary and include varying levels of completeness in upstream and downstream gene modules, we also organized each KEGG pathway into sub-modules. For example, the KEGG VEGF pathway was organized into several partially overlapping downstream submodules according to the KEGG pathway diagram, including: endothelial proliferation (15 genes), endothelial migration (9 genes) and endothelial survival (12 genes). For complete pathway and module definitions see Supplementary Table S2.

Statistical Analysis

We studied the interaction of OSA case status with pathway-specific risk for CAD in several steps. 1) We attempted to annotate each marker identified by Khera *et al* to one or more associated genes (Figure 1). SNV markers that were not annotated were discarded, resulting in a QC-filtered, gene-annotated, genome-wide set of markers. 2) We selected pathways of interest and defined pathway-based SNV-sets implied by the SNV-gene and gene-pathway assignments. 3) We computed pathways-specific genetic risk scores (PS-PRS), as a sum of allele counts weighted by per-allele effect size and direction across the pathway SNV-set. 4) We tested the associations of PS-PRS and OSA status, and their GxE interaction, on the composite CAD outcome, in a logistic generalized additive model, controlling for the mutual interaction of age and BMI by sex (via tensor-product splines) as well as smoking and its interaction with sex [44], self-reported white race, the first 5 genetic PCs, genotype platform (BiLEVE), asthma and chronic obstructive pulmonary disease (COPD). As two of the most common pulmonary diseases, COPD and asthma were controlled for as both are associated with both OSA and CAD, and can potentially affect gas exchange and exacerbate hypoxia. The ten pathways of interest described were allocated Type-1 error in the primary analysis, with Bonferroni-corrected significance level of 5E-3. In a secondary analysis, to further explore whether module and gene-level PRS showed consistent associations with CAD, as well as to suggestively localize effects estimated in the primary analysis, we looked at pathway sub-modules and individual genes. For each core-gene module we modeled GxE interaction for the gene-specific risk scores (GSRS) of the constituent genes, and for each KEGG pathway, we modeled sub-module PS-PGRSs and GSRSs and reported those with both nominally significant risk-amplifying main effects and nominally significant GxE interactions. We further performed the following sensitivity analyses: including comorbidity covariates potentially mediating the OSA-CAD relationship

(hypertension, Type 1 and Type 2 diabetes, along with their sex-interactions [44], and pulmonary fibrosis); excluding individuals with heart failure or stroke; sex-stratified analysis; and an analysis limited to self-reported white Europeans (the predominant ancestry in the CAD GWAS). Further detail is provided in the supplement.

RESULTS

Clinical characteristics

Of the 476,851 participants in our analytic sample, 4,974 (1.1%) were classified to have OSA and 26,699 (5.6%) to have CAD. The median (IQR) age was 58 (50, 63) years with a similar age distribution by OSA status (Table 1).

As expected, OSA cases tended to have a higher prevalence of several CAD risk factors and comorbidities. Individuals classified with OSA were predominantly male, had a higher mean BMI, and included more ever-smokers. Pulmonary diseases, including COPD and asthma also tended to have a higher prevalence among OSA cases (Table 1). There was a higher percentage of OSA cases genotyped on the BiLEVE platform, as opposed to the UKBB Axiom platform, consistent with the fact that the BiLEVE pilot study was enriched for pulmonary disease.

Characteristics of the PS-PRSs

Our annotation procedure resulted in the assignment of 70.3% of the 6.6M markers to one or more genes (average: 1.1 genes assigned per annotated variant). Pathway genes are shown in Supplementary Table S2. Gene-annotated CAD PRS data is available on request. After standardization on the cohort as a whole, participants with or without OSA had similar PS-PRS (Table 1).

The NF κ B signaling pathway (KEGG hsa04064: 100 genes) and TNF signaling pathway (KEGG hsa04668: 110 genes) share 29 genes, which induced a correlation of 0.39 in the pathway-specific risk scores (Figure 2). The curated 225-gene OSA/CAD pathway shares 22 genes with the KEGG HIF1 pathway, driving the observed correlation of 0.54. Similarly, because VEGF core-genes *VEGFA* and *FLT1* and HIF1 core-genes *HIF1A* and *ARNT* reside within the HIF1 KEGG pathway, they induce correlations of 0.44 and 0.32 respectively (Figure S3). Additional pairwise correlations for pathways, genes, and modules shown in Figures 2 and S3. Note that as pathways largely come from disjoint regions of the genome and, due to genetic recombination, allele counts from distal variants with low LD (e.g. in separate genes) will be nearly uncorrelated statistically. Hence the low observed correlation between participants' genetic risk scores in the majority of pathway pairs is expected due to statistical independence of the non-overlapping variants, and conversely, observed correlations are driven by the PS-PRS components attributable to the overlapping genes.

Estimated main effects of OSA, PRS and PS-PRSs

We found that OSA associated with higher CAD risk, with an approximate odds ratio (OR) of 1.5, and 95% confidence interval (CI) of (1.4, 1.6) across all models shown in Table 2. We also confirmed that the full genome-wide CAD PRS associated strongly with increased CAD risk in our model. For each standard deviation increase in the CAD PRS, the odds of CAD (in both OSA and OSA controls) increased by a factor of 1.70 (CI: 1.67, 1.72). Additionally, the PS-PRSs for each KEGG signaling pathway associated strongly with CAD, but had notably smaller estimated odds ratios (range 1.040 to 1.062). Core gene modules had still smaller odds ratios (range 1.019 to 1.032) (Table 2).

Estimated interaction effects between OSA, and PS-PRSs

In our primary analysis, we find evidence of effect-measure modification of genetic risk by OSA status. Specifically, at Bonferroni-corrected significance-level $5E-3$, we see significant qualitative effect modification in the VEGF KEGG pathway (interaction odds ratio 0.86, $p = 6E-04$), leading to differential genetic effects among participants with and without OSA (Table 2). Concretely, among OSA cases, we estimated the odds of CAD decreased, by a factor of 0.89 per standard deviation increase in the VEGF KEGG PS-PRS, contrasted with an increase of 1.04 among OSA controls. We find a qualitatively similar point estimate for GxE interaction in the VEGF core-genes PS-PRS. We do not find significant evidence of effect-measure modification between OSA and the KEGG HIF1 PS-PRS, TNF PS-PRSs, the NF κ B PS-PRSs, nor the full CAD PRS. However, we find a nominally significant trend for risk-amplifying effect-measure modification for both the HIF1 core-genes PS-PRS (interaction OR 1.1, $p = 4.46E-02$) and the curated 225 gene PS-PRS based on a literature-derived CAD- and OSA-involved genes (interaction OR 1.1, $p = 1.8E-02$).

Table 3 reports secondary analysis of module and gene-specific interactions. Adjusting for multiple testing, we find statistically significant effect-measure modification in one of the 30 pathway submodules (VEGF endothelial migration: *CDC42*, *MAPK11*, *MAPK12*, *MAPK13*, *MAPK14*, *MAPKAPK2*, *MAPKAPK3*, *PTK2*, *PXN*) and none of the 268 genes tested from the four KEGG pathways of interest. We see nominally significant effect-measure modification in several pathway modules and genes, with estimated effects generally consistent with pathways and modules that contain them (Table 3).

Sensitivity Analyses

We obtained similar results in sensitivity analyses. For the analyses of males-only, self-reported whites-only, and when adjusting for potentially mediating comorbidities, the KEGG

VEGF PS-PGRS had consistently Bonferroni-significant risk-inverting effect-measure modification, while the curated CAD- and OSA-related pathway had consistently nominally significant risk-amplifying effect modification; females-only was underpowered with only 71 joint OSA/CAD cases, nonetheless KEGG VEGF was nominally significant with risk-inverting effect.

DISCUSSION

Overview and context

Using biobank-scale data we tested the hypothesis that genetic variation in specific molecular pathways that modulate risk of CAD may influence the propensity for CAD differently in individuals with and without OSA—an exposure postulated to have complex effects on atherosclerosis, inflammation, and angiogenesis. We demonstrate novel interaction effects between OSA and PS-PRS within one pathway (VEGF) postulated *a priori* from current biological understanding to play a role in OSA-related hypoxia in CAD. The effect-measure modification of PS-PRS effects suggests that OSA can increase or attenuate genetic risk for CAD in a pathway-dependent manner, and, reciprocally, that subgroups with specific genetic profiles within certain CAD pathways may experience the influence of OSA differently. Our results also suggest heterogeneity of this effect modification across pathways, modules and genes.

Prior work has investigated phenotypic factors, biomarkers, and genetic variation to understand better and dissect the heterogeneous effects of OSA on CAD, including multiple variables, such as age, gender and comorbid diagnoses that may moderate associations of OSA and CVD, as well as blood-based biomarkers of CVD [45, 46]. Nonetheless, prior research using genetic variation to explore the divergent individual responses to hypoxic conditions has been

limited, despite suggestions advocating this approach to risk stratification of patients with OSA [47]. Our use of PS-PRS, applied to core gene modules and pathways, provides a novel means for interrogating pathway-level effect-measure modification in large-scale data to uncover biological heterogeneity.

Primary and secondary analyses

In our primary analysis, we see Bonferroni-significant evidence of risk-inverting effect-measure modification in the KEGG VEGF pathway. In secondary analysis also see suggestive evidence of risk-inverting effect-measure modification across multiple independent signals within this pathway. This qualitative effect-measure modification in the VEGF KEGG pathway suggests that among OSA cases, the direction of effect of the VEGF PS-PRS risk alleles, averaged across markers in the pathway, is reversed compared with OSA controls. For example, among OSA cases whose VEGF risk score is one standard deviation above the mean, we estimate that odds of CAD is decreased by approximately 14% compared to those with the mean VEGF risk score.

One explanation consistent with these data would be that OSA perturbs the tissue environment such that SNVs that increase risk in normoxia on average now function to reduce risk under hypoxia. For example, downregulation of a certain gene under normoxia may have been deleterious, but becomes beneficial under hypoxia (or vice versa). The 39 genes in the KEGG VEGF pathway mediate inflammation and angiogenesis, with genes regulating endothelial cell survival, proliferation, and migration as well as vascular permeability. Under intermittent hypoxia, angiogenesis may predominate, potentially conferring protection from clinically observed or fatal CAD events (ascertained in our CAD outcome) due to heart muscle collateralization.

The KEGG VEGF risk-inverting GxE effects appear to be largely driven by the VEGF endothelial migration submodule. However, there appear to be two additional congruent independent signals contributing the overall KEGG VEGF pathway effect: the VEGF core-genes risk-inverting effects, and a separate independent signal distributed among the VEGF endothelial permeability, endothelial survival and PI3K/AKT submodules, which are correlated to one another due to shared genes. This suggests the utility of pathway level analyses in circumstances where consistent effects are combined across many statistically independent genes or loci, which do not reach significance thresholds on their own, but contribute signal to detection of GxE at the pathway level.

By contrast we see nominally-significant suggestive evidence of risk-amplifying effect-modification in the HIF1 core genes module, apparently driven by the *ARNT* gene, as well as in the PS-PRS consisting of 225 curated OSA-CAD genes, suggesting a risk-amplifying GxE signal may be more widely distributed. If confirmed these effects are consistent with the idea that OSA-related physiological stressors may amplify the deleterious effects certain genetic variants on CAD. For example, *ARNT*, whose gene-product HIF1- β creates a heterodimer with HIF-1 α enabling nuclear activation of the HIF1 transcription factor, is a plausible key player in the regulatory network immediately upstream of a host of cellular responses to hypoxia. Although *ARNT* is not under direct hypoxic regulation, concentration of HIF1- β may serve as a rate-limiter on HIF1's transcription factor activity specifically under hypoxic conditions when HIF1- α concentration is maximized [48]. Hence, genetic effects on *ARNT* plausibly interact mechanistically with IH-induced effects on HIF1- α . Further investigation of OSA's and IH's potential for risk-amplifying effect modification appears warranted. If confirmed, differential positive and negative GxE interaction effects, increasing or decreasing individuals' prior genetic

risk in separate genes and pathways in OSA, would be consistent with the idea of a “double-edged” role of OSA in CAD.

Note that the KEGG HIF-1 pathway contains sub-modules with contrasting GxE effects including the KEGG HIF1 angiogenesis (risk-inverting) and the HIF1 core-genes (risk-amplifying) modules. As observed Supplementary Figure S3, the risk scores for these gene modules are both positively correlated with the KEGG HIF1 PS-PRS as a whole, being component terms in the sum. This observation suggests that the KEGG HIF1 signaling pathway may contain opposing signals which together attenuate overall HIF1 pathway effect modification. This situation illustrates the limitations of targeting large heterogeneous pathways in studying GxE effect modification and underscores the utility of also examining pathways limited to genes or sub-modules that better ensure specificity. Nonetheless, since pathway boundaries can be vague and subjective, further research is needed to investigate data-driven methods for aggregating co-regulated genes to modules and pathways that may respond similarly to stimuli such as IH.

A natural question is whether the strong biological relationships among the chosen pathways challenge our interpretation. Clearly the selected pathways implicated in both IH and CAD are co-regulated: notably VEGF is explicitly regulated by HIF1, but cross-talk and co-regulation among all four pathways including NFκB and TNF likely exists, and gene expression levels may be correlated as a result. Nevertheless, our analysis does not directly analyze gene expression but rather genetic variation and its perturbation of this signaling network due to risk alleles specific to each pathway that are largely statistically independent across pathways. Even though increased concentrations of HIF1 in OSA is expected to augment VEGF, a number of considerations could explain why the CAD effects of e.g. of *ARNT* risk alleles may (on average)

be exacerbated but the effects of VEGF pathway risk alleles may (on average) be reversed, including differential downstream effects.

These results open the way to varied experimental studies to probe the potentially novel molecular mechanisms that may lead to the observed qualitative effect-measure modification and underly the findings. Further study of genes and variants in this pathway and their effect on measured biomarkers and gene expression may provide clarification of the genetic mediators of CAD risk in OSA, and potentially point to molecular targets for ameliorating OSA-related CAD.

Additional Findings

Testing our primary hypothesis yielded additional interesting findings. We demonstrated that a genome-wide CAD PRS associated with CAD in both the OSA cases and controls was distributed similarly (in mean and standard deviation) and did not reveal significant effect-measure modification of the genome-wide CAD PRS. We also confirmed that the chosen PS-PRSs for CAD associate positively with CAD in OSA controls, and that we see similar distribution of PRS and PS-PRSs when stratifying by OSA status. Additionally, given that nominally significant point estimates of interaction between genetic risk for CAD and OSA differed qualitatively across pathways, this suggests analyses ignoring heterogeneity (such as only assessing the full PRS) may incorrectly lead to the conclusion that there is no evidence of GxE interaction. Lastly, we demonstrated a consistent positive association between OSA and CAD, after adjusting for covariates as well as PS-PRS main and GxE interaction effects, with an effect size (OR 1.5) similar to that previously reported in observational cohort studies [49].

Strengths and weaknesses

This study has several notable strengths. Our methodology offers the opportunity to detect distinct positive and negative interaction effects in separate pathways, which might

otherwise interfere with each other and prevent detection of a GxE interaction at the genome-wide PRS level. Furthermore, if our hypothesis holds, environments such as OSA-related hypoxemia will perturb gene activity across regulatory pathways or submodules containing thousands of SNVs. So, we expect aggregating SNV risk to genes, modules, or pathways acting together to produce similar downstream biological effects may enhance power and reduce the multiple testing burden, as compared with SNV-level interaction testing. So, PS-PRS may provide useful intermediate levels between individual genetic markers and genome-wide risk scores to study effect-measure modification. And unlike the challenges to studying these regulatory pathways with biological assays in relevant tissues and with appropriate timing, our approach extracts additional insights from existing GWAS and biobank data.

The study also has several limitations. First, we perform a cross-sectional analysis, due to the limited follow-up for incident CAD, as well as the difficulty determining the date of onset of sleep apnea. Cross-sectional analysis is commonly used in genetics, including the GWAS and PRS analyses on which our PS-PRS were based. While in the GxE context ideally the OSA exposure should be established as preceding the CAD outcome, the presence of OSA in the medical record generally is believed to reflect disease that occurred years earlier [50]. Second, UK Biobank has several sources of selection bias, including healthy volunteer bias, which may limit generalizability. Also, to overcome a limitation of the UK Biobank dataset in which sleep apnea is not subclassified as central or obstructive either in self-reported or medical records data, we attempted to reduce potential CSA cases, which are likely causally downstream of atherosclerosis and CAD, by requiring the presence of self-reported snoring. Third, the number of ascertained OSA cases is low compared to expected population prevalence estimates. Nonetheless, we expect our ascertainment of OSA to be specific (likely identifying more severe

and hypoxic individuals) if not sensitive, and we took additional steps to reduce misclassification of cases and controls using cardinal symptoms of OSA. Also, the use of a binary classification for OSA also did not allow us to characterize OSA according to severity of hypoxemia or sleep fragmentation, which precluded our ability to test dose-response associations or associations with specific OSA stressors.

Summary

The existence of large-scale genetic data such as the UK Biobank opens up new avenues in understanding genotype-phenotype relationships in the context of different environments, host factors and comorbid diseases. Genome-wide data on genetic risk variants and novel genomic risk scores enabled by new LD-aware PRS algorithms such as LDpred present opportunities to develop robust pathway-level characterization of genetic risk. Building on these novel PS-PRSs we deploy an interaction analysis that allows us to probe relationships between OSA, CAD genetics and CAD outcomes that localize to particular molecular pathways. In turn we can gain a deeper understanding of genetic disease architecture, including modification of pre-existing genetic risk by exposures such as OSA.

In conclusion, this study provides insight into the influence of OSA on CAD and identifies a novel mechanism that may explain the evident heterogeneity of CAD risk among individuals with OSA. Specifically, we find that OSA status can modify pre-existing genetic risk of CAD within the VEGF pathway, and more generally we provide a plausible approach to assess stratified genetic risk for CAD in OSA by illustrating a novel method for understanding pathway-specific genetic risk and its role in GxE interaction. Future larger PS-PGRS GxE studies may allow probing of dose-dependent effects of OSA-related intermittent hypoxia on a comprehensive set of CAD pathways, while further study of these pathways and genes with

additional –omics data or animal studies may further clarify our current findings. Having identified genetic pathways whose effect is modified by OSA status, we may also identify individuals with OSA who are most susceptible (or resilient) to CAD, disentangling their genetic risk, their OSA exposure risk, and the interaction between the two. Together these investigations may lead to personalized management of OSA and the ability to target treatment toward specific impacted pathways in those at heightened risk.

Acknowledgments

This research has been conducted using the UK Biobank Resource (application 6818). We would like to thank the participants and researchers from the UK Biobank who contributed or collected data.

Sources of Funding

National Heart, Lung, and Blood Institute, NIH; The University of Manchester (Research Infrastructure Fund).

Disclosures

Dr. Goodman reports grants from National Heart, Lung, and Blood Institute, NIH, during the conduct of the study; Dr. Cade has nothing to disclose; Dr. Shah has nothing to disclose; Dr. Huang has nothing to disclose; Dr. Dashti has nothing to disclose; Dr. Saxena has nothing to disclose; Dr. Rutter reports consulting fees and non-promotional lecture fees from Novo Nordis, outside the submitted work; Dr. Libby was an unpaid consultant to, or involved in clinical trials for Amgen, AstraZeneca, Baim Institute, Beren Therapeutics, Esperion, Therapeutics, Genentech, Kancera, Kowa Pharmaceuticals, Medimmune, Merck, Norvo Nordisk, Merck, Novartis, Pfizer, Sanofi-Regeneron. Dr. Libby is a member of scientific advisory boards for Amgen, Corvidia Therapeutics, DalCor Pharmaceuticals, Kowa Pharmaceuticals, Olatec Therapeutics, Medimmune, Novartis, and XBiotech, Inc during the conduct of this study; Dr. Libby's laboratory has received research funding in the last two years from Novartis; Dr. Libby also reports grants from the National Heart, Lung, and Blood Institute, the American Heart Association, the RRM Charitable Fund, and the Simard Fund during the conduct of this study; Dr. Libby is on the Board of Directors of XBiotech, Inc. Dr. Libby has a financial interest in Xbiotech, a company developing therapeutic human antibodies; in addition, Dr. Libby has a patent Use of Canakinumab pending to Ridker P, Thuren T, Bermann G, Libby P, inventors; Dr. Libby's interests were reviewed and are managed by Brigham and Women's Hospital and Partners HealthCare in accordance with their conflict of interest policies; Dr. Sofer has nothing to disclose. Dr. Redline reports grants from NIH during the conduct of the study; grants and personal fees from Jazz Pharma, personal fees from Eisai Inc, personal fees from Apnimed Inc, outside the submitted work; and Dr. Redline is the first incumbent of an endowed professorship donated to the Harvard Medical School by Dr. Peter Farrell, the founder and Board Chairman of ResMed, through a charitable remainder trust instrument, with annual support equivalent to the endowment payout provided to the Harvard Medical School during Dr. Farrell's lifetime by the ResMed Company through an irrevocable gift agreement.

References:

1. Javaheri, S., et al., *Sleep apnea: types, mechanisms, and clinical cardiovascular consequences*. Journal of the American College of Cardiology, 2017. **69**(7): p. 841-858.
2. Bauters, F., et al., *The link between obstructive sleep apnea and cardiovascular disease*. Current atherosclerosis reports, 2016. **18**(1): p. 1.
3. Loke, Y.K., et al., *Association of obstructive sleep apnea with risk of serious cardiovascular events: a systematic review and meta-analysis*. Circulation: Cardiovascular Quality and Outcomes, 2012. **5**(5): p. 720-728.
4. McEvoy, R.D., et al., *CPAP for prevention of cardiovascular events in obstructive sleep apnea*. New England Journal of Medicine, 2016. **375**(10): p. 919-931.
5. Peker, Y., et al., *Effect of positive airway pressure on cardiovascular outcomes in coronary artery disease patients with nonsleepy obstructive sleep apnea. The RICCADSA randomized controlled trial*. American journal of respiratory and critical care medicine, 2016. **194**(5): p. 613-620.
6. Barbé, F., et al., *Effect of continuous positive airway pressure on the incidence of hypertension and cardiovascular events in nonsleepy patients with obstructive sleep apnea: a randomized controlled trial*. Jama, 2012. **307**(20): p. 2161-2168.
7. Khan, S.U., et al., *A meta-analysis of continuous positive airway pressure therapy in prevention of cardiovascular events in patients with obstructive sleep apnoea*. European heart journal, 2018. **39**(24): p. 2291-2297.
8. Peker, Y., K. Wegscheider, and C. Eulenburg, *Reply: Effect of Continuous Positive Airway Pressure Therapy on Cardiovascular Outcomes: Risk Assessment*. American journal of respiratory and critical care medicine, 2017. **196**(5): p. 662-663.
9. Drager, L.F., et al., *Sleep apnea and cardiovascular disease: lessons from recent trials and need for team science*. Circulation, 2017. **136**(19): p. 1840-1850.
10. Pack, A.I., *Application of personalized, predictive, preventative, and participatory (P4) medicine to obstructive sleep apnea. A roadmap for improving care?* Annals of the American Thoracic society, 2016. **13**(9): p. 1456-1467.
11. Bonsignore, M.R., et al., *Personalised medicine in sleep respiratory disorders: focus on obstructive sleep apnoea diagnosis and treatment*. European Respiratory Review, 2017. **26**(146).
12. Schultz, A., et al., *Interindividual heterogeneity in the hypoxic regulation of VEGF: significance for the development of the coronary artery collateral circulation*. Circulation, 1999. **100**(5): p. 547-552.
13. Bielicki, P., et al., *Genetic determinants of cardiovascular disease in patients with obstructive sleep apnea (OSA)*. 2015, Eur Respiratory Soc.
14. Floras, J.S., *Sleep apnea and cardiovascular disease: an enigmatic risk factor*. Circulation research, 2018. **122**(12): p. 1741-1764.
15. Aronson, D., L. Lavie, and P. Lavie, *Does OSA upregulate cardioprotective pathways to an ischemic insult?* Chest, 2018. **153**(2): p. 295-297.
16. Mazzotti, D.R., et al., *Symptom subtypes of obstructive sleep apnea predict incidence of cardiovascular outcomes*. American journal of respiratory and critical care medicine, 2019. **200**(4): p. 493-506.
17. Libby, P. and G.K. Hansson, *From Focal Lipid Storage to Systemic Inflammation: JACC Review Topic of the Week*. J Am Coll Cardiol, 2019. **74**(12): p. 1594-1607.

18. Mann, D.L., et al., *Braunwald's heart disease : a textbook of cardiovascular medicine*. Tenth edition. ed. 2015, Philadelphia, PA: Elsevier/Saunders. volumes.
19. Morrow, D.A., *Myocardial infarction : a companion to Braunwald's heart disease*. 2017, St. Louis, Missouri: Elsevier. xxiii, 468 pages.
20. Dewan, N.A., F.J. Nieto, and V.K. Somers, *Intermittent hypoxemia and OSA: implications for comorbidities*. Chest, 2015. **147**(1): p. 266-274.
21. Ho, D.S.W., et al., *Machine learning SNP based prediction for precision medicine*. Frontiers in Genetics, 2019. **10**: p. 267.
22. Horne, B., et al., *Generating genetic risk scores from intermediate phenotypes for use in association studies of clinically significant endpoints*. Annals of human genetics, 2005. **69**(2): p. 176-186.
23. Dor, Y., R. Porat, and E. Keshet, *Vascular endothelial growth factor and vascular adjustments to perturbations in oxygen homeostasis*. American Journal of Physiology-Cell Physiology, 2001. **280**(6): p. C1367-C1374.
24. Lavie, L., et al., *Plasma vascular endothelial growth factor in sleep apnea syndrome: effects of nasal continuous positive air pressure treatment*. American journal of respiratory and critical care medicine, 2002. **165**(12): p. 1624-1628.
25. Yokoe, T., et al., *Elevated levels of C-reactive protein and interleukin-6 in patients with obstructive sleep apnea syndrome are decreased by nasal continuous positive airway pressure*. Circulation, 2003. **107**(8): p. 1129-1134.
26. Htoo, A.K., et al., *Activation of nuclear factor κ B in obstructive sleep apnea: a pathway leading to systemic inflammation*. Sleep and Breathing, 2006. **10**(1): p. 43-50.
27. Ryan, S., C.T. Taylor, and W.T. McNicholas, *Predictors of elevated nuclear factor- κ B-dependent genes in obstructive sleep apnea syndrome*. American journal of respiratory and critical care medicine, 2006. **174**(7): p. 824-830.
28. Jelic, S. and T.H. Le Jemtel, *Inflammation, oxidative stress, and the vascular endothelium in obstructive sleep apnea*. Trends in cardiovascular medicine, 2008. **18**(7): p. 253-260.
29. Jelic, S., et al., *Vascular inflammation in obesity and sleep apnea*. Circulation, 2010. **121**(8): p. 1014.
30. Xi, L. and T.V. Serebrovskaya, *Intermittent hypoxia and human diseases*. 2012: Springer Science & Business Media.
31. Zhang, X.-B., et al., *Vascular endothelial growth factor levels in patients with obstructive sleep apnea: a meta-analysis*. European Archives of Oto-Rhino-Laryngology, 2017. **274**(2): p. 661-670.
32. Labarca, G., et al., *Chronic intermittent hypoxia in obstructive sleep apnea: a narrative review from pathophysiological pathways to a precision clinical approach*. Sleep and Breathing, 2019: p. 1-10.
33. Sudlow, C., et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age*. PLoS medicine, 2015. **12**(3).
34. Bycroft, C., et al., *The UK Biobank resource with deep phenotyping and genomic data*. Nature, 2018. **562**(7726): p. 203-209.
35. Denny, J.C., et al., *PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations*. Bioinformatics, 2010. **26**(9): p. 1205-1210.

36. Khera, A.V., et al., *Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations*. Nature genetics, 2018. **50**(9): p. 1219-1224.
37. Vilhjálmsón, B.J., et al., *Modeling linkage disequilibrium increases accuracy of polygenic risk scores*. The american journal of human genetics, 2015. **97**(4): p. 576-592.
38. Nikpay, M., et al., *A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease*. Nature genetics, 2015. **47**(10): p. 1121.
39. Frankish, A., et al., *GENCODE reference annotation for the human and mouse genomes*. Nucleic acids research, 2019. **47**(D1): p. D766-D773.
40. Fantom Consortium, *A promoter-level mammalian expression atlas*. Nature, 2014. **507**(7493): p. 462.
41. Javierre, B.M., et al., *Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters*. Cell, 2016. **167**(5): p. 1369-1384. e19.
42. GTEx Consortium, J., 2016 #64}, *Genetic effects on gene expression across human tissues*. Nature, 2017. **550**(7675): p. 204-213.
43. Fishilevich, S., et al., *GeneHancer: genome-wide integration of enhancers and target genes in GeneCards*. Database, 2017. **2017**.
44. Appelman, Y., et al., *Sex differences in cardiovascular risk factors and disease prevention*. Atherosclerosis, 2015. **241**(1): p. 211-218.
45. Lebkuchen, A., et al., *Advances and challenges in pursuing biomarkers for obstructive sleep apnea: implications for the cardiovascular risk*. Trends in cardiovascular medicine, 2020.
46. Sánchez-de-la-Torre, M., et al., *Precision medicine in patients with resistant hypertension and obstructive sleep apnea: blood pressure response to continuous positive airway pressure treatment*. Journal of the American College of Cardiology, 2015. **66**(9): p. 1023-1032.
47. Lavie, L., *Intermittent hypoxia and obstructive sleep apnea: mechanisms, interindividual responses and clinical insights*, in *Atherosclerosis, Arteriosclerosis and Arteriolosclerosis*. 2019, IntechOpen.
48. Dengler, V.L., M.D. Galbraith, and J.M. Espinosa, *Transcriptional regulation by hypoxia inducible factors*. Critical reviews in biochemistry and molecular biology, 2014. **49**(1): p. 1-15.
49. Gottlieb, D.J., et al., *Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure: the sleep heart health study*. Circulation, 2010. **122**(4): p. 352-360.
50. Rahaghi, F. and R.C. Basner, *Delayed diagnosis of obstructive sleep apnea: don't ask, don't tell*. Sleep and Breathing, 1999. **3**(4): p. 119-124.

1 **Table 1. Sample characteristics**

	OSA control		OSA case	
N	471,877		4,974	
Outcome:				
CAD, N (%)	25,988	(5.50)	711	(14.30)
Covariates:				
BMI (median [IQR])	26.67	(24.09, 29.79)	31.61	(28.11, 36.39)
Age (median [IQR])	58	(50.00, 63.00)	58	(51.00, 63.00)
Male (%)	213,700	(45.30)	3,622	(72.80)
Smoking (%)				
current	36,568	(7.70)	519	(10.40)
never	259,572	(55.00)	2,162	(43.50)
occasionally	12,484	(2.60)	183	(3.70)
previous	163,253	(34.60)	2,110	(42.40)
Self-reported white (%)	397,269	(84.20)	4,054	(81.50)
Genetic PC1 (mean (SD))	-1.53	(53.33)	1.95	(60.96)
Genetic PC2 (mean (SD))	0.5	(27.38)	0.21	(27.29)
Genetic PC3 (mean (SD))	-0.17	(14.61)	0.82	(15.30)
Genetic PC4 (mean (SD))	0.13	(10.32)	-0.2	(12.07)
Genetic PC5 (mean (SD))	0.05	(7.57)	0.28	(7.42)
BiLEVE genotype platform (%)	48,528	(10.30)	619	(12.40)
COPD (%)	18,888	(4.00)	608	(12.20)
Asthma (%)	62,630	(13.30)	1,126	(22.60)
Standardized PRS and PS-PRS Exposures:				
KEGG hsa04066 HIF1 (mean (SD))	-0.0002	1.000	-0.011	0.985
KEGG hsa04370 VEGF (mean (SD))	-0.0014	1.000	0.029	0.982
KEGG hsa04064 NFκB (mean (SD))	0.0002	1.000	0.003	1.005
KEGG hsa04668 TNF (mean (SD))	0.0002	1.000	0.015	1.017
HIF1 core genes (mean (SD))	0.0018	1.000	-0.016	0.999
VEGF core genes (mean (SD))	0.0003	1.000	-0.010	0.991
NFκB core genes (mean (SD))	0.0011	1.000	-0.015	1.006
TNF core genes (mean (SD))	-0.0009	1.000	0.001	0.992
CAD PRS (mean (SD))	0.0000	0.998	0.034	1.015
OSA genes (mean (SD))	-0.0016	0.999	-0.005	0.997

1 Outcome, covariate and polygenic risk score distributions by OSA status. OSA: Obstructive
2 sleep apnea, CAD: coronary artery disease, BMI: body mass index, PC: principal component,
3 COPD: chronic obstructive pulmonary disease, KEGG: Kyoto encyclopedia of genes and
4 genomes, hsa: Homo sapiens pathway reference number, HIF1: hypoxia inducible factor 1,
5 VEGF: vascular endothelial growth factor, NFκB: nuclear factor kappa- beta, TNF: tumor
6 necrosis factor, PRS: polygenic risk score, PS-PRS: pathway-specific polygenic risk score.
7

1 Key Figure:

2

3 Estimated qualitative effect-measure modification by OSA on genetic risk of CAD, using

4 pathway-specific polygenic risk scores (PS-PRS) and gene-specific risk scores (GSRS), for the

5 given pathways and genes.²

² *HIF1α* and *ARNT* constitute the HIF1 core-genes module and had estimated risk-amplifying effect modification at nominal significance in primary analysis. HIF1 KEGG pathway not shown, but encompasses HIF1 core-genes as well as additional genes and modules shown in the graphic, and its estimated PS-PGRS effect modification did not meet nominal significance in primary analysis. Bonferroni-corrected significance level of 0.005 of based on 10 hypothesis tests in primary analysis and nominal level 0.05.

1 **Figure 1: Flow diagram showing the gene annotation process**

2

3

1 **Figure 2: Pairwise correlations among selected pathway-specific polygenic risk scores**

2

3 Pearson correlations between participants' primary analysis risk scores.

Table 2: Primary analysis of selected OSA pathways: Covariate-adjusted odds of CAD per s.d. of standardized PS-PRSs with effect modification by OSA status

PS-PRS	OSA effect		PS-PRS effect in OSA Controls			PS-PRS effect in OSA Cases			OSA x PS-PRS interaction effect	
	OR	CI	OR	CI	p-value	OR	CI	p-value	OR	p-value
A) Selected pathways: core-gene modules										
HIF1 core genes	1.50	(1.38, 1.64)	1.022	(1.009, 1.035)	1.06E-03	1.114	(1.025, 1.212)	1.10E-02	1.09	4.46E-02*
VEGF core genes	1.51	(1.38, 1.64)	1.032	(1.019, 1.045)	1.66E-06	0.952	(0.874, 1.037)	2.58E-01	0.92	6.70E-02
NFκB core genes	1.51	(1.38, 1.64)	1.031	(1.018, 1.045)	4.45E-06	0.981	(0.903, 1.066)	6.54E-01	0.95	2.45E-01
TNF core genes	1.51	(1.38, 1.64)	1.019	(1.006, 1.033)	4.44E-03	1.013	(0.931, 1.102)	7.59E-01	0.99	8.94E-01
B) Selected pathways: KEGG signaling pathways										
HIF1 KEGG	1.51	(1.38, 1.64)	1.062	(1.048, 1.076)	6.81E-20	1.089	(1.001, 1.185)	4.84E-02	1.03	5.66E-01
VEGF KEGG	1.51	(1.38, 1.64)	1.040	(1.027, 1.054)	4.30E-09	0.894	(0.821, 0.974)	1.02E-02	0.86	5.98E-04**
NFκB KEGG	1.51	(1.38, 1.64)	1.058	(1.044, 1.072)	3.26E-17	0.998	(0.919, 1.085)	9.71E-01	0.94	1.74E-01
TNF KEGG	1.51	(1.38, 1.64)	1.052	(1.038, 1.065)	1.65E-14	0.994	(0.916, 1.079)	8.86E-01	0.95	1.85E-01
C) Selected aggregated pathway comparators										
CAD PRS	1.52	(1.39, 1.67)	1.695	(1.672, 1.719)	<1.0E-99	1.664	(1.525, 1.816)	3.30E-30	0.98	6.80E-01
OSA genes	1.50	(1.37, 1.63)	1.102	(1.087, 1.116)	1.17E-46	1.222	(1.122, 1.330)	3.81E-06	1.11	1.83E-02*

In Table 2, each row depicts a separate PS-PRS model showing the OSA and PS-PRS main effects and the PS-PRS x OSA interaction effect in a model of the form $\text{logit}(\text{CAD}) = \text{OSA} + \text{PS-PRS} + \text{PS-PRS} \times \text{OSA} + \text{covariates}$. Three groupings representing different types of pathways are included: A) core-gene modules B) KEGG-defined signaling pathways, and C) aggregate pathways. These ten pathways are allocated Type-1 error in the primary analysis. Nominally significant (*) GxE results do not pass Bonferroni-corrected significance (**) level 5E-3. Covariates adjusted in a logistic generalized additive model: mutual interaction of age and BMI by sex (via tensor-product splines) as well as smoking and its interaction with sex, self-reported

- 1 white race, the first 5 genetic PCs, genotype platform (BiLEVE), asthma and chronic obstructive
- 2 pulmonary disease. OR: odds ratio; CI: confidence interval; OSA: Obstructive sleep apnea,
- 3 CAD: coronary artery disease, KEGG: Kyoto encyclopedia of genes and genomes, HIF1:
- 4 hypoxia inducible factor 1, VEGF: vascular endothelial growth factor, NFκB: nuclear factor
- 5 kappa- beta, TNF: tumor necrosis factor, PRS: polygenic risk score, PS-PRS: pathway-specific
- 6 polygenic risk score.

1 **Table 3: Secondary Analysis: Genes and Submodules of selected CAD pathways:**
2 **Covariate-adjusted odds of CAD per s.d. of standardized PS-PRSs, with effect modification**
3 **by OSA status**

A) Core-gene modules with all constituent genes										
	OSA		PS-PRS in OSA Controls			PS-PRS in OSA Cases			OSA x PS-PRS Interaction	
	OR	CI	OR	CI	p-value	OR	CI	p-value	OR	p-value
<i>HIF1 core genes</i>	1.50	(1.38, 1.64)	1.022	(1.009, 1.035)	1.06E-03	1.114	(1.025, 1.212)	1.10E-02	1.09	4.46E-02*
<i>HIF1A</i>	1.51	(1.38, 1.64)	1.005	(0.992, 1.018)	4.59E-01	0.978	(0.899, 1.063)	5.95E-01	0.973	5.21E-01
<i>ARNT</i>	1.50	(1.38, 1.64)	1.022	(1.009, 1.035)	1.10E-03	1.115	(1.025, 1.212)	1.08E-02	1.091	4.36E-02*
<i>VEGF core genes</i>	1.51	(1.38, 1.64)	1.032	(1.019, 1.045)	1.66E-06	0.952	(0.874, 1.037)	2.58E-01	0.92	6.70E-02
<i>VEGFA</i>	1.51	(1.38, 1.64)	1.018	(1.005, 1.031)	7.36E-03	0.941	(0.866, 1.022)	1.50E-01	0.924	6.52E-02
<i>FLT1</i>	1.51	(1.38, 1.64)	1.030	(1.017, 1.043)	7.73E-06	0.961	(0.882, 1.046)	3.56E-01	0.933	1.14E-01
<i>KDR</i>	1.51	(1.38, 1.64)	1.002	(0.989, 1.015)	7.56E-01	0.972	(0.894, 1.057)	5.09E-01	0.970	4.83E-01
<i>NFκB core genes</i>	1.51	(1.38, 1.64)	1.031	(1.018, 1.045)	4.45E-06	0.981	(0.903, 1.066)	6.54E-01	0.95	2.45E-01
<i>NFKB1A</i>	1.51	(1.38, 1.64)	1.005	(0.992, 1.018)	4.69E-01	1.010	(0.927, 1.100)	8.28E-01	1.005	9.16E-01
<i>NFKB1</i>	1.51	(1.38, 1.64)	1.006	(0.993, 1.019)	4.05E-01	0.938	(0.862, 1.019)	1.30E-01	0.932	1.05E-01
<i>RELA</i>	1.51	(1.38, 1.64)	1.025	(1.012, 1.039)	2.20E-04	1.003	(0.923, 1.089)	9.51E-01	0.978	6.04E-01
<i>NFKB2</i>	1.51	(1.38, 1.64)	1.018	(1.004, 1.031)	9.02E-03	1.076	(0.990, 1.170)	8.51E-02	1.058	1.95E-01
<i>RELB</i>	1.50	(1.38, 1.64)	1.015	(1.001, 1.028)	2.89E-02	0.914	(0.842, 0.992)	3.16E-02	0.901	1.37E-02*
<i>TNF core genes</i>	1.51	(1.38, 1.64)	1.019	(1.006, 1.033)	4.44E-03	1.013	(0.931, 1.102)	7.59E-01	0.99	8.94E-01
<i>TNF</i>	1.51	(1.38, 1.64)	1.013	(1.000, 1.027)	4.59E-02	1.022	(0.941, 1.110)	6.07E-01	1.009	8.43E-01
<i>TNFRSF1A</i>	1.51	(1.38, 1.64)	1.014	(1.000, 1.027)	4.32E-02	1.008	(0.927, 1.097)	8.46E-01	0.995	9.05E-01
<i>TNFRSF1B</i>	1.51	(1.38, 1.64)	1.014	(1.001, 1.027)	3.53E-02	1.010	(0.928, 1.098)	8.24E-01	0.996	9.18E-01

4

B) KEGG pathway sub-modules with suggestive GxE										
	OSA		PS-PRS in OSA Controls			PS-PRS in OSA Cases			OSA x PS-PRS Interaction	
	OR	CI	OR	CI	p-value	OR	CI	p-value	OR	p-value
HIF1 module: IH-induced HIF1- α degradation	1.51	(1.38, 1.64)	1.032	(1.019, 1.045)	1.82E-06	0.928	(0.853, 1.009)	8.05E-02	0.899	1.42E-02*
HIF1 module: angiogenesis	1.51	(1.38, 1.64)	1.032	(1.019, 1.046)	1.21E-06	0.947	(0.869, 1.032)	2.18E-01	0.918	5.25E-02
VEGF module: endothelial migration	1.51	(1.39, 1.65)	1.014	(1.001, 1.027)	3.66E-02	0.881	(0.810, 0.958)	3.05E-03	0.869	1.16E-03**
VEGF module: endothelial permeability	1.51	(1.38, 1.64)	1.020	(1.007, 1.034)	3.20E-03	0.913	(0.838, 0.994)	3.63E-02	0.895	1.17E-02*
VEGF module: endothelial survival	1.51	(1.38, 1.64)	1.013	(1.000, 1.027)	4.47E-02	0.911	(0.837, 0.990)	2.88E-02	0.899	1.35E-02*
HIF1&VEGF: shared PI3K/AKT module	1.51	(1.38, 1.64)	1.019	(1.006, 1.033)	4.79E-03	0.912	(0.837, 0.992)	3.23E-02	0.894	1.08E-02*

1

C) Top five genes from KEGG HIF1, VEGF, NFkB or TNF pathways suggestive of risk-inverting GxE										
	OSA		PS-PRS in OSA Controls			PS-PRS in OSA Cases			OSA x PS-PRS Interaction	
	OR	CI	OR	CI	p-value	OR	CI	p-value	OR	p-value
<i>RELB</i>	1.50	(1.38, 1.64)	1.015	(1.001, 1.028)	2.89E-02	0.914	(0.842, 0.992)	3.16E-02	0.901	1.37E-02*
<i>PIK3R3</i>	1.50	(1.38, 1.64)	1.02	(1.007, 1.033)	2.51E-03	0.928	(0.843, 1.021)	1.27E-01	0.910	5.63E-02
<i>IRF1</i>	1.51	(1.38, 1.64)	1.026	(1.013, 1.040)	9.50E-05	0.944	(0.869, 1.025)	1.71E-01	0.919	4.99E-02
<i>VEGFA</i>	1.51	(1.38, 1.64)	1.018	(1.005, 1.031)	7.36E-03	0.941	(0.866, 1.022)	1.50E-01	0.924	6.52E-02
<i>PLCG1</i>	1.51	(1.38, 1.64)	1.023	(1.010, 1.037)	4.25E-04	0.954	(0.877, 1.038)	2.71E-01	0.932	1.06E-01
D) Top five genes from KEGG HIF1, VEGF, NFkB or TNF pathways suggestive of risk-amplifying GxE										
	OSA		PS-PRS in OSA Controls			PS-PRS in OSA Cases			OSA x PS-PRS Interaction	
	OR	CI	OR	CI	p-value	OR	CI	p-value	OR	p-value
<i>ARNT</i>	1.50	(1.38, 1.64)	1.022	(1.009, 1.035)	1.10E-03	1.115	(1.025, 1.212)	1.08E-02	1.091	4.36E-02*
<i>MAPK1</i>	1.50	(1.38, 1.64)	1.015	(1.002, 1.028)	2.70E-02	1.095	(1.007, 1.191)	3.44E-02	1.079	7.92E-02
<i>PDHB</i>	1.50	(1.38, 1.64)	1.015	(1.002, 1.028)	2.69E-02	1.088	(1.000, 1.183)	5.05E-02	1.072	1.11E-01
<i>NFKB2</i>	1.51	(1.38, 1.64)	1.018	(1.004, 1.031)	9.02E-03	1.076	(0.990, 1.170)	8.51E-02	1.058	1.95E-01
<i>IL6R</i>	1.51	(1.38, 1.64)	1.033	(1.020, 1.046)	9.71E-07	1.091	(1.002, 1.187)	4.45E-02	1.056	2.13E-01

Table 3. Analysis of module and gene-level estimates of genetic effects in OSA cases and controls, presented as a secondary post-hoc analysis suggestively localizing primary analysis results. In section A, each core-gene module, defined as the nominal gene and its receptors, is accompanied by each of its constituent genes; in section B, each KEGG pathway is accompanied with those submodules that showed nominally significant main and interaction effects; in section C, genes with the largest risk-inverting and risk-amplifying estimated interaction effects, having nominally significant main effects, are presented. Core-gene modules and KEGG pathways were allocated type I error in the primary analysis. Secondary analysis is underpowered to comprehensively identify driver genes and modules. Nominally significant GxE results (*) do

1 not pass Bonferroni (**) thresholds for multiple testing for 30 pathway modules ($1.67E-3$), and
2 268 genes ($1.87E-4$) with membership in the KEGG HIF1, VEGF, NF κ B or TNF pathways.
3 Comprehensive secondary analysis gene- and module-level results for these KEGG pathways are
4 presented in Supplementary Table S4. OR: odds ratio; CI: confidence interval; OSA: Obstructive
5 sleep apnea, CAD: coronary artery disease, KEGG: Kyoto encyclopedia of genes and genomes,
6 HIF1: hypoxia inducible factor 1, VEGF: vascular endothelial growth factor, NF κ B: nuclear
7 factor kappa- beta, TNF: tumor necrosis factor, PRS: polygenic risk score, PS-PRS: pathway-
8 specific polygenic risk score.
9

Methods Supplement:

Selecting genomic variants and annotating genes: We obtained 6.6 million common SNVs (minor allele frequency > 0.01) available in UK Biobank imputed data which passed quality control filters in a prior GWAS and subsequent PRS analysis of CAD conducted by Khera et al. These SNVs were then given genetic annotations using a sequential multi-stage methodology as described above in the text and shown in Figure 1.

We ordered the quality of the annotation resources as follows: GenCode v34 exons [39], FANTOM5 [40], Promoter-Capture HiC experiments involving white blood cells as described in [41], GTEx v7 statistically significant eQTLs [42] derived from artery (aorta, coronary and tibial), heart (atrial appendage and left ventricle) and whole blood tissues, GeneHancer v4.4 [43] predicted gene annotations passing minimum enhancer score threshold of 0.3 and minimum gene score threshold of 4.0, and finally GenCode protein-coding regions (including all loci from 20kbp upstream of transcription start site to 10kbp downstream of transcription stop site), and noncoding transcript sequence regions. Within a single data set a SNV may be assigned to multiple genes, in which case both annotations are accepted.

Calculating genetic risk scores: After defining pathways and associated SNVs, we calculated a PS-PRS for each participant using a SNV-set representing each of the ten pathways (HIF-1, VEGF, NFκB, TNF represented by both a core-gene module and a KEGG pathway, the curated CAD/OSA pathway, and the full PRS). For each PS-PRS, we took the subset of SNVs annotated to that pathway and calculated a PS-PRS using the corresponding genetic effects from the Khera et al PRS. We standardized polygenic risk scores including individual PS-PRSs to mean zero and variance one in the total UKBB study population. To facilitate our secondary

analysis, we additionally calculated GSRS for each gene in the HIF-1, VEGF, NFκB, and TNF KEGG pathways.

Modeling and testing associations: For each PRS we created a single CAD outcome model and tested: A) the association of the PS-PRS with CAD, and B) the interaction of the PS-PRS with OSA status, controlling for demographics and comorbidities. We modeled the binary CAD outcome via generalized additive models (GAM) with logistic link, controlling for the covariates discussed above [44].

Statistical significance was assessed at level $\alpha = 0.05$. For our primary analysis testing interaction of PS-PRS and OSA status, we adopted the Bonferroni-adjusted significance threshold $\alpha = 0.005$ on ten interaction tests. In secondary analysis of GSRS, we report those genes that had risk-amplifying main effects on CAD (in OSA controls) and passed an $\alpha = 0.05$ main effect significance threshold. Statistical calculations were conducted using *R* version 3.5.2, and code used to perform these analyses will be posted to a publicly accessible repository (https://github.com/MatthewOGoodman/UKB_PSRxOSA_GxE_study_of_CAD).

1 **Supplementary Table S1:** Phenotype definitions and data sources

Phenotype	UKBB field: Code	Description
CAD:	UKBB 20002: 1075	Self-reported heart attack / MI
	UKBB 41202, 41204 derived phecode: 411.2	Myocardial infarction
	UKBB 42001	Myocardial infarction: algorithmically-defined
Sleep Apnea:	UKBB 41202, 41204 ICD10: G47.3	
	UKBB 20002: 1123	sleep apnoea: self-report of doctor diagnosed sleep apnea documented during an interview with a trained medical professional
Inclusion/Exclusion:		
Snoring	UKBB 1210	Self-reported snoring
Daytime sleepiness	UKBB 1220	Self-reported daytime sleepiness
Covariates:		
Sex	UKBB 22001	Genetic sex
BMI	UKBB 21001	Body Mass Index
Age	UKBB 21022	Age at recruitment
Smoking status	UKBB 1239	Current Tobacco Smoking
	UKBB 1249	Past Tobacco Smoking
	UKBB 20002: 1112	COPD: self-report

COPD (chronic obstructive pulmonary disease)	UKBB 41202, 41204 derived phecodes: 496, 496.1, 496.2, or 496.21	Chronic airway obstruction, Emphysema, Chronic bronchitis, Obstructive chronic bronchitis, Bronchiectasis
Asthma	UKBB 20002: 1111	Asthma
	UKBB 41202, 41204 derived phecodes: 495, 495.1, 495.11, or 495.2	Asthma, chronic obstructive asthma, asthma with exacerbation, chronic obstructive asthma with exacerbation
UKBB BiLEVE genotype platform	UKBB 22000	Binary covariate: whether the participant was genotyped on the pilot study BiLEVE genotype platform, or the UKBB Axiom platform.
Potentially Mediating Covariates:		
Hypertension	UKBB 20002: 1065 Derived phecode: 401	Hypertension
Likely Type 1 Diabetes	Algorithm of Eastwood et al.	Type 1 Diabetes
Likely Type 2 Diabetes	Algorithm of Eastwood et al.	Type 2 Diabetes
Pulmonary fibrosis	UKBB 20002: 1121	Pulmonary fibrosis
	UKBB 41202, 41204 derived phecodes: 502, 504, or 504.1	Postinflammatory pulmonary fibrosis, Other alveolar and parietoalveolar pneumonopathy, Idiopathic fibrosing alveolitis

- 1 Daytime sleepiness was recoded as “present” (moderate or severe) or “absent” (none or mild).
- 2 Smoking status was recoded as “never smoked,” “past smoker,” “current smoker
- 3 (occasionally),” and “current smoker (most days).” Diagnostic information for medical diagnoses
- 4 was obtained from the following UKBB data fields: Non-cancer illness, self-reported (20002),

- 1 Hospital Episode Inpatient data: Diagnoses - main ICD10 (41202), Diagnoses - secondary
- 2 ICD10 (41204), ICD10 codes were collapsed to medically interpretable groupings using the
- 3 phecode system using Phecode Map version 1.2 for the World Health Organization ICD10
- 4 codings.

Supplementary Table S2: Pathway and module definitions

Pathway Gene Membership

a. KEGG signaling pathways

i. HIF1 https://www.genome.jp/kegg-bin/show_pathway?hsa04066

ii. VEGF https://www.genome.jp/kegg-bin/show_pathway?hsa04370

iii. TNF https://www.genome.jp/kegg-bin/show_pathway?hsa04668

iv. NFκB https://www.genome.jp/kegg-bin/show_pathway?hsa04064

b. Core Gene modules:

i. HIF1 pathway core-gene module: *HIF1A*, *ARNT*

ii. VEGF pathway core-gene module: *VEGFA*, *FLT*, *KDR*

iii. TNF pathway core-gene module: *TNF*, *TNFRSF1A*, *TNFRSF1B*

iv. NFκB pathway core-gene module: *NFKB1A*, *NFKB1*, *RELA*, *NFKB2*, *RELB*

c. Curated OSA/CAD pathway:

(225 genes related to OSA and CAD, selected by literature review):

ACE, *ACPI*, *ACP2*, *ACP5*, *AGT*, *AGTR1*, *AKT1*, *AKT1S1*, *AKT2*, *AKT3*, *ALOX12*,

ALOX12B, *ALOX15*, *ALOX15B*, *ALOX5*, *APOE*, *ARG1*, *ARG2*, *ATF1*, *ATF2*, *ATF3*,

ATF4, *ATF5*, *ATF6*, *ATF6B*, *ATF7*, *BCS1L*, *CALM1*, *CAT*, *CAV1*, *CDH5*, *CUL3*,

CXCL8, *CXCR2*, *CYC1*, *DEPTOR*, *ENPP1*, *ENPP3*, *EPO*, *FGA*, *FGB*, *FGG*, *FLAD1*,

FOS, *FOSB*, *FOXO1*, *GCH1*, *GCLC*, *GCLM*, *GLRX*, *GPX1*, *GPX3*, *GPX4*, *GSR*,

GSS, *GSTA1*, *GSTA2*, *GSTA3*, *GSTA4*, *GSTA5*, *GSTK1*, *GSTM1*, *GSTM2*, *GSTM3*,

GSTM4, *GSTM5*, *GSTO1*, *GSTO2*, *GSTP1*, *GSTT2*, *GSTZ1*, *HES1*, *HES5*, *HEY1*,

HEY2, HIF1A, HMOX1, HSP90AA1, HSP90AB1, ICAM1, IL1B, IL1R1, IL1RN, IL6, IL6R, IL6ST, ITGA1, ITGA10, ITGA11, ITGA2, ITGA2B, ITGA3, ITGA4, ITGA5, ITGA6, ITGA7, ITGA8, ITGA9, ITGAD, ITGAE, ITGAL, ITGAM, ITGAV, ITGAX, ITGB1, ITGB2, ITGB3, ITGB4, ITGB5, ITGB6, ITGB7, ITGB8, JUN, JUNB, JUND, KDR, KEAP1, LYRM7, MAFF, MAFG, MAFK, MAPK1, MAPK10, MAPK11, MAPK12, MAPK13, MAPK14, MAPK15, MAPK3, MAPK4, MAPK6, MAPK7, MAPK8, MAPK9, MAPKAP1, MEF2A, MEF2C, MEF2D, MGST1, MGST3, MLST8, MPO, MTOR, NDUFA1, NDUFA10, NDUFA11, NDUFA12, NDUFA2, NDUFS1, NDUFS2, NDUFS3, NDUFS4, NDUFS6, NDUFS7, NDUFV1, NDUFV2, NFE2L2, NFKB1, NFKB2, NKRF, NOS2, NOS3, NOX1, NOX3, NOX4, NOX5, NQO1, PIK3C2B, PIK3C2G, PIK3CA, PIK3CB, PIK3CD, PIK3CG, PIK3R2, PIK3R3, PIK3R4, PIK3R5, PIK3R6, PRKAA1, PRKAA2, PRKAB1, PRKAB2, PRKAG1, PRKAG2, PRKAG3, PRKCA, PRKCB, PRR5, PTS, RFK, RHOA, RICTOR, ROCK1, ROCK2, RPTOR, SELE, SELL, SELP, SELPLG, SOD1, SOD2, SOD3, SPR, SREBF1, SREBF2, TELO2, TNF, TTC19, TTII, TXN, TXN2, TXNRD1, TXNRD2, TXNRD3, UQCC2, UQCC3, UQCRB, UQCRC2, UQCRQ, VCAM1, VEGFA, XDH

d. Submodules

1. KEGG HIF1 GF RTK HIF1a translation: *EGF, EGFR, ERBB2, IGF1, IGF1R, INS, INSR.*
2. KEGG HIF1 MTOR HIF1a translation: *EIF4E, EIF4E1B, EIF4E2, EIF4EBP1, MTOR, RPS6, RPS6KB1, RPS6KB2.*

- 1 **3.** KEGG HIF1 IH HIF1a degradation: *CYBB, EGLN1, EGLN2, EGLN3,*
2 *PLCG1, PLCG2, PRKCA.*
- 3 **4.** KEGG HIF1 HIF1a degradation ubiquitin: *CUL2, ELOB, ELOC, RBX1, VHL.*
- 4 **5.** KEGG HIF1 IH HIF1b activation: *ARNT, CAMK2A, CAMK2B, CAMK2D,*
5 *CAMK2G, CREBBP, CYBB, EP300, PLCG1, PLCG2.*
- 6 **6.** KEGG HIF1 RBC: *EPO, TF, TFRC.*
- 7 **7.** KEGG HIF1 angiogenesis: *ANGPT1, ANGPT2, ANGPT4, EGF, FLT1,*
8 *SERPINE1, TEK, TIMP1, VEGFA.*
- 9 **8.** KEGG HIF1 vascular tone: *EDN1, HMOX1, NOS2, NOS3, NPPA.*
- 10 **9.** KEGG HIF1 glucose inhibit citrate: *PDHA1, PDHA2, PDHB, PDK1,*
11 *SLC2A1.*
- 12 **10.** KEGG HIF1 glucose promote anaerobic metabolism: *ALDOA, ENO1, ENO2,*
13 *ENO3, ENO4, GAPDH, HK2, HK3, HKDC1, LDHA, PFKFB3, PFKL, PGK1.*
- 14 **11.** KEGG HIF1 glucose regulate proliferation: *BCL2, CDKN1A, CDKN1B.*
- 15 **12.** KEGG VEGF proliferation: *HRAS, KRAS, MAP2K1, MAP2K2, MAPK1,*
16 *MAPK3, NRAS, PLCG1, PLCG2, PRKCA, PRKCB, PRKCG, RAF1, SPHK1,*
17 *SPHK2.*
- 18 **13.** KEGG VEGF PGI2 production: *NFATC2, PPP3CA, PPP3CB, PPP3CC,*
19 *PPP3R1, PPP3R2, PTGS2.*
- 20 **14.** KEGG VEGF migration: *CDC42, MAPK11, MAPK12, MAPK13, MAPK14,*
21 *MAPKAPK2, MAPKAPK3, PTK2, PXN.*
- 22 **15.** KEGG VEGF endothelial permeability: *AKT1, AKT2, AKT3, NOS3, PIK3CA,*
23 *PIK3CB, PIK3CD, PIK3R1, PIK3R2, PIK3R3, RAC1, RAC2, RAC3, SRC.*

- 1 **16.** KEGG VEGF endothelial survival: *AKT1, AKT2, AKT3, BAD, CASP9,*
2 *PIK3CA, PIK3CB, PIK3CD, PIK3R1, PIK3R2, PIK3R3, SRC.*
- 3 **17.** KEGG HIF1 VEGF shared MAPK genes: *MAP2K1, MAP2K2, MAPK1,*
4 *MAPK3, MKNK1, MKNK2.*
- 5 **18.** KEGG HIF1&VEGF shared PI3K.AKT genes: *AKT1, AKT2, AKT3, PIK3CA,*
6 *PIK3CB, PIK3CD, PIK3R1, PIK3R2, PIK3R3.*
- 7 **19.** KEGG HIF1 NFkB-related genes: *IFNG, IFNGR1, IFNGR2, IL6, IL6R,*
8 *STAT3, TLR4.*
- 9 **20.** KEGG NFkB atypical pathway: *BCL2A1, BCL2L1, CSNK2A1, CSNK2A2,*
10 *CSNK2A3, CSNK2B, EGF, EGFR, ERBB2, NFKB1, NFKBIA, PLAUG, RELA.*
- 11 **21.** KEGG NFkB atypical upstream: *BAG4, IL1B, IL1R1, IRAK1, IRAK4,*
12 *MYD88, RIPK1, TNF, TNFRSF11A, TNFRSF1A, TNFSF11, TRADD, TRAF2,*
13 *TRAF5, TRAF6.*
- 14 **22.** KEGG TNF pathway JNK: *FOS, JUN, MAP2K4, MAP2K7, MAP3K7,*
15 *MAPK10, MAPK8, MAPK9, TAB1, TAB2.*
- 16 **23.** KEGG TNF pathway p38: *ATF2, ATF4, ATF6B, CREB1, CREB3, CREB3L1,*
17 *CREB3L2, CREB3L3, CREB3L4, CREB5, MAP2K3, MAP2K6, MAP3K7,*
18 *MAPK14, RPS6KA4, RPS6KA5, TAB1, TAB2.*
- 19 **24.** KEGG NFkB TNF pathway: *BIRC2, BIRC3, CHUK, IKBKB, IKBKG,*
20 *MAP3K14, MAP3K7, NFKB1, NFKBIA, RELA, RIPK1, TAB1, TAB2, TAB3,*
21 *TNF, TNFRSF1A, TRADD, TRAF2, TRAF5.*
- 22 **25.** KEGG NFkB pathway survival: *BCL2, BCL2A1, BCL2L1, CFLAR,*
23 *GADD45B, TRAF1, XIAP.*

1 Supplementary Figure S3: Correlations between pathways, modules and genes

2
3 Figure S3a) Pearson correlations between core-gene PS-PRSs and their constituent gene-specific
4 risk scores (GSRs). Components genes are listed below each core-genes pathway.

5

1
2 Figure S3b) Pearson correlations between selected HIF1 and VEGF sub-module, KEGG and
3 core-genes PS-PRSs in the secondary analysis. Pathway definitions are shown in Supplementary
4 Table S2.

1 **Supplementary Table S4:** Effect estimates of covariates for model of full PRS

	Est.	s.e.	p-value	OR	OR CI
(Intercept)	-3.49436	0.044507	0.00E+00	0.0304	(0.028, 0.033)
sex (male)	1.04334	0.045079	1.64E-118	2.8387	(2.599, 3.101)
smoking (never)	-0.93997	0.042274	1.56E-109	0.3906	(0.360, 0.424)
smoking (occasionally)	-0.31976	0.090599	4.16E-04	0.7263	(0.608, 0.867)
smoking (previous)	-0.63896	0.043239	2.05E-49	0.5278	(0.485, 0.575)
white	0.06102	0.025926	1.86E-02	1.0629	(1.010, 1.118)
Genetic PC1 (mean (SD))	0.00242	0.000192	2.71E-36	1.0024	(1.002, 1.003)
Genetic PC2 (mean (SD))	-0.00078	0.000318	1.41E-02	0.9992	(0.999, 1.000)
Genetic PC3 (mean (SD))	0.0072	0.000527	1.29E-42	1.0072	(1.006, 1.008)
Genetic PC4 (mean (SD))	0.00173	0.000691	1.23E-02	1.0017	(1.000, 1.003)
Genetic PC5 (mean (SD))	0.00857	0.000916	8.61E-21	1.0086	(1.007, 1.010)
BiLEVE	0.12948	0.01998	9.13E-11	1.1382	(1.095, 1.184)
COPD	0.79107	0.023188	4.38E-255	2.2058	(2.108, 2.308)
asthma	0.08522	0.019372	1.09E-05	1.089	(1.048, 1.131)
OSA	0.41954	0.046994	4.36E-19	1.5213	(1.387, 1.668)
CAD PRS	0.52797	0.00711	0.00E+00	1.6955	(1.672, 1.719)
sex by smoking (never)	0.34135	0.049628	6.06E-12	1.4069	(1.276, 1.551)
sex by smoking (occasionally)	-0.0056	0.102995	9.57E-01	0.9944	(0.813, 1.217)
sex by smoking (previous)	0.38252	0.050397	3.20E-14	1.4660	(1.328, 1.618)
OSA by CAD PRS	-0.01857	0.045087	6.80E-01	0.9816	(0.899, 1.072)

2

3

- 1 **Supplementary Table S5:** Comprehensive secondary analysis GxE testing results for KEGG
- 2 HIF1, VEGF, NFκB or TNF pathways. (See file Goodman_Supplementary_Tables.csv)
- 3 A) module-level GxE results
- 4 B) gene-level GxE results
- 5
- 6 **Supplementary Table S6:** Sensitivity Analysis results
- 7 (See file Goodman_Supplementary_Tables.csv)
- 8 C) Adjusting for potential mediators (hypertension, T1D, T2D and pulmonary fibrosis)
- 9 D) Males only
- 10 E) Females Only
- 11 F) White Europeans only