Genome-wide association meta-analysis of coronary artery disease and periodontitis reveals a novel shared risk locus

Matthias Munz1,2, Gesa M. Richter1, Bruno G. Loos4, Søren Jepsen5, Kimon Diveris6,7, Steven Offenbacher8, Alexander Teumer9, Birte Holtfreter10, Thomas Kocher10, Corinna Bruckmann11, Yvonne Jockel-Schneider12, Christian Graetz13, Loreto Munoz2,3,14, Anita Bhandari1,2,14, Stephanie Tennstedt2,3,14, Ingmar Staufenbiel15, Nathalie van der Velde16,17, André G. Uitterlinden16, Lisette C. P. G. M. de Groot16,18, Jürgen Wellmann19, Klaus Berger19, Bastian Krone20, Per Hoffmann21,22, Matthias Laudes23, Wolfgang Lieb24, Andre Franke25,26, Henrik Dommsch19, Jeanette Erdmann1,2,3,14 & Arne S. Schaefer1,5

Evidence for a shared genetic basis of association between coronary artery disease (CAD) and periodontitis (PD) exists. To explore the joint genetic basis, we performed a GWAS meta-analysis. In the discovery stage, we used a German aggressive periodontitis sample (AgP-Ger; 680 cases vs 3,973 controls) and the CARDIoGRAMplusC4D CAD meta-analysis dataset (60,801 cases vs 123,504 controls). Two SNPs at the known CAD risk loci ADAMTS7 (rs11634042) and VAMP8 (rs1561198) passed the pre-assigned selection criteria (P<0.05; PCAD < 5×10^{-8}; concordant effect direction) and were replicated in an independent GWAS meta-analysis dataset of PD (4,415 cases vs 5,935 controls). SNP rs1561198 showed significant association (PD[Replication]: P = 0.008 OR = 1.09, 95%
Coronary artery disease (CAD) has a well-established genetic basis and a strong inflammatory component (reviewed in1). An association between CAD and the common oral inflammatory disease condition, periodontitis (PD), was reported in several clinical and observational studies (previously reviewed1). Because of the high prevalence of CAD and PD this association is potentially of public health importance. Recent evidence indicates that the observed association between CAD and PD is independent of smoking2 and obesity1. However, it could be explained in parts by other shared risk factors like diabetes, and age. In this context, the knowledge of shared genetic risk variants could substantially contribute to the understanding of the mechanisms that underlie the epidemiological associations.

Recently, we demonstrated shared associations of two CAD risk loci, ANRIL1-7 and PLG5-9, with aggressive periodontitis (AgP), a severe early-onset form of PD. In addition, we observed suggestive evidence for shared association of a rare genetic variant at VAMP310, located at a chromosomal region that was earlier described to be associated with increased colonization of oral periodontal pathogens11. AgP, which has a global prevalence of 0.1%12, is characterized by a particular early age of disease onset (≤35 years of age), which is why patients with AgP generally do not suffer from late-onset diseases like CAD or diabetes. By contrast, chronic periodontitis (CP), a widespread form of PD, has a prevalence of >9% in adults aged ≥30 years13 and is mainly observed in the elderly. AgP and CP have a similar histopathology and can be considered as parts of the same disease spectrum, which attribute to the effects of different combinations of genetic risk loci that determine the individual immune response. Because AgP cases have a high heritability as reflected by the early-age of disease onset and are not confounded by risk factors such as diabetes and age, AgP is particularly suitable to explore the shared genetic basis of CAD and PD14. The identified variants can be subsequently validated for the relevance in the more moderate form of CP.

In the current study, we performed a meta-analysis using data from the “Coronary Artery Disease Genome-wide Replication and Meta-analysis plus The Coronary Artery Disease” (CARDIoGRAMplusC4D) consortium and based on genome-wide association studies (GWAS) datasets for AgP and CP. We provide evidence for the shared association of variant rs1561198 with CAD and PD, which has reported cis-effects on the expression of the adjacent gene VAMP8.

Materials and Methods

Participating studies. The discovery sample consisted of a GWAS of German AgP cases and controls, which was previously described15, and of a case-control GWAS meta-analysis of CAD from the CARDIoGRAMplusC4D consortium16. The replication was carried out with GWA studies of case-control samples of Dutch AgP15 and CP17, and CP case-control samples of German17 and of European American descent18. Summary statistics of all studies were based on the additive model.

The German AgP sample (AgP-Ger) included 680 cases and 3,973 controls. Cases were recruited across Germany by the biobank Poppen19, University-Hospital Schleswig-Holstein, Germany. Controls originate from North- and West-Germany and were recruited from the Competence Network “FoCus - Food Chain Plus”20, the Dortmund Gesundheitsstudie – DOGS21 and the Heinz Nixdorf Recall Studies 1–32. Genotyping of the cases was performed on Illumina Omni Bead Chips and the imputation was based on the 1000 Genomes Phase 3 reference panel23.

The German CP sample (CP-Ger) consisted of 993 cases and 1,419 controls from two independent cross-sectional studies SHIP and SHIP-TREND, recruited at the University Medicine Greifswald23–25. Cases and controls were defined by contrasting subjects within the first vs the third tertile of proportion of proximal sites with attachment loss (AL) ≥4 mm. To increase the statistical power by enriching the case sample with early-onset phenotypes, subjects aged >60 years were excluded. Cases and controls were genotyped either with the Affymetrix Genome-Wide Human SNP Array 6.0 or the Illumina Human Omni 2.5 array and imputed on the 1000 Genomes phase 1 version 3 reference. The meta-analysis was performed by using either the fixed-effects model or the random-effects model, depending on the level of statistical heterogeneity.

The Dutch AgP sample (AgP-NL) consisted of 171 cases of Dutch descent that were recruited from the ACTA (Academisch Centrum Tandheelkunde Amsterdam) and of 2,607 population representative controls from the B-Proof Study26, which were collected at Rotterdam and Wageningen. Genotyping of the cases and imputation was performed together with AgP-Ger.

The German CP (CP-Ger) sample consisted of 993 cases and 1,419 controls from two independent cross-sectional studies SHIP and SHIP-TREND, recruited at the University Medicine Greifswald23–25. Cases and controls were defined by contrasting subjects within the first vs the third tertile of proportion of proximal sites with attachment loss (AL) ≥4 mm. To increase the statistical power by enriching the case sample with early-onset phenotypes, subjects aged >60 years were excluded. Cases and controls were genotyped either with the Affymetrix Genome-Wide Human SNP Array 6.0 or the Illumina Human Omni 2.5 array and imputed on the 1000 Genomes phase 1 reference.

The European American CP (CP-EA) sample included 958 severe (sev) CP cases, 2,293 moderate (mod) CP cases and 1,909 controls from the Atherosclerosis Risk in Communities (ARIC) Study28. Genotyping was carried out using the Affymetrix Genome-Wide Human SNP Array 6.0 and the subsequent genotype imputation was performed on the HapMap Phase II reference with individuals of Northern and Western European (CEU) ancestry.
Discovery. In the discovery stage, we compared variants in GWAS of AgP-Ger and CAD regarding their effect direction and significance level. At first, variants with a minor allele frequency (MAF) < 0.05 were filtered out and only variants with genotypes available for both study samples were kept.

Then we selected variants showing the same effect direction in AgP-Ger and CAD and surpassing the pre-assigned restrictions $P_{\text{AgP-Ger}} < 0.05$ and $P_{\text{CAD}} < 5 \times 10^{-4}$. We chose these $P$-value thresholds to account for the huge difference in sample size (CAD sample ~35x larger than AgP-Ger sample). Subsequently, we pruned the remaining variants using intermediate linkage disequilibrium. We chose a pruning threshold of $r^2 \geq 0.2$ to account for association tails of variants arising from their low correlation with the truly associated haplotype block.

Replication meta-analysis. Variants that passed the discovery step were taken forward to replication stage and meta-analysed in additional PD GWAS samples AgP-NL, CP-Ger, CP-EA-mod and CP-EA-sev. By default, we applied the fixed effects model. However, for variants having a high amount of heterogeneity, i.e. where the $P$-value of Cochran’s $Q$ ($P(Q) < 0.05$ and heterogeneity index $I^2 > 0.5$), we applied the random effects model instead. Correction for multiple testing was performed using the method of Bonferroni. Additionally, we used the adjustment method for shared controls of Zaykin and Kozbur to check the $P$-value inflation when combining CP-EA-mod and CP-EA-sev.

Statistical power calculation for the discovery stage. The statistical power of the AgP-Ger sample at a significance level of 0.05 was calculated with Genetic Association Study (GAS) Power Calculator. When assuming an additive model and a prevalence rate of 0.1% for AgP in the general population, risk variants with minor allele frequencies of $>0.4$ can be detected at a probability (power) of 0.8 if the genotype relative risk (GRR) is $\sim 1.18$. The GRR is an estimator for the OR.

Functional annotation. Variants were annotated using the Genehopper database (DB). Genehopper DB integrates data of many public sources by applying periodically executed extraction, transformation and loading (ETL) processes. Specifically, we used integrated datasets of linkage disequilibrium (LD), expression quantitative trait loci (eQTL) mappings, topologically associated domain boundaries (TADs) and GWAS studies to annotate our findings.

Identified loci were annotated using LD information (correlation measures $r^2$ and $D^2$) from the European reference population (EUR) of 1000 Genomes Phase 3. EQTL mapping information was gathered from Genotype-Tissue Expression project (GTEx), Haploreg v4, GRASP v2, GEUVADIS, SCAN, seeQTL, and Blood eQTL Browser. Information about TAD boundaries was taken from Dixon et al. TADs are genomic regions defined by mainly cell-type independent interactome boundaries representing a spatial compartment in the genome. Physical interactions of regulatory DNA elements and gene promoters occur more frequently within a TAD. Thus, we refer to cis regulation, if a SNP with a putative effect on gene expression (expression SNP; eSNP) resides upstream or downstream from the gene but within the same TAD. In contrast, if a gene is located in a different TAD than the corresponding eSNP, we define this as a putative trans-regulatory effect, being most likely affected indirectly, e.g. via intermediate genes in a pathway. This dataset contained TADs with a length of ~853 kilo base pairs (kb) in average (maximum length $= 4.44$ mega base pairs [mb], minimum length $= 0.8$ kb). Variant consequence information was taken from Ensembl Variation DB and additionally, we annotated variants using combined annotation dependent depletion (CADD) score. The CADD score is calculated by applying the formula $-10 \times \log_2(n)$ on each variant in a ranking of single nucleotide variants which is created by combining several other variant annotation scores using machine learning techniques. Accordingly, a CADD score of $\geq 5.22$ indicates that the variant belongs to the 20% most deleterious substitutions in the human genome and a score $\geq 10$ indicates that the variant belongs to the 10% most deleterious substitutions. To elucidate the relationship to other traits and diseases, we extracted information about phenotype associations from the NHGRI-EBI Catalog.

Genetic risk score calculation. In an additional analysis, we assessed the genetic relation of CAD and PD by calculating a genetic risk score (GRS) based on known CAD risk variants and their corresponding effect sizes in CAD and by applying this score to the sample of AgP-Ger and respective controls (Supplementary Table 1). The GRS was calculated per individual using the formula $\text{GRS} = \sum w_i n_i$, with $w_i$ being the effect size (Odds ratio [OR]) and $n_i$ being the number of risk alleles of variant $x$. GRS of cases and controls of AgP-Ger were then compared using basic statistics.

Results

Discovery meta-analysis. The variant sets of AgP-Ger and CAD consisted of 6,416,838 and 9,455,779 variants. A total of 5,519,261 genetic variants were existent in both studies. In the discovery stage which included AgP-Ger and CAD, 276 variants in three distinct loci at 9p21.3, 15q25.1 and 2p11.2 surpassed our pre-assigned selection criteria (Supplementary Table 1).

The lead variant at 9p21.3 reached the highest level of significance in PD with $P = 1.23 \times 10^{-4}$ (OR = 1.26, 95% CI = [1.12–1.41]) and $P = 1.56 \times 10^{-32}$ (OR = 1.21, 95% CI = [1.17–1.41]) for lead variant rs10116277 in AgP and CAD, respectively (Table 1 and Fig. 1). This single nucleotide polymorphism (SNP) is located in the intronic region of the large non-coding antisense RNA $\text{CDKN2B-AS1}$ (alternative name $\text{ANRIL}$) which has been reported as risk factor for both AgP and CAD before.

The lead variant of the second locus at 15q25.1 was SNP rs11634042, intronic of the gene $\text{MORF4L1}$ (Mortality factor 4 like 1) and 21 kb upstream of $\text{ADAMT57}$ (ADAM metallopeptidase with thrombospondin type 1 motif, 7) with $P = 5.67 \times 10^{-3}$ (OR = 1.18, 95% CI = [1.05–1.32]) and $P = 2.17 \times 10^{-13}$ (OR = 1.08, 95% CI = [1.06–1.10]) for AgP and CAD.
At 15q25.1, SNP rs11634042 (intronic of ANRIL) had the lowest P-value in the replication meta-analysis, with an odds ratio of 1.19 (95% CI [1.05–1.34]). To address putative multiple independent association signals in these two loci, we first applied variant pruning on the 75 variants that passed our selection criteria in the discovery meta-analyses, using intermediate linkage disequilibrium of $r^2 \geq 0.8$; in the following we call these variant sets LD blocks) were selected for replication in PD (Supplementary Table 3).

Table 1. Three loci were identified in the discovery step (AgP-Ger, CAD); 9p21.3, 15q25.1 and 2p11.2. Since the locus at 9p21.3 is already a known shared risk locus for PD and CAD, we only took the loci at 15q25.1 and 2p11.2 forward to replication in PD (Pooled[Replication] = AgP-NL + CP-EA-mod + CP-EA-sev + CP-Ger). In the replication step, only locus 2p11.2 could be successfully replicated. At this locus, SNP rs1561198 showed the strongest association with PD in the overall PD sample (Pooled[PD] = AgP-Ger + AgP-NL + CP-EA-mod + CP-EA-sev + CP-Ger). A1 = Effect allele; A2 = Non-effect allele; EAF = Effect allele frequency; Cas = Cases; Con = Controls; OR = Odds ratio; CI = Confidence interval; P = P-value; CAD = Coronary artery disease; AgP = Aggressive periodontitis; CP = Chronic periodontitis; Ger = German; NL = Dutch; EA = European American; sev = severe; mod = moderate.

The third associated region at 2p11.2 showed association with $P = 4.57 \times 10^{-3}$ (OR = 1.19, 95% CI = [1.05–1.34]) and $P = 6.37 \times 10^{-10}$ (OR = 1.06, 95% CI = [1.04–1.08]) for variant rs1561198 in AgP-Ger and CAD, respectively. Variant rs1561198 is an intergenic SNP located 0.8 kb downstream of VAMP8 (vesicle-associated membrane protein 8) and 1.5 kb upstream of VAMP5 (vesicle-associated membrane protein 5). Also the loci at MORF4L1 and VAMP8 are already known susceptibility loci for CAD 16,45, however they have not yet been reported to be associated with PD.

Replication. Of the three loci that we identified in the discovery meta-analysis, ANRIL had repeatedly been replicated as a genetic risk factor of PD 7,46–48. Accordingly, we only selected the two novel suggestive risk loci at 15q25.1 and 2p11.2 for replication. To address putative multiple independent association signals in these two loci, we first applied variant pruning on the 75 variants that passed our selection criteria in the discovery meta-analyses, using intermediate linkage disequilibrium of $r^2 \geq 0.2$ (Supplementary Table 2). Only the two lead variants rs11634042 at 15q25.1 and rs1561198 at 2p11.2 remained after pruning, indicating an association of a single haplotype block at each locus. Therefore, only SNPs rs11634042 and rs1561198 as well as their high LD variants ($r^2 \geq 0.8$; in the following we call these variant sets LD blocks) were selected for replication in PD (Supplementary Table 3).

At 15q25.1, SNP rs11634042 (intronic of MORF4L1) had the lowest P-value in the replication meta-analysis, with $P = 0.95$ (OR = 1.0, 95% CI = [0.94–1.06]), showing no association with PD (Table 1 and Fig. 2).

At 2p11.2, SNP rs1561198 showed the strongest association (P = 0.007; OR = 1.09; 95% CI = [1.02–1.16]) among 15 out of a total of 38 strongly linked variants of this haplotype block, for which genotype data were available for all PD samples. The association of rs1561198 with PD was significant in the replication after correction for multiple testing, and after pooling all PD samples, SNP rs1561198 had a P-value of $P = 2.02 \times 10^{-4}$ (OR = 1.11; 95% CI = [1.05–1.17]).

A comparison of pooled P-value for CP-EA-mod and CP-EA-sev with and without adjustment for shared controls indicated a reasonable inflation of less than one potency (Supplementary Table 4).

In-silico characterization of selected variant effects. Next, we investigated the associated genetic region at rs1561198 for putative regulatory effects on other genes and associations with other traits. First, we grouped genes in cis and trans with respect to SNP rs1561198 and its 38 variants in the LD block using information about topologically associated domain boundaries (TADs). The LD block of rs1561198 is located within a TAD spanning 2 mega base pairs (mb) on chromosome 2. This TAD harbours 14 genes of which seven are protein coding (Supplementary Table 5). Examination of the 39 variants in the haplotype block of rs1561198 indicated study-wide significant cis- and trans-regulatory effects on multiple genes (Supplementary Table 6). Table 2 shows those study-wide significant eSNP effects in cis with $P < 10^{-10}$ for blood and gastrointestinal tissues that, to our knowledge, have been reported to date. For these cis-located genes the strongest eQTLs were found in blood for VAMP8 (Vesicle associated membrane protein 8) and -regulatory effects on multiple genes (Supplementary Table 6). Table 2 shows those study-wide significant eSNP effects in cis with $P < 10^{-10}$ for blood and gastrointestinal tissues that, to our knowledge, have been reported to date. For these cis-located genes the strongest eQTLs were found in blood for VAMP8 (Vesicle associated membrane protein 8) and VAMP5 (vesicle-associated membrane protein 5).
protein 8; $P = 9.8 \times 10^{-198}$), MAT2A (Methionine adenosyltransferase 2A; $P = 1.4 \times 10^{-76}$), USP39 ($P = 9.8 \times 10^{-38}$), VAMP5 ($P = 8.2 \times 10^{-32}$) and GGCX (Gamma-glutamyl carboxylase; $P = 3.7 \times 10^{-33}$).

Additionally, we assessed the relative pathogenicity of the variants in the haplotype block using the combined annotation dependant depletion score (CADD) to explore their potential of being causal for the observed association (Supplementary Table 7). Lead SNP rs1561198 was among 12 variants with CADD $\geq 5$, indicating a comparably high probability for having a causal effect (Fig. 3). The intergenic SNP rs2166529 ($r^2 = 0.82$ with rs1561198), was assigned with the highest CADD score (15.29). Spearman correlation of $r_s = 0.26$ ($P = 0.34$; Alternative hypothesis: true $r_s$ is not equal to 0) between CADD scores and $P$-values of the associations with PD indicated a low non-significant correlation.

Furthermore, we searched the LD block of rs1561198 for the presence of GWAS lead SNPs of other diseases and traits as listed in the NHGRI-EBI GWAS Catalog with $P < 10^{-5}$. According to the catalogue, the LD block is associated with CAD (rs7568458, $P = 4 \times 10^{-10}$), myocardial infarction (SNP rs10176176, $P = 3 \times 10^{-10}$) and prostate cancer (rs3731827, $P = 3 \times 10^{-3}$) with genome-wide significance (Supplementary Table 8). Moreover, the chromosomal region at rs1561198 $\pm 500$ kb is associated with additional phenotypes including...
In this study, we aimed to identify novel genetic risk factors that are shared between CAD and PD, in order to improve the current pathogenic understanding of both diseases and highlight possible common biological underpinnings. We provide evidence for an association with PD for SNP rs1561198, located at VAMP8, which is a known genetic risk variant of CAD. We were able to replicate our observation from the discovery sample in an independent case-control sample of PD.

A limitation of the study was the relatively small sizes of the available PD case-controls samples compared to the CARDioGRAMplusC4D sample. The CARDioGRAMplusC4D sample was ~12 times larger than the pooled PD samples. This is why we consider that the association of SNP rs1561198 reached genome-wide significance in CAD, whereas the same variant with a similar effect size had a significance level in the pooled PD sample several times lower. For the same reason we think, that the inter-sample heterogeneity is influenced to greater extent by genetic factors compared to the late onset diseases CAD and CP, or reflect the Winner's curse bias, which results in an overestimation of the effect size.

The power of our study is also demonstrated by the re-discovery of the known association at 9p21.3 (CDKN2B-AS1), which is well established for both CAD and PD. The lead-SNP of the association of this locus, rs10116277, had an OR = 1.06 (MAF = 52%) in AgP-Ger, which was comparable in size and direction to the CAD association (OR = 1.26). This SNP is among the variants with the highest OR of all currently known genetic loci of CAD and AgP (Supplementary Table 9). The other previously reported shared risk loci at VAMP3 (rs10864294) and PLG (rs4252120) were not re-discovered because they did not pass the pre-defined criterion F_{CAD} < 5 × 10^{-8}.

Our finding of shared risk variants at VAMP8 support and complement results of a previous report addressing the shared molecular mechanisms of both diseases. In this study, a transcriptome-wide shRNA knock-down approach demonstrated that CDKN2B-AS1 and VAMP3 expression is correlated on the RNA and protein level, and a rare variant upstream of VAMP3 was suggested to be associated with AgP and CAD. VAMP3 and VAMP8 are simultaneously expressed in various cell types, e.g. in the secretory granules of mast cells, where they form complexes with platelet syntaxin 4 during platelet secretion, to release inflammatory cytokines, or in mast cells, where they play the fusion of glucose transporter type 4 (GLUT4). VAMP3 and VAMP8 mediated membrane trafficking in platelets also plays an important role in thrombosis and wound healing, processes with established relevance for the etiology of CAD and PD. In addition, VAMP mediated platelet secretion has an antimicrobial role in the host response to bacterial infection.
PD is caused by bacterial pathogens and it was suggested that the risk of CAD is putatively increased by bacterial infections\textsuperscript{55,56}. Likewise, two clinical trials involving full-mouth mechanical debridement resulted in a transient deterioration of surrogate markers of CAD\textsuperscript{57,58}, which supports the notion that bacterial inoculation and other procedure-related inflammation resulting from mechanical debridement has immediate negative effects on risk for CAD. In addition, living oral pathogens were detected in atheromatous plaques from coronary arteries\textsuperscript{59,60}. In this context it is interesting that various studies determined that SNARE (soluble N-ethylmaleimide-sensitive factor attachment receptor) proteins like VAMP3 and VAMP8, are manipulated by specific pathogens to corrupt host membrane vesicular trafficking (reviewed in\textsuperscript{61}). E.g., \textit{Chlamydia} species mimic VAMP3 and VAMP8 to establish within the host cell\textsuperscript{62}, \textit{Shigella} species stimulate endocytosis of VAMP3 and VAMP8 in the internalization process of Shiga toxin\textsuperscript{63}, and virulence factors of \textit{Leishmania} parasites cleave VAMP3 and VAMP8 to evade phagocytosis\textsuperscript{64}. We also note that a GWAS on periodontal pathogen colonization reported a region within \textit{CAMTA1}, adjacent to \textit{VAMP3}, to be associated with increased quantities of oral pathogenic bacteria\textsuperscript{11,65}. Taken together, we note a possible yet speculative link of the observed associations with \textit{VAMP3} and \textit{VAMP8} and the increased risk of PD and CAD with effects of the putative causative variants on functions of coagulation, wound healing and thrombosis, as well as the invasion of some bacteria into host cells. However, the putative causal variant(s) of the \textit{VAMP8} associated haplotype block and their likely effects on gene regulation still need to be identified. Currently, their effects in the susceptibility for CAD and PD are speculative.

We conclude that our finding of a haplotype block at \textit{VAMP8}, which increases the risk for CAD and PD adds to the previously reported shared risk variants and indicates that the observed association of CAD and PD reported in prior epidemiological studies (previously reviewed\textsuperscript{1}) cannot be solely explained by shared environmental risk factors. In addition, the current study probably missed possible associations with a shared role in the genetic etiology of CAD and PD. This is because of a lack of statistical power of PD to identify shared risk variant of CAD and PD with ORs <1.1 and an incomplete coverage of common and rare variants in the GWAS data sets. To identify the full spectrum of pleiotropic variants, larger sample sizes are needed for future studies.

**Data Availability**

The CARDioGRAMplusC4D summary statistics dataset is freely available on the consortium website (http://www.cardiogramplusc4d.org). Similarly, summary statistics for CP-EU-mod and CP-EU-sev can also be downloaded (see respective publication). Genotype data for aggressive periodontitis samples and the chronic periodontitis sample with German descent are available upon request from the biobanks PopGen (https://www.epidemiologie.uni-kiel.de/biobanking) and ShIP (https://www.fvcm.med.uni-greifswald.de/index.html).

**References**

1. Hansson, G. K. Inflammation, Atherosclerosis, and Coronary Artery Disease. \textit{N. Engl. J. Med.} 352, 1685–1695 (2005).
2. Lockhart, P. B. \textit{et al.} Periodontal Disease and Atherosclerotic Vascular Disease: Does the Evidence Support an Independent Association? \textit{Circulation} 125 (2012).
3. Shungin, D. \textit{et al.} Using genetics to test the causal relationship of total adiposity and periodontitis: Mendelian randomization analyses in the Gene-Lifestyle Interactions and Dental Endpoints (GLIDE) Consortium. \textit{Int. J. Epidemiol.} 44, 638–650 (2015).
50. Polgár, J., Chung, S.-H. & Reed, G. L. Vesicle-associated membrane protein 3 (VAMP-3) and VAMP-8 are present in human platelets. Circ. Cardiovasc. Genet. 8, 159–167 (2015).

51. Bocheneck, G. et al. The large non-coding RNA ANRIL, which is associated with atherosclerosis, periodontitis and several forms of cancer, regulates ADIPOR1, VAMP5 and C11orf10. Hum. Mol. Genet. 22, 4516–4527 (2013).

52. Divaris, K. et al. Genome-wide association study of periodontal pathogen colonization. J. Dent. Res. 91, 215–285 (2012).

53. Susin, C., Haas, A. N. & Albandar, J. M. Epidemiology and demographics of aggressive periodontitis. Periodontol. 2000 65, 27–45 (2014).

54. Eke, P. I. et al. Update on Prevalence of Periodontitis in Adults in the United States: Nhanes 2009 to 2012. J. Periodontol. 86, 611–622 (2015).

55. Gibson, G. Rare and common variants: twenty arguments. Nat. Rev. Genet. 13, 135–45 (2012).

56. Munz, M. et al. A genome-wide association study identifies nucleotide variants at SIGLEC5 and DEFA1A3 as risk loci for periodontitis. Hum. Mol. Genet. 33, 272–279 (2017).

57. The CARDIoGRAMplusC4D Consortium. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat. Genet. 47, 1121–1130 (2015).

58. Teumer, A. et al. Genome-wide association study of chronic periodontitis in a general German population. J. Clin. Periodontol. 40, 977–985 (2013).

59. Divaris, K. et al. Exploring the genetic basis of chronic periodontitis: a genome-wide association study. Hum. Mol. Genet. 22, 2312–24 (2013).

60. Krawczak, M. et al. PopGen: Population-Based Recruitment of Patients and Controls for the Analysis of Complex Genotype-Phenotype Relationships. Public Health Genomics 9, 55–61 (2006).

61. Müller, N. et al. IL-6 blockade by monoclonal antibodies inhibits apolipoprotein (a) expression and lipoprotein (a) synthesis in humans. J. Lipid Res. 56, 1034–42 (2015).

62. Berger, K. DOGS. Bundesgesundheitsblatt - Gesundheitsforsch. - Gesundheitsschutz 55, 816–821 (2012).

63. Schmermund, A. et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL study. Am. Heart J. 144, 212–218 (2002).

64. Autum, A. et al. A global reference for human genetic variation. Nature 526, 68–74 (2015).

65. van Wijngaarden, J. P. et al. Rationale and design of the B-PROOF study, a randomized controlled trial on the effect of supplemental intake of vitamin B12 and folic acid on fracture incidence. BMC Geriatr. 11, 80 (2011).

66. John, U. et al. Study of Health In Pomerania (SHIP): a health examination survey in an East German region: objectives and design. Soz. Präventivmed. 46, 186–94 (2001).

67. Hensel, E. et al. Study of Health in Pomerania (SHIP): a health survey in an East German region. Objectives and design of the oral health section. Quintessence Int. 34, 370–8 (2003).

68. Voke, H. et al. Cohort Profile: The Study of Health in Pomerania. Int. J. Epidemiol. 40, 294–307 (2011).

69. Hill, C. et al. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am. J. Epidemiol. 129, 687–702 (1989).

70. Zaykin, D. V. & Rozbud, D. O. P-value based analysis for shared controls design in genome-wide association studies. Genet. Epidemiol. 34, 725–38 (2010).

71. Skol, A. D., Scott, L. J., Abecasis, G. R. & Boehnke, M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat. Genet. 38, 209–13 (2006).

72. Munz, M., Tönnes, S., Balke, W.-T. & Simon, E. Multidimensional gene search with Genehopper. Nucleic Acids Res. 43, 98–103 (2015).

73. GTex Consortium, G. et al. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348, 648–640 (2015).

74. Ward, L. D. & Kellis, M. haploReg v4: Transcriptome and genome sequencing uncovers functional variation in humans. Hum. Mol. Genet. 44, D877–D881 (2016).

75. Ward, L. D. & Kellis, M. haploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 40, D930–4 (2012).

76. Eicher, J. D. et al. GRASP v2.0: an update on the Genome-Wide Repository of Associations between SNPs and phenotypes. Nucleic Acids Res. 43, D799–804 (2015).

77. Lappalainen, T. et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature 501, 506–11 (2013).

78. Gamazon, E. R. et al. SCAN: SNP and copy number annotation. Bioinformatics 26, 259–62 (2010).

79. Xia, K. et al. seeQTL: a searchable database for human eQTLs. Bioinformatics 28, 451–2 (2012).

80. Westra, H.-J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. 45, 1238–45 (2013).

81. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–80 (2012).

82. Chen, Y. et al. Ensembl variation resources. BMC Genomics 11, 293 (2010).

83. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 46, 310–315 (2014).

84. Welte, D. et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 42, D1001–6 (2014).

85. Samani, N. J. et al. Genomewide association analysis of coronary artery disease. N. Engl. J. Med. 357, 443–53 (2007).

86. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat. Genet. 43, 339–44 (2011).

87. Schafer, A. S. et al. CDKN2BAS is associated with periodontitis in different European populations and is activated by bacterial infection. J. Med. Genet. 48, 38–47 (2011).

88. Ernst, F. D. et al. Replication of the association of chromosomal region 9p21.3 with generalized aggressive periodontitis (GAgP) using an independent case-control cohort. BMC Med. Genet. 11, 119 (2010).

89. Schafer, A. S. et al. Validation of reported genetic risk factors for periodontitis in a large-scale replication study. J. Clin. Periodontol. 40, 563–72 (2013).

90. Howson, J. M. M. et al. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. Nat. Genet. 49, 1113–1119 (2017).

91. Polgar, J., Chung, S.-H. & Reed, G. L. Vesicle-associated membrane protein 3 (VAMP-3) and VAMP-8 are present in human platelets and are required for grape-platelet secretion. Blood 100, 1081–3 (2002).
51. Paumet, F. et al. Soluble NSF attachment protein receptors (SNAREs) in RBL-2H3 mast cells: functional role of syntaxin 4 in exocytosis and identification of a vesicle-associated membrane protein 8-containing secretory compartment. J. Immunol. 164, 5850–7 (2000).
52. Zhao, P. et al. Variations in the requirement for v-SNAREs in GLUT4 trafficking in adipocytes. J. Cell Sci. 122, 3472–3480 (2009).
53. Sun, J. The Interaction Between Pathogens and the Host Coagulation System. Physiology 21, 281–288 (2006).
54. Yeaman, M. R. Platelets at the nexus of antimicrobial defence. Nat. Rev. Microbiol. 12, 426–437 (2014).
55. Beck, B. J. et al. Periodontal Disease and Coronary Heart Disease. Circulation 112, 19 LP–24 (2005).
56. Roivainen, M. et al. Infections, inflammation, and the risk of coronary heart disease. Circulation 101, 252–7 (2000).
57. Tonetti, M. S. et al. Treatment of Periodontitis and Endothelial Function. N. Engl. J. Med. 356, 911–920 (2007).
58. Minassian, C., D’Auito, F., Hingorani, A. D. & Sneath, L. Invasive Dental Treatment and Risk for Vascular Events. Ann. Intern. Med. 153, 499 (2010).
59. Gaetti-Jardim, E., Marcelino, S. L., Feitosa, A. C. R., Romito, G. A. & Avila-Campos, M. J. Quantitative detection of periodontopathogenic bacteria in atherosclerotic plaques from coronary arteries. J. Med. Microbiol. 58, 1568–1579 (2009).
60. Cavrini, F. et al. Molecular detection of Treponema denticola and Porphyromonas gingivalis in carotid and aortic atheromatous plaques by FISH: report of two cases. J. Med. Microbiol. 54, 93–96 (2005).
61. Wesołowski, J. & Paumet, F. SNARE motif: a common motif used by pathogens to manipulate membrane fusion. Virulence 1, 319–24 (2010).
62. Delevoye, C. et al. SNARE protein mimicry by an intracellular bacterium. PLoS Pathog. 4, e100022 (2008).
63. Renard, H.-F., Garcia-Castillo, M. D., Chambon, V., Lamaze, C. & Johanes, L. Shiga toxin stimulates clathrin-independent endocytosis of the VAMP2, VAMP3 and VAMP8 SNARE proteins. J. Cell Sci. 128, 2891–2902 (2015).
64. Matte, C. & Descoteaux, A. Exploitation of the Host Cell Membrane Fusion Machinery by Leishmania Is Part of the Infection Process. PLoS Pathog. 12, e1005962 (2016).
65. Rhodin, K. et al. Chronic periodontitis genome-wide association studies: gene-centric and gene set enrichment analyses. J. Dent. Res. 93, 882–90 (2014).

Acknowledgements

Data on coronary artery disease have been contributed by CARDIoGRAMplusC4D investigators and have been downloaded from www.CARDIOGRAMplusC4D.ORG. This work was supported by a research grant of the German Research Foundation DFG (Deutsche Forschungsgemeinschaft; G.Z.: SCH 1582/3-1). Funding was provided by the German Federal Ministry of Education and Research (BMBF) in the context of the e:Med program (e:AtheroSysMed and sysNFLAME), the FP7 European Union project CVgenes:target (261123) and a grant from the Fondation Leducq (CADgenomics: Understanding Coronary Artery Disease Genes, 12CVD02). This study was also supported through the Deutsche Forschungsgemeinschaft (DFG) cluster of excellence ‘Inflammation at Interfaces’. Collection of the AgP cases was additionally supported by the German Ministry of Education and Research through the POPGEN biobank project (01GR0468 and 01ELY1103). The Dortmund Health Study (DHS) is supported by the German Migraine & Headache Society (DMKG) and by unrestricted grants of equal share from Almirall, Astra Zeneca, Berlin Chemie, Boehringer, Boots Health Care, Glaxo-Smith-Kline, Janssen Cilag, McNeil Pharma, MSD Sharp & Dohme and Pfizer to the University of Muenster (collection of sociodemographic and clinical data). Blood collection in the Dortmund Health Study was done through funds from the Institute of Epidemiology and Social Medicine University of Muenster and genotyping supported by the German Ministry of Research and Education (BMBF, Grant No. 01ER0816). FOCUS was supported by the Federal Ministry of Education and Research BMBF (FKZ 0315540A). The HNR study is supported by the Heinz Nixdorf Foundation (Germany). Additionally, the HNR study was funded by the German Ministry of Education and Science and the German Research Council (DFG; Project SI 236/8-1, SI236/9-1, ER 155/6-1). The genotyping of the Illumina HumanOmni-1 Quad BeadChips of the HNR subjects was financed by the German Centre for Neurodegenerative Diseases (DZNE), Bonn. SHIP is part of the Community Medicine Research net (CMR, http://www.community-medicine.de) of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (Grants No. 01ZZ9603, 01ZZ0010, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by the Federal Ministry of Education and Research (grant 03IS2061A). This project was granted by BMBF-01 ZZ-9603/0 and BH was supported by GABA, Switzerland. Generation of genome-wide SNP data has been supported by the Federal Ministry of Education and Research (Grant No. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg, West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH.

Author Contributions

Study design: Matthias Munz, Steven Offenbacher, Thomas Kocher, André G. Uitterlinden, Lizette C.P.G.M. de Groot, Jeanette Erdmann, Arne S. Schaefer. Study management: Gesa M. Richter, Bruno G. Loos, Sören Jepsen, Kimon Divaris, Steven Offenbacher, Alexander Teumer, Birte Holtfrerter, Thomas Kocher, Corinna Bruckmann, Yvonne Jockel-Schneider, Christian Graetz, Ingmar Staufenbiel, Nathalie van der Velde, André G. Uitterlinden, Lizette C.P.G.M. de Groot, Jürgen Wellmann, Klaus Berger, Bastian Krone, Per Hoffmann, Matthias Laudes, Wolfgang Lieb, Andre Franke, Henrik Dommisch, Jeanette Erdmann, Arne S. Schaefer. Subject recruitment: Bruno G. Loos, Steven Offenbacher, Birte Holtfrerter, Thomas Kocher, Corinna Bruckmann, Yvonne Jockel-Schneider, Christian Graetz, Ingmar Staufenbiel, Nathalie van der Velde, André G. Uitterlinden, Lizette C.P.G.M. de Groot, Jürgen Wellmann, Klaus Berger, Bastian Krone, Per Hoffmann, Henrik Dommisch, Jeanette Erdmann, Arne S. Schaefer. Genotyping: Steven Offenbacher, Alexander Teumer, Christian Graetz, Matthias Laudes, Wolfgang Lieb, Andre Franke, Arne S. Schaefer. Methods and analysis: Matthias Munz, Kimon Divaris, Alexander Teumer, Nathalie van der Velde, Arne S. Schaefer. Interpretation of results: Matthias Munz, Gesa M. Richter, Arne S. Schaefer. Drafting of manuscript: Matthias Munz, Gesa M. Richter, Arne S. Schaefer. Critical review of manuscript: Gesa M. Richter, Kimon Divaris, Alexander Teumer, Birte Holtfrerter, Loreto Munoz, Stephanie Tennstedt, Anita Bhandari, Wolfgang Lieb, Jeanette Erdmann, Arne S. Schaefer.
Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-31980-8.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018