





## Antibody response to oral biofilm is a biomarker for acute coronary syndrome in periodontal disease

Mariliis Jaago<sup>1,2</sup>, Nadežda Pupina<sup>1</sup>, Annika Rähni <sup>1,2</sup>, Arno Pihlak<sup>1</sup>, Helle Sadam<sup>1,2</sup>, Nihal Engin Vrana<sup>3</sup>, Juha Sinisalo <sup>4</sup>, Pirkko Pussinen <sup>5</sup> & Kaia Palm <sup>1,2</sup>✉

Cumulative evidence over the last decades have supported the role of gum infections as a risk for future major cardiovascular events. The precise mechanism connecting coronary artery disease (CAD) with periodontal findings has remained elusive. Here, we employ next generation phage display mimotope-variation analysis (MVA) to identify the features of dysfunctional immune system that associate CAD with periodontitis. We identify a fine molecular description of the antigenic epitope repertoires of CAD and its most severe form - acute coronary syndrome (ACS) by profiling the antibody reactivity in a patient cohort with invasive heart examination and complete clinical oral assessment. Specifically, we identify a strong immune response to an EBV VP26 epitope mimicking multiple antigens of oral biofilm as a biomarker for the no-CAD group. With a 2-step biomarker test, we stratify subjects with periodontitis from healthy controls (balanced accuracy 84%), and then assess the risk for ACS with sensitivity 71–89% and specificity 67–100%, depending on the oral health status. Our findings highlight the importance of resolving the immune mechanisms related to severe heart conditions such as ACS in the background of oral health. Prospective validation of these findings will support incorporation of these non-invasive biomarkers into clinical practice.

<sup>1</sup>Protobios LLC, Mäealuse 4, 12618 Tallinn, Estonia. <sup>2</sup>Department of Chemistry and Biotechnology, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia. <sup>3</sup>Spartha Medical, 14B Rue de la Canardiere, 67100 Strasbourg, France. <sup>4</sup>Heart and Lung Center, Helsinki University Hospital, and Helsinki University, Helsinki, Finland. <sup>5</sup>Oral and Maxillofacial Diseases, University of Helsinki, FI-00014 Helsinki, Finland. ✉email: [kaia@protobios.com](mailto:kaia@protobios.com)

Coronary artery disease (CAD) is the leading cause of morbidity and mortality worldwide<sup>1</sup> caused by metabolic disorders in lipid oxidation promoting inflammatory alterations on the endothelium<sup>2,3</sup> and culminating in plaque rupture<sup>4,5</sup>. The heritability of CAD and its familial clustering are well established<sup>6</sup>. Genome-wide association studies (GWAS) have identified a number of causal CAD-associated genes and loci<sup>7</sup>. These findings highlight the largely polygenic nature of the inheritability of CAD<sup>8,9</sup>, rendering some individuals more susceptible or resilient to developing atherosclerosis<sup>10</sup>. CAD has been associated with the unhealthy lifestyle placing it among “immunoinflammatory” diseases<sup>11</sup>. While remarkable progress has been made in understanding the mechanisms of atherogenesis as robust methods of identifying high-risk atherosclerosis via genomics and imaging are at hand, highly sensitive and specific biomarkers for CAD have remained elusive. Importantly, as much as 30% of control populations are thought to unknowingly include subjects with CAD, impacting power and accuracy of clinical biomarker studies<sup>12–14</sup>.

Periodontitis, a major oral dysbiosis-driven inflammatory disease, is associated with increased risk of atherosclerotic cardiovascular diseases<sup>15</sup>. Up to 700 bacterial species have been identified in the oral cavity (Human Microbiome Project Consortium). Intriguingly, DNA of periodontal pathogens (e.g., *Porphyromonas gingivalis*) and live bacteria have been detected in atherosclerotic lesions<sup>16–18</sup>. The microbial composition of gut microflora of patients with CAD has been found to be more inflammatory than in healthy patients<sup>19</sup>. Similarly, the oral microbiome of CAD patients may be altered<sup>20</sup>. Host-microbe interaction in the periodontium can initiate or even aggravate atherosclerotic processes through the activation of innate immunity, bacteremia, and direct involvement of cytokines and inflammatory proteins of oral microbiota<sup>21–23</sup>.

Recent research suggests that abnormal changes to the gut microbiota flora may also contribute extensively to the progression of CAD<sup>24</sup>. As the microbiome plays a central role in the balance between immune activation and immune tolerance<sup>25</sup> and in the light of dysbiosis in microflora, it is no surprise that atherosclerosis has a strong autoimmune component<sup>26</sup>. First, CAD risk locus includes the major histocompatibility complex (MHC) containing a dense cluster of genes involved in inflammation, immunity, and self-recognition<sup>21,27</sup>. Furthermore, a

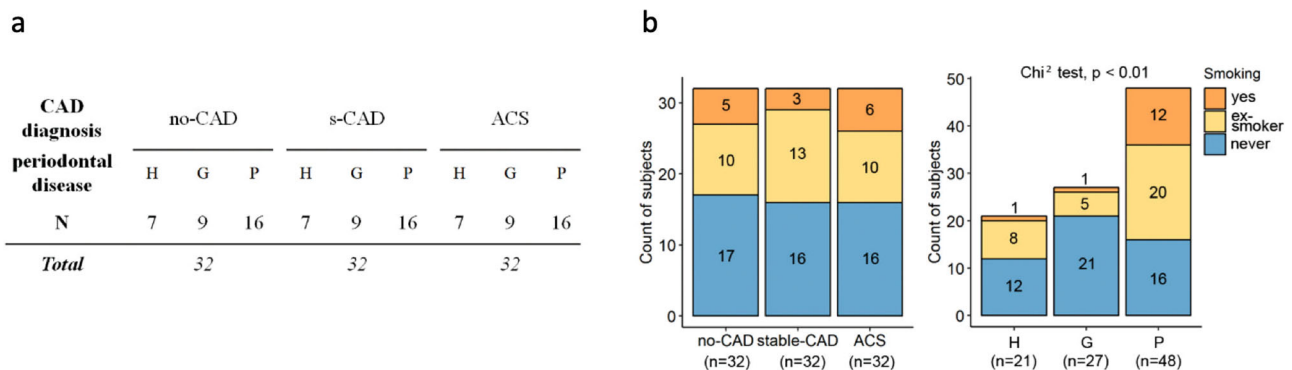
depletion of T or B cells leads to an attenuation of atherosclerosis<sup>22</sup>. Although T cells seem to be essential, B cells and antibodies play an accelerating and perpetuating role<sup>23</sup>. Thus, atherosclerosis is a chronic inflammatory disease with an autoimmune component<sup>26</sup>. Antibodies against oxidized low-density lipoproteins (oxLDL) positively correlated with the disease<sup>28</sup>. Besides oxLDL/ApoB, heat shock proteins (HSPs) and some foreign peptides from pathogens including cytomegalovirus (CMV), hepatitis C virus (HCV), HIV, human papillomavirus (HPV), and others have been proposed as atherosclerosis-relevant antigens<sup>29</sup>. However, the relation between antibodies and atherosclerotic disease burden and progression has remained unclear.

Here we used MVA<sup>30,31</sup>, an unbiased, high-throughput, comprehensive approach based on next-generation phage display, to identify biomarkers of periodontal conditions associated with stable coronary artery disease and progression to acute coronary syndrome.

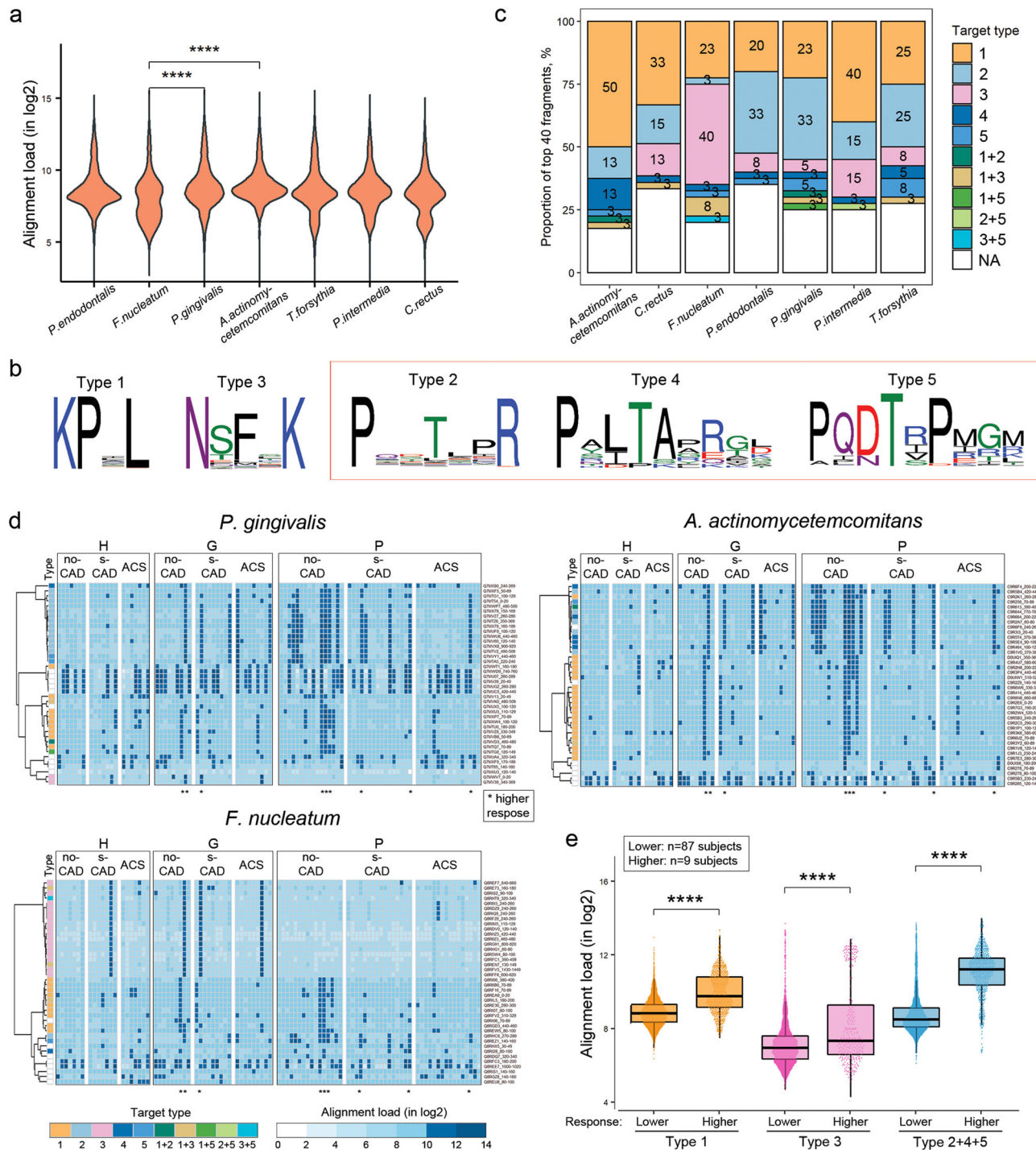
## Results

**Shared immunoreactivity to epitopes linked to periodontal pathogens.** We used MVA immunoprofiling analysis of sera samples of 96 individuals from the Corogene cohort<sup>32</sup> to identify peptide antigens related to antibody immune response in periodontal disease and CAD. Characteristics of the subjects according to their CAD and periodontal status are presented in Fig. 1 (Fig. 1a, Supplementary Table S1). Within the cohort, the proportion of ex- or active smokers was significantly higher in patients with periodontitis than in the rest of the cohort (Chi<sup>2</sup>,  $p < 0.01$ , Supplementary Table S1, Fig. 1b, Supplementary Fig. S2).

Altogether we identified 14.5 million distinct peptide epitopes from the MVA immunoprofiles across the whole study cohort and converged these to 8088 most abundant and shared antigenic epitopes by clustering analysis (Supplementary Fig. S3). Given the gum disease background of the samples<sup>33,34</sup>, we first examined whether we could detect from the immunoprofiles of the study samples the immunoreactivity to the 7 most common periodontal pathogens (*Porphyromonas gingivalis*, *Tannerella forsythia*, *Prevotella intermedia*, *Fusobacterium nucleatum*, *Campylobacter rectus*, *Aggregatibacter actinomycetemcomitans*, and *Porphyromonas endodontalis*). Data analysis showed that antibody response



**Fig. 1** Characteristics of the clinical cohort. **a** Subjects divided in three groups based on CAD diagnosis (no-CAD, stable-CAD (s-CAD), or ACS). Total, size of the group. Each of the three CAD groups included subjects with different periodontal health diagnosis: periodontally healthy (H), patients with gingivitis (G) – a transient gum inflammation –, or patients with periodontitis (P). N, size of the sub-group. **b** Significant association between periodontal diagnosis and smoking is observed in the study cohort. Frequency distribution graphs of subjects ( $n = 96$ ) in groups with coronary artery health condition (no-CAD, stable-CAD, or ACS) or periodontal condition (H, G, P). Statistically significant difference in prevalence of cigarette smoking was observed in periodontal condition groups, being highest among the periodontitis group subjects, but not within CAD groups (Chi<sup>2</sup> test,  $p$  value  $< 0.05$ ,  $n = 96$  independent subjects).  $x$ -axes – clinical subgroups;  $y$ -axes – number of subjects; color-fill – yes, active cigarette smoker (orange); ex-smoker, has quit cigarette smoking (yellow); never, no history of cigarette smoking (blue).



to antigens of *P. gingivalis* and *A. actinomycetemcomitans* was the highest, whereas *F. nucleatum* antigens were on average significantly less recognized (Mann–Whitney *U*, \*\*\*\* $p < 0.0001$ , two-sided) (Fig. 2a). By analyzing the sequences of the bacterial antigens, several dominant core epitopes shared by these antigens were identified, among which multiple types were derivatives of a common signature P.T.[P][R] (Types 2, 4 and 5), Fig. 2b, Supplementary Table S2). Interestingly, P..T.[P][R] patterns (i.e., Type 2, 4, or 5) were present in 45% and 30% of *P. gingivalis* and *A. actinomycetemcomitans* immunodominant antigens, respectively, whereas only 10% among *F. nucleatum* antigens (Fig. 2c). Shared antibody response patterns to bacterial antigen epitopes (target types, clustering on left) were observed as specific for

certain clinical heart and dental conditions (grouping on top) (Fig. 2d, Supplementary Fig. S4). Notably, the majority of subjects (5/9) with high immune reactivity to these pathogenic periodontal bacteria (indicated with dots under intensity plot) belonged to the no-CAD group (Fig. 2d, Supplementary Fig. S5). Overall, in subjects ( $n = 9$ ) with the highest immunoreactivity to these pathogens, P..T.[P][R] containing-epitopes (Types 2, 4, and 5) were the most prevalent and on average recognized at higher levels as compared to epitopes of Type 1 and 3 (Fig. 2e). Altogether, these data showed that immune response against a dominant core epitope P..T.[P][R], which was associated with common pathogenic oral bacteria, correlated with a differential risk of ACS.

**Fig. 2 Shared immunoreactivity to epitopes linked to periodontal pathogens detected by MVA from immunoprofiles of the study cohort.** Group-wide top 8088 epitopes from the MVA immunoprofiles were aligned onto proteomes of the most common periodontal pathogens ( $n = 7$ ). Based on how many epitopes aligned and how enriched were the epitopes in MVA immunoprofiles for individual sample separately, top 40 fragments with highest alignment loads were selected per pathogen. **a** Antibody response to top 40 antigens for *P. gingivalis* and *A. actinomycetemcomitans* was the highest, whereas antibody response to *F. nucleatum* antigens was found to be low. Pair-wise Mann-Whitney  $U$ , two-sided, \*\*\*\* $p < 0.0001$ ,  $p$  values not adjusted for multiple comparisons.  $y$ -axis - alignment load, representing how many MVA immunoprofile epitopes aligned onto protein sequences and how abundantly were they seen in MVA immunoprofiles (in  $\log_2$ ). **b** Unsupervised clustering identified most abundant epitopes with consensus sequences from alignments on target fragments. Type 1: KP.L in 1033 fragments; type 2: P..T.[P]R in 833 fragments; type 3: N[ST]F.K in 421 fragments; type 4: P[AYS][LI]TA.[REQ][GT][LDK] in 150 fragments; type 5: PQ[DN]T[RIV]P[MIR][GRT][MRK] in 107 fragments. Outline - type 2, 4, and 5 epitopes share similar core pattern P..T.[P][R]. **c** Proportions of antigenic epitope types across the antigens of the seven periodontal pathogens.  $y$ -axis - proportion (%) among top 40 antigen fragments (cumulative), data labels on bars - proportion (%),  $x$ -axis - oral pathogen species, target type (fill color) - epitope pattern type (from sequences in **b**). **d** Shared antibody response patterns to bacterial antigen epitopes (target types, clustering on left) were observed as specific for certain clinical heart and dental conditions (grouping on top). Immunoreactivity profiles against top 40 antigens in *F. nucleatum*, *P. gingivalis*, and *A. actinomycetemcomitans* are shown. Profiles were clustered row-wise according to epitope types (type, sequences in **b**). Subjects (in lanes) with highest immunoreactivity to top antigens (Supplementary Fig. S5) are marked with asterisks (\*) under intensity plots. Vertical lanes - individual samples ( $n = 96$ ), categorized based on CAD (no-CAD, s-CAD, ACS) and periodontal diagnoses (periodontally healthy (H), gingivitis (G), periodontitis (P)); rows - each row represents a distinct 20-aa antigenic region of protein primary sequence; blue color-scale - intensity of blue represents the alignment load of individual sample; target type - epitope type (sequences in **b**). **e** Distribution of subjects with low ( $n = 87$ ) or high ( $n = 9$ ) immune response (by Supplementary Fig. S5) to different targeted epitopes. Taken together, Types 2, 4, and 5 were more common to high-response subjects, as compared to other target types. Pair-wise Mann-Whitney  $U$ , two-sided, \*\*\*\* $p < 0.0001$ ,  $p$  values not adjusted for multiple comparisons.  $y$ -axis - alignment load in  $\log_2$ , response - high-response subjects ( $n = 9$ ) or low-response subjects ( $n = 87$ ), type - target epitope type (sequences in **b**).

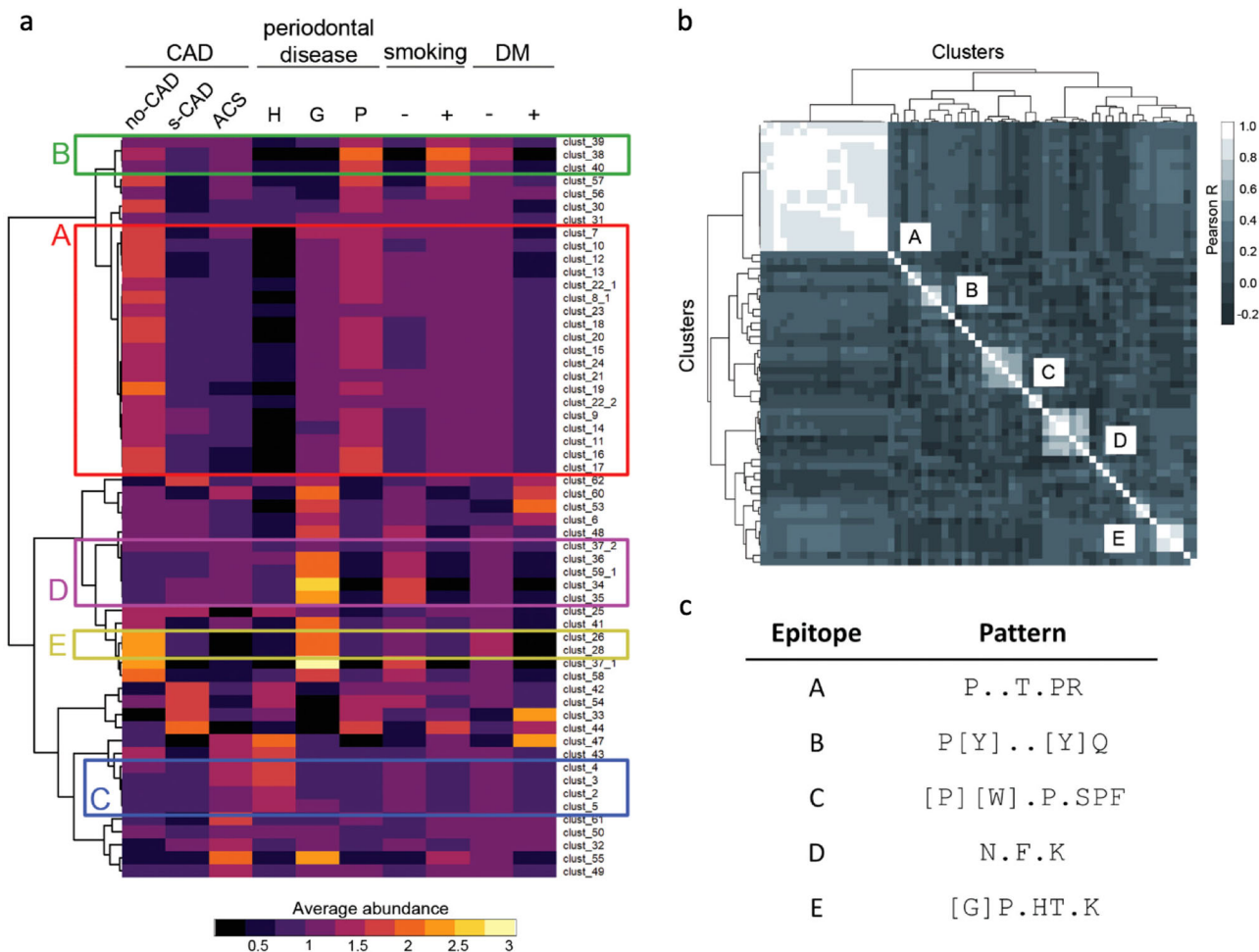
### Different clinical groups share common features in the immunoprofiles.

Given the findings that the immune response against oral pathogenic bacteria was associated with the clinical diagnosis, we next analyzed the difference of immune response to each of the 8088 epitopes in CAD or periodontal disease groups using ROC analysis with specific criteria: sensitivity  $\geq 50\%$  and specificity  $\geq 70\%$ , Kruskal-Wallis test  $p < 0.05$ . Clustered analysis based on sequence similarity arrived at 62 group-differentially targeted epitope clusters with shared core patterns (Supplementary Fig. S6, Supplementary Table S3). When we correlated the 62 epitope clusters based on average abundances in clinical diagnosis groups, clusters with similar response patterns were seen grouping together (Fig. 3a). Pearson correlation-based analysis united the 62 clusters further into five major epitopes (A to E), where the largest epitope A shared the core pattern P..T.PR (Fig. 3b, c, Supplementary Fig. S7). The epitope pattern P..T.PR includes P..T.[P][R], the one also observed as predominant among periodontal pathogens (Fig. 2b, Supplementary Table S2). We found that differential antibody response patterns against epitopes A to E were diagnosis group-specific (Fig. 3a, b, Supplementary Figs. S7 and S8). Specifically, stronger response to epitope A was specific to periodontitis and no-CAD cohorts (red outline), whereas stronger response to epitope B was detected in periodontitis and smoking subgroups (green outline) (Fig. 3a, b, Supplementary Fig. S8a, b). Stronger antibody response to epitope C, on the other hand, was characteristic to periodontally healthy patients but with an ACS diagnosis (blue outline), epitope D was more targeted in subjects with gingivitis (yellow outline), and epitope E in subjects with gingivitis but not in CAD (pink outline) (Fig. 3a, b, Supplementary Fig. S8c-e). In conclusion, these five major epitopes (A-E) were targeted by the strongest and differential antibody response across diagnosis groups.

**Microbial mimicry of the P..T.PR core epitope that is common to periodontitis encompasses the highly antigenic epitope of EBV VP26.** Our analyses highlighted a strong response to epitope A with the core pattern P..T.PR (Figs. 2d, 3c), which we have previously mapped to EBV VP26 protein encompassing 153-176aa<sup>30,31</sup>. When aligning all individual MVA immunoprofile-derived peptides of the current study cohort to the primary sequence of EBV VP26, we observed that subjects within the periodontitis group exhibited high immunoreactivity to this

C-terminal epitope (Mann-Whitney  $U$ , \* $p < 0.05$ , two-sided, Fig. 4a, b). Independent validation was performed using dot ELISA analysis, where sera samples from the current clinical cohort were exposed to phage particles displaying the C-terminal VP26 epitope sequence (Fig. 4c). Patients predicted as seropositive against the P..T.PR epitope by MVA (MVA+) were observed with significantly higher immunoreactivity to the displayed EBV VP26 epitope in the dot ELISA analysis (Fig. 4c, Mann-Whitney  $U$ , \*\*\*\* $p < 0.0001$ , two-sided). Therefore, these results further confirmed the MVA findings of specific seroreactivity mapping of epitope A (with the core pattern P..T.PR) to EBV VP26 (Fig. 4c). As we found that many antigens of the periodontal bacteria shared features of the epitope A (Fig. 2b, d, Supplementary Table S2) and could thus mimic the epitope of EBV VP26 antigen, we determined using annotation analysis that these could include a transmembrane protein signal peptidase I (100-129 aa, Uniprot accession Q7MTG1) and a cytosolic transcription termination factor Rho (160-189 aa, Q7MX79) of *P. gingivalis*, and isoleucine-tRNA ligase (770-789aa, C9R644) in *A. actinomycetemcomitans* (Supplementary Table S2). Also, epitope B was annotated to the tandem repeat ( $3 \times 13$  aa) in the C-terminal part (741-779 aa) of EBV nuclear antigen 6 (EBNA6) protein (Supplementary Fig. S9), in harmony with other studies reporting this region as highly immunogenic<sup>35,36</sup>. In conclusion, strong antibody response to EBV, in particular to VP26 and sequence-mimicking bacterial antigens, was distinguishing subjects with periodontitis from periodontally healthy controls.

**Biomarkers to predict ACS risk from periodontal disease.** To examine whether the delineated five epitopes (A-E) could act as biomarkers stratifying periodontitis and/or CAD conditions, multi-variable models were built and fitted using antibody response to epitopes as predictive biomarkers. Firstly, it was confirmed that immune response to epitopes A-E had no significant correlations with age or gender (Supplementary Fig. S10). The optimal model including immune response to epitopes A, B, and C to differentiate periodontitis from healthy showed a balanced accuracy of 81% for the training subset (80% of samples) and 84% for the validation subset (20% of samples) (Fig. 5a and Supplementary Fig. S11). As a whole, strong response to epitopes A and B was characteristic to the periodontitis group, whereas response to epitope C was more common



**Fig. 3** MVA immunoprofiles of subjects with different CAD conditions, periodontal disease severity, cigarette smoking behavior, and diabetes diagnosis. **a** 62 epitope clusters with group-specific MVA immunoprofile features across clinical classifiers. Although 19% of the subjects were diagnosed with diabetes (either type I or type II, not specified further), no significant association of diabetes with either CAD or periodontal diagnosis was found (Supplementary Fig. S2). Average abundance, calculated as mean of peptides containing the epitope within a given group and normalized with the mean values across all groups (color-coded from purple to yellow). Colored outlines with capital letters refer to epitopes in panel **c**. Clustering distance: Pearson correlation coefficient; clustering method: ward.D2. 57/62 clusters with high mean abundance (>150) across groups are shown. “-” under smoking designates subjects without any exposure to smoking, “+” depicts subjects with a history of exposure (ex-smoker) or currently actively smoking. CAD - coronary artery disease, no-CAD - no CAD diagnosis, s-CAD - stable CAD diagnosis, ACS - acute coronary syndrome, H -periodontally healthy controls, G - patients with gingivitis diagnosis, P - patients with periodontitis diagnosis, DM - diabetes mellitus. **b** Similar behavior-based clustering of 62 epitope clusters (x- and y-axis) using peptide abundance values (in log10) across the study cohort. Pearson R correlation indices (color-scale) were calculated and visualized in a correlation matrix. Five distinct large clusters were identified and defined as epitopes A to E. **c** Core consensus patterns of epitopes A to E, identified in Supplementary Fig. S7.

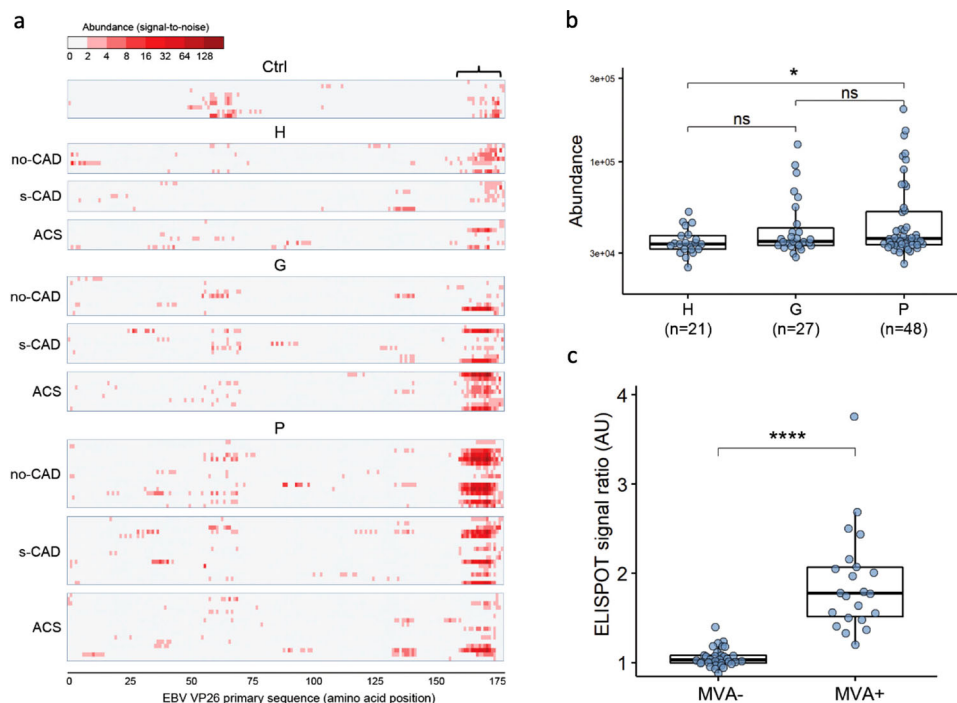
in the periodontally healthy group (Fig. 5a). Response to epitopes D and E did not provide additional useful information when discriminating between periodontitis diagnosis groups (Supplementary Fig. S12a).

Next, by characterizing immune features that were CAD-specific, we discovered that periodontally healthy subjects shared a strong response against epitope C that was high in ACS subgroup (with statistically significant trends, Mann-Whitney *U*, *p* = 0.097, two-sided, Fig. 5b). In subjects with gingivitis, a transient inflammation condition, response against epitope A was differentiating between CAD groups, being high in ACS as compared to no-CAD (Mann-Whitney *U*, \**p* < 0.05, two-sided, Fig. 5b). In the periodontitis group of subjects with chronic periodontal inflammation, antibody response to two independent markers (epitopes A and E) was identified as significant for the no-CAD cohort (Mann-Whitney *U*, \**p* < 0.05, \*\**p* < 0.01, two-sided, Fig. 5b). Response to epitopes B and D was not

differentially linked to no-CAD or ACS diagnosis (Supplementary Fig. S12b). ROC analysis was used to set thresholds for each of the 3 epitopes (A, C and E) in discriminating between clinical subgroups (Fig. 5c). As a result, when first classifying subjects based on their periodontal findings (epitopes A, B and C) and then assessing response to epitopes A, C, and E which differentiate between CAD diagnoses, it was possible to predict the ACS risk in periodontal disease and ascertain the no-CAD phenotype (Fig. 5c). In conclusion, MVA immunoprofiling provided useful blood-biomarkers to predict ACS risk from periodontal disease.

**Discussion**

Here, we report the detailed antibody epitope delineation study identifying a set of immunogenic features associated with periodontal pathogens and CAD.



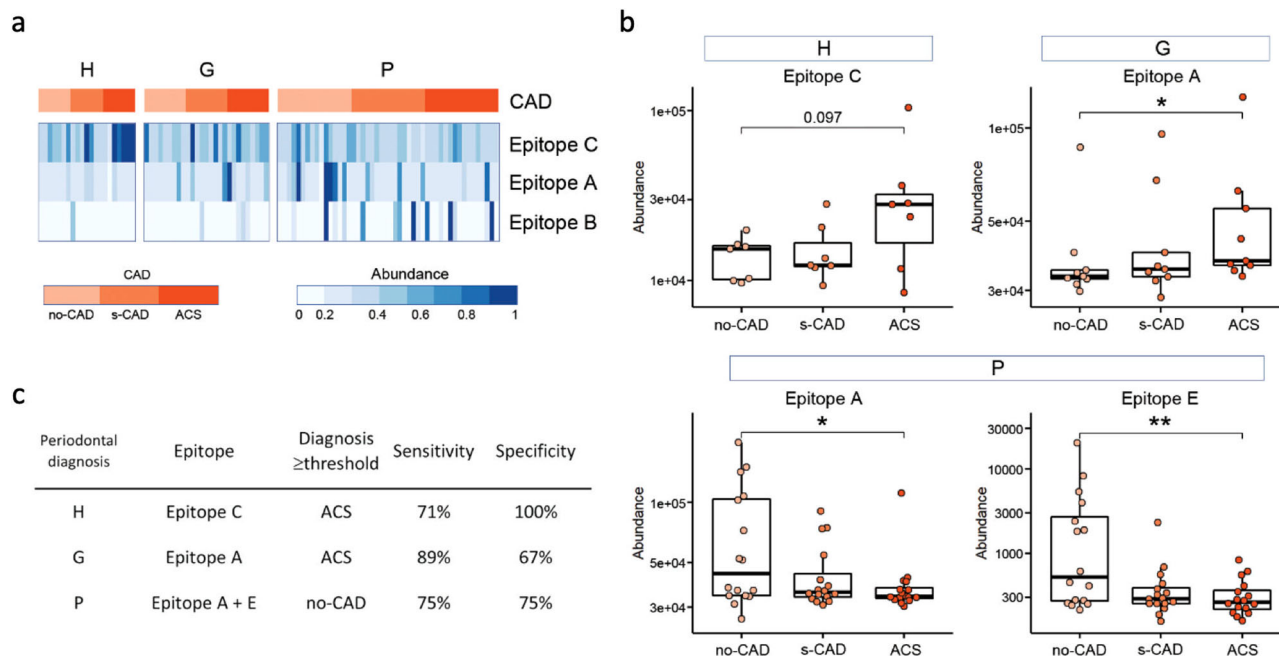
**Fig. 4 Immunoreactivity to epitope on VP26 EBV and alike mimicking features stratifies patients with periodontitis. a** Individual immunoreactivity profiles ( $n = 96$ ) of the study cohort against EBV protein VP26, shown for different clinical groups: periodontally healthy (H), gingivitis (G), or periodontitis (P). All peptides from individual immunoprofiles were aligned (with  $\geq 6$  matching amino acid positions) to primary sequence of EBV VP26 protein along with random reference. Peptide epitope abundance (as signal-to-noise ratio, red color-scale) is visualized per subject (in rows, separated into periodontal diagnosis groups) and per amino acid position (on x-axis). Ctrl – samples with negative EBV-viral capsid antigen serology ( $n = 9$ ). **b** Highly antigenic P..T.PR epitope is differentially targeted across periodontal groups shown as box plots of abundance of peptides containing the epitope (Mann-Whitney  $U$ , two-sided,  $p$  values not corrected for multiple comparisons,  $*p < 0.05$ , ns  $p > 0.05$ ,  $n = 96$  independent patients), Supplementary File S6. **a, b** Abbreviations and group sizes: H periodontally healthy ( $n = 21$  patients ( $3 \times 7$ )), G gingivitis ( $n = 27$  ( $3 \times 9$ )), P periodontitis ( $n = 48$  ( $3 \times 16$ )), no-CAD no CAD diagnosis ( $n = 32$ ), s-CAD stable CAD ( $n = 32$ ), ACS acute coronary syndrome ( $n = 32$ ). **c** Anti-P.DT.PR epitope-like immune response identified by MVA was in high correlation with data from independent dot ELISA validations. Phage particles displaying peptides with P.DT.PR or mutant P.DA.PR sequences were used for dot ELISA. Of the 52 randomly tested samples, those with a positive signal to anti-P.DT.PR response detected by MVA ( $>1200$  abundance of P.DT.PR containing peptides, MVA+) showed significantly higher signals in dot ELISA (y-axis, signal-to-background arbitrary unit (AU)) compared to samples with a negative response from MVA data (MVA-, Mann-Whitney  $U$  test, two-sided, \*\*\*\* $p < 0.0001$ ,  $n = 52$  independent subjects).

The remarkable heterogeneity in antigenic immune response between individuals has been noted previously, also by our recent studies<sup>30,31</sup>. One of the factors shaping the individual heterogeneity of immune response is associated with microbial symbiosis with the host and their antagonistic to mutualistic associations<sup>37</sup>. Among oral bacteria, these include health-associated early-stage, moderately pathogenic medium-stage and highly pathogenic late-stage colonisers of periodontal biofilm<sup>37,38</sup>. Our data reveal that many strongly targeted epitopes could potentially mimic and the antibodies could cross-react with the antigens of periodontal bacteria (Fig. 2, Supplementary Table S2, Supplementary Fig. S4). Other factors influencing individual variability in immune response to pathogens are age, lifestyle (diet, smoking, exercise etc), previous immune history (viral, bacterial), and the specific HLA-alleles that affect the presentation of major antigenic epitopes<sup>39</sup>. As some of these periodontal bacteria belong to orange or red complex groups of pathogenic oral species, the observed potential cross-reactivity of the antibody immune response (Fig. 2, Supplementary Fig. S4) could contribute to the dysbiosis of oral microbiota and thereby periodontal health.

We identified immunoprofiles stratifying the individuals with ACS from individuals with no CAD or stable CAD (Fig. 3). Our dataset has the advantage that we could identify and compare the presence of potential immunological markers in different CAD types, even in a limited study cohort. When the host's B cell

response is inefficient against re-surfacing of latent infections, this may cause endothelial inflammation which in turn can contribute to the formation of atherosclerotic plaques and their instability<sup>40</sup>. Furthermore, differences were also observed in antibody response against epitopes in specific pathogens, including herpesviruses, in gingivitis, smoking or diabetes subgroups (Fig. 3, Supplementary Fig. S8). Diabetes along with smoking are two big risk factors for periodontal disease<sup>41</sup>. Thus, despite the small scale of the study and varied pathology background, our results provide proof of principle that different stages of CAD may be identifiable by different features of systemic immunoprofiles, which also include antibody response against highly antigenic epitopes of oral microbiota and common viruses.

Our data show that patients with progressing gingivitis through to periodontitis have increasing levels of antibodies to the highly antigenic epitope mimicking EBV VP26 (Figs. 3c and 4a, b). On the other hand, within subjects with periodontitis, a strong response to the EBV VP26 epitope is characteristic to healthy subjects, but not to ACS patients, suggesting its protective role against ACS (Fig. 5b). A few studies have directly addressed the role of herpesvirus infections in susceptibility to secondary oral bacterial infections. Relatedly, it is known that *A. actinomycetemcomitans* and *P. gingivalis* might require support from active herpesviruses for periodontal destruction, whereas the stable periodontal lesions may be devoid of viruses<sup>42</sup>. The herpesviral–bacterial hypothesis of periodontitis proposes that the herpesvirus infection triggers a release of proinflammatory



**Fig. 5** Immune response to epitope biomarkers to predict the ACS risk in periodontal disease and ascertain the no-CAD phenotype. **a** Strength of immune response to three epitopes (in rows) across study subjects ( $n = 96$ , in lanes), grouped by periodontal (H, G, or P) and CAD diagnosis (no-CAD, s-CAD, or ACS) shown as intensity plots. Relative abundance of immunoprofile features is shown in blue color-scale. These epitope biomarkers in a 3-biomarker generalized logistic model differentiate P group from H (Supplementary Fig. S11). Vertical lanes – subjects ( $n = 96$ ); rows – epitopes; blue color-scale – normalized relative epitope-containing peptide abundance values (using 97.5<sup>th</sup> percentile values per feature and capped at value 1). **b** Different immunoprofile features (above boxplots) stratified ACS with periodontal diagnosis. y-axes: immunoreactivity to epitopes in individual immunoprofiles, shown as abundance of peptides containing the epitope. Mann-Whitney  $U$  test, two-sided,  $p$  values not corrected for multiple comparisons, \*\* $p < 0.01$ , \* $p < 0.05$ . H:  $n = 21$  independent patients; G:  $n = 27$  independent patients; P:  $n = 48$  independent patients. **c** Sensitivity and specificity measures of using biomarkers from B in separate periodontal groups (sub-group) to predict CAD diagnoses (either no-CAD or ACS). *Diagnosis  $\geq$  threshold* – diagnosis group into which patient was classified if immune response was over threshold. **a–c** Abbreviations and group sizes: H – periodontally healthy ( $n = 21$ ), G – gingivitis ( $n = 27$ ), P – periodontitis ( $n = 48$ ), no-CAD – no CAD diagnosis ( $n = 32$ ), s-CAD – stable CAD ( $n = 32$ ), ACS – acute coronary syndrome ( $n = 32$ ). Data in File S7.

cytokines to activate osteoclasts and matrix metalloproteinases (MMPs) to impair antibacterial immune mechanisms, causing an up-growth of periodontopathic bacteria<sup>42</sup>. Given that the extracellular matrix (ECM) breakdown represents a crucial factor in the periodontal pathophysiology, periodontitis confounded diagnostics of ACS has been proposed by measuring serum levels of MMP-9<sup>43</sup>, while serum MMP-8 and TIMP-1 levels were found to be associated with incident<sup>44</sup> and especially fatal cardiovascular events<sup>45</sup>. Overall, this indicates immune hyperresponsiveness to dysbiosis, typical to the pathogenesis of periodontitis<sup>46</sup>. On the other hand, periodontal herpesviruses themselves may disseminate via the systemic circulation to non-oral sites (including arteries) and thus represent a major link between periodontitis and cardiovascular diseases<sup>41</sup>. A recent meta-report described a world-wide association of EBV infection with periodontal disease<sup>47</sup>. Here we have investigated this observation further and fine-mapped the EBV VP26 epitope, against which the immunologic response could link periodontitis and associated CAD conditions. However, herein we show that it mimics antigens of periodontal bacteria (Fig. 2d, Supplementary Table S2). Microbial epitope similarity with pathogens, allergens and auto-antigens has been reported before, as it can elicit tolerogenic or inflammatory immune reactivity<sup>48</sup>. For example, different periodontopathogenic species share mimicry in their GroEL antigen, which in turn results in cross-reactivity with the human heat-shock proteins (HSP) expressed on the endothelial cell leading to endothelial dysfunction<sup>49</sup> and atherosclerosis<sup>50</sup>. This suggests that further studies of the herpesviral-bacterial-host epitope mimicry are warranted, in particular for improved diagnostics and therapy of dental health associated heart conditions.

Here we built a two-step biomarker model based on immune response to 4 epitopes that allows a) to classify subjects based on their periodontal diagnosis and b) to predict ACS risk and establish the no-CAD phenotype with 71–89% specificity (Fig. 5). Specific periodontitis-associated biomarkers for CAD could be beneficial in discerning ACS from other cardiac events. These data indicate that periodontitis, and ultimately putative progressing to ACS due to periodontitis, is the result of a partial response or lack of an efficient combined response against viral and bacterial infections landing on self-proteins.

Taken together, our findings clearly illustrate the power of MVA for the immunopathological analysis of oral health-related cardiac conditions, and we predict that the widespread use of this technology at scale will enhance the current understanding of chronic disease mechanisms, in particular cardiovascular diseases, and can lead to improved diagnostic accuracy and new markers.

**Limitations of the study.** We could not evaluate the prognostic value of these predicted biomarkers due to the clinical study design. Because of genetic and socioeconomic status variabilities in different study populations, it is hard to extrapolate the findings. This was a study on Finnish adults. New studies with other cohorts are needed. Also, the pathogenetic importance of the specific oral bacterial microbiota antigens remains firmly to be established.

## Methods

**Ethics statement.** The study was conducted in accordance with the guiding principles of the Declaration of Helsinki and the study participants gave written informed consent before enrollment. The study was approved by the ethics

committees of the Helsinki University Central Hospital (approval reference number 106/2007) and The National Institute for Health Development, Estonia (approval number 1045).

**Clinical cohort description.** The cohort ( $n = 96$ ) was selected from the initial Corogene study ( $n = 5294$ ) and divided into 3 subgroups: periodontally healthy (H,  $n = 21$ ), gingivitis positive (G,  $n = 27$ ), and periodontitis (P,  $n = 48$ ). The diagnostic features on clinical and radiographic findings have been described in detail earlier<sup>51,52</sup>. Periodontally healthy patients had no alveolar bone loss (ABL) and bleeding on probing (BOP) did not exceed 10%. Gingivitis was registered in patients without ABL but with BOP > 10%. Patients were diagnosed with periodontitis when the ABL exceeded the cervical third of the root. Coronary artery disease diagnosis (no coronary artery disease (no-CAD,  $n = 32$ ), stable coronary artery disease (stable-CAD,  $n = 32$ ), or acute coronary syndrome (ACS,  $n = 32$ )) was based on the degree of stenosis in the coronary arteries during the angiography, typical electrocardiographic changes, chest pain, and levels of cardiac biomarkers<sup>32</sup>. The age and gender proportions, along with other relevant clinical history, are in Supplementary Table S1.

**Statistics and reproducibility.** The statistical analyses performed during the study were accompanied by measures of statistical significance. The study was non-blinded and non-randomized and included  $n = 96$  independent study subjects. Group-wise parameters, such as median values, were visualized alongside intra-group range using violin- or boxplots. Reproducibility of Mimotope Variation analysis was confirmed by establishing the correlation coefficient of two replicates as  $R = 0.87$  ( $p < 0.0001$ ).

**Statistical analysis of clinical characteristics of samples.** Differences in proportions of genders, diabetes condition and smoking status were assessed in clinical sub-groups using Chi<sup>2</sup> test (MedCalc, 19.7.2, MedCalc Software Ltd, Ostend, Belgium; <https://www.medcalc.org>). For statistical analysis, two-sided Mann-Whitney  $U$  test was used for comparing two groups or two-sided Kruskal-Wallis test for >2 groups using R package “ggpubr”<sup>53</sup> and “ggplot2”<sup>54</sup>.

**Mimotope variation analysis.** Peptide antigens were selected from random peptide phage modified library (PhD12, NEB) with 10<sup>9</sup> different 12-mer peptide sequences<sup>30,31</sup>. Two  $\mu$ l of serum/plasma samples, previously precleared to plastic and *E. coli*/wt M13 phage lysates were incubated with 2.5  $\mu$ l library ( $\sim 5 \times 10^{11}$  phage particles) and immunoglobulin G (IgG) fraction was recovered using protein G-coated magnetic beads (S1430S, NEB). Captured phage DNA was analyzed by Illumina HiSeq sequencing of 50-bp single end reads using barcoded primers for sample multiplexing. Peptide abundance correlation coefficient (R) in two replicates by Pearson analysis was 0.87 ( $p < 0.0001$ ) (Supplementary Fig. S1) (R package “ggpubr”<sup>53</sup>). For further data analysis, sequencing errors and known artefacts were eliminated.

**Selecting peptides.** Group-enriched peptides (TopPeptide sets) were selected for clinical sub-groups (Supplementary Table S1). Peptides were selected for no-CAD ( $n = 292,667$ ), stable-CAD ( $n = 279,020$ ), and ACS ( $n = 308,445$ ), and periodontally healthy ( $n = 450,531$  peptides), gingivitis ( $n = 450,590$ ), and periodontitis ( $n = 342,261$ ) groups, using the criteria that these were to be identified in  $\geq 10\%$  individual samples of the group with abundance threshold  $\geq 10$  sequence counts in at least one sample.

**Sequence-based unsupervised clustering of peptide antigens.** Exhaustive sequence pattern search tool SPEXS2 was used (<http://egonelbre.github.io/spexs2/>) for sequence-based unsupervised clustering of peptide antigens. Starting from the 292,667 group-enriched peptides (TopPeptide set) identified for no-CAD group, all were used as input to SPEXS2 in random order with the search criteria: peptide coverage threshold:  $\geq 4$ ; motif coverage threshold:  $\geq 4$  fixed amino acid positions; hyper-geometric  $P$  value  $< 10^{-5}$ . Sequence pattern searches were performed in 2 iterative runs, where peptides from which a consensus was identified in the first run were excluded from the subsequent run. As a result, 4366 distinct motif consensus sequences were identified, which were contained in 29.0% or 84,873 of the original 292,667 peptides. Therefore, 4366 unique consensus motifs were identified for no-CAD (covering 29.0% of input peptides), similarly 2771 motifs were calculated for stable-CAD (27.0%), 4405 for ACS (30.5%), 6275 for periodontally healthy (4 iterative SPEXS2 runs due to greater starting peptide set) (42.8%), 9560 for gingivitis (4 runs) (44.4%), and 5936 for periodontitis group (4 runs) (27.7%) were defined.

**Selecting for group-differentiating motifs.** The epitope motifs contained high degree of redundancy, therefore stricter criteria (hyper-geometric  $P$  value  $< 10^{-8}$  or query/reference ratio  $\geq 10$ ) were imposed to select for characteristics with high significance and statistical power. Altogether 8088 unique motif sequences fit those criteria were selected for further analyses as the TopMotif set. Of 8088 distinct motif features, 995 were designated as group-differential. The 995 motifs satisfied all the criteria that 1) the average abundance value in each clinical sub-group was >3 greater than in another relevant sub-group, 2) the abundance-based separation of relevant clinical sub-groups was with  $\geq 50\%$  sensitivity and  $\geq 70\%$  specificity, and 3)

group-separation was statistically significant (Mann-Whitney  $U$  test,  $p$  value < 0.05, two-sided).

**Alignment profiles on periodontal bacteria.** The TopMotif set of 8088 motifs was aligned to proteomes of the periodontal pathogen species: *Porphyromonas gingivalis* (UniProt accession: UP000000588), *Tannerella forsythia* (UP000005436), *Prevotella intermedia* (UP000010099), *Fusobacterium nucleatum* (UP000002521), *Campylobacter rectus* (UP000003082), *Aggregatibacter actinomycetemcomitans* (UP000002569), and *Porphyromonas endodontalis* (UP000004295) (accessed 20-21.12.2019). Only exact alignments where all fixed amino acid positions of a motif (minimum of 4 positions) to match with the target were allowed. The pathogen database comprised of 15,928 proteins, of which 15,116 matched at least with one epitope motif. For each pathogen, the individual alignment load was calculated per 20-amino acid fragment of the protein, in two frameshifts (0 and -10 aa), yielding in 10 aa overlaps between considered fragments. Altogether we analyzed  $\sim 480,000$  distinct 20 amino acid fragments. The top 40 fragments for each bacterial species, from any frameshifts, with the highest total alignment loads were selected based on the formula:

$$\text{Alignment load (per 20aa fragment)} = \frac{\text{sum of abundance of motifs aligned}}{\text{count of motifs aligned}}$$

If higher alignment loads were encompassing two side-by-side regions (one with a frameshift), a longer 30-amino acid fragment was considered. Individual alignment loads (Supplementary Data 1) were compared across pathogen species using Kruskal-Wallis test for comparing all groups and two-sided Mann-Whitney  $U$  test for pair-wise comparisons. R packages “ggpubr” and “ggplot2” were used for calculation and visualization.

**Identifying target types of potential epitopes of periodontal bacteria.** Using the top 40 antigen fragments of periodontal pathogens, 12mer substrings were extracted for each aligned epitope motif (of the TopMotif set), whereas substrings shorter than 12mer were discarded (33 were found). Altogether 1691 substrings were selected, position weight matrices were built and the resulting enriched amino acid positions were visualized as sequence logos, all using a custom in-house tool with WebLogo<sup>55,56</sup> integration (parameters: no counts, distance cut-off 9, 10 minimum unique substrings in a cluster, similarity index 15). The analysis yielded 26 target types with consensus sequences, of which top 5 most prevalent are shown on Fig. 2b.

**Amino acid-based clustering of group-differential epitopes.** Sequence homology clustering analysis reduced 705 out of 995 epitope motifs into 62 clusters of similar epitopes ( $\geq 3$  identical amino acid positions) whereas 290 out of 995 epitope motifs either formed too small clusters (<3 motifs in a cluster) or were not similar enough with any other epitope (<3 identical amino acid positions). Additional R packages used in data analysis and visualization were: “readr”, “dplyr”<sup>57</sup>.

Based on the selected peptides, position weight matrices were built and the resulting enriched amino acid positions were visualized as sequence logos, all using a custom in-house tool with WebLogo<sup>55,56</sup> integration.

**Group-specific analysis of 62 clusters.** The average abundance for each of the 62 clusters with the defined sequence logo was calculated across clinical sub-groups. Abundance values in clinical sub-groups were normalized with average abundance across all groups (Supplementary Data 2). Clusters with average abundance values of <150 were left out of further analysis. Next, hierarchical clustering was performed based on the normalized abundance values using Pearson correlation coefficients for clustering distance and ward.D2 clustering method (R package “pheatmap”)<sup>58</sup>. R packages used in data analysis and visualization included “readr”, “dplyr”, “forcats”, “reshape2”<sup>54,57</sup>, and “viridis”<sup>59</sup>.

**Individual-based clustering of 62 clusters.** Abundance of peptides in individual immunoprofiles was calculated for each cluster. Using log10 values of the abundances, Pearson correlation coefficients R were calculated pair-wise for all 62 clusters (Supplementary Data 3). The clusters were subsequently grouped using R package “pheatmap” with ward.D2 as clustering method and Pearson correlation coefficients as clustering distance. Five distinct larger groups were identified and defined as epitopes A-E. For each of the epitopes A-E, the containing clusters were compared to determine consensus sequences describing the core epitopes (Supplementary Data 7).

**Alignment on EBV VP26.** Epitope A has been previously mapped to Epstein-Barr virus (EBV) protein VP26<sup>30</sup>. To validate the mapping in the current clinical cohort, epitopes from individual immunoprofiles ( $n = 96$ ) were aligned to the primary sequence of EBV protein VP26, a component of the viral capsid antigen (VCA) (UniProt accession code Q3KSU9, EBV strain GD-1, date accessed: 25.11.2020). Signal-to-random ratio for each sample was calculated across primary sequence of EBV VP26, where signal represented the count of aligned ( $\geq 6$  matching amino acid positions) unique peptides defined by MVA and random represented the count of aligned random peptides ( $\geq 6$  matching), generated by sequence scrambling of



peptides (Supplementary Data 4). As controls (Ctrl), peptides from EBV CA seronegative subjects ( $n = 9$ ) were analyzed similarly.

**ELISA.** CMV and EBV serostatus was measured from serum samples with ISO/IEC 17025 accredited methods. In brief, serological analyses were performed with anti-CMV IgG ELISA method (EUROIMMUN EI 2570-9601G) and with anti-EBV CA (capsid antigen) IgG ELISA method (EUROIMMUN EI 2791-9601G) according to manufacturer's specifications. Absorbance was measured at 450 nm with SpectraMax Paradigm (Molecular Devices).

**Dot ELISA.** MVA analysis findings were validated by dot ELISA analysis. For that, M13K phages displaying peptides with either the EBV VP26 epitope-containing sequence TLPMDTSPRAHW or the mutant sequence TLPMDASPRAHW as parts of the pIII protein were printed onto nitrocellulose (NC) slides (ArrayIt, US). Unspecific binding was reduced by blocking for 1 h at room temperature with 5% non-fat dried milk in 1xPBS-Tween20-0.05%. Human serum samples were pre-cleared to reduce unspecific binding to M13K phages, to *E. coli* bacterial proteins and to plastic. Preclearing step was performed in mix of 60  $\mu$ l 2.5% skimmed milk-1xPBS-0.05% Tween20 + 30  $\mu$ l Preabsorption Solution + 1:50 serum, at 4 °C overnight. Following this, the slides were incubated with either precleared serum 1:50 solutions ( $n = 54$ ) or with 1:5000 mouse anti-M13 antibody (#27-9420-01, GE healthcare) in 2.5% skimmed milk in 1xPBS-0.05% Tween20 (GE Healthcare) for 1 h at room temperature. The anti-M13K antibodies were used to quantify phages printed on NC slides. Multiple washes were performed with 5% skimmed milk in 1xPBS-0.05% Tween20. For visualization, the secondary antibodies used were 1:1000 rabbit anti-human HRP-conjugated antibody (#ab6759, Abcam) (for human serum samples) in 2.5% skimmed milk in 1xPBS-0.05% Tween20 or 1:1000 rabbit anti-mouse HRP-conjugated antibody (#ab6728, Abcam) (for anti-M13K) in 2.5% skimmed milk in 1xPBS-0.05% Tween20, incubated for 1 h at room temperature. After multiple washes, the presence of bound human sera/plasma IgG antibodies was detected via reaction with HRP substrate DAB chromogen diluted in substrate buffer (1:100). The slides were digitally scanned and the signals were quantified using ImageQuantTL (version 8.1) (Supplementary Data 5). Results of 27/54 samples were validated in an independent similar experiment (in total  $n = 2$  experiments) (including all samples which showed high signal-to-background ratios).

**Predicting periodontitis and CAD diagnoses.** Generalized linear model was fit to 80% of subjects' data (with 5x cross-validation) to classify different case-subgroups based on their immunoprofile features. Using model's prediction probabilities for subjects in the training set, receiver operating characteristic (ROC) analysis was performed. The model's area under the receiver operating characteristic curve (AUROC) was 0.843 with 95% CI (0.738...0.948). The model was validated on the validation subset (20% of samples).

**Data visualization.** Box and whisker plots were generated in the style of Tukey with R packages "ggpubr" or "ggplot2". Upper, middle and lower boxplot lines represent the 75th, 50th and 25th percentiles, while whiskers represent the largest or smallest value within 1.5 times interquartile range above the 75th percentile or below the 25th percentile, respectively. Individual data points without outliers are visualized.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The source data analyzed during the study was generated by MVA analysis. We provided the relevant data underlying the main findings in the Supplementary data. The whole datasets generated and/or analyzed during the current study are not publicly available due to containing sensitive clinical information but are available from the corresponding author on reasonable request.

## Code availability

The code used during the current study is not publicly available due to its proprietary nature, but detailed explanation of the analysis approaches is available from the corresponding author on reasonable request.

Received: 2 September 2021; Accepted: 8 February 2022;

Published online: 04 March 2022

## References

- Benjamin, E. J. et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. *Circulation* **139**, e56–e528 (2019).

- Madamanchi, N. R., Hakim, Z. S. & Runge, M. S. Oxidative stress in atherogenesis and arterial thrombosis: the disconnect between cellular studies and clinical outcomes. *J. Thromb. Haemost.* **3**, 254–267 (2005).
- Pant, S. et al. Inflammation and atherosclerosis-revisited. *J. Cardiovasc. Pharm. Ther.* **19**, 170–178 (2014).
- Hansson, G. K. Inflammation, atherosclerosis, and coronary artery disease. *N. Engl. J. Med.* **352**, 1685–1695 (2005).
- Catapano, A. L., Pirillo, A. & Norata, G. D. Vascular inflammation and low-density lipoproteins: is cholesterol the link? A lesson from the clinical trials. *Br. J. Pharm.* **174**, 3973–3985 (2017).
- Voight, B. F. et al. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* **8**, e1002793 (2012).
- Miller, C. L., Assimes, T. L., Montgomery, S. B. & Quertermous, T. Dissecting the causal genetic mechanisms of coronary heart disease. *Curr. Atheroscler. Rep.* **16**, 406 (2014).
- McPherson, R. & Hegele, R. A. The sum of its parts: the polygenic basis of coronary artery disease. *Can. J. Cardiol.* **32**, 1372–1374 (2016).
- Nikpay, M. et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet* **47**, 1121–1130 (2015).
- Vernon, S. T. et al. Utilizing state-of-the-art “omics” technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease. *Microcirculation* **26**, e12488 (2019).
- Kott, K. A. et al. Single-cell immune profiling in coronary artery disease: the role of state-of-the-art immunophenotyping with mass cytometry in the diagnosis of atherosclerosis. *J. Am. Heart Assoc.* **9**, e017759 (2020).
- Fernandez-Alvira, J. M. et al. Predicting subclinical atherosclerosis in low-risk individuals: ideal cardiovascular health score and fuster-BEWAT score. *J. Am. Coll. Cardiol.* **70**, 2463–2473 (2017).
- Nelson, C. P. et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat. Genet.* **49**, 1385–1391 (2017).
- Silverman, M. G. et al. Baseline subclinical atherosclerosis burden and distribution are associated with frequency and mode of future coronary revascularization: multi-ethnic study of atherosclerosis. *JACC Cardiovasc. Imaging* **7**, 476–486 (2014).
- Lockhart, P. B. et al. Periodontal disease and atherosclerotic vascular disease: does the evidence support an independent association?: a scientific statement from the American Heart Association. *Circulation* **125**, 2520–2544 (2012).
- Yumoto, H. et al. Sensitization of human aortic endothelial cells to lipopolysaccharide via regulation of Toll-like receptor 4 by bacterial fimbria-dependent invasion. *Infect. Immun.* **73**, 8050–8059 (2005).
- Kozarov, E. V., Dorn, B. R., Shelburne, C. E., Dunn, W. A. Jr. & Progluske-Fox, A. Human atherosclerotic plaque contains viable invasive *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. *Arterioscler. Thromb. Vasc. Biol.* **25**, e17–e18 (2005).
- Gaetti-Jardim, E., Marcelino, S. L., Feitosa, A. C. R., Romito, G. A. & Avila-Campos, M. J. Quantitative detection of periodontopathic bacteria in atherosclerotic plaques from coronary arteries. *J. Med. Microbiol.* **58**, 1568–1575 (2009).
- Jie, Z. et al. The gut microbiome in atherosclerotic cardiovascular disease. *Nat. Commun.* **8**, 845 (2017).
- Acharya, A. et al. Salivary microbiome in non-oral disease: a summary of evidence and commentary. *Arch. Oral. Biol.* **83**, 169–173 (2017).
- Consortium, C. A. D. et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat. Genet* **45**, 25–33 (2013).
- Grundtman, C. & Wick, G. The autoimmune concept of atherosclerosis. *Curr. Opin. Lipidol.* **22**, 327–334 (2011).
- Garcia-Rivas, G. et al. The role of B cells in heart failure and implications for future immunomodulatory treatment strategies. *ESC Heart Fail.* **7**, 1387–1399 (2020).
- Troseid, M., Andersen, G. O., Broch, K. & Hov, J. R. The gut microbiome in coronary artery disease and heart failure: current knowledge and future directions. *EBioMedicine* **52**, 102649 (2020).
- Lee, N. & Kim, W. U. Microbiota in T-cell homeostasis and inflammatory diseases. *Exp. Mol. Med.* **49**, e340 (2017).
- Kobiyama, K. & Ley, K. Atherosclerosis. *Circ. Res.* **123**, 1118–1120 (2018).
- Davies, R. W. et al. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. *Circ. Cardiovasc. Genet* **5**, 217–225 (2012).
- Tsiantoulas, D., Diehl, C. J., Witztum, J. L. & Binder, C. J. B cells and humoral immunity in atherosclerosis. *Circ. Res.* **114**, 1743–1756 (2014).
- Wolf, D. & Ley, K. Immunity and inflammation in atherosclerosis. *Herz* **44**, 107–120 (2019).
- Sadam, H. et al. Identification of two highly antigenic epitope markers predicting multiple sclerosis in optic neuritis patients. *EBioMedicine* **64**, 103211 (2021).
- Sadam, H. et al. Prostaglandin D2 receptor DP1 antibodies predict vaccine-induced and spontaneous narcolepsy type 1: large-scale study of antibody profiling. *EBioMedicine* **29**, 47–59 (2018).

32. Vaara, S. et al. Cohort profile: the Corogene study. *Int J. Epidemiol.* **41**, 1265–1271 (2012).
33. Liljestrand, J. M. et al. Association of endodontic lesions with coronary artery disease. *J. Dent. Res.* **95**, 1358–1365 (2016).
34. Liljestrand, J. M. et al. Immunologic burden links periodontitis to acute coronary syndrome. *Atherosclerosis* **268**, 177–184 (2018).
35. Falk, K. et al. Synthetic peptides deduced from the amino acid sequence of Epstein-Barr virus nuclear antigen 6 (EBNA 6): antigenic properties, production of monoreactive reagents, and analysis of antibody responses in man. *J. Med. Virol.* **46**, 349–357 (1995).
36. Loebel, M. et al. Serological profiling of the EBV immune response in Chronic Fatigue Syndrome using a peptide microarray. *PLoS One* **12**, e0179124 (2017).
37. Lamont, R. J., Koo, H. & Hajishengallis, G. The oral microbiota: dynamic communities and host interactions. *Nat. Rev. Microbiol.* **16**, 745–759 (2018).
38. Socransky, S. S. & Haffajee, A. D. Dental biofilms: difficult therapeutic targets. *Periodontol.* **2000** **28**, 12–55 (2002).
39. Brodin, P. & Davis, M. M. Human immune system variation. *Nat. Rev. Immunol.* **17**, 21–29 (2017).
40. Campbell, L. A. & Rosenfeld, M. E. Infection and atherosclerosis development. *Arch. Med Res* **46**, 339–350 (2015).
41. Tsuchida, S. Proteome analysis of molecular events in oral pathogenesis and virus: a review with a particular focus on periodontitis. *Int. J. Mol. Sci.* **21**, <https://doi.org/10.3390/ijms21155184> (2020).
42. Slots, J. Human viruses in periodontitis. *Periodontol.* **2000** **53**, 89–110 (2010).
43. Liukkonen, J. et al. Immunological and microbiological profiling of cumulative risk score for periodontitis. *Diagnostics* **10**, <https://doi.org/10.3390/diagnostics10080560> (2020).
44. Tuomainen, A. M. et al. Serum tissue-degrading proteinases and incident cardiovascular disease events. *Eur. J. Prev. Cardiol.* **21**, 806–812 (2014).
45. Kormi, I. et al. Matrix metalloproteinase-8 and tissue inhibitor of matrix metalloproteinase-1 predict incident cardiovascular disease events and all-cause mortality in a population-based cohort. *Eur. J. Prev. Cardiol.* **24**, 1136–1144 (2017).
46. Kinane, D. F., Stathopoulou, P. G. & Papapanou, P. N. Periodontal diseases. *Nat. Rev. Dis. Prim.* **3**, 17038 (2017).
47. Gao, Z., Lv, J. & Wang, M. Epstein-Barr virus is associated with periodontal diseases: a meta-analysis based on 21 case-control studies. *Medicine* **96**, e5980 (2017).
48. Carrasco Pro, S. et al. Microbiota epitope similarity either dampens or enhances the immunogenicity of disease-associated antigenic epitopes. *PLoS One* **13**, e0196551 (2018).
49. Wang, C. et al. Characterization of a natural mouse monoclonal antibody recognizing epitopes shared by oxidized low-density lipoprotein and chaperonin 60 of *Aggregatibacter actinomycetemcomitans*. *Immunol. Res.* **64**, 699–710 (2016).
50. Goulhen, F., Grenier, D. & Mayrand, D. Oral microbial heat-shock proteins and their potential contributions to infections. *Crit. Rev. Oral. Biol. Med.* **14**, 399–412 (2003).
51. Buhlin, K. et al. Periodontitis is associated with angiographically verified coronary artery disease. *J. Clin. Periodontol.* **38**, 1007–1014 (2011).
52. Hyvarinen, K. et al. A common periodontal pathogen has an adverse association with both acute and stable coronary artery disease. *Atherosclerosis* **223**, 478–484 (2012).
53. Kassambra, A. ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. <https://CRAN.R-project.org/package=ggpubr> (2020).
54. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. <https://ggplot2.tidyverse.org>. (Springer-Verlag New York, 2016).
55. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
56. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* **18**, 6097–6100 (1990).
57. Wickham, H., Averick, M., Bryan, J., Chang, W. & McGowan, L. Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 43 (2019).
58. pheatmap: Pretty Heatmaps. R package version 1.0.12. <https://doi.org/10.21105/joss.01686> (2019).
59. viridis: Default Color Maps from 'matplotlib'. R package version 0.5.1. <https://sjmgarnier.github.io/viridis/> (2018).

## Acknowledgements

We thank A.-H. Pool (Department of Neuroscience, University of Texas Southwestern Medical Center) for his expertise that supported this work. The members of Protobios team are thanked for their excellent technical assistance. This study was supported by research funding grants of Protobios (5.1-4/20/170, and PRG573) from the Estonian Ministry of Education and Estonian Research Council, respectively, and H2020-MSCA-RISE-2016 (EU734791) and H2020 PANBioRA (EU760921) projects from the European Union. K.P. was partially supported by the European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012). N.E.V. was supported by H2020 PANBioRA (EU760921) project from the European Union. P.P. was supported by the Sigrid Jusélius Foundation.

## Author contributions

M.J., N.P., A.R., A.P., H.S., and K.P. conceived and designed the study. J.S. and P.P. collected samples and acquired the clinical data. M.J., N.P., A.R., A.P., H.S., and K.P. analyzed and interpreted the data. N.E.V., J.S., P.P., and K.P. supervised the study. M.J. and K.P. drafted the manuscript. All authors revised and approved the final manuscript for submission.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-022-03122-4>.

**Correspondence** and requests for materials should be addressed to Kaia Palm.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Anam Akhtar.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**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) 2022