Patient-Specific Variations in Biomarkers across Gingivitis and Periodontitis

Radhakrishnan Nagarajan
University of Kentucky, rnagarajan@uky.edu

Craig S. Miller
University of Kentucky, craig.miller@uky.edu

Dolph Dawson
University of Kentucky, dolph.dawson@uky.edu

Mohanad Al-Sabbagh
University of Kentucky, malsa2@uky.edu

J. L. Ebersole
University of Kentucky, jeffrey.ebersole@uky.edu

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/biostatistics_facpub

Part of the Biostatistics Commons

Repository Citation

Patient-Specific Variations in Biomarkers across Gingivitis and Periodontitis

Notes/Citation Information
Published in *PLOS One*, v. 10, no. 9, article e0136792, p. 1-15.

RESEARCH ARTICLE

Patient-Specific Variations in Biomarkers across Gingivitis and Periodontitis

Radhakrishnan Nagarajan[1], Craig S. Miller[2,3], Dolph Dawson, III[2,4], Mohanad Al-Sabbagh[4], Jeffrey L. Ebersole[2]*

1 Division of Biomedical Informatics, Department of Biostatistics, College of Public Health, University of Kentucky, Lexington, KY, United States of America, 2 Center for Oral Health Research, College of Dentistry, University of Kentucky, Lexington, KY, United States of America, 3 Division of Oral Medicine, Department of Oral Health Practice, College of Dentistry, University of Kentucky, Lexington, KY, United States of America, 4 Division of Periodontics, Department of Oral Health Practice, College of Dentistry, University of Kentucky, Lexington, KY, United States of America

* Jeffrey.ebersole@uky.edu

Abstract

This study investigates the use of saliva, as an emerging diagnostic fluid in conjunction with classification techniques to discern biological heterogeneity in clinically labelled gingivitis and periodontitis subjects (80 subjects; 40/group) A battery of classification techniques were investigated as traditional single classifier systems as well as within a novel selective voting ensemble classification approach (SVA) framework. Unlike traditional single classifiers, SVA is shown to reveal patient-specific variations within disease groups, which may be important for identifying proclivity to disease progression or disease stability. Salivary expression profiles of IL-1ß, IL-6, MMP-8, and MIP-1α from 80 patients were analyzed using four classification algorithms (LDA: Linear Discriminant Analysis [LDA], Quadratic Discriminant Analysis [QDA], Naïve Bayes Classifier [NBC] and Support Vector Machines [SVM]) as traditional single classifiers and within the SVA framework (SVA-LDA, SVA-QDA, SVA-NB and SVA-SVM). Our findings demonstrate that performance measures (sensitivity, specificity and accuracy) of traditional classification as single classifier were comparable to that of the SVA counterparts using clinical labels of the samples as ground truth. However, unlike traditional single classifier approaches, the normalized ensemble vote-counts from SVA revealed varying proclivity of the subjects for each of the disease groups. More importantly, the SVA identified a subset of gingivitis and periodontitis samples that demonstrated a biological proclivity commensurate with the other clinical group. This subset was confirmed across SVA-LDA, SVA-QDA, SVA-NB and SVA-SVM. Heatmap visualization of their ensemble sets revealed lack of consensus between these subsets and the rest of the samples within the respective disease groups indicating the unique nature of the patients in these subsets. While the source of variation is not known, the results presented clearly elucidate the need for novel approaches that accommodate inherent heterogeneity and personalized variations within disease groups in diagnostic characterization. The proposed approach falls within the scope of P4 medicine (predictive, preventive, personalized, and participatory) with the ability to identify unique patient profiles that may predict specific disease trajectories and targeted disease management.

Introduction

Movement towards an era of personalized medicine and individualized clinical decisions in periodontology requires significant improvement in our ability to define risk and predict disease progression. While the medical field routinely makes clinical diagnoses based on signs and symptoms (*e.g.* pneumonia, diarrhea), decisions on patient management and treatment do not stop there. Modern medicine and its evolution towards P4 medicine (http://www.p4mi.org/) integrates clinical manifestations with biological assessments that enable the physician to focus on specific disease etiology and unique features of the patient in finalizing a treatment strategy. However, current practice in differential diagnosis and treatment options in dentistry often stops at the clinical and radiographic presentation of the patient, despite the expansion of information now available from biological fluids, plaque and tissue.

Gingivitis is a reversible condition associated with bacterial biofilms that generally resolves clinically in about one week after the reinstitution of oral hygiene procedures [1]. It is an often-overlooked disease, despite being the "gateway" to periodontitis for a significant portion of the population [2–5]. Gingivitis affects a large percentage of the U.S. population [6, 7] and is generally accepted that, if left untreated, may ultimately progress to periodontitis in a subset of individuals.[4, 8] In fact, a recent report by Lang et al. [2] focusing on gingivitis as a risk factor in periodontal disease identified that approximately 37% of teeth with consistent gingivitis progressed to periodontitis and tooth loss, while non-inflamed teeth showed a 99.5% survival. Additionally, the heterogeneity of these disease populations is evident in that various treatment studies of both gingivitis and periodontitis have shown that a portion of the gingivitis population does not respond well to standard mechanical therapy [9] nor do they return to biological health in a timely manner [10], and a portion of the periodontitis population remains somewhat refractory to treatment using standard approaches [9].

Differential host responses are the basis for various levels of susceptibility in chronic inflammatory diseases, such as periodontitis [11–13]. Engagement of host cell receptors with bacterial components activate immune and nonimmune cells to secrete an array of cytokines and inflammatory mediators, *eg.* interleukin(IL)-1ß, IL-6, and tumor necrosis factor (TNF)α, that results in production of matrix metalloproteinases (MMPs) that undermine the integrity of the gingival tissues[14]. These biomolecules of inflammation can be measured in oral fluids [15, 16], and saliva [17–19], supporting their potential for diagnostic utility regarding oral health and disease. This study selected 4 of these salivary analytes for evaluation based upon literature documenting their detection and likely role in periodontal disease pathology. IL-1ß has been identified in gingival crevicular fluid [GCF, [20]] and saliva [10, 21, 22] in elevated levels and has been related to disease extent/severity [23] and decreasing levels with successful therapy [24]. Increased IL-6 levels have also been shown in GCF [25] and saliva [26] in periodontitis, as well as in periodontitis tissues [27]. As with IL-1ß, treatment studies show decreased levels of IL-6 with improved clinical parameters [28]. MMP-8 (neutrophil collagenase) is increased at sites of inflammation in both GCF (20) and saliva [10, 21, 26] in periodontitis and has been suggested as a potential diagnostic molecule for the disease [16]. Finally, macrophage inflammatory protein (MIP)-1α (CCL3) is a chemokine produced by many cell types linked to inflammatory lesions [29, 30] and stimulates monocytes and/or osteoclast progenitor cells to become active osteoclasts [31]. Few studies are available regarding this biomolecule in saliva, although a single study indicated that MIP-1α is elevated in saliva of aggressive periodontitis patients [32] and we have shown increased concentrations in chronic periodontitis patients [33].

Recent studies have provided novel insights into changes in molecular signatures across the spectrum of periodontal disease [34–41]. Classification techniques have also been used for

investigating the predictive potential of molecular markers in discerning distinct disease groups [36]. However, a majority of existing approaches have clear limitations: (a) Existing techniques traditionally do not accommodate or provide insights into potential patient-specific variations within a disease group while discerning distinct disease groups. (b) Pre-defined thresholds and filters are often used as pre-processing in eliminating patient samples that do not conform to a given criteria in an attempt to homogenize the samples within a disease group. While helpful in reducing the dimensionality, such filters can eliminate natural variations within a disease group. (c) They aggregate the expression profiles across samples within each group. However, these in turn may minimize differences between distinct disease groups, as well as prevent identifying potential subpopulations with unique characteristics in the groups. (d) More importantly, a majority of the existing approaches use the same set of fixed markers as representative of each sample within and between the groups of interest and use a single classifier in discerning the disease groups of interest. Ensemble approaches that use a battery of classifiers as opposed to a single classifier have been shown to perform better [42–46]. In this report, we use a selective voting ensemble classification approach (SVA) [47, 48] in conjunction with a battery of classification techniques and salivary biomarkers (IL-1ß, IL-6, MIP-1α, MMP-8) to discern clinically-labeled gingivitis and periodontitis groups. The four salivary biomarkers are targeted because they correspond to the underlying biological processes that occur during gingivitis and periodontitis (i.e., inflammation, tissue destruction and bone remodeling) [49, 50]. Thus, their use in conjunction with the SVA approach is hypothesized to reveal patient-specific variations within disease groups and potentially important for identifying an increased risk for disease progression.

Materials and Methods

Human Subjects and Clinical Evaluation

Case-control studies were conducted at the University of Kentucky, College of Dentistry from 2009 through 2013 with protocols approved by the Institutional Review Board. Participants were recruited from the clinic populations of the College of Dentistry. Each participant was given verbal and written information that described the nature of the study, and each signed informed consent prior to enrollment of the study. Inclusion criteria included subjects older than 18 years of age who were in good general health (excluding the case definition) and had a minimum of 20 teeth. BOP (bleeding on probing), PPD (probing pocket depth), and CAL (clinical attachment level) were measured using a full mouth periodontal examination at six locations per tooth (mesial-buccal, mid-buccal, distal-buccal, mesial lingual, mid-lingual, and distal-lingual). After the measurement of PPDs, all sites were observed for bleeding on probing BOP [51].

Two oral disease groups were analyzed. The 40 subjects in the gingivitis group (age: 27.5 ±4.5; 50% female) had BOP at $\geq$20% of sites (6 sites per tooth), <10% of sites with PPD $\geq$4 mm, and no sites [52] $\geq$2 mm. An equal number of subjects (n = 40) in the periodontitis group (age: 40.8±10.5; 43% female) had BOP at >20% of sites, with >10% of sites with PPD $\geq$4 mm and CAL $\geq$2 mm. Individuals were excluded from either group if there was evidence of general health issues, chronic inflammatory conditions, pregnancy, use of antibiotics or other mucosal inflammatory disease [53]. While smoking has been clearly identified as a risk factor in the extent and severity of periodontitis [54, 55] it was not included as a feature in the classification approach. Population demographics of the cohorts based on patient-reported information indicated none of the gingivitis patients to be smokers and ~27% of the periodontitis patients as smokers.

Salivary Samples and Molecular Biomarkers

Whole unstimulated saliva was collected on ice between 9 AM and 3 PM from both groups prior to clinical evaluation as we have described previously [53, 56] according to a modification of the method described by Navazesh [57]. Aliquots were prepared and frozen at -80°C until analyzed. The MILLIPLEX MAP Kit (Billerica, MA, USA) was used to detect IL-1ß, IL-6, MMP-8, and MIP-1α as we have described previously [10].

Selective Voting Ensemble Classification (SVA). SVA approach for discerning gingivitis and periodontitis and extracting patient-specific profiles is detailed below. A comparison of the proposed SVA approach and traditional classifications is also depicted in Fig 1 for convenience.

Given. Salivary expression profiles of (m) genes across gingivitis (n_{gin}) and periodontitis (n_{per}) subjects represented by the matrix X_{mxn} where $n = n_{gin} + n_{per}$. Salivary expression profiles were log-transformed and normalized to zero-mean and unit variance. Clinical labels of the n samples based on the three clinical parameters (BOP, PD, CAL) in λ_{1xn} where

$$\lambda(k) = -1 \text{ implies the } k^{th} \text{ sample is gingivitis}$$

$$+1 \text{ implies the } k^{th} \text{ sample is periodontitis}$$

Initialize the voting matrices across the samples as follows:

$$v^g(r,k) \leftarrow 0; v^p(r,k) \leftarrow 0; k = 1 \ldots n, \ r = 1 \ldots N_b$$

Initialize the ensemble sets across the samples as follows:

$$\sigma^g(r,k) \leftarrow \emptyset; \sigma^p(r,k) \leftarrow \emptyset; k = 1 \ldots n, \ r = 1 \ldots N_b$$

where N_b represents the number of independent realizations.

Choose a classification technique (e.g. linear discriminant analysis).

Step 1: Set $r \leftarrow 1$

Partition the given samples X_{mxn} into training V_{mxk2} and test U_{mxk1} sets such that $k_1 + k_2 = n$ and $k_2 >> k_1$.

Step 2: Choose a pair of molecules $(i,j), i,j = 1 \ldots m, i \neq j$.

Step 3: Generate the classifier using the training set for the pair (i,j).

- If the kth test sample is voted as gingivitis then increment the vote $v^g(r,k) \leftarrow v^g(r,k)+1$. Store the pair (i,j) in the ensemble set $\sigma^g(r,k) \leftarrow \emptyset$.

- If the kth test sample is voted as periodontitis then increment the vote $v^p(r,k) \leftarrow v^p(r,k)+1$. Store the pair (i,j) in the ensemble set $\sigma^p(r,k) \leftarrow \emptyset$.

Step 4: Repeat Steps 2 and 3 for each pair of genes (i,j), $i,j = 1 \ldots m, i \neq j$.

Step 5: Repeat Steps 2–4 using the next window of k_1 samples as test set.

Step 6: Determine the classification labels $\lambda^*(k)$ of the kth sample by majority voting as follows
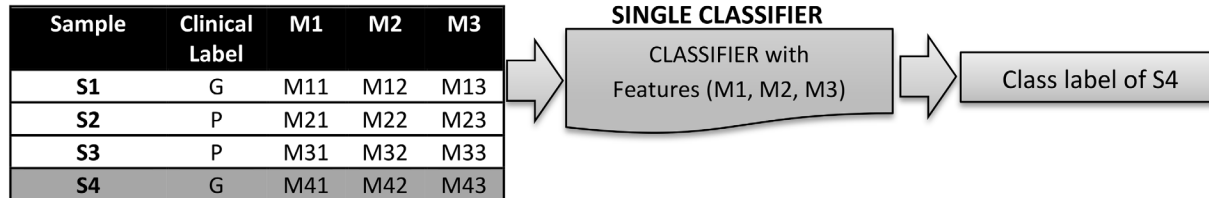
$$\lambda^*(k) = -1 \text{ if } v^g(r,k) > v^p(r,k) \qquad \text{Gingivitis}$$

$$+1 \text{ if } v^g(r,k) < v^p(r,k) \qquad \text{Periodontitis}$$

$$0 \text{ if } v^g(r,k) = v^p(r,k) \qquad \text{Tie}$$

(a) Traditional Classification using a Single Classifier and All Features

- **Training Samples** : S1, S2, S3
- **Test Sample** : S4
- **Feature Set** : M1, M2, M3
- **Classifier** : e.g. LDA, Linear Discriminant Analysis

Sample	Clinical Label	M1	M2	M3
S1	G	M11	M12	M13
S2	P	M21	M22	M23
S3	P	M31	M32	M33
S4	G	M41	M42	M43

SINGLE CLASSIFIER

CLASSIFIER with Features (M1, M2, M3) → Class label of S4

(b) SVA Using Ensemble of Classifiers, Pairs of Features and Selective Voting Strategy

- **Training Samples** : S1, S2, S3
- **Test Sample** : S4
- **Feature Set** : (M1, M2); (M1, M3); (M2, M3)
- **Classifier** : e.g. LDA, Linear Discriminant Analysis

ENSEMBLE OF CLASSIFIERS   **SELECTIVE VOTING**

Sample	Clinical Label	M1	M2
S1	G	M11	M12
S2	P	M21	M22
S3	P	M31	M32
S4	G	M41	M42

CLASSIFIER with Features (M1, M2) → Label of S4 is **P**

Sample	Clinical Label	M1	M3
S1	G	M11	M13
S2	P	M21	M23
S3	P	M31	M33
S4	G	M41	M43

CLASSIFIER with Features (M1, M3) → Label of S4 is **G**

Sample	Clinical Label	M2	M3
S1	G	M12	M13
S2	P	M22	M23
S3	P	M32	M33
S4	G	M42	M43

CLASSIFIER with Features (M2, M3) → Label of S4 is **P**

The ensemble sets are:
σ^G:{(M1,M3)}, i.e. 1 vote
σ^P:{(M1,M2), (M2,M3)} i.e. 2 votes.
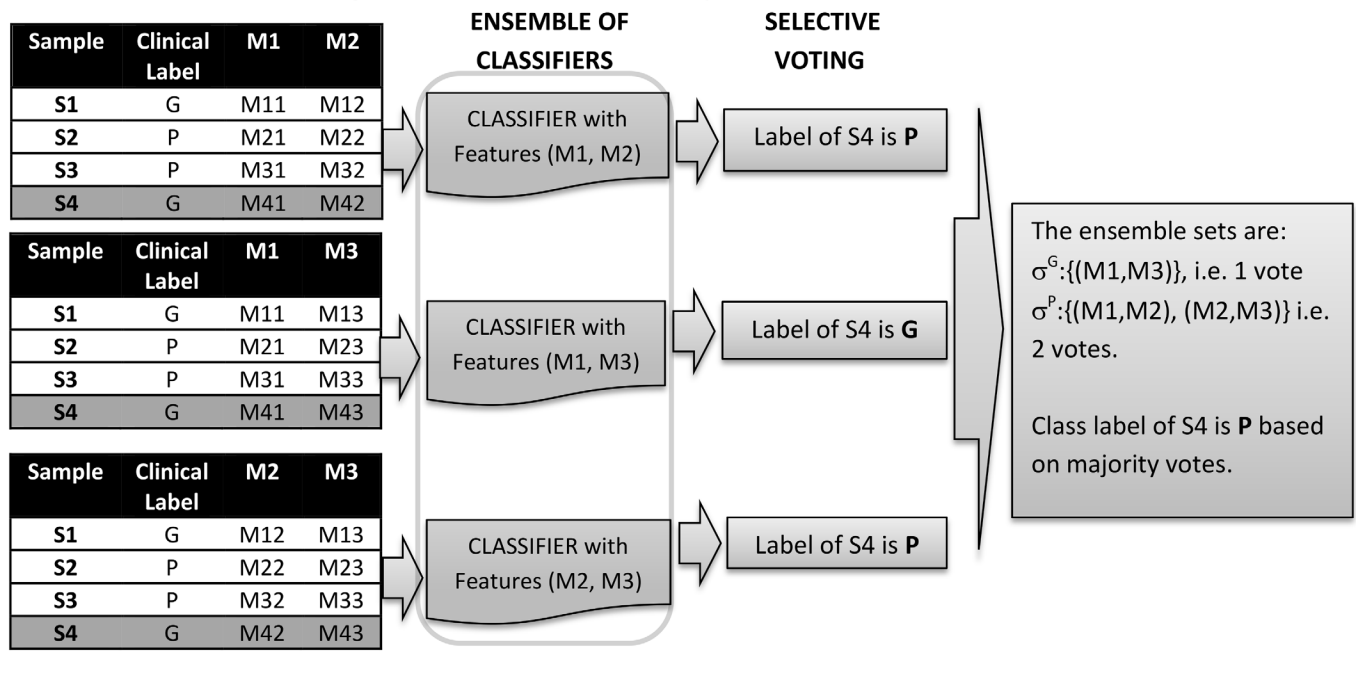
Class label of S4 is **P** based on majority votes.

Fig 1. Working Principle of the Selective Voting Approach. An example of traditional binary classification using all features and a single classifier is shown in **(a)**. SVA approach that selectively votes on samples using an ensemble of classifiers and pairs of features is shown in **(b)**. Three molecular markers (M1, M2, M3) and four samples (S1, S2, S3, S4) with clinical labels (G, P, P, G) corresponding to gingivitis (G) and periodontitis (P) are considered. Sample S4 is set aside as the test sample with (S1, S2, S3) as the training samples in the classification process. The single classifier approach implicitly assumes the markers to be identical and fixed across the samples and votes the test sample as either G or P. In contrast, SVA enables the pairs of features to selectively vote on the samples. For the above example we have $^3C_2 = 3$ potential classifiers {(M1, M2), (M1, M3), (M2, M3)}. While {(M1, M2), (M2, M3)} vote on the test sample as gingivitis represented by σ^G {(M1, M3)} vote on the test sample as periodontitis represented by σ^P. Based on majority votes (2 votes as periodontitis and 1 vote as gingivitis) the SVA class label of the test sample is deemed as periodontitis. More importantly, the test sample with clinical label gingivitis is classified as periodontitis by SVA. It is important to note that if we were to repeat the above process by setting aside the other gingivitis sample (S1) as test sample, the ensemble sets and votes need not necessarily be the same. This critical selective voting aspect of SVA enables personalization of the ensemble sets across samples within and between the classes (i.e. disease groups). Also, in contrast to the single classifier that deems the given test sample as G or P, SVA determines the proclivity of the test sample to G and P given by the normalized vote counts (1/3, 2/3).

If $r < N_b$ then set $r \leftarrow r+1$, Repeat Steps 1–6 by randomly permuting the columns of X_{mxn}.

Step 7: Proclivity and consensus map across the samples is determined as follows

Proclivity: The normalized vote count pair $(G_{votes}(k), P_{votes}(k)) \in [0,1]$ represents the proclivity of the kth sample to gingivitis and periodontitis where

$$G_{votes}(k) = \frac{1}{N_b} \sum_{r=1}^{N_b} \frac{v^g(r,k)}{m(m-1)/2}$$

$$P_{votes}(k) = \frac{1}{N_b} \sum_{r=1}^{N_b} \frac{v^p(r,k)}{m(m-1)/2}$$

Consensus Map: Represents the number of elements that are common between ensemble sets across a pair of samples (i,j). We define $\sigma^{g*}(k) = \cap_{r=1}^{N_b} \sigma^g(r,k)$ and $\sigma^{p*}(k) = \cap_{r=1}^{N_b} \sigma^p(r,k)$.

Consensus within the gingivitis group across a pair of samples (i,j) is given by:

$$\tau(i,j) = |\sigma^{g*}(i) \cap \sigma^{g*}(j)| \ (i,j), i,j = 1 \ldots n_{gin}, i \neq j.$$

Consensus within the periodontitis group across a pair of samples (i,j) is given by:

$$\tau(i,j) = |\sigma^{p*}(i) \cap \sigma^{p*}(j)| \ (i,j), i,j = 1 \ldots n_{per}, i \neq j.$$

Consensus between the gingivitis and periodontitis groups across samples (i,j) is given by:

$$\tau(i,j) = |\sigma^{g*}(i) \cap \sigma^{p*}(j)| \ (i,j), i = 1 \ldots n_{gin}; j = 1 \ldots n_{per}$$

In the present study, four different classification algorithms (**LDA**: Linear Discriminant Analysis; **QDA**: Quadratic Discriminant Analysis; **NBC**: Naïve Bayes Classifier and **SVM**: Support Vector Machines) were used independently as single classifiers as in Step 1 as a part of the SVA approach (i.e. SVA-LDA, SVA-QDA, SVA-NB, SVA-SVM). Performance of the classification algorithm with and without selective voting is determined assuming the clinical labels of the samples as the ground truth. In Step 1 the given n samples were divided into non-overlapping sets with 10 samples each. Each time, a different set is held out as the test data with the remaining samples used as the training data in the classification process. The cross-validation process was independently repeated ($N_b = 1000$) times by randomly permuting the columns so as to assign different samples to the test and training sets in each independent realization. In contrast to a traditional classifier that uses all markers as features, the proposed approach uses pairs of markers in Step 2 resulting in an ensemble of classifiers that vote on the classification label of a given sample, Step 6. Pairs of markers that vote consistently on a given sample across the ($N_b = 1000$) realizations were stored in the ensemble set σ^* of that sample in Step 7. Unlike traditional binary classification that deems a sample to either of the disease groups, SVA determines the proclivity of a sample to each of the disease groups given by the normalized vote counts $0 \leq (G_{votes}, P_{votes}) \leq 1$ in Step 7. Consensus between the ensemble sets of the samples

within and between the disease groups reflected by total number of common elements is stored in τ Step 7 and displayed as a heatmap in the present study.

Classifier Performance and Mismatch Samples

Measures such as (**ACC**: accuracy, **SEN**: sensitivity, **SPC**: specificity) have been used routinely to assess classifier performance. Ensemble vote counts from SVA can result in tie votes when a sample has equal proclivity to each of the disease groups, i.e. $G_{votes} = P_{votes}$. Therefore, prior to estimating the performance measures, samples with ties from SVA were randomly assigned labels from one of the classes. While we use the clinical labels of the samples as the ground truth it is important to appreciate that the clinical parameters (BOP, PPD, CAL) may not necessarily capture all intricate aspects of gingivitis and periodontitis. Therefore, samples whose clinical labels and classification labels do not match are termed as mismatch samples and not misclassified samples.

Results

Skewed Distributions of Expression Profiles

Clinical criteria (BOP, PPD) with published thresholds are traditionally used to classify a patient either as gingivitis or periodontitis. However, (%BOP, %PPD) across the clinically labelled gingivitis (n = 40) and periodontitis (n = 40) samples revealed marked variations in the magnitude of these parameters within each of the groups. Expression profiles of the salivary biomarkers (IL-1ß, IL-6, MMP-8, MIP-1α) were positively skewed accompanied by large standard deviations and considerable overlap between the disease groups ([Fig 2]). These findings may indicate the absence of an obvious fixed threshold separating these two disease groups and also the presence of considerable variability within each of the disease groups.

Performance Measures

Performance measures (ACC, SEN, SPC) estimated with the SVA approach and traditional single classifier approach using four different classification algorithms (LDA, QDA, NBC, SVM) across (N_b = 1000) independent realizations and leave-10-out cross-validation assuming clinical labels as ground truth were comparable ([Table 1]). Mismatch samples (i.e. classification labels and clinical labels did not match) across a majority of the (N_b = 1000) independent runs and across the four classification algorithms in the SVA approach consisted of 8 gingivitis (4, 15, 17, 18, 23, 25, 28, 38) and 9 periodontitis (41, 47, 51, 62, 64, 67, 70, 72, 78) samples. Three of the mismatch samples (41, 47, 51) were smokers. Thus smoking did not appear to be the primary driving factor behind the mismatch periodontitis samples.

Consensus Maps

Consensus map (τ) representing the magnitude of overlap in the ensemble sets between the samples in the two disease groups is shown in [Fig 3]. Since the consensus map is symmetric by definition only the upper-triangular part is shown. The consensus map consists of three distinct regions corresponding to consensus within the gingivitis and periodontitis groups represented by the triangular regions **P x P** and **G x G** in [Fig 3]. In addition, there is also a square region (**G x P**) that represents potential consensus in the ensemble sets between the gingivitis and periodontitis samples. Notably, pronounced dark streaks were observed across the mismatch gingivitis (4, 15, 17, 18, 23, 25, 28, 38) and mismatch periodontitis (41, 47, 51, 62, 64, 67, 70, 72, 78) and the rest of the samples within the respective disease groups in the P x P and G x G regions. While the mismatch samples exhibited lack of consensus within their respective
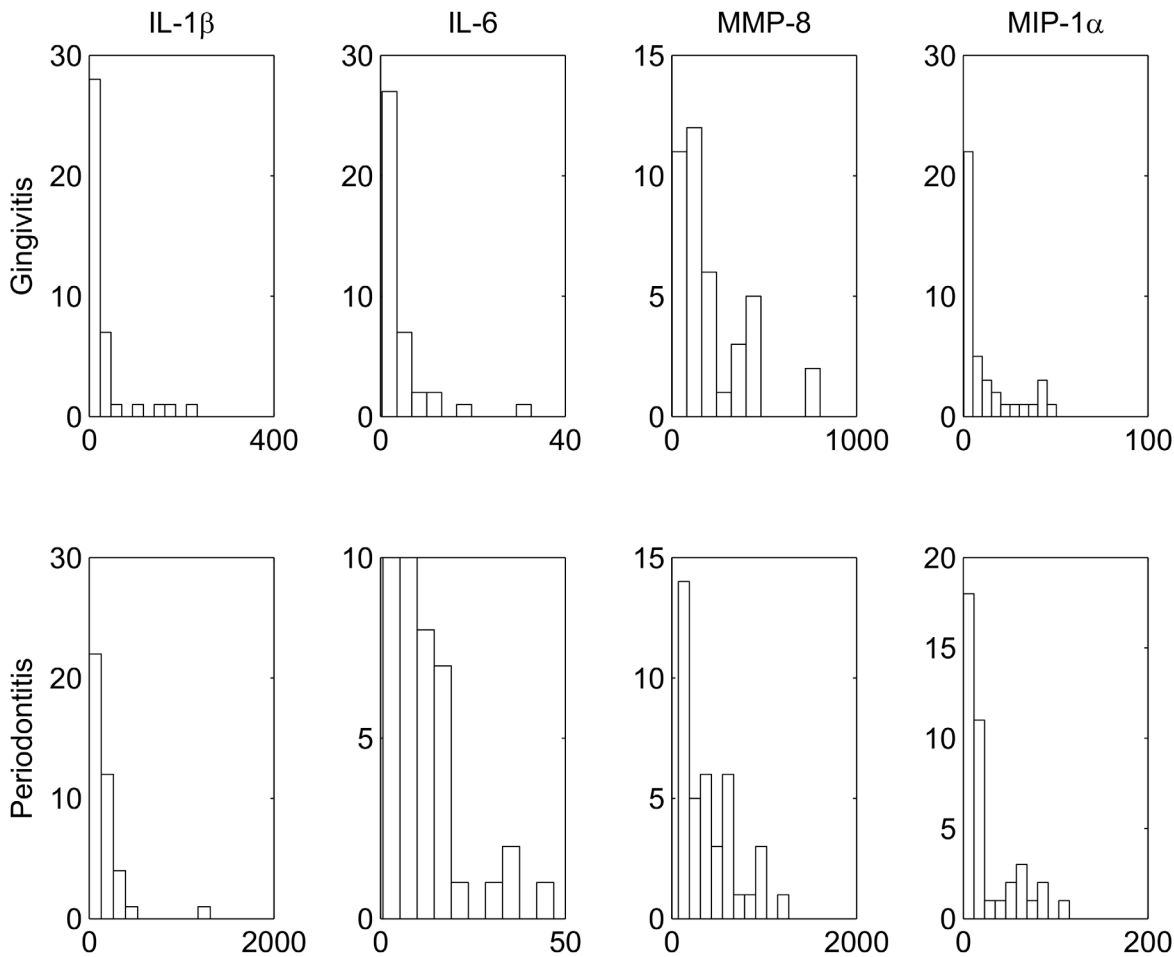
Fig 2. Distribution of the salivary biomarkers. Histograms representing the distribution of molecular expression profiles of (IL-1ß, IL-6, MMP-8, MIP-1α) across the gingivitis (n = 40, top row) and periodontitis (n = 40, bottom row) subjects. The mean and standard deviation of (IL-1ß, IL-6, MMP-8, MIP-1α) across gingivitis were (29.6±49.5, 3.9±5.9, 208.2±194.2, 10.9±14.5) whereas those across periodontitis were (157.6±217, 12.1±10.2, 397.9±302.1, 24.4 ±29.8).

doi:10.1371/journal.pone.0136792.g002

groups, a subset of these samples also exhibited marked consensus between the disease groups accompanied by marked transitions from dark streaks in G x G and P x P to bright streaks in G x P. This subset included gingivitis samples (4, 15, 17, 23, 28) indicating some of the gingivitis samples may share similarity to periodontitis. More importantly, these patterns were

Table 1. Classification Performance Metrics with Clinical Labels as Ground Truth.

| SVA Approach	Accuracy	Sensitivity	Specificity
SVA-LDA	0.76±0.02	0.75±0.02	0.77±0.03
SVA-QDA	0.76±0.02	0.77±0.02	0.75±0.03
SVA-NB	0.76±0.01	0.77±0.02	0.75±0.02
SVA-SVM	0.76±0.02	0.73±0.02	0.79±0.03
Traditional Single Classifier	**Accuracy**	**Sensitivity**	**Specificity**
LDA	0.78±0.01	0.78±0.02	0.78±0.02
QDA	0.76±0.02	0.80±0.03	0.72±0.02
NB	0.76±0.01	0.77±0.01	0.75±0.01
SVM	0.75±0.02	0.74±0.03	0.77±0.02

doi:10.1371/journal.pone.0136792.t001
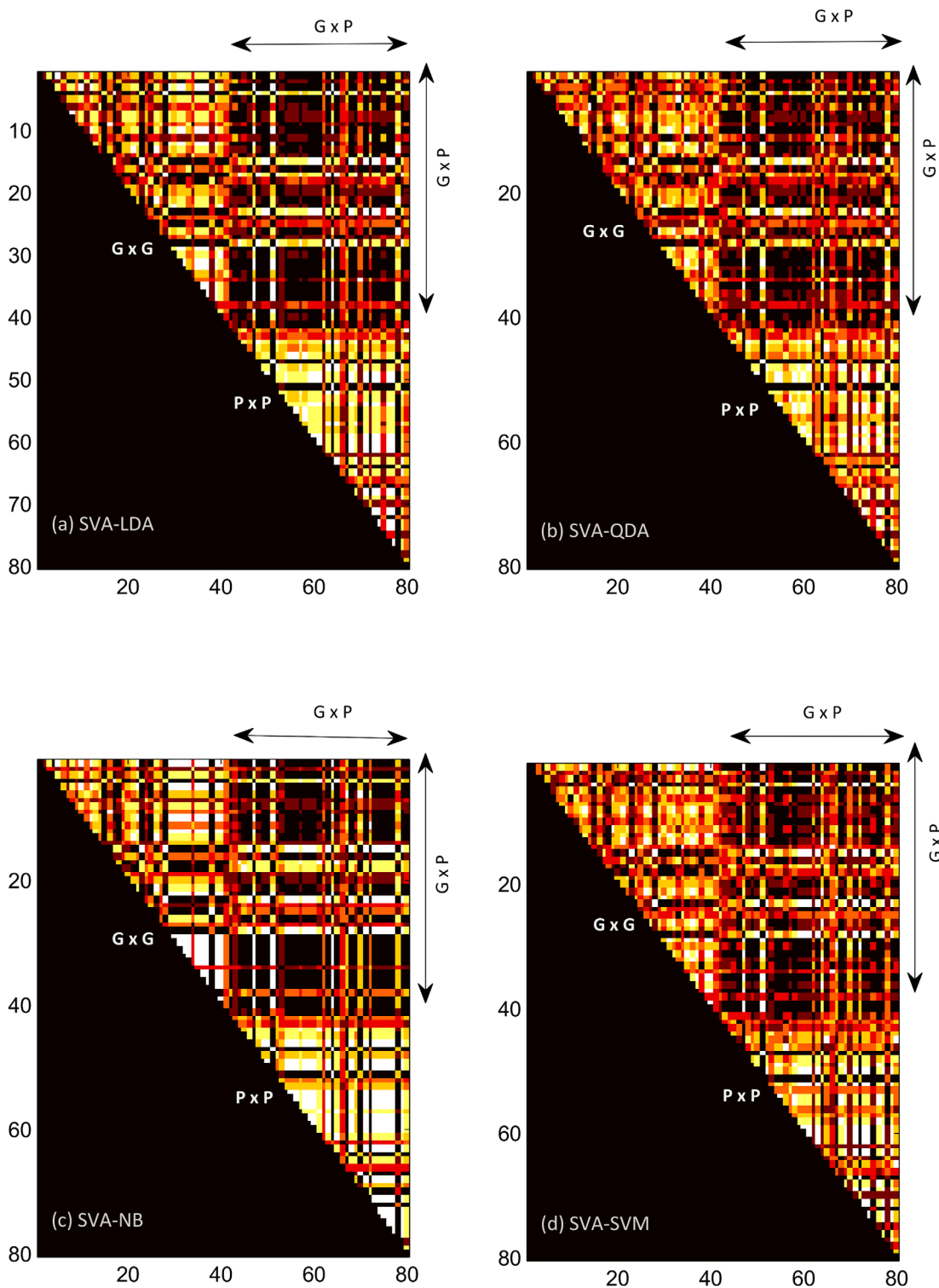
Fig 3. Heatmap Visualization. Heatmap visualization of the consensus map (τ) representing the consensus between the ensemble sets of the gingivitis (1…40) and periodontitis (41…80) samples. Significant overlap is represented by bright color and absence of overlap by dark color. Heatmap for the four different classification techniques (SVA-LDA, SVA-QDA, SVA-NB, SVA-SVM) are enclosed in the subplots (a-d) respectively. In each subplot, there are three distinct regions (**G x G, P x P, G x P**) corresponding to overlap within the gingivitis samples (triangle), within the periodontitis samples (triangle) and between the gingivitis and periodontitis samples (square). The misclassified gingivitis (4, 15, 17, 18, 23, 25, 28, 38) and periodontitis (41, 47, 51, 62, 64, 67, 70, 72, 78) samples are accompanied by pronounced dark streaks in each of the subplots.

doi:10.1371/journal.pone.0136792.g003

consistently observed across the four different classification algorithms (SVA-LDA, SVA-QDA, SVA-NB and SVA-SVM), Fig 3.

Proclivity

Proclivity of the samples to the gingivitis and periodontitis groups from their normalized vote counts (G_{votes}, P_{votes}) obtained using SVA and the four classification algorithms (SVA-LDA, SVA-QDA, SVA-NB and SVA-SVM) exhibited considerable variation across the samples as shown by the scatter plots (Fig 4). Since ($0 \leq G_{votes}$, $P_{votes} \leq 1$) each subplot represents a unit square. The line of separation between gingivitis and periodontitis is represented by the diagonal with samples on the diagonal having equal proclivity to each of these groups i.e. ($G_{votes} = P_{votes}$) Shift from the diagonal towards the abscissa and ordinate represent increasing proclivity to gingivitis and periodontitis respectively. While a majority of the mismatch samples exhibited similar proclivities irrespective of the choice of the classification algorithm (SVA-LDA, SVA-QDA, SVA-NB, and SVA-SVM), Fig 4, a few of them (gingivitis: 25, 38) and (periodontitis: 62, 67, 72) were sensitive to the choice of the classification algorithms.

Discussion

There remain numerous questions regarding the inherent risk of individuals to develop periodontitis, with extensive population variation in genetics and environmental contribution to disease that impact, time of onset, rapidity of progression, and extent of the disease throughout the mouth, combining to dictate the overall severity of the disease [9]. Demographic factors, genetic factors, medical history and risk-factors such as smoking have also been attributed to patient specific variations in gingivitis and periodontitis. These diverse factors in turn warrant developing novel approaches that provide insights into patient-specific variations within disease groups.

A substantial number of investigations over many decades have attempted to catalogue and utilize a range of biologic parameters involved in periodontal disease, focusing on the microbial milieu and the resulting armamentarium of host responses to control this colonization/infection [49]. However, two critical issues that routinely challenge the interpretation of these results is the lack of a clear understanding of the within subject variability and an ability to clearly differentiate the parameters of destructive inflammatory responses in the periodontium. The general strategy has been to utilize a "one size fits all" approach and create a fixed threshold to minimize overlap between clinical categories. In contrast to "biologically homogenizing" the samples within the clinical categories, the proposed approach (SVA) reveals within group heterogeneity of biologic responses as they relate to current clinical presentation from salivary molecular profiles. Thus these findings could be important with respect to personalized medical and dental care in that clinical decisions within this field will rely on recognizing the heterogeneity of patients and understanding patient-specific variations for improving health [58].

The underlying biology of gingivitis and periodontitis involves inflammation, connective tissue destruction and bone remodeling. Biomarkers reflective of these three processes [9, 59] in conjunction with sophisticated classification techniques will enable determining their diagnostic effectiveness and validity. Recent data from our lab and others indicate that salivary concentrations of IL-6, IL-8, albumin, calprotectin, PGE_2, MMP-8, and MIP-1α are elevated in patients who have gingivitis.[10, 60] Accordingly, these salivary analytes appear to serve as biomarkers of gingivitis, coupled with existing literature on biomarkers of periodontitis [52, 59–63], may help to discriminate those patients who are transitioning from health to gingivitis and on further to periodontitis. Moreover, these salivary markers showed that at least 40% of patients who have gingivitis do not return to biological health within 30 days after standard
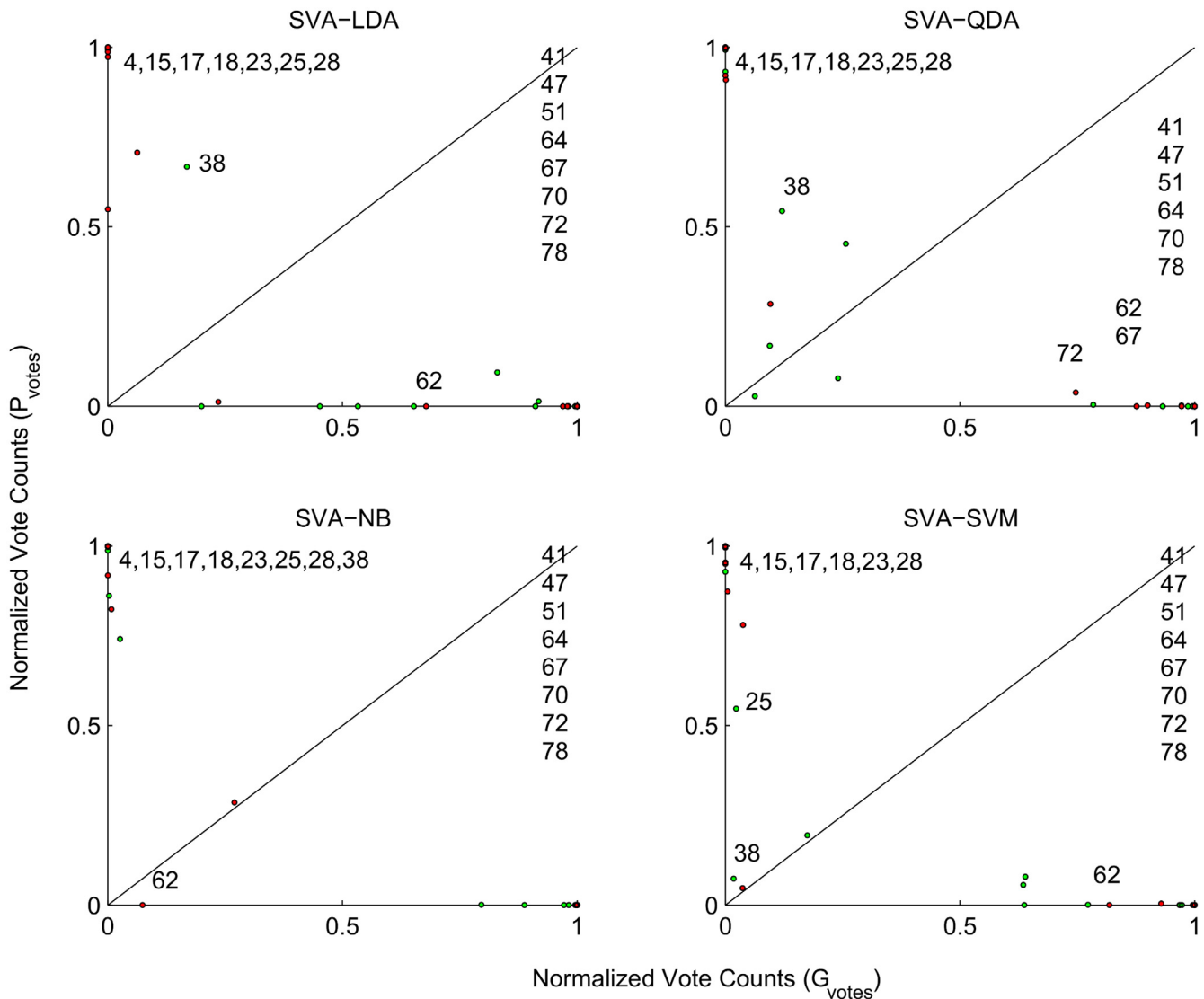
Fig 4. Proclivity (G_{votes}, P_{votes}) of the 80 samples to gingivitis and periodontitis groups obtained using the proposed ensemble selective voting approach is shown. The diagonal ($G_{votes} = P_{votes}$) represents the line of separation between the two groups with shift towards abscissa and ordinate representing increasing proclivity to gingivitis and periodontitis respectively. Gingivitis (4, 15, 17, 18, 23, 25, 28, 38, green) and periodontitis (41, 47, 51, 62, 64, 67, 70, 72, 78, red) samples whose clinical and class labels did not march across (SVA-LDA, SVA-QDA, SVA-NB, SVA-SVM) are identified in each of the subplots.

doi:10.1371/journal.pone.0136792.g004

treatment for gingivitis, even though clinical measures of their gingiva and periodontium indicate a return to clinical health; thus emphasizing the discrepancy between clinical and biological health [10]. Also, our current findings and numerous prior studies have clearly identified important limitations in existing approaches for periodontal disease assessment.

In this vein, we used these putative salivary biomarkers of gingivitis and periodontitis and SVA to evaluate variations in the biological presentations of a cohort of 80 gingivitis and periodontitis patients. Subsequently, we identified a subpopulation of approximately 20% in the cohort that was different biologically from the majority. This was confirmed using four different classification techniques. However, in contrast to traditional classification, SVA approach enabled the markers to vary across the subjects, within and between groups revealing patient-

specific variations. As such, this report may begin addressing a fundamental question in periodontology, that is, within the clinical heterogeneity of gingivitis, what patients are at greatest risk for progressing to destructive disease? However, the lack of temporal evaluation of the response to therapy or incidence of transition from gingivitis to periodontitis certainly limits our ability to address this question within the framework of the current study. An additional limitation of the study was that a targeted group of biomarkers were included primarily related to historical evidence of their ability to differentiate periodontitis from health [17, 62, 64]. A subsequent prospective longitudinal experimental design in conjunction with a broader panel of biomarkers [65, 66] would be necessary to demonstrate the prognostic ability of salivary biomarkers for identifying gingivitis patients at greatest risk for progressing to periodontitis. Finally, a more detailed study that explicitly accommodates the various factors that can contribute to patient-specific variations may be necessary to assess the individual contributions of these factors. The findings reported here could help jumpstart the development of personalized decisions in periodontal disease diagnosis and therapy.

Acknowledgments

Author Contributions

References

1. Pihlstrom BL, Michalowicz BS, Johnson NW. Periodontal diseases. Lancet. 2005; 366(9499):1809–20. PMID: 16298220.

2. Lang NP, Schatzle MA, Loe H. Gingivitis as a risk factor in periodontal disease. J Clin Periodontol. 2009; 36 Suppl 10:3–8. doi: 10.1111/j.1600-051X.2009.01415.x PMID: 19432625.

3. Schatzle M, Loe H, Lang NP, Burgin W, Anerud A, Boysen H. The clinical course of chronic periodontitis. J Clin Periodontol. 2004; 31(12):1122–7. doi: 10.1111/j.1600-051X.2004.00634.x PMID: 15560816.

4. Schatzle M, Faddy MJ, Cullinan MP, Seymour GJ, Lang NP, Burgin W, et al. The clinical course of chronic periodontitis: V. Predictive factors in periodontal disease. J Clin Periodontol. 2009; 36(5):365–71. doi: 10.1111/j.1600-051X.2009.01391.x PMID: 19419434.

5. Schatzle M, Loe H, Burgin W, Anerud A, Boysen H, Lang NP. Clinical course of chronic periodontitis. I. Role of gingivitis. J Clin Periodontol. 2003; 30(10):887–901. PMID: 14710769.

6. Albandar JM, Kingman A. Gingival recession, gingival bleeding, and dental calculus in adults 30 years of age and older in the United States, 1988–1994. Journal of periodontology. 1999; 70(1):30–43. Epub 1999/03/03. PMID: 10052768.

7. Tomar SL, Asma S. Smoking-attributable periodontitis in the United States: findings from NHANES III. National Health and Nutrition Examination Survey. J Periodontol. 2000; 71(5):743–51. PMID: 10872955.

8. Offenbacher S. Periodontal diseases: pathogenesis. Annals of periodontology / the American Academy of Periodontology. 1996; 1(1):821–78. PMID: 9118282.

9. Armitage GC. Learned and unlearned concepts in periodontal diagnostics: a 50-year perspective. Periodontol 2000. 2013; 62(1):20–36. doi: 10.1111/prd.12006 PMID: 23574462.

10. Syndergaard B, Al-Sabbagh M, Kryscio RJ, Xi J, Ding X, Ebersole JL, et al. Salivary Biomarkers Associated with Gingivitis and Response to Therapy. J Periodontol. 2014. doi: 10.1902/jop.2014.130696 PMID: 24502627.

11. Van Dyke TE, Sheilesh D. Risk factors for periodontitis. Journal of the International Academy of Periodontology. 2005; 7(1):3–7. PMID: 15736889.

12. Kornman KS, Page RC, Tonetti MS. The host response to the microbial challenge in periodontitis: assembling the players. Periodontol 2000. 1997; 14:33–53. PMID: 9567965.

13. Ebersole JL, Dawson DR 3rd, Morford LA, Peyyala R, Miller CS, Gonzalez OA. Periodontal disease immunology: 'double indemnity' in protecting the host. Periodontol 2000. 2013; 62(1):163–202. doi: 10.1111/prd.12005 PMID: 23574466.

14. Yucel-Lindberg T, Bage T. Inflammatory mediators in the pathogenesis of periodontitis. Expert reviews in molecular medicine. 2013; 15:e7. doi: 10.1017/erm.2013.8 PMID: 23915822.

15. Miller CS, King CP Jr., Langub MC, Kryscio RJ, Thomas MV. Salivary biomarkers of existing periodontal disease: a cross-sectional study. J Am Dent Assoc. 2006; 137(3):322–9. PMID: 16570465.

16. Sorsa T, Mantyla P, Ronka H, Kallio P, Kallis GB, Lundqvist C, et al. Scientific basis of a matrix metalloproteinase-8 specific chair-side test for monitoring periodontal and peri-implant health and disease. Ann N Y Acad Sci. 1999; 878:130–40. PMID: 10415725.

17. Miller CS, Foley JD, Bailey AL, Campell CL, Humphries RL, Christodoulides N, et al. Current developments in salivary diagnostics. Biomark Med. 2010; 4(1):171–89. Epub 2010/04/14. PMID: 20387312.

18. Giannobile WV, Beikler T, Kinney JS, Ramseier CA, Morelli T, Wong DT. Saliva as a diagnostic tool for periodontal disease: current state and future directions. Periodontology 2000. 2009; 50:52–64. Epub 2009/04/25. doi: 10.1111/j.1600-0757.2008.00288.x PMID: 19388953.

19. Prakasam S, Srinivasan M. Evaluation of salivary biomarker profiles following non-surgical management of chronic periodontitis. Oral diseases. 2014; 20(2):171–7. Epub 2013/03/19. doi: 10.1111/odi.12085 PMID: 23496245.

20. Teles R, Sakellari D, Teles F, Konstantinidis A, Kent R, Socransky S, et al. Relationships among gingival crevicular fluid biomarkers, clinical parameters of periodontal disease, and the subgingival microbiota. J Periodontol. 2010; 81(1):89–98. doi: 10.1902/jop.2009.090397 PMID: 20059421; PubMed Central PMCID: PMC2805280.

21. Salminen A, Gursoy UK, Paju S, Hyvarinen K, Mantyla P, Buhlin K, et al. Salivary biomarkers of bacterial burden, inflammatory response, and tissue destruction in periodontitis. J Clin Periodontol. 2014; 41(5):442–50. doi: 10.1111/jcpe.12234 PMID: 24460823.

22. Yoon AJ, Cheng B, Philipone E, Turner R, Lamster IB. Inflammatory biomarkers in saliva: assessing the strength of association of diabetes mellitus and periodontal status with the oral inflammatory burden. J Clin Periodontol. 2012; 39(5):434–40. doi: 10.1111/j.1600-051X.2012.01866.x PMID: 22420648.

23. Tobon-Arroyave SI, Jaramillo-Gonzalez PE, Isaza-Guzman DM. Correlation between salivary IL-1beta levels and periodontal clinical status. Arch Oral Biol. 2008; 53(4):346–52. Epub 2007/12/25. S0003-9969(07)00286-5 [pii]PMID: 18155182.

24. Sexton WM, Lin Y, Kryscio RJ, Dawson DR 3rd, Ebersole JL, Miller CS. Salivary biomarkers of periodontal disease in response to treatment. J Clin Periodontol. 2011; 38(5):434–41. doi: 10.1111/j.1600-051X.2011.01706.x PMID: 21480939; PubMed Central PMCID: PMC3095429.

25. Teles FR, Teles RP, Martin L, Socransky SS, Haffajee AD. Relationships among interleukin-6, tumor necrosis factor-alpha, adipokines, vitamin D, and chronic periodontitis. J Periodontol. 2012; 83(9):1183–91. doi: 10.1902/jop.2011.110346 PMID: 22181684.

26. Javed F, Ahmed HB, Saeed A, Mehmood A, Bain C. Whole salivary interleukin-6 and matrix metalloproteinase-8 levels in patients with chronic periodontitis with and without prediabetes. J Periodontol. 2014; 85(5):e130–5. doi: 10.1902/jop.2013.130514 PMID: 24171502.

27. Duarte PM, Miranda TS, Lima JA, Dias Goncalves TE, Santos VR, Bastos MF, et al. Expression of immune-inflammatory markers in sites of chronic periodontitis in patients with type 2 diabetes. J Periodontol. 2012; 83(4):426–34. doi: 10.1902/jop.2011.110324 PMID: 21859322.

28. Kardesler L, Buduneli N, Cetinkalp S, Lappin D, Kinane DF. Gingival crevicular fluid IL-6, tPA, PAI-2, albumin levels following initial periodontal treatment in chronic periodontitis patients with or without type 2 diabetes. Inflamm Res. 2011; 60(2):143–51. doi: 10.1007/s00011-010-0248-7 PMID: 20845058.

29. Giuliani N, Colla S, Rizzoli V. New insight in the mechanism of osteoclast activation and formation in multiple myeloma: focus on the receptor activator of NF-kappaB ligand (RANKL). Experimental hematology. 2004; 32(8):685–91. PMID: 15308315.

30. Terpos E, Politou M, Viniou N, Rahemtulla A. Significance of macrophage inflammatory protein-1 alpha (MIP-1alpha) in multiple myeloma. Leukemia & lymphoma. 2005; 46(12):1699–707. PMID: 16263571.

31. Choi Y, Woo KM, Ko SH, Lee YJ, Park SJ, Kim HM, et al. Osteoclastogenesis is enhanced by activated B cells but suppressed by activated CD8(+) T cells. Eur J Immunol. 2001; 31(7):2179–88. PMID: 11449372.

32. Fine DH, Markowitz K, Furgang D, Fairlie K, Ferrandiz J, Nasri C, et al. Macrophage inflammatory protein-1alpha: a salivary biomarker of bone loss in a longitudinal cohort study of children at risk for aggressive periodontal disease? J Periodontol. 2009; 80(1):106–13. Epub 2009/02/21. doi: 10.1902/jop.2009.080296 [pii]. PMID: 19228096.

33. Al-Sabbagh M, Alladah A, Lin Y, Kryscio RJ, Thomas MV, Ebersole JL, et al. Bone remodeling-associated salivary biomarker MIP-1alpha distinguishes periodontal disease from health. J Periodontal Res. 2012; 47(3):389–95. doi: 10.1111/j.1600-0765.2011.01445.x PMID: 22126530; PubMed Central PMCID: PMC3292624.

34. Jonsson D, Ramberg P, Demmer RT, Kebschull M, Dahlen G, Papapanou PN. Gingival tissue transcriptomes in experimental gingivitis. J Clin Periodontol. 2011; 38(7):599–611. doi: 10.1111/j.1600-051X.2011.01719.x PMID: 21501207; PubMed Central PMCID: PMC3413194.

35. Stoecklin-Wasmer C, Guarnieri P, Celenti R, Demmer RT, Kebschull M, Papapanou PN. MicroRNAs and their target genes in gingival tissues. J Dent Res. 2012; 91(10):934–40. doi: 10.1177/0022034512456551 PMID: 22879578; PubMed Central PMCID: PMC3446831.

36. Kebschull M, Papapanou PN. The use of gene arrays in deciphering the pathobiology of periodontal diseases. Methods Mol Biol. 2010; 666:385–93. doi: 10.1007/978-1-60761-820-1_24 PMID: 20717797.

37. Papapanou PN, Behle JH, Kebschull M, Celenti R, Wolf DL, Handfield M, et al. Subgingival bacterial colonization profiles correlate with gingival tissue gene expression. BMC microbiology. 2009; 9:221. doi: 10.1186/1471-2180-9-221 PMID: 19835625; PubMed Central PMCID: PMC2771036.

38. Offenbacher S, Barros SP, Paquette DW, Winston JL, Biesbrock AR, Thomason RG, et al. Gingival transcriptome patterns during induction and resolution of experimental gingivitis in humans. Journal of periodontology. 2009; 80(12):1963–82. doi: 10.1902/jop.2009.080645 PMID: 19961380.

39. Davanian H, Stranneheim H, Bage T, Lagervall M, Jansson L, Lundeberg J, et al. Gene expression profiles in paired gingival biopsies from periodontitis-affected and healthy tissues revealed by massively parallel sequencing. PLoS One. 2012; 7(9):e46440. doi: 10.1371/journal.pone.0046440 PMID: 23029519; PubMed Central PMCID: PMC3460903.

40. Kebschull M, Guarnieri P, Demmer RT, Boulesteix AL, Pavlidis P, Papapanou PN. Molecular differences between chronic and aggressive periodontitis. J Dent Res. 2013; 92(12):1081–8. doi: 10.1177/0022034513506011 PMID: 24122488; PubMed Central PMCID: PMC3834653.

41. Demmer RT, Behle JH, Wolf DL, Handfield M, Kebschull M, Celenti R, et al. Transcriptomes in healthy and diseased gingival tissues. J Periodontol. 2008; 79(11):2112–24. doi: 10.1902/jop.2008.080139 PMID: 18980520; PubMed Central PMCID: PMC2637651.

42. Zhou Z-H, Wu J, Tang W. Ensembling neural networks: Many could be better than all. Artificial Intelligence. 2002; 137(1–2):239–63.

43. Rokach L. Collective-agreement-based pruning of ensembles. Computational Statistics and Data Analysis. 2009; 53:1015–26.

44. Rokach L. Ensemble-based classifiers. Artificial Intelligence Review. 2010; 33:1–39.

45. Dietterich TG. Ensemble methods in machine learning. Multiple classifier systems: Springer; 2000. p. 1–15.

46. Dietterich TG. Ensemble learning. The handbook of brain theory and neural networks. 2002; 2:110–25.

47. Kodell RL, Zhang C, Siegel ER, Nagarajan R. Selective voting in convex-hull ensembles improves classification accuracy. Artificial intelligence in medicine. 2012; 54(3):171–9. doi: 10.1016/j.artmed.2011.10.003 PMID: 22064044; PubMed Central PMCID: PMC3666100.

48. Nagarajan R, Kodell RL. A selective voting convex-hull ensemble procedure for personalized medicine. AMIA Summits on Translational Science proceedings AMIA Summit on Translational Science. 2012; 2012:87–94. PMID: 22779058; PubMed Central PMCID: PMC3392048.

49. Benakanakere M, Kinane DF. Innate cellular responses to the periodontal biofilm. Frontiers of oral biology. 2012; 15:41–55. doi: 10.1159/000329670 PMID: 22142956.

50. Kinane DF, Bartold PM. Clinical relevance of the host responses of periodontitis. Periodontol 2000. 2007; 43:278–93. PMID: 17214845.

51. Novak MJ, Dawson DR 3rd, Magnusson I, Karpinia K, Polson A, Polson A, et al. Combining host modulation and topical antimicrobial therapy in the management of moderate to severe periodontitis: a randomized multicenter trial. J Periodontol. 2008; 79(1):33–41. doi: 10.1902/jop.2008.070237 PMID: 18166090.

52. Goncalves Lda R, Soares MR, Nogueira FC, Garcia C, Camisasca DR, Domont G, et al. Comparative proteomic analysis of whole saliva from chronic periodontitis patients. J Proteomics. 73(7):1334–41. Epub 2010/03/11. S1874-3919(10)00070-9 [pii]doi: 10.1016/j.jprot.2010.02.018 PMID: 20215060.

53. Ebersole JL, Schuster JL, Stevens J, Dawson D 3rd, Kryscio RJ, Lin Y, et al. Patterns of salivary analytes provide diagnostic capacity for distinguishing chronic adult periodontitis from health. Journal of

clinical immunology. 2013; 33(1):271–9. Epub 2012/08/29. doi: 10.1007/s10875-012-9771-3 PMID: 22926406; PubMed Central PMCID: PMCPmc3788639.

54. Genco RJ, Borgnakke WS. Risk factors for periodontal disease. Periodontology 2000. 2013; 62(1):59–94. Epub 2013/04/12. doi: 10.1111/j.1600-0757.2012.00457.x PMID: 23574464.

55. Hujoel PP. Destructive periodontal disease and tobacco and cannabis smoking. Jama. 2008; 299 (5):574–5. doi: 10.1001/jama.299.5.574 PMID: 18252889.

56. Thomas MV, Branscum A, Miller CS, Ebersole J, Al-Sabbagh M, Schuster JL. Within-subject variability in repeated measures of salivary analytes in healthy adults. J Periodontol. 2009; 80(7):1146–53. Epub 2009/07/01. doi: 10.1902/jop.2009.080654 PMID: 19563296.

57. Navazesh M. Methods for collecting saliva. Ann N Y Acad Sci. 1993; 694:72–7. Epub 1993/09/20. PMID: 8215087.

58. Mirnezami R, Kinross JM, Vorkas PA, Goldin R, Holmes E, Nicholson J, et al. Implementation of molecular phenotyping approaches in the personalized surgical patient journey. Annals of surgery. 2012; 255 (5):881–9. doi: 10.1097/SLA.0b013e31823e3c43 PMID: 22156927.

59. Buduneli N, Kinane DF. Host-derived diagnostic markers related to soft tissue destruction and bone degradation in periodontitis. J Clin Periodontol. 2011; 38 Suppl 11:85–105. doi: 10.1111/j.1600-051X. 2010.01670.x PMID: 21323706.

60. Lee A, Ghaname CB, Braun TM, Sugai JV, Teles RP, Loesche WJ, et al. Bacterial and salivary biomarkers predict the gingival inflammatory profile. J Periodontol. 2012; 83(1):79–89. doi: 10.1902/jop. 2011.110060 PMID: 21563952.

61. Salazar MG, Jehmlich N, Murr A, Dhople VM, Holtfreter B, Hammer E, et al. Identification of periodontitis associated changes in the proteome of whole human saliva by mass spectrometric analysis. J Clin Periodontol. 2013; 40(9):825–32. doi: 10.1111/jcpe.12130 PMID: 23790309.

62. Kinney JS, Morelli T, Braun T, Ramseier CA, Herr AE, Sugai JV, et al. Saliva/pathogen biomarker signatures and periodontal disease progression. J Dent Res. 2011; 90(6):752–8. doi: 10.1177/0022034511399908 PMID: 21406610; PubMed Central PMCID: PMC3144122.

63. Ramseier CA, Kinney JS, Herr AE, Braun T, Sugai JV, Shelburne CA, et al. Identification of pathogen and host-response markers correlated with periodontal disease. J Periodontol. 2009; 80(3):436–46. Epub 2009/03/04. doi: 10.1902/jop.2009.080480 PMID: 19254128.

64. Giannobile WV, McDevitt JT, Niedbala RS, Malamud D. Translational and clinical applications of salivary diagnostics. Advances in dental research. 2011; 23(4):375–80. doi: 10.1177/0022034511420434 PMID: 21917748; PubMed Central PMCID: PMC3172998.

65. Zhang L, Xiao H, Wong DT. Salivary biomarkers for clinical applications. Mol Diagn Ther. 2009; 13 (4):245–59. Epub 2009/08/29. 4 [pii]doi: 10.2165/11317160-000000000-00000 PMID: 19712004.

66. Hu S, Xie Y, Ramachandran P, Ogorzalek Loo RR, Li Y, Loo JA, et al. Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. Proteomics. 2005; 5(6):1714–28. Epub 2005/04/01. doi: 10. 1002/pmic.200401037 PMID: 15800970.