Support vector machine analysis is able to differentiate the microbial profiles of Aggressive and Chronic Periodontitis

Magda Feres[1], Yoram Louzoun[2,3], Simi Haber[2], Marcelo Faveri[1], Luciene C., Figueiredo[1], Liran Levin[4]

[1]Department of Periodontology Guarulhos University, SP, Brazil.

[2]Department of Mathematics Bar-Ilan University, Ramat-Gan, Israel.

[3]The Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, Israel.

[4]Professor and Head, Division of Periodontology, Faculty of Medicine and Dentistry, University of Alberta, Canada.

Short title: Automatic classification of periodontal disease

Keywords: plaque, oral health, prevention, periodontitis, mathematics

Correspondence and reprint requests to:

Prof. Liran Levin
University of Alberta, School of Dentistry | Faculty of Medicine & Dentistry
5-468 Edmonton Clinic Health Academy
11405 - 87 Avenue NW, 5th Floor
Edmonton AB T6G 1C9, CANADA
Phone: 780.492.9906 | Fax: 780.492.7536
E-mail: liran@ualberta.ca

Abstract

Aim: The aim of this study was to use a mathematical machine learning approach and the 40 bacterial species of the microbial complexes to compare the subgingival microbial profiles of Aggressive Periodontitis (AgP), Chronic Periodontitis (ChP) and periodontally healthy patients.

Method: Individual subgingival plaque samples were collected from nine non-contiguous interproximal sites in subject with AgP, ChP and healthy patients. A principal component analysis (PCA) was performed only for the visualization of the data. Two stages of Machine learning were performed. The data was split in each analysis to 50 % train and 50% test. A support vector machine (SVM) classifier was used with a Box constraint of 1 and a linear kernel. The analysis was divided into two parts. First we tested whether there was a general difference between the bacterial communities' composition in healthy and periodontally diseased patients. We then tested whether a difference existed between the two diseases studied.

Results: 435 subjects were included in the analysis, 53 periodontally healthy, 308 with chronic periodontitis and 74 with aggressive periodontitis. The results for the PCA (?) showed that the variance of the healthy samples in all PCR (?) directions was much smaller than the one of the periodontally diseased samples, suggesting that while the healthy cases are characterized by highly uniform bacterial levels, the periodontally diseased samples are much more diverse. The relative bacterial load could distinguish between AgP and ChP.

Conclusion: The results indicated that a SVM classifier using a panel of 40 bacterial species was able to distinguish between AgP and ChP. These results open new avenues for defining specific preventive and treatment protocols for these conditions.

Introduction

Periodontitis is an oral disease driven by deregulated inflammation induced by polymicrobial communities that form on subgingival tooth sites (Hajishengallis et al., 2016; Lamont & Hajishengallis 2015). The gingival sulcus and periodontal pocket form unique ecological niches for microbial colonization and the subgingival microbiota drives the inflammatory process that leads to periodontal tissue destruction.

The existence of different forms of periodontitis is a reality and over the years, different classification systems have been suggested for these conditions (Armitage 1999; Tonetti et al., 2005). In 1999, the American Academy of Periodontology (AAP) changed the term Adult Periodontitis to Chronic Periodontitis (ChP) and introduced the controversial term “Aggressive Periodontitis” (AgP) to define a group of destructive periodontal diseases with a rapid progression (Lang 1999). This definition aimed to encompass previous definitions of early-onset periodontitis, juvenile periodontitis and rapidly progressive periodontitis, using the terminology “aggressive”, rather unusual in the medical field. This put the emphasis on rate of disease progression, an information rarely available to the clinician (Levin et al., 2006; Nibali 2015). Although the classification systems have been continuously under debate, it is largely well accepted that disease in younger subjects is different from that in adults, and one possible explanation for these differences are different microbial profiles. If this is the case, a microbiological exam might, in theory, help in the differential diagnosis of AgP in young subjects with the more common ChP (Nibali 2015), and would have the potential to help in the

diagnosis and treatment of these infections (Albandar & Tinoco 2002; Nibali 2015).

Since the 1950s, the microbiota of the periodontal diseases has been studied, initially by culture methods, and afterwards by molecular techniques. The current knowledge about the microbiota associated with periodontal health or disease has been largely impacted by the evaluation of the 40 bacterial species that comprise the microbial complexes described by Socransky et al. in 1998 (Colombo et al., 2002; Faveri et al., 2009; Lopez et al., 2004; Socransky et al., 1998; Ximenez-Fyvie et al., 2006). Studies using different diagnostic techniques have defined four classical periodontal pathogens: the three species from the red complex, *Porphyromonas gingivalis*, *Tannerella forsythia and, Treponema denticola,* as well as *Aggregatibacter actinomycetemcomitans* (Costalonga & Herzberg 2014; Faveri et al., 2009; Feres et al., 2009; Socransky & Haffajee 2005; Socransky et al., 1998). In addition, several species belonging to those complexes have been associated with periodontal health, especially those from the genera *Actinomyces, Streptococcus and Capnocytophaga* (Abusleme et al., 2013; Kolenbrander et al., 2006; Socransky & Haffajee 2005). It is also important to note that there is moderate evidence in the literature to support the existence of newly identified periodontal pathogens or host-compatible species (Goncalves et al., 2016; Oliveira et al., 2016; Perez-Chaparro et al., 2014), but the role of these species as true pathogens or as markers for periodontal stability are yet to be established, specially by risk assessment and interventional studies. Therefore, the 40 bacterial species defined by Socransky et al. in 1998 are still considered a suitable biological marker for

studying the periodontal microbiota associated with periodontal health (PH) or disease.

Machine learning, such as support vector machines (SMV) is a discipline of computer science aimed on developing algorithms able to learn from experience instead of performing a predefined explicit routine. In essence, an SVM is a mathematical entity, an algorithm (or recipe) for maximizing a particular mathematical function with respect to a given collection of data. These approaches are becoming popular in a wide variety of biological application (Chan et al., 2016; Liu et al., 2013; Nakano et al., 2014; Noble 2006). A common biomedical application of SVM is the automatic classification of microarray gene expression profiles (Chan et al., 2016; Liu et al., 2013). Theoretically, an SVM can examine the gene expression profile derived from a tumor sample or from peripheral fluid and get to a diagnosis or prognosis (Chan et al., 2016; Golub et al., 1999). In addition, other biological applications of SVMs involve classifying protein, DNA sequences, microarray expression profiles and mass spectra (Noble 2004). Recently, Nakano et al. (2014) used SMV to diagnosis malodour from oral microbiota and methyl mercaptan levels in saliva. They reported that SMV achieved a high accuracy, with a sensitivity of 51.1% and specificity of 95.0%. These algorithms work by constructing a model from examples and then use it to make data-driven decisions. This approach is useful where design of an explicit algorithm is not feasible and never been used before to classify subjects with periodontitis.

The aim of this study was to use a SVM and the 40 bacterial species of the subgingival microbial complexes (Socransky et al., 1998) to compare the subgingival microbial profiles of AgP in young subjects, ChP and PH. The

hypothesis tested was that this analysis could create a model able to differentiate these clinical conditions.

Material and Methods

Subject population

435 subjects in age from 20-67 years who were considered to be periodontally healthy (n=53) or with chronic periodontitis (n=308) and aggressive periodontitis (n=74) were selected from the database of the department of periodontology of the Guarulhos University. Therefore, the analysis of this study was a data compiled from large clinical studies that evaluated the subgingival microbiota of periodontally healthy subjects and periodontitis patients. Those studies were conducted at Guarulhos University (São Paulo, SP, Brazil) from 2004 to 2015, and followed very similar protocols for selection of participants, sample collection and microbial analysis. These studies protocol were approved previously by Guarulhos University's Ethics Committee in Clinical Research.

Clinical examination

The clinical examination was performed always by trained and calibrated examiners. The Intra-examiner variability in all clinical studies ranges between 0.13mm to 0.21mm for PD and 0.22mm to 0.31mm for CAL. The examiners were able to provide reproducible measures under 0.5 mm in all studies included. Visible plaque (0/1), gingival bleeding (0/1), bleeding on probing (BOP, 0/1), suppuration (0/1), probing depth (PD, mm) and clinical

attachment level (CAL, mm) were measured at six sites per tooth (mesiobuccal, buccal, distobuccal, distolingual, lingual and mesiolingual) in all teeth excluding third molars at the baseline visit. PD and CAL measurements were recorded to the nearest millimetre using a North Carolina periodontal probe (Hu-Friedy, Chicago, IL, USA).

AgP, ChP or PH was diagnosed based on the periodontal classification of the American Academy of Periodontology (Armitage 1999). Subjects had at least 20 teeth and had to meet the following criteria in order to be included in this study: GAgP: ≤35 years of age; minimum of six permanent incisors and/or first molars with at least one site each with PD and CAL≥5 mm; minimum of six teeth other than first molars and incisors with at least one site each with PD and CAL≥5 mm; and familial aggregation (at least one other member of the family presenting or with a history of periodontal disease); GChP: ≥35 years of age; minimum of six teeth with at least one site each with PD and CAL≥5 mm; and at least 30% of the sites with PD and CAL≥4mm and presence of BOP; PH: no sites with PD and CAL ≥3 mm and <20% of sites exhibiting gingival bleeding and/or bleeding on probing.

Microbiological examination

Sample collection

Individual subgingival plaque samples were collected from nine non-contiguous interproximal sites per subject. For AgP and ChP groups, three sites at each of the following PD categories were sampled: ≤3 mm, between 4 and 6 mm, and ≥7 mm. Sites with PD≤3 mm were collected from PH group. The selected sites were randomized in different quadrants. After the clinical

parameters had been recorded, the supragingival plaque was removed and the samples were taken with individual sterile Gracey mini five curettes and immediately placed in separate polypropylene tubes containing 150 ml TE (10mM Tris-HCl, 1mM EDTA, pH 7.6). One hundred microlitres of 0.5M NaOH was added to each tube and the samples were dispersed using a vortex mixer.

Checkerboard DNA–DNA hybridization

After collection, the samples were immediately placed in separate Eppendorf tubes containing 0.15 ml of TE (10mM Tris-HCl, 1mM EDTA, pH 7.6) and 100 microliters of 0.5M NaOH was added to each tube. Subsequently, the samples were boiled for 10min and neutralized using 0.8ml of 5M ammonium acetate. The released DNA was then placed into the extended slots of a Minislot 30 apparatus (Immunetics, Cambridge, MA, USA), concentrated on a 15/15 cm positively charged nylon membrane (Boehringer Mannheim, Indianapolis, IN, USA) and fixed to the membrane by baking it at 120^0 C for 20min. The membrane was placed in a Miniblotter 45 (Immunetics) with the lanes of DNA at 90^o to the lanes of the device. Digoxigenin-labelled whole genomic DNA probes for 40 bacterial species were hybridized in individual lanes of the Miniblotter. After hybridization, the membranes were washed at high stringency and the DNA probes were detected using the antibody to digoxigenin conjugated with alkaline phosphatase and chemiluminescence detection. The last two lanes in each run contained standards at concentrations of 10^5 and 10^6 cells of each species. Signals were evaluated visually by comparison with the standards at 10^5 and 10^6 bacterial cells for the test species on the same membrane by a

calibrated examiner (k test = 93%). They were recorded as: 0, not detected; 1, $<10^5$ cells; 2, ~10^5 cells; 3, 10^5-10^6 cells; 4, ~10^6 cells and 5, >10^6 cells. The sensitivity of this assay was adjusted to allow detection of 10^4 cells of a given species by adjusting the concentration of each DNA probe (Mestnik et al., 2010; Socransky et al., 1994). The mean counts (10^5 cells) of individual bacterial species were averaged within each subject and then across subjects in each group. The percentage of the total DNA probe counts was determined initially in each site, then per subject and averaged across subjects in the two groups at each time point. The sum of the individual mean proportion was computed for each microbial complex described by Socransky et al. (1998).

Normalization and data analysis

The total concentration of each sample was normalized to 1. A Principal component analysis (PCA) analysis was performed only for the visualization of the data. Two stages of Machine learning were performed. In the first stage PH subjects were compared to all periodontally diseased patients. In the second stage ChP and AgP were compared. The data was split in each analysis to 50 % train and 50% test. A support vector machine (SVM) classifier was used with a Box constraint of 1 and a linear kernel. The result are presented as a Receiver Operating Characteristic (ROC) curve and the surface below the curve (AUC). Only bacterial expression correlated with an absolute correlation of more than 0.1 in the train set were used in the training. Samples from the same patient were categorically divided to be either in the train or the test set (i.e. no samples from the same patient were used in both the train and the test set).

The analysis was divided into two parts. First we tested whether there was a general difference between the bacterial communities' composition in healthy and periodontally diseased patients. We then tested whether a difference existed between the two diseases studied here. Note that we do not take into consideration the total bacterial load that may be affected by the experimental design. Instead, we analyze the relative levels of different bacteria species.

The fraction of sites with visible plaque, gingival bleeding, BOP and suppuration, as well as mean PD and CAL were computed for each subject and then averaged across subjects in the three groups. The significance of differences among groups was sought using the Kruskal-Wallis test and the Mann-Whitney U test. Chi-square test was employed to compare the differences in the frequency of gender.

Results

Demographic characteristics and clinical parameters of the studied population are presented in Table 1. The mean age of subjects with ChP was significantly higher (45.1+5.9) than the GAgP and PH groups (27.1+3.1 and 35.1+ 9.5; respectively). No difference was observed in the distribution of gender. The mean PD and CAL and the % of sites exhibiting BOP, GI and suppuration were significantly higher in the ChP and GAgP groups than in PH subjects. Periodontally healthy subjects and subjects with GAgP showed less visible plaque (28.1% and 49.5%, respectively) than subjects with ChP (84.6%, $p<0.05$).

The results for the Principal Component Analysis (PCA) showed that the variance of the healthy samples in all PCR directions was much smaller than the one of the periodontally diseased samples (Figure 1) suggesting that while the healthy cases are characterized by a highly uniform bacterial levels, the periodontally diseased samples are much more diverse. The diversity can be observed in the two conditions studied here. Thus the variability is within each periodontally involved population. Given the clear difference between the healthy and diseased populations, the classification of samples based on the relative bacterial load was tested. Indeed a linear SVM supervised classifier produces an AUC of over 0.97 on a test set between the diseased and healthy conditions.

More importantly, the relative bacterial load could distinguish between AgP and ChP with a high accuracy, with a sensitivity of XXX% and specificity of XXX%. (Is it possible to obtain these data?). While the difference between these two conditions is smaller than that between health and disease, as can be seen for example in the PCA based only on the diseased samples (Figure 2), applying, again, a linear SVM to the two diseased conditions produced an AUC (?) of 0.8 on a test set (Figure 3). Thus, not only were the bacterial load different between these two diseases, but this difference was enough to allow for a clear distinction between the two conditions. Specifically, a linear SVM was applied to the first 20 PCA vectors (representing over 95 % of the variability). The classifier defines a clear direction in the bacterial load concentration space with some bacteria correlated with each of the two diseases. The species *Porphyromonas gingivalis* followed by *Tanerella forsythia*, *Fusobacterium ssp. polymorphum*, *Treponema denticola* and

Prevotella denticola were the five species most correlated with ChP. In addition, several species from the orange complex were also associated with this condition, including *Fusobacterium periodonticum*, *Fusobacterium ssp. nucleatum*, *Prevotella nigrescens*, *Eubacterium nodatum*, and *Parvimonas micra*. On the other hand, *Propionibacterium acnes*, *Aggregatibacter actinomycetemcomitans*, *Streptococcus mitis*, *Streptococcus constellatus* and *Treponema socranskii* were the five species more strongly correlated with AgP.

Discussion

The hypothesis tested in this study that a SVM classifier using a panel of 40 bacterial species could differentiate between AgP and ChP, was confirmed.

The VSM analysis showed that the 3 red complex species (*P.gingivalis*, *T. forsythia* and *T. denticola*), as well as, some species from the orange complex (species of *Fusobacterium*, *P. intermedia*, *P. micra* and *E. nodatum*) had a high weight in the mathematic algorithms related to a chronic periodontitis diagnosis. On the other hand, *A. actinomycetemcomitans* was linked to AgP. These data are in agreement with previous studies showing that *A. actinomycetemcomitans* is an important pathogen in the etiology of AgP, and the pathogens from the red and orange complexes are more implicated in the etiology of ChP (Colombo et al., 2002; Moore & Moore 1994; Socransky et al., 1998). Nonetheless, when this same microbial panel was analyzed by conventional statistical approaches, some studies failed to find

major differences between ChP and AgP (Faveri et al., 2009; Rescala et al., 2010; Ximenez-Fyvie et al., 2006). This could be explained by limitations in the microbiological techniques, the number of samples analyzed or by the inability to clinically distinguish AgP from CP (Gajardo et al., 2005; Nibali 2015; Riep et al., 2009). Faveri et al. (2009) showed that the composition of the subgingival microbiota did not differ substantially among localized AgP, generalized AgP and ChP subjects. Similar results have been shown by other authors (Rescala et al., 2010; Ximenez-Fyvie et al., 2006).

To our knowledge, this is the first report in which AgP and ChP could be differentiated by the subgingival microbial profile. Nonetheless, it is important to bear in mind that the main criterion used to include subjects in the two periodontitis groups in this study was age, a parameter that it is not considered in the by the current classification of the American Academy of Periodontology (Armitage 1999). In fact, many clinical researchers in periodontology face the difficulty to select volunteers based on characteristics that are not normally available to the clinician. This is even more critical for the AgP. The three common features of aggressive periodontal diseases according to the Consensus Report of the AAP (Armitage 1999; Lang 1999) are: otherwise clinically healthy subjects; familial aggregation and rapid attachment loss and bone destruction. The first two characteristics are also observed in subjects with ChP, and the latter on is rarely available to the clinician. Determining the rate of attachment loss while selecting subjects for cross-sectional studies is unfeasible, leaving researchers with the alternative of using age as a discriminating factor, by estimating “rapid periodontal destruction” if the individual shows advanced disease in an early age. This

was the case of the database used in this study. Thus, it might be more accurate to say that the statistical model tested in this study is suitable to differentiate between advanced periodontitis in adults and in young individuals, than between ChP and AgP.

Though we understand the infectious nature of periodontitis, the microbial etiology remains an enigma in certain way, as defining the relationship between oral microbial consortia and disease has been precluded by our inability to study complex microbial interactions in the host (Khan et al., 2015). The continuous cataloguing of microbial species associated with disease and elucidation of the interspecies interactions in oral biofilm will contribute to our understanding of how these bacteria may act together and result in either health or disease (Khan et al., 2015). A recent review by Nibali (2015) stated that as we aim to understand host-associated factors and clinical differences between AgP and ChP, hoping to design more targeted management regimes for these conditions, an interesting insight could be given by studies comparing the microbial composition of these diseases. The use of advanced mathematical approaches, as the one used in the current report, might shed some further light on the differences between those two clinical conditions. Those methods could be applied in the future to further differentiate between sub-groups of the diseases (e.g., generalized, localized) and might open new avenues for using population science methods to further explore the potential of specific therapeutic interventions.

The results of the present study indicated that a SVM classifier using a panel of 40 bacterial species was able to distinguish between AgP and ChP.

These results open new avenues for defining specific preventive and treatment protocols for these conditions.

Acknowledgments

The authors declare there are no competing interest for the above manuscript. No funding was received for the presented work.

Table 1. Demographic characteristics and mean (±SD) full-mouth clinical parameters of the subjects in experimental groups.

Variables	Experimental groups			
	Periodontally healthy	Generalized aggressive periodontitis	Generalized chronic periodontitis	*p-value
Subjects (n)	53	308	74	
Age (years)	35.1±9.5^{A}	27.1±3.1^{B}	45.1±5.9^{C}	<0.001
Gender (male:female)	23:30	122:186	32:42	NS
Probing depth (mm)	1.9±0.5^{A}	4.2±1.1^{B}	4.1±1.3^{B}	<0.001
Clinical attachment level (mm)	0.7±0.4 ^{A}	3.8±1.2^{B}	3.8±1.1^{B}	<0.001
% sites with				
Plaque	28.1±7.7^{A}	49.5±14.2^{B}	84.6±12.2^{C}	<0.001
Gingival bleeding	2.0±1.0^{A}	33.2±12.9^{B}	34.1±22.1^{B}	<0.001
Bleeding on probing	2.1±0.8^{A}	46.1±14.2 ^{B}	45.2±27.2^{B}	<0.001
Suppuration	0±0^{A}	4.4±3.7^{B}	3.29±4.1 ^{B}	<0.001

The significance of differences among groups was assessed using the Kruskall-Wallis test (*). The significance of differences between pairs of comparisons was determined using Dunn multiple comparison test and the significances are represented by different capital letters. *SD, standard deviation, NS, not significant*.

Figure 1:

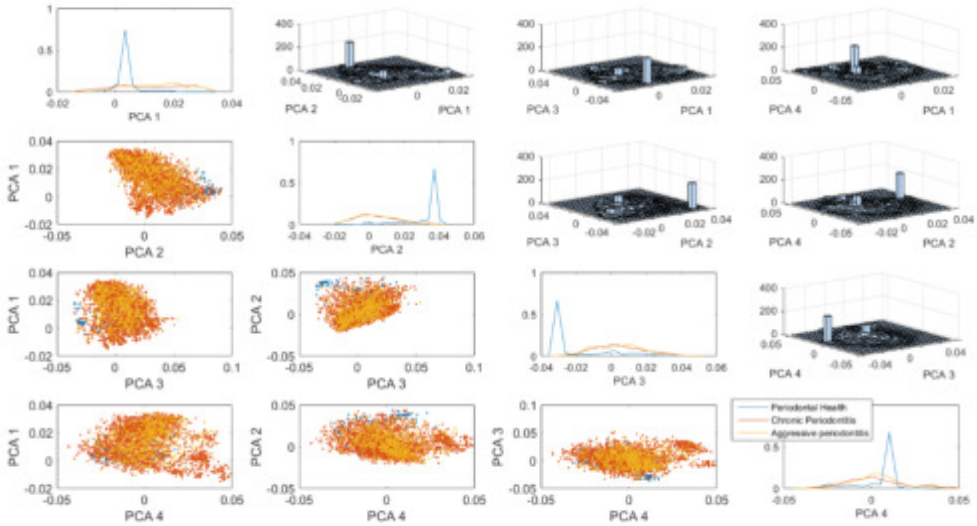

PCA 1
PCA 2
PCA 3
PCA 4
Periodontal Health
Chronic Periodontitis
Aggressive periodontitis

Figure 2:

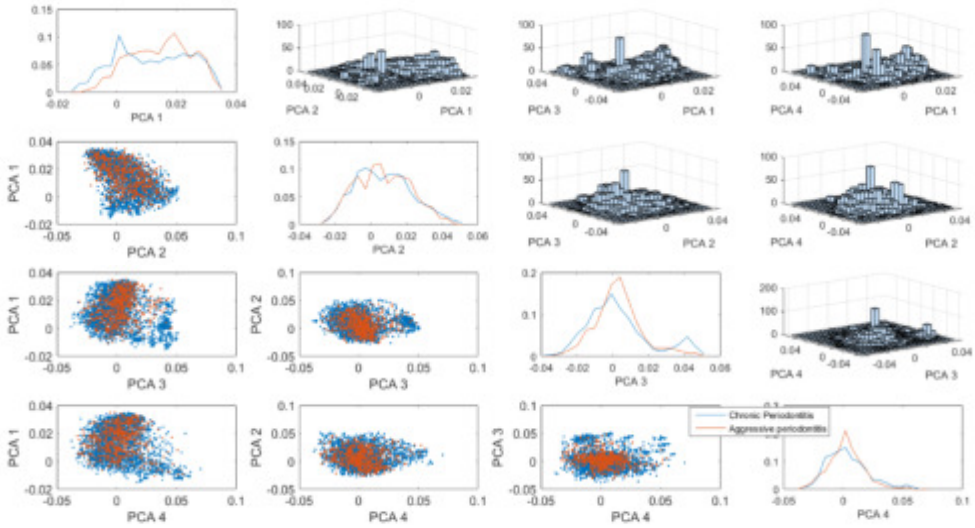

PCA 1
PCA 2
PCA 3
PCA 4
Chronic Periodontitis
Aggressive periodontitis

Figure 3:

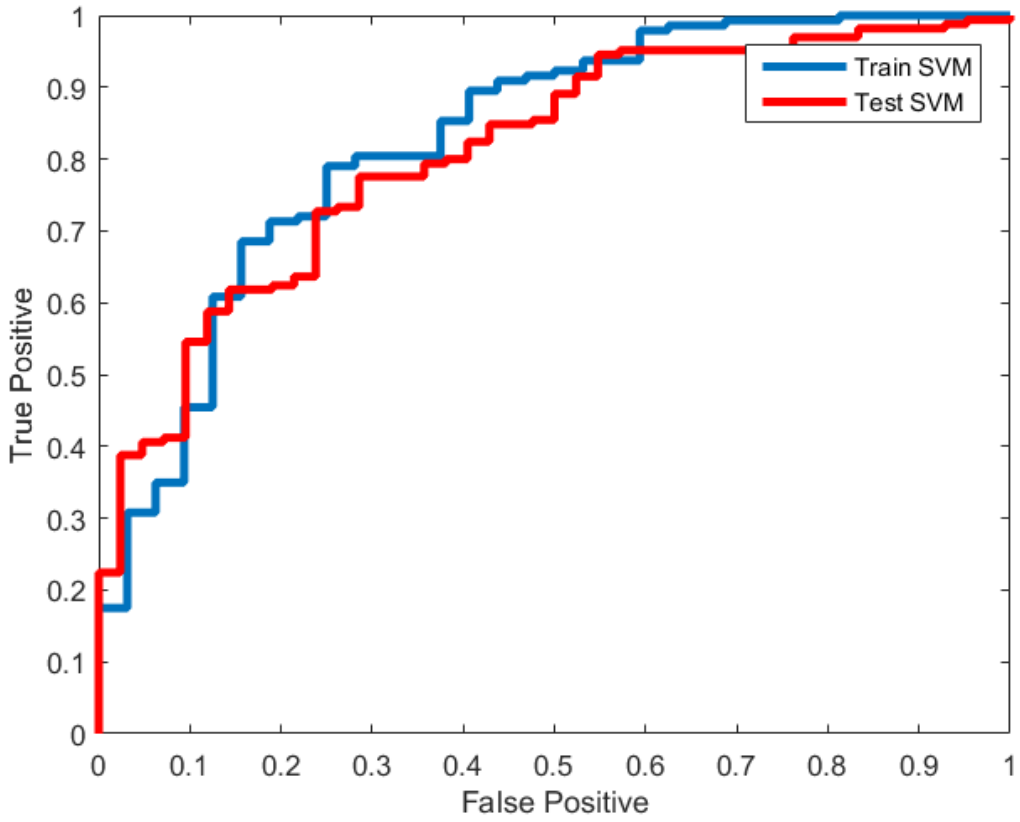
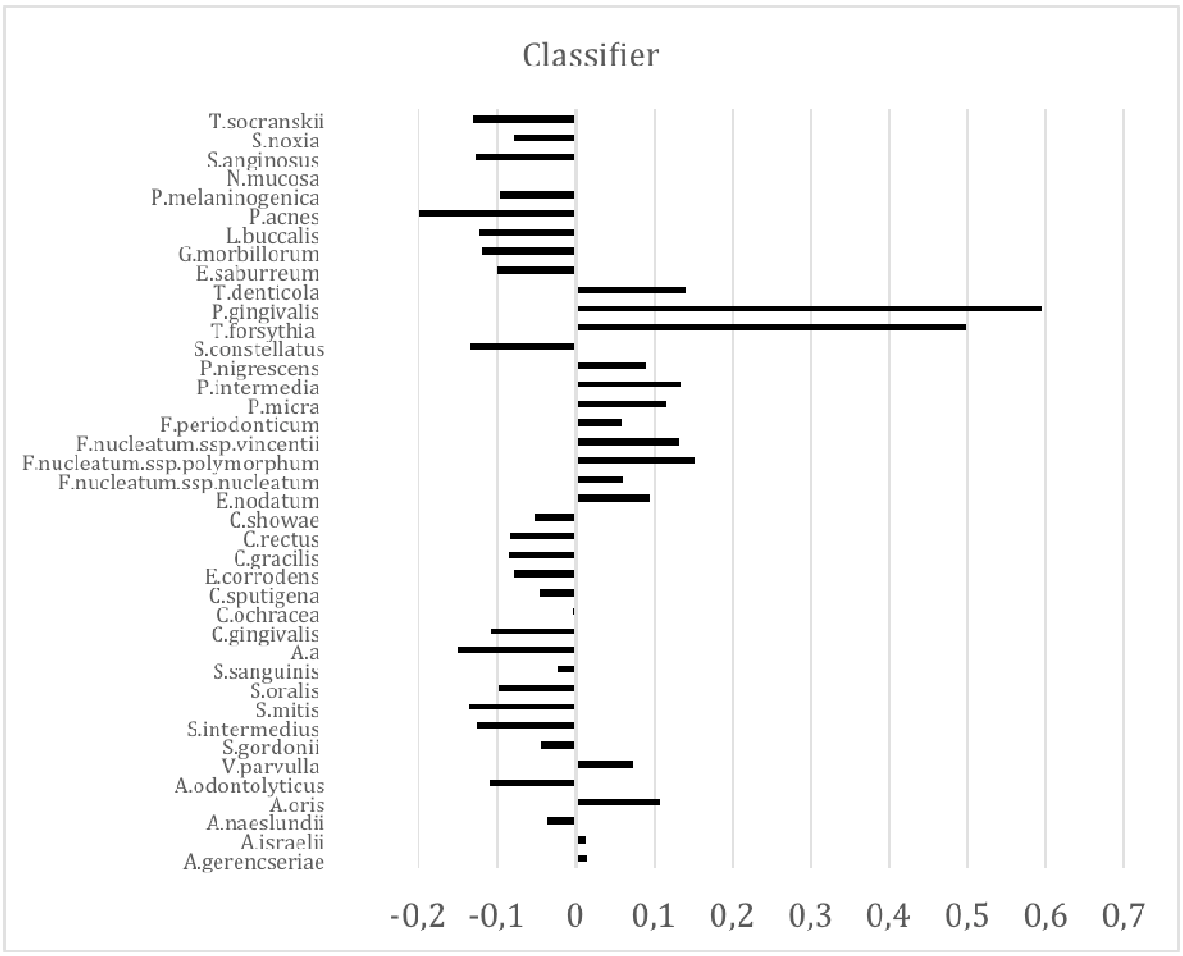

Train SVM
Test SVM
True Positive
False Positive
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 4:


Classifier
T.socranskii
S.noxia
S.anginosus
N.mucosa
P.melaninogenica
P.acnes
L.buccalis
G.morbillorum
E.saburreum
T.denticola
P.gingivalis
T.forsythia
S.constellatus
P.nigrescens
P.intermedia
P.micra
F.periodonticum
F.nucleatum.ssp.vincentii
F.nucleatum.ssp.polymorphum
F.nucleatum.ssp.nucleatum
E.nodatum
C.showae
C.rectus
C.gracilis
E.corrodens
C.sputigena
C.ochracea
C.gingivalis
A.a
S.sanguinis
S.oralis
S.mitis
S.intermedius
S.gordonii
V.parvulla
A.odontolyticus
A.oris
A.naeslundii
A.israelii
A.gerencseriae
-0,2
-0,1
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7

A.gerencseriae	0.354357
A.israelii	0.410579
A.naeslundii	-0.17185
A.oris	0.293528
A.odontolyticus	0.163547
V.parvulla	-0.49523
S.gordonii	0.117063
S.intermedius	-0.02785
S.mitis	-0.38548
S.oralis	0.031204
S.sanguinis	0.22704
A.a	-0.02117
C.gingivalis	-0.21593
C.ochracea	-0.35557
C.sputigena	-0.02281
E.corrodens	0.035917
C.gracilis	-0.25152
C.rectus	0.007709
C.showae	0.058355
E.nodatum	0.053016
F.nucleatum.ssp.nucleatum	-0.53602
F.nucleatum.ssp.polymorphum	0.268752
F.nucleatum.ssp.vincentii	-0.07817
F.periodonticum	0.1817
P.micra	0.92393
P.intermedia	-0.60472
P.nigrescens	-0.14788
S.constellatus	-0.23371
T.forsythia	0.933775
P.gingivalis	1.317112
T.denticola	0.219716

E.saburreum	-0.6419
G.morbillorum	-0.04267
L.buccalis	-0.37804
P.acnes	-0.08666
P.melaninogenica	-0.43854
N.mucosa	-0.12137
S.anginosus	-0.11136
S.noxia	-0.08018
T.socranskii	-0.14869

REFERENCES

Abusleme, L., Dupuy, A.K., Dutzan, N. et al. (2013) The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *ISME J* 7(5):1016-1025.

Albandar, J.M., and Tinoco, E.M. (2002) Global epidemiology of periodontal diseases in children and young persons. *Periodontol 2000* 29:153-176.

Armitage, G.C. (1999) Development of a classification system for periodontal diseases and conditions. *Ann Periodontol* 4(1):1-6.

Chan, W.H., Mohamad, M.S., Deris, S. et al. (2016) Identification of informative genes and pathways using an improved penalized support vector machine with a weighting scheme. *Comput Biol Med* 77:102-115.

Colombo, A.P., Teles, R.P., Torres, M.C. et al. (2002) Subgingival microbiota of brazilian subjects with untreated chronic periodontitis. *J Periodontol* 73(4):360-369.

Costalonga, M., and Herzberg, M.C. (2014) The oral microbiome and the immunobiology of periodontal disease and caries. *Immunol Lett* 162(2 Pt A):22-38.

Faveri, M., Figueiredo, L.C., Duarte, P.M., Mestnik, M.J., Mayer, M.P. and Feres, M. (2009) Microbiological profile of untreated subjects with localized aggressive periodontitis. *J Clin Periodontol* 36(9):739-749.

Feres, M., Gursky, L.C., Faveri, M., Tsuzuki, C.O. and Figueiredo, L.C. (2009) Clinical and microbiological benefits of strict supragingival plaque control as part of the active phase of periodontal therapy. *J Clin Periodontol* 36(10):857-867.

Gajardo, M., Silva, N., Gomez, L. et al. (2005) Prevalence of periodontopathic bacteria in aggressive periodontitis patients in a chilean population. *J Periodontol* 76(2):289-294.

Golub, T.R., Slonim, D.K., Tamayo, P. et al. (1999) Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* 286(5439):531-537.

Goncalves, C., Soares, G.M., Faveri, M. et al. (2016) Association of three putative periodontal pathogens with chronic periodontitis in brazilian subjects. *J Appl Oral Sci* 24(2):181-185.

Hajishengallis, G., Hajishengallis, E., Kajikawa, T. et al. (2016) Complement inhibition in pre-clinical models of periodontitis and prospects for clinical application. *Semin Immunol* 28(3):285-291.

Khan, S.A., Kong, E.F., Meiller, T.F. and Jabra-Rizk, M.A. (2015) Periodontal diseases: Bug induced, host promoted. *PLoS Pathog* 11(7):e1004952.

Kolenbrander, P.E., Palmer, R.J.Jr., Rickard, A.H., Jakubovics, N.S., Chalmers, N.I. and Diaz, P.I. (2006) Bacterial interactions and successions during plaque development. *Periodontol 2000* 42:47-79.

Lamont, R.J., and Hajishengallis, G. (2015) Polymicrobial synergy and dysbiosis in inflammatory disease. *Trends Mol Med* 21(3):172-183.

Lang, N., Bartold, M., Cullinan, M. et al. (1999) Consensus report: Aggressive periodontitis. *Ann Periodontol*. 4:53.

Levin, L., Baev, V., Lev, R., Stabholz, A., Ashkenazi, M. 2006. Aggressive periodontitis among young israeli army personnel. *J Periodontol* 77(8):1392-1396.

Liu, Y., Guo, J., Hu, G. and Zhu, H. (2013) Gene prediction in metagenomic fragments based on the svm algorithm. *BMC Bioinformatics* 14 Suppl 5:S12.

Lopez, N.J., Socransky, S.S., Da Silva, I., Japlit, M.R. and Haffajee, A.D. (2004) Subgingival microbiota of chilean patients with chronic periodontitis. *J Periodontol* 75(5):717-725.

Mestnik, M.J., Feres, M., Figueiredo, L.C., Duarte, P.M., Lira, E.A. and Faveri, M. (2010) Short-term benefits of the adjunctive use of metronidazole plus amoxicillin in the microbial profile and in the clinical parameters of subjects with generalized aggressive periodontitis. *J Clin Periodontol* 37(4):353-365.

Moore, W.E. and Moore, L.V. (1994) The bacteria of periodontal diseases. *Periodontol 2000* 5:66-77.

Nakano, Y., Takeshita, T., Kamio, N. et al. (2014) Supervised machine learning-based classification of oral malodor based on the microbiota in saliva samples. *Artif Intell Med* 60(2):97-101.

Nibali, L. (2015) Aggressive periodontitis: Microbes and host response, who to blame? *Virulence* 6(3):223-228.

Noble, W.S. (2004) Support vector machine applications in computational biology. In *Kernel methods in computational biology*. Schoelkopf, B., Tsuda, K. & Vert, J.-P (eds). Cambridge, MA.: MIT Press, pp. 71-92.

Noble, W.S. (2006) What is a support vector machine? *Nat Biotechnol* 24(12):1565-1567.

Oliveira, R.R., Fermiano, D., Feres, M. et al. (2016) Levels of candidate periodontal pathogens in subgingival biofilm. *J Dent Res* 95(6):711-718.

Perez-Chaparro, P.J., Goncalves, C., Figueiredo, L.C. et al. (2014) Newly identified pathogens associated with periodontitis: A systematic review. *J Dent Res* 93(9):846-858.

Rescala, B., Rosalem, W.Jr., Teles, R.P. et al. (2010) Immunologic and microbiologic profiles of chronic and aggressive periodontitis subjects. *J Periodontol* 81(9):1308-1316.

Riep, B., Edesi-Neuss, L., Claessen, F. et al. (2009) Are putative periodontal pathogens reliable diagnostic markers? *J Clin Microbiol* 47(6):1705-1711.

Socransky, S.S. and Haffajee, A.D. (2005) Periodontal microbial ecology. *Periodontol 2000 38*:135-187.

Socransky, S.S., Haffajee, A.D., Cugini, M.A., Smith, C. and Kent, R.L.Jr. (1998) Microbial complexes in subgingival plaque. *J Clin Periodontol* 25(2):134-144.

Socransky, S.S., Smith, C., Martin, L., Paster, B.J., Dewhirst, F.E. and Levin, A.E. (1994) "Checkerboard" DNA-DNA hybridization. *Biotechniques* 17(4):788-792.

Tonetti, M.S., Claffey, N., European Workshop in Periodontology group C. (2005) Advances in the progression of periodontitis and proposal of definitions of a periodontitis case and disease progression for use in risk factor research. Group c consensus report of the 5th european workshop in periodontology. *J Clin Periodontol* 32 Suppl 6:210-213.

Ximenez-Fyvie, L.A., Almaguer-Flores, A., Jacobo-Soto, V., Lara-Cordoba, M., Moreno-Borjas, J.Y. and Alcantara-Maruri, E. (2006) Subgingival microbiota of periodontally untreated mexican subjects with generalized aggressive periodontitis. *J Clin Periodontol* 33(12):869-877.