Int J Hum Genet, 18(2): 61-67 (2018)
DOI: 10.31901/24566330.2018/18.2.688

Gibbs Sampling Method Identifies Disrupted Pathways and Genes in Periodontitis

Qi-Zhi Zhang[1], Wei Wei[2], Xiu-Min Zhang[3], Xiao-Han Pan[4], Xiao-Qing Xu[4], Yan-Yan Gao[1], Si-Cong Chen[1] and Yan-Ying Zhao[4*]

[1]Outpatient department, The Second Hospital of Shandong University, Jinan, 250033, Shandong Province, China
[2]Department of Nuclear Medicine, The Second Hospital of Shandong University, Jinan, 250033, Shandong Province, China
[3]Gynecological clinic, The Second Hospital of Shandong University, Jinan, 250033, Shandong Province, China
[4]The Department of Health Management, The Second Hospital of Shandong University, Jinan, 250033, Shandong Province, China

KEYWORDS Gibbs Sampling. Key Genes. Markov Chain Monte Carlo. Pathways. Periodontitis. Probability

ABSTRACT Periodontitis is a chronic inflammatory disease triggered by the host immune response. The aim of this study is to explore the disturbed pathways and genes in periodontitis. Transcriptome data of healthy and diseased gingival tissues and human pathway data were recruited from public available database. Then, Gibbs sampling and Markov chain Monte Carlo (MCMC) algorithm were implemented to identify disturbed pathways and key genes. Disturbed pathways and key genes were identified under adjusted posterior value > 0.8. The researchers identified two disturbed pathways (cytokine-cytokine receptor interaction and hematopoietic cell lineage) and two key genes (TNFRSF17 and CXCL6). Gene expression analysis showed that all the disturbed pathways and key genes had increased expression levels in diseased gingival samples compared with healthy samples. The identified pathways and genes may play important role in periodontitis and could be considered as potential biomarkers for early detection and therapy for periodontitis.

INTRODUCTION

Periodontitis is a chronic inflammatory disease triggered by the host immune response. It is associated with the interactions between complex microbial biofilms and inflammatory mediators, leading to the tooth-supporting structure destruction (Hajishengallis 2014; Elhassan et al. 2017). Severe periodontitis can increase the risk of complex diseases, such as diabetes mellitus (Teeuw et al. 2017) , cardiovascular disease (Beukers 2017), and adverse pregnancy outcomes (Madianos et al. 2013). The underlying pathogenesis of periodontitis has not been fully elucidated. Thus, revealing the potential molecular mechanisms of periodontitis may be of great significance and importance for the prevention and treatment of periodontitis.

Previous studies have documented that periodontitis is a complex interplay of numerous parameters acting simultaneously and unpredictably (Hajishengallis 2014; Toker et al. 2017). Many genetic alterations have been observed to influence the likelihood of developing periodontitis (Song et al. 2015; Ni et al. 2017; Toker et al. 2017). However most disease development processes refer to the disruption of a set of genes, thus disclosing disrupted pathways may facilitate a better understanding of how gene perturbations account for the pathogenic procedure of disease. Current advances on high-throughput experimental technologies contribute to the revelation of disease-related genes and pathways involved in the etiology of periodontitis (Demmer et al. 2008; Papapanou et al. 2009).

Gibbs sampling is an extensively used algorithm in statistical inference, especially in Bayesian inference. Gibbs sampling is Markov Chain Monte Carlo (MCMC) algorithm for obtaining a sequence of observations which are approximat-

Address for correspondence:
Yan-Ying Zhao
The Department of Health Management,
The Second Hospital of Shandong University,
No.247 on Beiyuan Road, Jinan, 250033,
Shandong Province, China
Telephone: 86-0531-858875555
Fax: 0531-85875201
E-mail: Zhaoyanying321@126.com

ed from a specified multivariate probability distribution (Huelsenbeck and Ronquist 2001; Walsh 2004). According to the probability distributions, the investigators can identify disturbed pathways and genes in the pathology of complex diseases. In this study, the researchers attempted to implement Gibbs sampling method to identify disturbed pathways and genes in periodontitis based on MCMC algorithm.

Objective

In this study, the researchers try to identify disturbed pathways and genes in periodontitis by Gibbs sampling. Transcriptome data of healthy and diseased gingival tissues and human pathway data were firstly recruited from public available database. Then, all human pathways were converted into Markov chains, and Markov chain Monte Carlo (MCMC) algorithm was implemented to perform posterior inference to identify probability distributions of pathways in Gibbs sampling. Moreover gene expression variation was taken into account to adjust the probability. Disturbed pathways were identified under adjusted posterior value > 0.8. Then Gibbs sampling was implemented to key genes from disturbed pathways. This study provided a wide understanding of periodontitis.

METHODOLOGY

Preparatory Work

Transcriptome Data

The Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) is a public archive of functional genomics data submitted by the research community (Barrett et al. 2013). All data are freely available for reuse to the research community. To identify disturbed pathways and genes associated with periodontitis, the gene expression profiles of healthy and diseased gingival tissues were downloaded from GEO database, under the accession number of GSE10334 (Demmer et al. 2008). The dataset contains 90 systemically healthy non-smokers with moderate to advanced periodontitis (63 with chronic and 27 with aggressive periodontitis), providing a total of 247 individual tissue samples (183 from diseased and 64 from healthy gingival sites). General characteristics of the study participants were described in the study of Demmer et al. (2008). Affymetrix Human Genome U133 Plus2.0 arrays were used to gene chip hybridizations, including 54,675 probe sets. Through data preprocessing (Ma et al. 2006; Pepper et al. 2007), a total of 20,514 genes and their expression data were obtained for subsequent analysis.

Pathway Data

For the discovery of disturbed pathways associated with periodontitis, all human pathways were obtained from Kyoto Encyclopedia of Genes and Genomes (KEGG) database (www.genome.jp/kegg/). In KEGG database, there are a total of 287 pathways (covering 6894 genes). Generally, pathways with too few genes may have insufficient biological content. Thus pathways with gene number <5 were removed from this study, and a total of 280 pathways were obtained for further analysis.

Identifying Disturbed Pathways by Gibbs Sampling

It is well known that Gibbs sampling, a commonly used Bayesian statistical inference, is a Markov chain Monte Carlo algorithm for screening a number of observations that are approximated from a specified posterior probability distribution. For identifying disturbed pathways by Gibbs sampling, the KEGG pathways should be transformed to Markov chains. A Markov chain is a sequence of random variables whose distribution hinges on the value of the previous random variables, and presents an equilibrium distribution (Li et al. 2012). In this study, Gibbs sampling generates a Markov chains of samples, each of which is correlated with nearby samples.

After transforming the KEGG pathways to Markov chains, the posterior inference was performed to estimate the probability distributions of pathways. The posterior distribution was developed by the normalized product of the prior information and a sampling distribution. Gibbs sampling generates samples whose distribution transforms to posterior distribution. In this study, an empty object was firstly defined to carry out the Gibbs sampler, in which the KEGG pathway Markov chain data set was kept. Then Gibbs sampling was implemented to construct k-dimensional random vectors of n samples, and the k-dimensional vectors were initialized randomly

and the remaining one vector was selected by fixing k-1 elements of the vectors. After repeating this process k times, a new Markov chain was generated. Here, k = 10,000 and n = 280, that is 10,000 iterations for the Markov chains of 280 KEGG pathways. A satisfactory convergence of all parameters in all chains was obtained after 2000 iterations. Through Gibbs sampling, k probability of each pathway was obtained, that is, the posterior probability. The probabilityα was defined as follows:

$$\alpha = \frac{\sum_{i=2000}^{10000} p_i}{10000\text{-}2000+1},$$

where P_i was the posterior value in the i^{th} sample.

To improve result reliability, gene expression variation was taken into account to adjust the probability in Gibbs sampling. Based on the gene expression values between two conditions, the researchers calculated the pathway expression differences between periodontitis and healthy conditions using t-test. Then the pathways were ranked in ascending order according to the p-values. Based on the rank values, each pathway was given a correction coefficient c. The adjusted probability was calculated as follows:

$$\alpha_{adj} = a \times c$$

where $c = 1 - \frac{rank_i}{n}$, n was sample numbers, and $rank_i$ was the rank value of sample i according to the p-value.

After adjusting the posterior probability of all KEGG pathways based on gene expression variation, the disturbed pathways were identified under the criterion of adjusted probability > 0.8.

Identifying Key Genes by Gibbs Sampling

After obtaining the disturbed pathways, the researchers attempted to screen the key genes from disturbed pathways using Gibbs sampling. For identifying key genes by Gibbs sampling, the genes in the disturbed pathways were transformed to Markov chains. Then, the researchers performed the posterior inference to estimate the probability distributions of genes in the disturbed pathways. Similar to the identification of disturbed pathways by Gibbs sampling, k-dimensional random vectors of m samples was constructed. Here m was the gene numbers in the disturbed pathways. Similarly, gene expression variation was taken into account to adjust the probability. Based on the gene expression values between two conditions, the gene expression differences between periodontitis and healthy conditions were calculated using t-test. Then the probability was adjusted according to the gene expression differences. This procedure was similar to the section of identifying disturbed pathways by Gibbs sampling. Finally, the key genes were identified under the criterion of adjusted probability > 0.8.

RESULTS

Preparatory Work

The gene expression data of periodontitis were downloaded from GEO database, with accessing number GSE10334, containing 183 diseased and 64 healthy gingival samples. After data preprocessing, a total of 20,514 genes were obtained for subsequent analysis. Then human pathways were obtained from KEGG database, and a total of 280 pathways with at least 5 genes were retained for subsequent analysis.

Disturbed Pathways

Based on the gene expression data and KEGG pathways, Gibbs sampling evaluated the probabilities of these pathways using Markov chain Monte Carlo algorithm. The probability distribution of all KEGG pathways was illustrated in Figure 1. Under the criterion of probability > 0.8, the researchers obtained two disturbed pathways in this study, cytokine-cytokine receptor interaction (α=0.994) and hematopoietic cell lineage (α =0.834). Then the probability was adjusted by gene expression variation between periodontitis and healthy conditions. The disturbed pathways cytokine-cytokine receptor interaction and hematopoietic cell lineage showed adjusted probabilities of 0.990 and 0.819, respectively. The expression levels of cytokine-cytokine receptor interaction and hematopoietic cell lineage in periodontitis and healthy conditions were showed in Figure 2. The researchers could easily found that both disturbed pathways had increased expression levels in diseased gingival samples, relative to healthy samples.

Key Genes

To further identify the key genes in the disturbed pathways, the researchers performed

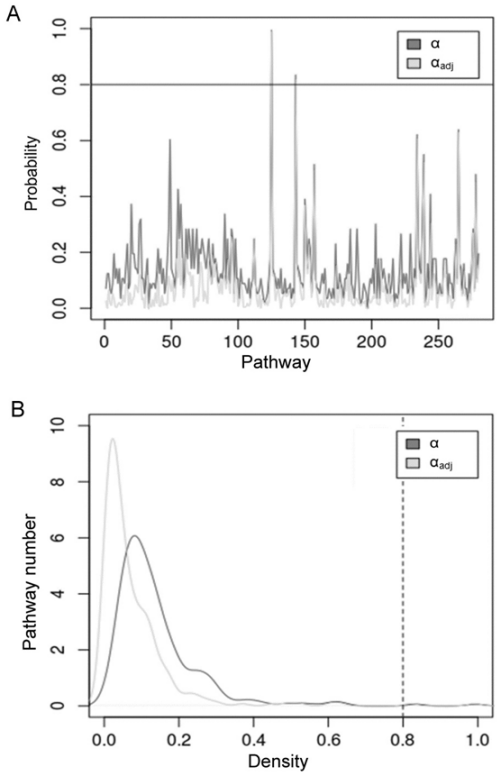

Fig. 1. A: The probabilities distribution of all human pathways. B: The density distribution of pathways. The horizontal solid line and vertical dotted line represented the threshold value of posterior value of 0.8. The dark grey line is the initial probabilities, and the light grey line is the adjusted probabilities

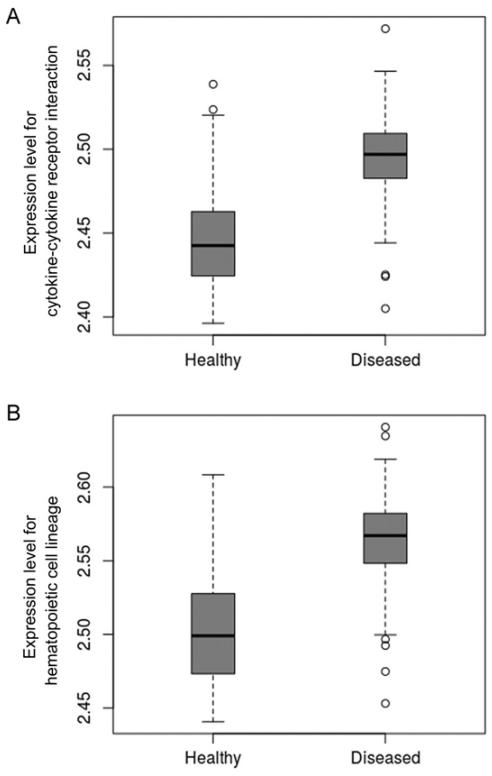

Fig. 2. A: The expression levels of cytokine-cytokine receptor interaction in diseased and healthy gingival samples. B: The expression levels of hematopoietic cell lineage in diseased and healthy gingival samples

Gibbs sampling to estimate the probability distributions of genes in the disturbed pathways. The researchers analyzed the gene composition of two disturbed pathways, and found that cytokine-cytokine receptor interaction contained 253 genes and hematopoietic cell lineage contained 83 genes. After removing genes absent from expression profile and repetitive genes, a total of 304 genes were remained for Gibbs sampling. The probability distribution of genes in disturbed pathways were illustrated in Figure 3. Under the criterion of probability > 0.8, the researchers identified two key genes associated with periodontitis, including TNFRSF17 (α = 0.931) and CXCL6 (α= 0.987). Then gene expression variation was employed to adjust the probability. The key genes TNFRSF17 and CXCL6 had adjusted probabilities of 0.864 and 0.854, respectively. Gene expression analysis showed that both TNFRSF17 and CXCL6 showed increased expression levels in diseased gingival samples compared with healthy samples, as shown in Figure 4.

DISCUSSION

In this study, the researchers identified two disturbed pathways involved in periodontitis via Gibbs sampling. Cytokine-cytokine receptor interaction showed an adjusted probability of 0.990. Cytokines are a broad of small proteins, such as chemokines, interferons, interleukins and lymphokines, that play important roles in host responses to infection, immune responses, and inflammation (Elahi et al. 2017; Kuwabara et al. 2017; Waldmann 2017). Because of biological

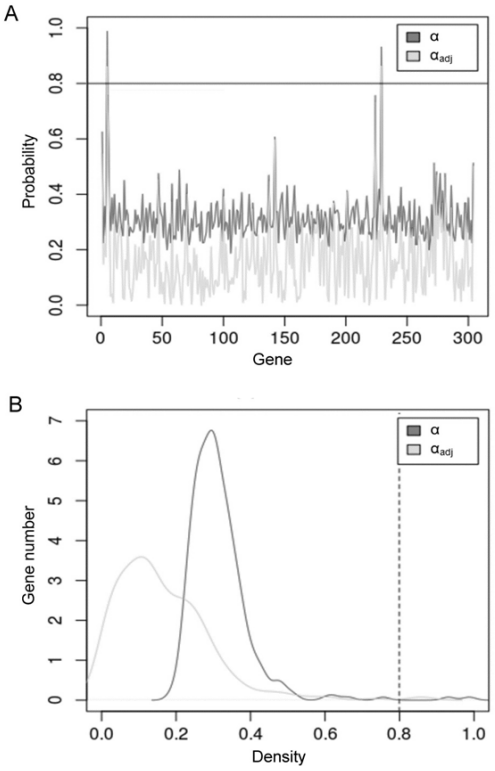

Fig. 3. A: The probabilities distribution of genes in the disturbed pathways. B: The density distribution of genes in the disturbed pathways. The horizontal solid line and vertical dotted line represented the threshold value of posterior value of 0.8. The dark grey line is the initial probabilities, and the light grey line is the adjusted probabilities

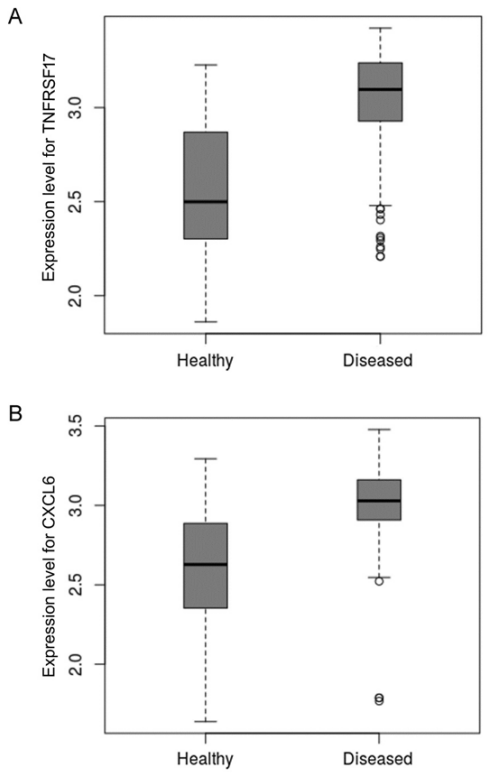

Fig. 4. The expression levels of TNFRSF17 in diseased and healthy gingival samples. B: The expression levels of CXCL6 in diseased and healthy gingival samples.

activities of cytokines in inflammatory tissue destruction, researches on the pathogenesis of periodontitis have focused on cytokines (Lavu et al. 2017). Numerous studies have been performed to elucidate the association of cytokines with periodontitis, and confirmed that several cytokines are involved in the progress of periodontitis, such as tumor necrosis factor alpha, interferon gamma, and interleukins (Bascones et al. 2005; Schenkein et al. 2007; Cintra et al. 2016; Stadler et al. 2016). It is well known that cytokines act through cytokine receptors (Schreiber and Walter 2010). Cytokine receptors modify cytokines biological activity and regulate cytokine-mediated biological events in an antagonistic or agonistic way. Cytokine-cytokine receptor interaction is a proper way to understand cytokine-specific immunological responses (Oppenheim 2001). Demmer et al. (2008) also indicated that cytokine-cytokine receptor interaction pathway was differentially regulated in diseased gingival tissues. Song et al. (2015) reported that the up-regulated genes in periodontitis were mainly involved in staphylococcus aureus infection and cytokine-cytokine receptor interaction. Moreover, the researchers identified two key genes TNFRSF17 and CXCL6 associated with periodontitis using Gibbs sampling. Both of TNFRSF17 and CXCL6 were involved in cytokine-cytokine receptor interaction pathway, implying that cytokine-cytokine receptor interaction disorder played critical roles in the development of periodontitis.

TNFRSF17, also known as B-cell maturation antigen, is a cell surface receptor of the tumor necrosis factor receptor superfamily and binds

B-cell activating factor (Laabi et al. 1994). It is expressed in B lineage cells resulting in potent B cell activation and NF-êB and MAPK8/JNK activation (Shen et al. 2016). TNFRSF17 plays important roles in B cell development and autoimmune response, and has been implicated in various diseases. A gene network-based microarray analysis showed that TNFRSF17 was most up-regulated in generalized aggressive periodontitis (Guzeldemir-Akcakanat et al. 2016). Moreover, TNFRSF17 showed an increased expressed in aged periodontitis approximately10-fold compared with healthy adult tissues (Ebersole et al. 2016). CXCL6, also known as granulocyte chemotactic protein 2, is a small cytokine that belongs to the CXC chemokine family, which is involved in neutrophil recruitment and migration. Kebschull et al. (2009) indicated that CXCL6 correlated with the severity of periodontitis and might act as a hitherto unrecognized functional adjunct to IL-8 in diseased gingival tissues. Epigenetic modification of inflammatory gene CXCL6 was also speculated to be related to the occurrence of aggressive periodontitis (Schulz et al. 2016). Cytokine-cytokine receptor interaction and the associated genes may provide novel targets for better understanding of the local adaptive immune response to periodontitis.

The present study identified hematopoietic cell lineage pathway as a disturbed pathway in periodontitis. Hematopoietic cell lineage is involved in blood-cell development progresses, which can differentiate into a multi-lineage committed progenitor cell: a common lymphoid progenitor or a common myeloid progenitor. Cells undergoing these differentiation process express a stage- and lineage-specific set of surface markers. Lv et al. (2015) indicated that hematopoietic cell lineage pathway had higher alter score in Porphuromonas gingivalis-infected periodontitis cells, relative to Aggregatibacter actinomycetemcomitans-infected periodontitis cells. Few studies give the direct relationship between hematopoietic cell lineage pathway and periodontitis. Further elucidation and characterization may identify new mechanisms of periodontitis pathogenesis and future opportunity for therapeutic intervention of periodontitis.

CONCLUSION

In this work, the researchers identified two disturbed pathways (cytokine-cytokine receptor interaction and hematopoietic cell lineage) and two key genes (TNFRSF17 and CXCL6) in periodontitis utilized an efficient Gibbs sampling.

RECOMMENDATIONS

This study contributed to the understanding of underlying pathogenesis and the identified pathways as well as the key genes might be potential biomarkers for early detection and therapy for periodontitis. Furthermore, related experiments are urgently needed to explore and verify the roles of the disturbed pathways and key genes in periodontitis development.

REFERENCES

Barrett T, Wilhite SE, Ledoux P et al. 2013. NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Res,* 41(Database Issue): D991-D995.

Bascones A, Noronha AS, Gomez M et al. 2005. Tissue destruction in periodontitis: Bacteria or cytokines fault? *Quintessence Int,* 36(4): 299-306.

Beukers NG, van der Heijden GJ, van Wijk AJ, Loos BG 2017. Periodontitis is an independent risk indicator for atherosclerotic cardiovascular diseases among 60 174 participants in a large dental school in the Netherlands. *J Epidemiol Community Health,* 71(1): 37-42.

Cintra LT, Samuel RO, Azuma MM et al. 2016. Multiple apical periodontitis influences serum levels of cytokines and nitric oxide. *J Endod,* 42(5): 747-751.

Demmer RT, Behle JH, Wolf DL et al. 2008. Transcriptomes in healthy and diseased gingival tissues. *J Periodontol,* 79(11): 2112-2124.

Ebersole JL, Kirakodu SS, Novak MJ et al. 2016. Transcriptome analysis of B cell immune functions in periodontitis: Mucosal tissue responses to the oral microbiome in aging. *Front Immunol,* 7: 272.

Elahi S, Thompson DR, Van Kessel J et al. 2017. Protective role of passively transferred maternal cytokines against bordetella pertussis infection in newborn piglets. *Infection and Immunity,* 85(4): IAI.01063-01016.

Elhassan AT, Alfakry H, Peeran SW 2017. Reasons to seek periodontal treatment in a Libyan community. *Dentistry and Medica Research,* 5(2): 38-42.

Guzeldemir-Akcakanat E, Sunnetci-Akkoyunlu D, Orucguney B et al. 2016. Gene-expression profiles in generalized aggressive periodontitis: A gene network-based microarray analysis. *J Periodontol,* 87(1): 58-65.

Hajishengallis G 2014. Immunomicrobial pathogenesis of periodontitis: Keystones, pathobionts, and host response. *Trends Immunol,* 35(1): 3-11.

Huelsenbeck JP, Ronquist F 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics,* 17(8): 754-755.

Kebschull M, Demmer R, Behle JH et al. 2009. Granulocyte chemotactic protein 2 (gcp-2/cxcl6) complements

interleukin-8 in periodontal disease. *J Periodontal Res,* 44(4): 465-471.
Kuwabara T, Ishikawa F, Kondo M et al. 2017. The role of IL-17 and related cytokines in inflammatory autoimmune diseases. *Mediators of Inflammation,* 2017(12): 3908061.
Laabi Y, Gras MP, Brouet JC et al. 1994. The BCMA gene, preferentially expressed during B lymphoid maturation, is bidirectionally transcribed. *Nucleic Acids Res,* 22(7): 1147-1154.
Lavu V, Venkatesan V, Venugopal P et al. 2017. Clinical relevance of cytokines gene polymorphisms and protein levels in gingival cervical fluid from chronic periodontitis patients. *Iran J Immunol,* 14(1): 51-58.
Li HD, Xu QS, Liang YZ 2012. Random frog: An efficient reversible jump Markov chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification. *Anal Chim Acta,* 740: 20-26.
Lv J, Zhu YX, Liu YQ et al. 2015. Distinctive pathways characterize A. actinomycetemcomitans and P. gingivalis. *Mol Biol Rep,* 42(2): 441-449.
Ma L, Robinson LN, Towle HC 2006. ChREBP• Mlx is the principal mediator of glucose-induced gene expression in the liver. *Journal of Biological Chemistry,* 281(39): 28721-28730.
Madianos PN, Bobetsis YA, Offenbacher S 2013. Adverse pregnancy outcomes (APOs) and periodontal disease: Pathogenic mechanisms. *J Periodontol,* 84(4 Suppl): S170-S180.
Ni X, Jia BC, Yu HD et al. 2017. Comprehensive analysis of interleukin-8 gene polymorphisms and periodontitis susceptibility. *Oncotarget,* 8(30): 48996-49004.
Oppenheim JJ 2001. Cytokines: Past, present, and future. *Int J Hematol,* 74(1): 3-8.
Papapanou PN, Behle JH, Kebschull M et al. 2009. Subgingival bacterial colonization profiles correlate with gingival tissue gene expression. *BMC Microbiol,* 9: 221.
Pepper SD, Saunders EK, Edwards LE et al. 2007. The utility of MAS5 expression summary and detection call algorithms. *BMC Bioinformatics,* 8: 273.
Schenkein HA, Barbour SE, Tew JG 2007. Cytokines and inflammatory factors regulating immunoglobulin production in aggressive periodontitis. *Periodontol,* 45: 113-127.
Schreiber G, Walter MR 2010. Cytokine-receptor interactions as drug targets. *Curr Opin Chem Biol,* 14(4): 511-519.
Schulz S, Immel UD, Just L et al. 2016. Epigenetic characteristics in inflammatory candidate genes in aggressive periodontitis. *Hum Immunol,* 77(1): 71-75.
Shen X, Guo Y, Qi J et al. 2016. Binding of B-cell maturation antigen to B-cell activating factor induces survival of multiple myeloma cells by activating Akt and JNK signaling pathways. *Cell Biochem Funct,* 34(2): 104-110.
Song L, Yao J, He Z et al. 2015. Genes related to inflammation and bone loss process in periodontitis suggested by bioinformatics methods. *BMC Oral Health,* 15: 105.
Stadler AF, Angst PD, Arce RM et al. 2016. Gingival crevicular fluid levels of cytokines/chemokines in chronic periodontitis: A meta-analysis. *J Clin Periodontol,* 43(9): 727-745.
Teeuw WJ, Kosho MXF, Poland DCW et al. 2017. Periodontitis as a possible early sign of diabetes mellitus. *BMJ Open Diabetes Research & Care,* 5(1): e000326.
Toker H, Gorgun EP, Korkmaz EM 2017. Analysis of IL-6, IL-10 and NF-kappaB gene polymorphisms in aggressive and chronic periodontitis. *Cent Eur J Public Health,* 25(2): 157-162.
Waldmann TA 2017. Cytokines in cancer immunotherapy. *Cold Spring Harbor Perspectives in Biology,* 2017: a028472.
Walsh B 2004. Markov Chain Monte Carlo and Gibbs Sampling. *Lecture Notes for EEB 581.*

Paper received for publication on March 2018
Paper accepted for publication on May 2018