

RESEARCH

Open Access



Transcriptomics and methylomics in chronic periodontitis with tobacco use: a pilot study

Young-Dan Cho^{1,2†}, Pil-Jong Kim^{3†}, Hong-Gee Kim³, Yang-Jo Seol¹, Yong-Moo Lee¹, Young Ku¹, In-Chul Rhyu¹ and Hyun-Mo Ryoo^{2*}

Abstract

Background: Accumulating evidence suggests that tobacco smoking affects the susceptibility to and severity of chronic periodontitis. Epigenetics may explain the role of smoking in the development and progress of periodontal disease. In this study, we performed transcriptomic and methylomic analyses of non-periodontitis and periodontitis-affected gingival tissues according to smoking status.

Methods: Human gingival tissues were obtained from 20 patients, including non-smokers with and without periodontitis ($n = 5$ per group) and smokers with and without periodontitis ($n = 5$ per group). Total RNA and genomic DNA were isolated, and their quality was validated according to strict standards. The Illumina NextSeq500 sequencing system was used to generate transcriptome and methylome datasets.

Results: Comprehensive analysis, including between-group correlation, differential gene expression, DNA methylation, gene set enrichment, and protein-protein interaction, indicated that smoking may change the transcription and methylation states of extracellular matrix (ECM) organization-related genes, which exacerbated the periodontal condition.

Conclusions: Our results suggest that smoking-related changes in DNA methylation patterns and subsequent alterations in the expression of genes coding for ECM components may be causally related to the increased susceptibility to periodontitis in smokers as they could influence ECM organization, which in turn may have an effect on disease characteristics.

Keywords: DNA methylation, Epigenomics, Extracellular matrix, Periodontal disease, Smoking, Transcriptome

Background

Periodontal diseases are typical inflammatory conditions caused by bacterial infection and promoted by environmental factors or other modifying factors [1]. Chronic periodontitis presents a destructive periodontal disease that leads to alveolar bone resorption [2]. Tobacco smoking is considered a major risk factor, and many studies have demonstrated that smoking alters the development and progression of periodontitis [3–5]. Other risk factors that modify the host response to the challenge of bacterial

infection and may induce periodontitis are alcohol, diet, pollution, and drugs [6].

An increasing number of recent studies have focused on the role of epigenetic events in the development of various diseases [7, 8]. Unlike genetics which analyzes changes in the DNA sequence, epigenetics represents the study of cellular or physiological phenotypic trait variations caused by environmental or external factors, which modulate gene expression without altering DNA sequence [9]. Epigenetic modifications include chemical alteration of DNA and associated proteins such as histones, which leads to chromatin remodeling and plays an important role in regulating gene expression [10]. Among these effects, DNA methylation is a common epigenetic mechanism observed in human cells [11].

* Correspondence: hmryoo@snu.ac.kr

†Equal contributors

²Department of Molecular Genetics, School of Dentistry, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea
Full list of author information is available at the end of the article

DNA methylation carried out by DNA methyltransferases typically occurs in CpG dinucleotide-rich regions termed “CpG islands,” which are mainly located in gene promoters, and is associated with gene silencing [12]. The methylated sites interact with the methyl-CpG-binding domain proteins (MBDs) which in turn recruit histone deacetylase-containing complexes and induce histone condensation. Moreover, histones can also be directly modified by methylation. Both mechanisms block the binding of transcription factors to gene loci; however, while histone modification is transient, DNA methylation exhibits a more stable nature of gene regulation [13].

Some reports have suggested that the CpG methylation status of inflammation-related genes (e.g., *IL-2* and *IL-8*) is implicated in gene expression in chronic periodontitis [12, 14, 15]. Most epigenetic studies of periodontitis used low-throughput-level screening and were focused on host response to bacterial infection [16, 17], while the impact of environmental or external factors on DNA methylation was somewhat neglected. The aim of this study was to test a hypothesis that tobacco smoking could change the epigenetic state of periodontal cells and modulate the susceptibility to or severity of periodontitis. To determine whether this is the case, we performed comprehensive genome-wide high-throughput analysis using RNA sequencing (RNA-seq) combined with reduced representation bisulfite sequencing (RRBS) in smokers and non-smokers with and without periodontitis.

Methods

Patient selection

The study population included 20 generally healthy patients without and with chronic periodontitis (Table 1). The smoker group (*n* = 10) comprised people who smoked for at least 5 years; individuals who had never smoked were included in the non-smoker group (*n* = 10). Statistical analyses were performed using SPSS 12.0 (SPSS Inc., Chicago, IL, USA). For analysis of patient age and smoking condition, a two-tailed Wilcoxon rank sum test was conducted. For analysis of sex, two-tailed Fisher’s exact test was used. Visual inflammation signs, probing

pocket depth (PPD), bleeding on probing (BOP), and alveolar bone resorption in radiography were used to diagnose periodontitis. Gingival tissue of patients without any clinical inflammatory signs, BOP, and alveolar bone loss, and with PPD ≤ 4 mm served as control (non-periodontitis group). Chronic periodontitis was determined by the presence of chronic inflammatory signs, PPD ≥ 6 mm, BOP, and alveolar bone loss (chronic periodontitis group). Exclusion criteria were acute or aggressive periodontitis, any severe systemic disease that could affect the periodontal condition, current pregnancy or lactation, and the use of systemic antibiotics or anti-inflammatory medications within 6 months prior to baseline.

Collection of gingival tissue samples

Gingival tissues were resected from maxillary or mandibular molar sites in the process of crown lengthening for the non-periodontitis group or open flap debridement for the chronic periodontitis group. All biopsy sites were subjected to surgical procedures for the first time, and the specimens were obtained by the same experienced periodontal surgeon. After local anesthesia, buccal and palatal/lingual scalloped internal incision was made around the teeth away from the free gingival margin. The size of collected gingival sample was approximately 0.5 × 1.0 cm, which included the epithelium and connective tissue. Tissues were immediately transferred to the laboratory, where half of each sample was used for DNA and another half for RNA isolation.

RNA isolation, quality control, and sample preparation for RNA-seq

Total RNA was isolated from gingival tissue using the RNeasy®-Mini kit (Qiagen, Valencia, CA, USA). RNA quality was evaluated using the RNA 6000 NanoLabChip Kit for the Bioanalyzer system (Agilent Technologies, Santa Clara, CA, USA). All samples had an RNA integrity value > 8. A cDNA library was generated for each sample, and its quality was evaluated using the Bioanalyzer system.

Table 1 Background of the study population

	Non-smoker			Smoker			<i>P</i>
	Non-periodontitis NN, (N = 5)	Periodontitis NP, (N = 5)	<i>P</i>	Non-periodontitis SN, (N = 5)	Periodontitis SP, (N = 5)	<i>P</i>	
Age (years)	44 (41–55)	53 (52–55)	0.525*	52 (47–54)	56 (52–56)	0.205*	0.676*
Men:women	3:2	1:4	0.444†	5:0	4:1	0.206†	0.350†
Cigarettes/day	–	–	–	17.5 (5–25)	20.5 (8–22)	0.917*	–
Years smoking	–	–	–	16.5 (16.5–24)	27 (17–27.5)	0.295*	–

Values are median (interquartile range)

**P* values are calculated by two-tailed Wilcoxon rank sum test

†*P* values are calculated by two-tailed Fisher’s exact test

DNA isolation, quality control, and sample preparation for reduced representation bisulfite sequencing (RRBS)

Genomic DNA (gDNA) was isolated from gingival tissue using the DNeasy® Blood & Tissue Kit (Qiagen), and its quality was evaluated using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The ratio of the absorbance at 260 and 280 nm (OD260/280) was used as an indicator of sample purity: DNA was considered pure at OD260/280 of 1.8–2.0. For additional quality check, agarose gel electrophoresis was performed to exclude such effects as RNA interference and DNA nicking or damage. Half of each gDNA sample was subjected to bisulfite conversion using the EpiTect® Bisulfite kit (Qiagen).

RNA sequencing

In total, 10 non-periodontitis and 10 chronic periodontitis samples yielded a sufficient amount of RNA (300 ng) for subsequent analysis. RNA sequencing libraries were prepared using the Illumina TruSeq RNA library kit, and paired-end sequencing with 100 bp reads was performed using the Illumina Nextseq500 platform, resulting in 46.1 M reads per sample in average. For removing adapters, and trimming and quality control, we used Trimmomatic [18]. PE-phred33 ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 MINLEN:75 2. STAR15 [19] was utilized to align the fastq files to the human GRCh37/hg19 reference genome with the following settings: maximum intron size, 500 kb; minimum intron size, 20; four mismatches allowed. HTSeq (version HTSeq-0.6.1) [20] was performed to produce raw counts.

Comparison of gene expression profiles between groups

To assess the homogeneity of gene expression profiles between groups, Spearman's rank correlation coefficient (SRCC) of log-transformed gene expression for each sample was calculated using the `cor` function in R (<http://www.r-project.org>). The median and interquartile ranges of SCC for all between-group comparisons were calculated, and scatter plots of the worst SRCC in each group were obtained as representative plots. SRCC values were compared by Wilcoxon rank sum test, and $P < 0.05$ was considered statistically significant.

Differential expression analysis using DESeq2

DESeq2 in the R-package was used to identify differentially expressed (DE) genes in each condition [21]. In DESeq2, the Wald test specifically developed for raw RNA-seq count data with a negative binomial model was used to identify differentially expressed genes. Genes with $P < 0.05$ were used for further enrichment analysis.

DNA methylation analysis

The gDNA methylation profiles were analyzed using reduced representation bisulfite sequencing. Briefly, we treated gDNA with sodium bisulfite using the EpiTech® Bisulfite kit (Qiagen), digested with the *MspI* restriction enzyme, and selected fragments averaging 100 to 250 bp. We multiplexed four samples per lane and sequenced the libraries using single-end 100-bp reads and the Illumina Nextseq500 platform; as a result, an average of 25.1 M reads per sample were obtained. We performed initial quality control for fastq files using FastaQC. The reads were aligned using the Bowtie2 end-to-end alignment mode allowing four mismatches; hg19 was used as a reference genome [22]. BS-seeker2 [23] was used for calling the methylation status of the individual CpG sites. The resultant text files of BS-seeker2 were cleaned up to obtain the methylation percentage of individual CpG sites for each case. In a 1000-bp window (from +500 to -500 bp) based on transcription start sites of all genes identified using RNA-Seq, CpG methylation percentages were obtained to assess the effect of methylation near transcription start sites on gene expression. For the non-smoker/non-periodontitis (NN), non-smoker/periodontitis (NP), smoker/non-periodontitis (SN), and smoker/periodontitis (SP) groups (Table 1), the statistical significance of CpG methylation for each gene was calculated to identify statistically significant methylation near gene loci. Genes with $P < 0.05$ were used for further enrichment analysis.

Correlation between methylation and gene expression

To analyze the relationship between methylation and expression of genes that showed statistically significant differences in both parameters, regression analysis was performed. To determine whether differential DNA methylation occurs in subjects with periodontitis compared to normal subjects, a regression model with methylation (x_{i1}) and periodontitis (x_{i2}) as the predictor variables and gene expression (y_i) as the response variable was fitted by R function `aov`:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + \epsilon_i$$

where i represents each gene. This model was compared to the null linear model having only methylation (x_{i1}) as the predictor variable:

$$y_i = \beta_0 + \beta_1 x_{i1} + \epsilon_i$$

An analysis of variance (ANOVA) of the two models was conducted to elucidate the effect of periodontitis on the regression model.

Gene set enrichment analysis

To analyze biological significance of the obtained methylation data, we performed gene set enrichment analysis (GSEA) of the Gene Ontology Biological Processes (GO-BP). The GO term source to calculate GSEA was obtained using the GO.db library (version 2.1) of R. The GSEA test of each group was performed by manipulating the clusterProfiler library of R [24]. Gene sets with $q < 0.05$ were considered to be statistically significant.

Protein-protein interaction

Based on the assumption that the products of methylated genes might directly or indirectly interact with other proteins, we obtained a list of genes associated with protein-protein interaction (PPI) by performing PPI analysis for DE genes exhibiting DNA methylation in non-smoker and smoker cases. We extracted PPI-affected genes in the database from iRefIndex 13.0 [25] by manipulating the iRefR library which bridges the iRefIndex database to an R environment. The gene list obtained by PPI was analyzed by GSEA to identify gene sets with potential relation to both DE genes and DNA methylation.

Results

Study participants

Twenty patients, 13 men and 7 women, were included in this study (Table 1). The median age of non-smokers in the NN and NP groups was 44 and 53 years, respectively, and that of smokers in the SN and SP groups was 52 and 56 years, respectively. In general, there were more men than

women in the smoker group ($P = 0.350$), and patients in the periodontitis group were slightly older compared to the healthy group ($P = 0.676$), but the difference was not significant. Furthermore, the mean number of cigarettes per day and smoking duration tended to be lower in the SN group compared to the SP group, but the difference was not significant ($P = 0.917$ and $P = 0.295$, respectively).

Correlation of gene expression with smoking

In the correlation analysis of the RNA-seq data, we considered smoking (Fig. 1a) and periodontal condition as factors (Fig. 1b). When smoking status was considered, the median SRCC between the two non-smoker groups (0.968) was higher than that between the two smoker groups (0.957) or between non-smokers and smokers (0.958) (Fig. 1a). When Wilcoxon rank sum test was performed for each comparison correlation values, comparisons between the correlation of the non-smoking group and that of other groups were found to be statistically significant ($P < 0.01$). When periodontal status was considered (Fig. 1b), the two periodontitis groups exhibited the highest mean correlation coefficient (0.962). In the Wilcoxon rank sum test, no group comparison was found to be statistically significant.

Differential gene expression pattern in periodontitis with a smoking factor

The number of DE genes without controlling for the false discovery rate (FDR) identified between NN and SN cases was 1841 (Additional file 1: Table S1). Among the DE genes, 1105 were upregulated and 736 were downregulated

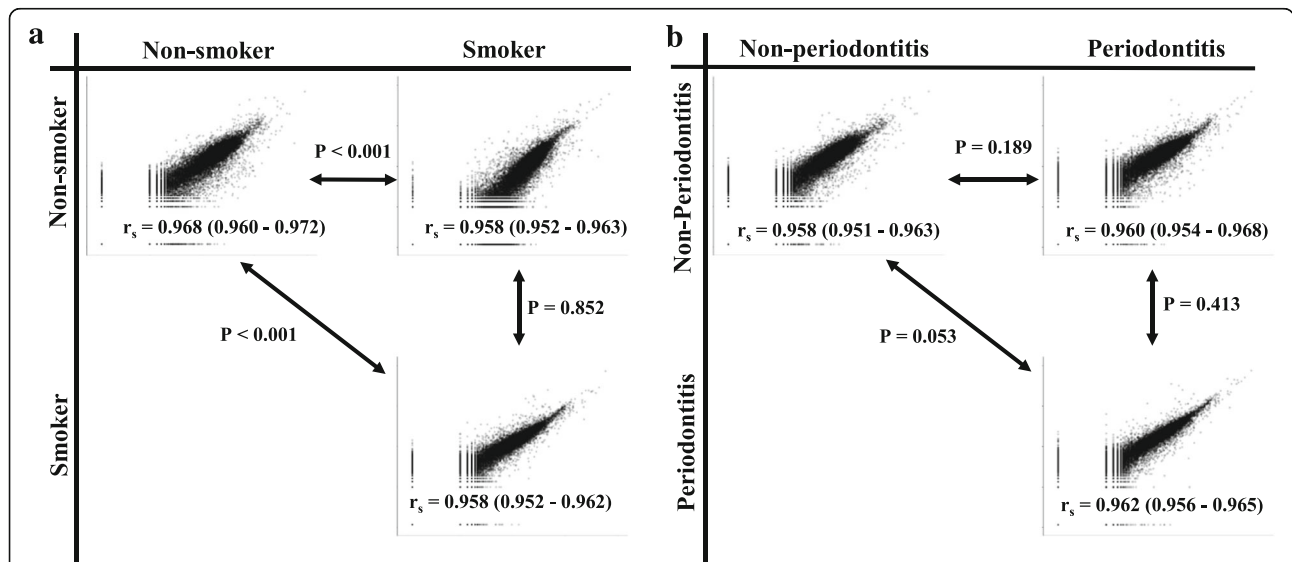


Fig. 1 Representative correlation plots for mRNA expression. **a** Non-smoker vs. smoker groups. **b** Non-periodontitis vs. periodontitis groups. The r_s values in the scatter plot are shown as the median and interquartile of Spearman's rank correlation coefficients between the cases of the groups. P values near the line with arrows at both ends are P values determined by the Student Wilcoxon rank sum test to compare the Spearman rank correlation coefficients between two groups shown in the x and y axes

in the SN group. In comparison, 2901 DE genes without controlling for the FDR were observed between NP and SP cases (Additional file 1: Table S2); among them, 1298 were upregulated and 1603 were downregulated in the SP group. In the NN and SN groups, statistically significant DE genes were enriched in the GO-BP categories of “skin development” (GO:0043588, $P = 1.23E^{-06}$, $q = 2.54 E^{-03}$; e.g., *Staphylococcus aureus* infection, basal cell carcinoma, or melanoma) (Fig. 2a), whereas in the NN and NP groups, DE genes were mostly enriched in the category “adaptive immune response” (GO:0002250, $P = 1.70E^{-05}$, $q = 4.48 E^{-02}$; e.g., natural killer cell-mediated cytotoxicity) (Fig. 2b). Notably, in smokers, the genes associated with “ECM

organization” (GO:0030198) were increased in periodontitis (SP compared to NP or SN, $P = 1.61E^{-16}$ and $P = 4.97E^{-07}$, $q = 2.95E^{-11}$ and $q = 9.33E^{-04}$, respectively; Fig. 2c, d). Collectively, these results show the differential gene expression pattern with the smoking effect in different periodontal conditions (Fig. 2e).

Methylation analysis and combined analysis of transcriptome and methylome data

The number of genes that were differently methylated (DM) in a 500-bp window of the transcription start site between NN and SN cases was 84 when FDR control was not used. Among the DM sites, 36 were

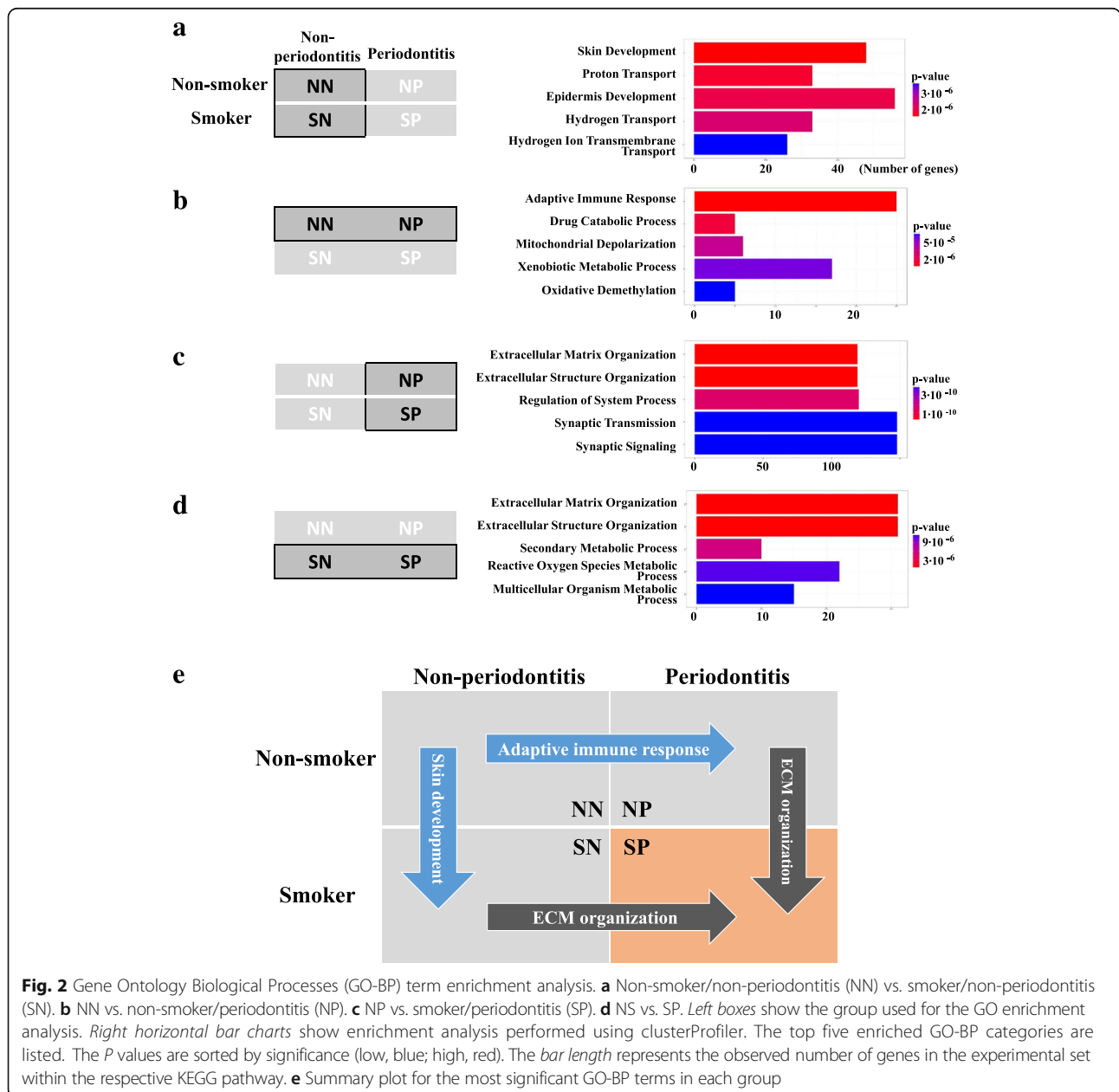


Fig. 2 Gene Ontology Biological Processes (GO-BP) term enrichment analysis. **a** Non-smoker/non-periodontitis (NN) vs. smoker/non-periodontitis (SN). **b** NN vs. non-smoker/periodontitis (NP). **c** NP vs. smoker/periodontitis (SP). **d** NS vs. SP. *Left boxes* show the group used for the GO enrichment analysis. *Right horizontal bar charts* show enrichment analysis performed using clusterProfiler. The top five enriched GO-BP categories are listed. The *P* values are sorted by significance (low, blue; high, red). The *bar length* represents the observed number of genes in the experimental set within the respective KEGG pathway. **e** Summary plot for the most significant GO-BP terms in each group

hyper-methylated and 48 were hypo-methylated in SN (Additional file 1: Table S3). The number of DM sites between NP and SP cases was 96 when no FDR control was conducted; among them, 30 were hyper-methylated and 66 were hypo-methylated in SP (Additional file 1: Table S4). The numbers of DE and DM genes were 7 and 16, respectively, without overlap (Additional file 1: Table S5). The regression model of methylation and expression of genes had a negative slope (-4.17) without a significant difference in the model with periodontitis as a response variable ($P = 0.684$). Among the DE and DM genes showing a negative slope, three genes (*DHRS7B*, *LIG1*, and *SPATA2L*) were identified between NN and SN cases and seven genes (*EGFL7*, *GPRC5C*, *IGF2*, *KLF9*, *PODN*, *RAD54L*, and *STARD9*) were identified between NP and SP cases (Fig. 3).

PPI analysis

PPI analysis performed for 17 genes exhibiting inverse correlation of DNA methylation with gene expression between non-smokers (NN and NP) and smokers (SN and SP) (Additional file 1: Table S6) identified 157 genes as potentially affected by PPI (Additional file 1: Table S7). Among them, 50 genes were differentially expressed and 107 genes showed no change between the non-smoker and smoker groups (Additional file 1: Table S7). Fisher’s exact test indicated that this ratio was statistically significant ($P = 0.040$). Among the GO-BP terms of PPI-affected genes (Fig. 4, yellow circle), “blood coagulation and fibrin clot formation” was the most significantly enriched ($GO:0007596$, $P = 1.76E^{-10}$, $q = 4.04E^{-07}$). The top 10 GO-BP terms of PPI-affected genes (Fig. 4, yellow circle) in common with DE genes for SN vs. SP (Fig. 4, pink circle) or NN vs. SP (Fig. 4, blue circle) were “ECM

organization” ($GO:0030198$, $P = 6.45E^{-08}$, $q = 2.57E^{-05}$) and “extracellular structure organization” ($GO:0043062$, $P = 6.73E^{-08}$, $q = 2.57E^{-05}$) (Fig. 4). The top 10 GO-BP terms of PPI-affected genes did not overlap with those of NN vs. NP.

Discussion

To the best of our knowledge, this is the first high-throughput study investigating both DNA methylation and gene expression to evaluate the effect of smoking on chronic periodontitis. Our results indicate that smoking is closely associated with ECM organization-related genes, which intensify the periodontal condition as indicated by DE and DM patterns, and PPI-affected genes. As smoking might contribute to periodontal disease susceptibility through microbial infection [26–28] and the development of malignancy [29], many variations of mRNA expression associated with a different DNA methylation status might be important in the context of biological and clinical aspects of this disease.

By correlation analysis, we revealed statistically significant differences between the groups depending on the smoking status, when non-smokers vs. non-smokers exhibited the highest Spearman’s rank correlation coefficient value (Fig. 1a). The difference in SRCC can be explained by changes in gene expression caused by smoking (e.g., years of smoking or number of cigarettes smoked per day) [30, 31], gender-related variations [32], or personal habits [33]. Different internal and external stress factors may trigger cellular responses which may result in the damage of macromolecular structures and functions, promoting a variety of pathological conditions [34].

Periodontitis is a chronic inflammatory disease leading to the destruction of periodontal tissues, including alveolar bone [35]. Previously, De Souza et al. [14] have reported variations in DNA methylation of immunity-related genes between healthy individuals and periodontitis patients [14]. Consistent with these results, our RNA-seq data indicated that the GO-BP category related to adaptive immune response are highly represented; however, the difference in the methylation status between healthy and periodontitis group was not significant (Fig. 2b). One of the important host factors in periodontitis is the activity of matrix metalloproteinases, which are responsible for collagen and ECM degradation of periodontal tissues [36]. Consistent with this notion, we revealed differential transcriptional regulation of ECM-related genes in the periodontitis group (Fig. 2) and changes in the DNA methylation pattern in the smoker group (Fig. 3). Notably, mRNA expression analysis indicated that the GO-BP term description of NP vs. SP (Fig. 2c) and SN vs. SP (Fig. 2d) identified the same term, ECM organization, indicating the expression, assembly, and arrangement of ECM

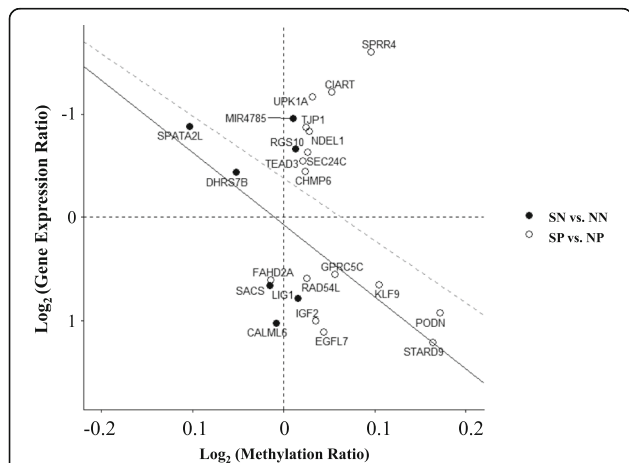
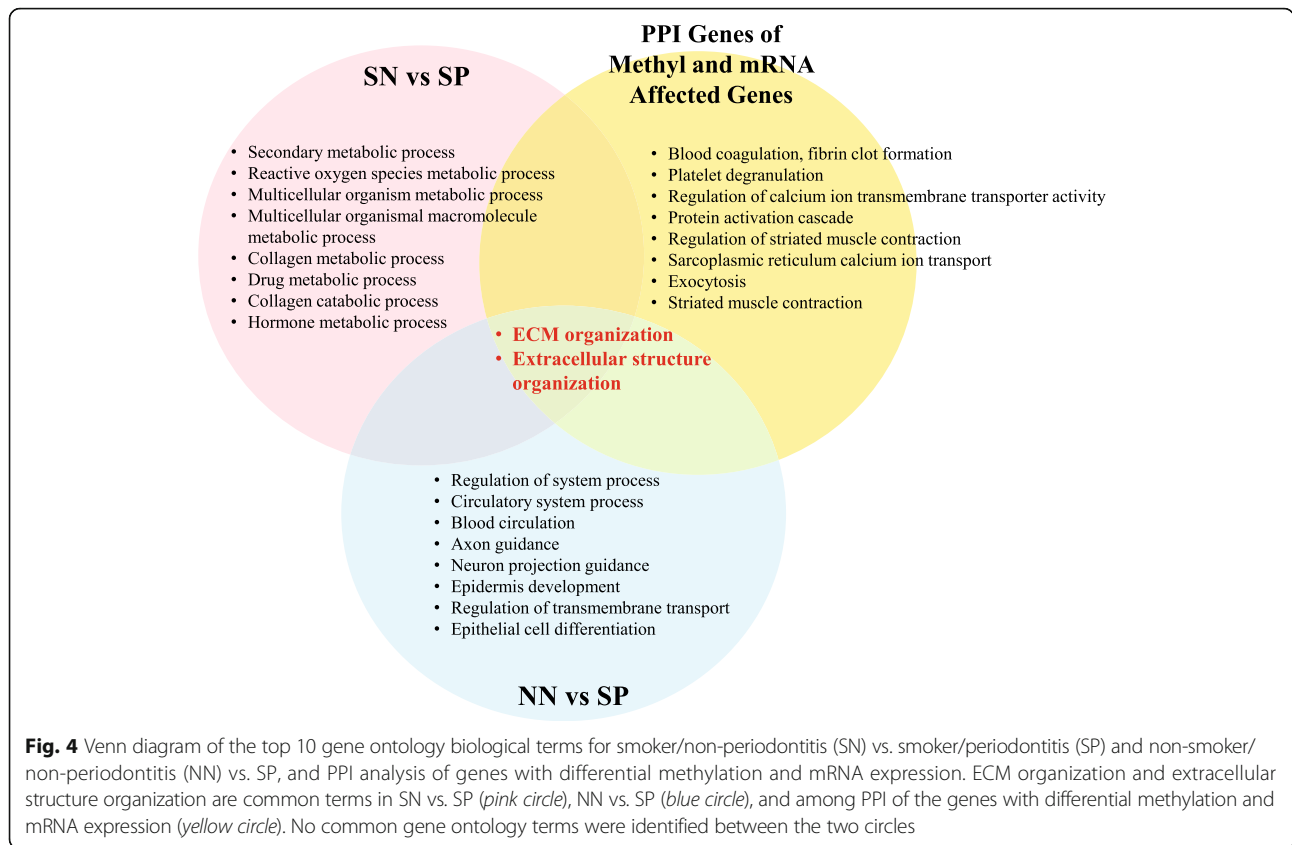


Fig. 3 Scatter plot of DNA methylation versus gene expression. The x and y axes present the log₂-scaled ratios of DNA methylation and gene expression, respectively. Solid and dashed lines indicate the regression of smoker/periodontitis (SP) vs. smoker/non-periodontitis (SN) and of SN vs. non-smoker/non-periodontitis (NN)



constituents according to the web application AmiGO [37]. Smoking may exert a variety of biological effects on the ECM [35, 38], which may account for the enrichment of this GO-BP term by smoking in periodontitis patients (from NP to SP) as well as by periodontitis in smokers (from SN to SP) (Fig. 2e). Methylation analysis indicated that smoking affected both DNA methylation and gene expression, which showed negative correlation (Fig. 3). Several novel biomarkers for smoking-associated blood DNA methylation have been identified using the Illumina 27K array [39, 40] and Illumina Infinium Human-Methylation 450K BeadChip array [41]. Among them, seven CpGs (F2RL3 [cg03636183], AHRR [cg21161138 and cg05575921], 2q37.1 [cg21566642, cg01940273, and cg05951221], and 6p21.33 [cg06126421]) were common among most DM sites [42]. However, we could not identify common biomarkers based on our data (Additional file 1: Table S3 and S4), which can be due to low sample size and the fact that we used gingival tissue. We assume that the changes in DNA methylation that affect gene expression were localized to positions near transcription start sites; however, methylation near transcription start sites could not explain all epigenetic aspects, suggesting the need for further studies recommended in the “Encyclopedia of DNA Elements (ENCODE)” project,

which aims to identify comprehensive functional elements in the human genome.

The GO terms ECM organization and extracellular structure organization were also found to be common among SN vs. SP and NN vs. SP groups and among genes that had PPI relationship with methylated DE genes (Fig. 4). However, among the GO terms identified through the combined analysis of transcriptome and methylome data for NN vs. SN or NP vs. SP, no common GO terms were obtained by comparing NN vs. NP or SN vs. SP. PPI extends the effect of the identified DE and DM genes to that of indirectly related genes through protein-protein cooperation. In this study, PPI broadened our perspective from single genes to more complex functional units (Fig. 4). The category showing the greatest PPI effect, blood coagulation and fibrin clot formation, is considered to be associated with wound healing in periodontal disease (Fig. 4, yellow circle) [43]. Although based on these results, we cannot conclude that DNA methylation caused by smoking directly affects the periodontitis status; the genes identified as related to those with increased DNA methylation may disturb ECM organization through PPI, which in turn would affect periodontitis in smokers.

Collectively, it is conceivable that both smoking-related DNA methylation and gene expression profiles reflect an increased sensitivity of periodontitis patients to environmental hazards such as smoking. Our results suggest that the smoking factor may contribute to weakening of periodontal tissue healing potential and accelerate tissue destruction through effects on DNA methylation and expression of ECM organization-related genes. Although DNA methylation caused by smoking may not directly affect periodontitis, the genes regulated by smoking-mediated epigenetic modifications can influence ECM organization, which in turn may have an impact on the disease characteristics.

Some genes were differentially expressed between groups; however, the conclusions may be premature because of study limitations such as small sample size and other unaccounted epigenetic factors (histone modification, microRNA, etc.). In this study, we could not identify specific target genes based on DNA methylation and mRNA expression patterns, as the results did not reach statistical significance. Instead, we detected the trend of gene expression and methylation shifts depending on smoking and periodontitis statuses by analyzing the correlation differences between groups, putative functions of DE/DM genes based on GO-BP terms, and protein-protein interactions. Our study provides a foundation for further studies using larger cohorts and incorporating a broader spectrum of epigenetic modifications, which are required to address potential mechanisms underlying the observed patterns and to discover diagnostic biomarkers for the susceptibility of patients with periodontitis to environmental hazards.

Conclusions

In summary, our data indicate that smoking may negatively affect the healing potential of periodontal tissue and accelerate its destruction through effects on DNA methylation and expression of ECM organization-related genes.

Additional file

Additional file 1: Table S1. Differentially expressed (DE) genes in SN vs. NN. **Table S2.** Differentially expressed (DE) genes in SP vs. NP. **Table S3.** Differentially methylated (DM) genes in SN vs. NN. **Table S4.** Differentially methylated (DM) genes in SP vs. NP. **Table S5.** Differentially expressed (DE) and methylated (DM) genes in smokers vs. non-smokers. **Table S6.** Differentially expressed (DE) and methylated (DM) genes. **Table S7.** Genes involved in protein-protein interaction (PPI). (XLSX 309 kb)

Abbreviations

BOP: Bleeding on probing; DE: Differentially expressed; DM: Differentially methylated; ECM: Extracellular matrix; FDR: False discovery rate; GO-BP: Gene Ontology Biological Processes; GSEA: Gene set enrichment analysis; PPD: Probing pocket depth; PPI: Protein-protein interaction; RNA-seq: RNA sequencing; RRBS: Reduced representation bisulfite sequencing; SRCC: Spearman's rank correlation coefficient

Acknowledgements

Not applicable

Funding

This research was supported by grants from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (HI14C2681), Basic Science Research Program (2016R1A1A3A04004838), Basic Research Lab Program (2017R1A4A1014584), and MSIP/IITP (2017-0-00398) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning.

Availability of data and materials

Data cannot be available because it is patient's health information.

Authors' contributions

YDC, PJK, and HMR conceived and designed the experiments. YDC and PJK performed the experiments and analyzed the data. HGK, YJS, YML, YK, ICR, and HMR contributed reagents/materials/analysis tools. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

This study was performed after the approval by the Institutional Review Board of Seoul National University Dental Hospital (CRI15003). All participants received an explanation of the study and signed an informed consent form.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Periodontology, School of Dentistry, Seoul National University, 101 Daehak-ro, Jongno-gu, Seoul 03080, South Korea.

²Department of Molecular Genetics, School of Dentistry, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea. ³Department of Dental Services Management and Informatics, School of Dentistry, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, South Korea.

Received: 17 April 2017 Accepted: 3 August 2017

Published online: 10 August 2017

References

- AlJehani YA. Risk factors of periodontal disease: review of the literature. *Int J Dent.* 2014;182513.
- Williams RC. Periodontal disease. *N Engl J Med.* 1990;322:373–82.
- Kubota M, Tanno-Nakanishi M, Yamada S, Okuda K, Ishihara K. Effect of smoking on subgingival microflora of patients with periodontitis in Japan. *BMC Oral Health.* 2011;11:1.
- Vouros ID, Kalpidis CD, Chadjipantelis T, Konstantinidis AB. Cigarette smoking associated with advanced periodontal destruction in a Greek sample population of patients with periodontal disease. *J Int Acad Periodontol.* 2009;11:250–7.
- Zini A, Sgan-Cohen HD, Marcenes W. Socio-economic position, smoking, and plaque: a pathway to severe chronic periodontitis. *J Clin Periodontol.* 2011;38:229–35.
- Shchipkova AY, Nagaraja HN, Kumar PS. Subgingival microbial profiles of smokers with periodontitis. *J Dent Res.* 2010;89:1247–53.
- Besingi W, Johansson A. Smoke-related DNA methylation changes in the etiology of human disease. *Hum Mol Genet.* 2014;23:2290–7.
- Ospelt C, Gay S. Epigenetics in inflammatory systemic diseases. *Internist (Berl).* 2014;55:124–7.
- Eichler EE, Flint J, Gibson G, Kong A, Leal SM, et al. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet.* 2010;11:446–50.

10. Larsson L, Castilho RM, Giannobile WV. Epigenetics and its role in periodontal diseases: a state-of-the-art review. *J Periodontol*. 2015;86:556–68.
11. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet*. 2008;9:465–76.
12. Oliveira NF, Damm GR, Andia DC, Salmon C, Nociti FH Jr, et al. DNA methylation status of the IL8 gene promoter in oral cells of smokers and non-smokers with chronic periodontitis. *J Clin Periodontol*. 2009;36:719–25.
13. Gomez RS, Dutra WO, Moreira PR. Epigenetics and periodontal disease: future perspectives. *Inflamm Res*. 2009;58:625–9.
14. De Souza AP, Planello AC, Marques MR, De Carvalho DD, Line SR. High-throughput DNA analysis shows the importance of methylation in the control of immune inflammatory gene transcription in chronic periodontitis. *Clin Epigenetics*. 2014;6:15.
15. Martins MD, Jiao Y, Larsson L, Almeida LO, Garaicoa-Pazmino C, et al. Epigenetic modifications of histones in periodontal disease. *J Dent Res*. 2016;95:215–22.
16. Hart TC, Kornman KS. Genetic factors in the pathogenesis of periodontitis. *Periodontol*. 1997;14:202–15.
17. Laine ML, Crielaard W, Loos BG. Genetic susceptibility to periodontitis. *Periodontol*. 2012;58:37–68.
18. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20.
19. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
20. Anders S, Pyl PT, Huber W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–9.
21. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
23. Guo W, Fiziev P, Yan W, Cokus S, Sun X, et al. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics*. 2013;14:774.
24. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284–7.
25. Razick S, Magklaras G, Donaldson IM. iRefIndex: a consolidated protein interaction database with provenance. *BMC Bioinformatics*. 2008;9:405.
26. Bergström J, Eliasson S. Noxious effect of cigarette smoking on periodontal health. *J Periodontol Res*. 1987;22:513–7.
27. Fiorini T, Musskopf ML, Oppermann RV, Susin C. Is there a positive effect of smoking cessation on periodontal health? A systematic review. *J Periodontol*. 2014;85:83–91.
28. Khan S. Effect of smoking on periodontal health. *Dis Mon*. 2011;57:214–7.
29. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*. 1999;91:1194–210.
30. Spitz MR, Shi H, Yang F, Hudmon KS, Jiang H, et al. Case-control study of the D2 dopamine receptor gene and smoking status in lung cancer patients. *J Natl Cancer Inst*. 1998;90:358–63.
31. Noble EP. The DRD2 gene, smoking, and lung cancer. *J Natl Cancer Inst*. 1998;90:343–5.
32. Paul S, Amundson SA. Differential effect of active smoking on gene expression in male and female smokers. *J Carcinog Mutagen*. 2014;5:1000198.
33. Davies GE, Soundy TJ. The genetics of smoking and nicotine addiction. *S D Med*. 2009;Spec No:43–9.
34. Poljšak B, Milisav I. Clinical implications of cellular stress responses. *Bosn J Basic Med Sci*. 2012;12:122–6.
35. Sapna G, Gokul S, Bagri-Manjrekar K. Matrix metalloproteinases and periodontal diseases. *Oral Dis*. 2014;20:538–50.
36. Birkedal-Hansen H. Role of matrix metalloproteinases in human periodontal diseases. *J Periodontol*. 1993;64:474–84.
37. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, et al. AmiGO: online access to ontology and annotation data. *Bioinformatics*. 2009;25:288–9.
38. Tipton DA, Dabbous MK. Effects of nicotine on proliferation and extracellular matrix production of human gingival fibroblasts in vitro. *J Periodontol*. 1995;66:1056–64.
39. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet*. 2011;88:450–7.
40. Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, et al. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet*. 2012;21:3073–82.
41. Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*. 2013;8:e63812.
42. Ambatipudi S, Cuenin C, Hernandez-Vargas H, Ghantous A, Le Calvez-Kelm F, et al. Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. *Epigenomics*. 2016;8:599–618.
43. Polimeni G, Xiropaidis AV, Wikesjö UM. Biology and principles of periodontal wound healing/regeneration. *Periodontol*. 2006;41:30–47.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

