Research paper

# RNA-sequencing study of peripheral blood monocytes in chronic periodontitis ☆

Yao-Zhong Liu [a],*, Pooja Maney [b],**, Jyoti Puri [b], Yu Zhou [a], Melody Baddoo [c], Michael Strong [c], Yu-Ping Wang [d], Erik Flemington [c], Hong-Wen Deng [a]

[a] Center of Genomics and Bioinformatics, Dept. of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, United States
[b] Dept. of Periodontics, School of Dentistry, Louisiana State University Health Sciences Center, New Orleans, LA, United States
[c] Tulane Cancer Center, Tulane University School of Medicine, New Orleans, LA, United States
[d] Dept. of Biomedical Engineering, Tulane University School of Science and Engineering, United States

## ARTICLE INFO

## ABSTRACT

*Background:* Monocytes are an important cell type in chronic periodontitis (CP) by interacting with oral bacteria and mediating host immune response. The aim of this study was to reveal new functional genes and pathways for CP at monocyte transcriptomic level.
*Methods:* We performed an RNA-sequencing (RNA-seq) study of peripheral blood monocytes (PBMs) in 5 non-smoking moderate to severe CP (case) individuals vs. 5 controls. We took advantage of a microarray study of periodontitis to support our findings. We also performed pathway-based analysis on the identified differentially expressed (DEx) transcripts/isoforms using DAVID (Database for Annotation, Visualization and Integrated Discovery).
*Results:* Through differential expression analyses at both whole gene (or whole non-coding RNA) and isoform levels, we identified 380 DEx transcripts and 5955 DEx isoforms with a PPEE (posterior probability of equal expression) of <0.05. Pervasive up-regulation of transcripts at isoform level in CP vs. control individuals was observed, suggesting a more functionally active monocyte transcriptome for CP. By comparing with the microarray dataset, we identified several CP-associated novel genes (e.g., FACR and CUX1) that have functions to interact with invading microorganisms or enhance TNF production on lipopolysaccharide stimulation. DAVID analysis of both the RNA-seq and the microarray datasets leads to converging evidence supporting "endocytosis", "cytokine production" and "apoptosis" as significant biological processes in CP.
*Conclusions:* As the first RNA-seq study of PBMs for CP, this study provided novel findings at both gene (e.g., FCAR and CUX1) and biological process level. The findings will contribute to better understanding of CP disease mechanisms.

© 2016 Elsevier B.V. All rights reserved.

## 1. Background

Chronic periodontitis (CP) (Armitage, 1999a) is an inflammatory disease that affects the supporting structures of teeth. It is one of the most prevalent chronic diseases and a major cause of loss of permanent teeth (Eke et al., 2012), which severely affects the quality of life in old people.

While oral bacteria are necessary for initiating pathological cascades of CP (Ezzo and Cutler, 2003), it is host immune response and inflammatory process that is responsible for periodontal tissue damage (Taubman et al., 2005), where monocytes play a central role. Monocytes, by differentiating into dendritic cells (DCs) in periodontium (Pulendran et al., 2001), trigger T cell differentiation into type 1, type 2 or type 17 T-helper (Th) cells, which in turn direct cellular, humoral and innate immune responses, respectively (Mosmann and Coffman, 1989; Korn et al., 2009). Monocytes, by interacting with oral bacteria, initiate inflammatory processes in periodontium and produce a series of inflammatory factors, such as PGE2, IL-1-β, TNF-α and MMP (Offenbacher et al., 1993).

Although the important pathological cascades mediated by monocytes mainly occur in periodontium, peripheral blood monocytes

(PBMs), as functional continuum and the sole source of the monocytes in the periodontium, play a critical role in CP pathogenesis. It was found that in CP patients, PBMs had a "specific functional profile" that favors the Th2-cell response over Th1-cell response (Sigusch et al., 1998; Petit et al., 1999; Tew et al., 1983; Boyatzis and Seymour, 1986). Therefore, functional aberrance of PBMs may predispose an individual to a destructive host immune response leading to CP.

While the immune response and inflammatory processes are well known as the key mechanisms for CP, the detailed functional genes and pathways underlying the disease at the PBMs transcriptome level are still unclear. Despite two recent genome-wide gene expression studies of CP (Davanian et al., 2012; Kebschull et al., 2014) the studies were performed on the gingival tissue that may have a different transcriptomic landscape from PBMs, and hence the findings from those studies may not be relevant to functional genes contributing to CP at the PBMs level. Another study (Papapanou et al., 2007) was conducted on PBMs on CP patient using the microarray technology. However, the interesting findings from the study may need further confirmation from an independent study, such as the one presented here.

Here, to shed light on those important genes and pathways for CP, we used the latest next generation sequencing technology and performed the first RNA-sequencing (RNA-seq) analysis of PBMs for CP. Our findings contributed to further understanding of CP pathogenesis by having identified a number of novel yet functionally relevant genes and biological processes as specific key players for the disease.

## 2. Methods

### 2.1. Study participant characteristics

The study was approved by Institutional Review Board of Louisiana State University Health Science Center (LSUHSC). Signed informed-consent documents were obtained from all study participants before entering the study.

The study participants (cases and controls) were recruited at LSUHSC School of Dentistry (LSUHSC SOD) periodontics clinic. All case individuals (n = 5) were non-smoking African American (AA) males diagnosed with generalized moderate to severe CP as per the 1999 world workshop classification (Armitage, 1999b).

Specifically, cases with >30% sites with clinical attachment loss (CAL) of ≥3–4 mm were diagnosed as generalized "moderate" chronic periodontitis and cases with >30% sites with CAL ≥5 mm were diagnosed as generalized "severe" chronic periodontitis.

Control individuals (n = 5) were periodontally healthy based on clinical and radiographic examinations. None of the controls exhibited CAL ≥ 3 mm, probing depths ≥4 mm or radiographic evidence of bone loss. All control individuals were also non-smokers and gender and ethnically matched to the cases.

The average age of case individuals was 44.6 years with a standard deviation of 7.73 years. The average age of control individuals was 36.4 years with a standard deviation of 12.97 years. There was no statistical difference in age between case and control individuals ($p = 0.26$) according to the t test.

Exclusion criteria for all participants were as follows: 1) Recent treatment for periodontal disease (within a year); 2) current and/or long-term use of NSAIDs (non-steroidal anti-inflammatory drugs) which could affect the outcome of the study; 3) former or current smokers, since smoking is one of the most influential factors in periodontitis; 4) diabetes mellitus; 5) treatment with corticosteroid therapy currently or for more than 6 months duration at any time; 6) treatment with anticonvulsant therapy or anti-depressants currently or for more than 6 months duration at any time; 7) treatment with antibiotics currently or for more than 6 months duration at any time; 8) evidence of other metabolic or inherited bone diseases such as hyper or hypoparathyroidism, Paget's disease, osteomalacia, osteogenesis imperfect or

others; 9) hyperthyroidism; and 10) history of or current diagnosis with aggressive periodontitis.

### 2.2. Monocyte isolation

Monocyte isolation from whole blood was performed with a monocyte-negative isolation kit (Miltenyi Biotec Inc., Auburn, CA) following the manufacturer's recommendation. The kit is an indirect magnetic labeling system for isolating untouched monocytes from human peripheral blood mononuclear cells (PBMCs). Non-monocytes, i.e. T cells, NKcells, B cells, dendritic cells and basophils, are indirectly magnetically labeled using a cocktail of biotin-conjugated antibodies against CD3, CD7, CD16, CD19, CD56, CD123 and glycophorin A, and anti-biotin MicroBeads. Isolation of highly pure unlabeled monocytes is achieved by depletion of the magnetically labeled cells. Based on our extensive experience of monocyte isolation using this kit (Deng et al., 2011; Shen et al., 2013), the isolated monocytes can achieve a purity of at least 85%.

### 2.3. Total RNA extraction

Total RNA from monocytes was extracted using Qiagen RNeasy Mini kit (Qiagen, Inc., Valencia, CA). We used Agilent Bioanalyzer (Agilent, Santa Clara, CA) to control the RNA quality before RNA-seq experiment, where RNA integrity number (RIN) was confirmed at no less than 7.0 before RNA-seq experiment.

### 2.4. RNA-seq

RNA sequencing was performed at JP Sulzberger Columbia Genome Center. We used poly-A pull-down to enrich mRNAs from total RNA samples (200 ng–1 μg per sample) and proceeded on library preparation by using Illumina TruSeq RNA prep kit (Illumina Inc., San Diego, CA).

Specifically, we purified the polyA containing mRNA molecules using oligo-dT attached magnetic beads with two rounds of purification. During the second elution of the polyA RNA, the RNA was fragmented and primed for cDNA synthesis. We then performed reverse transcription on the RNA fragments to synthesize first strand cDNA using reverse transcriptase and random primers. The RNA template was then removed and a 2nd strand cDNA was synthesized. End repair was then performed to convert the overhangs into blunt ends using the End Repair Mix. After 3′ ends adenylation, multiple indexing adapters were ligated to the ends of the ds cDNA and those DNA fragments with adapter molecules were enriched with PCR reactions that constructed a library for downstream sequencing experiments. The library was validated using Agilent Bioanalyzer (Agilent, Santa Clara, CA) to ensure the size and purity of the sample by observing a band at approximately 260 bp.

Libraries were then sequenced using an Illumina HiSeq 2500 instrument (Illumina Inc., San Diego, CA). We multiplexed samples in each lane, which yielded targeted number of single-end 100 bp reads for each sample, as a fraction of 180 million reads for the whole lane.

We used the RTA software (Illumina Inc., San Diego, CA) for base calling and CASAVA software (version 1.8.2) (Illumina Inc., San Diego, CA) for converting the BCL files to fastq format, coupled with adaptor trimming. The read depth for each sample was 30 million reads.

The raw RNA-seq data were deposited to GEO with an accession number of GSE61490.

### 2.5. Differential expression analysis of RNA-seq data

We used RSEM v1.2.11 (Li and Dewey, 2011) to analyze the RNA-seq data for alignment and expression calculation. Under RSEM, we used EBSeq (Leng et al., 2013) to identify differentially expressed (DEx) transcripts at whole gene (or whole non-coding RNA) level or at isoform level.

EBSeq assumes the expected count ($X$) for a gene or an isoform is distributed as Negative Binomial (NB), i.e., $X \sim NB(r; p)$, and determines the posterior probability of differential expression of a gene or an isoform through the Bayes' rule. To achieve that, EBSeq models a prior distribution of the data using $\beta$ distribution. Under the null hypothesis (EE: equal expression), the data $X_{g_i}^{C1,C2} = X_{g_i}^{C1}, X_{g_i}^{C2}$ (where $g_i$ denote a gene or isoform $i$, $C1$ condition 1 and $C2$ condition 2) can be represented by a prior predictive distribution $f_0^{I_g}(X_{g_i}^{C1,C2})$ (which is constructed based on beta functions). Under the alternative (DE: differential expression), $X_{g_i}^{C1,C2}$ follows another prior predictive distribution $f_1^{I_g}(X_{g_i}^{C1,C2})$. Given a latent variable $Z_{g_i}$, where $Z_{g_i} = 1$ represents that the gene/isoform $g_i$ is DE and $Z_{g_i} = 0$ represents that $g_i$ is EE, and also $Z_{g_i} \sim Bernoulli(p)$. Based on that, the marginal distribution of $X_{g_i}^{C1,C2}$ and $Z_{g_i}$ is $(1-p)f_0^{I_g}(X_{g_i}^{C1,C2}) + pf_1^{I_g}(X_{g_i}^{C1,C2})$.

Therefore, the posterior probability of DE for $g_i$ is defined by Bayes' rule:

$$\frac{pf_1^{I_g}\left(X_{g_i}^{C1,C2}\right)}{(1-p)f_0^{I_g}\left(X_{g_i}^{C1,C2}\right) + pf_1^{I_g}\left(X_{g_i}^{C1,C2}\right)}.$$

For more details, please refer to the supplement to the paper on EBSeq (Leng et al., 2013).

### 2.6. An online microarray dataset to provide supportive evidence

We took advantage of an online microarray dataset from GEO (GEO accession #: GSE6751) (Papapanou et al., 2007) as an independent dataset to support the findings from our RNA-seq data. The dataset contains 15 CP patients, whose PBM expression was analyzed 1 week before periodontal treatment (time point #1), at treatment initiation (time point #2), 6 weeks after the treatment (time point #3) and 10 weeks after the treatment (time point #4) (Papapanou et al., 2007). At 10 weeks after treatment, considerable improvement in periodontal status, including bleeding sites per individual (decreased from an average of 81% before treatment to an average of 25% after treatment) and deep pockets per individual (decreased from an average of 69.2 before treatment to an average of 9.2 after treatment), was achieved (Papapanou et al., 2007). The treatment involved only periodontal surgery and tooth extractions, if necessary, without local or systemic antibiotics.

We took advantage of the 15 individuals' PBM samples before treatment and the samples 10 weeks after treatment as a "paired" cohort with difference in periodontal status. Given the considerable improvement of periodontal status 10 weeks after treatment vs. before treatment, we treat the individuals before treatment as "periodontally unhealthy status" and the individuals 10 weeks after treatment as "periodontally healthy status". Comparing the two statuses in gene expression may shed lights on those genes differentially expressed in terms of "difference in periodontal health". Such differentially expressed genes may provide supporting evidence for those differentially expressed genes as identified in our RNA-seq data, where two groups of individuals are also different in terms of periodontal health.

The PBM gene expression was assayed using Affymetrix U133 Plus 2.0 arrays. We analyzed the raw cel files for each individual before the periodontal treatment (time point #1) and 10 weeks after the treatment (time point #4). The cel files were normalized using RMA (Irizarry et al., 2003). We then performed differential expression analysis using paired $t$ test through the Bioconductor's LIMMA (linear models for microarray data) package (Smyth, 2005) to compare the PBM gene expression before vs. after periodontal treatment. Genes and their differential expression $p$ values as assessed above were used as the basis for supporting those DEx transcripts identified in our RNA-seq data.

### 2.7. Enrichment analysis using DAVID

GO (gene ontology) and other pathway/functional term enrichment analysis was performed on a number of group of genes (e.g., DEx transcripts or isoforms identified in CP vs. control individuals) using DAVID (Database for Annotation, Visualization and Integrated Discovery) v6.7 software package (http://david.abcc.ncifcrf.gov/) (dW et al., 2009).

## 3. Results

### 3.1. Quality control (QC) analysis

The mean number of reads over the 10 samples is 26,549,771, with a range from 21,697,093 reads to 31,754,419 reads. The average mapping rate (ratio of number of uniquely mapped reads over total number of reads) across the 10 samples is 75.93%, with a range from 69% to 84%.

The sequence QC results on base quality are quite excellent. According to the FastQC software (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/), all of the 10 samples pass the QC (with a green light tick) for both the "per base sequence quality" and "per sequence quality scores", two important metrics for base quality QC. Specifically, the average quality score per read is either 37 or 38 for all the samples, which is very good as the maximum quality score is 40.

Our sequence read depth turns out to be sufficient for quantifying lowly expressed transcripts as we have detected a large number of transcripts with a low count. As shown in our data, we have observed >4400 transcripts with a count of 10–50 for at least one sample. (To preclude false expression signals, these transcripts do not include those whose counts are either 0 or lower than 10 in any sample.) Among these, there are >600 transcripts with an average count of <50 across the 10 samples, representing transcripts with an overall low expression for all the 10 samples. Note that although the level of expression is very low these >600 transcripts are consistently detected in all the samples. Hence it is unlikely that these detected transcripts are false signals (e.g., due to mistaken alignment).

### 3.2. RNA-seq analysis at the whole gene level

At the whole gene (or whole non-coding RNA) level, we identified a total of 380 DEx transcripts in CP vs. control individuals, which achieved a PPEE (posterior probability of equal expression) of <0.05 and a PPDE (posterior probability of differential expression) of >0.95. Among these transcripts, 228 were up-regulated and 152 were down-regulated in CP vs. control individuals.

As a large number of DEx transcripts achieved a PPEE of 0 or near 0 (or PPDE of 1 or near 1), to rank the DEx transcripts, we used the measure of PostFC (posterior fold change). Listed in Table 1 are the top 10 up-regulated transcripts with the largest PostFC in CP vs. control individuals and the top 10 down-regulated transcripts with the smallest PostFC in CP vs. controls.

Among the transcripts with the most extreme PostFC in CP vs. control individuals (Table 1), representing those transcripts that may be strongly regulated in terms of CP status, LST1 was previously found to be up-regulated in response to LPS and bacterial infection and in blood of rheumatoid arthritis (RA) patients as compared to controls (Mulcahy et al., 2006). This gene's known function is consistent with the context of CP pathogenesis, where LPS and bacterial infection initiates the pathologic cascade of the disease. Another up-regulated gene, LILRA5 was found to be expressed by synovial tissue macrophages in RA patients and induce pro-inflammatory cytokines (Mitchell et al., 2008). RNF5, down-regulated in CP patients vs. controls, was found to influence susceptibility to bacterial infection (Kuang et al., 2012). Another down-regulated transcript, IER3, was reported in a study, where deficiency of IER3 expression in mice may result in aberrant immune regulation and enhanced inflammation (Arlt and Schafer, 2011).

**Table 1**
The DEx transcripts identified at the whole gene level with the most extreme fold changes.

| Ensembl ID | Transcript name | PPEE[1] | PPDE[2] | PostFC[3] |
|---|---|---|---|---|
| *Up-regulated transcripts* | | | | |
| ENSG00000235040 | MTCO3P1 | 0 | 1 | 323 |
| ENSG00000213029 | SPHAR | 0 | 1 | 306 |
| ENSG00000262610 | LILRA5 | 0 | 1 | 266 |
| ENSG00000266198 | PRKAB2 | 1.3E − 9 | 0.99 | 238 |
| ENSG00000111971 | LY6G5C | 0 | 1 | 189 |
| ENSG00000234514 | LST1 | 2.78E − 8 | 0.99 | 157 |
| ENSG00000206549 | PRSS50 | 0 | 1 | 133 |
| ENSG00000263001 | GTF2I | 0 | 1 | 133 |
| ENSG00000198406 | BZW1P2 | 0 | 1 | 120 |
| ENSG00000253356 | RP11-90P5.2 | 0 | 1 | 112 |
| *Down-regulated transcripts* | | | | |
| ENSG00000204619 | PPP1R11 | 0 | 1 | 0.0018 |
| ENSG00000228405 | RNF5 | 0 | 1 | 0.0019 |
| ENSG00000230230 | TRIM26 | 0 | 1 | 0.0029 |
| ENSG00000254133 | RP11-1198D22.3 | 0 | 1 | 0.0033 |
| ENSG00000132692 | BCAN | 0 | 1 | 0.0042 |
| ENSG00000237051 | GABBR1 | 0 | 1 | 0.0065 |
| ENSG00000206478 | IER3 | 0 | 1 | 0.0089 |
| ENSG00000224859 | ZNRD1 | 0 | 1 | 0.0104 |
| ENSG00000206285 | B3GALT4 | 0 | 1 | 0.0107 |
| ENSG00000228987 | HLA-DRA | 1.3E − 07 | 0.99 | 0.0113 |

Note:
1. PPEE: The posterior probability (estimated by EBSeq) that a transcript is equally expressed.
2. PPDE: The posterior probability (estimated by EBSeq) that a transcript is differentially expressed
3. PostFC: The posterior fold change (cases over controls) for a transcript, which is defined as the ratio of posterior mean expression estimate of a transcript in CP over that in control individuals.

As another interesting finding, almost half of the transcripts (n = 8) presented in Table 1 are located in the MHC (major histocompatibility complex) regions, which include 2 up-regulated transcripts, LST1 (Rollinger-Holzinger et al., 2000) and LY6G5C (Mallya et al., 2006), and 6 down-regulated transcripts, HLA-DRA (Lee et al., 1982), ZNRD1 (Fan et al., 2000), GABBR1 (Hsu et al., 2012), RNF5 (de Jong et al., 2012), TRIM26 (de Jong et al., 2012), and PPP1R11 (Amadou et al., 1995). The co-localization to the MHC regions of these strongly regulated transcripts in terms of CP status once again suggests immune aberration as an important component of CP pathogenesis.

### 3.3. Pathway-based analysis using DAVID for those transcripts identified at the whole gene level

Using DAVID, we analyzed the 380 DEx transcripts that achieved a PPEE of <0.05. The top ten enriched pathways/functional terms are listed in Table 2. These terms achieved significant enrichment p values even after adjusted with the Bonferroni method.

Among these ten terms, two are related to immune response (GO:0006955 and SP_PIR_keyword immune response), three are related to immunoglobulin/major histocompatibility complex (MHC)

(IPR003006, IPR003597 and GO:0042611) and one is related to antigen processing and presentation (GO:0019822).

### 3.4. Differential expression analysis of the microarray dataset

Using Bioconductor's LIMMA package (GK, 2014), we identified 2354 probe sets, corresponding to 2092 annotated genes, which were DEx in periodontally unhealthy (one week before treatment) vs. periodontally healthy statuses (10 weeks after periodontal treatment), among which 1310 genes were found to be up-regulated and 782 genes down-regulated before vs. after treatment at a p value of <0.05.

### 3.5. Supporting evidence from the microarray data on those transcripts identified at the whole gene level

At the whole gene level, we identified in the RNA-seq data 380 DEx transcripts, which achieved a PPEE (posterior probability of equal expression) of <0.05 and a PPDE (posterior probability of differential expression) of >0.95. We selected a subset with a higher significance, which were 210 DEx transcripts that achieved a PPDE = 1 (and a PPEE <4.70E − 10), for finding supporting evidence from the microarray dataset. Those transcripts in the list of 210 DEx transcripts were checked for differential expression signals in the microarray data. Through this analysis, we identified 11 genes that achieved a PPEE of 1.78E − 15 or lower in the RNA-seq dataset and a p value of <0.05 in the microarray dataset. The 11 genes also had the same direction of regulation across the two datasets. The direction of regulation in the RNA-seq dataset is defined based on the ratio of posterior mean expression in CP patients over controls, with a ratio of >1 as up-regulation and a ratio of <1 as down-regulation. The direction of regulation in the microarray dataset is defined based on the ratio of mean expression in the periodontally unhealthy status (before treatment) over periodontally healthy status (after treatment), with a ratio of >1 as up-regulation and a ratio of <1 as down-regulation. Table 3 shows the information of the 11 genes supported by the microarray dataset.

Among the 11 DEx transcripts with supporting evidence from the microarray data (Table 3), FCAR achieved the most significant p value (7.78E − 4) in the microarray study (Papapanou et al., 2007). In addition, multiple probe sets (207674_at, 211307_s_at, 211816_x_at) from this gene were associated with periodontal health status in the microarray study (Papapanou et al., 2007). According to OMIM description (entry #: 147045), FCAR "interacts with aggregated IgAs, such as IgA coated on the surface of an invading microorganism, and mediates several immunologic defense processes such as phagocytosis, antibody-dependent cell-mediated cytotoxicity, and stimulation of the release of inflammatory mediators" (Herr et al., 2003). This gene's expression in monocyte or mononuclear cells has been associated with several autoimmune diseases, such as autoimmune thyroid disease (Heul-Nieuwenhuijsen et al., 2010), multiple sclerosis (Achiron et al., 2004) and ankylosing spondylitis (Montenegro et al., 2000). Given FCAR's roles in immunologic defense of microorganism and the gene's

**Table 2**
Top 10 enriched functional terms identified by DAVID for the DEx transcripts identified at the whole gene level.

| Category | Term | Enrichment p value | Bonferroni p value | Fold enrichment |
|---|---|---|---|---|
| INTERPRO | IPR003006: Immunoglobulin/major histocompatibility complex, conserved site | 9.61E − 10 | 4.85E − 07 | 11.74 |
| UP_SEQ_FEATURE | Domain: Ig-like C1-type | 7.97E − 10 | 7.05E − 07 | 20.96 |
| SMART | SM00407: IGC1 | 1.84E − 08 | 2.06E − 06 | 12.04 |
| SP_PIR_KEYWORDS | Signal | 1.14E − 08 | 4.05E − 06 | 1.92 |
| INTERPRO | IPR003597: Immunoglobulin C1-set | 1.32E − 08 | 6.65E − 06 | 12.71 |
| UP_SEQ_FEATURE | Signal peptide | 1.50E − 08 | 1.33E − 05 | 1.91 |
| GOTERM_BP_FAT | GO:0019882~antigen processing and presentation | 1.01E − 08 | 1.44E − 05 | 10.99 |
| GOTERM_CC_FAT | GO:0042611~MHC protein complex | 7.39E − 08 | 1.79E − 05 | 12.67 |
| SP_PIR_KEYWORDS | Immune response | 7.27E − 08 | 2.58E − 05 | 6.03 |
| GOTERM_BP_FAT | GO:0006955~immune response | 8.41E − 08 | 1.21E − 04 | 3.19 |

**Table 3**
The DEx transcripts identified at the whole gene level and supported by the PBM microarray dataset.

| Discovery dataset (the RNA-seq dataset) | | | | Supportive dataset (the microarray dataset) | | |
|---|---|---|---|---|---|---|
| Transcript name | PPEE | PPDE | Direction of regulation[1] | Probe set | p value | Direction of regulation[2] |
| FCAR | 1.78E − 15 | 1 | Up-regulation | 207674_at | 7.78E − 04 | Up-regulation |
| | | | | 211307_s_at | 3.60E − 02 | Up-regulation |
| | | | | 211816_x_at | 4.88E − 02 | Up-regulation |
| CLCN5 | 0 | 1 | Up-regulation | 232127_at | 9.82E − 03 | Up-regulation |
| CUX1 | 0 | 1 | Up-regulation | 202367_at | 2.14E − 02 | Up-regulation |
| | | | | 227069_at | 2.19E − 02 | Up-regulation |
| | | | | 214743_at | 2.98E − 02 | Up-regulation |
| | | | | 225935_at | 4.12E − 02 | Up-regulation |
| RNASE3 | 0 | 1 | Up-regulation | 206851_at | 1.29E − 02 | Up-regulation |
| REL | 7.77E − 16 | 1 | Up-regulation | 206036_s_at | 1.75E − 02 | Up-regulation |
| PXN-AS1 | 0 | 1 | Down-regulation | 227332_at | 1.95E − 02 | Down-regulation |
| SGMS2 | 0 | 1 | Up-regulation | 242963_at | 2.02E − 02 | Up-regulation |
| GGT1 | 0 | 1 | Up-regulation | 215603_x_at | 2.20E − 02 | Up-regulation |
| HLA-DOA | 0 | 1 | Up-regulation | 217001_x_at | 2.94E − 02 | Up-regulation |
| VNN2 | 0 | 1 | Up-regulation | 205922_at | 3.01E − 02 | Up-regulation |
| ME1 | 0 | 1 | Up-regulation | 204059_s_at | 3.06E − 02 | Up-regulation |

Note:
1. Direction of regulation in the RNA-seq dataset is defined based on the ratio of posterior mean expression in CP patients over controls, with a ratio of >1 as up-regulation and a ratio of <1 as down-regulation.
2. Direction of regulation in the microarray dataset is defined based on the ratio of mean expression 1 week before over 10 weeks after periodontal treatment (i.e., periodontally unhealthy over periodontally healthy statuses), with a ratio of >1 as up-regulation and a ratio of <1 as down-regulation.

importance in autoimmune diseases, FCAR may also play a key role in CP pathogenesis.

In addition to FCAR, other 6 DEx genes supported by the microarray dataset (Table 3) may also be involved in CP pathogenesis as they have a close link with immune response or immune-related diseases. CUX1 was found to enhance TNF production on LPS stimulation (Ikeda et al., 2013). RNASE3 is a mediator in host immune response to parasites, bacteria and viruses and may also cause side-effects on the host's own tissues (Topic and Dodig, 2011). CLCN5 was found to play a role in immunopathogenesis of ulcerative colitis (Alex et al., 2010). REL is a transcription factor that regulates innate and adoptive immunity. In particular, REL is essential for IL-12 and IL-23 transcription by macrophages and dendritic cells, which is crucial for T-cell differentiation and effector functions (Visekruna et al., 2012). VNN2's up-regulation in mononuclear cells was associated with both inflammatory bowel disease and RA (Bovin et al., 2007). HLA-DOA, as a MHC class II gene, has been associated with asthma (Yucesoy et al., 2014), RA (Reynolds et al., 2010) and type I diabetes (Santin et al., 2009).

### 3.6. Supporting evidence at the pathway level from the microarray data for those transcripts identified at the whole gene level

Using DAVID, we analyzed the 380 DEx transcripts (PPEE <0.05) identified at the whole gene level from the RNA-seq data. We did the same analysis using DAVID on those 2092 DEx genes ($p < 0.05$) as detected in the microarray dataset. The identified pathways and functional terms that achieved a Bonferroni enrichment $p$ value of <0.05 in the RNA-seq data were checked for enrichment signals in the DAVID analytical results of the microarray data. Through this analysis, we identified one functional term "SP_PIR_Keyword: immune response", which achieved an enrichment $p$ value of 7.27E − 8 (Bonferroni $p$ value = 2.58E − 5) in the RNA-seq data and an enrichment $p$ value of 2.24E − 2 in the microarray data.

In the RNA-seq data, the component genes that fall into this term include HLA-DRA, HLA-H, HLA-DMB, HLA-DPB1, MICB, HLA-DOB, HLA-DOA, C1QB, HLA-F, TMED7, LILRB4, TLR9, TAP2, HLA-DQA2, LST1 and LILRA5. For the microarray data, the component genes that fall into this term include CD7, IL23A, LIME1, DCLRE1C, POLR3G, TLR3, POLR3B, IL2, HLA-DOA, NLRX1, POLR3E, IFIH1, TLR10, HFE, CD300LB, INPPL1, CNPY3, TLR7, MYD88, TLR6, TLR5, LILRA1, TLR4, CR1, HLA-DRB1, CLU, CLEC7A, TLR8, CLEC5A, LY86, CLEC4E, CLEC4A and TLR2.

### 3.7. Differential expression analysis at the isoform level using the RNA-seq data

We identified 5955 DEx at the isoform level, which achieved a PPEE of <0.05 and a PPDE of >0.95. Due to the large number of DEx isoforms identified, to control for false positives, we focused on a subset of isoforms that achieved a higher significance for differential expression, which are 2344 DEx transcripts that achieved a PPDE of 1 (and a PPEE from 0 to 4.89E − 10).

The majority of the 2344 isoforms (~85%) were up-regulated in CP vs. control individuals. These isoforms are from 1816 distinct genes or non-coding RNAs. We used DAVID to analyze these isoforms and the top 10 most enriched functional terms are listed in Table 4.

### 3.8. Supporting evidence from the microarray data on the DEx isoforms identified from the RNA-seq data

Due to the large number (n = 2344) of DEx isoforms that achieved a PPDE of 1, we focused on the subset of transcripts (with a higher significance) that achieved both a PPDE of 1 and PPEE of 0 (1843 isoforms) to find supporting evidence from the microarray data. For these isoforms, we checked for the DEx signals at the gene level from the microarray data. The analysis identified supporting evidence for 213 isoforms from 177 distinct genes or non-coding RNAs, which achieved a $p$ value of less than 0.05 in the microarray data. Again they all had the same direction of regulation in terms of CP status across the two datasets. Almost all of the 213 isoforms/genes were up-regulated in CP cases vs. controls (or periodontally unhealthy status vs periodontally healthy status); among these isoforms, only those from 7 genes were down-regulated in CP cases vs. controls, which are URGCP, RPS20, FAM98A, XBP1, G3BP1, NFAT5 and ZNF207.

For these 213 isoforms/genes supported by the microarray dataset, we used DAVID to map them to functional terms and pathways. The top 10 most enriched functional terms/pathways were all GO terms (listed in Table 5). Almost all of these GO terms were significant even after Bonferroni correction (Bonferroni enrichment $p$ values from 5.83E − 3 to 5.06E − 2). Among these terms are endocytosis (GO:0006897), cytokine production (GO:0001816) and apoptosis (GO:0006915).

Thirteen genes (DEx in both RNA-seq and microarray datasets) belong to the GO term "GO:0006897~endocytosis", which are DENND1A,

**Table 4**
Top 10 enriched functional terms identified by DAVID for the DEx isoforms.

| Category | Term | Enrichment p value | Bonferroni p value | Fold enrichment |
|---|---|---|---|---|
| SP_PIR_KEYWORDS | ubl conjugation | 6.66E − 16 | 4.32E − 13 | 2.22 |
| GOTERM_CC_FAT | GO:0031981~nuclear lumen | 7.98E − 13 | 4.64E − 10 | 1.58 |
| GOTERM_CC_FAT | GO:0031974~membrane-enclosed lumen | 2.56E − 12 | 1.49E − 09 | 1.49 |
| GOTERM_CC_FAT | GO:0043233~organelle lumen | 2.58E − 12 | 1.50E − 09 | 1.50 |
| SP_PIR_KEYWORDS | Nucleotide-binding | 2.86E − 12 | 1.86E − 09 | 1.56 |
| GOTERM_CC_FAT | GO:0070013~intracellular organelle lumen | 5.87E − 12 | 3.42E − 09 | 1.49 |
| SP_PIR_KEYWORDS | Protein transport | 3.79E − 11 | 2.46E − 08 | 2.10 |
| GOTERM_BP_FAT | GO:0045184~establishment of protein localization | 1.13E − 10 | 4.04E − 07 | 1.73 |
| GOTERM_BP_FAT | GO:0046907~intracellular transport | 1.18E − 10 | 4.22E − 07 | 1.80 |
| GOTERM_BP_FAT | GO:0015031~protein transport | 1.21E − 10 | 4.35E − 07 | 1.74 |

RUFY1, CORO1C, ASGR2, APP, DAB2, PICALM, CD36, AP1S2, CLEC7A, THBS1, CLCN5 and RIN3. All of these genes/isoforms were up-regulated in CP patients vs. controls (or periodontally unhealthy vs. periodontally healthy statuses).

Seven genes (DEx in both RNA-seq and microarray datasets) belong to the GO term "GO:0001816~cytokine production", which are NLRC4, G6PD, MYD88, NFAT5, TLR4, NLRP3 and PTAFR. Except for NFAT5, all of the genes were up-regulated in CP patients vs. controls (or periodontally unhealthy vs. periodontally healthy statuses).

Twenty genes (DEx in both RNA-seq and microarray datasets) belong to the GO term "GO:0006915~apoptosis", which are ARHGEF2, SGK1, DNM1L, XIAP, UBE4B, CIDEB, STK17B, TRIO, NLRP3, BCL2L13, NCSTN, TNFRSF1A, PEA15, NLRC4, APP, GSN, HIPK3, BNIP3L, NLRP12, and THBS1. All of these genes were up-regulated in CP patients vs. controls (or periodontally unhealthy vs. periodontally healthy statuses). Among these genes, ten are categorized by DAVID as apoptosis induction genes (BCL2L13, BNIP3L, NLRC4, NLRP3, ARHGEF2, ARP, CIDEB, NCSTN, STK17B, TRIO), whereas only 5 were categorized as anti-apoptosis genes (BNIP3L, XIAP, HIPK, PEA15 and THBS1).

### 3.9. Supporting evidence from the microarray data at the pathway level for the DEx isoforms identified in the RNA-seq data

We used DAVID to map those 2344 DEx isoforms, which achieved a PPDE of 1 (and a PPEE from 0 to 4.89E − 10) in the RNA-seq dataset. For those functional terms/pathways that achieved a Bonferroni-corrected enrichment p value of <0.05, we cross-checked the enrichment signals against the DAVID analytical results from the microarray data.

Through this analysis, we identified 7 functional terms that achieved a Bonferroni-corrected enrichment p value of <0.05 in both the RNA-seq data and the microarray data. These functional terms include 6 GO terms and 1 SP_PIR_Keywords term (listed in Table 6). Among these terms, 3 are related to apoptosis: GO:0042981~regulation of apoptosis, GO:0043067~regulation of programmed cell death and GO:0010941~regulation of cell death, and other 4 are GO:0031982~vesicle, GO:0031410~cytoplasmic vesicle, GO:0010627~regulation of protein kinase cascade and actin-binding (SP_PIR_KEYWORDS).

In the RNA-seq data, 121 genes belong to the GO apoptosis-related terms. In the microarray data, 137 genes belong to the GO apoptosis-related terms. By comparing the 121 vs. 137 component genes across the two datasets, 22 genes (29 isoforms from RNA-seq data vs. 27 probe sets from microarray data) were shared between the two datasets (Table 7). These isoforms/probe sets were all up-regulated in CP vs. control individuals (or periodontally healthy vs. unhealthy statuses). According to DAVID analysis, among the 22 genes, 15 are classified as those for apoptosis induction (TRIO, ARHGEF2, BCL6, TLR4, STK17B, NLRP3, BCL2L13, IKBKG, BNIP3L, APP, NLRC4, RXRA, CIDEB, NCSTN, BID), whereas only 8 as those for anti-apoptosis (HIPK3, BCL6, BNIP3L, PEA15, XIAP, NUP62, THBS1, MYD88) (Table 7). (Note that BCL6 and BNIP3L are classified as both apoptosis induction and anti-apoptosis genes.)

## 4. Discussion

We performed the first RNA-seq study of PBMs for CP. Although we had a small sample size, the DEx transcripts as detected by the RNA-seq analysis were biologically relevant; the top 10 enriched functional terms identified by DAVID (dW et al., 2009) for the DEx transcripts identified at the whole gene level (Table 2) were almost all related to immune response, antigen processing and presentation and immuno-globulin/MHC. As host immune response is well recognized as the key basis for CP pathogenesis, this result (Table 2) lent strong support to the biological validity of DEx transcripts identified in our RNA-seq study.

We performed differential expression analysis of the RNA-seq data at both whole gene (or whole non-coding RNA) level and isoform level. We used a PBMs microarray dataset (Papapanou et al., 2007) to provide supporting evidence for our RNA-seq analyses results. We also used DAVID (dW et al., 2009) to map the DEx transcripts/isoforms identified in the RNA-seq data and supported by the microarray datasets to major functional terms (e.g., GO terms) in order to gain an overall picture of the functions of the identified genes.

Overall, the majority of the transcripts DEx at the whole gene level and supported by the microarray dataset are immune system or immune disease related genes, which were all up-regulated in CP vs. control individuals. Through analysis of the DEx transcripts at the functional

**Table 5**
Top 10 enriched functional terms identified by DAVID for the DEx isoforms supported by the microarray dataset.

| Category | Term | Enrichment p value | Bonferroni p value | Fold enrichment |
|---|---|---|---|---|
| GOTERM_BP_FAT | GO:0016044~membrane organization | 3.86E − 06 | 5.83E − 03 | 4.08 |
| GOTERM_BP_FAT | GO:0006897~endocytosis | 4.94E − 06 | 7.44E − 03 | 5.40 |
| GOTERM_BP_FAT | GO:0010324~membrane invagination | 4.94E − 06 | 7.44E − 03 | 5.40 |
| GOTERM_CC_FAT | GO:0031982~vesicle | 5.42E − 06 | 1.35E − 03 | 3.13 |
| GOTERM_CC_FAT | GO:0031988~membrane-bounded vesicle | 6.36E − 06 | 1.59E − 03 | 3.36 |
| GOTERM_BP_FAT | GO:0001816~cytokine production | 9.73E − 06 | 1.46E − 02 | 13.91 |
| GOTERM_CC_FAT | GO:0031410~cytoplasmic vesicle | 1.02E − 05 | 2.54E − 03 | 3.12 |
| GOTERM_CC_FAT | GO:0016023~cytoplasmic membrane-bounded vesicle | 1.53E − 05 | 3.83E − 03 | 3.30 |
| GOTERM_BP_FAT | GO:0006915~apoptosis | 2.80E − 05 | 4.15E − 02 | 3.04 |
| GOTERM_BP_FAT | GO:0012501~programmed cell death | 3.43E − 05 | 5.06E − 02 | 2.99 |

**Table 6**
The functional terms identified by DAVID on the DEx isoforms and supported by the microarray dataset.

| Category | Term | Discovery dataset (the RNA-seq dataset) | Supportive dataset (the microarray dataset) |
| --- | --- | --- | --- |
| | | Enrichment Bonferroni p value | Enrichment Bonferroni p value |
| GOTERM_CC_FAT | GO:0031982~vesicle | 2.82E − 05 | 2.11E − 04 |
| GOTERM_CC_FAT | GO:0031410~cytoplasmic vesicle | 4.10E − 04 | 4.08E − 04 |
| GOTERM_CC_FAT | GO:0010627~regulation of protein kinase cascade | 1.59E − 04 | 3.62E − 02 |
| GOTERM_CC_FAT | GO:0042981~regulation of apoptosis | 2.29E − 02 | 3.39E − 03 |
| GOTERM_CC_FAT | GO:0043067~regulation of programmed cell death | 3.71E − 02 | 5.97E − 03 |
| GOTERM_CC_FAT | GO:0010941~regulation of cell death | 4.43E − 02 | 7.40E − 03 |
| SP_PIR_KEYWORDS | Actin-binding | 8.33E − 04 | 4.19E − 02 |

term (or pathway) level using DAVID (dW et al., 2009), the term that was shared between our RNA-seq and the microarray datasets (Papapanou et al., 2007) is "immune response (a SP_PIR_Keyword)". Taking the above evidence, our findings confirmed immune response as the key mechanism for CP pathogenesis. Many genes identified in this study, in spite of their functional relevance to immune response, inflammation and other aspects of periodontitis pathogenesis, have not been linked to the disease in previous studies. Hence our findings also furnished novel genetic clues to the field and contributed to better delineation of the disease mechanism at the genetic/genomic level.

Additional new information was revealed by isoform level analysis of the RNA-seq data. First, the majority (~85%) of the DEx isoforms were up-regulated in CP patients vs. controls. This observation suggests that PBMs may be more functionally active at transcriptomic level in CP status than in healthy status, consistent with the key functional roles that PBMs may play in the pathologic process of CP.

Second, through DAVID (dW et al., 2009) analysis on the isoforms supported by the microarray dataset (Papapanou et al., 2007), we identified several GO terms, including endocytosis (GO:0006897), cytokine production (GO: 0001816) and apoptosis (GO: 0006915). Essentially all of the 40 component genes of the three GO terms were up-regulated in CP patients vs. controls, which may suggest enhanced monocytic endocytosis, cytokine production and apoptosis as the key potential mechanisms for CP.

The microarray dataset (Papapanou et al., 2007) is not a perfect dataset for supporting our RNA-seq study. However, to our knowledge, this is the only dataset that used PBMs to study gene expression changes related to periodontitis. As findings from gene expression studies are

**Table 7**
The component genes of the apoptosis GO term shared by the RNA-seq and the microarray datasets.

| Gene | Discovery (RNA-seq) dataset | | | | Supportive (microarray) dataset | | | Categorization according to DAVID |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Ensembl ID | PPEE | PPDE | PostFC | Probe set ID | p value | FC | |
| HIPK3 | ENST00000456517 | 0 | 1 | 183.15 | 210148_at | 1.90E − 03 | 2.32 | Anti-apoptosis |
| TRIO | ENST00000512070 | 0 | 1 | 60.94 | 209010_s_at | 9.04E − 03 | 1.48 | Apoptosis induction |
| TRIO | ENST00000537187 | 0 | 1 | 119.87 | 209010_s_at | 9.04E − 03 | 1.48 | Apoptosis induction |
| TRIO | ENST00000512070 | 0 | 1 | 60.94 | 244527_at | 3.25E − 02 | 1.48 | Apoptosis induction |
| TRIO | ENST00000537187 | 0 | 1 | 119.87 | 244527_at | 3.25E − 02 | 1.48 | Apoptosis induction |
| ARHGEF2 | ENST00000313667 | 0 | 1 | 205.24 | 207629_s_at | 1.17E − 02 | 1.77 | Apoptosis induction |
| BCL6 | ENST00000232014 | 0 | 1 | 49.61 | 228758_at | 1.26E − 02 | 3.16 | Apoptosis induction/ anti-apoptosis |
| TLR4 | ENST00000394487 | 0 | 1 | 20.66 | 221060_s_at | 1.50E − 02 | 2.87 | Apoptosis induction |
| TLR4 | ENST00000394487 | 0 | 1 | 20.66 | 232068_s_at | 3.88E − 02 | 8.60 | Apoptosis induction |
| TLR4 | ENST00000394487 | 0 | 1 | 20.66 | 224341_x_at | 4.06E − 02 | 12.61 | Apoptosis induction |
| STK17B | ENST00000409228 | 0 | 1 | 167.60 | 243797_at | 1.87E − 02 | 3.34 | Apoptosis induction |
| ACTN4 | ENST00000390009 | 0 | 1 | 11.75 | 200601_at | 1.87E − 02 | 1.76 | N/A |
| ACTN4 | ENST00000495553 | 0 | 1 | 22.24 | 200601_at | 1.87E − 02 | 1.76 | N/A |
| ACTN4 | ENST00000588618 | 0 | 1 | 111.48 | 200601_at | 1.87E − 02 | 1.76 | N/A |
| ACTN4 | ENST00000424234 | 0 | 1 | 197.24 | 200601_at | 1.87E − 02 | 1.76 | N/A |
| NLRP3 | ENST00000336119 | 0 | 1 | 4.53 | 216015_s_at | 2.16E − 02 | 2.44 | Apoptosis induction |
| NLRP3 | ENST00000391827 | 0 | 1 | 116.22 | 216015_s_at | 2.16E − 02 | 2.44 | Apoptosis induction |
| NLRP3 | ENST00000336119 | 0 | 1 | 4.53 | 207075_at | 4.35E − 02 | 4.97 | Apoptosis induction |
| NLRP3 | ENST00000391827 | 0 | 1 | 116.22 | 207075_at | 4.35E − 02 | 4.97 | Apoptosis induction |
| BCL2L13 | ENST00000317582 | 0 | 1 | 11.41 | 223664_x_at | 2.42E − 02 | 1.40 | Apoptosis induction |
| NLRP12 | ENST00000391775 | 0 | 1 | 169.64 | 1554952_s_at | 2.73E − 02 | 3.89 | N/A |
| NLRP12 | ENST00000391775 | 0 | 1 | 169.64 | 223944_at | 3.77E − 02 | 4.32 | N/A |
| IKBKG | ENST00000263518 | 0 | 1 | 1.07 | 209929_s_at | 2.86E − 02 | 1.57 | Apoptosis induction |
| BNIP3L | ENST00000523515 | 0 | 1 | 114.90 | 221479_s_at | 2.95E − 02 | 2.45 | Apoptosis induction/ anti-apoptosis |
| PEA15 | ENST00000488858 | 0 | 1 | 87.49 | 200788_s_at | 3.06E − 02 | 3.54 | Anti-apoptosis |
| APP | ENST00000439274 | 0 | 1 | 268.66 | 214953_s_at | 3.24E − 02 | 2.43 | Apoptosis induction |
| XIAP | ENST00000371199 | 0 | 1 | 1.37 | 243026_x_at | 3.55E − 02 | 1.72 | Anti-apoptosis |
| NLRC4 | ENST00000360906 | 0 | 1 | 16.54 | 1552553_a_at | 3.69E − 02 | 8.24 | Apoptosis induction |
| NUP62 | ENST00000597723 | 0 | 1 | 73.06 | 207740_s_at | 3.72E − 02 | 1.73 | Anti-apoptosis |
| NUP62 | ENST00000597029 | 0 | 1 | 101.74 | 207740_s_at | 3.72E − 02 | 1.73 | Anti-apoptosis |
| THBS1 | ENST00000560894 | 0 | 1 | 4.23 | 201108_s_at | 4.02E − 02 | 4.29 | Anti-apoptosis |
| THBS1 | ENST00000559746 | 0 | 1 | 24.85 | 201108_s_at | 4.02E − 02 | 4.29 | Anti-apoptosis |
| RXRA | ENST00000481739 | 0 | 1 | 12.21 | 202449_s_at | 4.19E − 02 | 3.99 | Apoptosis induction |
| CIDEB | ENST00000258807 | 0 | 1 | 1.42 | 221188_s_at | 4.20E − 02 | 1.88 | Apoptosis induction |
| NCSTN | ENST00000535857 | 0 | 1 | 65.56 | 208759_at | 4.59E − 02 | 2.05 | Apoptosis induction |
| BID | ENST00000399765 | 0 | 1 | 2.28 | 211725_s_at | 4.63E − 02 | 2.44 | Apoptosis induction |
| MYD88 | ENST00000463956 | 0 | 1 | 124.08 | 209124_at | 4.91E − 02 | 2.55 | Anti-apoptosis |

tissue/cell type dependent, it is critical that our RNA-seq study on PBMs is compared with a study using PBMs also. In terms of study design, the microarray study (Papapanou et al., 2007) did not compare periodontitis patients with healthy controls as in our RNA-seq study. However, the microarray study data do allow us to compare a periodontally unhealthy status (at the baseline) with a periodontally healthy status (at 10 weeks when periodontal tissue is much healthier compared with the baseline) through a paired study design (Papapanou et al., 2007). From the perspective of periodontal health, our RNA-seq study had a similar study goal, which is also to compare periodontally unhealthy status (in diseased individuals) with periodontally healthy status (in healthy individuals), although this comparison was done cross-sectionally. Therefore, although we acknowledge the limitation of the microarray study (Papapanou et al., 2007) as a supportive dataset for our RNA-seq study, due to the above reasons, we are confident that the comparison between the two datasets is necessary and valuable and can capture "converging" evidence for transcripts associated with periodontitis, which may represent much solid scientific findings than those from the RNA-seq study alone. Indeed, the value of using the microarray dataset is evidenced by the functional relevance of the identified genes shared by the two datasets (Table 3).

This study has two major limitations. First, it has a small sample size, which leads to a limited statistical power and more genes/transcripts associated with CP may be identified with a larger sample size. Second, the sequencing depth (100 bp single ended reads) is also limited and more isoforms, especially those isoforms with a low expression, may be detected with a higher sequencing depth. Therefore, the findings from this study may need to be further validated with independent studies of larger sample sizes and higher sequencing depths. Such studies are also needed to identify more genes and isoforms important to CP.

## 5. Conclusions

In summary, through RNA-seq analysis, we identified a large number of transcripts DEx in CP vs. control individuals at both whole gene and isoform levels. At the whole gene level, the identified transcripts are enriched in immune response related functional terms (Table 2). Comparison analysis using the microarray dataset for these transcripts also identified a number of novel genes related to immune response (Table 3), such as FCAR and CUX1, which were not reported for CP before. New, interesting information emerged from the isoform level analyses, including the pervasive up-regulation of PBMs gene expression, and enhanced endocytosis, cytokine production and apoptosis in CP vs. control status (Tables 5 and 6). Overall, our findings highlighted the significant role of PBMs for CP pathogenesis and identified a number of novel genes and potentially new mechanisms for the disease. These findings may represent novel biomarkers/signatures for clinical diagnosis and treatment of CP and provide the basis for further biological and functional investigations.

## Competing interests

The author(s) declare that they have no competing interests.

## Authors' contributions

YZL: RNA-seq and microarray data analysis and manuscript writing
PM: coordination of subject recruitment, manuscript writing
JP: subject recruitment
YZ: monocyte isolation, RNA extraction
MB: RNA-seq data analysis
MS: RNA-seq data analysis
YPW: manuscript writing
EF: coordination of RNA-seq data analysis, manuscript writing
HWD: manuscript writing

## References

Achiron, A., Gurevich, M., Magalashvili, D., Kishner, I., Dolev, M., Mandel, M., 2004. Understanding autoimmune mechanisms in multiple sclerosis using gene expression microarrays: treatment effect and cytokine-related pathways. Clin. Dev. Immunol. 11, 299–305.

Alex, P., Ye, M., Zachos, N.C., Sipes, J., Nguyen, T., Suhodrev, M., Gonzales, L., Arora, Z., Zhang, T., Centola, M., Guggino, S.E., Li, X., 2010. Clcn5 knockout mice exhibit novel immunomodulatory effects and are more susceptible to dextran sulfate sodium-induced colitis. J. Immunol. 184, 3988–3996.

Amadou, C., Ribouchon, M.T., Mattei, M.G., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., Avoustin, P., Pontarotti, P., 1995. Localization of new genes and markers to the distal part of the human major histocompatibility complex (MHC) region and comparison with the mouse: new insights into the evolution of mammalian genomes. Genomics 26, 9–20.

Arlt, A., Schafer, H., 2011. Role of the immediate early response 3 (IER3) gene in cellular stress response, inflammation and tumorigenesis. Eur. J. Cell Biol. 90, 545–552.

Armitage, G.C., 1999a. Development of a classification system for periodontal diseases and conditions. Ann. Periodontol. 4, 1–6.

Armitage, G.C., 1999b. Development of a classification system for periodontal diseases and conditions. Ann. Periodontol. 4, 1–6.

Bovin, L.F., Brynskov, J., Hegedus, L., Jess, T., Nielsen, C.H., Bendtzen, K., 2007. Gene expression profiling in autoimmune diseases: chronic inflammation or disease specific patterns? Autoimmunity 40, 191–201.

Boyatzis, S., Seymour, G.J., 1986. Effect of age and periodontal disease status in man on the spontaneous proliferation of peripheral blood lymphocytes. Arch. Oral Biol. 31, 749–755.

Davanian, H., Stranneheim, H., Bage, T., Lagervall, M., Jansson, L., Lundeberg, J., Yucel-Lindberg, T., 2012. Gene expression profiles in paired gingival biopsies from periodontitis-affected and healthy tissues revealed by massively parallel sequencing. PLoS One 7, e46440.

de Jong, S., van Eijk, K.R., Zeegers, D.W., Strengman, E., Janson, E., Veldink, J.H., van den Berg, L.H., Cahn, W., Kahn, R.S., Boks, M.P., Ophoff, R.A., 2012. Expression QTL analysis of top loci from GWAS meta-analysis highlights additional schizophrenia candidate genes. Eur. J. Hum. Genet. 20, 1004–1008.

Deng, F.Y., Lei, S.F., Zhang, Y., Zhang, Y.L., Zheng, Y.P., Zhang, L.S., Pan, R., Wang, L., Tian, Q., Shen, H., Zhao, M., Lundberg, Y.W., Liu, Y.Z., Papasian, C.J., Deng, H.W., 2011. Peripheral blood monocyte-expressed ANXA2 geneis involved in pathogenesis of osteoporosis in humans. Mol. Cell. Proteomics (-PMCID: PMC3226411).

dW, Huang, BT, Sherman, RA, Lempicki, 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.

Eke, P.I., Dye, B.A., Wei, L., Thornton-Evans, G.O., Genco, R.J., 2012. Prevalence of periodontitis in adults in the United States: 2009 and 2010. J. Dent. Res. 91, 914–920.

Ezzo, P.J., Cutler, C.W., 2003. Microorganisms as risk indicators for periodontal disease. Periodontol. 32, 24–35.

Fan, W., Wang, Z., Kyzysztof, F., Prange, C., Lennon, G., 2000. A new zinc ribbon gene (ZNRD1) is cloned from the human MHC class I region. Genomics 63, 139–141.

GK, Smyth, 2014. Limma: linear models for microarray data. In: Gentleman, R., Carey, V., Dudoit, S., RA, Irizarry, Huber, W. (Eds.), Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer, New York, pp. 397–420.

Herr, A.B., Ballister, E.R., Bjorkman, P.J., 2003. Insights into IgA-mediated immune responses from the crystal structures of human FcalphaRI and its complex with IgA1-Fc. Nature 423, 614–620.

Heul-Nieuwenhuijsen, L., Padmos, R.C., Drexhage, R.C., de Wit, H., Berghout, A., Drexhage, H.A., 2010. An inflammatory gene-expression fingerprint in monocytes of autoimmune thyroid disease patients. J. Clin. Endocrinol. Metab. 95, 1962–1971.

Hsu, W.L., Tse, K.P., Liang, S., Chien, Y.C., Su, W.H., Yu, K.J., Cheng, Y.J., Tsang, N.M., Hsu, M.M., Chang, K.P., Chen, I.H., Chen, T.I., Yang, C.S., Goldstein, A.M., Chen, C.J., Chang, Y.S., Hildesheim, A., 2012. Evaluation of human leukocyte antigen-A (HLA-A), other non-HLA markers on chromosome 6p21 and risk of nasopharyngeal carcinoma. PLoS One 7, e42767.

Ikeda, T., Kumagai, E., Iwata, S., Yamakawa, A., 2013. Soluble CD26/dipeptidyl peptidase IV enhances the transcription of IL-6 and TNF-alpha in THP-1 cells and monocytes. PLoS One 8, e66520.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., Speed, T.P., 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4, 249–264.

Kebschull, M., Demmer, R.T., Grun, B., Guarnieri, P., Pavlidis, P., Papapanou, P.N., 2014. Gingival tissue transcriptomes identify distinct periodontitis phenotypes. J. Dent. Res. 93, 459–468.

Korn, T., Bettelli, E., Oukka, M., Kuchroo, V.K., 2009. IL-17 and Th17 cells. Annu. Rev. Immunol. 27, 485–517.

Kuang, E., Okumura, C.Y., Sheffy-Levin, S., Varsano, T., Shu, V.C., Qi, J., Niesman, I.R., Yang, H.J., Lopez-Otin, C., Yang, W.Y., Reed, J.C., Broday, L., Nizet, V., Ronai, Z.A., 2012. Regulation of ATG4B stability by RNF5 limits basal levels of autophagy and influences susceptibility to bacterial infection. PLoS Genet. 8, e1003007.

Lee, J.S., Trowsdale, J., Bodmer, W.F., 1982. cDNA clones coding for the heavy chain of human HLA-DR antigen. Proc. Natl. Acad. Sci. U. S. A. 79, 545–549.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., Kendziorski, C., 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics 29, 1035–1043.

Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. 12, 323.

Mallya, M., Campbell, R.D., Aguado, B., 2006. Characterization of the five novel Ly-6 super-family members encoded in the MHC, and detection of cells expressing their poten-tial ligands. Protein Sci. 15, 2244–2256.

Mitchell, A., Rentero, C., Endoh, Y., Hsu, K., Gaus, K., Geczy, C., McNeil, H.P., Borges, L., Tedla, N., 2008. LILRA5 is expressed by synovial tissue macrophages in rheumatoid arthritis, selectively induces pro-inflammatory cytokines and IL-10 and is regulated by TNF-alpha, IL-10 and IFN-gamma. Eur. J. Immunol. 38, 3459–3473.

Montenegro, V., Chiamolera, M., Launay, P., Goncalves, C.R., Monteiro, R.C., 2000. Impaired expression of IgA Fc receptors (CD89) by blood phagocytic cells in ankylosing spon-dylitis. J. Rheumatol. 27, 411–417.

Mosmann, T.R., Coffman, R.L., 1989. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annu. Rev. Immunol. 7, 145–173.

Mulcahy, H., O'Rourke, K.P., Adams, C., Molloy, M.G., O'Gara, F., 2006. LST1 and NCR3 ex-pression in autoimmune inflammation and in response to IFN-gamma, LPS and mi-crobial infection. Immunogenetics 57, 893–903.

Offenbacher, S., Heasman, P.A., Collins, J.G., 1993. Modulation of host PGE2 secretion as a determinant of periodontal disease expression. J. Periodontol. 64, 432–444.

Papapanou, P.N., Sedaghatfar, M.H., Demmer, R.T., Wolf, D.L., Yang, J., Roth, G.A., Celenti, R., Belusko, P.B., Lalla, E., Pavlidis, P., 2007. Periodontal therapy alters gene expression of peripheral blood monocytes. J. Clin. Periodontol. 34, 736–747.

Petit, M.D., Wassenaar, A., van der Velden, U., van Eden, W., Loos, BG, 1999. Depressed re-sponsiveness of peripheral blood mononuclear cells to heat-shock proteins in peri-odontitis patients. J. Dent. Res. 78, 1393–1400.

Pulendran, B., Palucka, K., Banchereau, J., 2001. Sensing pathogens and tuning immune re-sponses. Science 293, 253–256.

Reynolds, R.J., Kelley, J.M., Hughes, L.B., Yi, N., Bridges Jr., S.L., 2010. Genetic association of htSNPs across the major histocompatibility complex with rheumatoid arthritis in an African-American population. Genes Immun. 11, 94–97.

Rollinger-Holzinger, I., Eibl, B., Pauly, M., Griesser, U., Hentges, F., Auer, B., Pall, G., Schratzberger, P., Niederwieser, D., Weiss, E.H., Zwierzina, H., 2000. LST1: a gene with extensive alternative splicing and immunomodulatory function. J. Immunol. 164, 3169–3176.

Santin, I., Castellanos-Rubio, A., Aransay, A.M., Gutierrez, G., Gaztambide, S., Rica, I., Vicario, J.L., Noble, J.A., Castano, L., Bilbao, J.R., 2009. Exploring the diabetogenicity of the HLA-B18-DR3 CEH: independent association with T1D genetic risk close to HLA-DOA. Genes Immun. 10, 596–600.

Shen, H., Qiu, C., Li, J., Tian, Q., Deng, H.W., 2013. Characterization of the DNA methylome and its interindividual variation in human peripheral blood monocytes. Epigenomics 5, 255–269.

Sigusch, B., Klinger, G., Glockmann, E., Simon, H.U., 1998. Early-onset and adult periodon-titis associated with abnormal cytokine production by activated T lymphocytes. J. Periodontol. 69, 1098–1104.

Smyth, G.K., 2005. Limma: linear models for microarray data. In: Gentleman, R., Carey, V., Dudoit, S., Irizarry, R., Huber, W. (Eds.), Bioinformatics and Computational Biology So-lutions using R and Bioconductor. Springer, pp. 397–420.

Taubman, M.A., Valverde, P., Han, X., Kawai, T., 2005. Immune response: the key to bone resorption in periodontal disease. J. Periodontol. 76, 2033–2041.

Tew, J.G., Burmeister, J.A., Palcanis, K.G., Ranney, R.R., 1983. Spontaneous lymphocyte pro-liferation and the periodontal status of young adults. J. Periodontal Res. 18, 534–540.

Topic, R.Z., Dodig, S., 2011. Eosinophil cationic protein—current concepts and controver-sies. Biochem. Med. (Zagreb) 21, 111–121.

Visekruna, A., Volkov, A., Steinhoff, U., 2012. A key role for NF-kappaB transcription factor c-Rel in T-lymphocyte-differentiation and effector functions. Clin. Dev. Immunol. 2012, 239368.

Yucesoy, B., Johnson, V.J., Lummus, Z.L., Kashon, M.L., Rao, M., Bannerman-Thompson, H., Frye, B., Wang, W., Gautrin, D., Cartier, A., Boulet, L.P., Sastre, J., Quirce, S., Tarlo, S.M., Germolec, D.R., Luster, M.I., Bernstein, D.I., 2014. Genetic variants in the major histo-compatibility complex class I and class II genes are associated with diisocyanate-induced asthma. J. Occup. Environ. Med. 56, 382–387.