ORAL DISEASES
Leading in Oral, Maxillofacial, Head & Neck Medicine

Structural and functional microbial patterns in cohabitating family members with history of periodontitis

Journal:	*Oral Diseases*
Manuscript ID	ODI-08-20-OM-8760.R2
Manuscript Type:	Short Communication
Date Submitted by the Author:	n/a
Complete List of Authors:	Diaz-Faes, Lucia; University of Granada, Periodontology, Stomatology Soriano-Lerma, Ana; University of Granada, Physiology Magán-Fernández, Antonio; University of Granada, Periodontology, Stomatology López, Maria; University of Granada, Periodontology, Stomatology Gijon, Juan; University of Granada, Periodontology, Stomatology García-Salcedo, Jose Antonio; GENYO Soriano, Miguel; GENYO Mesa, Francisco; Universidad de Granada Facultad de Odontología, Periodontology, Stomatology
Keywords:	High-Throughput Nucleotide Sequencing., Microbiota, Chronic Periodontitis, Aggressive Periodontitis

Structural and functional microbial patterns in cohabitating family members with history of periodontitis.

Authors: Lucía Díaz-Faes*[1], Ana Soriano-Lerma*[2, 3], Antonio Magan-Fernandez[1], María López[1], Juan Gijon[1], Jose Antonio García-Salcedo[3, 4], Miguel Soriano[#3, 5], Francisco Mesa[#1].

* Lucía Díaz-Faes and Ana Soriano-Lerma have equally contributed to this work.

[#]Miguel Soriano and Francisco Mesa have equally contributed to this work as senior authors.

Affiliations:

1. Department of Periodontics, Faculty of Dentistry, University of Granada, Granada, Spain.
2. Department of Physiology (Faculty of Pharmacy, Campus Universitario de Cartuja), Institute of Nutrition and Food Technology "José Mataix", University of Granada, E-18071 Granada, Spain.
3. GENYO. Centre for Genomics and Oncological Research: Pfizer / University of Granada / Andalusian Regional Government, PTS Granada, E-18016, Granada, Spain.
4. Microbiology Unit, Biosanitary Research Institute ibs. GRANADA, University Hospital Virgen de las Nieves, E-18014, Granada, Spain.
5. Center for Intensive Mediterranean Agrosystems and Agri-food Biotechnology (CIAMBITAL), University of Almeria, E-04001, Almería, Spain.

Running title: Microbial patterns in periodontitis.

Keywords: Chronic Periodontitis; Aggressive Periodontitis; Microbiota; High-Throughput Next Generation Sequencing.

Acknowledgements: This work was financially supported by the Ministry of Economy and Competitiveness of Spain co-financed with European Regional Development Funds (References: CGL2015-71709-R) and research group BIO344 (Junta de Andalucía, Spain). A.S.L. was supported by a fellowship from the Ministry of Education, Culture and Sport (FPU 17/05413). JAGS was supported by 3TR, IMI2 (H2020-JTI-IMI2-2018) and by Instituto de Salud Carlos III under the frame of EuroNanoMed III (AC18/00008). The authors declare that they have no competing interests.

Corresponding author
Antonio Magan-Fernandez, DDS, MSc., PhD.
Facultad de Odontología, Campus de Cartuja s/n, Universidad de Granada. E-18071 GRANADA, Spain. Tel +34-958240654 E-mail: amaganf@ugr.es

Periodontitis is characterized by microbially-associated, host-mediated inflammation resulting in loss of periodontal attachment(Tonetti et al., 2018). Although specific genetic associations have not yet been consistently demonstrated, family aggregation supports the importance of genetic influence on disease risk(Fine et al., 2019), with aggressive periodontitis (AgP) having a better established heritable component compared to chronic periodontitis (CP)(Munz et al., 2019, Offenbacher et al., 2016, Divaris et al., 2013). Whether such genetic influence involves bacterial transmission within families is, however, currently unknown.

It is unclear whether currently available clinical parameters are sufficient to assess the grade of periodontitis(Tonetti et al., 2018), and considering the high throughput yield of data provided by next generation sequencing techniques (NGS), a special focus on periodontitis-associated pathogens and their functionality might be adopted to identify new biomarkers, increase diagnostic accuracy and enhance disease management.

The objectives of this pilot study were to assess within-family transmission of bacterial genera related to periodontitis and to analyse the structure and functioning of subgingival microbial communities in AgP, CP and healthy subjects.

Subjects were recruited from the University Dental Clinic (Department of Periodontics, University of Granada, Andalucía, Spain). Written informed consent was obtained from all participants, being the study approved by the Human Research Ethics Committee of the

University of Granada (1480/CEIH/2020). Study inclusion criteria were families with periodontal disease defined as: a) Stages III or IV and grade C (formerly known AgP), b) Stages I to IV and grades A or B (formerly known CP), or c) Periodontally healthy individuals. At least one offspring fulfilling condition a) was required to be eligible. Exclusion criteria included periodontal or antibiotic treatment in the 6 months prior to the study, immunosuppression, or any systemic condition affecting the microbiome.

Dental and clinical histories were obtained, and a periodontal examination was performed. Subgingival plaque samples were obtained from the sites showing the largest probing depth and bleeding in each quadrant, according to Mombelli´s technique(Mombelli et al., 1991). After the removal of supragingival plaque with an ultrasonic device, areas of interest were isolated and two sterile no.30 paper tips (Maillefer, Ballaigues, Switzerland) were inserted consecutively for 10 seconds, being subsequently frozen (-20ºC).

DNA isolation, 16S rRNA amplicon sequencing and bioinformatic analysis were as previously described(Soriano-Lerma et al., 2020a). Redundant, non-chimera FASTA files were taxonomically classified using the Ribosomal Database Project Bayesian classifier and database(Wang et al., 2007). Abundance was expressed as a percentage with respect to the total number of sequences in each sample. Functional analysis was carried out using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States(Douglas et al., 2018) on high-throughput 16S rRNA gene sequencing data. Principal Component Analysis (PCA) and Linear discriminant analysis effect size (LEfSe)(Segata et al., 2011) were implemented as previously described(Soriano-Lerma et al., 2020b). Venn diagrams were plotted using specific software (Mothur v1.42.0, University of Michigan Medical School, Ann Arbor, MI, USA)(Schloss et al., 2009). SourceTracker analysis was performed using QIIME

(Python 3.5) (Knights et al., 2011). Mann Whitney U test was implemented in SPSS v.20.0 (SPSS Inc., Chicago, IL, USA).

Three families composed of four members (father, mother, son and daughter) whose ages ranged between 17 and 59 years old, were consecutively enrolled in the study. Subjects were classified into three groups: AgP, CP, and healthy patients (H), as shown in Table S1.

Analysis of within-family transmission of bacterial genera associated with AgP and CP revealed that in family 1 (Figure 1a), the two periodontally affected members showed different dominating genera: *Fretibacterium* and *Rikenellaceae_rc9_gut_group* in the daughter, and *Pseudomonas* in the mother. Operational taxonomic units (OTU) analysis and Venn diagrams revealed that, in fact, only 47 species were shared between daughter and mother, out of 134 in both samples (Figure 1a). The number of species transmitted from mother to daughter was estimated in 0% (Table S2). Strikingly, *Rikenellaceae_rc9_gut_group* has been described to harbour three bacterial species commonly found in the gastrointestinal tract of several animals and related to periodontal disease in dogs(Hardham et al., 2005). Further enquiries revealed that some of the daughter's habits included kissing the dog's mouth, thus providing a possible explanation for the transmission. To the best of our knowledge, this is the first study showing the simultaneous presence of *Rikenellaceae_rc9_gut_group* in animals and humans, and its association with periodontal disease. Similar results were obtained in family 2 (Figure 1b), where the dominating microbial genera in the progenitors, *Neisseria* in the mother and *Enterobacter* in the father, were not considerably abundant in the offspring. OTU analysis and Venn diagrams detected 25 shared species among all members in family 2 (Figure 1b), out of 349. The transmission from parents to offspring was greater in the case of the daughter, although more than 50% of species were estimated to proceed from an unknown origin (Table

S2, Supplementary Figure 1). All identified taxa in both families have been clearly associated with polymicrobial infections and periodontal disease(Patini et al., 2018, Vieira Colombo et al., 2016, Ardila Medina, 2010). Family 3 was the only one showing a similar microbial composition within periodontally affected members, sharing their main microbial genera (*Fretibacterium, Fusobacterium, Porphyromonas, Prevotella* and *Treponema)* and species (Figure 1c); however, the healthy son differed in the relative abundance of microbial genera such as *Rothia*, *Streptococcus* and *Propionibacterium,* gram-positive facultative bacteria forming plaque biofilms, not necessarily related to periodontitis(Diaz et al., 2006). A reduced number of species were shared between the periodontally healthy son and the other members in the family, while 70 species were shared between daughter and parents (Figure 1c). Consequently, a higher percentage of bacterial transmission was estimated in the case of the daughter (Table S2, Supplementary Figure 2), while the son showed 70% of species from an unknown origin. An overview of the total number of sequences per sample along with bacterial genera whose relative abundance was higher than 0.1% is included in Table S3.

PCA was carried out to determine whether microbial communities present in AgP and CP patients were similar, being bacterial genera with relative abundance higher than 0.1% included as variables. The healthy subject showed a distinctive microbiota profile, since it was clearly separated from patients suffering either AgP or CP along the X axis (PC1). However, no clear separation could be observed between AgP and CP samples along the Y axis (PC2), since both ellipses were overlapped to a great extent (Figure 2a). For the most part, microbial composition at the genus level was similar for both conditions. At the species level, OTU analysis identified 237 shared species between AgP and CP patients, out of a total number of 454 (Figure 2a). In this context, host immune response could play a considerable role in the grade of periodontitis.

PCA was then performed to assess microbial functional patterns in AgP and CP. KEGG microbial pathways classified at level 3 were included as variables, showing in this case a clearer differentiation between AgP and CP along the Y axis (PC2) (Figure 2b) with the severity of the disease increasing towards the top part of the graphic. Hence, functional microbial variables could be useful differentiating features to discern between AgP and CP and were therefore identified through LeFSe analysis (Table S4). Mann Whitney U test was applied on individual genes within the most abundant and differentially distributed pathways (Figure 3) to provide guidance about potential indicators for periodontal state that might be investigated in more detail and in a bigger cohort of patients (Table S5). Distinction between AgP and CP might be achieved through the determination of each causative bacterial species. However, our approach using functional biomarkers (such as KEGG orthologs) might be a useful alternative since it allows a more specific differentiation of the two conditions. Some microbial pathways are often shared between different bacterial species and therefore, its analysis facilitate diagnostic procedures.

NGS techniques provides researchers and clinicians with valuable information about key microbial taxa involved in several pathological processes, being a useful approach to identify potential diagnostic, prognostic and therapeutic targets(Segata et al., 2011). Overall, this pilot study revealed no evident within family transmission of bacterial genera related to AgP or CP, thus suggesting a predominant role of bacterial-independent genetic heritability in family aggregation. It also puts forward the idea of studying microbial functionality and genes as potential diagnostic biomarkers to discern between AgP and CP, since they seem to better reflect differences between conditions compared to genera detection.

Study limitations include the small sample size and reduced statistical power, and the existence of confounders, which might affect structural and functional analysis of the subgingival microbiome. Selection and information skews should be taken into consideration and further research would be needed to investigate these findings in detail.

To sum up, our pilot study showed no clear intrafamilial transmission considering the structure of subgingival microbial communities. A trend regarding differences at the structural and functional level was identified in the subgingival microbiota of the healthy subject compared to periodontal patients, while AgP and CP affected subjects differed to a greater extent in terms of bacterial functionality.

REFERENCES

Ardila Medina C (2010). Asociación potencial entre enterobacterias presentes en periodontitis y enfermedades sistémicas. *Acta odontológica venezolana* **48**: 108-113.

Diaz PI, Chalmers NI, Rickard AH, Kong C, Milburn CL, Palmer RJ, Jr. and Kolenbrander PE (2006). Molecular characterization of subject-specific oral microflora during initial colonization of enamel. *Appl Environ Microbiol* **72**: 2837-48.

Divaris K, Monda KL, North KE, Olshan AF, Reynolds LM, Hsueh WC, Lange EM, Moss K, Barros SP, Weyant RJ, Liu Y, Newman AB, Beck JD and Offenbacher S (2013). Exploring the genetic basis of chronic periodontitis: a genome-wide association study. *Hum Mol Genet* **22**: 2312-24.

Douglas GM, Beiko RG and Langille MGI (2018). Predicting the Functional Potential of the Microbiome from Marker Genes Using PICRUSt. *Methods in Molecular Biology*. Humana Press Inc., pp. 169-177.

Fine DH, Armitage GC, Genco RJ, Griffen AL and Diehl SR (2019). Unique etiologic, demographic, and pathologic characteristics of localized aggressive periodontitis support classification as a distinct subcategory of periodontitis. *J Am Dent Assoc* **150**: 922-931.

Hardham J, Dreier K, Wong J, Sfintescu C and Evans RT (2005). Pigmented-anaerobic bacteria associated with canine periodontitis. *Vet Microbiol* **106**: 119-28.

Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R and Kelley ST (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature methods* **8**: 761-3.

Mombelli A, McNabb H and Lang NP (1991). Black-pigmenting gram-negative bacteria in periodontal disease. I. Topographic distribution in the human dentition. *J Periodontal Res* **26**: 301-7.

Munz M, Richter GM, Loos BG, Jepsen S, Divaris K, Offenbacher S, Teumer A, Holtfreter B, Kocher T, Bruckmann C, Jockel-Schneider Y, Graetz C, Ahmad I, Staufenbiel I, van der Velde N, Uitterlinden AG, de Groot L, Wellmann J, Berger K, Krone B, Hoffmann P, Laudes M, Lieb W, Franke A, Erdmann J, Dommisch H and Schaefer AS (2019). Meta-analysis of genome-wide association studies of aggressive and chronic periodontitis identifies two novel risk loci. *Eur J Hum Genet* **27**: 102-113.

Offenbacher S, Divaris K, Barros SP, Moss KL, Marchesan JT, Morelli T, Zhang S, Kim S, Sun L, Beck JD, Laudes M, Munz M, Schaefer AS and North KE (2016). Genome-wide association study of biologically informed periodontal complex traits offers novel insights into the genetic basis of periodontal disease. *Hum Mol Genet* **25**: 2113-2129.

Patini R, Staderini E, Lajolo C, Lopetuso L, Mohammed H, Rimondini L, Rocchetti V, Franceschi F, Cordaro M and Gallenzi P (2018). Relationship between oral microbiota and periodontal disease: a systematic review. *Eur Rev Med Pharmacol Sci* **22**: 5775-5788.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ and Weber CF (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537-41.

Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS and Huttenhower C (2011). Metagenomic biomarker discovery and explanation. *Genome Biol* **12**: R60.

Soriano-Lerma A, Magan-Fernandez A, Gijon J, Sanchez-Fernandez E, Soriano M, Garcia-Salcedo JA and Mesa F (2020a). Short-term effects of hyaluronic acid on the subgingival microbiome in peri-implantitis: A randomized controlled clinical trial. *J Periodontol* **91**: 734-745.

Soriano-Lerma A, Perez-Carrasco V, Sanchez-Maranon M, Ortiz-Gonzalez M, Sanchez-Martin V, Gijon J, Navarro-Mari JM, Garcia-Salcedo JA and Soriano M (2020b). Influence of 16S rRNA target region on the outcome of microbiome studies in soil and saliva samples. *Sci Rep* **10**: 13637.

Tonetti MS, Greenwell H and Kornman KS (2018). Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. *J Periodontol* **89 Suppl 1**: S159-S172.

Vieira Colombo AP, Magalhaes CB, Hartenbach FA, Martins do Souto R and Maciel da Silva-Boghossian C (2016). Periodontal-disease-associated biofilm: A reservoir for pathogens of medical importance. *Microb Pathog* **94**: 27-34.

Wang Q, Garrity GM, Tiedje JM and Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261-7.

FIGURE LEGENDS

Figure 1. Column diagrams representing the relative percentage of most abundant genera in all members of each family (a, b, c; left part). Venn diagrams at distance 0.03 showing the number of bacterial species shared between members in each family (a, b, c; right part). Number of shared species between individuals is represented in each respective overlapping area; abbreviations represent each respective family (F1: family 1, F2: family 2 and F3: family 3), and each respective member (S: son; D: daughter; M: mother; F: father). (a) Family 1 (b) Family 2 (c) Family 3.

Figure 2. Principal component analysis (PCA) (a, b; left part) and Venn diagram at distance 0.03 showing the number of bacterial species shared between AgP and CP patients (a, right part). PCA includes plots for the bacterial genera with a relative abundance higher than 0.1% (a, left part) and for KEGG microbial pathways classified at level 3 (b). Abbreviations represent each respective family (F: family 1, F2: family 2 and F3: family 3), and each respective member (S: son; D: daughter; M: mother; F: father). Samples are represented by its identification and colour ellipses are added to represent clustering around oral health conditions (AgP, CP and H). (a) PCA considering bacterial genera with relative abundance higher than 0.1% (left part). Venn diagram at distance 0.03 showing, in the overlapping area, the number of bacterial species shared between AgP and CP patients (right part) (b) PCA considering KEGG microbial pathways classified at level 3.

Figure 3. KEGG microbial pathways classified at level 3 differentially distributed between AgP and CP. Mean relative abundance and standard deviation for each condition are represented

in horizontal columns and error bars, respectively. For clearer visualization, only the most abundant pathways have been included in the plot.

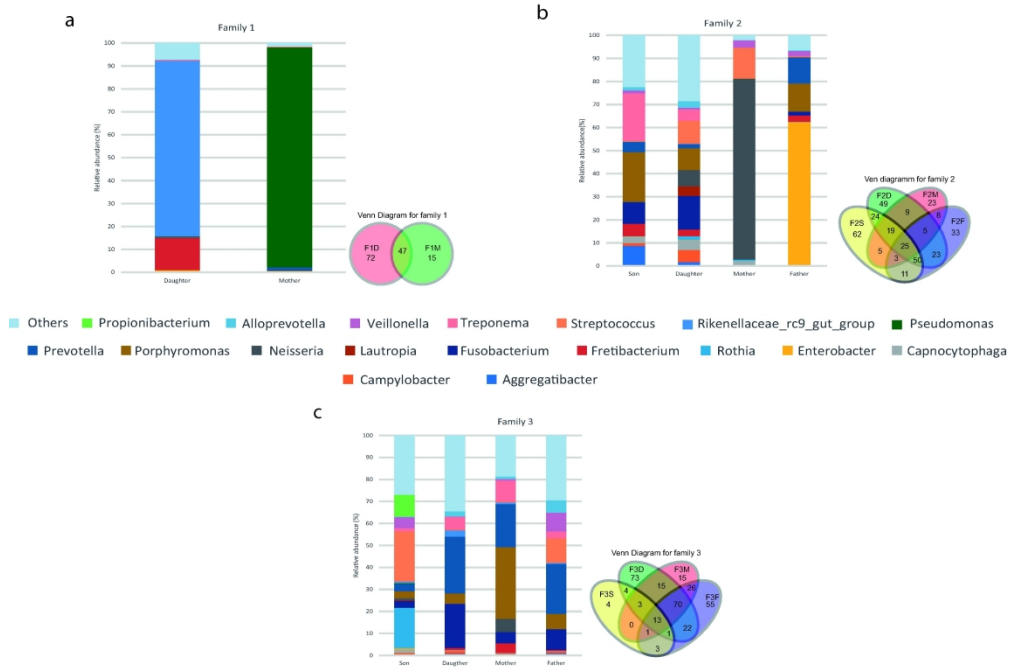


Figure 1. Column diagrams representing the relative percentage of most abundant genera in all members of each family (left part). Venn diagrams at distance 0.03 showing the number of bacterial species shared between members in each family (right part). Number of shared species between individuals is represented in each respective overlapping area; abbreviations represent each respective family (F1: family 1, F2: family 2 and F3: family 3), and each respective member (S: son; D: daughter; M: mother; F: father). (a) Family 1 (b) Family 2 (c) Family 3.

153x102mm (600 x 600 DPI)

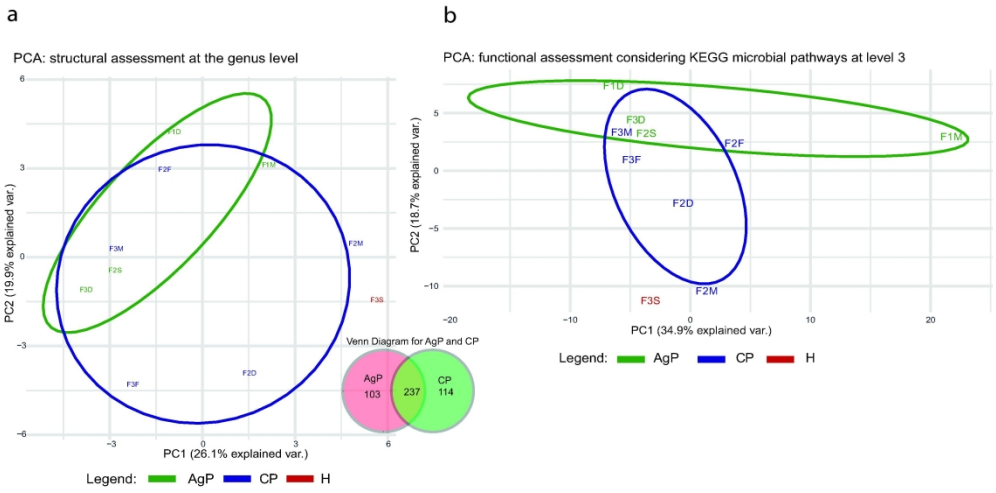


Figure 2. Principal component analysis (PCA) (left part) and Venn diagram at distance 0.03 showing the number of bacterial species shared between AgP and CP patients (a, right part). PCA includes plots for the bacterial genera with a relative abundance higher than 0.1% (a, left part) and for KEGG microbial pathways classified at level 3 (b). Abbreviations represent each respective family (F: family 1, F2: family 2 and F3: family 3), and each respective member (S: son; D: daughter; M: mother; F: father). Samples are represented by its identification and colour ellipses are added to represent clustering around oral health conditions (AgP, CP and H). (a) PCA considering bacterial genera with relative abundance higher than 0.1% (left part). Venn diagram at distance 0.03 showing the number of bacterial species shared between AgP and CP patients (right part); see legend in Figure 1 (b) PCA considering KEGG microbial pathways classified at level 3.

152x78mm (600 x 600 DPI)

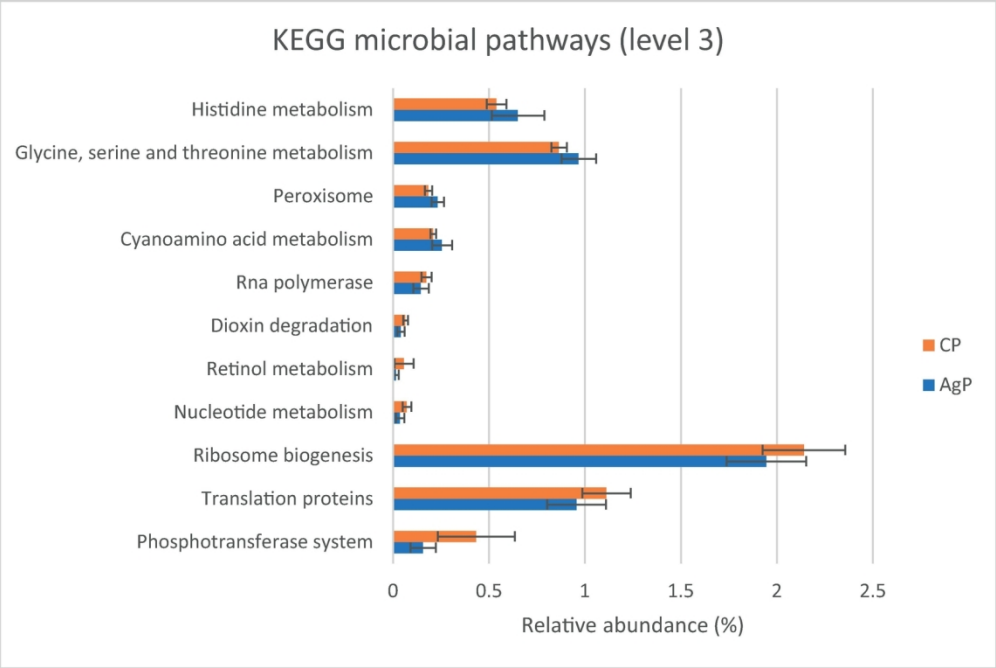
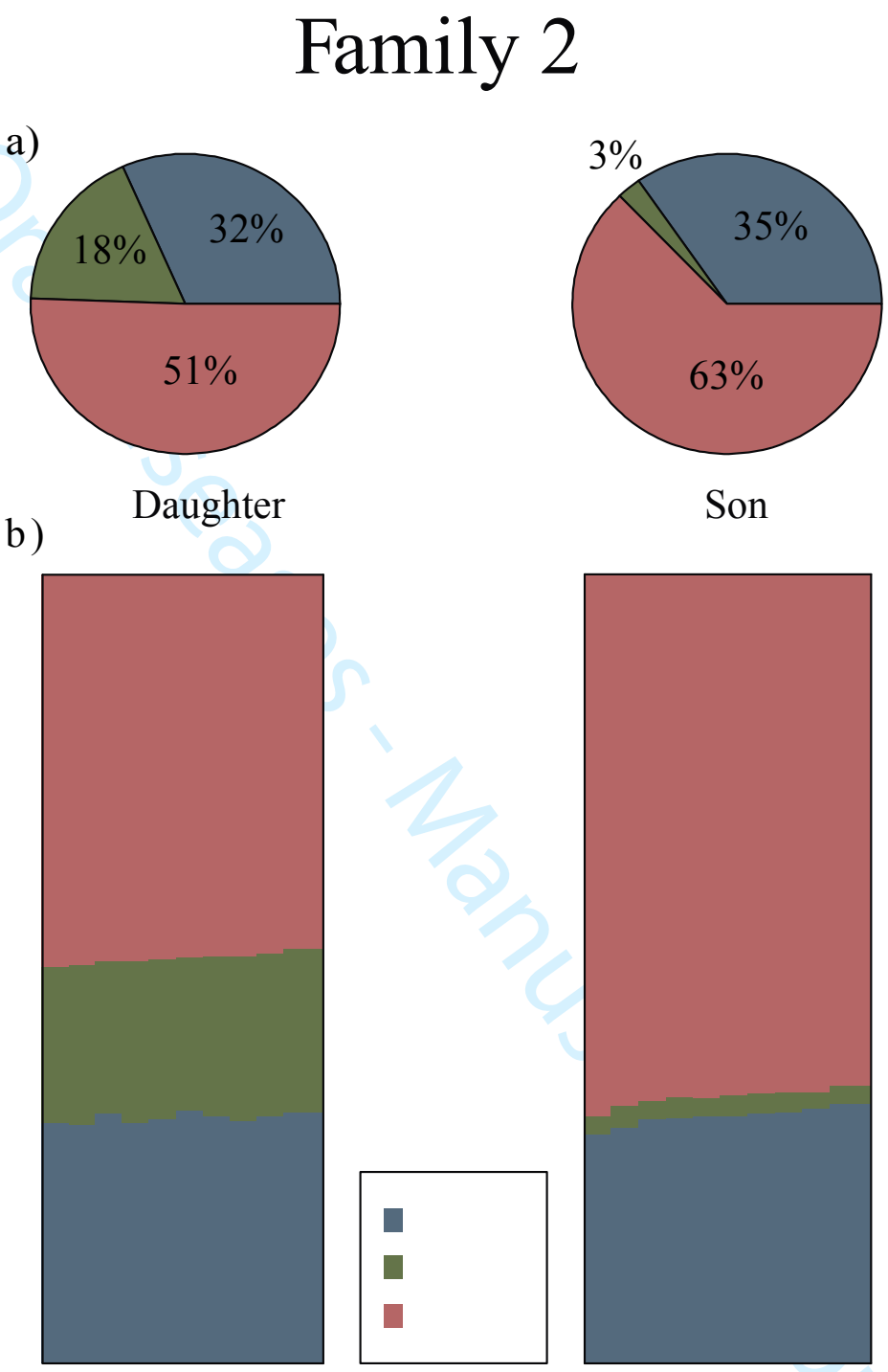
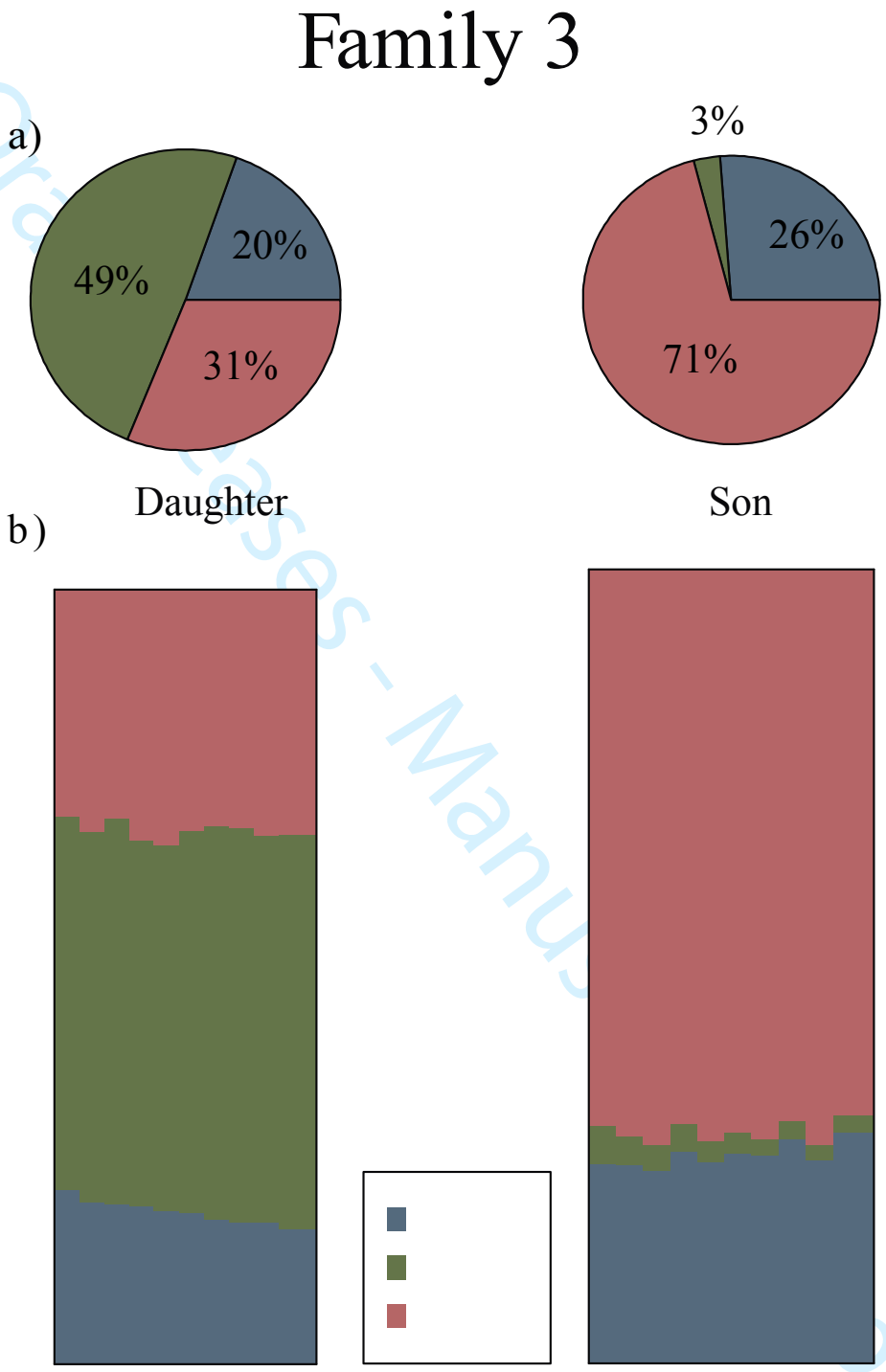


Figure 3. KEGG microbial pathways classified at level 3 differentially distributed between AgP and CP. Mean relative abundance and standard deviation for each condition are represented in horizontal columns and error bars, respectively. For clearer visualization, only the most abundant pathways have been included in the plot.

150x100mm (300 x 300 DPI)



Supplementary Figure 1. SourceTracker proportion estimates for family 2 (a,b). (a) Source environment mean proportions for 100 draws from Gibbs sampling. (b) Visualization of the 100 Gibbs draws; each column shows the mixture from one draw, with columns ordered to keep similar mixtures together. SourceTracker analysis was implemented in QIIME (Python 3.5).



Supplementary Figure 2. SourceTracker proportion estimates for family 3 (a,b). (a) Source environment mean proportions for 100 draws from Gibbs sampling. (b) Visualization of the 100 Gibbs draws; each column shows the mixture from one draw, with columns ordered to keep similar mixtures together. SourceTracker analysis was implemented in QIIME (Python 3.5).

Supplementary Table 1. Description of periodontal conditions in the enrolled subjects

Family 1	
Father	H*
Mother	AgP (Stage III, grade C)
Daughter	AgP (Stage III, grade C)
Son	H*
Family 2	
Father	CP (Stage III, grade A)
Mother	AgP (Stage III, grade C)
Daughter	CP (Stage II, grade B)
Son	AgP (Stage III, grade C)
Family 3	
Father	CP (Stage III, grade B)
Mother	CP (Stage III, grade B)
Daughter	AgP (Stage III, grade C)
Son	H

H: healthy subjects; AgP: patients suffering aggressive periodontitis; CP: patients suffering chronic periodontitis

*Amplification failed in these samples, possibly due to reduced amount of bacterial DNA in healthy subjects

Supplementary Table 2. SourceTracker estimation regarding the percentage of bacterial species transmitted from parents to offspring.

		Father	**Mother**	**Unknown**
Family 1	**Daughter**	Not available	0%	100%
Family 2	**Daughter**	32%	18%	51%
	Son	35%	3%	63%
Family 3	**Daughter**	20%	49%	31%
	Son	26%	3%	71%

SourceTracker analysis was implemented in QIIME (Python 3.5)

Supplementary Table 3. Total number of sequences per sample and relative abundance of bacterial genera included in the statistical analysis (see text). Abbreviations represent each respective family (F1: family 1, F2: family 2 and F3: family 3), and each respective member (S: son; D: daughter; M: mother; F: father).

Taxon	**F1D**	**F1M**	**F2S**	**F2D**	**F2M**	**F2F**	**F3S**	**F3D**	**F3M**	**F3F**
Total number of sequences	37439	44488	103179	38991	104341	104530	342	25028	21029	29257
Prevotella	4.9	6.8	9.1	4.5	0.7	16.7	3.2	25.8	19.5	22.7
Porphyromonas	1.1	4.1	13.5	5.7	0.0	13.8	3.2	4.7	32.6	6.9
Pseudomonas	0.0	78.3	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0
Neisseria	0.0	0.5	0.6	9.9	59.1	0.0	0.9	0.0	6.0	0.1
Streptococcus	1.2	1.8	0.8	8.4	18.7	1.4	22.8	0.4	0.5	11.1
*Rikenellaceae*_rc9_gut_group	57.1	1.0	1.0	0.2	0.0	0.2	0.9	2.9	0.8	0.5
Fusobacterium	0.4	0.1	7.9	12.8	0.1	4.0	3.2	20.1	5.2	9.4
Treponema	1.0	0.5	19.8	6.0	0.0	2.0	1.5	5.8	9.5	3.3
Fretibacterium	15.8	0.9	4.3	2.6	0.0	5.0	0.0	1.0	4.4	0.5
*Enterobacteriaceae*_unclassified	2.2	0.0	0.0	0.0	0.0	26.7	0.6	0.0	0.0	0.0
Veillonella	0.3	0.0	1.3	0.7	3.2	2.2	5.0	0.0	0.6	8.4
Rothia	0.0	0.0	0.1	1.6	0.5	0.0	18.4	0.0	0.0	0.1
Bacteroidales_unclassified	2.3	0.1	4.9	0.8	0.0	0.4	2.0	8.7	0.9	0.6
Bacteria_unclassified	0.9	0.5	4.6	2.8	0.5	0.8	0.0	3.1	2.3	2.5
Enterobacter	1.2	0.0	0.0	0.0	0.0	15.0	0.0	0.0	0.0	0.0
Clostridiales_unclassified	2.1	0.1	3.5	2.7	0.0	0.3	0.6	3.4	2.1	0.5
Alloprevotella	0.1	0.0	1.5	2.9	0.0	0.3	0.3	2.4	1.1	5.7
Capnocytophaga	0.1	0.0	2.9	3.2	5.7	0.2	1.5	0.3	0.3	0.1
Haemophilus	0.0	0.3	0.1	3.9	4.2	0.0	2.0	0.0	0.1	0.0
Tannerella	0.3	0.0	2.8	0.9	0.0	2.8	0.0	2.8	0.6	0.2
Campylobacter	0.0	0.0	1.4	4.5	0.2	0.3	0.9	1.3	0.5	0.9
Propionibacterium	0.0	0.0	0.0	0.0	0.0	0.0	9.9	0.0	0.0	0.0

Aggregatibacter	0.0	0.0	4.5	1.6	0.0	0.0	0.3	0.8	0.2	0.9
Filifactor	0.3	0.2	0.6	0.2	0.0	0.5	0.0	1.5	3.7	0.7
Actinomyces	0.0	0.0	0.1	1.9	0.1	0.1	3.8	0.2	0.1	1.1
Leptotrichia	0.0	0.0	0.6	1.2	1.4	0.2	0.9	0.7	0.2	1.1
Gemella	0.0	1.5	0.4	1.0	1.2	0.1	1.2	0.0	0.2	0.7
*Candidatus_Saccharibacteria*_unclassified	0.0	0.0	0.2	1.3	0.0	0.1	0.9	0.9	0.3	2.1
Parvimonas	2.7	0.2	2.1	0.2	0.0	0.2	0.0	0.1	0.1	0.2
Peptostreptococcus	0.1	0.2	2.8	0.4	0.0	0.7	0.0	1.2	0.3	0.3
Catonella	1.1	0.0	0.2	0.4	0.0	0.2	0.3	1.6	0.6	1.0
Selenomonas	0.4	0.0	0.6	1.0	0.0	0.6	0.3	0.4	0.4	1.4
*Lachnospiraceae*_unclassified	2.0	0.6	0.1	0.1	0.0	0.1	0.0	0.5	0.8	0.9
Lautropia	0.0	0.0	0.1	4.4	0.1	0.0	0.3	0.0	0.0	0.0
Escherichia/Shigella	0.0	0.0	0.0	0.0	0.0	0.0	4.4	0.0	0.0	0.0
Dialister	0.1	0.0	0.2	0.5	0.0	0.6	0.0	1.0	0.9	1.1
Corynebacterium	0.0	0.0	0.2	1.2	0.0	0.1	1.8	0.2	0.0	0.1
*Veillonellaceae*_unclassified	0.0	0.0	0.2	0.3	0.0	0.1	0.0	0.2	0.1	2.5
*Neisseriaceae*_unclassified	0.0	0.0	0.1	1.2	1.0	0.0	0.6	0.0	0.3	0.0
*Prevotellaceae*_unclassified	0.0	0.0	0.1	0.1	0.0	0.1	0.0	0.3	0.7	1.4
Odoribacter	0.2	0.0	0.0	0.0	0.0	2.0	0.0	0.5	0.0	0.0
Firmicutes_unclassified	0.1	0.0	0.3	0.3	0.1	0.1	0.0	0.3	0.7	0.6
Eikenella	0.0	0.0	0.3	1.1	0.5	0.0	0.0	0.4	0.0	0.0
Lachnoanaerobaculum	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.1	0.0	1.5
Morococcus	0.0	0.0	0.2	0.8	0.9	0.0	0.3	0.0	0.0	0.0
Mycoplasma	0.0	0.0	1.5	0.1	0.0	0.1	0.0	0.0	0.3	0.2
*Porphyromonadaceae*_unclassified	0.0	0.0	0.6	0.2	0.0	0.1	0.6	0.4	0.3	0.1
Megasphaera	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8
*Leptotrichiaceae*_unclassified	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.0	0.0	1.3
Schwartzia	0.0	0.0	0.0	0.4	0.0	0.1	0.0	0.1	0.7	0.2

Desulfobulbus	0.1	0.0	1.0	0.0	0.0	0.1	0.0	0.2	0.1	0.1
*Pasteurellaceae*_unclassified	0.0	0.1	0.1	0.9	0.1	0.0	0.0	0.0	0.1	0.1
Lactobacillus	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.1
Granulicatella	0.0	0.1	0.0	0.7	0.2	0.0	0.0	0.0	0.2	0.1
*Peptococcaceae_1*_unclassified	0.3	0.3	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0
Olsenella	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.5	0.0	0.6
Eubacterium	0.2	0.0	0.3	0.1	0.0	0.0	0.0	0.4	0.1	0.2
Kingella	0.0	0.0	0.5	0.2	0.4	0.0	0.0	0.0	0.0	0.0
Atopobium	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	1.0
Anaerovorax	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.5	0.3	0.2
*Propionibacteriaceae*_unclassified	0.0	0.0	0.1	0.9	0.0	0.0	0.0	0.0	0.0	0.0
Sneathia	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.4	0.3	0.0

Supplementary Table 4. KEGG microbial pathways classified at level 3 differentially distributed between AgP and CP.

KEGG microbial pathways (level 3)	Condition	*p*-value	LDA (log10)
Histidine metabolism	AgP	0.086	-2.757
Glycine, serine and threonine metabolism	AgP	0.086	-2.627
Peroxisome	AgP	0.014	-2.481
Cyanoaminoacid metabolism	AgP	0.050	-2.426
Linoleic acid metabolism	AgP	0.014	-2.369
Beta-lactam resistance	AgP	0.050	-2.287
Ethylbenzene degradation	AgP	0.086	-2.203
Betalain biosynthesis	AgP	0.049	-2.174
RNA polymerase	CP	0.086	2.208
Dioxin degradation	CP	0.086	2.230
Retinol metabolism	CP	0.050	2.346
Nucleotide metabolism	CP	0.050	2.368
Ribosome biogenesis	CP	0.086	2.938
Translation proteins	CP	0.086	3.051
Phosphotransferases system	CP	0.050	3.193

Each group is always compared with the other during the implementation of the algorithm. The condition for which each pathway is enriched is included under the "Condition" tab. *P*-value and effect size-related parameter LDA are included in the subsequent columns. Negative and positive values for LDA have been included to discern between pathways for AgP and CP.

Supplementary Table 5. KEGG orthologs showing statistical significance or statistical tendency within the most abundant and differentially distributed pathways for each condition.

KO_number	Mean (AgP)	Standard deviation (AgP)	Mean (CP)	Standard deviation (CP)	*p*-value*	KEGG pathway
K00128	6.086	6.844	1.426	1.121	0.086	Histidine metabolism
K01620	3.746	0.691	2.319	1.240	0.050	Glycine, serine and threonine metabolism
K02760	2.980	1.195	4.692	0.934	0.050	Phosphotransferases system
K02761	3.203	1.327	4.862	0.916	0.086	Phosphotransferases system
K02810	1.062	0.516	2.036	0.775	0.050	Phosphotransferases system
K02793	1.536	0.744	2.771	0.621	0.050	Phosphotransferases system
K02795	2.446	1.225	4.402	1.199	0.086	Phosphotransferases system
K02794	2.429	1.242	4.357	1.176	0.086	Phosphotransferases system
K02796	2.312	1.126	4.255	1.059	0.086	Phosphotransferases system
K02759	2.833	1.379	4.450	0.951	0.050	Phosphotransferases system
K02821	2.600	1.335	5.463	2.999	0.050	Phosphotransferases system

Mean relative abundance and standard deviation for AgP and CP, *p*-value and each respective KEGG pathway have been included in consecutive columns.
*Mann Whitney U test in SPSS v.20.0 (SPSS Inc., Chicago, IL, USA).