Salivary biochemistry of the healthy oral ecosystem
The studies presented in this thesis were funded by Top Institute Food and Nutrition, a public-private partnership in pre-competitive research. The research work was carried out at the department of Oral Biochemistry of the Academic Centre for Dentistry Amsterdam (ACTA), VU University of Amsterdam and University of Amsterdam, The Netherlands.

Printing of this thesis was financially supported by:

ISBN: 978-94-6332-085-6

Cover design by: GVO drukkers & vormgevers BV
Printed and bound by: GVO drukkers & vormgevers BV

Copyright © Andrei Prodan, Amsterdam 2016

All rights reserved. No part of this book may be reproduced, stored in retrievable systems, or transmitted in any form or by any means, mechanical, photo-copying, recording, or otherwise, without the prior permission of the holder of the copyright.
Salivary biochemistry of the healthy oral ecosystem

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Tandheelkunde
op diensdag 17 januari 2017 om 11:45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Andrei Prodan
geboren te Deva, Roemenië
promotor: prof.dr. E.C.I. Veerman

copromotoren: dr. H.S. Brand

prof.dr.ir. B.J.F. Keijser
### Table of Contents

**CHAPTER 1**
Introduction

**CHAPTER 2**
Interindividul variation, correlations, and sex-related differences in the salivary biochemistry of young healthy adults

**CHAPTER 3**
A study of the variation in the salivary peptide profiles of young healthy adults acquired using MALDI-TOF

**CHAPTER 4**
On the ecosystemic network of the salivary microbiome in healthy young adults

**CHAPTER 5**
Effect of experimental gingivitis induction and erythritol on the salivary metabolome and functional biochemistry of systemically healthy young adults

**CHAPTER 6**
Effect of experimental gingivitis induction and erythritol on the salivary MALDI-TOF-MS peptide profiles of systemically healthy young adults

**CHAPTER 7**
Summarizing discussion

Acknowledgements
CHAPTER 1

Introduction
The oral ecosystem in health and disease: homeostasis and dysbiosis

A healthy mouth is necessary for a number of important functions such as chewing, speech, taste and social interactions [1]. The human oral cavity displays a unique set of environmental conditions and is home to a complex community of microorganisms [1-5]. This resident microflora is constantly modulated by intrinsic, host-related factors (e.g. saliva, host immune cells) and extrinsic factors (e.g. food particles, xenobiotics, oral hygiene measures). The oral ecosystem is more stable compared to those from other ecological niches in or on the human body [6, 7]. Its microbial composition is better maintained across time under normal conditions and is more resilient to challenges (such as treatment with antibiotics) compared to the gut microbiome [6]. The oral microbiota shows less deviation and recovers faster than the gut, skin, or vaginal microflora [6, 7].

In spite of the aforementioned relative stability of the oral ecosystem, dysbiosis may occur and exert potentially severe negative effects on oral health. Dental caries and periodontitis – the most frequent oral diseases – do not fit the model of single microbe pathogenesis, but may rather be viewed as resulting from severe dysbiosis occurring within the mouth [4].

Dental caries is a multifactorial disease involving excessive demineralization and subsequent localized destruction of tooth hard tissue [8]. Cycles of demineralization and remineralization continuously occur under normal conditions, but certain factors may permanently upset this equilibrium and tilt the process towards an irreversible degradation of tooth structure (Fig. 1).
Foremost among cariogenic factors is a diet rich in easily fermentable carbohydrates (e.g., sucrose). It favors the selection of acidogenic and aciduric bacteria and therefore shifts the oral microbiome towards a cariogenic type [10]. The bacterial biofilm attached to the surface of the tooth is directly responsible for cariogenesis. When provided with fermentable sugars, it produces and releases acids, causing a localized drop in pH and accelerating demineralization. Poor oral hygiene is also a major risk factor [11]. Dental caries is among the most prevalent chronic diseases, and therefore has a significant detrimental effect on well-being and imposes a considerable economic burden [12, 13].

Periodontitis is a bacterial biofilm-induced chronic inflammatory disease [14]. It can eventually destroy the connective tissue and alveolar bone surrounding and supporting the teeth (i.e., the periodontium) and ranks as the number one cause of tooth loss worldwide [15]. The tissue damage is caused not only by the action of a dysbiotic bacterial biofilm, but also by the host immune reaction to the microbial challenge, which may in fact be the primary effector [15]. An increase in the production of inflammatory cytokines disrupts bone homeostasis and is a major contributor to the loss of periodontium [14, 15]. Periodontitis is often preceded by gingivitis, a reversible inflammation of the gingiva caused by the accumulation of supragingival plaque [16].
It can be ascertained that a healthy mouth depends on oral host-microbiome homeostasis. Homeostasis is defined by an intricate network of relationships and interdependencies. The next chapters will examine some of the main components of the system (i.e. salivary proteins and peptides, the oral microbiome, and the oral metabolome), determine their boundaries, and describe the interactions between them when at health.

**Saliva and oral health**

Saliva is a mucoserous fluid secreted by three pairs of major salivary glands (parotid, submandibular and mandibular), as well as by minor glands on the lower lip, tongue, palate and cheeks [17-19]. Saliva production begins within specialized acinar cells contained inside the glands, after which salivary secretions are transported and further modified along the salivary ducts [18]. Whole saliva is a complex mixture composed of secretions from the salivary glands, gingival crevicular fluid, food debris, and components derived from oral microflora as well as from exfoliating oral mucosal cells and host immune cells [18, 20].

Saliva is essential for oral health. This can be inferred by observing the consequences of insufficient salivary production. Salivary secretion impairment may be caused by Sjögren’s syndrome (an autoimmune condition), radiation therapy, but it most commonly appears as a side-effect of xerostomic medications [21]. Regardless of etiology, the effect is highly damaging: oral soreness and discomfort, loss of taste, difficulty in swallowing, increased caries, candidiasis [21].

One of the primary roles of saliva is to minimize and repair acid-induced demineralization of the tooth enamel [18]. The hydroxyapatite component of dental enamel begins to demineralize when the pH of saliva drops below 5.0 – 5.5. Saliva minimizes the effect of acid challenges through four different buffer systems: bicarbonate/carbonate, phosphate, urea, and via the side chains of the salivary proteins [1, 18]. However, the main buffering mechanism in saliva is the bicarbonate/carbonate system [22]. Salivary urea is metabolized by oral bacteria with production of ammonia and subsequent pH increase [23, 24]. Saliva is supersaturated in calcium and phosphate ions which remineralize the enamel when the salivary pH rises back to normal values and are crucial for the long-term preservation of the tooth integrity [25].

Saliva bathes and lubricates oral surfaces, providing both protection against mechanical wear and a superficial layer shielding against acids, toxins and bacterial proteases [26, 27]. The main component enabling this function is mucin MUC5B, a very large (>1 MDa), oligomeric and highly glycosylated protein, with covalently attached sugar moieties comprising 80% of its molecular mass [1, 22]. It is a major component of whole unstimulated saliva, making up 20-30% of the total protein content [22]. The functionality of MUC5B is directly linked to its structure. MUC5B’s large size and filament-like form, together with a large hydrophilic carbohydrate component, are responsible for the
characteristic viscoelastic, coating, and lubricant properties of saliva [1]. The other salivary mucin, MUC7, is smaller (around 200 kDa), monomeric, but also rich in carbohydrate side chains. MUC7 contributes to the aggregation and clearance of bacteria, as it is able to bind to a variety of oral bacteria [1, 22].

Saliva actively aids digestion. Amylase is one of the most abundant proteins in saliva, and its primary function is to initiate the digestion of starch [1, 28]. However, amylase is also a component of the dental pellicle and mediates bacterial adhesion to the tooth surface [28, 29].

Saliva contains a number of antimicrobial proteins and enzymes [1]. Secretory-Immunoglobulin A is the predominant immunoglobulin in saliva, where it binds and aggregates bacteria, viruses, and toxins [18, 30]. Lactoferrin is a multifunctional protein: it sequesters iron – a limiting factor for bacterial growth – thus providing an important bacteriostatic effect, but also releases bactericidal peptides when subjected to proteolysis, and may have other iron-independent antimicrobial functions [1, 31]. Lysozyme is an enzyme that hydrolyzes peptidoglycan, thus cleaving Gram-positive bacterial cell walls leading to cell lysis and death [1, 18, 32]. It is also thought to have non-enzymatic killing effects on Gram-negative bacteria [33]. The enzyme chitinase has a similar effect on yeast cell walls [34]. Cystatins, though not antimicrobial, may have a protective effect by inhibiting the damaging action of some bacterial proteases [35]. Many of the salivary antimicrobial proteins and enzymes act through several different mechanisms and have not only additive, but also synergistic effects, forming a complex defensive functional network [1].

Salivary albumin is not secreted by the glands, but originates from plasma via gingival crevicular fluid and mucosal leaking. It is therefore a useful marker for oral lesions and bleeding.

**Saliva as a research medium: opportunities and caveats**

As a research medium, saliva presents a number of advantages. Salivary collection is quick, simple, non-invasive, and poses no risk of adverse effects for the donor [36]. Indeed, several studies have focused on saliva in order to develop biomarkers and diagnostic tests for diseases ranging from periodontal disease and Sjögren’s syndrome to oral and breast cancers [36, 37]. Saliva is a prime candidate as a medium for the study of oral health, due to its capacity to supply quick and accessible biological information derived from multiple levels of the oral ecosystem.

However, using saliva as a research and diagnostic material involves some particular challenges. Salivary secretion is a process influenced by both the sympathetic and the parasympathetic branches of the nervous system [18, 19]. Salivary flow rate and composition can be drastically altered by stimuli such as mechanical movement (chewing,
Chapter 1

speaking) or gustatory triggers [19]. Moreover, other factors such as age, various medications, length of time elapsed since the last stimulation, and phase of the circadian rhythm can also have a significant effect [38-41]. It is therefore very important to carefully control all these factors at the time of saliva sampling in order to obtain meaningful and reproducible data.

Another caveat lies within the sequence of procedures used to process and store saliva prior to analysis [20, 42, 43]. The varying turbidity of sampled saliva caused by the presence of cellular debris and potential food residues means that centrifugation is generally used to clarify saliva and remove these unwanted components as well as other insoluble aggregates. The exact conditions of clarification may influence salivary composition [20, 42].

Finally, saliva is prone to rapid changes and degradation if improperly processed and stored. The presence of numerous proteases in saliva makes it inherently unstable at room temperature. It is therefore crucial that saliva samples are collected and processed on ice and subsequently frozen at -80°C as quickly as possible [20, 42].

**Oral bacteria and oral health**

Although the oral cavity is home to bacteria, archaea, fungi, protozoa, and viruses, most research has generally focused on the bacteria, with far less information available on other microorganisms [2, 44]. In this thesis, ‘microbiome’ and ‘microflora’ are subsequently used to refer strictly to bacteria.

There are currently around 700 species/phylotypes of oral bacteria catalogued in the Human Oral Microbiome Database (www.homd.org) [45]. Depending on the methodology, the actual total number of species is estimated to be several thousand [46]. Although there is substantial variation between individuals with regard to the composition of the oral microbiome, studies have found that some taxa form an oral ‘core microbiome’ [3, 7, 47]. These predominant taxa include Firmicutes (e.g. genus *Streptococcus, Veillonella*), Proteobacteria (e.g. *Neisseria*, *Haemophilus*), Actinobacteria (*Corynebacterium*, *Actinomyces*), Bacteroidetes (*Prevotella, Porphyromonas*), and Fusobacteria (*Fusobacterium*) [3].

Commensal oral microflora play an important part in maintaining oral health [2, 48]. First of all, they inhibit colonization by pathogens, utilizing several mechanisms for colonization resistance [2]. They also convert dietary nitrate to nitrite, which is subsequently converted to nitric oxide, an anti-hypertensive molecule important for vascular health [49].

The composition and behavior of oral bacteria is shaped by the conditions inside the oral cavity: constant presence of liquid water (from saliva), extreme short-term temperature fluctuations, and wide variations in the availability of carbon and nitrogen
substrates [50]. An important characteristic of oral bacteria is their tendency to form biofilms (e.g. dental plaque) [9, 51]. Oral biofilms are species-diverse and structurally and functionally organized [9]. The biofilms form in a sequence of stages: pellicle formation (from salivary components), reversible adhesion of initial colonizers (Streptococcus gordonii, Streptococcus mitis, Streptococcus oralis, Streptococcus sanguinis), followed by stronger, adhesion/receptor-mediated bacterial attachment [9, 50]. Subsequently, secondary colonizers also attach and exopolysaccharides are synthesized, strengthening the biofilm structure. There is extensive metabolic interaction, cell-signaling and cross-talk between the different bacterial species within a biofilm [52]. Fragments of the biofilm may detach from the surface and get swept up by saliva [9].

Oral bacteria use saliva as a primary nutrient substrate [53]. The major salivary glycoproteins (e.g. mucins) are an important source of carbon. Although the sugar groups on these proteins are highly diverse and thus require an extensive enzymatic apparatus in order to be fragmented, oral bacteria collaborate and work sequentially in order to catabolize them [53, 54]. The overall composition of available substrates and the functional profile of the bacterial communities metabolizing these substrates determine a distinctive signature of small molecules in saliva – the salivary metabolome.

The oral cavity contains several ecological sub-niches (e.g. teeth, tongue, cheeks, hard and soft palates), either on the surfaces of the teeth or on mucosa [3]. Each sub-niche may offer different conditions (pH, oxygen and nutrient availability, redox potential) and select for a specific type of bacterial community [3, 51]. In turn, the metabolic activity of the resident microbiota can modify conditions at the respective site [55]. Saccharolytic bacteria in supragingival sites produce acids and lower local pH. Non-mutans streptococci and Actinomyces dominate in healthy supragingival plaque. Some conditions (e.g. high frequency of carbohydrate exposure) can select for more acid-tolerant bacteria and more efficient acid producers such as Streptococcus mutans, lactobacilli or Bifidobacterium, disturbing enamel mineralization homeostasis and leading to irreversible demineralization and dental caries [9, 51]. On the other hand, asaccharolytic bacteria in subgingival sites secrete nitrogenous catabolites derived from gingival crevicular fluid, thereby raising local pH [55]. Fusobacterium and Prevotella in subgingival plaque increase pH and stimulate the flow of gingival crevicular fluid [55]. This favors acid-intolerant, proteolytic species associated with periodontitis [55].

The oral microbiome is characterized by high taxonomic diversity, complex networks of inter-bacterial and bacterial-host interactions, and functional complementarity and redundancy. These traits may partly explain the remarkable stability and resilience of the oral ecosystem [44]. However, the mechanisms associated with homeostasis and therefore with oral health have yet to be fully understood [44].
Chapter 1

Objectives

This thesis aimed to describe the boundaries and interrelationship between the salivary proteins, the salivary microbiome, and the salivary metabolome of a healthy oral ecosystem, and to examine the changes occurring within these levels of the ecosystem when a challenge is applied (i.e. experimental gingivitis). The overall purpose was to increase our understanding of the oral ecosystem and to gain insights into the processes involved in maintaining oral health.
The studies encompassed in this thesis were performed as part of the ‘Novel strategies to promote oral health’ project within the framework of TI Food and Nutrition. Data were acquired from two separate clinical studies: a cross-sectional observational study aimed at estimating the boundaries of a healthy oral ecosystem (Chapters 2, 3, and 4) and a challenge intervention, randomized clinical trial exploring the dynamic interactions in the oral ecosystem during the induction of mild gingival inflammation (Chapters 5 and 6).
Chapter 1

References


Introduction


Interindividual variation, correlations, and sex-related differences in the salivary biochemistry of young healthy adults


Abstract

A cross-sectional observational study was conducted to evaluate interindividual biochemical variation in unstimulated whole saliva in a population of 268 systemically healthy young students aged 18-30 yr (150 males and 118 females) with no of apparent caries lesions or periodontal disease. Salivary flow rate, protein content, pH, buffering capacity, mucins MUC5B and MUC7, albumin, secretory-IgA, cystatin S, lactoferrin, chitinase, amylase, lysozyme and proteases were measured using Enzyme-Linked Immunosorbent Assays (ELISA) and enzymatic activity assays. Significant differences were found between males and females. Salivary pH, buffering capacity, protein content, MUC5B, secretory-IgA, and chitinase activity were all lower in females compared to males, while MUC7 and lysozyme activity were higher in females ($p < 0.05$). There was no significant difference between sexes in salivary flow rate, albumin, cystatin S, amylase and protease activity. Principal Component Analysis (PCA) and Spectral Clustering (SC) were used to assess inter-variable relationships within the dataset and to identify subgroups. SC analysis identified 2 clusters of participants, which were subsequently described. This study provides a comprehensive overview of the distribution and interrelations of a set of important salivary biochemical variables in the young systemically healthy young adult population screened for the absence of apparent caries lesions and periodontal disease. It highlights significant gender differences in salivary biochemistry.
Introduction

Saliva is crucial for the maintenance of oral health. It initiates the digestion of carbohydrates, provides lubrication as a countermeasure to tooth wear, stabilizes oral pH and helps the re-mineralization of enamel [1, 2]. Saliva prevents pathogenic microorganisms from colonizing the mouth and subsequently causing disease [2-4]. It aggregates and clears planktonic bacteria, whereas the protective coating it forms on tooth surfaces may serve as the foundation for oral biofilm (plaque) formation [1, 3]. Many of the individual proteins and biochemical processes that mediate these functions have been thoroughly studied [1]. However, salivary components may perform multiple, sometimes overlapping tasks and form a functional network characterized by interactions and redundancies [5-7]. No single salivary constituent has a decisive effect on oral health by itself. It is rather the properties of the salivary functional network as a whole that determine whether the oral ecosystem maintains homeostasis or shifts towards a disease state [2].

The starting point in understanding the action of saliva as a biological system in relation to oral health is to investigate what the constitution of this system is when it is functioning properly: i.e. the salivary biochemistry of individuals with healthy mouths. Yet, despite extensive salivary research over the last decades, this work has not resulted in a clear picture [8]. There are many factors that can influence the results of clinical studies focusing on saliva analysis. This makes it difficult to compare and compile the results of separate studies into an integrated model. Age and sex differences of the participants, (oral) health status, medication, physical activity and level of oral self-care can all influence salivary composition [2, 9, 10]. Even the circadian cycle - and therefore the time of day when saliva is collected - has a significant effect [11, 12]. Sample processing and storage conditions can also affect results. Saliva is predisposed to proteolysis and salivary protein aggregation may occur during sample freezing and thawing [13, 14]. Centrifugation, often used to remove insoluble aggregates and oral bacteria, can also remove some of the salivary proteins to a varying extent [15].

The primary aim of this study was to provide a detailed description of the distribution, interindividual variation and interrelations of a broad set of salivary proteins and other biochemical variables in a systemically healthy young adult population. A secondary aim was to identify and visualize subgroups of individuals based on their salivary protein biochemistry. Unstimulated whole saliva was collected from 268 young adults screened for good oral health. The levels of ten major salivary proteins and enzymes with known relevance for oral health were measured, as well as salivary flow rate, pH, buffered pH and total protein content.
Materials and Methods

Clinical study structure, study population and exclusion criteria

The study was carried out within the framework of the Top Institute Food and Nutrition project "Estimating the boundaries for a healthy oral ecosystem in young individuals".

Whole unstimulated saliva was collected in a cross-sectional single-center observational clinical study at the Academic Center for Dentistry Amsterdam (ACTA). The study population comprised a convenience sample of systemically healthy young adults aged 18-30 yr old without periodontitis. Participants were students of universities and colleges in and around Amsterdam, The Netherlands. They were invited for screening when they had visited their general dentist the previous year and had been considered to be without oral or dental problems. The volunteers were screened for suitability according to the criteria as proposed for the Dutch Periodontal Screening Index (DPSI) [16]. They were included if they had a DPSI ≤ 3-. The exclusion criteria were: presence of overt dental caries, inter-proximal restorations between the first and second, or second and third upper molars, apparent oral lesions, a habit of smoking, infections, recent use of antibiotics, use of anti-inflammatory drugs or other prescribed medication which could interfere with the outcome of this study (except for oral contraceptives). Participants were instructed not to eat, drink, chew gum or perform strenuous physical exercise prior to the appointment, and not to brush their teeth in the morning of the appointment or the evening before.

The menstrual cycle phase (menstrual, proliferative, follicular, luteal or secretory) was noted for female participants at the date of saliva sampling.

The study was conducted in accordance with the Declaration of Helsinki (2008) of the World Medical Association and approximating Good Clinical Practice guidelines. The study protocol was reviewed and approved by the Medical Ethics Committee of the Academic Medical Centre of Amsterdam (2012_210#B2012406) and registered at the Dutch Trial Register (NTR3649). All participants signed an informed consent form.

Unstimulated saliva sampling and pH measurement

All saliva samples were collected between 9 and 10 a.m. Participants were instructed to allow saliva to accumulate in the floor of the mouth without stimulation such as orofacial movements and to spit at 30 s intervals into pre-weighed 30-ml polypropylene tubes (Sterilin, Newport, U.K.) which were kept on ice. The collection period was 5 min. The tubes containing the unstimulated saliva samples were weighed and salivary flow rate was calculated assuming a saliva density of 1.0 g ml\(^{-1}\). Output was calculated as flow rate
multiplied by protein concentration. Samples were homogenized by vortexing for 20 s. Salivary pH was measured immediately after acquisition and homogenizing using a Eutech pH 5+ microelectrode pH meter (ThermoScientific, West Palm Beach, U.S.A.). The pH buffering capacity of saliva was assayed as follows: 20 µl of sample were mixed with 40 µl of 0.005 N HCl, briefly vortexed and incubated at room temperature for 30 s before measurement of the buffered pH [17]. Samples were clarified by centrifugation for 10 min at 4°C and 10,000 g to remove epithelial cell debris, bacteria and food residues. The resulting clarified saliva was diluted 1:1 with a 500 mM NaCl solution to a final concentration of 250 mM NaCl, aliquoted and stored at -80°C. It had been verified beforehand that diluting saliva with a NaCl solution would not interfere with ELISAs or enzyme activity assays. The dilution prevented protein aggregation and precipitation during saliva freezing and storage and lowered viscosity allowing for more precise sample manipulation and improved reproducibility (PRODAN et al. manuscript in preparation). A sufficient number of aliquots of each sample were produced to avoid exposing samples to multiple freezing and thawing cycles.

**Total protein content determination**

Total protein content was measured with a Pierce BCA Protein Assay Kit (ThermoScientific, West Palm Beach, USA) in polystyrene 96-well microplates (Greiner Bio-One, Friechhausen, Germany) according to the manufacturer's specifications. Bovine serum albumin (BSA) was used as a standard. Optical readouts for the BCA assay and for all ELISAs performed in this study were obtained with a Multiscan FC microplate photometer (ThermoScientific, West Palm Beach, U.S.A.).

**Mucins MUC5B and MUC7 quantification**

High-binding polystyrene 96-well microplates (Greiner Bio-One, Frieckenhausen, Germany) were used for all ELISAs. For mucin ELISA, 100 µl of each saliva sample was diluted 1:100 and coated onto microplates by incubating overnight at 4°C. Each sample was 2-fold serially-diluted in coating buffer (0.1 M Na₂CO₃, pH = 9.6). MUC5B levels were then determined using mAb F2, as previously described [18]. MUC7 levels were determined using a polyclonal rabbit anti-MUC7 Ab [19]. All microplates contained a reference sample consisting of pooled saliva from 10 volunteers. The concentration of MUC5B and MUC7, respectively, in the reference sample was expressed as 1 Arbitrary Unit (AU) ml⁻¹.
**Chapter 2**

**Albumin quantification**

Microplates were coated with a rabbit polyclonal anti-human albumin Ab (Sigma-Aldrich, St. Louis, U.S.A.) overnight at room temperature [9]. Samples were subsequently 2-fold serially-diluted and incubated for 2 h at 37°C. Captured salivary albumin was detected using a HRP-conjugated rabbit anti-human albumin Ab (GeneTex, Inc., Irvine, CA, U.S.A.). Human serum albumin (Sigma-Aldrich, St. Louis, U.S.A.) was used as a standard.

**Lactoferrin quantification**

Microplates were coated with polyclonal rabbit anti-human lactoferrin Ab (Sigma-Aldrich, St. Louis, U.S.A.). Purified human lactoferrin (Sigma-Aldrich, St. Louis, U.S.A.) was used as a standard. Captured lactoferrin was assayed with a HRP-conjugated rabbit anti-human lactoferrin Ab (RayBiotech, Norcross, GA, U.S.A.) [20].

**Secretory-IgA quantification**

Microplates were coated overnight at room temperature with monoclonal rabbit anti-human secretory-IgA Ab (Sigma-Aldrich, St. Louis, U.S.A.). The microplates were then incubated with the saliva samples and a standard of purified human secretory-IgA (Nordic-MUbio, Susteren, The Netherlands) for 2 h at 37°C. Captured secretory-IgA was detected with HRP-conjugated goat anti-human IgA Ab (Sigma-Aldrich, St. Louis, U.S.A.) [21].

**Cystatin S quantification**

Samples were coated directly onto microplates overnight at 4°C together with a pooled saliva reference sample. Cystatin S was detected with a monoclonal mouse anti-human cystatin S Ab developed and characterized as described by HENSKENS et al. [22]. Captured antibodies were detected with polyclonal HRP-conjugated rabbit anti-mouse Ab (Sigma-Aldrich, St. Louis, U.S.A.).

**Enzymatic activity assays**

Black 96-well polypropylene microplates (Greiner Bio-One, Frieckenhausen, Germany) were used for all fluorescence-based enzymatic activity assays. Amylase activity was measured with an EnzChek Ultra Amylase Assay kit (ThermoScientific, West Palm Beach, U.S.A.) according to the manufacturer's specifications. Chitinase activity was quantified by adding 50 µl saliva to a substrate solution of 4-methylumbelliferyl β-D-N,N′,N″-
triaectylchitotrioside (Sigma-Aldrich, St. Louis, U.S.A.) to a final reaction volume of 200 µl and a substrate concentration of 40 nM. The increase in fluorescence was subsequently acquired for 15 min at 37°C [23]. Lysozyme activity was measured using an EnzChek Lysozyme Activity kit (ThermoScientific, West Palm Beach, U.S.A.) according to manufacturer's specifications.

Protease activity was measured based on the cleavage of two Fluorescence Resonance Energy Transfer (FRET) substrates designated BikKam9 and BikKam15, developed and previously described in detail by KAMAN et al. [24]. The BikKam9 substrate is highly specific for proteases of the periodontal pathogen Porphyromonas gingivalis, whereas BikKam15 was used to assess the overall protease activity in saliva. Fluorescence was recorded at 2 min intervals in a BMG Fluostar Galaxy microplate reader (MTX Lab Systems, Inc., Vienna, U.S.A.).

All ELISAs and enzymatic activity assays were performed in duplicate.

Data analysis and statistical methods

The data were statistically analyzed and Principal Component Analysis (PCA) was performed using SPSS 21.0 software (IBM, Armonk, NY, U.S.A.). The normality of the distributions of salivary biochemical variables was assessed with one-sample Kolmogorov-Smirnov tests as well as by visual examination of the individual variable distribution histograms. In case of a non-normal distribution non-parametric tests were chosen for subsequent analysis. Spearman's test was used for correlations and Mann-Whitney U tests were used to compare distributions. The statistical significance level used was 0.05. The Benjamini-Hochberg False Discovery Rate (FDR) was used to correct for multiple comparisons [25]. The FDR was set at 0.05.

Spectral Clustering

Spectral Clustering (SC) analysis was performed on the final dataset containing the data from all biochemical assays. SC was performed in open-source Python 2.7 using the neighborhood co-regularized SC algorithm developed by TSIVTSIVADZE et al. [26] based on the SC method published by VON LUXBURG [27]. For this purpose, the data from each biochemical assay were scaled to equal ranges and a similarity matrix was calculated based on the Euclidean distances between each pair of participants (i.e. on the similarity of the overall salivary biochemistry of each pair of participants). A co-occurrence matrix was subsequently calculated based on the clustering results, quantifying the tendency of any two participants to fall within the same cluster over many k-means clusterings using varying parameters. Matlab R2012b (MathWorks, Natick, MA, U.S.A) was used to produce the co-occurrence clustering plots. After visual examination of the SC plots the number of clusters
Chapter 2

was determined and participants were mathematically assigned to the clusters using a probabilistic decomposition algorithm [28].
Results

Prior to the study, 336 potential participants were screened at the clinic in a separate session. Of these, 46 (23 males and 23 females) were excluded based on the inclusion and exclusion criteria (14 subjects had interproximal dental restorations, 13 used medications which could interfere with the outcome of this study, nine subjects showed evidence of overt caries, four subjects had a DPSI score of $\geq 3^+$, three subjects did not visited their general dentist within the previous year, one presented an oral piercing, one subject was a smoker and one had a schedule conflict), 10 took part in a pilot study (not included in the final data) and 12 dropped out (11 due to schedule conflicts and one reported pregnancy). In total, 268 participants completed the study (150 males, 118 females), with a mean age of 22.6 yr (on the day of their appointment) and a range of 18-32 yr. Two of the subjects were unable to provide any saliva during the sampling period. Data obtained from the remaining 266 subjects (148 male and 118 female) is presented in this article. Due to insufficient amounts of saliva obtained from some participants several assays could not be performed on all samples. Across the whole study, the average percentage of samples available for analysis for an individual assay was 97%, with a minimum of 90% for the MUC7 measurement.

Table 1 provides a summary of the mean outcomes of the salivary biochemistry analyses. With the exception of salivary pH, the distributions of all biochemical variables showed significant deviations from normality. This inference was supported by visual inspection of the distribution histograms of each variable. There was no measurable cleavage of the Bikkam9 protease substrate by proteases specific to P. gingivalis in any of the samples. Therefore, only the results of the general salivary protease activity assay are presented.
Table 1. Salivary biochemistry results for the study population.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Median</th>
<th>Mean ± SD (N)*</th>
<th>95% Confidence Interval for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td></td>
<td>7.05</td>
<td>7.04 ± 0.28 (256)</td>
<td>7.00 - 7.07</td>
</tr>
<tr>
<td>Buffered pH</td>
<td></td>
<td>6.30</td>
<td>6.16 ± 0.63 (256)</td>
<td>6.09 - 6.24</td>
</tr>
<tr>
<td>Flow rate</td>
<td>ml min⁻¹</td>
<td>0.34</td>
<td>0.38 ± 0.22 (266)</td>
<td>0.37 - 0.41</td>
</tr>
<tr>
<td>Protein content</td>
<td>mg ml⁻¹</td>
<td>0.92</td>
<td>0.99 ± 0.34 (264)</td>
<td>0.95 - 1.03</td>
</tr>
<tr>
<td>MUC5B</td>
<td>AU ml⁻¹</td>
<td>1.0</td>
<td>1.2 ± 1.1 (265)</td>
<td>1.2 - 1.4</td>
</tr>
<tr>
<td>MUC7</td>
<td>AU ml⁻¹</td>
<td>3.8</td>
<td>5.8 ± 6.3 (241)</td>
<td>5.0 - 6.6</td>
</tr>
<tr>
<td>Albumin</td>
<td>µg ml⁻¹</td>
<td>20</td>
<td>37 ± 54 (266)</td>
<td>30 - 43</td>
</tr>
<tr>
<td>Secretory-IgA</td>
<td>µg ml⁻¹</td>
<td>261</td>
<td>305 ± 186 (266)</td>
<td>283 - 328</td>
</tr>
<tr>
<td>Cystatin S</td>
<td>AU ml⁻¹</td>
<td>1.0</td>
<td>2.5 ± 3.6 (265)</td>
<td>2.1 - 2.9</td>
</tr>
<tr>
<td>Lactoferrin</td>
<td>µg ml⁻¹</td>
<td>0.9</td>
<td>1.0 ± 0.7 (246)</td>
<td>1.0 - 1.1</td>
</tr>
<tr>
<td>Chitinase activity</td>
<td>AU ml⁻¹</td>
<td>19</td>
<td>32 ± 40 (263)</td>
<td>27 - 37</td>
</tr>
<tr>
<td>Amylase activity</td>
<td>U ml⁻¹</td>
<td>27</td>
<td>41 ± 34 (266)</td>
<td>37 - 45</td>
</tr>
<tr>
<td>Protease activity</td>
<td>AU ml⁻¹</td>
<td>44</td>
<td>47 ± 26 (259)</td>
<td>44 - 50</td>
</tr>
<tr>
<td>Lysozyme activity</td>
<td>U ml⁻¹</td>
<td>1505</td>
<td>1650 ± 1260 (246)</td>
<td>1492 - 1809</td>
</tr>
</tbody>
</table>

*N is the number of samples assayed for the respective variable out of the total sample population of 268 participants.

Principal Component Analysis (PCA) performed on the salivary biochemistry dataset indicated that salivary total protein content was the main contributor to the 1st principal component (PC1) (Fig. 1). Indeed, most of the salivary proteins assayed correlated to some extent with total protein content, as shown in Table 2. However, the strength of these correlations varied from ρ = 0.73 for secretory-IgA, to a not significant correlation for cystatin S (Table 2). There was a significant negative correlation between salivary flow rate and total protein content (ρ = -0.47, \( p < 0.001 \)). These two relationships were mirrored in the negative correlations between flow rate and most individual protein concentrations. With the exception of MUC7, the negative correlations between individual proteins and flow rates were smaller than the corresponding positive correlations between the respective proteins and total protein content (Table 2).
Figure 1. PCA loadings plot. PC1 and PC2 account for 35% and 14% of total variance, respectively (49% cumulative).

Table 2. Spearman’s rank correlation coefficients between individual salivary proteins, total protein content and flow rate, respectively. Variables are arranged in decreasing order of the strength of their correlation with total protein content.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Total protein content</th>
<th>Flow rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secretory-IgA</td>
<td>0.73</td>
<td>-0.64</td>
</tr>
<tr>
<td>Lactoferrin</td>
<td>0.61</td>
<td>-0.34</td>
</tr>
<tr>
<td>Albumin</td>
<td>0.52</td>
<td>-0.38</td>
</tr>
<tr>
<td>Amylase activity</td>
<td>0.48</td>
<td>NS*</td>
</tr>
<tr>
<td>Protease activity</td>
<td>0.44</td>
<td>-0.31</td>
</tr>
<tr>
<td>MUC5B</td>
<td>0.34</td>
<td>-0.21</td>
</tr>
<tr>
<td>Lysozyme activity</td>
<td>0.32</td>
<td>NS</td>
</tr>
<tr>
<td>Chitinase activity</td>
<td>0.29</td>
<td>-0.24</td>
</tr>
<tr>
<td>MUC7</td>
<td>0.29</td>
<td>-0.38</td>
</tr>
<tr>
<td>Cystatin S</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

*NS – not significant, $p > 0.05$. All other correlations shown have $p < 0.001$. 
Chapter 2

To assess intrinsic relationships between salivary proteins independent of the effect of total protein content and flow rate, a partial correlation analysis was performed in which the two aforementioned variables were controlled for. This revealed positive partial correlations between MUC7 and lysozyme ($\rho = 0.43$), and albumin and chitinase ($\rho = 0.55$), as well as a negative correlation between chitinase and lysozyme ($\rho = -0.46$) (all with $p < 0.001$). These relationships are supported by the PCA loadings plot (Fig. 1). Albumin and chitinase are grouped close together, whereas chitinase and lysozyme are situated on opposite regions of the plot, separated by the 2nd principal component. There was also a negative correlation between lysozyme activity and salivary pH ($\rho = -0.44$, $p < 0.001$), illustrated in Fig. 2. This correlation was maintained when it was controlled for the effects of flow rate and protein content ($\rho = -0.48$, $p < 0.001$).

Figure 2. Scatter plot of lysozyme activity versus salivary pH. Upper and lower lines indicate 95% confidence interval.

Several significant differences were observed when the distributions of salivary components in males were compared to those of females (Tables 3 and 4). Marked differences were seen in the distribution of salivary pH and buffered pH values (Fig. 3). Salivary pH values were normally distributed for both sexes, but the mean salivary pH of females was 0.23 pH units lower compared to males, while the mean buffered pH was 0.50 units lower. Also, more females had buffered pH values in the extreme low range: 15 females had buffered pH below 5.0, compared to only 4 males. There were no significant differences in salivary biochemistry between groups of women at different phases of their menstrual cycle, when compared using a one-way ANOVA test.
### Table 3. Salivary variables stratified according to gender. Percentage differences are shown for variables with distributions in males found to be significantly different compared to females using a Mann-Whitney U test.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Median Male</th>
<th>Median Female</th>
<th>Mean ± SE Male</th>
<th>Mean ± SE Female</th>
<th>Percentage difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td></td>
<td>7.13</td>
<td>6.94</td>
<td>7.13 ± 0.02</td>
<td>6.90 ± 0.02</td>
<td>0.23*</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Buffered pH</td>
<td></td>
<td>6.51</td>
<td>6.07</td>
<td>6.40 ± 0.04</td>
<td>5.90 ± 0.06</td>
<td>0.50†</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Flow rate</td>
<td>ml min⁻¹</td>
<td>0.36</td>
<td>0.35</td>
<td>0.40 ± 0.03</td>
<td>0.36 ± 0.03</td>
<td>NS</td>
<td>0.121</td>
</tr>
<tr>
<td>Protein content</td>
<td>mg ml⁻¹</td>
<td>1.00</td>
<td>0.82</td>
<td>1.05 ± 0.03</td>
<td>0.91 ± 0.03</td>
<td>15%*</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MUC5B</td>
<td>AU ml⁻¹</td>
<td>1.15</td>
<td>0.92</td>
<td>1.40 ± 0.08</td>
<td>1.22 ± 0.11</td>
<td>15%*</td>
<td>0.028</td>
</tr>
<tr>
<td>MUC7</td>
<td>AU ml⁻¹</td>
<td>3.3</td>
<td>4.1</td>
<td>5.41 ± 0.55</td>
<td>6.36 ± 0.61</td>
<td>18%†</td>
<td>0.020</td>
</tr>
<tr>
<td>Albumin</td>
<td>µg ml⁻¹</td>
<td>22.5</td>
<td>16.0</td>
<td>36.4 ± 4.0</td>
<td>37.8 ± 5.6</td>
<td>NS</td>
<td>0.051</td>
</tr>
<tr>
<td>Lactoferrin</td>
<td>µg ml⁻¹</td>
<td>0.97</td>
<td>0.74</td>
<td>1.09 ± 0.05</td>
<td>0.98 ± 0.07</td>
<td>NS†</td>
<td>0.042</td>
</tr>
<tr>
<td>Cystatin S</td>
<td>AU ml⁻¹</td>
<td>1.1</td>
<td>0.8</td>
<td>2.64 ± 0.31</td>
<td>2.28 ± 0.31</td>
<td>NS</td>
<td>0.113</td>
</tr>
<tr>
<td>S-IgA</td>
<td>U ml⁻¹</td>
<td>287</td>
<td>229</td>
<td>324 ± 14</td>
<td>282 ± 18</td>
<td>15%*</td>
<td>0.002</td>
</tr>
<tr>
<td>Amylase activity</td>
<td></td>
<td>29.5</td>
<td>23.8</td>
<td>44.4 ± 3.0</td>
<td>37.2 ± 2.8</td>
<td>NS</td>
<td>0.199</td>
</tr>
<tr>
<td>Chitinase activity</td>
<td>mU ml⁻¹</td>
<td>22.6</td>
<td>17.2</td>
<td>38.7 ± 3.8</td>
<td>24.6 ± 2.6</td>
<td>57%*</td>
<td>0.019</td>
</tr>
<tr>
<td>Protease activity</td>
<td>AU ml⁻¹</td>
<td>47.2</td>
<td>42.0</td>
<td>48.2 ± 2.3</td>
<td>47.1 ± 2.3</td>
<td>NS</td>
<td>0.813</td>
</tr>
<tr>
<td>Lysozyme activity</td>
<td>U ml⁻¹</td>
<td>1342</td>
<td>1660</td>
<td>1541 ± 112</td>
<td>1794 ± 112</td>
<td>16%†</td>
<td>0.026</td>
</tr>
</tbody>
</table>

SE - standard error.

* Distributions significantly different, with mean higher in males.
† Distributions significantly different, with mean higher in females.
‡ Although p < 0.05, not significant after FDR adjustment.
### Table 4. Salivary variables stratified according to gender, results expressed in units of output per min (calculated as flow rate x concentration). Percentage differences are shown for variables with distributions in males found to be significantly different compared to females using a Mann-Whitney U test.

<table>
<thead>
<tr>
<th>Unit</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Percentage difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protein content</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mg min(^{-1})</td>
<td>0.35</td>
<td>0.29</td>
<td>0.39</td>
<td>0.30</td>
<td>26%*</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>MUC5B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AU min(^{-1})</td>
<td>0.39</td>
<td>0.27</td>
<td>0.48</td>
<td>0.40</td>
<td>19%*</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>MUC7</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AU min(^{-1})</td>
<td>1.30</td>
<td>1.51</td>
<td>1.83</td>
<td>2.05</td>
<td>NS</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>Albumin</strong></td>
<td>µg min(^{-1})</td>
<td>7.4</td>
<td>6.6</td>
<td>13.2 ± 2.2</td>
<td>10.5 ± 1.4</td>
<td>NS</td>
</tr>
<tr>
<td><strong>Lactoferrin</strong></td>
<td>µg min(^{-1})</td>
<td>0.32</td>
<td>0.26</td>
<td>0.38 ± 0.02</td>
<td>0.33 ± 0.03</td>
<td>NS</td>
</tr>
<tr>
<td><strong>Cystatin S</strong></td>
<td>AU min(^{-1})</td>
<td>0.40</td>
<td>0.28</td>
<td>1.00 ± 0.11</td>
<td>0.80 ± 0.12</td>
<td>NS</td>
</tr>
<tr>
<td><strong>S-IgA</strong></td>
<td>µg min(^{-1})</td>
<td>106</td>
<td>79</td>
<td>107 ± 3</td>
<td>82 ± 3</td>
<td>26%*</td>
</tr>
<tr>
<td><strong>Amylase activity</strong></td>
<td>U min(^{-1})</td>
<td>12.0</td>
<td>9.9</td>
<td>17.7 ± 1.5</td>
<td>12.8 ± 1.0</td>
<td>32%*</td>
</tr>
<tr>
<td><strong>Chitinase activity</strong></td>
<td>mU min(^{-1})</td>
<td>7.7</td>
<td>5.1</td>
<td>12.3 ± 1.1</td>
<td>8.3 ± 1.0</td>
<td>37%*</td>
</tr>
<tr>
<td><strong>Protease activity</strong></td>
<td>AU min(^{-1})</td>
<td>15.9</td>
<td>14.5</td>
<td>17.6 ± 0.9</td>
<td>16.5 ± 1.1</td>
<td>NS</td>
</tr>
<tr>
<td><strong>Lysozyme activity</strong></td>
<td>U min(^{-1})</td>
<td>424</td>
<td>576</td>
<td>589 ± 46</td>
<td>696 ± 54</td>
<td>NS</td>
</tr>
</tbody>
</table>

\(^{1}\)SE - standard error.

* Distributions significantly different, with mean higher in males.
Figure 3. Distribution histograms of salivary pH (A) and buffered pH (B) divided according to sex. Males are represented in blue on the left side, females in green on the right side of both histograms. The correlation between pH and buffered pH values was $\rho = 0.83$, $p < 0.001$.

The Spectral Clustering co-occurrence plot in Fig. 4 offers an intuitive visualization of the clustering of the participants based on salivary biochemistry. Fig. 5 shows a general overview of the variation of each salivary biochemical variable across the sample population. Fig. 5 connects the clustering with the biochemical variation and provides an indication of the salivary biochemical variables that are mainly responsible for the clustering seen in Fig. 4.
Figure 4. Spectral Clustering co-occurrence plot. Participants are ordered along both X- and Y-axis according to the co-occurrence matrix (i.e. the more similar the salivary biochemistry profiles of any two participants, the higher their tendency to cluster together and the closer they are placed on the axis). Co-occurrence values range from 0 (for participants who never cluster together) to 1.0 (for participants who always cluster together). The horizontal bar delimits the two clusters (I – red and II - blue).

Figure 5. Visual overview of the biochemical assay results and their co-variation within the study population. The X-axis contains the 14 biochemical variables. The Y-axis contains
266 participants, arranged according to the co-occurrence matrix (same as in Fig. 3). Results for each individual assay are scaled to equal range (from 0 to 1) and color coded (legend on the right side of the figure). The vertical bar on the left delimits the two clusters (I – red and II - blue).

The first, small cluster visible in the upper-left corner of Fig. 4 (cluster I) contained 22 participants (8.3% of total number of participants, 9 males, 13 females). Fig. 5 shows that this cluster was differentiated by low pH and buffered pH values together with high lysozyme activity and MUC7 content compared to the remainder of the population. The inter-cluster differences observed in Fig. 4 were confirmed by Mann-Whitney U tests with \( p < 0.001 \) for all 4 biochemical variables. The participants accounting for the remaining 91.7% of the sample population made up a large cluster (cluster II, lower-right side of Fig. 3), within which 3 weakly differentiated sub-clusters could be discerned. The 1st sub-cluster (descending down the main diagonal, top-left to bottom-right) tends to group participants with average pH, high buffered pH, high amylase and lysozyme activity; the 2nd sub-cluster participants with average pH and buffered pH and low flow rate; the 3rd sub-cluster, participants with high pH and buffered pH, high flow rate and low protein content. However, these 3 sub-clusters were not sufficiently differentiated to allow for rigorous subject assignment and statistical comparisons.
Discussion

The purpose of this study was to obtain an overview of the distribution and interrelations of a set of salivary biochemical variables with known relevance for oral health, measured under strictly defined conditions in systemically healthy young adults. The participants were screened and included in case of absence of apparent caries lesions and periodontal disease. They were non-smokers. The rationale was to obtain accurate reference values of the biochemistry of saliva associated with oral health or, in other words, to construct a profile of the salivary biochemical network at a state of homeostasis. Another goal was to detect possible subgroups of individuals in the sample population based on their salivary biochemistry.

The means of the distributions of biochemical variables measured in the present study are generally within the range of values reported in previous studies, where similar conditions and analytical methods were employed [5, 15, 29-31]. Several previous investigations have measured various smaller subsets of these biochemical variables. However, methodological issues such as broadly defined study populations, diverse saliva sampling methods and different analytical methods and measurement units interfere with compiling the results across different studies. In some cases, the results are not comparable, for example where previous studies have measured the amount of a particular enzyme antigen present in saliva rather than the respective enzyme activity [32]. In other cases, differences in inclusion criteria, saliva collection and/or processing protocols may account for diverging results. This underlines the advantage of investigating a wider array of functional salivary proteins and enzymes within the framework of a single study in order to circumvent methodological bias and obtain a reliable overview.

Significant correlations were found between salivary biochemical variables. Total protein content and salivary flow rate were negatively correlated ($\rho = -0.47$, $p < 0.001$). Most of the salivary proteins and enzymes quantified correlated positively with total protein content and negatively with flow rate, although the size of these correlations varied considerably (Table 1). These findings were partly in line with results from a previous study by RUDNEY et al. that focused on a subset of salivary antimicrobial proteins (secretory-IgA, lactoferrin and lysozyme) [15]. The present data support their finding that secretory-IgA and lactoferrin correlate highly with salivary protein content. However, the size of the correlation observed for lysozyme was much smaller compared to the one reported in the aforementioned study, with a correlation coefficient of 0.32 as opposed to 0.62. This may be explained by the fact that for the present study lysozyme activity was measured, while RUDNEY et al. assessed the amount of lysozyme antigen using an ELISA technique. Also, their addition of ethylenediaminetetraacetic acid (ETDA) to the samples may have affected the results, as suggested by the authors themselves [15].
Quantifying the cumulative anti-microbial effect of lysozyme is complicated by the fact that besides muramidase activity (which kills bacteria by lysing cell walls) lysozyme also displays a non-enzymatic killing effect attributed to the activation of bacterial autolysins [33]. Interestingly, a negative correlation was found between lysozyme activity and salivary pH ($\rho = -0.44, p < 0.001$, Fig. 2). Since pH was controlled in the lysozyme enzymatic activity assay used in the present study, differences in salivary pH could not have influenced the assay results. The size of the correlation did not change after controlling for flow rate and protein content, thereby adjusting for the possibility of a confounding influence of these two factors. The inverse correlation between pH and lysozyme was mirrored in the cluster structure seen in Fig. 4 and Fig 5, with cluster I differentiated by low pH and high lysozyme activity compared to the relatively high pH and low lysozyme activity of cluster II. The mechanism behind this correlation remains unclear. It may be linked to the relative contribution of the different salivary glands to total salivary secretion. Sublingual saliva contains higher levels of lysozyme compared to parotid saliva and has a lower pH [34]. A higher contribution of sublingual saliva relative to parotid saliva would therefore increase lysozyme activity while slightly decreasing salivary pH.

Another interesting relationship was the negative correlation between lysozyme activity and chitinase activity ($\rho = -0.46, p < 0.001$). These two enzymes have somewhat complementary roles: while lysozyme cleaves bacterial cell walls, chitinase lyses fungal walls and can act against the oral pathogen Candida albicans [35]. The inverse relationship might suggest that these enzymes act together and that the effects of the two are cumulative, whereby a decrease in the level of one could be compensated by a concomitant increase of the other.

Significant differences between sexes were found for several of the biochemical variables examined (Tables 3 and 4). Although the mean salivary flow rate was lower in females than in males, the difference in flow rate between the two genders did not reach statistical significance, in contrast with several other studies [2, 36, 37]. Strikingly, the salivary pH distribution in females was significantly shifted towards a more acidic pH as compared to males. A similar shift towards lower values was seen for buffered pH, together with a higher number of females with extreme low buffered pH values (Fig. 3). Females made up 79% of the group of 19 participants with a buffered pH below 5.0, with the aforementioned group showing an interesting clustering effect around a buffered pH of 4.7 – 4.8 (Fig. 3). This convergence could imply that within this group of participants the first two salivary buffering systems (carbonate and phosphate) were insufficient, leading to the protein buffering system to act as a last resort. Protein buffering is the only buffering system present in saliva that is effective at pH values below 5.0 [38]. To our knowledge, sex-related salivary pH differences have not been previously reported, although lower buffering capacity in females has been recorded [39, 40]. Many studies have found higher caries prevalence in females [41, 42]. The causative factors most often invoked are social and family roles, dietary habits, genetic variations, hormonal fluctuations and lower
salivary flow rates [41, 42]. In the light of the results of the present study, lower salivary pH and buffering capacity may also be taken into consideration as potential contributors to this higher caries susceptibility.

The observed sex-related biochemical differences in saliva could be explained by physiological factors, primarily the influence of sex hormones on the salivary glands. Firstly, females have smaller salivary glands, which may contribute to the difference in salivary biochemical profiles [37]. Secondly, different gene expression profiles have been found in female salivary gland tissue as compared to males, including the expression of several genes associated with saliva secretion. This presumably is a consequence of the action of estrogen on the salivary glands estrogen receptors [43, 44]. Salivary gland sexual dimorphism and diverging gene expression profiles could explain the differences seen in salivary pH and protein composition, although the limitations of the present observational study prevent it from pinpointing definitive causative factors.

In summary, this study gives a detailed overview of the distribution and interrelations of a set of salivary biochemical variables relevant to oral health and highlights some significant biochemical differences between male and female salivary biochemistry. These findings may help advance our understanding of saliva as a functional biochemical network and its relationship to oral health.

Acknowledgements

The project was funded by TI Food and Nutrition, a public-private partnership on precompetitive research in food and nutrition. The public partners are responsible for the study design, data collection and analysis, decision to publish, and preparation of the manuscript. The private partners have contributed to the project through regular discussion. We wish to thank Nienke Hennequin-Hoenderdos, Dagmar Else Slot and Eveline van der Sluijs for their excellent contribution in developing the clinical protocol, and for coordinating and monitoring the clinical part of this study, Kamran Nazmi for his expert assistance in optimizing the biochemical assays, as well as Floris Bikker and Jan Bolscher for kindly providing protease substrates and antibodies.
References


CHAPTER 3

A study of the variation in the salivary peptide profiles of young healthy adults acquired using MALDI-TOF


PLOS ONE 2016, 11(6):e0156707
Abstract

A cross-sectional observational study was conducted to evaluate the inter-individual variation in the MALDI-TOF MS peptide profiles of unstimulated whole saliva in a population of 268 systemically healthy adults aged 18-30 yr (150 males and 118 females) with no apparent caries lesions or periodontal disease. Using Spectral Clustering, four subgroups of individuals were identified within the study population. These subgroups were delimited by the pattern of variation in 9 peaks detected in the 2-15 kDa m/z range. An Unsupervised Feature Selection algorithm showed that P-C peptide, a 44 residue-long salivary acidic proline-rich protein, and three of its fragments (Fr. 1-25, Fr. 15-35 and Fr. 15-44) play a central role in delimiting the subgroups. Significant differences were found in the salivary biochemistry of the subgroups with regard to lysozyme and chitinase, two enzymes that are part of the salivary innate defense system (P < 0.001). These results suggest that MALDI-TOF MS salivary peptide profiles may relate information on the underlying state of the oral ecosystem and may provide a useful reference for salivary disease biomarker discovery studies.
**Introduction**

Saliva is crucial for the maintenance of oral health [1]. Salivary components provide lubrication, stabilize oral pH, aid remineralization of dental enamel and modulate growth and adherence of oral bacteria to tooth surfaces [2, 3].

The diagnostic potential of saliva has been increasingly explored during the last decades. A number of features make saliva an attractive medium for biomarker discovery [4-6]. First of all, saliva sampling is a non-invasive, safe, and cost-effective option compared to the collection of other body fluids. Secondly, salivary biomarkers not only give insight into the health status of the oral cavity, but can also convey information regarding systemic health. This is due to the presence of numerous serum-derived compounds which enter the oral fluid either directly (from the salivary glands) or indirectly (either via the gingival crevicular fluid or through inflamed gingiva or damaged parts of the oral mucosa) [5, 7, 8].

Saliva as a potential source of biomarkers poses its own unique set of challenges. Saliva is a mixture of secretions from three pairs of major glands (parotid, submandibular and sublingual) as well as numerous minor glands, each having a characteristic protein composition [9, 10]. As a result, the composition of whole saliva depends heavily on the manner in which saliva was collected and is also affected by factors such as age, sex, medication, circadian rhythm, physical activity, and oral hygiene procedures prior to collection [9, 11-13]. Careful standardization of saliva collection is therefore crucial for obtaining reproducible results [7]. Other complicating factors are related to the nature of salivary proteins and peptides. Firstly, virtually all the major salivary protein families display a large degree of phenotypic variation due to genetic polymorphisms, alternative RNA splicing and various post-translational modifications [9, 14]. Moreover, secreted saliva is exposed to proteolytic activity from both endogenous proteases (originating from the salivary glands or mucosal cells) and exogenous proteases (produced by the oral microflora) [9]. As the vast majority of salivary biomarkers are proteins / peptides, these additional levels of complexity hamper interpretation of the data.

Comparison of a healthy control group to a diseased group is common in studies that attempt to discover salivary biomarkers [15, 16]. This raises the question: what is the inherent biological variation of potential peptide biomarkers in the saliva of a healthy population? Is the variation among individuals related to specific sets of salivary peptides that can cluster individuals into discrete subgroups? Knowledge of the variation in peptide profiles in healthy saliva is therefore particularly relevant for salivary biomarker discovery.

The aim of this study was to examine inter-individual variation in MALDI-TOF MS salivary peptide profiles within a population of systemically healthy young adults and to identify potential subgroups.
Materials and Methods

Clinical study structure, study population and exclusion criteria

The study was carried out within the framework of the Top Institute Food and Nutrition project "Estimating the boundaries of a healthy oral ecosystem in young individuals" [17].

Whole unstimulated saliva was collected in a cross-sectional single-center observational clinical study at the Academic Center for Dentistry Amsterdam (ACTA). The protein biochemistry of the same saliva sample set has already been analyzed and described previously [17]. The study population comprised a convenience sample of systemically healthy young adults aged 18-30 yr old without periodontitis or apparent caries lesions. Participants were students of universities and colleges in and around Amsterdam, The Netherlands. They were invited for screening when they had visited a dentist the previous year and had been considered to be without oral or dental problems. The volunteers were screened for suitability according to the criteria of the Dutch Periodontal Screening Index (DPSI) [18]. They were included if they had a DPSI ≤ 3-. The following exclusion criteria were used: overt dental caries, inter-proximal restorations between the first and second or second and third upper molars, apparent oral lesions, infections, a habit of smoking, recent use of antibiotics, use of anti-inflammatory drugs or other prescribed medication which could interfere with the outcome of this study (except for oral contraceptives). Participants were instructed to abstain from eating, drinking, chewing gum or performing strenuous physical exercise prior to the appointment, and not to brush their teeth in the 24 hours prior to the appointment.

The study was conducted in accordance with the Declaration of Helsinki (2008) of the World Medical Association and approximated Good Clinical Practice guidelines. The study protocol was reviewed and approved by the Medical Ethics Committee of the Academic Medical Centre of Amsterdam (2012_210#B2012406) and recorded at the Dutch Trial Register (NTR3649). All participants signed an informed consent form.

Unstimulated saliva sampling and salivary biochemical analysis

All saliva samples were collected between 9 and 10 a.m. Participants were instructed to accumulate saliva in the floor of the mouth without stimulation by orofacial movements and to spit at 30 s intervals into pre-weighed 30-ml polypropylene tubes (Sterilin, Newport, U.K.) which were kept on ice. The collection period was 5 min. The tubes containing unstimulated saliva samples were weighed and salivary flow rate was calculated assuming a saliva density of 1.0 g ml⁻¹. Saliva samples were homogenized by vortexing for 20 s. Salivary pH and buffered pH were measured immediately after saliva collection as previously described [17]. Samples were clarified by centrifugation for 10 min at 4°C and
10,000 g to remove epithelial cell debris, bacteria and food residues. The resulting clarified saliva sample was diluted 1:1 with a 500 mM NaCl solution to a final concentration of 250 mM NaCl, aliquoted and stored at -80°C. The dilution prevented protein aggregation and precipitation during saliva freezing and storage and lowered viscosity allowing for more precise sample manipulation and improved reproducibility [17]. A sufficient number of aliquots of each sample were produced to avoid exposing samples to multiple freezing and thawing cycles. Salivary total protein content, mucins MUC5B and MUC7, lactoferrin, secretory-IgA, albumin, amylase, chitinase, proteases and lysozyme were measured as described previously [17].

**MALDI-TOF MS spectra acquisition**

Prior to MALDI-TOF MS analysis the saliva samples were desalted using C18 ZipTips (Merck Millipore, Darmstadt, Germany) as follows: ZipTips were wetted twice with 10 µL acetonitrile (ACN), equilibrated 3 times with 0.1% trifluoroacetic acid (TFA), loaded with 10 µL of saliva sample diluted 1:50 with double-distilled water, and washed 3 times with 0.1% TFA. The purified peptides were then eluted and spotted directly onto a MALDI target plate with 1 µL of matrix solution (10 mg α-cyano-4-hydroxy cinnamic acid in 1 ml of ACN/water 1:1 (v/v) with 2.5% TFA). The spots were allowed to air-dry and spectra were acquired using an Autoflex III MALDI-TOF mass spectrometer (Bruker, Bremen, Germany). Spectra were recorded in linear mode at a mass range of 2 – 15 kDa with a 200 Hz laser at 355 nm. Each spectrum was an average of 2000 laser shots. The mass spectrometer was calibrated using a Bacterial Test Standard (Bruker, Bremen, Germany).

Samples were analyzed in duplicate, in 12 sessions spread over a period of one month. For duplicate analysis, separate aliquots of each saliva sample were analyzed independently (from storage to ZipTip-ing, spotting and spectral acquisition). Visual quality control was performed for all raw duplicate spectra. Samples were re-analyzed in new duplicates if either one of the duplicate spectra from a sample showed excessive noise or large baseline drift, or if the duplicate spectra exhibited dissimilarities. Spectra from 49 of the 261 samples analyzed (19%) failed initial quality control and were re-analyzed.

Spectra of a reference saliva sample were acquired in quadruplicate at each session in order to aid the subsequent sample spectral alignment and to assess the reproducibility of the assay. Peak intensity data from a set of 11 reference peaks in the mass-to-charge (m/z) range 2.0 to 7.6 kDa were used to calculate the within-session, between-session and overall coefficients of variation (CV). The mean within-session CV was 6% (range 4-12) while mean between-session CV was 10% (range 4-21, with the highest CV for the highest m/z peak considered, at 7.6 kDa). Overall CV was 12% (range 6-21).
MALDI-TOF MS spectra processing

The raw MALDI-TOF MS spectra were processed in Matlab R2012b using the Mathworks bioinformatics tool box (MathWorks, Natick, MA, U.S.A). The workflow consisted of spectra resampling followed by baseline subtraction, smoothing and normalization of the total area under the curve (i.e. normalizing based on the total amount of sample protein ionized per spectrum). Normalization for total area under the curve insures that differences in saliva sample protein concentration, or in the protein concentration of the spotted sample, are compensated for. Reference spectra were used to align batches of spectra analyzed on different days and to compensate for inter-session instrument drift. Duplicate spectra were then averaged into one sample spectrum. Subsequently, peak detection was performed, followed by peak binning (peak coalescing) using a hierarchical clustering algorithm to calculate a common m/z reference peak vector. The final result was a 3-dimensional database of salivary peptide profiles consisting of sample ID’s, peak m/z values, and peak intensities. Peak identification was attempted using MALDI-MS/MS or by matching the peak m/z ratios to literature values from previous studies [19, 20]. The spectra processing workflow is exemplified in S1 Fig.

MALDI-TOF MS/MS

Identification of MALDI-TOF peaks by MALDI-TOF MS/MS was attempted for all peaks <4,000 Da using a Laser Induced Fragmentation in Time (LIFT) protocol on a Autoflex III MALDI-TOF mass spectrometer (Bruker, Bremen, Germany). Spectra were averages of 8,000 laser shots. The mass window was ±13 Da around the precursor ion mass. The reflector detector voltage was 1.8 kV. For each individual precursor ion mass, the respective extraction delay times in both ion sources were calculated by the instrument control software.

Spectral Clustering and Feature Selection

Spectral Clustering analysis was performed using the dataset that contained the processed data from all MALDI-TOF MS spectra. Spectral Clustering was performed in open-source Python 2.7 using the neighborhood co-regularized Spectral Clustering algorithm developed by Tsivtsivadze et al. [21] based on the Spectral Clustering method published by von Luxburg [22]. For this purpose, the data for each MALDI-TOF MS peak were scaled to equal ranges and a similarity matrix was calculated based on the Euclidean distances between each pair of participants (i.e. on the similarity of the overall peptide profiles of each pair of participants). A co-occurrence matrix was subsequently calculated based on the
clustering results, quantifying the tendency of any pair of participants to fall within the same cluster over many k-means clusterings using varying parameters. The co-occurrence clustering plots were constructed using Matlab R2012b (MathWorks, Natick, MA, U.S.A). After visual examination of the Spectral Clustering plots the number of clusters was determined and participants were mathematically assigned to the clusters using a probabilistic decomposition algorithm [23].

Unsupervised Feature Selection was implemented in open-source Python 2.7 code using the Unsupervised Multi-View Feature Selection via Co-Regularization algorithm [24]. Data from all features in the input dataset (i.e. intensities corresponding to each m/z value) were scaled to equal range. The output of Unsupervised Feature Selection was a list of features ranked based on the values of the weighting coefficients’ vector norm of the regularized spectral regression problem. The highest ranking scores based on the vector norm values correspond to the features that were the most important in determining the data clustering [24].

**Data analysis and statistical methods**

The data were analyzed statistically using SPSS 21.0 software (IBM, Armonk, NY, U.S.A.). Pearson’s product-moment was used to assess correlations and independent sample t-tests were used to compare means. The statistical significance level used was 0.05. The Benjamini-Hochberg False Discovery Rate (FDR) procedure was used to correct for multiple comparisons [25]. The FDR was set at 0.05. Cohen’s d was used to quantify effect size [26].
Results

Prior to the study, 336 potential participants were screened at the dental clinic in a separate session. Of these, 46 (23 males and 23 females) were excluded based on pre-defined inclusion and exclusion criteria as reported in a previous publication (16). Of the remaining 290 subjects, 10 took part in a pilot study (not included in the final data) and 12 dropped out (11 due to schedule conflicts and one reported pregnancy). In total, 268 participants completed the study (150 males, 118 females), with a mean age of 22.6 yr (on the day of their appointment) and a range of 18-32 yr. Of these, 7 were unable to provide sufficient saliva during the sampling period. Data obtained from the remaining 261 participants (145 male and 116 female) are presented in this article.

A total of 129 peaks were detected in the MALDI-TOF MS spectra. Of these, 49 peaks were present in less than 2% of the samples (i.e. in less than 5 participants out of 261). These low frequency peaks were filtered out and all further analyses were performed on the remaining 80 peaks. The complete dataset is available in S2 Dataset. Table 1 lists the peaks for which either definitive or putative identities were assigned. One peak was identified by means of MALDI-TOF MS/MS: a peak at 2043 Da was identified as a P-C peptide fragment containing amino acids 15 through 35 (Fr. 15-35). The MS/MS spectrum is shown in S3 Fig. Another 19 peaks were identified by matching their m/z values to those reported in previous studies [19, 27].
Table 1. Putative identities for peaks in the MALDI-TOF MS spectra of saliva from healthy volunteers.

<table>
<thead>
<tr>
<th>m/z [Da]</th>
<th>Present in [%] of individuals</th>
<th>Peak identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2043</td>
<td>92</td>
<td>P-C peptide Fr. 15-35&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>2185</td>
<td>95</td>
<td>P-C peptide (2+ charge peak)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>2523</td>
<td>94</td>
<td>P-C peptide Fr. 1-25&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>2921</td>
<td>94</td>
<td>P-C peptide Fr. 15-44&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3039</td>
<td>42</td>
<td>Histatin 3 Fr. 1-24 (Histatin 5)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3200</td>
<td>18</td>
<td>Histatin 3 Fr. 1-25&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>3376</td>
<td>5</td>
<td>α-defensin 2&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3447</td>
<td>8</td>
<td>α-defensin 1&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3497</td>
<td>88</td>
<td>α-defensin 3&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3546</td>
<td>8</td>
<td>Histatin 3&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>4250</td>
<td>23</td>
<td>P-C peptide Des Q44&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>4371</td>
<td>69</td>
<td>P-C peptide (1+ charge peak)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>4557</td>
<td>8</td>
<td>P-B Des 1-12&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>4921</td>
<td>26</td>
<td>Histatin 1&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>5390</td>
<td>8</td>
<td>Statherin diposphorylated&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>5954</td>
<td>79</td>
<td>Cystatin B Fr. 1-53&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>7528</td>
<td>12</td>
<td>II-2 basic proline rich protein&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>7618</td>
<td>66</td>
<td>II-2 basic proline rich protein, phosphorylated&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>10444</td>
<td>17</td>
<td>S100A12 (calgranulin C)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>11172</td>
<td>26</td>
<td>aPRP-4, diposph./PRP-3 diposphorylated&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Identification by MALDI-TOF MS/MS (S3 Fig).
<sup>b</sup>Castagnola et al. 2012 [27].
<sup>c</sup>Cabras et al. 2010 [19].

The Spectral Clustering co-occurrence plot (Fig 1) offers a visualization of the clustering of the participants based on their salivary peptide profiles. The clusters were designated I through IV from top-left to bottom-right of Fig 1. Cluster III contained the majority of the individuals from the study population (213, 81.6% of total), while the other three clusters were considerably smaller (cluster I – 7, 2.7% of total; cluster II – 16, 6.1 % of total; cluster IV – 25, 9.6% of total). Fig 2 gives an overview of the peptide profile variation across the sample population and highlights the 9 peaks that were found by Unsupervised Feature Selection to determine the clustering (Table 2). Fig 3 shows the intensities of each of these 9 peaks in the 4 clusters. Example spectra from each cluster are shown in S4 Fig.
Figure 1. Spectral Clustering co-occurrence plot. Participants are ordered along both X- and Y-axis according to the co-occurrence score (i.e. the more similar the peptide profiles of any two participants, the higher their tendency to cluster together and the closer they are placed on the axis). Co-occurrence score values range from 0 (for participants who never cluster together) to 1.0 (for participants who always cluster together). The horizontal bar delimits the four clusters.

Figure 2. Overview of the peptide profile variation within the study population. The X-axis contains the 80 peaks ordered by their m/z ratio, (lowest to highest, from left to right). The
Y-axis contains the 261 participants, arranged according to the co-occurrence matrix (same as in Fig 1). Data for each individual peak were scaled to equal range (from 0 to 1) and color coded (legend on the right side of the figure). The vertical bar on the left delimits the four clusters. Red arrows on the upper margin highlight the 9 peaks shown by Unsupervised Feature Selection to be determining the clustering structure (Fig 1).

**Table 2.** Peaks identified by Unsupervised Feature Selection to be determining the subgroups found by Spectral Clustering. Peaks are listed in decreasing order of their effect on the clustering.

<table>
<thead>
<tr>
<th>No.</th>
<th>Putative peak identity</th>
<th>Peak m/z [Da]</th>
<th>Peak present in [%] of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P-C peptide, Fr. 15–44&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2921</td>
<td>94</td>
</tr>
<tr>
<td>2</td>
<td>P-C peptide, Fr. 1-25&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2523</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>unidentified peak</td>
<td>5980</td>
<td>41</td>
</tr>
<tr>
<td>4</td>
<td>P-C peptide (1+ charge)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4371</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>unidentified peak</td>
<td>2725</td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td>Cystatin B, Fr. 1-53&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5954</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>P-C peptide, Fr. 15-35&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2043</td>
<td>92</td>
</tr>
<tr>
<td>8</td>
<td>P-C peptide (2+ charge)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2185</td>
<td>95</td>
</tr>
<tr>
<td>9</td>
<td>II-2 basic proline-rich protein, phosphorylated&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7610</td>
<td>96</td>
</tr>
</tbody>
</table>

<sup>a</sup>Castagnola et al. 2012 [27].

<sup>b</sup>Identified by MALDI-TOF MS/MS.
Figure 3. Boxplots stratified on the 4 clusters showing peak intensities across the sample population for the 9 peaks determining the clustering structure. X-axis contains the 4 clusters. Y-axis contains the peak intensity, scaled from 0 (minimum detection level) to 100 (maximum detected intensity of any peak across all samples and peaks). Peaks are listed in decreasing order of their importance for the clustering, same as in Table 2. (A) P-C peptide, Fr. 15-44. (B) P-C peptide, Fr. 1-25. (C) unidentified peak (m/z = 5980). (D) P-C peptide (1+ charge). (E) unidentified peak (m/z = 2725). (F) Cystatin B, Fr. 1-53. (G) P-C peptide, Fr. 15-35. (H) P-C peptide (2+ charge). (I) II-2 basic proline-rich protein, phosphorylated.

Principal Component Analysis (PCA) scores plots in Fig 4 illustrate the degree to which the 9 peaks found by Unsupervised Feature Selection account for the overall grouping of individuals based on their salivary peptide profiles.
Figure 4. PCA scores plot. (A) PCA based on all 80 peaks compared to (B) PCA scores plot based solely on the 9 peaks resulting from the Feature Selection procedure. Individuals are color labeled according to their cluster. The proportion of variance explained was as follows: (A) PC1 – 17.4%, PC2 – 8.8% (26.2% cumulative). (B) PC1 – 39.5%, PC2 – 26.3% (65.8% cumulative).

Of the 9 peaks determining the clustering of individuals based on their peptide profile (Table 2), 5 peaks were identified as either the intact form or fragments of P-C peptide. A number of significant correlations were found between these 5 peaks. First of all, there was a positive correlation \( r = 0.72 \) between the two peaks corresponding to the 1+ and 2+ charged forms of intact P-C peptide. Secondly, there were positive correlations between the three peaks corresponding to the different P-C peptide fragments: between Fr. 1-25 and Fr. 15-44 \( (r = 0.75) \), between Fr. 1-35 and Fr. 15-44 \( (r = 0.51) \), and between Fr. 15-35 and Fr. 15-44 \( (r = 0.52) \). Finally, there was a negative correlation between the intact P-C peptide (1+ charge) and Fr. 15-35 \( (r = -0.63) \). All correlations had \( p < 0.001 \).

A set of nine salivary proteins with known relevance for oral health were also measured from aliquots of the same saliva samples from which the peptide profiles were acquired [17]. Salivary lysozyme and chitinase differed significantly between clusters when compared using a one-way ANOVA (both with \( p < 0.001 \)) (Tables 3 and 4). No significant
inter-cluster differences were found for mucins MUC5B and MUC7, salivary lactoferrin, albumin, secretory-Immunoglobulin A, cystatin S or amylase.

**Table 3.** Salivary lysozyme activity and chitinase activity stratified by Spectral Clustering clusters (Values expressed in Units of activity ml⁻¹, mean ± standard error).

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Lysozyme</th>
<th>Chitinase</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (n = 7)</td>
<td>400 ± 180</td>
<td>126.6 ± 55.4</td>
</tr>
<tr>
<td>II (n = 16)</td>
<td>1,738 ± 86</td>
<td>28.3 ± 2.3</td>
</tr>
<tr>
<td>III (n = 213)</td>
<td>2,434 ± 322</td>
<td>17.4 ± 4.2</td>
</tr>
<tr>
<td>IV (n = 25)</td>
<td>349 ± 79</td>
<td>61.0 ± 9.2</td>
</tr>
<tr>
<td>Overall (n = 261)</td>
<td>1,638 ± 80</td>
<td>32.8 ± 2.5</td>
</tr>
</tbody>
</table>

**Table 4.** Differences between Spectral Clustering clusters in salivary lysozyme and chitinase activity compared using one-way ANOVA and Games-Howell post-hoc tests. The magnitude of the inter-cluster differences is quantified using Cohen’s $d$ measure of effect size (i.e. the difference between the two means divided by the standard deviation of the data).

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Cohen’s $d$</th>
<th>$p$-value (post-hoc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme I - II</td>
<td>1.1</td>
<td>0.002</td>
</tr>
<tr>
<td>I - III</td>
<td>1.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>I - IV</td>
<td>0.1</td>
<td>0.993</td>
</tr>
<tr>
<td>II - III</td>
<td>0.5</td>
<td>0.198</td>
</tr>
<tr>
<td>II - IV</td>
<td>1.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>III – IV</td>
<td>2.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Chitinase I - II</td>
<td>1.9</td>
<td>0.388</td>
</tr>
<tr>
<td>I - III</td>
<td>1.3</td>
<td>0.284</td>
</tr>
<tr>
<td>I - IV</td>
<td>0.7</td>
<td>0.793</td>
</tr>
<tr>
<td>II - III</td>
<td>0.3</td>
<td>0.075</td>
</tr>
<tr>
<td>II - IV</td>
<td>0.9</td>
<td>0.008</td>
</tr>
<tr>
<td>III – IV</td>
<td>1.3</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Discussion

The aim of this study was to examine the inter-individual variation in MALDI-TOF MS salivary peptide profiles and to identify potential subgroups within a population of 268 systemically healthy young adults. All volunteers were non-smokers and were screened and included in the study only in the absence of apparent caries lesions and periodontal disease. The rationale was to examine the natural variation in salivary peptide profiles of healthy individuals, a population that is commonly included as a control group in studies aiming to discover salivary biomarkers. The results of the present study suggest that individuals in a healthy population may be clustered into several sub-groups based on a limited number of MALDI-TOF MS peaks (Table 2). Furthermore, 5 of the 9 peaks found to determine the clustering were related to a single peptide. These peaks were identified as intact P-C peptide in its 1+ and 2+ charged form and 3 different fragments of P-C peptide.

P-C peptide is a member of the acidic proline-rich protein (aPRP) class of the salivary proline-rich proline (PRP) family [28]. aPRPs are secreted by the parotid as well as by the submandibular / sublingual glands and account for 20-30% of the total salivary protein content [29, 30]. aPRPs show an extensive heterogeneity and include numerous polymorphic isoforms with high sequence homology [29, 31]. aPRP precursors of P-C peptide are encoded on 2 genes: PRH1 (alleles PIF-s and Db-s) and PRH2 (alleles PRP-1 and PRP-2) [32]. Prior to secretion, these precursors are partially cleaved by proprotein convertases present in the salivary glands yielding the 44 residue-long P-C peptide (the C-terminal fragment of the precursors) and different N-terminal fragments, depending on the precursor isoform (PRP-3, PRP-4, PIF-f or Db-f) [30, 32]. The glandular convertases are unidentified metalloproteases [31, 33]. The cleavage is not complete and a proportion of the precursors are secreted intact. Interestingly, the degree of pre-secretion conversion differs between individuals but is stable for a given individual, with no day to day variation [34].

P-C peptide precursors (PRP1, PRP2, Pif-f and Db) have specific functionalities which change depending on whether they are cleaved or not, with subsequent biological consequences [33]. A 30 residue region at the N-terminal of the intact precursors is rich in aspartate, glutamine and proline and contains several serine phosphate residues. This domain allows these aPRPs to attach to recently cleaned tooth surfaces, become part of the dental pellicle and mediate hydroxyapatite crystal growth [35-38]. Moreover, the attachment causes a conformational change which exposes a bacterial binding site in the C-terminal region, such that intact precursors can also mediate the adherence of bacteria onto the tooth surface [35, 39]. Cleavage of the precursors enhances their attachment ability while removing the bacterial binding site together with the P-C peptide [33, 39]. In addition, free P-C peptide may protect against tannins, harmful polyphenolic compounds found in some plant-based foods [36]. P-C peptide effectively binds and precipitates tannins while its intact precursors do not [36].
The nomenclature of the PRP family can be confusing [40]. It has changed over time, and in some cases different names have been used in parallel for the same protein. This is particularly true for the P-C peptide, which has been designated “IB-8b” in some studies and “peptide Tz” in others, and is sometimes still classified as part of the basic PRPs rather than the aPRPs [36, 40-42]. However, all three designations refer to same peptide [43]. “P-C peptide” dominates in more recent literature and therefore this designation has been used in the present study [30, 43, 44].

The pattern of variation of P-C peptide and its different fragments in saliva may provide valuable information about the underlying state of the oral ecosystem. Un-cleaved precursors are converted post-secretion by bacterial endoproteases in the oral cavity, thus adding to the amount of P-C peptide already secreted from the glands [31]. P-C peptide itself is degraded by bacterial proteases yielding a mix of different fragments. Therefore, the relative amount of P-C peptide present in saliva and its observed pattern of fragmentation are influenced both by processes taking place within the salivary glands as well as by the particular (proteolytic) microbial profile in the oral cavity. At least 7 different P-C peptide fragments have been reported in saliva [29, 31]. The peaks of intact P-C peptide with 1+ and 2+ charge were found earlier top-down MALDI-TOF studies of salivary peptide profiles [43, 45]. A more recent study, also using top-down MALDI-TOF (with prior fractionation by nano-HPLC) found all 3 of the P-C peptide fragments selected as markers for the clusters defined in the present study (Fr. 1-25, Fr. 15-35, and Fr. 15-44) [46]. Another study, using similar techniques to examine the salivary peptide profiles of infants, found 2 of the P-C peptide fragments (Fr. 1-25 and Fr. 15-44), as well as the peaks for both the 1+ and 2+ charge intact P-C peptide [47]. All 4 peaks showed significant increases at the age of 6 months compared to the initial age of 3 months. Some intact P-C peptide is found in almost all saliva samples [29, 30]. However, the pattern of P-C peptide fragments seems to differ in healthy saliva compared to saliva from individuals with high caries experience and from patients with Sjögren’s syndrome [29]. Fr. 1-25 and Fr. 15-44 were predominant in the saliva of healthy individuals [29]. The peaks corresponding to these 2 fragments were both present in the saliva of 94% of the individuals in the present study, screened to be healthy. A previous work comparing the salivary peptide profiles of Sjögren Syndrome patients to those of healthy controls found that 2 peaks corresponding to α-defensins 1 and 2 (m/z values of 3447 Da and 3376 Da, respectively) were significantly higher in Sjögren Syndrome patients [48]. The peaks of α-defensins 1 and 2 were only detected in 8% and 5%, respectively, of the healthy adult population sampled in the current study (Table 1). P-C peptide has also been highlighted in a study examining changes in the salivary peptide profiles of children affected by type 1 diabetes compared to healthy controls [19]. The study found that intact P-C peptide was significantly higher in healthy children, while P-C peptide fragments (Fr. 1-25 and Fr. 15-44) were higher in type 1 diabetic children [19].
P-C peptide and P-C peptide fragments accounted for 5 of the 9 peaks found by Unsupervised Feature Selection to define the 4 clusters observed in the present study population (Fig 1, Table 2). Fig 4 illustrates the reliability of the Unsupervised Feature Selection results. A scores plot from a PCA analysis performed exclusively on data from the 9 selected peaks accurately maintained the grouping seen when including data from all 80 peaks.

Significant differences were found when comparing individuals belonging to the 4 different clusters with regard to salivary lysozyme and chitinase activity (Tables 3 and 4). Lysozyme is a major component of the salivary innate defense system [1, 2]. It acts by cleaving peptidoglycan, thereby killing Gram-positive bacteria via cell wall lysis [1, 2]. Similarly, chitinase contributes to the protection of the oral cavity from pathogenic yeast such as Candida albicans whose cell wall contains chitin [1, 49]. Differences in the salivary biochemistry of clusters based on salivary peptide profiles support the possibility that these profiles relate information on the underlying state of the oral ecosystem.

MALDI-TOF MS is a valuable tool for the profiling of biological samples. It is a fast, high-throughput method to obtain a molecular “fingerprint” and it can detect intact molecular species and post-translational modifications [50]. However, it also has some important limitations [51]. First of all, it is semi-quantitative. MALDI-TOF MS does not allow absolute quantification and can only quantify relative differences between samples for particular peaks. Secondly, a definitive identification of peaks by MALDI-TOF MS/MS is challenging, and for some peaks only putative identities may be obtained by matching m/z values [52]. Other top-down proteomic methods such as HPLC-ESI-MS are superior in this regard but provide lower throughput [53]. A few previous studies using top-down proteomic platforms (integrating multiple, high resolution mass-spectrometry systems and sample fractionation strategies) have compiled datasets of peptide identities, structures and m/z values that provide valuable reference points for other studies on saliva [19, 27]. These works were also used in the present MALDI-TOF MS study in the attempt to match the m/z values of unidentified peaks to known salivary peptides. While the preparation of the saliva samples in these studies (involving acidification with 0.2% TFA (1:1) prior to sample centrifugation) was different compared to the present study, this should not affect peptides in the mass range being discussed (2-12 kDa). The TFA acidification precipitates several of the large, abundant salivary proteins (e.g. amylase, mucins) while leaving the smaller proteins and peptides solubilized for analysis [43].

In summary, this study gives an overview of the peptide profile variation in the saliva of 268 healthy young adults and examines the 4 subgroups of individuals found in the study population based on these profiles. Salivary P-C peptide and its fragmentation pattern had an important role in delimiting these subgroups. Significant differences were found in the underlying functional salivary biochemistry of the subgroups with regard to the activity of lysozyme and chitinase, two enzymes relevant for oral health which are involved in the salivary innate defense system.
Acknowledgements

We wish to thank Nienke Hennequin-Hoenderdos, Dagmar Else Slot and Eveline van der Sluijs for their excellent work developing the clinical protocol and for coordinating and monitoring the clinical part of this study. We are grateful to Hein Trip for his contribution in setting up the MALDI-TOF MS assay protocol.
References


52. Chaiyarit P, Taweechaisupapong S, Jaresitthikunchai J, Phaonakrop N, Roytrakul S. Comparative evaluation of 5–15-kDa salivary proteins from patients with

Supporting Information

**S1 Fig.** Exemplification of the MALDI-TOF spectra processing workflow. (A) Raw spectrum. (B) Base line subtraction and normalization for total area under the curve. (C) De-noising. (D) Peak detection.
**S2 Dataset.** Peaks (with m/z labels). SPSS data file containing MALDI-TOF peaks m/z and intensities for each study participant.

**S3 Fig.** MALDI-TOF MS/MS spectrum for P-C peptide Fr. 15-35. Peptide amino acid sequence: GPPPPPGKPQGPPQGGRPQ.
S4 Fig. Example spectra from each of the 4 clusters. (A) Cluster I. (B) Cluster II. (C) Cluster III. (D) Cluster IV.
On the ecosystemic network of the salivary microbiome in systemically healthy young adults

Egija Zaura#, Bernd W. Brandt#, Andrei Prodan, Joost M. Teixeira de Mattos, Sultan Imangaliyev, Jolanda Kool, Jean Marie Buijs, Ferry LPW Jagers, Nienke L. Hennequin-Hoenderdos, Elena Nicu, Dagmar E. Slot,
Henk S. Brand, Enno C.I. Veerman, Michiel Kleerebezem, Bruno G. Loos, Fridus G.A. van der Weijden, Wim Crielaard, and Bart J.F. Keijser

# equal contribution

Submitted to: International Society for Microbial Ecology (ISME) Journal
Abstract

Although the mechanisms that underlie the development of oral diseases have been studied for decades, our understanding of oral health and of the boundaries of a healthy oral ecosystem is still limited. Here we present a (partial) quantification of the ecobiological relations between the salivary microbiome, salivary metabolome, and the host-related factors including salivary protein composition. Unstimulated saliva was collected from 268 systemically healthy adults after overnight fasting. The heterogeneity of the microbiome (16S rRNA gene sequencing) and metabolome (LC/MS-MS and GC/MS) as well as the mutual interrelations between these datasets and host-related biological and environmental parameters were assessed. Individuals were grouped into five microbiome- and four metabolome-based clusters that significantly related to biochemical parameters of saliva. Low salivary pH and high lysozyme activity were associated with high proportions of streptococcal phylotypes and increased membrane lipid degradation products. Gender-specific differences in the microbiome and metabolome were observed relating to salivary pH and dietary protein intake. Microbial taxa showed either strongly positive or negative correlations with proteolytic and saccharolytic salivary metabolites. This dichotomy together with the relation with specific host parameters suggests the presence of highly specialized ecotypes of a healthy oral ecosystem. An over-specialization either toward proteolytic or saccharolytic ecotype may indicate a shift toward a dysbiotic state.
Introduction

The oral ecosystem consists of hundreds of bacterial taxa and other microorganisms that interact with each other [1, 2]. The importance of understanding the role of the oral microbiome in maintaining oral and general health has recently been recognized, resulting in vast number of studies describing the bacterial component of this ecosystem [3, 4].

Saliva, besides its functions related to oral health and wellbeing (e.g., lubrication, taste experience and wound healing) is the main source of nutrients for oral microbial communities [5]. Changes in the local environment (e.g., prolonged periods of low pH) contribute to the transition of a healthy oral ecosystem from homeostasis towards dysbiosis (e.g., driving the pathogenesis of dental decay or caries)[6]. Nevertheless, the complex relations between the components of a healthy oral ecosystem so far have not been addressed.

Here, we aimed at describing the boundaries of a healthy oral ecosystem by deciphering the relations between salivary bacterial composition (microbiome), salivary metabolites (metabolome) and the host-related biological and environmental parameters in a systemically healthy young adult population.
Materials and Methods

This cross-sectional observational study was carried out on 268 systemically healthy young adults (aged 18-32 years) after overnight fasting and refraining from all oral hygiene procedures for 24h before the appointment. The inclusion and exclusion criteria of this study (Dutch Trial Register number NTR3649), have been described previously [7, 8]. Detailed materials and methods are available as Supplementary Methods. The study was conducted according to the Declaration of Helsinki [9]. The study protocol was reviewed and approved by the Medical Ethics Committee of the Academic Medical Center of Amsterdam (2012_210#B2012406), and registered at the Dutch Trial Register (NTR3649).

Saliva collection and processing

Unstimulated saliva was collected as described previously [8] in two sessions of 5 min with a 5 min interval between the two samples. Sample collection occurred between 9 and 10 am, before any food or drink intake that day, by drooling into a sterile vial cooled on ice. The first sample was immediately aliquoted into two vials and stored at -80° C for microbiological and metabolome analyses, respectively, while the second sample was processed for biochemical analyses [8]. Assessments included salivary flow rate measurements, pH and buffered pH measurements, determination of the total protein content, salivary mucins (MUC5B and MUC7), albumin, lactoferrin, secretory-IgA, cystatin S and enzymatic activities of amylase, chitinase, lysozyme and protease.

Dietary assessment

A food frequency questionnaire (FFQ) was developed to assess habitual dietary intake in the previous month. The FFQ was based on consumption data of 20-40 year old participants of the Dutch National Food Consumption Survey of 2007-2010 [10] and the Dutch food composition database [11]. The FFQ included questions on frequency and amount of intake of 130 food items. Selection of food items and calculation of the weighted average nutrient composition of the food items were performed using the validated Dutch FFQ-tool™ [12].

16S rRNA gene amplicon sequencing and data processing

Microbial DNA was extracted and processed for amplicon sequencing of the V4 hypervariable region of the 16S rRNA gene on the Illumina MiSeq platform (see Supplementary methods). The sequences were processed with mothur v.1.31.2 [13]. The
data were stored on a dedicated server at TNO and are available upon request via https://dashin.eu/download/TIFN_WP1_saliva/.

**Metabolome assessment**

Saliva samples were extracted, prepared and processed at Metabolon (Durham, NC, USA) as described previously [14] (Supplementary methods). The normalized metabolome dataset provided by Metabolon was range-scaled between 0 and 10. Metabolites with a single value were omitted from statistical analysis, resulting in a dataset with 493 metabolites.

**Statistical analyses**

Spectral clustering (SC) was performed using the neighborhood co-regularized SC algorithm [15], based on the SC method [16] (Supplementary methods).

Principal component analysis (PCA), permutational analysis of variance (PERMANOVA), Shannon diversity index and Chao-1 estimate of species richness were calculated using PAST software v. 3.04 [17].

Inter-cluster differences in host parameters and in microbiome alpha diversity were assessed using the Kruskal-Wallis and Mann-Whitney test and the differences in gender by the Pearson Chi-square test in SPSS version 21. Benjamini & Hochberg False Discovery Rate correction of the $p$-values for multiple comparisons was performed in R.

The LDA Effective Size (LEfSe) biomarker discovery tool [18] was used with the “one against all” strategy for multiclass analysis and logarithmic LDA score threshold of 2, $p < 0.05$.

Assessment of significant patterns of microbial co-occurrence or mutual exclusion was performed using CoNet [19] and visualized in Cytoscape v. 3.3.0.

High-dimensional feature selection by Significance Analysis of Microarrays (SAM) analysis [20] in the TIGR MultiExperiment Viewer (111.tm4.org/mev.html) was used to compare scaled metabolite data among metabolome sample clusters and to assess the differences in host-related SC clusters and microbiome SC clusters by metabolites.

Elastic Net regression [21] with Stability Selection [22] was used to select features (OTUs or metabolites) that were related to the host environmental variables.

The relation between the microbiome and the metabolome datasets was assessed using Spearman Correlation with Bonferroni-corrected $P$-values. The concordance between the two datasets was assessed using Procrustes Analysis [23] and the Mantel test [24] implemented in QIIME [25] v.1.8.0, both with 10000 random permutations.
Results

In this study, we aimed at quantifying the heterogeneity of and the mutual interrelationship between the bacterial composition of unstimulated saliva (microbiome), salivary metabolites (metabolome), and host-related biological and environmental parameters in a systemically healthy young adult population. The salivary biochemical parameters [8] and clinical parameters [7] have been previously reported. Here, we first assessed the boundaries of the salivary microbiome and metabolome and then related these datasets to different host parameters and to each other.

Demographics and dietary habits of the study population

Out of the total of 336 participants screened, 268 systemically healthy individuals (150 males, 118 females, mean age 22.6 years, SD 2.6, range 18-32) were eligible and included in the study (Supplementary table S1). The food questionnaire was completed by 255 individuals of whom 16 (6.3%) reported a vegetarian lifestyle (Supplementary table S2).

Heterogeneity of salivary microbiome

Of the 268 saliva samples, one sample was lost during sample processing. In total, sequencing produced 11.66 million reads of which 78.3% passed through the processing pipeline, resulting in 34,211 reads per sample (SD 10,164, range 13,434 – 67,729). The sequences were clustered into 2,758 operational taxonomic units (OTUs). After filtering the dataset, 912 OTUs, classified into 95 genera or higher taxa belonging to 11 microbial phyla, remained in the dataset (Supplementary table S3). After subsampling at 12,000 reads/sample, an individual sample contained on average 345.5 OTUs (SD 70.4; median 364; range 175 – 500). Twenty OTUs (53% of the reads) were present in all samples, while 63 OTUs (82% of the reads) were found in 95% of the samples (Supplementary table S3). The largest ubiquitous OTUs were classified as *Streptococcus dentisani/infantis/mitis/oralis* (OTU4, 13% of all reads), *Veillonella atypica/dispar* (OTU1, 9.4% of reads), *Haemophilus parainfluenzae* (OTU2, 8% of reads) and *Prevotella melaninogenica* (OTU21, 6.3% of reads). To assess the heterogeneity within the salivary microbiome dataset, Spectral Clustering (SC) was performed. This identified three distinct sample co-occurrence clusters (SC1: N=191; SC2: N=59, SC3: N=17) (Figure 1A). The microbiome taxonomic profiles within the largest cluster (SC1) still showed high heterogeneity, particularly due to differences in relative abundance of genus *Neisseria* (Figure 1B). Therefore, SC 1 was further subdivided into three sub-clusters (SC1.1: N = 48; SC1.2: N = 63; SC1.3: N = 80). The resulting five clusters were all statistically significantly different among each other (Fig. 1C).
Figure 1. Heterogeneity of salivary microbiome: A) Spectral Clustering co-occurrence plot of microbiome samples ordered along X- and Y-axis according to the co-occurrence matrix: the more similar the microbiome profiles, the closer they are on the axis. Co-occurrence values range from 0 (samples never cluster) to 1 (samples always cluster together) after multiple k-means clustering assignments. Machine-learning tools defined three main sample clusters (SC1, SC2 and SC3) where the largest SC was subdivided into three subclusters (SC1.1, SC1.2, SC1.3). Below the graph are Spectral Clustering labels according to salivary microbiome, metabolome and host parameter (‘Host’ SC) datasets. B) Microbiome profiles at the genus level. C) Principal Component Analysis (PCA) of microbiome samples. All five microbiome sample clusters differed significantly among each other (PERMANOVA). D) 33 most significantly discriminatory OTUs between the microbiome clusters (LDA > 3.5, \( p < 0.05 \), LEfSe). E) Estimated species richness (Chao-1) per individual sample, ordered as in 1A-B. SC1.2 samples showed significantly lower diversity than other samples (\( p < 0.001 \)).

To determine which OTUs contributed to the observed differences among the sample clusters, the LDA Effective Size (LEfSe) biomarker discovery tool was used (Supplementary table S4, Fig. 1D). The smallest cluster, SC3 (\( N = 17 \)), had the highest
proportion of OTUs classified as *Veillonella atypica/Veillonella dispar* (OTU1) and *Prevotella histicola* (OTU35, OTU94) and the lowest proportion of *Streptococcus, Rothia* and *Granulicatella* among the five sample clusters (Figure 1D). Cluster SC1.3 was discriminated by the highest proportion of *Neisseria subflava* (OTU23) and *Haemophilus parainfluenzae* (OTU2); cluster SC1.2 by *Streptococcus salivarius/Streptococcus vestibularis* (OTU14), *Streptococcus australis/Streptococcus parasanguinis* (OTU15) and *Granulicatella adiacens* (OTU41); cluster SC1.1 by *Prevotella* sp. HOT313 (OTU31) and *Paraprevotella/Alloprevotella* sp. HOT308 (OTU234); and cluster SC2 by the highest proportion of *Streptococcus infantis/Streptococcus mitis/Streptococcus oralis* group (OTU4), *Streptococcus gordonii* (OTU86) and *Rothia aeria* (OTU8) (Fig. 1D).

The samples that clustered into SC1.2 showed the lowest alpha diversity: species richness (number of OTUs/sample), the Shannon Diversity and Chao-1 index were significantly lower in SC1.2 samples compared to the samples from any other cluster (Mann-Whitney test, $P<0.001$). For instance, the estimated species richness (Chao-1) of the SC1.2 samples was 340 (SD=92), while SC1.1 had 508 (SD 69), SC1.3 – 491 (SD 72), SC2 – 525 (SD 63) and SC3 – 490 (SD 52) taxa (Fig. 1E).

Next, the association networks of the most prevalent taxa per SC cluster (Fig. 2) were assessed. The bacterial co-occurrence network of cluster SC3 had the highest complexity due to the highest number of co-occurring taxa (66) and the highest average number of neighbors (2.88), followed by SC2 (50 taxa and 2.84 neighbors), SC1.3 (43 taxa and 2.65 neighbors), SC1.2 (36 taxa and 1.83 neighbors) and SC1.1 (31 taxa and 1.87 neighbors). The most frequent co-occurrences were found between the taxa belonging to the same genus. The highest number of neighbors belonging to different genera was found in SC1.3, where *Prevotella salivae* co-occurred with *Streptococcus salivarius/vestibularis, Leptotrichia, Selenomonas, Megasphaera micronuciformis, Veillonella atypica/dispar* and *Atopobium parvulum*. Depending on the SC cluster, different prevotellae species, *Megasphaera micronuciformis*, Fusobacteria or *Veillonella atypica/dispar* displayed the most co-occurrences with the other taxa. Interestingly, streptococci, although highly abundant in most of the sample clusters, did not form complex networks with other genera.
Figure 2. Salivary microbiome co-occurrence analysis output per individual microbiome SC sample cluster. The size of the nodes is related to the relative abundance of the taxa; the
color of the node indicates the connectivity to the other nodes (red – low number of neighbors, green – high number of neighbors). Analysis was performed using CoNet in Cytoscape. Taxa at the species level are obtained from OTUs made taxonomically non-redundant using HOMD database. Taxonomic names marked with * have been truncated for legibility.

**Heterogeneity of salivary metabolome**

From the 268 saliva samples, 266 metabolome profiles were obtained (two samples had insufficient volumes for the metabolome analysis). The entire salivary metabolome dataset contained 535 metabolites, 79 of which were present in all samples and 275 in at least 75% of the samples (Supplementary table S5). On average, 337 metabolites (SD 52, range 214-430) were identified in an individual saliva sample. The peptide pathways contained the highest number of metabolites, while metabolites belonging to amino acid and lipid pathways contributed most to the metabolites that were found across all samples (Supplementary table S5). Of the 535 metabolites, 396 were identified to a known biological pathway.

As for the microbiome data, spectral clustering (SC) was applied to assess the heterogeneity within the metabolome dataset. The SC co-occurrence analysis resulted in four clusters, where the majority of samples (74%) showed high co-occurrence of metabolites and clustered together (SC1, N = 198), followed by three smaller clusters with 44 (SC2), 19 (SC3) and 5 samples (SC4), respectively (Fig. 3A). The PCA analysis and PERMANOVA confirmed the spectral clustering outcome (Fig. 3B). Samples in SC4 had the highest average number of metabolites (417 metabolites/sample, SD 6.8), followed by SC3 (409, SD 11) and SC2 (386, SD 14), while samples in the SC1 had the lowest number of metabolites per sample (315, SD 41) compared to the other clusters ($p < 0.0001$).
Figure 3. Heterogeneity of the salivary metabolome: A) Spectral Clustering co-occurrence matrix clustering the metabolome dataset into four clusters: SC1 (N = 198 samples), SC2 (N = 44), SC3 (N = 19) and SC4 (N = 5). B) Principal Component Analysis (PCA) of metabolome samples. All metabolome sample clusters differed significantly among each other (PERMANOVA). C) Significantly positively and significantly negatively associated metabolites between samples belonging to the SC1 and the other clusters. Of the 217 negatively associated metabolites, only the 6 with the highest fold change are shown. D) OTUs of the salivary microbiome discriminating between the metabolome clusters (LEfSe, $p < 0.001$; LDA > 3).
Significance Analysis of Microarrays (SAM) identified 222 metabolites that discriminated the large sample cluster (SC1) from the other three clusters (Supplementary table S6). Only 6 of the 222 metabolites – unidentified metabolites X19870 and X13230, phosphoethanolamine, glycerol, citrate and urea – were significantly higher in SC1 (Figure 3C). As expected from the PCA analysis, the metabolites associated with the three smaller clusters showed a gradient in their abundance (Fig. 3C).

The relation between the microbiome and the biochemical salivary parameters

As reported previously [8], fourteen host-related biochemical parameters of relevance for oral health were analyzed from the same saliva sample set. These included salivary flow rate, pH and various salivary enzymes. Spectral Clustering of these host factors resulted in two distinct sample clusters, where the smallest cluster (SC1, N=22, 8.3%) was differentiated by low salivary pH and low buffered pH and by high lysozyme activity and salivary mucin MUC7 content compared to the SC2 (N=245), which comprised the rest of the study population (Fig. 4A).
Salivary microbiome, metabolome and host environment

Figure 4. Salivary biochemistry data A) stratified into Spectral clusters SC1 (N = 22) and SC2 (N = 245) (Fig. 4 with permission from Prodan et al., 2015) and compared with the microbiome data B) where 12 most discriminatory OTUs between the two biochemistry clusters (LEfSe, $p < 0.05$, LDA > 3) are shown. In total, 42 OTUs were significantly discriminatory (LDA > 2). C) Abundance of the top four discriminatory OTUs per biochemistry cluster (red – SC1, blue – SC2). D) Metabolites that significantly discriminated two biochemistry clusters (SAM; delta 0.8; 100 permutations, 0 false significant genes).

Of the 912 microbiome OTUs, 99 significantly discriminated between the two host-related salivary biochemistry-based SC clusters (Supplementary table S7). *Streptococcus* OTU4 (*Streptococcus dentisani*/infantis/ mitis/oralis), *Haemophilus parainfluenzae* (OTU2), *Granulicatella adiacens* (OTU41) and *Rothia dentocariosa* (OTU29) were the most positively associated taxa with cluster SC1 samples (the ‘low pH cluster’), while *Veillonella* OTU1 (*Veillonella atypica*/Veillonella dispar), *Prevotella* OTU10 (*Prevotella sp. HOT 310*) and *Fusobacterium periodonticum* (OTU22) predominated in the SC2 samples (Figure 4B-C). Additionally, the microbiomes of SC1 samples had a significantly
lower Shannon Diversity index (3.498, SD 0.39) compared to SC2 samples (3.716, SD 0.27) ($p = 0.01$).

Next, the relation between the individual host-related biochemical salivary parameters and the microbiome sample spectral clusters was assessed. Samples belonging to cluster SC3 differed most from samples in the other clusters, with significantly higher salivary chitinase and albumin and significantly lower lysozyme activity, while SC1.2 and SC2 had significantly lower salivary pH and buffered pH compared to the other clusters (Fig. 5A).
Figure 5. Five out of 14 measured host-related salivary parameters where significant differences among the samples belonging to the different A) microbiome and B) metabolome clusters were observed. The lines indicate significantly different clusters ($p < 0.05$, FDR-corrected for multiple comparisons).
For cluster-independent analysis, Elastic Net regression was used to relate the salivary microbiome data with the individual salivary biochemical parameters. Salivary pH, buffered pH and lysozyme activity could be predicted with good accuracy based on microbiome composition (Fig. 6). Among the 10 most stable OTUs that predicted salivary buffered pH, eight OTUs (belonging to Streptococcus, Actinomyces, Granulicatella adiacens, Rothia dentocariosa and Prevotella oris) decreased in abundance with increasing pH, while only two OTUs (classified as Capnocytophaga granulosa and Neisseria) were more abundant at higher pH (Supplementary table S8, Figure 6B). Salivary lysozyme activity was positively associated with mainly Gram-positive taxa such as streptococci and actinomyces, and negatively associated with Gram-negatives such as prevotellae, selenomonae and neisseriae (Supplementary table S9, Fig. 6D).

**Figure 6.** Results of Elastic Net regression analysis on salivary microbiome dataset in predicting A-B) buffered salivary pH and C-D) salivary lysozyme activity, where distribution of the four most stable and abundant OTUs that predicted B) buffered pH and D) lysozyme activity are shown divided into the quartiles of the respective variable. Buffered pH was divided into the following quartiles: Q1: pH4.1-5.9, Q2: pH5.9-6.3, Q3:
pH 6.3-6.6, Q4: pH 6.6-7.3. Lysozyme activity was divided into Q1: 28-464, Q2: 464-1514, Q3: 1514-2491, Q4: 2491-5635.
The relation between the metabolome and the biochemical salivary parameters

Only eight of the 493 metabolites discriminated between the two clusters based on biochemistry data: phosphoethanolamine was more abundant in cluster SC1 (the ‘low pH’ cluster), while creatinine, 2-amino adipate, 3,4-hydroxyisovaleriate, 2-piperidinone, alpha hydroxyisovaleriate, deoxycarnitine and unidentified metabolite X12944 were significantly more abundant in the biochemistry SC2 (Fig. 4D).

Next, the relation between the biochemical salivary parameters and the metabolome sample clusters was assessed. Similarly to metabolites, the host-related parameters followed a gradient: pH, buffered pH, chitinase activity, albumin and MUC5B increased while lysozyme activity decreased from SC1 through SC4 when different metabolome sample clusters were compared (Fig. 5B).

Based on the metabolites in saliva, Elastic Net regression predicted eight of the fourteen host parameters: salivary pH ($r^2 = 0.28$, NMRSE = 12.8), buffered pH ($r^2 = 0.37$, NMRSE = 13.9), albumin concentration ($r^2 = 0.25$, NMRSE = 9.5), total protein ($r^2 = 0.296$, NMRSE = 12), secretory-IgA ($r^2 = 0.45$, NMRSE = 7.4), and the activity of lysozyme ($r^2 = 0.35$, NMRSE = 13.9), amylase ($r^2 = 0.28$, NMRSE = 13.9), and proteases ($r^2 = 0.23$, NMRSE = 10.58). For instance, albumin concentration was positively associated with pro-inflammatory metabolites (arachidonate, dihomo-linoleate and 1-stearylglycerophosphoinositol) and negatively with anti-inflammatory palmitoyl ethanolamide, while low pH, low buffered pH and high lysozyme activity were associated with membrane lipid degradation products phosphoethanolamine, ethanolamine and glycerol.

Gender and saliva

As reported previously, salivary biochemical parameters exhibited gender-related differences: salivary pH, buffering capacity, protein content, MUC5B, secretory-IgA, and chitinase activity were all lower in females compared to males, while MUC7 and lysozyme activity were higher in females [8]. Here, we assessed whether salivary microbiome and metabolome also showed gender-related differences.

LEfSe identified 65 OTUs that differentiated between males and females (Supplementary table 10), of which 44 OTUs (including 19 streptococcal OTUs) were significantly more abundant in females. Male salivary microbiomes had a higher abundance of OTUs classified as Veillonella, Prevotella and Megasphaera. Females had a significantly higher number of estimated taxa (Chao-1) than males (females mean 484, SD 93 versus
male mean 450, SD 108) \( (p = 0.013) \), however there was no difference in Shannon Diversity index between males (mean 3.695, SD 0.27) and females (3.699, SD 0.30). The salivary microbiome SC clusters did not show any relation with gender (results not shown).

The SC clusters based on metabolome data did differ by gender of the subjects \( (p = 0.009) \): there were significantly more males than females in the smaller metabolome sample clusters: SC1 contained 66.9% of all males and 83.8% of all females, SC2 - 20.3% and 12%, SC3 - 9.5% and 4.3% of all males and females, respectively, while SC4 contained only males. Additionally, the SAM analysis revealed that 205 metabolites were differentially more abundant by gender (Supplementary table S11). Of these, only three metabolites, namely glycerol, caffeine, and 3-hydroxyindolin-2-one, were overrepresented in females. The top five most abundant metabolites in males compared to females were glycerate, tryptophylserine, unidentified metabolite X19863, proline and alanylphenylalanine.

While males showed a lower number of microbial taxa, they had significantly more metabolites in their saliva (mean number of metabolites 349, SD 48) than females \( (321, SD 52) \) \( (p = 0.000013) \).

**Diet and saliva**

Total dietary protein correlated positively with isovalerycarnitin \( (r = 0.221, p < 0.0001) \) and negatively with glycerol \( (r = -0.221, p < 0.0001) \). Interestingly, dietary protein also correlated with buffered pH \( (r = 0.266, p = 0.00003) \).

**Relation between salivary metabolome and microbiome**

First, the relationship between different metabolome sample clusters and microbial taxa was assessed. Of the 912 OTUs, 207 were significantly discriminatory among the metabolome clusters (Supplementary table S12). Of these, 30 OTUs discriminated one of the metabolome clusters from the others at an LDA score 3 or above. The largest cluster (SC1, N=198 samples) had a significantly higher abundance of OTU4 \( \text{(Streptococcus dentisani/infantis/mitis/oralis group)} \), OTU2 \( \text{(Haemophilus parainfluenzae)} \) and OTU27 \( \text{(Porphyromonas sp. HOT 278 and HOT279)} \), while other clusters were dominated by OTUs classified as veillonellae and prevotellae (Fig. 3D).

Then, the differences in metabolome between the five microbiome sample clusters were analyzed. Of the 493 metabolites, 64 were significantly discriminatory (SAM analysis, median FDR 0%, delta 0.0033) (Supplementary figure S1), the majority of which were significantly more abundant in the samples from the smallest microbiome cluster \( \text{(SC3, N = 17)} \) compared to the other samples. Among these, the largest difference between SC3 and the other samples was observed in the higher abundance of the dipeptides.
histidylhistidine, leucylserine, and leucyltyrosine, of phenylacetate, and of two unidentified metabolites - X18037 and X18165.

The relationship between the microbiome and metabolome datasets was assessed by examining bivariate correlations between the top 400 most abundant OTUs and the 493 metabolites (without a priori categorizing the samples into clusters). Of the top 400 OTUs, 167 correlated significantly with one or more (maximum 232) metabolites and of the 493 metabolites, 282 correlated with at least one (maximum 96) OTU (Bonferroni-corrected $p < 0.05$; $r < -0.2$ or $r > 0.2$). Interestingly, a clear dichotomy in the relation of these OTUs with metabolites was observed: the taxa had either strong positive or strong negative associations with metabolites (Supplementary figure S2). The most significant correlations ($r < -0.5$ or $r > 0.5$) centered around two OTUs: one of the largest OTUs (13% of all reads) - OTU4 (Streptococcus dentisani/infantis/mitis/oralis group) - and a relatively small OTU (0.8% of all reads) - OTU66 (Megasphaera micronuciformis) (Fig. 7). OTU4 associated negatively with 228 metabolites (35 with $r < -0.5$) and positively with only four metabolites (urea: $r = 0.45$, phosphoethanolamine: $r = 0.39$, glycerol: $r = 0.33$ and glycerol-3-phosphate-G3P: $r = 0.32$). In contrast, OTU66 correlated negatively only with three metabolites (urea: $r = -0.34$, phosphoethanolamine: $r = -0.33$ and choline phosphate: $r = -0.32$) and positively with 224 metabolites (13 with $r > 0.5$). Among the metabolites the strongest network was observed for 3-phenylpropionate (hydrocinnamate), which had four negative and 29 positive correlations with OTUs. Of these, seven strongly correlating edges ($r > 0.5$) were with OTUs classified as Prevotella (2 OTUs), Alloprevotella, Fusobacterium periodonticum, Leptotrichia, Peptostreptococcus stomatis and Eubacterium sulci (Fig. 7).
Figure 7. The network of the most significant Spearman correlations (rho < -0.5 or > 0.5) between the top 400 most abundant OTUs and the 493 metabolites. OTUs are shown with blue circle, the diameter of which is proportional to the abundance. Metabolites are shown with yellow circle. Positive correlations are indicated with green lines, whereas negative correlations are red. Taxonomic names marked with * have been truncated for legibility.

Furthermore, the significance of the correlation of the two data sets was tested using the Mantel test on the Bray-Curtis distance matrices of the metabolome and the microbiome data. The matrices correlated significantly ($p = 0.001$), though with a relatively low correlation coefficient ($r = 0.199$).

Additionally, the PCA plots of both datasets were superimposed using Procrustes analysis, where the majority of the metabolome samples clustered at the center of the plot (Supplementary figure S3). Although Procrustes output was significant ($p < 0.0001$), the sum of the squared deviations was very high ($M^2 = 0.905$), indicating low similarity between the principal components of the two datasets.
Discussion

We aimed at increasing the understanding of a healthy oral ecosystem. For this, unstimulated saliva samples were collected from 268 systemically healthy young adults. The same individual samples provided microbiome, metabolome and host-related salivary data. The salivary microbiome appeared to be heterogeneous, forming five distinct sample clusters. The individuals grouped into salivary microbiota clusters differed not only in microbial community structure and diversity, but also in salivary pH, buffered pH, albumin levels and activity of lysozyme and chitinase. No direct relation with the metabolome profiles could be observed. In contrast to the marked differences in microbiome composition, the majority of the individuals had highly similar salivary metabolome profiles. Interestingly, we observed a clear dichotomy in the correlations between salivary bacteria and salivary metabolites. Dipeptides and metabolites of amino acid catabolism displayed positive correlations with *Prevotella*, *Fusobacterium*, and *Veillonella* species, while showing negative correlations with saccharolytic bacterial species. We also found that the microbiome and the metabolome of saliva was gender-related, and that dietary protein intake correlated with salivary pH.

The most intriguing finding of this study was the strong dichotomy in salivary OTU-metabolite associations. Species positively associated with metabolites of amino acid metabolism were key members of microbiome clusters SC1.1, SC1.3, and SC3. A relatively high resting pH was found in the samples from these clusters, in correspondence with deaminase and decarboxylase activity under amino acid metabolism [26]. Conversely, microbiota clusters SC2 and SC1.2 were dominated by *Streptococcus* species and showed a negative correlation with dipeptides and other metabolites linked to amino acid metabolism. Instead, a positive correlation was found with glycerol-3-phosphate, choline phosphate, and phosphoethanolamine, all related to phospholipid metabolism.

We propose that the five salivary microbiota clusters identified in these systemically healthy individuals represent different ecological states or ‘ecotypes’ of the oral ecosystem, with different ecological properties and different levels of specialization. We found clusters that are functionally adapted to proteolysis and amino acid fermentation, as well as clusters functionally adapted for saccharolysis (Fig. 8). Based on our findings, a proteolytic adaptation is reflected in a higher dominance of proteolytic bacteria, and a higher resting pH. Cluster SC3 appears to display signs of advanced ecological specialization towards protein and amino acid metabolism, indicative of an increased potential for a dysbiotic shift towards an early inflammatory state. A striking observation for cluster SC3 is the extensive microbiota community network linkage, together with the higher levels of salivary albumin (which is not secreted by the salivary glands, but originates from gingival crevicular fluid and blood plasma), low levels of lysozyme and high levels of chitinase. On the other side of the spectrum, we found an ecological state that is more adept to saccharolytic functions, represented by clusters SC2 and SC1.2. These
states are dominated by streptococcal species. Even in a resting state and in the prolonged absence of external carbohydrate sources after an overnight fasting, the salivary pH in these individuals is lower compared to the clusters adapted to proteolysis. Interestingly, the saccharolytic cluster SC1.2 showed a marked reduction in species diversity, having the lowest number of ecological network connections, as well as a relatively high lysozyme level. We propose that cluster SC1.2 represents an early phase of acidogenic adaptation, specialized in fast sugar metabolism and prone to a more cariogenic state [26]. It is clear that longitudinal studies are required to establish the stability of these five ecosystem states, as well as their relation with the maintenance of oral health.

Figure 8. Proposed ecological states or ecotypes of the oral ecosystem and the positioning of the microbiome clusters according to these states. Dichotomy in bacteria-metabolite associations and the relation with salivary parameters is depicted in saccharolytic (left side) or proteolytic (right side) adaptations of the ecosystem. Based on the observed associations, the sample clusters SC1.2 and SC3 are positioned toward the dysbiotic state of the system, while clusters SC2, SC1.1 and 1.3 are positioned at the healthy state of the system.

Similar to ‘enterotypes’ in the gut [27, 28], the presence of ‘oral ecotypes’ suggests a role of host factors in shaping the microbial community structure. This was evidenced by our
findings on associations between the salivary biochemical parameters and microbial composition. One of the factors that related to the microbial community composition was the pH of saliva. This was not unexpected, since change in environmental pH is a well-known driver of microbial community activity and ecological shifts [29, 30].

Other interesting associations were found between salivary lysozyme activity and microbial composition. Lysozyme is an antimicrobial protein secreted by the salivary glands, as well as by neutrophils and macrophages, which is capable of lysing bacterial cell wall peptidoglycan [31]. Therefore, it was surprising to find a positive correlation between salivary lysozyme activity and streptococci, which are Gram-positive and would be targeted by lysozyme. Additional to its enzymatic antimicrobial activity, lysozyme is involved in aggregation of bacteria, which are subsequently cleared by swallowing [32]. It adsorbs to oral bacteria in a strain- and species-specific manner, where, among the species tested, the lowest adsorption of lysozyme was observed to different *Streptococcus mitis* strains [33]. Interestingly, the highly abundant OTU with the strongest positive relation with lysozyme activity in saliva was classified as *Streptococcus dentisani/infantis/mitis/oralis* or mitis group streptococci (OTU4). These mitis group streptococci are known as primary colonizers of oral surfaces and are associated with oral health [4]. Higher abundance of mitis group streptococci was not only associated with a higher lysozyme activity, but also with a lower salivary pH. In fact, lysozyme activity correlated negatively with salivary pH [8]. Since the optimum activity of lysozyme depends on pH and ionic strength [34], lysozyme activity, measured in vitro at pH 7, might indicate a compensatory over-expression of lysozyme at salivary pH below the functional optimum of this enzyme.

The simultaneous assessment of microbiome and metabolome allowed a direct comparison between these two datasets. As opposed to the heterogeneity observed in the salivary microbiome, a majority of the samples had highly similar metabolome profiles. This implies that overnight fasting saliva metabolism in different individuals carrying different microbiota is driven by similar processes and fits well to previously described functional redundancy of human microbial communities [35]. However, about one fourth of the individuals had clearly different (e.g. more proteolytic activity-driven) and more diverse metabolome profiles. These profiles related with several biochemical parameters of saliva, such as higher pH and higher albumin concentration, and with microbial communities dominated by anaerobes and Gram-negatives, the taxa associated with mature oral microbial communities and gingival inflammation [36]. Although no significant relation between clinical status (such as plaque levels or gingival bleeding) and salivary metabolome and microbiome was found (results not shown), the positive relation between salivary albumin and pro-inflammatory metabolites in these saliva samples suggests presence of inflammation [37].

Both microbiome and metabolome of saliva was gender-associated. Analysis of the host-related parameters had already shown that gender significantly affected a number of salivary biochemical parameters [8]. Males, generally more susceptible to periodontal
disease [38], presented with significantly higher salivary pH than females, who are more prone to dental decay [39]. Low salivary pH in females has been attributed to physiological factors such as the influence of sex hormones on salivary gland gene expression and a smaller salivary gland size in females [8]. Higher abundance of streptococci found in female saliva could be an indirect result of an ecological pressure due to this physiologically lower pH [40]. However, higher abundance of streptococci could directly contribute to a lower salivary pH, since streptococci are the main acid producers in a healthy oral cavity [41].

Even more striking were the gender differences in the salivary metabolome of the study participants. Males predominated in the smaller metabolome clusters and had a significantly higher number of metabolites in their saliva. A previous metabolome study using NMR spectroscopy and comparing 20 males with 20 females found that nearly all metabolites that were significantly different by gender were higher in concentration in males [42]. Interestingly, the two most gender-discriminative OTUs – the female-associated mitis group streptococcus (OTU4) and the male-associated *M. micronuciformis* (OTU66), were the two central OTU-metabolite network players. Since the total salivary protein concentration was higher in males [8], this could reflect in higher amino acid and peptide abundance in male metabolomes, selecting for the more proteolytic microbiome observed here.

Diet did not appear to influence the microbial microbiome or the metabolome of saliva, although this study included only a limited number of non-omnivores. A recent study specifically designed to address the influence of diet concluded that metabolome, but not the microbiome, was affected by diet [43]. Since details on controlling for timing of food consumption were not reported, saliva might have been collected considerably shorter after the last food intake compared to the current study with overnight fasting, confirmed by low peripheral blood glucose values during sample collection visit. We did however find a positive relation between total dietary protein intake and salivary pH. This indicates a potential role of diet in modifying salivary pH, one of the strongest environmental factors in the oral cavity.

In conclusion, unstimulated saliva of systemically healthy adults is microbially heterogeneous and this heterogeneity is not related to salivary metabolites, but to biochemical host parameters of saliva. Host-related parameters, including gender, appear to affect the oral ecosystem in multiple ways. The observed clear dichotomy in the bacteria-metabolite associations and the relation with specific host parameters suggest a presence of highly specialized ecotypes of a healthy oral ecosystem. An over-specialization either toward proteolytic or saccharolytic ecotype may indicate a shift toward a dysbiotic state.
Acknowledgments

The project was funded by TI Food and Nutrition, a public-private partnership on precompetitive research in food and nutrition. The public partners are responsible for the study design, data collection and analysis, decision to publish, and preparation of the manuscript. The private partners have contributed to the project through regular discussion.

We are thankful to Wendy de Wit, Elly van Deutekom-Mulder, and Carolien Bosch-Tijhof (ACTA, Department of Preventive Dentistry) for their assistance and help during the clinical, laboratory and data analysis stages of the study.
References


12. Molag M. Towards transparent development of food frequency questionnaires: scientific basis of the Dutch FFQ-TOOLTM: a computer system to generate, apply and process ffqs: Wageningen Universiteit (Wageningen University); 2010.


tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. Anal Chem. 2009;81.


Supplementary files:

Supplementary methods. This file contains detailed description of methods of this study.

Supplementary tables:

Supplementary table S1. Blood chemistry values, demographic and clinical parameters of study volunteers and salivary biochemical parameters.

Supplementary table S2. Results of the dietary assessment, based on food frequency questionnaire, completed by 255 of 268 individuals.

Supplementary table S3. Relative abundance of 912 OTUs per individual sample.

Supplementary table S4. Significantly discriminatory OTUs between the microbiome sample clusters. From the 912 OTUs, 443 discriminated significantly a sample cluster from the other clusters (LDA > 2, \( p < 0.05 \), LEfSe).

Supplementary table S5. Prevalence of salivary metabolites per metabolic pathway category. The prevalence is split into the overall prevalence (presence in at least a single sample), the prevalence in the majority (75%) of the 266 samples and the prevalence in all samples.

Supplementary table S6. Metabolites (\( N = 222 \)) that significantly discriminated between the metabolome SC 1 and the other clusters identified by Significance Analysis of microarrays (SAM). Of these, 6 metabolites were significantly higher in the SC 1.

Supplementary table S7. Salivary OTUs that significantly discriminated the two SC sample salivary biochemical parameter clusters (Prodan et al., 2015) (LEfSe, LDA > 2, \( p < 0.05 \)).

Supplementary table S8. Ten most stable OTUs that predicted salivary buffered pH using Elastic Net regression.

Supplementary table S9. OTUs that predicted salivary lysozyme activity using Elastic Net regression.

Supplementary table S10. Significantly discriminatory OTUs between males and females in salivary microbial profiles (LEfSe, LDA > 2, \( p < 0.05 \)).

Supplementary table S11. Metabolites (\( N = 205 \)) that significantly discriminated between males and females in salivary metabolome, using SAM analysis
Supplementary figures:

Figure S1. Metabolites that significantly discriminated among microbiome SC sample clusters, using SAM analysis.
Figure S2. Correlations between microbial taxa and metabolites. Average linkage hierarchical clustering, Spearman Rank Correlation matrix of Spearman r values of the significant OTUs (columns) vs significant metabolites (rows). Color scale visualizes r values from low (blue), to zero (black) to high (yellow) of the correlations. OTUs belonging to the same major genera are highlighted in color: *Prevotella* – yellow, *Veillonella* – green, *Streptococcus* – red, *Neisseria* – blue.
Figure S3. Relation between the microbiome and metabolome datasets: Superimposed PCA plots of microbiome and metabolome datasets (Procrustes analysis, $p < 0.0001$, $M^r = 0.905$). Grey edge segment – microbiome sample; black edge segment – metabolome sample.
CHAPTER 5

Effect of experimental gingivitis induction and erythritol on the salivary metabolome and functional biochemistry of systemically healthy young adults

Andrei Prodan, Sultan Imangaliyev, Henk S. Brand, Martijn N.A. Rosema, Evgeni Levin, Wim Crielaard, Bart J.F. Keijser, and Enno C.I. Veerman

Metabolomics 2016, 12(10):147
Abstract

Introduction Understanding the changes occurring in the oral ecosystem during development of gingivitis could help improve prevention and treatment strategies for oral health. Erythritol is a non-caloric polyol proposed to have beneficial effects on oral health.

Objectives To examine the effect of experimental gingivitis and the effect of erythritol on the salivary metabolome and salivary functional biochemistry.

Methods In a two-week experimental gingivitis challenge intervention study, non-targeted, mass spectrometry-based metabolomic profiling was performed on saliva samples from 61 healthy adults, collected at five time-points. The effect of erythritol was studied in a randomized, controlled trial setting. Fourteen salivary biochemistry variables were measured with antibody- or enzymatic activity-based assays.

Results Bacterial amino acid catabolites (cadaverine, N-acetylcadaverine, and α-hydroxyisovalerate) and end-products of bacterial alkali-producing pathways (N-α-acetylnornithine and γ-aminobutyrate) increased significantly during the experimental gingivitis. Significant changes were found in a set of 13 salivary metabolite ratios composed of host cell membrane lipids involved in cell signaling, host responses to bacteria, and defense against free radicals. An increase in mevalonate was also observed. There were no significant effects of erythritol. No significant changes were found in functional salivary biochemistry.

Conclusions The findings underline a dynamic interaction between the host and the oral microbial biofilm during an experimental induction of gingivitis.
Introduction

Oral health is of paramount importance for functions like chewing, tasting, speech, and social interaction [1]. Oral diseases such as caries, gingivitis, and periodontitis have high prevalence rates, reduce quality of life, and can have a profound negative effect on general health [2, 3]. The oral cavity is a specialized ecological niche with thousands of microbial species, and the site of complex interactions between oral microbiota, salivary components, and host immune factors [4]. The commensal oral microflora has adapted to the human host during a long evolutionary relationship. Several studies have uncovered considerable cross-talk between oral microflora and host, mediated by collaborative cell sensing, immunomodulation, and discriminatory systems [5-7]. However, an inadequate level of oral hygiene, increased consumption of carbohydrates, and excessive host inflammatory responses - among other factors - can alter the composition and the functional profile of the oral microflora and shift this ecosystem away from homeostasis and towards a state of disease [4, 8]. Understanding these changes in the oral ecosystem is an essential step in the development of prevention strategies and improved treatments for oral health.

Metabolomics is the study the small molecules (<2000 Da) found in living cells or environments [9]. Metabolomics is complementary to transcriptomics and proteomics, with the advantage of being further ‘downstream’ of gene expression and thus more closely reflecting the activity of the biological system at a functional level [10].

Studies of the changes in metabolomic profiles caused by various life-style, pharmaceuticals, and diseases (e.g. diabetes, cancer) have offered valuable biological insights and uncovered biomarkers for diagnostic purposes [9, 11]. Metabolomics is also increasingly used in oral sciences. The salivary metabolite profile of healthy individuals has already been catalogued [12]. Several studies have found a periodontitis-specific metabolomic fingerprint in the saliva and gingival crevicular fluid of periodontitis patients, characterized by increased concentrations of degradation products of macromolecules and intensified host-bacterial interactions [13-15]. However, no mass-spectrometry-based studies of salivary metabolome changes occurring in healthy individuals during the development of gingivitis have yet been published. The changes in the salivary metabolome during different stages of oral plaque overgrowth and oral biofilm maturation, and during the corresponding reaction of the host defense systems to these challenges, could yield important clues to the biological mechanisms that maintain a healthy mouth.

Erythritol ((2R,3S)-butane-1,2,3,4-tetraol) is a non-caloric polyol used as a sweetener. Its safety has been well documented: it is classified as GRAS (Generally Recognized As Safe) by the U.S. Food and Drug Administration and is considered safe for use as a food additive (E968) by the European Food Safety Authority [16]. Erythritol is non-cariogenic, inhibits Streptococcus mutans strains in vitro, and lowers plaque levels,
caries rates, and caries progression in in vivo trials [17-20]. Erythritol may therefore exert a stabilizing effect on the oral homeostasis.

The primary aim of this study was to examine the changes in the salivary metabolome of young healthy adults during a two-week experimental gingivitis challenge intervention, and to explore the potential effect of a daily intake of erythritol on these changes during a controlled, randomized clinical study. Non-targeted, multi-platform, mass spectrometry (MS)-based metabolic profiling was performed on saliva samples from the study participants. The secondary aim was to look at potential changes in host salivary biochemistry. An array of targeted functional biochemistry analyses (antibody-based ELISAs and enzymatic activity assays) were used to measure salivary proteins with known relevance for oral health: mucin MUC5B, albumin, secretory-IgA, lactoferrin, amylase, lysozyme, chitinase, and proteases, alongside salivary pH, buffered pH, flow rate, and total protein content.
Materials and Methods

Study design, study population and inclusion/exclusion criteria

The study was carried out within the framework of the TI Food and Nutrition project ‘An exploratory study on the dynamic (microbial, biochemical and immunological) interactions of the oral ecosystem during induction of mild gingival inflammation. Dynamics of a healthy oral ecosystem’.

Whole unstimulated saliva was collected in a single-center, challenge intervention, randomized study at the Academic Centre for Dentistry Amsterdam (ACTA). The study population was a convenience sample of systemically healthy adults. The CONSORT flow diagram of the study is shown in Fig. 1.

The challenge intervention was based on a full-mouth modified experimental gingivitis protocol [21]. The first stage was a two-week baseline period in which the treatment group started the use of erythritol (Day -14 to Day 0). All participants were then requested to refrain from any form of oral hygiene for two weeks (Day 1 to Day 14) resulting in plaque accumulation and induction of (mild) inflammation. The two weeks of plaque accumulation (the challenge intervention) were followed by a one week resolution phase. At the onset of the resolution phase all participants were provided with a standardized manual toothbrush, fluoride toothpaste, and specific instructions to re-initiate tooth brushing. The duration of the resolution phase was one week (Day 14 to Day 21). The study design is illustrated in Fig. 2.

One group of participants was requested to take six doses of erythritol per day during the entire duration of the study (five weeks): two weeks prior, two weeks during, and one week after the challenge intervention (Day -14 through Day 21). The erythritol was administered orally, in the form of tablets containing approximately 90% erythritol and 10% maltodextrin, 0.5% magnesium stearate and 0.1% menthol. Each tablet contained 2.0 g of erythritol, amounting to a daily dose of 12.0 g. The No-Observed-Adverse-Effect Level (NOAEL) for laxative effects for erythritol is 0.71 g/kg/bw/day, much higher than the dose administered in this study [16]. Participants were instructed to take one tablet after breakfast, lunch, and dinner, and additionally during the mid-morning, mid-afternoon, and mid-evening. Participants allocated to the erythritol group were specifically instructed not to chew the tablets, but to keep them as long as possible in their mouth, for an optimal exposure of the oral cavity to erythritol. The planned treatment (erythritol) group / control group sample size ratio was 1.00 / 2.00.

Unstimulated saliva samples were collected from all participants at seven time-points across the challenge intervention period: at Day-14, Day 0 (beginning of the challenge intervention), Day 2, Day 5, Day 9, Day 14 (end of the challenge intervention and beginning of the resolution phase), and Day 21 (end of the resolution phase). The team
collecting and analyzing the samples was blinded to the erythritol group or control group status of the participants or samples. Saliva samples were split into separate aliquots for metabolome analysis and for each targeted salivary biochemistry assay. Volunteers were included in the study if they were aged between 18 and 55 yr, systemically healthy (as assessed by a medical questionnaire), had a minimum of 20 natural teeth (with first and second molars present), and had visited a dentist for a regular check-up within the last year. The exclusion criteria were: a dental pocket probing depth of ≥ 5 mm with bleeding on probing and attachment loss ≥ 2 mm (Dutch Periodontal Screening Index score 3+/4) [22], >40% Bleeding on Probing (BOP), dental neglect (such as overt dental caries), smokers (more than one cigarette per day for a least one year), removable partial dentures or night guard, oral/peri-oral piercings, apparent oral lesions, abuse of drugs/alcohol, dental student or dental professional, pregnancy, use of antibiotics during the previous three months, prescribed medication (excepting contraceptives), and evidence of systemic disease.

The study was conducted in accordance with the Declaration of Helsinki (2013) of the World Medical Association and approximating Good Clinical Practice guidelines [23]. The study protocol was reviewed and approved by the Medical Ethics Committee of the Free University of Amsterdam Medical Centre (2014_505) and was registered at the Dutch Trial Register (NL51111.029.14). All participants signed an informed consent form.

Saliva sample collection and metabolomics

Unstimulated saliva was collected between 09:00 and 12:00. Participants were instructed to allow saliva to accumulate on the floor of the mouth and to spit at 30 s intervals into pre-weighed 30 ml polypropylene tubes (Sterilin, Newport, UK), which were kept on ice. The collection period was 5 min. Participants were instructed not to eat or drink two hours prior to the appointment. Samples were immediately frozen and stored at -80 °C.

Non-targeted metabolite profiling of saliva samples was performed on a fee-for-service basis at Metabolon® (Durham, NC, USA). Samples were extracted and prepared for analysis using Metabolon’s standard solvent extraction method [24]. Samples were prepared using the automated MicroLab STAR® system (Hamilton Company). To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five aliquots: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one sample reserved for backup. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before
preparation for analysis. All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18 - 2.1 x 100 mm, 1.7 µm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, with 6.5 mM ammonium bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10 mM ammonium formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS/MS scans using dynamic exclusion. The scan range varied slightly between methods but covered 70-1000 m/z.

Raw data were extracted, peak-identified and quality controlled using Metabolon’s hardware and software. Compounds were identified by comparison with Metabolon’s reference library (retention times, MS and MS/MS spectra) [25]. A data normalization step was performed to correct variation resulting from instrument inter-day tuning differences. Each compound was corrected in run-day blocks by registering the medians to equal one (1.00) and normalizing each data point proportionately. Overall process variability was determined by calculating the median Relative Standard Deviation (RSD) for all endogenous metabolites present in 100% of technical replicates of a pooled sample. The RSD was 10%. Missing values were assumed to be caused by the compound in question being below the detection limit in the respective samples and were imputed with the minimum detected value for that compound across all samples (after the normalization step). Due to differences in ionization potentials between different compounds, normalized peak areas were scaled by the median value for each compound (i.e. each value for a particular compound in a sample was expressed as a multiple of the median value for that compound across all samples in which it was detected) [25]. Metabolome data was acquired for the five time-points corresponding to the experimental gingivitis challenge (Day 0, Day 2, Day 5, Day 9, and Day 14). Due to practical limitations, no metabolome data was
obtained for Day -14 and Day 21, and therefore these two time-points were excluded in the salivary metabolome analysis.

**Targeted functional salivary biochemistry assays**

Salivary pH, buffered pH, flow rate, total protein content, mucin MUC5B, albumin, secretory-IgA, lactoferrin, amylase, lysozyme, chitinase, and proteases were measured as described in a previous study [26]. In brief, total protein content was assayed using a Pierce BCA Protein Assay Kit (Thermo Scientific). MUC5B, albumin, secretory-IgA, and lactoferrin were analyzed using enzyme-linked immunosorbent assays (ELISAs). Amylase, lysozyme, chitinase, and proteases were quantified with enzyme activity assays. Reference samples of pooled saliva were analyzed on each ELISA / enzymatic activity plate to assess assay reproducibility. The mean RSD for the array of targeted salivary biochemistry assays was 8.1% (range 4.2% – 11.8%). Functional salivary biochemistry data were obtained from all seven time-points in the study (Day -14 to Day 21).

**Statistical analysis**

Exploratory data analysis was performed with Principal Component Analysis (PCA) in SPSS version 21. The effects of the challenge intervention, of the erythritol treatment, as well as potential interaction effects between the two were examined with mixed-design ANOVA models implemented in Scientific Python using the “pyvttbl” package, available for download at https://code.google.com/archive/p/pyvttbl/ [27]. Metabolites absent in more than 33% of samples were excluded from the mixed-design ANOVA analysis. The primary output of the mixed-design ANOVA models were \( p \)-values for the within-subject effect, the between-subject effect, and the interaction between the main effects, respectively. The within-subject (repeated-measures) factor was time (i.e. five time-points for metabolites / seven time-points for the salivary functional biochemistry variables). The between-subject factor was erythritol (erythritol vs. control). A Huyhn-Feldt correction was used to compensate for any deviation from sphericity in the data [28].

Metabolite data were log10-transformed (examination of histograms showed that the data were approximately log-normal distributed). Mixed-design ANOVA was performed on each individual metabolite present in more than two thirds of the 305 samples from which metabolome data were acquired, as well as on all metabolite ratios constructed from these metabolites. Ratio data were also log10-transformed to insure that ratio values were independent of the metabolite order and to half the multiple comparison burden, as \( \log(A/B) = -\log(B/A) \) [29, 30]. Multiple comparison adjustment was performed using the Bonferroni correction. The \( p \)-value significance threshold for individual metabolites was calculated at \( 1.2 \times 10^{-4} \). The \( p \)-value significance threshold for metabolite ratios was \( 6.0 \times 10^{-7} \). Additionally, for each ratio a \( p \)-gain statistic was calculated to assess whether the
additional information gained from the ratio was significant compared to the information from the respective individual metabolites [30]. The $p$-gain was defined as the change in the $p$-value obtained when using the metabolite ratios compared to the smallest of the two $p$-values obtained when analyzing the two respective metabolites separately [29, 30]. Metabolite ratios were considered significant if they had a $p$-value $< 1.2 \times 10^{-7}$ as well as a $p$-gain >408. Significant effects were examined with post-hoc Šídák tests.

Mixed-design ANOVA was also performed on data from each of the 12 targeted salivary biochemistry variables measured, with the addition of two extra time-points (Day -14 and Day 21) compared to the metabolome data. The $p$-value significance threshold for the functional salivary biochemistry variables (using the Bonferroni correction) was calculated at $4.2 \times 10^{-3}$.

The statistical significance level (alpha) was 0.05. Standardized effect sizes are reported as partial eta-square ($\eta^2$) as defined by Cohen: partial $\eta^2 = \frac{SS_{between}}{SS_{between} + SS_{error}}$ [31]. Partial $\eta^2$ for a specific effect expresses the proportion the variance explained by that effect out of the variance remaining after excluding variance explained by other effects [32]. Plotted 95% confidence intervals were calculated according to Loftus and Mason’s procedure for within-subject and mixed-design studies [33].
Chapter 5

Results

Study work flow and demographics

Of the 63 volunteers enrolled in the study, 61 participants completed the entire study and were included in the final analysis (Fig. 1). The study design is illustrated in Fig. 2. The erythritol group (N = 20) and the control group (N = 41) were balanced with regard to age and sex (Table 1). The final treatment (erythritol) group / control group sample size ratio was 1.00 / 2.05.

![CONSORT flow diagram of the study - Dutch Trial Register (NL51111.029.14).](image)

**Figure 1.** CONSORT flow diagram of the study - Dutch Trial Register (NL51111.029.14).
Figure. Study design.

Table 1. Age and gender distribution of participants.

<table>
<thead>
<tr>
<th></th>
<th>Males (N = 21)</th>
<th>Females (N = 40)</th>
<th>All (N = 61)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Control (N = 41)</td>
<td>15 males, 26 females</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>45</td>
<td>26.0 ± 6.9</td>
</tr>
<tr>
<td>Treatment (erythritol) (N = 20)</td>
<td>6 males, 14 females</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>23</td>
<td>21.2 ± 1.5</td>
</tr>
<tr>
<td>All (N = 61)</td>
<td>21 males, 40 females</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>47</td>
<td>25.2 ± 6.2</td>
</tr>
</tbody>
</table>
**Salivary metabolites**

The analysis identified a total of 497 metabolites (mass range 75 - 862 Da) in the 305 saliva samples obtained from the 61 participants at five time-points during the experimental gingivitis challenge. The mean number of metabolites identified per sample was 424.1 (SD = 28.4). The complete salivary metabolite dataset is given in Online Resource 1. A number of 446 metabolites out of the 497 total were found in more than 50% of samples, while 291 metabolites were found in more than 95 percent of samples. Only 186 metabolites were detected in 100% of samples. An overview of the distribution of metabolites between the main classes is shown in Table 2. The best represented classes of metabolites were amino acids and amino acid derivatives (29.6% of total) and lipids (28.4%). Among the 141 lipids, the most common sub-classes were phospholipids (34) and sphingolipids (24).

**Table 2.** Overview of metabolites organized by super-pathway.

<table>
<thead>
<tr>
<th>Super-pathway</th>
<th>Total metabolites detected from super-pathway</th>
<th>Mean number of metabolites per sample from super-pathway ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino Acids</td>
<td>147</td>
<td>135.8 ± 6.5</td>
</tr>
<tr>
<td>Lipids</td>
<td>141</td>
<td>130.1 ± 6.4</td>
</tr>
<tr>
<td>Xenobiotics</td>
<td>60</td>
<td>34.2 ± 4.6</td>
</tr>
<tr>
<td>Nucleotides</td>
<td>50</td>
<td>42.5 ± 4.8</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>34</td>
<td>26.8 ± 3.2</td>
</tr>
<tr>
<td>Peptides</td>
<td>34</td>
<td>27.2 ± 5.3</td>
</tr>
<tr>
<td>Cofactors and Vitamins</td>
<td>22</td>
<td>18.6 ± 2.7</td>
</tr>
<tr>
<td>Energy metabolism</td>
<td>9</td>
<td>9.0 ± 0.2</td>
</tr>
</tbody>
</table>

To indicate the compliance in the treatment group and the efficacy of the erythritol treatment delivery, Fig. 3 shows the erythritol concentration at the different time-points in the control and treatment groups, respectively. The mean salivary erythritol concentration was approximately two orders of magnitude higher in the treatment group compared to the control.

PCA analysis was performed to explore the metabolomic data and to assess whether any patterns were discernable based on either within-subject (challenge intervention) or between-subject (erythritol) factors. The PCA scores plots are shown in Fig. 4. No grouping was apparent for either of the two factors.
Effect of experimental gingivitis on the salivary metabolome

Figure 3. Mean erythritol concentrations in the erythritol and control groups, respectively, at each of the five time-points measured during the 14-day experimental gingivitis challenge. Erythritol concentrations are expressed as multiples of the median concentration of erythritol (across all samples in which it was detected). Error bars show 95% confidence intervals of the mean.

Figure 4. PCA scores plots based on data from all 497 metabolites detected. a) Samples labeled according to control and treatment (erythritol) groups; b) Samples labeled according to the time-point in the experimental gingivitis challenge intervention. PC1 and PC2 accounted for 40.3% and 6.6% of variance, respectively (46.9% cumulative).
Data from each of the 408 salivary metabolites detected in more than two thirds of the samples was modeled in a mixed-design ANOVA. The full list of mixed-design ANOVA analyses results for both within-subject (challenge intervention) and between-subject (erythritol) effects are given in Online Resource 2. Table 3 lists the ten salivary metabolites for which the effect of the challenge intervention was significant after adjusting for multiple comparisons using the Bonferroni correction. There was no significant effect of erythritol on any of the metabolites after multiple comparison adjustment.
**Table 3.** List of salivary metabolites found to change significantly between time-points during the experimental gingivitis challenge intervention (after Bonferroni correction for multiple comparisons). *p*-values of the within-subject effect of the mixed-design ANOVA (*N* = 61), relative changes in median values compared to Day 0, and effect sizes are listed for each ratio.

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Super-pathway / Pathway</th>
<th><em>p</em>-value <em>a</em></th>
<th>Max relative change in median value (compared to Day 0)</th>
<th>Effect size (partial η²) <em>b</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>N-α-acetylornithine</td>
<td>Amino acid / Urea cycle; Arginine and proline metabolism</td>
<td>5.8 x 10⁻¹⁰</td>
<td>+ 75%</td>
<td>0.23</td>
</tr>
<tr>
<td>cadaverine</td>
<td>Amino acid / Lysine metabolism</td>
<td>8.6 x 10⁻⁹</td>
<td>+ 69%</td>
<td>0.16</td>
</tr>
<tr>
<td>N-acetylcadaverine</td>
<td>Amino acid / Lysine metabolism</td>
<td>1.5 x 10⁻⁶</td>
<td>+ 58%</td>
<td>0.16</td>
</tr>
<tr>
<td>γ-aminobutyrate</td>
<td>Amino acid / Glutamate metabolism</td>
<td>2.0 x 10⁻⁸</td>
<td>+ 33%</td>
<td>0.19</td>
</tr>
<tr>
<td>mevalonate</td>
<td>Lipid / Mevalonate metabolism</td>
<td>3.3 x 10⁻⁸</td>
<td>+ 51%</td>
<td>0.15</td>
</tr>
<tr>
<td>cytosine</td>
<td>Nucleotide / Pyrimidine metabolism</td>
<td>1.6 x 10⁻⁷</td>
<td>+ 47%</td>
<td>0.18</td>
</tr>
<tr>
<td>octanoylcarnitine</td>
<td>Lipid / Fatty acid metabolism (acyl carnitine)</td>
<td>5.3 x 10⁻⁷</td>
<td>- 32%</td>
<td>0.14</td>
</tr>
<tr>
<td>α-hydroxyisovalerate</td>
<td>Amino acid / Leucine, isoleucine and valine metabolism</td>
<td>3.1 x 10⁻⁵</td>
<td>+ 42%</td>
<td>0.12</td>
</tr>
<tr>
<td>hexanoylcarnitine</td>
<td>Lipid / Fatty acid metabolism (acyl carnitine)</td>
<td>4.5 x 10⁻⁵</td>
<td>- 22%</td>
<td>0.11</td>
</tr>
<tr>
<td>pseudouridine</td>
<td>Nucleotide / Pyrimidine metabolism, uracil containing</td>
<td>5.8 x 10⁻⁵</td>
<td>- 28%</td>
<td>0.13</td>
</tr>
</tbody>
</table>

*a* The Bonferroni *p*-value threshold for significance was 1.2 x 10⁻⁴.

*b* Approximate interpretation for η² effect size: η² ≈ 0.02 – small, η² ≈ 0.13 – medium; η² ≈ 0.26 – large effect [31]
Figure 5. Median values (non-log10-transformed data) per time-point for metabolites found to change significantly during the challenge intervention:

a) N-α-acetylornithine;
b) cadaverine;
c) N-acetylcadaverine;
d) γ-aminobutyrate;
e) mevalonate;
f) cytosine;
g) octanoylcarnitine;
h) α-hydroxyisovalerate;
i) hexanoylcarnitine;
j) pseudouridine.

Error bars indicate 95% confidence intervals of the medians (N = 61).
Salivary metabolite ratios

Data from each ratio between the 408 salivary metabolites detected in more than two thirds of samples was modeled in a mixed-design ANOVA (83,028 ratios). For 760 ratios the effect of the challenge intervention had significant P-values after adjusting for multiple comparisons. Out of these, 166 ratios had P-gains >408 and are listed in Online Resource 3. However, of these 166 ratios, 135 ratios included at least one member of the 10 metabolites found to significantly change in the univariate mixed-design ANOVA analyses. Of the remaining 31 ratios, 13 ratios involved only metabolites known to be cell membrane lipids (Table 4).

A total of 62 membrane lipids were detected (excluding fatty-acid-free degradation products): 25 sphingolipids, 20 glycerophosphocholines (GPC), 15 glycerophosphoethanolamines (GPE), one glycerophosphoserine (GPS), and one glycerophosphoglycerol (GPG). Ten of these lipids were plasmalogens, a class of lipid where one of the two fatty acids is attached by an ether rather than an ester bond. Thirteen lipids (eight sphingolipids and five GPEs) were involved in the 13 ratios changing. Of the 5 GPEs involved in ratio changes, four were plasmalogens (the exception being 1,2-dioleoyl-GPE).

None of the metabolite ratios showed a significant effect of erythritol after Bonferroni correction.

![Figure 6](image.png)

**Figure 6.** Median values (non-log10-transformed data) per time-point for two of the metabolite ratios changing significantly during the challenge intervention: a) Ratio R1 [1-(1-enyl-palmitoyl)-2-arachidonoyl-GPE (P-16:0/20:4) : lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)]; b) Ratio R13 [stearoyl sphingomyelin (d18:1/18:0 : sphingomyelin (d18:1/15:0)]. Error bars indicate 95% confidence intervals of the medians.
Table 4. Membrane lipid ratios found to change significantly between time-points during the experimental gingivitis challenge intervention. *p*-values for the within-subject effect of the mixed-design ANOVA, *p*-gains, and effect sizes are listed for each ratio.

<table>
<thead>
<tr>
<th>No</th>
<th>Metabolite M₁</th>
<th>Metabolite M₂</th>
<th>p-value a</th>
<th>p-gain b</th>
<th>Effect size (partial η²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R01</td>
<td>1-(1-enyl-palmitoyl)-2-arachidonoyl-GPE (P-16:0/20:4)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>2.9 x 10⁻⁹</td>
<td>1.5 x 10⁻⁸</td>
<td>0.20</td>
</tr>
<tr>
<td>R02</td>
<td>1-(1-enyl-palmitoyl)-2-linoleoyl-GPE (P-16:0/18:2)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>1.8 x 10⁻⁷</td>
<td>2.4 x 10⁻⁵</td>
<td>0.20</td>
</tr>
<tr>
<td>R03</td>
<td>1-(1-enyl-palmitoyl)-2-oleoyl-GPE (P-16:0/18:1)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>1.2 x 10⁻⁸</td>
<td>3.6 x 10⁻⁸</td>
<td>0.20</td>
</tr>
<tr>
<td>R04</td>
<td>1-(1-enyl-stearoyl)-2-arachidonoyl-GPE (P-18:0/20:4)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>2.4 x 10⁻⁷</td>
<td>1.8 x 10⁻⁵</td>
<td>0.16</td>
</tr>
<tr>
<td>R05</td>
<td>1,2-dioleoyl-GPE (18:1/18:1)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>1.0 x 10⁻⁷</td>
<td>4.2 x 10⁻⁸</td>
<td>0.16</td>
</tr>
<tr>
<td>R06</td>
<td>behenoyl sphingomyelin (d18:1/22:0)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>3.7 x 10⁻⁷</td>
<td>1.2 x 10⁻⁵</td>
<td>0.17</td>
</tr>
<tr>
<td>R07</td>
<td>N-palmitoyl-sphinganine (d18:0/16:0)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>3.0 x 10⁻⁹</td>
<td>1.5 x 10⁻⁸</td>
<td>0.20</td>
</tr>
<tr>
<td>R08</td>
<td>N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>4.3 x 10⁻⁹</td>
<td>1.0 x 10⁻⁸</td>
<td>0.21</td>
</tr>
<tr>
<td>R09</td>
<td>palmitoyl dihydrosphingomyelin (d18:0/16:0)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>5.8 x 10⁻⁸</td>
<td>7.6 x 10⁻⁵</td>
<td>0.20</td>
</tr>
<tr>
<td>R10</td>
<td>palmitoyl sphingomyelin (d18:1/16:0)</td>
<td>lactosyl-N-palmitoyl-sphingosine (d18:1/16:0)</td>
<td>1.9 x 10⁻⁷</td>
<td>2.3 x 10⁻⁵</td>
<td>0.18</td>
</tr>
</tbody>
</table>
Effect of experimental gingivitis on the salivary metabolome

| R11 | stearoyl sphingomyelin (d18:1/18:0) | lactosyl-N-palmitoyl-sphingosine (d18:1/16:0) | 1.0 x 10^{-7} | 4.4 x 10^{3} | 0.19 |
| R12 | palmitoyl dihydrosphingomyelin (d18:0/16:0) | sphingomyelin (d18:1/15:0) | 9.3 x 10^{-8} | 4.4 x 10^{6} | 0.18 |
| R13 | stearoyl sphingomyelin (d18:1/18:0) | sphingomyelin (d18:1/15:0) | 1.3 x 10^{-7} | 3.0 x 10^{6} | 0.17 |

^aThe Bonferroni p-value threshold for significance for metabolite ratios was 6.0 x 10^{-7}.

^b Metabolite ratios had to have a p-gain > 408 in order to be considered significant.

**Targeted salivary biochemistry assays**

Data from each of the 12 functional salivary biochemistry variables measured were modeled in a mixed-design ANOVA. All results from the analyses of functional salivary biochemistry data are listed in Online Resource 4. Of the 12 variables examined (salivary buffered pH, flow rate, total protein content, mucin MUC5B, albumin, secretory-IgA, lactoferrin, amylase, lysozyme, chitinase, and proteases), only lysozyme (p = 0.02) had p-values <0.05 for the effect of the experimental gingivitis challenge. However, neither the effect of the experimental gingivitis challenge intervention nor the effect of erythritol was significant in any of the functional salivary variables analyzed after adjusting for multiple comparisons.
Discussion

The results of this study show some significant changes in the salivary metabolome of young healthy adults during a two-week experimental gingivitis challenge intervention. These changes were observed in both the control and the treatment groups, with no significant effect of erythritol found on any of the metabolites analyzed.

Of the ten metabolites found to change significantly during the experimental gingivitis challenge (Table 3), seven metabolites showed increasing trends (N-α-acetylornithine, cadaverine, N-acetylcadaverine, γ-aminobutyrate, mevalonate, cytosine, and α-hydroxyisovalerate), while the other three (octanoylcarnitine, hexanoylcarnitine, and pseudouridine) decreased (Fig. 5).

Cadaverine, N-acetylcadaverine, and α-hydroxyisovalerate are known products of bacterial amino acid catabolism. Oral bacteria can produce cadaverine and N-acetylcadaverine from lysine, while α-hydroxyisovalerate is derived from valine [34, 35]. This apparent increase in amino acid catabolism by oral bacteria could be a consequence of the accumulation of dental plaque after the cessation of oral hygiene measures. As the oral biofilm is left undisturbed and becomes more mature and organized, its metabolic efficiency increases [36]. An accumulation of catabolic products is therefore in line with the expected intensification of plaque metabolism.

The increase in both γ-aminobutyrate and N-α-acetylornithine may be explained in the context of an adaptive response of oral bacteria to lower pH in the biofilm’s vicinity. As it matures, the oral biofilm may shift towards a more acidogenic functional profile and become more efficient in fermenting sugars and therefore lowers the pH of its surroundings [37]. However, not all bacterial species in dental plaque are acidoduric (e.g. due to lacking acid-resistant ATPases), and in fact some employ specific metabolic pathways to produce alkali in order to resist an increase in acidic stress [38]. One of the main alkali-producing pathways in bacteria is the Arginine Deiminase System (ADS), which plays a major role in oral biofilm pH homeostasis [39, 40]. The ADS pathway, present in common oral bacteria such as Streptococcus gordonii, Streptococcus parasanguis, Streptococcus rattus, Streptococcus sanguis and Lactobacillium fermentum, catabolizes L-arginine (originating either from salivary secretions or dietary sources) to produce ammonia, with ornithine as a by-product [40]. Together with a subsequent acetylation step by bacterial acetyltransferases, the additional ornithine resulting from an up-regulation of ADS could account for the observed increase in N-α-acetylornithine.

Another metabolic route used by bacteria to increase their environmental pH is glutamate decarboxylation, present in some species of Bacteroides, Fusobacterium, Eubacterium, and Lactobacillus, among other Gram-positive bacteria [41, 42]. In this reaction cycle, extracellular glutamate is imported into the bacterial cell through a glutamate / γ-aminobutyrate antiporter. The imported glutamate is subsequently
Effect of experimental gingivitis on the salivary metabolome
decarboxylated and the resulting non-ionized γ-aminobutyrate is returned to the extracellular environment through the aforementioned antiporter [41]. The net outcome is an increase in bacterial intracellular pH and an extracellular increase in γ-aminobutyrate.
In the present study, N-α-acetylornithine and γ-aminobutyrate show highly similar trends, with considerable increases during the first five days of the experimental gingivitis before leveling off (Fig. 5, a and d). This could indicate that an up-regulation of bacterial acid-adaptation pathways occurred in the first few days of the challenge intervention.

Mevalonate is a precursor in the mevalonate pathway, where mevalonate is converted to sterol terpenoids (e.g. cholesterol) as well as non-sterol terpenoids (e.g. heme-A, isopentenyl tRNA and ubiquinone) [43]. The mevalonate pathway has anti-inflammatory and immunomodulatory effects [43, 44]. It is particularly important for modulating regulatory T-cell proliferation and function, maintaining a tolerogenic rather than proinflammatory phenotype [44, 45]. The >50% increase in median mevalonate levels observed during the first days of the experimental gingivitis challenge (Table 3, Fig. 5 e) may point towards an early anti-inflammatory response by host epithelial tissues to changes in the oral plaque.

Octanoyl- and hexanoylcarnitine were both found to decrease during the study (Table 3, Fig. 5 g, i). No significant changes were seen in the other nine detected acyl-carnitines. Carnitine and acyl-carnitines are involved in mitochondrial beta-oxidation of fatty acids [46]. However, bacteria may utilize acyl-carnitines as carbon and/or nitrogen sources [47]. It is unclear why these two particular acyl-carnitines decreased, although it may be linked to preferential use by oral bacteria.

Cytosine increased during the experimental gingivitis challenge intervention, showing a marked change between Day 2 and Day 5 (Table 3, Fig. 5 f) while pseudouridine - another nucleotide – decreased (Fig. 5 j). Whereas cytosine is one of the two canonical pyrimidines in DNA, pseudouridine is a modified pyrimidine found in eukaryotic ribosomal and transfer RNA, but not in messenger RNA [48]. Pseudouridine is a stable end-product of RNA catabolism and has been considered a potential marker of cell turnover [48]. However, in the context of the oral environment, the mechanisms behind the increase in cytosine and the decrease in pseudouridine are unclear. No other nucleotides showed significant changes.

Previous studies have discovered metabolite ratio biomarkers associated with smoking, diabetes, and other health-related conditions [30, 49, 50]. The analysis of metabolite ratios is advantageous. It cancels out systematic experimental variation and, perhaps most importantly, when the two members of the ratio are connected by a biochemical pathway, it can yield valuable additional information on the status of respective pathway [51].

A set of 13 metabolites ratios involving membrane lipids were found to change significantly during the study, with all ratios showing the same trend: an increase from Day 0 to Day 2 followed by a continuous decrease from Day 2 through Day 14 (Table 4, Fig. 6).
Of the diversity of membrane lipid classes detected in this study, only sphingolipids (eight sphingomyelins and three ceramides) and GPEs (one GPE and four GPE-plasmalogens) were involved in the changing ratios.

Sphingomyelins are a class of lipids containing an aliphatic amino alcohol connected to a sphingoid backbone and a fatty acid N-linked via an amide. They are the most abundant lipids in human cell membranes, yet absent in most bacteria [52]. Hydrolysis of sphingomyelins yields ceramides (N-acyl-sphingosines) and free choline. However, ceramides can also be converted back into sphingomyelins [53]. Human cells maintain a controlled balance between sphingomyelins and ceramides, as the relative contribution of each will have a significant effect on the properties of the lipid bilayer [53]. Moreover, these membrane lipids also mediate cell signaling processes and may release second messengers (e.g. arachidonic acid)) [53]. Glycosylated sphingolipids, such as lactosyl-N-palmitoyl-sphingosine (involved in 11 of the 13 ratios found to change during the experimental gingivitis challenge intervention), are involved in cell-to-cell communication and are part of cell surface receptors for bacterial toxins [54-56]. Both sphingomyelins and ceramides are implicated in the intracellular activation of signaling pathways and in the release of cytokines, and therefore have an important role in the host response to microbial challenges [57]. Interestingly, another salivary metabolomics study found palmitoyl sphingomyelin to be elevated in saliva of periodontitis patients compared to healthy controls [14].

Four plasmalogens are involved in four of the 13 changing lipid ratios (see Table 4). Plasmalogens are membrane glycerophospholipids containing a fatty alcohol with a vinyl ether bond at the sn-1 position. Plasmalogens represent up to 20% of the total phospholipids in human cell membranes [58]. While several anaerobic bacteria also contain plasmalogens (e.g. members of Clostridia, Maegasphera, and Veillonella), the composition of the four plasmalogens involved in the changing ratios (e.g. arachidonic / linoleic acid components) suggest that they originate from the host [58, 59]. Besides influencing the fluidity of the cell membrane, plasmalogens have been proposed to act as defensive antioxidants on the cell membrane surface. The vinyl ether bond makes plasmalogens more susceptible and they are preferentially oxidized over other membrane lipids when exposed to free radicals [60]. However, plasmalogen-derived oxidative products cannot further propagate lipid oxidation and the damage to the membrane is therefore limited [61]. Indeed, reactive oxygen species (ROS)-producing human cells have plasmalogen-enriched membranes [62].

The set of membrane lipids ratios found to be significantly changing - involving almost exclusively sphingomyelins, ceramides and plasmalogens - may be linked to increased interactions between host tissues and oral bacteria, brought by changes in the functional profile and metabolism of the oral biofilms during the experimental gingivitis challenge intervention. While the complexity and the inter-connected nature of the functions of many of these membrane lipids make a detailed interpretation of individual
ratios difficult, these findings cumulatively point to a dynamic adjustment of the host cells in response to the effects of the changes in the oral biofilm metabolism.

The study found no significant effects of a daily intake of erythritol on the salivary metabolome during the challenge intervention. While a previous prospective clinical trial found significant decreases in dental plaque and in dental caries rates, the duration of erythritol intake in that study was three years, much longer than the five weeks in the present study [17, 18]. Perhaps any effect that erythritol may have on the oral ecosystem requires an exposure duration longer than the one used in this study. Moreover, the design of the present study was aimed at observing changes during the induction of gingivitis, not caries progression as in the case of the aforementioned prospective clinical trial.

None of the salivary functional biochemistry variables analyzed in this study showed any changes during the challenge intervention, pointing towards a relatively stable long-term profile of salivary secretions which was not significantly influenced by the changes occurring in the mouth during the experimental gingivitis challenge intervention.

One of the main difficulties in salivary metabolomic analysis is that a large number of metabolites may have both human and bacterial origin [4]. While this may complicate interpretation of results, salivary metabolomics remains a powerful tool for gaining biological insights into the mechanisms involved in the maintenance of oral health.

**Concluding remarks**

This study aimed to examine the changes in the salivary metabolome of healthy young adults during a two-week experimental gingivitis challenge intervention. A number of significant changes in salivary metabolites and metabolite ratios were found. An increase in bacterial amino acid catabolites (cadaverine, N-acetylcadaverine, and α-hydroxyisovalerate) pointed towards an intensified metabolism of the oral biofilm, while increases in end-products of bacterial alkali-producing pathways (N-α-acetylornithine and γ-aminobutyrate) may be a response of less aciduric oral bacteria faced with an increasingly acidogenic oral biofilm. Synchronous changes in a set of 13 salivary metabolite ratios involving sphingomyelins, ceramides and plasmalogens, all host cell membrane lipids known to be involved in cell signaling, host responses to bacteria, and defense against free radicals could be a consequence of host adaptations to changes in the oral biofilms. An increase in mevalonate, a key member in a pathway involved in anti-inflammatory responses, was also observed. Taken together, the findings of this study underline a dynamic interaction between host and oral microbial biofilms during induction of experimental gingivitis.
References


'136


Effect of experimental gingivitis induction and erythritol on the salivary MALDI-TOF MS peptide profiles of systemically healthy young adults

Abstract

Introduction Understanding the changes occurring in the oral ecosystem during development of gingivitis could help to improve prevention and treatment strategies for oral health. Erythritol is a non-caloric polyol proposed to have beneficial effects on oral health.

Objectives To examine the effect of experimental gingivitis induction and the effect of erythritol on the salivary peptide profiles of healthy adults.

Methods In a two-week experimental gingivitis challenge intervention study, MALDI-TOF-MS profiling was performed on saliva samples from 61 healthy adults, collected at seven time-points. The effect of erythritol was studied in a randomized, controlled trial setting. Effects were analyzed using mixed-design ANOVA models.

Results 83 salivary peptide peaks in the mass range 2 – 15 kDa were quantified. There were no significant effects of the experimental gingivitis or of erythritol on any of the peaks analyzed (all effect p-values > 0.05, before adjusting for multiple comparisons).

Conclusions No significant effects of experimental gingivitis induction and/or of a daily dose of erythritol on the MALDI-TOF salivary peptide profiles of healthy adults were observed.


Introduction

Oral health is integral part of the general health [189]. Gingivitis is an oral health condition, a local inflammatory response caused by an oral bacterial biofilm [190]. Gingivitis is regarded as a transitional stage between oral health and oral disease, as it does not affect the supporting structures of the teeth and is generally reversible with improved oral hygiene [191]. Despite this, gingivitis is a potential risk factor in the development of periodontitis, one of the major dental diseases [190, 192]. Biomarkers that can give information on the state of the oral cavity with regard to gingivitis could be useful tools for understanding the changes occurring in the oral ecosystem during gingivitis development, potentially leading to improved oral health prevention strategies. Saliva is a biological fluid that can be collected easily and non-invasively and has been explored as a source of biomarkers for oral as well as systemic diseases [193, 194].

Erythritol is a non-caloric polyol used as a sweetener. Its safety has been well documented: it is classified as GRAS (Generally Recognized As Safe) by the U.S. Food and Drug Administration and is considered safe for use as a food additive (E968) by the European Food Safety Authority [144]. Erythritol is non-cariogenic and has been shown to lower plaque levels and caries rates in a prospective clinical trial [145, 146]. The positive effect of erythritol with regard to caries may be due it potential stabilizing effect on the oral microflora: in-vitro studies have found that erythritol inhibits the growth of Streptococcus mutans strains [148]. Erythritol may therefore have a protective effect on the oral homeostasis.

A previous study examining variation in salivary MALDI-TOF peptide profiles found that individuals could be clustered based on their peptide profiles into four subgroups [195]. Salivary P-C peptide and its fragmentation pattern had an important role in the clustering. Individuals from the four subgroups showed significant differences in their underlying salivary biochemistry with regard to the activity of lysozyme and chitinase, two enzymes relevant for oral health which are involved in the salivary innate defense system [195]. Therefore, it is possible that changes in salivary peptide profiles during an induced shift in the oral ecosystem could yield information on the underlying mechanisms contributing to the stability and resilience of that system, and thus ultimately to oral health.

The aim of this study was to examine the changes in the salivary peptide profiles of young healthy adults acquired using MALDI-TOF MS during a two-week experimental gingivitis challenge intervention, and to assess the effect of a daily intake of erythritol on these changes during a controlled, randomized clinical study.
Materials and methods

Study design

This study was part of the TI Food and Nutrition project ‘An exploratory study on the dynamic (microbial, biochemical and immunological) interactions of the oral ecosystem during induction of mild gingival inflammation. Dynamics of a healthy oral ecosystem’.

Whole unstimulated saliva was collected in a single-center, challenge intervention, randomized study at the Academic Centre for Dentistry Amsterdam (ACTA). The study population was a convenience sample of systemically healthy adults. The study design and inclusion/exclusion criteria have been previously described in detail [196]. In brief, the challenge intervention was based on a full-mouth modified experimental gingivitis protocol [149]. The first stage was a two-week baseline period in which the treatment group started the use of erythritol (Day -14 to Day 0). All participants were then requested to refrain from any form of oral hygiene for two weeks (Day 1 to Day 14) resulting in plaque accumulation and induction of (mild) inflammation. The two weeks of plaque accumulation were followed by a one week resolution phase. At the onset of the resolution phase all participants were provided with a standardized manual toothbrush, fluoride toothpaste, and specific instructions to re-initiate tooth brushing.

One group of participants (the treatment / erythritol group) was requested to take six doses of erythritol per day during the entire duration of the study (five weeks), amounting of 12.0 g/day of erythritol. The erythritol group (n = 20) and the control group (n = 41) were balanced with regard to age and sex.

Unstimulated saliva samples were collected from all participants at seven time-points across the challenge intervention period: at Day -14, Day 0, Day 2, Day 5, Day 9, Day 14 and Day 21.

The study was conducted in accordance with the Declaration of Helsinki (2013) of the World Medical Association and approximating Good Clinical Practice guidelines [151]. The study protocol was reviewed and approved by the Medical Ethics Committee of the Free University of Amsterdam Medical Centre (2014_505) and was registered at the Dutch Trial Register (NL51111.029.14). All participants signed an informed consent form.

Sample collection and MALDI-TOF-MS analysis

Unstimulated saliva samples were collected and processed as previously described [100]. In brief, participants were instructed to spit for 5 min at 30 s intervals into ice-chilled 30-ml polypropylene tubes (Sterilin, Newport, U.K.). After vortexing, samples were clarified by centrifugation for 10 min at 4°C and 10,000 g. The resulting clarified saliva was diluted with a NaCl solution to a final concentration of 250 mM NaCl and stored at -80°C.
Samples were desalted prior to MALDI-TOF MS analysis using C18 ZipTips (Merck Millipore, Darmstadt, Germany). The purified peptides were spotted directly onto a MALDI target plate with 1 µL of matrix solution (10 mg α-cyano-4-hydroxy cinnamic acid in 1 ml of ACN/water 1:1 (v/v) with 2.5% trifluoroacetic acid). Spectra were recorded in linear mode at a mass range of 2 – 15 kDa with a 200 Hz laser at 355 nm on an Autoflex III MALDI-TOF mass spectrometer (Bruker, Bremen, Germany). Details of the MALDI-TOF protocol have been previously described [195]. Samples were analyzed in duplicate. Visual quality control was performed for all raw spectra. Samples were re-analyzed if either one of the duplicate spectra showed excessive noise or large baseline drift, or if the duplicate spectra exhibited dissimilarities. Spectra of a reference saliva sample were acquired in quadruplicate at each session in order to aid the subsequent sample spectral alignment and to assess the reproducibility of the assay. The mean coefficient of variation of the assay was 18%.

Raw MALDI-TOF MS spectra were processed in Matlab R2012b using the Mathworks bioinformatics tool box (MathWorks, Natick, MA, U.S.A). The workflow consisted of spectra resampling followed by baseline subtraction, smoothing, peak detection and finally peak binning (peak coalescing) using a hierarchical clustering algorithm.

**Statistical analysis**

The effects of the challenge intervention, of the erythritol treatment, as well as potential interaction effects between the two were examined with mixed-design ANOVA models implemented in Scientific Python using the “pyvttbl” package, available for download at https://code.google.com/archive/p/pyvttbl/ [154]. Peaks were excluded from the analysis if they were present in <2% of the total number of samples. P-values were calculated for the within-subject effect (the challenge intervention), the between-subject effect (erythritol), and the interaction between the main effects, respectively. A Huyhn-Feldt correction was used to compensate for any deviation from sphericity in the data [155]. The significance level (alpha) was 0.05. The Benjamini-Hochberg False Discovery Rate (FDR) procedure was used to correct for multiple comparisons [73]. The FDR was set at 0.05.
Results

Of the 63 enrolled volunteers, 61 participants completed the entire study and were included in the final analysis. Sample volumes from one participant from the control group were insufficient. Therefore, MALDI-TOF MS data from a total number of 60 participants were included in the final analysis (40 from the control group; 20 from the erythritol group). The age of the participants was 25.2 ± 6.2 (mean ± SD). A number of 83 MALDI-TOF MS peaks were present in more than 2% of samples (i.e. in more than 8 out of 420 samples) and were therefore included in the mixed-design ANOVA analysis. The complete dataset is provided as Supplementary Information 1.

There was a large overlap between the peaks found in this study and the peaks found in the saliva of young healthy adults in a previous, cross-sectional study, using an identical MALDI-TOF protocol [195]. The cross-sectional study found 80 peaks, 69 of which were also found in the present study (83% overlap). The frequency of the 69 overlapping peaks (the proportion of samples in which the respective peaks were found) in the two studies were similar (Supplementary Information 2).

Most of the 14 peaks that did not overlap (i.e. 12 out of 14) had very low frequencies, being present in <5% of samples. The degree of concordance between the peaks detected in the two studies of salivary MALDI-TOF is illustrated in Fig. 1.

![Venn diagram illustrating the overlap (yellow) between the peaks found in a previous cross-sectional study (red circle on the left, 80 peaks) and the present study (green circle on the right, 83 peaks).](image)

Figure 1. Venn diagram illustrating the overlap (yellow) between the peaks found in a previous cross-sectional study (red circle on the left, 80 peaks) and the present study (green circle on the right, 83 peaks).
The mixed-design ANOVA analysis found no significant effects of either the experimental gingivitis challenge intervention or of erythritol on any of the 83 peaks included in the analysis (all $p$-values > 0.05, before adjusting for multiple comparisons). The detailed results of the mixed-design ANOVA are provided as Supplementary Information 3.
Discussion

The present study found no significant changes in the salivary peptide profiles of healthy adults during the two weeks period of gingivitis induction or after the one week resolution phase. Also, there was no significant effect of erythritol on the salivary peptide profiles. There was good agreement between the list of peaks detected in the present study and the peaks detected using an identical analytical protocol in a previous cross-sectional study of salivary peptide profiles, performed on a different study population (Fig. 1). The majority of none-overlapping peaks (found in one study, but not the other) were present at low levels, close to the detection limit of the assay, in <5% of the samples analyzed in the respective studies. The present assay therefore appears to be able to consistently detect and measure approximately 70 salivary peptide peaks.

While a previous prospective clinical trial found significant decreases in dental plaque and in dental caries rates, the duration of erythritol intake in that study was three years, much longer than the total of five weeks in the present study [145, 146]. Salivary erythritol concentrations during the present study indicate good levels of compliance with regard to erythritol intake. Mean salivary concentrations in the erythritol group during the experimental gingivitis induction phase (between Day 0 and Day 14) were approximately two orders of magnitude higher than in the control group [196]. However, it may be that any potential effect of erythritol on the oral ecosystem requires a much longer exposure duration than the one used in this study. Moreover, the design of the present study was aimed at observing changes during the induction of gingivitis, not caries progression as in the case of the aforementioned prospective clinical trial.

There could be several reasons for the non-significant effect of the experimental gingivitis induction on the salivary peptide profiles. Firstly, it may be that none of the peptides quantified was influenced by the induction or by the resolution of the 2-week experimental gingivitis challenge. One of the limitations of MALDI-TOF MS is the limited mass range at which it is effective (in this study, between 2 – 15 kDa), therefore giving only a partial coverage of the collection of peptides present in saliva. However, an antibody- and enzymatic activity-based analysis of the same samples set looking at eight larger salivary proteins with known functional relevance for oral health also found no significant changes during the induction of experimental gingivitis [196]. Secondly, it may be that the effect, if it did in fact occur, was too small to detect with the present study sample size (N = 61).

In conclusion, this study found no significant effect of experimental gingivitis induction or of erythritol on the MALDI-TOF salivary peptide profiles of healthy adults.
References


CHAPTER

7

Summarizing discussion
General discussion

The oral cavity is the site of complex interactions between the oral microbiota, salivary components and the host immune system [1-3]. Co-evolution of the oral bacteria and the human host has created an ecosystem that exhibits significant temporal stability and resilience to challenges [4-6]. However, factors such as diet, smoking and an insufficient level of oral hygiene can shift the ecosystem towards dysbiosis [7]. Dysbiosis can manifest through either compositional or functional changes in the oral microflora. This can lead to dental decay or induce a maladaptive inflammatory response which destroys the periodontium and eventually may cause tooth loss (i.e. periodontitis) [7-10]. Maintaining oral homeostasis is therefore crucial for the preservation of oral health [11]. However, the mechanisms acting to maintain homeostasis are not well understood [11].

The aim of this thesis was to describe the molecular and bacterial composition of the oral ecosystem in a healthy state, to examine the network of interrelations between salivary proteins and peptides, oral bacteria and the oral metabolome, and to assess the changes occurring in the ecosystem when it is faced with a specific challenge (i.e. induction of experimental gingivitis).

The objective of Chapters 2, 3 and 4 was to obtain insights into the natural variation in saliva in a systemically and orally healthy population. These chapters contain data acquired from samples collected during a cross-sectional, observational clinical study - Dutch Trial Register (NTR3649) - conducted at the Academic Center for Dentistry Amsterdam (ACTA) in 2013. The study protocol is described in detail in Chapter 2.

In Chapter 2 the analysis of a set of salivary proteins and enzymes relevant for oral health was described. The study population were systemically healthy young adults (N = 268) screened for suitability according to the criteria as proposed for the Dutch Periodontal Screening Index (DPSI). Volunteers were included if they had a DPSI < 3-. Exclusion criteria were: presence of overt dental caries, inter-proximal restorations between the first and second, or second and third upper molars, apparent oral lesions, a habit of smoking, infections, recent use of antibiotics, and use of anti-inflammatory drugs. Concentrations of mucins MUC5B and MUC7, secretory-IgA, lactoferrin, cystatin S, and albumin were determined using ELISAs. Lysozyme, chitinase, amylase, and proteases were measured with enzymatic activity assays. Salivary pH, buffered pH, flow rate and total protein content were also measured. The subset of health-related salivary functional proteins and enzymes analyzed is not exhaustive. Other salivary components which play important roles in oral homeostasis like agglutinin, statherin and lactoperoxidase were not determined. A selection of analytical targets had to be made, taking into account limited sample volumes and available resources.
First, the network of correlations between the analyzed proteins and enzymes was examined. While most proteins correlated negatively with salivary flow rate and positively with each other and with total protein content, there were some intriguing, flow rate-independent, negative partial correlations between specific components, such as between lysozyme and chitinase. While lysozyme possesses bactericidal activity, chitinase lyses fungal cell walls [12, 13]. This network of positive and negative correlations between functional salivary proteins may suggest that feedback and/or balancing mechanisms are present in the regulation of salivary composition.

Using Spectral Clustering, we found two subgroups within the study population. The smaller subgroup (N = 22) was differentiated from the rest of the study population by lower salivary pH and buffered pH, and higher MUC7 levels and lysozyme activity. In addition, a number of significant differences were observed between males and females. Salivary pH, buffering capacity, protein content, MUC5B, secretory-IgA, and chitinase activity were all lower in females compared to males, while MUC7 and lysozyme activity were higher in females (P < 0.05, FDR set to 5%). We speculated that these differences could be due to physiological factors, primarily the influence of sex hormones on gene expression in salivary gland tissues, as well as sexual dimorphism in the size of the glands between males and females [14-16]. These sex-related physiological differences in saliva may be a factor contributing to the higher caries prevalence observed in women compared to men [17, 18].

The data from the salivary functional protein biochemistry analyses were also used in Chapter 4 to construct the salivary ecosystemic network.

In Chapter 3 we analyzed the salivary peptide profiles of the 268 healthy volunteers using MALDI-TOF mass spectrometry (MS). Salivary peptides were purified and concentrated using Zip-Tips and spectra were acquired for m/z values between 2 – 15 kDa. Spectral Clustering identified four subgroups within the study population. An Unsupervised Feature Selection algorithm revealed that these clusters were differentiated by the variation in nine peptide peaks. Five of the nine peaks (two different ionization forms and three different fragments) were linked to the same molecular entity, an acidic proline-rich protein named P-C peptide. P-C peptide is formed from precursors (PRP1, PRP2, Pif-f and Db) which show different functionality compared to the free P-C peptide [19]. The precursors possess a domain which facilitates attachment to the tooth surface. They play a role in mediating enamel hydroxyapatite crystal growth, as well as in the formation of the dental pellicle and subsequent bacterial adherence [20-22]. On the other hand, P-C peptide (which lacks the aforementioned binding domain) is thought to have a protective function against dietary tannins [23]. The relative amount of precursors and P-C peptide, and the P-C peptide fragmentation profile are determined by proteolytic events taking place both pre-secretion (inside the salivary glands and ducts, under the action of host proteases), as well as post-secretion (in the mouth, due to microbial proteases). Thus,
the pattern of these peptides may provide information on the physiological state of the salivary glands as well as on the functional composition of the oral microflora. This hypothesis is supported by the observation that the four clusters of volunteers discriminated by their salivary peptide profiles differ significantly with regard to two antimicrobial enzymes - lysozyme and chitinase (P < 0.001). These findings suggest that MALDI-TOF MS salivary peptide profiles may relate information on the underlying state of the oral ecosystem.

In Chapter 4 the ecosystemic network of saliva of healthy volunteers was examined. Microbial DNA was extracted from saliva and processed for amplicon sequencing of the V4 hypervariable region of the 16S rRNA gene on an Illumina MiSeq platform. Salivary metabolome data were obtained from three separate platforms (two LC/MS-MS and one GC/MS) [24]. The salivary microbiome data were integrated with the salivary metabolome data and the salivary functional protein data (previously described in Chapter 2). The heterogeneity of the microbiome, of the metabolome, and the mutual interrelations between these datasets and the host-related biological and environmental parameters were assessed.

The salivary microbiome appeared heterogeneous, forming five sample clusters. These microbiome clusters related to characteristic biochemical properties of saliva. The presence of ‘oral ecotypes’ and the links observed between the microbiome clusters and salivary functional proteins suggest a role of host factors in shaping the microbial community. One of the factors that related to the oral microflora composition was salivary pH, a known driver of microbial community activity and ecological shifts [25, 26]. Other interesting associations were found between salivary lysozyme activity and microbial composition. Lysozyme is an antimicrobial protein capable of lysing the bacterial cell wall peptidoglycan [27]. It was therefore surprising to find a positive correlation between salivary lysozyme activity and streptococci, which are Gram-positive. However, lysozyme is thought to also possess additional, Gram-independent mechanisms of action, such as aggregation of bacteria, which are subsequently cleared by swallowing [28]. A previous study showed that lysozyme adsorbs to oral bacteria in a strain-specific manner [29]. The lowest adsorption of lysozyme was observed for *Streptococcus mitis* strains [29]. Interestingly, the highly abundant OTU with the strongest positive relation with lysozyme activity in saliva was classified as *Streptococcus dentisani/infantis/mitis/oralis* or mitis group streptococci (OTU4). These mitis group streptococci are known as primary colonizers of oral surfaces and are associated with oral health [30]. Higher abundance of mitis group streptococci was not only associated with higher lysozyme activity, but also with lower salivary pH. Since the optimum activity of lysozyme depends on pH and ionic strength [31], lysozyme activity, measured *ex vivo* at pH 7, might indicate a compensatory over-expression of lysozyme at a physiological salivary pH below the functional optimum of the enzyme.
The large inter-individual heterogeneity seen in the salivary microbiome did not translate to the salivary metabolome profiles, which were much more similar. This may imply that metabolism in overnight fasted saliva of different individuals carrying different microbiota is driven by similar processes, in line with the previously described functional redundancy of human microbial communities [32]. However, about one fourth of the individuals showed different (e.g. more proteolytic activity-driven) and more diverse metabolome profiles. These profiles related with several biochemical parameters of saliva, such as higher pH and higher albumin concentration, and with microbial communities dominated by anaerobes and Gram-negative taxa associated with mature oral microbial communities and gingival inflammation [33]. The positive relation between salivary albumin and pro-inflammatory metabolites in these saliva samples may suggest the presence of inflammation [34].

A clear dichotomy was observed when relating the microbial taxa with the metabolites. Either strong positive (such as for OTU66, *Megasphaera micronuciformis*), or strong negative correlations (such as for OTU4, *Streptococcus mitis* group) with the majority of the metabolites were observed. This suggests that during overnight fasting, when saliva is the main substrate for oral bacteria [35], there are metabolite ‘producers’ such as megasphaerae and metabolite ‘consumers’ such as streptococci. The complex microbial interactions required to break down salivary glycoproteins were confirmed by visualization of the complex networks of megasphaerae co-occurring with other taxa compared to the isolated position of streptococci.

Both the microbiome and the metabolome of saliva were gender-associated. The higher abundance of streptococci in female saliva could be a result of an ecological pressure due to physiologically lower pH [36]. On the other hand, the higher abundance of streptococci could directly contribute to a lower salivary pH, since streptococci are the main acid producers in a healthy oral cavity [37].

Brought together, these findings indicate that a healthy oral ecosystem is based on concerted interactions of its heterogeneous and multifactorial components. Data described in Chapters 2, 3, and 4 were obtained from a cross-sectional, observational study. As such, mechanistic explanations for observed correlations and trends may be suggested but causal links cannot be clearly drawn solely on observational data. Deciphering the mechanisms which enable the stability of this ecosystem will only be possible through longitudinal studies.

**Chapters 5** and **6** contain data acquired from samples collected during a challenge intervention, randomized study - Dutch Trial Register (NL51111.029.14) – conducted at the Academic Centre for Dentistry Amsterdam (ACTA) in 2015. The study population were systemically healthy adults (N = 61). The exclusion criteria were: a dental pocket probing depth of $\geq 5$ mm with bleeding on probing and attachment loss $\geq 2$ mm (DPSI 3+/4) [38], >
40% bleeding on probing, overt dental caries, smokers, prescribed medication, and evidence of systemic disease.

The challenge intervention was based on a full-mouth modified experimental gingivitis protocol [39]. The study design was composed of a two-week baseline period, followed by a two-week gingivitis induction period in which all participants were asked to refrain from any form of oral hygiene, and finally a one week resolution phase. The aim of the study was to observe the changes occurring in the oral ecosystem as a result of plaque accumulation and subsequent induction of inflammation. The rationale was to gain insights into the early stages of a shift of the oral ecosystem away from homeostasis by measuring the changes occurring in the salivary microbiome, metabolome, and functional biochemistry.

The treatment group (N = 20) was asked to take erythritol every day during the entire duration of the study. The control group (N = 41) did not take erythritol. Unstimulated saliva samples were collected from all participants at seven time-points during the study. The study protocol is described in detail in Chapter 5.

In Chapter 5 the changes occurring in the salivary functional biochemistry and the salivary metabolome during a 2 week-long experimental gingivitis induction were analyzed. The effect of erythritol, a polyol proposed to have a beneficial effect for oral health, was also examined in a randomized trial setting. Salivary functional proteins were measured using ELISAs and enzyme activity assays as previously described in Chapter 2 [40]. Salivary metabolome data were acquired on four different LC-MS/MS platforms, each optimized for a subset of metabolites with different physiochemical characteristics. Univariate mixed-design ANOVA models were constructed for each variable analyzed, with the challenge intervention (experimental gingivitis) as a within-subject factor and erythritol as a between-subject factor. FDR was used to adjust for multiple comparisons (FDR set to 5%).

Bacterial amino acid catabolites (cadaverine, N-acetylcadaverine, and α-hydroxyisovalerate) were found to increase, pointing towards an intensification of plaque metabolism. End-products of alkali-producing bacterial pathways (both γ-aminobutyrate and N-α-acetylornithine) also increased, suggesting that some bacteria within the oral biofilm employed these pathways in order to adapt to a presumably lower pH in the maturing dental plaque. Increased bacterial amino acid catabolites and increased products of bacterial alkali-producing pathway suggest an up-regulation in oral plaque metabolism as the biofilm matured.

Mevalonate increased by 50%. The mevalonate pathway is known to be involved in anti-inflammatory and immunomodulatory effects [41, 42]. Significant changes were also found in a set of salivary metabolite ratios involving 13 host cell membrane lipids (eight sphingolipids and five glycerophospholipids) potentially involved in cell signaling, host responses to bacteria, and defense against free radicals [43]. The increase in
mevalonate – occurring within the first 2 days of experimental gingivitis induction –
together with the changes in lipid ratios may be indicative of an early anti-inflammatory
reaction by the host in response to changes in the quantity and functional profile of oral
plaque.

No significant effects of the experimental gingivitis induction were found on any
of the functional salivary proteins and enzymes analyzed. Similarly, no significant effects
of erythritol on the salivary metabolome or on the functional salivary proteins were found.
While a previous prospective clinical trial found significant decreases in dental plaque and
in dental caries rates, the duration of erythritol intake in that study - 3 years - was much
longer [44, 45]. It is possible that any effect that erythritol may have on the oral ecosystem
is dependent on longer exposure than used in the present study. Moreover, the present study
was designed to observe changes during the induction of gingivitis, not to monitor caries
progression.

In Chapter 6 the salivary peptide profiles of healthy volunteers were examined during a 2-
week induction of experimental gingivitis. MALDI-TOF spectra were acquired as
previously described in Chapter 3 [46]. The effect of erythritol was also assessed in a
randomized trial setting. Mixed-design ANOVA models were constructed to analyze
changes in each peptide peak.

No significant changes were found in any of the 80 peaks analyzed, with no
significant effects of the experimental gingivitis challenge or of the erythritol treatment.

Future research

The studies described in the present thesis provide an extensive - yet by no means complete
- description of the oral ecosystem and the network of interactions between its different
components (Chapters 2, 3 and 4), as well as a glimpse into the dynamics of this system
when subjected to a challenge (Chapters 5 and 6). It is clear, however, that our
understanding of the intricate complexity underlying the oral ecosystem remains limited.

An obvious next step in the continuation of the analyses presented here would be the
integration of the data from the experimental gingivitis induction study (i.e. linking salivary
microbiome data with metabolome and functional biochemistry data). This could provide
more insights into the mechanisms involved in oral plaque maturation and adaptation and
into the corresponding changes in plaque-host interactions. It is becoming more and more
apparent that examining the composition of the oral microbiota is not sufficient, and that
the key lays in examining microbiota functionality. Indeed, results presented in this thesis
(Chapter 4) suggest that oral microbiomes with large differences in composition may in fact
have very similar functional profiles. It is this functionality that needs to be thoroughly
understood in order to be able to effectively manipulate the system into a desired direction
or to determine the optimal way to prevent an unwanted shift. This may be achieved through the integration in future longitudinal and prospective cohort studies, combining metabolomic and metatranscriptomic analyses with clinical measurements and targeted functional biochemistry assays. In such studies, particular focus should be given to microbial pathways dealing with carbohydrate metabolism, protein and peptide degradation, and alkali production, as they appear to be of primary importance in determining oral ecosystem homeostasis [47]. The ultimate goal is to reach a level of understanding that can be successfully applied in the clinic: one can imagine personalized therapies for periodontitis patients or caries-prone individuals, designed to restore the specific dysfunction diagnosed in their oral ecosystem. The analytical technology needed is already available, yet rigorous study design and improved data processing methods and statistical approaches are needed to fully benefit from the wealth of data that can be produced.
References


Acknowledgements

Someone once said: “No man is an island”. I have been guided and helped along these 4 years of pursuing a PhD by many wonderful people, some of whom I would like to thank below:

ENNO, my supervisor, for guiding me through the PhD while allowing me to learn to manage my own research. I am also grateful for his example of a healthy, fact-based, no-nonsense approach to research and science.

HENK, my daily supervisor, for his always energetic, optimistic and cheerful presence, and the continuous care and dedication that he shows towards all his students.

BART and WIM, for their vision and management of the TIFN Oral Health project, of which my PhD project was an integral part.

FLORIS, TOON, KAMRAN, JAN, PETRA, MARIANNE and SABRINA, and all my department colleagues who have helped and taught me throughout my research at ACTA.

EGIJA, BERND and ELENA, with whom it was a pleasure to collaborate and who have given me useful insights into microbial genomics and immunology.

SULTAN, MARLEEN and MARCELA, my fellow PhD students and team members, with whom I’ve shared this experience and the struggles and joys that came with it.

AD DE JONG and ARMAND PAAUW, for their time and patience and the fruitful collaboration at TNO Rijswijk.

ILONA and CATHERINE, for their support during these years.