Journal of
PERIODONTAL RESEARCH

J Periodont Res 2013; 48: 392–403

JOURNAL OF PERIODONTAL RESEARCH
doi:10.1111/jre.12025

Quantitative analysis of the intra- and inter-subject variability of the whole salivary proteome

N. Jehmlich[1], K. H. D. Dinh[1], M. Gesell-Salazar[1], E. Hammer[1] L. Steil[1], V. M. Dhople[1] C. Schurmann[1], B. Holtfreter[2] T. Kocher[2], U. Völker[1]

[1]Department of Functional Genomics, Interfaculty Institute of Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany and [2]Unit of Periodontology, Department of Restorative Dentistry, Periodontology and Endodontology, University Medicine Greifswald, Ernst-Moritz-Arndt-University, Greifswald, Germany

Jehmlich N, Dinh HDK, Salazar MG, Hammer E, Steil L, Dhople VM, Schurmann C, Holtfreter B, Kocher T, Völker U. Quantitative analysis of the intra- and inter-subject variability of the whole salivary proteome. J Periodont Res 2013; 48: 392–403. © 2012 John Wiley & Sons A/S. Published by John Wiley & Sons Ltd

Background and Objective: Interest in human saliva is increasing for disease-specific biomarker discovery studies. However, protein composition of whole saliva can grossly vary with physiological and environmental factors over time and it comprises human as well as bacterial proteins.

Material and Methods: We compared intra- and inter-subject variabilities using complementary gel-based (two-dimensional difference gel electrophoresis, 2-D DIGE) and gel-free (liquid chromatography tandem mass spectrometry, LC-MS/MS) proteomics profiling of saliva. Unstimulated whole saliva of four subjects was examined at three different time-points (08.00 h, 12.00 h and 17.00 h) and variability of the saliva proteome was analyzed on two successive days by LC-MS/MS.

Results: In the 2-D DIGE experiment, the median coefficient of variation (CV) for intra-subject variability was significantly lower (CV of 0.39) than that for inter-subject variability (CV of 0.57; CV of technical replicates 0.17). LC-MS/MS data confirmed the significantly lower variation within subjects over time (CV of 0.37) than the inter-subject variability (CV of 0.53; CV of technical replicates 0.11), and that the inter-subject variability was not time-dependent.

Conclusion: Both techniques revealed similar trends of variations on technical, intra- and inter-subject level but provided peptide and protein focused information and should thus be used as complementary approaches. The data presented indicate that 2-D DIGE as well as LC-MS/MS approaches are suitable for biomarker screening in saliva.

Prof. Dr. Uwe Völker, Ernst-Moritz-Arndt-University Greifswald, Interfaculty Institute of Genetics and Functional Genomics, Department of Functional Genomics, Friedrich-Ludwig-Jahn-Str. 15a, 17487 Greifswald, Germany
Tel: +49 3834 86 5870
Fax: +49 3834 86 795871
e-mail: voelker@uni-greifswald.de

Key words: whole saliva; intra- and inter-subject variability; proteome; 2-D DIGE; LC-MS/MS

Accepted for publication September 26, 2012

Human saliva is one of the major body fluids and constitutes an easily accessible source for non-invasive collection of biomaterial for clinical diagnostics (1–4). In the last 15 years, the use of saliva as a diagnostic fluid has drawn much attention, because in the future salivary tests may provide a new resource for fully automated test systems to measure proteins or peptides, which may reflect a profile of various systemic diseases (5–7). The differential composition of the salivary protein pattern was extensively investigated in Sjögren's syndrome (8), diabetes mellitus (9,10), periodontitis (5,11), oral (12,13) or lung cancer (14). In addition, the protein pattern between whole saliva and parotid and parotid and submandibular–sublingual gland secretions were compared (15). In-depth proteomic characterization of human saliva is complicated by the broad dynamic range of protein abundances, the varying degree of post-translational modifications of individual proteins,

and the mainly nutrition-dependent temporal and dynamic changes of proteins, making selection of appropriate sample collection times a challenge (16–19). A few abundant proteins dominate the total salivary protein pattern and low abundance proteins can only be covered when diverse protein/peptide fractionations and large-scale MS approaches are combined (20–22). Thus, the number of identified proteins in whole saliva was increased using a large amount of starting material (75 mg of protein) from 1050 proteins in 2007 (6) to about 2340 proteins in 2009 using a three-dimensional peptide fractionation analysis platform (23).

However, such workflows are not compatible with protein biomarker screenings of biomaterial from patient cohorts where the volume of collected saliva is limited and the MS analysis of several hundred subjects is required. Yet another challenge of human salivary proteome analysis is the physiological variability associated with the heterogeneous composition of saliva caused by inter-subject variability and the above-mentioned nutrition-dependent variations (24). Thus, precise knowledge of the intra- and inter-subject variability in the human salivary proteome is a prerequisite for using saliva as a resource in screening for diagnostic and prognostic markers (25,26). Therefore, current proteomics workflows need to be further improved in terms of sensitivity, throughput, quantitation and reduction in experimental MS analysis time (27).

In this study, we compared the intra- and inter-subject variabilities in whole saliva analyzed by two different proteomic techniques, namely gel-based (two-dimensional difference gel electrophoresis, 2-D DIGE) and gel-free (liquid chromatography tandem mass spectrometry, LC-MS/MS) approaches. Both 2-D DIGE and LC-MS/MS have been very successfully employed to identify proteins involved in biological systems, but in recent years, MS-based approaches have gained more weight because of their easy use and far greater sensitivity. However, the MS approaches display a peptide-centric view and might miss important protein-based information, e.g. co-occurrence of post-translational modification on the same molecule. To evaluate if both methods are suitable for protein biomarkers discovery experiments we examined the variability of the protein pattern in whole saliva of four individuals at three different time-points (08.00 h, 12.00 h and 17.00 h) and further analyzed the variability on two successive days by LC-MS/MS. The salivary protein variability was either monitored by LC-MS/MS applying label-free quantitation or 2-D DIGE in combination with protein identification using matrix-assisted laser desorption/ionization tandem MS (MALDI-MS/MS).

Material and methods

Saliva collection and preparation

Four human volunteers, two female (subjects A and B) and two male (subjects C and D), age 33–54 years ($n = 4$), were recruited for this study. Participants were asked to refrain from eating and drinking 2 h before saliva collection and they were free of fever and/or cold. No medical examination was carried out, and we named the saliva in this study 'normal'. The saliva samples were collected at five different time-points for each volunteer. Participants gave informed consent for the collection and analysis of their saliva.

All saliva samples were collected with a Salivette® (Sarstedt, Nümbrecht, Germany). The commercially available Salivette® was used because of: (i) easy, reproducible handling; (ii) representation of a global proteome picture in whole saliva; and (iii) suitability for large population-based studies (28). A plain cotton roll was chewed for 1 min and mechanically stimulated saliva collected. The rolls with the absorbed saliva were placed into the Salivette® and immediately centrifuged at 11,600 g for 20 min at 4°C yielding clear saliva. To inhibit protein degradation a protease inhibitor cocktail (v/v 1 : 5000, Sigma-Aldrich, St. Louis, MO, USA) was added before centrifugation. The collected saliva volume ranged from 0.6 to 2 mL with an average volume of 1.3 ± 0.6 mL. Saliva was stored at −80°C until used for further analysis.

Saliva samples were thawed and centrifuged at 16,200 g for 30 min at 4°C to remove food remnants, insoluble material and cell debris. Supernatants were transferred into new collection tubes. Proteins were precipitated using trichloroacetic acid (TCA) at a final concentration of 10% (v/v) and dithiothreitol (0.12% w/v). After vortexing and incubation on ice for 15 min, precipitated protein was concentrated by centrifugation (16,200 g, 15 min, 4°C). Protein pellets were washed twice with ice-cold 100% acetone and dried with a SpeedVac for 10 min; afterwards the protein pellet was solubilized in 8 M urea and 2 M thiourea. Protein concentrations were determined using a Bradford assay (Bio-Rad, Hercules, CA, USA) and yielded an average of 0.87 mg ± 0.28 mg [standard deviation (SD), $n = 20$] protein per mL saliva initially collected. Aliquots of the protein extracts were stored at −80°C until analysis.

The protein precipitation method (TCA) and protein estimation assay (Bradford) were tested for robustness, reliability and recovery of saliva proteome preparation. Different volumes of pooled saliva were precipitated with TCA and the protein amounts were measured using a Bradford assay. The fitted linear regression approximation of data points [$R^2 = 0.992$, coefficients of variation (CV) ranges between 0.04 and 0.10] indicated the robustness of TCA precipitation of different volumes. To test the protein recovery by TCA precipitation with different protein concentrations, we spiked pooled saliva samples with known amounts of excess albumin (0.5, 1, 2 and 4 mg) (data not shown). The yield was calculated as the percentage of total saliva protein recovered by the precipitation. The recovery ranged from 36% (+ 4 mg) to 61% (+ 0.5 mg) in agreement with commonly observed efficiencies of organic protein precipitation methods where approximately 40–60% protein were lost (29).

Gel-based approach (two-dimensional difference gel electrophoresis), image and statistical analysis

Saliva protein samples were labeled for the DIGE experiment according to the manufacturer's instruction incorporating a dye swap for discrimination of dye-specific effects (GE Healthcare, Munich, Germany) (30). In all experiments, the Cy2-dye was used to label an internal standard composed of all samples analyzed in the respective experiment. For two-dimensional protein gel electrophoresis (2-DE) two labeled samples (Cy3 and Cy5, each 50 μg) and the corresponding internal standard (Cy2, 50 μg) were mixed and incubated in rehydration solution (8 M urea, 2 M thiourea, 2% w/v CHAPS, 28 mM dithiothreitol 1.3% v/v pharmalytes, pH 4–7 and bromophenol blue), separated according to their pI on immobilized pH gradient strips (24 cm, GE Healthcare) with a pH range from 4 to 7 (first dimension) and on 12.5% sodium dodecyl sulfate –polyacrylamide gel electrophoresis in low fluorescent glass plates (GE Healthcare) (second dimension) as described previously (31). Each sample was analyzed with four technical replicates. Images of the three different channels were acquired using a Typhoon 9400 laser scanner (GE Healthcare) at excitation/emission wavelength of 488/520 (Cy2), 532/670 (Cy3) and 633/670 nm (Cy5).

After scanning, the Delta-2D software package (v3.4, Decodon, Greifswald, Germany) was used to match all gel images. After background subtraction, the spot volumes were calculated and normalized against the spot volume of the internal standard (Cy2) as described (30).

Statistical analysis was performed with the GeneSpring software (GeneSpringGX, v7.3.1, Agilent Technologies, Waldbronn, Germany) as described earlier (32). Briefly, for normalization, the background-corrected volume of each spot was divided by the median of all spots of the same image. These median normalized volumes of the Cy3 and Cy5 channels were divided by the corresponding median normalized Cy2 channel volume of the same spot on the gel. For the analysis of changes in spot intensities depending on time as well as for calculation of technical, intra- and inter-subject variability, the spot volumes were median-normalized across the 2-D DIGE project. Spots were considered significantly different in intensity between different time-points when normalized spot intensities differed more than two-fold with an analysis of variance (ANOVA) score of $P < 0.05$ (multiple testing correction according to reference 33).

Matrix-assisted laser desorption/ionization tandem mass spectrometry analysis of 2-dimensional spots

Two-hundred sixty 2-D spots from a representative 2-D gel were manually excised, peptide extracts prepared and spotted on to a MALDI target using an Ettan Spot Handling Workstation (GE Healthcare) according to a previously described standard protocol (34). The MALDI time-of-flight MS (MALDI-TOF MS/MS) of spotted peptide solutions was carried out using a 4800 Proteomics Analyzer (ABI SCIEX, Darmstadt, Germany) (35). MS analysis was performed with the following settings: 10 MS/MS spectra with 100 shots per MS/MS spectrum were accumulated using a random search pattern and a mass range of m/z 804 up to 4000. An internal calibration was automatically performed as a one- or two-point calibration for self-digested trypsin fragments at m/z 1045.5 and m/z 2211.1.

MS/MS analysis was performed for the five strongest peaks of the MS spectrum. For one main spectrum, 20 MS/MS spectra with 125 shots per MS/MS spectrum were accumulated using a random search pattern. The internal calibration was automatically performed as a one-point calibration if the monoisotopic arginine $(M+H)^+$ m/z at 175.119 or lysine $(M+H)^+$ m/z at 147.107 reached a signal-to-noise ratio of at least 20. The peak lists were created using the GPS Explorer software (Applied Biosystems, Carlsbad, CA, USA) with the following settings: mass range from 60 Da to a mass that was 20 Da lower than the precursor mass; peak density of 10 peaks/200 Da; minimal area of 100 and maximal 100 peaks per precursor; minimal signal/noise ratio of 7. For protein identification, we used the MASCOT algorithm via the GPS Explorer software package version 3.6 (Applied Biosystems) with a human-specific database (SwissProt version 55.1), and proteins were considered to be identified when the Mowse score (MASCOT) exceeded 55, which corresponds to a P-value of 0.05. Protein identifications based on peptide mass fingerprint data were confirmed by at least one protein-specific peptide fragmentation (MS/MS) (Table S1).

Gel-free approach (LC-MS/MS) and data analysis

Reverse phase separation of tryptic peptides before mass spectrometric analyses was performed on a nano UPLC (Acquity UPLC system, Waters, Milford, MA, USA) with a 10 cm nanoAcquity (100 μm i.d., 1.7 μm C18) analytical column at a flow rate of 300 nL/min with a binary buffer system consisting of 0.1% acetic acid, 2% acetonitrile in water (buffer A) and 0.1% acetic acid in 100% acetonitrile (buffer B). Peptide separation was achieved using a linear gradient of buffer B from 5 up to 25% within 63 min. MS data were generated using the Orbitrap Velos MS equipped with a nanoelectrospray ion source (PicoTip Emitter, New Objective, Woburn, MA, USA). After a first survey scan (r = 60,000) MS/MS data were recorded for the 20 highest mass peaks in the linear ion trap at a collision-induced energy of 35%. The exclusion time to reject masses was set to 60 s and the minimal ion signal for MS/MS was 2000.

Proteins were identified and quantified via Rosetta Elucidator software (Ceiba Solutions, Seattle, WA, USA). Data were searched against a forward-reverse Swiss-Prot database limited to human entries (v2010-08, n = 40,716, including common contamination sequences) using the SEQUEST algorithm v2.7 (Sorcerer v4.04, Sage-N

Research Inc, Milpitas, CA, USA). Search parameters were 10 p.p.m. parent mass tolerance and 1 Da for fragment ion mass tolerance. Carbamidomethylation of cysteines was specified as a fixed modification and methionine oxidation as a variable modification. Peptides were annotated on a false positive rate of 1% calculated by PeptideTeller (PeptideProphet) embedded in Elucidator. Only proteins with at least two significant peptides were considered for identification. The following settings for Elucidator were used for quantification: (i) feature detection and alignment; (ii) feature annotation; (iii) median normalization by a feature set containing search results to discriminate features arising from single-charged contaminants; (iv) combination of signal intensities of technical replicates; (v) filtering for unique peptides; and (vi) statistical analysis using a two-tailed two-sample *t*-test ($P < 0.05$) (Table S2).

Statistical analysis

Statistical analyses were performed based on spot volumes (gel-based approach) and protein intensities (gel-free approach). The technical replicate values were averaged, median normalized and $\log_2$ transformed. ANOVA was calculated using either time-point or subject as factor. Pearson's correlation coefficients of spot volumes or protein intensities were calculated for all time-points and subjects as well as for each subject stratified by every time-point. Heat maps were generated by the function 'heatmap.2' embedded in the R package 'gplots' and hierarchical cluster analyses were performed with the function 'hclust' from the R package 'stats' either based on squared Pearson's correlation coefficients or SD (36).

Results and discussion

Gel-based (two-dimensional difference gel electrophoresis) analysis

Saliva proteome composition was monitored by 2-D DIGE followed by Delta-2D differential analysis and subsequent mass spectrometry. A total of 606 protein spots were detected across all gels in the present study and were retained for comparative analysis. In Fig. 1, representative gels (pI 4–7) from subject A at three time-points (08.00 h, 12.00 h and 17.00 h) illustrate intra-subject comparisons, and 2-D gels of three additional subjects at collection time 08.00 h exemplify the inter-subject variability. The 2-D gel patterns were characterized by compact, dense and intense spots in the high molecular mass range. In addition, those spots feature extensive 2-D spot chains across the selected pI range (4–7). Global 2-D spot patterns were reproducible and highly similar between subjects and among the collection times. In total, 260 2-D spots were excised from a preparative 2-D gel (Fig. S1) and analyzed by MALDI-MS/MS (Table S1). In the 260 2-D spots, 106 unique proteins (SwissProt/Uniprot annotations) were identified with > 95% confidence. The number of distinct proteins was much smaller than the number of spots because the gel pattern was dominated by a few abundant proteins such as immunoglobulins, albumin or alpha-amylase, which were widespread over the 2-D gels. Immunoglobulin

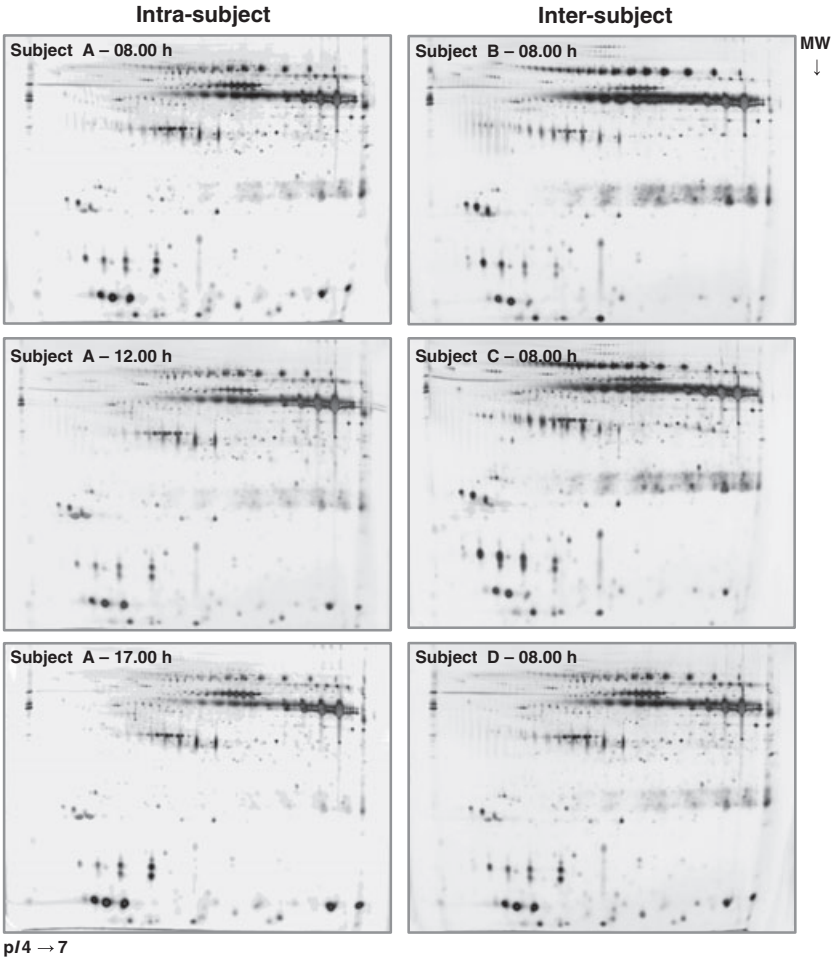

Fig. 1. Representative two-dimensional fluorescence (2-D DIGE, Cy3) images from one subject (A) at different collection times on 1 day (intra-subject, left panel) and 2-D images from three more subjects (B–D) at the same collection time at 08.00 h (inter-subject, right panel). For 2-DE analysis, two-labeled whole saliva proteome samples (Cy3 and Cy5, each 75 μg) and the corresponding internal standard (Cy2, 75 μg) were focused on 24 cm immobilized pH gradient strips with pH range 4–7 (first dimension) and afterwards separated on 12.5% sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

proteins (e.g. immunoglobulin kappa chain C region, immunoglobulin alpha-1 chain C region or polymeric immunoglobulin receptor precursor) accounted for 78 2-D spots (30% of $n = 260$), serum albumin was identified in 25 2-D spots (10%) and alpha-amylase was present in 20 different 2-D spots (8%).

Intra- and inter-subject variability of the gel-based analysis

A principal components analysis (PCA) plot of the 2-DE global proteome pattern is shown in Fig. 2 (left panel). This plot displays the first two principal components. Each point represents four technical 2-D gel replicates.

Generally, subjects A and C were clearly separated from the other two subjects, B and D. Particularly for subject A all three saliva collection points were located closer together than to any other collection point of another subject. The large distance of the point representing subject C, 12.00 h from the other two might be explained by food intake only 10 min before the whole saliva was collected. To quantitatively assess sample variation, we determined CV for each of the three levels (technical, intra- and inter-level) using the normalized intensities of protein spots. For the technical variation in 2-D DIGE, the median CV was 0.17. The variation of the intra-subject proteome for the three time-points (08.00 h, 12.00 h and 17.00 h) was 0.39 and the inter-subject variation was calculated with a CV of 0.57. The approximate quantitative contribution of each variation to the total variability can be estimated because variances are additive among each other. Therefore, within our data set, the technical variability contributed to 9%, the intra-subject variability to 38% and inter-subject variability to 53% of the total variability, respectively (37).

Our analysis of the intra-subject variability in the human saliva proteome revealed significant changes in intensity by a factor of at least 2 (fold change = 2) at different time-points for 168 (28%) to 246 (41%) of 606 protein spots. In total, 142 (23%) protein spots displayed significant changes in intensity in all four subjects ($P < 0.05$). However, in contrast to the other three subjects in subject C more protein spots displayed changes in intensity (397 corresponding to 66% of 606 protein spots) suggesting that food intake immediately before (within 10 min) sampling had profound effects on the proteome pattern. For the inter-subject variability, the analysis revealed 437 (72%) significant variations of protein spot intensity at time-point 08.00 h, 461 (76%) at 12.00 h and 486 (80%) at 17.00 h, respectively.

Interestingly, the majority of spots that display inter-subject variation, namely 336 (about 73%), exhibited significant intensity variations at all three examined time-points suggesting that inter-subject variability appeared to be independent of sampling time.

Whole saliva analyses by 2-DE have been described before (12,26). Here, we use 2-D DIGE analysis of human whole saliva to monitor the intra- and inter-subject variation in four subjects on three different time-points over one day. Quantifying intra-subject variability at 38% and inter-subject variability at 53%, we support the notion that intra-subject variability has weaker effects on the salivary protein patterns than inter-individual differences. 2-D analysis showed that most changes (73%) were related to inter-subject variability. Thus, even if a large number of salivary proteins vary in abundance and probably display varying degrees of modification, reflected in change of intensity patterns of protein spots representing the same protein [i.e. post-translational modifications such as glycosylation (16)] during the day, inter-subject variability was more pronounced. Therefore, this probably facilitates meaningful sampling in population-based and large clinical studies, where sampling time cannot be fixed to a certain time-point. To gain information about the heterogeneity and distribution of spot volumes of the analyzed bioset, spot volume values ($\log_2$) were pair-wise compared in scatter plots for each subject and time-point (Fig. 3). This presentation clearly indicated that subject-related distribution (inter-subject) was more widespread and therefore this variation was stronger than the time-related distribution (intra-subject).

We then used heat map analysis based on the squared Pearson correlation coefficients (R^2) (Fig. 4) to intuitively display information about

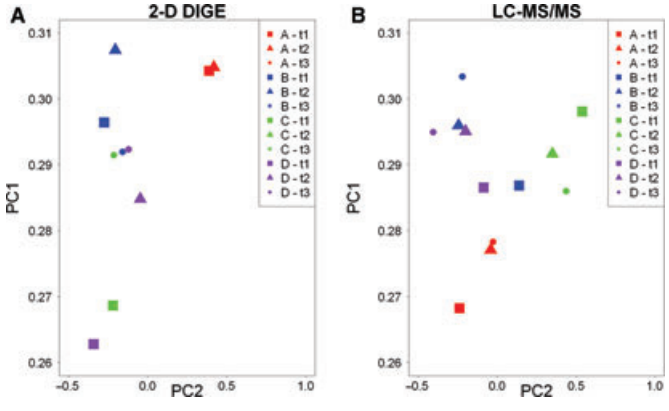

Fig. 2. Principal components analysis (PCA) plot based on the covariance matrix of the two-dimensional difference gel electrophoresis (2-D DIGE, left panel) and liquid chromatography tandem mass spectrometry (LC-MS/MS, right panel) analysis. Subjects and time-points are represented by the same colors and symbols, respectively. The *y*-axis represents the first principal component (PC1), the *x*-axis the second component (PC2). Collection times: t1, 08.00 h; t2, 12.00 h; and t3, 17.00 h. (left panel) PCA plot with each point representing the mean of four 2-D gel replicates (Cy3 and Cy5) from one subject at a given time-point. (right panel) PCA-plot derived from LC-MS/MS data (summarization level are proteins). Each data point represents three technical replicate measurements.

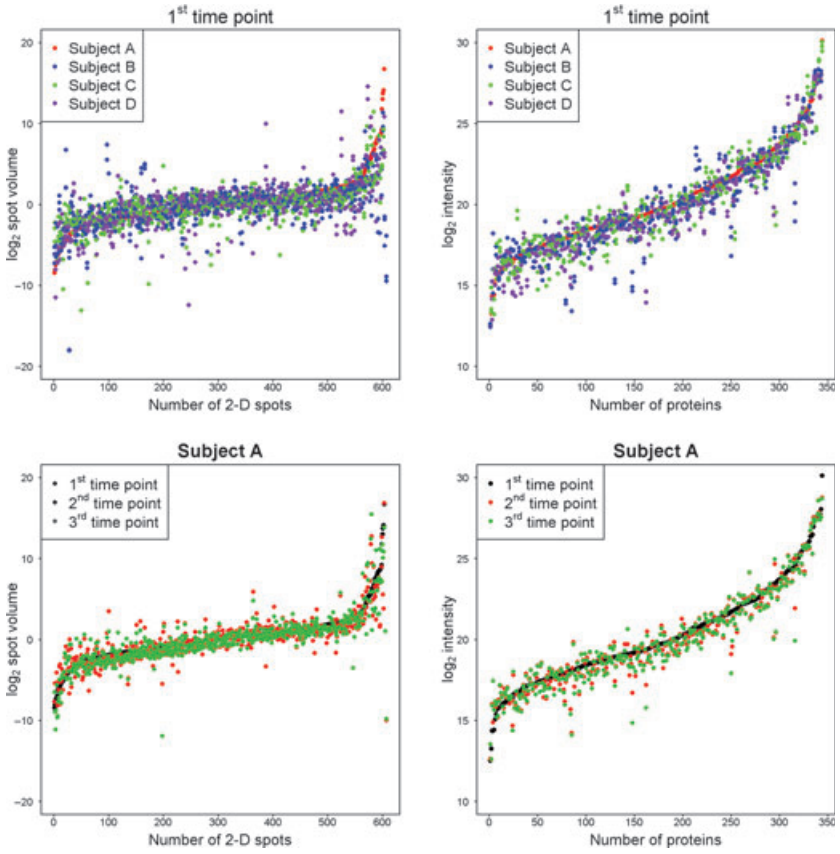

Fig. 3. Representative scatter plot distribution of the spot volumes [two-dimensional difference gel electrophoresis (2-D-DIGE, left panel), gel-based analysis] and protein intensities [liquid chromatography tandem mass spectrometry (LC-MS/MS, gel-free analysis, right panel]. Data were transformed to $\log_2$ values and exemplarily displayed in the upper panel for the first time-point (08.00 h, upper panel) for all subjects (A–D) and in the lower panel for subject A for each of the three different time-points at one day (first, 08.00 h; second, 12.00 h; and third, 17.00 h). Color coding of subjects and time-points are indicated in the insets of each plot.

the quality of the data. The squared correlation coefficients ranged from 0.27 to 0.81 ($n = 30$) with an average of 0.58. The weakest coefficient with 0.27 was observed for the comparison between subject B (time-point 12.00 h) and subject C (time-point 12.00 h) indicating the high impact of the above-mentioned food intake.

Generally, the intra-subject squared correlation coefficients (intra-$R^2 = 0.67$) showed greater values than inter-subject comparisons (inter-$R^2 = 0.52$) (Table S3). Thus, we confirmed our initial observation of a greater impact of inter-subject variability by analyzing Pearson's correlation coefficients. Homogeneous distribution of the spot volumes is exemplified in subject A (P1), whereas subject C (P3) is the most heterogeneous subject indicated by the blue quadrats in the analyzed bioset (Fig. 4, left heat map).

The close clustering of samples from the same subject in the heat maps (Fig. 4) also reiterated the fact that inter-subject variations had apparently a more pronounced effect than intra-variability of sampling during the course of the day.

To gain more information about the diversity of the whole saliva proteome pattern, we estimated the SD of the average spot volumes ($\log_2$). For each subject the corresponding time-points were displayed in one plot where the spot volumes were ordered by the first time-point (08.00 h, Fig. S1). For subject C the data clearly displayed wider spreading over a large range of values, particularly for the second time-point (12.00 h), which was different to the other subjects that showed more homogeneous values. Additionally, for each time-point, subjects A, B, C and D were displayed in one plot to reveal if the time-points of saliva sample collection have relevance for proteome variance (Fig. S2). Although subject C, which displayed wider variation at the second time-point, was included in the SD analysis, the values of the three time-points were comparable. This pattern was confirmed by results of ANOVA analysis where observed variance was partitioned into components attributable to different sources of variation, namely subject and time (Table S4). Because the variance attributed to the time-points and subjects indicated no significant change compared to the overall variance (P-values = 0.57 and 0.10, respectively), these statistical results support our conclusion that saliva collection and further proteome analysis can be performed irrespective of collection time.

Gel-free (liquid chromatography-tandem mass spectrometry) analysis

LC-MS/MS analyses were performed in triplicate for each of the 12 different saliva samples (four subjects with three time-points) resulting in a total of 36 LC-MS/MS runs of 100 min each and total measurement time of about 1 week. Measurement time for the triplicate analysis was about 5 h and the total protein amount was < 2 μg, which are key points for biomarker screening studies with samples derived from population-based studies where sample material is often limited due to time constraints during the epidemiological fieldwork.

LC-MS/MS analysis resulted in the identification of 681 unique proteins from 3106 sequence-unique peptides using the SEQUEST algorithm; 392 proteins were identified with at least two peptides per protein. The total number of protein identifications varied between the subjects (separate LC-MS/MS measurements) from 160 to 287 proteins (≥ 2 peptides per

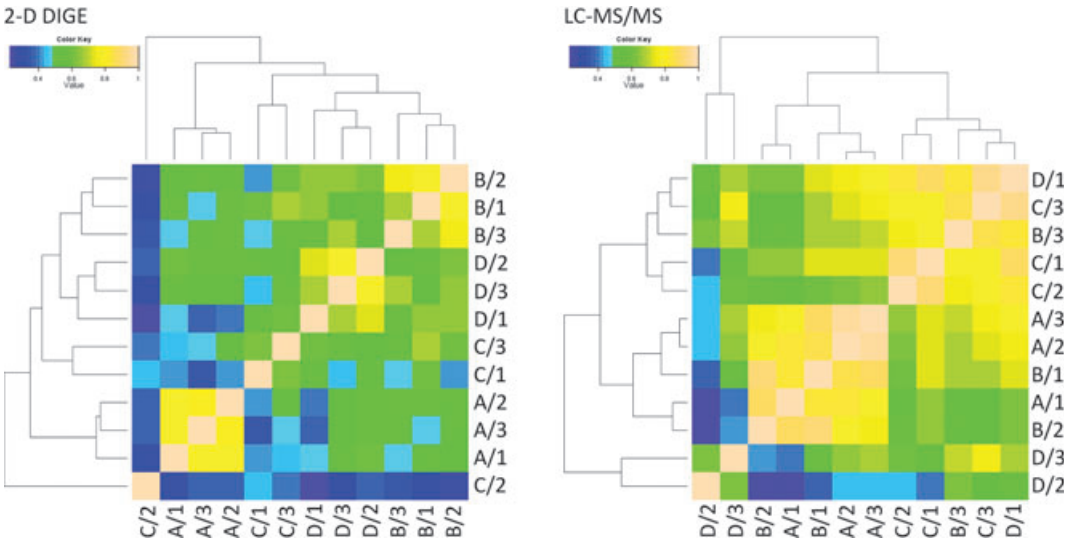

Fig. 4. Heat maps based on squared Pearson's correlation coefficient (R^2) of the gel-based (two-dimensional difference gel electrophoresis, 2-D DIGE) and gel-free (liquid chromatography tandem mass spectrometry, LC-MS/MS) analyses. The matrix is color coded as shown in the upper left inlet. Subjects: A–D, Time-points: 1 = 08.00 h; 2 = 12.00 h; and 3 = 17.00 h.

protein) with an average of 212 proteins per LC-MS/MS run.

Intra- and inter-subject variability of the gel-free analysis

Using shotgun LC-MS/MS analysis, we monitored the protein variability of adults at different collection times on the same day and at one time-point for two successive days. Direct label-free quantitation was applied using peptide intensities as indicators for protein abundance and variability. Relevant mass spectrometric information for the entire data set is included in Table S2.

A PCA plot of the data is shown in Fig. 2 (right panel). Generally, we observed the same trend as for the 2-D DIGE approach: subjects A and C were separated from the other two subjects B and D. Particularly for subject A, all three saliva collection points were located closer together than to any other collection point of another subject. Subject C was separated from the others. CV was determined for each of the three levels (technical, intra- and inter-level) of sample variation. The median CV values did not significantly vary when three (day 1 only) or five time-points (including days 1 and 2) were considered. Therefore, the median CV values for the total LC-MS/MS data set are presented. For technical triplicate measurements, the median of CV was 0.11, which is an expected value for repeated LC-MS measurements. The variation of the intra-subject proteome over five time-points (08.00 h, 12.00 h, 17.00 h, days 1 and 2) was 0.37. The inter-subject variation was calculated with a CV of 0.53. Note that the calculations of CVs were on a protein-by-protein basis. We further calculated the approximate quantitative contribution to the total variability. In our data set, the technical variability contributed to 5%, the intra-subject variability to 45% and the inter-subject variability to 50% of the total variability, respectively.

Protein intensity values were pairwise compared in scatter plots for each subject and time-point to gain more detailed information about the heterogeneity and distribution of protein intensities (Fig. 3, right panel). The plots clearly indicate also for gel-free analysis that subject-related distribution (inter-subject) was more heterogeneous than time-related distribution (intra-subject).

The squared Pearson correlation coefficients (R^2) were also investigated for the LC-MS/MS data set. The coefficients ranged from 0.71 to 0.99 (n = 70) with an average of 0.89. The weakest correlation with 0.71 was observed for the comparison between subject A (time-point 08.00 h) and subject C (time-point 08.00 h).

In the same way as for the gel-based approach, we observed for gel-free analysis that the average of the intra-subject squared correlation coefficients (intra-R^2 = 0.92) showed greater correlation than inter-subject comparisons (inter-R^2 = 0.86) (Table S3). In addition, we used heat maps based on R^2 values to visualize the data. Samples of one subject clustered closer together than those from the same time-point supporting the view that inter-subject variation had the strongest impact on the variability of the data set (Fig. 4, right panel).

To assess the stability of the saliva proteome pattern, we estimated the SD from the average protein intensities. The intra-subject SDs clustered closer than those of the inter-subject samples indicating a high rank correlation for the intra-subject comparison (Fig. S3). Subject C displayed wider spreading of values over a wide range of values and differed from other subjects that remained more homogeneous. These deviations are displayed in Figure S1 and the

mean SD of subjects A, B and D were comparable whereas the SD of subject C was clearly larger. Additionally, for each time-point all four subjects were displayed to reveal if the time-points of saliva sample collection were of relevance for the proteome variance. Although subject C was included in the analysis, the deviation of the five time-points was comparable and indicated for the LC-MS/MS experiment that inter-subject deviation was not time-dependent (Fig. S2). A hierarchical clustering of the SD of protein intensities of the gel-free experiment indicated the highest heterogeneity for subject C (mean SD of 0.85). This analysis also demonstrated that the majority of SDs was relatively low, ranging from 0.38 for subject A to 0.53 for subject C and that a large proportion of proteins displayed rather stable values.

Analysis of relative abundance revealed that only relatively few proteins dominated the whole saliva proteome. Together, the top 30 abundant proteins contributed up to 85% of the total saliva protein and the majority of them were secreted proteins (Table 1).

The individual protein CV values of the 30 most abundant proteins ranged from 0.08 to 0.58 for the intra-subject variability and from 0.21 to 0.92 for the inter-subject variabilities indicating that abundant proteins also vary in level at different time-points and between subjects. The SD ranking over all proteins ($\geq$ 2 peptides, n = 346) spanned from the most 'stable' protein (immunoglobulin kappa chain C) with a rank of 1 to more 'unstable' proteins such as protein S100-A9 with only a rank of 309.

This observation was confirmed by ANOVA: time-points (P-value = 0.20) had no significant influence of the variance between subjects. Thus, these data confirmed the observation of 2-D DIGE analysis that saliva proteome analysis of inter-subject variability was largely independent of sampling time (Table S4). However, the subjects (P = 2.8E-3) have a significant influence for the different time-points revealing high inter-subject variability of the saliva proteome.

The next step was to verify the time-independency of saliva collection, and therefore we collected whole saliva from our four subjects on two different days. These saliva proteome samples (n = 8) were only analyzed in triplicates by LC-MS/MS resulting in 24 LC-MS/MS runs of 100 min each. The resulting LC-MS/MS raw files were analyzed alone and together with the previous sample set (one day with three collection points) using Elucidator. LC-MS/MS analysis of this subset yielded in the identification of 622 unique proteins from 2923 unique peptides using the SEQUEST algorithm [$\geq$ 1 peptide per protein, high confidence < 1% false discovery rate (FDR)]; 372 proteins were identified with at least two peptides per protein. The total number of protein identifications varied between the subjects (separate LC-MS measurements) from 183 to 253 proteins ($\geq$ 2 peptides per protein), with an average of 220 proteins per run.

The squared Pearson correlation coefficient of the protein intensities indeed reached a high average value of 0.96 for intra-subject and was comparable with the intra-subject value of at least > 0.90 of the three time-points of the analysis of the first day.

The inter-subject value of 0.88 was clearly lower but had nearly the same value as the three time-points per day analysis with 0.85 and supported the importance of inter-subject variability.

In this study, gel-based (2-D DIGE) and gel-free (LC-MS/MS) turned out to be feasible and robust techniques for the detection of intra-subject or inter-subject variabilities in whole saliva studies. However, keeping efficient workflows and labor intensity in mind one would probably prefer gel-free approaches in larger-scale studies. The daily variation in human saliva has been described in detail in previous publications using different methodologies (multiplexing iTRAQ, intact protein LC-ESI-TOF MS or 1-D PAGE) (38–42). All studies presented similar observations that greater variation was readily apparent between subjects than within daily profiles of a single subject. Quintana *et al.* (26) analyzed the global inter-subject variability of 2-DE proteome profiles of 'normal' adults. The collection time and gender had close to no effect on saliva proteome profiles; however, the inter-individual variability was likewise significant. The intra- and inter-subject variations in our analyzed whole saliva sample set are supported by previous findings. Our results are in good agreement with the variation to other human body fluids such as human urine, where the intra-subject variability contributed to 45% and inter-subject variability to 47% of the total variability (37). Similar findings, though with the focus on food-related enzymatic activities showed that the saliva composition for intra-subject variability was smaller than inter-subject variability (43).

In this study, saliva samples were collected with a Salivette® (Sarstedt, Nümbrecht, Germany). We are well aware of the fact that different sampling devices might generate slightly differing proteome coverage. However, using gel-based as well as mass spectrometry centered proteomics approaches we have been able to shown excellent technical reproducibility of the whole workflow, including sampling [technical CVs of 0.17 (2-D DIGE) and 0.11 (LC-MS/MS)], thus proving its suitability for large population-based studies that require easy and reproducible handling and a representative proteome coverage.

The application of the designed approach using 2-D DIGE and/or LC-MS/MS for salivary biomarker screening will be feasible because inter-subject variability was observed independent of time of sampling at three time-points. Our finding supports the concept that saliva sampling does not need to be performed at an exactly predetermined time of the day, which is an important prerequisite for including saliva sample screenings in population-based cohort studies.

Conclusion

Here, we present a study to assess the proteome variability in whole saliva

Table 1. TOP 30 protein list of the whole saliva proteome data set. Proteins were ranked according to their intensities and the contributed protein amount was calculated. Proteins were considered as a new protein group when unique peptides were identified. The TOP 30 proteins account for up to 80% of the protein amount; however, the mean standard deviation of protein intensities over the time has a notable variation

Protein name	Gene name	Protein description	Subcellular localization	Length (AA)	Protein amount (% of total intensity)	Protein amount (% cumulative)	CV [intra-subject variability]	CV [inter-subject variability]	Mean SD	RankSD
P02768	ALB	Serum albumin OS = Homo sapiens	Secreted	609	15.0	15.0	0.27	0.35	0.446	116
P01833	PIGR	Polymeric immunoglobulin receptor OS = Homo sapiens	Cell membrane, Secreted, [more entries]	764	9.9	24.9	0.11	0.37	0.239	3
P04745	AMY1A	Salivary alpha-amylase OS = Homo sapiens	Secreted	511	6.8	31.7	0.36	0.39	0.470	144
P01834	IGKC	Immunoglobulin kappa chain C region OS = Homo sapiens	NA	106	6.3	37.9	0.08	0.30	0.204	1
P12273	PIP	Prolactin-inducible protein OS = Homo sapiens	Secreted	146	5.7	43.7	0.13	0.35	0.276	14
P01036	CST4	Cystatin-S OS = Homo sapiens	Secreted	141	5.4	49.0	0.29	0.46	0.467	140
P01876	IGHA1	Immunoglobulin alpha-1 chain C region OS = Homo sapiens	NA	353	4.9	53.9	0.14	0.29	0.255	9
P25311	AZGP1	Zinc-alpha-2-glycoprotein OS = Homo sapiens	Secreted	298	4.1	58.0	0.19	0.22	0.260	10
P01037	CST1	Cystatin-SN OS = Homo sapiens	Secreted	141	3.3	61.2	0.27	0.44	0.455	126
Q9HC84	MUC5B	Mucin-5B OS = Homo sapiens	Secreted	5762	2.8	64.0	0.39	0.51	0.616	229
P01591	IGJ	Immunoglobulin J chain OS = Homo sapiens	Secreted	159	2.2	66.2	0.14	0.46	0.307	23
P09228	CST2	Cystatin-SA OS = Homo sapiens	Secreted	141	1.9	68.1	0.25	0.76	0.496	160
A8K2U0	A2ML1	Alpha-2-macroglobulin-like protein 1 OS = Homo sapiens	Secreted	1454	1.9	70.0	0.44	0.44	0.610	225
Q9UBC9	SPRR3	Small proline-rich protein 3 OS = Homo sapiens	Cytoplasm	169	1.8	71.7	0.49	0.48	0.682	258
Q6P5S2	C6orf58	UPF0762 protein C6orf58 OS = Homo sapiens	Secreted	330	1.4	73.2	0.16	0.55	0.434	108
P23280	CA6	Carbonic anhydrase 6 OS = Homo sapiens	Secreted	308	1.3	74.5	0.43	0.63	0.806	303
P04080	CSTB	Cystatin-B OS=Homo sapiens	Cytoplasm, Nucleus	98	1.3	75.8	0.38	0.45	0.562	202
P02788	LTF	Lactotransferrin OS = Homo sapiens	Secreted	710	0.9	76.7	0.41	0.62	0.701	269
Q9UGM3	DMBT1	Deleted in malignant brain tumors 1 protein OS = Homo sapiens	Secreted	2413	0.9	77.6	0.32	0.72	0.486	155
P01857	IGHG1	Immunoglobulin gamma-1 chain C region OS = Homo sapiens	Secreted	330	0.9	78.5	0.26	0.46	0.477	151
P06702	S100A9	Protein S100-A9 OS = Homo sapiens	Secreted, Cytoplasm	114	0.9	79.4	0.58	0.92	0.855	309
P63261	ACTB	Actin, cytoplasmic 1 OS = Homo sapiens	Cytoplasm, Cytoskeleton	375	0.8	80.2	0.20	0.39	0.346	51
P04406	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase OS = Homo sapiens	Cytoplasm, Cytosol, Nucleus, Cytoplasm	335	0.8	80.9	0.23	0.32	0.362	65
P06733	ENO1	Alpha-enolase OS = Homo sapiens		434	0.7	81.7	0.18	0.21	0.262	11

Table 1. (continued)

Protein name	Gene name	Protein description	Subcellular localization	Length (AA)	Protein amount (% of total intensity)	Protein amount (% cumulative)	CV [intra-subject variability]	CV [inter-subject variability]	Mean SD	RankSD
			Cytoplasm, Cell membrane, Cytoplasm							
P02787	TF	Serotransferrin OS = Homo sapiens	Secreted	698	0.7	82.3	0.21	0.29	0.393	82
P28325	CST5	Cystatin-D OS = Homo sapiens	Secreted	142	0.7	83.0	0.35	0.79	0.691	265
P61769	B2M	Beta-2-microglobulin OS = Homo sapiens	Secreted	119	0.6	83.6	0.25	0.49	0.343	49
P10599	TXN	Thioredoxin OS = Homo sapiens	Nucleus, Cytoplasm, Secreted	105	0.6	84.2	0.33	0.43	0.499	162
P80188	LCN2	Neutrophil gelatinase-associated lipocalin OS = Homo sapiens	Secreted	198	0.5	84.7	0.12	0.29	0.266	12
P22079	LPO	Lactoperoxidase OS = Homo sapiens	Secreted	712	0.5	85.2	0.28	0.33	0.420	101

samples without extensive sample concentration, prefractionation or depletion of most abundant proteins. We show for both gel-based 2-D DIGE and LC-MS/MS approaches that the median CV for the intra-subject variability was clearly lower than the inter-subject variability. The determination of proteome variability in whole saliva should be useful for assay development of saliva as a diagnostic and/or prognostic biomarker fluid. Determining the variation for each protein and ranking them accordingly might provide an overview of the suitability of specific proteins for biomarker screening. However, only case–control studies using saliva samples will provide the necessary information about intra-group variation and inter-group differences.

We were able to identify 160–287 proteins ($\geq$ 2 peptides per protein) with an average of 212 proteins in one LC-MS/MS run of about 1 h LC gradient analysis time using only about 500 ng of the digested whole saliva protein extract with high conservative confidence settings of < 1% FDR.

This is in contrast to numerous previous reports pooling samples to achieve enough material for extensive prefractionation for cataloguing the saliva proteome. The highly sensitive, robust and time-saving approach of LC-MS/MS analysis make gel-free approaches to a method of choice in large-scale proteome analysis of whole saliva. We present a robust protocol without any prefractionation steps that enables a reasonable depth of proteome coverage while minimizing overall analysis time.

This analysis needs only a small amount of protein and is therefore applicable for population-based studies where the amount of patient samples and analyzing time are limited. Thus, high-resolution LC-MS/MS (gel-free) appears to be sufficient to facilitate biomarker screenings based on the whole saliva.

Conflict of interest

There are no conflicts of interest to declare.

Acknowledgement

The authors would like to thank Jette Anklam and Ulrike Lissner for their excellent technical assistance and Alexander Teumer for support in the bioinformatic analyses. This work is part of the research project Greifswald Approach to Individualized Medicine (GANI_MED). The GANI_MED consortium is funded by the Federal Ministry of Education and Research and the Ministry of Cultural Affairs of the Federal State of Mecklenburg–West Pomerania (03IS2061A). GABA, Switzerland, provided unlimited education funds to support BH.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. For identification of protein spots of whole saliva a pool of 400 μg protein was separated by 2-DE and stained with coomassie brilliant blue.

Figure S2. Subject ordered standard deviation of spot volumes of the 2-DIGE experiment.

Figure S3. Subject ordered standard deviation of protein intensities of the LC-MS/MS experiment.

Table S1. MALDI protein identification_n106.xlsx.

Table S2. LC–MS protein identification list with corresponding proteins.

Table S3. Pearson correlation coefficient (R^2) of protein intensities.

Table S4. Table of analysis of variance (ANOVA) values of the gel-based and gel-free approaches.

References

1. Spielmann N, Wong D. Saliva: diagnostics and therapeutic perspectives. *Oral Dis* 2011;**17**:345–354.
2. Castagnola M, Picciotti PM, Messana I *et al*. Potential applications of human saliva as diagnostic fluid. *Acta Otorhinolaryngol Ital* 2011;**31**:347–357.
3. Hu S, Loo JA, Wong DT. Human saliva proteome analysis and disease biomarker discovery. *Expert Rev Proteomics* 2007; **4**:531–538.
4. Punyadeera C, Dimeski G, Kostner K, Beyerlein P, Cooper-White J. One-step homogeneous C-reactive protein assay for saliva. *J Immunol Methods* 2011; **373**:19–25.
5. Goncalves LR, Soares MR, Nogueira FC *et al*. Comparative proteomic analysis of whole saliva from chronic periodontitis patients. *J Proteomics* 2010;**73**: 1334–1341.
6. Hu S, Loo JA, Wong DT. Human saliva proteome analysis. *Ann N Y Acad Sci* 2007;**1098**:323–329.
7. Pfaffe T, Cooper-White J, Beyerlein P, Kostner K, Punyadeera C. Diagnostic potential of saliva: current state and future applications. *Clin Chem* 2011;**57**:675–687.
8. Giusti L, Baldini C, Bazzichi L *et al*. Proteome analysis of whole saliva: a new tool for rheumatic diseases–the example of Sjogren's syndrome. *Proteomics* 2007;**7**:1634–1643.
9. Lopez ME, Colloca ME, Paez RG, Schallmach JN, Koss MA, Chervonagura A. Salivary characteristics of diabetic children. *Braz Dent J* 2003;**14**:26–31.
10. Rao PV, Reddy AP, Lu X *et al*. Proteomic identification of salivary biomarkers of type-2 diabetes. *J Proteome Res* 2009; **8**:239–245.
11. Haigh BJ, Stewart KW, Whelan JR, Barnett MP, Smolenski GA, Wheeler TT. Alterations in the salivary proteome associated with periodontitis. *J Clin Periodontol* 2010;**37**:241–247.
12. Dowling P, Wormald R, Meleady P, Henry M, Curran A, Clynes M. Analysis of the saliva proteome from patients with head and neck squamous cell carcinoma reveals differences in abundance levels of proteins associated with tumour progression and metastasis. *J Proteomics* 2008; **71**:168–175.
13. Hu S, Arellano M, Boontheung P *et al*. Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res* 2008;**14**:6246–6252.
14. Xiao H, Zhang L, Zhou H, Lee JM, Garon EB, Wong DTW. Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. *Mol Cell Proteomics* 2012; **11**: M111.012112.
15. Walz A, Stuhler K, Wattenberg A *et al*. Proteome analysis of glandular parotid and submandibular-sublingual saliva in comparison to whole human saliva by two-dimensional gel electrophoresis. *Proteomics* 2006;**6**:1631–1639.
16. Bandhakavi S, Van RS, Tawfik PN *et al*. Hexapeptide libraries for enhanced protein PTM identification and relative abundance profiling in whole human saliva. *J Proteome Res* 2011;**10**:1052–1061.
17. Ruhl S. The scientific exploration of saliva in the post-proteomic era: from database back to basic function. *Expert Rev Proteomics* 2012;**9**:85–96.
18. Krief G, Deutsch O, Gariba S, Zaks B, Aframian DJ, Palmon A. Improved visualization of low abundance oral fluid proteins after triple depletion of alpha amylase, albumin and IgG. *Oral Dis* 2011;**17**:45–52.
19. Hardt M, Witkowska HE, Webb S *et al*. Assessing the effects of diurnal variation on the composition of human parotid saliva: quantitative analysis of native peptides using iTRAQ reagents. *Anal Chem* 2005;**77**:4947–4954.
20. Ghafouri B, Tagesson C, Lindahl M. Mapping of proteins in human saliva using two-dimensional gel electrophoresis and peptide mass fingerprinting. *Proteomics* 2003;**3**:1003–1015.
21. Guo T, Rudnick PA, Wang W, Lee CS, Devoe DL, Balgley BM. Characterization of the human salivary proteome by capillary isoelectric focusing/nanoreversed-phase liquid chromatography coupled with ESI-tandem MS. *J Proteome Res* 2006;**5**:1469–1478.
22. Xie H, Onsongo G, Popko J *et al*. Proteomics analysis of cells in whole saliva from oral cancer patients via value-added three-dimensional peptide fractionation and tandem mass spectrometry. *Mol Cell Proteomics* 2008;**7**:486–498.
23. Bandhakavi S, Stone MD, Onsongo G, Van RS, Griffin TJ. A dynamic range compression and three-dimensional peptide fractionation analysis platform expands proteome coverage and the diagnostic potential of whole saliva. *J Proteome Res* 2009;**8**:5590–5600.
24. Helmerhorst EJ, Oppenheim FG. Saliva: a dynamic proteome. *J Dent Res* 2007; **86**:680–693.
25. Esser D, Alvarez-Llamas G, de VM, Weening D, Vonk RJ, Roelofsen H. Sample stability and protein composition of saliva: implications for its use as a diagnostic fluid. *Biomark Insights* 2008; **3**:25–27.
26. Quintana M, Palicki O, Lucchi G *et al*. Inter-individual variability of protein patterns in saliva of healthy adults. *J Proteomics* 2009;**72**:822–830.
27. Schulz BL, Cooper-White J, Punyadeera CK. Saliva proteome research: current status and future outlook. *Crit Rev Biotechnol* 2012; **??**: ??–??.
28. Topkas E, Keith P, Dimeski G, Cooper-White J, Punyadeera C. Evaluation of saliva collection devices for the analysis of proteins. *Clin Chim Acta* 2012; **413**:1066–1070.
29. Afkarian M, Bhasin M, Dillon ST *et al*. Optimizing a proteomics platform for urine biomarker discovery. *Mol Cell Proteomics* 2010;**9**:2195–2204.
30. Hammer E, Bien S, Salazar MG *et al*. Proteomic analysis of doxorubicin-induced changes in the proteome of

HepG2cells combining 2-D DIGE and LC-MS/MS approaches. *Proteomics* 2010; **10**:99–114.

31. Thiele T, Steil L, Gebhard S *et al*. Profiling of alterations in platelet proteins during storage of platelet concentrates. *Transfusion* 2007;**47**:1221–1233.
32. Steil L, Thiele T, Hammer E *et al*. Proteomic characterization of freeze-dried human plasma: providing treatment of bleeding disorders without the need for a cold chain. *Transfusion* 2008;**48**:2356–2363.
33. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med* 1990;**9**:811–818.
34. Eymann C, Dreisbach A, Albrecht D *et al*. A comprehensive proteome map of growing Bacillus subtilis cells. *Proteomics* 2004;**4**:2849–2876.
35. Karanam NK, Grabarczyk P, Hammer E *et al*. Proteome analysis reveals new mechanisms of Bcl11b-loss driven apoptosis. *J Proteome Res* 2010;**9**:3799–3811.
36. R_Development_Core_Team. *R: A language and environment for statistical computing, R Development Core Team*. Vienna: R Foundation for Statistical Computing, 2008.
37. Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. *J Proteome Res* 2011;**10**:637–645.
38. Gusman H, Leone C, Helmerhorst EJ *et al*. Human salivary gland-specific daily variations in histatin concentrations determined by a novel quantitation technique. *Arch Oral Biol* 2004;**49**:11–22.
39. Messana I, Cabras T, Inzitari R *et al*. Characterization of the human salivary basic proline-rich protein complex by a proteomic approach. *J Proteome Res* 2004;**3**:792–800.
40. Millea KM, Krull IS, Chakraborty AB, Gebler JC, Berger SJ. Comparative profiling of human saliva by intact protein LC/ESI-TOF mass spectrometry. *Biochim Biophys Acta* 2007;**1774**:897–906.
41. Morales-Bozo I, Urztua-Orellana B, Dominguez P, Aguilera S, Lopez-Solis R. Patterns and variability in electrophoretic polypeptide profiles of human saliva in a healthy population. *J Physiol Biochem* 2006;**62**:179–188.
42. Neyraud E, Sayd T, Morzel M, Dransfield E. Proteomic analysis of human whole and parotid salivas following stimulation by different tastes. *J Proteome Res* 2006;**5**:2474–2480.
43. Neyraud E, Palicki O, Schwartz C, Nicklaus S, Feron G. Variability of human saliva composition: Possible relationships with fat perception and liking. *Arch Oral Biol* 2012;**57**:556–566.