

RESEARCH

Open Access

Analysis of age and gender associated *N*-glycoproteome in human whole saliva

Shisheng Sun^{1†}, Fei Zhao^{1†}, Qinzhe Wang^{1†}, Yaogang Zhong¹, Tanxi Cai², Peng Wu², Fuquan Yang^{2*} and Zheng Li^{1*}

Abstract

Background: Glycoproteins comprise a large portion of the salivary proteome and have great potential for biomarker discovery and disease diagnosis. However, the rate of production and the concentration of whole saliva change with age, gender and physiological states of the human body. Therefore, a thorough understanding of the salivary glycoproteome of healthy individuals of different ages and genders is a prerequisite for saliva to have clinical utility.

Methods: Formerly *N*-linked glycopeptides were isolated from the pooled whole saliva of six age and gender groups by hydrazide chemistry and hydrophilic affinity methods followed by mass spectrometry identification. Selected physicochemical characteristics of salivary glycoproteins were analyzed, and the salivary glycoproteomes of different age and gender groups were compared based on their glycoprotein components and gene ontology.

Results and discussion: Among 85 *N*-glycoproteins identified in healthy human saliva, the majority were acidic proteins with low molecular weight. The numbers of salivary *N*-glycoproteins increased with age. Fifteen salivary glycoproteins were identified as potential age- or gender-associated glycoproteins, and many of them have functions related to innate immunity against microorganisms and oral cavity protection. Moreover, many salivary glycoproteins have been previously reported as disease related glycoproteins. This study reveals the important role of salivary glycoproteins in the maintenance of oral health and homeostasis and the great potential of saliva for biomarker discovery and disease diagnosis.

Keywords: Saliva, Glycoproteome, Glycoproteins, Age, Gender, Hydrazide chemistry, Hydrophilic affinity, Mass spectrometry

Background

Whole saliva is a slightly cloudy colorless liquid which is mainly comprised of the secretions of the parotid, submandibular, sublingual and minor salivary glands, and it exhibits multiple host defense functions in the maintenance of oral health [1]. Compared to other body fluids such as blood, cerebral spinal fluid (CSF) and urine, the collection of whole saliva is easy, noninvasive, safe, simple and cost-effective. Moreover, the significant overlap in protein content between saliva and plasma also suggests that saliva could be a potentially attractive fluid for

disease biomarker discovery as well as a diagnostic alternative to blood tests [2,3]. However, a thorough understanding of whole saliva is a prerequisite for human saliva to have diagnostic utility.

For this reason, several salivary studies have focused on the proteomic composition of human whole saliva. In the published proteomic study of whole saliva, in 2004, Hu *et al.* [4] identified 64 non-redundant proteins using 2D-gel electrophoresis (2-DE) coupled with mass spectrometry (MS) analysis. In the following years, more advanced technologies for pre-separation and MS identification such as 2D-LC-MS/MS were used, and the numbers of identified salivary proteins increased to more than 1,500 [5-8].

The analysis of the salivary glycoproteome has also been conducted. Glycosylation is a common posttranslational modification which plays an important role in

* Correspondence: fqyang@bp.ac.cn; zhengli@nwu.edu.cn

†Equal contributors

²Laboratory of Proteomics, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, P.R. China

¹Laboratory of Functional Glycomics, College of Life Sciences, Northwest University, Xi'an 710069, P.R. China

many cellular processes, e.g., protein conformation, folding, transport, targeting, and stability [9]. Many biological processes such as cell growth, differentiation, cell–cell communication, immune response, and microbial pathogenesis are also effected by glycosylation [10–13]. Moreover, glycosylation changes in glycoproteins have been identified in various diseases and can be used as biomarkers for disease diagnosis or prognosis [14,15]. Therefore, it is not surprising that there has been an increasing effort in applying glycoproteomic technologies to identify additional disease biomarkers from specific organs or in bodily fluids. These methods include lectin affinity, hydrazide chemistry, hydrophilic affinity, boronic acid affinity, size exclusion chromatography, and titanium dioxide-based enrichment [16–22]. Ramachandran, *et al.* was the first group to publish a study related to the salivary glycoproteome. They identified 84 formerly *N*-glycosylated peptides corresponding to 45 glycoproteins using the hydrazide capture technique coupled with mass spectrometry analysis [23]. Larsen *et al.* used the TiO₂ enrichment strategy to isolate the whole salivary sialome, and 97 *N*-linked glycosylation sites were identified [20]. Ramachandran, *et al.* subsequently extended the salivary glycoprotein catalogue using the modified hydrazide capture method, and they identified a total of 156 formerly *N*-glycosylated peptides representing 77 unique *N*-glycoproteins in salivary fluid [24]. Most recently, using a novel hexapeptide library method, Bandhakavi, *et al.* significantly increased the number of salivary glycoproteins to 192 [25]. However, all these salivary *N*-glycoproteins were either from adult donors or the specific age and gender annotations were missing.

It is known that the composition of saliva changes with the physiological states of the human body [26]. The rate of production and the concentration of saliva differ before and after meals. Additionally, there are known differences in the production and concentration of saliva according to age and gender. It has been reported that the structure and function of salivary glands change with age and gender [27,28]. Animal research as well as human studies have revealed gender differences in the composition and production rate of saliva [29–31]. Moreover, women are more frequently affected by autoimmune diseases, such as Sjogrens syndrome and systemic lupus erythematosus, which affect salivary gland function [32]. Therefore, the first step toward the development of saliva diagnostic tests requires an understanding of the protein expression profiles of healthy individuals of different ages and genders.

In this study, two glycopeptide extraction methods, hydrazide chemistry and hydrophilic affinity, were employed coupled with mass spectrometry analysis to profile the whole salivary *N*-glycoproteomes of six different age and gender groups. The composition and biological functions of the whole salivary glycoproteomes among various age and

gender groups were compared to determine differentiating trends. The results can facilitate an improved understanding of the significant functions of salivary glycoproteins in oral health and homeostasis as well as enhance the great potential of saliva for biomarker discovery and disease diagnosis.

Results and discussion

Protein concentrations and composition of human whole saliva

Six pooled saliva samples of different age and gender groups, including two gender groups of children (5–7 years old, mean age of 6.2 years, 28 males and 25 females), two gender groups of young adults (21–25 years old, mean age of 24.5 years, 13 males and 13 females) and two gender groups of the elderly (65–90 years of age, mean age of 71.7 years, 16 males and 14 females), were used for the study. After the collection and pooling of human whole saliva, the protein concentrations and composition of pooled saliva from the six age and gender groups were measured by BCA protein assay and SDS-PAGE, respectively. The overall trend was that the mean protein concentration of human whole saliva increased with age (Figure 1), in agreement with the results of a previous report [33]. The increasing rate was higher in females than in males. The increased age-related salivary protein concentration may be due to a combined effect of both the reduction of the production rate of whole saliva [34] and the increase of total protein content [35] as people age. The protein composition of saliva from different age and gender groups also showed a visible difference via gel electrophoresis (Additional file 1: Figure S1). The conspicuous differences of total protein content in human whole saliva among different ages and

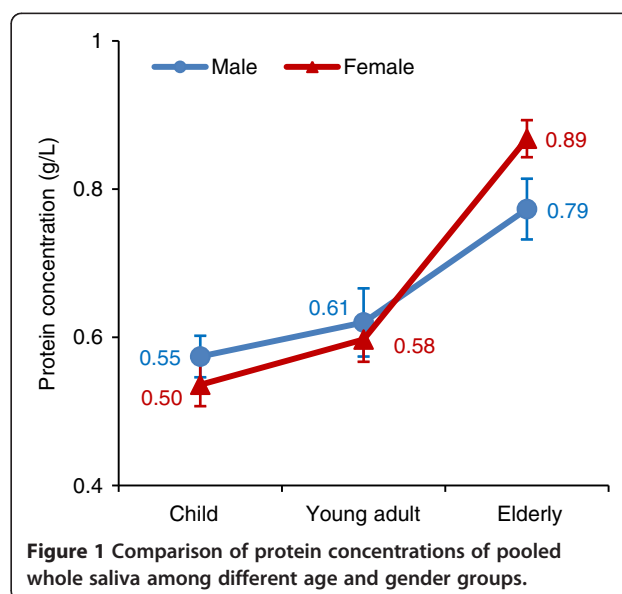


Figure 1 Comparison of protein concentrations of pooled whole saliva among different age and gender groups.

genders and the abundance of glycosylation on salivary proteins indicated the significance of the research regarding age- and gender-associated changes in the glycoproteome of human whole saliva.

Identification of salivary *N*-glycoproteins

Two different methods, hydrazide chemistry and hydrophilic affinity, were employed for the isolation of formerly *N*-glycopeptide from six pooled saliva samples. To obtain a better understanding of the age- and gender-associated glycoprotein composition of whole saliva, an equivalent amount of salivary proteins (0.5 mg) from each group was used for *N*-linked glycopeptide isolation by each method. A total of 156 nonredundant formerly *N*-linked glycopeptides (*N*-deglycopeptides) representing 164 unique *N*-glycosylation sites and 85 *N*-linked glycoproteins were identified from human whole saliva by the two glycopeptide isolation methods (Additional file 2: Table S1). Of these, 119 glycopeptides (76%) from 70 glycoproteins (82%) were annotated as “known” *N*-glycopeptides in the UniProt database and 23 glycopeptides (15%) from 11 glycoproteins (13%) were annotated as “potential”. The other 14 *N*-glycopeptides (9%) were “novel” *N*-glycopeptides identified in this study (Additional file 2: Table S1, Additional file 3: Figure S2). After glycoprotein identification, some physiochemical characteristics of salivary glycoproteins, mainly their isoelectric point (pI) values and molecular weights (MWs), were statistically analyzed (Figure 2). The pI values of the identified salivary glycoproteins were between 4 and 12 (Figure 2A), with most clustering around 5–7, and 8–9. Approximately 3/4

(74.1%) of the identified salivary *N*-glycoproteins had pI values ≤ 7 . The MW of most salivary *N*-glycoproteins (81.2%) ranged from 10–90 kDa and more than half (61.2%) ranged from 10–60 kDa (Figure 2B), which is similar to the MW distribution of the salivary proteome [3]. The results from the above analysis indicate that most salivary glycoproteins are acidic with low molecular weight.

Comparison of salivary *N*-glycoproteomes among different age groups

To further understand the age-associated alterations of the human salivary *N*-glycoproteome, a comparison of human salivary *N*-glycoproteomes was conducted among different age groups after mass spectrometry identification. Note that for some glycoproteins, the differences among different age and gender groups detected in this study might just reflect the concentration changes of the glycoproteins in the whole saliva among age and gender groups.

Among the males, there were 45 (82), 54 (93), and 69 (110) *N*-glycoproteins (*N*-deglycopeptides) identified from the children, young adult, and elderly groups, respectively (Figure 3A). Thirty five *N*-glycoproteins were commonly identified in all three age groups; twelve, five and two glycoproteins were not identified in the children, young adult and elderly groups, respectively, while three, five, and seventeen *N*-glycoproteins were uniquely identified in the whole saliva of the children, young adult, and elderly groups, respectively.

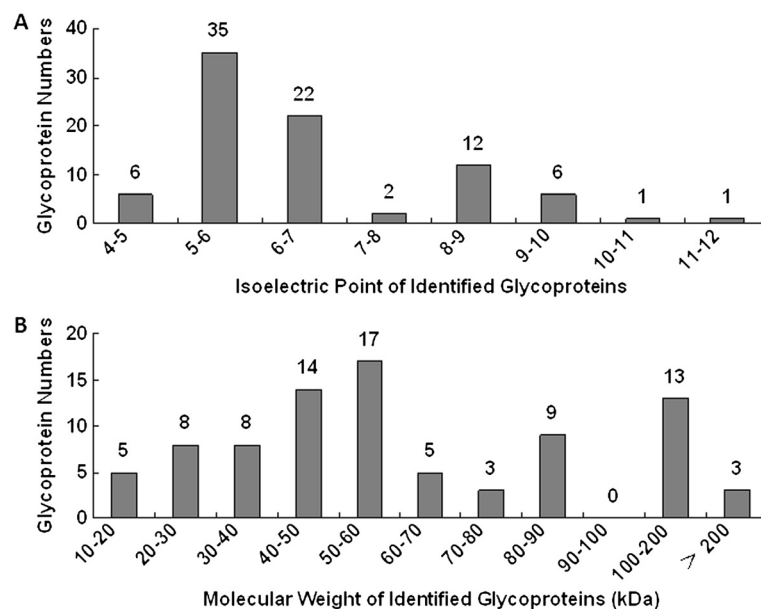
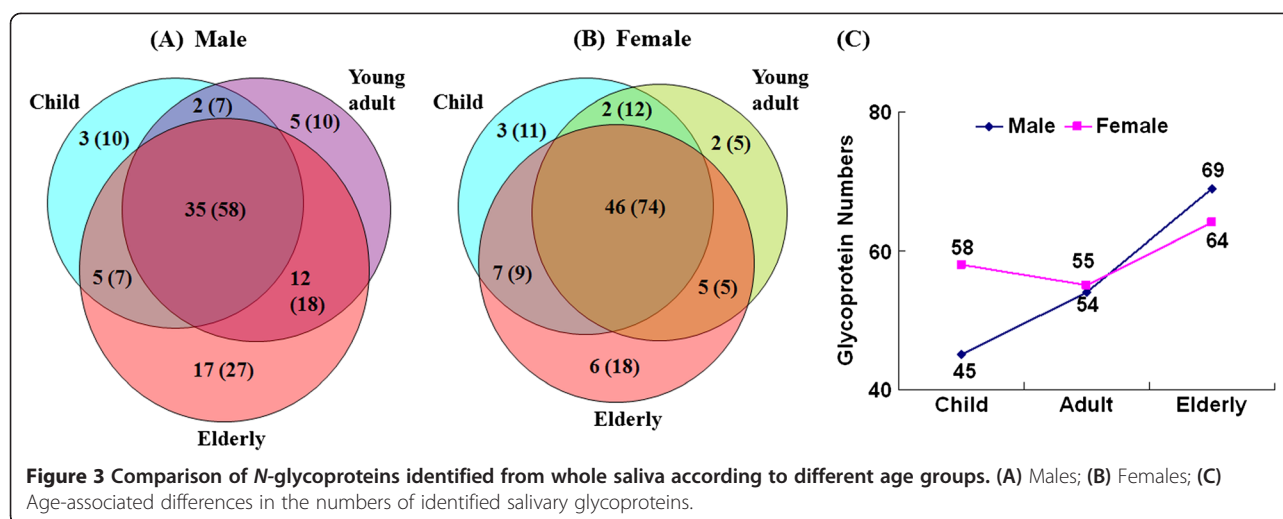


Figure 2 Isoelectric point value (A) and molecular weight (B) distribution of *N*-glycoproteins and *N*-glycopeptides identified from human saliva.



In the female groups, the number of *N*-glycoproteins (*N*-deglycopeptides) identified from the children, young adult, and elderly groups was 58 (106), 55 (96), and 64 (106), respectively (Figure 3B). Of these, 46 *N*-glycoproteins were identified in all three age groups; five, seven and two glycoproteins were not identified in the children, young adult and elderly groups, respectively. Similarly, there were three, two, and six *N*-glycoproteins that were uniquely identified from the whole saliva of the children, young adult, and elderly groups, respectively. Based on these results, there was an increasing trend in the number of salivary *N*-glycoproteins that was associated with increasing age, and the increase in the number

of *N*-glycoproteins was higher in the male groups than in the female groups, except for the female adult group which exhibited a small decrease compared to the female children group (Figure 3C).

The results also showed that eight salivary glycoproteins were likely to be associated with ageing (Table 1). For example, kallikrein-11 (KLK11) was only identified in the whole saliva of the male and female groups of children. Kallikrein-11 is a protein that belongs to a subgroup of serine proteases with diverse physiological functions [36]. Kallikreins are responsible for the coordination of various physiological functions including blood pressure, semen liquefaction and skin desquamation [37]. Golgi membrane

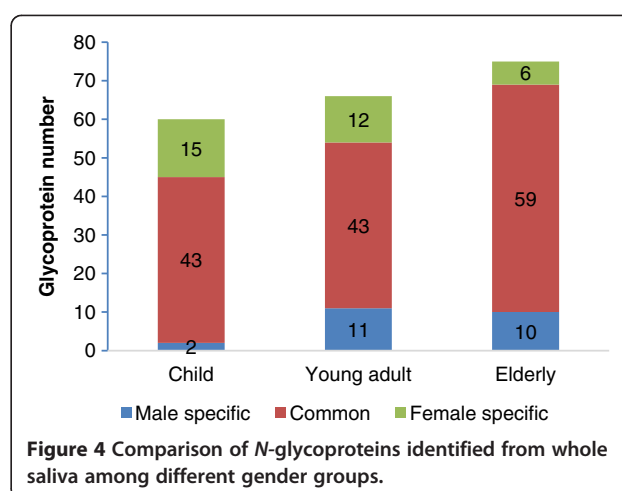
Table 1 Age- and gender-associated salivary glycoproteins

Accession	Glycoprotein name	Age or gender group
Age-associated		
IPI00002818	Isoform 1 of Kallikrein-11	Children
IPI00759659	Isoform 2 of Golgi membrane protein 1	Young adults
IPI00000877	Hypoxia up-regulated protein 1	Elderly
IPI00022371	Histidine-rich glycoprotein	Elderly
IPI00022255	Olfactomedin-4	Elderly
IPI00291410	Isoform 1 of Long palate, lung and nasal epithelium carcinoma-associated protein 1	Children, young adults
IPI00374315	UPF0762 protein C6orf58	Children, young adults
IPI00218413	Biotinidase	Young adults, elderly
Gender-associated		
IPI00006114	Pigment epithelium-derived factor	Male young adults, male elderly
IPI00006690	Eosinophil peroxidase	Male young adults, male elderly
IPI00305461	Inter-alpha (Globulin) inhibitor H2, isoform CRA_a	Female young adults, female elderly
IPI00029260	Monocyte differentiation antigen CD14	Female young adults, female elderly
IPI00291866	Plasma protease C1 inhibitor	All female groups, male elderly
IPI00478003	Alpha-2-macroglobulin	All female groups, male elderly
IPI00152154	Mucin-7	All female groups, male children

protein 1 (GOLM1) was only identified in the whole saliva of the young adult groups. GOLM1 is a type II Golgi transmembrane protein. Its biological function is still unknown, but GOLM1 is hypothesized to be involved in the cellular response to viral infection [38]. Hypoxia up-regulated protein 1 (HYOU1), histidine-rich glycoprotein (HRG), and Olfactomedin-4 were only identified in the whole saliva of the elderly groups. HYOU1 belongs to the heat shock protein 70 family and it plays an important role in cytoprotective cellular mechanisms against hypoxia/ischemia-induced neuronal death [39,40]. HRG can bind several ligands such as heparin, heme, heparan sulfate, plasminogen, thrombospondin, and divalent metal ions. It can act as an adapter protein and it is involved in many biological processes such as immune complex formation and pathogen clearance, coagulation, angiogenesis, fibrinolysis, cell chemotaxis and cell adhesion. HRG mediates the clearance of necrotic cells by enhancing the phagocytosis of necrotic cells in a heparan sulfate-dependent pathway, and it is also involved in the regulation of tumor angiogenesis and tumor immune surveillance [41]. OLFM4 has been found to be up-regulated in gastrointestinal cancer [42] and many inflammatory diseases [43,44]. Biotinidase was not detected in the whole saliva of the groups of children, and two salivary *N*-glycoproteins, BPI fold-containing family B member 1 (BPIFB1) and UPF0762 protein C6orf58, were not detected in the elderly groups. BPIFB1 is a secreted protein which may play a role in innate immunity in the mouth, nose and lungs and it has a lipid binding function [45]. The biological function of UPF0762 protein C6orf58 remains unknown. In addition, biotinidase, which functions to catalyze biotin release from biocytin, was not detected in the groups of children. Based on this analysis, many salivary glycoproteins associated with ageing, particularly those that were uniquely identified or not identified in the elderly groups, are involved in the immune response and oral cavity protection. These results indicate the important role of salivary glycoproteins in the maintenance of oral health and homeostasis.

Comparison of salivary *N*-glycoproteomes among different gender groups

A comparison of salivary *N*-glycoproteins (*N*-deglycopeptides) identified in different gender groups with the same age range is shown in Figure 4. Although children and young adult groups had same numbers (43) of common *N*-glycoproteins identified between males and females, the numbers of *N*-glycoproteins uniquely identified in the young adult group increased. The children group contained two unique *N*-glycoproteins that were specific to males and 15 that were specific to females, whereas the young adult group had 11 unique *N*-glycoproteins that were specific to males and 12 that were specific to females. Contrarily, compared to the young



adult groups, the numbers of commonly identified *N*-glycoproteins between males and females increased in the whole saliva of the elderly groups, while the number of unique *N*-glycoproteins decreased slightly. There were 59 commonly identified *N*-glycoproteins between the male and female elderly groups, while 10 *N*-glycoproteins were uniquely identified in the male elderly group and six were uniquely identified in the female elderly group. These data indicated that the young adult groups had a slightly larger difference in the numbers of salivary *N*-glycoproteins between genders among the three age groups and the number of unique glycoproteins identified in the female groups decreased with age.

The results showed that some glycoproteins might also be associated with gender (Table 1). For example, Pigment epithelium-derived factor (PEDF) and eosinophil peroxidase (EPX) were only identified in the whole saliva of the male young adult and elderly groups. Inter-alpha (Globulin) inhibitor H2, isoform CRA_a and monocyte differentiation antigen CD14 (CD14) were identified in the female young adult and elderly groups. Plasma protease C1 inhibitor (SERPING1) was not identified in the male children and young adult groups and mucin 7 was not identified in the male young adult and elderly groups. Many of these glycoproteins, especially some female-specific glycoproteins, are also immune associated proteins. For example, CD14 mediates the innate immune response to bacterial lipopolysaccharide [46]; SERPING1 may potentially play a crucial role in regulating important physiological pathways including complement activation and blood coagulation [47]; A2M is involved in complement and coagulation cascade pathways [48]; and mucin-7 may have a protective capacity in promoting the clearance of bacteria in the oral cavity [49]. With normal aging, the physiological states of the human body and microbial communities in the oral cavity may change significantly, which might cause changes

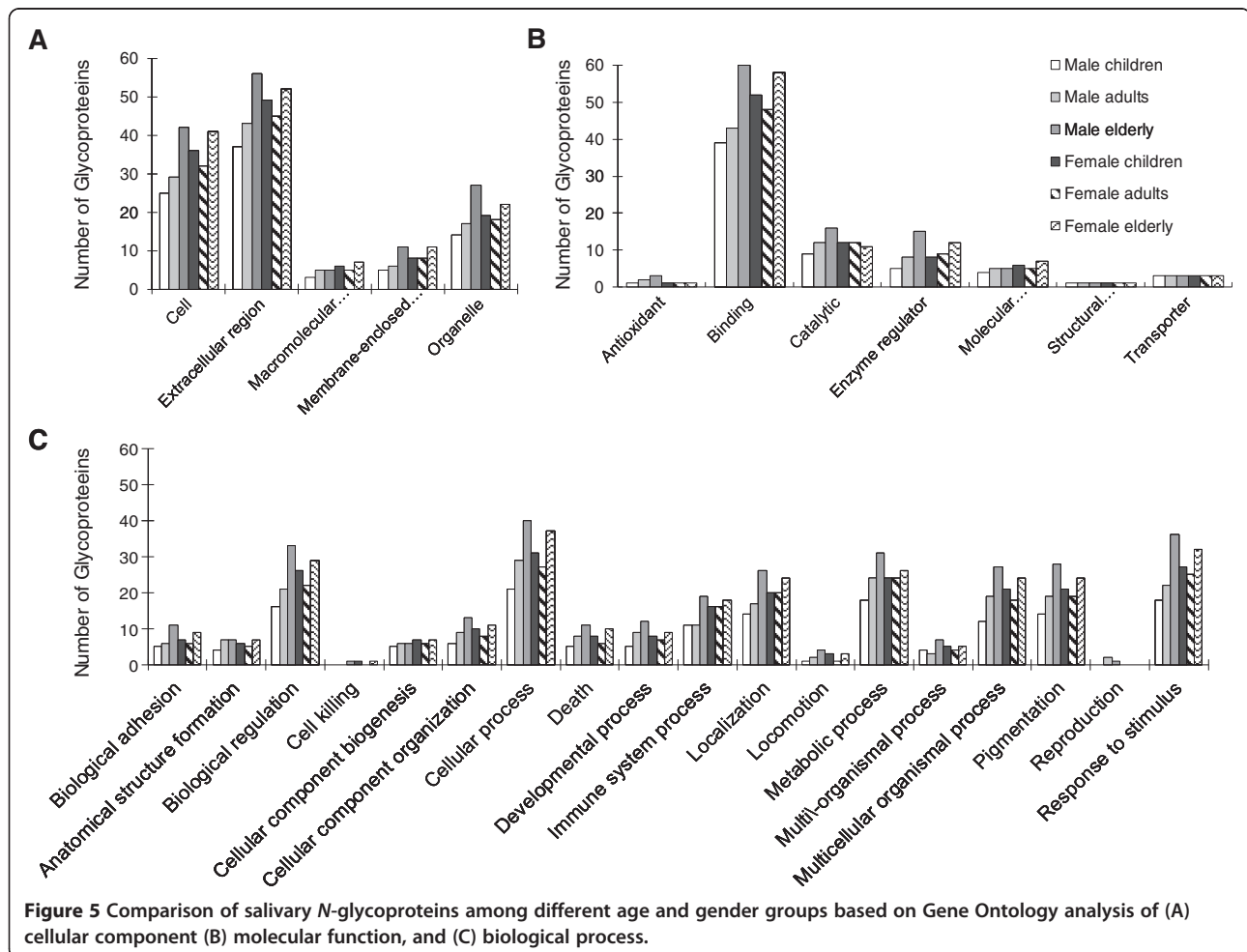
in the salivary glycoproteome [23]. The majority of the proteins that were commonly identified in all six age and gender groups, including prolyl-rich protein, statherin, cysteine containing nitric acid protein, mucin, amylase, and salivary peroxidase are the basic protein components of saliva. Conversely, many proteins that were specifically expressed in different age and gender groups might function to regulate the physiological states of the human body and adapt to the specific microbial community in the oral cavity. Therefore, these age- and gender-associated glycoproteins might play a dominant role in the maintenance of oral health [1] and the immune response [23].

Gene ontology (GO) analysis and disease association

To further study the biological function of the salivary glycoproteome in the maintenance of oral health, the salivary *N*-glycoproteins from different age and gender groups were further compared based on their GO annotation of cellular component, molecular function, and biological process terms (Figure 5). The majority of the salivary *N*-glycoproteins were located in the extracellular region (41%), cell (30%) and organelle (18%); thus, the

number of glycoproteins from these cellular locations had the most markedly increasing trend corresponding with age, especially in males (Figure 5A). The GO distribution based on their biological functions showed that 62% of the salivary *N*-glycoproteins were involved in binding, 14% in catalytic activity and 13% in enzyme regulator activity. The numbers of salivary *N*-glycoproteins with these functions showed a noticeably increasing trend with age in the male groups. Conversely, the number of *N*-glycoproteins that function as structural molecules (IPI00855918) and transporters (IPI00022488, IPI00166729, IPI00299547) remained constant in all six different age and gender groups. The number of glycoproteins with molecular transducer activity also remained nearly constant among the different groups (Figure 5B).

With regard to biological processes (Figure 5C), the numbers of salivary *N*-glycoproteins belonging to most of the categories increased with age in the male groups, especially the *N*-glycoproteins that are involved in biological adhesion, biological regulation, cellular component organization, cellular process, death, developmental process, immune system process, localization, locomotion, metabolic process, multi-organismal process, multicellular organismal process, pigmentation, reproduction, response to stimulus



and response to stimulus biological processes. Interestingly, one *N*-glycoprotein involved in cell killing (IPI00022395) and two proteins involved in reproduction (IPI00103633, IPI00006114) were also identified in the whole salivary proteome of adults and/or old males. Some exceptions to the age-related trend in the number of *N*-glycoproteins were determined. In particular, the numbers of salivary *N*-glycoproteins identified in the female young adult groups decreased slightly compared to the numbers identified in the female children, but the overall trend was also an increasing tendency with age. Of these, the number of immune-related glycoproteins increased with age in both male and female groups. However, in the younger age group, the number of immune, biological regulation, inhibition of enzyme activity and protein binding related salivary glycoproteins was significantly higher in females than in males. This difference was reduced with age and the level of these functional glycoproteins was largely similar between genders in the elderly group. These results may demonstrate the significant functions of salivary glycoproteins in oral health and homeostasis.

To assess the potential of salivary glycoproteins for disease biomarker discovery and diagnostic efforts, the association of salivary glycoproteins with human systematic diseases was also illustrated. Among the 85 proteins identified in the human whole saliva, 44 proteins (51.8%) were associated with human disease in a genetic association database and an OMIM disease database based on DAVID functional annotation (Additional file 4: Table S2). For example, nine glycoproteins including alpha-2-macroglobulin (IPI00478003) and myeloperoxidase (IPI00007244) are related to Alzheimer's disease. Nineteen, 14, 17 and 15 salivary *N*-glycoproteins are associated with metabolic, neurological, immune, and cardiovascular diseases, respectively. These results indicate the great potential of saliva for biomarker discovery and disease diagnosis.

Comparison of two glycopeptides isolation methods

A combination of hydrazide chemistry and hydrophilic affinity methods was used to increase the identification coverage of salivary glycoproteomes. Among 85 salivary *N*-glycoproteins (156 formerly *N*-glycopeptides) identified in the study, 43 *N*-glycoproteins (72 formerly *N*-glycopeptides) were identified by both methods, 39 *N*-glycoproteins (74 *N*-glycopeptides) were identified by hydrazide chemistry method uniquely, and 3 *N*-glycoproteins (10 *N*-glycopeptides) were identified uniquely by hydrophilic affinity method (Additional file 2: Table S1). Although some complementarity existed in two methods in the study, the hydrazide chemistry method showed higher specificity and identification rate of salivary *N*-glycoproteins than hydrophilic affinity method. This may be due to the inherent characteristic of these two methods: hydrazide resin captures glycoproteins/

glycopeptides by covalent bonding, and thus non-specific adsorbed proteins can be thoroughly removed by intensely washing without any loss of glycopeptides. While Sepharose cell-4B captures glycoproteins/glycopeptides with hydrophilic interaction. The washing needs to be much mild, or many glycopeptides may be lost in the washing process. Besides, O-glycopeptides will also be released from Sepharose cell-4B in the elution process which may interfere with *N*-deglycopeptide identification by mass spectrometry. All the above reasons might result in the relatively low identification rate of the hydrophilic affinity method.

Conclusion

In this study, the formerly *N*-linked glycopeptides of whole saliva were isolated, identified and compared among six different age and gender groups. The results showed that most salivary glycoproteins are acidic with low molecular weight. The number of salivary *N*-glycoproteins had an increasing age-associated trend, and the rate of increase was higher in the male groups than in the female groups. Fifteen salivary glycoproteins could be associated with gender or age. Based on their biological functions and gene ontology, many of the glycoproteins, especially those that were uniquely identified or not identified in the elderly groups, were involved in immune response and oral cavity protection. Moreover, more than half of the identified salivary glycoproteins were associated with human disease pathways. The data reveal the important role of salivary glycoproteins in the maintenance of oral health and homeostasis and the great potential of saliva for biomarker discovery and disease diagnosis.

Materials and methods

Human whole saliva collection

(1) Subject selection. Fifty-three healthy children (28 males and 25 females) between 5 and 7 years of age (mean age of 6.2 years), 26 young adults (13 males and 13 females) between 21 and 25 years of age (mean age of 24.5 years), and 30 elderly individuals (16 males and 14 females) between 65 and 90 years of age (mean age of 71.7 years) were selected from a primary school, our laboratory and a home for the aged. The subjects were randomized prior to their participation in the study. (2) Saliva collection. Whole, un-stimulated saliva was collected from subjects in the morning, 2 h after the last intake of food. The donors were asked to rinse their mouth with normal saline immediately before collection. The whole saliva was collected and placed on ice before being centrifuged at 12,000 rpm at 4°C for 60 min. The supernatant was collected and protease inhibitor (1 µL/ml whole saliva) was added to minimize protein degradation. Equal volumes of individual saliva samples were pooled to construct different age and gender pools that were as homogeneous as possible to reduce the individual variance,

and the protein amount was measured using a BCA protein analysis kit (Pierce, Rockford, IL). The mixed saliva was stored at -80°C .

The collection of human whole saliva with informed consent was approved by the Human Ethics Committee of Northwest University and conducted in accordance with the ethical guidelines of the Declaration of Helsinki.

SDS-PAGE

The mixed human whole saliva proteins were analyzed by SDS-PAGE. Ten μL of the samples were mixed with $2 \times$ loading buffer, boiled for 5 min at 95°C , and applied onto a discontinuous 10% polyacrylamide gel. After SDS-PAGE, the bands on the gels were visualized via silver staining.

Extraction and trypsin digestion of whole salivary proteins

Whole saliva containing 1 mg of proteins was concentrated by a 3kD Amicon Ultra centrifugal filter device (Millipore, Bradford, MA, USA) with denaturing buffer [50] (8 M urea in 0.1 M NH_4HCO_3 solution, pH 8.3) at 12,000 g for 20 min at 4°C (repeated three times) to exchange the solvent and denature the salivary proteins. The salivary proteins were then reduced by 5 mM DTT at 60°C for 60 min and alkylated by 20 mM iodoacetamide at room temperature in the dark for 30 min. The solution was diluted 5-fold with 0.1 M ammonia bicarbonate (pH 8.3) and the proteins were digested with 20 μg sequencing-grade modified trypsin (trypsin: protein, 1:50, w/w) overnight at 37°C . Digestion was terminated by acidifying the sample mixture with TFA to pH <3. The peptides were desalted by a Sep-Pak[®] Vac C18 cartridge (Waters, Milford, MA) and eluted in 0.4 mL of 80% ACN/0.1% TFA. The peptides were then divided into two equal aliquots. One aliquot was used for formerly *N*-glycopeptide isolation by the hydrazide chemistry method, while the other aliquot was used for formerly *N*-glycopeptide isolation by the hydrophilic affinity method.

Formerly *N*-linked glycopeptide isolation using hydrazide chemistry

The *N*-glycopeptides of salivary proteins were enriched by the hydrazide method according to previously described protocols [51,52]. Briefly, half of the peptides (~ 0.5 mg) from each group were oxidized by 10 mM sodium periodate at room temperature for 1 h in the dark. The oxidized peptide samples were diluted 16-fold with 0.1% TFA and purified by C18 column. The peptides were eluted directly into hydrazide resin (Bio-Rad, Hercules, CA) and incubated overnight at room temperature with shaking.

The resin was washed three times each with 80% ACN, 1.5 M NaCl and D.I. water. The formerly *N*-linked glycopeptides were released from resin via 2 μL PNGase F (New England Biolabs, Beverly, MA) in 100 μL of

25 mM ammonium bicarbonate at 37°C overnight with shaking. The supernatant and wash solutions were combined and dried via SpeedVac.

Glycopeptide enrichment by hydrophilic affinity method

The remaining half of the tryptic peptides (~ 0.5 mg) was mixed with agarose-hydrophilic resin (Sigma-Aldrich, St. Louis, MO) to bind the glycopeptides to the resin. After gently shaking for 45 min, the microcentrifuge tube was centrifuged (9,000 g, 5 min), and the resulting supernatant was removed. The resin was then washed three times with 80% ACN to remove non-glycosylated peptides. Finally, the glycopeptides bound to the resin were eluted twice with 200 μL H_2O . The elution solutions were combined and were concentrated to ~ 50 μL by vacuum centrifugation. The *N*-glycans were removed from *N*-glycopeptides via 2 μL PNGase F overnight at 37°C with shaking in 25 μL of 100 mM ammonium bicarbonate solution. The solution was dried via SpeedVac.

LC-MS/MS analysis

The lyophilized peptides were resuspended in 25 μL of 0.1% FA, and 8 μL was used for each LC-MS/MS run. LC-MS/MS analysis of peptides was conducted using an LC Packing nano-LC system (Agilent 1200 series) with a nanoelectrospray chip interface (Agilent) and a quadrupole TOF mass spectrometer (Agilent 6530 Accurate-Mass Q-TOF LC/MS, USA). The samples were first loaded onto an HPLC-Chip (G4240-62010, Zorbax 300SB C18 particles) for nano-LC separation at a flow rate of 300 nL/min. The eluents used for the LC were (A) 1% ACN/0.1% FA and (B) 90% ACN/0.1% FA. A gradient was utilized from 3% B to 10% B in 10 min, from 10% B to 45% B in 70 min, from 45% B to 95% B in 10 min, and held at 95% B for 15 min. Then the column was re-equilibrated for 15 min before the next run. Due to the statistical fluctuations of peptide precursor selection during the MS/MS acquisition, two LC-MS/MS assays were run for each sample to facilitate a proper proteome comparison.

Data mining and analysis

Protein identification was accomplished utilizing the MASCOT database search engine (v.2.3.02, Matrix Science, London). The MS/MS spectra were used to search the IPI human 3.74 database in which trypsin and up to one miscleavage were specified. Carbamidomethylation (C) was set as a fixed modification, while oxidation (M) and deamination (N) were set as variable modifications. A peptide tolerance of 20 ppm and a tolerance of ± 0.7 Da for the fragment ions were used. The peptide identification was filtered by a Mascot Score above 30 with $p < 0.05$. *N*-Glycopeptide identification were filtered by a deamidated (N) site at N-X-S/T motif (X is any amino acid except Pro) to reduce potential false positive identification [53].

Bioinformatic analysis of salivary glycoproteins

Gene Ontology (GO) analysis of the identified glycoproteins was conducted by Blast2GO [54], and comparative GO analysis between different age and gender groups was conducted by WEGO [55]. The whole analysis process was conducted according to standard operating procedures. The molecular weight and isoelectric point of the identified glycoproteins were obtained by MAS-COT. Disease-associated salivary glycoproteins were identified by DAVID [56] functional enrichment analysis according to standard procedures.

Additional files

Additional file 1: Figure S1. SDS-PAGE analysis of human whole saliva from different age and gender groups. Line 1: Marker, Lines 2: pooled saliva of male children, Lines 3: pooled saliva of female children, Lines 4: pooled saliva of male young adults, Lines 5: pooled saliva of female young adults, Lines 6: pooled saliva of male elderly, Lines 7: pooled saliva of female elderly. Lines 2–7 were loaded with the same volume of pooled saliva.

Additional file 2: Table S1. N-linked glycoproteins/glycopeptides and glycosylation sites identified from human whole saliva obtained from male and female patients of different ages.

Additional file 3: Figure S2. UniProt-based annotation of N-glycopeptides and N-glycoproteins identified from human saliva. (A) N-glycopeptides. (B) N-glycoproteins.

Additional file 4: Table S2. Disease associated glycoproteins identified in human whole saliva.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SS, FZ, and ZL were involved in the study conception and design. FZ and YZ collected the saliva samples. FZ, QW, SS and YZ carried out the experiments. FY, TC, SS, and PW were involved in the mass spectrometry analysis. SS, FZ and ZL wrote the manuscript. SS, FZ and QW were involved in the data analysis and bioinformatic analysis. All authors read and approved the final manuscript.

Acknowledgments

This work is supported by NFSC (Grant No. 81372365) and the international S&T cooperation program (2009DFA32730) from the Chinese Ministry of Science and Technology. The authors thank Dr. Hua Zhang of the Shaanxi Provincial People's Hospital for her excellent technical assistance.

Received: 1 October 2013 Accepted: 9 April 2014

Published: 5 June 2014

References

1. Edgar WM: Saliva: its secretion, composition and functions. *Br Dent J* 1992, **172**:305–312.
2. Malamud D: Saliva as a diagnostic fluid. *Dent Clin N Am* 2011, **55**:159–178.
3. Loo JA, Yan W, Ramachandran P, Wong DT: Comparative human salivary and plasma proteomes. *J Dent Res* 2010, **89**:1016–1023.
4. Hu S, Xie Y, Ramachandran P, Ogorzalek Loo RR, Li Y, Loo JA, Wong DT: Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. *Proteomics* 2005, **5**:1714–1728.
5. Xie S, Rhodus NL, Griffin RJ, Carlis JV, Griffin TJ: A catalogue of human saliva proteins identified by free flow electrophoresis-based peptide separation and tandem mass spectrometry. *Mol Cell Proteomics* 2005, **4**:1826–1830.
6. Wilmarth PA, Riviere MA, Rustvold DL, Lauten JD, Madden TE, David LL: Two-dimensional liquid chromatography study of the human whole saliva proteome. *J Proteome Res* 2004, **3**:1017–1023.
7. 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.
8. Fang X, Yang L, Wang W, Song T, Lee CS, DeVoe DL, Balgley BM: Comparison of electrokinetics-based multidimensional separations coupled with electrospray ionization-tandem mass spectrometry for characterization of human salivary proteins. *Anal Chem* 2007, **79**:5785–5792.
9. Helenius A, Aebi M: Intracellular functions of N-linked glycans. *Science* 2001, **291**:2364–2369.
10. Collins BE, Paulson JC: Cell surface biology mediated by low affinity multivalent protein-glycan interactions. *Curr Opin Chem Biol* 2004, **8**:617–625.
11. Lin X: Functions of heparan sulfate proteoglycans in cell signaling during development. *Development* 2004, **131**:6009–6021.
12. Sun S, Wang Q, Zhao F, Chen W, Li Z: Glycosylation site alteration in the evolution of influenza A (H1N1) viruses. *PLoS One* 2011, **6**:e22844.
13. Sun S, Wang Q, Zhao F, Chen W, Li Z: Prediction of biological functions on glycosylation site migrations in human influenza H1N1 viruses. *PLoS One* 2012, **7**:e32119.
14. Roth J: Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem Rev* 2002, **102**:285–304.
15. Durand G, Seta N: Protein glycosylation and diseases: blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring. *Clin Chem* 2000, **46**:795–805.
16. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, Kasai K-i, Takahashi N, Isobe T: Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol* 2003, **21**:667–672.
17. Sparbier K, Wenzel T, Kostrzewa M: Exploring the binding profiles of ConA, boronic acid and WGA by MALDI-TOF/TOF MS and magnetic particles. *J Chromatogr B* 2006, **840**:29–36.
18. Zhang H, Li XJ, Martin DB, Aebersold R: Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 2003, **21**:660–666.
19. Wada Y, Tajiri M, Yoshida S: Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal Chem* 2004, **76**:6560–6565.
20. Larsen MR, Jensen SS, Jakobsen LA, Heegaard NHH: Exploring the sialome using titanium dioxide chromatography and mass spectrometry. *Mol Cell Proteomics* 2007, **6**:1778–1787.
21. Alvarez-Manilla G, Atwood, Guo Y, Warren NL, Orlando R, Pierce M: Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *J Proteome Res* 2006, **5**:701–708.
22. Sun S, Yang G, Wang T, Wang Q, Chen C, Li Z: Isolation of N-linked glycopeptides by hydrazine-functionalized magnetic particles. *Anal Bioanal Chem* 2010, **396**:3071–3078.
23. Ramachandran P, Boontheung P, Xie Y, Sondej M, Wong DT, Loo JA: Identification of N-linked glycoproteins in human saliva by glycoprotein capture and mass spectrometry. *J Proteome Res* 2006, **5**:1493–1503.
24. Ramachandran P, Boontheung P, Pang E, Yan W, Wong D, Loo J: Comparison of N-linked glycoproteins in human whole saliva, parotid, submandibular, and sublingual glandular secretions identified using hydrazide chemistry and mass spectrometry. *Clin Proteomics* 2008, **4**:80–104.
25. Bandhakavi S, van Riper SK, Tawfik PN, Stone MD, Haddad T, Rhodus NL, Carlis JV, Griffin TJ: Hexapeptide libraries for enhanced protein PTM identification and relative abundance profiling in whole human saliva. *J Proteome Res* 2010, **10**:1052–1061.
26. de Almeida PV, Gregio AM, Machado MA, de Lima AA, Azevedo LR: Saliva composition and functions: a comprehensive review. *J Contemp Dent Pract* 2008, **9**:72–80.
27. Inoue H, Ono K, Masuda W, Morimoto Y, Tanaka T, Yokota M, Inenaga K: Gender difference in unstimulated whole saliva flow rate and salivary gland sizes. *Arch Oral Biol* 2006, **51**:1055–1060.

28. Pedersen W, Schubert M, Izutsu K, Mersai T, Truelove E: **Clinical science** Age-dependent decreases in human submandibular gland flow rates as measured under resting and post-stimulation conditions. *J Dent Res* 1985, **64**:822–825.
29. Lukacs JR, Largaespada LL: **Explaining sex differences in dental caries prevalence: Saliva, hormones, and “life-history” etiologies.** *Am J Hum Biol* 2006, **18**:540–555.
30. Percival RS, Challacombe S, Marsh PD: **Flow rates of resting whole and stimulated parotid saliva in relation to Age and gender.** *J Dent Res* 1994, **73**:1416–1420.
31. Heintze U, Birkhed D, Bjorn H: **Secretion rate and buffer effect of resting and stimulated whole saliva as a function of age and sex.** *Swed Dent J* 1983, **7**:227–238.
32. Ornoy A, Yacobi S, Matalon ST, Blank M, Blumenfeld Z, Miller RK, Shoenfeld Y: **The effects of antiphospholipid antibodies obtained from women with SLE/APS and associated pregnancy loss on rat embryos and placental explants in culture.** *Lupus* 2003, **12**:573–578.
33. Ambatipudi KS, Lu B, Hagen FK, Melvin JE, Yates JR: **Quantitative analysis of Age specific variation in the abundance of human female parotid salivary proteins.** *J Proteome Res* 2009, **8**:5093–5102.
34. Navazesh M, Mulligan RA, Kipnis V, Denny PA, Denny PC: **Comparison of whole saliva flow rates and mucin concentrations in healthy Caucasian young and aged adults.** *J Dent Res* 1992, **71**:1275–1278.
35. Ben-Aryeh H, Shalev A, Szargel R, Laor A, Laufer D, Gutman D: **The salivary flow rate and composition of whole and parotid resting and stimulated saliva in young and old healthy subjects.** *Biochem Med Metab Biol* 1986, **36**:260–265.
36. Yousef GM, Scorilas A, Diamandis EP: **Genomic organization, mapping, tissue expression, and hormonal regulation of trypsin-like serine protease (TLSP PRSS20), a new member of the human kallikrein gene family.** *Genomics* 2000, **63**:88–96.
37. Borrego CA, Diamandis EP: **The emerging roles of human tissue kallikreins in cancer.** *Nat Rev Cancer* 2004, **4**:876–890.
38. Kladney RD, Bulla GA, Guo L, Mason AL, Tollefson AE, Simon DJ, Koutoubi Z, Fimmel CJ: **GP73, a novel Golgi-localized protein upregulated by viral infection.** *Gene* 2000, **249**:53–65.
39. Tamatani M, Matsuyama T, Yamaguchi A, Mitsuda N, Tsukamoto Y, Taniguchi M, Che YH, Ozawa K, Hori O, Nishimura H: **ORP150 protects against hypoxia/ischemia-induced neuronal death.** *Nat Med* 2001, **7**:317–323.
40. Ozawa K, Kuwabara K, Tamatani M, Takatsuji K, Tsukamoto Y, Kaneda S, Yanagi H, Stern DM, Eguchi Y, Tsujimoto Y: **150-kDa oxygen-regulated protein (ORP150) suppresses hypoxia-induced apoptotic cell death.** *J Biol Chem* 1999, **274**:6397–6404.
41. Simantov R, Febbraio M, Crombie R, Asch AS, Nachman RL, Silverstein RL: **Histidine-rich glycoprotein inhibits the antiangiogenic effect of thrombospondin-1.** *J Clin Invest* 2001, **107**:45–52.
42. Yasui W, Oue N, Aung PP, Matsumura S, Shutoh M, Nakayama H: **Molecular-pathological prognostic factors of gastric cancer: a review.** *Gastric Cancer* 2005, **8**:86–94.
43. Shinozaki S, Nakamura T, Iimura M, Kato Y, Iizuka B, Kobayashi M, Hayashi N: **Upregulation of Reg 1 α and GW112 in the epithelium of inflamed colonic mucosa.** *Gut* 2001, **48**:623–629.
44. Mannick E, Schurr J, Zapata A, Lentz J, Gastanaduy M, Cote R, Delgado A, Correa P, Correa H: **Gene expression in gastric biopsies from patients infected with *Helicobacter pylori*.** *Scand J Gastroenterol* 2004, **39**:1192–1200.
45. Bingle CD, Craven CJ: **PLUNC: a novel family of candidate host defence proteins expressed in the upper airways and nasopharynx.** *Hum Mol Genet* 2002, **11**:937–943.
46. Kelley SL, Lukk T, Nair SK, Tapping RI: **The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket.** *J Immunol* 2013, **190**:1304–1311.
47. Aulak K, Davis A, Donaldson V, Harrison R: **Chymotrypsin inhibitory activity of normal C1-inhibitor and a P1 Arg to His mutant: Evidence for the presence of overlapping reactive centers.** *Protein Sci* 1993, **2**:727–732.
48. Armstrong PB, Quigley JP: **α -2-macroglobulin: an evolutionarily conserved arm of the innate immune system.** *Dev Comp Immunol* 1999, **23**:375–390.
49. Reddy MS, Levine MJ, Paranchych W: **Low-molecular-mass human salivary mucin, MG2: structure and binding of *Pseudomonas aeruginosa*.** *Crit Rev Oral Biol M* 1993, **4**:315–323.
50. Sun S, Zhou J-Y, Yang W, Zhang H: **Inhibition of protein carbamylation in urea solution using ammonium-containing buffers.** *Anal Biochem* 2014, **446**:76–81.
51. Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H: **Solid-phase extraction of N-linked glycopeptides.** *Nat Protocols* 2007, **2**:334–339.
52. Sun S, Zhang B, Aiyetan P, Zhou J-Y, Shah P, Yang W, Levine DA, Zhang Z, Chan DW, Zhang H: **Analysis of N-glycoproteins using genomic N-glycosite prediction.** *J Proteome Res* 2013, **12**:5609–5615.
53. Palmisano G, Melo-Braga MN, Engholm-Keller K, Parker BL, Larsen MR: **Chemical deamidation: a common pitfall in large-scale N-linked glycoproteomic mass spectrometry-based analyses.** *J Proteome Res* 2012, **11**:1949–1957.
54. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21**:3674–3676.
55. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J: **WEGO: a web tool for plotting GO annotations.** *Nucleic Acids Res* 2006, **34**:W293–W297.
56. Huang DW, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protocols* 2008, **4**:44–57.

doi:10.1186/1559-0275-11-25

Cite this article as: Sun et al.: Analysis of age and gender associated N-glycoproteome in human whole saliva. *Clinical Proteomics* 2014 **11**:25.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

