Published in final edited form as:

J Proteome Res. 2012 November 2; 11(11): 5492–5502. doi:10.1021/pr3007254.

Proteomic and Bioinformatic Profile of Primary Human Oral Epithelial Cells

Santosh K. Ghosh^{1,*}, Elizabeth Yohannes², Gurkan Bebek², Aaron Weinberg¹, Bin Jiang¹, Belinda Willard⁵, Mark R. Chance^{2,4}, Michael T. Kinter⁶, and Thomas S. McCormick^{3,*}

¹Department of Biological Sciences, Case Western Reserve University, Cleveland, OH 44106

²Center for Proteomics and Bioinformatics, Case Western Reserve University, Cleveland, OH 44106

³Department of Dermatology, Case Western Reserve University, Cleveland, OH 44106

⁴Department of Physiology & Biophysics, Case Western Reserve University, Cleveland, OH 44106

⁵Departments of Research Core Services, Cleveland Clinic, Cleveland, OH 44195

⁶Free Radical Biology and Aging Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104

Abstract

Wounding of the oral mucosa occurs frequently in a highly septic environment. Remarkably, these wounds heal quickly and the oral cavity, for the most part, remains healthy. Deciphering the normal human oral epithelial cell (NHOEC) proteome is critical for understanding the mechanism(s) of protection elicited when the mucosal barrier is intact, as well as when it is breached. Combining 2D gel electrophoresis with shotgun proteomics resulted in identification of 1662 NHOEC proteins. Proteome annotations were performed based on protein classes, molecular functions, disease association and membership in canonical and metabolic signaling pathways. Comparing the NHOEC proteome with a database of innate immunity-relevant interactions (InnateDB) identified 64 common proteins associated with innate immunity. Comparison with published salivary proteomes revealed that 738/1662 NHOEC proteins were common, suggesting that significant numbers of salivary proteins are of epithelial origin. Gene ontology analysis showed similarities in the distributions of NHOEC and saliva proteomes with regard to biological processes, and molecular functions. We also assessed the inter-individual variability of the NHOEC proteome and observed it to be comparable with other primary cells. The baseline proteome described in this study should serve as a resource for proteome studies of the oral mucosa, especially in relation to disease processes.

*Corresponding author: (T.S.M.) Phone: 216-368-0238, Fax: 216-368-0212, tsm4@case.edu; (S.K.G) Phone: 216-368-6586, Fax: 216.368.0145, skg12@case.edu.

Supporting Information Available: Figure S1: Gel band image of 1D SDS-PAGE showing the fractionations that were cut and run on mass spectrometry for proteins identification Figure S2: [A] pI distribution and [B] Cellular locations (as determined by Ingenuity Knowledge base, Ingenuity ® Systems, Redwood City, CA) of proteins identified by LC MS/MS Figure S3: Comparison of NHOEC proteome with proteins associated with innate immunity: Venn diagram comparing the NHOEC proteome versus innate immunity related proteins from the innateDB database (<http://www.innatedb.ca/>). Sixty-four (64) common proteins were further analyzed. The proteins associated with immune and inflammatory responses and also with ubiquitination and antigen presentation pathways are marked[Fx= Functions, CP= Canonical pathways]. Table S1: List of identified proteins in NHOEC [Proteins common to saliva are shaded with grey]. Table S2: Panther protein family/subfamily, GO molecular functions, GO biological processes and cellular components of identified proteins in NHOEC. Table S3: Functional annotations of Infectious disease related proteins of NHOEC.

Keywords

Primary human oral epithelial cells; proteomics; inter-individual variability; innate immunity; saliva

INTRODUCTION

The human oral cavity is confronted by constant challenges that range from mechanical trauma to potential microbial colonization from more than 700 species of bacteria, fungi and viruses¹. The resilient nature of the oral epithelium is indeed impressive and is only compromised when chronic disturbances to the oral mucosal barrier ensue, often precipitated by poor oral hygiene leading to severe periodontal disease, or when the host becomes immuno compromised. Normal human oral epithelial cells (NHOECs), which line the oral cavity, present a dynamic barrier with a repertoire of innate defenses, some of which we and others have previously described^{2–10}.

As the first line of defense to microbial challenges, NHOECs express antimicrobial peptides (AMPs), chemokines and cytokines that act to curtail or destroy potential invading organisms. The repertoire of epithelial defense consists of numerous proteins that can either; 1) exacerbate an inflammatory response (cytokines); 2) recruit professional phagocytic cells (chemokines) or; 3) destroy microbes directly without inducing a secondary immune response (certain AMPs).

In order to further understand the dynamics of protein change following epithelial cell challenge, it is necessary to understand the nature of the host proteins; i.e., their associated function and signaling pathways within oral epithelia. Studies using primary NHOECs are limited; however, emerging work suggests that NHOECs may stimulate the host innate immune response^{11,12}. Although each protein response can be considered individually and modeled by recapitulation of the *in vivo* conditions with *in vitro* challenge, recently, there has been growing interest in exploring the entire protein repertoire (proteome) using advanced proteomics technology^{13,14}.

Numerous proteomic studies have used *in vitro* cell culture models of primary and immortalized cell lines. Although the latter are more readily accessible and do not exhibit inherent inter-donor variability, they present disadvantages including the potential to immortalize phenotypic characteristics that are not representative of the actual protein profiles observed in primary tissue. For this reason, primary cell cultures are considered to be more representative of “normal” cellular function and in studies related to human diseases, primary cells are the preferred option as they provide responses considered to more closely mimic *in vivo* conditions.

The goal of the present study was to employ proteomic approaches to identify and characterize the proteome of NHOECs. We identified and characterized the oral epithelial cell proteome using both 2-dimensional electrophoresis/mass spectrometry (2DE/MS) and shotgun proteomic techniques. We also compared the proteome of NHOECs with previously published salivary proteomes. Because primary epithelial cells are expected to exhibit inter-individual variability at the proteome level we also assessed the inter-personal variation associated with NHOECs.

MATERIALS AND METHODS

Ethics Statement

Volunteers have been used as the source of material for the described work outlined in this manuscript. Healthy human gingival tissue overlying impacted third molars was collected after written informed consent was provided by study participants and/or their legal guardians. University Hospitals Case Medical Center Institutional Review Board (IRB) Protocol #19981017 approved this study.

Human subjects

Study Volunteers recruited for the study were non-smokers. No diagnosis of gingivitis or periodontitis; i.e., inflammation of the gingival tissue or alveolar bone loss, was observed at the biopsy sites and hence the subjects were considered “healthy”. To examine the inter-individual subject variability, three age and sex matched subjects were chosen. The ages of the three subjects (A, B and C) were 17, 17 and 16 respectively.

Cells and culture conditions

Healthy oral tissue overlying impacted third molars of normal adults were extracted and used to isolate HOECs, as described previously^{2, 3}. Cells were cultured in EpiLife growth medium (Cascade Biologics, Portland, OR) and maintained at 37 °C in 5% CO₂. Epithelial monolayers were grown to 80–90% confluence in serum-free EpiLife media (Cascade Biologics Inc, Portland, OR)

2D/MS approach

(a) Sample preparation—Epithelial cell monolayers were harvested using trypsin, pelleted by centrifugation, treated with protease inhibitor cocktail (Roche, IN) and subsequently re-suspended in Tris buffered saline (TBS). Cell pellets were stored at –20°C until further processing. For the initial protein isolation, the cell pellets were dissociated by hypotonic shock by adding water (1ml) followed by 5 freeze/thaw cycles over approximately 8h for initial protein extraction. The cell lysate was separated by centrifugation at 10,000g in a table-top micro centrifuge, the supernatant was then carefully collected and total protein content was measured (BioRad, Hercules, CA).

(b) 2DE—For the analytical gels, an aliquot of the homogenate containing 100µg of protein was precipitated by addition to 800µL ice-cold acetone followed by centrifugation (13,000g, 15 min). The isolated protein pellet was re-suspended in 100µL 1% SDS and re-precipitated with acetone to completely desalt the sample. The protein pellet was then dissolved in 300µL IPG (Immobilized pH Gradient) buffer containing 6M urea, 2M thiourea, 1% CHAPS, 1% triton X100, 50mM DTT, 1% ampholytes, and 1% bromophenol blue. Seventeen cm IPG strips were actively rehydrated at 50V overnight and a three step (rapid ramp to 250V in 15min; rapid ramp to 8000V in 30 min; at 8000V for 4.5h) isoelectric focusing program was used to separate the proteins. The current was limited to 50µA per strip and the program typically gave approximately 35kVh. All isoelectric focusing steps were carried out using a Protean IEF cell (Bio-Rad, Hercules, CA). Immediately after the isoelectric focusing, the strips were equilibrated in two 15min steps in 6M urea, 50mM tris pH-8.8, 1% SDS and 30% glycerol with either 50mM DTT (first step) or 75mM iodoacetamide (second step) in 2mL volumes. The strips were then transferred to a 12.5% SDS-Page gel (BioRad Criterion) and proteins were separated at 200V for 1h. The gels were removed from the cassettes, fixed in 50% ethanol/10% acetic acid for at least 1h, and stained with SyproRuby (Invitrogen, Carlsbad, California) at room temperature, overnight.

(c) Protein identification—Each protein gel spot was cut from the gel using a 2-mm diameter punch. The excised gel plugs were washed in 50% ethanol/5% acetic acid for 4h, dehydrated in acetonitrile, and dried in a Speedvac (Labconco Corporation, Kansas City, MO). An in-gel digestion was then carried out by adding 5µL of 10ng/µL trypsin in 10mM ammonium bicarbonate and incubating the sample at room temperature, overnight. Peptides were extracted from the gel matrix in 60µL of 50% acetonitrile/5% formic acid, then evaporated to <5µL in a SpeedVac and reconstituted to 25µL with 1% acetic acid for analysis. Tandem mass spectra for all the digests were acquired using a Thermo Fisher LTQ ion trap mass spectrometer (Thermo Scientific, San Jose, CA) interfaced with a Surveyor HPLC system (Thermo Scientific, San Jose, CA). For the analysis, 2µL aliquots of the digest were manually injected onto a 10cm long reverse phase capillary column (Phenomenex Jupiter C18 material packed into a 75 µm id New Objective Picofrit tip). The column was eluted with a linear gradient of acetonitrile in 50mM acetic acid (2% to 70% in 45min) at a flow rate of 200nL/min. The column eluate was sprayed directly into the mass spectrometer using a nanoelectrospray source. The mass spectrometer was operated in the data-dependent mode acquiring a mass spectrum and 5 CID spectra in an approximately 5 sec cycle.

The search program Mascot version 2.0 (Matrix Sciences, London, UK) used all CID spectra recorded in each analysis to search the human Reference Sequence database (NCBI). Key search parameters were: enzyme trypsin, monoisotopic mass, peptide mass tolerance 3Da, MS/MS tolerance 1.5Da, 2 missed cleavages, fixed modifications- carbamidomethylcysteine, variable modifications – oxidized methionine; peptide charge 1+, 2+, and 3+. All primary identifications were based on at least 5 matched peptides with ion scores >25 and gave standard Mascot scores >300. Secondary identification was accepted if at least 2 matching peptides were detected with ion scores >25. All identifications, including any secondary identification, were verified by manual inspection and reconciliation of at least 2 matching CID spectra.

Shotgun proteomic approach for protein identification

(a) Unfractionated sample preparation—Oral epithelial cell pellets from three independent normal subjects were lysed with RIPA buffer (Thermo Scientific, Rockford, IL). After the supernatant was collected and cleaned by precipitation using a 2D-clean up kit (GE Healthcare), protein precipitant was isolated by centrifugation (10,000g, 10 min). The isolated protein pellet was then re-suspended in 8M urea, 50mM ammonium bicarbonate. Once the protein concentration was determined, 30 µg of protein was diluted into 2M urea with 50 mM ammonium bicarbonate reduced, alkylated and digested with modified trypsin (Promega) overnight. The digests were dried by SpeedVac and stored at –80°C until used for the downstream LC-MS/MS analysis.

(b) SDS PAGE fractionation—Epithelial cell pellets from an additional three normal subjects were lysed with a sample buffer containing 4% sodium dodecyl sulfate (SDS), 30 µg of protein lysate reduced, alkylated, and fractionated into sub-molecular weight using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The gel bands, as shown in Supplementary Figure S1, were cut into eight fractions and each fraction was de-stained, reduced, alkylated, and in gel digested with modified trypsin (Promega). The digests were dried by SpeedVac and stored at –80°C until used for the downstream LC-MS/MS analysis.

(c) Reverse Phase LC-MS/MS analysis—LC-MS/MS data for approximately 600 nanograms of each sample, either unfractionated or fractionated, was acquired on an LTQ-FT mass spectrometer (Thermo Electron, San Jose, CA) equipped with Dionex Ultimate

3000 LC system (Dionex, Sunnyvale, CA). Peptides were desalted in a trap column (300 μm $\times$ 5 mm, packed with C18 PepMap100, 5 μm , 100 $\text{\AA}$ (Dionex, Sunnyvale, CA)) and subsequently resolved in a reversed phase column (75 μm $\times$ 15 cm nano column, packed with C18 PepMap100, 3 μm , 100 $\text{\AA}$ (Dionex, Sunnyvale, CA) using a gradient of 7 to 40% mobile phase B (0.1% formic acid and acetonitrile (ACN)). Liquid chromatography was carried out at ambient temperature at a flow rate of 300 nl/min using a gradient mixture of 0.1% formic acid in water (solvent A) and 0.1% formic acid in 85% ACN (solvent B). Peptides eluting from the capillary tip were introduced into the LTQ source in nano-electrospray mode with a capillary voltage of 2.4 kV. A full scan was obtained for eluted peptides in the range of 315–1800 amu followed by nine data-dependent MS/MS scans. MS/MS spectra were generated by collision-induced dissociation of the peptide ions at normalized collision energy of 35% to generate a series of b- and y-ions as major fragments.

(d) Peak annotation and Database Search—For protein identification in the Elucidator system, peak list files (.mgf) were created for tandem MS/MS spectra using Deamonextract_msn algorithm. Prior to carrying out search, IPI Human database of 86392 sequences (release 04/02/2010) was downloaded in FASTA format via file transfer protocol (FTP) from the website of UniProt into the local server (<http://promas/mascot>). The resulting peak list (.dta) files were then used to interrogate sequences (86392 entries) present in IPI human by running Mascot algorithm of Mascot version 2.2.1 (Matrix Science, London, UK, version 2.2.0). Mascot searches were performed with maximum peptide and fragment ion mass tolerance of 10 ppm and 0.8 Da respectively, with variable methionine oxidation, and cysteine carbamidomethylation, and 1 missed cleavage site were also allowed in the search parameters. The search results are then imported back into Elucidator to run validation on the search results using peptide/protein Teller algorithm. The search result data was then filtered at 0.8 minimum probability and predicted error of 0.02; with this stringent criteria 1640 proteins were accepted as correct identification.

Proteome functional classification analyses

To classify proteins by gene ontology (GO) molecular function processes, cellular component, and pathway terms both Panther (www.pantherdb.org) and the Ingenuity Pathway Analysis Tool (IPA Tool; Ingenuity® Systems, Redwood City, CA, USA) were used. The complete dataset containing protein identifiers was uploaded into the applications. Panther uses binomial statistics whereas IPA utilizes Fischer's exact test to calculate a p-value that determines the probability that each biological function assigned to the data set is due to chance alone.

Inter-individual proteome variability analysis

The 2D gel images of proteins isolated from three different age and sex matched subjects were recorded using a BioRadFx scanning densitometer and preset SyproRuby conditions at 50 μm resolution. The image files were exported as TIFF files for comparative analysis using REDFIN (Ludesi, Malmo, Sweden). Protein spots that localized to the same position were matched and the Coefficient of Variation (as %CV) of intensities of each of the similar spots across three gels was calculated.

RESULTS AND DISCUSSION

The primary aim of this study was to utilize proteomic approaches to identify and characterize the proteome of normal human oral epithelial cell (NHOEC) derived proteins. NHOECs were grown from explants of human gingiva overlying either the maxillary or mandibular tuberosity behind the second molar and expanded in culture for approximately 2 weeks. We obtained fibroblast-free cultures through two steps, first, separation of the tissue

with dispase and second, is the use of selective Keratinocyte Growth Media (KGM) as the culture media. Dispase has been previously shown to specifically cleave the basement membrane zone via cleavage of fibronectin and type IV collagen, but not laminin¹⁵. The selective function of dispase leads to a clean separation of the surface epithelium from the underlying fibrous connective tissue¹⁵. The expansion of cells in KGM with low calcium concentrations (0.15 mM) further ensures the selective growth of epithelial cells. Indeed, in our study, using expanded epithelial cells which have been grown in culture for two weeks, we did not find any collagen or fibroblast-specific protein 1 (FSP1), rather we found a series of keratins indicating that expanded epithelial cells used in the present study was purely of epithelial origin

Protein identification by 2DE/MS and proteomic coverage

In total, 2DE followed by fluorescent SYPRO Ruby staining resolved more than 500 protein spots. Protein spots which were most abundant (in terms of spot intensity) across the gels, were excised, washed to remove the stain, and the gel slices were subjected to LC-MS/MS analysis. Protein identification was carried out using LC-tandem MS peptide sequencing with at least 2 peptides detected and sequenced for each primary identification. Using this strategy, we identified 237 abundant proteins. Elimination of redundant identifications yielded 181 proteins with unique gene identifiers (GI).

The proteins identified in our 2D experiments were mapped in relation to their pI. The distribution of pI values over a range of 4–9 is shown in Supplementary Figure S2. Whole proteome pI values have been previously demonstrated to correlate with subcellular localization¹⁶. In eukaryotic cells, cytoplasmic proteins exhibit a distinct clustering around pI 5–6; integral membrane proteins cluster primarily around pI 8.5–9.0 and nuclear proteins are generally evenly distributed throughout the pI range¹⁶. Thus, the distribution of the predicted pI values of oral epithelial cells indicates that the vast majority of our identified proteins are cytosolic (cluster at pI 6) and a small number of proteins are membrane associated (pI range 8.0–9.0) [Figure S2A]. Indeed, 75% of the identified proteins that exhibited pI values in the range associated with a cytosolic profile were also identified as cytosolic using the Ingenuity knowledge base [Figure S2B].

Keratin proteins are often identified as contaminants in tryptic digests due to handling of the samples before and after gel electrophoresis. This can make specific identification of keratin proteins difficult. However, keratin contamination usually involves the identification of peptides from various keratin proteins with low abundance. In this study, several 2D gel bands contain keratin peptides that map to distinct keratin proteins with an abundance that is greater than normally observed in 2D gel bands that have keratin contamination. Identified keratin proteins included Keratin 2 (Mw 65), Keratin 6A (Mw 60), Keratin 9 (Mw 62), Keratin 14 (Mw 51), Keratin 16 (Mw 51), Keratin 17 (Mw 48) and Keratin 19 (Mw 44). Amongst the keratin proteins identified in this 2DE/MS study, keratins 14 and 19 are known to be basal cell markers^{17,18} and K16 is known to be present in the spinous layer¹⁸.

Protein identification by shotgun proteomics

After establishing and validating the 2DE techniques, we sought to increase our coverage of the oral epithelial cell proteome. To that end shotgun proteomic analysis was performed on un-fractionated and SDS-PAGE fractionated (8 fractions, supplementary Figure S1) lysates from primary NHOECs of three additional individuals for each type of analysis. A combined approach of 2DE/MS and shotgun to identify total proteins from saliva has previously been published by Hu et al¹⁹. Using these techniques we have identified 365 and 1553 proteins, respectively; 278 proteins were common to both techniques. Shotgun proteomics using SDS-PAGE fractionated samples gave better coverage than shotgun using unfractionated

lysates. To create a comprehensive catalog of the Normal Human Oral Epithelial Cell (NHOEC) proteome, proteins identified by all three methods were combined, resulting in a total of 1662 uniquely identified proteins [Figure 1, A]. The list of these 1662 proteins is tabulated in supplementary table S1. Patel et al²⁰ previously identified 154 proteins from normal oral epithelial cells microdissected from formalin-fixed paraffin-embedded (FFPE) tissue samples. In FFPE samples, during the fixation process, proteins undergo degradation and cross-linking, making conventional protein analysis problematic. Moreover, the analysis of the FFPE proteome may be hampered by the chemical content of the extraction buffers needed to effectively extract proteins (e.g., detergents, reductants, denaturants, and salts)²¹. When compared to the NHOEC proteome generated using expanded primary epithelial cells, only 11% (17 out of 154) of FFPE sample proteins were identified in our samples [Figure 1, B].

Proteome annotation based on functions

The NHOEC proteins (1662) identified in this study were first sorted into protein types based on annotations in the PANTHER (Protein ANalysis THrough Evolutionary Relationships) protein classification database [<http://www.pantherdb.org/>]. The pie chart in Figure 2 summarizes the result. The three most abundant protein classes are nucleic acid binding proteins (20.4%), transferases (11%), cytoskeletal proteins (7.6%) and enzyme modulators (7.2%). In addition, we also identified hydrolases, isomerases, phosphatases, proteases, transcription factors and chaperons, indicating diversity in protein types within the oral epithelial cell proteome. Protein family/subfamily, GO molecular functions and GO biological processes and cellular component (as assessed by PANTHER) of the identified NHOEC proteins are shown in supplementary table S2.

To obtain a functional overview of the total identified proteins, we utilized Ingenuity (Ingenuity Systems) to retrieve known functions of each identified protein. Figure 3, A shows the top 20 most significant functions. The major significant categories include, among others, post-translational modification, protein synthesis, protein folding, as well as cellular death, growth and proliferation. We found that 17% (239 out of 1662) of the identified proteins are associated with programmed cell death (apoptosis), a process causing the elimination of specific cells without disturbing surrounding tissue structure or function^{22,23}. Indeed, in the normal oral mucosal surface, there is rapid renewal of epithelial cells, a mechanism hypothesized to facilitate exfoliation and clearance of infected cells^{24–26}. In the oral cavity, cell death plays a vital role in a wide variety of biological phenomena, including development, differentiation, remodeling of inflamed tissue, homeostasis, and regulation of inflammatory responses^{27,28}. Infected gingival epithelial cells attempt to eliminate challenging micro-organisms by inducing cell death to protect neighboring cells from infection²⁹. Emerging evidence supports the concept that bacterially-modulated apoptosis is a relevant process in the pathogenesis of periodontal diseases^{28, 30–32}. Since the oral epithelium is constantly challenged by mixed bacterial flora, it is hypothesized that a high rate of epithelial cell turnover enables greater bacterial clearance and thus limits the invasion of the gingival tissues^{28, 31, 33}.

In order to further analyze the proteome of NHOECs, we determined if our identified proteins were associated with specific diseases [Figure 3, B]. Interestingly, identified proteins within the oral epithelial cell proteome were found to be most significantly associated with dermatological disease and conditions followed by genetic disorder, infectious diseases, cancer, gastrointestinal, hematological, immunological and respiratory diseases. Cancer turned out to be the most highly associated disease within the NHOEC proteome with the most reported proteins identified from our analysis. This may be biased by far more cancer annotations than annotations for any other diseases. Nevertheless, in our

analysis, we found that more than one third (592 of 1662) of the identified oral epithelial cell proteins are associated with cancer and not surprisingly, more than 90% of oral cancer is squamous epithelial cell in origin (<http://oralcancerfoundation.org/>).

We also found 311 proteins within the oral epithelial cell proteome to be associated with infectious disease [Figure 3, B]. The majority of these proteins are associated with HIV by Ingenuity pathway analysis [Supplementary table S3]. Stratified squamous epithelium represents the most abundant cell type exposed to HIV-1 during sexual contact and breast feeding^{34, 35}. Using Ingenuity, we found 168 proteins within NHOECs to be associated with HIV infection [Supplementary table S3]. This is in agreement with recent studies by Vacharaksaet al³⁶ showing that oral keratinocytes support non-replicative infection and transfer of harbored HIV-1 to permissive cells. These proteins are of interest in examining HIV-associated oral complications, as epidemiological reports indicate an increase in oral lesions in chronically infected HIV subjects on highly active antiretroviral therapy (HAART)³⁷. Indeed, we recently reported a predominance of down-regulated proteins which are necessary to attenuate cellular stress, to induce pro-and anti-inflammatory responses, and proteins which regulate ROS in HIV-infected compared to HIV-control subjects³⁸. Interestingly, we identified 6 (CRYAB, PGM1, PSMB6, RNH1, TUBB4 and XRCC5) proteins out of the 168 HIV infection-associated NHOEC proteins were common to those previously reported to be down regulated in our study of the proteome of HIV on-HAART epithelia³⁸.

Canonical Pathway Analyses

Mining a proteome currently requires time and effort in seeking all the available information, which is further hampered by a wide heterogeneity of vocabulary used for protein name and function. We sought to categorize protein identifications using a pathway enrichment analysis based on a database that describes signaling pathways and network relationships (Ingenuity Systems). The identified proteins were analyzed for their membership in collated canonical and metabolic signaling pathways.

One hundred and ten (110) different pathways were linked to the NHOEC proteome. Thirty-five (35) pathways were observed to be significant as assessed by Ingenuity. Out of these 35 pathways, 17 pathways are shown in Figure 4 with greater than 25 associated proteins of NHOEC. The highest number of proteins within the NHOEC proteome (>80) is associated with eukaryotic initiation factor-2 (eIF2) signaling. In response to different environmental stresses, phosphorylation of eIF2 rapidly reduces protein synthesis, which lowers energy expenditure and facilitates reprogramming of gene expression to remediate stress damage³⁹. Thus, an NHOEC proteome enriched with the eIF2 signaling cascade might play a major role in managing environmental stress/damage within the oral cavity.

Our canonical pathway analysis also indicates that a significant number of identified proteins are associated with actin cytoskeleton (49 proteins) and integrin signaling (46 proteins). Extensive bacterial host invasion studies have revealed that the internalization of invasive bacteria into epithelial tissues involves a concerted interplay of many bacterial and epithelial cell proteins^{40–42}. Yilmaz et al⁴³ showed that the oral pathogen *Porphyromonas gingivalis* induces formation of integrin-associated focal adhesions with subsequent remodeling of the actin and tubulin cytoskeleton in NHOECs. Infection of a human immortalized gingival keratinocyte (HIGK) cell line with differences *Streptococcus gordonii*, *Fusobacterium nucleatum*, and *Aggregatibacter actinomycetemcomitans*, respectively, has also been shown to cause significant changes in cytoskeletal signaling¹². ERK/MAPK and PI3K/AKT signaling pathways are also observed to involve >25 NHOEC identified proteins. These key signaling pathways may be of great importance for epithelial cell defense and adaptability to extracellular challenges such as microbes, wounding or disease. Indeed,

MAPK signaling is key to various responses to infection and in the maintenance of oral health¹². Recently, it was reported that, in oral epithelial cells, HSV-1-induced PI3K/Akt activation is involved in the regulation of apoptosis and viral gene expression⁴⁴. Numerous proteins within the NHOEC proteome are involved in 14-3-3 protein mediated signaling. Recently, the 14-3-3 proteins were shown to be integrators of diverse signaling cues that impact cell fate and oncogenesis⁴⁵. 14-3-3 ζ (KCIP-1, protein kinase C inhibitor protein-1) has been associated with increased cell signaling involved in inflammation, cell proliferation and abrogation of apoptosis during oral carcinogenesis⁴⁶. The NRF2 (Nuclear factor-erythroid 2 p45-related factor 2) signaling pathway also utilizes numerous NHOEC proteins. NRF2 is a transcription factor that counteracts the damage caused by electrophiles and reactive oxygen species (ROS) and likely plays a protective role within the oral epithelium. Tissues of the oral cavity are the first to encounter xenobiotics and oxidative stress and we have found 28 proteins within the NHOEC proteome database to be associated with xenobiotic metabolism [Figure 4].

Most interestingly, 4.3% (73 of 1662) of total identified NHOEC proteins are associated with the ubiquitination pathway, which ranked second among canonical pathways within the NHOEC proteome as determined by Ingenuity analysis. Recently Bhoj and Chen⁴⁷ demonstrated protein ubiquitination to be a key mechanism that regulates the immune response. We further compared the NHOEC proteome with proteins related to innate immunity. The innate immunity related proteins were uploaded from the InnateDB (<http://www.innatedb.ca>)⁴⁸. InnateDB is a publicly available database of the genes, proteins, experimentally-verified interactions and signaling pathways involved in the innate immune response to microbial infection. After comparing the InnateDB related proteins (825 proteins) with NHOEC proteins (1662 proteins) we identified 64 common proteins [Supplementary Figure S3] that are associated with innate immunity within the NHOEC proteome. Further analysis revealed that the proteins in common are associated with immune and inflammatory responses and also with ubiquitination and antigen presentation pathways [Supplementary Figure S3].

Comparison of the oral epithelial cell proteome with salivary proteomes

We were also interested in comparing the NHOEC proteome with that of human saliva. For this purpose we pooled salivary proteomes of eight published independent studies^{49–56}. A total of 1747 (combined into a unique dataset with no overlapping proteins) salivary proteins was compared with the NHOEC protein dataset [Figure 5, A]. Out of 1662 NHOEC proteins, 44% were common with salivary-derived proteins (shaded with grey in supplementary table S1), suggesting that a significant number of proteins of epithelial origin are released into the saliva. NHOEC and saliva proteomes were further compared based on their annotation in GO terms of biological processes [Figure 5, B] and molecular functions [Figure 5, C]. Not surprisingly, the NHOEC and salivary proteomes showed similar distributions in the GO biological processes and molecular functions. In terms of biological processes both saliva and NHOEC proteomes are enriched for metabolic processes. However, compared to the salivary proteome, the NHOEC proteome contains more proteins associated with metabolic processes. With respect to molecular function in GO terms, in both NHOEC and saliva, the largest group of the identified proteins has catalytic activity and the second largest group of proteins is associated with binding as determined by PANTHER.

Because substantial numbers of epithelial proteins were common to saliva, we further categorized those common proteins, in terms of cellular location, and compared with cellular locations of proteins unique to NHOEC (not found in saliva), all NHOEC proteins, all salivary proteins and proteins unique to saliva (not found in NHOEC). Figure 6 describes the proportions (as %) of proteins in different cellular locations (analyzed using Ingenuity

Knowledge base) for each category of proteins. As expected for body fluids, the cellular locations indicated that total and unique salivary proteins are over-represented with extracellular proteins when compared with the NHOEC proteome. In contrast, total and unique NHOEC proteins are over-represented with nuclear proteins when compared with the salivary proteome. Moreover, it appears that more of the cytoplasmic proteins compared to nuclear or extracellular proteins tend to secret/accumulate in saliva, which could result from epithelial cell death. However, specific transport pathways may also exist.

Inter-individual plasticity of primary HOEC proteome

A healthy human proteome is characterized by significant plasticity associated with maintenance of functional resources of the organism⁵⁷. Numerous studies have demonstrated that a proteomic profile is characterized by significant intra- and inter-individual variability, and quite often natural (“normal”) variability of some proteins can be comparable to changes observed in pathological processes⁵⁸. In the present study, we sought to determine inter-individual plasticity of the NHOEC proteome by 2DE analysis of proteins isolated from three separate age and sex matched healthy individuals. The 2D gel profiles for these samples are shown in Figure 7. A total of 553 protein spots that localized to the same position were matched across three gels. For each spot in a given gel, there is a corresponding band volume which can be compared across the three gels. The Coefficient of Variation (CV) of the standard abundance for the 553 protein spots among the three samples was calculated to evaluate the variation in protein expression levels among the three individuals. CV values ranged from 1 to 128% through all the spots, with mean and median CV values of 37% and 33%. When compared to the CV of either plasma proteins⁵⁹ (inter individual CV 45%) or mammalian cell lines⁶⁰ (average CV was 39–47%), the CV of NHOEC proteins appears to be within the normal range for interpersonal variation. It should be noted that within any 2D-gel electrophoresis experiment both methodological and/or technical variations may also occur. Molloy et al.⁵⁹ has defined the degree of technical variation from the process of two-dimensional electrophoresis as 20–30% coefficient of variation. On the other hand, biological variation observed experiment-to-experiment showed a broader degree of variation depending upon the sample type. For biomarker profiling, the ability to reproducibly measure the myriad of protein expression changes across numerous multivariate experiments is a particular challenge. Therefore, it is essential to have an appreciation of the scope of biological and technical variation and design the experiments in such a way that this variation is taken into consideration and minimized where possible. Given an average percent coefficient of variation due to technical or intra-individual variation is approximately 30%, the percent coefficient of variation for actual inter-individual variation will be less than what we observed in the present study. Out of 553 spots identified, 295 spots (53%) have a coefficient of variation of greater than 30; the variation of these proteins is likely due to “inter-individual variations”.

Considering the low sample sizes for coefficient of variation analyses, this was a conservative estimation regarding the plasticity of the normal human oral epithelial cell proteome. Advanced analysis with large sample sizes and identification of proteins with high CV values will be very important in future studies. Such analysis will help to discriminate changes associated with particular diseases versus natural proteome plasticity.

CONCLUDING REMARKS

The present study establishes the initial baseline proteome for NHOECs. Our data suggests that the NHOEC proteome is associated with a high number of proteins involved in the apoptosis pathway, perhaps due to the high level of cellular turnover associated with NHOECs. The identification of 64 proteins associated with innate immunity in NHOECs supports the ability of oral tissue to act as a primary active barrier to infection. The baseline

proteome established here for NHOECs is a valuable resource that can now be used to compare normal NHOECs with those isolated from diseased oral tissues; i.e., cancers, oral warts, Kaposi sarcoma etc. The comparison of the NHOEC proteome with the salivary proteome indicates that significant numbers of epithelial proteins are found in the reported saliva proteome and are likely of HOEC origin. Whether or not these proteins accumulate due to cellular shedding or necrosis of epithelial tissue remains to be determined. The current proteomic approach aims to obtain better insight into the proteomic pattern of primary oral epithelial cells commonly used for *in vitro* models. This information is highly relevant to researchers using epithelial models, to potentially infer data for the *in vivo* situation. Establishing a baseline proteome provides us with a solid foundation for comparative studies with other epithelial tissues as well as the ability to examine the dynamic response of NHOECs when challenged in *ex vivo* experiments such as infection and injury models. Based on the preliminary studies regarding inter-individual variation of primary oral epithelial cells proteome it is understandable that additional studies regarding normal plasticity of the healthy human oral epithelial cell proteome profile will extend our knowledge regarding the physiology of oral epithelial cells. Proteins characterized by high variability should not be considered as potential biomarkers. On the contrary, proteins and peptides, exhibiting insignificant dispersion in the population of healthy individuals and possessing concentration stable in time may give important information about oral health conditions during dramatic changes in their levels.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health grants R01DE016334 (A.W.) and P01DE019759 (A.W./T.S.M.). We thank Dr. J.R. Blakemore, Dr. E.K. Schneider, Dr. W.S. Blood and Dr. F. Faddoul for providing us with normal human oral tissue.

References

1. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, et al. The human oral microbiome. *Journal of bacteriology*. 2010; 192:5002–5017. [PubMed: 20656903]
2. Krisanaprakornkit S, Weinberg A, Perez CN, Dale BA. Expression of the peptide antibiotic human beta-defensin 1 in cultured gingival epithelial cells and gingival tissue. *Infect Immun*. 1998; 66:4222–4228. [PubMed: 9712771]
3. Weinberg A, Krisanaprakornkit S, Dale BA. Epithelial antimicrobial peptides: review and significance for oral applications. *Critical reviews in oral biology and medicine: an official publication of the American Association of Oral Biologists*. 1998; 9:399–414. [PubMed: 9825219]
4. Krisanaprakornkit S, Kimball JR, Weinberg A, Darveau RP, Bainbridge BW, et al. Inducible expression of human beta-defensin 2 by *Fusobacterium nucleatum* in oral epithelial cells: multiple signaling pathways and role of commensal bacteria in innate immunity and the epithelial barrier. *Infect Immun*. 2000; 68:2907–2915. [PubMed: 10768988]
5. Shugars DC, Wahl SM. The role of the oral environment in HIV-1 transmission. *J Am Dent Assoc*. 1998; 129:851–858. [PubMed: 9685760]
6. Shugars DC. Endogenous mucosal antiviral factors of the oral cavity. *The Journal of infectious diseases*. 1999; 179(Suppl3):S431–435. [PubMed: 10099113]
7. Ross KF, Herzberg MC. Calprotectin expression by gingival epithelial cells. *Infect Immun*. 2001; 69:3248–3254. [PubMed: 11292747]
8. Kapas S, Bansal A, Bhargava V, Maher R, Malli D, et al. Adrenomedullin expression in pathogen-challenged oral epithelial cells. *Peptides*. 2001; 22:1485–1489. [PubMed: 11514033]

9. Feng Z, Weinberg A. Role of bacteria in health and disease of periodontal tissues. *Periodontology* 2000. 2006; 40:50–76. [PubMed: 16398685]
10. McCormick TS, Weinberg A. Epithelial cell-derived antimicrobial peptides are multifunctional agents that bridge innate and adaptive immunity. *Periodontology* 2000. 2010; 54:195–206. [PubMed: 20712640]
11. Darveau RP. Periodontitis: a polymicrobial disruption of host homeostasis. *Nature reviews. Microbiology*. 2010; 8:481–490.
12. Handfield M, Baker HV, Lamont RJ. Beyond good and evil in the oral cavity: insights into host-microbe relationships derived from transcriptional profiling of gingival cells. *Journal of dental research*. 2008; 87:203–223. [PubMed: 18296603]
13. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422:198–207. [PubMed: 12634793]
14. Lin D, Tabb DL, Yates JR 3rd. Large-scale protein identification using mass spectrometry. *Biochim Biophys Acta*. 2003; 1646:1–10. [PubMed: 12637006]
15. Stenn KS, Link R, Moellmann G, Madri J, Kuklinska E. Dispase, a neutral protease from *Bacillus polymyxa*, is a powerful fibronectinase and type IV collagenase. *J Invest Dermatol*. 1989; 93(2): 287–90. [PubMed: 2546994]
16. Schwartz R, Ting CS, King J. Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. *Genome Res*. 2001; 11:703–709. [PubMed: 11337469]
17. Lindberg K, Rheinwald JG. Three distinct keratinocyte subtypes identified in human oral epithelium by their patterns of keratin expression in culture and in xenografts. *Differentiation; research in biological diversity*. 1990; 45:230–241.
18. Presland RB, Dale BA. Epithelial structural proteins of the skin and oral cavity: function in health and disease. *Critical reviews in oral biology and medicine: an official publication of the American Association of Oral Biologists*. 2000; 11:383–408. [PubMed: 11132762]
19. Hu S, Xie Y, Ramachandran P, Ogorzalek Loo RR, et al. 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. [PubMed: 15800970]
20. Patel V, Hood BL, Molinolo AA, Lee NH, Conrads TP, et al. Proteomic analysis of laser-captured paraffin-embedded tissues: a molecular portrait of head and neck cancer progression. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2008; 14:1002–1014. [PubMed: 18281532]
21. Azimzadeh O, Barjaktarovic Z, Aubele M, Calzada-Wack J, Sarioglu H, Atkinson MJ, et al. Formalin-fixed paraffin-embedded (FFPE) proteome analysis using gel-free and gel-based proteomics. *J Proteome Res*. 2010; 9(9):4710–20. [PubMed: 20604508]
22. Cohen JJ, Duke RC, Fadok VA, Sellins KS. Apoptosis and programmed cell death in immunity. *Annu Rev Immunol*. 1992; 10:267–93. [PubMed: 1590988]
23. Tonetti MS, Cortellini D, Lang NP. In situ detection of apoptosis at sites of chronic bacterially induced inflammation in human gingiva. *Infect Immun*. 1998; 66:5190–5195. [PubMed: 9784521]
24. Gao LY, Kwaik YA. The mechanism of killing and exiting the protozoan host *Acanthamoeba* polyphaga by *Legionella pneumophila*. *Environ Microbiol*. 2000; 2(1):79–90. [PubMed: 11243265]
25. Häcker G, Kirschnek S, Fischer SF. Apoptosis in infectious disease: how bacteria interfere with the apoptotic apparatus. *Med Microbiol Immunol*. 2006; 195(1):11–9. [PubMed: 16086183]
26. Mao S, Park Y, Hasegawa Y, Tribble GD, James CE, et al. Intrinsic apoptotic pathways of gingival epithelial cells modulated by *Porphyromonas gingivalis*. *Cell Microbiol*. 2007; 9:1997–2007. [PubMed: 17419719]
27. Haslett C, Savill JS, Whyte MK, Stern M, Dransfield I, et al. Granulocyte apoptosis and the control of inflammation. *Philos Trans R Soc Lond B BiolSci*. 1994; 345:327–333. [PubMed: 7846130]
28. Tonetti MS, Cortellini D, Lang NP. In situ detection of apoptosis at sites of chronic bacterially induced inflammation in human gingiva. *Infect Immun*. 1998; 66:5190–5195. [PubMed: 9784521]

29. Menaker RJ, Jones NL. Fascination with bacteria-triggered cell death: the significance of Fas-mediated apoptosis during bacterial infection in vivo. *Microbes Infect.* 2003; 5:1149–1158. [PubMed: 14554257]
30. Arakawa S, Nakajima T, Ishikura H, Ichinose S, Ishikawa I, et al. Novel apoptosis-inducing activity in *Bacteroid esferis* thus: a comparative study with three serotypes of *Actinobacillus actinomycetemcomitans*. *Infect Immun.* 2000; 68:4611–4615. [PubMed: 10899863]
31. Jarnbring F, Somogyi E, Dalton J, Gustafsson A, Inge B. Quantitative assessment of apoptotic and proliferative gingival keratinocytes in oral and sulcular epithelium in patients with gingivitis and periodontitis. *J Clin Periodontol.* 2002; 29:1065–1071. [PubMed: 12492905]
32. Bantel C, Eisenach JC, Duflo F, Tobin JR, Childers SR. Spinal nerve ligation increases alpha2-adrenergic receptor G-protein coupling in the spinal cord. *Brain Res.* 2005; 1038:76–82. [PubMed: 15748875]
33. Smith JL, Bayles DO. The contribution of cytolethal distending toxin to bacterial pathogenesis. *Crit Rev Microbiol.* 2006; 32:227–248. [PubMed: 17123907]
34. Baron S, Poast J, Richardson CJ, Nguyen D, Cloyd M. Oral transmission of human immunodeficiency virus by infected seminal fluid and milk: a novel mechanism. *The Journal of infectious diseases.* 2000; 181:498–504. [PubMed: 10669332]
35. Syrjanen S. PL7 Oral viral infections that could be transmitted oro-genitally. *Oral Dis.* 2006; 12:2.
36. Vacharaksa A, Asrani AC, Gebhard KH, Fasching CE, Giacaman RA, et al. Oral keratinocytes support non-replicative infection and transfer of harbored HIV-1 to permissive cells. *Retrovirology.* 2008; 5:66. [PubMed: 18637194]
37. Greenspan D, Canchola AJ, MacPhail LA, Cheikh B, Greenspan JS. Effect of highly active antiretroviral therapy on frequency of oral warts. *Lancet.* 2001; 357:1411–1412. [PubMed: 11356441]
38. Yohannes E, Ghosh SK, Jiang B, McCormick TS, Weinberg A, et al. Proteomic signatures of human oral epithelial cells in HIV-infected subjects. *PLoS One.* 2011; 6(11):e27816. [PubMed: 22114700]
39. Palam LR, Baird TD, Wek RC. Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J Biol Chem.* 2011; 286(13):10939–49. [PubMed: 21285359]
40. Knodler LA, Celli J, Finlay BB. Pathogenic trickery: deception of host cell processes. *Nat Rev Mol Cell Biol.* 2001; 2:578–588. [PubMed: 11483991]
41. Cossart P, Sansonetti PJ. Bacterial invasion: the paradigms of enteroinvasive pathogens. *Science.* 2004; 304:242–248. [PubMed: 15073367]
42. Patel JC, Galan JE. Manipulation of the host actin cytoskeleton by *Salmonella*--all in the name of entry. *Curr Opin Microbiol.* 2005; 8:10–15. [PubMed: 15694851]
43. Yilmaz O, Young PA, Lamont RJ, Kenny GE. Gingival epithelial cell signalling and cytoskeletal responses to *Porphyromonas gingivalis* invasion. *Microbiology.* 2003; 149:2417–426. [PubMed: 12949167]
44. Hsu MJ, Wu CY, Chiang HH, Lai YL, Hung SL. PI3K/Akt signaling mediated apoptosis blockage and viral gene expression in oral epithelial cells during herpes simplex virus infection. *Virus Res.* 2010; 153(1):36–43. [PubMed: 20620179]
45. Morrison DK. The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. *Trends Cell Biol.* 2009; 19(1):16–23.
46. Matta A, Bahadur S, Duggal RD, Gupta S. Over-expression of 14-3-3zeta is an early event in oral cancer. *BMC Cancer.* 2007; 7:169. [PubMed: 17764575]
47. Bhoj VG, Chen ZJ. Ubiquitylation in innate and adaptive immunity. *Nature.* 2009; 458:430–437. [PubMed: 19325622]
48. Lynn DJ, Winsor GL, Chan C, Richard N, Laird MR, et al. InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Molecular systems biology.* 2008; 4:218. [PubMed: 18766178]
49. Vitorino R, Lobo MJ, Ferrer-Correia AJ, Dubin JR, Tomer KB, et al. Identification of human whole saliva protein components using proteomics. *Proteomics.* 2004; 4:1109–1115. [PubMed: 15048992]

50. Denny P, Hagen FK, Hardt M, Liao L, Yan W, et al. The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *Journal of proteome research*. 2008; 7:1994–2006. [PubMed: 18361515]
51. Fang X, Yang L, Wang W, Song T, Lee CS, et al. 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. [PubMed: 17614365]
52. Ramachandran P, Boontheung P, Xie Y, Sondej M, Wong DT, et al. Identification of N-linked glycoproteins in human saliva by glycoprotein capture and mass spectrometry. *Journal of proteome research*. 2006; 5:1493–1503. [PubMed: 16740002]
53. Walz A, Stuhler K, Wattenberg A, Hawranke E, Meyer HE, 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. [PubMed: 16402355]
54. Wilmarth PA, Riviere MA, Rustvold DL, Lauten JD, Madden TE. Two-dimensional liquid chromatography study of the human whole saliva proteome. *Journal of proteome research*. 2004; 3:1017–1023. [PubMed: 15473691]
55. Xie H, 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. [PubMed: 16103422]
56. Guo T, Rudnick PA, Wang W, Lee CS, Devoe DL, et al. Characterization of the human salivary proteome by capillary isoelectric focusing/nanoreversed-phase liquid chromatography coupled with ESI-tandem MS. *Journal of proteome research*. 2006; 5:1469–1478. [PubMed: 16739998]
57. Nelsestuen GL, Zhang Y, Martinez MB, Key NS, Jilma B, Verneris M, Sinaiko A, Kasthuri R. Plasma protein profiling: Unique and stable features of individuals. *Proteomics*. 2005; 5:4012–4024. [PubMed: 16130168]
58. Pakharukovaa NA, Pastushkovaa L Kh, Moshkovskii SA, Larina IM. Variability of the Healthy Human Proteome. *Biochemistry (Moscow) Supplement Series B: Biomedical Chemistry*. 2011; 5(3):203–212.
59. Anderson NL, Anderson NG. A two-dimensional gel database of human plasma proteins. *Electrophoresis*. 1991; 12:883–906. [PubMed: 1794344]
60. Molloy MP, Brzezinski EE, Hang J, McDowell MT, VanBogelen RA. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics*. 2003; 3:1912–1919. [PubMed: 14625853]

Figure 1.

[A] Proteome identification via multiple approaches. The Venn diagram shows the number of proteins that are identified from multiple approaches and compares these set of proteins across the three pipelines. In total, 1662 proteins were identified. [B] Comparison of proteins identified in our study [NHOEC proteome] from expanded primary epithelial cells and proteins identified in normal oral epithelial cells from laser captured microdissected formaldehyde fixed paraffin embedded tissues by Patel et al (2008) [22].

\$watermark-text

Figure 3. [A] Top 20 most significant molecular and cellular functions and [B] the most significant disease associations of the identified proteins from oral epithelial cells. Y-axis represents the significance [$-\log(p\text{-value})$]. The significance calculated for each function returned in functional analysis is a measurement of the likelihood that the function is associated with the dataset by random chance. The numbers of associated proteins within each of the categories are also shown within each category of protein.

Figure 4.

Canonical pathway analysis: The number of proteins associated with canonical signaling pathways as defined by Ingenuity Pathway analysis is shown. Seventeen pathways with 25 NHOEC identified proteins are shown.

Figure 5. [A] Venn diagram showing the NHOEC and saliva proteome overlaps [B], [C] GO distributions of NHOEC and saliva proteome [B] biologic process, [C] molecular function.

Figure 6. Cellular distribution of NHOEC, salivary, unique NHOEC, unique salivary proteome and proteome common to NHOEC & saliva.

Figure 7. Inter-individual variations of oral epithelial cell proteome

Tissue from three individuals was harvested and used to isolate the epithelial cells as described in the methods. **Left Panel:** Black and white images of the Sypro Ruby stained 2D protein gels from three individuals (A-C). The gel images are presented with a 20% enhancement of the darkness curve (input = 80, output = 100) by the image processing program to enhance the visibility of the low abundance spots. **Right Panel:** Distribution of calculated coefficient of variation (%CV) of 553 individual spots across three gels.