Cancer and Clinical Oncology; Vol. 5, No. 1; 2016
ISSN 1927-4858 E-ISSN 1927-4866
Published by Canadian Center of Science and Education

Altered Salivary Protein Profiles among Individuals Diagnosed with Cervical Dysplasia and Cervical Carcinoma in Situ

Charles Streckfus[1]

[1] University of Texas School of Dentistry at Houston, Houston, Texas, USA

Correspondence: Charles Streckfus, University of Texas School of Dentistry at Houston, Houston, Texas, USA Tel: 1-713-486-4531. E-mail: charles.streckfus@uth.tmc.edu

Received: February 26, 2016 Accepted: March 21, 2016 Online Published: March 22, 2016

doi:10.5539/cco.v5n1p63 URL: http://dx.doi.org/10.5539/cco.v5n1p63

Abstract

Objective: The objective of this study was to determine if protein-by-products secondary to cervical cancer oncogenes appear in the saliva.

Methods: Four pooled (n=10 subjects/pool) stimulated whole saliva specimens from women were analyzed. One pooled specimen was from healthy women, while other pooled specimens were from women diagnosed with CIN 2 moderate cervical dysplasia (n=10), CIN 3 severe cervical dysplasia (n=10) and the other pooled group from women diagnosed with cervical carcinoma *in situ* (n=10). Differential expression of proteins was measured by isotopically tagging proteins in each groups and comparing them to the healthy control group. Saliva from each of the pooled samples was trypsinized and the peptide digests labeled with the appropriate iTRAQ reagent. Labeled peptides from each of the digests were combined and analyzed by reverse phase (C18) capillary chromatography on an Applied Biosystems QStar LC-MS/MS mass spectrometer equipped with an LC-Packings HPLC.

Results: The results of the salivary analyses yielded approximately 133 proteins in the saliva specimens. Forty eight proteins were differentially expressed between the healthy control pool, precancerous and cancerous conditions.

Conclusions: The study suggests that saliva is a fluid suffused with solubilized by-products of oncogenic expression and that these proteins may be modulated secondary to precancerous and cancerous conditions and that these proteins may be useful in the study of cervical cancer progression, treatment efficacy and the tailoring of individualized patient care.

Keywords: saliva, cervical cancer, cancer modeling, individualized patient care, signaling, exosomes

1. Introduction

Invasive cervical cancer is the second most common cause of death from cancer among women (Bosch et al., 1995). In the United States, it is estimated that the number of new cases of cervix uteri cancer is approximately 7.7 per 100,000 women per year. The number of deaths was 2.3 per 100,000 women per year. These rates are age-adjusted and based on 2008-2012 cases and deaths as reported by the NCI Surveillance, 2014, Epidemiology and End Results Program.

Cervical lesions are divided into different stages of aberrant cell growth. The initial stages are the precancerous conditions designated as cervical intraepithelial neoplasia (CIN) which exhibit alteration of the cervical epithelial lining with the presence of abnormal cell growth (Garbett et al., 2014). CIN status is determined by the extent in which the cervical epithelial lining and is divided into three stages based on the histological features, nuclear changes and the extent of epithelial involvement. The grades are mild, moderate and severe with the moderate and severe conditions having a higher likelihood to progress into carcinoma in situ (CIS) or invasive carcinoma (Garbett et al., 2014).

In order to further enhance cytological assessments, liquid based Hybrid Capture II technology for HPV has been incorporated into PAP smear assessment (Cochand-Priollet et al., 2005; Higareda-Almaraz, Enríquez-Gasca., Hernández-Ortiz, Resendis-Antonio & Encarnación-Guevara, 2011). However HR-HPV cannot differentiate between CIN lesions having a higher likelihood of progressing from those that do not resulting in many women being over treated (Higareda-Almaraz, Enríquez-Gasca., Hernández-Ortiz, Resendis-Antonio, &

Encarnación-Guevara, 2011). Consequently, additional methods are still required in order to determine which cervical lesions may progress to IC.

Collectively, it is the hypothesis of this preliminary investigation to determine if there are alterations in salivary proteins as a consequence of CIN and CIS. This information could be used as a tool to further the understanding of cervical cancer pathogenesis and clinical decision making.

2. Methods and Materials

2.1 Proteomic Design

The investigators analyzed pooled, stimulated whole saliva specimens. Each pooled specimen within a cohort consisted of ten individual patient saliva specimens from a bank of control and cancer specimens frozen at -80°C. One pooled saliva specimen consisted of saliva from 10 healthy subjects; another pooled specimen was a pooled saliva specimen from 10 subjects diagnosed with moderate cervical dysplasia, severe dysplasia and with cervical carcinoma *in situ*. In other words each pooled saliva specimen corresponded to 10 patients from each cervical lesion considered in this study.Except for the control group, the dysplasia and cancer cohorts were as determined by the pathology report. All subjects were closely matched for age and race and were non-tobacco users. All disease categories were HPV positive.

All participating volunteers were explained their participation rights and signed an IRB consent form. The saliva specimens and related patient data are non-linked and bar coded in order to protect patient confidentiality. This study was performed under the UTHSC IRB approved protocol# HSC-DB-05-0394.

2.2 Saliva Collection and Sample Preparation

Stimulated whole salivary gland secretion is a reflex response occurring during the mastication of a bolus of food. Usually, a standardized bolus (1 gram) of paraffin or a gum base (generously provided by the Wrigley Co., Peoria, IL) is given to the subject to chew at a regular rate. The individual, upon sufficient accumulation of saliva in the oral cavity, expectorates periodically into a preweighed disposable plastic cup. This procedure is continued for a period of five minutes. The volume and flow rate is then recorded along with a brief description of the specimen's physical appearance (Streckfus et al., 2008). The cup with the saliva specimen is reweighed and the flow rate determined gravimetrically. The authors recommend this salivary collection method with the following modifications for consistent protein analyses (Streckfus et al., 2008). A protease inhibitor cocktail from Sigma Co (St. Louis, MI, USA) is added along with enough dithiothreitol from a 1 M stock solution to bring its concentration 1 mM. The treated samples were centrifuged for 10 minutes at top speed in a table top centrifuge. The supernatant was divided into 1 ml aliquots and frozen at -80°C.

2.3 Bottom-up Mass Spectrometry Using iTRAQ Labeling

A thorough explanation for the top-down mass spectrometry using iTRAQ labeling can be found in detail in previous publications (Streckfus et al., 2008). Briefly, the saliva samples were thawed and immediately centrifuged to remove insoluble materials. The supernatant was assayed for protein using the Bio-Rad protein assay (Hercules, CA, USA) and an aliquot containing 100 µg of each specimen was precipitated with six volumes of -20ºC acetone. The precipitate was resuspended and treated according to the manufacturer's instructions. Protein digestion and reaction with iTRAQ labels was carried out as previously described and according to the manufacturer's instructions (Applied Biosystems, Foster City, CA). Briefly, the acetone precipitable protein was centrifuged in a tabletop centrifuge at 15,000 x g for 20 minutes. The acetone supernatant was removed and the pellet resuspended in 20μL dissolution buffer. The soluble fraction was denatured and disulfides reduced by incubation in the presence of 0.1% SDS and 5 mM TCEP (tris-(2-carboxyethyl) phosphine) at 60ºC for one hour. Cysteine residues were blocked by incubation at room temperature for 10 minutes with MMTS (methyl methane-thiosulfonate).

Afterwards, trypsin was added to the mixture to a protein: trypsin ratio of 10:1. The mixture was incubated overnight at 37ºC. The protein digests were labeled by mixing with the appropriate iTRAQ reagent and incubating at room temperature for one hour. On completion of the labeling reaction, the four separate iTRAQ reaction mixtures were combined. Since there are a number of components that can interfere with the LC-MS/MS analysis, the labeled peptides are partially purified by a combination of strong cation exchange followed by reverse phase chromatography on preparative columns. The combined peptide mixture is diluted 10 fold with loading buffer (10 mM KH_2PO_4 in 25% acetonitrile at pH 3.0) and applied by syringe to an ICAT Cartridge-Cation Exchange column (Applied Biosystems, Foster City, CA) column that has been equilibrated with the same buffer. The column is washed with 1 ml loading buffer to remove contaminants.

To improve the resolution of peptides during LC-MS/MS analysis, the peptide mixture is partially purified by elution from the cation exchange column in three fractions. Stepwise elution from the column is achieved with sequential 0.5 ml aliquots of 10 mM KH_2PO_4 at pH 3.0 in 25% acetonitrile containing 116 mM, 233 mM and 350 mM KCl respectively. The fractions are evaporated by Speed Vac to about 30% of their volume to remove the acetonitrile and then slowly applied to an Opti-Lynx Trap C18 100 μL reverse phase column (Alltech, Deerfield, IL) with a syringe. The column was washed with 1 ml of 2% acetonitrile in 0.1% formic acid and eluted in one fraction with 0.3 ml of 30% acetonitrile in 0.1% formic acid. The fractions were dried by lyophilization and resuspended in 10 μL 0.1% formic acid in 20% acetonitrile solution. Each of the three fractions was analyzed by reverse phase nano-LCMS/MS on an API QSTAR XL mass spectrometer (ABS Sciex Instruments).

2.4 Bioinformatics & Statistical Analysis

The Swiss-Prot database was employed for protein identification while the PathwayStudio® bioinformatics software package was used to determine Venn diagrams were also constructed using the Venn Diagram Plotter NIH software program (https://omics.pnl.gov/software/venn-diagram-plotter). Pathways were retrieved from three databases: DAVID (Huang et al., 2009), KEGG (Kanehisa et al., 2014), and the NCI's Protein Interaction Database (Schaefer et al., 2009). Gene ontologies were determined by employing the GO and AmiGO databases (Schaefer et al., 2009).

The accumulated MSMS spectra were analyzed by ProQuant and ProGroup software packages (Applied Biosystems) using the SwissProt fasta database for protein identification. The ProQuant analysis was carried out with a 75% confidence cutoff with a mass deviation of 0.15 Da for the precursor and 0.1 Da for the fragment ions. The ProGroup reports were generated with a 95% confidence level for protein identification. Additionally, univariate t-test were used to calculate p-values for the protein comparisons.

3. Results

The bottom-up salivary proteomic analysis revealed a total of 133 proteins 48 (36%) of which were differentially expressed in at least one of the three groups of women. The results are provided in Table 1 giving both the p-value and the ratio for each protein. In this study, the ratio represents the protein value of the disease state divided by the healthy control protein value. Among the three groups 31 (65%) proteins were up-regulated and 17 (35%) were down-regulated.

Table 1. Ratios and p-values for differentiated salivary proteins across the three cohorts of individuals

Gene ID	UniProt	Protein Name	CIN-2		CIN-3		CIS	
			Ratio	p-value	Ratio	p-value	Ratio	p-value
ADA32	Q8TC27	ADAM 32	0.5237	0.0017	0.6923	0.0163	0.7322	0.0687
ENOA	P06733	Alpha-enolase	0.8016	0.4261	0.7408	0.3288	1.4349	0.0037
ANXA1	P04083	Annexin A1	2.0351	0.0003	2.0638	0.0022	1.3091	0.0546
ALK1	P03973	Antileukoproteinase 1	2.9400	0.032	1.8682	0.0931	1.3263	0.0243
COBA1	P12107	Collagen alpha-1	0.7874	0.5355	0.6688	0.0914	0.5704	0.0054
SPRR3	Q9UBC9	Cornifin beta	0.8026	0.4132	0.7535	0.3308	1.9157	0.0001
CYTA	P01040	Cystatin A	1.161	0.0281	1.041	0.7512	1.3136	0.0551
CYTB	P04080	Cystatin B	1.0503	0.6084	0.7575	0.033	0.9593	0.6527
CYTC	P01034	Cystatin C	0.796	0.1433	0.5707	0.1233	0.7657	0.0051
CYTT	P09228	Cystatin SA	0.491	0.0001	1.1473	0.1187	0.6215	0.0001
CYTN	P01037	Cystatin SN precursor	0.5301	0.0001	1.1563	0.0294	0.6662	0.0001
K2C1	P04264	cytoskeletal 1	2.0725	0.0112	2.2194	0.0212	1.5669	0.0706
K1C13	P13646	cytoskeletal 13	6.2089	0.0001	5.9058	0.0001	1.0462	0.243
K1C14	P02533	cytoskeletal 14	5.3359	0.0001	6.5546	0.0001	1.0927	0.0858
K22O	Q01546	cytoskeletal 2	7.466	0.0009	5.3974	0.0035	1.1865	0.0599
K2C4	P19013	cytoskeletal 4	5.4646	0.0001	5.9058	0.0001	1.112	0.0078
K2C5	P13647	cytoskeletal 5	2.6533	0.0001	3.4924	0.0001	0.748	0.0015

K2C6A	P02538	cytoskeletal 6A	4.336	0.0001	5.7765	0.0001	1.0797	0.0592
DMBT1	Q9UGM3	DMBT1	1.8276	0.0027	1.3987	0.0442	1.2193	0.0154
FABPE	Q01469	Fatty acid-binding protein	0.6976	0.0841	0.5486	0.0237	1.0733	0.6338
FGRL1	Q8N441	FGFR-1	0.2128	0.0007	0.4391	0.0017	0.6446	0.0123
HPT	P00738	Haptoglobin precursor	1.0724	0.1989	1.0092	0.8633	2.6383	0.0001
HBA	P69905	Hemoglobin subunit alpha	1.1837	0.1354	1.4634	0.0072	3.7891	0.0001
IGHA1	P01876	Ig alpha-1 chain C region	1.5492	0.0001	1.2326	0.0049	0.9099	0.182
IGHA2	P01877	Ig alpha-2 chain C region	1.6536	0.0017	1.2296	0.0593	1.1415	0.221
IGHG1	P01857	Ig gamma-1 chain C region	1.4084	0.0003	1.0791	0.3557	1.4797	0.0001
IGHG2	P01859	Ig gamma-2 chain C region	1.1294	0.0597	0.9542	0.5763	1.5838	0.0037
KAC	P01834	Ig kappa chain C region	1.4171	0.0002	1.1102	0.0876	1.1216	0.0711
LAC	P01842	Ig lambda chain C regions	1.5102	0.0015	1.0102	0.926	0.9741	0.7148
LV3B	P80748	Ig lambda chain V-III region	1.3278	0.0422	1.162	0.0902	1.2645	0.0227
IGJ	P01591	Immunoglobulin J chain	1.2987	0.0543	1.000	0.9995	0.7708	0.039
TRFL	P02788	Lactotransferrin precursor	0.8186	0.472	0.6958	0.1345	0.6385	0.0148
LUZP1	Q86V48	Leucine zipper protein 1	0.3644	0.1605	0.4651	0.0512	0.6236	0.0429
LCN1	P31025	Lipocalin-1	1.1458	0.0743	0.7156	0.0009	1.3294	0.0001
MUC5B	Q9HC84	Mucin-5B	2.0361	0.0003	1.7545	0.0007	1.0232	0.8456
DEF3	P59666	Neutrophil defensin 3	3.857	0.0001	2.6127	0.0001	1.166	0.0119
PIGR	P01833	Poly.-immunoglob. Recept.	1.0415	0.7137	0.8355	0.0779	0.8764	0.0096
S10A7	P31151	S100-A7	0.7963	0.0743	0.737	0.016	0.7609	0.0257
S10A8	P05109	S100-A8	3.4601	0.0001	3.3518	0.0001	1.4209	0.0001
S10A9	P06702	S100-A9	3.8593	0.0001	3.1049	0.0001	1.3805	0.0001
AMYS	P04745	Salivary alpha-amylase	0.5847	0.0001	0.5542	0.0001	0.775	0.0001
ALBU	P02768	Serum albumin precursor	0.9514	0.1104	0.9896	0.7099	1.9365	0.0001
SPLC2	Q96DR5	SPLUNC	0.7395	0.0147	1.223	0.0431	2.1095	0.0001
THIO	P10599	Thioredoxin	1.023	0.8184	1.1162	0.1566	1.4268	0.0001
TRFE	P02787	Transferrin	1.0047	0.9582	1.0873	0.2001	1.9953	0.0001
TRY1	P07477	Trypsin-1	0.8391	0.0033	0.8841	0.0303	1.1449	0.1926
ZN248	Q8NDW4	Zinc finger protein 248	1.5336	0.1127	1.2782	0.0555	0.8428	0.0464
ZA2G	P25311	Zinc-alpha-2-glycoprotein	0.9099	0.0417	1.1264	0.0023	1.2934	0.0001
Up-regulated proteins			31		31		31	
Down-regulated proteins			17		17		17	

Note: CIN-2 = moderate cervical dysplasia; CIN-3 severe cervical dysplasia; CIS cervical carcinoma *in situ.*

Figure 1 and Table 2 illustrate the overlap of proteins among the three groups of women. Fourteen proteins (29%) were common to all three cohorts. CIN-2 and CIN-3 had only two unique significantly differentiated proteins while the CIS cohort exhibited 10 unique salivary proteins. As presented in Table 1, the CIS group expressed nine unique proteins that were significant. The proteins that were unique to CIS are ENOA, COBA1, SPRR3, CYTC, HPT, TRFL, ALBU, THIO and TRFE.

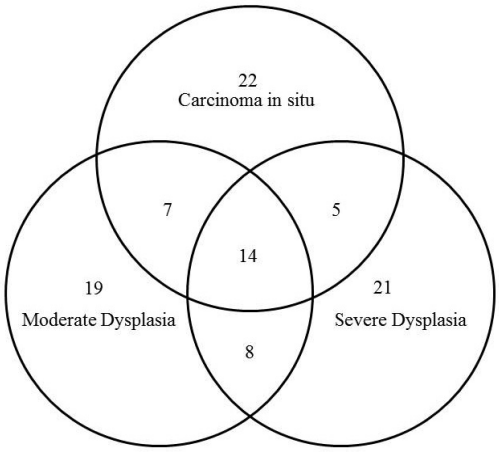

Figure 1. Overlap of salivary proteins among the three cohorts

Table 2. Overlap of proteins among the three groups of women

Status	CIN-2	CIN-3	CIS
Moderate Dysplasia	--	8	7
Severe Dysplasia	8	--	5
Cancer in Situ	7	5	--
Significant Proteins	31	30	36
No Significance	17	18	12
Unique Proteins	2	2	10

Table 3 lists the fourteen proteins that are common to all three groups with the majority of the salivary proteins (74%) being up-regulated. The information listed in the "Trends" column in Table 3 is a graphical representation of the protein ratios across CIN-2, CIN-3 and CIS. For example, protein annexin A-1 has a CIN-2 ratio of 2.0351, a CIN-3 ratio of 2.0638 and a CIS ratio of 1.3091. Across the three cohorts for Annexin A-1 the CIN-2 and CIN-3 are approximately the same whereas the CIS ratio is approximately 47% lower than the CIN-2 and CIN-3 protein ratios; therefore, the trend appears as "‾\" in that column.

Table 3. Lists the fourteen proteins that are common to all three disease groups

Gene ID	UniProt	Protein Name	CIN-2		CIN-3		CIS		Trend
			Ratio	p Value	Ratio	p Value	Ratio	p Value	
ANXA1	P04083	Annexin A1	2.0351	0.0003	2.0638	0.0022	1.3091	0.0546	
CYTN	P01037	Cystatin SN	0.5301	0.0001	1.1563	0.0294	0.6662	0.0001	
K22O	Q01546	Cytoskeletal 2	7.4660	0.0009	5.3974	0.0035	1.1865	0.0599	
K2C4	P19013	Cytoskeletal 4	5.4646	0.0001	5.9058	0.0001	1.1120	0.0078	
K2C5	P13647	Cytoskeletal 5	2.6533	0.0001	3.4924	0.0001	0.7480	0.0015	
K2C6A	P02538	Cytoskeletal 6A	4.3360	0.0001	5.7765	0.0001	1.0797	0.0592	
DMBT1	Q9UGM3	DMBT1	1.8276	0.0027	1.3987	0.0442	1.2193	0.0154	
FGRL1	Q8N441	FGFR-1	0.2128	0.0007	0.4391	0.0017	0.6446	0.0123	
DEF3	P59666	Neutrophil defen. 3	3.8570	0.0001	2.6127	0.0001	1.1660	0.0119	
S10A8	P05109	S100-A8	3.4601	0.0001	3.3518	0.0001	1.4209	0.0001	
S10A9	P06702	S100-A9	3.8593	0.0001	3.1049	0.0001	1.3805	0.0001	
AMYS	P04745	alpha-amylase	0.5847	0.0001	0.5542	0.0001	0.7750	0.0001	
SPLC2	Q96DR5	SPLUNC	0.7395	0.0147	1.2230	0.0431	2.1095	0.0001	
ZA2G	P25311	Zinc α-2-glycoprot.	0.9099	0.0417	1.1264	0.0023	1.2934	0.0001	

A heat map was generated from data reflecting protein expression of p-values across the three categories of women exhibited in Table 4. Spaces that are white are not statistically significant while the varying shades of gray represent significant p-values. The darker the shade of gray among data the more statistically significant is the p-value.

Table 4. Is a graphical representation (heat map) of the data where the individual values contained in a matrix are represented as shades of white and gray

UniProt	Protein Name	CIN-2	CIN-3	CIS
		p-value	p-value	p-value
Q8TC27	ADAM 32	0.0017	0.0163	0.0687
P06733	Alpha-enolase	0.4261	0.3288	0.0037
P04083	Annexin A1	0.0003	0.0022	0.0546
P03973	Antileukoproteinase 1	0.032	0.0931	0.0243
P12107	Collagen alpha-1	0.5355	0.0914	0.0054
Q9UBC9	Cornifin beta	0.4132	0.3308	0.0001
P01040	Cystatin A	0.0281	0.7512	0.0551
P04080	Cystatin B	0.6084	0.033	0.6527
P01034	Cystatin C	0.1433	0.1233	0.0051
P09228	Cystatin SA	0.0001	0.1187	0.0001
P01037	Cystatin SN precursor	0.0001	0.0294	0.0001
P04264	cytoskeletal 1	0.0112	0.0212	0.0706
P13646	cytoskeletal 13	0.0001	0.0001	0.2430
P02533	cytoskeletal 14	0.0001	0.0001	0.0858
Q01546	cytoskeletal 2	0.0009	0.0035	0.0599
P19013	cytoskeletal 4	0.0001	0.0001	0.0078
P13647	cytoskeletal 5	0.0001	0.0001	0.0015
P02538	cytoskeletal 6A	0.0001	0.0001	0.0592
Q9UGM3	DMBT1	0.0027	0.0442	0.0154
Q01469	Fatty acid-binding protein	0.0841	0.0237	0.6338
Q8N441	FGFR - 1	0.0007	0.0017	0.0123
P00738	Haptoglobin precursor	0.1989	0.8633	0.0001
P69905	Hemoglobin subunit alpha	0.1354	0.0072	0.0001
P01876	Ig alpha-1 chain C region	0.0001	0.0049	0.1820
P01877	Ig alpha-2 chain C region	0.0017	0.0593	0.2210
P01857	Ig gamma-1 chain C region	0.0003	0.3557	0.0001
P01859	Ig gamma-2 chain C region	0.0597	0.5763	0.0037
P01834	Ig kappa chain C region	0.0002	0.0876	0.0711
P01842	Ig lambda chain C regions	0.0015	0.926	0.7148
P80748	Ig lambda chain V-III region LOI	0.0422	0.0902	0.0227
P01591	Immunoglobulin J chain	0.0543	0.9995	0.039
P02788	Lactotransferrin precursor	0.472	0.1345	0.0148
Q86V48	Leucine zipper protein 1	0.1605	0.0512	0.0429
P31025	Lipocalin-1	0.0743	0.0009	0.0001
Q9HC84	Mucin-5B	0.0003	0.0007	0.8456
P59666	Neutrophil defensin 3	0.0001	0.0001	0.0119
P01833	Polymeric-immunoglobulin receptor	0.7137	0.0779	0.0096
P31151	S100-A7	0.0743	0.016	0.0257
P05109	S100-A8	0.0001	0.0001	0.0001
P06702	S100-A9	0.0001	0.0001	0.0001
P04745	Salivary alpha-amylase precursor	0.0001	0.0001	0.0001
P02768	Serum albumin precursor	0.1104	0.7099	0.0001
Q96DR5	SPLC2	0.0147	0.0431	0.0001
P10599	Thioredoxin	0.8184	0.1566	0.0001
P02787	Transferrin	0.9582	0.2001	0.0001
P07477	Trypsin-1	0.0033	0.0303	0.1926
Q8NDW4	Zinc finger protein 248	0.1127	0.0555	0.0464
P25311	Zinc-alpha-2-glycoprotein	0.0417	0.0023	0.0001

Note: White equals no statistical significance; light gray is slightly significant $p>0.001$; medium gray is $p>0.0001$; dark grey is highly significant $p<0.0001$.

Table 5 illustrates the basic functional categories of the proteins along with their biochemical activity respectively. Collectively, cellular function is categorized as follows: 1) metabolic proteins (44%), anti-inflammatory and immunoresponse proteins (27%), cytoskeletal proteins (19%) and a small number of miscellaneous proteins (10%). Additionally, Table 5 categorizes the proteins according to their biochemical activity. The major biochemical functions are binding, protein binding, cellular structure and catalytic activity.

Table 5. Functional analysis of the proteins associated with the varying disease states

GOID	TERM	NUMBER	SIZE	FREQUENCY
GO:0005488	binding	41	47	87.23%
GO:0005515	protein binding	36	47	76.60%
GO:0005198	structural molecule activity	11	47	23.40%
GO:0003824	catalytic activity	10	47	21.28%
GO:0030234	enzyme regulator activity	9	47	19.15%
GO:0003676	nucleic acid binding	8	47	17.02%
GO:0016787	hydrolase activity	6	47	12.77%
GO:0004871	signal transducer activity	5	47	10.64%
GO:0016209	antioxidant activity	4	47	8.51%
GO:0004872	receptor activity	4	47	8.51%
GO:0005215	transporter activity	4	47	8.51%
GO:0016491	oxidoreductase activity	2	47	4.26%
GO:0008565	protein transporter activity	1	47	2.13%
GO:0015075	ion transmembrane transporter activity	1	47	2.13%
GO:0004386	helicase activity	1	47	2.13%
GO:0016829	lyase activity	1	47	2.13%
GO:0016301	kinase activity	1	47	2.13%
GO:0016740	transferase activity	1	47	2.13%

4. Discussion

The literature did not reveal any manuscripts using saliva as a fluid for studying cervical cancer progression. As a consequence, the author will compare the results of this investigation to studies analyzing cell lines, tissues and other body fluids (Garbett et al., 2014; Thiede et al., 2013).

The results of this study support the hypothesis that salivary proteins are altered secondary to the presence of neoplastic lesions of the cervix. The entire list of proteins is too long to discuss in this manuscript so the author has selected a few examples to illustrate how the findings support the hypothesis. For example, salivary proteins such as ENOA (Capello et al., 2011), ANXA1 (Wang et al., 2008; Bae et al., 2005), K1C13 (Carrilho et al., 2004), K1C14 (Southern et al., 2001), S100A9 (Qin et al., 2010) and ZA2G (Hassan et al., 2008) to name just a few of the proteins have been reported in the cervical cancer literature and they appear altered in saliva secondary to CIS.

In Table 1 and Table 4, ENOA was up-regulated and significantly differentiated at the $p<0.004$ level which corresponds to the findings of Bae et al., 2005 and was further reported by Capello, Ferri-Borgogno, Cappello & Novelli (2011). ENOA is a glycolytic enzyme that catalyzes the conversion of 2-phosphoglycerate into phosphoenolpyruvate and is the frequently deregulated in various types of cancers (Capello et al., 2011). The overexpression of ENOA is associated with the Warburg effect which is a “hallmark” of many malignancies. Interestingly, ENOA was unchanged among the CIN-2 and CIN-3 cohorts, but was up-regulated among the CIS grouping. This suggests a demarcation between dysplastic and cancerous tissues that warrants further investigation.

ANXA1 is a member of the annexin superfamily which binds to phospholipid membranes producing anti-inflammatory and anti-proliferative effects (Wang et al., 2008; Bae et al., 2005). In the presence of cancer, the protein is found to be down-regulated. As illustrated in Table 1 and Table 4, ANXA decreases in

up-regulation across cohorts with tumor progression and is marginally significant among the CIS cohort. This finding is supported by Wang et al. (2008) where they found up-regulation with increasing cancer progression.

Of particular interest is cytokeratin proteins K1C13 and K1C14 across the three cohorts of women. Among the CIN-2 and CIN-3 cohorts K1C13 and K1C14 are highly up-regulated and significant at the $p<0.0001$ level; however, among the CIS cohort these cytoskeletal protein are not significant. Table 1 illustrates the relative decrease in ratios across the three cohorts from the dysplasia to carcinoma in situ to the point where the proteins are not significant among the CIS group. Research by Carrilho (2004) and Southern (2001) support this finding as they found a deregulation of these genes among individuals diagnosed with invasive cervical carcinoma. This finding, albeit preliminary may help our understanding of cervical tumorigenesis.

S100A8 (Jin et al., 2012) and S100A9 (Qin et al., 2010) proteins are known as a damage-associated molecular pattern molecules due to their pro-inflammatory actions. Moreover, these chemotaxic proteins are involved in various processes including calcium homeostasis, cellular migration, and energy metabolism. S100A9 reportedly is overexpressed in a variety of cancers (Yong & Moon., 2007) and are implicated in the metastatic process; however, our findings tend to agree more with the clinicopathological data published by Zhu et al. (2013) which showed that the expression of S100A9 gradually decreased during the development of cervical squamous carcinoma. The S100A8 & S100 A9, as shown in Table 1 are, indeed, significantly over-expressed for the CIS cohort, but the ratios are approximately three-times less than those shown for the CIN-2 & CIN-3 groups. As suggested by Zhu et al. (2013), S100A9 may play a dual role in tumor progression, depending on the cancer type.

Zinc-alpha2-glycoprotein is a 40-kDa single chain polypeptide assigned to the chromosome 7q22.1. And is involved in carcinogenesis and differentiation. Similar to the S100 A8 and S100 A9 proteins, recent reports showed ZA2G as overexpressed in some tumors, but lost or reduced in other tumors (Abdul-Rahman et al., 2007). In this study, ZA2G is overexpressed and is contrary to our findings in the salivary profile of in situ breast cancer patients where ZA2G was down regulated (Streckfus et al., 2008). One can only speculate, but the difference may be attributed to the two different types of carcinoma. In this study the carcinomas are squamous cell where as in the breast cancer study the tumors were adenocarcinomas (Streckfus et al., 2008). The differences are not clear and require further research.

This is a preliminary study and has its limitations. For example, the results produced a panel of high abundance proteins found in saliva. Low abundance proteins along with those bound to albumin and other carrier proteins warrant investigation. The salivary marker panel also needs to be validated among a larger cohort of subjects taking into account socioeconomic background, HPV status, tumor staging and other potential confounding variables. Analysis of a larger number of cervical cancer specimens and the correlation with biological phenotypes with gene expression patterns may also identify clinically meaningful characteristics. As a footnote, the investigator protein profiled squamous cell carcinomas located in the head and neck. In that study there were 29 differentiated salivary proteins. When comparing the head & neck salivary protein profiles with the CIS salivary pattern there was a 69% percent protein overlap of the differentiated proteins between the two sites i.e. cervix and head/neck demonstrating the heterogeneity of carcinoma from a molecular perspective.

In conclusion, the author has identified differences in protein expression between CIN2, CIN3 and CIS saliva specimens. We have identified a number of aberrantly regulated genes that were previously described in cervical cancers, and a list of genes that were not previously associated with cervical cancer development. What has been demonstrated is the importance in analyzing carcinomas as the tumors progress in staging. In the proceeding paragraphs, examples have been presented showing the changes in protein amplification as the tumor progresses. Some proteins increased during tumorigeneses while others decreased in value across stages and may no longer be significantly differentiated. Taken together the results suggest that alterations in salivary protein expression in cervical tumors may yield clues to their pathogenesis. This could provide important clues to develop novel markers as well as strategies for efficient prevention and therapy for cervical cancer.

Acknowledgements

The authors would also like to thank Dr. William Dubinsky of the UT School of Medicine for the LC-MS/MS salivary mass spectrometry analyses.

Declaration of Interest

The authors have no declarations of interest to report.

References

Abdul-Rahman, P. S., Lim, B. K., & Hasim, O. H. (2007). Expression of high-abundance proteins in sera of patients with endometrial and cervical cancers: analysis using 2-DE with silver staining and lectin detection methods. *Electrophoresis, 28*(12), 1989-1996.

Bae, S. M., Lee, C. H., Cho, Y. L., Nam, K. H., Kim, Y. W., Kim, C. K., … Ahn, W. S. (2005). Two-dimensional gel analysis of protein expression profile in squamous cervical cancer patients. *Gynecologic Oncology, 99*(1), 26-35.

Bosch, F. X., Manos, M. M., Muñoz, N., Sherman, M., Jansen, A. M., & Peto, J. (1995). Prevalence of Human papillomavirus in cervical cancer: a worldwide perspective. *Journal National Cancer Institute, 11*, 796-802.

Capello, M., Ferri-Borgogno, S., Cappello, P., & Novelli, F. (2011). A-enolase: a promising therapeutic and diagnostic tumor target. *FEBS Journal, 278*, 1064-1074. http://dx.doi.org/10.1111/j.1742-4658.2011.08025.x

Carrilho, C., Alberto, M., Buane, L. & David, L. (2004). Keratins 8, 10, 13, and 17 are useful markers in the diagnosis of human cervix carcinomas. *Human Pathology, 35*(5), 546-551.

Cochand-Priollet, B., Cartier, I., de Cremoux, P., Le Galès, C., Ziol, M., … Coste, J. (2005). Cost-effectiveness of liquid-based cytology with or without hybrid-capture II HPV test compared with conventional Pap smears: a study by the French Society of Clinical Cytology. *Diagnostic Cytopathology, 33*(5), 338-343.

Garbett, N. C., Merchant, M. L., Helm, C. W., Jenson, A. B., Klein, J. B. & Chaires, J. B. (2014). Detection of cervical cancer biomarker patterns in blood plasma and urine by differential scanning calorimetry and mass spectrometry. *PLoS ONE*, *9*(1), e84710. http://doi.org/10.1371/journal.pone.0084710

Hassan, M. I., Waheed, A., Yadav, S., Singh, T. P., & Ahmad, F. (2008). Zinc alpha 2-glycoprotein: a multidisciplinary protein. *Molecular Cancer Research*, *6*(6), 892-906.

Higareda-Almaraz, J. C., Enríquez-Gasca Mdel, R., Hernández-Ortiz, M., Resendis-Antonio, O., & Encarnación-Guevara, S. (2011). Proteomic patterns of cervical cancer cell lines, a network perspective. *BMC Systems Biology, 5*, 96. http//www.biomedcentral.com/1752-0509/5/96

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protocols, 4*(1), 44-57.

Jin, L., Shen, Q., Ding, S., Jiang, W., Jiang, L., & Zhu, X. (2012). Immunohistochemical expression of Annexin A2 and S100A proteins in patients with bulky stage IB-IIA cervical cancer treated with neoadjuvant chemotherapy. *Gynecology Oncology, 126*(1), 140-146. http://dx.doi.org/10.1016/j.ygyno.2012.04.005.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Research*, *42*, D199–D205.

Nagaraj, N., Wisniewski, J., Geiger, T., Cox, J., Kircher, M., … Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology, 7*, 548.

NCI Surveillance, Epidemiology and End Results Program. (2014). Retrieved from http://seer.cancer.gov/statfacts/html/cervix.html

Qin, F., Song, Y., Li, Z., Zhao, L., Zhang, Y., & Geng, L. (2010). S100A8/A9 induces apoptosis and inhibits metastasis of CasKi human cervical cancer cells. *Journal of Pathology and Oncology Research, 16*(3), 353-360. http://dx.doi.org/ 10.1007/s12253-009-9225-2.

Schaefer, C., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., & Buetow, K. (2009). PID: The Pathway Interaction Database. *Nucleic Acids Research*, *37*, D674-679.

Southern, S. A., McDicken, I. W., & Herrington, C. S. (2001). Loss of cytokeratin 14 expression is related to human papillomavirus type and lesion grade in squamous intraepithelial lesions of the cervix. *Human Pathology, 32*(12), 1351-1355.

Streckfus, C. F., Mayorga-Wark, O., Arreola D., Edwards, C., Bigler, L. R., & Dubinsky, W. P. (2008). Breast cancer related proteins are present in saliva and are modulated secondary to ductal carcinoma *in situ* of the breast. *Cancer Investigations, 26*(2), 159-167. http://dx.doi.org/10.1080/07357900701783883.

Thiede, B., Koehler, C. J., Strozynski, M., Treumann, A., Stein, R., Zimny-Arndt, U., … Jungblut, P. R. (2013). High resolution quantitative proteomics of HeLa cells protein species using stable isotope labeling with amino acids in cell culture (SILAC), two-dimensional gel electrophoresis(2DE) and nano-liquid

Chromatography Coupled to an LTQ-Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics, 12*(2), 529 –538.

Wang, L. D.,Yang, Y. H., Liu, Y., Song, H. T., Ly Zhang, L.Y., & Li, P. I. (2008). Decreased expression of Annexin A1 during the progression of cervical neoplasia. *Journal of International Medical Research, 36*, 665-672.

Yong, H. Y., & Moon, A. (2007). Roles of calcium-binding proteins, S100A8 and S100A9, in invasive phenotype of human gastric cancer cells. *Archives Pharmacological Research, 30*, 75–81.

Zhu, H., Liu, L., Liu, H., Wu, T., Wu, Y., Zeng, S., & Zeng, L. (2013). Expression of galectin-7 and S100A9 and development of cervical squamous carcinoma. *Zhong Nan Da Xue Xue Bao Yi Xue Ban, 38*(9), 888-895. http://dx.doi.org/ 10.3969/j.issn.1672-7347.2013.09.004