Homi Bhabha National Institute

Ph. D. PROGRAMME

| 1. Name of the Student:          | Bihari Lal Soni                                      |
| 2. Name of the Constituent Institution: | Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer. |
| 3. Enrolment No. :               | LIFE09200704007 (3rd September 2007)                 |
| 4. Title of the Thesis:          | Global protein profiling during rat lingual carcinogenesis and validation of differentiator proteins in human tongue |
| 5. Board of Studies:             | Life Sciences.                                       |
SYNOPSIS

1. Introduction:

Oral squamous cell carcinoma (OSCC) remains a major cancer in the Indian subcontinent, comprising more than 30% of all cancers. The most commonly involved sites of tumor development in the Indian population are buccal mucosa and tongue. The major risk factors for oral cancer are chewing tobacco either alone or with allied products and alcohol consumption. Precancerous lesions of leukoplakia and sub mucous fibrosis are also prevalent in India due to these habits. The five year survival rate of OSCC has not changed in the last few decades.

In patients, the molecular analysis of multiple steps is hampered by the unavailability of biopsies of all the stages of carcinogenesis. However, animal models of carcinogenesis allow the reproducible isolation of all stages, including normal tissues, which are then amenable to pathological, genetic and biochemical analyses. To this end, 4 Nitro-quinoline 1 oxide (4NQO) induced rat model of carcinogenesis remains the preferred model for studies related to oral carcinogenesis because it mimics molecular and pathological changes observed in humans.

Proteomics has grown as a powerful tool for biomarker discovery. iTRAQ-based LC MS-MS is a powerful tool which is utilized in large number of proteomics studies to understand the difference between protein expression profiles of normal vs. diseased samples.

In spite of the fact that a large number of molecules have been identified as potential early diagnostic and prognostic markers for oral cancer, none of them has reached the clinics. Possible reasons could be, 1. Most of the studies do not specify which sub sites were studied and 2. Very few studies have attempted sequential analysis. In
order to sequentially dissect the molecular events during different stages of carcinogenesis, it was proposed to carry out proteomic analysis on samples obtained at sequential stages of rat lingual carcinogenesis. Thus the work in this thesis is towards dissecting sequentially molecular alterations occurring at a single subsite i.e. tongue using a rodent model and validating the observations in cancer of human tongue. The objectives were thus as follows:

**OBJECTIVES:**

1. Establishment of rat lingual cancer model induced by carcinogen

2. Identification of differentially expressed proteins at different stages of lingual cancer development in a rat model using quantitative proteomics

3. Validation of results obtained from quantitative proteomic study

4. Correlation of the data with human samples.

**2. Materials and Methods:**

**2.1 Establishment of rat lingual cancer model induced by carcinogen (4NQO)**

The study was approved by the Institutional animal ethics committee. Approximately 5-6 weeks old Sprague Dawley rats (SD rats) were given 4NQO in drinking water at a concentration of 30 ppm. Rats were divided in three groups and each group was kept for 80, 120, 160 and 200 days respectively. All the animals in a group were sacrificed at each time point, the tongue examined for lesions and the respective tissues were collected and stored at -80°C.
2.2. Histology

5-8µm thick sections of the tissues cut from paraffin blocks were stained with hematoxyline and eosin, and were examined under upright microscope (Axio imager Z1, Zeiss). The pathological status of the tissues was defined by the pathologist.

2.3. Quantitative proteomic studies of rat tissues

2.3.1.1: 2-Dimensional gel Electrophoresis

Total cell lysate of tongue tissue was prepared in Urea lysis buffer. Protein estimation was done by using RC-DC kit (Sigma, USA). Two hundred µg of lysate proteins were resolved on the first dimension using 17 cm strips of either pI range 3-10 or 4-7 (Bio-Rad).

After completion of IEF, the strips were equilibrated in Equilibration buffer I and II for 15 min. respectively. Each strip was placed on to 12% SDS polyacrylamide gel and resolved as per Laemmli protocol7.

The gels were washed and stained with silver essentially according to Fulzele et. al. 20138. The stained proteins on the gels were scanned and proteins which were differentially expressed were subjected to mass spectrometry.

2.3.1.2 Mass spectrometry analysis

Differentially expressed protein spots were cut out from the gel, destained in destaining solution and were subjected to in-gel digestion with Trypsin (20ng/gel piece). The peptides were extracted, reconstituted in 1% Trifluoroacitic acid (TFA) and analyzed on the MALDI TOF-TOF Ultraflex-II from Brucker Daltonics, Germany.
2.3.2: iTRAQ analysis

Hundred µg of protein was pooled from each group of normal (n=10), hyperplasia (n=5), papilloma (n=5) and tumor (n=5) tissues and digested with proteomics grade Trypsin. Peptides generated from normal, hyperplasia, papilloma and carcinoma tissues were labeled with reporter ions of m/z 114, 115, 116 and 117 respectively as per manufacturer’s protocol. Labeled samples were then pooled and subjected to strong cationic exchange chromatography (SCX). SCX fractions were subjected to nanoflow LC system (Agilent 1200 Series) interfaced with LTQ-Orbitrap Velos mass spectrometer. Spectra obtained by mass spectrometer were analyzed by Proteome Discoverer software (Thermo Scientific).

2.4. Bioinformatics analysis of proteomics Data

All differentially expressed proteins were subjected to Gene Ontology (GO) analysis.
2.5. Validation of results obtained by quantitative proteomic study

Some of the differentially expressed proteins identified using either 2DE or iTRAQ-LC-MS/MS analysis of rat tissue samples were further validated by Immunohistochemistry (IHC) and/or RT-PCR. Human tongue samples were also used for validation of novel proteins.

2.5.1 Antibodies

The following antibodies were used

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Dilution</th>
<th>Clone</th>
<th>Company</th>
<th>Catalog no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vimentin</td>
<td>1:400</td>
<td>V9 clone, Mouse monoclonal</td>
<td>Sigma</td>
<td>V 6630</td>
</tr>
<tr>
<td>Transglutaminase 3</td>
<td>1:8000</td>
<td>Mouse monoclonal</td>
<td>SantaCruze</td>
<td>sc-101366</td>
</tr>
<tr>
<td>Periostin</td>
<td>1:50</td>
<td>Rabbit polyclonal</td>
<td>SantaCruze</td>
<td>sc 49480</td>
</tr>
<tr>
<td>Coronin 1a</td>
<td>1:1000</td>
<td>Rabbit polyclonal</td>
<td>Covance</td>
<td>PRB-5002-100</td>
</tr>
<tr>
<td>Tenascin N</td>
<td>1:100</td>
<td>Rabbit polyclonal</td>
<td>HPA</td>
<td>HPA-026764</td>
</tr>
<tr>
<td>Fascin</td>
<td>1:200</td>
<td>Mouse monoclonal</td>
<td>Thermo Fischer</td>
<td>MA1-20912</td>
</tr>
<tr>
<td>Galectin 7</td>
<td>1:300</td>
<td>Mouse monoclonal</td>
<td>SantaCruze</td>
<td>Sc 166222 (H8)</td>
</tr>
</tbody>
</table>

2.5.2 Immunohistochemistry

Formalin-fixed, paraffin-embedded, 5 μm thick rat tissue sections were mounted on poly-L-lysine coated glass slide. IHC was carried out with respective antibody as per standard protocol. Diaminobenzidine was employed as the chromogen and slides were counterstained with Mayor's hematoxylin.
2. 5.3 RNA isolation and qRT-PCR

To validate the proteomics data qRT-PCR was performed whenever the respective antibodies were not available. Total cellular RNA was extracted from the tissue by Tri-reagent (Sigma-Aldrich, USA) as per manufacture’s protocol. RNA was estimated by measuring absorbance at 260 nm and 280 nm using nanodrop (ND-1000 Spectrophotometer, Wilmington, USA). cDNA synthesis was carried out as per the manufacturer’s protocol (Fermentas, Thermo Scientific, Waltham, MA). Obtained cDNA was used as template for qRT-PCR. Master Mix SYBR Green (Applied Biosystems, Bedford, MA) was used with 5 nM of forward and reverse primers. Real-time quantitative PCR was performed with the ABI PRISM7700 Sequence Detection System. Beta actin gene was used as endogenous control. All amplifications were done in triplicate. Results were expressed as relative gene expression using the \(2^{-\Delta\Delta Ct}\) method.\(^9\)

2 Results

2.1 Development of rat model of carcinogenesis

The animals treated with milliQ water, acetone or 80 days 4NQO did not reveal any alterations at the dorsum of the tongue. Hyperplasia/atypical hyperplasia was observed after 120 days and papilloma/atypical papilloma was observed after 160 days at the dorsum of the tongue. Squamous cell carcinoma developed at the dorsum of the tongue in 200 days in rats treated with 4NQO.
3.2 Differential Proteomics

3.2.1: 2DE gel electrophoresis

In the initial study using 2DE gel electrophoresis, five differentially expressed proteins were identified. These included three upregulated proteins; fatty acid binding protein 5, keratin 6 A and serum albumin precursor protein and two down regulated proteins; galectin 7 and transglutaminase 3.

Due to limitations in proper resolution in 2DE gel and poor identification of proteins by MALDI-TOF-TOF, we employed iTRAQ-LC-MS/MS based quantitative proteomics technology for better proteome coverage.

3.2.2: Proteomics study using iTRAQ technology

Four plex iTRAQ LC-MS proteomics analysis at each stage during the tongue tumorogenesis induced by 4NQO in Sprague Dawley rats resulted in identification of 2,223 proteins from the rat tongue SCC of which 415 proteins were found to be differentially expressed in comparison to normal (untreated tissues). Of these 415 proteins, 194 proteins were up-regulated while 221 proteins were down-regulated in SCC of tongue tissues. Table 2 describes the details of differentially expressed proteins at each stage.

<table>
<thead>
<tr>
<th>Stages</th>
<th>No. of Up regulated proteins</th>
<th>No. of Down regulated proteins</th>
<th>Total proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperplasia</td>
<td>35</td>
<td>74</td>
<td>109</td>
</tr>
<tr>
<td>Papilloma</td>
<td>155</td>
<td>178</td>
<td>333</td>
</tr>
<tr>
<td>SCC</td>
<td>194</td>
<td>221</td>
<td>415</td>
</tr>
<tr>
<td>Total</td>
<td>384</td>
<td>473</td>
<td>857</td>
</tr>
</tbody>
</table>

Among the differentially expressed proteins, 5 proteins were sequentially upregulated while 10 proteins were sequentially down regulated from hyperplasia to SCC.

Similarly, sequentially up or down regulation of 62 and 51 proteins respectively was observed from papilloma to carcinoma tissues.
3.3 Validation of differentially expressed proteins

3.3.1 Proteins from gel based study

From the five differentially expressed protein spots identified in the 2DE based study galectin7 was validated because it showed down-regulation as opposed to available literature showing its upregulation in human SCC\textsuperscript{10}. Galectin7 was validated by IHC and found to be down-regulated in rat tongue SCC.

3.3.2 Proteins from iTRAQ based study

Several proteins like Vimentin, K14, K17, MMP9, TGM3 and Periostin, had been reported earlier in human OSCC and were also found to be differentially expressed in rat tongue SCC. The analysis, in addition, detected number of novel proteins which have not been reported previously in human OSCC. In this study some of the known candidate proteins whose differential expression in human oral carcinomas has been shown previously by us and others were validated by IHC or qRT-PCR. Vimentin, Fascin, Periostin and Transglutaminase3 were validated by IHC while Cornulin was validated by qRT-PCR. Vimentin, Fascin and Periostin were found to be sequentially up-regulated while Transglutaminase 3 and Cornulin were found to be sequentially down regulated.

3.3.3 Validation of Novel molecules

Four novel molecules identified by iTRAQ were also validated for their expression in rat tissues. These include Tenascin N and Coronin 1a by Immunohistochemistry and showed sequential upregulation. Trichohyalin and Thrombospondin 2 were validated by real time PCR. Trichohyalin was sequentially down regulated while Thrombospondin 2 was sequentially upregulated. The expression of Tenasin N and Coronin1a were evaluated in human tongue tissues (normal (n= 14), Leukoplakia (n=...
10) and Tumor (n=32)) to determine if the observations from the rat model are also valid in human tongue.

4. Bioinformatics by GO analysis

Bioinformatics analysis was carried out to classify proteins based on subcellular localization and biological function using Gene Ontology (GO) annotations.

These results will be discussed at length in thesis.

5. Conclusions and Future perspective

This is the most extensive quantitative proteomic study in rat model of 4NQO induced oral carcinogenesis carried out to date. Through this model several known proteins like vimentin, fascin, transglutaminase3, peristin and cornulin were identified thereby supporting the use of the model for evaluating markers for different steps of the carcinogenesis process. The model has also enabled the identification of novel molecules like Tenascin N, Coronin1a, Trichohyalin, and Thrombospondin2. Using this model, it has been possible to show sequential alterations in expression pattern during rat tongue carcinogenesis. Furthermore, the observations could also be extrapolated from the rat model data to human system indicating the fact that this model has potential to be used for biomarker discovery for human oral cancer. The clinical utility of the novel proteins will be now evaluated on a large scale on human tissues of SCC of tongue at different stages i.e. from T1 to T4, and leukoplakia of tongue with the ultimate aim of establishing these proteins as predictive markers for human oral cancer.
References:


Publications in Refereed Journal:

Accepted: Bihari Lal Soni, Arivusudar Marimuthu, Harsh Pawar, Sharada S. Sawant, Anita Borges, Ranganathan Kannan, Pandey, A., Arvind D. Ingle, Harsha, H. C., and Vaidya M.M.

"Quantitative proteomic analysis of different stages of rat lingual carcinogenesis" Accepted for publication in Clinical Communications-Oncology.

Signature of Student: 

Date: 16.04.2014

Doctoral Committee:

<table>
<thead>
<tr>
<th>Sr. N</th>
<th>Name</th>
<th>Designation</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dr. Vinay Kumar</td>
<td>Chairman</td>
<td></td>
<td>17.04.2014</td>
</tr>
<tr>
<td>2</td>
<td>Dr. Milind Vaidya</td>
<td>Convener</td>
<td></td>
<td>16.04.2019</td>
</tr>
<tr>
<td>3</td>
<td>Dr. G.B. Maru</td>
<td>Member</td>
<td></td>
<td>16.04.2014</td>
</tr>
<tr>
<td>5</td>
<td>Dr. Rukmini Govekar</td>
<td>Member</td>
<td></td>
<td>16/04/14</td>
</tr>
<tr>
<td>6</td>
<td>Dr. Anita Borges</td>
<td>Invitee</td>
<td></td>
<td>15/04/14</td>
</tr>
</tbody>
</table>

Forwarded through

Dr. S.V. Chiplunkar

Director, ACTREC &
Chairperson, Academics and Training Programme,

ACTREC

Dr. S. V. Chiplunkar
Director
Advanced Centre for Treatment, Research & Education in Cancer (ACTREC)
Tata Memorial Centre

Dr. K.S. Sharma
Director Academics
Tata Memorial Centre

Prof. K.S. Sharma
DIRECTOR - ACADEMICS, TMC
Mumbai - 400 012