Contents

ABBREVIATIONS i-v
LIST OF FIGURES vi-ix
LIST OF TABLES x-xiv
ABSTRACT xv-xvi

CHAPTER 1: INTRODUCTION 1

1.1 CERVICAL CANCER: INCIDENCE AND MORTALITY 2
1.2 CERVICAL CANCER: DISEASE SITE AND STRUCTURE 3
1.3 DETECTION AND CLASSIFICATION OF CERVICAL CANCER 4
  1.1.1 Detection and cytological classification of cervical smears 4
  1.1.2 Invasive cervical cancer staging 6
  1.1.3 Types of cervical cancer 9
1.4 HUMAN PAPILLOMAVIRUS : THE MAJOR ETIOLOGIC AGENT OF CERVICAL CANCER 10
  1.4.1 History of Human Papillomavirus (HPV) association with cervical cancer 10
  1.4.2 HPV Types and their prevalence 10
  1.4.3 The HPV16 genome 11
  1.4.4 Productive infection by HPV16 and transformation of cervical epithelium 13
    (A) Maintenance Phase 14
    (B) Differentiation-dependent Phase 14
1.5 FACTORS ASSOCIATED WITH CERVICAL CANCER PATHOGENESIS

1.5.1 Viral Factors

(A) Physical status of HPV16 genome
(B) HPV16 load in cervical tissues
(C) The role of HPV16 oncoprotein E7 in cervical cancer development

1.5.2 Host Factors

(A) Polycomb repressive complexes and cancers
(B) HOX Cluster members (coding and noncoding) and cancers

1.5.3 Environmental factors

1.6 PREVENTION, MANAGEMENT AND TREATMENT OF CERVICAL CANCER

1.6.1 Pap Screening and HPV testing

1.6.2 Vaccination

(A) Preventive
(B) Therapeutic

1.6.3 Treatment

CHAPTER 2: BACKGROUND AND OBJECTIVES

2.1 ORIGIN OF THE STUDY

2.2 BACKGROUND OF THE STUDY

2.2.1 Viral factors necessary for cervical cancer pathogenesis

(A) Impact of Human Papillomavirus 16 (HPV16) load on cervical cancer development
(B) Role of HPV16 integration on cervical cancer pathogenesis
(C) Role of HPV16 oncoprotein E7 in cervical
2.2.2 Host genetic and epigenetic factors in cervical cancer pathogenesis

(A) Alteration of global gene expression in cancers
(B) HOX cluster encoded transcription factors and long noncoding transcripts in cancers
(C) Chromatin remodelling Polycomb Repressive Complex 2 (PRC2) in cancer

2.3 OBJECTIVES OF THE STUDY

CHAPTER 3: METHODS

3.1 STUDY DESIGN

3.2 SUBJECTS AND SAMPLES

3.2.1 Statement of ethics

3.2.2 Subject recruitment for identification of various categories of cervical samples

3.2.3 Malignant tissues from cervical cancer patients

3.2.4 Sample collection and storage

3.2.5 Histopathological analysis of cervical tissue biopsies

3.3 MOLECULAR CHARACTERIZATION OF SAMPLES

3.3.1 DNA isolation

3.3.2 Estimation of DNA concentration and purity

(A) Nanodrop based DNA quantification

(B) Fluorometer based DNA quantification

3.3.3 Screening of cervical tissue DNA samples for HPV

(A) Detection for presence of HPV infection

(B) Testing cervical tissue DNA samples for the presence of HPV16
3.3.4 Estimation of HPV16 load and the E2 status in cervical cancer samples
   (A) Determination of E6 copy numbers 42
   (B) Determination of E2 copy numbers 44
3.3.5 RNA isolation 45
3.3.6 Estimation of RNA concentration and purity 46
   (A) Fluorometer based RNA quantitation 46
   (B) Assessment of RNA quality by Agilent Bioanalyzer 2100 46
3.3.7 cDNA preparation 47
3.3.8 Identification of integration status of HPV16 in cervical samples 48
   (A) APOT assay 49
   (B) Taqman assays 49

3.4 GENE EXPRESSION ANALYSIS OF VIRAL AND HOST TRANSCRIPTS 50
3.4.1 Quantitation of HPV16 E7 expression by Taqman based qRT-PCR assay 50
3.4.2 Quantitation of host transcript expression by SYBR Green based qRT-PCR assay 51
3.4.3 Quantitation of miRNA expression by SYBR Green based qRT-PCR assay 55

3.5 MICROARRAY BASED GLOBAL GENE EXPRESSION PROFILING 56
3.5.1 Gene expression assay work plan 56
3.5.2 Microarray data analysis 57
3.5.3 Pathway Analysis 59

3.6 IMMUNOBLOT ANALYSIS OF VIRAL AND HOST PROTEINS 60

3.7 CLONING AND EXPRESSION OF HPV16 E7 INTO MAMMALIAN CELLS 60
3.7.1 Cloning of HPV16 E7 into mammalian expression
vector pcDNA3.1(+)

3.7.2 Cell culture
3.7.3 Transfection of plasmids into cell lines

3.8 SEQUENCING OF HOTAIR GENE TO IDENTIFY CERVICAL CANCER ASSOCIATED SEQUENCE VARIATIONS

3.8.1 Deep sequencing of HOTAIR gene
(A) PCR based amplification of the target region and library preparation
(B) Quality checking and identification of variations in HOTAIR gene by Ion Torrent algorithm

3.8.2 Sanger sequencing based validation of findings from deep sequencing of HOTAIR gene

3.9 RNA IMMUNOPRECIPITATION

3.10 CHROMATIN IMMUNOPRECIPITATION-QPCR

3.11 IN SILICO ANALYSIS

3.11.1 RPISeq
3.11.2 catRAPID
3.11.3 ImageLab (BioRad)
3.11.4 RNAsnp
3.11.5 SNPsfold
3.11.6 lncRNASNP
3.11.7 Ingenuity Pathway Analysis (IPA)
3.11.8 ENCODE database

3.12 STATISTICAL ANALYSIS

3.12.1 Hypothesis testing
3.12.2 Kolmogorov-Smirnov test (K-S test)
3.12.3 Mann-Whitney U test
3.12.4 Two-sample t-test
3.12.5 Linear regression
3.12.6 Test for association (Chi-square test)
3.12.7 Multiple testing corrections
CHAPTER 4: RESULTS

4.1 GENE EXPRESSION PROFILING OF CERVICAL SAMPLES TO IDENTIFY GENES AND BIOLOGICAL PATHWAYS ASSOCIATED WITH CERVICAL CANCER PATHOGENESIS

4.1.1 Demographic characteristics of the study subjects

4.1.2 Hierarchical clustering analysis of the cervical samples

4.1.3 Identification of differentially expressed genes among the cervical cancer cases in comparison to the histopathologically normal samples (HPV negative control samples and HPV16 positive non-malignant samples)

4.1.4 Identification of significantly altered biological processes among the cervical cancer cases in comparison to the histopathologically normal samples

4.1.5 Identification of significantly altered pathways among the cervical cancer cases in comparison to the histopathologically normal samples

4.1.6 Identification of differentially expressed genes among the cervical cancer cases in comparison to the histopathologically normal samples: HPV negative controls and HPV16 positive non-malignant samples

4.2 DIFFERENTIAL EXPRESSION OF HOX CLUSTER MEMBERS AMONG CERVICAL CANCER CASES

4.2.1 Microarray based gene expression profiling revealed deregulation of several HOX cluster members among cervical cancer cases, in comparison to the HPV negative control samples and HPV16 positive non-malignant samples

4.2.2 Quantitative real-time PCR (SYBR green assay) based confirmation of deregulated expression of HOX cluster members
4.2.3 Correlation analysis of expression of HOX Cluster members with HPV16 E7 expression and viral load 108
4.2.4 Correlation analysis between the expression levels of HOX Cluster members and the Epithelial Mesenchymal Tansition (EMT) markers, E-Cadherin and Vimentin 110

4.3 DIFFERENTIAL EXPRESSION OF LONG NONCODING RNAs (lncRNAs) OF THE HOX CLUSTER 113
4.3.1 Demographic characteristics of the study subjects 113
4.3.2 Quantitative real-time PCR (SYBR green assay) based identification of HOX cluster lncRNA deregulation 113

4.4 lncRNA HOTAIR PLAYS A CRUCIAL ROLE IN REGULATING GLOBAL GENE EXPRESSION IN CERVICAL CANCER PATHOGENESIS 117
4.4.1 Classification of cervical cancer samples into two subgroups based on HOTAIR expression levels 117
4.4.2 Correlation of HOTAIR expression with HPV16 E7 expression and viral load among cervical cancer cases 120
4.4.3 Negative correlation exists between HOTAIR and HOXD10 expression among cervical cancer cases 123
4.4.4 Effect of HOTAIR deregulation on global gene expression 124
4.4.5 Impact of sequence variations in HOTAIR gene on expression and function of HOTAIR in cervical cancer pathogenesis 150
   (A) Deep sequencing based identification of HOTAIR sequence variations 152
   (B) Sanger sequencing based validation of the results 155
   (C) Impact of rs2366152C on HOTAIR expression among cervical cancer cases 159
   (D) In silico approach revealed impact of rs17720428, rs12312094 and rs2366152 on HOTAIR function and expression 160
(E) miR-22-5p shows increased expression in cervical cancer samples and correlates negatively with HOTAIR and HPV16 E7 expression.

4.5 IDENTIFICATION OF THE ROLE OF HPV16 E7 IN REGULATING HOX CLUSTER GENE EXPRESSION BY REGULATING PRC2-LSD1 COMPLEX ACTIVITY THROUGH HOTAIR

4.5.1 Altered expression of HOTAIR interacting partners and PRC2-complex members, EZH2 and SUZ12

4.5.2 HPV16 E7 possesses the ability to interact physically with HOTAIR

4.5.3 HPV16 E7 possesses the ability to regulate HOX gene expression by modulating H3K4me3 and H3K27me3 marks created by HOTAIR through recruitment of PRC2-LSD1 complex

CHAPTER 5: DISCUSSION

5.1 CERVICAL CANCERS SHOW DISTINCT GENE EXPRESSION SIGNATURES IN COMPARISON WITH HPV NEGATIVE CONTROLS AND HPV16 POSITIVE NON-MALIGNANT SAMPLES

5.2 HOX CLUSTER MEMBERS (CODING AND NON-CODING) PLAY A CRUCIAL ROLE IN REGULATING CERVICAL CANCER PATHOGENESIS

5.3 HOTAIR SERVES AS AN IMPORTANT BIOMARKER FOR STRATIFICATION OF CERVICAL CANCER SAMPLES

5.4 HPV16 E7 FUNCTIONS AS THE MASTER REGULATORY MOLECULE IN DRIVING CERVICAL CANCER PATHOGENESIS BY REGULATING SEVERAL HOST ENCODED MOLECULES (CODING AND NON-CODING)
CONCLUSION 192
FUTURE DIRECTIONS 194
BIBLIOGRAPHY 195
LIST OF PUBLICATIONS 216
CONFERENCES ATTENDED AND AWARDS RECEIVED 217
ANNEXURE-I