LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Investigation of oral and oropharyngeal squamous cell carcinoma in young adults

A thesis submitted for the Degree of Doctor of Philosophy

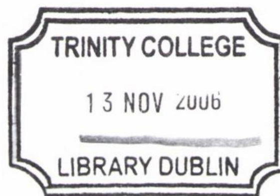
By

Esther O'Regan
BDS FDS

Trinity College
University of Dublin

Michelmas Term
2005

Supervisors: Dr. Mary Toner FRCPath
& Dr Orla Sheils, FAMLS, PhD


TRINITY COLLEGE
13 NOV 2006
LIBRARY DUBLIN
THESIS
8024

Declaration

I declare that this thesis is my own work, and has not been submitted previously for a PhD degree at this or any other university. I agree that the library may lend or copy this thesis on request

Esther O'Regan

For my parents Michael and Kay

Contents

Acknowledgements i

List of abbreviations iii

Publications vi

Summary ix

Chapter 1

General Introduction **1**

1.1	Oral and oropharyngeal cancer: historical context	2
1.2	Anatomy of oral and oropharyngeal cancer	4
1.2.1	Site distribution in young adults	7
1.3	Histology of oral and oropharyngeal SCC	7
1.3.1	Tumour differentiation in young adults	10
1.4	Prognosis	10
1.4.1	Prognosis in young adults	11
1.5	Aetiology	12
1.5.1	Tobacco	13
1.5.2	Alcohol	14
1.5.2.1	Tobacco and alcohol in young adults	15
1.5.3	Diet and nutrition	16
1.5.3.1	Diet and nutrition in young adults	17
1.5.4	Marijuana	17
1.5.4.1	Marijuana in young adults	18
1.5.5	Genetic susceptibility	18

1.5.5.1	Genetic susceptibility in young adults	19
1.5.6	Human papillomavirus	20
1.5.6.1	Human papillomavirus in young adults	22
1.5.7	Occupational	22
1.6	Molecular Pathology of oral and pharyngeal cancer	23
1.6.1	Genetic instability	23
1.6.1.1	Chromosome 3	23
1.6.1.2	Chromosome 7	24
1.6.1.3	Chromosome 11	24
1.6.2	Oncogenes	24
1.6.2.1	Cyclin D1	25
1.6.2.2	c-Myc	25
1.6.3	Tumour Suppressor Genes	26
1.6.3.1	CDKN2A ($p16^{INK4}$)	26
1.6.3.1.1	p16 in young adults	28
1.6.3.2	p53	28
1.6.3.2.1	p53 in young adults	29
1.7	Comparative genomic hybridisation	30
1.8	Microarrays	31
1.9	Genomic DNA arrays	31
1.10	Expression arrays	32
1.11	Aims and objectives	33
1.12	References	34

Chapter 2

Methodology

2.1	Introduction	54
2.2	**Section 1**	**54**
2.2	Samples	54
2.2.2	Fresh tissue harvesting and archiving	56
2.2.3	Formalin fixed paraffin embedded tissue	56
2.3	**Section 2**	**57**
2.3.1	Tissue preparation prior to nucleic acid extraction	57
2.3.2	Frozen section protocol	57
2.3.3	Fixing and staining for H+E	58
2.3.4	Tissue homogenisation	58
2.3.5	Mortar and pestle homogenisation	59
2.3.6	QIAshredder®	59
2.3.7	Microfuge pellet mixer	60
2.4	**Section 3**	**60**
2.4.1	Arcturus PixCell II Laser Capture Microdissection system	60
2.4.2	Setting the LCM parameters	62
2.5	**Section 4**	**63**
2.5.1	Nucleic acid extraction	63
2.5.2	DNA extraction from frozen sections	63
2.5.3	RNA extraction using Qiagen RNeasy spin columns	65
2.6	**Section 5**	**69**
2.6.1	Assessment of nucleic acids	69

2.6.2	Quantifying RNA and DNA Samples	70
2.6.3	Assessing of nucleic acid integrity	70
2.6.4	Assessing RNA quality/quantity	71
2.7	**Section 6**	**74**
2.7.1	Reverse transcription of RNA using cDNA archive kit	74
2.8	**Section 7**	**75**
2.8.1	TaqMan® PCR	75
2.8.2	Real-time quantitative TaqMan® RT-PCR	78
2.8.3	Choice of template DNA/positive control	80
2.9	**Section 8**	**81**
2.9.1	Nucleic acid sequence-based amplification (NASBA)	81
2.9.2	PreTect HPV Proofer Kit (Norchip AS)	84
2.9.3	NASBA with molecular beacon 'real-time' detection	84
2.10	**Section 9**	**85**
2.10.1	Immunohistochemistry using the $p16^{INK4A}$ Research Kit	85
2.10.2	Interpretation of antigen expression in biopsy tissues	87
2.11	**Section 10**	**88**
2.11.1	Array CGH Background	88
2.11.2	Labeling Protocol	89
2.11.3	DNase Digestion and Labeled Probe Purification	90
2.11.4	Checking the Labeled DNA	91
2.11.5	Hybridization Protocol	91
2.11.6	Image analysis	95
2.12	**Section 11**	**97**
2.12.1	Applied Biosystems expression microarray system	97

2.12.2 Reverse transcription-In vitro transcription (RT-IVT) 98

2.12.2.1 RT-IVT labelling protocol 100

2.12.3 Chemiluminescent detection 103

2.12.3.1 Chemiluminescence detection protocol 104

2.12.4 Array controls 108

Chapter 3

Head and neck squamous cell carcinoma in a cohort of young Irish adults

3.1 Summary 120

3.2 Introduction 121

3.2.1 Cancer in Ireland 121

3.2.2 Head and neck cancer in Ireland 123

3.3 Aims 125

3.4 Materials and Methods 126

3.5 Results 127

3.5.1 Age and gender 129

3.5.2 Site of tumour and history of risk factors 130

3.5.3 Stage, grade and haematological status 131

3.6 Discussion 132

3.7 Conclusion 137

3.8 References 138

Chapter 4

Microarray comparative genomic hybridisation analysis of oral and oropharyngeal squamous cell carcinoma

4.1	Summary	144
4.2	Introduction	145
4.2.1	Comparative Genomic Hybridisation	146
4.2.2	Microarrays	147
4.2.3	Vysis Genosensor Array 300	147
4.3	Aims	148
4.4	Materials and Methods	149
4.4.1	Patients	151
4.4.2	Laser Microdissection and DNA extraction	151
4.4.3	CGH	151
4.4.3.1	Validation of technique using Control DNA	151
4.4.3.2	Validation of system using sex mismatch test/reference	151
4.5	Results	154
4.5.1	Average number of aberrations	155
4.5.2	Specific aberrations	159
4.6	Discussion	162
4.6.1	Frequency of aberrations	163
4.6.2	Copy gain/loss profile differences	164
4.7	Conclusion	167
4.8	References	169

Chapter 5

Human papillomavirus detection and p16^{INK4} expression in SCC of the oral cavity and oropharynx

5.1	Summary	176
5.2	Introduction	177
5.2.1	p16	177
5.2.2	HPV	178
5.3	Aims	180
5.4	Section 1 Analysis of p16^{INK4} protein and expression levels	181
5.4.1	Materials and Methods	181
5.4.1.1	Immunohistochemistry	181
5.4.1.2	TaqMan PCR for p16	184
5.4.2	Results	186
5.4.2.1	Site	186
5.4.2.2	Immunohistochemistry	186
5.4.2.3	mRNA expression of p16	190
5.5	Section 2 HPV detection in oral and pharyngeal carcinoma	195
5.5.1	Materials and Methods	198
5.5.1.2	DNA and RNA extraction	198
5.5.1.3	TaqMan PCR for detection of HPV16/18 DNA	198
5.5.1.4	Nucleic acid sequence based amplification (NASBA)	200
5.5.2	Results	203
5.5.2.1	Comparison of HPV DNA and mRNA detection	203
5.5.2.2	Correlation between HPV status and clinical features	204

5.5.2.3 Correlation between p16^{INK4A} status and HPV status 205

5.6 Discussion 207

5.6.1 p16 208

5.6.2 HPV 210

5.6.3 P16 and HPV 213

5.7 Conclusion 214

5.8 References 215

Chapter 6

Expression arrays

6.1 Summary 225

6.2 Introduction 226

6.2.1 Expression array studies 226

6.2.2 Applied Biosystems Human Genome Survey Microarray 227

6.2.3 Expression profiles of head and neck cancer 227

6.3 Aims 230

6.4 Materials and Methods 231

6.4.1 Samples 231

6.4.2 RNA Extraction, *In vitro* transcription 231

6.4.3 Chemiluminescence Reaction 233

6.4.4 Analysis of Microarray Data 234

6.4.4.1 Replicates 234

6.4.4.2 Normalisation 234

6.4.4.3 Analysis by Spotfire 235

6.4.4.4 Filtering genes 235

6.4.4.5 Cluster analysis 236

6.4.4.6 Treatment comparison 236

6.4.4.7 Adjusting for fold change noise 239

6.4.4.8 Analysis of gene lists 240

6.5 Results 240

6.5.1 RNA analysis 240

6.5.2 Post RT-IVT cRNA analysis 243

6.5.3 Microarray image capture 245

6.5.4 Statistical analysis 246

6.5.4.1 Normalisation 246

6.5.4.2 Treatment comparisons 247

6.5.4.3 Distinction Analysis 252

6.5.4.4 Smokers versus nonsmokers 253

6.6 Discussion 254

6.7 Conclusion 257

6.8 References 258

Chapter 7

Validation of gene expression array data

7.1 Summary 263

7.2 Introduction 264

7.3 Aims 265

7.4 Material and Methods 265

7.4.1 Specimens 265

7.4.2 Assays 266

7.4.3 CDNA generation and TaqMan PCR 267

7.4.4 Statistical methods 267

7.4.4.1 Comparing expression array and TaqMan PCR 267

7.4.4.2 Comparing CGH array and expression array 269

7.5 Results 269

7.5.1 Young nonsmokers vs old smoker 270

7.5.2 Smokers versus nonsmokers 273

7.5.3 CGH array and expression array data 273

7.6 Discussion 275

7.6.1 DOC-1 277

7.6.2 IL-12 and STAT 4 278

7.6.3 MCM10 and CDC6 279

7.6.4 Other genes 281

7.6.5 P16 and RB1 282

7.6.6 CDKN2B 283

7.7 Conclusion 284

7.8 References 285

Chapter 8

General Discussion

8.1 Introduction 289

8.2 Review of results and conclusions 291

8.3 An emerging approach – pathway analysis 294

8.4 'Every clarification breeds new questions' Arthur Bloch 297

8.5 References 299

Figures

Figure 1.1	Anatomy of oral cavity and oropharynx	4
Figure 1.2	Grades of squamous cell carcinoma	9
Figure 1.3	Disruption of the Rb cell cycle regulatory pathway	27
Figure 2.1	Overview of LCM Process	61
Figure 2.2	Example of cells captured by LCM	62
Figure 2.3	Agarose gel electrophoresis of total RNA	71
Figure 2.4	Agilent Bioanalyser 2100 'chip	72
Figure 2.5	Example of an electropherogram	73
Figure 2.6	TaqMan PCR	76
Figure 2.7	Increased fluorescence activity due to cleaved probe	77
Figure 2.8	Example of an amplification plot	79
Figure 2.9	Schematic Representation of NASBA	83
Figure 2.10	Agarose Gel Electrophoresis of purified probe	91
Figure 2.11	Genosensor array 300 Report page	96
Figure 2.12	RT-IVT DIG-labelling of sample mRNA	99
Figure 2.13	Chemiluminescent detection of DIG-labelled cRNA	103
Figure 2.14	Controls used in Applied Biosystems microarrays	109
Figure 2.15	Fluorescence oligos used to image microarray features	114
Figure 3.1	Incidence of cancers in Ireland by site, 1998-2000	122
Figure 3.2	Mortality rates from cancer in Ireland, 1998-2000	122
Figure 3.3	Incidence/mortality rates for oral/pharyngeal/laryngeal cancer in Ireland 1994-2000	124
Figure 3.4	Male: female ratio in each age category	129

Figure 3.5	Age & gender distribution of HNSCC patients	130
Figure 3.6	Distribution of primary HNSCC in under 40 yr olds and over 40 yr olds	131
Figure 4.1	Comparative genomic hybridisation	148
Figure 4.2	Microarray visualization: 9p21.3(MTAP) deletion (red) and 17q21 (BRCA 1) amplification (green).	154
Figure 4.3	Mean numbers gains and losses in each cohort.	156
Figure 4.4	Average number of aberrations in various age groups	156
Figure 4.5	Average no. of aberrations in smokers vs nonsmokers	157
Figure 5.1	Grade 3 Immunohistochemical staining for $p16^{INK4A}$	187
Figure 5.2	Grade 2 Immunohistochemical staining for $p16^{INK4A}$	187
Figure 5.3	p16 immunohistochemistry: normal oral mucosa	187
Figure 5.4	p16INK4 detection in oral and pharyngeal SCC	190
Figure 5.5	TaqMan RT-PCR standard curves	191
Figure 5.6	$p16^{INK4}$ TaqMan amplification plot	192
Figure 5.7	Normalised $p16^{INK4}$ mRNA expression vs p16 protein expression and CDKN2A status	192
Figure 5.8	Comparison of $p16^{INK4}$ expression between old smokers and young nonsmokers	193
Figure 5.9	Normalised p16 expression vs tumour site	194
Figure 5.10	HPV 16 DNA Genomic Map	195
Figure 5.11	Action of E6 on p53	196
Figure 5.12	Action of E7 on Rb	197

Figure 5.13	Amplification plot for p16	200
Figure 5.14	PreTect Analysis Software. Positive samples (green)	202
Figure 5.15	HPV status in oral and pharyngeal SCC	204
Figure 5.16	Comparison of HPV and p16INK4 status	206
Figure 6.1	Overview of Chemiluminescence chemistry	233
Figure 6.2	Agarose gel electrophoresis of 2 RNA samples	241
Figure 6.3	Agarose gel electrophoresis of a cRNA sample	243
Figure 6.4	Captured image of sample number 12166 microarray	245
Figure 6.5	MA plots two samples	246
Figure 6.6	Heatmap of hierarchical cluster	248
Figure 6.7	Comparison of genes showing increased expression in old smokers compared to young nonsmokers	249
Figure 6.8	Profile chart of genes showing increased and decreased expression in young non-smokers and old smokers with clear distinction patterns	252
Figure 6.9	Panther Classification of transcripts	253
Figure 7.1	Expression levels of IL-12 and STAT4 in young non-smokers vs old smokers	270
Figure 7.2	MCM10 and CDC6 in young nonsmokers vs old smokers	271
Figure 7.3	DOC-1 & CDK5RAP in young nonsmokers vs old smokers	271
Figure 7.4	TA-NFKBH and TRRAP in young non-smokers versus old smokers	272
Figure 7.5	Expression levels of RBA600, p16 and RB1	272
Figure 7.6	Expression levels of p15 in all samples	273
Figure 7.7	Comparison of CGH and cDNA arrays	274

Figure 7.8	IL-12 and STAT 4 signalling pathways	279
Figure 7.9	Regulation of DNA replication	280
Figure 7.10	CDKN2A and CDKN2B pathways	284
Figure 8.1	Signalling pathways used in SCC in young non-smokers versus old smokers	295

Tables

Table 1.1	Frequency of oral SCC according to site	6
Table 1.2	Oropharyngeal SCC according to site	7
Table 1.3	Relationship of oral cancer TNM stage to 5-year survival rate	11
Table 2.1	Samples used in each experiment	55
Table 2.2	LCM Settings	62
Table 3.1	Correlation between age at diagnosis of HNSCC and clinicopathological features	128
Table 4.1	Clinicopathological details of all samples	150
Table 4.2	Validation by sex mismatch.	153
Table 4.3	Analysis of numbers of aberrations by clinical details	158
Table 4.4	The most frequent copy number imbalances in the old group and comparison with the young group	160
Table 4.5	The most frequent copy number imbalances occurring in the 'young' group and comparison with the old group	161
Table 5.1	Clinicopathological details of all 25 samples	183
Table 5.2	Immunohistochemistry results	189
Table 5.3	Statistical analyses of $p16^{INK4}$ mRNA expression	193
Table 5.4	Primers and probes used in HPV detection and typing	200
Table 5.5	Statistical analyses of HPV results using Mann-Whitney U test and Fisher Exact Test	205
Table 6.1	Sample Details	232
Table 6.2	UV spectrophotometric analysis of SCC samples	242

Table 6.3	UV spectroscopic analysis of RT-IVT outputs	244
Table 6.4	Genes with increased expression in young nonsmokers relative to old smokers after t test (P<0.001). Average fold change >2	250
Table 6.5	Decreased expression in young nonsmokers vs old smokers	251
Table 6.6	Genes showing increased expression in nonsmokers with oral and oropharyngeal cancer relative to smokers with oropharyngeal cancer	254
Table 6.7	Genes showing decreased expression in non-smokers with oropharyngeal cancer relative to smokers with oropharyngeal cancer	254
Table 7.1	Twelve genes selected for the validation study	266

Acknowledgements

I'd like to begin by thanking my supervisors:

Dr. Mary Toner – my mentor and my friend, who has spent so many hours with me on various planes, trains and automobiles brainstorming this thesis among other topics. My thanks to Dr. Orla Sheils, for her help, advice and for her generosity in supplying us students with frequent coffees and pastries! I would like to thank Professor John O'Leary for constant encouragement and kind words and for all the guidance over the past few years.

I would like to acknowledge support of a Cancer Research Ireland grant, which has allowed me to build firm foundations for head and neck cancer research in Dublin. Thanks to the Dublin Dental School and Hospital for supporting through the past 5 years. Many thanks to Applied Biosystems for providing excellent training and assistance throughout the project.

Thanks also to Professor Conrad Timon, his patients, and his team, for making all this work possible. Thanks to the CPL (especially Jan Walker and Ronan Ward) for their invaluable expertise.

Now a word of thanks for the basement crew. A million thanks to Paul Smyth, who has been a patient teacher, a sound advisor and a pretty good Santa Claus over the past few years. Thanks to Sue (the best reading room decorator ever) for the laughs, and for my daily celebrity gossip updates. Keep working on the Cork accent! Thanks to Amanda for making those long array days much more bearable. Richie (Hungry) Flavin thanks for the songs!!! Jean, thanks for the coffee breaks and chats. Thanks to the research lab in the Coombe Womens Hospital (Cara, Martina, Niamh, Cynthia) for all the help, advice and the chat over the years. Thanks to Anya for keeping me

motivated. Many many thanks to my family (Dad, Mom, Karen, Michael and Oonagh) for listening to me and providing me with constant support in everything I do.

Finally, I wish to thank my husband Stephen, firstly for listening to me on the good days and the bad, and providing advice, assistance and a shoulder to cry on every step of the way. I would not and could not have done this without you.

List of Abbreviations

AJCC	American Joint Committee on Cancer Staging
APES	3-aminopropyltriethoxysaline
BAC	Bacterial Artificial Chromosome
BLAST	Basic Local Alignment Search Tool
β-ME	Beta-Mercaptoethanol
CDC6	cell division cycle 6 homologue
CDKN2A	Cyclin dependant kinase inhibitor 2A
CL	Chemiluminescence
Cy 3/5	Cyanine 3/5
CYP1A1	Cytochrome P-450
CGH	Comparative Genomic Hybridisation
DAB	diaminobezaminidine
DIG-dUTP	digoxigenin-dUTP
DNA	Deoxyribonucleic acid
DOC-1	Downregulated in ovarian cancer
EtOH	Ethanol
EVA	ethylene vinyl acetate
EGFR	Epidermal Growth Factor Receptor
FAM	6-carboxyfluorescein
FFPE	Formalin fixed paraffin embedded
FGFR	Fibroblast Growth Factor Receptor
FHIT	Fragile Histidine triad
FISH	fluorescence in situ hybridisation

FWER	Family wise error Rate
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GSTM1	glutathione-S-transferase
H+E	haematoxylin and eosin
HMM	Hidden Markov Model
HNSCC	Head and neck squamous cell carcinoma
HPV	Human Papillomavirus
IARC	International Agency for Research on Cancer
IVT	In vitro transcription
LCM	Laser Capture Microdissection
LOH	loss of heterozygosity
MAPK	mitogen-activated protein kinase
MCM	minichromosome maintenance
MGB	Minor groove binding
NASBA	Nucleic acid sequence-based amplification
NTC	No template control
PAC	P1-derived artificial chromosomes
RB	Retinoblastoma
RNA	Ribonucleic acid
rRNA	ribosomal RNA
RT-PCR	Reverse transcriptase polymerase chain reaction
SCC	Squamous cell carcinoma
STAT4	Signal Transducer and activator of transcription
TAMRA	6-carboxy-N, N, N´, N´-tetramethylrhodamine
TET	6-carboxy-4, 7,2´, 7´-tetrachlorofluorescein

TNM	Tumour Node Metastasis
Tm	Melting temperature
TRRAP	transformation/transcription domain-associated protein
TSG	Tumour Suppressor Gene
UICC	International Union Against Cancer
UNG	Uracil-DNA glycosylase
UTP	Uridine triphosphate
W.H.O	World Health Organisation

Publications

Array CGH analysis of head and neck squamous cell carcinoma in young adults. Head Neck. 2006 Apr;28(4):330-8. PMID: 16470878

O'Regan EM, Toner M, Smyth P, Finn S, Timon C, Flavin R, Cahill S, O'Leary J, Sheils O.

Head and neck squamous cell carcinoma in young Irish adults. Br J Oral Maxillofac Surg. 2006 Jun;44(3):203-6.. PMID: 16005553

O'Regan EM, Sheils OM, Timon C, Codd M, Stassen L, O'Leary J, Toner M.

Delay in diagnosis and its effect on outcome in head and neck cancer. Br J Oral Maxillofac Surg. 2005 Aug; 43(4): 281-4.

McGurk M, Chan C, Jones J, O'Regan E, Sherriff M.

Array CGH analysis of gamma irradiated human thyrocytes. 2004. Finn SP, Smyth P, O'Regan EM, Cahill S, Flavin R, O'Leary JJ, Sheils O. Virchows Archive 445; 396-404

Abstracts

Absence of p16 deletions in young patients with head and neck cancer.

O'Regan E, Sheils O, Smyth P, Finn SP, Flavin R, Timon C, Toner M, O'Leary.

Presented at the USCAP annual meeting, San Antonio, Texas, 2005. This presentation was awarded the **Stowell Orbison award** for best poster 2005. Abstract published in Modern Pathology.

Genome wide investigation of head and neck SCC in young adults. O'Regan E, Sheils O, Smyth P, Finn SP, Flavin R, Timon C, Toner M, O'Leary.

Presented at the USCAP annual meeting in Vancouver, March 2004.

This presentation was awarded the **North American Head and Neck Society Award** for best presentation at this meeting. Abstract published in Modern Pathology.

Analysis of head and neck squamous cell carcinoma in young patients using microarray comparative genomic hybridisation. O'Regan E, Sheils OM, Smyth P, Finn SP, Cahill S, Flavin R, Toner M, O'Leary.

Platform presentation at the summer meeting of the Pathological Society of Great Britain and Ireland, Amsterdam, 2004. Abstract published in Journal of Pathology.

Investigation of copy gain and loss in young adults with HNSCC. O'Regan E, Sheils O, Smyth P, Finn S, Cahill S, Flavin R, Toner M, O'Leary J. Platform presentation at the International Association of Oral Pathology, Madrid, 2004. Abstract published in Journal of Oral Pathology and Oral Medicine.

Book Chapters

High Density SNP and cDNA array analysis. Prosser J, Picton S, Howells D, Sherlock J, Martin C, Smyth P, Finn SP, O'Regan E, Sheils O, O'Leary J. In The science and art of laboratory diagnosis. Crocker and Burnett (2nd Edition, John Wiley)

CGH array analysis of human tissues. O'Regan E, Finn SP, Martin C, Smyth P, Sheils O, O'Leary J. In The science and art of laboratory diagnosis. Crocker and Burnett (2nd Edition, John Wiley)

Laser Capture Microdissection. Sheils O, Smyth P, Martin C, O'Regan E, Finn S, O'Leary. In: The science and art of laboratory diagnosis. Crocker and Burnett (2nd Edition, John Wiley) 2005

Summary

While oral and pharyngeal squamous cell carcinoma is classically a disease of older male smokers, this tumour type can occur in young patients with minimal or no exposure to the traditional risk factors. The prognosis remains poor for all of these patients irrespective of age or smoking status, despite improvements in treatment modalities.

The purpose of this project was to use new molecular techniques to enhance the understanding we have of oral and oropharyngeal SCC, paying particular attention to the cancer occurring in young patients (under 40 years old).

The project began by comparing features of oral and pharyngeal SCC in a group of Irish patients with results published from other countries. Concurring with previous reports, this study found a highly significant association between both the age and smoking status of the HNSCC patient and site of tumour occurrence, thus setting the scene for the subsequent molecular analysis performed on tumour samples from a selection of these patients.

To begin the molecular analysis of these young patients, a comprehensive microarray comparative genomic hybridisation analysis was performed. Our experiments on DNA from patients with oral and oropharyngeal SCC have shown that far less genomic aberrations occurred in the young non-smoker samples compared with the old smokers. A gene of interest that is very often deleted in head and neck cancer is p16. Our studies showed that this gene was not deleted in any of the samples from young patients, whereas over 50% of the older cohort showed a p16 deletion using microarray CGH. This finding prompted further research into the p16 tumour suppressor protein. HPV 16 and 18 are well recognised in the cervix as

oncogenic viruses exerting their effect on the cell cycle, resulting in overexpression of p16. Our investigations involved efforts to detect HPV 16/18 in oral and oropharyngeal SCCs, and to correlate the findings with p16 overexpression and p16 deletions. In brief, alteration or overexpression of p16 did not seem to feature strongly in the development of cancer in young nonsmokers. Regarding the HPV prevalence in young nonsmokers, previously thought to be a causative agent, no female non-smoker had evidence of HPV16 integration. HPV prevalence was highest in the oropharyngeal cancers, and p16 immunohistochemistry staining mirrored this. These findings strongly suggest that HPV 16/18 and subsequent p16 alterations do not have significant roles to play in the development of oral/oropharyngeal cancer in young adult nonsmokers.

The application of expression microarrays in this study were in an effort to find a group of genes that distinguish between the SCCs in the old smokers and the SCCs occurring in the young adult nonsmokers. While several genes were found to be differentially expressed between the cohorts (most notably cdc6 and mcm10), overall there were very few. This is not surprising retrospectively, as all the SCCs in all the cohorts are phenotypically (clinical and microscopic descriptions) similar leading us to conclude that tumours that show genomic divergence (i.e CGH microarray studies) may have converging profiles at a transcriptomic level. This thesis has demonstrated at a molecular level some of the differences and the similarities between the typical oral and oropharyngeal SCC occurring in old smokers and the increasing number of young adults particularly nonsmokers that are developing this disease.

Chapter 1

General Introduction

1.1 Oral and oropharyngeal cancer: historical context

Knowing that the addictive effect of tobacco can have both physical and psychological elements, it is perhaps ironic that one of the most famous oral cancer sufferers of all was the father of modern psychoanalysis. In 1923, when he was 67 years old, Freud's cigar smoking began to produce outward indications of the painful cancer that would kill him. Freud developed a growth in his mouth that later turned out to be cancer of the soft palate. An operation was eventually performed - the first of thirty-three operations for cancer of the jaw and oral cavity which he endured during the sixteen remaining years of his life. Freud died of oral cancer in 1939, at the age of eighty-three.

Almost four hundred and fifty people in Ireland were diagnosed with oral and pharyngeal cancer in 2000 (All Ireland Cancer Statistics, 2005). Thirty thousand Americans will be diagnosed with oral or pharyngeal cancer this year (Jemal, 2004). It will cause over 8,000 deaths, killing roughly 1 person per hour, 24 hours per day. Of those 30,000 newly diagnosed individuals, only half will be alive in 5 years. This is a number that has not significantly improved in decades. Although oral cancer has a very low public profile, and lacks public awareness, it may be alarming to hear that the death rate for oral cancer is higher than that of cervical cancer, Hodgkin's disease, cancer of the brain, liver, testes, kidney, or skin cancer (malignant melanoma). In fact, oral cancer is the 8th most common cancer worldwide in men (The WHO World Oral Health Report, 2003). Worldwide about 389,000 new cases occurred in 2000, and these cancers are responsible for some 200,000 deaths each year (World Cancer Report IARC Press Lyon, 2003).

Oral and pharyngeal squamous cell carcinomas (SCC) typically affect men and over 75% of these patients use tobacco, either smoking or chewing it. While overall mortality rates for oral and oropharyngeal SCCs are stable or declining, an alarming rise in the incidence of tongue SCC in young adults has been reported in the US and in Europe (Schantz, 2002; Annertz, 2002; Depue, 1986; Davis, 1987; Franceschi, 1994). Moreover, reports suggest an absence of traditional risk factors in a significant portion of these young people, especially amongst females (Byers, 1975; Carniol, 1982; McGregor, 1983; Newman, 1983; McKenzie, 2000; Llewellyn, 2004). As high as 40% of females under 45 years old with oral and oropharyngeal cancer have been found to be nonsmokers (Llewellyn, 2004), suggesting that factors other than tobacco may be contributing to the development of cancer in this subset.

Potential factors accounting for this increase in incidence of oral SCC in this age group are diverse and could reflect changes in dental care, diet, food processing, nutritional supplementation, sexual habits and their subsequent effect on biological processes. However, there remains no clear evidence to support the significance of any single determinant. The recent trend towards increasing numbers of oral SCC in young people warrants further investigation into whether other risk factors such as genetic predisposition or viral infections and sexual practices can account for the development of oral cancer in these young people (Scully, 2002). Despite the paucity of research into this subsection of oral SCC, there seems evident that the pattern of oral cancer is changing and it can no longer be assumed by health care professionals that the stereotypical patient with oral cancer is an elderly male smoker.

1.2 Anatomy of oral cancer and oropharyngeal cancer

According to the American Joint Committee on Cancer Staging (AJCC manual, 2002) the oral cavity technically extends from the skin-vermilion junction of the lips to the junction of the hard and soft palate above and to the line of the circumvallate papillae below. The oropharynx, in turn, extends from the plane of the hard palate superiorly to the plane of the hyoid bone inferiorly and is continuous with the oral cavity through an opening termed the faucial isthmus, however, the inferior limit is contentious and most clinicians accept the tip of the epiglottis as being the lower border of the oropharynx. (Figure 1)

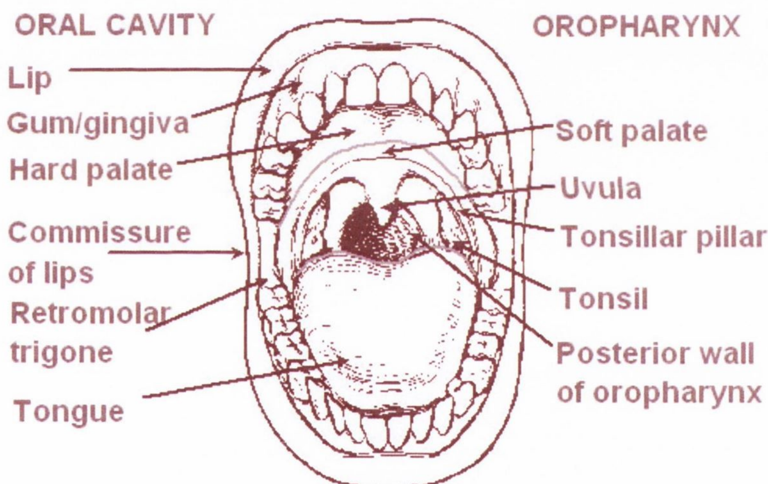


Figure 1.1 Anatomy of oral cavity and oropharynx (Barnes, 2003)

The oral cavity can be divided into eight areas namely labial mucosa, buccal mucosa, lower alveolar ridge, upper alveolar ridge, retromolar trigone, floor of mouth, hard palate and anterior tongue, some of which are outlined in Figure 1.1.

The **buccal mucosa** includes all the mucosal lining of the inner surface of the cheeks. (AJCC, 1992). The soft tissue that clinically circumscribes the teeth is

referred to as the free and attached **gingiva**, while the **alveolar mucos**a covers the maxillary and mandibular ridges in edentulous areas. The **retromolar trigone** consists of a strip of mucosa covering the ascending ramus of the mandible immediately posterior to the last molar tooth and ends at the apex of the tuberosity of the maxilla **(Fig 1.1)**. The **floor of the mouth** is a semicircular region, the lateral and anterior borders of which are formed by the lingual surface of the mandible. The posterior boundaries are composed of the base of the tonsillar pillars and ventral surface of the tongue and it is divided in the midline by the lingual frenum.

The **tongue** is responsible for moving food through the oral cavity and oropharynx, and for assisting in speech, and providing taste and sensation. It has an anterior portion that lies in the oral cavity and a posterior component that lies in the oropharynx. The anterior two-thirds extend back to the circumvallate papillae, and the base of the tongue extends from the circumvallate papillae to the junction with the base of the epiglottis. Seventy-five per cent of all lingual SCCs arise on the anterior portion of the tongue (Decroix, 1981). The **palate** forms the roof of the mouth, and is divided into the hard palate in the oral cavity and the soft palate posteriorly in the oropharynx.

Location	Frequency (%)
Lip	10.4
Tongue	30.1
Floor of Mouth	27.6
Gingiva	16.7
Palate	3.2
Buccal mucosa	12.0

Table 1.1 Frequency of oral SCC according to site (Strong, 1981)

The **oropharynx** has six main areas: base of tongue, inferior surface of the soft palate and uvula, anterior and posterior fauces, the glossotonsillar sulci, the pharyngeal tonsils and the pharyngeal walls. The **base of tongue**, described previously, extends posteriorly to the base of the epiglottis and anteriorly to the circumvallate papillae.

The inferior surface of the **soft palate** is continuous with the tonsillar area, which is bordered by the anterior and posterior faucial pillars. The superior border of the **pharyngeal wall** is at the level of the soft palate while the lower border is at the level of the vallecula.

Location	Frequency (%)
Tonsil	46
Base of tongue	34
Soft palate-uvula	10
Pharyngeal wall	10

Table 1.2 Oropharyngeal SCC according to site (Fletcher GH, 1967)

Compared with oral cavity malignancies, those of the oropharynx tend to be more aggressive and less well differentiated, with more frequent cervical metastases. This difference can, in part, be explained on the basis that the oropharynx is less receptive for self-inspection and early detection. (Barnes, 2002)

1.2.1 Site distribution in young adults

Early reports of oral and pharyngeal SCC in young people (Depue, 1986; McGregor, 1983; Sankaranarayanan, 1989; MacFarlane, 1987), noted that the tongue was by far the most common site of SCC in those less than 40 years of age.

1.3 Histology of squamous cell carcinoma of the oral cavity and pharynx

A squamous cell carcinoma is a malignant epithelial tumour with squamous differentiation characterised by the formation of keratin or the presence of intercellular bridges or both. The diagnosis of oral or pharyngeal squamous cell

carcinoma is based on histological examination of haematoxylin and eosin (H+E)-stained biopsy specimen. Tumour characteristics are graded as being well, moderately or poorly differentiated. Well differentiated lesions, also referred to as 'low-grade', show more keratin production and little cellular pleomorphism, whereas the poorly differentiated type ('high-grade') demonstrate little or no keratin formation and more cellular pleomorphism. The moderately differentiated have intermediate levels of these characteristics (Barnes, 2002). Bryne et al. introduced a multifactorial malignancy grading of only the deep invasive margins of oral SCC, which proved to be of high prognostic value (Bryne, 1991)

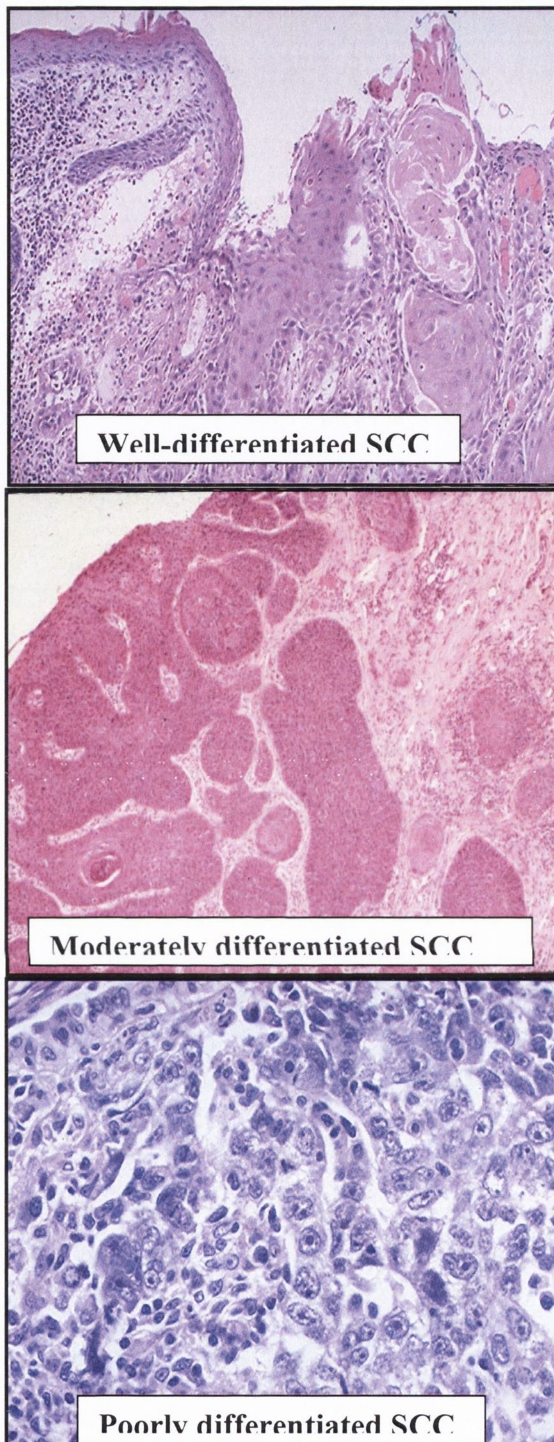


Figure 1.2 Grades of squamous cell carcinoma

This classification system has been used since 1920, but it has been criticised because it relies on subjective interpretation. Moreover it does not predict future

clinical behaviour, specifically, the likelihood of tumour recurrence or metastasis, and, therefore, the patient's prognosis (Burkhardt, 1993).

1.3.1 Tumour differentiation in young adults

Byers (1975) reported that cancer in young adults tends to be more frequently anaplastic resulting in a more aggressive behaviour and poor prognosis. More recently however, a number of studies concentrating on the development of oral and pharyngeal SCC in young adults have noted a predominance of well-differentiated tumours in the younger population (Kuriakose, 1992, Ramirez-Amador, 1995).

1.4 Prognosis

Patients with oral SCC have shown no statistically significant improvement in their survival probability since the 1960s (Parker, 1996). The major determinants of survival include:

- Patient factors at diagnosis: Age, gender, socio-economic status, nutritional status, intercurrent disease and continuing presence of risk factors
- Tumour factors: site, size, differentiation, extent of regional and distant metastases
- Management factors: treatment modality, quality and extent of supportive care

The overall five-year prognosis for oral cancer at all stages remains at less than 50% (Silverman, 1990). This is widely attributed to the late stages (III and IV) at which

most oral and pharyngeal SCCs are diagnosed. The relationship of oral tumour TNM to five-year survival is tabulated below (Table 1.3)

TNM Stage	5 year survival rate
Stage 1	85%
Stage 11	66%
Stage 111	41%
Stage IV	9%

Table 1.3 Relationship of oral cancer TNM stage to 5-year survival rate (Silverman, 1990)

Also contributing to the unfavourable prognosis for oral and oropharyngeal carcinoma patients is the frequent development of second primaries within the upper aerodigestive tract. The risk of development of a second primary carcinoma is estimated to be between 3.7% and 5% per year (Day, 1992). The development of new primary tumours is attributed to field cancerisation. First described by Slaughter (1953), this concept suggests that all epithelial surfaces exposed to tobacco are at risk of developing malignancies.

1.4.1 Prognosis in young adults

Existing published data suggest that there is a trend towards better survival in early disease (Friedlander, 1998). However, there is little agreement in the literature regarding the outcome of younger patients diagnosed with SCC of the oral cavity

and oropharynx compared with the outcome of older patients. Pitman (2000) showed that oral and pharyngeal cancer in younger patients typically presents at an earlier stage and therefore is associated with a higher 5-year survival rate than is the case with older adults. In the recent analysis of the cancer surveillance database in the U.S, Schantz (2002) found that 56% of the young adults with head and neck cancer presented with localised disease, compared with 37% and 43% localised disease in the 40 to 64 year olds and those over 65 years old, respectively. Likewise, Lacey (2000), found young adults have a more favourable outcome. On the contrary, Manuel (2003) argued that it was pathological node status, type of primary surgery, surgical clearance and selection of appropriate treatment that influenced survival, irrespective of age.

1.5 Aetiology

Analytical epidemiology has provided considerable insight into factors that contribute to the development of head and neck SCC. In order to judge the significance of these factors, a framework based on a modification of Koch Postulates is provided by Wynder and Day (1961), and are as follows:

(i) The factor has to increase the risk of cancer

(ii) After its removal or reduction in a given population, the rate of cancer should decline after a suitable latent period.

1.5.1 Tobacco

Ernest Wynder (1956), the father of tobacco control, published the first report demonstrating the causal relationship between tobacco and head and neck cancer, when he reported a 10-fold increase in the risk of developing head and neck cancer in smokers. Since then tobacco smoking has been the most intensively investigated environmental cause of cancer (World Cancer Report, 2003). Among the developed countries, tobacco is considered responsible for 24 percent of all male deaths and 7 percent of female deaths, but this figure rises to over 40 percent for males in some of these countries (Johnson NW, 1999). The geographical variation regarding oral and pharyngeal cancer mortality rates has historically been linked to heavy alcohol and tobacco use in these communities, with several European countries, France, Hungary and the Czech Republic ranking especially high. (Banoczy, 2001). Much of the world tobacco is consumed without combustion and smokeless tobacco use is prevalent in the Indian region, Northern Europe and South America. It is placed in contact with mucous membranes, through which nicotine is absorbed providing the pharmacological effect.

Carcinogenic agents in tobacco smoke include N-Nitrosamines, arsenic, and polycyclic hydrocarbons, many of them requiring metabolic activation in order to exert an effect in both early and late steps in the process of carcinogenesis. Once the reactive forms of these procarcinogens are generated, they may become covalently bound to DNA in various tissues leading to miscoding events, and thus to mutations. Individual risk may be affected by various genotypic polymorphisms of cytochrome P-450 (CYP1A1) and glutathione-S-transferase (GSTM1) which may result in loss

of one's ability to detoxify carcinogens and reactive oxygen species, leading to a higher risk of development of HNSCC, even with a low dose of cigarette smoking (Katoh, 1999).

Oral and pharyngeal cancer risk in relation to tobacco smoking follows the basic principles of chemical carcinogenesis: risk is determined by the dose of carcinogen, the duration of administration and the intensity of exposure. The fact that most of the individuals who develop oral and oropharyngeal SCC are older has been attributed to prolonged exposure and the cumulative effects of tobacco and alcohol, the primary risk factors (Boyle, 1992). In males that smoke but do not drink alcohol, the risk of developing cancer of the oral cavity is about double that for males that do not smoke or drink alcohol. Elevations of ten-fold or more are evident for cancer of the larynx and five-fold or more for oesophageal cancer (World Cancer Report 2003). It must be noted, however, that a small number of patients that never smoke still develop oral and oropharyngeal cancer.

1.5.2 Alcohol

Beverages containing alcohol have been an integral part of most human societies and cultures since ancient times. On a global scale, the consumption of alcoholic beverages by adults as calculated from official figures is equivalent to 4 litres of alcohol per year (WHO, 1999). However, unofficial consumption is estimated to account for an additional amount corresponding to 20-100% of the official figures depending on the country. There is strong regional variability in consumption levels with a minimum in Southern Asia and Africa, while central and southern Europeans consume the highest levels. Like tobacco use, there has been a decrease in alcohol

consumption in more developed countries over recent years, and an increase in consumption in less developed countries.

Although alcohol may act as an independent risk factor, smokers tend to consume larger quantities of alcohol than do nonsmokers (Boyle, 1995). Multiple types of alcoholic beverages have been associated with oral cancer, indicating the primary role of ethanol or its primary metabolite acetaldehyde. In addition to a systemic effect, it has been suggested that ethanol can act directly to damage the mucosa of the upper gastrointestinal tract. It has also been proposed that alcohol might act as co carcinogen by

(1) Increasing the permeability of the mucosa to other carcinogens
(2) Induction of enzymes involved in the metabolism of tobacco smoke carcinogens both systemically and locally (Blot, 1999)

The effect of type of beverage on the risk of developing cancer of the oral cavity and pharynx remains controversial in the literature. While some studies have found an elevated risk for consumption of spirits (Castellsague, 2003), others have shown an excess risk for wine (Barra, 1990).

1.5.2.1 Tobacco and alcohol in young adults

Oral cancer is rare in young adults, and little knowledge of the specific risk factors in this cohort exists. Although case control studies have implicated tobacco and /or alcohol as the major risk factors for oral and pharyngeal cancer in adults of any age (Blot, 1988; IARC World Cancer Report, 2003), recent evidence suggests an absence of traditional risk factors in a significant proportion of younger patients, especially females (Kuriakose, 1992; Mackenzie, 2000; Newman, 1983; Carniol;

1982). Moreover, the time span for carcinogens such as tobacco and alcohol to exert a detrimental effect in these younger patients is relatively short. A comprehensive study by Llewellyn (2004) found that although the exposure to risk factors in these young patients were similar to older patients, it is questionable whether exposure to factors of risk would be long enough in duration to become aetiological agents of oral cancer. They also studied cannabis, and alcohol consumption and found that although elevated risks were found in young patients in relation to cannabis and/or alcohol use, these were not significant.

1.5.3 Diet and nutrition

A number of macro and micronutrient deficiencies or imbalances have long been implicated as factors contributing to the development of cancer of the oral cavity and pharynx. Increased risks accompanying high intakes of dairy foods, fats (Toporcov, 2004), cholesterol (McLaughlin, 1988), starchy food (Martinez, 1969) and maize (Franceschi, 1990) have been reported, while it is generally recognised that fruits and most likely vegetables, are likely to provide protection against oral and pharyngeal cancer (Bosetti, 2003).

Although no general biological mechanism has been fully established as an explanatory factor for the modulating effects of fat on carcinogenesis, there are convincing indications that qualitative and quantitative differences in dietary lipids can significantly alter the membrane fatty acid composition of both normal and neoplastic cells, thus modifying the physical-chemical environment of hormone receptors and/or enzymes of the tumour cells (Woutersen, 1999)

Growing evidence exists that a Mediterranean diet has a beneficial influence on health, and that such a diet has a favourable effect on oral and pharyngeal cancer risk, with strong suggestions that olive oil may convey a protective effect against these cancers (Bosetti C 2003, Franceschi S, 1999). With respect to micronutrients there is evidence that Vitamin C is associated inversely with the risk of oral cancer but convincing data is sparse regarding the protective effect of other micronutrients such as Vitamin E, thiamine, Vitamin B_6, folic acid, niacin, potassium and iron, (Negri, 2000, Petridou, 2002)

1.5.3.1 Diet and nutrition in young adults

There are few dietary studies that have focused on young adults exclusively or have analysed data according to age. A study by Llewellyn (2004), however, found a highly significant reduction in risk among females consuming three or more portions of fresh fruit and vegetables daily, suggesting that in younger patients, childhood diet may be a more pertinent factor to investigate than diet in adulthood.

1.5.4 Marijuana

Marijuana is the second most commonly smoked substance in the U.S, with as many as 31% of the U.S population 12 years or older having used it (Johnston LD, 1999). One must bear in mind that recreational users of marijuana often also enjoy alcohol and tobacco, and tobacco usually forms part of the marijuana smoking mix. It is thus impossible to discern an independent risk for the smoking of cannabis products

themselves. Marijuana smoke, however, does contain even higher concentrations of irritants and carcinogens than tobacco smoke and there is growing evidence that marijuana smoke may be a dose dependant risk factor for development of oral and pharyngeal cancer (Zhang, 1999; Schantz, 2002; Donald, 1986). However, a recent large, population- based study in Washington State by Rosenblatt (2004) found no association between marijuana use and oral cancer risk.

1.5.4.1 Marijuana and young adults

There is little in the literature implicating marijuana as a causative factor in oral and oropharyngeal cancer in young patients. Schantz (2002) reported a sharp increasing trend in the incidence of tongue cancer attributed to those born between 1938 and 1947, and suggested that marijuana was a contributing factor considering the sharp increase in prevalence of marijuana use in the U.S in teenagers and young adults in the mid to late 1960s, (i.e. those born between 1938 and 1947).

1.5.5 Genetic Susceptibility

Factors other than tobacco, acting either in concert with tobacco or alone, may contribute to cancer development. These may include host factors that control the metabolism of carcinogens, traits that relate to gender and race, genetic determinants found in individuals with various cancer family syndromes, as well as syndromes associated with DNA repair deficiency.

Most carcinogenic compounds within tobacco will not act directly on host DNA. Rather, these compounds must be metabolised by enzymes to intermediates. The

most extensively studied xenobiotic enzyme superfamily is the p450 system. Its role in laryngeal cancer was studied by Brandenberg (1978) who found that individual risk may be affected by various genotypic polymorphisms of cytochrome P-450 (CYP1A1) and glutathione-S-transferase (GSTM1) which may result in loss of one's ability to detoxify carcinogens and reactive oxygen species, resulting in a higher risk of development of HNSCC, even with a low dose of cigarette smoking.
The cancer family syndromes, which are characterised by both a high incidence of cancer, as well as specific types of cancer within a family unit, are relevant in the development of head and neck cancers. Conditions carrying increased risk of HNSCC include epithelial differentiation disorders, such as dyskeratosis congenita, and DNA repair deficiency syndromes such as Bloom's syndrome, Fanconi anaemia, ataxia telangiectasia and xeroderma pigmentosum (Foulkes, 1996). There is also now increasing epidemiological evidence from case control studies of head and neck cancer patients that a family history of head and neck cancer is a risk factor for this disease (Foulkes, 1995)

1.5.5.1 Genetic susceptibility in young adults

Generally, onset of cancer at an early age is thought to be an indicator of hereditary cancer (Lynch, 1995), and studies have shown higher familial risks for same site cancer in first-degree relatives where the index cancer was diagnosed at a young age (Goldgar, 1994). Mork *et al* conducted a population-based study to investigate the familial risk of developing head and neck cancer before the age of 45 (Mork, 1999), and found that young females with head and neck cancer were highly likely to have a first degree relative with cancer of the upper aerodigestive tract. Unfortunately,

due to lack of adjustment for other risk factors, it was not possible to discern whether the raised risk was due entirely to tobacco/alcohol consumption or to an interaction between genetic susceptibility and other risk factors.

1.5.6 Human Papillomavirus

The mechanism of HPV carcinogenesis was first identified in cervical cancer. Worldwide, greater than 95% of cervical cancers are related to HPV infection, with types 16 and 18 being implicated in over 70% of cases (Munoz, 2003).
High-risk HPVs are known to be tumorigenic in human epithelial tissues. Two viral oncoproteins of high risk HPVs, E6 and E7, promote tumour progression by inactivating the p53 gene and retinoblastoma tumour suppressor gene product, respectively (ZurHausen, 1991) Accordingly, these viral oncoproteins are capable of transforming primary human keratinocytes from either genital or upper respiratory tract epithelia and disrupting cell-cycle regulatory pathways in the genetic progression of HNSCC.

The link between oral squamous cell carcinoma and HPV seems logical, given the viral propensity for epithelial cell involvements. Both epidemiologic and molecular data suggest that HPV infection of the upper airway may promote head and neck malignancies (Gillison, 1999; Franceschi, 1996). The majority of studies of head and neck lesions have focused on HPV 16 and 18, since these are known to confer high risk in cervical cancer. Reports of the prevalence of HPV prevalence vary markedly, with the percentage of tumours containing HPV varying from 14% to 79% (Franceschi, 1996). Such variance may reflect several factors, including investigator

bias, geographic location of the population studied, number of cases, and the methods employed to detect the viral genome.

Recent studies support an aetiologic role for HPV in at least a subset of head and neck cancers, especially in tonsillar carcinoma (Gillison, 2000). Of course, identification of a virus does not prove a causal relationship, since it may have been activated by the disease rather than the converse. Thus, HPV-DNA detection is insufficient evidence for a causal role. Viral integration into the host genome, would be evidence against HPV being merely a secondary invader, and would strongly suggest a causal role in carcinogenesis. Identification of E6 and E7, the two major viral oncogenes expressed in tumour tissue that are known to stimulate cell proliferation and interfere with tumour suppressor proteins would further substantiate an oncogenic role for HPV in oral and oropharyngeal cancer. Just a few studies have demonstrated HPV E6 or E7 gene expression in OSCC and tonsillar carcinomas (Andle, 1998; Wang, 1998). Conversely, because HPV infection can be focal and transient, the lack of detection does not rule out viral involvement.

The means by which HPV is transmitted to the upper airway is unclear. Although oral HPV infection is rare in newborn children of infected mothers, infections do increase after onset of sexual activity (Kellokoski, 1992). While HPV presence in head and neck cancers has not yet been strongly linked to specific sexual practices such as oral sex, there are reports of links between HPV positivity and number of sexual partners in more than one study (Schwartz, 1998). A recent comprehensive review of the literature on HPV and its role in head and neck cancer by Ha (2004)

concluded that while HPV may be a contributing factor in a subset of oral malignancies, it is not in any way as significant as it is in cervical cancers.

1.5.6.1 HPV in young adults

Studies on HPV and its role in oral and pharyngeal cancer have rarely focused on age. Cruz (1996), found a significant association between HPV presence and age, with patients under 60 years of age showing a far higher prevalence of HPV, while Van Rensburg (1996) in their study of under 40 year olds concluded that HPV was not important in the development of oral SCC.

1.5.7 Occupational

The first reports of associations between risk of cancer and employment in particular appeared during the 18th century (Pott, 1775), since then around 25 chemicals, groups of chemicals or mixtures for which exposures are mostly occupational, have been established as human carcinogens (World Cancer Report, 2003).

The lack of well-established occupational risk factors for head and neck cancer may be as informative as their existence. This suggests the requirement of more chronic exposures as a determinant of disease. The median duration for tobacco use among head and neck cancer patients is about 40 years (Blot, 1988). Employment at any particular trade rarely extends over such a prolonged interval. Notwithstanding this, several occupational exposures have been related to laryngeal carcinoma, including exposure to nickel, sulphuric acid and the most extensively investigated of all, asbestos (Maier, 1991).

1.6 Molecular Pathology of oral and pharyngeal cancer

1.6.1 Genetic instability

All methods of analysis have indicated that, although oral and pharyngeal squamous cell carcinomas have very complex karyotypes, these karyotypes also show regions of consistent gain and loss, therefore not just the result of random, widespread, genomic damage. A variety of genetic alterations and aberrant karyotypes in tumours of the oral cavity and pharynx have been characterised extensively. It remains unclear how these genetic alterations interfere with various steps in tumour progression. A high level of genomic imbalance found in head and neck cancers indicates that chromosome aberrations represent a very important mechanism of genetic instability. Molecular studies have shown that a number of chromosomal regions are consistently altered, such as gains on 3q, 5p, 8q, 11q13, and losses on 3p and 9p (Hermsen, 1997; Gebhart, 2003; Wolff, 1998).

1.6.1.1 Chromosome 3

Overall, the most frequent DNA copy number changes are detected on chromosome 3, (Maestro, 1993; Latif, 1992). Possible tumour suppressor genes (TSG) on 3p include the FHIT (Fragile Histidine triad) gene located within 3p14.2. Hypothetically, the initial event is triggered by extrinsic factors like carcinogen exposure or a viral infection, which is enhanced by the alteration of an oncogene or

TSG. Bearing this in mind, it is interesting that the FHIT gene contains a fragile site that coincides with a spontaneous HPV16 integration site (Wilke, 1996).

1.6.1.2 Chromosome 7

Consistent gains of 7p may relate to the frequent overexpression of the epidermal growth factor receptor (EGFR) and activation of this gene is thought to be an early event in head and neck carcinogenesis (Grandis, 1993). Over expression of c-erb-B2, which is an EGFR-like oncogene located on chromosome 17, has been observed in 75% of oral cancer patients and correlated to shorter survival periods (Xia, 1997).

1.6.1.3 Chromosome 11

CGH studies of oral tumours have revealed frequent amplification in 11q13 region, which has been correlated with poor prognosis and also associated with aggressive histologic appearance, late stage tumours and recurrence (Bockmuhl, 1996, Akervall, 1995). Several candidate genes are present in this area, including Cyclin D1 and Int-2 (FGF3).

1.6.2 Oncogenes

A gene whose protein products have been found to be important for normal cell growth signalling and whose over-expression or mutation leads to unchecked cell growth and tumorigenesis is defined as an oncogene. They are broadly represented by growth factors, or growth factor receptors (FGF3, EGFR, c-erbB2), intracellular

signal transducers (ras, raf), transcription factors (myc, jun, fos), cell cycle regulators (cyclin D1), and those involved in apoptosis (bcl-2). Mutations, chromosomal translocation, gene amplification or retroviral insertion can result in activation of a proto-oncogene allowing it to function as an oncogene.

1.6.2.1 cyclin D1

Cyclin D1 gene regulates initiation of DNA synthesis and G1/S transition of cells. Its deregulation has been reported in 25%-75% of oral cancers and as previously mentioned it has been associated with aggressive tumour behaviour and poor prognosis (Akervall, 1995)

1.6.2.2 Myc

The c-myc proto-oncogene encodes a transcription factor for genes involved in growth and differentiation. C-myc expression normally declines with cell cycle progression and is shut off completely following full differentiation and growth arrest. It is also thought to play a role in apoptosis. Oncogenic activation leads to up-regulation of MYC, which has been detected in many human cancers including oral and pharyngeal cancers (Saranath, 1989)

1.6.3 Tumour Suppressor Genes

Tumour suppressor genes (TSG) serve as transducers of negative growth signals. They are involved in cell cycle regulation including cell cycle arrest and apoptosis. TSGs are inactivated by several mechanisms including point mutations, deletions and binding with cellular and viral proteins.

1.6.3.1 p16

The prime candidates for the frequent deletion of 9p in oral cancer are the p16 gene and its homologues p15, and p14 that are overlapping genes located within 9p21. In combination with the cyclin D1 gene at 11q13 and the retinoblastoma gene at 13q14, p16 takes part in the same cell cycle regulatory pathway in which the kinase inhibitor p16 and the pRB protein act as tumour suppressors and cyclin D1 as an oncogene. Disruption of the $p16^{INK4A}$-Rb cell cycle regulatory pathway results in unrestricted proliferation eventually contributing to the malignant transformation of cells. Not surprisingly then, loss of $p16^{INK4A}$ function represents a common pathway to tumourigenesis. $p16^{INK4A}$ can be inactivated by a variety of different mechanisms including mutation, altered splicing, homozygous deletion and promoter hypermethylation. The majority of these genetic alterations result in loss of $p16^{INK4A}$ protein expression. However, overexpression of $p16^{INK4A}$ protein in tumours has also been described, and has been shown to correlate with inactivation of Rb function through disruption of a negative feed back mechanism (Tam, 1996). This suggests that $p16^{INK4A}$ overexpression might be an indicator of Rb inactivation.

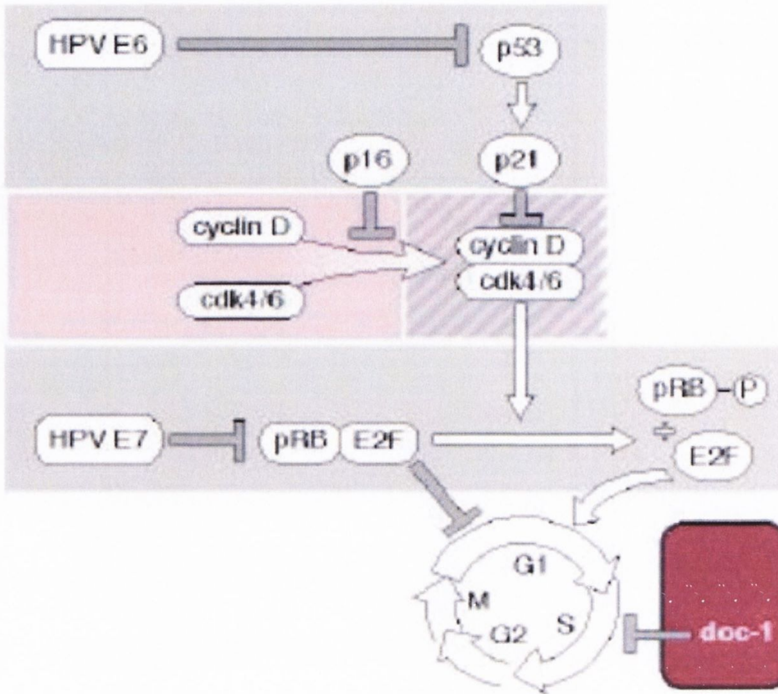


Figure 1.3 Disruption of the Rb cell cycle regulatory pathway

Disruption of the p16^{INK4A}-Rb cell cycle regulatory pathway by HPV E7 results in unrestricted proliferation eventually contributing to the malignant transformation of cells.

A marked overexpression of p16^{INK4A} is observed in precancerous and malignant cervical lesions. This phenomenon is believed to be a consequence of functional inactivation of Rb by HPV E7 protein (Sakaguuchi, 1996). It has been suggested that p16^{INK4A} may be directly induced by the transcription factor E2F that is released from pRb after binding of HPV E7. In concordance with this a number of recent studies have reported a significant increase in p16^{INK4A} protein expression in a number of high-risk HPV associated cancers including HPV associated oral cancers

(Fregonesi, 2003). Cyclin D1 (CCND1) and p16 are frequently simultaneously altered in the deregulation of the G1 checkpoint, and CGH studies have shown simultaneous alterations on 9p, 11q and 13q in head and neck cancers (Bockmuhl, 1996). Inactivation of p16 ranging from 25% to 83% has been reported in head and neck cancer (EL-Naggar, 1997, Gonzalez, 1997).

$p14^{ARF}$ is a member of the p53 tumor suppressor pathway. Proteosome mediated degradation of p53 is controlled by the cellular oncogene mdm2 via a negative feedback mechanism. $p14^{ARF}$ binds directly to mdm2 causing its degradation and the concurrent stabilisation and accumulation of p53. $p14^{ARF}$ tumor suppressor function results in both G1 arrest and G2/M arrest (Scott, 1998).

1.6.3.1.1 p16 and young adults

In 2003 EL-Mofty published his investigations of p16 protein expression in young adults with oral and pharyngeal cancer, and found that in tonsillar tumours, where HPV was also detected, p16 immunohistochemical staining was strong and diffuse, whereas oral tumours lacked HPV and lacked p16 protein expression (El-Mofty, 2003).

1.6.3.2 p53

The p53 protein was identified in 1979 (Lane, 1979; Linzer, 1979), and was initially considered to be an oncogene because it was often over expressed in tumour cells. However, after sequencing p53 from a normal cell it became evident that the p53

protein expressed in most tumours was mutated, and that these were loss-of-function mutations, suggesting that p53 was a tumour suppressor gene, like Rb. Mutation of this gene, located on chromosome 17, is the most reported genetic aberration in human cancer. Normal p53 functions as a cell cycle checkpoint, and modulates G1-arrest, DNA repair and apoptosis, therefore cells with mutated p53 are predisposed to further genetic alterations because of inadequate DNA repair, escape from apoptosis and manifestation of additional DNA damage in subsequent cell cycles.

In head and neck cancer, there is an association between carcinogenic factors, such as alcohol and tobacco use, and p53 overexpression or p53 mutation (Field, 1991; Field, 1992). The clinical relevance of p53 overexpression in head and neck cancer has been subject to debate, with studies reporting improved survival in patients with p53 overexpression (Sauter, 1992), while a lack of correlation between p53 overexpression and clinicopathological parameters or survival has been reported by other groups (Field, 1993). With regard to mutation of p53, again there are mixed reports of a correlation with clinical parameters with Bradford (1997) finding an association between p53 mutation and poor prognosis in laryngeal cancer patients, while Nylander (1995) did not find any relationship between p53 mutation and survival in head and neck cancer patients. A negative p53 regulator protein, MDM2, has been detected in as many as 40% of oral tumours, providing an alternative mechanism of p53 protein dysregulation in some cancers (Matsumura, 1996).

1.6.3.2.1 p53 and young adults

In 1997, Sorensen reported results of both phenotypic and genotypic analysis of the p53 gene in patients less than 40 years old with oral cancers (Sorensen, 1997). They

found a complete lack of p53 mutations in young nonsmokers and non-drinkers, suggesting that a different mechanism responsible for oral carcinogenesis may be occurring in young nonsmokers developing oral cancer. Regezi (1999), however, compared p53, p21, Rb, and MDM2 protein expression in under 35 year olds and over 75 year olds and detected no significant differences in expression in these tumours between the young group and the old.

1.7 Comparative Genomic Hybridisation

Genomic instability reflects the propensity and susceptibility of the genome to acquire multiple alterations. In addition to fundamental cytogenetic studies to detect chromosomal rearrangements in oral and pharyngeal cancer, other advanced techniques such as comparative genomic hybridisation (CGH) and fluorescence in situ hybridisation (FISH) are also frequently utilised to detect patterns of chromosomal imbalances and gross chromosomal regions involved in structural rearrangements.

CGH involves using disease or tumour specific DNA labelled with a green fluorochrome, mixed (1:1) with normal (diploid) DNA labelled with a red fluorochrome. This mixture is hybridised to normal metaphase human preparations on a glass slide.

The labelled DNA fragments compete for hybridisation to their locus of origin on the metaphase spread of chromosomes, and hybridisation of tumour DNA is represented by green fluorescence, while hybridisation of normal DNA is represented by red fluorescence. The relative amounts of tumour and reference DNA bound at any given locus are dependent on the relative abundance of those

sequences in the two DNA samples. The resulting green: red fluorescence ratio is the quantitative representation of loss or gain of genetic material. In short, gene copy number gain (amplification) produces an elevated green to red ratio, and copy number loss (deletion) at a specific locus result in a reduced green to red ratio.

All CGH work on oral cancer so far in the literature has been directed at the typical oral cancer patient i.e. the 60-70 year-old male smoker, and to date there are no references to array CGH studies of young patients with oral cancer (Hermsen, 1997).

1.8 Microarrays

In 1995, Schena (1995) developed the microarray, where cDNA clones of gene specific hybridisation targets from plants were used to quantitatively measure expression of the corresponding plant genes. The success of this novel scientific tool led to use of microarrays for human genome analysis, thus permitting placement of a representation of the entire human genome on a single slide.

Microarrays are orderly arrangements of individual nucleic acid samples, which are immobilised in the form of a grid on a solid surface such as glass, chromium, silicon etc. Microarray analysis is now a foundational technology allowing the analysis of DNA sequence variation, gene expression, protein levels, tissues, and cells in an extensive parallel format.

1.9 Genomic DNA arrays

Recently the principle of micro-array based CGH was introduced (Pinkel, 1998), and this combination of CGH on a chip has avoided some of the limitations of

conventional CGH, thus providing us with a platform for aberration detection that has improved sensitivity, improved resolution, better reproducibility and higher throughput.

1.10 Expression arrays

The ability to monitor gene expression at the transcript level has become far easier with the development of microarrays, which allow monitoring of gene expression for tens of thousands of genes in parallel. Expression profiles from microarrays experiments are rapidly adding to our understanding of the complex functions and interactions of these genes. Recently, global expression profiling using microarrays in oral cancer has been used to explore, classify, and predict the biological processes underlying this malignancy (Kuo 2002, 2003)

1.11 Aims and objectives

The overall aim of this study was to identify differences between the increasing number of young adults that develop oral and pharyngeal cancer and the typical oral and pharyngeal cancer patients (i.e. older adult smokers).

The specific objectives were

1. To examine the variation in site of occurrence, smoking status, alcohol consumption, haematologic status and tumour stage and grade and nodal status across age groups in Irish patients with oral and pharyngeal cancer

2. To investigate the prevalence of human papillomavirus, and the expression of p16 in young adults with oral and pharyngeal cancer and compare this to older adults

3. To investigate global DNA copy gain and loss in oral and pharyngeal cancer detected in young adults, and to compare this to the profile in older adults

4. To analyse the expression profile of oral and pharyngeal cancer in young adults in order to evaluate pathobiology of the disease in this subset of cancers with a view to aiding earlier identification.

1.12 References

All Ireland Cancer Statistics Second Report 1998-2000. 24th March 2005. http://www.allirelandnci.org/publications/report2/allireland1998-2000.pdf

Ahmedin J, Ram C, Murray T, Ghafoor A, Samuels A, Ward E, Feuer E, Thun M. American Cancer Society Surveillance Report 2004. CA Cancer J Clin 2004; 54:8-29

Akervall JA, Jin Y, Wennerberg JP, Zatterstorm UK, Kjellen E, Mertens F. Chromosomal abnormalities involving 11q13 are associated with poor prognosis in patients with HNSCC. Cancer 1995; 76:853-9

Almadori G, Paludetti G, Cerullo F, Ottaviani F, D'Alatri L. Marijuana smoking as a possible cause of tongue carcinoma in young patients. Journal of Laryngology and Otology 1990; 104: 896-899.

American Joint Committee on Cancer. Manual for Staging of cancer, 4th ed. Philadelphia: JB Lippincott, 1992:28.

American Joint Committee on Cancer: Lip and oral cavity AJCC Cancer Staging Manual. 6th Ed. New York, NY: Springer, 2002, pp 23-32. .

Andle T, Kahn T, Pfuhl A, Nicola T, Erber R, Conradt C. Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinomas lacking retinoblastoma cell cycle control. Cancer Res 1998; 58:5-13

Annertz K, Anderson H, Biorklund A. Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002; 101:95-9.

Banoczy J, Gintner Z, Dombi C. Tobacco use and oral leukoplakia. Journal of Dental Education 2001; 4: 322-327.

Barnes L. Surgical pathology of the head and neck. Marcel Dekker; 2nd edition, 2001.

Blot WJ, Mc Laughlin JK, Winn DM, Austin DF, Greenberg DS, Preston-Martin S, Bernstein L, Schoenberg JB, Stemhagen A, Fraumeni JF Jr. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Res 48:3282-3287, 1988.

Blot WJ. Invited commentary: more evidence of increased risks of cancer among alcohol drinkers. Am J Epidemiol 1999; 150:1138- 40.

Bockmuhl U, Schwendel A, Dietel M, Petersen I. Distinct patterns of chromosomal alterations in high and low-grade head and neck squamous cell carcinomas. Cancer Research 1996; 56:5325-5329

Bosetti C, Gallus S, Trichopoulou A, Talamini R, Franceschi S, Negri E, La Vecchia C. Influence of the Mediterranean diet on the risk of cancers of the upper aerodigestive tract. Cancer Epidemiology, Biomarkers and Prevention. 2003; 12: 1091-1094

Boyle P, Macfarlane GJ, Zheng T, Maisonneuve P, Evstifeeva T, Scully C. Recent advances in epidemiology of head and neck cancer. Curr Opin Oncol. 1992 Jun; 4(3): 471-7.

Boyle P, Macfarlane GJ, Blot WJ, Chiesa F, Lefebvre JL, Azul AM, de Vries N, Scully C. European School of Oncology Advisory report to the European Commission for the Europe Against Cancer Programme: oral carcinogenesis in Europe. Eur J Cancer B Oral Oncol. 1995; 31B(2): 75-85.

Brandenberg JH, Kellerman G. Aryl hydrocarbon hydroxylase inducibility in laryngeal carcinoma. Arch Otolaryngol 1978; 49:187-192.

Bradford CR, Zhu S, Poore J for the Department of Veterans Affairs Laryngeal Cancer Cooperative Group. P53 mutation as a prognostic marker in advanced laryngeal carcinoma. Arch Otolaryngol Head Neck Surg. 1997; 123:605-609

Bryne M. Prognostic value of various molecular and cellular features in oral squamous cell carcinomas: a review. J Oral Pathol and Med. 1991:20; 413-420

Burkhardt A. Advanced methods in the evaluation of premalignant lesions and carcinomas of the oral mucosa. J Oral Pathol 14:751-778, 1985

Byers RM. Squamous cell carcinoma of the oral tongue in patients less than thirty years of age. Am J Surg. 1975 Oct; 130(4): 475-8.

Carniol PJ, Fried MP. Head and neck carcinoma in patients under 40 years of age. Ann Otol Rhinol Laryngol. 1982; 91: 152-5.

Castellsague X, Quintana MJ, Martinez MC, Nieto A, Sanchez MJ, Juan A, Monner A, Carrera M, Agudo A, Quer M, Munoz N, Herrero R, Franceschi S, Bosch X. The role of tobacco and type of alcoholic beverage in oral carcinogenesis. Int J Cancer 2004; 108:741-749.

Cruz IB, Snijders PJ, Steenbergen RD, Meijer CJ, Snow GB, Walboomers JM, van der Waal I. Age-dependence of human papillomavirus DNA presence in oral squamous cell carcinomas. Eur J Cancer B Oral Oncol. 1996; 32B(1): 55-62.

Davis S, Severson RK. Increasing incidence of cancer of the tongue in the United States among young adults. Lancet 1987; 2:910-11

Day G, Blot W. Second primary tumours in patients with oral cancer. Cancer 70:14-19, 1992.'

Decroix Y, Ghossein NA. Experience of the Curie Institute in treatment of cancer of the mobile tongue. Cancer 47:496-502, 1981

Depue RH. Rising mortality from cancer of the tongue in young white males. N Engl J Med 1986; 315:647

Donald PJ. Marijuana smoking possible cause of head and neck carcinoma in young patients Otolaryngol Head and Neck Surg. 1986; 94:517

El-Mofty SK, Lu DW. Prevalence of human papillomavirus type 16 DNA in squamous cell carcinoma of the palatine tonsil, and not the oral cavity, in young patients: a distinct clinicopathologic and molecular disease entity. Am J Surg Pathol. 2003; 27(11): 1463-70.

El-Naggar AK, Lai S, Clayman G, Lee J-KJ, Luna MA, Goepfert H, Batsakis JG. Methylation, a major mechanism of p16/CDKN2 gene inactivation in head and neck squamous carcinoma. Am J Pathol 151:1767-1774

Field JK, Spandidos DA, Stell PM. Overexpression of p53 gene in head and neck cancer, linked with heavy smoking and drinking. Lancet 1992; 339:502-503

Field JK, Pavelic ZP, Spandidos DA. The role of p53 tumour suppressor gene in squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg 1993; 119:1118-1122.

Field JK, Spandidos DA, Malliri A. Elevated p53 expression correlates with a history of heavy smoking in squamous cell carcinoma of the head and neck. Br J Cancer, 1991; 64:573-577.

Fletcher GH, Jesse RH, Healey JE, Jr, Thoma GW Jr. Oropharynx. In: Cancer of the head and neck (MacComb WS, Fletcher GH, eds). Baltimore: Williams & Wilkins, 1967:179-212.

Foulkes WD, Brunet JS, Sieh W, Black MJ, Shenouda G, Narod SA. Familial risks of squamous cell carcinoma of the head and neck: retrospective case-control study. BMJ 1996; 313:716-21.

Foulkes WD, Brunet JS, Kowalski LP, Narod SA, Franco EL. Family history is a risk factor for squamous cell carcinoma of the head and neck in Brazil: a case-control study. Int J Cancer 1995; 63: 769-73

Franceschi S, Levi F, Lucchini F. Trends in cancer mortality in young adults in Europe, 1966-1989. Eur J Cancer 1994; 30A: 2096-2118

Franceschi S, Favero A, Conti E, Talamini R, Volpe R, Negri E, Barzan L, La Vecchia C. Food groups, oils and butter, and cancer of the oral cavity and pharynx. BJC 1999 80; 614-620

Franceschi S, Bidoli E Baron AE, La Vecchia C. Maize and risk of cancers of the oral cavity, pharynx, and oesophagus. J. Nat. Cancer. Inst. 1990; 82: 1407-1411

Franceschi S, Munoz N, Bosch XF, Snijders PJ, Walboomers JM. Human papillomavirus and cancers of the upper aerodigestive tract: a review of epidemiological and experimental evidence. Cancer Epidemiol Biomarkers Prev 1996; 5:567-75

Fregonesi PAG, Teresa DB, Duarte RA, Neto CB, de Oliveira MRB, Soare CP. P16 immunohistochemical overexpression in premalignant and malignant oral lesions infected with human papillomavirus. The Journal of Histochemistry and Cytochemistry. 2003; 51:1291-1297

Friedlander PL, Schantz SP, Shaha AR, Yu G, Shah JP. Squamous cell carcinoma of the tongue in young patients: a matched-pair analysis. Head Neck 1998:20:363-368

Gebhart E, T. Liehr, E. Wolff, J. Wiltfang, S. Koscielny, J. Ries. Loss of 9p21 is embedded in a complex but consistent pattern of genomic imbalances in oral squamous cell carcinomas. Cytogenetic and Genome Research 2003; 101:106-112

Gillison M, Koch WM, Shah KV. Human papillomavirus and head and neck squamous cell carcinoma: are some head and neck cancers sexually transmitted disease? Curr Opinion Oncol 1999; 11:191-9

Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L. Evidence of a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst. 2000; 92(9): 675-677, 709-720.

Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. J Natl Cancer Inst. 1994; 86(21): 1600-8.

Gonzalez MV, Pello MF, Lopez-Larrea C, Suarez C, Menendez MJ, Coto E. Deletion and methylation of the tumour suppressor gene p16/CDKN2 in primary head and neck squamous cell carcinoma. J Clin Pathol 1997; 50:509-512

Grandis JR, Tweardy DJ. Elevated levels of transforming growth factor alpha and epidermal growth factor receptor mRNA are early markers of carcinogenesis in head and neck cancer. Cancer Res. 1993; 53:3579-3584.

Ha PK, Califano JA. The role of Human papillomavirus in oral carcinogenesis. Critical reviews in Oral Biology and medicine. 2004; 15:188-196.

Harris EL. Association of oral cancers with alcohol consumption: exploring mechanisms. J Natl Cancer Inst 1997; 89:1656-7.

Hermsen MAJA, Joenje H, Arwert F. Assessment of chromosomal gains and losses in oral squamous cell carcinoma by CGH. Oral Oncology 1997; 33:414-18

Herrero R, Castellsague X, Pawlita M, Lissowska J, Kee F, Balaram P, Rajkumar T. Human papillomavirus and oral cancer. The International Agency for Research n Cancer multicenter study. J Natl Cancer Inst 2003; 95:1772-1783.

Hodge KM, Flynn MB, Drury T. Squamous cell carcinoma of the upper aerodigestive tract in nonusers of tobacco. Cancer 1985; 55:1232-5

Johnston LD, Gerstein D. Initiation of use of alcohol, cigarette, marijuana, cocaine, and other substances in U.S birth cohort since 1919. Am J Public Health 1998:88; 27-33.

Johnson NW. Oral Cancer. FDI World Dental Press Ltd. 1999; 7-9, 29-33

Jovanovic A, Schulten EAJM, Kostense PJ, Snow GB, van der Waal I. Tobacco and alcohol related to the anatomic site of oral squamous cell carcinoma. J Oral Path Med 1993; 22: 459-62

Katoh T, Kaneko K, Kohshi K *et al*. Genetic polymorphisms of tobacco and alcohol related metabolism enzymes and oral cancer. Int J Cancer 1999; 83(5): 606-9

Kellokoski JK, Syrjanen SM, Chang F, Yliskoski M, Syrjanen KJ. Southern blot hybridisation and PCR detection of oral human papillomavirus infections in women with genital HPV infections J Oral Pathol Med 1992; 21:459-64

Kuo WP, Janssen TK, Park PJ, Lingen MW, Hasina R, Ohno-Machado L. Gene expression levels in different stages of progression in oral squamous cell carcinoma. Proc AMIA Symp 2002:415-9

Kuo WP, Hasina R, Ohno-Machado L, Lingen MW. Classification and identification of genes associated with oral cancer based on gene expression profiles. A preliminary study. NY State Dent J 2003; 69:23-6

Lacey PD, Piccirillo JF, Merritt MG, Zequeira MR. Head and neck squamous cell carcinoma: better to be young. Otolaryngol Head Neck Surg. 2000; 122:255-258

Lane DP, Crawford LV. T-antigen is bound to a host protein in SV40 transformed cells. Nature 1979; 278:261-263

Lang JC, Tobin EJ, Knobloch TJ, Schuller DE, Bartynski KJ, Mountain RE, Nicholson R, De Young BR, Weghorst CM. Frequent mutation of p16 in squamous cell carcinoma of the head and neck. Laryngoscope 1998; 108:923-928

Latif F, Fivash M, Glenn G, Tory K, Orcutt M, Hampsch K, Delisio J, Lerman M, Cowan J, Beckett M. Chromosome 3p deletions in head and neck carcinomas: statistical ascertainment of allelic loss. Cancer Res, 1992; 52: 1451-1456.

Lilischkis R, Sarcevic B, Kennedy C, Warlters A, Sutherland RL. Cancer associated missense and deletion mutations impair p16 CDK inhibitory activity. International Journal of Cancer 1996; 66:249-254.

Linzer DI, Levine AJ. Characterisation of a 54k Dalton cellular SV40 tumour antigen present in SV40 transformed cells and uninfected embryonal carcinoma cells. Cell 1979; 17:43-52

Llewellyn CD, Johnson NW, Warnakulasuriya KA. Risk factors for oral cancer in newly diagnosed patients aged 45 years and younger: a case-control study in Southern England. Journal of Oral Pathology and Medicine. 2004; 33(9): 525-32.

Lynch HT, Fusaro RM, Lynch J. Hereditary cancer in adults. Cancer Detection and prevention 1995; 19 219-33

MacFarlane GJ, Bole P, Scully C. Rising mortality from cancer of the tongue in young Scottish males. The Lancet 1987, 912.

MacKenzie J, Ah-See K, Thakker N. Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncology 2000; 37:401-18.

Maden C, Beckmann AM, Thomas DB, McKnight B, Sherman KJ, Ashley RL. Human papillomaviruses, herpes viruses and the risk of oral cancer in men. Am J Epidemiol 1992; 135:1093- 102.

Maestro R, Gasparotto D, Vukosavljevic T, Barzan L. Sulfaro S, Boiocchi M. Three discrete regions of deletion at 3p in head and neck cancers. Cancer Res, 1993; 53:5775-5779

Maier H, Dietz A, Gewelke V *et al*. Occupational exposure to hazardous substances and risk of cancer in the area of the oral cavity, oropharynx, hypopharynx, and larynx. A case-control study. Laryngol Rhinol Otol. 1991; 70:93-98.

Manuel S, Raghavan SKN, Pandey M, Sebastian P. Survival in patients under 40 years with squamous cell carcinoma of the tongue. Int J Oral Maxillofacial Surg 2003; 32:167-173.

Martinez I. Factors associated with cancer of the oesophagus, mouth, and pharynx in Puerto Rico. J Nat Cancer Inst 1969; 42: 1069-1094.

Matsumura T, Yoshihama Y, Kimura T, Shintani S, Alcalde RE. P53 and MDM2 expression in oral squamous cell carcinoma. Oral Oncology 1996; 53:308-12

McGregor GI, Davis N, Robins RE. Squamous cell carcinoma of the tongue and lower oral cavity in patients under 40 years of age. Am J Surg. 1983 Jul; 146(1): 88-92

McLaughlin JK, Gridley G, Block G, Winn, DM, Preston-Martin S, Schoenberg JB, Greenberg RS, Stemhagen A, Austin DF, Ershow AG, Blot WJ, Fraumeni JFJr. Dietary factors in oral and pharyngeal cancer. J. Nat Cancer Inst. 1988; 80: 1237-1243

Mercante G, Bacciu A, Ferri T, Bacciu S. Gastroesophageal reflux as a possible co-promoting factor in development of the squamous cell carcinoma of the oral cavity, the larynx and of the pharynx. Acta Otorhinolaryngol Belg. 2003; 57 (2): 113-7.

Mork J, Moller B, Glattre E. Familial risk in head and neck squamous cell carcinoma diagnosed before the age of 45: a population-based study. Oral Oncol. 1999; 35(4): 360-7.

Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 2003; 348(6): 518–27.

Negri E, Franceschi S, Bosetti C, Conti E, Parpinel M, La Vecchia C. Selected micronutrients and oral and pharyngeal cancer. Int Journal of Cancer 2000; 86:122-127.

Newman AN, Rice DH, Ossoff RH, Sisson GA. Carcinoma of the tongue in persons younger than 30 years of age. Arch Otolaryngol 1983; 109:309-4.

Nylander K, Nilsson P, Mehle C, Roos G. p53 mutations, protein expression and cell proliferation in squamous cell carcinomas of the head and neck. Br J Cancer. 1995; 71:826-830

Parker SL, Tong T, Bolden S, Wingo PA. Cancer Statistics, 1996. CA Cancer J Clin 46:5-27, 1996.

Neville BW, Damm DD, Allen CM, Bouquot JE. Oral and Maxillofacial Pathology. Philadelphia: WB Saunders, 1995:295-304

Petridou E, Zavras AI, Lefatzis D, Dessypris N, Laskaris N, Dokianis G, Sigas J, Douglas CW, Diehl SR, Trichopoulos D. The role of diet and specific micronutrients in the etiology of oral carcinoma. Cancer 2002; 94: 2981-2988.

Pindborg JJ, Reichart PA, Smith CJ, van der Waal I. Histological typing of cancer and precancer of the oral mucosa. (2ed) Berlin:Springer, 1997. International Histological Classification of tumours.

Pitman KT, Johnson JT, Wagner RL, Myers EN. Cancer of the tongue in patients less than forty. Head Neck 2000; 22:297-302

Pott P, ed. Chirurgical Observations. London. Hawes, Clarke and Collins. 1775

Ramirez-Amador V, Esquivel-Pedraza L, Ochoa-Carrillo FJ, Cuapio-Ortiz A, Frias-Mendivil M, Menses-Garcia A, Sanchez-Mejorada G. Cancer of the mobile tongue in Mexico. A retrospective study of 170 patients. Eur J Cancer Oral Oncol 1995; 31(B): 37-40

Regezi JA, Dekker NP, McMillan A, Ramirez-Amador V, Menses-Garcia A, Ruiz-Godoy Rivera L.M, Chrysomaly E, Ng IOL. P53, p21, Rb, MDM2 proteins in tongue carcinoma from patients <35 versus >75 years. Oral Oncology, 1999; 379-383

Rich AM, Radden BG. Squamous cell carcinoma of the oral mucosa: review of 244 cases in Australia. J Oral Pathol 1984; 13: 459-71

Rosenblatt KA, Daling JR, Chen C, Sherman, KJ, Schwartz SM. Marijuana use and risk of oral squamous cell carcinoma. Cancer Research 2004; 64:4049-4054

Sakaguchi M, Fujii Y, Hirabayashi H, Yoon HE, Komoto Y, Oue T, Kusafuka T, Okada A, Matsuda H. Inversely Correlated expression of p16 and Rb protein in Non-Small Cell Lung Cancers: An immunohistochemical study. Int. J. Cancer 1996; 65: 442-445.

Saranath D, Panchal RG, Nair R, Mehta AR, Sanghavi V, Sumegi J *et al*. Oncogene amplification in squamous cell carcinoma of the oral cavity. Jpn J Cancer Res 1989; 80:430-7

Sauter ER, Ridge JA, Gordon J, Eisenberg BL. P53 overexpression correlates with increased survival in patients with squamous cell carcinoma of the tongue base. Am J Surg 1992; 164:651-653

Schantz SP, Guo-Pei Y. Head and Neck Cancer Incidence Trends in young Americans, 1973-1997, with a special analysis for tongue cancer. Arch Otolaryngol Head Neck Surg. 2002; 128:268-274.

Schena M, Shalon D, Davis RW, Brown PO. 'Quantitative monitoring of gene expression patterns with a complimentary DNA microarray', Science, 1995; 270: 467-70.

Schwartz SM, Daling JR, Doody DR, Wipf GC, Carter JJ, Madeleine MM. Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection. J Natl Cancer Inst 1998; 90: 1626-36

Scott FJ, Bates S, James MC, McConnell BB, Starborg M, Brooks S, Palerma I, Ryan K, Hara E, Vousden KH, Peters G. The alternative product from the human CDKN2A locus, $p14^{ARF}$, participates in a negative feedback loop with p53 and MDM2. EMBO 1998; 15: 5001-5014.

Scully C. Oral squamous cell carcinoma: from an hypothesis about a virus, to concern about possible sexual transmission. Oral Oncology 2002; 38:277-34

Serrano M. Hannon GJ, Beach D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature 1993; 366: 704-707

Shanmugaratnam K: Histological typing of tumours of the upper respiratory tract and ear, 2nd ed. Berlin: Springer, 1991:28-33

Silverman S Jr, Gorsky M. Epidemiologic and demographic update in oral cancer: California and national data –1973-1985. J Am Dent Assoc 120:495-499, 1990.

Slaughter DP, Southwick HW, Smejkal W. 'Field Cancerisation in oral stratified squamous epithelium: clinical implications of multicentric origin. Cancer 6:963-968, 1953.

Sorensen DM, Lewark TM, Haney JL, Meyers AD, Krause G, Franklin WA. Absence of p53 mutations in squamous carcinomas of the tongue in non-smoking and nondrinking patients younger than 40 years. Arch Otolaryngol Head Neck Surg 1997; 123: 503-506.

Strong EW, Spiro RH. Cancer of the oral cavity. In: Cancer of the head and neck (Suen JY, Myers EN, eds). New York: Churchill-Livingstone, 1981:301-341.

Tam SW, Shay JW, Pagano M. Differential expression and cell cycle regulation of the cyclin-dependant kinase 4 inhibitor p16. Cancer Research 1994; 54:5816-5820

Toporcov TN, Antunes JLF, Tavares MR. Fat food habitual intake and risk of oral cancer. Oral Oncology 2004; 40:925-931

Trichopoulos D, Willett WC. Introduction. Nutrition and cancer. Cancer causes control. 1996; 7:3-4

Van Rensburg EJ, Engelbrecht S, Van Heerden WFP, Raubenheimer EJ, Schoub BD. Human papillomavirus DNA in oral squamous cell carcinomas from an African population sample. Anticancer Research 1996; 16:969-74

Wang J, Li J, Huang H, Fu Y. Detection of the E7 transform gene of human papillomavirus type 16 in human oral squamous cell carcinoma. Chin J Dent Res 1998; 1: 35-7

Wilke CM, Hall BK, Hoge A, Paradee W, Smith DI, Glover TW. FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. Hum Mol Genet 1996; 5:187-195.

Wolff E, Girod S, Liehr T, Vorderwulbecke U, Ries J, Steininger H, Gebhart E. Oral squamous cell carcinomas are characterized by a rather uniform pattern of genomic imbalances detected by comparative genomic hybridisation. Oral Oncol. 1998; 34(3): 186-90.

World Cancer Research Fund and American Institute for Cancer Research. Food, nutrition and the prevention of cancer. Mouth and pharynx. In: Cancer, nutrition and food. Menasha WI: World Cancer Research Fund and American for Cancer Research, 1997:96-106, 599-600.

World Cancer Report IARC Press Lyon 2003

World Health Organisation (1999) Global Status Report on alcohol, Geneva, WHO.

Woutersen P, Clifford CK, Milner JA. Diet and cancer prevention. Eur J Cancer 2001; 37: 948-65.

Xia W, Lau YK, Zhang HZ. Strong correlation between c-erbB-2 overexpression and overall survival of patients with oral SCC. Clin Cancer Res 1997; 3: 3-9

Zhang Z, Morganstern H, Spitz M, Tashkin DP, Yu G, Marshall J, Hsu TC, Schantz SP. Marijuana use and increased risk of SCC of head and neck. Cancer Epidemiology, Biomarkers and Prevention. 1999; 8: 1071-1078

Zurhausen H. Viruses in human cancers. Science. 1991; 254:1167-1174.

Chapter 2

Methodology

2.1 Introduction

This chapter is a comprehensive account of all the methodologies employed in this thesis, along with background information for the newer technologies. Several of the technologies are employed in a number or all of the chapters e.g. TaqMan RT-PCR/PCR. Where this occurs the specific primer and probe sequences and optimised reaction conditions are included in the relevant chapters.

2.2 Section 1

2.2.1 Sample selection

All samples were selected from patients in St. James's Hospital, Dublin. While some tests required merely paraffin sections, other required RNA from fresh tissue, resulting in some patient sample differences from chapter to chapter.

Table 2.1 summarises the samples used in each chapter.

Sample number	Chapter 4 CGH arrays	Chapter 5 p16 IHC	Chapter 5 HPV	Chapter 6 Expression arrays	Chapter 7 TaqMan
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					

Table 2.1 Green boxes represent the samples that were used in each experiment. The blue boxes represent the samples that were not used.

2.2.2 Fresh tissue harvesting and archiving

All surgically removed tissue from patients with oral and pharyngeal cancer with appropriate consent was considered for tissue banking. To ensure no compromise of the diagnostic potential of the sample, the specimen was examined by a pathologist prior to banking.

1) A slice of tumour/diseased tissue was taken with a new scalpel blade. The amount of tissue harvested depended on the size of the specimen or tumour.

2) Each slice was divided into separate smaller fragments (4mm x 4mm.), and the fragments were placed in Eppendorf or cryovial tubes labelled (in pencil) with the specimen number and the letter T for tumour.

3) The Eppendorfs were immersed in liquid Nitrogen for 2 minutes and then immediately stored in the freezer at -80°C.

4) Separate fragments were placed in cryovials/eppendorfs containing RNAlater®, a proprietary solution used for storing RNA without degradation. Tissues were stored in RNAlater RNA Stabilization Reagent for up to 1 day at 37°C, 7 days at 18–25°C, 4 weeks at 2–8°C, or for archival storage at –20°C or –80°C.

2.2.3 Formalin fixed paraffin embedded tissue

Formalin fixed (10% buffered) and paraffin embedded tumour samples were selected from the pathology files of St. James's Hospital, Dublin. Haematoxylin & Eosin (H&E) stained slides of all tumour samples were reviewed by a pathologist.

2.3 Section 2

2.3.1 Tissue preparation prior to nucleic acid extraction

Various preparation methods were used depending on how the tissue was stored and which nucleic acid was being extracted.

2.3.2 Frozen section cutting Protocol

The frozen section technique was employed for a number of reasons:

- To establish the presence of tumour in a sample of tissue.
- To allow for laser capture microdissection prior to nucleic acid extraction

The frozen tissue fragment was placed on a corkboard in a large drop of sterile ultra pure H_2O. (Use of embedding compound OCT gives less favourable results). The corkboard was placed on a freezing microtome chuck.

The tissue, cork and chuck were immersed in liquid nitrogen and frozen.

The chuck was then transferred to the cold plate of the microtome. The block was "faced off" initially at a cutting depth of 30μm.

After facing off, ribbon sections were cut at 8μm thickness.

2.3.3 Fixing and staining for H&E

After cutting the frozen sections the slides were fixed in alcohol and stained rapidly as follows (protocol provided by J. Winters, NIH, Bethesda, Maryland 20892).

Alcohol 70% - 15 dips.

H2O – 1 dip.

Haematoxylin – 15 dips.

H2O – 1 dip.

Alcohol 70% - 10 dips.

Alcohol 95% - 15 dips.

Eosin – 15 dips.

Alcohol 95% - 15 dips.

Alcohol 100% (1) dips.

Alcohol 100% (2) – 15 dips.

Xylene (1) until clean.

Xylene Hold.

2.3.4 Tissue Homogenisation

Laser capture microdissected (LCM) cells require no special homogenisation procedures. In contrast, homogenisation of the large fragments of banked oral and pharyngeal SCC tissue was a critical step, particularly prior to RNA extraction. A

variety of homogenisation techniques were evaluated. The best homogenisation procedure for snap frozen tissue was as follows.

2.3.5 Mortar and pestle homogenisation of tissue

The banked frozen or RNA later stabilised tissue was removed from the -80°C freezer and placed on dry ice during the extraction process. The mortar and pestle was autoclaved and treated with an RNase inhibitor (RNA Zap®) prior to homogenisation. The mortar and pestle was pre-cooled with liquid nitrogen. Tissue was transferred directly from the dry ice to the mortar and pestle and immersed in liquid nitrogen. The liquid nitrogen may be allowed to evaporate, but the tissue should not be allowed to thaw.

The tissue was then thoroughly homogenised by grinding and the resultant powder and small fragments transferred to an RNase free Eppendorf tube, containing the appropriate lysis buffer (600µL). After vortexing, the sample was further homogenised using a Qiagen Shredder column.

2.3.6 QIAshredder®

Qiagen shredders are a commercially available spin column based shearing system. They can be used for further homogenisation of tissue initially homogenised by mortar and pestle and for cultured cells as the primary form of homogenisation.

Lysates were pipetted directly onto a QIAshredder spin column placed in 2 ml collection tube, and centrifuged for 2 min at maximum speed. The flow through was used for subsequent RNA extraction.

2.3.7 Microfuge pellet mixer.

A Microfuge pellet mixer comprises plastic rods, which fit snugly into 1.5ml Eppendorf type tubes, and which can be attached to a hand held motorized rotor. They were used to partially homogenize formalin fixed paraffin embedded (FFPE) tissue samples prior to further digestion using Proteinase K.

2.4 Section 3

2.4.1 Arcturus PixCell II Laser Capture Microdissection system

LCM is based on the selective adherence of visually targeted cells and tissue fragments to a thermoplastic ethylene vinyl acetate (EVA) membrane activated by a low energy infrared laser pulse. The system consists of an inverted microscope, a solid state near infrared laser diode, a laser control unit, a joystick controlled microscope stage with a vacuum chuck for immobilisation, a CCD camera, and a colour monitor. The LCM microscope is connected to a computer with additional laser control and image archiving.

In brief, the cap is positioned on the desired area of the dehydrated tissue section. Laser activation leads to focal melting of the EVA membrane. The melted polymer expands into the section and fills the extremely small hollow spaces present in the tissue. The polymer resolidifies within milliseconds and forms a composite with the tissue, allowing selective removal of the cells. The selected cells are harvested by simple lifting of the cap, which is then transferred to a micro centrifuge tube containing the appropriate extraction lysis buffer (Figure 2.1). The maximum temperatures reached by the tissue are in the range of 90°C for several milliseconds. The low energy of the infrared laser also avoids potentially damaging photochemical effects (Fend, 2000).

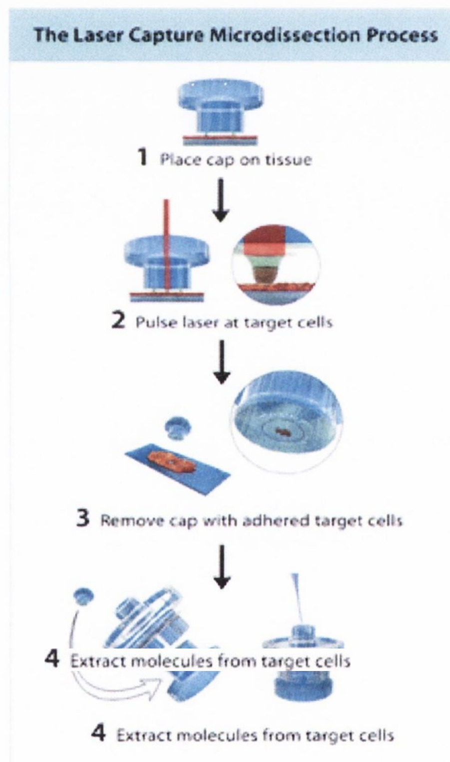


Figure 2.1 Overview of LCM Process

2.4.2 Setting LCM Parameters

The laser spot size can be adjusted to three sizes (<7.5µm, 15µm and 30µm). Further adjustable parameters include: pulse power, pulse width, threshold voltage and pulse duration. Laser pulse power and duration determine the size of the LCM transfer. Table 2.1 shows the best performing settings for LCM of HNSCC tissue.

Spot Size	Power	Duration	Threshold Voltage
<7.5 □m	40mW	450 ms	285MmV
~15□m	25mW	1.5ms	285MmV
~30□m	20mW	5ms	285MmV

Table 2.1 LCM Settings

The total number of pulses in each case was approximately 1000. For a 30µm spot size this yielded a tissue volume in the range of 10^{-7} to 10^{-6} μm^3. Examples of laser captured oral SCC cells are seen in Figure 2.2.

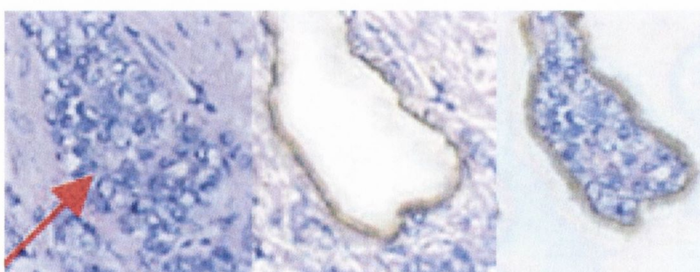


Figure 2.2. Example of cells captured by LCM. Red arrow indicates the malignant epithelial cells being captured

2.5 Section 4

2.5.1 Nucleic acid extraction protocols

Throughout this thesis various analyses were carried out on nucleic acids from a variety of thyroid tissue types stored in different formats (e.g. fresh frozen, FFPE, etc.). The extraction protocol employed was largely dictated by the quantity and quality of nucleic acid required and by the format of the starting tissue.

The extraction of pure, intact RNA is important in a variety of molecular biological techniques, and is essential for gene expression analysis. Ribonucleases however can cause difficulty with RNA isolation because they are very stable, active enzymes that require no cofactors for enzymatic activity. Cell lysis in an environment that causes denaturation of ribonucleases is therefore essential. So too is adequate decontamination of work surface areas, solutions and plastic disposables. DNA extraction need not be as stringent, as simple resuspension with a chelating agent confers nuclease protection to the molecule, however it is wise to be consistent with precaution.

2.5.2. DNA extraction from frozen sections using Puregene® Extraction Kit

Cell Lysis and RNA digestion

Add 0.3ml cell lysis solution to the laser captured cells.

Add 1.5 µl Proteinase K solution (20µg/ml) to the lysate, mix well and incubate at 55°C overnight. (With constant mixing)

Allow cooling to room temperature. Add 1.5µl RNase A solution (4µg/ml) to the cell lysate.

Mix by inverting x 25 and incubate at 37°C for 1 hour.

Protein precipitation

Cool sample to room temperature and add 100µl of the protein precipitation solution.

Vortex vigorously for 20 seconds to mix the protein precipitation solution evenly with cell lysate and then centrifuge at 13,000-16,000 x g for 3 minutes.

DNA precipitation

Pour off the supernatant containing the DNA into 1.5ml microfuge tube containing 0.3ml 100% isopropanol. Mix by inverting x 50 times. If DNA yields are expected to be low add a carrier such as glycogen (1µl, 20mg/ml).

Centrifuge at 13,000-16,000 x g for 5 minutes.

Pour off the supernatant and drain the tube on clean absorbent paper.

Add 0.3ml 70% alcohol and invert tube to wash the pellet. Centrifuge at 13,000 - 16,000 x g for 1 min. Carefully pour off the alcohol.

Invert and drain the tube on clean absorbent paper and allow to air dry for 10-15 minutes.

DNA hydration

Add 20μl of the DNA hydration solution (This will give a concentration of 100ng/ml if the total yield is 2μg of DNA)

Rehydrate DNA by incubating the sample for 1 hour at 65°C.

Proceed to PCR assay or store DNA at -20°C. For future use the DNA may be aliquoted into 1μl volumes and stored at -20°C.

2.5.3 Extraction of RNA from tissue samples using Qiagen RNeasy spin columns.

Preparation for RNeasy column RNA extraction

A maximum amount of 30 mg tissue can generally be processed with RNeasy mini columns. For oral and pharyngeal SCC, the binding capacity of the column (100 μg RNA) and the lysing capacity of Buffer RLT will not be exceeded by these amounts.

To process, do not allow tissue to thaw during weighing or handling prior to disruption in Buffer RLT.

β-Mercaptoethanol (β-ME) must be added to Buffer RLT before use. β-ME is toxic and should be dispensed in a fume hood and wearing appropriate protective clothing. Add 10 μl β-ME per 1 ml Buffer RLT. Buffer RLT is stable for 1 month after addition of β-ME.

Buffer RPE is supplied as a concentrate. Before using for the first time, add 4 volumes of ethanol (96–100%), as indicated on the bottle, to obtain a working solution.

All centrifugation steps are performed at 20–25°C in a standard microcentrifuge. The tissue or cells should be homogenized as discussed in Section 2.3.

1) After homogenization the tissue lysate should be centrifuged for 3 min at maximum speed in a microcentrifuge. Then carefully transfer the supernatant to a new microcentrifuge tube by pipetting. Use only this supernatant (lysate) in subsequent steps.

2) Add 1 volume (usually 600 μl for tissue) of 70% ethanol to the cleared lysate, and mix immediately by pipetting. Do not centrifuge. Continue without delay with step 3.

Note! A precipitate may form after the addition of ethanol, but this will not affect the RNeasy procedure.

3) Apply up to 700 μl of the sample, including any precipitate that may have formed, to an RNeasy mini column placed in a 2 ml collection tube (supplied). Close the tube gently, and centrifuge for 15 s at >8000 x *g* (>10,000 rpm). Discard the flow through. Reuse the collection tube.

If the volume exceeds 700 μl, load aliquots successively onto the RNeasy column, and centrifuge as above. Discard the flow-through after each centrifugation step.

<u>Protocol for on column DNase digestion</u>

Although the RNeasy silica-membrane technology efficiently removes most of the DNA without DNase treatment further DNA removal may be necessary for certain

RNA applications that are sensitive to very small amounts of DNA (e.g., TaqMan RT-PCR analysis with a low-abundant target) and cDNA microarray.

Preparation for DNase digestion

Prepare DNase 1 stock solution before using the RNase-Free DNase Set for the first time. Dissolve the solid DNase I (1500 Kunitz units) in 550 μl of the RNase-free water provided. Take care that no DNase 1 is lost when opening the vial. Mix gently by inverting the tube.

DNase 1 is especially sensitive to physical denaturation and must not be vortexed. Mixing should only be carried out by gently inverting the tube. For long-term storage of DNase 1, remove the stock solution from the glass vial, divide it into single-use aliquots, and store at –20°C for up to 9 months. Thawed aliquots can be stored at 2–8°C for up to 6 weeks. Do not refreeze the aliquots after thawing.

DNase digestion

Carry out lysis, homogenization, and loading onto the RNeasy mini column as indicated in the protocols above.

1) Pipet 350 μl Buffer RW1 into the RNeasy mini column, and centrifuge for 15 s at ≥8000 x *g* (≥10, 000 rpm) to wash. Discard the flow-through.

2) Add 10 μl DNase I stock solution (see above) to 70 μl Buffer RDD. Mix by gently inverting the tube. Buffer RDD is supplied with the RNase-Free DNase Set.

3) Pipet the DNase 1 incubation mix (80 μl) directly onto the RNeasy silica-gel membrane, and place on the bench top (20–30°C) for 15 min.

Note: Make sure to pipet the DNase 1 incubation mix directly onto the RNeasy silica-gel membrane. DNase digestion will be incomplete if part of the mix sticks to the walls or the O-ring of the RNeasy column.

4) Pipet 350 μl Buffer RW1 into the RNeasy mini column, and centrifuge for 15 s at ≥8000 x *g*. Discard the flow-through. * Continue with the first Buffer RPE wash step in the relevant protocol.

Washing and eluting RNA

1) Transfer the RNeasy column into a new 2 ml collection tube (supplied). Pipet 500 μl Buffer RPE onto the RNeasy column. Close the tube gently, and centrifuge for 15 s at >8000 x *g* (>10,000 rpm) to wash the column. Discard the flow-through. Reuse the collection tube in the next step.

2) Add another 500 μl Buffer RPE to the RNeasy column. Close the tube gently, and centrifuge for 2 min at >8000 x *g* (>10,000 rpm) to dry the RNeasy silica-gel membrane.

To eliminate any chance of possible Buffer RPE carryover, continue first with step 3.

It is important to dry the RNeasy silica-gel membrane since residual ethanol may interfere with downstream reactions. This centrifugation ensures that no ethanol is carried over during elution.

Note: Following the centrifugation, remove the RNeasy mini column from the collection tube carefully so the column does not contact the flow-through as this will result in carryover of ethanol.

3. Place the RNeasy column in a new 2 ml collection tube and discard the old collection tube with the flow-through. Centrifuge in a microcentrifuge at full speed for 1 minute.

4. To elute, transfer the RNeasy column to a new 1.5 ml collection tube. Pipet30–50 µl RNase-free water directly onto the RNeasy silica-gel membrane. Close the tube gently, and centrifuge for 1 min at ≥8000 x *g* (≥10, 000 rpm) to elute.

5. If the expected RNA yield is >30 µg, repeat the elution step (step 11) as described with a second volume of RNase-free water. Elute into the same collection tube.

To obtain a higher total RNA concentration, this second elution step may be performed by using the first elute (from step 4). The yield will be 15–30% less than the yield obtained using a second volume of RNase-free water, but the final concentration will be higher.

2.6 Section 5

2.6.1 Assessing RNA and DNA samples

All array experiments, both CGH and expression arrays required accurate quantitation of DNA and RNA yields along with assessment of nucleic acid integrity.

2.6.2 Quantifying RNA and DNA Samples By UV Spectrophotometry

The wavelength of maximum absorption of light by both DNA and RNA is 260 nm (λ_{max}=260 nm), which is conveniently distinct from the λ_{max} of protein (280nm). The absorption properties of nucleic acids can be used for detection, quantitation, and assessment of purity. Quantitation of nucleic acids is possible by measuring the absorbance at 260 nm and by factoring in the appropriate extinction coefficient (RNA extinction coefficient = 40, DNA extinction coefficient = 50). The background absorbance at 320 nm should be close to zero and can be used to correct for background absorbance (e.g. A260 – A 320, True absorbance at 260 nm). The purity of nucleic acids may be estimated by determination of the ratio of absorbance at 260 and 280 nm A_{260}/A_{280}. Pure double stranded DNA has an A_{260}/A_{280} of 1.8, and pure RNA one of around 2.0. Protein with λ_{max} of 280 has an A_{260}/A_{280} of less than 1. Hence, if a DNA sample has an A_{260}/A_{280} greater than 1.8, this is suggestive of RNA contamination, whereas one less than 1.8 suggests protein contamination (Turner, 2000).

To calculate RNA/DNA concentration of each sample: (A260-A320) x Extinction coefficient (40 for RNA and 50 for DNA) x dilution factor (*i.e.*100/5) x final sample volume = RNA/DNA yield in µg.

2.6.3 Assessing of nucleic acid integrity by Agarose Gel Electrophoresis

Agarose is a polysaccharide derived from seaweed, which forms a solid gel when dissolved in aqueous solution at concentrations between 0.5 and 2% (w/v). When an

electric field is applied to an agarose gel in the presence of a buffer solution, DNA/RNA fragments move through the gel towards the positive electrode at a rate, which is dependent on their size and shape. The nucleic acid is visualized after electrophoresis by staining with Ethidium bromide and illuminating the gel under UV light (Figure 2.3)

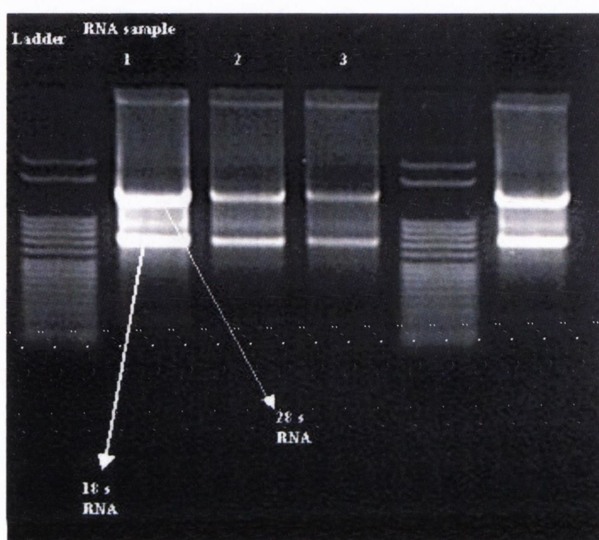


Figure 2.3 Agarose gel electrophoresis of total RNA. High quality RNA has 28S rRNA band: 18S rRNA ratio equal to 2:1 intensity and there is little or no genomic DNA carry-over (high MW).

2.6.4 Assessing RNA quality and quantity using the Agilent Bioanalyser 2100

While agarose gel electrophoresis can be used to monitor the quality of RNA preparations, it is a messy, labour intensive and time-consuming technique requiring relatively large amounts of sample. The Bioanalyser from Agilent is an alternative

system for characterization of total or mRNA samples and determination of their concentrations.

This system uses lab-on-a-chip technology to provide improved analysis of DNA, RNA, proteins and cells through:

- Automation for improved accuracy and reproducibility, simple, robust protocols
- Rapid visualisation of sample quality and quantity (up to 12 samples in 30 minutes)
- Easy detection of RNAse degradation
- High sensitivity with minimal sample quantities (As little as 5 ng/μl RNA is required per analysis)
- Reduced use and waste of hazardous chemicals

RNA 6000 Nano Analytical Specifications

Specification	Total RNA Assay	mRNA Assay
Quantitative range	25–500 ng/μl	25–250 ng/μl
Qualitative range	5–500 ng/μl	25–250 ng/μl
Maximum sample buffer strength	10 mM Tris-EDTA	10 mM Tris-EDTA
Reproducibility of quantitation	10% CV	10% CV

Figure 2.4 Agilent Bioanalyser 2100 'chip

RNA LabChip® kits contain chips and reagents designed for sizing and analysis of RNA fragments. Each RNA LabChip® contains an interconnected set of micro

channels that sieves nucleic acid fragments by size as they are driven through it by means of electrophoresis.

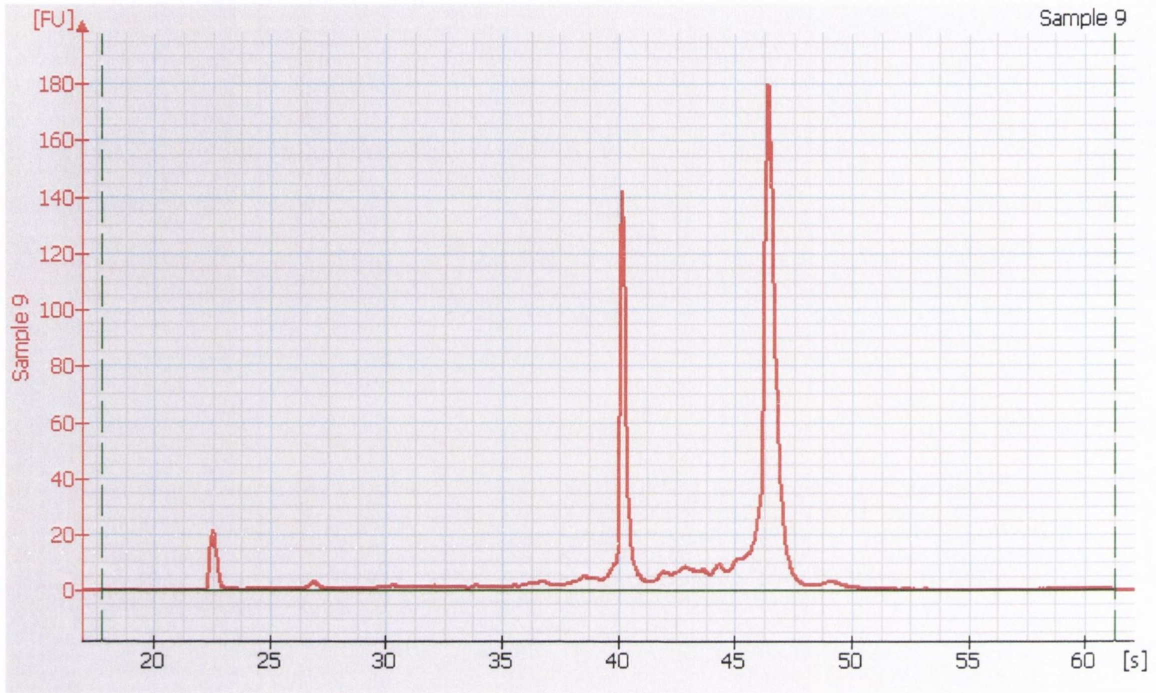


Figure 2.5 Example of an electropherogram. Major features for a successful total RNA run are 2 ribosomal peaks (with successful sample preparation) and 1 marker peak

2.7 Section 6

2.7.1 Reverse transcription of RNA using cDNA archive kit

The High-Capacity cDNA Archive Kit (Roche, Nutley, NJ) contains reagents for reverse transcription (RT) of total RNA to single-stranded cDNA.

Protocol

Preparing Reaction Master Mix

1 Allow the kit components to thaw on ice.

2 Calculate the volume of components needed by using the following volumes per reaction.

10X Reverse Transcription Buffer 10μl

25X dNTPs 4μl

10X random primers 10μl

MultiScribe™ Reverse Transcriptase, 50 U/μL 5μl

Nuclease-free H2O 21μl

Total per Reaction 50μl

3. Pipette 50 μL of 2X RT master mix into each microcentrifuge tube.

4. Pipette 50 μL of RNA sample into each tube, pipetting up and down two times to mix.

4 Place the tubes on ice until GeneAmp® PCR System 9600 thermal cycler is set up.

Performing the RT reaction

1 Program the thermal cycler conditions as follows: 25°C for 10 minutes followed by 37°C for 120 minutes

2 Set the reaction volume to **100 μL**.

3 Load the reaction plate into the thermal cycler.

4 Start the reverse transcription run.

2.8 Section 7

2.8.1 TaqMan® PCR

A TaqMan® PCR-based system was selected for mRNA quantitation and DNA genotyping in this study. The Taqman® system was chosen for a number reasons. Firstly Taqman® PCR and Reverse-Transcriptase (RT) PCR requires only a few nanograms of target DNA/RNA. This is highly significant since the amount of DNA/RNA extractable from formalin fixed and paraffin embedded archival material is low. Secondly, Taqman PCR and RT-PCR products are small (generally less than 200bp) and thus can be used to amplify partially degraded or fragmented DNA/RNA such as that obtained from FFPE material.

TaqMan® PCR exploits the 5′ nuclease activity of Amplitaq Gold® DNA Polymerase to cleave a TaqMan probe during PCR. The TaqMan probe contains a reporter dye at the 5′ end of the probe and a quencher dye at the 3′ end of the probe. During the reaction, cleavage of the probe separates the reporter dye and the quencher dye,

resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye, shown in Figure 2.6.

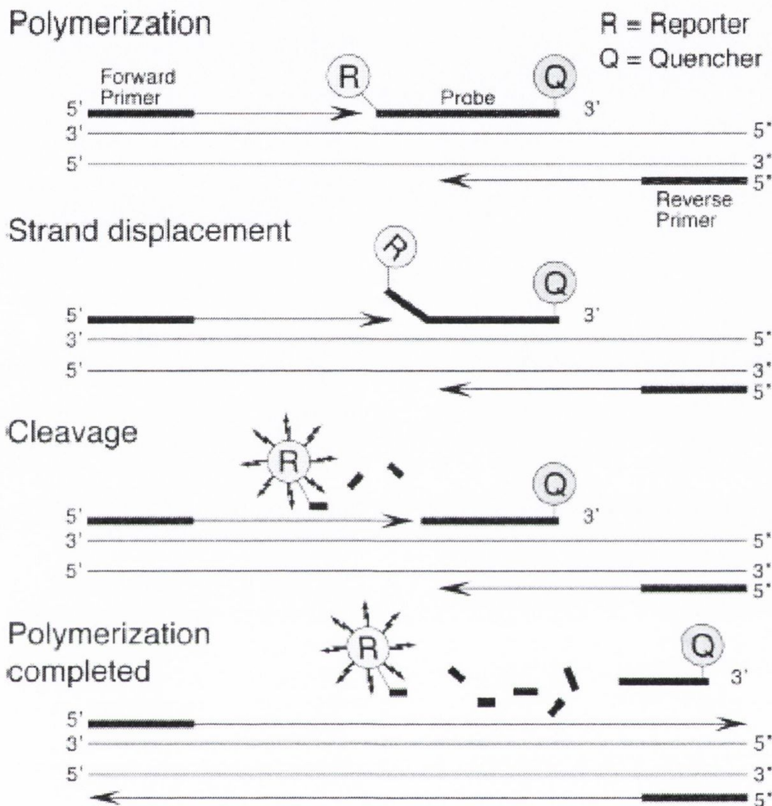


Figure 2.6 The forklike-structure-dependent, polymerisation-associated, 5′–3′ nuclease activity of AmpliTaq Gold DNA Polymerase during PCR

When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence primarily by Förster-type energy transfer (Lakowicz, 1983). During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites.

The 5′-3′-nucleolytic activity of the AmpliTaq Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridises to the target. The probe fragments are then displaced from the target, and polymerisation of the strand continues. The 3′ end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product.

The probe consists of an oligonucleotide with a 5′-reporter dye and a 3′-quencher dye. A fluorescent reporter dye, such as FAM (6-carboxyfluorescein), is covalently linked to the 5′ end of the oligonucleotide. TET (6-carboxy-4,7,2′,7′-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), and VIC are also used as reporter dyes. In older TaqMan® probes each of the reporters is quenched by TAMRA (6-carboxy-N,N,N′,N′-tetramethylrhodamine) at the 3′ end.

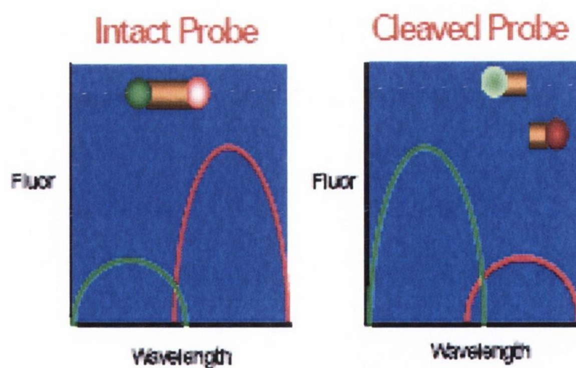


Figure 2.7 Increased fluorescence activity due to cleaved probe

Newer TaqMan® MGB probes are recommended for genotyping/allelic discrimination (AD) purposes and when conventional probe exceed 30bp. MGB probes contain:

A non-fluorescent quencher (NFQ) at the 3′ end - the SDS instruments can measure the reporter dye contributions more precisely because the quencher does not fluoresce.

A minor groove binder at the 3′ ends - The minor groove binder increases the melting temperature (Tm) of probes, allowing the use of shorter probes (Kutyavin, 2000)

2.8.2 Real-time quantitative TaqMan® RT-PCR

Real-time RT-PCR is the ability to monitor the progress of the PCR as it occurs (i.e., in real time). Data is therefore collected throughout the PCR process, rather than at the end of the PCR. This completely revolutionizes the way one approaches PCR-based quantitation of DNA and RNA. In real-time RT-PCR, reactions are characterized by the point in time during cycling when amplification of a target is first detected rather than the amount of target accumulated after a fixed number of cycles. The higher the starting copy number of the nucleic acid target, the sooner a significant increase in fluorescence is observed. In contrast, an endpoint assay (also called a "plate read assay") measures the amount of accumulated PCR product at the end of the PCR cycle. RT-PCR can be one-step or two-step in nature.

In the initial cycles of PCR, there is little change in fluorescence signal. This defines the baseline for the amplification plot. An increase in fluorescence above the baseline indicates the detection of accumulated target. A fixed fluorescence threshold can be set

above the baseline. The parameter C_T (threshold cycle) is defined as the fractional cycle number at which the fluorescence passes the fixed threshold (see Fig 2.8).

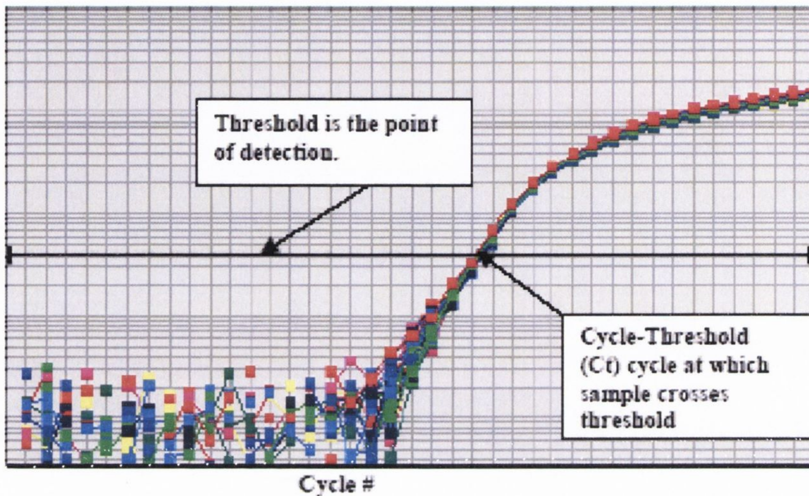


Figure 2.8 Example of an amplification plot

A sample of known concentration used to construct a standard curve. By running standards of varying concentrations, you create a standard curve from which you can extrapolate the quantity of an unknown sample. The standards used depend on whether absolute or relative quantitation is to be used. Relative quantitation was chosen as the method of quantitation in all the experiments described in this thesis.

It is easy to prepare standard curves for relative quantitation as quantity is expressed relative to some basis sample, such as the calibrator. For all experimental samples, target quantity is determined from the standard curve and divided by the target quantity

of the calibrator. Thus, the calibrator becomes the 1X sample, and all other quantities are expressed as an n-fold difference relative to the calibrator. As an example, in a study of drug effects on expression, the untreated control would be an appropriate calibrator.

For quantitation normalized to an endogenous control, standard curves are prepared for both the target and the endogenous reference. For each experimental sample, the amount of target and endogenous reference is determined from the appropriate standard curve. Then, the target amount is divided by the endogenous reference amount to obtain a normalized target value. Again, one of the experimental samples is the calibrator, or 1X sample. Each of the normalized target values is divided by the calibrator normalized target value to generate the relative expression levels. For the quantitation of gene expression, researchers have used β-actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ribosomal RNA (rRNA), or other RNAs as an endogenous control.

In the p16 TaqMan experiments, relative quantitation by standard curve was used (described in detail in Chapter 5). For the HPV DNA TaqMan experiments, an end-point read was used, where HPV DNA was deemed present or absent.

2.8.3 Choice of template DNA/positive control

DNA extracted from the HPV positive cervical carcinoma cell lines Caski and Hela were employed as positive controls.

2.9 Section 8

2.9.1 Nucleic acid sequence-based amplification (NASBA)

Note: While RNA extraction was performed by the student (E O'Regan), the amplification process was kindly performed by Hanne Skomedal (NorChip AS, Klokkarstua, Norway). The data was then analysed by E O'Regan.

Reverse transcription followed by polymerase chain reaction (RT-PCR) is frequently used for highly sensitive detection of gene expression. However the presence of contaminating genomic DNA in RNA samples can frequently give rise to false positive results. In real-time RT-PCR assays PCR probes are when possible designed to span an exon-exon junction with PCR primers placed on either side of the intron. As a result false positive results caused by contaminating genomic DNA can be avoided. However, this strategy is not possible for intronless genes of eukaryotes and gene expression studies in prokaryotes.

In contrast to PCR, nucleic acid sequence-based amplification (NASBA) can directly amplify RNA and is selective for single-stranded RNA. Double-stranded DNA is not denatured during the reaction and consequently not amplified (Burchill, 2002). As a result knowledge of intron-exon boundaries is not essential and contaminating DNA should not affect the efficiency of the reaction or generate false positive results. For these reason NASBA was chosen for the detection of "intronless" viral targets.

NASBA is an isothermal nucleic acid amplification method, which amplifies RNA in a manner analogous to the replication of retroviruses (Hibbitts, 2003). A NASBA reaction is based on the concurrent activity of three enzymes: avian myeloblastosis virus reverse transcriptase (AMV-RT), RNase H and T7 RNA polymerase together with two oligonucleotide primers to produce amplification. The NASBA process occurs at 41°C and results in the exponential amplification of products within 90 minutes. Single-stranded RNA of opposite sense to the original target is the major amplification product. Since the process takes place at 41°C, double stranded DNA is not denatured and consequently not amplified. In contrast to polymerase chain reaction (PCR), NASBA amplifies RNA directly and does not require an extra reverse transcription reaction. Figure 2.9 gives an overview of the reaction principle. Detection of NASBA products can be detected by either probe-capture hybridisation and electrochemiluminesence (ECL) or detection can be carried out in ‘real-time’ using molecular beacons.

Molecular beacons are single stranded oligonucleotides designed to have a loop region containing a probe sequence complementary to the target nucleic acid and a stem region, which is formed by the annealing of complementary arm sequences that are located on either end of the probe sequence. One arm of the stem is labeled with a fluorescent dye and the other arm is labelled with a non-fluorescent quencher. In this form the probe does not fluoresce, however when the molecular beckon hybridises to its target it undergoes a conformational change, which separates the quencher from the fluorophore resulting in emission of fluorescence (Hibbitts, 2003).

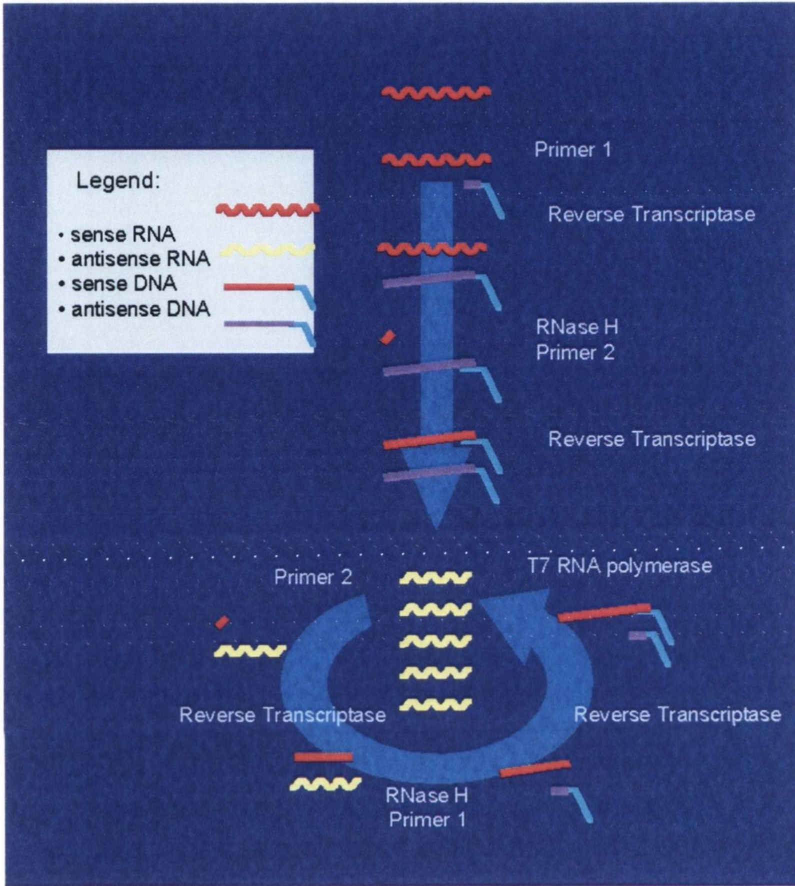


Figure 2.9 Schematic Representation of NASBA

2.9.2 PreTect HPV Proofer Kit (Norchip AS)

The PreTect HPV-Proofer Kit (NorChip AS, Norway) is a molecular test kit for assessing the presence of HPV mRNA. The test is based on Real-Time NASBA technology. The NASBA reaction generates single-stranded RNAs, to which specific molecular beacon probes can hybridise simultaneously, producing a fluorescent signal. The formation of the newly generated RNA molecules is determined in real-time by continuous monitoring of fluorescence in a fluorescent reader. The PreTect HPV Proofer Kit also contains a primer and probe set for U1A, which is employed as an extraction control. Molecular beacons are labeled with either FAM of TEXAS RED fluorescent dyes. Data analysis is carried out using PreTect Analysis Software (PAS).

2.9.3 NASBA with molecular beacon 'real-time' detection

NASBA reactions were carried out according to the PreTect HPV Proofer User Guide. Briefly 80μl of reagent diluent, 24μl primer & molecular beacon solution and 14μl KCL stock solution were added to each lyophilised reagent sphere and vortexed. The 3' sequence of primer one complements the target RNA and the 5' terminal contains a T7 RNA polymerase promoter sequence. The second primer is complementary to the cDNA sequence that is produced by extension from primer 1 at the 3' end. The molecular beacon probe is a fluorescent labelled single-stranded oligonucleotide, which is complementary to the target RNA sequence. The amplification solution was aliquoted

into 10μl reactions in a 96 well microtiterplate and 5μl of extracted nucleic acid was added. The appropriate positive (CaSki and HeLa RNA mix) and negative (water) controls were included in each assay. The reactions were incubated at 65°C for 5 minutes to destabilize secondary structures in RNA. Following incubation at 41°C for 5 minutes 5μl of NASBA enzyme solution was added to initiate amplification. (Enzyme spheres were supplied lyophilised and were reconstituted in enzyme diluent at least 20 minutes prior to use). The microtiterplate was then immediately transferred to a Lambda Fluoro 320 reader (MWG Biotech, Germany) and incubated at 41°C for 90 minutes to complete amplification by the combined simultaneous activity of AMV-RT, RNase H and T7 RNA polymerase. Fluorescence was monitored in 'real-time' over the 90-minute amplification period. Analysis of results was carried out using PreTect Analysis Software (PAS) (NorChip AS).

2.10 Section 9

2.10.1 Immunohistochemistry using the p16^{INK4A} Research Kit (DAKO, MTM laboratories AG)

1) Cut 4μm sections from formalin fixed and paraffin embedded biopsy samples and mount on 3-aminopropyltriethoxysaline (APES) coated glass slides. Allow sections were allowed to dry overnight at 37°C.

2) Dewax sections by passage through xylene (two 5 minute washes) and then rehydrate in graded alcohol (100% EtOH → 80% EtOH → 60% EtOH→ 40% EtOH →. H_2O).

3) For antigen unmasking/retrieval use Epitope Retrieval Solution (supplied) at 95°C-99°C for 40 minutes. Firstly, fill coplin jars with Epitope Retrieval Solution and place them in a water bath. Heat the water and the Epitope Retrieval Solution to 95°C-99°C. Then immerse the tissue sections in the Epitope Retrieval Solution and bring temperature of the water and Epitope Retrieval solution back up to 95°C-99°C. Cool to room temperature for 20 minutes.

4) Decant off the Epitope Retrieval Solution rinse sections wash buffer (supplied). Soak slides in the wash buffer for 5 minutes prior to staining. Decant off excess buffer and clean off liquid surrounding the tissue in order to keep reagents within the prescribed area.

5) Incubate the tissue sections with 200μl of Peroxidase-Blocking Reagent (supplied) for 5 minutes at room temperature. Wash carefully with wash buffer and place in a fresh buffer bath for 5-10 minutes.

6) Decant off excess buffer and incubate sections with 200μl of freshly diluted Anti-Human $p16^{INK4A}$ reagent or Negative Control reagent for 30 minutes at room temperature. Wash tissue sections and place in a fresh buffer bath for 5-10 minutes.

7) Decant off excess buffer and incubate sections with 200μl of Visualisation reagent for 30 minutes at room temperature. Wash gently with wash buffer and place in a fresh buffer bath for 5-10 minutes.

8) Develop slides with 200µl Substrate-Chromagen Solution (diaminobezaminidine DAB) for 10 minutes. Stained lightly with haematoxylin, mount in DPX and coverslip.

9) Negative and positive controls are included in each staining run. Sections of a known p16 positive invasive squamous cell carcinoma were employed as a positive control. Sections of the same invasive case incubated with Negative Control reagent were employed as a negative staining control. This negative control reagent consists of culture supernatant containing monoclonal mouse IgG2a antibody to Aspergillus Niger glucose oxidase, an enzyme that is neither present nor inducible in mammalian tissues.

2.10.2 Interpretation of antigen expression in biopsy tissues

All formalin fixed and paraffin embedded sections, which demonstrated either strong nuclear or cytoplasmic staining was considered positive. A certified pathologist then graded all sections qualitatively according to the following arbitrary scale: 0 (no positive staining of tumour cells), 1 (<10% positive staining of tumour cells), 2 (>10% but <50% positive staining of tumour cells) and 3 (>50% positive staining of tumour cells).

2.11 Section 10

2.11.1 Array CGH Background

Microarray based genomic analysis is a novel technique intended for rapid examination of human DNA for changes in copy number of specific sequences. The assay involves labeling of sample DNA with Cy3 (Perkin Elmer/NEN #NEL576) fluorophore. This is mixed with whole genomic reference DNA that is labeled with Cy5 (Perkin Elmer/NEN #NEL577) fluorophore and co-hybridized to a microarray in the presence of human Cot 1 DNA to suppress hybridization of labeled probe to repeat sequences. The microarray contains target clone DNA (P1, PAC or BAC clones) representing regions that are important in cytogenetics and oncology. See Appendix 1 for a complete list of annotated clones. DNA clones comprising the desired target sequences are arrayed in target spots of approximately 75-125 μm in diameter, whereby 3 target spots represent each clone.

Following hybridization and removal of unhybridised probe, target spots are counter-stained with a blue fluorophore included in the Array DAPI mounting solution and analyzed using the GenoSensor Reader System™. The GenoSensor Reader System is a large-field multicolor fluorescence imaging system which captures an image of the hybridized chip in 3 color planes: Cy 3, Cy5 and DAPI blue. The included software automatically identifies each spot and, by analysis of the set of Cy3/C5 ratios on all targets, calculates the ratio most representative of the modal DNA copy number of the sample DNA. For each target, the normalized ratio, relative to the modal DNA copy

number is calculated. This normalized ratio of a target indicates the degree of gain or loss of copy number compared with the sample's modal copy number. Detection of copy ratio changes will be highly dependent on the purity of DNA in the extracted tissue specimen. Even a very highly amplified gene will not appear as such if tumor DNA represents only a small fraction of total extracted tissue DNA. For this reason LCM was performed to ensure truly homogenous cell population

2.11.2 Labeling Protocol

1) In 1.5 ml DNase free snap lock tubes mix the following reagents for test and reference labeling reactions:

Cy-3 (Test)	Cy-5 (Reference)
Test DNA (25 ng/μL) 4.0 μL	
	Reference DNA (25 ng/μL) - 4.0 μL
TE Buffer 41.6 μL 41.6 μL	TE Buffer 41.6 μL 41.6 μL
2.5X Random Priming Mix 40 μL 40 μL	2.5X Random Priming Mix 40 μL 40 μL

2) Denature at 100°C, 10 minutes and then immediately place tubes on ice for 10 mins.

3) Spin down condensate inside tube, place tubes back on ice and add the following reagents test and reference tubes:

GenoSensor Array 300 Nucleotide Mix 10 μL
Klenow 2 μL
Cy-3 dCTP (1 mM) 2.4 μL (Test only)
Cy-5 dCTP (1 mM) - 2.4 μL (Reference only)

4) Mix gently by vortexing and quickly spin. Incubate at 37°C, 2 hours in the dark, then place tubes on ice.

2.11.3 DNase Digestion and Labeled Probe Purification

Add the following to the random priming reaction (100µL).

1. DNase Reaction Buffer -17 µL and 1:20 diluted DNase Amp Grade (prepared fresh on ice)- 3 µL.

Then:

1) Incubate at 15°C, 1 hour in the dark and then place tubes on ice.

2) Quench reactions with 6 µL of Stop Buffer and then vortex.

3) Prepare a MicroSpin S-200 HR column (Amersham) as described in the manufacturer's protocol (Procedure A: General Protocol).

4) Slowly apply 126 µL of labeled probe to the column and spin the column according to the manufacturer's protocol.

5) Add 0.1 volume of 3 M Sodium Acetate (12 µL) to the column.

6) Add 1 µL of precipitation reagent, vortex briefly.

7) Add 2.5 volumes of cold (-20°C) 100% Ethanol (350 µL), vortex briefly.

8) Incubate at – 20°C, 1 hour. Then centrifuge at 16,000 rcf, 30 minutes at 4°C

9) Remove supernatant and then air-dry pellets

10) Resuspend pellets in 4.0 µL of 10 mM Tris pH 8.0, vortex and incubate at room temperature for 30 min prior to setting up gel, hybridization, or storage.

11) Quickly spin the tubes to collect the sample at the bottom of the tube.

2.11.4 Checking the Labeled DNA

Check the labeled probe on a 2% agarose gel. If there is no DNA trailing above 200 base pairs, do not use the probe in array hybridization. See Figure 2.10

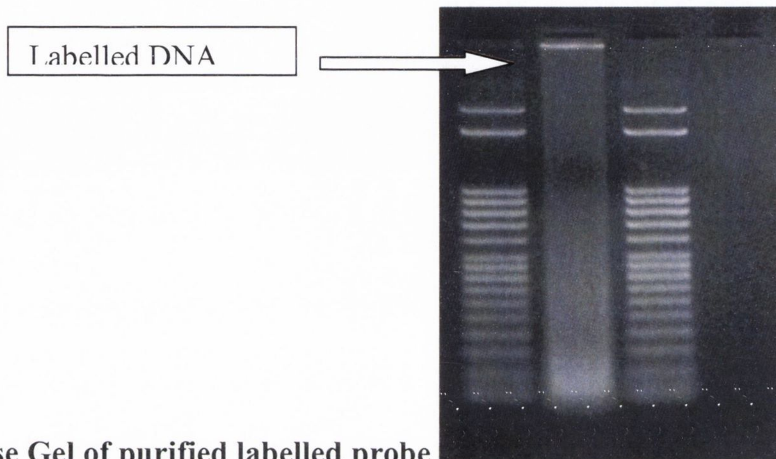


Figure 2.10 Agarose Gel of purified labelled probe

2.11.5 Hybridization Protocol

<u>Preparing the Reagents</u>

20X SSC Mix thoroughly 66 g (entire bottle) of 20X SSC in 250 mL purified H2O.

2X SSC/50% formamide wash. Mix thoroughly 105 mls formamide, 21 mls 20X SSC and 84 mls purified H2O. Pour enough formamide wash solution into 3 Coplin jars (labeled ‘1’, ‘2’, ‘3’) to cover the arrays but not so the labels are immersed.

1X SSC Add 50 mL 20X SSC to 950 mL purified H2O. Pour wash solution into each of four Coplin jars with lids. Label the jars ‘4’, ‘5’, ‘6’, and ‘7’.

Array DAPI Solution Store the Array DAPI solution at -20°C in the dark. Allow solution to warm to room temperature prior to pipetting.

Preparing the Hybridization Solution

Pre-warm the hybridization buffer to 37°C for 30 min, vortex and spin prior to use to ensure that the solution is uniform. Note: It is extremely important that the hybridization buffer is well dissolved. Due to high concentration of Cot-1 DNA in hybridization buffer, the solution may appear cloudy or opalescent. Some cloudiness in the buffer after pre-warming is acceptable as long as there are no visible clumps of precipitate. Quickly spin the tube to collect the solution at the bottom of the tube. Hold the hybridization buffer at 37°C until immediately before use.

Combine the following in an Eppendorf microfuge tube, vortex and spin:

1. Microarray Hybridization Buffer 25 μL.
2. Test DNA Probe (Cy-3 dCTP) 2.5 μL
3. Reference DNA Probe (Cy-5 dCTP) 2.5 μL

Hybridization

1. Place the microcentrifuge tube containing the hybridization mixture into an 80°C water bath and incubate for 10 minutes to denature the DNA. Remove from the water bath and centrifuge at 12,000 – 16,000 rcf for 5 seconds.

2. Quickly transfer the microcentrifuge tubes containing the hybridization mixture to a 37°C incubator or covered heating block, incubate in the dark for 1 hour.

3. Remove the necessary number of microarrays from their protective packaging using the tear marks. Place the microarrays in a 37°C dry air incubator for 30 min prior to use.

4. Place a paper towel folded in half on the bottom of a sealable (air-tight) box. Add 14-16 ml of 50% formamide/ 2XSSC wash solution to saturate the paper towel. Place the box in 37°C dry air incubator for 30 min – 1-hour prior to use.

5. Remove the cover protecting the microarray hybridization area.

6. Mix gently and quickly spin the hybridization solution. Add the hybridization mixture onto the corner of the array. Do not touch the pipette tip on the DNA array area or introduce air bubbles (leave the microarray on the slide warmer).

7. Using forceps, carefully remove a clean Hybridization Coverslip from the bag. Holding the Hybridization Coverslip at an angle, contact the hybridization solution on the microarray with the painted side of the Hybridization Coverslip touching the solution and slowly lower the Hybridization Coverslip to ensure that no air bubbles are introduced.

8. Place microarray in pre-warmed box in 37°C incubator

9. Repeat steps 6 -8 for the remaining microarrays

10. Hybridize arrays for 60-72 hours

Washing the Microarrays

Optimally, washes should be done in reduced light if possible. Array labels are resistant but not impervious to formamide wash solution and immersion should be as limited as practical. Replace wash solutions after washing about 30microarrays, or two weeks of use. **Adequate washing is vital to assay performance**. **The microarrays should not**

be allowed to dry at any step upon completion of the hybridization. Never apply force when washing chips.

Place formamide washes (#1-3) in 40°C water bath allowing them to equilibrate approximately 30 minutes. Prior to washing arrays use a thermometer to check washes are at 40°C±1°. Place Coplin jars (#4 -7) containing 1X SSC and another jar (#8) containing distilled water at room temperature.

1. Remove one microarray from the box in the incubator.

Using fine tip forceps, remove the Hybridization Coverslip by grabbing the overhanging edge and gently lifting up. Immediately immerse the microarray in 50% 2XSSC/formamide wash solution at 40°C (Coplin jar #1) and agitate briefly.

2. Repeat for remaining microarrays (up to 5), one at a time. Incubate the microarrays in Coplin jar #1 for 10 minutes

3. After incubation, agitate each microarray in turn and transfer to Coplin jar #2 (formamide wash at 40°C). Incubate microarrays in Coplin jar #2 for 10 minutes

4. After incubation, agitate microarrays and transfer, in succession as above, to Coplin jar #3 (formamide wash at 40°C). Incubate microarrays in Coplin jar #3 for 10 minutes

5. Agitate as above and transfer to Coplin jar #4 (1X SSC at room temperature). Incubate microarrays in Coplin jar #4 for 5 minutes

6. Repeat step 5 for Coplin jar #5, #6 and #7.

7. After removing array from last wash, rinse in distilled H2O for 1-2 seconds.

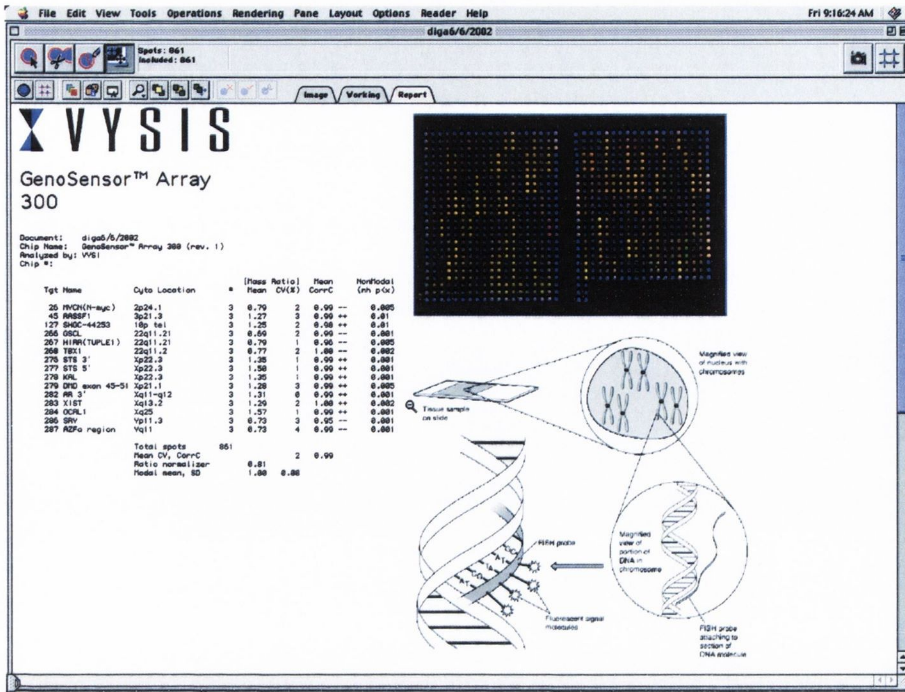
8. Briskly shake chip twice to get rid of the excess water ensuring the array area itself remains wet.

Quickly apply coverslip-containing DAPI mounting solution to the wet chip. Chips should be stored in the dark for 45 minutes prior to reading on the Genosensor instrument.

2.11.6 Image analysis

The Genosensor Reader Software automatically captures 3 images of each microarray, specific for the DAPI (counterstain), the Test DNA (green), and the reference DNA (red), respectively. The software uses the blue image to identify target spots and their location in the grid.

Once all spots are identified, the program analyzes each pixel within each spot for its intensity in each of the remaining colour planes. An algorithm is employed to determine the local background for each of these colours, which is subtracted from the intensity of each colour. The background-corrected intensities are then used to determine the ratio of Test Intensity (green) to Reference Intensity (red), which after further normalization can be used to estimate the relative abundance of a specific target sequence in the test DNA. The software generates a report page for each array. (Figure 2.11)



Document: diga6/6/2002
Chip Name: GenoSensor™ Array 300 (rev. 1)
Analyzed by: VYSI
Chip #:

Tgt	Name	Cyto Location	#	[Mass Ratio] Mean	CV(%)	Mean CorrC		NonModal (nh p<x)
26	MYCN(N-myc)	2p24.1	3	0.79	2	0.99	--	0.005
45	RASSF1	3p21.3	3	1.27	3	0.99	++	0.01
127	SHGC-44253	10p tel	3	1.25	2	0.98	++	0.01
266	GSCL	22q11.21	3	0.69	2	0.99	--	0.001
267	HIRA(TUPLE1)	22q11.21	3	0.79	1	0.96	--	0.005
268	TBX1	22q11.2	3	0.77	2	1.00	--	0.002
276	STS 3'	Xp22.3	3	1.35	1	0.99	++	0.001
277	STS 5'	Xp22.3	3	1.50	1	0.99	++	0.001
278	KAL	Xp22.3	3	1.35	1	0.99	++	0.001
279	DMD exon 45-51	Xp21.1	3	1.28	3	0.99	++	0.005
282	AR 3'	Xq11-q12	3	1.31	0	0.99	++	0.001
283	XIST	Xq13.2	3	1.29	2	1.00	++	0.002
284	OCRL1	Xq25	3	1.57	1	0.99	++	0.001
286	SRY	Yp11.3	3	0.73	3	0.95	--	0.001
287	AZFa region	Yq11	3	0.73	4	0.99	--	0.001
		Total spots	861					
		Mean CV, CorrC			2	0.99		
		Ratio normalizer		0.81				
		Modal mean, SD		1.00	0.06			

Figure 2.11 Genosensor array 300 Report page

2.12 Section 11

2.12.1 Applied Biosystems expression microarray system

The Applied Biosystems Expression Array System is based on a microarray design that represents the whole human genome, utilises current transcript data and relies entirely upon gene annotations that have been validated by experts in human curation. Each probe is part of a relational database that includes both Celera Genomics annotations and those in the public domain. Combined with specially developed chemiluminescent chemistries, this complete system delivers greater probe and detection sensitivity than previous generations of microarray systems. In addition annotation information for all of the 31,097 human genes that are represented on the microarray is included in an Oracle® database that is provided with the 1700 system. The manufacturers suggest that the result is a complete system that is capable of rapid and accurate analysis of microarray data for gene expression research. 'Follow on' experiments from microarrays can be achieved by linking to quantitative real time PCR TaqMan® probe based assays that enable microarray data validation, absolute quantitation of transcript production and investigation of alternative splicing events.

The Applied Biosystems Expression Array System consists of an Analyser (Applied Biosystems 1700 Chemiluminescent Analyser) that can image arrays in chemiluminescence, to survey and measure the gene expression at very low levels and

in fluorescence, to locate and auto-grid features. The 1700 is equipped with a state-of-the-art high resolution, large format CCD camera. The cooled CCD is back-illuminated for high efficiency and has very low read noise. This, coupled with the low background from chemiluminescence, results in very high sensitivity.

The microarrays are sealed in a pre-assembled cartridge and contain oligonucleotides with a feature diameter of <180 μm, and a space of >45 μm (edge-to-edge) between each feature. The oligonucleotides target transcripts in each gene of the human genome. Oligonucleotide probes are designed to ensure maximal specificity. Prior to microarray manufacture, all probes undergo analysis by mass spectrometry for quality control. All Applied Biosystems microarrays utilize 60-mer oligonucleotides (oligos) as DNA probes. Oligos of this length offer the best combination of sensitivity and specificity when compared to microarrays containing either shorter oligos or cDNA probes. 60-mer oligos offer the good single-base hybridisation specificity that is expected from shorter oligos, and the strong sensitivity of longer fragments expected from cDNA arrays.

2.12.2 Reverse transcription-In vitro transcription (RT-IVT)

Excellent results may be achieved from as little as 500ng of starting total RNA by using the Applied Biosystems Chemiluminescent RT-IVT Labelling Kit. The chemistry of the RT-IVT Kit exploits the Eberwine (Van Gelder, 1990) linear amplification procedure, which converts mRNA into cDNA and then into amplified RNA, in a manner that allows up to 10^6 amplification of the starting material. The reverse transcriptase

incorporates deoxynucleotides and digoxigenin-dUTP (DIG-dUTP) in the synthesis of single-stranded cDNA from sample RNA and RT Labelling Control RNA and is a modified version of M-MLV reverse transcriptase with no RNase H activity. It also provides longer cDNA transcripts and higher yields than the wild type enzyme. Multiple transcription rounds result in the production of DIG-labelled cRNA (see Fig 2.12). The resultant digoxigenin-labelled cDNA or cRNA is specifically hybridised to the Applied Biosystems microarray.

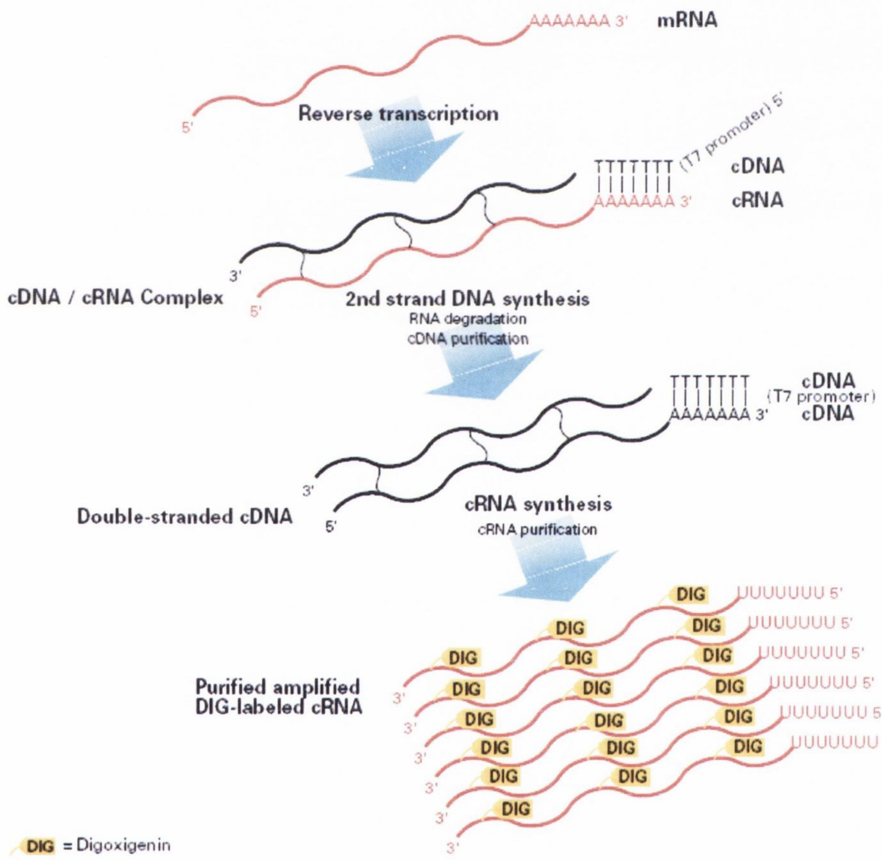


Figure 2.12 RT-IVT DIG-labelling of sample mRNA

2.12.2.1 RT-IVT labelling protocol

RT

1) Pipette the following components into 0.2ml MicroAmp reaction tube on ice: 2μl T7-Oligo (dT) primer, 4μl Control RNA, RNA sample (1-10μg) and nuclease-free water to 15μl. Heat the RNA and primer mixture in a thermal cycler to 70°C for 5min followed by a 4°C hold. After the run, place the tube on ice.

2) Add the following components to the reaction tube on ice and mix thoroughly by pipetting: 2μl 10X 1st Strand Buffer Mix and 3μl RT Enzyme Mix. Perform reverse transcription in a thermal cycler: 25°C 10min, 42°C 2hr, 70°C 15min and 4°C hold.

3) After the run, place the tube on ice and add the following components to the cDNA mixture: 30μl 5X 2nd Strand Buffer Mix, 5μl 2nd Strand Enzyme Mix and 95μl nuclease-free water. Perform second strand synthesis in a thermal cycler: 16°C 2hr, 70°C 15min and 4°C hold.

Purifying cDNA

1) In a 1.5ml nuclease-free microcentrifuge tube, combine: 150μl DNA Binding Buffer and 150μl 2nd strand synthesis reaction mix. Insert a DNA purification column into a 2.0-ml receptacle tube. Transfer the reaction-DNA Binding Buffer mixture (300μl) to the column and centrifuge at 10,000rpm for 1min.

2) Remove the column from the tube, discard the liquid, and reinsert the column into the tube. Add 700μl DNA Wash Buffer to the column and centrifuge at 10,000rpm for 1min.

3) Remove the column from the tube, discard the liquid, and reinsert the column into the tube. Add 700µl DNA Wash Buffer to the column and centrifuge at 10,000rpm for 1min.

4) Remove the column from the tube, discard the liquid, and then reinsert the column into the tube. Close the tube, then centrifuge the empty column and tube at 10,000rpm for 1 minute. Transfer the column to a new 1.5ml elution tube.

5) Pipette 30µl of DNA Elution Buffer onto the fibre matrix at the bottom of the column and allow to stand at room temperature for 1min. Centrifuge the column and tube at 10,000rpm for 1min. Repeat elution step twice more for a final elution volume of 90µl.

IVT labelling

Calculate the volume of cDNA output required, based on the amount of RNA input:

Amount of RNA input	**Volume of cDNA output required**
≤2µg	90µl
>2µg	150 ÷ (µg of total RNA input)

If the volume of cDNA output required is:

Greater than 24 µl: Use a vacuum concentrator to concentrate the required volume of cDNA output to 24 µl.

Less than 24µL: Add nuclease-free water to the required volume of cDNA output until the total volume is 24 µL.

Add the following IVT components to the 24 µl cDNA output at room temperature: 8µl 5X IVT Buffer Mix, 4µl DIG-UTP (approximately 14 nmol) and 4µl IVT Enzyme Mix.

Perform IVT in the thermal cycler: 37°C 9hr followed by a 4°C hold.

Purifying cRNA

1) In a new 1.5ml nuclease-free microcentrifuge tube, combine and then vortex briefly to mix: 20µl nuclease-free water and 40µl IVT reaction.

2) Add and mix by pipetting: 200µl RNA Binding Buffer and 140µl 100% ethanol.

Insert an RNA purification column into a 2ml receptacle tube, add the IVT reaction-RNA Binding Buffer-ethanol mixture (400µl) to the column, and centrifuge at 10,000rpm for 1min. Discard the flow-through.

3) Add 500µl of RNA Wash Buffer to the column and centrifuge at 10,000rpm for 1min.

Discard the flow-through. Repeat this. Discard the flow-through. Close tube, centrifuge at 10,000rpm for an additional minute. Transfer the column to a new 1.5ml elution tube.

4) Pipette 50µl of RNA Elution Buffer onto the fibre matrix at the bottom of the column and incubate at room temperature for 2min. Centrifuge at 10,000rpm for 1min for an elution volume of 50µl.

5) Pipette 50µl of RNA Elution Buffer onto the fibre matrix at the bottom of the column and incubate at room temperature for 2min. Centrifuge at 10,000rpm for 1min for a final elution volume of 100µl.

6) Resulting labelled cRNA should be stored on ice while quantity and quality are assessed using UV spectroscopy and agarose gel electrophoresis respectively. cRNA can be stored at -20°C for up to months or -80°C for long term storage

2.12.3 Chemiluminescent detection

The Applied Biosystems Chemiluminescence Detection Kit is used to visualize features that have digoxigenin-labelled cDNA or cRNA bound to the oligonucleotide probes. Visualization is achieved by incubating the microarray with an anti-digoxigenin alkaline phosphatase conjugate. Alkaline phosphatase hydrolyses a chemiluminescent substrate and emits light at a wavelength of ~458nm. The signal intensity is proportional to the mRNA level expressed in the cells (see Fig 2.13).

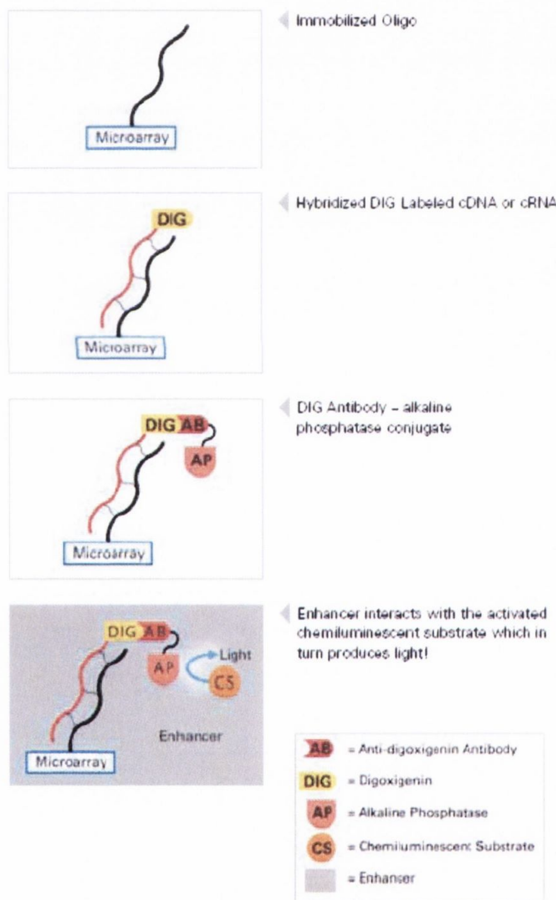


Figure 2.13 Chemiluminescent detection of bound DIG-labelled cRNA

A motorized chassis moves the microarray into the optical path of the analyser to detect chemiluminescent and fluorescent signals. System software can then relate the resultant intensities to gene expression, accurate feature registration, and data quality control. When the microarray is loaded into the 1700 analyser, the temperature is raised to 35°C in order for the enzymatic reaction to reach a steady state. Multiple images are then taken to bring the microarray into focus and to measure, in turn, the fluorescent and chemiluminescent signals

The microarray is imaged in both short (5 seconds) and long (25 seconds) read times to extend the linear dynamic range of the chemiluminescent signals (>1,000-fold). There are two imaging areas for each microarray. The total time required to image a microarray, including the pre-incubation and focus steps, is approximately 12 minutes. Light production on the microarray reaches a steady state within the first five minutes, while the microarray is being focused and brought up to 35°C. The chemiluminescent reaction emits light at a steady state for at least 60 minutes thereafter. The absence of an excitation background, together with the highly reproducible photon emissions, makes the signal-to-noise ratio produced from equivalent hybridisations superior to that found in alternative array systems.

2.12.3.1 Chemiluminescence detection protocol

Pre-hybridisation

Prepare pre-hybridisation mixture in a nuclease-free tube and vortex to mix: 150µl nuclease-free water, 330µl Hybridisation Buffer, 100µl Hybridisation Denaturant and

420µl Blocking Reagent. Transfer pre-hybridisation mixture into each microarray cartridge and incubate in a 55°C oven at 100prm for 1hr.

Fragmenting cRNA

Combine components into a 0.2mL MicroAmp® reaction tube on ice, then mix by pipetting: 10µl cRNA Fragmentation Buffer and 90µl 10µg DIG-labelled cRNA and nuclease-free water.

Heat the tube in a thermal cycler at 60°C for 30min.

Add 50µL of cRNA Fragmentation Stop Buffer, mix by pipetting and place on ice.

Hybridisation

For each microarray, prepare hybridisation mixture in a nuclease-free microcentrifuge tube: 100µl nuclease-free water, 170µl Hybridisation Buffer, 30µl Hybridisation Controls, 50µl Hybridisation Denaturant and 150µl fragmented cRNA targets.

Vortex the hybridisation mixture, then centrifuge the tube briefly.

Quickly transfer hybridisation mixture into each microarray cartridge, drying the port with lint-free tissue prior to sealing. Return cartridges to the oven and incubate at 55°C at 100rpm for 16hr.

Hybridisation washes

Add 600ml hybridisation wash buffer 1 to a clean wash tray.

Remove microarrays from their cartridges, decant liquid and submerge in wash buffer.

Agitate on the rocking platform shaker (tilt angle 10°; tilt speed 30 tilts back and forward per min) for 5min.

Add 600ml hybridisation wash buffer 2 to a clean wash tray. Transfer microarray to new tray, draining them in the process. Agitate on the rocking platform shaker for 5min.

Drain the buffer from the wash tray.

Add 400ml CL rinse buffer to the wash tray, making sure that all microarrays are submerged in buffer. Agitate on the rocking platform shaker for 5min.

Drain the buffer from the wash tray.

Add 400ml CL rinse buffer to the wash tray, making sure that all microarrays are submerged in buffer. Agitate on the rocking platform shaker for 5min.

Remove wash tray from the rocking platform shaker.

Microarrays may be left in CL rinse buffer at room temperature for up to 1hr.

Antibody binding

Combine components for the CL blocking buffer/antibody mixture in a nuclease-free tube and mix well by inversion. Do not vortex (N.B. per microarray): 2.8ml nuclease free water, 200μl Chemiluminescence Rinse Buffer Concentrate, 1ml Blocking Reagent and 15μl Anti-digoxigenin-AP.

Remove one microarray from the wash tray, shake and tap it gently, then place it vertically on the rocking platform shaker.

Immediately add 4ml CL blocking buffer/antibody mixture to the microarray.

Cover the arrays with the wash tray cover and agitate on the rocking platform shaker for 20min at room temperature.

Antibody washes

Add 500ml CL rinse buffer to a clean wash tray.

Decant CL blocking buffer/antibody mixture, shake and tap the microarray gently, then place it in the wash tray.

Cover the wash tray and agitate on the rocking platform shaker for 10min.

Drain the buffer from the wash tray.

Add 500ml CL rinse buffer to the wash tray and agitate on the rocking platform shaker for 10min.

Drain the buffer from the wash tray.

Add 500ml CL rinse buffer to the wash tray and agitate on the rocking platform shaker for 10min.

Remove wash tray from the rocking platform shaker. Microarrays may be left in the CL rinse buffer at room temperature for up to 3hr.

Chemiluminescent reaction

The chemiluminescent reaction is time-dependent. Perform this procedure with only one microarray at a time.

Place 100ml CL enhancing rinse buffer in a 10.2 X 12.7cm (4 X 5inch) tray.

Remove one microarray from the wash tray, decant the CL rinse buffer, shake and tap the microarray gently and place it in the CL enhancing rinse buffer.

Agitate on the rocking platform shaker for 10 min.

Remove the microarray from the small tray, decant CL enhancing rinse buffer, shake and tap the microarray gently, then place it vertically on the rocking platform shaker.

Quickly add 4ml of Chemiluminescence Enhancing Solution to the microarray.

Agitate on the rocking platform shaker for 20 min.

Place 100ml CL enhancing rinse buffer in a second small tray.

Decant the enhancing solution from the microarray, then shake and tap the microarray gently.

Place the microarray in the second small tray and agitate on the rocking platform shaker for 5min.

Remove the microarray from the small tray, decant the CL enhancing rinse buffer, then shake and tap the microarray gently.

Add 3.5ml of Chemiluminescence Substrate to the microarray.

The chemiluminescent reaction is time dependent. After you perform this step, proceed with performing CL detection immediately on the ABI 1700 Chemiluminescent Analyser.

2.12.4 Array controls

Labelling kit controls are included with both the Applied Biosystems Chemiluminescent RT Labelling Kit and the RT-IVT Labelling Kit. These controls provide quality information on RT and RT-IVT Kit enzyme activity and DIG-label incorporation efficiency for each experiment (see Fig 2.14).

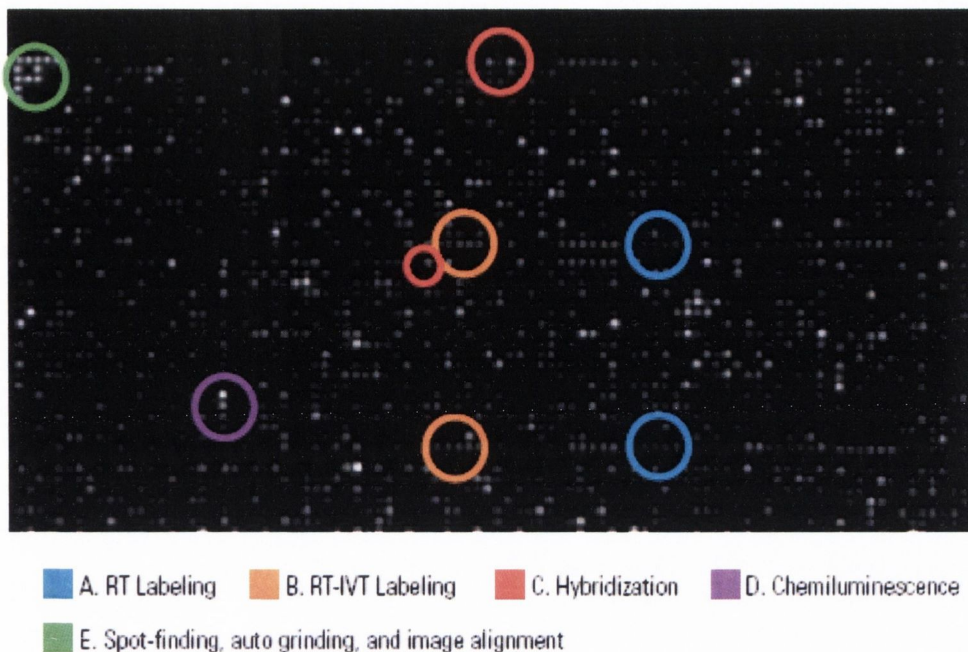


Figure 2.14 Examples of controls used in Applied Biosystems microarrays

<u>Assay Background Controls</u>

Purpose: The assay background controls provide information about non-specific background signals from all probes on the microarray.

Definition: The assay background controls consist of oligos designed to have low cross reactivity with the genome being tested. The design strategy:

1. Applied Biosystems performed BLAST® (Basic Local Alignment Search Tool) searches with random sequences.

2. Sequences were ranked according to low predicted reactivity with the genome sequence of interest.

3. Sequences were refined further based upon empirical data generated at Applied Biosystems. The microarray contains 196 features, 2 copies of 98 unique 60-mer probes

placed throughout the array. Function Non-specific binding to the assay background controls results in signal. Non-specific binding can result from cross-hybridisation, binding of alkaline phosphatase enzyme, or other chemical processes that lead to background chemiluminescent signal. The software calculates the amount of background chemiluminescent signal and uses the correction to measure all other chemiluminescent signals on the microarray.

Blank Features

Purpose: The amount of signal produced at the blank features provides information about:

- The background signal for every region of the microarray
- Cross-talk from adjacent spots

Definition: 1633 spots on the microarray are blank. The features are located between the image zones and between the 19x19 subgrids on the microarray. Function Non-specific binding of DNA or RNA, antibodies, alkaline phosphatase enzyme, or other chemical or mechanical processes can result in background signal. Signals from the blank features indicate the level of non-specific signal generated from the assay. The software calculates the amount of background chemiluminescent and fluorescent signals from the blank features and uses the correction to measure all other chemiluminescent and fluorescent signals on the microarray.

Control Ladders

Purpose: The control ladders are used to:

• Perform quality checks of array spotting and chemistry

• Compare reproducibility of spotting across batches and many arrays.

• Perform quality checks of chemiluminescent detection chemistry

• Demonstrate sensitivity and dynamic range of chemiluminescent detection chemistry.

Definition: The **fluorescent** fiducial ladder consists of probes labelled with LIZ® dye and spotted on the microarray in a five-fold dilution series: 625X, 125X, 25X, 5X, and 1X. The **chemiluminescent** fiducial ladder consists of DIG-labelled probes spotted on the microarray in a five-fold dilution series: 625X, 125X, 25X, 5X, and 1X. There are two copies of each ladder in each image zone.

Function: Because the labelled probes are attached to the microarray, fiducial control ladders are independent of labelling and hybridisation. Signal indicates that the chemiluminescent reaction was successful. Signal intensity variability indicates spotting and attachment efficiencies across various batches of manufactured arrays.

Hybridisation Controls

Purpose: The hybridisation controls can be used to:

• Indicate successful hybridisation

• Indicate hybridisation stringency

• As a spatial normalization control

Definition: The hybridisation controls consist of:

• Three DIG-labelled 60-mer oligo control targets supplied with the CL detection kit:

– HYB_Control_1_Ct, HYB_Control_2_Ct and HYB_Control_3_Ct

• Three unlabelled probes spotted on the microarray:

HYB_Control_1_Cp, HYB_Control_2_Cp, HYB_Control_3_Cp

Function: The DIG-labelled oligo targets are added to the microarray with the DIG-labelled cDNA or cRNA targets. The three hybridisation controls on the microarray are designed to hybridize to all of the DIG-labelled oligo targets:

- Presence of signal indicates hybridisation occurred.
- Signal strength indicates hybridisation stringency.

Internal Control

Purpose: The internal control can be used to:

- Grid all features on the microarray
- Normalise the CL signal

Definition: This control consists of:

- Internal control target (ICT) supplied with the CL detection kit: 24mer oligo labelled with LIZ dye
- Internal control probe (ICP) on the microarray: 24mer co spotted at every feature in the microarray that contains a 60mer gene expression probe

Function: The ICT is added to the hybridisation mixture. When hybridisation occurs, the ICTs hybridize to every ICP to produce a fluorescent signal.

IVT Controls

Purpose: When using the RT-IVT labelling kit, the IVT controls indicate whether the *in vitro* transcription labelling reaction worked and how well the reaction worked. The IVT controls are not used when using the RT labelling kit.

Definition: The IVT controls consist of:

• Three synthetic double-stranded cDNA with a T7 promoter and bacterial control gene sequences:

bioB — 1000-nt ds-cDNA , *bioC* — 750-nt ds-cDNA and *bioD* — 600-nt ds-cDNA

• A total of 120 features on the microarray:

– Fifteen probes on the microarray: five for each bacterial control gene, *BioB*, *BioC*, and *BioD*.

– Each probe is spotted eight times: four times in each image zone.

Function: The synthetic control cDNAs are added to the RNA sample when using the RT-IVT labelling kit. The control cDNA undergoes *in vitro* transcription in the presence of DIG to produce DIG-labelled cRNA. The DIG-labelled control cRNAs hybridize to the probes on the microarray and, after a successful chemiluminescent reaction, generate signal.

<u>Landmark Fiducials</u>

Purpose: The landmark fiducial controls are used by the system software to:

•Confirm that the placements of a subset of the features are in the correct position after manufacturing

• Determine the orientation of plates and of the array

• Define the edges and corners of the array

• Align the fluorescent and chemiluminescent images

Definition: The landmark fiducial controls consist of probes on the microarray:

• Probes labelled with LIZ dye and probes labelled with DIG

Function: The fluorescent landmark fiducial controls provide strong signals independent of labelling, hybridisation, and the chemiluminescent reaction.

After a successful chemiluminescent reaction, the chemiluminescent landmark fiducial controls provide strong signals independent of labelling and hybridisation. Also present during hybridisation is a 24-mer oligonucleotide labelled with the fluorescent LIZ® dye. This oligonucleotide is complementary to one that is co-deposited at microarray features during manufacture of Applied Biosystems microarrays. The fluorescent signal, which has a close spatial correlation with the chemiluminescent signal, locates all features on the microarray, even in the absence of gene expression products.

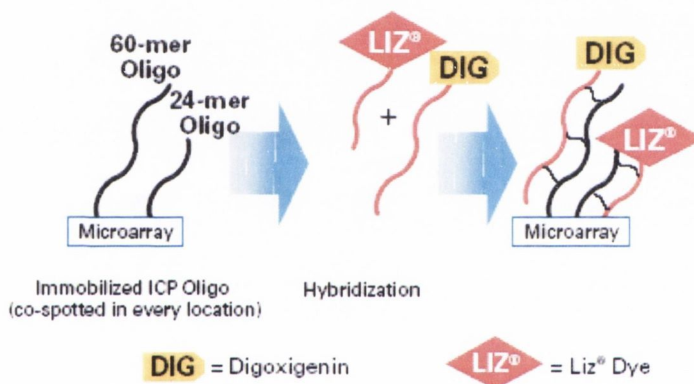


Figure 2.15 Fluorescence oligos used to image microarray features

RT Controls

Purpose: The RT controls indicate:

• With the RT labelling kit, whether the reverse transcription labelling reaction worked and how well the reaction worked.

• With the RT-IVT labelling kit, whether the reverse transcription reaction worked and how well the reaction worked.

Definition: The RT controls consist of:

• Three synthetic mRNAs with bacterial control gene sequences:

– *lys*— 1000-nt mRNA with poly (A) tail

– *phe* — 1400-nt mRNA with poly(A) tail

– *dap* — 1900-nt mRNA with poly(A) tail

• A total of 120 features on the microarray:

– Fifteen probes on the microarray: five for each bacterial control gene *DAP*, *LYS*, and *PHE*.

– Each probe is spotted eight times: four times in each image zone.

Function: The synthetic mRNAs are added to the reverse transcription reaction with the RNA sample when using the RT labelling kit or the RT-IVT labelling kit:

• With the RT labelling kit, the control RNA undergoes reverse transcription in the presence of DIG-dUTP to produce DIG-labelled cDNA.

• With the RT-IVT labelling kit, the control RNA undergoes reverse transcription and second strand synthesis into double-stranded cDNA. The control sequences then undergo *in vitro* transcription in the presence of DIG-dUTP to produce DIG-labelled cRNA. The DIG-labelled control cDNAs (RT labelling kit) or DIG-labelled control cRNAs (RT-IVT labelling kit) hybridize to the probes on the microarray and, after a successful chemiluminescent reaction, generate signal.

Spatial Normalization Controls

Purpose: The software uses the spatial normalization control signal to:

• Compare the signals within one image zone and calculate the percent coefficient of variation (%CV) within the image zone

• Compare the signals across the array and calculate the percent coefficient of variation (%CV) across the array

Definition: The spatial normalization controls consist of:

• Three DIG-labelled oligo control targets supplied with the CL detection kit

• HYB_Control_3_Cp, unlabelled probe spotted on the microarray

Function The unlabelled probe is spotted on at least one corner of each 19x19 subgrid on the microarray (117 features total).

The DIG-labelled oligo targets are added to the microarray with the DIG-labelled cDNA or cRNA targets. The spatial normalization control probe on the microarray is designed to hybridize to all of the DIG-labelled oligo targets.

2.15 References

Burchill SA, Perebolte L, Johnston C, Top B, Selby P. Comparison of the RNA-amplification based methods RT-PCR and NASBA for the detection of circulating tumour cells. British Journal of Cancer. 2002 Jan 7; 86(1): 102-9.

Fend F, Raffield M. 2000. Laser Capture microdissection in pathology. J Clin Pathol 53:666-672

Hibbitts S, Rahman A, John R, Westmoreland D, Fox JD. Development and evaluation of NucliSens basic kit NASBA for diagnosis of parainfluenza virus infection with 'end-point' and 'real-time' detection. Journal of Virology Methods. 2003 Mar; 108(2): 145-55.

Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanov EA, Belousov ES, Singer MJ, Walburger DK, Lokhov SG, Gall AA, Dempcy R, Reed MW, Meyer RB, Hedgpeth J. 3'-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. Nucleic Acids Res. 2000; 28(2): 655-61.

Lakowicz JR, Keating S. Binding of an indole derivative to micelles as quantified by phase-sensitive detection of fluorescence. J Biol Chem. 1983 May 10; 258(9): 5519-24

Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. 1990. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A. 87(5): 1663-7

Chapter 3

Head and neck squamous cell carcinoma in a cohort of young Irish adults

3.1 Summary

Classically, head and neck squamous cell carcinoma (HNSCC) is a disease of the older adult population, with exposure to tobacco and excess alcohol contributing to the vast majority HNSCC. While cancers of the pharynx and the larynx remain relatively stable, there have been recent reports of an increasing incidence of tongue cancer in young people. Quite a number of these young patients develop tongue cancer in the apparent absence of the primary risk factor, i.e. tobacco smoking.

The focus of this chapter was to study a cohort of HNSCC patients in the Republic of Ireland. It compares gender distribution, site distribution, smoking history, haemoglobin status, tumour stage, grade and nodal status of one hundred and thirty HNSCC patients from St. James Hospital Dublin, Ireland, thirty of whom are less than forty years old.

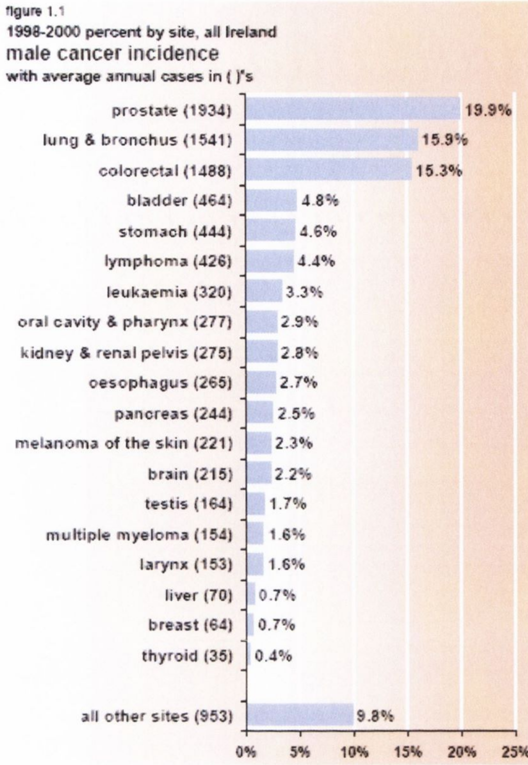
Concurring with other reports, this study finds a highly significant association between both the age and smoking status of the HNSCC patient and site of tumour occurrence. The pre-surgical haemoglobin status of HNSCC patients is assessed and although 15% were anaemic, there is no significant difference between the occurrence of anaemia in young and old HNSCC patients.

3.2 Introduction

3.2.1 Cancer in Ireland

Cancer is a major cause of death in Ireland. More people die from cancer than from heart disease, stroke, respiratory disease, or injuries and poisonings.

Over 19000 new cases of cancer are diagnosed each year, and cancer causes more than 11000 deaths. By and large, the burden of cancer weighs more heavily on men than on women. Specifically, the incidence rate for men is 20% higher than for women, while the mortality rate for men is more than 40% higher. Nevertheless, for women the incidence and mortality rates in Ireland are significantly higher than in the EU, while for men, they are equivalent to those in the EU. For both men and women, Ireland's incidence rates are lower than the US. However, for both men and women, Ireland's mortality rates are higher (National Cancer Registry in Ireland)



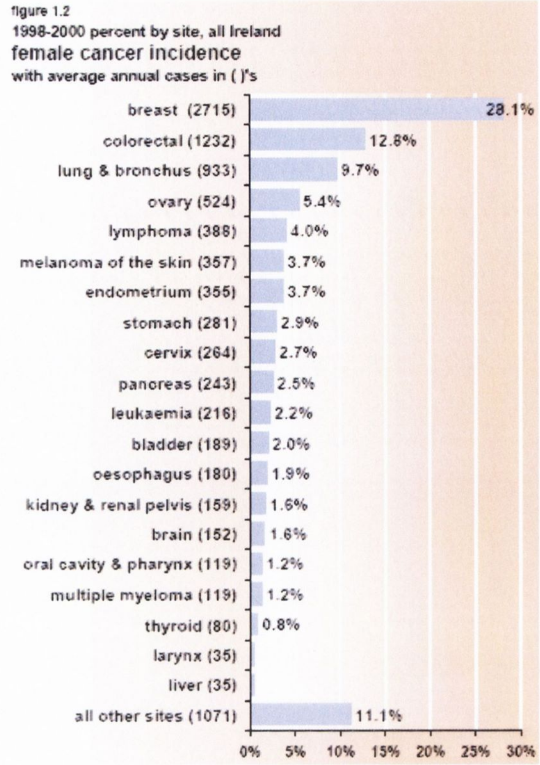
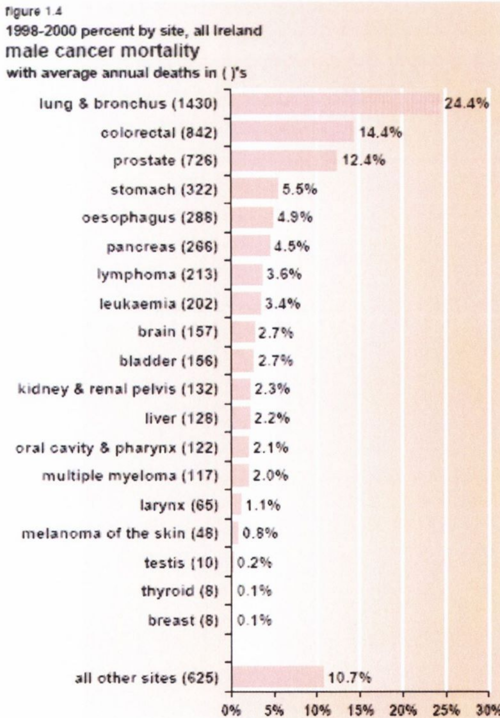


Figure 3.1 Incidence of cancers in Ireland by site, 1998-2000
http://www.allirelandnci.org/publications/report2/allireland1998-2000.pdf



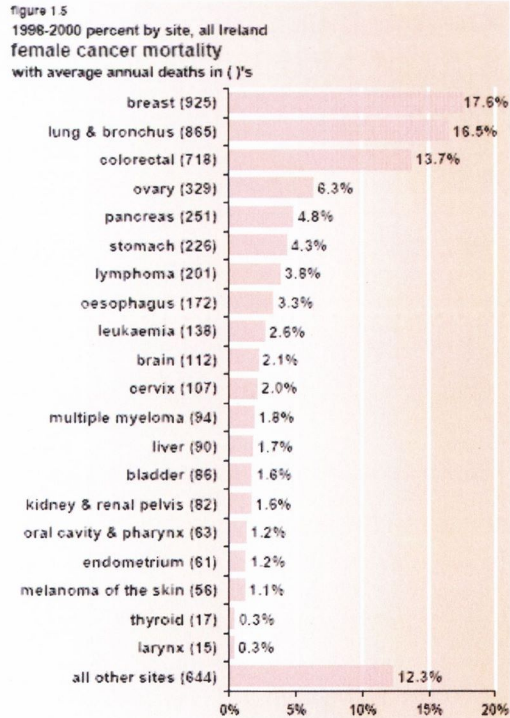
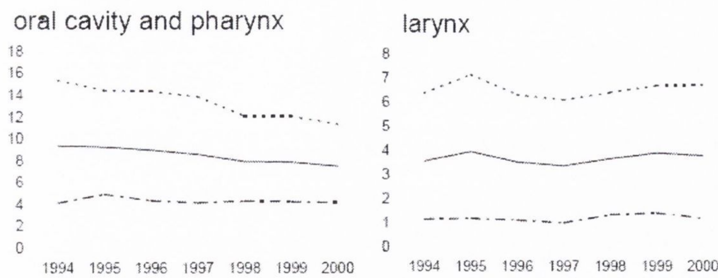


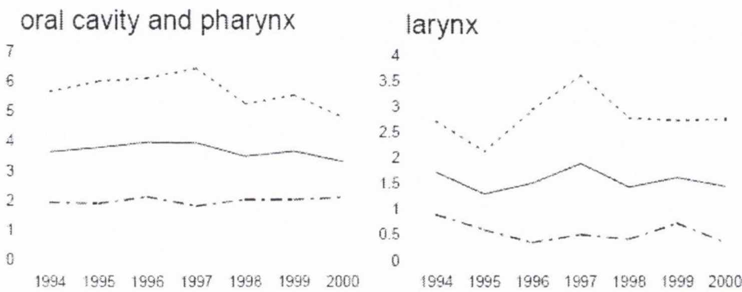
Figure 3.2 Mortality rates from cancer in Ireland by site, 1998-2000
http://www.allirelandnci.org/publications/report2/allireland1998-2000.pdf

3.2.2 Head and Neck Cancer in Ireland

Head and neck squamous cell carcinoma (HNSCC) is a major health problem worldwide accounting for 6% of all cancer (Vokes, 1993). Despite advances in treatment modalities, the prognosis still remains poor, with very little improvement in five-year survival from HNSCC over the last four decades. This tumour is typically seen in middle aged to elderly males and shows a strong association with tobacco and alcohol use, with these two risk factors contributing to approximately 75% of all oral and pharyngeal cancers in the United States (Blot, 1998). In most countries HNSCC incidence and mortality rates have been stable or increasing in the last four decades (World Cancer Report, 2003). From 1994 – 2000, the National Cancer Registry in Ireland reported very little change in incidence and mortality rates in these cancers (All Ireland Cancer Report, 2004) (Figure 3.2)



Incidence rates in Ireland 1994-2000



Mortality rates in Ireland 1994-2000

Figure 3.3 Incidence and mortality rates for oral, pharyngeal and laryngeal cancer in Ireland 1994-2000

The results of many studies suggest that since the 1960s HNSCC, particularly oral tongue cancer is increasing in young adults internationally (Schantz, 2002; Llewellyn, 2003; MacFarlane, 1992; Plesko, 1994). In a number of studies published on young patients with HNSCC it has been noted that there appears to be a distinct sub-set of younger patients, particularly females, who reported little or no exposure to the major risk factor i.e. tobacco (Llewellyn, 2004; McGregor, 1983; Tsukuda, 1993). Other putative aetiological agents include drug abuse, viral infections, sexual practices and diet, however, as discussed in Chapter 1, there is no good evidence of a strong association between any of these factors and the development of HNSCC in young patients. Chronic iron deficiency anaemia leads to mucosal atrophy and it has

been suggested that this may lead to an increased susceptibility to carcinogen exposure (Maier, 1994). There is a paucity of literature on this issue and its role in HNSCC cannot be discounted.

3.3 Aims

The objective of this chapter was to examine the variation in site of occurrence, smoking status, alcohol consumption, haemoglobin status and tumour stage and grade and nodal status across age groups in Irish patients with HNSCC.

3.4 Materials and methods

The medical records of all patients forty years of age or younger that were diagnosed with HNSCC between 1993 and 2003 at St. James's Hospital, Dublin, Ireland were reviewed. From the same HNSCC database, 100 patients over 40 years of age were randomly selected using a table of random sampling numbers.

The threshold for ones definition of 'young' has varied in the literature with a cut-off of 35, 40 and 45 years being used (Llewellyn, 2003; Llewellyn, 2004; McGregor 1983; Byers, 1975). In order to make the study comparable with the majority of the published literature, the cut-off age of 40 years was chosen. Young patients with SCC of the oral cavity, oropharynx, hypopharynx, and larynx diagnosed between 1993 and 2003 were included (ICD-10- C01-06, C09, 10, C12-14, C32). SCC of the lip was excluded, because many carcinomas at this site are skin rather than mucosal in origin, with a strong association with exposure to U.V. light. Nasopharyngeal carcinoma was excluded because of the well described association of EBV with this specific tumour.

In all 130 cases, there was a biopsy-confirmed diagnosis of SCC. For each patient, medical records were consulted and the following data obtained: age, sex, social history, site, stage, grade and nodal status of tumour. The use of tobacco was classified by type (cigarette smoking, cigar smoking, or smokeless tobacco use), and number of pack years (average number of packs per day times number of years). Only those that had never smoked were considered non-smokers, while any ex-smokers in the study were grouped with the smokers and their pack years recorded.

Similarly, with regard to alcohol intake, only those that never drank alcohol were

considered non-drinkers. Family history of head and neck cancer was noted if recorded in the chart. The haemoglobin level of each patient prior to surgery was assessed by reviewing the relevant reports. If the haemoglobin was below 11.5g/dl for a female and 13.5g/dl for a male, the patient was deemed anaemic. Each tumour was staged according to the TNM staging criteria of the International Union Against Cancer (UICC). Statistical analysis was carried out using the JMP® statistical package (SAS Institute Inc.). Chi-squared tests or Fisher's exact tests were carried out as appropriate. P values <0.05 were regarded as statistically significant.

3.5 Results

653 patients were diagnosed with a head and neck squamous cell carcinoma between 1993 and 2003 in St. James's Hospital and the Dublin Dental School and Hospital. This represents approximately 30% of HNSCC diagnosed in Ireland over this period. Thirty of the tumours were found to occur in the under forty-year old age group. These lesions accounted for 4.3 % of the total cases of HNSCC diagnosed in this hospital. A statistician was consulted regarding statistical power calculations. The cohort size suggested was unrealistic considering the rarity of the disease, therefore it was decided to forfeit statistical power in favour of the valuable preliminary data that would result from a small cohort study.

Using random tables, 100 patients over 40 years of age with HNSCC were selected from the database. In total, 130 (i.e 30 under 40 year olds and 100 over 40 year olds) histologic reports and patient charts were reviewed.

Clinicopathological feature	<40 years old N=30	>40 years old N=100	P value
Sex			
Male	**19**	**79**	
Female	**11**	**21**	**0.09**
Tobacco smoking			
Yes	**16**	**91**	
No	**14**	**9**	**<0.0001**
Alcohol			
Yes	**24**	**87**	
No	**4**	**11**	**0.74**
Site			
Alveolar ridge	**0**	**1**	
Buccal	**1**	**0**	
Floor of mouth	**2**	**10**	
Tongue	**23**	**19**	**<0.001**
Maxillary	**0**	**1**	
Palate	**0**	**2**	
Gingiva	**1**	**0**	
Retromolar	**0**	**5**	
Pharynx	**1**	**32**	
Larynx	**2**	**30**	
T stage			
T1	**15**	**26**	
T2	**8**	**31**	
T3	**5**	**32**	
T4	**2**	**11**	**0.08**
Nodal status			
Positive	**15**	**36**	
Negative	**15**	**64**	**0.20**
Differentiation			
Well	**5**	**5**	
Moderately	**21**	**74**	
Poorly	**4**	**21**	**0.09**
Anaemia status			
Yes	**3**	**16**	
No	**27**	**84**	**0.56**

Table 3.1 Correlation between age at diagnosis of HNSCC and clinicopathological features

3.5.1 Age and Gender

Thirty of the patients (male: female ratio = 1.7:1) were between the ages of 17 and 40 years (mean = 33). Because of the variability in the literature regarding the age cut-off (i.e 35, 40, 45 years old), the gender ratio of the under 35-year-old group was noted separately. In the under 35-year-old group the male: female ratio was 1.3:1. This ratio changed to 1.7:1 when the 35-40 year old group was included. Figure 3.4 illustrates the change in gender ratio through different age groups. Twenty-nine patients were Irish (Caucasian), and one female patient was a recent immigrant from the Philippines (Asian). All the over forty year olds were Caucasian. The mean age of the over forty year old group (N=100) was 61 (range 41-85) and the male: female ratio was 4.8:1 (Figure 3.5). The overall mean age of males (mean = 54) did not differ significantly from the mean age of the females (mean = 53) (Figure 3.5).

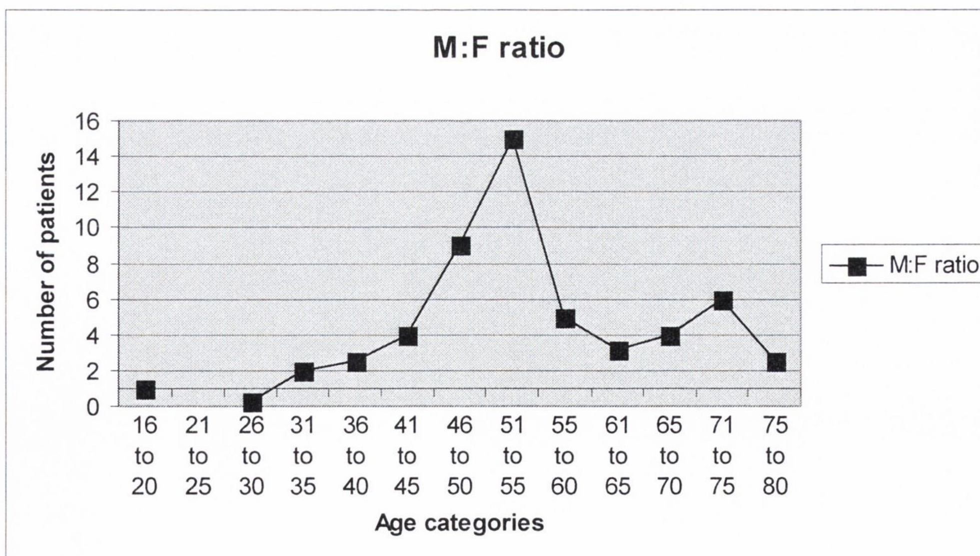


Figure 3.4 Male: female ratio in each age-group

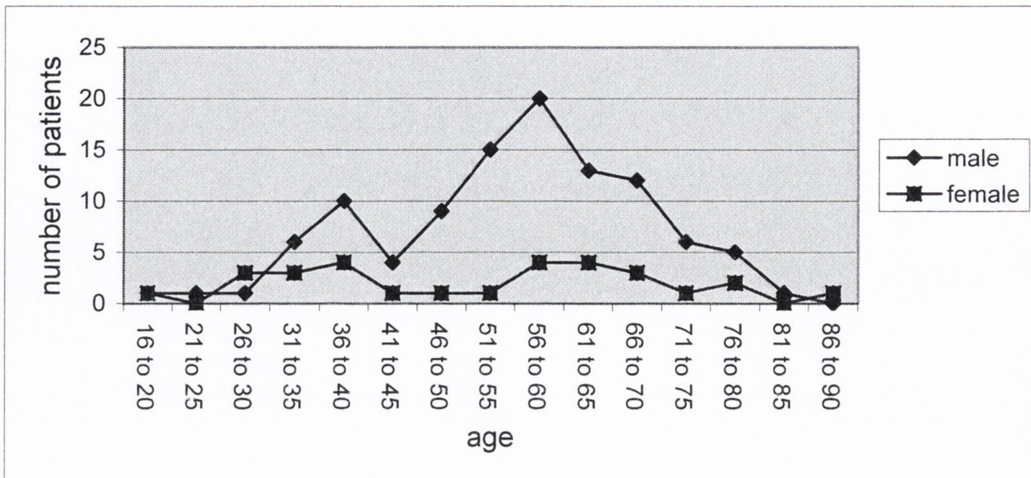


Figure 3.5 Age and gender distribution of cohort of HNSCC patients

3.5.2 Site of tumour and history of risk factors

A comparison of clinicopathological features of HNSCCs in under forty year olds and in over 40 year olds is summarised in Table 3.1. Carcinomas of the anterior two-thirds of the tongue were significantly associated with the young patients, whereas the older group were more likely to have pharyngeal or laryngeal tumours ($p<0.001$) (Figure 3.6). Eighty-two per cent of the 130 patients smoked or had smoked in the past and the frequency of tobacco smoking was significantly higher in the older age group ($p<0.0001$). Pack years were not consistently recorded in the medical notes and so results for amount of tobacco intake were inconclusive. Similarly the type of tobacco was not always specified, and so no valid results were obtained. No significant difference was found between the two groups with regard to alcohol consumption ($p= 0.74$). Family history of head and neck cancer was

recorded in one case, a seventeen year old male with both a grandmother and an aunt having had head and neck cancer.

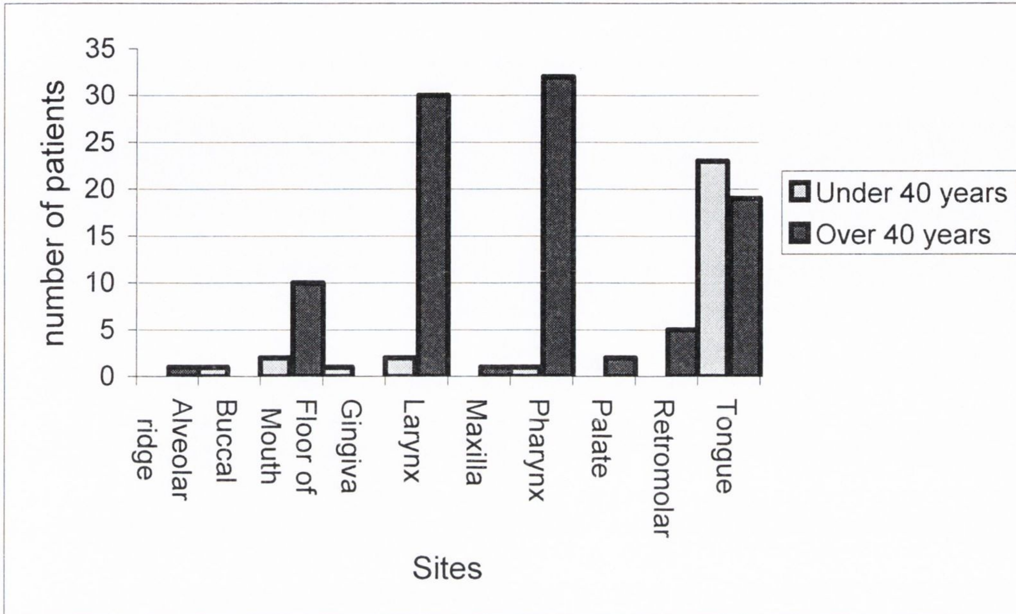


Figure 3.6 Distribution of primary HNSCC in under 40 yr olds and over 40 yr olds

3.5.3 Stage, grade and haematological status

On examination of the T Stage at diagnosis, it was found that although not significant, there was a trend towards lower stages at diagnosis in the younger group ($p=0.08$). Also approaching significance was the grade of tumour at diagnosis, with a higher proportion of the younger cohort having well-differentiated tumours ($p=0.09$). Nodal status was similar in both groups. Of the entire sample, 15% had anaemia on diagnosis, however there was no significant association between anaemia and age or sex ($p=1$, $p=0.5$ respectively).

3.6 Discussion

Recent data has indicated an increase in HNSCC occurring in young patients. Significant increases in incidence and mortality amongst younger males have been reported in the United Kingdom over the last 30 years, and similar trends have been found across Western and Eastern Europe (Schantz, 2002; Llewellyn, 2003; MacFarlane 1992; Plesko, 1994). These alarming trends have also been reported in the United States of America, Davis (1987) noted that the largest increase in tongue cancer since 1973 was among male 30-39 year olds, the rate increasing by 3 fold in this age group, and Schantz (2002) found that although the incidence rates for all the other cancers in younger Americans remained relatively stable, the incidence of tongue cancer increased significantly.
The focus of this chapter was to examine a cohort with HNSCC in the Republic of Ireland, and evaluated the association between age at diagnosis of the disease and a number of clinical and pathologic factors including site of tumour development and risk factors (Table 3.1).

In the Irish cohort examined in this study, a male: female ratio of 1.7:1 was reported in the under 40 year olds. Past publications on head and neck cancer and young people have chosen 35, 40, or 45 as upper threshold for ‘young patients’ and this has resulted in variable gender ratio results. To highlight awareness of this issue, in this chapter a male: female ratio of 1.3:1 in the under 35 year old group was reported. It was noted that this ratio changed to 1.7:1 when the 35-40 year old group was included. In contrast to these ratios, an average ratio of 4.8:1 was found in the over 40 year old. Figure 3.4 and Figure 3.5 illustrate how gender ratio changes as

age at diagnosis increases. While the worldwide male-female ratio of occurrence of HNSCC in over 40 year olds varies from 2-15:1 depending on both geographic location and anatomical sub site (World Cancer Report, 2003; Manuel, 2003) the gender distribution in young HNSCC patients appears to differ markedly from that of the older patients. McGregor *et al* (1983) reported a female preponderance of 3:1 in the under 30-year-old age group, and this was supported by Byers (1975), who found a significant number of those diagnosed with SCC of the head and neck under 30 years of age were female.

With regard to site of development of carcinomas in the head and neck, the mouth and anterior two-thirds of the tongue generally predominate in developing countries, in all age groups, whereas pharyngeal cancers tend to occur more often in developed countries (World Cancer Report, 2003; Lipkin, 1985). The over forty year olds in this study showed a predominance of pharyngeal and laryngeal cancer (Figure 3.6), and only 19% of the tumours occurred in the anterior two-thirds of the tongue in this age-group. Concurring with recent data being generated on young patients with HNSCC, in which the tongue has been reported as the dominant site (Schantz, 2002), 76% of the young patients in this study developed cancer of the anterior two-thirds of the tongue. While some previous reports have not done so, anterior two-thirds of the tongue was separated from the posterior one-third cohort in this thesis. This is an important segregation because of the different histological appearances of SCC in each of the sites (posterior one-third tumours are more likely to be basaloid SCC and less differentiated), and the possible association of HPV with posterior tongue tumours.

Tobacco smoking and alcohol use have long been implicated as the major risk factors for HNSCC in the developed world, with an estimated 41% of HNSCC in males attributed to smoking (World Cancer Report, 2003). There have been several reports of an absence of the more traditional risk factors in a significant proportion of young patients that are developing HNSCC, with Mackenzie *et al* (2000) concluding that 32% of their 109 under 40 years of age patients exhibited no risk factors, while 37% had one risk factor. Llewellyn *et al* (2004) used a detailed questionnaire format to analyse under 45 years olds with HNSCC regarding their exposure to various risk factors, including tobacco, alcohol, cannabis, betel quid, and fruit/ vegetable consumption. 26% showed little if any exposure to any major risk factors, while in 75% of the cases tobacco use and excessive alcohol consumption were recorded. The remarkable difference between the results of these two studies might be explained by the use of a higher cut-off (i.e 45 years old) in Llewellyns study.

Of the 17 % of never smokers in this study, we found that the majority of them were under forty years of age, and a striking 100% of the under 30 year olds within the group had never smoked. This finding is supported by other groups who have looked at the under 30-year-olds and found an absence of risk factors in 52% and 61% of them respectively (Sankaranarayanan, 1989; Newman, 1983). This lack of conventional risk factors in the younger group has prompted several groups to postulate that other factors may have an intrinsic role in the development of HNSCC in this setting.

Tobacco chewing and use of oral snuff are widespread in the United States. However, these are unlikely to be implicated in the increasing number of HNSCC

occurring in this young population, because the importation, manufacture and sale of smokeless tobacco is banned in Ireland. So far, there does not appear to be a trend towards the illegal use of these substances.

Several U.S case studies have reported an association between marijuana smoking and development of HNSCC in young patients (Schantz, 2002; Donald, 1986; Zhang, 1999). We were unable to obtain any definitive information regarding cannabis usage in this study.

Case-control studies have shown that a diet deficient in fruit and vegetables is consistently found in HNSCC patients and it has also been shown that frequent consumption of vegetables, citrus fruit, fish and vegetable oils are major features of a low-risk diet for cancer of the oral cavity and pharynx (MacFarlane, 1995; Bosetti, 2003). Notwithstanding this, Llewellyn (2003) highlights in a comprehensive review that there are no dietary evaluation studies focusing on young patients. Due to the retrospective nature of this study, we were unable to assess the dietary status of the group.

Anaemia in an untreated cancer patient occurs due to a number of reasons, related or unrelated to the malignancy. Defective marrow production may be caused by impaired production of erythropoietin, marrow infiltration, myelofibrosis, or marrow necrosis. Low iron availability may result from nutritional deficiency (especially iron and folic acid) that patients with large tumours are prone to have, while haemorrhage or haemolysis may cause increased red blood cell loss (Ludwig, 1998). How anaemia and the development of HNSCC influence each other is not

clear. It has also been suggested that chronic iron deficiency anaemia may lead to mucosal atrophy thus increased susceptibility to carcinogen exposure and subsequent development of cancer, and a greater incidence in tongue tumours has been reported in iron-deficient animals (Maier, 1994; Prime, 1983). A comprehensive European study on anaemia and cancer found that the prevalence of anaemia in untreated cancer patients was 39.7% (Ludwig, 2004). On review of haemogloblin tests performed at the time of diagnosis, 17 of the 130 patients (13%) were diagnosed with anaemia prior to surgery for the HNSCC. No significant gender difference was found between the 13% of males and the 12% of females with anaemia and similarly the percentage of anaemic young patients was only slightly less than the percentage found in the older group (10% and 14% respectively). Despite its recognition as a putative predisposing factor for HNSCC, very few conclusive studies have been performed on anaemia and oral cancer and further work needs to be carried out to investigate the relationship of iron deficiency anaemia to oral carcinogenesis.

3.7 Conclusion

While the majority of HNSCC patients are over 50 years old, an alarming rise in the incidence of tongue SCC in young adults has been reported in the US and in Europe. This chapter highlights the need for dedicated research to be carried out to investigate the possibility of a different biology of HNSCC occurring in under forty year olds compared to conventional HNSCC. The reason for the development of HNSCC in young patients without typical risk factors remains unclear, but because of the worldwide reports of increasing numbers of young people developing this disease, it is of paramount importance that we continue full pace in our search for a causative agent of HNSCC development on the younger population. While all studies focusing on young adults with HNSCC have small numbers, it is still important to pursue this line of research, because of the devastating effect diagnosis and treatment of HNSCC has on this age group. Moreover, the absence of traditional risk factors in this group suggests that this indeed may be a disease quite different from the conventional HNSCC.

3.8 References

All Ireland Cancer Statistics Second Report 1998-2000. 24th March 2005. http://www.allirelandnci.org/publications/report2/allireland1998-2000.pdf

Blot WJ, McLaughlin JK, Winn DM, Austin DF, Greenberg DS, Preston-Martin S, Bernstein L, Schoenberg A, Stemhagen A, Fraumeni JFJr. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Research 1988; 48(11): 3282-7

Bosetti C, Gallus S, Trichopoulou A et al. Influence of the Mediterranean diet on the risk of cancers of the upper aerodigestive tract. Cancer Epidemiology Biomarkers Prevention. 2003; 12(10): 1091-4.

Byers RM. SCC of the oral tongue in patients less than 30 years of age American Journal Surgery 1975; 130(4): 475-478

Davis S, Severson RK. Increasing incidence in tongue cancer in the U.S among young adults Lancet. 1987; 2(8564): 910-1

Donald PJ. Marijuana smoking--possible cause of head and neck carcinoma in young patients. Otolaryngology Head Neck Surgery. 1986; 94(4): 517-21.

Lipkin A, Miller RH, Woodson GE. Squamous cell carcinoma of the oral cavity, pharynx, and larynx in young adults. Laryngoscope. 1985; 95(7 Pt 1): 790-3.

Llewellyn CD, Johnson NW, Warnakulasuriya KA. Risk factors for oral cancer in newly diagnosed patients aged 45 years and younger: a case-control study in Southern England. Journal of Oral Pathology and Medicine. 2004; 33(9): 525-32.

Llewellyn CD, Johnson NW, Warnakulasuriya KAAS. Risk factors for squamous cell carcinoma of the oral cavity in young people. A comprehensive literature review. Oral Oncology 2001; 37:401-418

Llewellyn CD, Linklater K, Bell J, Johnson NW, Warnankulasuriya KA. Squamous cell carcinoma of the oral cavity in patients aged 45 years and under: a descriptive analysis of 116 cases diagnosed in the South East of England from 1990 to 1997. Oral Oncology. 2003 Feb; 39(2): 106-14.

Ludwig H, Eritz E. Anaemia in cancer patients. Seminal Oncology 1998; 25: 2-6

Ludwig H, Van Belle S, Barrett-Lee P, Birgegard G, Bokemeyer C, Gascon P, Kosmidis P, Krzakowski M, Nortier J, Olmi P, Schneider M, Schrijvers D. The European Cancer Anaemia Survey (ECAS): A large, multinational, prospective survey defining the prevalence, incidence, and treatment of anaemia in cancer patients. European Journal of Cancer. 2004; 40(15): 2293-306.

MacFarlane GJ, Boyle P, Scully C, Oral cancer in Scotland: changing incidence and mortality British Medical Journal 1992; 305 6862:1121-1123

MacFarlane GJ, Zheng T, Marshall JR et al. Alcohol, tobacco, diet and the risk of oral cancer: a pooled analysis of three case-control studies. Oral Oncology. 1995;Vol.31B, No.3, pp.181-187

MacKenzie J, Ah-See K, Thakker N et al. Increasing incidence of oral cancer amongst young persons: what is the aetiology? Oral Oncology 2000; 36(4): 387-389

Maier H, Weidauer H, Zoller J, Seitz HK, Flentje M, Mall G, Born IA. Effect of chronic alcohol consumption on the morphology of the oral mucosa. Alcohol Clinical Experimental Research. 1994; 18(2): 387-91.

Manuel S, Raghavan SK, Pandey M, Sebastian P. Survival in patients under 45 years of age with squamous cell carcinoma of the oral tongue. International J Oral Maxillofacial Surgery 2003; 32(2): 167-173

Mat

McGregor GI, Davis N, Robins RE. Squamous cell carcinoma of the tongue and lower oral cavity in patients under 40 years of age American Journal of Surgery 1983; 146(1): 88-92.

Newman, AN, Rice DH, Ossoff RH, Sisson GA. Carcinoma of the tongue in persons younger than 30 years of age. Archives of Otolaryngology. 1983; 109(5): 302-4.

Plesko I, MacFarlane GJ, Evstifeeva TV, Obsitnikova A, Kramarova E. Oral and Pharyngeal cancer incidence in Slovakia. International Journal Cancer 1994; 56: 1-6

Prime SS, MacDonald DG, Rennie JS. The effect of iron deficiency on experimental oral carcinogenesis in the rat. British Journal Cancer 1983; 47(3) 413-8

Sankaranarayanan R, Mohideen MN, Nair MK, Padmanabhan TK. Aetiology of oral cancer in patients less than or equal to 30 years of age. British Journal of Cancer. 1989; 59(3): 439-40.

Schantz SP, Yu GP. Head and neck cancer incidence trends in young Americans, 1973-1997, with a special analysis for tongue cancer. Archives of Otolaryngology Head Neck Surgery 2002; 128(3): 268-274

Stewart B.W and P. Kleihues. World Cancer Report. IARC press. Lyon 2003: 232-236.

Tsukuda M, Ooishi K, Mochimatsu I, Sato H. Head and neck carcinomas in patients under the age of forty years. Japanese Journal of Cancer Res. 1993; 84(7): 748-52.

Vokes EE, Weichselbaum RR, Lippman SM, Hong WK. Head and Neck cancer. New England Journal of Medicine 1993; 328(3): 184-19.

Zhang Z, Morgenstern H, Spitz M et al. Marijuana use and increased risk of squamous cell carcinoma of the head and neck. Cancer Epidemiology Biomarkers Prevention. 1999; 8(12): 1071-8.

Chapter 4

Microarray comparative genomic hybridisation analysis of oral and oropharyngeal squamous cell carcinoma

4.1 Summary

Genetic alterations have been recognised as important events in the carcinogenesis of head and neck SCC. A high level of genomic imbalance found in head and neck cancers indicates that chromosome aberrations represent a very important mechanism of genetic instability in these tumours. Comparative genomic hybridisation (CGH) studies have shown that a number of chromosomal regions are consistently altered, such as gains on 3q, 5p, 8q, 11q13, and losses on 3p and 9p. The technique of CGH has been improved by the introduction of array CGH and by the use of laser capture microdissection to guarantee pure populations of tumour cells.

In this chapter, array CGH profiles of laser microdissected oral tumours are described. Profiles of young patients with oral SCC were compared with profiles of old patients. In concurrence with the literature, tumours from the older cohort, irrespective of smoking status, manifested deletions involving 3p, and 9p21 and gains involving 3q, 5q, 7p, 8q, 11q and 20q. The most significant finding in this study was that the young nonsmokers showed very few changes overall, and the aberrations were not in the sites classically associated with oral cancer. Deletion of CDKN2A ($p16^{INK4}$) was completely absent in the young group, irrespective of smoking, while present in 50% of the older cohort. This study shows that there is far less genomic instability in young non-smokers with oral cancer than found in the typical oral cancer patient. These observations indicate that oral cancer presenting at a younger age particularly in nonsmokers has a different genomic profile to the classically described oral cancer.

4.2 Introduction

One of the key features in the pathogenesis of many solid cancers, especially head and neck cancer, is chromosomal instability, with gene gain or loss reflecting this genetic instability. Genetic alterations have been recognised as important events in oral carcinogenesis. Previous studies have analysed chromosomal alterations in head and neck and/or oral cancer (Hermsen, 1997; Brzoska, 1995; Bockmuhl, 1996) and a high level of conformity of specific patterns of DNA copy number gains and losses has been noted consistently, such as gains at 3q, 5q, 8q, and 11q, and losses at 3p and 9p (Wolff, 1999). Of the many genetic alterations found in oral or head and neck cancers, loss of heterozygosity (LOH) at 9p and 3p are believed to be among the earliest events in oral cancer development (Califano, 1996; Gollin 2001).

Although many molecular studies of head and neck SCC and oral SCC have been published, there is a paucity of specific studies involving younger patients. Jin (1999), in a study of 36 adults under 40 years of age found that the frequency of loss of heterozygosity (LOH) at markers on chromosomes 3, 9 and 17p was comparable with those reported in older patients. This is not surprising considering that most of the young people in this study had relatively long histories of tobacco and alcohol consumption. All CGH work on oral cancer so far in the literature has been directed at the typical oral cancer patient, and to date there are no references to array CGH studies of young patients who lack exposure to the usual risk factors.

It was hypothesised in this part of the thesis that by focusing on those who lack exposure to the usual risk factors a different pattern or even an absence of chromosomal aberration might be revealed, thus furthering our understanding of the development of oral cancer in young adults with no risk factors.

4.2.1 Comparative Genomic Hybridisation

Comparative genomic hybridization, first developed by Kallioniemi (1992) is a technique that uses only a small amount of DNA, and allows the survey of the genome and detects losses and gains in all chromosomes at once, without the need for cell culture. It has been widely used to screen for multiple chromosomal aberrations in a variety of cancers (Struski, 2002), and can help identify genes involved in cell growth and tumourigenesis. The technique is described in detail in Chapter 2.

The sensitivity of CGH depends on the level and size of the copy number changes thus a copy number increase of 50% should be detectable if the region is 2 megabase pairs (Mb) or larger, whereas a region (amplicon) of 250 kilobase pairs (kb) would need a 400% copy number increase to be detectable (Joos, 1996). This limitation prevents precise localisation of sequences of interest. When a deletion is 100% (no copies present), a resolution of 1-2Mb can be achieved. The sensitivity of CGH can be adversely affected by contamination of tumour material with normal cells that surround it therefore a tumour component representing a minimum of 70% of the DNA extracted from the tissue is highly desirable. The use of laser capture microdissection or another microdissection technique is indispensable by ensuring near 100% tumour content for CGH experiments.

4.2.2 Microarrays

The microarray, developed in 1995 by Schena is a foundational technology that allows the analysis of DNA sequence variation, gene expression, protein levels, tissues, and cells in an extensive parallel format. Microarrays are simply orderly arrangements of individual nucleic acid samples, which are immobilised in the form of a grid on a solid surface such as glass, chromium, silicon etc. Recognised limitations of conventional CGH have been overcome by coupling it to microarray technology. Instead of using metaphase chromosomes, the use of arrayed DNA fragments, such as large-insert genomic clones or cDNA clones, as hybridization targets has dramatically increased the resolution of CGH (Pinkel, 1998). Array-based CGH not only offers a higher resolution (down to ~ 100kb), but also better reproducibility, higher throughput, and allows direct mapping of aberrations to the genomic sequence. However, there are limitations. Array CGH is limited to a given number of genome loci and for this reason does not currently give information that is homogenous across the entire genome.

4.2.3 Vysis GenoSensor Array 300

The GenoSensor System simultaneously screens for gene copy number changes in two hundred and eighty seven targets spotted in triplicate, thus permitting the screening of proto-oncogenes, tumor suppressor genes, microdeletion syndrome gene regions and subtelomeric regions. In addition, the array includes clones which mark known areas of loss of heterozygosity in cancer and unique Subtelomeric sequences relevant to other genetic diseases. The targets are in the form of large –

insert genomic clones (Bacterial Artificial Chromosome (BAC) and P1-derived artificial chromosomes (PAC)). Appendix 1 details the clone names, cytogenetic location and Genbank accession number of these clones. Although the array provides an average of 40Mb coverage of each chromosome, it must be emphasised that this coverage is not homogenous along the chromosome and effectively amounts to 1% coverage of the entire genome. For this reason Vysis array 300 results and conventional CGH results are not easily comparable. However, the Vysis array remains a powerful tool for investigating the pathobiology of cancer.

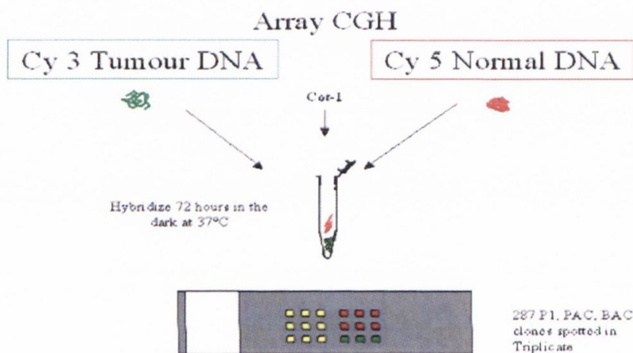


Figure 4.1 Array comparative genomic hybridisation

4.3 Aims

The aim of this chapter is firstly to identify genomic aberrations in oral and oropharyngeal cancer, using the Vysis Genosensor Array 300 CGH microarray platform, and compare the findings with the literature. The second aim is to compare the profile of genomic aberrations in four study cohorts;

1. Under 40 year old (young) nonsmokers**, 2**. Under 40 year old (young) smokers

3. Over 40 year old (old) nonsmokers, **4**. Over 40 year old (old) smokers

4.4 Materials and Methods

4.4.1 Patients

The St. James's Hospital and Federated Dublin Voluntary Hospitals Ethics Committee granted ethical approval for this study and informed consent was obtained from all patients prior to surgical treatment for oral and oropharyngeal cancer.

Fresh tumour samples were obtained from surgical resections of 20 previously untreated oral and oropharyngeal SCCs. Smoking status was ascertained from the clinical notes. While attempts were made to categorise smoking status using pack years, some pack year figures were not recorded in the notes. Smokers were defined as those that currently smoked tobacco, while nonsmokers were defined as those who never smoked tobacco. Ex-smokers were defined as having smoked in the past and thus were considered as part of the smoking cohort for the purpose of this study. Patients who were treated with radiotherapy prior to surgical resection were excluded. The study was restricted to squamous cell carcinomas of the oral cavity (Table 4.1). The acquired tumour tissue was snap frozen and stored at –80° until analysis. A consultant pathologist verified the presence of SCC on a frozen section, according to standard criteria. For each case 5 frozen sections were cut, and stained with Haematoxylin and Eosin.

	Under 40 years old	Over 40 years old
Number of patients		
Male	6	8
Female	4	2
Tumour site		
Tongue	7	7
Floor of mouth	1	1
Soft Palate	1	1
Alveolar ridge	1	1
Smoking status		
Smoker	4 (mean pack year = 16.5)*	6 (mean pack year = 35)*
Non-smoker	6	4
Stage		
T1	2	1
T2	5	4
T3	3	4
T4	0	1
N0	4	3
N+	6	7
Tumour differentiation		
Well	1	0
Moderate	8	7
Poor	1	3

Table 4.1 Clinicopathological details of all samples

4.4.2 Laser Capture Microdissection and DNA Extraction.

Frozen sections were cut and stained with H+E. Laser capture microdissection was performed using the Arcturus PixCell2™ Laser Capture microdissection system to ensure near 100% tumour cell content. Nucleic acid extraction was performed using Gentra protocols as described in Chapter 2.

4.4.3 CGH

Array CGH was performed as described in Chapter 2

4.4.3.1 Validation of technique using a control DNA (COSH)

CoSH is a control and validation DNA provided by Vysis. The DNA consists of a mixture of DNAs extracted from 3 tumor cell lines (COLO 320 HSR, SJSA-1 and BT-474). The mixture of cell lines has known amplifications of MYC, GLI and MDM2. Array CGH performed on COSH revealed amplifications of these genes in accordance with the data from Vysis. CGH of COSH was used as an external validation of operator technique for each batch of experiments

4.4.3.2 Validation of sensitivity of the system using sex mismatched test/reference

There is no internal control clone or group of clones with the Vysis Genosensor system that can be used for internal experimental validation. One potentially useful

way to incorporate an internal control is to use reference DNA of the sex opposite to that of the sample being arrayed (i.e. Male Sample, XY: Female Reference XX). For the sex markers there are five of 11 X chromosome loci and two Y chromosome loci that can be reliably used to determine the confidence in the results. The reliable loci on the X chromosome are STS 3' (Xp22.3), STS5' (Xp22.3), KAL (Xp22.3), AR 3' (Xq11-q12) and XIST (Xq13.2). The 2 Y loci on the array are SRY (Yp11.3) and AZFa (Yq11). Other markers are not always detected due to some homology between X and Y for those specific clones (Personal Communication, Dr. Teresa Ruffalo, Vysis) (Table 4.2).

To determine the variations in the ratios of the spots in normal control DNA, comparative hybridisations using test and reference DNA from 3 normal samples were performed. This approach has been used previously by many other groups (Albrecht, 2004; Schraml, 2003). Mean ratio (1.01) and standard deviation (0.09) of the 3 normal CGH arrays was calculated and a value of the mean ratio +/- two standard deviations was set as the cut-off level for normal gene copy number. Ratios higher than 1.19 and lower than 0.81 were considered to have copy gain and copy loss respectively. Any value outside this cut-off was regarded as a gain or loss. This correlated closely with Vysis recommended figures of 0.8 and 1.2 for a loss and a gain respectively.

Case number		#12166	#7558
Test/Reference		Female/Male XX/XY	Male/Female XY/XX
Target number/Name	Chromosome	T/R Ratio	T/R Ratio
276 STS 3'	X	1.24	0.69
277 STS 5'	X	1.48	0.67
278 KAL	X	1.30	0.70
282 AR 3'	X	1.84	0.51
283 XIST	X	1.25	0.73
286 SRY	Y	0.39	1.51
287 AZFa region	Y	0.47	1.62

Table 4.2 Validation by sex mismatch. T/R-Normalised Test to Reference ratio

Internal validation was performed in two of the cases (Cases numbers 12166 and 7558), Table 4.2. In Case 12166 (Female Test and Male reference) 100% of the expected gains on the X chromosome and 100% of the expected deletions on the Y chromosome were detected. In Case 7558 (Male Test and Female Reference) 100% of the expected losses and 100% of the expected gains were detected. In cases where the sexes were matched occasional aberrations were noted

4.5 Results

There were 20 cases that met inclusion criteria. Ten (50%) samples were taken from oral cancer in nonsmokers and ten (50%) were from smokers. All samples were from primary SCC of the oral cavity, with 14 (70%) located in the anterior two-thirds of the tongue.

CGH analysis revealed that genomic imbalances were a feature of every tested sample, although to varying extents across the sample cohorts.

Gain of genomic material was consistently more frequent in all groups than loss (Figure 4.3). The report generated by the Genosensor Array 300 tabulates many variables including the assigned target number, target clone name, cytogenetic location, the number of spots included, and the normalized test to reference intensity. Figure 4.2 shows the rendered image of the array surface generated by the Genosensor.

Figure 4.2 Microarray visualization showing a 9p21.3(MTAP) deletion (red) and 17q21 (BRCA 1) amplification (green). Only deletions or copy gains occurring in triplicate spots are considered to be valid

4.5.1 Average number of aberrations

The over 40-year-old cohort had a significantly higher number of aberrations per tumour (mean 99, range 32 – 164) than the younger group (mean 57, range 8-111) (p=0.02), with the young nonsmokers showing the lowest and the old smokers showing the highest number of aberrations (p=0.04) (Figure 4.3). A linear trend was evident on further breakdown of the age groups, with an increase in the average number of aberrations correlating directly with an increase in the age of the patient (Figure 4.4). On comparing smokers versus nonsmokers irrespective of age, the mean number of aberrations in smokers was higher than the mean number of aberrations in nonsmokers (Figure 4.5).

Regarding the T stage of tumours, the number of aberrations increased with increasing T stage, however this was not statistically significant (Table 4.3). At any one target, the frequency of gains in the old group ranged from 0% to 67%. The twenty-five most frequently aberrated loci in the classic HNSCC cohort (over 40 year old smoker) are tabulated below (Table 4.3).

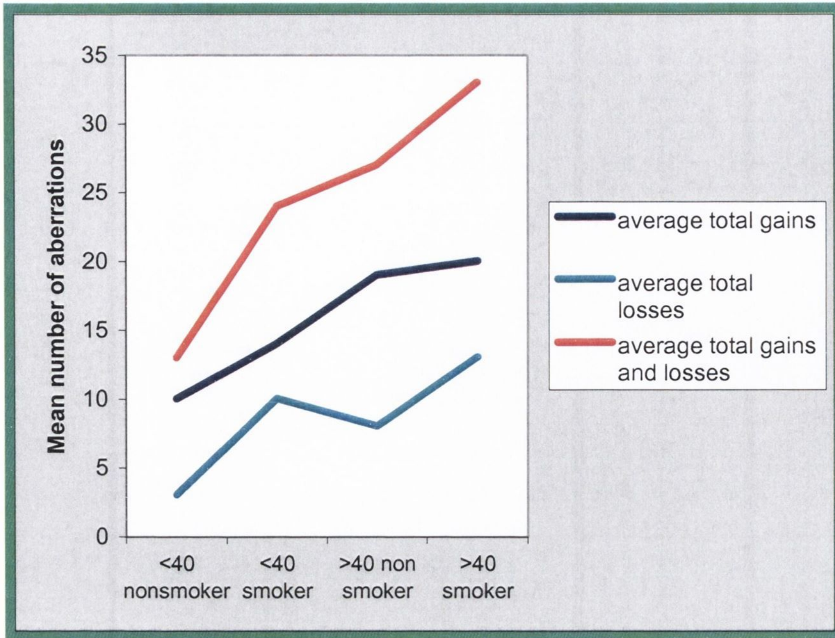


Figure 4.3 mean numbers of gains and losses in each of the four cohorts. A graph of the mean numbers of aberrations in each group shows a significant difference between the young nonsmokers and the old smokers three groups (p=0.04)

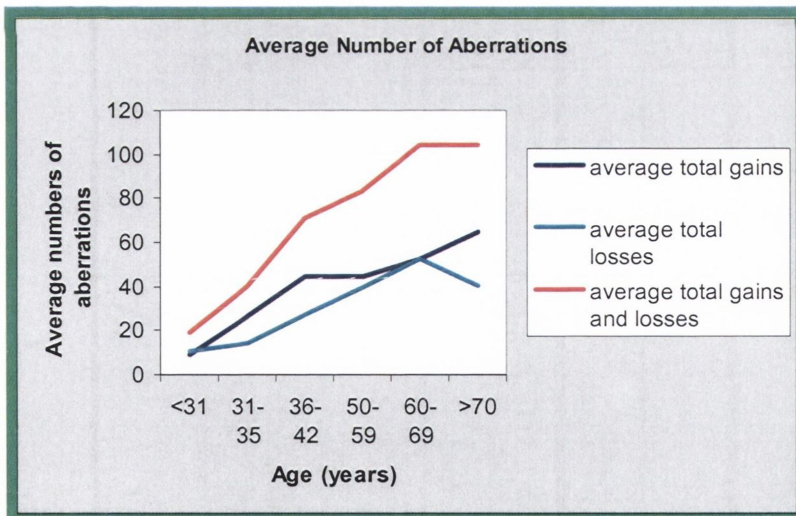


Figure 4.4: Average number of aberrations in various age groups. A graph of the mean numbers of aberrations in each group shows an increasing mean number of aberrations with age

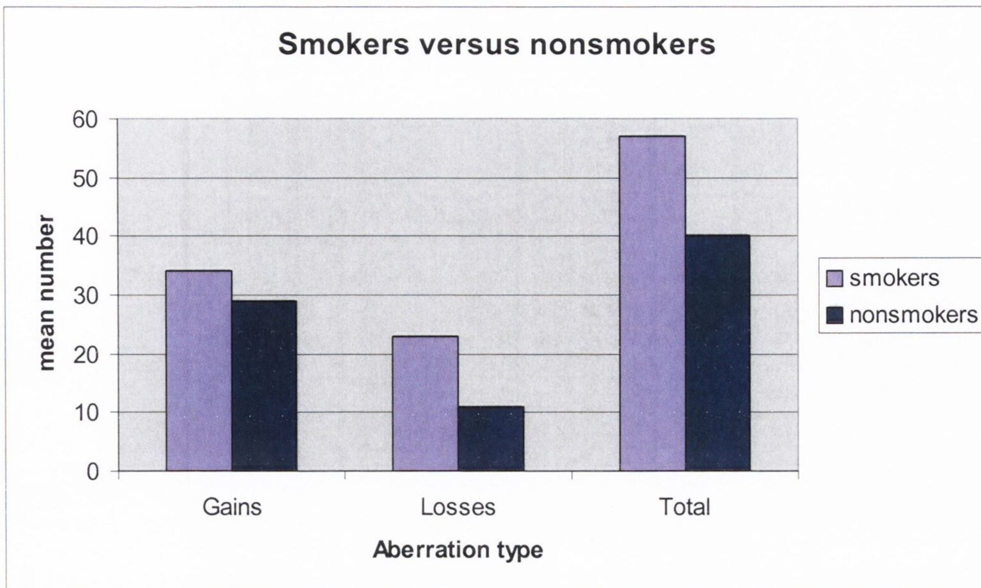


Figure4.5: Average number of aberrations in smokers compared to nonsmokers

Table 4.3 Analysis of numbers of aberrations by clinical details

	Mean number of amplifications	Mean number of deletions	Total mean number of aberrations	p value[a]
Age				
<40	70	48	118	p=0.02
>40	104	90	194	
Smoking status				
Non-smoker	68	55	123	P=0.2
Smoker	106	83	189	
Age & smoking				
<40 non-smoker	25	20	45	p=0.04
<40 smoker	45	28	73	
>40 non-smoker	43	35	78	
>40 smoker	61	55	116	
Stage				
T1	37	22	59	p=0.1
T2	43	30	73	
T3	49	44	93	

[a]P for Mann Whitney U test

4.5.2 Specific aberrations

Gains most commonly found in the older cohort were in clones mapping to 1p, 5p, 7p, 11p (subtelomeric regions) and to 5q33, 11q13 (*FGF4*), and 22q13 (*PDGFB*), while clones mapping to 3p (subtelomeric region), 3p24 (*THRB*) and 9p21 ($p16^{INK4}$) were among the most frequently lost (Table 4.4). This profile of frequent gain and loss in the older group mirrors that found in the literature on oral and head and neck SCC. Deletions of $p16^{INK4}$ (50%), 9p11.2 (70%), and 5q11.2 (*MSH3*) (60%) were found exclusively in the old group. Overall, there was a significant difference found in the occurrence of deletions at 9p21, 9p11.2, 5q11.2 and 5q21 between the two groups (Fishers exact $p<0.01$)

The profile of the young cohort was markedly different from the older cohort in terms of copy gain and loss. The most frequent copy number changes were found in clones mapping to 1p and 3p (subtelomeric), and to 3q (*MFI2),* 11p (*INS*) and 11q13 (FGF4), 11q13-q14 (*PAK1*), 14q32.1 (*TCL1A*) and 22q11.2 (*TBX1*). See Table 4.5

Chromosome	Gain/Loss	Clone name (NCBI cytogenetic locus)	>40 yrs	<40 yrs
Chr 1	Gain	1ptel (subtelomeric)	60%	50%
Chr 2	Gain	2qtel (subtelomeric)	70%	20%
Chr 3	Loss	3p (subtelomeric)	70%	50%
	Loss	3p (subtelomeric)	50%	40%
	Loss	*THRB* (3p24.3)	60%	20%
	Loss	*FHIT* (3p14.2	50%	30%
	Loss	P44s10 (marker)	60%	20%
	Loss	*ROBO1* (3p12-3p13)	70%	20%
	Gain	*TERC* (3q26)	50%	40%
	Gain	*MFI2* (3q29)	60%	50%
Chr 5	Loss	*MSH3* (5q11.2-5q13.2)	50%	0%
	Loss	*APC* (5q21-5q22)	67%	0%
	Gain	CSF1R (marker)	70%	20%
Chr 7	Gain	*EGFR* (7p12.3-p12.1)	50%	50%
Chr 8	Gain	*FGFR1* (8p11.2-p11.1)	60%	20%
	Gain	*MYC* (8q24.11-q24.13)	80%	30%
Chr 9	Loss	*P16* (9p21)	50%	0%
	Loss	AFM137XA11 (9p11.2)	70%	0%
Chr 11	Gain	*INS* (11p subtelomeric)	80%	60%
	Gain	*P57* (KIP2)(11p15.5)	50%	30%
	Gain	*CCND1* (11q13)	67%	30%
	Gain	*FGF4* (11q13)	70%	50%
Chr 14	Gain	*TCL1* (14q32.1)	60%	50%
	Gain	*AKT* (14q32.32)	80%	40%
Chr 22	Gain	*PDGFB* (22q13.1)	70%	40%

Table 4.4 The most frequent copy number imbalances in the old group and comparison with the young group

Chromosome	Gain/loss	Clone name (locus name NCBI)	<40 years	>40 years
Chr 1	Gain	1p (subtelomeric)	50%	60%
Chr 3	Loss	3p (subtelomeric)	50%	70%
	Gain	*MFI2* (3q29)	50%	60%
Chr 7	Gain	*EGFR* (7p12.3)	50%	50%
Chr 11	Gain	*INS* 11p (subtelomeric)	60%	80%
	Gain	*FGF4* (11q13)	50%	70%
	Gain	*PAK1* (11q13)	50%	40%
Chr 13	Gain	*TCL1A* (14q32)	50%	60%
Chr 22	Gain	*TBX1* (22q11.2)	50%	40%

Table 4.5 The most frequent copy number imbalances occurring in the young group and comparison with the old group

4.6 Discussion

In 1992, Kallioniemi developed a molecular cytogenetic method, capable of detecting and mapping relative DNA sequence copy number variation across the whole genome in one single experiment (Kallioniemi, 1992). CGH uses a relatively small amount of DNA, and allows analysis of the whole genome detecting losses and gains in all chromosomes at once, without the need for cell culture (Hermsen, 1996). Both conventional and array CGH has been previously applied in studies of oral cancer, and has demonstrated a non-random pattern of genomic aberrations including deletions of material from 3p, 5q, 9p and gains involving 3q, 5p, 7p, 8q, 11q, 17q, and 20q (Gebhart, 2003; Garnis, 2003; Garnis, 2004). This is the first CGH study in the literature comparing the profile of young patients with older patients with oral cancer. Jin et al (1999) used examined oral SCC samples from 36 young adults, looking at specific LOH found in older adults with oral SCC. This paper noted a similar incidence of LOH at chromosome 3, 9 and 17p to that found in older patients. However, this paper did not address the presence or absence of risk factors. This may be a critical factor in determining the profile of those young patients developing oral SCC.

The aims of this present study were two-fold. The first objective was to map gains and losses in DNA from cases of oral cancer, and compare our findings with those in the published literature. Secondly, to compare the profile of young patients with oral cancer with the profile found in older patients with oral cancer, paying special attention to the use of tobacco. The present study shows for the first time by CGH

analysis that the genomic profile of oral cancer in young patients especially the non-smokers, differs markedly from the profile of the more typical oral cancer patient.

4.6.1 Frequency of aberrations

Chromosomal imbalance is thought to be the major type of genetic instability in solid tumours (Kinzler, 1997). In this study, the old group, irrespective of smoking status, had high numbers of aberrations, occurring at loci well described in the literature.

This suggests three possibilities:

1. Although the old nonsmokers are not directly exposed to the carcinogens, prolonged exposure in a secondary manner to similar carcinogens may be a factor contributing to the development of oral cancer in this group. However due to lack of clinical details regarding lifestyle of the patients in this study, it is not possible to draw firm conclusions on the effects of environmental carcinogens.
2. The possibility that some of these aberrations are simply a result of aging cannot be out ruled. However, the aberrations found in the old nonsmokers closely resemble the chromosomal changes seen in all the smokers and those changes quoted in the literature.
3. There may be another stimulus, environmental or otherwise, that may give rise to a common group of alterations in both old smokers and old nonsmokers.

All the smokers, irrespective of age, had a high frequency of copy gain and loss. There were common molecular changes found in both young and old smokers in this study suggesting that molecular events due to the carcinogens in tobacco are similar in both young and old smokers with cancer. It must be noted that the smokers in the under 40 year old group all had a long history of tobacco use, with some disclosing a smoking history of 20 years. Therefore it is not surprising that the molecular profile in all smokers, irrespective of age, were similar. While the overall number of aberrations were high in all groups using tobacco, there were some differences in site copy gain and loss, such as CDKN2A deletion, discussed below.

The most remarkable finding was that young nonsmokers contrasted dramatically with the other groups, by having far fewer aberrations (Figure 4.3). The young nonsmokers had considerably fewer aberrations than the other three groups suggesting that chromosomal imbalance is not the main mode of genetic instability in this group.

4.6.2 Copy gain/loss profile differences

Over the years, classical cytogenetic analysis of head and neck tumors has been very productive, resulting in the identification of key genomic regions involved in the development and progression of HNSCC. Studies on oral cancer and HNSCC have shown that a number of chromosomal regions are consistently altered, such as gains on 3q, 5p, 8q, 11q13, and losses on 3p, 9p among others (Hermsen, 1997; Wolff, 1998; Gebhart, 2003; Gollin, 2001).

Array CGH analysis of oral cancer cases in this study found that the most frequent alterations were mapped to chromosome 3 with gain of the q arm often occurring simultaneously with loss of the p arm. The older group in this study showed a profile similar to that of the 'classic' oral cancer patient in the literature, including deletions on 3p, 9p and gains of 3q, 5q, 7p, and 11q13. In concordance with the literature on oral cancer, gain of the q arm was a less frequent event than loss of the p-arm (Squire, 2002).

Having established that the profile of the old group resembled closely that found in the literature, the second aim was to compare the genomic profile of the young group with the old group. The changes occurring most frequently in the young group are tabulated in Table 4.5. Taking the young group as a whole, irrespective of smoking status, the most striking finding was the absence of two aberrations most commonly found in the literature, namely loss of *p16* and loss of *APC* (Table 4.4).

Deletion of *CDKN2A* ($p16^{INK4}$) has been associated in the literature with smoking. Sanchez-Cespedes in a study of carcinomas of the lung, found that deletion of the *p16* gene locus was only observed in tumours from smokers (Sanchez-Cespedes, 2001). Considering the prevalence of non-smokers in the young cohort in the present oral cancer study, the absence of deletions of *p16* in this group is not that surprising. In light of absence of *p16* deletions in young patients in this study, it is possible that *p16* deletions do not play a comparable role in oral SCC occurring in young adults as in classic oral SCC seen in older adults. Of course this finding does not discount the role of p16 in oral and pharyngeal cancer. *p16* may be inactivated by other methods not explored in this paper such as methylation. Of course methylation is an

alternative mechanism for many genes and so this theory can be applied to absence of loss of other genes also, such as APC.

Notwithstanding this striking absence of p16 deletion in young tumour samples, it must be highlighted that overall the young smokers did have a similar mean number of aberrations to all the other smokers, irrespective of age.

Loss of APC (5q21) is completely absent in the young nonsmokers. This is in contrast to the frequent occurrence (67%) of loss of APC in the old smokers.

The contribution of down-regulated APC expression to the development of human OSCC has been reported to be about 30%, and hypermethylation of the gene promoter CpG island seems to be a significant mechanism of inactivation of the APC gene in oral carcinogenesis (Uesugi et al, 2005)

The group that differed from the other cohorts, in all aspects, was the young non-smoker cohort. Having established that there was less genomic instability in the young nonsmokers, it was noted that the genes most commonly altered in this group were also seen to be altered in the old group; therefore, there was no genomic aberration that was exclusively gained or lost in the young group.

In this study, far fewer alterations were found on chromosome 3 in the young non-smokers in comparison to the other cohorts. The diminished role for chromosome 3 aberrations found in this group of oral cancers contradicts Partridge *et al* (1994) who found that loss of heterozygosity (LOH) on Chr 3p was not confined to patients exposed to heavy smoking or alcohol consumption and was present at an early stage in the disease suggesting that this genetic alteration may be a fundamental event for all oral cancer.

The data generated in this present study indicate that the aberration profile in the young non-smokers is remarkably different from the profile found in the more typical oral cancer patient.

The young non-smokers appear to have very few gains and losses at the sites established in the literature as commonly altered in oral cancer. This suggests that deletions of tumour suppressor genes such as *p16*, and deletions of short arm of chromosome 3, which are alterations commonly found in oral cancer, may have a diminished role to play in the development of oral cancer in young people.

With a notable absence of traditional risk factors in a significant portion of young people developing oral cancer (Mackenzie, 2000; Llewellyn, 2004), other possible risk factors have been explored, for example occupation (Moulin, 1986), diet (Franceschi, 1999), genetic and familial factors (Friedlander, 2001) and viruses (Gillison, 2000), but these areas require further exploration in order to establish the causative factors underlying oral cancer in the young.

4.7 Conclusion

This work represents the first study applying age discriminated oral cancers to a CGH array. A marked difference in the number of aberrations is evident, with the young non-smokers showing considerably lower numbers of aberrations than the other groups. The aberration profiles of the young group, in particular the young non-smokers, differs distinctly from the old group.

Such significant difference in genomic profiles, with absence of deletions of tumour suppressor gene *CDKN2A (p16)* suggests that a different molecular mechanism for oral carcinogenesis in these young non-smoking patients may be in place, however, this alternative mechanism responsible for the development of oral cancer in young people remains to be elucidated. The pattern of oral cancer is changing, and it can no longer be assumed that only old male smokers are likely to develop this disease.

4.8 References

Albrecht B, Hausmann M, Zitzelsberger H, Stein H, Siewert JR, Hopt U, Langer R, Hofler H, Werner M, Walch A. Array-based comparative genomic hybridisation for the detection of DNA sequence copy number changes in Barrett's adenocarcinomas. J of Pathol 2004; 203(3): 780-8

Bockmuhl U, Wolf G, Schmidt S, Schwendel A, Jahnke V, Dietel M, Petersen I. Genomic alterations associated with malignancy in head and neck cancer. Head Neck. 1998; 20(2): 145-51

Brzoska PM, Levin NA, Fu KK et al. Frequent novel DNA copy number increase in squamous cell head and neck tumors. Cancer Res. 1995; 55(14): 3055-9

Califano J, van der Riet P, Westra W, Nawroz H, Clayman G, Piantadosi S, Corio R, Lee D, Greenberg B, Koch W, Sidransky D. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res. 1996; 56(11): 2488-92

Franceschi S, Favero A, Conti E, Talamini R, Volpe R, Negri E, Barzan L, La Vecchia C. Food groups, oils and butter, and cancer of the oral cavity and pharynx. BJC 1999; 80: 614-620

Friedlander PL. Genomic instability in head and neck cancer patients. Head Neck. 2001; 23(8): 683-91

Garnis C, Baldwin C, Zhang L, Rosin MP, Lam WL. Use of complete coverage array comparative genomic hybridisation to define copy number alterations on chromosome 3p in oral squamous cell carcinomas. Cancer Res. 2003; 63(24): 8582-5

Garnis C, Campbell J, Zhang L, Rosin MP, Lam WL OCGR array: an oral cancer genomic regional array for comparative genomic hybridisation analysis. Oral Oncol. 2004; 40(5): 511-9

Gebhart E, Liehr T, Wolff E, Wiltfang J, Koscielny S, J. Ries. Loss of 9p21 is embedded in a complex but consistent pattern of genomic imbalances in oral squamous cell carcinomas. Cytogenetic and Genome Research 2003; 101:106-112

Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Viglione M, Symer DE, Shah KV, Sidransky D. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst. 2000; 92(9): 709-20

Gollin SM. Chromosomal alterations in squamous cell carcinomas of the head and neck: Window to the biology of disease. Head & Neck. 2001; 23(3): 238-253

Hermsen M, Guervos MA, Meijer G et al. New Chromosomal regions with high-level amplifications in SCCs of the larynx and pharynx, identified by CGH. Journal of Pathology 2001; 194:177-182

Hermsen MA, Meijer GA, Baak JP, Joenje H, Walboomers JJ. CGH- A new tool in cancer pathology. Human Pathology 1996; 27(4): 342-922.

Hermsen MAJA, Joenje H, Arwert F. Assessment of chromosomal gains and losses in oral squamous cell carcinoma by CGH. Oral Oncology 1997; 33:414-18

Jin YT, Myers J, Tsai ST, Goepfert H, Batsakis JG, el-Naggar AK. Genetic alterations in oral squamous cell carcinoma of young adults. Oral Oncol. 1999; 35(3): 251-6

Joos S, Scherthan H, Speicher MR, Schlegel J, Cremer T, Lichter P. Detection of amplified DNA sequences by reverse chromosome painting using genomic tumor DNA as probe. Hum Genet. 1993; 90(6): 584-9

Kallioniemi A, Kallioniemi OP, Sudar D et al. Comparative genomic hybridisation for molecular cytogenetic analysis of solid tumours. Science. 1992; 258(5083): 818-21

Kinzler KW, Vogelstein B. Cancer susceptibility genes. Gatekeepers and caretakers. Nature 1997; 24:761-763

Llewellyn CD, Johnson NW, Warnakulasuriya KA. Risk factors for oral cancer in newly diagnosed patients aged 45 years and younger: a case-control study in Southern England. J Oral Pathol Med. 2004; 33(9): 525-32

Mackenzie J, Ah-see K, Thakker N, Sloan P, Maran A.G. Birch J, Macfarlane G.J. Increasing incidence of oral cancer amongst young persons what is the aetiology? Oral Oncology 2000; **36:** 387-389

Mork J, Lie AK, Glattre E, Hallmans G, Jellum E, Koskela P, Moller B, Pukkala E, Schiller JT, Youngman L, Lehtinen M, Dillner J. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. N Engl J Med. 2001; 344(15): 1125-31

Moulin JJ, Mur JM, Wild P, Perreaux JP, Pham QT. Oral cavity and laryngeal cancers among man-made mineral fiber production workers. Scand J Work Environ Health. 1986; 12(1): 27-31

O'Regan E, Smyth P, Finn S, Martin C, Sheils O, O'Leary J. Comparative Genomic Hybridisation. The Science of Laboratory Diagnosis. Crocker and Burnett; Wiley, 2005.

Partridge M, Kiguwa S, Langdon JD. Frequent deletion of chromosome 3p in oral squamous cell carcinoma. Eur J Cancer B Oral Oncol. 1994; 30B(4): 248-51

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High-resolution analysis of DNA copy number variation using comparative genomic hybridisation to microarrays. Nat Genet. 1998; 20(2): 207-11

Sanchez-Cespedes M, Decker PA, Doffek KM, Esteller M, Westra WH, Alawi EA, Herman JG, Demeure MJ, Sidransky D, Ahrendt SA. Increased loss of chromosome 9p21 but not p16 inactivation in primary non-small cell lung cancer from smokers. Cancer Res. 2001; 61(5): 2092-6

Schantz SP, Yu GP. Head and neck cancer incidence trends in young Americans, 1973-1997, with a special analysis for tongue cancer. Arch Otolaryngol Head Neck Surg 2002; 128(3): 268-274

Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray Science. 1995; 270(5235): 467-70

Schraml P, Schwerdtfeger G, G, Burkhalter F, Raggi A, Schmidt D, Ruffalo T, King W, Wilber K, Mihatsch MJ, Moch H. Combined array comparative genomic hybridisation and tissue microarray analysis suggest PAK1 at 11q13.5-q14 as a critical oncogene target in ovarian carcinoma. Am J Pathol. 2003 163(3): 985-92

Squire JA, Bayani J, Luk C, Unwin L, Tokunaga J, Macmillan C, Irish J, Brown D, Gullane P, Kamel-Reid S. Molecular cytogenetic analysis of head and neck squamous cell carcinoma: By comparative genomic hybridisation, spectral karyotyping, and expression array analysis. Head Neck. 2002; 24(9): 874-87

Struski S, Doco-Fenzy Martine, Cornillet-Lefebvre P. Compilation of published comparative genomic hybridisation studies. Cancer Genetics Cytogenetics 2002; 135: 63-90.

Uesugi H, Usawa K, Kawasaki K, Shimada K, Moriya T, Tada A, Shiiba M, Tazada H. Status of reduced expression and hypermethylation of the APC tumor suppressor gene in human oral squamous cell carcinoma. Int J Mol Med. 2005 15(4):597-602

Wolff E, Girod S, Liehr T, Vorderwulbecke U, Ries J, Steininger H, Gebhart E. Oral squamous cell carcinomas are characterized by a rather uniform pattern of genomic imbalances detected by comparative genomic hybridisation. Oral Oncol. 1998; 34(3): 186-90.

Chapter 5

Human papillomavirus detection and p16^{INK4} expression in SCC of the oral cavity and oropharynx in relation to age and smoking

5.1 Summary

Alterations of p16, a tumour suppressor gene, have been consistently reported in head and neck cancer. These alterations occur in the form of deletions, mutations and methylation. Human papillomavirus (HPV) is thought to inactivate retinoblastoma, thus resulting in an increase in p16 expression.

After finding a lack of p16 deletion in young head and neck cancer patients in Chapter 4, it remains to be seen whether this absence of deletion manifests as low level p16 expression in the young patients with cancer. HPV has been implicated in the past in the development of head and neck cancer, and so this chapter investigates not only p16 expression but also the presence of HPV DNA and integrated DNA in young patients with oral and oropharyngeal cancer compared with older patients.

By using immunohistochemistry and PCR, this study concludes that mRNA expression of high-risk HPV16 does occur in a number of oral and oropharyngeal SCC, and may be a contributing factor in a subset of these malignancies but it is not a necessity, as it is in cervical cancer. Expression of p16 is significantly associated with presence of HPV, which is strongly suggestive of viral integration and subsequent malfunction of p16. This study also notes that HPV mRNA expression occurs more often in oropharyngeal tumours than cancers of the oral cavity.

Perhaps most interesting of all was that none of the young female nonsmokers in this study showed viral integration or positive p16 protein expression. This strongly suggests that HPV, previously proposed as a key risk factor in young nonsmokers, does not play an important role in the development of SCC in this cohort.

5.2 Introduction

5.2.1 p16^{INK4}

p16INK4a belongs to the family of cell cycle regulators called cyclin dependant kinase inhibitors, which bind to 'cyclin-CDK' complexes and cause cell cycle arrest in the G1 phase. p16 exerts its anti-proliferative effects by binding to and inhibiting the actions of CDK4 and 6. The INK4a gene is located at the chromosome locus 9p21, a region that undergoes frequent hemi- and homozygous deletion in human cancers. Several lines of evidence suggest that INK4A is a bona fide tumour suppressor gene at 9p21. Somatic mutations or allelic deletions of this INK4a have been described in a variety of cancers including head and neck SCCs.
A prevalence of p16^{INK4} alterations ranging from 25% to 83% has been reported in head and neck cancer (Reed, 1996; Matsuda, 1996; El-Naggar, 1997).
The p16^{INK4} (CDKN2A) gene can be inactivated by a number of mechanisms, but in head and neck cancer, the predominant methods are methylation and deletion of the gene rather than gene mutations, which are less common. Immunohistochemistry methods have shown low p16^{INK4} expression in 55% to 89% of oral and oropharyngeal SCC, similar to that found in normal non-malignant mucosa (El Naggar, 1999; Pandey, 1998), however overexpression of p16^{INK4} has also been reported in head and neck cancers (Andl, 1998; Gillison, 2000; Lang, 2002). It is thought that HPV infection, via inactivation of retinoblastoma gene, accounts for the observed high levels of p16^{INK4} expression through disruption of a negative feedback mechanism.

Differences between smoking and non-smoking patients with regard to the expression and genetic status of p16^{INK4} in head and neck SCCs have been previously reported, with some evidence of a link between p16 promoter methylation and smoking (Lea, 2004; Wong, 2000). This evidence concurs with reports of an association between p16^{INK4} overexpression in tumours of nonsmokers (Haas; 2002). Higher levels of p16^{INK4} expression in the oropharynx (especially tonsil) compared to the oral cavity have also been reported, but this not surprising as the incidence of HPV is also higher in an oropharyngeal tumour site especially tonsil (Haas, 2002; Fouret, 1997; Andl, 1998; Koch, 1999).

These studies, however, did not take into account the age and gender of the patients, and because nonsmokers with oral and oropharyngeal SCC are very often young females with tongue cancer, age and gender are important factors to be included in the analysis of p16^{INK4} status. Evidence of an increase in p16^{INK4} protein expression in oral and oropharyngeal SCCs of young patients (Li, 2004) has been published, but no details regarding risk factors and gender were recorded in this study. To our knowledge, there are no published reports analysing p16^{INK4} expression in relation to all 4 factors: age, gender, risk factors and site.

5.2.2 HPV

Marked variations in reports of HPV prevalence in oral and pharyngeal SCCs have made it difficult to assess the causal association in HPV in these malignancies (Ha, 2004; Miller, 1996). The significance of the role of HPV in the progression to malignancy in oral SCCs still remains unclear. HPV may be a contributing factor in

a subset of oral malignancies, however it is clear that its role in oral carcinoma differs from cervical cancer, where HPVs have been established as a primary cause (Cuzick, 2000).

While tobacco smoking is the primary cause of head and neck cancer, it is also a well established co-factor in the development of cervical cancer. In fact, a number of reports state that tobacco smoking increases the risk of cervical cancer among HPV positive women (Plummer, 2003).

The fact that nicotine and tobacco specific carcinogens have been detected in the cervical mucus of smokers further strengthens the hypothesis of a synergistic action between cigarette smoking and HPV for the development of cervical cancer (Prokopczyk, 1997). Chemical tobacco-related carcinogens may exert a direct mitogenic effect causing DNA damage. Some authors hypothesize that exposure to tobacco may affect the ability of the host to mount an effective local immune response against viral infections, since it has been shown that smoking may reduce the number of Langerhans' cells and other markers of immune function (Poppe, 1995).

As already discussed in Chapter 1, there is a possibility that the HPV detected in the oropharyngeal and oral SCCs may be sexually transferred, and that the HPV infection may then interact with alcohol and tobacco exposure in tumor promotion (Gillison, 2003). Schwartz (1998) found that in males, oral SCC risk increased with self-reported decreasing age at first intercourse, increasing number of sex partners, and a history of genital warts. Despite this finding, the detection level of oncogenic HPV in exfoliated oral tissue was the same in both cases and controls.

Most interesting was its strong joint association of cigarette smoking and HPV 16 with oral cancer. In contrast to Schwartz' report, another USA study found that HPV-related genital lesions, oro-genital sexual behaviour, and number of sexual partners, did not differ between patients with OSCC and controls (Smith, 1998).

Oral and oropharyngeal SCC of the head and neck occurring in the younger age groups has long been thought to constitute an etiologically distinct group, and HPV has come under scrutiny as a possible risk factor. Nonetheless, the relationship between HPV and age is controversial. There are reports of a link between HPV positivity and older age at diagnosis (Cruz, 1996), while others have found a strong association with HPV detection and younger age (Zhang, 2004; Li, 2004). It has yet to be determined whether the links between HPV positivity and younger age reflect sexual practices (i.e orogenital sexual contact) or genetic predisposition.

With the growing evidence of the contribution of HPV to oropharyngeal carcinogenesis and with evidence of an increase in young female nonsmokers developing oral and oropharyngeal cancer (Llewellyn, 2003; Schantz, 2002) it may be helpful to investigate the prevalence of integrated HPV in these tumours.

5.3 Aims

Having found an absence of p16 deletion in young nonsmokers by using microarray CGH (Chapter 4), further investigations of this tumour suppressor gene were warranted, as p16 deletion or methylation are thought to be common events in head and neck cancer overall.

This chapter has been divided into two sections. The first section describes the protein and mRNA expression pattern of $p16^{INK4A}$ in the same cohort of young and

old patients with oral and pharyngeal carcinomas that were tested in Chapter 4. In order to further investigate the link between $p16^{INK4}$ and HPV, the second section describes the prevalence of both high risk HPV DNA and HPV mRNA in the samples and their association with $p16^{INK4A}$ expression. An overall discussion of both sections is included at the end of the chapter.

While this chapter focuses on expression of p16 in relation to HPV, methylation status of p16 was also of interest. However, due to restrictions, it was decided that this avenue would not be explored for this thesis.

5 .4 Section 1 Analysis of $p16^{INK4}$ protein and mRNA expression levels

This first section of this chapter describes experiments designed to detect the protein and mRNA expression patterns of $p16^{INK4A}$ in oral and pharyngeal SCCs, and to correlate this with the p16 results from the CGH study (Chapter 4). The results from all the $p16^{INK4}$ experiments are correlated with the clinical data from each sample e.g. site, age, gender, risk factors and TNM stage.

5.4.1 Materials and methods

5.4.1.1 Immunohistochemistry

In this study $p16^{INK4A}$ protein expression was examined by immunohistochemical analysis of 25 oral and pharyngeal squamous cell carcinomas. Cases (all from the pathology archives of St. James Hospital, Dublin) were formalin fixed and paraffin embedded. The immunohistochemical distribution of $p16^{INK4A}$ in these samples was evaluated using the DAKO $p16^{INK4A}$ Research Kit (DakoCytomation, Denmark

A/S). This kit employs a monoclonal anti-p16^{INK4A} antibody Clone E6H4, (MTM laboratories AG, Germany). Immunohistochemical analysis was carried out as outlined in Chapter 2. The clinical details (Age, gender, and site etc) of each of the 25 samples are tabulated below (Table 5.1).

Controls were included in each assay, comprising positive controls in the form of sections of cervical mucosa with evidence of CIN3, negative controls in which the primary antibody was replaced with PBS, and a control of normal oral tissue. The immunoreaction cut-off was established by quantifying nuclear and cytoplasmic positivity in normal tissue, which ranged from zero to 5%.

Sample	Age	Gender	Smoking	Site	Stage	Grade
1	Under 40	M	No	A	T2N0	Well
2	Under 40	M	No	T	T2	Mod
3	Under 40	F	No	T	T2N2	Mod
4	Under 40	F	No	T	T1N1	Mod
5	Under 40	M	No	FOM	T1N0	Mod
6	Under 40	F	No	T	T2N2	Mod
7	Under 40	F	No	T	T2N2	Mod
8	Under 40	M	Yes	T	T3N0	Mod
9	Under 40	M	Yes	FOM	T2N1	Mod
10	Under 40	M	Yes	T	T3N0	Mod
11	Under 40	M	Yes	T	T3N2	Mod
12	Under 40	M	Yes	T	T2N2	Poor
13	Over 40	F	No	T	T3M3	Mod
14	Over 40	F	No	T	T1Nx	Poor
15	Over 40	M	No	OP	T2N2	Mod
16	Over 40	M	No	T	T3N2	Mod
17	Over 40	M	Yes	TB	T4N2	Mod
18	Over 40	M	Yes	TB	T3N1	Mod
19	Over 40	M	Yes	TB	T2N2	Mod
20	Over 40	M	Yes	TB	T3N0	Poor
21	Over 40	M	Yes	T	T2N0	Mod
22	Over 40	F	Yes	A	T2N2	Poor
23	Over 40	M	Yes	OP	T3N2	Poor
24	Over 40	M	Yes	OP	T4	Poor
25	Over 40	M	Yes	OP	T3N0	Mod

Table 5.1 Clinicopathological details of all 25 samples. A: Alveolar Ridge, T: Tongue, FOM: Floor of Mouth, OP: Oropharynx, TB: Tongue Base

Following staining, tissue sections were qualitatively graded by a certified pathologist (MT) according to the following arbitrary scale: 0 (no positive staining), 1 (<10% positive staining), 2 (>10% and <50% positive staining), and 3 (>50% positive staining). The positive samples were then assessed further for nuclear or cytoplasmic staining. Excel Analyse-it® was employed to carry out statistical analysis of results.

5.4.1.2 TaqMan PCR for $p16^{INK4}$

In this study $p16^{INK4A}$ mRNA expression was examined in samples from 20 of the 25 cases of oral and pharyngeal carcinomas and 2 normal oral mucosa samples. Fresh tissue was unavailable in 5 of the 25 cases used in the IHC tests. All material was collected fresh, snap frozen and stored at minus 80 degrees (as described in **Chapter 2**). Diagnosis was confirmed on H&E stained sections in all cases. After a frozen section was performed to verify presence of tumour, RNA extraction from freshly frozen tissue was carried out using the Qiagen Systems RNA Isolation Kit as outlined in Chapter 2. RNA from all samples including UHR (Universal Human Reference RNA, Stratagene) was converted to cDNA using the Applied Biosystems cDNA Archiving Kit (Applied Biosystems, Foster City, CA), which includes reverse transcriptase and random primers. The method is further described in Chapter 2. To check the linearity of the detection system, the UHR cDNA was serially diluted for standard curve creation in six serial 5-fold dilutions of the UHR sample and these dilutions were run with $p16^{INK4}$ primer and probes and GAPDH primers and probe.

A p16^{INK4A} PDAR (Pre-Developed TaqMan Assay Reagents) assay from Applied Biosystems was chosen for analysis of p16^{INK4A} mRNA expression. Each PDAR assay is composed of 20X forward and reverse primers and 6FAM dye-MGB labelled probe. The p16^{INK4A} PDAR assay was designed against Homo sapiens cyclin-dependent kinase inhibitor 2A (melanoma, p16^{INK4}, inhibits CDK4) (CDKN2A), transcript variant 1, mRNA (Accession: NM_000077). Start and stop sites of the p16^{INK4A} amplicon are 384 and 443 respectively. Primer and probe sequences were not available (copyright protected). The p16^{INK4A} 58 base pair amplicon traverses the exon-exon boundary of exon 2 and exon 3 ensuring amplification of mRNA only.

Each 25µl PCR reaction contained 1x PCR Mastermix, 1x p16^{INK4A} primer and probe mix and 100ng RNA. Cycling conditions were as follows: 2 minutes at 50°C, 10 minutes at 95°C and 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Amplification and detection were carried out in the ABI PRISM 7900 Sequence Detection System (PE Applied Biosystems). A total of six negatives (no template controls) were included in each TaqMan Run.

A relative quantitation assay was used to analyse changes in p16 gene expression in the oral and pharyngeal carcinomas relative to a reference sample. Relative quantitation of gene expression was performed using the standard curve method. A GAPDH standard curve was generated by performing a serial dilution on control human RNA (Stratagene UHR) and analysing as per the test samples. Likewise, a p16 standard curve was produced by using 5-fold dilutions of control human RNA. (Figure 5.5). The target (p16) quantity was then determined from the standard curve

and divided by the target quantity of the calibrator (normal oral mucosa). A total of six negatives (no template controls) were included in each TaqMan run.

The significance of different levels of $p16^{INK4A}$ mRNA expression in samples of oral and pharyngeal carcinomas was assessed using the Mann-Whitney U-Test, which assesses pair wise differences between two cohort groups.

5.4.2 Results

5.4.2.1 Site

The oral carcinomas in this cohort were most commonly located on the tongue (52%), floor of mouth (4%) and alveolus (4%), while the most common oropharyngeal sites were the base of tongue (16%) and soft palate and other sites in the pharynx (16%)

5.4.2.2 Immunohistochemistry: $p16^{INK4}$

Six of the twenty-five samples (24%) showed positive staining. The positive cases varied in strength of staining, with 2/25 cases showing exceptionally strong staining. Four out of six of the positive cases had both cytoplasmic and nuclear staining, and the remaining two lacked nuclear positivity. Overall the $p16^{INK4}$ staining was diffuse (i.e over 80% of the tumour cells stained with the p16 antibody). In one case, dysplastic areas adjacent to invasive carcinoma showed positivity (Figure 5.1, 5.2, 5.3)

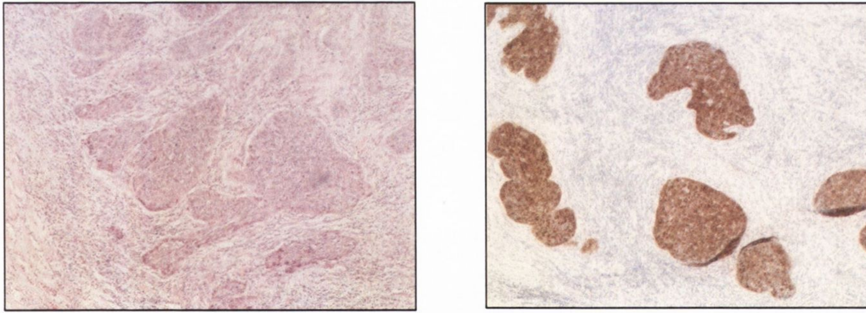


Fig 5.1 Score 3 Strong immunohistochemical staining for p16^{INK4A} in SCC.

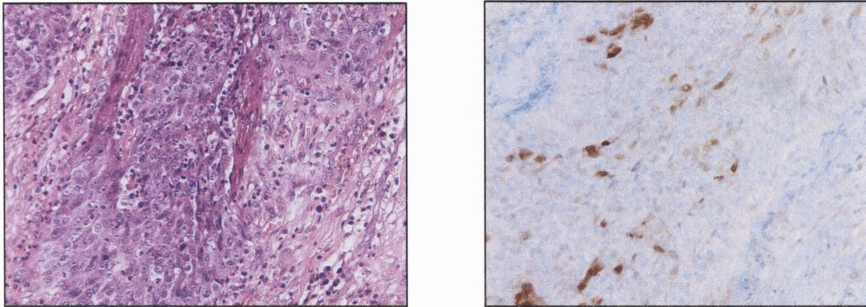


Figure 5.2 Score 2 >10% and <50% positive staining

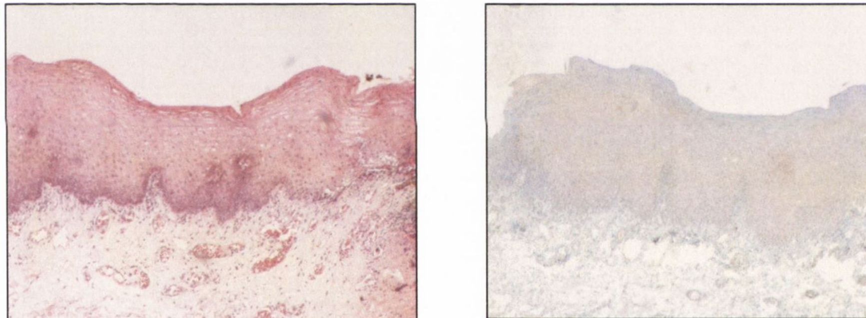


Figure 5.3 p16 immunohistochemical staining in normal oral mucosa

Table 2 shows the data of p16^{INK4} IHC expression in relation to site, gender, smoking history and tumour grade and stage. A statistically significant association was found between site of tumour and p16^{INK4} staining, with tongue base SCCs and floor of mouth SCCs demonstrating the highest level of p16^{INK4} staining. All the p16^{INK4} (100%) positive samples were from males with oral and oropharyngeal SCCs, and 66% of the positive samples were smokers. With regard to age, exactly

half of the p16^{INK4} positive cases were under 40 years old. Figure 5.1 illustrates the similar profiles of p16 protein expression across 4 cohorts, namely young nonsmokers, young smokers, old nonsmokers and old smokers. It is striking that while a significant absence of p16 gene deletion is evident in the young cohorts (CGH results), the p16 protein expression remains more or less the same in the young and the old cohorts, suggesting that mechanisms other than deletion may be influencing p16 protein expression in this group.

Clinicopathological feature	$p16^{INK4}$ positive	$p16^{INK4}$ negative	P value
Age			
Young	3	9	1
Old	3	10	
Sex			
Male	6	12	
Female	0	7	0.2
Tobacco smoking			
Yes	4	10	
No	2	9	0.9
Site			
Alveolar ridge	0	2	
Floor of mouth	2	0	
Tongue	1	12	0.01
Tongue base	2	2	
Soft palate	1	0	
Oropharynx NOS	0	3	
T stage			
T1	1	2	
T2	2	9	
T3	2	7	
T4	1	1	0.7
Nodal status			
Positive	3	12	0.5
Negative	3	4	
Differentiation			
Well	0	1	0.8
Moderately	5	14	
Poorly	1	4	

Table 5.2 Immunohistochemistry results

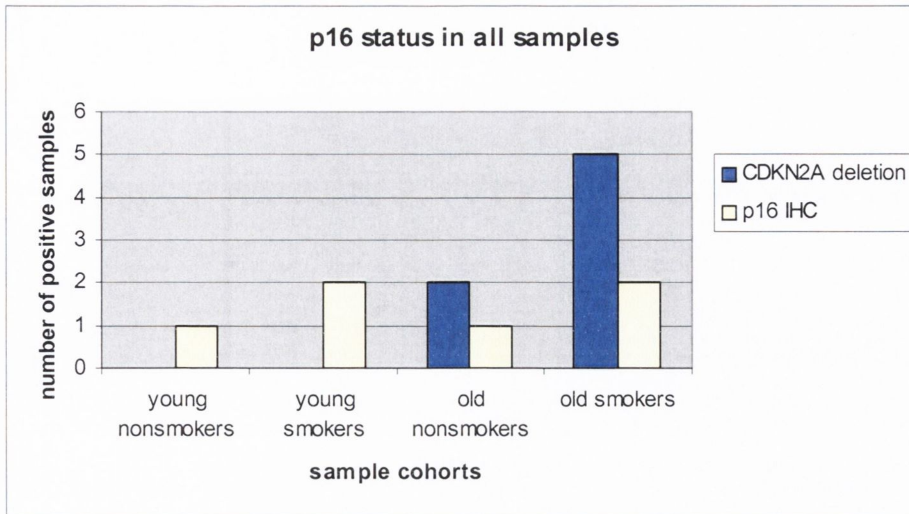


Figure 5.4 p16 detection in oral and pharyngeal SCC

5.4.2.3 Results of mRNA expression of p16INK4

All samples included in the study detected positive for GAPDH. TaqMan™ RT-PCR standard curves for both GAPDH and p16^{INK4} are represented in **Fig 5.5**. An example of the differential expression levels of p16 is illustrated in an amplification plot in **Fig 5.6.** P16^{INK4} expression was detected in all 20 cases of oral and pharyngeal carcinoma, and the p16 expression index varied from 0.019- 41.013 (p16 copies/GAPDH copies).

Mann Whitney U tests showed that a high p16 RNA expression index was highly associated with positive p16 protein expression (p=0.002) **Fig 5.7**.

Although age was not significantly associated with p16 overexpression, there was certainly a trend towards overexpression in under 40 year olds and under expression in over 40 year olds (p=0.08). When young nonsmokers and old smokers were

compared, no statistical significance was evident (p=0.5), with young nonsmokers showing higher levels of $p16^{INK4}$ expression than old smokers **Fig 5.8.**

Overall, there was no significant effect of site over $p16^{INK4}$ mRNA expression level (p=0.23), although there was a slight predominance of overexpression in the pharyngeal tumours when compared to the $p16^{INK4}$ expression index of the oral cavity tumours (60% and 50% respectively) **Fig 5.9.** Pairwise comparisons using Mann-Whitney U test are displayed in **Table 5.2**

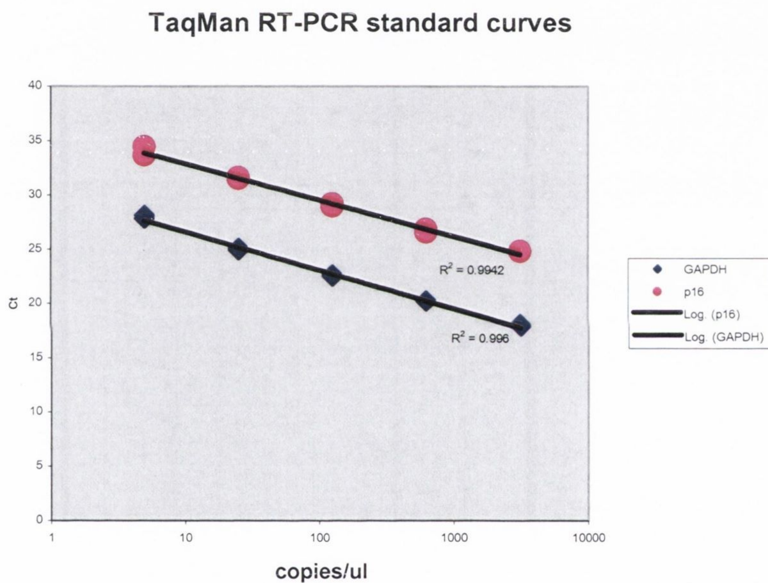


Figure 5.5 TaqMan RT-PCR standard curves

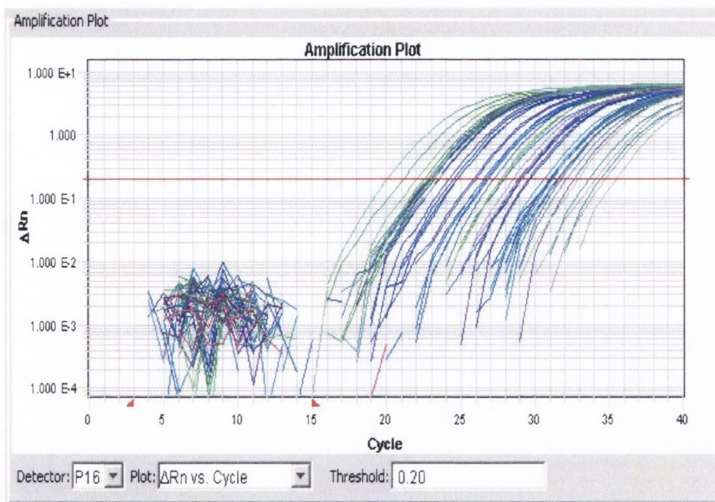


Figure 5.6 p16^{INK4} TaqMan amplification plot

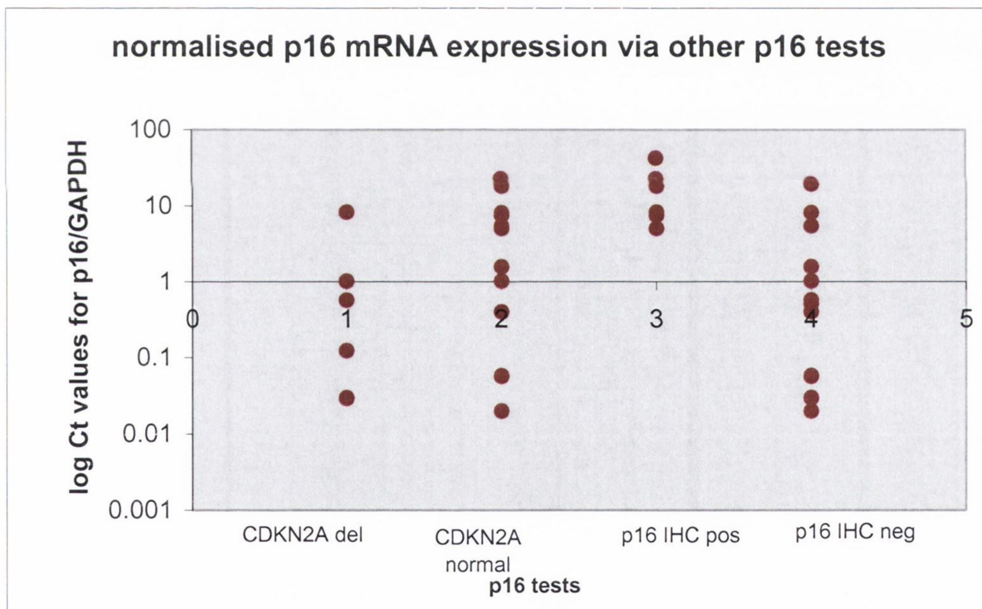


Fig 5.7 Normalised p16^{INK4} mRNA expression vs p16 protein expression and CDKN2A status

Test	p Value
Site	0.23
Gender	0.7910
Age	0.081
Smoking status	0.55
P16 protein IHC	0.002
CDKN2A gene copy number	0.17

Table 5.3 statistical analyses of p16^{INK4} mRNA results using MannWhitney U Test

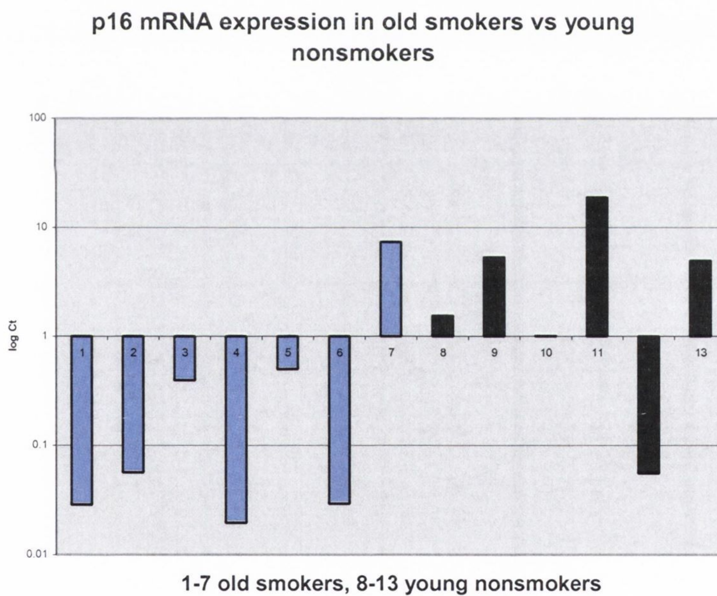


Figure 5.8 Comparison of p16^{INK4} expression between old smokers and young nonsmokers

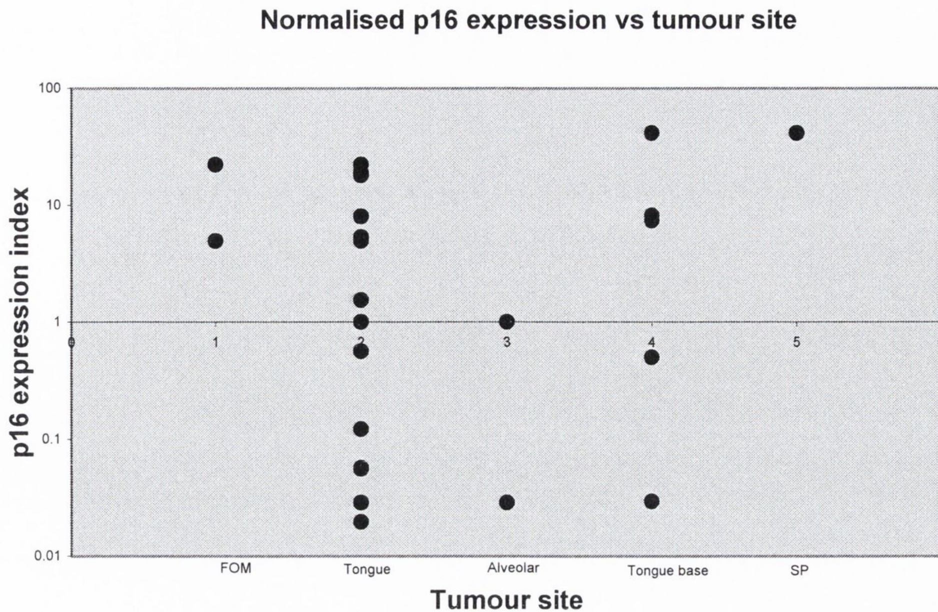


Figure 5.9 Normalised p16 expression vs tumour site

In summary,

- Tongue base SCCs and floor of mouth SCCs demonstrated a higher level of p16^{INK4} staining than any other site.
- All the p16^{INK4} (100%) positive samples were from males with oral and oropharyngeal SCCs.
- 66% of the positive samples were smokers.
- 50% of the p16^{INK4} positive cases were under 40 years old.

Section 2

5.5. HPV detection in oral and pharyngeal carcinoma

The second section of this chapter described HPV DNA 16 and 18 and HPV 16/18 mRNA detection in 20 oral and pharyngeal carcinomas. The results of these two tests were correlated with the $p16^{INK4}$ tests **(Section 1)** and clinical data on each sample.

The HPV genome typically consists of nine open-reading frame sequences, located on only one of the strands of DNA, and is divided into seven early-phase genes (E), and two late-phase genes (L). The early genes serve to regulate the transcription of DNA, while the late genes encode for proteins involved in viral spread, such as capsid proteins **(Figure 5.10)**

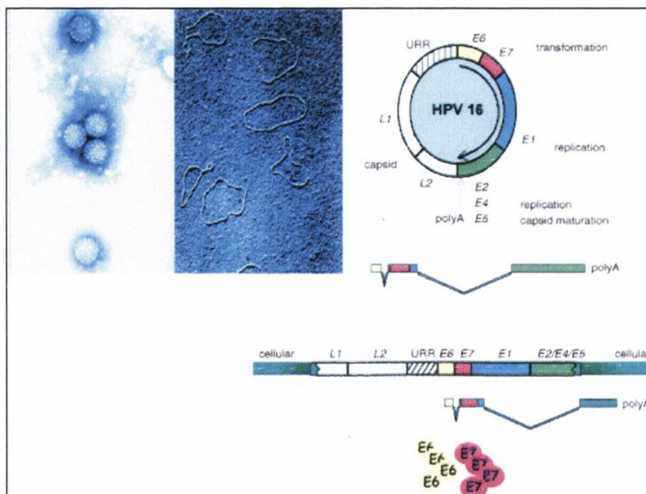


Figure 5.10 HPV 16 DNA Genomic Map

The E1 and E2 gene products are more specifically involved in regulating the transcription and replication of viral proteins. These different gene regions and gene products provide the basis on which molecular detection methods have been created. Upon integration the alteration of these genes can lead to unchecked transcription of E6 and E7 (Baker, 1987).

E6

The E6 gene is able to inactivate p53, which normally activates the transcription of p21, an inhibitor of kinases that promote mitosis. E6 can also activate telomerase independent of p53, leading to immortalisation of the infected cell.

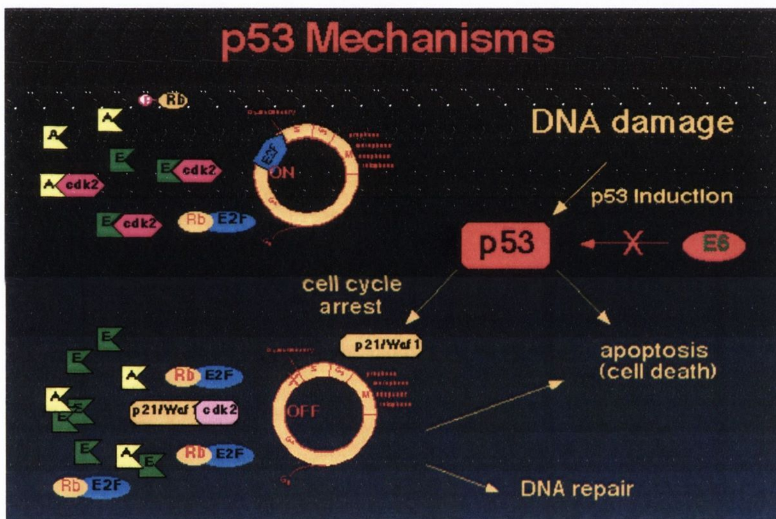


Figure 5.11 E6 forms a complex with the cellular ubiquitin protein ligase E6-AP, which is then able to bind and ubiquinate p53. **http://www.telemedicine.org/warts/cutmanhpv.htm**

E7

The HPV protein E7 can bind to retinoblastoma gene product (pRb), a cellular tumour suppressor, rendering it unable to bind and regulate E2F, a cell-cycle

activator, which then can freely activate the cell cycle to progress uncontrollably (Figure 5.12)

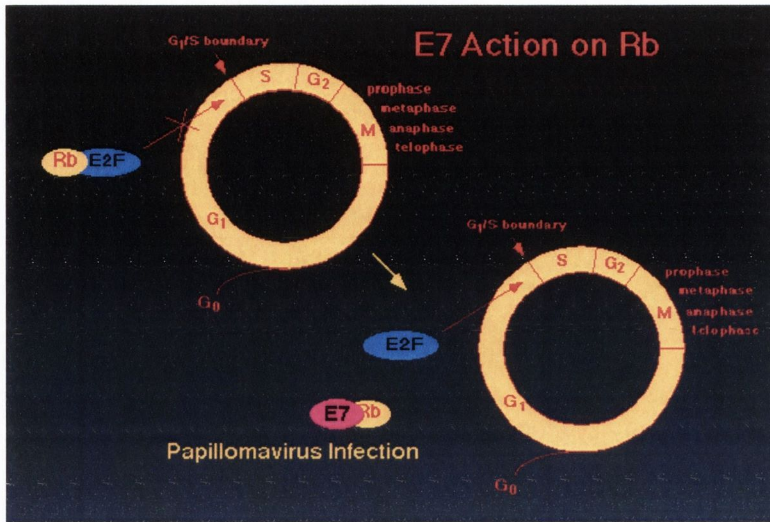


Figure 5.12 E2F acts as both an activator (G1 through S) and a repressor (M through G1) at different points in the cell cycle. When E7 binds pRB, pRB cannot bind to E2F and thus E2F freely activates the cell cycle.

http://www.telemedicine.org/warts/fig13e7.gif

Identification of E6 and E7, the two major viral oncogenes expressed in tumour tissue that are known to stimulate cell proliferation and interfere with tumour suppressor proteins, would further substantiate an oncogenic role for HPV in oral and oropharyngeal cancer. As predicted from the ability of HPV E6 to abrogate p53 function, HPV-positive oropharyngeal tumours are less frequently associated with p53 mutations than HPV-negative cases. They also appear to be less associated with alcohol and tobacco abuse (Gillison, 2000)

5.5.1 Materials and Methods

Fresh tissue from twenty one oral and pharyngeal carcinomas was acquired. These samples were from the same twenty tumours that were used in both the CGH study (Chapter 4) and the p16^{INK4} study **(See Table 5.1).** The HPV testing was not performed on 5/25 tumour samples in the p16 study due to limited availability of fresh tissue.

5.5.1.2 DNA & RNA extraction

DNA was extracted using the Gentra Puregene kit as outlined in Chapter 2. TaqMan Real-Time PCR for the endogenous reference gene β-actin was used as an extraction control to validate the presence and quality of extracted DNA (**as outlined in Chapter 2**). RNA extraction from the fresh tissue samples was carried out using the RNeasy Mini Kit (Qiagen). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an endogenous contol gene to validate the presence and quality of extracted RNA. TaqMan Real-Time RT-PCR was carried out as detailed in Chapter 2.

5.5.1.3 HPV Specific TaqMan PCR for detection of HPV16/18 DNA

There are many methods by which HPV can be detected, and every method has its strengths and weaknesses. In fact, the different prevalence rates of HPV may well reflect different technical approaches.

TaqMan PCR was employed for specific HPV DNA typing. Specific primer and probe sequences for HPV type 16 was synthesised following the design of Swan *et al.,* 1997. These primers and probe were designed within the L1 open reading frame of HPV16. PCR reaction conditions and cycling parameters were carried out as outlined by Swan *et al.,* 1997. HPV 18 primers and probe were designed by Murphy (2003) using Primer Express Software versions 1.5-1.7 (Applied Biosystems, Lincoln Centre drive, Foster City, CA 94404, USA) according to criteria described in the Primer Express Manual (Murphy, 2003). HPV18 TaqMan primer and probe sequences were designed against the L1 open reading frame of human papillomavirus type 18 complete genome sequence (Accession# AY262282). All primer and probe sequences were compared against the National Centre for Biotechnology Information's BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/). This confirmed the validity and appropriate design. Primer and probe sequences of both assays are included in **Table 5.4.** In brief, each 25μl TaqMan PCR amplification reaction contained 1x Universal Mastermix, 100nM probe, 200nM forward and reverse primers and 100ng DNA. After template denaturation for 10 minutes at 95°C, amplification conditions were as follows: for HPV 16: 40 cycles (each) of 30 seconds at 94°C, 10 seconds at 60°C and 2 minutes at 65°C (Swan, 1997), and for HPV18: 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Amplification and detection were carried out using an ABI PRISM 7000/7700 Sequence Detection System (Applied Biosystems). Interpretation of results was performed using Applied Biosystems Sequence Detector Software (Figure 5.13)

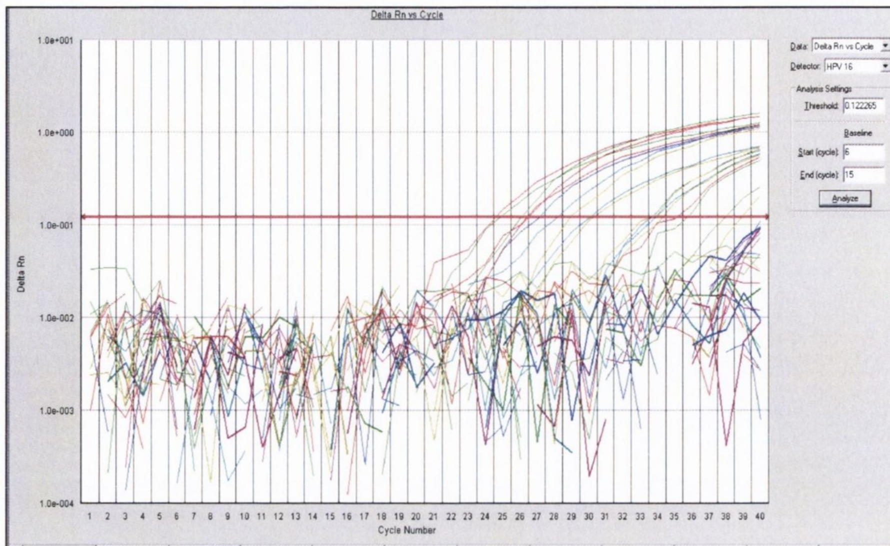


Fig 5.13 Amplification plot for p16

HPV Type		Primer & Probe Sequence
HPV 16	Forward primer	CCTTATTGGTTACAACGAGCAC
	Reverse primer	GCGTCCTAAAGGAAACTGATATA
	Probe	CCCCAGGAGGCACACTAGAAGAT
HPV 18	Forward primer	TCTGTTGCTATTACCTGTCAAAAGGA
	Reverse primer	TCCTTTAAATCCACATTCCAAAACT
	Probe	CTGCACCGGCTGAAAATAAGGATCCC

Table 5.4 Primers and probes used in HPV detection and typing.

5.5.1.4 Nucleic acid sequence based amplification (NASBA)

While RNA extraction was performed by the student (E O'Regan), the amplification process was kindly performed by Hanne Skomedal (NorChip AS, Klokkarstua, Norway). The data was then analysed by E O'Regan.

Reverse transcription followed by polymerase chain reaction (RT-PCR) is the most common method employed for the detection of mRNA in clinical samples. However

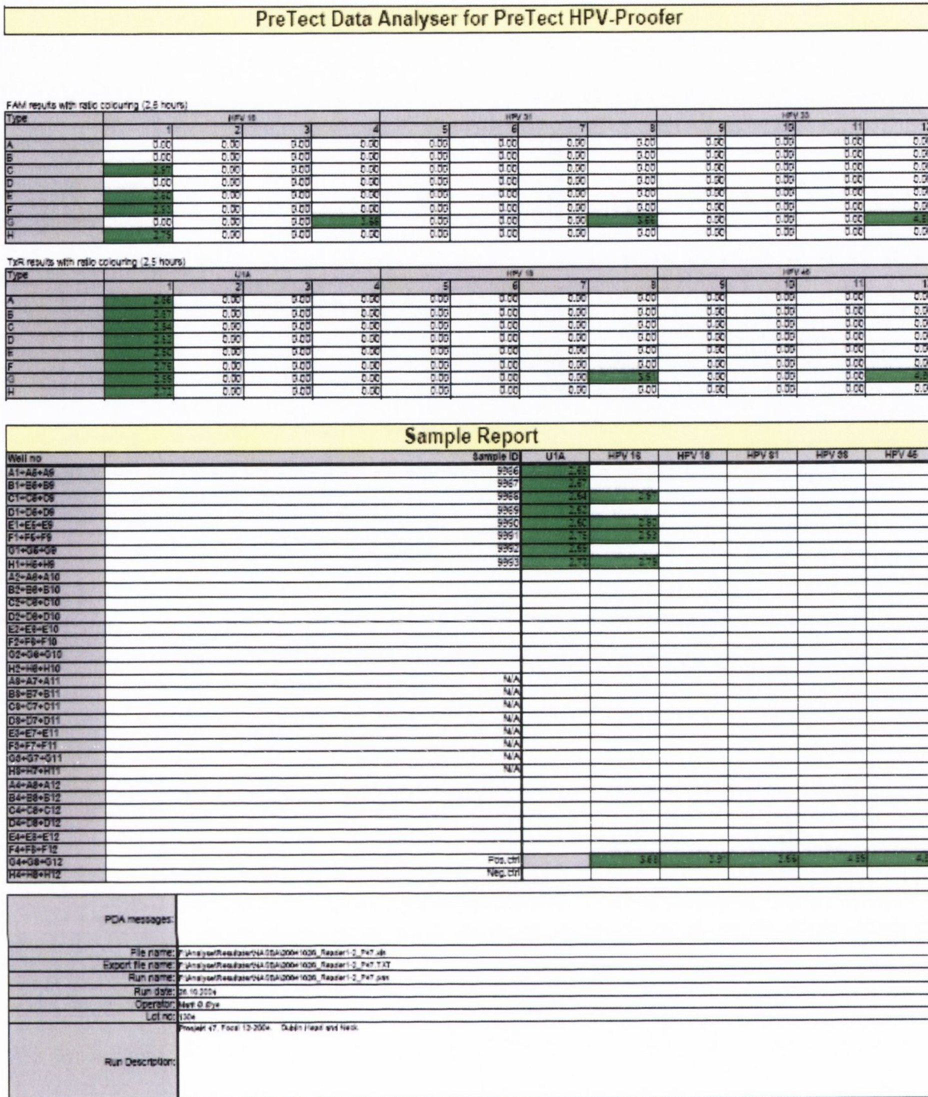
RT-PCR is not ideally suited to the detection of HPV E6/E7 mRNA transcripts. With real-time RT-PCR assays, PCR probes should be designed to span an exon-exon junction with PCR primers placed on either side of the intron. As a result, false positive results caused by contaminating genomic DNA in RNA samples can be avoided. However, this strategy is not amenable to the detection of HPV E6/E7 mRNAs given there are no introns.

Nucleic acid sequence based amplification (NASBA) is an alternative method for the amplification of mRNA. The NASBA reaction occurs at 41°C therefore double stranded DNA is not denatured and consequently not amplified (Burchill, 2002; Heim, 1998). As a result the presence of intron-exon boundaries or DNase digestion steps are superfluous considerations and contaminating HPV DNA should not affect the efficiency of the reaction or generate false positive results. For these reasons NASBA is ideally suited for the detection of HPV E6/E7 mRNA targets.

It must be noted here that while detection of E6/E7 mRNA may be highly suggestive of integration of HPV, it would be necessary to prove that the open reading frame of E2 were disrupted in order to guarantee integration.

The PreTect HPV-Proofer (NorChip AS, Norway) molecular test kit was employed for HPV mRNA detection and typing. This test is based on Real-Time NASBA technology. Nucleic acid sequence based amplification (NASBA) reaction uses the combined activity of three enzymes: avian myeloblastosis virus reverse transcriptase (AMV-RT), RNase H and T7 RNA polymerase. Together with two oligonucleotide primers they generate a logarithmic amplification of an RNA template without temperature cycling. Single-stranded RNA of opposite sense to the original target is the major amplification product. The PreTect HPV-Proofer Kit incorporates specific HPV molecular beacon probes, which hybridise to the newly generated RNA

molecules. The formation of the newly generated RNA molecules can therefore be determined in real-time by continuous monitoring of fluorescence in a Lambda Fluoro 320 reader (MWG Biotech, Germany). The PreTect HPV-Proofer Kit includes specific primer and molecular beacon probe sets for detection and simultaneous typing of two of the high risk HPV types (16, 18). The PreTect HPV Proofer Kit also contains a primer and probe set for U1A, which is used as an extraction control. Molecular beacons are labelled with either FAM or TEXAS RED fluorescent dyes. NASBA based detection of HPV mRNA was carried out as detailed in **Chapter 2**. Data analysis was carried out using PreTect Analysis Software (PAS) **Figure 5.14**.



PreTect Data Analyser for PreTect HPV-Proofer

FAM results with ratio colouring (2.5 hours)

Type	HPV 16				HPV 31				HPV 33			
	1	2	3	4	5	6	7	8	9	10	11	12
A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C	2.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E	2.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F	2.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G	0.00	0.00	0.00	3.66	0.00	0.00	0.00	3.66	0.00	0.00	0.00	4.85
H	2.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TxR results with ratio colouring (2.5 hours)

Type	U1A				HPV 18				HPV 45			
	1	2	3	4	5	6	7	8	9	10	11	12
A	2.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B	2.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C	2.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D	2.62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E	2.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F	2.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G	2.59	0.00	0.00	0.00	0.00	0.00	0.00	3.91	0.00	0.00	0.00	4.80
H	2.72	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Sample Report

Well no	Sample ID	U1A	HPV 16	HPV 18	HPV 31	HPV 33	HPV 45
A1+A5+A9	9966	2.66					
B1+B5+B9	9967	2.67					
C1+C5+C9	9968	2.54	2.97				
D1+D5+D9	9969	2.62					
E1+E5+E9	9960	2.60	2.80				
F1+F5+F9	9961	2.76	2.53				
G1+G5+G9	9962	2.59					
H1+H5+H9	9963	2.72	2.79				
A2+A6+A10							
B2+B6+B10							
C2+C6+C10							
D2+D6+D10							
E2+E6+E10							
F2+F6+F10							
G2+G6+G10							
H2+H6+H10							
A3+A7+A11	N/A						
B3+B7+B11	N/A						
C3+C7+C11	N/A						
D3+D7+D11	N/A						
E3+E7+E11	N/A						
F3+F7+F11	N/A						
G3+G7+G11	N/A						
H3+H7+H11	N/A						
A4+A8+A12							
B4+B8+B12							
C4+C8+C12							
D4+D8+D12							
E4+E8+E12							
F4+F8+F12							
G4+G8+G12	Pos. ctrl		3.66	3.91	2.56	4.85	4.80
H4+H8+H12	Neg. ctrl						

PDA messages:	
File name:	[illegible]
Export file name:	[illegible]
Run name:	[illegible]
Run date:	26.10.2004
Operator:	[illegible]
Lot no:	[illegible]
Run Description:	Prosjekt 47, Focal 12-2004. Dublin Head and Neck

Figure 5.14 PreTect Analysis Software. Positive samples highlighted in green.

5.5.2 Results

5.5.2.1 Comparison of HPV DNA and mRNA detection

Out of the 21 oral and pharyngeal carcinoma samples, eight were positive for HPV 16 DNA, and 0/21 were positive for HPV 18 DNA. Four of the 21 samples were positive for both HPV 16 DNA and mRNA, and all were negative for HPV 18 mRNA.

Comparison of HPV16 DNA and mRNA detection results are illustrated in **Figure 5.15.** As expected, all HPV 16 mRNA positive samples were also positive for HPV DNA.

5.5.2.2 Correlation between HPV status and clinical features

HPV DNA was associated with site but not significantly. It was detected in 60% of the oropharyngeal SCCs and 31% of the oral SCCs (p=0.52). The HPV16 mRNA showed a similar site predilection, with 75% of the positive cases originating in oropharyngeal tumours (p=0.05). 100% of the cases that showed integration of HPV 16 were male, while 75% were smokers. Only one of the cases showing integrated HPV was under 40 years old.

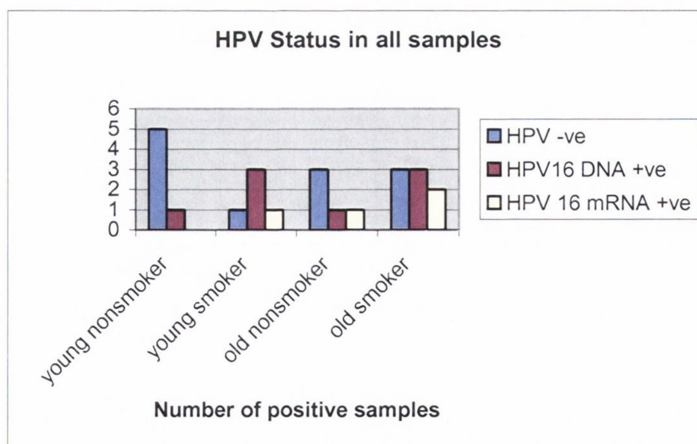


Figure 5.15 HPV status in oral and pharyngeal SCC

5.5.2.3 Correlation between p16^{INK4A} status and HPV status

- Overall the p16 status correlated very well with the HPV status (Table 5.5). HPV DNA and p16 immunohistochemistry: Taking the presence of HPV DNA sample and comparing it with p16 antibody staining, a statistically significant association was found. ($p < 0.03$). 83% of cases positive for p16^{INK4A} staining were also positive for HPV Type 16 DNA.
- E6/E7 mRNA and p16 IHC: An even stronger association was found between the samples that showed integration of HPV and p16^{INK4} antibody staining, where 100% of the mRNA positive cases demonstrated positive staining for the p16^{INK4} protein ($p=0.006$).
- E6/E7 mRNA and p16 mRNA: It was also found that p16 mRNA (measured by TaqMan PCR) expression was significantly associated with the presence of integrated HPV ($p=0.01$).

	Test	HPV DNA P values	HPV mRNA P values
P16 IHC	**Fishers Exact**	**0.03**	**0.006**
P16 mRNA	**Mann Whitney U**	**0.11**	**0.01**
Site	**Fishers Exact**	**0.59**	**0.05**
Smoker	**Fishers Exact**	**0.16**	**0.5**
Gender	**Fishers Exact**	**0.81**	**0.31**
Age	**Fishers Exact**	**1**	**0.5**

Table 5.5 Statistical analyses of results using Mann-Whitney U test and Fisher Exact Test

Figure 5.16 illustrates the prevalence of HPV across all four patient cohorts, namely young nonsmokers, young smokers, old nonsmokers and old smokers and its association with p16 expression. It is clear from this figure that deletion of p16 is an event occurring exclusively in the older cohorts, and that HPV DNA/RNA was more frequent in smoking cohorts. Where HPV DNA was detected in the nonsmokers, p16 overexpression was also found.

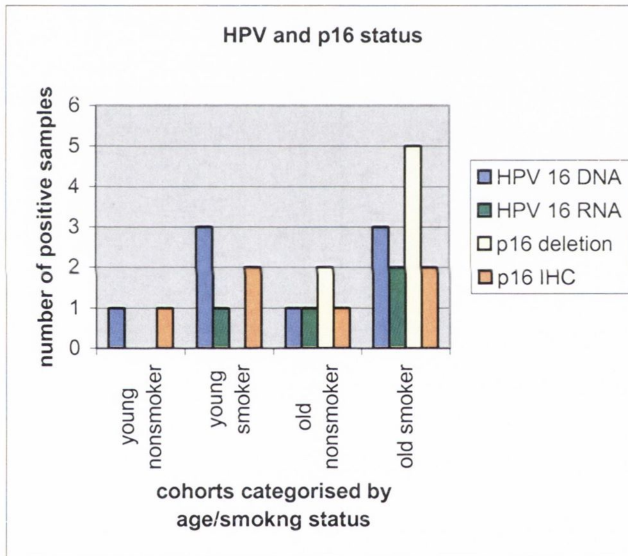


Figure 5.16 Comparison of HPV and p16INK4 status

5.6 Discussion

$p16^{INK4A}$ can be disrupted by a variety of mechanisms such as mutation, homozygous deletion (Kamb, 1994) and/or promoter hypermethylation, which result in reduced, or absent expression of $p16^{INK4A}$ (Merlo, 1995). In primary head and neck tumorigenesis, $p16^{INK4}$ inactivation is predominantly caused by homozygous deletions and methylation (Reed 1996; Gonzales 1994; El-Naggar, 1999). Overexpression, however, of $p16^{INK4}$ has also been reported in head and neck cancers (Andl, 1998; Gillison, 2000; Lang et al 2002) and in many cases it is likely that HPV infection, via inactivation of retinoblastoma, accounts for this high level of $p16^{INK4}$ expression through disruption of a negative feedback mechanism (Fregonesi, 2003). There have also been reports of $p16^{INK4}$ overexpression, independent of HPV infection (Henshall, 2001). In most published studies, immunohistochemistry has been used to identify p16 expression. The immunoreaction cut-off in most of these studies is established by quantifying nuclear and cytoplasmic positivity in normal tissue, which ranged from zero to 5%.

The aim of this study was to examine the association of $p16^{INK4}$ expression with the presence of HPV in oral and pharyngeal SCCs. Four different methods were employed in this chapter:

- TaqMan PCR to examine p16 mRNA expression
- Immunohistochemistry to examine p16 protein expression
- PCR to identify HPV DNA
- NASBA to identify mRNA expression of HPV

5.6.1 p16

Immunohistochemical analysis of $p16^{INK4A}$ protein expression was examined using mouse monoclonal antibodies. All samples were formalin fixed and paraffin embedded. Staining intensity was assessed using a 0 to 3 scoring system. P16INK4A mRNA expression was examined using quantitative TaqMan RT-PCR. TaqMan PCR was also employed for specific HPV typing.

As a component of nuclear protein, p16 should have been located only in the nuclei. However, we detected p16 in the cytoplasm as well as the nucleus in some cases. This has been noted in previous studies, but the significance remains unclear. What we can deduce is that this phenomenon is not due to technical error as control samples were completely negative. This localization issue does imply a complex relationship of p16 with tumour progression, and further studies are required to shed light on this phenomenon. The presence of p16 in the cytoplasm may result from a type of posttranslational modification or, more simply, overproduction of the protein may force its transfer into the cytoplasm.

In concurrence with the literature, the majority (76%) of the samples of oral and pharyngeal carcinoma in this study, showed no p16 protein staining. This is consistent with the concept low expression of CDK inhibitors reflects a reduced tumour suppressor role of CDK inhibitors and thus may play an important role in neoplastic transformation in oral and pharyngeal cancer. Other possible explanations for the absence of $p16^{INK4A}$ expression in these lesions could be methylation of the

$p16^{INK4A}$ promoter resulting in silencing of the $p16^{INK4A}$ gene or mutation of the p16 gene.

The epigenetic phenomenon of methylation has been described in a variety of cancers (Ferreux, 2003; Moyohanen, 1998), and a number of studies have been published regarding the frequency of $p16^{INK4A}$ promoter hypermethylation in oral carcinogenesis (Shintani, 2001; Riese, 1999), with Shintani et al reporting aberrant methylation in 50% of the oral SCCs. There are conflicting reports of an association between hypermethylation and smoking. While several groups have showed that methylation of p16 is significantly involved in the pathogenesis of cervical carcinoma (Lea, 2004), and head and neck carcinoma (Wong, 2003), there are also reports of more promoter methylation of the $p16^{INK}$ gene in nonsmokers compared to smokers in lung carcinomas (Sanchez-Cespedes, 2003). Because our study did not assess methylation status, we can only speculate that promoter methylation may explain some of the p16 negative cases.

Homozygous deletion of the p16INK4 gene has, in the past, been detected in invasive and preinvasive head and neck cancers (Shintani, 2001; Ambrosch, 2001) and indeed the inactivation of the gene by this method is thought to be an early event in head and neck carcinogenesis (Gonzalez, 1994; Van der Riet, 1994). The inactivation of the p16INK4 tumour suppressor gene can result in the loss of positive immunostaining for the protein (Reed, 1996, Shintani, 2001).

In this study, it has been shown that 80% of the tumours that demonstrated a p16 deletion (Chapter 4) showed underexpression of p16 mRNA, and lack of p16 protein expression. One sample, however, did have a p16 deletion, a high level of p16 mRNA expression and strong staining for the p16 protein. This may be explained by

the fact that in the case of p16 deletions, there may still be a functionally active p16 allele left, which can be transcribed and translated as normal. Interestingly in this case, HPV mRNA was detected, therefore inactivation of the RB gene and subsequent overexpression of p16 may have occurred despite deletion of one allele.

Mutations of the p16 gene are relatively uncommon in oral and oropharyngeal tumours, however there is a possibility that p16 mutations may have occurred. Mutational analysis of p16 were not assessed in this thesis.

It was expected that a concurrent increase in $p16^{INK4A}$ at the mRNA level would be observed in the 6 cases that demonstrated $p16^{INK4A}$ protein overexpression. Results of the relative quantitation TaqMan RT-PCR analysis mirrored the IHC results, with the positive p16 protein samples demonstrating high levels of mRNA expression. However, a number of the p16 protein negative cases also showed high levels of p16 mRNA expression. This might be explained by the presence of a mutation outside the area amplified in the RT-PCR reaction, or another alteration affecting transcription or translation of the p16 gene.

Overall, p16 mRNA expression significantly correlates with p16 protein expression. Tumour site was significantly associated with p16 protein positivity with tongue base showing the highest number of p16 positive tumours. There was no significant correlation between age, gender and smoking history, although there was a trend towards an association between low levels of p16 mRNA expression and over forty year olds (p=0.08)

5.6.2 HPV

Normal individuals have a wide range of reported prevalence rates for HPV, which are likely dependent on sample collection methods and detection methods employed.

Recent large studies reported a prevalence of 1-2% in normal individuals (Lambropoulos, 1997, Smith, 1998), however smaller studies have reported the prevalence of HPV in normal oral mucosa as high as 27% (Nagpal, 2002). Molecular epidemiological studies have demonstrated that the vast majority of HPV infections in humans are transient and clinically non-significant. Consequently, testing for HPV mRNA expression rather than merely the presence of HPV DNA may be a more relevant clinical indicator for the development of oral cancer.

A number of cases in this study were HPV 16 DNA positive, but lacked HPV 16 mRNA detection and lacked p16 expression, suggesting that it is HPV 16 mRNA presence rather than mere presence of HPV DNA in a lesion that gives rise to abnormal overexpression of p16.

Site: Recent studies support an aetiologic role for a number of high risk HPV types in at least a subset of head and neck cancers, especially in tonsillar carcinoma (Herrero, 2003; Gillison, 2000). The majority of studies involving oral lesions do not specify subsites, and there is often no distinction made between the oral cavity and the oropharynx. In light of the general agreement that oropharyngeal carcinomas are frequently associated with HPV (Gillison, 2000), it is imperative that this distinction is made.

In this study, HPV mRNA was more often detected in the oropharyngeal samples. In fact, while 60% of the oropharyngeal cases showed HPV mRNA, E6/E7 mRNA expression was detected in only 6% of the oral samples. This finding supports previous studies that have found a higher prevalence of both HPV DNA and HPV mRNA expression in tonsillar SCCs and other sites in the oropharynx (Andl, 1998). It is possible, therefore, that HPV has a role to play in the development of oropharyngeal carcinomas and far less of a role to play in oral cavity SCCs. Since

most of the young patients in this study have oral SCC, this may explain the lower prevalence of HPV mRNA detection in this cohort. These findings highlight the need for careful site categorisation in oral and oropharyngeal SCC studies.

Gender: All of the cases that showed HPV integration were male. This represented 28.5% of all the males in the study, compared to 0% of the females. Other studies have also reported a higher prevalence of HPV in males with head and neck cancer (Ibeita, 2005; Zhang, 2004)

Smoking status: Based on the evidence provided by the largest epidemiological studies that using sensitive detection methods allowed for the effects of HPV, it can be concluded that, among HPV positive women, high parity, long-term OC use, smoking, and co-infection with other sexually transmitted agents are the most consistently identified environmental co-factors likely to influence the risk of progression from cervical HPV infection to invasive cervical cancer (Castellague, 2002). In the case of oral and oropharyngeal cancer the literature is mixed and while Herrero et al did find a higher prevalence of HPV infection in never-smokers (Herrero, 2003), most studies have found equal or marginally lower rates of HPV detection in oral/oropharyngeal cancer in smokers. In this study, 75% of the cases showing HPV16 mRNA were smokers (Koch, 1999; Dahlstrom, 2003).

These findings coupled with the site predilection of HPV16 mRNA positive tumours for the oropharynx, suggest that it is unlikely that the HPV has an important role to play in the tumours occurring in the young female nonsmokers.

It is noteworthy that DNA methylation provides an additional means to regulate HPV transcription directly. It has been reported that HPV16 methylation had been involved in HPV oncogene expression in cervical lesions and the early late switch (Kalantari, 2004). While this would certainly be worthwhile to investigate the role of

methylation of HPV in oral cancers, it was not the aim of this thesis to investigate how HPV is regulated in oral cancer, but rather whether HPV was associated with oral cancer in young adults.

5.6.3 HPV and p16

In the present study, p16 positivity patterns were found to correlate with HPV DNA detection ($p<0.01$) and even more so with HPV mRNA detection ($p<0.001$)
This concurs with previous studies that found a strong association between high-risk HPV types and p16 overexpression in oral cancer (Fregonesi, 2003). Interestingly, in the over forty year olds, HPV DNA and mRNA was detected in 100% of the p16 positive cases, whereas in the younger patients, p16 staining was found in two cases that were negative for HPV mRNA. Milde-Langosch (2001) reported $p16^{INK4A}$ overexpression in 41% of HPV negative carcinomas. These findings may indicate that a non-HPV E7 mediated mechanism of $p16^{INK4A}$ upregulation may also exist. Loss of transcriptional repression in the presence of inactivating mutations in the pRb gene is the most well defined mechanism of a non-HPV related mechanism of $p16^{INK4A}$ upregulation. Indeed, $p16^{INK4A}$ expression may in some cases be independent of pRB. Henshall (2001), suggest that continued proliferation, despite the growth inhibitory effects of increased levels of $p16^{INK4A}$ expression may be possible when cells also overexpress cyclin E, when or $p27^{Kip}$ is unavailable (Henshall, 2001).

5.7 Conclusion

We have examined HPV status with p16 mRNA and protein expression in oral and pharyngeal SCCs in order to contribute to the emerging literature suggesting that HPV infection and p16 overexpression are in some way related to the development of oral and pharyngeal cancer. In summary, this study concludes that mRNA expression of high-risk HPV16 does occur in a number of oral and oropharyngeal SCC, and may be a contributing factor in a subset of these malignancies but it is not a necessity, as it is in cervical cancer. There is a highly significant association between p16 mRNA and protein expression in tumours that show HPV mRNA expression, which is strongly suggestive of viral integration and subsequent malfunction of p16. This finding supports the use of p16 as a marker of high-risk HPV16 presence in head and neck cancer. This study also notes that HPV mRNA expression occurs more often in oropharyngeal tumours than cancers of the oral cavity. Perhaps most interesting of all was that none of the young female nonsmokers showed viral mRNA expression or positive p16 protein expression. This strongly suggests that HPV, previously proposed as a key risk factor in young nonsmokers, does not play an important role in the development of SCC in this cohort. Also of note is the overexpression of p16 mRNA in two of these young female nonsmokers, suggesting that while HPV mRNA may not be a risk factor, some other factor is stimulating overexpression of p16 in some of these patients. These findings, along with the absence of p16 gene deletions in this cohort (Chapter 4) suggest that alternative mechanisms targeting p16 and related pathways should be investigated further in this group of patients.

5.8 References

Ambrosch P, Schlott T, Hilmes D, Ruschenburg I. P16INK4 Alterations and retinoblastoma protein expression in SCC and neighboring dysplasia from the upper aerodigestive tract. Virchows Archive 201; 438:343-349

Andl T, Kahn T, Pfuhl A, Nicola T, Erber R, Conradt C, Klein W, Helbig M, Dietz A, Weidauer H, Bosch FX. Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinoma lacking retinoblastoma cell cycle control. Cancer Res 1998; 58:5-13.

Baker CC, Phelps WC, Lindgren V, Braun MJ, Gonda MA, Howley PM. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. J Virol 1987; 61:962-971

Cairns P. Rates of p16INK4 (MTS1) mutations in primary tumours with 9p loss. Science 1994; 265: 415±6.

Castellsague X, Bosch FX, Munoz N. Environmental co-factors in HPV carcinogenesis. Virus Res. 2002; 89(2):191-9

Cruz IBF, Snijders PJF, Steenbergen RDM, Meijer CJLM, Snow GB, Walboomers JMM, van der Waal I. Age-dependence of human papillomavirus DNA presence in oral SCCs. Oral Oncology, 1996; 32b: 55-62

Cuzick J. Role of HPV testing in clinical practice. Virus Res. 2002; 89(2): 263-9

Franco EL, Duarte-Franco E, Ferenczy A. Cervical cancer: epidemiology, prevention and the role of human papillomavirus infection. CMAJ 2001; 164(7): 1017-25.

Dahlstrom KR, Adler-Storthz K, Etzel CJ, Liu Z, Dillon L, El-Naggar AK, Spitz MR, Schiller JT, Wei Q, Sturgis EM. Human papillomavirus type 16 infection and squamous cell carcinoma of the head and neck in never-smokers: a matched pair analysis. Clin Cancer Res. 2003; 9(7):2620-6.

Danahey DG, Tobin EJ, Schuller DE, et al. p16 mutation frequency and clinical correlation in head and neck cancer. Acta Otolaryngol. 1999; 119:285-288.

De Roda Husman AM, Walboomers JMM, Van den Brule AJC, Meijer CJLM, Snijders PJF. The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by polymerase chain reaction. J Gen Virol 1995, 76, 1057-1062

El-Naggar AK, Lai S, Clayman G, Lee L-JK, Luna MA, Goepfert H, Batsakis JG. Methylation, a major mechanism of p16INK4 gene inactivation in head and neck SCC. Am J Pathol 1997; 151: 1767-1774

El-Naggar AK, Lai S, Clayman GL, et al. Expression of p16, Rb, and cyclin D1 gene products in oral and laryngeal squamous carcinoma: biological and clinical implications. Hum Pathol. 1999; 30: 1013-1018.

Fregonesi PAG, Teresa DB, Duarte RA, Neto CB, Brancini de Oliveira MR, Soares CP. P16INK4 immunohistochemical overexpression in premalignant and malignant oral lesions infected with human papillomavirus. J Histochem Cytochem. 2003; 51(10):1291-7.

Gillison ML, Koch WM, Shah KV. Human papillomavirus in head and neck squamous cell carcinoma: are some head and neck cancers a sexually transmitted disease? Curr Opin Oncol 1999;11(3):191–9.

Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L *et al.* Evidence of a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst. 2000; 92(9): 675-677, 709-720.

Gonzales MV, Pello MF, Lopez-Larrea C, Squarez C, Menendez NJ, Coto E. Loss of heterozygosity and mutation analysis of the p16INK4 and p53 genes in SCC of the head and neck. Clin Cancer Res 1994; 1:1043-9

Ha PK, Califano JA. The role of Human papillomavirus in oral carcinogenesis. Critical reviews in Oral Biology and medicine. 2004; 15:188-196.

Henshall SM, Quinn DI, Lee CS, Head DR, Golovsky D, Brenner PC, Delprado W, Stricker PD, Grygiel JJ and Sutherland RL.Overexpression of the cell cycle inhibitor p16INK4 in high-grade prostatic intraepithelial neoplasia predicts early relapse in prostate cancer patients. Clin Cancer Res 2001; 7:544-550.

Herrero R, Castellsague X, Pawlita M, Lissowska J, Kee F, Balaram P, Rajkumar T. Human papillomavirus and oral cancer. The International Agency for Research n Cancer multicenter study. J Natl Cancer Inst 2003; 95:1772-1783.

Ibieta BR, Lizano M, Fras-Mendivil M, Barrera JL, Carrillo A, Ma Ruz-Godoy L, Mohar A. Human papilloma virus in oral squamous cell carcinoma in a Mexican population. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2005; 99(3):311-5.

Jares P, Fernandez PL, Nadal A, et al. p16MTS1/CDK4I mutations and concomitant loss of heterozygosity at 9p21-23 are frequent events in squamous cell carcinoma of the larynx. Oncogene. 1997; 15:1445-1453.

Kalantari M, Calleja-Macias I, Tewari D, Hagmar B, Lie K, Barrera-Saldana H, Wiley D, Bernard H. Conserved Methylation Patterns of Human Papillomavirus Type 16 DNA in Asymptomatic Infection and Cervical Neoplasia. J Virol. 2004 78(23): 12762–12772.

Lambropoulos AF, Dimitrakopoulos J, Frangoulides E, Katopodi R, Kotsis A, Karakasis D. Incidence of human papillomavirus 6, 11, 16, 18, 33 in normal oral mucosa of a Greek population. Eur J Oral Sci 1997; 105: 294-297

Lang JC, Brochers J, Danahey D, Smith S, Stover DG, Agrawal A, Malone JP, Schuller DE, Weghorst CM, Holinga AJ, Lingam K, Patel CR, Esham B. Mutational status of overexpressed p16INK4 in head and neck cancer: evidence for germline mutation of p16INK4/p14. Int J Oncol. 2002 ;21(2):401-8.

Lea JS, Coleman R, Kurien A, Schorge JO, Miller D, Minna JD, Muller C. Aberrant p16 methylation is a biomarker for tobacco exposure in cervical squamous cell carcinogenesis. American Journal of Obstetrics and Gynecology (2004) 190, 674e9

Llewellyn CD, Linklater K, Bell J, Johnson NW, Warnakulasuriya KA. Squamous cell carcinoma of the oral cavity in patients aged 45 years and under: a descriptive analysis of 116 cases diagnosed in the South East of England from 1990 to 1997. Oral Oncol. 2003 Feb; 39(2): 106-14.

Matsuda H, Konishi N, Hiasa Y, Hayashi I, Tsuzuki T, Tao M, Kitahori Y. Alterations of p16INK4, p53, ras genes in oral squamous cell carcinomas and premalignant lesion. J Oral Pathol Med 1996; 25:232-238

McHaig McKaig RG, Baric RS, Olshan AF. Human papillomavirus and head and neck cancer: epidemiology and molecular biology. Head Neck 1998; 20:250-65

Mork J, Lie AK, Glattre E, Hallmans G, Jellum E, Koskela P, Moller B, Pukkala E, Schiller JT, Youngman L, Lehtinen M, Dillner J. Human papillomavirus infection as a risk factor for squamous cell carcinoma of the head and neck. N Engl J Med, 2001; 344: 1125-1131

Murphy N, Ring M, Killalea AG, Uhlmann V, O'Donovan M, Mulcahy F, Turner M, McGuinness E, Griffin M, Martin C, O'Leary JJ. P16INK4 as a marker for cervical dyskaryosis: CIN and cGIN in cervical biopsies and ThinPrep smears. J Clin Pathol 2003; 56:56-63

Nagpal JK, Patnaik S, Das BR. Prevalence of high-risk human papilloma virus types and its association with P53 codon 72 polymorphism in tobacco addicted oral squamous cell carcinoma (OSCC) patients of Eastern India. Int J Cancer. 2002; 97(5):649-53.

Nakanishi H, Uzawa K, Yokoe H et al. Rare mutations of the growth suppressor genes involved in negative regulation of the cell cycle: p15, p16INK4 and p18 genes in oral squamous cell carcinoma. Int J Oncol 1997; 11:1129-33

Pande P, Mathur M, Shukla N, et al. pRb and p16 protein alterations in human oral tumorigenesis. Oral Oncol. 1998; 34:396-403.

Plummer M, Herrero R, Franceschi S, Meijer CJ, Snijders P, Bosch FX, de Sanjose S, Munoz N; IARC Multi-centre Cervical Cancer Study Group. Smoking and

cervical cancer: pooled analysis of the IARC multi-centric case--control study. Cancer Causes Control. 2003;14(9):805-14.

Poppe WA, Ide PS, Drijkoningen MP, Lauweryns JM, Van Assche FA. Tobacco smoking impairs the local immunosurveillance in the uterine cervix. An immunohistochemical study. Gynecol Obstet Invest 1995; 39: 34–8.

Prokopczyk B, Cox JE, Hoffmann D, Waggoner SE. Identification of tobacco-specific carcinogen in the cervical mucus of smokers and nonsmokers. J Natl Cancer Inst 1997; 89: 868–73

Reed AL, Califano J, Cairins P, Westra WH, Jones RM, Koch W, Ahrendt S et al. High frequency of p16INK4 (CDKN2/MST-1/INK4A) inactivation in head and neck squamous cell carcinoma. Cancer Res 1996, 56:3630-3633

Riese U, Dahse R, Fiedler W, Theuer C, Koscielny S, Ernst G et al. Tumour suppressor gene p16 (CDKN2A) mutation status and promoter inactivation in head and neck cancer. Int J Mol Med 1999; 4:61-65.

Sartor M, Steingrimsdottir H, Elamin F, et al. Role of p16/MTS1, cyclin D1 and RB in primary oral cancer and oral cancer cell lines. Br J Cancer. 1999; 80:79-86.

Schantz SP, Yu GP. Head and neck cancer incidence trends in young Americans, 1973-1997, with a special analysis for tongue cancer. Arch Otolaryngol Head Neck Surg. 2002 Mar; 128(3): 268-74.

Schwartz SM, Daling JR, Doody DR, Wipf GC, Carter JJ, Madeleine MM et al. Oral cancer risk in relation to sexual history and evidence of HPV infection. J Natl Cancer Inst. 1998; 90: 1626-36

Scully C, Field J.K., Tanzawa H. Genetic aberrations in oral or head and neck squamous cell carcinoma 2: chromosomal aberrations. Oral Oncology, 2000; 36: 311-27

Shintani S, Nakahara Y, Mihara M, Ueyama A, Matsumura T. Inactivation of the p14, p15 and p16INK4 gene is a frequent event in human oral squamous cell carcinomas. Oral Oncol 2001; 37; 498-504

Shope R. Infectious papillomatosis of rabbits. J Experimental Medicine 1933; 58: 607-24

Smith EM, Hoffman HT, Summersgill KS, Kirchner HL, Turek LP, Haugen TH. Human papillomavirus and risk of oral cancer. Laryngoscope 1998; 108: 1098-1103

Syrjanen K, Syrjanen S, Lamberg M, Pyrhonen S, Nuutinen J. Morphological and immunohistochemical evidence suggesting human papillomavirus (HPV) involvement in oral squamous cell carcinogenesis. Int J Oral Surg 1983; 12: 418-424.

Van der Riet P, Nawroz H, Hruban RH, Corio R, Tokino K, Koch W et al. Frequent loss of chromosome 9p21-22 early in head and neck cancer progression Cancer Res 1994; 54:1156-8

Van Rensburg EJ, Engelbrecht S, Van Heerden WFP, Raubenheimer EJ, Schoub BD. Human papillomavirus DNA in oral squamous cell carcinomas from an African population sample. Anticancer Research 1996; 16:969-74

Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J. Pathology 1999; 189: 12-19.

Zhang Z-y, Sdek P, Cao J, Chen W-t. Human papilloma virus type 16 and 18 DNA in oral squamous cell carcinoma and normal mucosa. International Journal of Oral and Maxillofacial surgery 2004; 33: 71-74

Zochbauer-Muller C, Lam S, Toyooka S, Virmani AK, Toyooka K, Seidl S, Minna JD, Gazdar A. Aberrant methylation of multiple genes in the upper aerodigestive tract epithelium in heavy smokers. Int J Cancer 2003; 612-618

Chapter 6

Expression arrays

6.1 Summary

While oral and pharyngeal SCC is classically a disease of older male smokers, this tumour type can occur in young patients with minimal or no exposure to the traditional risk factors. Findings from Chapter 4 suggest that the genomic profile of the tumours from these young patients is very different from that of cancers of the typical oral and oropharyngeal cancer patient i.e the older male smoker.

Changes in the physiological state of cells and tissues generally lead to specific changes in gene expression patterns, thus microarray analysis of oral and pharyngeal SCCs can provide valuable information regarding tumour biology. In this chapter the expression patterns of oral and oropharyngeal tumours is investigated.

The development of cDNA microarray technology allows us to perform genome-wide screening of differentially expressed genes between clinical subtypes of oral and oropharyngeal SCCs.

RNA was extracted from 20 oral and oropharyngeal SCC samples and 1 sample of normal oral mucosa. A novel chemiluminescence expression array system was used to study the expression of 31,600 genes in the samples. Chemiluminescence provides improved performance by ensuring high signal to noise values in combination with low detection limits.

Using a distinction calculation method it was possible to identify the most informative up and down regulated candidate biomarker genes, which distinguish young nonsmokers from old smokers. ANOVA tests revealed a number of statistically significant candidate genes whose expression distinguished between young nonsmokers and old smokers.

A group of genes that distinguish the SCC found in young nonsmokers from that found in old smokers has been identified. This is the first expression microarray evidence of the different patterns in young nonsmokers with oral and oropharyngeal cancer compared to the old smokers.

6.2 Introduction

6.2.1 Expression microarray studies

With the recent avalanche of data from microarray studies, there has been a fundamental transformation in cancer research, which has changed from focused studies on individual genes and proteins to global studies of cells, tissues or the entire organism. By examining the expression of thousands of genes simultaneously microarrays provide a snapshot of the genome-wide transcriptional processes. Predicting who will develop cancer and how this disease will behave and respond to therapy after diagnosis are some of the potential benefits of this technology.

Microarrays can be used to investigate problems in cell biology in various ways; and the different experimental approaches fall between two extremes. At one end, the investigator is interested only in finding the single change in gene expression that might be the key to a given alteration in phenotype, and in some ways this is like looking for a needle in a haystack, and could be thought of as an entirely local approach to analysis of gene-expression changes. At the other extreme, the aim is to look at overall patterns of gene expression in order to understand the architecture of genetic regulatory networks, a global approach that could ultimately lead to complete description of the transcription-control mechanisms in a cell. Definition of

a genome wide expression pattern for disease states will enhance our understanding of cell biology and may provide a link between molecular genetics and clinical practice (Van't Veer LJ, 2003; West, 1998; Mazzanti 2004).

6.2.2 Applied Biosystems Human Genome Survey Microarray

The Applied Biosystems Human Genome Survey Microarray is a comprehensive gene expression-profiling system that combines high-sensitivity chemiluminescence detection, and complete, annotated, fully curated human genome transcript data.
It allows researchers to survey and measure gene expression over the entire human genome in a single experiment, thus enables the determination of differential gene expression over the entire human genome in various populations, tissues, developmental stages and disease states.
The array provides 31,700 probes for the interrogation of 27,868 genes. Of this high quality gene set, 28% have been curated exclusively by Celera Genomics. The current version (March, 2004) of the Celera Discovery System (CDS) catalogues 31,566 human genes. The oligonucleotide probes on the array are 60 base pairs (60 mers) and are designed to report an aggregate gene activity level from alternatively spliced transcripts of a gene, and they are designed not to discriminate between the abundance levels of alternative transcripts.

6.2.3 Expression profiles of head and neck cancer

Until now classification of cancer has been primarily based on histopathological assessment with clinicopathological correlation. Past studies that attempted to

characterize and predict tumor behavior in head and neck carcinoma, including oral carcinoma, placed great emphasis on the examination of genetic loci that are thought to be associated with tumor suppressor genes. There also have been examinations of individual genes or proteins (Califano, 1997), such as p53, p16, c-erb-B2 and H-ras, K-ras, and N-ras (Oh, 1997). Studies in gene expression profiling of tumor samples have revealed the great transcriptional heterogeneity of cancer. Altogether, they have shown the advantages of the microarray to discriminate between RNA expression levels of different genes. Furthermore, among classically indistinguishable tumors, these studies have allowed the classification of new clinically and biologically important subclasses that may represent different diseases requiring different strategies. With the advent of DNA microarrays, genomic-scale differential gene expression profiles can be obtained.

A number of recent studies have reported gene expression profiles of HNSCC patients using microarrays (Leethanakul, 2000; Alevizos, 2001; Al Moustafa, 2002; Belbin, 2002; El-Naggar, 2002; Mendez, 2002). These studies use a variety of array platforms, tissue types, and experimental designs. Most of these studies have focused predominantly on the overall expression profiles of HNSCC or oral cancer in comparison to normal tissue (Leethanakul, 2000; Alevizos, 2001; Al Mustafa, 2002). Ginos et al (2004) correlated a gene expression signature with risk for local treatment failure. In a similar way, El Naggar (2002) used expression microarrays to distinguish between conventional SCCs and variant SCCs.

Mendez et al (2002) aimed to identify a gene expression signature that would differentiate early-stage from late-stage disease or metastatic from non-metastatic

cancers. They concluded that gene expression signatures do not correlate well with stage of disease. This lack of correlation can be explained by the hypothesis that changes in genetic expression occur relatively early in carcinogenesis. This implies that early tumors already may contain genetic changes that allow more advanced tumors to grow, invade, and metastasize.

While the histomorphology of the oral and oropharyngeal SCCs occurring in young patients with no risk factors is essentially the same as that occurring in older patients with risk factors, we have already found different genomic profiles in these two groups, suggesting that the molecular signature of these two groups are not exactly the same.

Since the gene expression profile of a cell determines its function, phenotype, and response to external stimuli, comparing the signatures of these two groups may identify some of the reasons why increasing numbers of young nonsmokers are developing SCCs of the oral cavity and oropharynx. We hypothesised that genome-wide identification of those gene expression changes that are attributable to underlying gene copy number alterations would highlight transcripts that are actively involved in the causation or maintenance of the malignant phenotype. We compared the genome profiles identified using CGH microarray technology (chapter 4) with the expression profiles identified with the cDNA microarrays to: (*a*) determine the global impact that gene copy number variation plays in oral and oropharyngeal cancer development and progression; and (*b*) identify and characterise those genes whose mRNA expression is most significantly associated with amplification of the corresponding genomic template.

6.3 Aims

The aim of this chapter is to identify gene expression signatures that differentiate between the oral and oropharyngeal SCCs occurring in young nonsmokers and those conventional SCCs occurring in old smokers using the new Applied Biosystems chemiluminescent-based expression microarray platform.

6.4 Materials and Methods

6.4.1 Samples

Samples were specifically restricted to oral and oropharyngeal. The term 'oropharyngeal' represents the oropharynx i.e samples from base of tongue, soft palate, fauces and pharyngeal wall from the plane of the hard palate superiorly to the plane of the hyoid bone inferiorly. All specimens for microarray analysis were collected fresh from theatre at St. James's Hospital and immediately snap frozen in liquid nitrogen or immersed in RNA Later. All specimens were collected prospectively. Diagnosis of the lesion was performed by a consultant pathologist, on biopsy prior to resection. The presence of lesional tissue in the banked frozen samples was confirmed by frozen section examination prior to nucleic acid extraction. The clinical details for each sample can be found in Table 6.1.

6.4.2 RNA Extraction, *In vitro* transcription

RNA extraction was performed as outlined in chapter 2 (Section 2.5.3) using the Qiagen RNeasy protocol. Amplification of RNA was performed using the Eberwine (1996) procedure in order to generate enough RNA to hybridise to the array. Between 1μg –10 μg of starting total RNA is recommended for input to the reverse transcription-in-vitro transcription reaction. Generally between 3μg-5μg total RNA was used. The integrity and quantity of extracted nucleic acid were verified by gel electrophoresis and by UV spectrophotometry respectively. The specific details and protocols for RNA labelling and hybridisation are outlined in Chapter 2 (Section 2.13).

Number	Age	Gender	Smoking status	Site	Array I.D.
1	30	F	No	T	HA006DJ
2	35	F	No	T	HA00205
3	32	F	No	T	HA006DA
4	37	F	No	T	HA000LP/L8
5	39	M	No	A	HA006DH
6	39	M	Yes	T	HA000LW
7	38	M	Yes	T	HA006DB
8	34	M	Yes	FOM	HA000LC
9	39	M	Yes	T	HA00232
10	60	M	Yes	A	HA001CK/DN
11	60	M	Yes	TB	HA001BJ
12	43	M	Yes	TB	HA0023D
13	68	M	Yes	TB	HA00124
14	70	M	No	P	HA006DU
15	58	M	No	T	HA006DV
16	82	F	No	T	HA000L1
17	69	F	No	T	HA001FD
18	65	M	Yes	TB	HA000LF
19	62	M	Yes	T	HA001ZT
20	67	M	Yes	T	HA001FC
21	29	F	No	T	HA001BU

Table 6.1

Sample Details: Tongue, TB: Tongue base, FOM: floor of mouth, A: alveolus, P: palate

6.4.3 Chemiluminescence Reaction

Chemiluminescence is the generation of electromagnetic radiation as light by the release of energy from a chemical reaction. While the light can, in principle, be emitted in the ultraviolet, visible or infrared region, those emitting visible light are the most common. An overview of the chemiluminescence process is seen in Figure 6.1.

High signal strength and absence of background noise in chemiluminescent technology allows detection of extremely low levels of gene expression, compared with alternative fluorescent techniques.

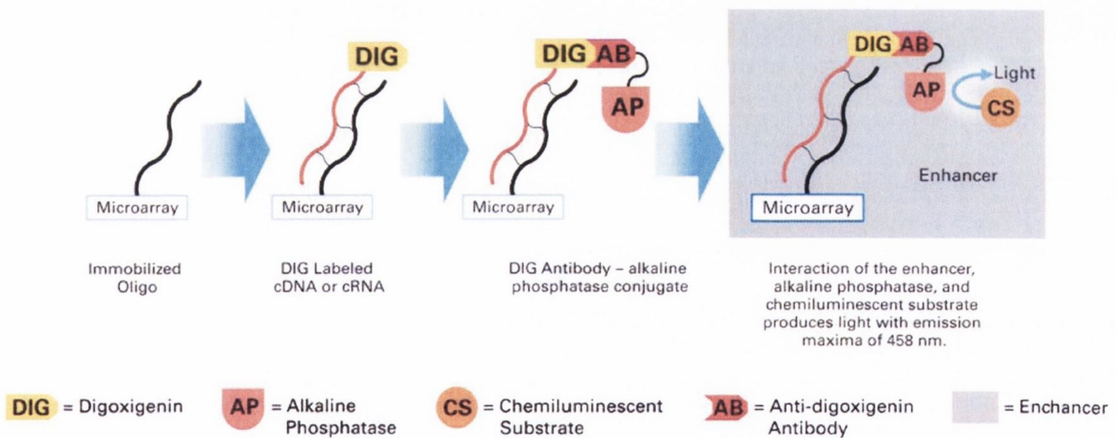


Figure 6.1 Overview of Chemiluminescence chemistry

6.4.4 Analysis of Microarray Data

6.4.4.1 Replicates

In any experiment, both random and systematic bias tends to occur. Arrays and samples have different qualities (purity, reaction efficiency) that lead to different overall expression values on different chips. One cannot remove the random bias but it is possible to minimise it by using replicates.

Technical replicates between slides refers to replication in which the target mRNA is from the same pool, that is, from the same extraction. Aliquots of mRNA from a single extraction were used as technical replicates.

Samples labelled '12166' and '12166 rep' are technical replicates of RNA from the same oral SCC sample and 'LP' and 'LP rep' are technical replicates of the same oral SCC sample.

The term 'biological replicates' usually refers to hybridisation that involves mRNA from a different extraction – for example, from different samples of cells from a particular cell line or tissue. In this respect, this form of control is intimately connected with the statistical extrapolation from sample to population, and therefore is recommended as the principal source of control. The term can also mean the target mRNA comes from different individuals as in this study, and this form of biological replication involves a much greater degree of variation in measurements.

6.4.4.2 Normalisation

The process by which data from different channels are equalised before analysis is known as normalisation. The normalisation process does not alter the content of the

data but rather corrects for a number of technical variations between and within single hybridisations, namely quantity of starting RNA and labelling and detection efficiencies for each sample. Normalisation is an attempt to correct for systematic bias in data.

There is a variety of normalisation schemes in use and various forms of global normalization are applicable to the data presented in this chapter. Global normalisation makes an assumption that given a large enough sample the average signal intensity (corresponding to gene intensity level) doesn't change. Global normalisation methods that can be used include Median, Quantile, and Trimmed Mean. MA plots (Log2 ratios plotted (M) against average intensities (A)) are plots used to identify signal dependent biases and MA plots can be used to compare the effect of different normalisation methods.

6.4.4.3 Analysis by Spotfire

Raw signal intensities were exported from the ABI 1700 to a proprietary data analysis platform (Spotfire, Decision Site for Functional Genomics).

Additional statistical calculations were performed using the R statistical package (http://www.r-project.org/). The data was filtered to remove data with a Signal/Noise ratio < 3 as recommended by the manufacturers.

6.4.4.4 Filtering genes

A very basic filtering was applied initially, where flagged genes and those with a signal-to-noise ratio (S/N) of less than 3 were deemed undetectable and removed

from further analysis. All control elements on the array such as hybridisation controls, RT-IVT controls, etc. were also removed prior to further analysis, after being screened to ensure the data met the quality control requirements.

6.4.4.5 Cluster Analysis

Cluster analysis is based on a metric for measuring distance between expression profiles. This metric is generally based on the complete set of genes measured on the microarray or on a subset showing greatest variability across the arrays. It is not based on information about which genes are informative for discriminating among predefined classes. Consequently, cluster analysis is not very powerful for distinguishing classes that differ with regard to a relatively small number of genes. Hierarchical clustering was performed initially on all the samples and then on the differential gene lists that were subsequently generated to determine whether technical replicates and cell line types grouped appropriately. Clustering was performed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) with Euclidian distance as the similarity measure. Average value was used as the ordering function.

6.4.4.6 Treatment comparison

Recent developments in transcriptome-oriented biotechnologies have made possible the comparative analysis of thousands of gene expression changes in parallel. With

the advent of microarray technology the case of n<<p (thousands of genes and a very small number of biological replicates) has been adopted.

One very useful type of study involves comparison of expression profiles obtained from differential predefined classes of specimens i.e Class Comparison. There are three objectives in Class Comparison studies. Firstly, to determine whether the expression profiles differ among the classes. Secondly, to identify which genes are differentially expressed among the classes and the third objective is to develop a multivariate predictor of class membership based on the level of expression of selected genes.

The Treatment Comparison tool in Spotfire uses a *t*-test (for 2 groups) or ANOVA (more than 2 groups). For comparison of two conditions (e.g. young nonsmokers and old smokers), a *t*-test is the method of choice. The *t*-test looks at the mean and variance of the sample and control distributions and calculates the probability that the observed difference in mean occurs when the null hypothesis is true. The null hypothesis states that the mean of the two distributions is equal.

An alternative way of looking for biomarkers is to use the distinction calculation (Golub, 1999) The distinction value is a measure of how distinct the difference in expression level is between different parts of a profile. This is particularly useful when looking for genetic biomarkers. An example would be a gene that shows high expression for a certain type of tumor cells but not for healthy cells.

The aim is to look for genes that show a distinct difference in expression between the groups. While there should be a large difference in expression between the groups there should also be as little variation as possible within the groups. The

profiles are compared to an idealised expression pattern in which the expression level is uniformly high in the first group and uniformly low in the second group, etc. The calculated distinction values are a measure of how similar each profile is with this ideal profile.

Profiles with high positive distinction values have high expression values in the first group and low expression values in the second group. Profiles with high negative distinction values have low expression values in the first group and high expression values in the second group.

In the context of t-tests, setting a cut-off for differential expression is problematic because one has to balance the false positives (Type 1 error) and the false negatives (Type 2 error). Until now, statistical procedures have mostly relied on the multiple testing framework in order to control false positive conclusions. In this framework, two quantities have been considered: the Family Wise Error Rate (FWER) and the False Discovery Rate (FDR). However, controlling the FWER in multiple testing settings may not always be appropriate. Indeed, in large-scale hypothesis generating studies such as microarray experiments, this criterion becomes so conservative that the probability of detecting any true association is, in some cases, almost nil. The FDR is the expected proportion of erroneously rejected null hypotheses among the rejected ones. The main reason for controlling the FDR is that it controls a quantity that is relevant and leads to more powerful procedures than those relying on the FWER.

Many researchers tend to dismiss the issue of multiplicity in microarray data analysis, arguing that these experiments merely serve for screening and their

purpose is to supply the researcher with an initial pool of candidates. Therefore, statistical considerations that limit the power to generate candidate hypotheses should not be taken at this stage. This argument has been deemed acceptable in the sense that the severe loss of power involved in protection against false positives is not justified when one is planning on further validation.

6.4.4.7 Adjusting for fold-change noise

When comparing two samples for fold-change differences, results were adjusted for fold-change noise. This was accomplished using the following:

Plot: Log 2 (Sample 1/Replicate 1) against (Log2 (Sample 1 x Replicate 1 / 2)

From this, genes that appeared upregulated in one comparison and downregulated in the other, or vice versa were then eliminated. This method, coupled with adjusted p-values from a t-test results in a list of genes that are significantly up- or downregulated between the two samples. This list was further reduced by restricting it to those genes with a greater than two-fold change.

6.4.4.8 Analysis of gene lists

Final gene lists were manipulated and displayed using Spotfire software. For additional analyses, gene lists were exported and uploaded into PANTHER (panther.celera.com). The PANTHER database was designed for high-throughput analysis of protein sequences. One of the key features is a simplified ontology of protein function, which allowed browsing of the database by biological functions. Biologist curators have associated the ontology terms with groups of protein sequences rather than individual sequences. Statistical models (Hidden Markov Models, or HMMs) are built from each of these groups. The advantage of this approach is that new sequences can be automatically classified as they become available. To ensure accurate functional classification, HMMs are constructed not only for families, but also for functionally distinct subfamilies. Multiple sequence alignments and phylogenetic trees, including curator-assigned information, were available for each family.

6.5 Results

6.5.1 RNA analysis

The quality and quantity of RNA extracted from 20 fresh tumour samples (and one sample of normal mucosa) was examined prior to its inclusion in microarray analysis. Agarose gel and UV spectroscopy results for a selection of samples are shown in Fig 6.2 and Table 6.2.

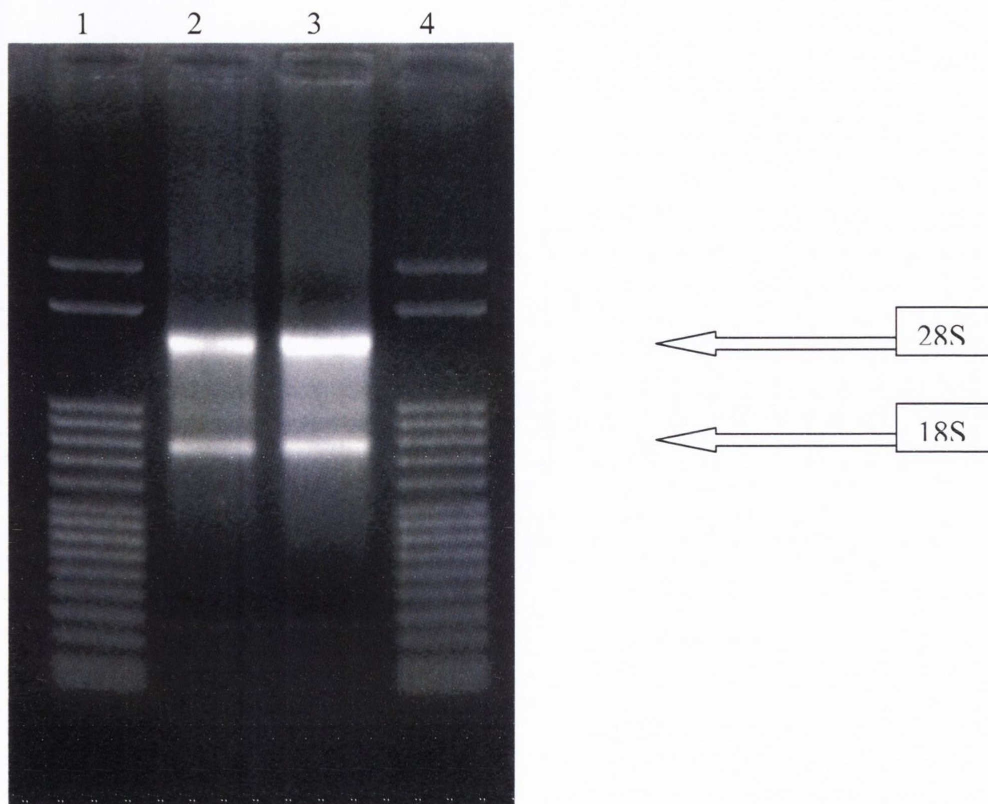


Figure 6.2 Agarose gel electrophoresis of 2 RNA samples

Typical RNA extracts electrophoresed on a 2% agarose gel containing ethidium bromide. Lanes 1 and 4: 50bp DNA ladder (used for reference only and not for sizing purposes), Lane 2: sample #12166 RNA extract, Lane 3: sample #400 RNA extract

Sample Number	RNA Concentration (ng/μl)	A260:280
1	674.8	2.078
2	908.62	2.136
3	793.05	2.186
4	901.33	2.179
5	804.20	2.104
6	225.33	2.177
7	297.61	2.165
8	544.32	2.153
9	566.95	2.1
10	785.67	2.099
11	512.67	2.01
12	1196	1.9
13	282.00	2.0
14	797	2.15
15	1025	2.09
16	504	1.9
17	559.50	2.1
18	519.66	2.1
19	468.44	2.1
20	356.60	2.266
21	1145.70	2.0

Table 6.2 UV spectrophotometric analysis of SCC samples

6.5.2 Post RT-IVT cRNA analysis

cRNA outputs from the RT-IVT reaction were analysed prior to any hybridisation. UV spectroscopy was once again used to determine nucleic acid concentration (Table 6.3) and agarose gel electrophoresis was used to determine the approximate size distribution of the labelled cRNA targets (see Fig 6.3 for example).

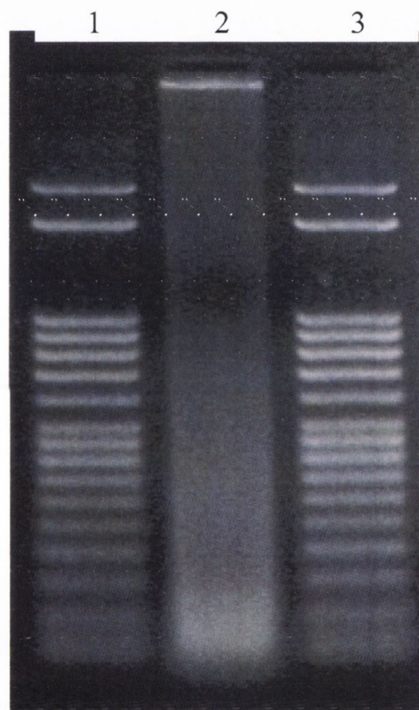


Figure 6.3 Agarose gel electrophoresis of a cRNA sample

A typical cRNA output from the RT-IVT kit electrophoresed on a 2% agarose gel containing ethidium bromide. Lanes 1 and 3: 50bp DNA ladder, lane 2: sample #12166 cRNA.

Sample Name	**CRNA Concentration (ng/μl)**
111124	**566**
2	**512**
3	**220.98**
4	**265.39**
5	**837.37**
6	**224.33**
7	**324**
8	**544.32**
9	**566.95**
10	**220.71**
11	**512.67**
12	**955.29**
13	**255.67**
14	**988.76**
15	**537**
16	**351.52**
17	**451.50**
18	**359.38**
19	**982.52**
20	**316.91**
21	**1021.90**

Table 6.3 UV spectroscopic analysis of RT-IVT outputs

6.5.3 Microarray image capture

Images of processed microarrays were captured and processed using the ABI 1700 reader and software system. Arrays were captured in two halves and an example of one half is displayed in Fig 6.4.

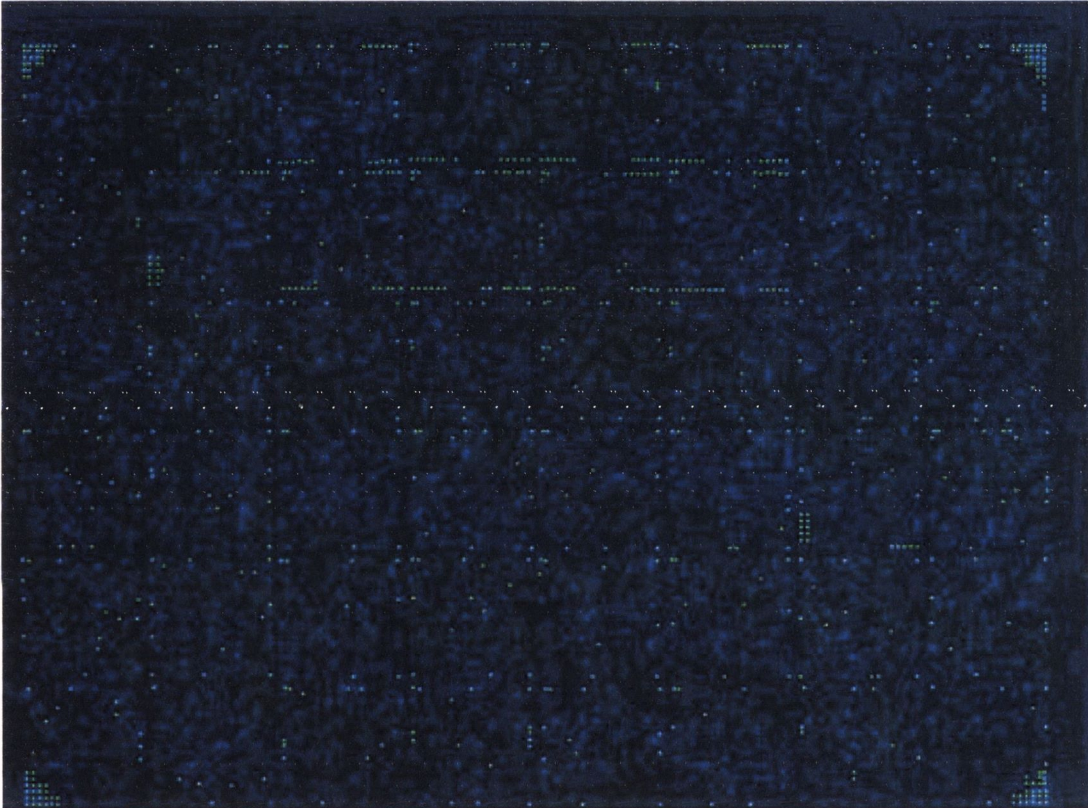


Figure 6.4 Captured image of sample number 12166 microarray

ABI 1700 software captures the full microarray image in two halves. The green spots represent the various array and other controls on half of the microarray grid

6.5.4 Statistical analysis

6.5.4.1 Normalisation

5% trimmed mean method was selected as an appropriate normalisation method, as it is more powerful when analysing low signal intensity genes.

Figure 6.5 shows an MA plot for the technical replicates. Both plots are symmetrical around zero and show no appreciable difference. This is entirely in keeping with what would be expected for technical replicates.

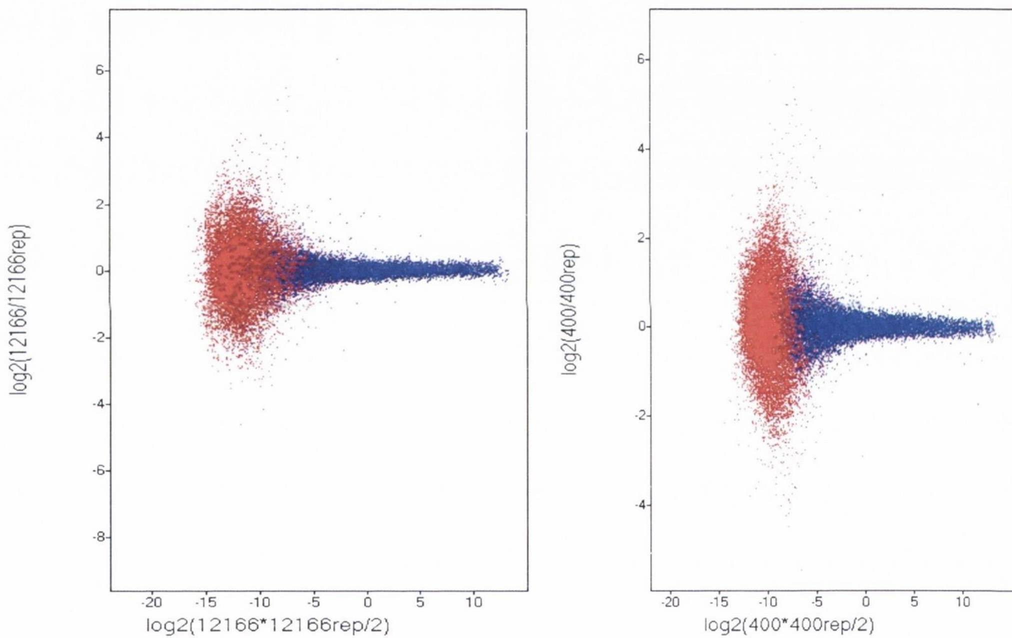


Figure 6.5 MA plots for sample number 12166 (left), sample number 400 (right)

6.5.4.2 Treatment comparisons

Comparing young nonsmokers to old smokers

Initial analysis of the data consisted of using a t-test to compare average normalized signal values for young nonsmokers and old smokers groups, with a p value setting of **0.01**. In addition, only genes showing at least **2-fold expression change** between the two groups were selected. This analysis gives us a relatively small list of genes (N=131) that distinguish between the young nonsmokers and the old smokers. Initial clustering of these 131 genes generates a heat map and cluster dendrogram, which separates the young nonsmokers and the old smokers (Figure 6.6)

Gene profile charts showing those genes significantly up- or down-regulated in one comparison are shown in Figure 6.7. For data reduction purposes, further analysis of this gene group was performed using a p value set at 0.001, generating a list of 37 genes showing increased expression in young nonsmokers relative to old smokers. A further 58 genes showed decreased expression in young nonsmokers (Table 6.5)

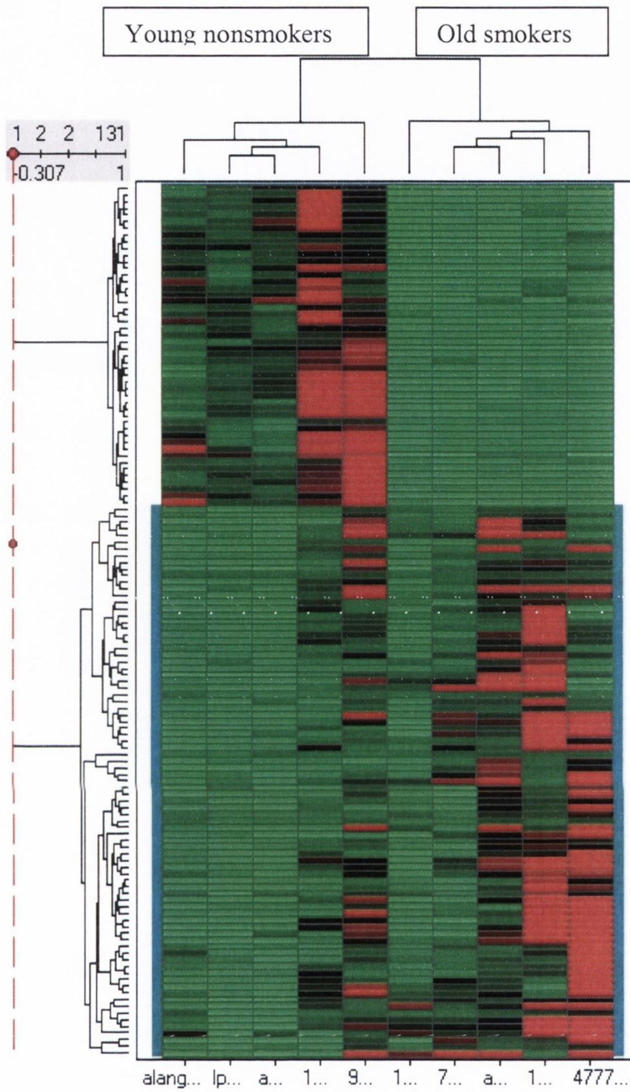


Figure 6.6 To visualise the gene expression data, hierarchical clustering was performed using genes that satisfied stringent filtering criteria ($p<0.01$, >2 fold expression change) yielding a discriminate set of 131 genes.

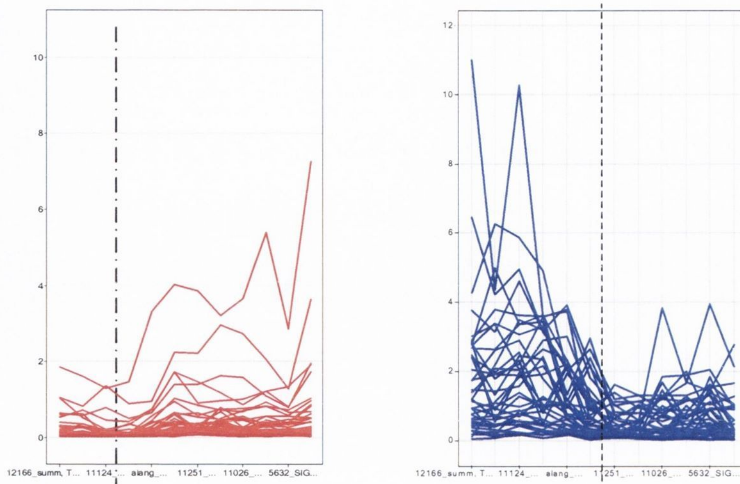


Figure 6.7 displays genes that are significantly up-regulated (blue) and down-regulated (red) in young nonsmokers vs. oldsmokers. The groups being compared are divided by a dashed line in each visualisation.

Gene ID	Symbol	Gene Name
hCG23484	CREM	Camp responsive element modulator
hcCG9791	WFDC1	Wap4 disulphide core domain 1
hCG2039142	FCERIg	Fc fragment of IgE
hCG24573	CH25H	Cholesterol 25 hydroxylase
1642458	MGC12921	Hypothetical protein MGC12921
96142	PRG1	Proteoglycan1 secretory granule
39101	C20orf32	Chromosome 20 open reading frame 32
1981647	GS2	GS2 gene
28240	DOC1	Down regulated in ovarian cancer 1
15615	NK4	Natural killer cell transcript 4
16800	BAG2	BCL2 associated atanogene 2
1749202	CCL3L1	Chemokine (C-C motif) like 3 ligand 1
25217	MAGP2	Microfibril associated glycoprotein 2
41052	RGS2	Regulator of G-Protein signalling 2
1775764	C5R1	Complement component 5 receptor 1
1641555	STAT 4	Signal transducer activator of transcription 4
18349	TAO1	Thousand and one amino acid protein 1
2036728	PCSK5	Protein convertase subtilisin/kexin type 5
2039559	POPDC3	Popeye domain containing 3
20429	FLJ36119	Hypothetical protein
1818318	UGTB11	UDP glycosyltransferase 2 family
hCG1646859	FLJ23420	hypothetical protein FLJ23420
hCG2032251	FLJ25393	hypothetical protein FLJ25393
hCG20429	FLJ36119	hypothetical protein FLJ36119
hCG1642458	MGC12921	hypothetical protein MGC12921
hCG1981647	DXS1283E	GS2 gene

Table 6.4 Genes with increased expression in young nonsmokers relative to old smokers after t test (P<0.001). Average fold change >2.

GENE ID	Symbol	Gene name
hCG2003601	C6orf197	Chromosome 6 open reading frame 197
hCG37929	ZFYVE16	Zinc finger FYVE containing domain 16
hCG23604	TA-NFKB	T cell activation NFkappaB-like protein
hCG1643836	CDC6	CDC6 cell division cycle 6
hCG40094	NY-REN-58	NY-REN-58
hCG15700	SSH1	Slingshot 1
hCG21840	DRIM	Down regulated in metastasis
hCG41324	TRRAP	Transformation/transcriptiondomain associated protein
hCG20768	GPHN	gephyrin
hCG40780	MCM10	Minichromosome maintenance deficient 10
hCG23744	MPHOSPH1	M phase phosphoprotein 1
hCG2039640	C10orf6	Chromosome10 open reading frame 6
hCG1984462	RBAF600	Retinoblastoma associated factor 600
hCG1811093	NRF1	Nuclear respiratory factor 1
hCG29362	CDK5RAP3	CDK5 regulatory subunit associated factor 3
hCG1685939	IL12A	Interleukin12A
hCG25641	PPARG-CR1	PPAR gamma coactivator related 1
hCG20423	CCNL2	CyclinL2
hCG25242	MASTL	Microtubule associated serine/threonine kinase-like
hCG1640632	TPD52	Tumour protein D52
hCG20383	MTA2	Metastasis associated like-1
hCG25998	TCF19	Transcription factor 19
hCG25264	H2AFY2	H2A histone family Y2
hCG16355	ZFP276	Zinc finger protein 276
hCG24011	MGC9718	Zinc finger protein 9718
hCG1985219	SHMT1	Serine hydroxymethyltransferase 1
hCG2032998	RPL7A	Ribosomal protein L7A
hCG32574	KIAA 1287	KIAA 1287 protein
hCG1998970	KIAA 1641	KIAA 1641 protein
hCG17719	ANLN	anillin, actin binding protein
hCG27532	atherin	atherin;LOC90378
hCG2024897	CAP350	centrosome-associated protein 350
hCG22023	C14orf160	chromosome 14 open reading frame 160
hCG1785860	FUT10	fucosyltransferase 10(alpha (1,3) fucosyltransferase)
hCG18322	INPP5E	inositol polyphosphate-5-phosphatase, 72 kDa

Table 6.5 Decreased expression in young nonsmokers vs old smokers (p<0.001)

6.5.4.3 Distinction Analysis

Figure 6.8 shows a profile chart of 200 genes generated after using the Distinction Analysis method. Using the distinction method a larger group of genes may be identified as potential biomarkers. Young nonsmokers lie to the left. The blue illustrates the genes that are upregulated in young nonsmokers relative to old smokers and the red illustrates those genes that are downregulated in young nonsmokers relative to the old smokers. The diagram illustrates the concept of distinction calculation, which relies on an analysis of how different parts of a profile appear to be.

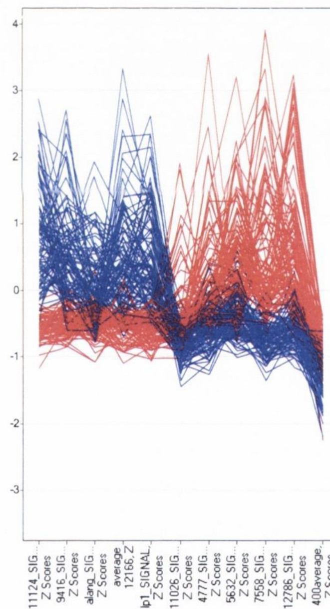


Figure 6.8 Profile chart of genes showing increased and decreased expression in young nonsmokers and old smokers with clear distinction patterns

Figure 6.9 shows a pie chart documenting the cellular processes upregulated in young nonsmokers after a distinction calculation has been applied (Panther

classification). WNT signalling pathway and JAK-STAT signalling pathways are the most frequently upregulated processes in young nonsmokers relative to old smokers.

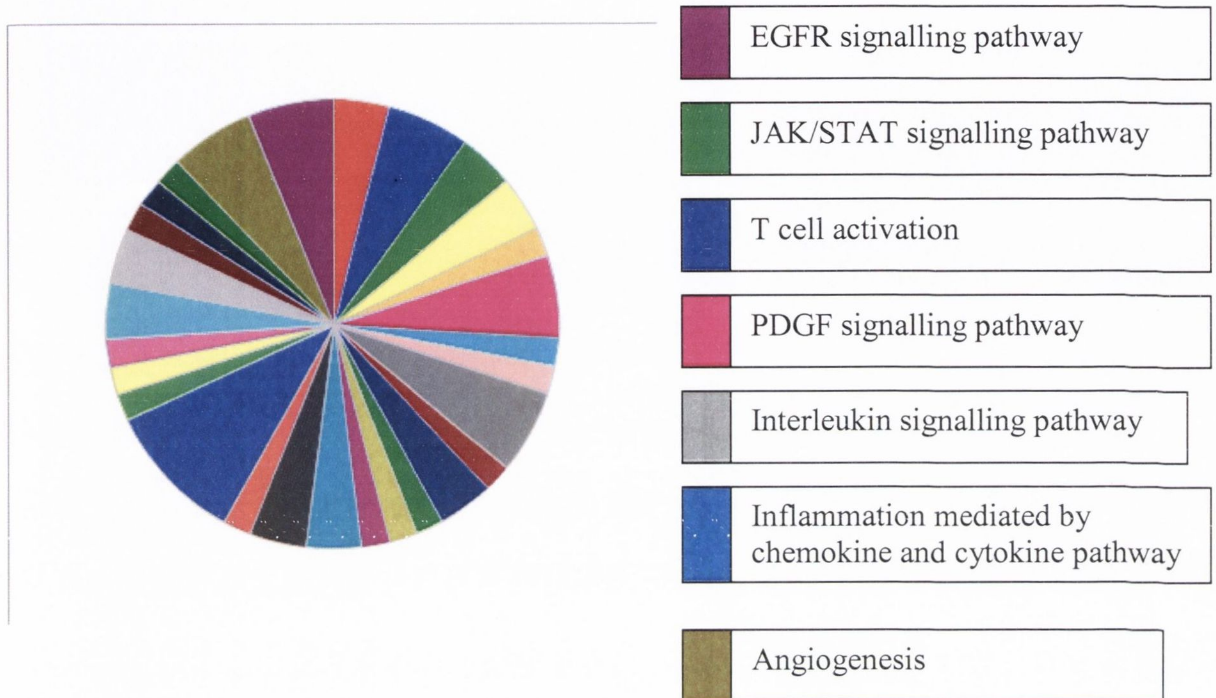


Figure 6.9 Panther Classification of transcripts

6.5.4.4 Smokers versus nonsmokers

The results of this treatment comparison can be found in Tables 6.6 and 6.7, which show the genes that were upregulated and downregulated in nonsmokers relative to smokers, using a p-value <0.001. Of note is the significant overexpression of CDKN2A in nonsmokers compared with smokers.

Gene I.D	Symbol	Name
hCG20343	**ZNF 180**	**Zinc finger protein 180**
hCG42702	**RPL3L**	**Ribosomal protein L3-like**
hCG29297	**CCL13**	**Chemokine (C-C motif) ligand 13**
hCG28310	**CDKN2B**	**Cyclin dependant kinase inhibitor 2B**
hCG18762	**CYP39A1**	**Cytochrome p450, family39, subfamilyA, polypeptide1**
hCG40728	**NCF2**	**Neutrophil cytosolic factor 2**
hCG28240	**DOC1**	**Downregulated in ovarian cancer 1**
hCG1749202	**CCL3L1**	**Chemokine (C-C motif) ligand 3L1**
hCG15424	**MYL3**	**Myosin, ligt polypeptide 3**
hCG27314	**OSF1**	**Osteoclast stimulating factor 1**
hCG27669	**CCL18**	**Chemokine (C-Cmotif) ligand 18**
hCG22330	**IF127**	**Interferon *a*-inducible protein 27**

Table 6.6 Genes showing increased expression in nonsmokers with oral and oropharyngeal cancer relative to smokers with oropharyngeal cancer after t-test (p<0.001). Average fold change >2

Gene I.D	Symbol	Name
hCG27700	**SH-y**	**SMCY homolog y-linked**
hCG40094	**NY-REN-58**	**NY-REN-58 antigen**
hCG1987333	**Cyorf15B**	**Chromosome y orf 15B**
hCG33025	**PCSK9**	**Proprotein convertase subtilisin**

Table 6.7 Genes showing decreased expression in nonsmokers with oropharyngeal cancer relative to smokers with oropharyngeal cancer after t-test (p<0.001). Average fold change >2

6.6 Discussion

There are a number of studies published on head and neck cancer and expression microarrays, with several groups focusing on correlation between gene expression signatures and metastatic potential (Ginos, 2004, Vigneswaren, 2005, Roepman 2005). Others have preferred to investigate the transcriptome of cancer in the head and neck and compare it to normal tissue at this site (Leethanakul, 2004, Hwang, 2003, Lemaire, 2003).

The majority of patients with oral and oropharyngeal cancer are over 50 years of age and have prolonged exposure to tobacco and alcohol. However, there is a subset of patients with oral and oropharyngeal that are under 40 years of age with no apparent risk factors. Results from Chapter 4 (Microarray CGH) have demonstrated that there is a significant difference between the genomic profile of the typical oral cancer patient (over 50 year old male smoker) and this emerging subset of young nonsmokers. To identify any impact these genomic differences have on gene expression patterns, the transcriptome of head and neck cancer in both these cohorts was then analysed.

The aim of this study was to identify correlations between gene expression signatures and two parameters (Age/Smoking status) with the goal of identifying a group of genes that distinguish between the type of cancer occurring in young nonsmokers and that occurring in old smokers.

Using statistical and data filtering criteria, 131 genes differentially expressed between oral and oropharyngeal cancer in young nonsmokers and old smokers were identified. Selecting differentially expressed genes between two types of samples or

under two experimental conditions is the most commonly used approach in statistical analysis. The simplest most straightforward, and hence, widely implemented ones are the classic paired or grouped t-tests, which were employed here.

The first gene list generated from the analysis was from comparing young nonsmokers with old smokers. Overall analysis identified relatively few genes that distinguished between young nonsmokers and old smokers. This is not surprising considering the major similarities between the two tumour cohorts. While they differ in age and smoking status, they do however have similar histological and clinical appearances. Powell (2003) identified a limited number of genes that differed between lung tumours in smokers and nonsmokers, but found that by comparing normal tissue to tumour tissue in these cohorts, far more genes differed.

In this comparison, a number of genes of interest have been identified. Many of the genes identified many of which are entirely novel in the context of head and neck neoplasia. This is to be expected since there are no published studies focusing on the genes that distinguish between head and neck cancer in young nonsmokers and old smokers. However, it is reassuring to identify a number of genes previously cited in the literature as being relevant in cancer, namely DOC-1 (Rimon, 2004; Mok, 1994), STAT4 (Callo, 2003), MCM and Cdc6 (Murphy, 2005; Robles 2002). These will be discussed in detail in Chapter 7. The Panther system allowed for analysis of functional correlate of the genes detected. Figure 6.9 shows that 'inflammation mediated by cytokines and chemokines' and 'Jak-Stat' pathways were the Panther families containing the highest number of upregulated genes in the young nonsmokers. From this data, it is tempting to suggest that the increased expression

of cytokines and Stat proteins may be responsible for the occurrence of cancer in young nonsmokers. Of course without further validation, these interesting findings are of uncertain significance. We will address these results again in Chapter 7.

Comparing all the nonsmokers with all the smokers yielded a very interesting finding, insofar as CDKN2B (p15) was overexpressed in nonsmokers. Considering the association of its homologue (p16) with head and neck cancer, this finding warrants further exploration.

6.7 Conclusion

In this chapter, candidate markers distinguishing between oral and oropharyngeal cancer occurring in young nonsmokers and old smokers have been identified. It must be noted, however, that despite possible differences in the aetiology of the disease, the carcinogenesis process itself is remarkably similar at the expression levels of all tumour samples in this study. Perhaps important genetic changes occur post-transcriptionally. From this study and from the microarray CGH study (Chapter 4), it seems likely that further genome studies or post-transcriptional analysis may yield more fruitful findings than expression studies. Also, a larger sample size would be required to achieve a larger discriminatory database.

6.8 References

Al Moustafa AE, Alaoui-Jamali MA, Batist G, Hernandez-Perez M, Serruya C, Alpert L, Black MJ, Sladek R, Foulkes WD. Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. Oncogene. 2002; 21(17): 2634-40.

Alevizos I, Mahadevappa M, Zhang X, Ohyama H, Kohno Y, Posner M, Gallagher GT, Varvares M, Cohen D, Kim D, Kent R, Donoff R, Todd R, Yung C, Warrington J, Wong D. Oral cancer in vivo gene expression profiling assisted by laser capture microdissection and microarray analysis. Oncogene. 2001; 20(43): 6196-204.

Alizadeh A, Eisen M, Davis R, Ma C, Lossos I, Rosenwald A, Boldrick J, Sabet H, Tran T, Yu X, Powell J, Yang L, Marti G, Moore T, Hudson J Jr, Lu L, Lewis D, Tibshirani R, Sherlock G, Chan W, Greiner T, Weisenburger D, Armitage J, Warnke R, Levy R, Wilson W, Grever M, Byrd J, Botstein D, Brown P, Staudt L. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000; 403(6769): 503-11.

Belbin TJ, Singh B, Barber I, Socci N, Wenig B, Smith R, Prystowsky M, Childs G Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays. Cancer Research. 2002; 62(4): 1184-90.

Bonferroni CE. 1936. Teoria statistica delle classi e calcalo delle probabilitá. Pubblicazioni del R Instituto Superiore di Scienze Economiche e Commerciali di Firenze, 8:3-62

Califano J, van d Westra W, Nawroz H, et al. Genetic progression model for head and neck cancer: implications for field cancerisation. Cancer Research. 1996; 56: 2488–2492.

Calo V, Migliavacca M, Bazan V, Macaluso M, Buscemi M, Gebbia N, Russo A. STAT proteins: from normal control of cellular events to tumorigenesis. Journal of Cell Physiology. 2003; 197(2): 157-68

Eberwine J. Amplification of mRNA populations using a RNA generated from immobilized oligo (dT)-T7 primed cDNA. Biotechniques 1996; 20(4); 584-591

El-Naggar AK, Kim HW, Clayman GL, Coombes MM, Le B, Lai S, Zhan F, Luna MA, Hong WK, Lee JJ. Differential expression profiling of head and neck squamous carcinoma: significance in their phenotypic and biological classification. Oncogene. 2002; 21(53): 8206-19.

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286(5439): 531-7

Hwang D, Alevizos I, Schmitt W, Misra J, Ohyama H, Todd R, Mahahdevappa M, Warrington J, Stephanopoulos G, Wong D, Stephanopoulos G. Genomic dissection for characterisation of cancerous oral epithelium using transcription profiling. Oral Oncology 2003; 39: 259-268

Leethanakul C, Knezevic V, Patel V, Amornphimoltham P, Gillespie J, Shillitoe E, Emko P, Park M, Emmert-Buck M, Strausberg R, Krizman D, Gutkind J. Gene discovery in oral squamous cell carcinoma through the Head and Neck Cancer Genome Anatomy Project. Oral Oncology. 2003; 39(3): 248-58.

Leung YF, Cavalieri D. 2003. Fundamentals of cDNA microarray analysis. Trends in Genetics 19(11): 649-659

Mendez E, Cheng C, Farwell DG, Ricks S, Agoff SN, Futran ND, Weymuller EA Jr, Maronian NC, Zhao LP, Chen C. Transcriptional expression profiles of oral squamous cell carcinomas. Cancer. 2002; 95(7): 1482-94.

Mok SC, Wong KK, Chan RK, Lau CC, Tsao SW, Knapp RC, Berkowitz RS. Molecular cloning of differentially expressed genes in human epithelial ovarian cancer. Gynaecologic Oncology. 1994; 52(2): 247-52

Murphy N, Ring M, Heffron CC, Martin CM, McGuinness E, Sheils O, O'Leary JJ Quantitation of CDC6 and MCM5 mRNA in cervical intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix. Mod Pathol. 2005

Oh Y, Mao L. Biomarkers in head and neck carcinoma review. Current Opinion in Oncology. 1997; 9: 247–225.

Rimon E, Sasson R, Dantes A, Land-Bracha A, Amsterdam A. Gonadotropin-induced gene regulation in human granulose cells obtained from IVF patients: Modulation of genes coding for growth factors and their receptors and genes involved in cancer and other diseases. International Journal of Oncology 2004; 24: 1325-1338

Robles LD, Frost AR, Davila M, Hutson AD, Grizzle WE, Chakrabarti R. Down-regulation of Cdc6, a cell cycle regulatory gene, in prostate cancer. Journal of Biological Chemistry. 2002; 277(28), 25431-25438

Roepman P, Wessels LF, Kettelarij N, Kemmeren P, Miles AJ, Lijnzaad P, Tilanus MG, Koole R, Hordijk GJ, van der Vliet PC, Reinders MJ, Slootweg PJ, Holstege FC. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. Nat Genet. 2005; 37(2): 182-6.

Schena M, Shalon D, Davis RW, Brown PO, 1995, 'Quantitative monitoring of gene expression patterns with a complimentary DNA microarray'. Science; 270: 467-70.

Vigneswaran N, Wu J, Sacks P, Gilcrease M, Zacharias W. Microarray gene expression profiling of cell lines from primary and metastatic tongue squamous cell carcinoma: possible insights from emerging technology. J Oral Pathol Med. 2005; 34(2): 77-86.

Chapter 7

Validation of gene expression array data

7.1 Summary

The expression signatures detected in Chapter 6 suggest that several specific genes may be differentially regulated across subsets of oral and oropharyngeal cancers. And that these genes could be used to help identify the cause of tumourigenesis in young nonsmokers. Using labelled cDNA from 21 tumour samples, a novel chemiluminescence expression array system was used to study the expression of 31,600 genes in each sample. A number of genes were identified that distinguished between the oral/oropharyngeal cancer samples in young nonsmokers and the samples from old smokers. The identification of altered gene expression profiles with microarrays is best viewed as the first step in the determination of potential disease-associated genes; however, the false-positive rate can be high, particularly with small sample sets. Quantitative PCR is a rapid and highly sensitive technique for accurate quantification of microarray results and is now considered the "technique of choice" for validating gene expression changes identified with microarray expression profiling technologies.

In order to validate the expression array findings in Chapter 6, cDNA was generated from the same 21 RNA samples used in Chapter 6 and real-time PCR analysis was performed for 12 genes and compared to the results from the expression arrays study. Quantitative PCR confirmed the direction of fold change for nearly all the samples.

7.2 Introduction

In the fast-paced world of genomic research, the process of experimentally validating microarray data is often regarded as a side issue in many laboratories. Indeed, unlike many early papers in the field that typically included a series of validation experiments, there are now an increasing number of microarray-related papers in the literature have apparently forgone performing such 'wet-bench' validations altogether. In one respect, this could be viewed as an indication that the microarray platform has successfully evolved to a new state of accepted technological maturity with regards to reliability and consistency. There is some truth to this, particularly with the increasing availability of inexpensive commercially fabricated microarrays, which are typically of high quality and display minimal chip-to-chip variance. Nevertheless, when dealing with novel platforms such as that used in this study, it is still important to establish the 'robustness' and scientific validity of a particular microarray finding.

Perhaps the most significant reason as to why experimental validations remain essential is that, depending on the particular technological platform used, microarray data is inherently noisy. The large numbers of gene expression measurements obtained in a typical microarray experiment, by virtue of their sheer numbers, can often yield significant numbers of false positive and negative results. Microarray technology is still very much an evolving field, and as such, it is essential to have the resources to perform at least a measure of independent validation of the microarray result.

After hybridisation and subsequent primary analysis of the expression microarrays, the initial global data set generated was then subjected to various forms of statistical and computational analysis (discussed in Chapter 6). The end result of this analysis was the identification of a relatively small number of genes that exhibited a particular behaviour of interest, such as being differentially regulated across a particular class distinction (e.g. young nonsmokers vs old smokers). This target set of 131 genes was considered too large to allow experimental validation of every single one. Therefore for initial validation, 12 genes of interest i.e genes with potential relevance to head and neck cancer were selected.

7.3 Aims

As can be appreciated from the introductory comments above, there is a clear need to validate at least some of the potential targets resulting from the microarray experiments before discussing their significance. This chapter aims to validate twelve genes of interest, selected from the target list generated from the microarray study.

7.4 Material and Methods

7.4.1 Specimens

All 21 samples from Chapter 6 were used in the validation study. The same 21 RNA samples used in the microarray study were employed in this validation study.
Of note, DNA from 20 of these 21 samples was used in the microarray CGH study (Chapter 4).

7.4.2 Assays

The PANTHER system workspace (Chapter 6) is integrated with the Applied Biosystems e-commerce site. Twelve corresponding Applied Biosystems TaqMan® Gene Expression assays corresponding to the gene I.Ds of the twelve selected genes were ordered from this site. Nine genes were selected from the young non-smoker/old smoker list and one gene was selected from the smoker/non-smoker list. Two further genes were selected because of their documented relevance to head and neck cancer. The TaqMan® Gene Expression Assays are provided in a single-tube, 20X format. Table 7.1 lists the probe sequences.

Assay I.D	Gene symbol	Probe sequences
Hs00231372_m1	STAT 4	AGGAAATGAGGGCTGTCACATGGTG
Hs00168405_m1	IL-12	CATGCTCCAGAAGGCCAGACAAACT
Hs00218560_m1	MCM10	CTGAAAGAGGCCGAGGCTGAGATGC
Hs00154374_m1	CDC6	CAGGAAACGTCTGGGCGATGACAAC
Hs00233365_m1	CDKN2B	GGCGCGCGATCCAGGTCATGATGAT
Hs00706279_s1	DOC-1	ACAGAAATCAGTGCCAAGCATGCGA
Hs00262018_m1	TA-NFKBH	GGGCCCTGAGGGGCTCCGGCAGCTG
Hs00602528_g1	CDK5RAP3	GGAAAGATGGAGGACCATCAGCACG
Hs00268883_m1	TRRAP	GAGAGACTCTAAGGCCCCTCGCCTA
Hs00378210_m1	RBAF600	AGCCACCAGGCTGACAGATAAGGCA
Hs00233365_m1	CDKN2A	GTCCCTCAGACATCCCCGATTGAAA
Hs00153108_m1	RB1	CTGCCCTGGGCGAGATCACATGGTC

Table 7.1 Twelve genes selected for the validation study.

7.4.3 cDNA generation and TaqMan PCR

The High-Capacity cDNA Archive Kit (Roche, Nutley, NJ) was used for reverse transcription (RT) of total RNA to single-stranded cDNA (Chapter 2). Generally 5ug of total RNA was used. Twelve genes were selected from the target lists generated in Chapter 6 (Figure 7.1). A TaqMan® PCR-based system was selected for mRNA quantitation. For the real-time quantitative PCR reaction, the Applied Biosystems 7900 detection system was used, with mastermix (Applied Biosystems, Foster City, CA) and the selected primer/probe mix (as above). **TaqMan PCR is described in detail in Chapter 2.** The integrity of the extracted RNA was confirmed by demonstrating amplification of GAPDH housekeeping gene. PCR products were detected in real time.

7.4.4 Statistical Methods

7.4.4.1 Comparing expression array and TaqMan PCR

A sample of a known concentration can be used to construct a standard curve. By running standards of varying concentrations, a standard curve is created from which the quantity of an unknown sample can be extrapolated. The standards used depend on whether absolute or relative quantitation is to be used. Relative quantitation is chosen as the method of quantitation in all the validation experiments described in this chapter.

It is easy to prepare standard curves for relative quantitation as quantity is expressed relative to some basis sample, such as the calibrator. For all experimental samples, target quantity is determined from the standard curve and divided by the target quantity of the calibrator. Thus, the calibrator becomes the 1X sample, and all other quantities are expressed as an n-fold difference relative to the calibrator.

For quantitation normalized to an endogenous control, standard curves are prepared for both the target and the endogenous reference. For each experimental sample, the amount of target and endogenous reference is determined from the appropriate standard curve. Then, the target amount is divided by the endogenous reference amount to obtain a normalized target value. Again, one of the experimental samples is the calibrator, or 1X sample. Each of the normalized target values is divided by the calibrator normalized target value to generate the relative expression levels. For the quantitation of gene expression, researchers have used β-actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ribosomal RNA (rRNA), or other RNAs as an endogenous control.

To prepare the TaqMan PCR results for comparison with the expression microarray results, a relative quantitation method was used. For each assay, the Ct for each sample was divided by the GAPDH for each sample, and a calibrator (in this case normal tissue sample) was subtracted from this. This gave us a ΔΔCt result for each sample of each assay.

Expression array: The normalised signal intensities of each sample for each of the twelve genes was divided by the calibrator (again the normal tissue sample was used as the calibrator). For each sample, the ΔΔCt and the signal/calibrator were plotted against each other giving us plots for each assay.

7.4.4.2 Comparing CGH array and expression array

Genomic aberrations do not always manifest as transcriptome changes. A comparison between genomic changes from Chapter 4 and the transcriptome changes from Chapter 6 was performed in order to identify which genomic aberrations were functional. High-level amplifications in well-known genes were selected and compared directly with signal intensities from the expression array study.

7.5 Results

To confirm the expression detected by the microarrays, quantitative real-time PCR analysis comparing fold change between young nonsmokers and old smokers was conducted. Selected assays were ordered from the Applied Biosystems TaqMan® Gene Expression Assays. For microarrays, gene expression fold change was calculated as a ratio of tumor sample to normal tissue after 5% trimmed mean normalisation. For TaqMan data, fold change was calculated using relative quantitation, or Ct method, where $\Delta\Delta Ct = \Delta Ct$ tumor- ΔCt normal (ΔCt is GAPDH normalized Ct). Fold changes seen by TaqMan assays and by the Expression Array System are similar, with RT-PCR results confirming the direction of fold change in nearly all the samples. Most of the plots showed a similar trend between TaqMan data and expression array data, so that when microarray values increased, they were matched by an increase in the correlating TaqMan value.

7.5.1 Young nonsmokers vs old smoker

Nine genes were selected from the target list generated from the microarray study by comparing young nonsmokers with old smokers. These genes were selected on the basis of potential scientific relevance. The following figures illustrate the correlation between the expression array data and the TaqMan PCR data in old smokers and young nonsmokers (Figure 7.1, 7.2, 7.3). The plots for CDK5RAP IL-12 and TA-NFKBH illustrate some contradictory results, with one young nonsmoker (number 11124) exhibiting high levels of expression with the TaqMan result, but relatively low levels with the array data. Assays for p16 and Rb1 were also selected considering the association of both these genes with head and neck cancer.

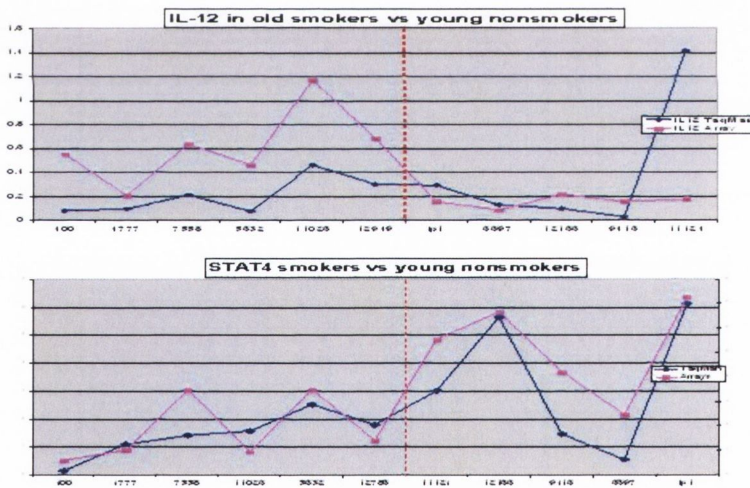


Figure 7.1 Expression levels of IL-12 and STAT4 in old smokers (left of red line) vs young nonsmokers (right of red line).

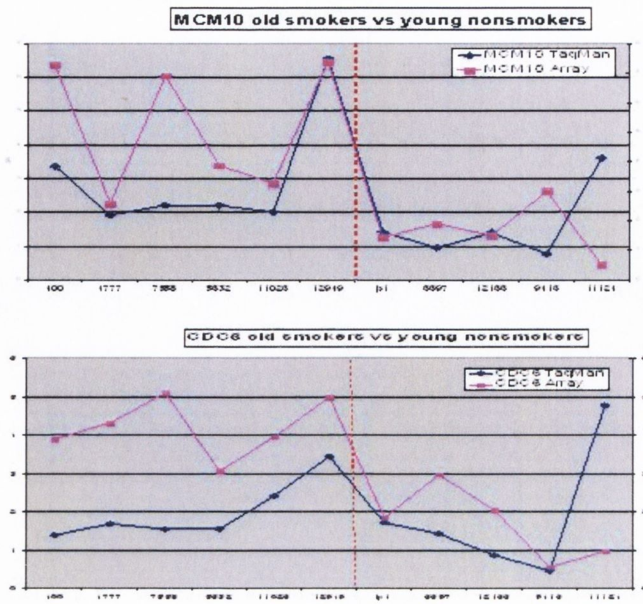


Figure 7.2 Expression levels of MCM10 and cdc6 in old smokers (left of red line) vs young nonsmokers (right of red line).

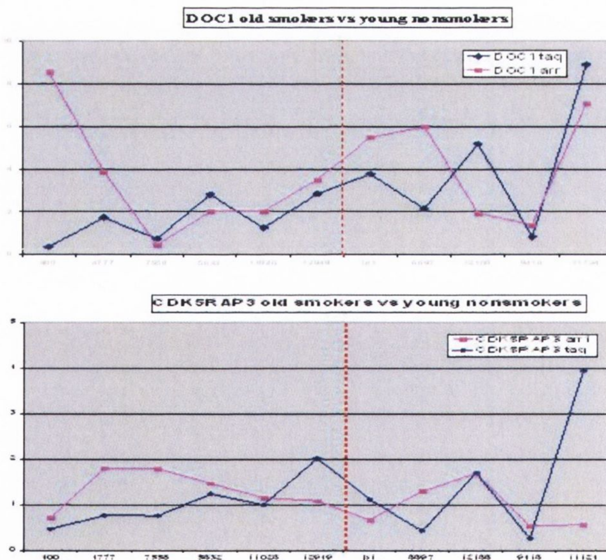


Figure 7.3 DOC-1 and CDK5RAP in old smokers (left of red line) vs young nonsmokers (right of red line)

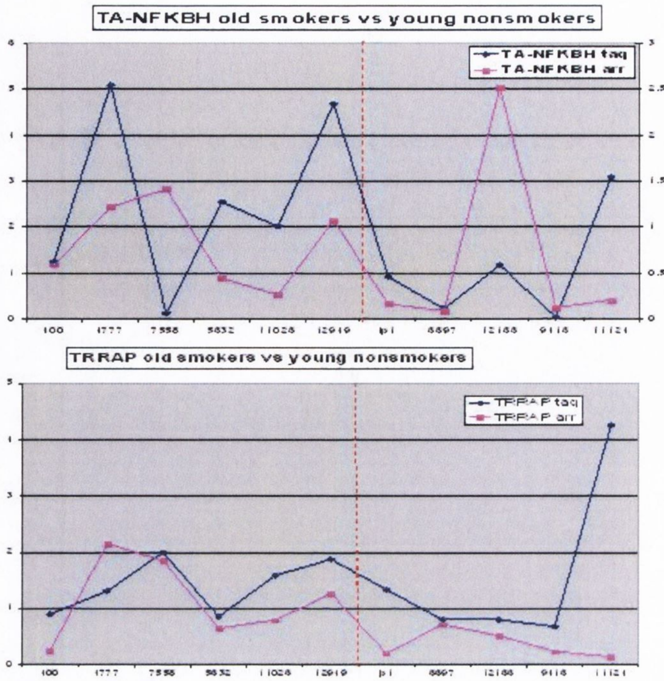


Figure 7.4 TA-NFKBH and TRRAP in old smokers (left of red line) vs young nonsmokers (right of red line)

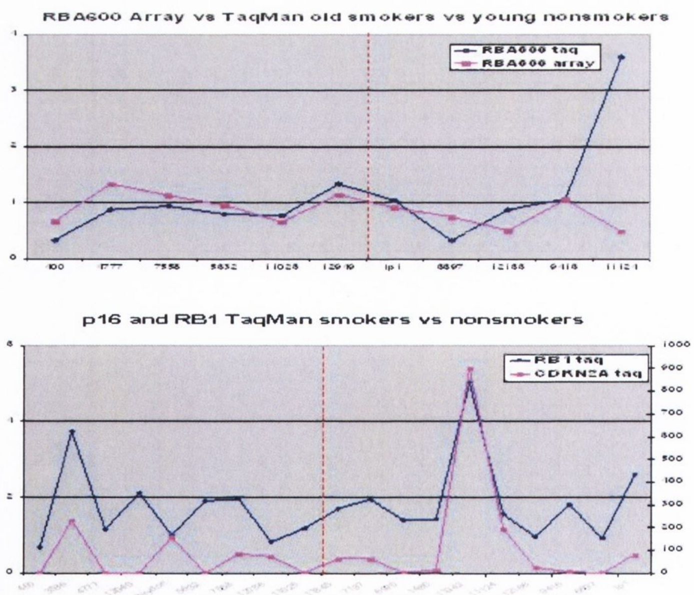


Figure 7.5 Expression levels of RBA600, p16 and RB1. For RBA600 old smokers are left of red line, young nonsmokers are right of red line. For p16 and RB1 smokers are left of red line and nonsmokers are right of red line

7.5.2 Smokers versus nonsmokers

Irrespective of age, the expression microarray profiles of smokers and nonsmokers were compared, and a number of target genes were identified. Not surprisingly, some overlap was observed between the young nonsmoker/old smoker list and the nonsmoker/smoker list. A gene that found to be significantly overexpressed in nonsmokers compared with smokers was CDKN2B (p15). Considering the association between its homologue p16 and head and neck cancer, p15 was selected for validation (Figure 7.6).

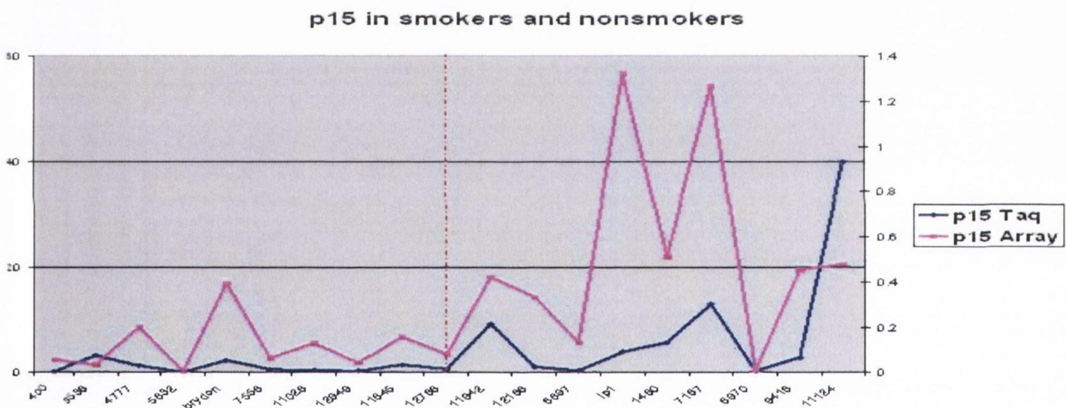


Figure 7.6 Expression levels of p15 in all samples. Smokers left of the red line, nonsmokers right of the red line.

7.5.3 CGH array and expression array data

While it was noted that genomic aberrations did not necessarily manifest as transcriptome changes, it was reassuring to find that several well known genes with

high levels copy gain from the CGH microarray experiment also had high levels of expression. The results illustrate a considerable influence of copy number on gene expression patterns. Overall high levels of overexpression were found in genes that demonstrated high levels of copy gain (i.e. >2.5). Figure 7.7 illustrates the patterns of copy gain and overexpression in CCND1 and EGFR. Some low-level copy number gains and losses also had a significant influence on expression levels of genes in the regions affected, but these effects were more subtle on a gene-by-gene basis than those of high-level amplifications.

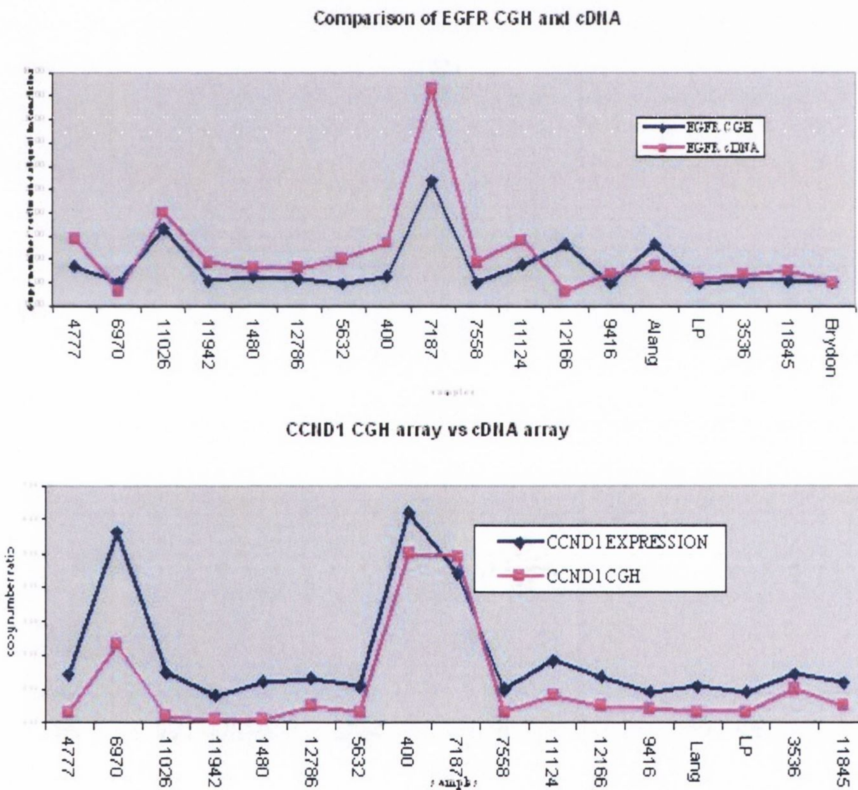


Figure 7.7 Comparison of CGH and cDNA arrays for genes with high level gains

7.6 Discussion

Gene expression profiling at the RNA level has been significantly facilitated by microarray analysis and quantitative real-time RT-PCR. Microarray analysis offers the advantage of profiling expression levels of hundreds or thousands of genes simultaneously using a single RNA preparation. Although real-time PCR analysis provides precise quantification over a wider dynamic range of expression levels, it is not suited for simultaneous analysis of large numbers of genes. Therefore, array analysis is often used as a tool to screen for target genes that are differentially expressed between biological samples, and quantitative real-time RT-PCR then provides precise quantification and validation of the microarray results.

When dealing with novel microarray platforms such as that used in this study, it is important to establish the 'robustness' and scientific validity of a particular microarray finding. Quantitative PCR is a rapid and highly sensitive technique for accurate quantification of microarray results and is now considered the "technique of choice" for validating gene expression changes identified with microarray expression profiling technologies. Quantitative real-time PCR (Q-PCR) with TaqMan® probes capitalizes on the fact that there is a quantitative relationship between the amount of starting target sample and the amount of PCR product at any given PCR cycle number. The Applied Biosystems real-time quantitative PCR solution, using ABI PRISM® Sequence Detection Systems and 5' nuclease chemistry with TaqMan® probes exploits the exonuclease activity of AmpliTaq® DNA polymerase by using a cleavable fluorescent probe in combination with forward and reverse PCR primers. The 5' nuclease assay is sensitive (ability to detect

1 copy in 10-100 cells), reproducible (reliably low CVs), and has a very large dynamic range.

The aim of this chapter was to validate a number of the targets found to be differentially expressed between various cohorts. Eight genes were selected for validation. Seven of these genes were from the target list generated by comparing young nonsmokers with old smokers, namely STAT4, IL-12, MCM10, CDC6 and DOC-1. One gene was selected from the smokers versus nonsmokers list (CDKN2B).

In almost every plot, there were some contradicting values, where the microarray point might be low, but the TaqMan point was high, or vice versa. This variability could be explained by a number of factors, including array-to-array variability, statistical analysis of microarray data, inherent false positive and false negative rates arising from laboratory processes, and general limits of microarray technology which can lead to results that fail validation by downstream techniques. The large number of genes analysed on a microarray statistically leads to random fluctuations, producing some misleading results. The large quantity of data generated from a single experiment requires statistical analysis to determine which transcripts are producing positive signals. However, different software applications, or even the same software application programmed with different parameters will produce different sets of results. General laboratory processes contribute to the need to validate microarray results. Artifacts can be introduced in both the RNA extraction steps, pre-amplification of RNA, and the labeling reactions. Hybridisation of labelled cDNA to microarrays is, like many experiments, a technique dependent

process, meaning that different researchers may carry out the experiments differently, leading to varying results.

7.6.1 DOC-1

The expression microarray study identified increased expression of DOC-1 in young nonsmokers. The nomenclature here is confusing, with the name DOC-1 representing two genes - down regulated in ovarian cancer 1 (DOC-1) and deleted in oral cancer (DOC-1). The ABI 1700 platform is however meticulous in its gene identification system and the DOC-1 in this study is downregulated in ovarian cancer (Gene I.D 28240)

DOC-1 is a protein whose gene transcript is present in normal surface epithelial cells of the ovary, but consistently downregulated in ovarian cancer (Mok, 1994). Rimon *et al* observed the downregulation of DOC-1 in cells that were stimulated with gonadotropins, and suggested that gonadotropin stimulation may modulate DOC-1 downregulation thus elevating the risk for specific cancers (Rimon, 2004). Schwarze found that DOC-1 was up regulated in senescent cell lines, but downregulated in tissue from prostate cancer (Schwarze, 2002).

Considering that 80% of the young nonsmoking cohort in this microarray study are female, it interesting that all but one of these young female nonsmokers have significant overexpression of DOC-1 compared to the old smokers. DOC-1 is thought to be a tumour suppressor gene and it may be that DOC-1 functions like p16, by controlling the activity of another gene. DOC-1 overexpression could be as a result of inactivation of another gene leading to disruption of a negative feedback

mechanism. Of course DOC-1 could also be mutated in these cases resulting in its overexpression.

7.6.2 IL-12 and STAT4

The role of IL-12 in cellular immunity is largely mediated by the STAT-4 transcription factor. As illustrated in Figure 7.8, Interleukin-12 activates the transcription factor signal transducer and activator of transcription 4 (STAT4).

It is now known that many cytokines, hormones and growth factors utilise STAT signaling pathways to control a remarkable variety of biological responses, including development, differentiation, cell proliferation, and survival. Given the roles of STAT proteins in these fundamental cellular processes, which are often altered in cancer, there is accumulating evidence defining a critical role for STAT proteins in oncogenesis (Burdelya, 2002). Members of this relatively small family of proteins serve as both transducers of cytoplasmic signals and nuclear transcription factors, thereby directly converting a stimulus at the cell surface to an alteration in the genetic program. Moreover, STAT proteins can cross-talk with other central signaling pathways, such as the mitogen-activated protein kinase (MAPK) family of proteins and the nuclear receptor signaling pathways. On the other hand, other proteins are known to interact with and modulate STAT signaling pathways, including histone acetyl-transferases.

In this study, STAT4 was found to be significantly up regulated in young nonsmokers compared with old smokers. Paradoxically, IL-12 was significantly

downregulated in the young nonsmokers. However, this apparent contradiction could be explained by the fact that other mediators, irrespective of the status of IL-12, such as nitric oxide, can influence STAT4 activation.

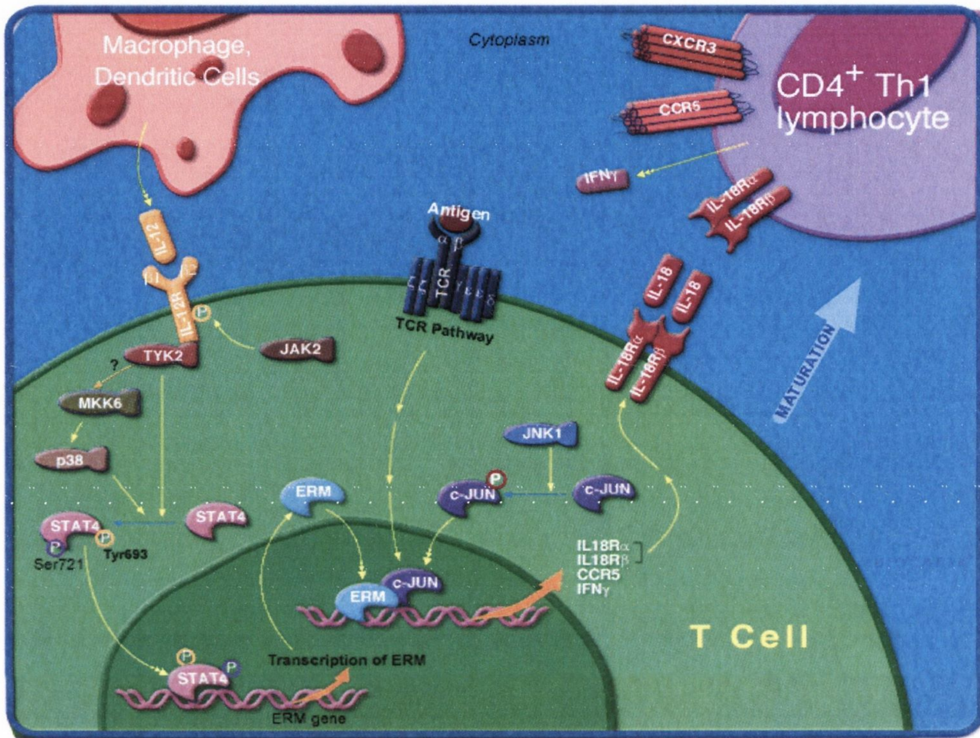


Figure 7.8 IL-12 and STAT4 dependant signalling pathways. http://www.biocarta.com/pathfiles/h_IL12Pathway.asp

7.6.3 MCM10 and CDC6

The MCM (mini chromosome maintenance) and Cdc (cell division cycle) proteins are important regulators in the process of eukaryotic DNA replication. These proteins are detectable throughout the proliferative phases of the cell cycle but are lost in differentiation and quiescence, making them specific indicators of cell proliferation.

The crucial role of MCM and Cdc6 in the cell cycle is mirrored by their potential in cancer diagnosis. In fact, it has been recently shown that Cdc6 is downregulated in prostate cancer, and that downregulation of Cdc6 is associated with an aggressive prostate cancer phenotype (Robles, 2002) (Figure 7.9). In contrast, MCM and cdc6 upregulation were found to occur as a consequence of cervical carcinoma progression (Murphy, 2005). The presence of proteins of the prereplicative complex in the nuclei of dysplastic and neoplastic cells is consistent with the notion that these cells remain in cycle, as a result of deregulation of normal controls over cell proliferation. The present study found MCM10 and CDC6 to be significantly downregulated in young nonsmokers compared with old smokers.

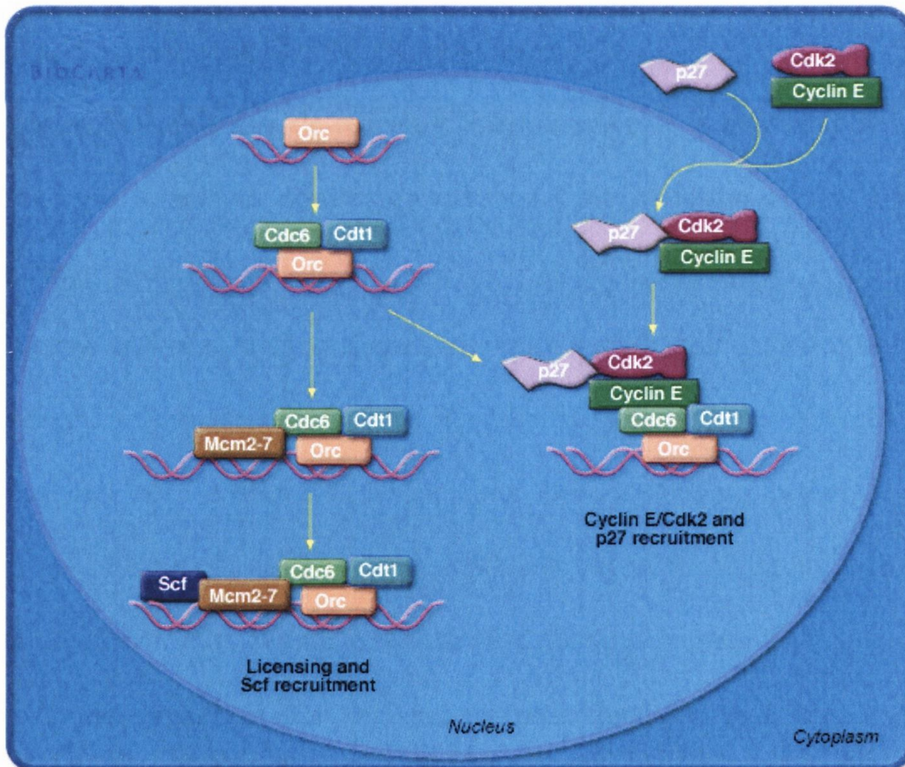


Figure 7.9 Regulation of DNA replication.

http://www.biocarta.com/pathfiles/m_mcmPathway.asp

7.6.4 Other genes

CDK5RAP3, a regulator of Cyclin-dependent kinase 5 (CDK5) activity, is primarily associated with brain development but it is also implicated in lens and muscle differentiation. TRRAP (transformation/transcription domain-associated protein) is a 434 kDa that is utilised by transcription factors important in human cancer, including p53, c-MYC and E2F, to activate transcription of their downstream target genes (McMahon, 1998). p53 binds directly to TRRAP and recruits it to the promoter of the p53 target gene mdm2. TRRAP recruitment results in increased histone acetylation at the mdm2 promoter and blocking TRRAP expression blocks the ability of p53 to activate mdm2 transcription. The biological function of this molecule and how it controls proliferation are still unclear and little has been published on the association of TRRAP with cancer.

RBA600 (retinoblastoma associated factor 600) plays a role in cell cycle regulation and studies have shown an association between overexpression of this gene and neuroblastoma tumourigenesis.

On microarray analysis of these genes in patients with oral and oropharyngeal cancer, this study found a significant differential expression level between young nonsmokers and old smokers, with the latter cohort showing higher levels of expression of TRRAP, CDK5RAP3 and RBA600. On PCR validation, there were some contradictory results, with one non-smoker showing consistently high levels of expression of these genes on PCR analysis, contradicting the array results.

There are a number of suggestions for this disparity, the most likely being:

1. Slight sample degradation between the expression array experiment and the TaqMan experiment

2. The oligonucleotide probes on the array are 60 base pairs (60 mers) and are designed to report an aggregate gene activity level from alternatively spliced transcripts of a gene, and they are designed not to discriminate between the abundance levels of alternative transcripts. TaqMan probes are far more discriminative and there are a number of splice variants to choose from when selecting the Taqman probe, thus it is highly probable that the probe location only targeted one variant resulting in different values.

7.6.5 p16 and RB1

The regulation of retinoblastoma is known to be intimately related p16 expression. In fact, p16 protein overexpression is thought to be a marker for inactivation of retinoblastoma gene. In this study, a remarkable correlation existed between overexpression of p16 mRNA and overexpression of RB1 mRNA.

This apparent contradiction could be explained by the fact that CDK regulation of the retinoblastoma gene occurs at the protein level. Therefore, while the mRNA levels may still mirror each other, overexpression of p16 protein is usually accompanied by underexpression of Rb protein. Moreover, it has been noted in the literature that high levels of p16 expression were not necessarily the result of effects on the retinoblastoma gene (Rb) but could be in response to a dysfunction elsewhere within this pathway, however the correlation in expression levels between the two genes is too significant to be a mere coincidence.

Of note is the detection of HPV mRNA in the two cases with the highest levels of p16 mRNA and RB1 mRNA and furthermore, these cases also strongly expressed

the p16 antibody. Further studies on RB1 antibody are planned in order to shed more light on these results.

7.6.6 CDKN2B

The genes encoding p15 and its closely related human homologue p16 are located adjacent to one another on chromosome 9p21 (Figure 7.10). The primary function of p15 is inactivation of the cdk4 and cdk6 kinases, by direct binding and possibly by displacement of cyclin D, which prevents phosphorylation of Rb and leads to G1 arrest. Mutational inactivation of INK4B (p15) is observed in tumours, but rarely without concomitant mutation of INK4A (p16) (Lang, 2002). Loss of expression of p15 is observed in some tumours due to hypermethylation of the INK4B promoter (Chang, 2004), and in others due to deletion (Swellam, 2004). Transcription of INK4B is stimulated by TGF-β or interferon-α, and increased p15 expression is one way in which TGF-β affects cell cycle arrest (Weinberger, 2004). Cells show elevated p15 expression upon differentiation whereas proliferative growth is associated with reduced or absent expression of p15.

In this study, p15 expression was significantly higher in nonsmokers compared with smokers. There are a number of explanations for this differential expression. Firstly, p15 may simply be overexpressed in the nonsmokers. Like p16 regulation, this overexpression could be caused by a defective negative feedback mechanism, or by a mutated form of p15. However, the differential expression level may be due to methylation and subsequent underexpression of p15 in the smoking cohort (Lang, 2002) rather than overexpression in the nonsmoking cohort.

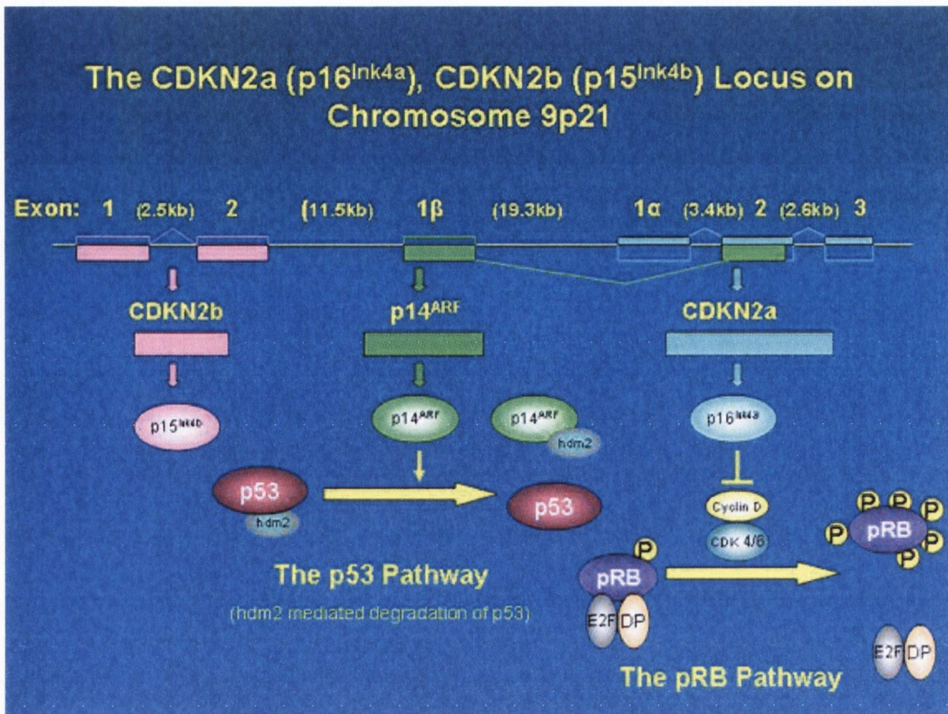


Figure 7.10 CDKN2b and CDKN2A arose from a gene duplication event and are consequently very similar (77% amino acid sequence identity in humans). From University of Vermont, CDKN2A Database Project (biodesktop.uvm.edu/perl/p16).

7.7 Conclusion

In this chapter, data collected from the expression array study has been validated. Future work will involve analysis of the identified differentially expressed genes in a larger sample cohort. The true value of the genes of interest identified in this chapter and Chapter 6 can only be established by applying the findings to a much larger number of patients.

7.8 References

Belinsky S, Swafford D, Middleton S, Kennedy C, Tesfaigzi J. Deletion and differential expression of p16INK4a in mouse lung tumours. Carcinogenesis. 1997; 18(1): 115-20

Burdelya L, Catlett-Falcone R, Levitzki A, Cheng F, Mora L, Sotomayor E, Coppola D, Sun J, Sebti S, Dalton W, Jove R, Yu H. Combination Therapy with AG-490 and Interleukin 12 Achieves Greater Antitumor Effects than Either Agent Alone. Molecular Cancer Therapeutics 2002; (1), 893-899,

Chang HW, Ling GS, Wei WI, Yuen AP. Smoking and drinking can induce p15 methylation in the upper aerodigestive tract of healthy individuals and patients with head and neck squamous cell carcinoma. Cancer. 2004: 101(1): 318-332

Janoueix-Lerosey I, Novikov E, Monteiro M, Gruel N, Schleiermacher G, Loriod B, Nguyen C, Delattre O. Gene expression profiling of 1p35-36 genes in neuroblastoma. Oncogene. 2004; 23(35): 5912-22

Lang JC, Borchers J, Danahey D, Smith S, Stover DG, Agrawal A, Malone JP, Schuller DE, Weghorst CM, Holinga AJ, Lingam K, Patel CR, Esham B. Mutational status of overexpressed p16 in head and neck cancer: evidence for germline mutation of p16/p14ARF. International Journal Oncology. 2002; 21(2): 401-8.

McMahon SB, Van Buskirk HA, Dugan KA, Copeland TD, Cole MD. The novel ATM-related protein TRRAP is an essential cofactor for the c-Myc and E2F oncoproteins. Cell. 1998; 94(3): 363-74.

Mok SC, Wong KK, Chan RK, Lau CC, Tsao SW, Knapp RC, Berkowitz RS. Molecular cloning of differentially expressed genes in human epithelial ovarian cancer. Gynaecologic Oncology. 1994; 52(2): 247-52

Murphy N, Ring M, Heffron CC, Martin CM, McGuinness E, Sheils O, O'Leary JJ Quantitation of CDC6 and MCM5 mRNA in cervical intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix. Mod Pathol. 2005 Feb 4 epub

Rimon E, Sasson R, Dantes A, Land-Bracha A, Amsterdam A. Gonadotropin-induced gene regulation in human granulose cells obtained from IVF patients: Modulation of genes coding for growth factors and their receptors and genes involved in cancer and other diseases. International Journal of Oncology 2004; 24: 1325-1338

Robles LD, Frost AR, Davila M, Hutson AD, Grizzle WE, Chakrabarti R. Down-regulation of Cdc6, a cell cycle regulatory gene, in prostate cancer. Journal of Biological Chemistry. 2002; 277(28), 25431-25438

Schwarze S, DePrimo S, Grabert L, Fu V, Brooks J, Jarrard D. Novel Pathways Associated with Bypassing Cellular Senescence in Human Prostate Epithelial Cells. The Journal of Biological chemistry 2002; 17:14877–14883,

Swellam M, El-Aal AA, AbuGabel KM. Deletions of p15 and p16 in schistosomal bladder cancer correlate with transforming growth factor-alpha expression. Clinical Biochemistry. 2004; 37(12): 1098-104.

Weinberger P, Yu Z, Haffty B, Kowalski D, Harigopal M, Sasaki C, Rimm D and Psyrri A. Prognostic Significance of p16 Protein Levels in Oropharyngeal Squamous Cell Cancer. Clinical Cancer Research 2004(10), 5684-5691,

Chapter 8

General Discussion

8.1 Introduction

Cancer of the oral cavity and oropharynx kills more people every year than malignant melanoma, cancer of the cervix or cancer of the uterus. Oral and oropharyngeal cancer occurs predominantly in older males (>60 years of age), following decades of smoking and alcohol abuse. However, the myth that young people do not get oral cancer needs to be dispelled. Between 5 and 10 per cent of cases will occur in younger patients, many of who do not have the identified risk factors for head and neck cancer, such as smoking and alcohol (Llewellyn, 2004). Female patients are also over represented in the younger group compared with the more typical group of oral cancer patients (McKenzie, 2000; Llewellyn, 2004).

The first published report of oral cancer in young adults was in 1970 by Jones, who described oral cancer in two young adults. Jones continued researching this disease cohort and published again in 1989, having identified 11 young adults with oral cancer over a ten-year period. Jones noted in these patients, that the previously identified risk factors, and the male/female ratio were quite different from the older patients (Jones, 1970; Jones, 1989). Since 1970, over 30 studies have reported on oral cancer in young patients. Reports from the U.S., Scotland, England, and Scandinavia have all noted an increasing trend in tongue in young adults (Schantz, 2002; Annertz, 2002; Depue, 1986; Davis, 1987). Most studies have focused on social habits and the clinical course of the disease (Sasaki, 2005; Lipkin, 1985; Llewellyn, 2004; Manuel, 2003), while few have looked at molecular events in oral cancer in these young adults (Lingen, 2000; Sorensen, 1997).

The overall aim of this study was to investigate molecular events of oral and oropharyngeal cancer in a group of young patients and to compare the findings with the more typical oral cancer patient.

8.2 Review of results and conclusions

Most of the previous literature concerning oral and oropharyngeal cancer in young adults points to the unusual characteristics of these patients – the disproportionate number of women, the lack of heavy smoking and drinking, and the predilection for the tongue (McGregor, 1983; Carniol, 1982). It was clearly important, at an early stage in this thesis, to assess an Irish cohort and establish as a baseline that trends seen in other parts of the world were reflected in the Irish population.
Concurring with the literature, the gender ratio, tumour site and the tobacco smoking status were dramatically different in the young cohort compared to the older patient. The male: female ratio was far higher in the older patient (4.8:1 in the older patient compared with 1.7:1 in the younger). The majority of tumours in the young patients occurred in the tongue (78%), whereas the older group showed a predominance of laryngeal and pharyngeal cancers. While the majority of the older cohort had a history of tobacco smoking, the young group (especially the under thirty year olds) were never-smokers.

Chromosome aberrations represent very frequent and important events in head and neck cancer with studies reporting consistent alterations in certain chromosomal sites, such as gains on 3q, 5p, 8q, 11q13, and losses on 3p, 9. The aim of Chapter 4 was to broaden the search for alternative genetic alterations in tumours of young

nonsmokers, by using microarray CGH. Four oral and oropharyngeal cancer cohorts were established;

over 40 year old smokers (typical profile of oral and pharyngeal cancer patients)

over 40 year old nonsmokers

under 40 year old smokers

under 40 year old nonsmokers

The 'older group' in this study showed a profile similar to that of the 'classic' oral cancer patient in the literature, including deletions on 3p, 9p and gains of 3q, 5q, 7p, and 11q13. The aberration profile detected in the young nonsmokers was markedly different to that detected in the old smokers. In fact, the young nonsmokers showed far less genomic instability of the chromosomes consistently aberrated in HNSCC in smokers. Another significant finding was the complete absence of deletion of the p16 gene in the young cohorts, irrespective of smoking history. This is a gene that is commonly deleted in classic oral cancer. This further raises the possibility of alternative pathways being brought into play in the development of oral and oropharyngeal cancer in young adults.

Of interest was the high number of aberrations identified in all the smokers, irrespective of age. While 40 years is a very reasonable age cut-off for distinguishing between young and old, one has to note that 35-40 year old smokers may still have been smoking for 20 years, thus it is not surprising that they closely resemble the old smokers in their genomic profile. Perhaps, future studies should focus more on under 30 year olds. This does raise the issue of study size, with far less statistical power being produced from a very small sample number.

Also of interest was the similarity between the old nonsmokers profile and the profiles of the smoker cohorts. Both the number of aberrations and site of the

aberrations were quite similar which raises the possibility that prolonged exposure in a secondary manner to similar carcinogens may be a factor contributing to the development of oral cancer in the old non-smoking group. The possibility that some of these aberrations are simply a result of aging cannot be out ruled. However, the aberrations found in the old nonsmokers closely resemble the chromosomal changes seen in all the smokers and those changes quoted in the literature. One might ask is the old nonsmoker a suitable cohort. While it is true to say that maybe dysplastic or normal samples may be seen as a better comparative group, the difficulty in collection of these samples has meant that a compromise was necessary and the next most suitable sample groups were as described above.

By the end of Chapter 4, it was clear that the young nonsmoker cohort were quite different to all the others, most strikingly in the absence of p16 deletions. Both low p16 expression and overexpression of p16 have been reported previously in oral and oropharyngeal cancer (El-Naggar, 1999; Pandey, 1998, Andl, 1998; Gillison, 2000), and the overexpression of p16 had often been linked to HPV infection. Several research groups support the aetiologic role for HPV16/18 in at least a subset of head and neck cancers, and some have suggested it as the causative agent in young nonsmokers. In view of the different patterns of p16 aberrations revealed in Chapter 4, the objectives of Chapter 5 were to elaborate on this.

In brief, alteration or overexpression of p16 did not seem to feature strongly in the development of cancer in young nonsmokers. Regarding the HPV prevalence in young nonsmokers, previously thought to be a causative agent, no female nonsmokers had evidence of HPV16 integration. Interestingly, concurring with very recent reports, HPV prevalence was highest in the oropharyngeal cancers, and p16 immunohistochemistry staining mirrored this. Considering the propensity of the

young female group for tumours occurring in the tongue, it is not surprising that HPV was not detected.

Overall the Chapter 5 findings strongly suggest that HPV 16/18 and subsequent p16 alterations do not have significant roles to play in the development of oral/oropharyngeal cancer in young adult nonsmokers.

It is important to note that the clinical and histologic phenotypes of all the tumour samples used in this thesis are similar, i.e they all are squamous cell carcinomas.

Leading on from the CGH studies, expression microarray experiments were performed on the same sample cohort. The expression microarray studies revealed a group of genes that were differentially expressed in samples of SCC from young nonsmoking adults compared with old smoking adults.

It is important to bear in mind when working with expression microarray platforms that so many different algorithms can be used to generate gene lists. The analysis methods used in Chapter 6 are simple and are built around two group comparisons using ANOVA. Multiple gene lists were generated with a number of genes sparking immediate interest. It must be acknowledged that the results of any study of small sample size can only be suggestive, and thus point to the need for further validation.

Signal transducer and activator of transcription (STAT) proteins play a critical role in regulating physiological responses to stimulation by cytokines and growth factors. In HNSCC, there is compelling evidence that constitutive activation of STAT is linked to cancer development and growth (Song, 2000). While initial investigations have established that both STAT3 and STAT1 have a significant role to play in HNSCC progression there is to date no evidence that STAT4 is linked to head and

neck cancer (Grandis, 2000). The results of the expression array study in Chapter 6 indicate that STAT4 is differentially expressed in SCC in young nonsmokers compared with old smokers. STAT4 is a central mediator in generating inflammation during protective immune responses and immune-mediated diseases. The results in Chapter 6 may suggest that STAT4 has a principle role to play in carcinogenesis in SCC in young adults. This remains to be investigated further. More significantly, perhaps, it does however raise the possibility that the immunoregulation of young adults developing head and neck cancer may be quite different from that of classic head and neck cancer patients.

CDC6 (Cell division cycle 6 homolog) is part of the initiator protein complex of DNA replication. CDC6 has emerged as a potential marker for cervical neoplasia (Murphy, 2004). The MCM2-MCM7 complex is an essential component of the prereplication complex (pre-RC), which is recruited by the CDC6 and CDT1 proteins to origins of DNA replication during G (1) phase. Recently MCM8 has also been reported as a crucial component of the pre-RC (Volkening, 2005). In Chapter 6 of this thesis, MCM10 upregulation mirrored that of CDC6, and this presents the opportunity to further investigate the role of other MCMs in DNA replication.

8.3 An emerging approach – pathway analysis

Genes never act alone in a biological system – they are working in a cascade of networks. As a result, analysis of microarray data with specific attention to molecular pathways could lead to a higher level of understanding of the system. Using the Panther website and Panther analysis algorithms

(http://panther.appliedbiosytems.com), the expression data was mapped onto metabolic pathways in order to identify and evaluate specific metabolic pathways most relevant to the specific expression experiments. In particular, by comparing the pathways most active in young nonsmokers compared with old smokers, it was possible to identify biologically relevant molecules (Figure 8.1)

Signalling pathways in young nonsmokers	**Pathways common in both cohorts**	**Signalling pathways in old smokers**
JAK/STAT Apoptosis EGFR Alzheimer disease Endothelin	PDGF Interleukin Inflammation mediated by cytokines	FAS TGF-beta Wnt

Figure 8.1 Signalling pathways used in oral and oropharyngeal SCC

After the RT-PCR studies, STAT4 and CDC6 are currently the genes of interest that will be immediately followed up in larger study groups. The discovery of p15 as a differentially overexpressed gene in nonsmokers may also be relevant and requires further validation.

The CGH array experiments resulted in the emergence of two distinct profiles, with the genomic profile of old smokers with SCC showing copy gain and loss of all the typical genes/loci described previously in the literature, while that of the young non-smoker lacked aberrations of these genes/loci. In particular p16 previously shown to be frequently deleted in head and neck cancer was never deleted in young nonsmokers in this cohort. One would expect that the expression microarray results would mirror closely this striking distinction between the two cohorts. However, such a marked distinction is not clearly evident in the expression array studies. This

is difficult to reconcile, however, there are a number of possible explanations that may help to rationalise this.

1. The genomic profiles differ mainly on the basis of deletions; chromosome 3, 5 and 9 deletions were evident in old smokers and not in young nonsmokers. There may be very little difference in expression levels between those with p16 deletions and without. Therefore the differential expression may not be large enough to qualify in the top 100 or so genes found to be differentially expressed using the statistical algorithm applied in this study.

2. While genes were found to be differentially aberrated in the old smokers by CGH studies, these very same genes may in some way be modified other than a copy gain and loss in the young nonsmokers resulting in similar signal levels in the expression array, thus not differentially expressed. This highlights a limitation of the microarray CGH study.

Fresh tissue acquisition is a challenging area, from initial ethical approval to tissue handling and to successful RNA extraction. To make matters even more challenging, the acquisition of fresh tissue from a rare group (young nonsmokers with oral and oropharyngeal SCC) was critical for the fulfilment of this study. While larger numbers would have resulted in increased statistical power of this study, it must be highlighted that studies restricted in number still have their merit and until a tumour bank is well established, small cohort studies like this can be performed and the results validated on the more readily available paraffin embedded archival samples.

8.4 'Every clarification breeds new questions' Arthur Bloch

It is clear from an epidemiological perspective that there is a trend towards earlier development of head and neck cancer over the past forty years. Since the early 1970s, there have been reports of an increase in tongue cancer development in young adults. Many of these reports highlight the absence of a smoking history in these young adults. While it can be recognised that oral cancer is increasing in young adults and it appears to be a disease with a different molecular profile to the classic oral cancer, a specific question remains to be answered:

What causes head and neck cancer in young adults that do not smoke?

Perhaps the causative agent is one of the many potential risk factors in head and neck cancer. This thesis has investigated HPV as a potential causative agent, however there are many others, such as dietary factors (Cho, 2003).

A likely theory is that an environmental agent/s acts in tandem with a host susceptibility factor. Considering the very different genomic profiles identified in young adult nonsmokers with cancer, compared with old smokers with cancer, it is possible that some host factor may enhance the carcinogenic effect of environmental agents. Further higher resolution analysis of the genomic profile might help to unravel some of these theories. In the quest to develop a better understanding of the reasons why oral and oropharyngeal cancer are increasingly being discovered in young adults, we plan to use single nucleotide polymorphism microarrays.

Single nucleotide polymorphisms or SNPs are DNA sequence variations that occur when a single nucleotide in the genome sequence is altered. For example a SNP

might change the DNA sequence AAGGCTAA to ATGGCTAA. SNPs, which make up about 90% of all human genetic variation, occur every 100 to 300 bases along the 3-billion-base human genome. Variations in DNA sequence can have a major impact on how humans respond to disease; environmental insults such as bacteria, viruses, toxins, chemicals and drugs. This makes SNPs very valuable tools in biomedical research. SNPs are also evolutionarily stable making them easier to follow in population studies. Defining and understanding the role of genetic factors in disease will also allow researchers to better evaluate the role of non-genetic factors e.g. such as diet and lifestyle have on disease.

'It is not the answer that enlightens, but the question' Descouvertes

Throughout this thesis, the focus has been on four primary cohorts, namely tumours from young nonsmokers, young smokers, old nonsmokers and old smokers. Clearly the principal finding in this thesis is that squamous cell carcinomas developing in young nonsmokers are markedly different, not in any recognisable phenotypic way, but undoubtedly at a genetic level. Of course this is intriguing, and possibly begs more questions than it answers. Given that smoking appears not to be causative of the tumours in these people, it seems plausible if not very probable that there is a factor or factors, be they environmental or genetic that underlie this predisposition. The real challenge ahead in this area is to identify either a previously unrecognised environmental pathogen, if one truly exists, or discover a previously unrecognised genetic susceptibility that predisposes to the development of this cancer. By asking and hopefully in time answering these questions, we will learn more about not only the pathogenesis of this particular cancer, but also the biology of cancer in general.

8.5 References

Althuis MD, Brogan DR, Coates RJ, Daling JR, Gammon MD, Malone KE, Schoenberg JB, Brinton LA. Hormonal content and potency of oral contraceptives and breast cancer risk among young women. Br J Cancer. 2003; 88(1): 50-7

Andl T, Kahn T, Pfuhl A, Nicola T, Erber R, Conradt C, Klein W, Helbig M, Dietz A, Weidauer H, Bosch FX. Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinoma lacking retinoblastoma cell cycle control. Cancer Res 1998; 58:5-13.

Annertz K, Anderson H, Biorklund A. Incidence and survival of squamous cell carcinoma of the tongue in Scandinavia, with special reference to young adults. Int J Cancer 2002; 101:95-9.

Carniol PJ, Fried MP. Head and neck carcinoma in patients under 40 years of age. Ann Otol Rhinol Laryngol. 1982 91(2 Pt 1): 152-5.

Cho E, Spiegelman D, Hunter DJ, Chen WY, Colditz GA, Willett WC. Premenopausal dietary carbohydrate, glycemic index, glycemic load, and fiber in relation to risk of breast cancer. Cancer Epidemiol Biomarkers Prev. 2003 Nov; 12(11): 1153-8.

Davis S, Severson RK. Increasing incidence of cancer of the tongue in the United States among young adults. Lancet 1987; 2:910-11

Depue RH. Rising mortality from cancer of the tongue in young white males. N Engl J Med 1986; 315:647

El-Naggar AK, Lai S, Clayman GL, et al. Expression of p16, Rb, and cyclin D1 gene products in oral and laryngeal squamous carcinoma: biological and clinical implications. Hum Pathol. 1999; 30: 1013-1018

Grandis JR., Zeng Q, Drenning SD. Epidermal growth factor receptor--mediated stat3 signaling blocks apoptosis in head and neck cancer. Laryngoscope 2000, 110, 868-874. J Clin Pathol. 2005; 58(5): 525-534.

Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L *et al.* Evidence of a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst. 2000; 92(9): 675-677, 709-720.

Jones BJ, Lampe HB, Cheung HW. Carcinoma of the tongue in young patients. J Otolaryngol. 1989; 18(3): 105-8

Jones BJ. Oral carcinoma in the young patient. With a report of two cases. Br J Oral Surg. 1970; 8(2): 159-64.