



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.

Non-Coding RNA Expression in Cancer Stem Cell Progenies

Derived from Tumours +/- BRAF V600E Mutation

Gary Sommerville

A thesis submitted to the University of Dublin

for the degree of

Doctor of Philosophy

Based on research performed at

Trinity College Dublin

Department of Histopathology

St. James's Hospital

Dublin 8

Ireland

Supervisor: Professor Orla Sheils

Co-Supervisor: Professor John O' Leary

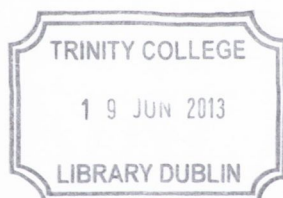
May 2013

I. Declaration

This thesis is submitted by the undersigned for the degree of Doctor of Philosophy at the University of Dublin. It has not been submitted previously as an exercise for a degree at this or any other University. I declare that this thesis is entirely my own work, except where acknowledged in the text. I give permission to the Library to lend or copy this thesis upon request.

G. Sommerville

Gary Sommerville M.Sc.



Thesis 10159

II. Summary

Cancer has placed a huge burden on the global health system and rising rates in particular cancer types such as lung cancer and melanoma due to smoking and increased UV ray exposure respectively has highlighted the need to combat this disease. Intense research has shed light on the genetic processes that contribute to tumour formation and progression. The BRAF gene is commonly mutated and is thought to be implicated in nearly one third of all human cancers and so represents an important target for novel therapies to combat tumorigenesis. The V600E mutation confers constitutive kinase activity, independent of mitogen initiation, within the MAP-ERK pathway which can lead to aberrant cell signalling. Cancer stem cells have been identified in many human malignancies in a variety of different tissue types ranging from prostate to thyroid. These cells possess characteristics similar to normal human stem cells and are believed to contribute towards tumorigenesis through self-renewal and differentiation. Additionally, persistence of cancer stem cells post-treatment has been proposed as an explanation for metastasis and tumour repopulation. MicroRNAs (miRNAs) are recently discovered post-transcriptional regulators that target messenger RNA transcripts (mRNAs), usually resulting in translational repression. The aim of this project was to investigate if cancer stem cells could be isolated from various different cancer types, both BRAF wild type and BRAF V600E mutated, and to establish if they have stemness characteristics. Additionally we sought to investigate the downstream effects of BRAF mutation on methylation status of cancer-related genes and the contribution of de-regulated non-coding RNAs towards tumorigenesis. Having SNP genotyped cells lines to establish BRAF status, holoclones were generated from cell lines using a high salt soft agar assay. Quantitative Real Time TaqMan[®] PCR methods were used to investigate the expression of stemness genes in holoclone populations. Once a stemness signature was established in holoclones, we proceeded to next generation sequencing via the SOLiD[™] 4 platform to highlight dysregulated non-coding RNAs and novel transcripts in cancer stem cells with high capacity for self-renewal. We also investigated the methylation pattern of cancer genes in the same cell lines. This study seeks to understand the association of BRAF mutation and dysregulation of different cellular processes resulting in tumorigenesis and to identify potential targets in the treatment of various human malignancies.

III. Acknowledgements

I would like to express my sincere thanks to Professor Orla Sheils for her constant guidance and mentorship in all aspects of my research over the last three years. You helped better me as a scientist and for that I am eternally grateful. I would also like to thank my co-supervisor Professor John O' Leary for his support during this endeavour.

The opportunity to carry out my research at the Department of Histopathology would not have been possible without the generous financial support of the Department and the Trinity College Dublin Postgraduate Research Studentship.

I would like to thank my colleagues in the laboratory, both past and present, for making it a stimulating and exciting environment to work in: Dr. Paul Smyth for his advice and support throughout the course of my research, Dr. Victoria McEneaney, Dr. Daragh Crowley, Dr. Cathy Spillane, Dr. Barry Murphy, Dr. Sinead Aherne, Emma Dorris, Lisa Keogh and Yvonne Salley. I am also indebted to Dr. Gordon Blackshields who provided the bioinformatical support for this project. To the new additions to the lab, Steven Busschots and Louise Flynn, you have joined a great department and I wish you all the best in your research.

Finally, my family and friends deserve special thanks for their encouragement and support throughout the course of my Ph.D.

IV. Presentations / Publications

“Non-Coding RNA and BRAF V600E Mutation in Thyroid Cancer”, Irish Association for Cancer Research Conference, Dublin, February 2013

“Stemness Gene Expression in Cancer Stem Cell Progenies +/- BRAF V600E Mutation”, Poster Presentation, United States & Canadian Academy of Pathologists Conference, Vancouver, March 2012

“Stemness Gene Expression in Cancer Stem Cell Progenies +/- BRAF V600E Mutation”, Oral Presentation, International Cancer Conference, Trinity College Dublin, September 2011

“Stemness Gene Expression in Cancer Stem Cell Progenies +/- BRAF V600E Mutation” School of Medicine Postgraduate Research Day, Trinity College Dublin, September 2011

Non-Coding RNA Profiling in Cancer Stem Cell Progenies +/- BRAF V600E Mutation, Manuscript in Preparation

V. Table of Contents

I.	Declaration	i
II.	Summary	ii
III.	Acknowledgements	iii
IV.	Presentations / Publications	iv
V.	Table of Contents	v
VI.	List of Figures	x
VII.	List of Tables	xv
VIII.	List of Abbreviations	xvii
Chapter 1 Introduction		1
1.1	Cancer Genetics	2
1.2	Melanoma	4
1.3	Thyroid Cancer	5
1.4	Ovarian Cancer	6
1.5	Colorectal Cancer	7
1.6	The BRAF Gene	9
1.7	Cancer Stem Cells	12
1.8	Cancer Stem Cell Morphologies	15
1.9	Stemness Genes & Associated Pathways	16
1.10	MicroRNAs	20
1.11	MicroRNA Biogenesis	21
1.12	MicroRNAs in Cancer	24
1.13	SnoRNAs	25
1.14	Methylation in Cancer	26
1.15	SOLiD™ Next Generation Sequencing	28
1.16	Aims & Objectives	29
Chapter 2 Materials & Methods		31
2.1	Introduction	32

2.2 Cell Culture	32
2.2.1 Cell Lines	32
2.2.2 Propagation of Cells from Liquid Nitrogen Storage	34
2.2.3 Subculturing	34
2.2.4 Cell Counting with a Haemocytometer	34
2.2.5 Cryopreservation	35
2.2.6 High Salt Agar Assay	35
2.3 Nucleic Acid Extraction, Quantitation and Quality Control	36
2.3.1 Genomic DNA Purification	36
2.3.2 RNA Extraction	36
2.3.3 Nucleic Acid Quantitation	36
2.3.4 Nucleic Acid Quality Control	37
2.3.4.1 Analysing RNA by the Agilent 2100 Bioanalyser	37
2.3.4.2 Analysing Small RNA by the Agilent 2100 Bioanalyser	37
2.3.4.3 Analysing DNA by the Agilent 2100 Bioanalyser	37
2.4 Polymerase Chain Reaction	37
2.4.1 Reverse Transcription of RNA to cDNA	37
2.4.2 TaqMan® PCR	38
2.4.3 TaqMan® SNP Genotyping and Allelic Discrimination Assay	38
2.4.4 Real-time quantitative TaqMan® RT-PCR	40
2.4.4.1 Relative Gene Expression	41
2.4.4.2 TaqMan® Gene Expression Assays	42
2.5 Protein Analysis	43
2.5.1 Immunofluorescence Staining of Cells by Confocal Microscopy	43
2.5.2 Confocal Microscopy	44
2.6 Methylation Analysis	44
2.6.1 Performing Bisulfite Conversion	45
2.6.2 Methylation Primer Design	45
2.6.3 PCR Amplification	46
2.6.4 Bisulfite Sequencing	46

2.7 Next Generation Sequencing of Small RNAs via SOLiD™ 4	47
2.8 Sample Preparation	47
2.9 Library Preparation	48
2.9.1 Hybridisation, Overnight Ligation and Reverse Transcription	48
2.9.2 cDNA Size Selection	48
2.9.3 In-Gel PCR and Multiplex Barcoding	49
2.9.4 Preparation of Multiplexed Library	49
2.10 Multiplexed Library Quantitation	50
2.11 Templated Bead Preparation	51
2.11.1 Full Scale Emulsion PCR	51
2.11.2 Emulsion Breaking and Bead Washing	52
2.11.3 SOLiD™ Bead Quantitation	53
2.11.4 SOLiD™ Templated Bead Enrichment	54
2.11.5 P2-Enriched Bead Isolation	55
2.11.6 3' End Modification	55
2.11.7 Modified SOLiD™ Bead Quantitation	55
2.12 SOLiD™ 4 Next Generation Sequencing Run	56
2.12.1 Workflow Analysis (WFA)	56
2.12.1.1 Preparation of SOLiD™ XD Slides	56
2.12.1.2 WFA	56
2.12.2 SOLiD™ 4 Sequencing Run	57
Chapter 3 Gene Expression in Holoclones Derived From Cell Lines +/- BRAF	
V600E Mutation	63
3.1 Introduction	64
3.2 TaqMan® SNP Genotyping of Cell Line Panel	65
3.2.1 Melanoma	65
3.2.2 Thyroid	67
3.2.3 Ovarian	69
3.2.4 Colorectal	71

3.3 High Salt Agar Assay	73
3.3.1 Melanoma	73
3.3.2 Thyroid	74
3.3.3 Ovarian	75
3.3.4 Colorectal	76
3.4 TaqMan® Gene Expression in Holoclones	78
3.4.1 Melanoma	78
3.4.2 Thyroid	82
3.4.3 Ovarian	86
3.4.4 Colorectal	90
3.4.5 Grouped Differential Gene Expression	93
3.4.6 Discussion	102
3.5 Immunofluorescence of NANOG protein	108
3.5.1 HT-29	108
3.5.2 COLO320	112
3.5.3 Discussion	113
Chapter 4 Non-Coding RNA Expression in Holoclones Derived From Cell Lines	
+/- BRAF V600E Mutation	114
4.1 Introduction	115
4.2 SOLiD™ 4 Next generation sequencing and annotation of miRNAs	116
4.3 Differential miRNA expression profiling in holoclone populations	119
4.3.1 Melanoma	119
4.3.2 Thyroid	120
4.3.3 Ovarian	123
4.3.4 Colorectal	124
4.3.5 Discussion	125
4.4 Overrepresented Gene Ontology (GO) terms in holoclone populations	129
4.4.1 Melanoma	130
4.4.2 Thyroid	134
4.4.3 Ovarian	138

4.4.4 Colorectal	142
4.4.5 Discussion	146
4.5 Differential snoRNA expression in holoclone populations	148
4.5.1 Melanoma	149
4.5.2 Thyroid	150
4.5.3 Ovarian	151
4.5.4 Colorectal	152
4.5.5 Discussion	153
Chapter 5 Methylation of RARβ2 and MLH1 in a Cell Line Panel +/- BRAF V600E Mutation	158
5.1 Introduction	159
5.2 PCR of Bisulfite Converted DNA from Cell Line Panel	161
5.3 Methylation Sequencing Results	163
5.3.1 MLH1	163
5.3.2 RAR β 2	164
5.3.3 Discussion	164
Chapter 6 Discussion	169
viii. Bibliography	176
ix. Appendix I. Supplementary Methods	198
x. Appendix II. Supplementary Figures	223
xi. Appendix III. Supplementary SOLiD TM Data	240

VI. List of Figures

Figure 1.5	Illustration depicting the accumulation of genetic abnormalities that take place in the progression of normal colon cells to adenoma and finally to colon carcinoma
Figure 1.6	Illustration depicting the MAPK-ERK cell signalling pathway
Figure 1.7	Illustration depicting the current hypothesis of the involvement of cancer stem cells in tumour progression
Figure 1.8	Green Fluorescent Protein images of the 3 distinct colony morphologies (holoclone, meroclone and paraclone) observed in the prostate cancer cell line PC-3
Figure 1.9	A schematic of the Transforming Growth Factor beta (TGF- β) signalling pathway
Figure 1.11	Illustration depicting the miRNA biogenesis pathway and miRNA-Mediated Gene Regulation in Animal Cells
Figure 1.14(a)	Illustration depicting gene inactivation as a result of methylation
Figure 1.14(b)	Illustration of the bisulfite conversion of genomic DNA
Figure 2.4.3(a)	Allelic discrimination using the 5' nuclease assay
Figure 2.4.3(b)	Typical allelic discrimination output plot illustrating a homozygous wild type, homozygous mutant and heterozygous result
Figure 2.8(a)	Workflow involved in creating Small RNA libraries for next generation sequencing via the SOLiD TM 4 platform
Figure 2.8(b)	Formula used to calculate miRNA percentage in a small RNA sample
Figure 2.9.2	cDNA size selection after gel electrophoresis
Figure 2.9.4	Typical Agilent 2100 Bioanalyser size profile of a successfully amplified library
Figure 2.11.1(a)	Aqueous phase and oil phase prior to emulsification
Figure 2.11.1(b)	Emulsion after amplification
Figure 2.11.3(a)	SOLiD TM Bead Concentration Chart

Figure 2.11.3(b)	SOLiD™ Bead Concentration Workflow
Figure 2.12.2(a)	SOLiD™ Sequencing Chemistry
Figure 2.12.2(b)	Dibase encoding and how it relates to calling the actual template sequence
Figure 3.2.1(b)	Allelic Discrimination plot illustrating the BRAF status of melanoma cell lines COLO794 and SK-Mel 28
Figure 3.2.2(b)	Allelic Discrimination plot illustrating the BRAF status of thyroid cell lines TPC-1 and 8505C
Figure 3.2.3(b)	Allelic Discrimination plot illustrating the BRAF status of ovarian cell lines A2780 and ES-2
Figure 3.2.4(b)	Allelic Discrimination plot illustrating the BRAF status of colorectal cell lines COLO320 and HT-29
Figure 3.4.1(a)	Grouped differential gene expression in SK-Mel 28 BRAF V600E mutated holoclones relative to parental cells
Figure 3.4.1(b)	Grouped differential gene expression in COLO794 BRAF wild type holoclones relative to parental cells
Figure 3.4.2(a)	Grouped differential gene expression in 8505C BRAF V600E mutated holoclones relative to parental cells
Figure 3.4.2(b)	Grouped differential gene expression in TPC-1 BRAF wild type holoclones relative to parental cells
Figure 3.4.3(a)	Grouped differential gene expression in ES-2 BRAF V600E mutated holoclones relative to parental cells
Figure 3.4.3(b)	Grouped differential gene expression in A2780 BRAF wild type holoclones relative to parental cells
Figure 3.4.4(a)	Grouped differential gene expression in HT-29 BRAF V600E mutated holoclones relative to parental cells
Figure 3.4.4(b)	Grouped differential gene expression in COLO320 BRAF wild type holoclones relative to parental cells
Figure 3.4.5(a)	Differential gene expression of NANOG in BRAF V600E and BRAF Wild Type holoclones relative to parental cells

- Figure 3.4.5(b) Differential gene expression of Oct4 in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.4.5(c) Differential gene expression of ALDH1 in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.4.5(d) Differential gene expression of E-Cadherin in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.4.5(e) Differential gene expression of SNAI2 in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.4.5(f) Differential gene expression of SHH in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.4.5(g) Differential gene expression of TGF- β in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.4.5(h) Differential gene expression of β -Catenin in BRAF V600E and BRAF Wild Type holoclones relative to parental cells
- Figure 3.5.1(a) Confocal images of HT-29 parental cells (A) and HT-29 holoclone (B) stained for NANOG protein
- Figure 3.5.1(b) XY and ZY views illustrating the localisation of NANOG protein in a HT-29 holoclone
- Figure 3.5.1(c) Z-Stacks illustrating NANOG expression in a HT-29 holoclone.
- Figure 3.5.1(d) Confocal images illustrating NANOG protein localisation in HT-29 holoclone
- Figure 3.5.2(e) Confocal images illustrating NANOG protein localisation in HT-29 holoclone (no DAPI and no TRITC staining)
- Figure 3.5.2(a) Confocal images of COLO320 parental cells (A) and COLO320 holoclone (B) stained for NANOG protein
- Figure 4.2(b) Workflow Analysis (WFA) for SOLiD next generation sequencing
- Figure 4.4.1(a) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated SK-Mel 28 melanoma cancer cell line

- Figure 4.3.5(a) MAPK-ERK signalling pathway in which several genes are targeted by the miR-26 family members miR-26a and miR-26b
- Figure 4.4.1(b) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type COLO794 melanoma cell line
- Figure 4.4.2(a) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated 8505C thyroid cancer cell line
- Figure 4.4.2(b) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type TPC-1 thyroid cancer cell line
- Figure 4.4.3(a) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated ES-2 ovarian cancer cell line
- Figure 4.4.3(b) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type A2780 ovarian cancer cell line
- Figure 4.4.4(a) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated HT-29 colorectal cancer cell line
- Figure 4.4.4(b) Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type COLO320 colorectal cancer cell line
- Figure 4.5.5(a) Predicted secondary structure of SNORD113/SNORD114 snoRNA family
- Figure 4.5.5(b) Predicted secondary structure of SNORD71
- Figure 4.5.5(c) Predicted secondary structure of SNOORD85
- Figure 5.1(a) Pilot study illustrating the percentage of methylation in cancer related genes across a panel of cell lines
- Figure 5.2(a) 2% agarose gel of successfully amplified control β -Actin in bisulfite converted DNA isolated from cancer cell line panel.

- Figure 5.2(b) 2% agarose gel of successfully amplified MLH1 in bisulfite converted DNA isolated from cancer cell line panel.
- Figure 5.2(c) 2% agarose gel of successfully amplified RAR β 2 in bisulfite converted DNA isolated from cancer cell line panel.
- Figure 5.3.3(a) MLH1 and RAR β 2 methylation percentage in a cell line panel with (indicated in blue) and without (indicated in red) BRAF V600E mutation

VII. List of Tables

Table 2.3.5.2	List of TaqMan® Gene Expression Assays illustrating gene symbol, name and ID
Table 3.4.1(a)	SNP genotyping data for melanoma cell lines COLO794 and SK-Mel 28
Table 3.4.2(a)	SNP genotyping data for thyroid cell lines TPC-1 and 8505C
Table 3.4.3(a)	SNP genotyping data for ovarian cell lines A2780 and ES-2
Table 3.4.4(a)	SNP genotyping data for colorectal cell lines COLO320 and HT-29
Table 4.2(a)	Uniquely barcoded samples and total number of sequence reads generated for each library.
Table 4.3.1(a)	Fold change of miR-103a in SK-Mel 28 holoclones
Table 4.3.1(b)	Fold change of miR-26 transcripts in COLO794 holoclones
Table 4.3.2(a)	Fold change of miR-103 in 8505C holoclones
Table 4.3.2(b)	Fold change of the miR-302 family members in 8505C holoclones
Table 4.3.2(c)	Fold change of miR-26 family members in TPC-1 holoclones
Table 4.3.2(d)	Fold change of the miR-302 family members in TPC-1 holoclones
Table 4.3.3(a)	Fold change of miR-103a transcripts in ES-2 holoclones
Table 4.3.3(b)	Fold change of miR-26 family members in A2780 holoclones
Table 4.3.4(a)	Fold change of miR-103a transcripts in HT-29 holoclones
Table 4.3.4(b)	Fold change of a miR-26a transcript in COLO320 holoclones
Table 4.3.5(b)	Fold change of the miR-302 family 8505C and TPC-1 thyroid holoclones
Table 4.5.1(a)	Fold change of SNORD114 and SNORD113 in SK-Mel 28 holoclones
Table 4.5.1(b)	Fold change of SNORD85 and SNORD71 in COLO794 holoclones
Table 4.5.2(a)	Fold changes of SNORD114 transcripts in 8505C holoclones
Table 4.5.2(b)	Fold changes of SNORD85 in TPC-1 holoclones

Table 4.5.3(a)	Fold changes of a SNORD114 transcript in ES-2 holoclones
Table 4.5.3(b)	Fold change of SNORD85 and SNORD71 in A2780 holoclones
Table 4.5.4(a)	Fold change of SNORD85 and SNORD71 in COLO320 holoclones
Table 5.1(a)	Methylation results of gene panel in BRAF wild type and BRAF V600E mutant thyroid cancer cell lines
Table 5.3.3(b)	Log ₂ fold changes of consistently overexpressed snoRNAs in SK-Mel 28, 8505C and TPC-1 holoclones

VIII. List of Abbreviations

The following abbreviations are used:

AKT	V-akt Murine Thymoma Viral Oncogene Homolog 1
ALDH1	Aldehyde Dehydrogenase 1
APC	Adenomatous Polyposis Coli
ATC	Anaplastic Thyroid Cancer
BRAF	V-Raf Murine Sarcoma Viral Oncogene Homolog B1
BRCA1	Breast Cancer 1 Early Onset
cDNA	Complementary DNA
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CO ₂	Carbon Dioxide
CRC	Colorectal Cancer
CSC	Cancer Stem Cell
CTLA4	T-Lymphocyte-Associated Antigen 4
DHH	Desert Hedgehog
DITC	Dacarbazine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
EDTA	Ethylenediaminetetraacetic Acid
EMT	Epithelial-Mesenchymal Transition
ES	Embryonic Stem Cell
FAM	6-carboxyfluorescein
FBS	Fetal Bovine Serum
FDA	Food & Drug Administration
FIGO	International Federation of Gynecology and Obstetrics
gDNA	Genomic DNA
GI	Gastrointestinal
GTF	General Transfer Format
GSTP1	Pi-Class Glutathione S-Transferase Gene

HHT	Huntington
IHH	Indian Hedgehog
lncRNA	Long Non-coding RNA
MAPK-ERK	Mitogen-Activated Protein Kinase-Extracellular Signal-Regulated Kinase
miRNA	MicroRNA
MgCl ₂	Magnesium Chloride
MLH1	MutL Homologue
MMR	Mismatch Repair
mRNA	Messenger Ribonucleic Acid
Oct4	Octamer-Binding Transcription Factor 4
PCR	Polymerase Chain Reaction
piRNA	Piwi-interacting RNA
PTC	Patched
PTEN	Phosphatase and Tensin Homolog
RAS	Rat Sarcoma Viral Oncogene Homolog
RISC	RNA Induced Silencing Complex
RNA	Ribonucleic Acid
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute
rRNA	Ribosomal RNA
SEM	Standard Error Mean
SHH	Sonic Hedgehog
siRNA	Small Interfering RNA
SMO	Smoothed
SNAI2	Snail Homolog 2
snoRNA	Small nucleolar RNA
snoRNP	Small Nucleolar Ribonucleoprotein
SNP	Single Nucleotide Polymorphism
SOLiD	Sequencing by Oligonucleotide Ligation and Detection
TBE	Tris Borate EDTA
TE	Tris EDTA

TGF- β	Transforming Growth Factor Beta
TKR	Tyrosine Kinase Receptor
TNM	TNM Classification of Malignant Tumours
TP53	Tumour Protein 53
tRNA	Transfer RNA
TSH	Thyroid Stimulating Hormone
UTR	Untranslated Region
UV	Ultraviolet

A cell is regarded as the true biological atom

George Henry Lewes (1817 – 1878)

Chapter 1

Introduction

1.1 Cancer Genetics

The human genome is a complex sequence of 3 billion base pairs distributed unevenly across 22 autosomal chromosome pairs and a sex determinative pair, X for females or Y for males (Genome Reference Consortium, 2012). At present it is thought the human genome within every human cell codes for up to 20-25,000 distinct protein-coding genes, much lower than what was expected until completion of the Human Genome Project in 2003 (International Human Genome Sequencing Consortium, 2004). The rest of the sequences make up non-coding RNA genes, regulatory sequences, introns, and non-coding DNA. Because the human genome is the result of evolution over millions of years, its genetic mechanism has been finely tuned and like other organisms is constantly changing in response to environmental pressures.

Despite the complexity of the human building blocks of life, errors are common and while most are corrected, some carry on unrepaired and have no phenotypic effects whereas others can be detrimental with severe consequences. Many disorders affecting many different components of human biology occur as a result of genetic alterations. This can range from mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene causing the predominantly pulmonary disorder Cystic Fibrosis to Huntington's disease, a neurodegenerative genetic disorder caused by a mutation in the Huntington (HTT) gene which detrimentally affects the muscular system (Nayak, 2012; Kalathur et al., 2012).

Similarly, genetics plays a role in many forms of human cancer and can result from inherited "germline" mutations passed down from previous generations or acquired "sporadic" mutations as a result of environmental factors such as harmful ultraviolet (UV) rays or tobacco (Rasool et al., 2012). Therefore, cancers can be classed as sporadic or familial depending on the type of genetic abnormality. Different types of genes that carry out different functions can be linked to cancer and can fall into different categories.

Tumour suppressor genes have a protective function and control growth rate in cells by monitoring turnover, repairing mismatched DNA, and by playing a role in apoptosis (Brand et al., 2000).

When a tumour suppressor gene acquires a mutation, be it a germline or sporadic one, cells can bypass nuclear checkpoint control and growth is uncontrolled (Morandell & Yaffe, 2012). These cells can go on to proliferate and form a tumour which may stay localised or metastasise to distant sites. Breast Cancer 1 Early Onset (BRCA1) is an example of a tumour suppressor gene and germline mutations can increase a woman's risk of developing hereditary breast or ovarian cancers (Liede et al., 2004). The BRCA1 encoded protein forms the BRCA1-associated genome surveillance complex with other proteins to maintain genomic stability by recognising abnormal DNA structures and damaged DNA (Jhanwar et al., 2003). Another tumour suppressor gene commonly mutated in human cancers is the gene that encodes tumour protein 53 (TP53) with most mutations occurring sporadically as a result of environmental factors. Wild type p53 functions as a tumour suppressor in many tumour types; by inducing growth arrest or apoptosis depending on the physiological circumstances and cell type (Pei et al., 2012). Many tumours carry missense mutations with the affected cells accumulating excessive amounts of the mutant p53 protein. Evidence suggests that p53 mutations also endow the mutant protein with new activities that can contribute actively to various stages of tumour progression and to increased resistance to anticancer treatments (Oren & Rotter, 2010).

Oncogenes are genes that turn a healthy normal cell into a cancerous one and are generally more susceptible to sporadic mutations rather than germline. The Ras superfamily of small GTPase genes comprises a group of molecular switches that regulate adversity of cellular functions and are considered oncogenes (Macara et al., 1996). Ras receives signals from a protein complex bound to an activated cell surface receptor and induces activation of downstream Raf proteins in the MAPK-ERK signalling pathway. As this pathway is vital in the processes of cell growth and division, constitutively active Ras signalling, as a result of mutations can ultimately lead to cancer (Goodsell, 1999). Ras is the most commonly mutated oncogene in human cancer with mutations that permanently activate Ras, found in 20-25% of all human tumours and up to 90% in some specific types of cancer (Downward, 2003).

DNA mismatch repair (MMR) genes, function by recognizing and repairing erroneous DNA insertions, deletions and other mis-incorporations that may occur during the processes of DNA replication and recombination (Iyer et al., 2006). They also function by repairing some forms of DNA damage that occur in cells that must be repaired before cell division can take place and mutations can be germline or sporadic (Larrea et al., 2010). If an MMR gene is mutated it cannot carry out its function and repair errors that may be present in DNA sequences, resulting in division of cells carrying mutations that may lead to cancer, especially in the case of tumour suppressor genes or oncogenes. Mutations in the DNA MMR gene human MutL homologue (hMLH1) has been shown to be associated with hereditary non-polyposis colon cancer (Bronner et al., 1994).

Despite everything that is currently known about the molecular functioning of cancer associated genes, many cancers cannot be linked with a single specific gene and it is likely that multiple different genes are involved in the development of most cancers.

1.2 Melanoma

The potential for metastasis and treatment resistance associated with melanoma makes it one of the most aggressive forms of human cancers (Tsao, 2012). Melanoma arises when melanocytes located mainly in the epidermis that produce the dark pigment melanin, become malignantly transformed (Volkovova et al., 2012). Incidence rates are on the increase in Ireland, possibly as a result of the social and environmental consequences of “package holiday” sun exposure and the increasing use of artificial tanning beds which subject users to intense UV ray exposure (Kanavy & Gerstenblith, 2011). The disease is most common in Caucasian populations living in sunnier climates such as North America, Australia, Latin America and Northern Europe, as a result of higher ultraviolet light (UV) exposure (Parkin et al., 2002). Harmful UV rays can have many detrimental effects such as DNA damage inducing genetic mutations including loss of function of key tumour suppressor genes and activation of proto-oncogenes, oxidative stress, immunosuppression and inflammation which all contribute to the development of melanoma (DeGrujil et al., 2012; Yoshiki et al., 2012).

The outcome for patients with melanoma depends on how advanced the cancer is at diagnosis and also the subtype. Superficial Spreading Melanoma, nodular, and lentigo maligna make up 90% of all diagnosed malignant melanomas with acral lentiginous and some other rare types together making up the other 10% (Saldanha et al., 2006). For many melanoma patients with early stage cancers, excisional biopsies or surgery may be sufficient to eradicate disease and halt cancer progression. However more advanced melanomas may require radiotherapy and chemotherapy drugs such as dacarbazine (DTIC) and vinblastine or cisplatin (Su et al., 2011).

Current treatments for advanced melanoma include biological therapies such as Vemurafenib and the monoclonal antibody Ipilimumab. Vemurafenib (marketed as Zelboraf) functions by targeting cells harbouring the BRAF V600E mutation, which is present in up to 70% of melanoma patients (Sala et al., 2008). The drug interrupts the B-Raf/MEK transition step in the B-Raf/MEK/ERK pathway causing apoptosis in cells with the V600E mutation (Ravnan & Matalaka, 2012). Ipilimumab is targeted at cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) (Trinh & Hwu, 2012). It functions by blocking the CTLA-4 inhibitory signal, prolonging anti-tumour immune responses against cells which can have impaired antigen presentation mechanisms that can allow them to evade immune system surveillance (Tarhini & Iqbal, 2010).

1.3 Thyroid Cancer

The thyroid is one of the largest glands in the human endocrine system and is located in the throat. It functions by releasing the tyrosine-based thyroid hormones triiodothyronine (T_3) and thyroxine (T_4), which act on cells throughout the body to control cell growth and metabolism and maintain tissue homeostasis (Danzi & Klein, 2012). Many benign disorders are linked to dysfunction within the thyroid gland and consequent hormone imbalance. However, thyroid cancer is the most common endocrine malignancy and presents in several morphologic variants. The incidence of thyroid cancer is increasing in Ireland and globally and the rate of increased incidence is particularly noteworthy among females (The National Cancer Registry Ireland, 2012).

Like melanoma, there are different phenotypic and morphologic sub-types of thyroid cancer with ranging prognoses. Most commonly occurring are Papillary, Follicular, Medullary and Anaplastic thyroid cancers (Shih et al., 2012). Papillary thyroid cancer has been linked with mutations in *ret* and *BRAF* oncogenes while follicular thyroid cancer is typically associated with *RAS* mutations. At present, the main treatments for thyroid cancer are surgery and radiotherapy supplemented with thyroid hormones (recombinant human TSH) (Duntas et al., 2003). Total thyroidectomy or lobectomy usually involves resection of any associated lymph nodes in the neck (Derbel et al., 2011). Targeted radiotherapy that uses a radioactive form of Iodine 131 ($I-131$) is used for papillary or follicular thyroid cancers (Silberstein, 2012). Chemotherapy is not generally used as a first treatment for thyroid cancer because surgery and radiotherapy tend to be more effective. Anaplastic thyroid cancer (ATC) is the most aggressive form of thyroid cancer and has a very poor prognosis with the overall median survival limited to months (Pasiaka, 2003). It displays significant propensity for widespread metastasis and exhibits resistance to conventional cancer treatments (Cornett et al., 2007). ATC displays poor survival rates with a median survival of less than 6 months (Lim et al., 2012). In the future, molecular-based gene targeted therapies currently being investigated, may offer hope and improved prognosis to patients suffering from advanced forms of thyroid cancer.

1.4 Ovarian Cancer

Ovarian cancer has the worst prognosis among gynaecologic malignancies, with the majority of patients succumbing to their disease within 5 years of diagnosis. Patients typically present with advanced disease, which while initially receptive to chemotherapeutic intervention, typically recurs and recurrent disease is generally chemo-resistant (Davidson et al., 2012). In Ireland, ovarian cancer is the fourth most common cancer in women and each year over 300 women are diagnosed with the disease (The Irish Cancer Society, 2012). More than 90% of ovarian cancers are classified as epithelial and are believed to arise from the ovarian epithelium but a recent school of thought advocates for carcinogenesis that is initiated in the fallopian tubes (Piek et al., 2008). Borderline tumours are not truly malignant given their low rate of growth and low potential to invade or metastasize however; they retain the ability to become malignant subject to molecular progression (Tinelli et al., 2006). They are staged using the same FIGO system used for ovarian cancers (Baak et al., 1986).

Abnormal cells can sometimes break away from borderline tumours and colonise elsewhere in the body, usually within the abdomen.

Like many different types of human malignancies, treatments for ovarian cancer differ depending on stage at presentation. A total hysterectomy is recommended for stage 1 ovarian cancer (Kim et al., 2012). Biopsies are taken from the abdomen and also from areas where the malignancy might have locally spread such as pelvic lymph nodes, the diaphragm and the tissue lining the abdomen and pelvis. After surgery, adjuvant chemotherapy may be chosen for stage 1c or a high grade (grade 3) (Burges & Schmalfeldt, 2011). Stage 2, 3 or 4 cancers are more advanced and usually require surgery and adjuvant chemotherapy. In some cases it may not be possible to remove all of the cancer surgically, so patients undergo neoadjuvant chemotherapy before surgery to try and shrink the cancer and make it easier to debulk (Menczer et al., 2011). As with earlier stage disease, biopsies are taken at locations where the cancer is most likely to have spread, including the diaphragm, the lining of pelvic and abdominal cavities and lymph nodes in the pelvis and abdomen (Heinen & Pérez, 2012). For very advanced ovarian cancers, surgery may not be an option so chemotherapy is administered to shrink and slow the tumour growth and in some cases radiotherapy to relieve symptoms.

1.5 Colorectal Cancer (CRC)

Colorectal cancer (CRC), also known as colon cancer or bowel cancer occurs in the large intestine (colon) or the rectum (end of the colon) and if left untreated, can invade into the muscle layers underneath, and then through the bowel wall (Zlobec & Lugli, 2009).

According to the Irish Cancer Society, CRC is the second most common cancer in Ireland. In 2009, there were 2,271 people diagnosed with the disease and it is also the second most common cause of cancer death in Ireland (The Irish Cancer Society, 2012)

CRC arises through cumulative effects of inherited genetic predispositions and environmental factors (Al-Sohaily et al., 2012). Cancer of the GI tract frequently occurs as a result of mutations, inherited or acquired, in the Wnt signalling pathway. The Wnt signalling pathway is responsible for controlling cell-cell communication and is vital for embryonic

development and is dysregulated in many types of cancer (Lorenowicz & Korswagen, 2009). The Adenomatous Polyposis Coli (APC) gene is one of the most commonly mutated genes in CRC and produces the APC protein which prevents an over-accumulation of β -catenin protein (Grover et al., 2012). β -catenin can accumulate in the cytoplasm in the absence of APC and translocates to the nucleus, where it activates gene transcription. In many cases of APC-mediated CRC, synthesis of truncated APC products occurs causing the stabilization of β -catenin resulting in genes involved in cell renewal, differentiation and survival to become over expressed (Chandra et al., 2012). Like many other cancer types, a mutation in the TP53 gene can facilitate malignant transformation (Rodrigues et al., 1990). The TP53 gene encodes the p53 protein which functions by monitoring cell division and kills cells that possess Wnt pathway defects.

Rat sarcoma viral oncogene homolog (RAS) genes are amongst the most frequently mutated genes in human cancers and in particular, KRAS is mutated in 40-50% of colorectal cancers (Poulogiannis et al., 2012). RAS encoded proteins such as KRAS, function in the MAPK-ERK signalling pathway as a small GTPase that acts as a molecular switch between homeostatic inputs and downstream signal transduction proteins (Poulogiannis et al., 2012). A single amino acid substitution is responsible for an activating mutation in KRAS which change the function of normal proto-oncogenic RAS proteins to oncogenic RAS proteins. This alters the expression of downstream transcription factors and subsequently stimulates cell proliferation whilst repressing apoptosis (Rusconi et al., 2012). The transforming protein that results is implicated in various malignancies, including lung adenocarcinoma, ductal carcinoma of the pancreas and in conjunction with other genetic abnormalities in the colorectal carcinoma sequence (Jaas, 2006).

It is now commonly known that epigenetic changes can also play a role in the development of CRC and in particular changes to mismatch repair genes. Microsatellite instability (MSI) is the molecular fingerprint of the deficient mismatch repair system, which characterizes ~15% of colorectal cancers (Sinicrpe & Sargent, 2012). MSI can frequently develop from epigenetic silencing of MLH1 in sporadic tumours and this occurs when hypermethylation of CpG islands takes place in an MMR gene's promoter region.

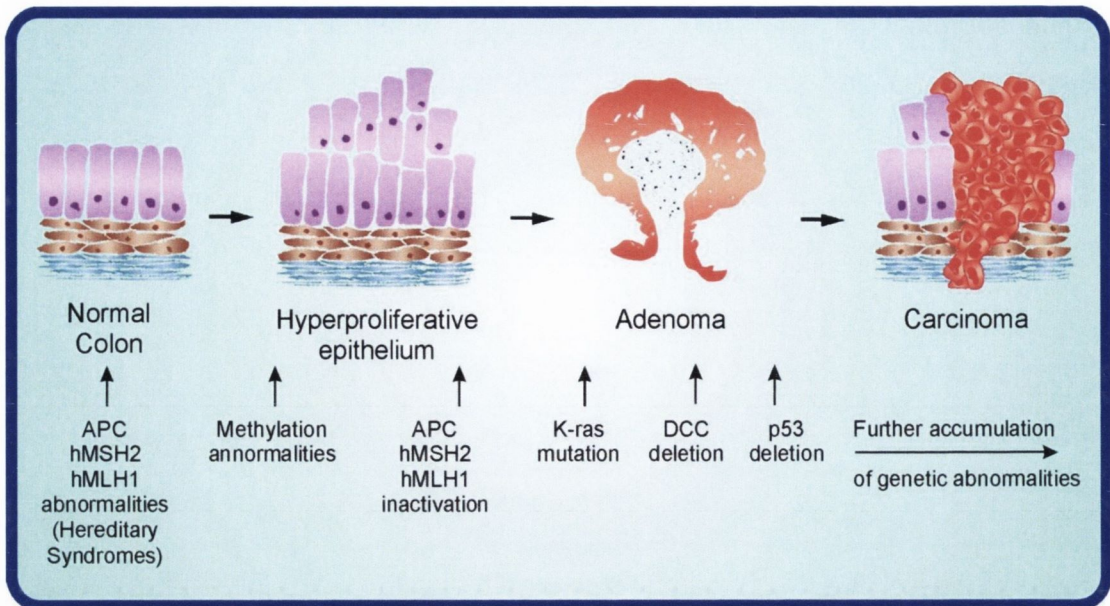


Figure 1.5: Illustration depicting the accumulation of genetic abnormalities that take place in the progression of normal colon cells to adenoma and finally to colon carcinoma (Pathology Outlines, 2012). In figure 1.5, mutations in genes such as APC or epigenetic changes in MMR genes like hMLH1 facilitate the transformation of normal colon cells into hyper proliferative epithelia. Inactivation of other MMR genes such as MutS homolog 2, colon cancer, non-polyposis type 1 (hMSH2), by DNA methylation contributes to adenoma formation. Mutations in K-ras, p53 deletions and other genetic abnormalities play a pivotal role in the progression of colon adenoma into colon carcinoma.

In cancer diagnosis and management TNM staging is used to determine how advanced a CRC is at presentation and what treatment options will be chosen (Terhaar sive Droste et al., 2010). Like most cancers, when CRC is detected early surgery can be curative but when the tumours are more advanced and metastasis is present, treatments are generally aimed at prolonging life and palliative care (Cunningham et al., 2010).

1.6 The BRAF gene

V-Raf murine sarcoma viral oncogene homolog B1, otherwise known as BRAF, encodes a protein belonging to the raf/mil family of serine/threonine protein kinases that are activated by RAS in a GTP-dependent manner (Vakiani & Solit, 2011). Located on

Chromosome 7q34, B-Raf is not ubiquitously expressed in mammalian systems but is expressed at higher levels in tissues and certain cell types such as thyroid follicular cells (Kondo et al., 2007).

The encoded 84kDa B-Raf protein is involved in the regulation of the MAPK-ERK signalling pathway, which affects cell growth, proliferation and apoptosis (Brzezianska & Pastuszek-Lewandoska, 2011). Membrane bound tyrosine kinase receptors (TKR) are activated by ligands (mitogens or growth factors) and recruit the protein complex to the tyrosine domains of the receptor. Phosphorylation events induce the conversion of inactive Ras Guanine DiPhosphate (RasGDP) into active Ras Guanine TriPhosphate (RasGTP) by the protein SOS. The conversion of inactive Ras into its active state facilitates the release of B-raf from the 14-3-3 regulatory protein and subsequent activation of downstream kinases. The B-Raf protein activates MAPKK (MEK) and MEK phosphorylates MAPK (ERK), each of which is required for maintaining responsiveness to growth factor stimulation. Mutations in the MAPK-ERK pathway are common, occurring in up to 30% of all human cancers with BRAF gene mutations thought to be specifically implicated in up to 8% of all solid tumours (Cantwell-Dorris et al., 2011). The most frequent mutation observed in BRAF is a single point mutation located at codon 600 in exon 15. The V600E mutation results from a transversion mutation T1799A causing an amino acid substitution of Valine (V) to Glutamic Acid (E) in the resultant protein (Zafon, & Obiols, 2009). This confers constitutive activity to the BRAF kinase which normally requires phosphorylation by upstream Ras kinases to activate it and initiate cell signalling.

The BRAF V600E mutation has been detected with different prevalence rates in papillary thyroid, melanomas, colorectal and ovarian cancers amongst others and so represents an important target for therapies to combat tumour development (Cahill et al., 2007; Kumar et al., 2004; Estep et al., 2007). Highlighting this observation, the FDA approved Vemurafenib in 2011 for the treatment of late stage melanoma. The drug functions by interrupting the B-Raf/MEK step in the MAPK-ERK pathway where the BRAF gene harbours the V600E mutation (Luke & Hodi, 2011). Approximately 50% of advanced melanomas possess the BRAF V600E mutation, and so Vemurafenib is only effective in melanoma patients that

harbour this specific mutation (Nissan & Solit, 2011.). However, recent clinical trials by Plexicon have explored the combination of Vemurafenib and the MEK inhibitor GDC-0973 in melanoma patients. In the future, targeting this pivotal signalling pathway may be used in conjunction with other molecular based therapies to improve treatment of other aggressive cancers.

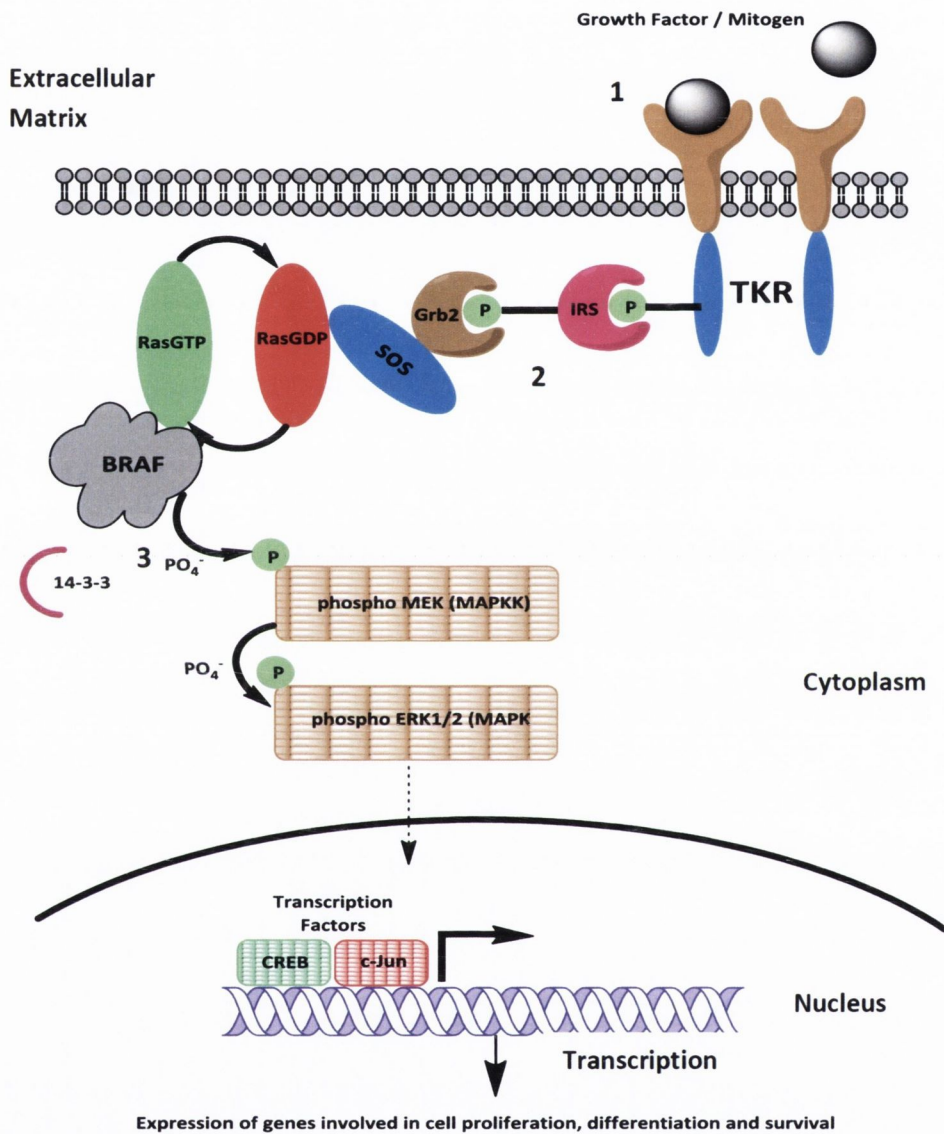


Figure 1.6: Illustration depicting the MAPK-ERK cell signalling pathway. MAPK-ERK signalling pathway is activated upon ligand binding to a membrane bound TKR dimer (1). The activated receptor recruits a protein complex (2) to its intracellular domain and phosphorylation activates the Ras kinase. Once activated, Ras facilitates the release of BRAF from the 14-3-3 regulatory protein and activation (3). BRAF subsequently

activates downstream MEK by phosphorylation followed by ERK1/2 (4) which then translocates to the nucleus. Here it interacts with transcription factors to initiate transcription of genes involved in cell proliferation and survival (5).

1.7 Cancer Stem Cells

A cell's potency signifies its ability to differentiate into different cell types. The three main classes of cell potencies that occur in a typical cell lineage are Totipotency, Pluripotency and Unipotency (Wagers & Weissman, 2004). A totipotent cell such as a zygote has the ability to divide and produce all the differentiated cells in an organism. A single totipotent cell arises from the fertilization of an egg by a sperm and divides to produce multiple identical totipotent cells (Ratajczak et al., 2008). After four divisions these totipotent cells differentiate into cells that make up the inner cell mass of a blastocyst. This mass is composed of pluripotent stem cells that have the potential to differentiate into any of the three germ layers made up of unipotent cells: endoderm, mesoderm, or ectoderm (Ratajczak et al., 2008; Kooreman & Wu, 2010). Endoderm germ layers can be found in the lungs and gastrointestinal tract, mesoderm layers make up muscle and blood whereas ectoderm layers produce epidermal tissues and nervous system components (de Santa Barbara et al., 2003; Kelly, 2012; Adams & Bronner-Fraser, 2009).

Stem cells are distinguishable from other cell types by two important characteristics: (1) they are undifferentiated cells capable of renewing themselves through cell division and (2) they can be induced to become tissue- or organ-specific cells with special functions under certain physiological or experimental conditions (de Cuevas & Matunis, 2011). Some organs have regularly dividing stem cells to facilitate damage repair such as the kidney and heart whereas in other organs, stem cells will only divide under special conditions (Morigi et al., 2004; Unno et al., 2012) Two types of stem cells exist in humans: embryonic stem cells and adult "somatic" stem cells (Fortier, 2005). Embryos at the blastocyst stage approximately 14 days post fertilization are rich in embryonic stem (ES) cells which have nearly unlimited self-renewal capacity and are required to differentiate into the different cell types needed for growth and development (Wobus & Boheler, 2005). Transcription factors such as Oct4 and NANOG are highly expressed in ES cells as they are required for the regulation of genes

needed to maintain ES cells in an undifferentiated, pluripotent state capable of self-renewal (Chan et al., 2011). Adult stem cells are undifferentiated cells, found sequestered among differentiated cells in tissues or an organ that self-renew and differentiate into some or all of the major specialized cell types of that specific tissue or organ. The main role of adult stem cells is one of maintenance and damage repair (Williams & Hare, 2011).

Research has suggested that small populations of cancer cells with characteristics similar to normal ES cells exist in solid tumours and cancer cell lines (Bomken et al., 2010). Currently, two schools of thought exist among the scientific community with regard to progression of solid tumours. The stochastic hypothesis states that any tumour cell has the capacity to form and maintain the tumour mass (Schwarz-Cruz-y-Celis & Meléndez-Zajgla, 2011). The more recently formed hierarchical hypothesis suggests the bulk of cancer cells have limited regenerative capability, but a subpopulation can consistently proliferate and are tumourigenic upon transplantation into animal models which is considered the gold standard assay (Sakashita et al., 2007). This hypothesis may explain why drugs can lead to rapid tumour shrinkage and initial optimism, but poor long-term survival with disease recurrence or development of chemoresistance. Cancer stem cells are innately more aggressive and have resilience against conventional anti-cancer agents (Rosen & Jordan, 2009). Studies suggest tumour progression associated Epithelial-Mesenchymal Transition (EMT) in cancer cells may confer them stem-like properties (Mani et al., 2008).

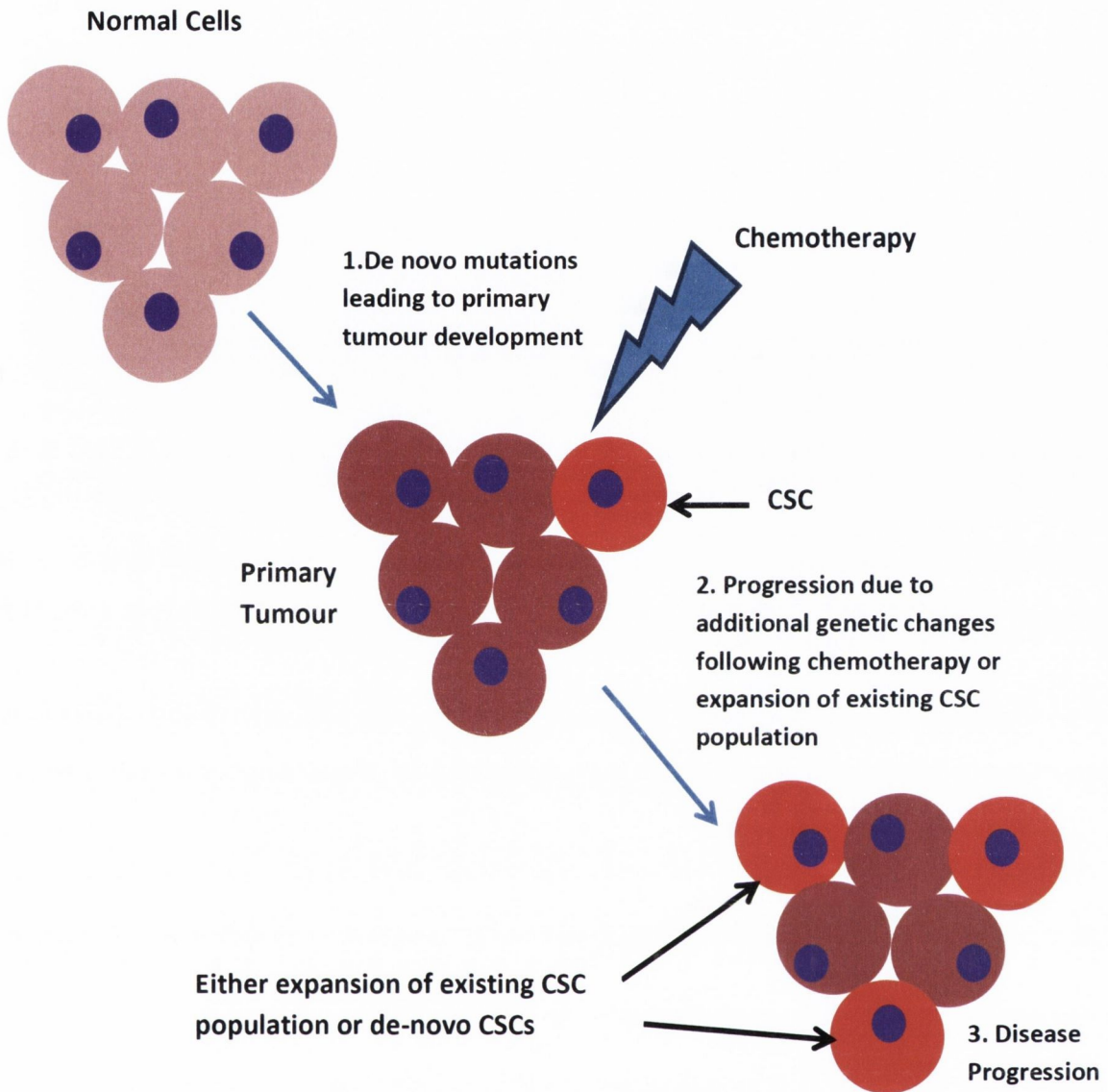


Figure 1.7: Illustration depicting the current hypothesis of the involvement of cancer stem cells in tumour progression. Normal cells acquire de novo mutations (1) leading to the development of a primary tumour which contains a small subpopulation of CSCs. Within a tumour, a small proportion of CSCs exist which possess capacity to self-renew and to cause heterogeneous lineages of cancer cells that comprise the tumour. These cells exhibit the properties of stem cells and are tumourigenic. It is suggested that they cause revival of cancers after treatment with conventional methods such as chemotherapy. Treatments such as chemotherapy only target the bulk of the tumour but the hypothesis suggests that CSCs have greater chemoresistance potential and may

survive. By acquiring additional genetic changes as a result of chemotherapy (2), these cancer stem cells may give rise to differentiated clones which could account for tumour recurrence (3) and so represent a possible novel mechanism for potential treatment of various cancer types. To date, cancer stem cells are largely uncharacterised, so investigating their genetic profiles may yield novel targets which to which therapeutic interventions may be aimed at.

1.8 CSC Morphologies

Model studies into the importance of cancer stem cells in tumour (re)population has shed light on 3 morphologically distinguishable types observed in cell lines; Meroclones, Paraclones and Holoclones (Kalirai et al., 2011). Research suggests meroclones and paraclones contain more mature and differentiated cells whereas holoclones have the strongest self-renewal ability out of the 3 types and thus the greatest potential for tumour re-population (Pfeiffer & Schalken, 2009). Holoclones are groups of small tightly packed cells with the greatest replicative capacity. These tightly packed clones of cells generally contain both stem cells and progenitors. Paraclones are groups of larger loosely packed cells with a very short replicative life span (Li et al., 2008). However meroclones contain a combination of cells of different proliferative potential and is a transitional stage between a holoclone and paraclone (Li et al., 2008). Holoclones have a high colony-forming efficiency whereas paraclones and meroclones have a much lower colony-forming capacity (Larcher et al., 2007). Previous studies have established that only a small percentage of cells (ranging from 0.1-0.3%) have the ability to generate holoclones within cell lines however the majority of cells have the ability to form paraclones and meroclones (Li et al., 2008).



Figure 1.8: Green Fluorescent Protein images of the 3 distinct colony morphologies (holoclone, meroclone and paraclone) observed in the prostate cancer cell line PC-3 (Li et al., 2008).

Recently there has been an upsurge in research into the molecular characteristics of CSCs with data emerging on intracellular and cell surface proteins whose expression changes in CSC phenotypes. It is clear that a greater understanding of these processes is imperative to understand carcinogenesis, disease progression, recurrence and chemoresistance.

1.9 Stemness Genes & Associated Pathways

Certain genes are required by stem cells to maintain in an undifferentiated, pluripotent state capable of self-renewal. The expression of many of these genes has been investigated in ES cells but also CSCs. They represent the range of properties that cancer stem cells possess; epithelial-mesenchymal transition, self-renewal and proliferation and highlights the need to learn more about them as they may offer a potential for future therapeutic interventions.

NANOG

The NANOG gene encodes a 305 amino acid homeobox protein that functions as transcription regulator involved in inner cell mass and ES cell proliferation and self-renewal (Miyanari & Torres-Padilla, 2012). The NANOG gene is believed to be a key factor in maintaining pluripotency and works alongside other genes such as SOX2 and Oct4 to establish ES cell identity (Wang et al., 2012). Overexpression of NANOG in human embryonic stem cells enables their propagation for multiple passages during which the cells remain pluripotent. Knockdown experiments of NANOG have been shown to promote differentiation, thereby demonstrating a role for these factors in human embryonic stem

cell self-renewal (Chen et al., 2009). Recent studies have shown that NANOG is highly up regulated in a variety of different cancer stem cells indicating that these cell subpopulations may have pluripotent characteristics and have self-renewal capabilities (Guo et al., 2012).

Oct4

Oct4 (octamer-binding transcription factor 4) also known as POU5F1, is a homeodomain transcription factor encoded by the POU5F1 gene (Sterneckert et al., 2012). Oct4 has a critical involvement in the self-renewal of undifferentiated embryonic stem cells and is frequently used as a marker for undifferentiated cells (Hammachi et al., 2012). Severe under or over-expression of Oct4, results in differentiation and so its expression level must be tightly regulated. Similar to NANOG, gene knockdown of Oct4 has shown to promote differentiation, indicating its important role in human embryonic stem cell self-renewal (Zafarana et al., 2012). New research has shown that Oct4 is commonly up regulated in cancer stem cells in conjunction with other stem cell markers such as ALDH1 and CD133 (Kim & Nam, 2011).

ALDH1

Aldehyde dehydrogenase 1 family (ALDH1) encodes a protein belonging to the aldehyde dehydrogenase family which are the next enzymes after alcohol dehydrogenase in the major pathway of alcohol metabolism (Douville et al., 2009). ALDH1 has been identified as an important enzyme in the protection of normal hematopoietic stem cells, and is now also widely used as a marker to identify and isolate various types of normal stem cells and CSCs (Ma & Allan, 2011). It has been observed that other ALDH isozymes in addition to ALDH1, are important for multiple biological activities including drug resistance, cell differentiation, and oxidative stress response and therefore offers an attractive potential biomarker for the identification of cancer stem cells in cancer patients (Moreb, 2008).

SNAI2

Located on chromosome 8, the snail homolog 2 (SNAI2) gene also known as Slug, encodes a zinc finger protein that is a member of the Snail family of C2H2-type zinc finger transcription factors (Shirley et al., 2010). SNAI2 has demonstrated to play a specific and critical role for in the pathogenesis of mesenchymal tumours as it functions during epithelial-mesenchymal transitions (EMT) (Cobaleda et al., 2007). SNAI2 controls key aspects of stem cell function in both mouse and humans, suggesting that similar mechanisms control normal development

and cancer stem cell properties (Pérez-Mancera et al., 2005). Studies have reported that loss of SNAI2 in different cancer types severely inhibits invasion and decreases cell proliferation in aggressive cancer types (Becker et al., 2002).

E-Cadherin

The CDH1 gene encodes the Cadherin 1, type 1, E-cadherin (epithelial) protein also known as E-cadherin and is a member of the cadherin superfamily and functions as a calcium-dependent cell-cell adhesion glycoprotein (Tian et al., 2011). Loss of E-cadherin function or expression has been implicated in cancer progression and metastasis (Buda & Pignatelli, 2011). Snail has been shown to repress E-cadherin and is implicated in many different cancer types as it's down regulation decreases the strength of cellular adhesion within a tissue, resulting in an increase in cellular motility (Becker et al., 2002). This in turn may allow cancer cells to cross the basement membrane and invade surrounding tissues. Studies investigating E-cadherin function in murine stem cells observed that E-cadherin is highly expressed in mouse embryonic stem cells, and interference with E-cadherin expression causes differentiation (Li et al., 2012).

SHH

The Sonic Hedgehog (SHH) gene encodes a protein which is a member of the hedgehog mammalian signalling pathway family along with other family members desert hedgehog (DHH) and Indian hedgehog (IHH) (Lauth & Toftgård, 2011). It functions as a ligand that induces signalling in the pathway upon binding to the patched (PTC) receptor, which functions in association with smoothened (SMO), to activate the transcription of target genes (Tao et al., 2011).

Despite this pathway being critical in processes such as central nervous system patterning and vertebrate organogenesis in embryo development, it still remains important throughout adulthood (Choy & Cheng, 2012). Studies have shown SHH is involved in the regulation of cell division in adult stem cells but also plays a role in CSCs and so has been implicated in development of some cancers (Song et al., 2011). SHH has been shown to play a role in pancreatic CSCs in conjunction with cell surface markers CD44 and CD24 by conferring properties of self-renewal and multi-lineage differentiation (Lee et al., 2008).

TGF- β

TGF- β is a protein involved in the well characterised Transforming Growth Factor Beta (TGF- β) signalling pathway involved in many human cellular processes including cell growth, cell differentiation and apoptosis (Sun, 2004).

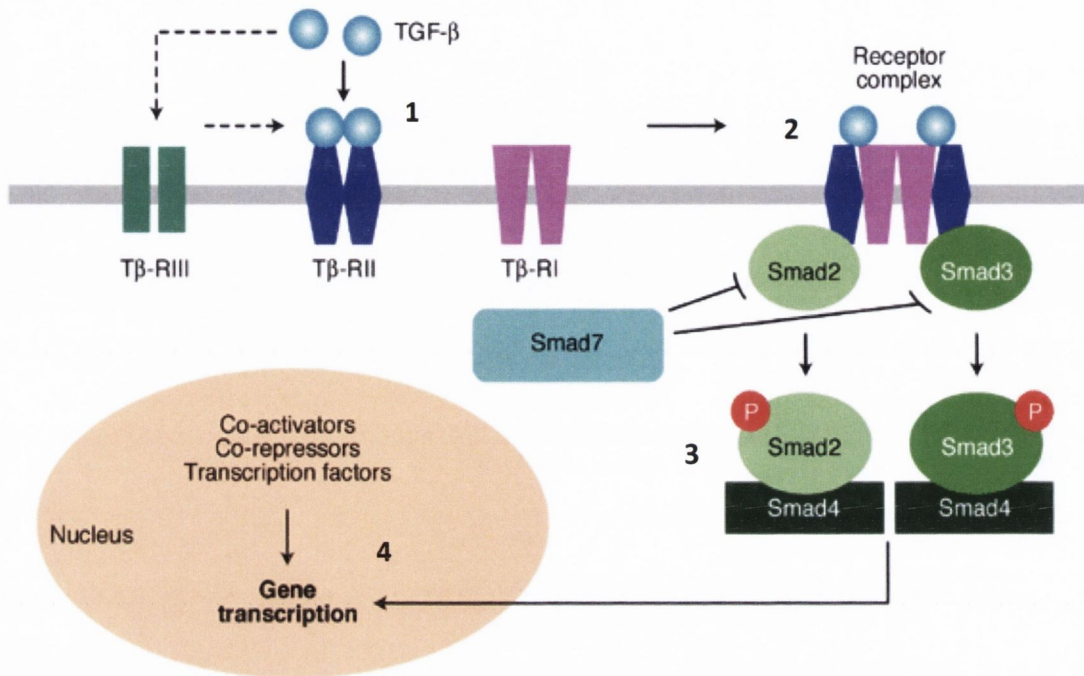


Figure 1.9: A schematic of the Transforming Growth Factor beta (TGF- β) signalling pathway. Figure 1.9 depicts the role played by TGF- β in the TGF- β cell signalling pathway. TGF- β functions as a ligand and binds to the receptor TGF- β R2 (1) which then recruits and phosphorylates TGF- β R1, leading to activation of Smad2 and Smad3 by phosphorylation (2). Activated Smad2 and Smad3 form heterodimers with Smad4 (3) and translocate to the nucleus and together with co-activators, co-repressors and other transcription factors, the Smad complex regulates gene expression (4) (Verrecchia & Mauviel, 2002).

The TGF- β pathway has been shown to function as both a tumour suppressor and promoter and play a role in directing ES cell fate, however its involvement is poorly understood so far (Galvin-Burgess & Vivian, 2011). Recent studies have highlighted the function of TGF- β regulated noncoding RNAs in driving EMT and promoting cancer stem cell self-renewal (Singh & Settleman, 2010).

β-Catenin

Catenin (cadherin-associated protein), beta 1, 88kDa (β-Catenin) is encoded by the CTNNB gene and is part of a complex of proteins that constitute adherens junctions (AJs) (Baum & Georgiou, 2011). AJs are necessary for the creation and maintenance of epithelial cell layers by regulating cell growth and adhesion between cells. The β-Catenin protein also anchors the actin cytoskeleton and may be responsible for transmitting the contact inhibition signal that causes cells to stop dividing once the epithelial sheet is complete (Valenta et al., 2012). β-catenin is the central nuclear effector of the Wnt signalling pathway. Upon activation of the pathway, β-catenin translocates to the nucleus where it induces transcriptional activation of target genes that function during both normal and malignant development. Constitutive activation of the Wnt pathway leads to inappropriate nuclear accumulation of β-catenin and gene transactivation, an important step in cancer progression (Jamieson et al., 2012). Numerous inhibitory agents targeting specific signalling pathways, such as the Wnt/β-catenin pathway, are currently under evaluation in early clinical trials. By targeting tumour radioresistance, one of the features of CSCs, it could be potentially overcome to improve outcome for patients with solid malignancies (Monchamont et al., 2012).

1.10 MicroRNAs

MicroRNAs (miRNAs) are a class of small endogenous non-protein coding single RNA strands measuring 21-25 nucleotides in length. They regulate the translation of messenger RNA into protein (He, & Hannon, 2004). The notion of miRNA was first postulated 19 years ago in the lab of Victor Ambros in Dartmouth College, through developmental timing studies of the heterochronic lin-4 gene in the nematode *Caenorhabditis elegans* larvae (Lee et al., 1993). Lin-4 encodes miRNA with antisense which represses lin-14 and lin-28 protein synthesis in early larval stages. Subsequent studies identified a second gene let-7, which controls developmental transition from the L4 stage to adult stage in *C.elegans* and was shown to function by inhibiting lin-41 and lin-57 translation (Reinhart et al., 2000). Despite this earlier research, the term microRNA was only adopted for the first time in a collection of three papers in the October 2001 edition of *Science* (Ruvkun, 2001). Further research revealed that miRNAs display high evolutionary conservation in animals, plants and more recently humans (Zhang et al., 2007).

These RNA molecules are involved in a biological process known as RNA interference (RNAi), a term coined by Fire and Mello in Nature 1998 (Fire et al., 1998). MiRNAs can induce RNAi along with other non-coding RNA molecules found in Eukaryotes such as Small Interfering RNA (siRNA).

Experimental evidence has shown that more than 30% of all human genes are subject to miRNA-mediated regulation (Nilsen, 2007). Due to their short sequences, each miRNA has the ability to regulate many different genes, and conversely many miRNAs can regulate the same gene. These effects include regulation of kinase pathways such as MAP kinase and also regulatory changes to signalling pathways such as the Wnt (Wingless) pathway, identified in the common fruit fly *Drosophila melanogaster* (Rijsewijk et al., 1987). These molecules function as regulators of gene expression and have been shown to be involved in human carcinogenesis by acting as tumour suppressors or oncogenes (Tie & Fan, 2011). A miRNA that is down regulated in cancer and targets an oncogene might act as a tumour suppressor, whereas an up regulated miRNA that targets a tumour suppressor or a gene important for differentiation might act as an oncogene. MiRNAs act as post-transcriptional repressors of their target genes when bound to complementary sequences in the 3' untranslated region (UTR) of the target mRNA resulting in translational repression. In this way, genes can be regulated or "silenced" as a result of miRNA-target mRNA binding in the RNA Induced Silencing Complex (RISC) (Tie & Fan, 2011).

1.11 miRNA Biogenesis

MicroRNA structure and function differs to that of other RNA interference molecules as they must undergo thorough post-transcriptional modification before reaching maturity and occurs via a miRNA biogenesis pathway.

Firstly, RNA polymerase II transcribes miRNA into pri-miRNA which represents a primary spliced transcript possessing a 5' cap and a polyA tail (Kwak et al., 2010). The Drosha endonuclease processes pri-miRNA into short stem-loop structures known as pre-miRNA which are generally 70 nucleotides long and involve extended single stranded RNA in the form of 1-4 nucleotide 3' overhangs, 25-30 base pair stems and relatively small loops (Schwarz et al., 2003). Cleavage by the enzymatic activity of Drosha takes place 22

nucleotides downstream of the terminal loop. Exportin-5 is responsible for the export of pre-miRNA from the nucleus to the cytoplasm and has been shown to bind directly and specifically to correctly processed pre-miRNAs (Lee et al., 2011). Once in the cytoplasm, pre-miRNA stem loops are cleaved into two short complementary RNA molecules by an RNase III endonuclease called Dicer, which occurs approximately 19 base pairs from the Drosha cut site (Lee et al., 2011).

One of these strands known as the “guide” strand is selected by the RNase argonaute protein based on 5’ end stability and is incorporated into the RNA Induced Silencing Complex (RISC). The strand with lower stability base pairing of the 2-4 nucleotide at the 5’ end of the duplex preferentially associates with RISC and subsequently becomes the active miRNA (Lee et al., 2003) whereas the other “anti-guide” strand is degraded as a RISC complex substrate. Once integrated into RISC, the miRNA base-pairs with a complementary mRNA molecule inducing argonaute protein-mediated mRNA degradation. RISC is ultimately responsible for the gene silencing effects of miRNA expression and RNA interference (Lee et al., 2003). Perfect complementarity between miRNA and mRNA 3’ UTR sequences will result in the degradation of target mRNA (a feature mainly seen with siRNA). However, in the majority of cases of human RNA interference by miRNA, only partial complementarity exists in the 3’ UTR causing inhibition of protein translation (Bartel, 2004).

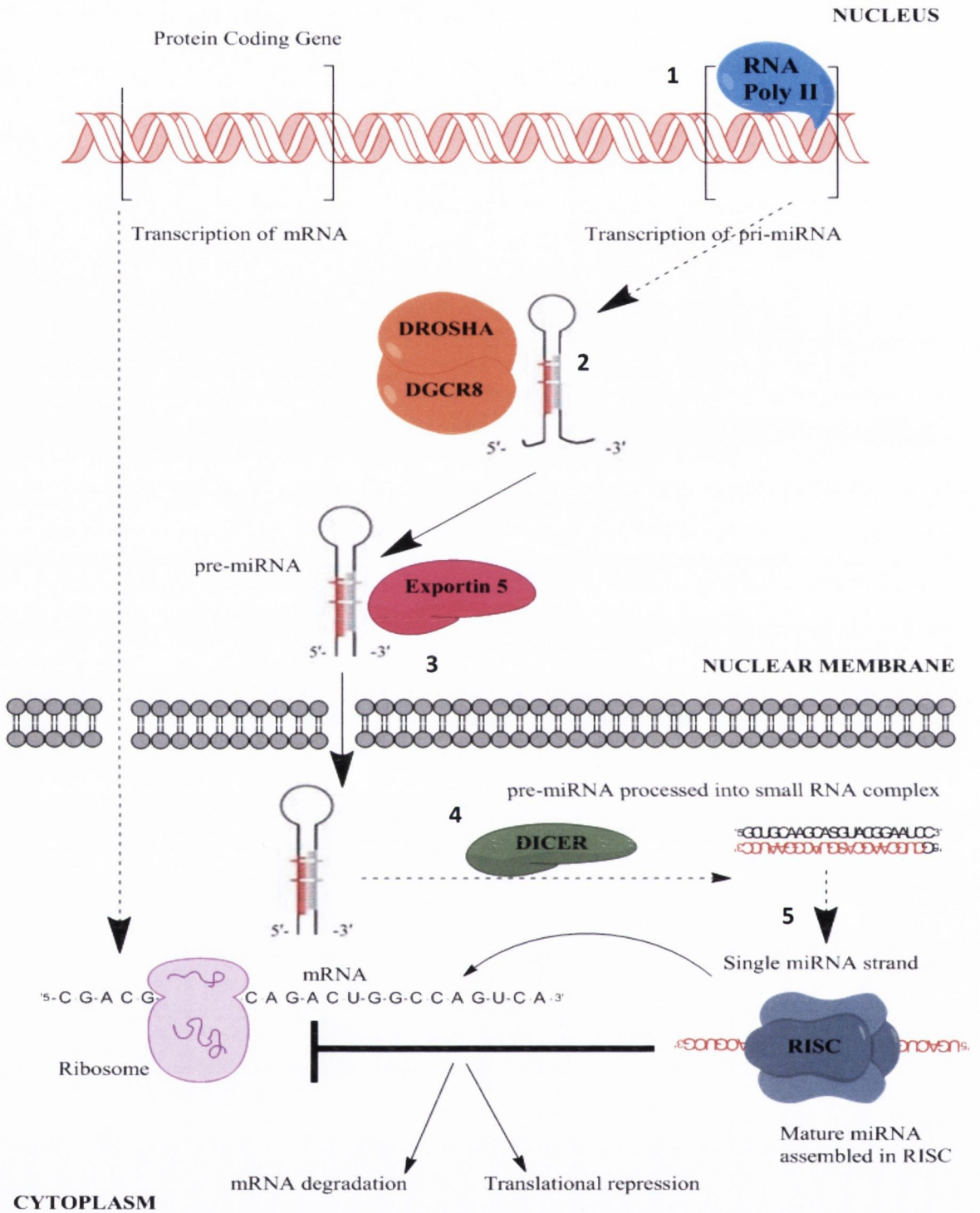


Figure 1.11: Illustration depicting the miRNA biogenesis pathway and miRNA-Mediated Gene Regulation in Animal Cells. In figure 1.11, RNA polymerase II transcribes a pri-miRNA (1) in the cell nucleus before being processed by Drosha into a 70nt long pre-miRNA (2). Exportin-5 then transports the pre-miRNA from the nucleus into the cell cytoplasm (3) where the stem loops are cleaved into two short complementary strands

by Dicer (4). RNase argonaute proteins select a “guide strand” and incorporate it into RNA Induced Silencing Complex (RISC) where it base-pairs with a complementary mRNA molecule inducing argonaute protein-mediated mRNA degradation (5). This miRNA-mRNA interaction results in degradation of the target transcript of translational repression.

1.12 miRNAs in Cancer

MiRNAs offer attractive targets for future therapeutic interventions for the treatment of many types of human diseases and in particular, malignancies that have a strong underlying genetic component driving them. It is known that miRNAs function as regulators of gene expression, and they have been shown to be involved in human carcinogenesis by acting as tumour suppressors or oncogenes as described in section 1.11 (Lynam-Lennon et al., 2009). By manipulating miRNAs involved in disease pathogenesis and limiting off target effects, the process of tumour development might be slowed or halted completely when used in conjunction with conventional established therapies for the treatment of cancer. In the field of RNAi biology, many researchers use miRNA mimics (pre-miRs) and miRNA antagonists (anti-miRs) to investigate their effects on the different hallmarks of aggressive cancer such as invasiveness, metastasis and chemoresistance (Schoof et al., 2012; Piao et al., 2012). MiRNAs that confer growth advantages to cancer cells may be knocked down by anti-miR molecules to make them more susceptible to apoptosis by increasing their sensitivity to chemotherapy drugs such as taxol (Meng et al., 2012).

The issue of miRNAs in CSCs is also beginning to be addressed with studies looking at how miRNA antagonism can hinder the development of potential tumour initiating cells. Antagonism of miR-21 has been shown to reverse EMT and CSC phenotype in breast cancer through targeting PTEN, via inactivation of AKT and ERK1/2 pathways (Han et al., 2012). Conversely, delivering pre-miRs into cancer cells that mimic miRNAs targeting genes involved in migration that become up regulated in cancer, could potentially be carried out to investigate if the manipulation can disrupt the invasive potential of the cells. Recent studies have attempted this by targeting miRNAs that regulate the matrix metalloproteinase (MMP) family of proteins (Pytliak et al., 2012). MMPs function by degrading extracellular

proteins and are involved in many cancer processes such as angiogenesis and cell migration (Pytliak et al., 2012). Delivery of pre-miRs that mimic miR-21 has been shown to affect cell migration and invasion in gastric cancer by targeting MMP-2, MMP-9 and MMP-14 via the RECK tumour suppressor gene (Zhang et al., 2008).

Functional analysis of miRNAs has highlighted one of the complex mechanisms of genetic regulation in human biology and how this becomes deregulated in many different cancer types. The intricate interaction between regulator (miRNA) and gene target (mRNA transcript) has shed new light on one of the many forms of genetic alterations that takes place during carcinogenesis. Despite representing potential for future therapeutic intervention further down the line, more must be understood about the effect RNAi has on tumourigenesis before considering these molecules as a viable treatment option for cancer patients.

1.13 SnoRNAs

Small nucleolar RNAs (snoRNAs) are a class of RNA molecules that guide chemical modifications of other RNAs such as ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs). In higher eukaryotes, snoRNAs reside in introns from which they are released during pre-mRNA processing of the hosting genes through nuclease action (Brown et al., 2008). A snoRNA molecule acts as a guide for individual modifications in a target RNA. In order to carry out modification, each snoRNA associates with at least four protein molecules in an RNA/protein complex referred to as a small nucleolar ribonucleoprotein (snoRNP) (Phipps et al., 2011). The associated proteins are in the correct physical location to catalyse the chemical modification of the target base once the snoRNP recognises and binds to a target RNA.

The two main classes of snoRNA are the C/D box snoRNAs and the H/ACA box snoRNAs and have different functions (Khanna & Stamm, 2010). C/D box snoRNAs play a role in methylation, in which they assist in the attachment or substitution of a methyl group onto various substrates. Human rRNA contains approximately 115 methyl group modifications (Maden & Hughes, 1997). The majority of these are the attachment of a methyl group to a ribose group known as 2'O-ribose-methylations. H/ACA box snoRNAs are involved in

pseudouridylation whereby they assist the conversion of the nucleoside uridine to a different isomeric form pseudouridine (Ψ). Mature human rRNAs contain approximately 95 Ψ modifications (Maden & Hughes, 1997).

SnoRNAs assist chemical modifications of other RNA species that are pivotal to a cell and so their dysregulation may play a role in human disease. Research into non-coding RNAs has implicated snoRNAs in a range of human diseases including the cardiovascular and central nervous systems (Rederstorff & Hüttenhofer, 2010). Similarly, recent studies have attempted to highlight the role of snoRNAs in a range of human cancers. SnoRNAs such as SNORA42 has been shown to function as an oncogene as its overexpression has been identified in lung cancer (Mei et al., 2012). In breast cancer, snoRNAs SNORD12, SNORD12B, and SNORD12C have been shown to exhibit tumour suppressive properties through their interaction with the ZFAS1 gene (Askarian-Amiriet al., 2011). As the evidence of non-coding RNA involvement in cancer grows, the need to investigate the role played by snoRNAs in tumourigenesis will continue to be of interest.

1.14 Methylation in Cancer

The DNA methylation pattern is an additional layer of information that is superimposed on the DNA code and that determines many phenotypic attributes (methylSEQr™ Bisulfite Conversion Kit protocol). While the DNA code is stable, the DNA methylation patterns change in response to spatial, temporal, and environmental cues. The DNA methylation of cytosines in the promoter regions has been shown to produce selective gene inactivation (Figure 1.14(a)). These cytosine residues are methylation-susceptible and occur next to guanine residues as CpG di-nucleotides. CpG islands (that is, clusters of CpGs with a GC content of > 55% in ~500bp regions) are often found in the regulatory regions of genes.

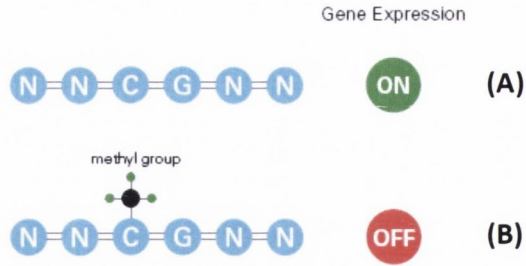


Figure 1.14(a): Illustration depicting gene inactivation as a result of methylation (Applied Biosystems methylSEQr™ Bisulfite Conversion Kit protocol, Appendix CD).

In figure 1.14(a), gene expression takes place when a gene’s promoter region is free of methyl groups bonded to the DNA sequence (A). The addition of a methyl group to a cytosine (C) residue in a gene’s promoter DNA sequence renders the gene inactive as it is unable to be transcribed by polymerases (B). The host cell subsequently loses expression of this gene as a result of methylation.

Genome-wide demethylation has been suggested to be a step in carcinogenesis. Evidence for this notion comes from the frequently observed global DNA hypomethylation in tumour cells, and from a recent study suggesting that defects in DNA methylation might contribute to the genomic instability of tumours (Chen et al., 1998). The methylation state of some genes can be used as a biomarker for tumorigenesis. For example, pi-class glutathione S-transferase gene (GSTP1) promoter hypermethylation appears to be a promising diagnostic indicator of prostate cancer (Nakayama et al., 2004).

Epigenetically mediated gene silencing occurs gradually and begins with a subtle decrease in transcription, fostering a decrease in protection of the CpG island from the spread of flanking heterochromatin and methylation into the island. This loss results in gradual increases of individual CpG sites, which vary between copies of the same gene in different cells (Jones & Baylin, 2002). Methylated Cs can be identified by treating genomic DNA (gDNA) with sodium bisulfite, which converts non-methylated C to U. Methylated C is protected from the bisulfite conversion. After PCR, the sequence of the bisulfite-converted DNA has C residues only if the C residues were already methylated. All other Cs are detected as thymines (T). Thus, an unmethylated gDNA sample produces no Cs in the sequencing data. A comparison of sodium bisulfite-treated DNA sequences with sequences obtained

from untreated gDNA allows a precise identification of all methylated Cs within a long stretch of DNA.

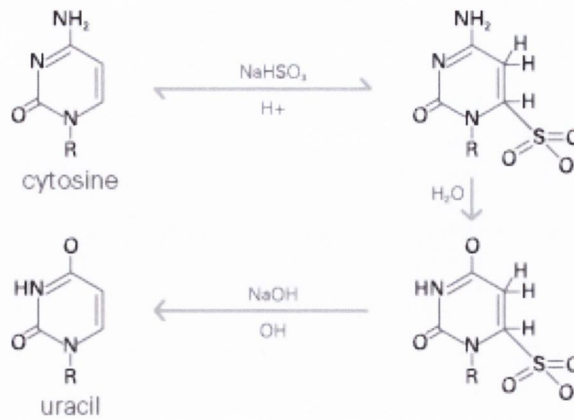


Figure 1.14(b): Illustration of the bisulfite conversion of genomic DNA. Sodium bisulfite (NaHSO_3) deaminates an unmethylated cytosine to produce Uracil. Methylated cytosines are protected from this conversion.

1.15 SOLiD™ 4 Next Generation Sequencing

The SOLiD™ 4 System is a revolutionary genetic analysis platform which provides parallel sequencing of clonally-amplified DNA fragments linked to beads. This technology provides accuracy, ultra-high throughput and application flexibility to advance research. The SOLiD™ 4 System allows analysis of the different classes of small RNA such as miRNA, snoRNA, piwi-interacting RNA (piRNA) and long non-coding (lncRNA).

The ultra-high throughput SOLiD™ 4 System has many benefits when it comes to whole genome sequencing of small RNAs such as:

- Conserve strandedness of cDNA, allowing the user to know exactly which strand a particular transcript originated from.
- Generation of up to greater than 700 million mappable sequence reads per run for a fragment library enabling high sensitivity for analysing low RNA expression levels and measuring accurate fold changes at the same low expression levels.
- Enables the user to multiplex and sequence small RNA libraries simultaneously, reducing the cost of analysis per sample.
- Detect allele specific expression patterns

The SOLiD™ Small RNA Solution is a robust method for hypothesis-neutral, whole genome analysis of expression patterns that enables the discovery of novel RNA without the probe bias of microarrays.

As this is the only such instrument in Ireland and offered this project a unique opportunity to apply state-of-the-art technology to address fundamental questions of carcinogenesis. The overall aim of this work was to highlight validated and also novel miRNAs in cancer stem cell populations as potential diagnostic and prognostic biomarkers across of wide range of tumour types.

1.16 Aims & Objectives

- Establish a cell line panel comprised of a wild type and V600E mutated BRAF Melanoma, Thyroid, Ovarian and Colorectal cell lines confirmed by BRAF SNP Genotyping
- Generate cancer stem cell progenies in the form of holoclones, from parental cells using a positive selection high salt agar assay
- Investigate expression of a panel of genes involved in various different biological processes involved in carcinogenesis such as stemness, EMT and cell signalling
- Investigate protein expression of the most consistently up regulated gene in holoclones via confocal microscopy
- Interrogate the non-coding RNA profile of BRAF wild type and V600E mutant cell lines and derived holoclone populations using the Applied Biosystems SOLiD™ 4 Next Generation sequencing platform
- Examine the methylation status of the genes RAR β 2 and MLH1 in the cell line panel, which have been shown in previous studies to be hypermethylated in cancers possessing a V600E mutated BRAF gene

The hypotheses of this thesis are:

1. That cancer stem cells are present in a variety of different cancers and may hinder effective treatment of different tumour types due to their more aggressive phenotype and chemoresistance potential.
2. That CSCs in holoclones derived from parent cell lines will display different gene expression and miRNA expression profiles than their founder line.
3. That BRAF V600E mutant CSCs will display different gene expression, miRNA and snoRNA profiles than CSCs generated from BRAF wild type cell lines.
4. That BRAF V600E mutated cell lines will display different MLH1 and RAR β 2 methylation profiles than BRAF wild type cell lines.

BRAF V600E mutation has been shown to have many downstream effects on cell signalling and its prevalence varies across cancer types. Our group has previously demonstrated that BRAF V600E mutation may cause dysregulation of non-coding RNAs, facilitating aberrant gene transcription and cell growth in carcinogenesis.

By comprehensively examining non-coding RNA profiles of wild type and BRAF V600E mutated cell lines we will map the involvement of BRAF mutation in tumour progression. Cancer stem cells present an attractive target for developing therapies to combat tumour (re)population and so by isolating cancer stem cells in wild type and BRAF V600E cell lines, we will also identify dysregulated genes and pathways central to the attainment and maintenance of stemness characteristics. This work proposes to generate greater insights into BRAF mediated biological pathways and molecular mechanisms that regulate cell fate, development, and disease progression.

Chapter 2
Materials & Methods

2.1 Introduction

This chapter is a summation of all the methodologies employed in this thesis, accompanied by background information on some of the newer techniques. Comprehensive descriptions of the more detailed protocols can be found in the Appendix accompanying this thesis.

2.2 Cell Culture

2.2.1 Cell Lines

Cell lines were cultured in appropriate cell culture media in a Class 1 laminar flow and incubated in a humidified incubator at 37°C & 5% CO₂.

Melanoma

The BRAF wild type cell line COLO794 (ECACC Number 94072237) was purchased from European Collection of Cell Cultures (ECACC) and was established from a 14 year old male patient with a malignant melanoma. The melanoma showed rapid growth and metastases in the patient and the biopsy was taken before therapy from a subcutaneous metastasis. COLO794 cells were grown in RPMI 1640, containing 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:2 to 1:3.

The BRAF V600E mutated cell line SK-Mel 28 (ATCC Number HTB-72) is a malignant melanoma cell line established from a 51 year old male and was purchased from American Tissue Culture Collection (ATCC). SK-Mel 28 cells were grown in Eagle's Minimum Essential Medium containing 10% FBS, 2% Sodium Bicarbonate, 1% Sodium Pyruvate, 1% L-Glutamine and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:3 to 1:6.

Thyroid

The BRAF wild type TPC-1 is a papillary thyroid carcinoma cell line kindly obtained as a gift from Dr. Mariwil Wong, Mt Zion Medical Centre of UCSF after receiving permission to use the cells from Dr. Nobuo Satoh, Cancer Research Institute, Kanazawa University. Cells were grown in RPMI 1640, containing 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:3 to 1:6.

The BRAF V600E mutated cell line 8505C (DSMZ Number ACC 219) is a cell line derived from a human anaplastic thyroid carcinoma and was purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ). The cell line was established from the primary tumour of a 78-year-old woman with thyroid carcinoma (undifferentiated carcinoma, histologically a largely papillary adenocarcinoma with some spindle, polygonal and giant cells). 8505C cells were grown in Minimum Essential Media supplemented with 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:3 to 1:6.

Ovarian

The BRAF wild type A2780 (ECACC Number 93112519) is an ovarian carcinoma line established from tumour tissue from an untreated patient and was purchased from ECACC. A2780 cells were grown in RPMI 1640, containing 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:3 to 1:6.

The BRAF V600E mutated cell line ES-2 (ATCC Number CRL-1978) is a poorly differentiated ovarian clear cell carcinoma line and was purchased from ATCC. It was established from a surgical tumour specimen taken from a 47 year old black woman. ES-2 cells were grown in McCoy's 5A media supplemented with 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:4 to 1:8.

Colorectal

The BRAF wild type cell line COLO320 (ATCC Number CCL-220) is a human colon carcinoma cell line derived from a sigmoid colon carcinoma from a 55 year old Caucasian male and was purchased from ATCC. COLO320 cells were grown in RPMI 1640, containing 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:5 to 1:10.

The BRAF V600E mutated cell line HT-29 (ATCC Number HTB-38) is a human colon cell line derived from a colorectal adenocarcinoma. HT-29 cells were grown in Dulbecco's Modified Eagle's Medium containing 10% FBS and 1% 20,000U/ml Penicillin/Streptomycin and had a recommended subcultivation ratio of 1:3 to 1:8.

2.2.2 Propagation of Cells from Liquid Nitrogen Storage

Complete media (with FBS and Penicillin/Streptomycin) was heated to 37°C in a water bath for approximately 30 mins. A cryovial of cells was removed from liquid nitrogen storage and placed in a 37°C water bath for 2 minutes to thaw out. 5ml of warmed complete media was aliquoted into a 15ml tube. 500µl of warmed media was added to the thawed cells then added to the 15ml tube. Cells were centrifuged for 5 minutes at 200g and the supernatant was decanted. Cells were resuspended in 1ml of warmed media & a cell count was performed if required. The appropriate amount of cells was added to a T75 tissue culture flask along with the required amount of media and was incubated at 37°C with 5% CO₂ overnight. After 24 hours (if cells had adhered) cells were washed with 10ml of 1X phosphate buffered saline pH 7.4 (PBS) and media was replenished.

2.2.3 Subculturing

Spent media was removed from each flask and decanted into a designated biological waste bottle. Cells were washed gently with 10ml 1X PBS to remove dead cells and debris which were decanted into a waste bottle. 3ml of 0.25% Trypsin/EDTA was added to the flask and incubated at 37°C for 5-10 minutes (depending of adherence strength). Flasks were examined by phase microscopy to assess cell detachment and if needed were given gentle taps to encourage detachment of cells from the plastic surface. When >90% of cells had detached, complete culture media was added to neutralise trypsin activity. Cell suspension was transferred to a new 15ml tube and centrifuged for 4mins at 1000rpm to pellet the cells. Supernatant was removed and the cell pellet was resuspended in 2ml fresh culture media. New flasks were seeded with 25% (500ul for 1:4 split ratios) of the suspended cells.

2.2.4 Cell Counting with a Haemocytometer

Live cells were stained with Trypan Blue and counted in 2 grids on a haemocytometer. The average number counted was multiplied by the dilution factor (e.g. 10) to equal 10^4 , and was used to calculate the final cell number in 1ml (e.g. $200 \times 10^4 = 2 \times 10^6$ or 2 million).

2.2.5 Cryopreservation

Media was removed and cells were washed twice with 1X PBS and 3ml of 0.25% Trypsin-EDTA was added to detach cells from flask. Cells were incubated at 37°C for 5 minutes and tapped gently to encourage cell detachment before incubating for another 2 minutes. Flask was examined under microscope briefly to confirm detachment of cell monolayer. 3ml of fresh culture media was added to the flask to inactivate trypsin and before being centrifuged at 300g for 5 mins. Supernatant was removed and cell pellet was resuspended in the small amount of media that remained in the tube. Cryoprotectant (90% complete media, 10% DMSO) was added drop wise onto the cell suspension and mixed gently. 1ml of cell/DMSO mix was aliquoted slowly into each cryovial. Cryovials were stored at -80°C overnight before transferring to liquid nitrogen storage.

2.2.6 High Salt Agar Assay

1% agarose and 1% Sodium chloride (NaCl) was added to distilled H₂O water and autoclaved. 30ml of the autoclaved salt agar mix was plated into high adherence petri dishes and were allowed to set in a laminar flow hood before adding 20ml of appropriate complete media onto the of salt agarose layer. Adherent cells were washed with 1X PBS and trypsinised to get into suspension as per normal protocol. A cell count was performed and then 1×10^6 of suspended cells were pipetted onto the layer of culture media. Petri dishes were incubated at 37°C and 5% CO₂. Dishes were inspected using a Brightfield microscope with a 10X objective lens every 4-5 days to examine condition of culture media and to visualise holoclone development. Visible clumps of dead cells and debris were removed and fresh culture media was added when appropriate (observed colour change as per normal cell culture practices). Depending on cell line (average time 4 weeks), spheroids that appeared to be embedded in the agarose were examined microscopically to confirm holoclone morphology. Holoclones were harvested individually with a clean pipette tip and placed in a microcentrifuge tube with a small amount of warm culture media and centrifuged at 1000rpm for 5 mins. Supernatant was removed and cells underwent another spin in 1X PBS to remove any residual agar. Supernatant was decanted and cell pellets were used for immediate RNA extractions or stored at -80°C for future extractions.

2.3 Nucleic Acid Extraction and Quantitation

2.3.1 Genomic DNA Purification

Genomic DNA was extracted from cultured cells for TaqMan® SNP Genotyping using the Qiagen DNeasy Blood and Tissue Kit. For maximum DNA yield, the elution step was repeated once to give an increased overall DNA yield. The full protocol for Genomic DNA purification using the Qiagen DNeasy Blood and Tissue Kit can be found in Appendix I.

2.3.2 RNA Extraction

The Qiagen miRNeasy Mini Kit was used to isolate total RNA for TaqMan® Gene Expression Assays and to purify small RNA for SOLiD™ Small RNA Sequencing. The miRNeasy Kit provides a separate specialized protocol for enrichment of miRNA and other small RNAs (less than ~200 nt) separate fractions. Elutions were carried out twice to ensure maximum RNA yield was obtained. The full protocol for RNA purification using the Qiagen miRNeasy Mini Kit can be found in Appendix I.

2.3.3 Nucleic Acid Quantitation

Nucleic acid samples were checked for concentration and quality using the NanoDrop® ND-1000 spectrophotometer. The NanoDrop® ND-1000 is a full-spectrum (220-750nm) spectrophotometer that measures 1µl samples with high accuracy and reproducibility. For each sample the following results were recorded:

- Concentration (ng/µl): sample concentration in ng/µl
- 260/280: ratio of sample absorbance at 260 and 280nm. The ratio of absorbance at 260 and 280nm is used to assess the purity of DNA and RNA. A ratio of ~1.8 is generally accepted as “pure” for DNA; a ratio of ~2.0 is generally accepted as “pure” for RNA.

2.3.4 Nucleic Acid Quality Control

2.3.4.1 Analysing RNA by the Agilent 2100 Bioanalyser

Large RNA (above 200nt) integrity and concentration were analysed using the 2100 Bioanalyser protocol found in Appendix I.

2.3.4.1 Analysing Small RNA by the Agilent 2100 Bioanalyser

Small RNA (below 200nt) populations and concentration were analysed using the 2100 Bioanalyser protocol found in Appendix I. Two distinct regions could be defined arbitrarily with the Small RNA region from 0 to 150nt, and the microRNA region (miRNA) from the 10 to 40nt.

2.3.4.1 Analysing DNA by the Agilent 2100 Bioanalyser

Small RNA libraries were analysed using to Agilent DNA 1000 Kit. Library integrity and concentration was analysed using the 2100 Bioanalyser and the software's smear analysis program allowed regions of interest to be investigated. The protocol for analysing DNA using the 2100 Bioanalyser can be found in Appendix I.

2.4 Polymerase Chain Reaction

2.4.1 Reverse Transcription of RNA to cDNA

The High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, CA, USA) was used to convert total RNA to single stranded cDNA. The full manufacturer's supplied protocol can be found in Appendix I. Briefly, the kit reagents (10X reverse transcription buffer, 25X dNTP mixture (100mM), 10X Multiscribe[®] reverse transcriptase (50U/ μ l)) were prepared into a 2X master mix in nuclease-free water. RNA samples were diluted in nuclease-free water to ensure concentrations of both parent and holoclone samples were equal. Diluted RNA was then combined 1:1 with the 2X mastermix to prepare a 1X mixture with a final concentration of 2.5U/ μ l Multiscribe[®] enzyme in a final volume of 20 μ l/sample. Samples were briefly centrifuged and the reverse transcription was run in a thermal cycler under the following conditions: 25°C x 10 mins, 37°C x 120 mins, 85°C x 5 mins, 4°C x ∞ .

2.4.1 TaqMan® PCR

TaqMan® PCR was employed for SNP genotyping of purified cell line genomic DNA and also mRNA quantitation of parent and holoclone samples in this study.

TaqMan PCR and RT-PCR products are small (generally less than 200 base pairs) and thus can be used to amplify partially degraded or fragmented DNA/RNA such as that obtained from FFPE material.

2.4.2 TaqMan® SNP Genotyping and Allelic Discrimination Assay

In order to establish a cell line panel based on BRAF status, cell lines underwent allelic discrimination using a custom TaqMan® SNP Genotyping Assay designed in-house by Dr. Paul Smyth, TCD.

The 40X custom TaqMan® SNP Genotyping Assay contained:

BRAF Forward Primer 5' CAT GAA GAC CTC ACA GTA AAA ATA GGT GAT 3'

BRAF Reverse Primer 5' GGA TCC AGA CAA CTG TTC AAA CTG A 3'

BRAF-P^{WT} VIC-5' CCA TCG AGA TTT CAC TGT AG 3'

BRAF-P^{MUT} FAM-5' CCA TCG AGA TTT CTC TGT AG 3'

The assay was mixed with 2X TaqMan® Gene Expression Mastermix and gDNA template in nuclease-free water to a final concentration of 1X with 10ng of gDNA per reaction. Non-template controls (NTC), homozygous wild type, homozygous mutant and heterozygous controls were included in duplicate in each plate at 10ng/reaction. Q-PCR endpoint assay was run on a 7900 Prism® RT-PCR machine (Applied Biosystems, CA, USA) with cycling parameters: 50°C x 2 mins, 95°C x 10 mins and 40 cycles of: 92°C x 15 secs, 60°C x 60 secs. A post-read allelic discrimination assay was run and the results were autocalled using a 95% confidence interval.

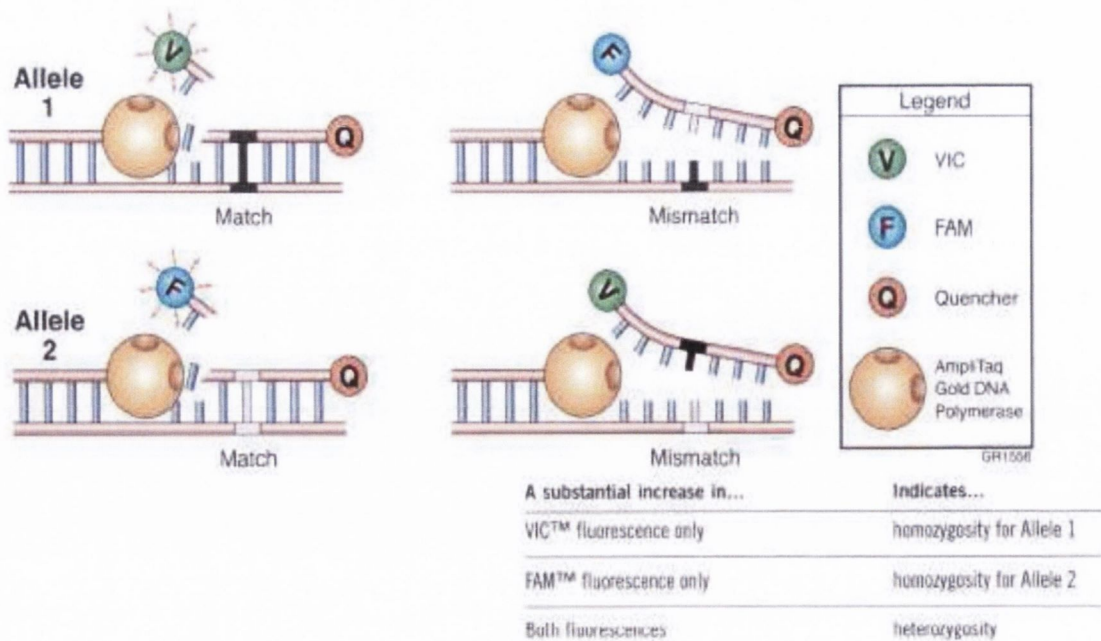


Figure 2.4.3(a): Allelic discrimination using the 5' nuclease assay.

Three factors contribute to the discrimination based on a single mismatch.

First, the mismatch has a disruptive effect on hybridisation. A mismatched probe will have a lower T_m than a perfectly matched probe. Proper choice of annealing/extension temperature during PCR will favour hybridisation of an exact-match probe over a mismatched probe. Second, the assay is performed under competitive conditions with both probes present in the same reaction tube. Therefore, mismatched probes are prevented from binding due to stable binding of exact match probes.

Third, the 5' end of the probe must start to be displaced before cleavage occurs. The 5' nuclease activity of Taq polymerase recognises a forked structure with a displaced 5' strand of at least 1 to 3 nucleotides (Lyamichev et al., 1993). Once a probe starts to be displaced, complete dissociation occurs faster with a mismatch than an exact match. This means there is less time for cleavage to occur with a mismatch probe. Thus, the presence of a mismatch promotes dissociation rather than cleavage of the probe.

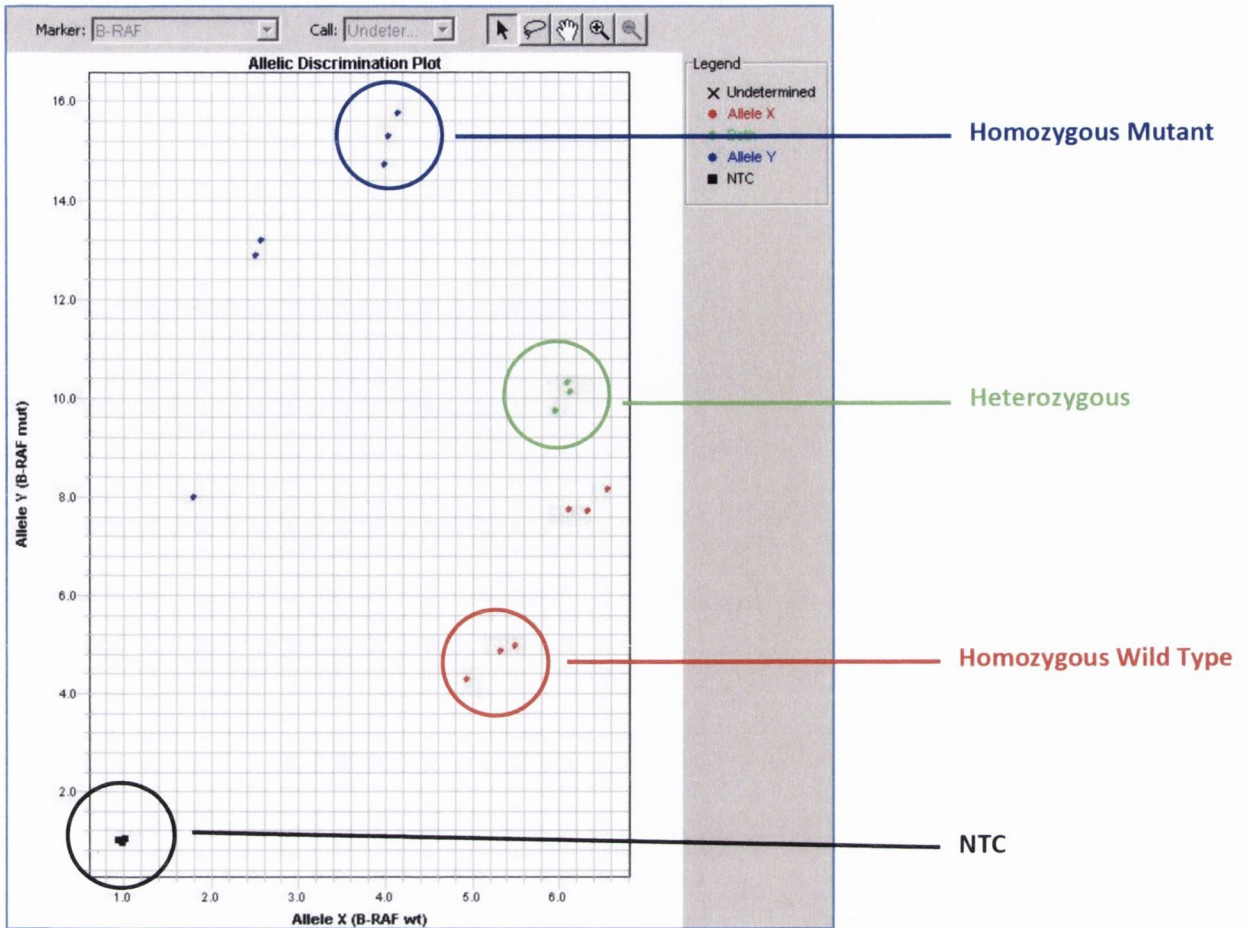


Figure 2.4.3(b): Typical allelic discrimination output plot illustrating a homozygous wild type, homozygous mutant and heterozygous result

2.4.3 Real-time quantitative TaqMan® RT-PCR

Real-time RT-PCR allows monitoring of 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 revolutionises the way one approaches PCR-based quantitation of DNA and RNA. In real-time RT-PCR, reactions are characterised 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 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 CT (threshold cycle) is defined as the fractional cycle number at which the fluorescence passes the fixed threshold.

2.4.3.1 Relative Gene Expression

The ABI 7500Fast Real Time PCR System supports real-time relative quantification of nucleic acids using the Comparative CT method. Relative quantification describes the change in expression of the target gene relative to a reference group such as an untreated control or a sample at time zero in a time-course study. In the comparative CT method, relative quantity (RQ) is expressed relative to calibrator sample that is used as the basis for comparative results. Therefore, the calibrator is the 1X sample and all other quantities are expressed as an n-fold difference relative to the calibrator. The Comparative CT Method (Livak & Schmittgen, 2001) uses the following arithmetic formula to achieve results for relative quantification: **2^{-ΔΔCT}**

Where ΔCT is achieved by subtracting the endogenous control CT from your samples' CT, therefore $\Delta CT = CT(\text{sample}) - CT(\text{endogenous control})$. To work out ΔΔCT you must compare expression levels relative to a calibrator sample, thus $\Delta\Delta CT = \Delta CT(\text{sample}) - \Delta CT(\text{calibrator})$. Using the 2-ΔΔCT method, the data are presented as the fold change in gene expression normalised to an endogenous reference gene and relative to the untreated control.

Relative quantification on the 7500Fast instrument is accomplished through the use of the polymerase chain reaction and the fluorogenic 5' nuclease assay. After selection of endogenous control, probe and primer sets are designed or chosen for both target and control sequences. Samples containing cDNA, master mix, and assays targeting specific nucleic acid sequences are loaded onto the 7500Fast. During the run, the instrument records the emission resulting from the cleavage of TaqMan® probes in the presence of the target sequence. After the run, the SDS software processes the raw fluorescence data to produce threshold cycle (CT) values for each sample. This software computes relative quantities from the CT values of the calibrator sample and the data from the unknown samples within the gene expression profile. Relative quantification values (gene fold

changes) can be achieved manually by collecting CT values and inputting them into the 2- $\Delta\Delta$ CT formula; alternatively the SD software can collect data and produce the RQ values.

2.4.4.2 TaqMan® Gene Expression Assays

The following genes were analysed for their expression levels in holoclones relative to their parental cells by Real-time quantitative TaqMan® RT-PCR:

Gene Symbol	Gene Name	TaqMan Assay ID	Notes
CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)	Hs00153277_m1	Endogenous Control
NANOG	Nanog homeobox	Hs02387400_g1	
POU5F1	POU class 5 homeobox	Hs00999632_g1	
ALDH1A1	aldehyde dehydrogenase 1 family, member A1	Hs00946916_m1	
SNAI2	snail homolog 2 (Drosophila)	Hs00950344_m1	
SHH	sonic hedgehog	Hs00179843_m1	
TGFBR1	transforming growth factor, beta receptor 1	Hs00610320_m1	
CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa	Hs00355049_m1	
CDH1	cadherin 1, type 1, E-cadherin (epithelial)	Hs01023894_m1	

Table 2.4.4.2: List of TaqMan® Gene Expression Assays illustrating gene symbol, name and ID

2.5 Protein Analysis

2.5.1 Immunofluorescence Staining of Cells by Confocal Microscopy

Parental cells and holoclones derived from 2 cell lines, COLO320 and HT-29 were stained for NANOG protein expression and examined via confocal microscopy.

Cell Preparation:

Cells were plated into 8 well chamber slides and left in an incubator at 37°C until a confluent monolayer formed. Cells were then rinsed twice with 1X PBS.

Paraformaldehyde/Triton Fixation:

Cells were incubated at 37°C for 30mins in 4% paraformaldehyde in 1X PBS and permeabilized with 0.50% TRITONX-100 in 1X PBS for 10mins. Cells were stained with a 50 mg/ml fluorescent phalloidin conjugate solution in 1X PBS (containing 1% DMSO from the original stock solution) for 40 minutes at room temperature and then washed three times with 1X PBS to remove unbound phalloidin conjugate.

Blocking:

Chamber slides were incubated with blocking buffer (1% BSA in 1X PBS) for 30mins at 37°C to minimize non-specific absorption of the antibodies

Incubation with Primary Antibody

Blocking buffer was removed and primary antibody was diluted to 1.0-10 µg/ml in blocking buffer (optimal concentration can depend on several variables, such as the affinity of the antibody and the abundance of the antigen). 150-200µl of the primary antibody solution was distributed on each chamber and incubated for 1 hour at room temperature in a humidified chamber. The antibody solution was removed by aspiration and chamber slides were washed three times in 1X PBS, 5 minutes each wash.

Preparation for Microscopy:

50µl of mounting media was added to the slide and a cover slip was gently lowered onto the slide. The edges of each cover slip were sealed with regular transparent nail polish and allowed to dry for 3 minutes.

2.5.2 Confocal Imaging

Images were captured at X40 magnification and X4 zoom. Scans were performed at 1µm interval depths through the fixed cells, and single or merged images are presented as XY single planes through the mid-section of the cells unless otherwise specified.

Fluorescent Stains:

Fluoroshield with DAPI (Sigma Aldrich, USA)

Excitation: 360 Emission: 450 Colour: Blue

Rhodamine Phalloidin, Tetramethylrhodamine B isothiocyanate (TRITC) (Sigma Aldrich, USA)

Excitation: 552 Emission: 570 Colour: Red/Orange

Anti NANOG FITC: Catalog number: 347197

Excitation: 495 Emission: 525 Colour: Green

2.6 Methylation Analysis

Overview

In-house pilot studies have reported different methylation patterns in Wnt signalling genes across a panel of wild type and V600E mutated BRAF thyroid cell lines. The degree of methylation was expressed as the percentage of CpGs showing any methylation as a function of the total CpGs analysed. MLH1 and RARβ2 exhibited 23 out of 23 and 10 out of 10 methylated CpGs respectively in a BRAF V600E mutated cell line. This was a comprehensive increase of methylation levels over cell lines with wild type BRAF so we sought to investigate if BRAF mutation was inducing methylation of the cancer related genes MLH1 and RARβ2 across a panel different cell lines possessing wild type and mutant BRAF.

Genomic DNA was extracted from a panel of cell lines and before undergoing bisulfite conversion. Converted DNA was amplified using specific customized primers using an optimised protocol before undergoing bisulfite sequencing.

2.6.1 Performing Bisulfite Conversion

Genomic DNA was isolated using the Qiagen DNeasy Blood & Tissue Kit following the protocol outlined in section 2.3.1. DNA concentration and purity was measured by the Nano-Drop® 1000 spectrophotometer following the protocol outlined in section 2.3.3.

The Applied Biosystems MethyLSeqr™ Bisulfite Conversion kit was used to bisulfite convert 700ng of purified genomic DNA. The full ABI MethyLSeqr™ Bisulfite Conversion protocol can be found in Appendix I.

2.6.2 Methylation Primer Design

PCR primers were designed to amplify specific areas in the promoter regions of the genes RARβ2 and MLH1. In collaboration with ABI, primers for these regions were designed and included M13 tails for sequencing upon completion of PCR.

Primer Sequences

MLH1

Forward - 5' TTT TTTTTT AGG AGT GAA GGA GGT TA 3'

Reverse - 5' ATA AAA CCC TAT ACC TAA TCT ATC 3'

Amplicon Size –221bp

RARβ2

Forward - 5'AAG TAG TAG GAA GTG AGT TGT TTA GA 3'

Reverse - 5' CCAAATTCTCCTTCCAAATAA 3'

Amplicon Size –220bp

2.6.3 PCR Amplification

PCR was carried out using AmpliTaq Gold reagents (Applied Biosystems, CA, USA). A mastermix of the kit reagents was made up of the following: 10x TaqGold Buffer, 2.5mM dNTPs, 25 mM MgCl₂, TaqGold Polymerase, Primer Pair 2.5um each, and bisulfite-converted DNA (6ng/μl) diluted in nuclease-free water to give a final reaction volume of 10μl. Reactions were run in a thermal cycler under the following conditions: 95°C for 5mins, 45 cycles of: 95°C for 30secs, 60°C for 2mins, 72°C for 45secs, followed by 60°C hold for 10mins and a 4°C hold for ∞.

PCR Clean Up

PCRs were cleaned prior to sequencing by treatment with EXO-SAP-it (USB). The PCRs were diluted as needed, by adding 10μL of water to each well and using 0.5μl per 10μl sequencing reaction.

2.6.4 Bisulfite Sequencing

Sanger Sequencing on PCR products was carried out externally by Source Bioscience Lifesciences. PCR products were quantified using a NanoDrop® ND-1000 spectrophotometer following the protocol in Section 2.3.3 and were diluted accordingly to 1ng/100bp as recommended. M13 forward and reverse stock primers were used in the sequencing reaction at a concentration of 3.2pmol/μl.

Reaction Mix

BDT v1.1	4
M13 primer 3.2 uM	0.5
PCR template, diluted	0.5
<u>Water</u>	<u>5</u>
Total	10μl

Sequencing reactions were purified using Big Dye XTerminator to remove unincorporated dye terminators, and the sequencing products were run on a 3730 sequencer.

2.7 Next Generation Sequencing of Small RNAs via SOLiD™ 4

2.8 Sample Preparation

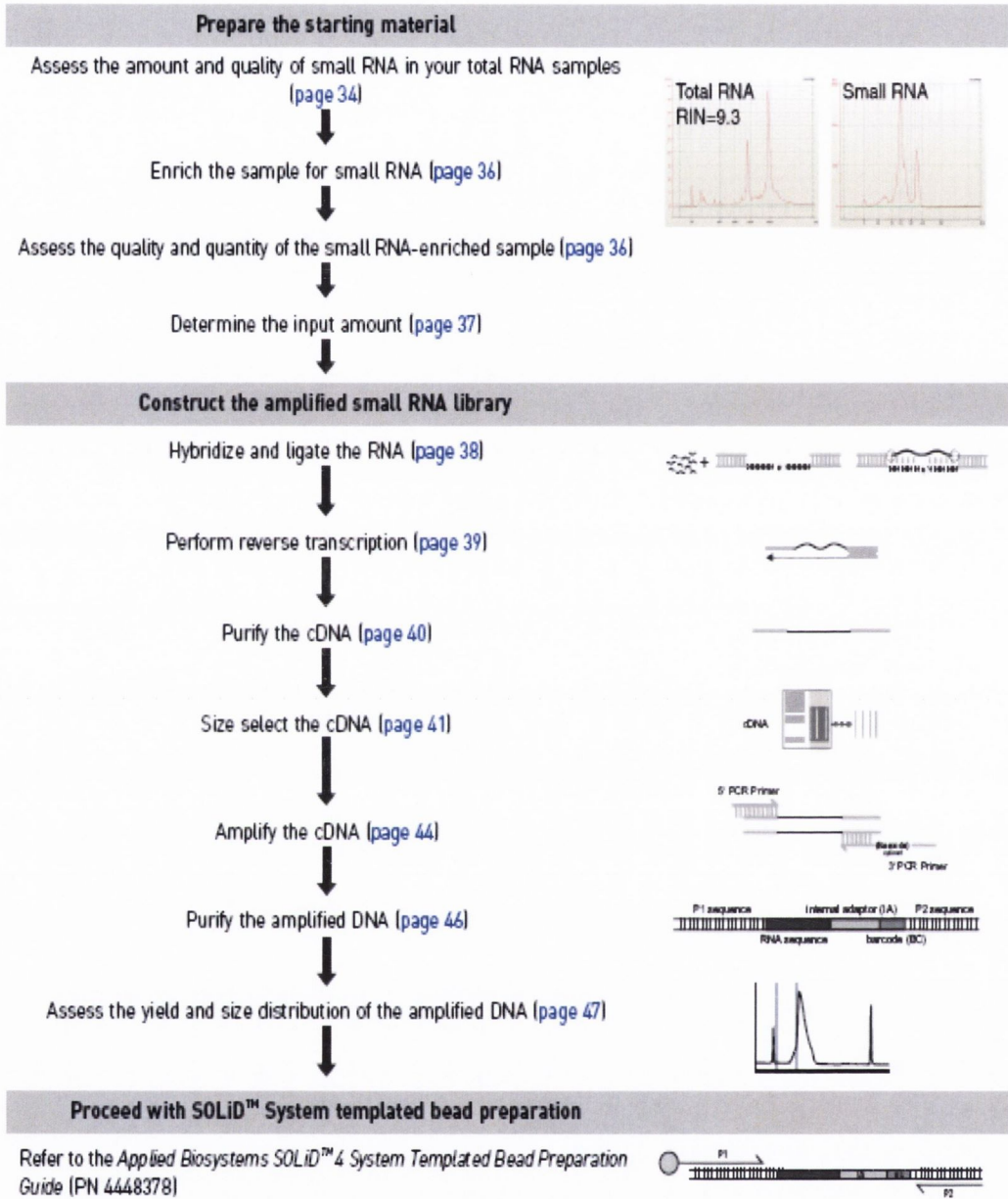


Figure 2.8(a): Workflow involved in creating Small RNA libraries for next generation sequencing via the SOLiD™ 4 platform (ABI SOLiD™ Total RNA-Seq Kit Protocol)

Purified small RNA samples were diluted to 20µl in water and 1µl per well was analysed using the Small RNA Kit (Agilent, USA) on the 2100 Bioanalyzer™ Instrument (Agilent, USA) to determine the percentage of miRNA in a small RNA sample.

$$\% \text{ miRNA} = \left(\frac{\text{mass of miRNA (10–40 nts) from the Small RNA Chip}}{\text{mass of enriched small RNA from the Small RNA Chip}} \right) \times 100$$

Figure 2.8(b): Formula used to calculate miRNA percentage in a small RNA sample

Samples were then diluted in nuclease-free water such that the small RNA sample was less than 100ng/μl with miRNA content between 1-100ng/μl.

2.9 Library Preparation

2.9.1 Hybridisation, Overnight Ligation and Reverse Transcription

SOLiD adapter mix was added to hybridisation solution and small RNA template at a ratio of 2:3:3 and run in a thermal cycler at 65°C for 10 mins and 16°C for 5 mins. The ligation enzyme (2μl per sample) and 2X ligation buffer (10μl per sample) were added to the hybridisation sample and mixed by pipetting. The ligation reaction was run in a thermal cycler with heated lid turned off for 16 hours at 16°C.

19μl of reverse transcription master mix (4μl 10X RT Buffer, 2μl 2.5mM dNTP, 2μl SOLiD RT Primer, 11μl water) was added. The solutions were mixed well by pipetting and briefly centrifuged before being heated at 70°C for 5 mins in a thermal cycler and then snap cooled on ice. To each reaction 1μl ArrayScript™ Reverse Transcriptase was added followed by a gentle vortex to mix the contents. Samples were briefly centrifuged and incubated at 42°C for 30 mins to allow the RT reaction to convert the adaptor-ligated RNA to cDNA. The RT reaction was then purified using the MinElute® PCR Purification Kit (Qiagen, USA).

2.9.2 cDNA Size Selection

DNA samples were size selected using gel electrophoresis. 5μl of cDNA was mixed with 5μl Novex 2X TBE-Urea Sample Buffer (Invitrogen, USA). A 40ng/μl 10bp DNA ladder (Invitrogen, USA) was mixed 1:1 in Novex 2X TBE-Urea Sample Buffer (Invitrogen, USA). Samples and ladder were heat denatured at 95°C for 3 mins and then snap cooled on ice. A Novex 10% TBE-Urea gel was locked into position in an X-Cell Sure Lock™ Mini-Cell (Invitrogen, USA) and filled with 1X TBE running buffer (Invitrogen, USA). 10μl of ladder or sample was added into each well and electrophoresis was run at 180V for approximately 45 mins. The gel was

incubated with gentle agitation in a staining solution (0.01% solution of SYBR Gold® nucleic acid stain (Invitrogen, USA) in 1X TBE running buffer) for 10 mins at room temperature. Each gel was visualised on a UV transilluminator and the band between 60-80nts was quickly and carefully excised for each sample. Each gel piece was then cut vertically into 4 equal sections with the 2 middle pieces transferred into a 0.2ml PCR tube and the 2 outside pieces stored at -20°C.

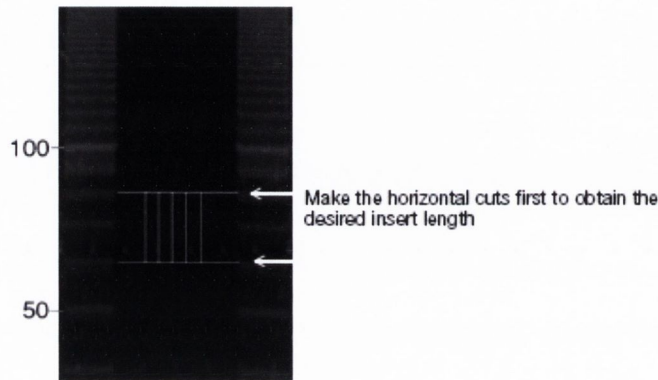


Figure 2.9.2: cDNA size selection after gel electrophoresis

2.9.3 In-Gel PCR and Multiplex Barcoding

DNA was amplified using a 5' primer which incorporates a 5' sequence that is required for downstream binding to microbeads for emulsion PCR (P1 adaptor) and a 3' sequence that incorporates a unique barcode to allow multiplexing of sample libraries and a universal adaptor at the 5' end required for downstream processes (P2 adaptor). For each gel piece, a 100µl PCR master mix was added composed of; 10µl 10X PCR Buffer, 8µl 2.5mM dNTP, 2µl 5' Primer, 1.2µl AmpliTaq® DNA Polymerase, 2µl Barcoded 3' Primer, 76.8µl water. Each PCR reaction was run in a thermal cycler under the following conditions: 95°C for 5 mins, 18 cycles of 95°C for 30 secs, 62°C for 30 secs, 72°C for 30 secs; and 72°C for 7 mins. The PCR reactions were purified using the PureLink™ PCR Micro Kit (Invitrogen, USA).

2.9.4 Preparation of Multiplexed Library

The Agilent DNA 1000 Kit and Chip was used to assess the DNA libraries on the 2100 Bioanalyser as outlined in Section 2.3.4.1. The post-read smear analysis was used to determine the quality of each library. The proportion of miRNA in relation to the total

library was established by calculating the ratio of 120–130bp DNA (desired size): 25–150bp DNA.

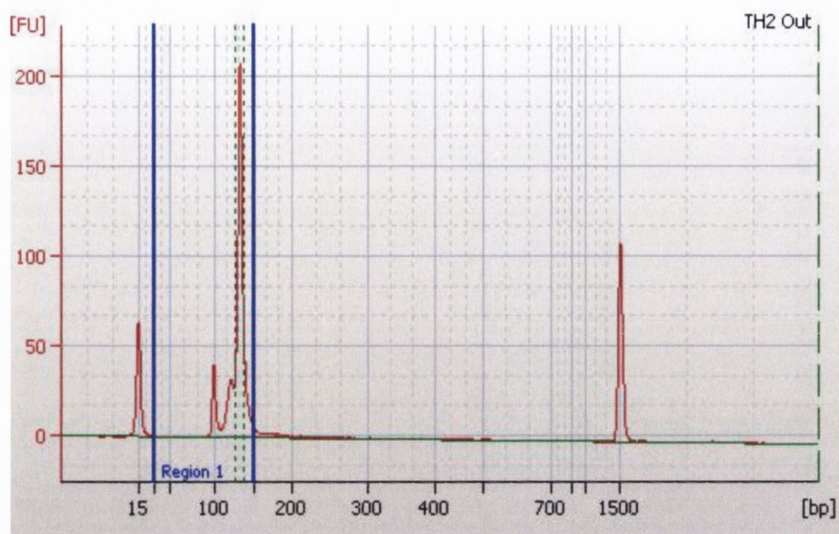


Figure 2.9.4: Typical Agilent 2100 Bioanalyser size profile of a successfully amplified library

Barcoded libraries were calculated in nM:

$$\text{(Library Concentration (ng/}\mu\text{l))} \times \text{(Average base pair size} \times \text{650)}$$

Barcoded libraries were pooled by diluting each library to the lowest library's nM concentration and then combining in equal quantities to obtain a multiplexed library.

2.10 Library Quantitation

Multiplexed libraries were quantitated by Quantitative RT-PCR using the SOLiD™ Library TaqMan® Quantitation Kit (Applied Biosystems, USA), with the manufacturer's protocol supplied in Appendix I. The multiplexed library was quantified and diluted 1/100, 1/1000, 1/10,000 to prepare the input sample within range of the standard curve for library quantitation. Kit-supplied library standards were diluted to 100pM (standard 1) and serially diluted to 0.01pM (standard 5). Per well, 10μl SOLiD Library qPCR mix, 1μl TaqMan assay (Ac00010015_a1), 0.04μl 25μM ROX Reference Dye and 3.96μl nuclease-free water was added. 5μl of qPCR library standard (1-5), sample or water (NTC) was added per well and all wells were performed in triplicate. A fast RT-PCR was run using the 7500HT Fast RT-PCR Thermal Cycler under the following conditions: 95°C for 20 secs, 40 cycles of 95°C for 3 secs, 60°C for 30 secs. A standard curve was generated and the absolute quantitation of samples

determined via regression analysis. Library concentrations were calculated in pM and diluted to 500pM in 1X TE buffer to proceed to emulsion PCR.

2.11 Template Bead Preparation

The full scale template bead protocol (located on Appendix CD) was followed for 0.4pM and 0.8pM libraries, with 4 PCRs per concentration. To determine whether 0.4pM or 0.8pM is optimal, a Work Flow Analysis (WFA) sequencing run was performed. The manufacturer's supplied SOLiD™ System Templated Bead Preparation Guide can be found on the Appendix CD but the steps are briefly described below.

2.11.1 Full Scale Emulsion PCR

The oil phase for ePCR was prepared by gently adding 4.5% Emulsion Stabilizer 1 to 1% Emulsion Stabilizer 2, then adding 94.5% emulsion oil and vortexing to a uniform consistency. The oil phase is then degassed prior to the dispensation of 9ml oil phase to an ePCR tube. The aqueous phase was prepared to a final concentration of 1X PCR Buffer, 14mM dNTP mix, 25 mM Magnesium Chloride, 40nM PCR primer 1, 3μM PCR Primer 2, 0.4pM or 0.8pM small RNA library template, 0.54U/μl AmpliTaq Gold DNA Polymerase in nuclease free water. The aqueous phase per PCR is prepared in a total volume of 5,600μl: 5440μl PCR Reagents and 160μl prepared and sonicated SOLiD™ P1 DNA Beads. The oil phase was placed into a fixed-setting Ultra-Turrax® tube device (IKA, China) and engaged in spinning. An X-Stream® Pipettor (Eppendorf, Germany) was used to add 5.6ml aqueous phase to the spinning oil phase and spun for 5 mins to create the emulsion. A Repeater® Plus Pipettor (Eppendorf, Germany) was used to pipette 150μl emulsion per well of a 96-well MicroAmp PCR plate and sealed with MicroAmp clear adhesive film. ePCR was run on a 9700 thermal cycler under the following conditions: 95°C for 5 mins; 40 cycles of 93°C for 15 secs, 62°C for 30 secs, 72°C for 75 secs, 72°C for 7 mins; and 4°C for ∞ with reaction volume set to 50μl and a ramp speed of 9600. Following the ePCR run the bottom of the PCR plates were checked for broken emulsions, which appear as brown/amber specks. If more than 3 wells were broken, the ePCR was to be repeated.

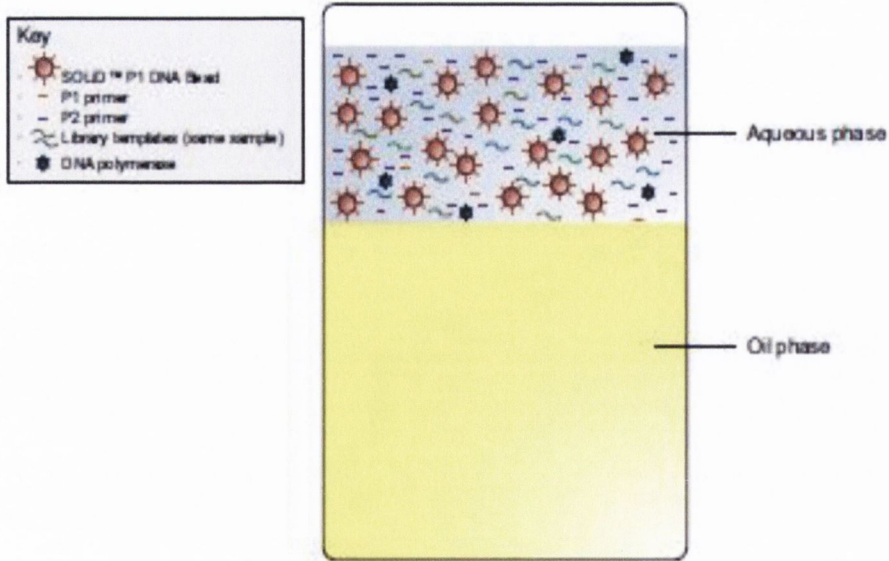


Figure 2.11.1(a): Aqueous phase and oil phase prior to emulsification

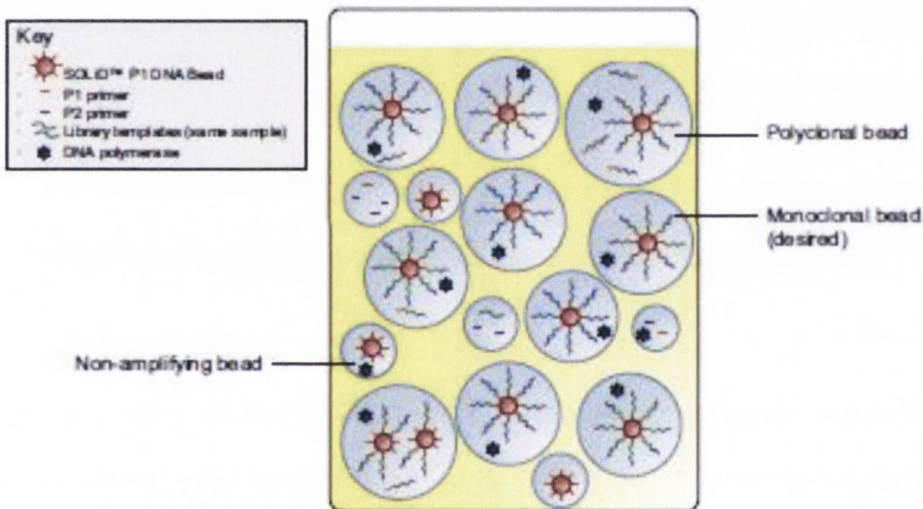


Figure 2.11.1(b): Emulsion after amplification

2.11.2 Emulsion Breaking and Bead Washing

An emulsion collection tray was attached to the PCR plate, inverted and centrifuged at 550g for 2 mins at room temperature with a high acceleration and low deceleration centrifuge setting. The PCR plate was carefully removed from the collection tube and, in a fume hood, 10ml of 2-Butanol was added to the collection tray using a 10ml serological pipette. The solution was mixed via pipetting until homogenous and then transferred to a 50ml tube. Another 6ml of 2-Butanol was added to the collection tube to wash the tray and remove any

remaining emulsion then transferred to the 50ml tube. The tube was then vortexed and centrifuged at 300g for 5 mins to pellet the beads. Supernatant was discarded and the tube inverted onto a paper towel to maximise oil removal.

Pellets were resuspended in 600µl 1x bead wash buffer, incubated for 2 mins then transferred to a LoBind 1.5ml tube. The 50ml tube was rinsed with a further 600µl 1X bead wash buffer and added to the 1.5ml tube. The tube was vortexed and centrifuged at 21,000g for 1 min. The oil phase was removed as much as possible without disturbing the aqueous layer. The remaining supernatant was carefully removed, the pellet resuspended in 1X bead wash buffer, vortexed and pulse-spun before being transferred to a fresh tube. The bottom of the original tube was rinsed with 1X bead buffer and transferred to the new tube. A further 1ml 1X bead buffer was added to the new tube and vortexed to mix. The tube was centrifuged for 21,000g for 1 min and supernatant was discarded. Beads were resuspended in 200µl 1X TEX Buffer and placed on a magnetic rack. When the solution cleared, the supernatant was decanted and samples were removed from the magnetic rack before adding 200µl 1X TEX Buffer.

2.11.3 SOLiD™ Bead Quantitation

Surplus sonicated template beads were counted using a haemocytometer. Beads were diluted in 1X TEX Buffer to 6 standard concentrations ranging from $2 \times 10^5/\mu\text{l}$ – $1.2 \times 10^6/\mu\text{l}$. The standards were read 3 times each on the NanoDrop® ND-1000 spectrophotometer under the cell culture tab using 2µl per read and a standard curve was generated which was used to measure all future bead concentrations. The ePCR amplified template bead sample was sonicated to declump the sample and 1ml dilution of 1:10 beads: 1X TEX Buffer sample was prepared. The colour of the sample was then compared to the SOLiD Bead Concentration Chart (Figure 2.11.3(a)) and diluted, if required, to within the 7.5×10^5 – 1.25×10^6 range. The beads were then quantified in triplicate on the NanoDrop®, the results averaged and the concentration calculated from the standard curve. The diluted and undiluted beads were then combined.

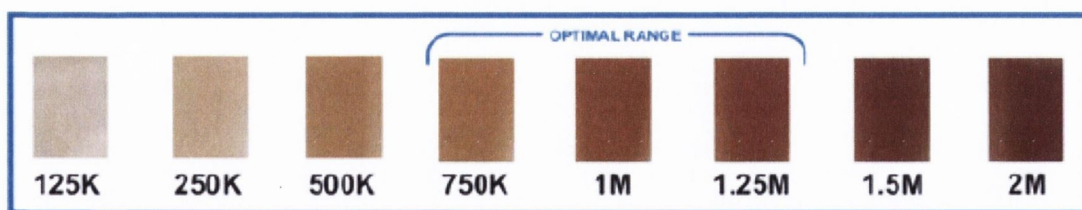


Figure 2.11.3(a): SOLiD™ Bead Concentration Chart

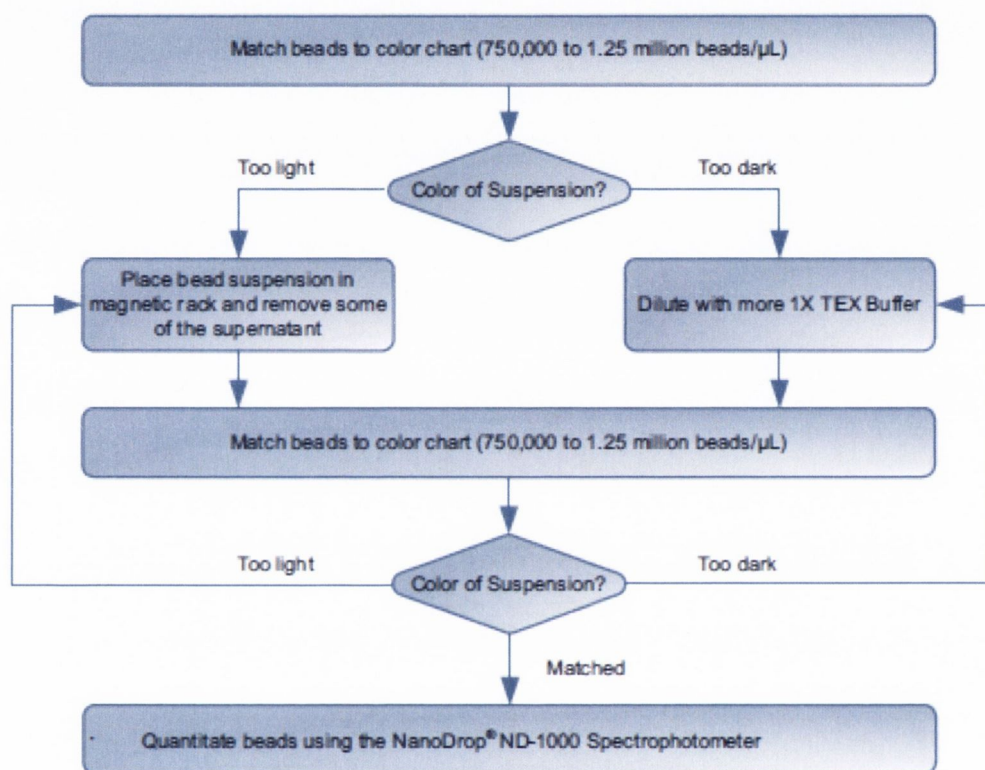


Figure 2.11.3(b): SOLiD™ Bead Concentration Workflow

2.11.4 SOLiD™ Templated Bead Enrichment

Beads were placed onto a magnetic rack until the solution cleared and the supernatant was discarded. Beads were resuspended in 300μl Denaturing Buffer Solution (10% denaturant, 90% Denaturing Buffer), incubated for 1 min, placed on the magnetic rack for 1 min and the supernatant discarded. This step was repeated twice. Beads were then washed in 1X TEX Buffer 3 times, then resuspended in 150μl 1X TEX Buffer, transferred to a 500μl LoBind tube and sonicated to declump the beads.

Prepared enrichment beads in 150μl 1X low salt binding buffer (per ePCR) were added to the template beads. Beads were vortexed, pulse-spun, sonicated, pulse-spun then

incubated at 61°C for 15 mins, with a vortex and pulse-spin every 5 mins. The beads were snap-cooled on ice for 2 mins then gently layered onto 60% glycerol in a 1.5ml LoBind tube and centrifuged for 3 mins at 21,000g. The template beads were separated onto the top layer and untemplated beads are sedimented on the bottom of the tube. The top layer was carefully removed to a 2ml LoBind tube containing 1ml 1X TEX Buffer. The tube was filled to 2mls with 1X TEX Buffer, vortexed and centrifuged at 21,000g for 1 min and the beads pelleted. Supernatant was removed and beads were resuspended in 400µl 1X TEX Buffer.

2.11.5 P2 Enriched Bead Isolation

Samples were centrifuged at 21,000g for 1 min, supernatant was discarded and resuspended in denaturing buffer solution. Samples were placed in a magnetic rack for 1 min and the supernatant discarded. This step was repeated until the supernatant appeared clear. Beads were washed in 1X TEX Buffer 3 times, then resuspended in 200µl 1X TEX Buffer and transferred into a 1.5ml LoBind tube. The 2ml tube was rinsed with 200µl 1X TEX Buffer, which was added to the new 1.5ml tube. Beads were sonicated to declump the sample and placed in a magnetic rack until the supernatant was clear and was then discarded. Beads were resuspended in 400µl 1X TEX Buffer.

2.11.6 3' End Modification

Sonicated beads were isolated in a magnetic rack and resuspended in 1X terminal transferase reaction buffer (terminal transferase buffer, cobalt chloride and water) and transferred into a fresh 1.5ml tube. Beads were magnetically isolated and again resuspended in 100µl terminal transferase buffer, this step was repeated once. Beads were isolated magnetically and resuspended in 178µl terminal transferase buffer and 20µl 1mM Bead Linker Solution. Beads were sonicated and 2µl terminal transferase (20U/µl) was added. Samples were incubated on a tube rotator for 2 hours at 37°C, gently rotating. Samples were placed on the magnetic rack for 1 min and the supernatant discarded. The beads were washed in 1X TEX Buffer before being resuspended in 400µl TEX.

2.11.7 Modified SOLiD™ Bead Quantitation

3' end modified beads were quantified as per section 2.11.3.

2.12 SOLiD™ 4 Next Generation Sequencing Run

The manufacturer's supplied protocol for SOLiD sequencing can be found on the Appendix CD and is briefly outlined in the sections below.

2.12.1 Workflow Analysis (WFA)

The WFA determined the optimal library concentration and the bead efficiency. WFA determined whether the 0.4pM or 0.8pM library should be used and the fraction of successfully amplified beads as a proportion of the total beads. This was then used to determine the optimal input amount of beads for the full sequencing run. To perform the WFA, each library was deposited at a density of 1.5×10^7 beads per quadrant of the sequencing run slide and then subjected to a single SOLiD™ ligation cycle.

2.12.1.1 Preparation of SOLiD™ XD Slides

1.5×10^7 declumped beads were transferred to a fresh 1.5ml LoBind tube then stored and isolated on a magnetic rack. Supernatant was discarded and beads resuspended in 400 μ l SOLiD™ XD Deposition Buffer. Beads were placed on a magnetic rack and isolated and supernatant was discarded which was repeated twice. SOLiD™ XD Slides were brought to room temperature. Beads were resuspended in the correct volume of SOLiD™ XD Deposition Buffer dependent upon slide carrier. The slide carrier was prepared and the XD slide assembled into the deposition chamber. Beads were sonicated twice to declump and immediately pipetted into the entry porthole. Deposition chambers were centrifuged at 167g for 5 mins and incubated at 37°C for 1 hour. The slide carrier was removed from the deposition chamber and gently washed twice with slide storage buffer. The carrier was loaded onto the cleaned, prepared flow cell and secured into place. The flowcell was locked into the scan position, the instrument doors were closed and the flowcell loaded with buffer to prevent it from drying out.

2.12.1.2 WFA

The manufacturer's supplied protocol (Appendix CD) provides the computational steps required to set up a WFA run using the SOLiD Instrument Control Software (ICS) using a 4-spot WFA mask, with each spot representing a well and each well containing template beads

from a single ePCR. The focal range was detected and confirmed by visualising the beads. The WFA run was started by selecting the start run tab for the appropriate flowcell. The WFA run was monitored as per the manufacturer's supplied protocol and the library concentration (0.4pM or 0.8pM) with the higher titration matrix was used for full sequencing. The volume of beads deposited was calculated based on WFA results as follows:

- (1) Calculate the P2 positive (P2⁺) beads/μl in each ePCR sample (each mask/well)

P2 positive per panel [from WFA]

X

426 panels Input volume of beads used for WFA run (μl)

- (2) For a single well slide, deposition of 708 million beads is optimal. Approximately 20% of beads are lost in the washing and deposition preparation steps. Hence the input volume of bead solution is calculated as follows:

708 million P2 positive beads /μl x 120%

2.12.2 SOLiD™ 4 Sequencing Run

The steps for the sequencing run are the same as those for WFA, using a single well deposition slide and adjusting the bead volume as per 2.12.1.2. The ICS is used to set-up a multiplex sequencing run as barcoded libraries were used. During this step, barcodes were matched to libraries. The focal plane is detected as before and a sequencing run is performed using barcode tag (BC Tag MM10) primers. This initial step differs only by strip reagents used and the selected sequencing run on the ICS. When this initial run was finished and saved, the strip reagents were changed (Fragment library 35bp read (F3 tag MM35) primers) and instrument reagents were checked to ensure sufficient quantities remain. The focal plane was again detected and the full sequencing run set up and run using the ICS. The run was monitored as per the manufacturer's supplied protocol (Appendix CD). Primary analysis of the sequencing run was performed using the ICS software as per protocol (Appendix CD). Following assessment of sequencing run success, full computational analysis of the SOLiD Next Generation Sequencing data was analysed bioinformatically.

The SOLiD sequencing platform generates two sequence data files for each barcoded library. These FASTA formatted files contain colorspace-encoded base-calls, and the corresponding PHRED-scaled quality scores, respectively, for each sequenced read. Base-call accuracy was verified and enhanced using ABI's proprietary SOLiD Accuracy Enhancement Tool (SAET). SAET attempts to correct mis-calls within reads prior to mapping, thereby increasing the number and quality of mappable reads in each sample.

Read-Mapping

Read-mapping was performed using the SHRiMP (Short Read Mapping Program) alignment package. SHRiMP natively supports colorspace reads, supports local alignment, and has built-in parameter settings optimised specifically for when dealing with miRNA-sized DNA fragments. Reads from each input sample were first mapped against sources of potential contamination, to filter out fragments of unwanted RNA species, mitochondrial DNA, and adapter sequences. The remaining reads were then mapped against the collection of human non-coding RNA species, downloaded from ENSEMBL. Aligned reads were saved in .bam file format.

Gene-Counting

Gene counting was performed using HTSeq, a software tool written in Python to process data from high-throughput sequencing assays. HTSeq calculates the number of aligned reads contained within a BAM file that map to certain genomic features, i.e. a set of genes, and presents these values as raw gene counts. The list of genes and their corresponding genomic locations was obtained in GTF format from ENSEMBL.

Expression Analysis

Differential expression analysis was performed using the Bioconductor library DESeq. DESeq tests for differential expression by use of the negative binomial distribution and a shrinkage estimator for the variance of the distribution. DESeq takes as input a table of raw gene counts as produced by HTSeq. Each column in the table corresponds to a single input sample. The rows correspond to the list of genes that have been quantified in the previous step. Differential expression analysis was carried out successively on each pair of samples

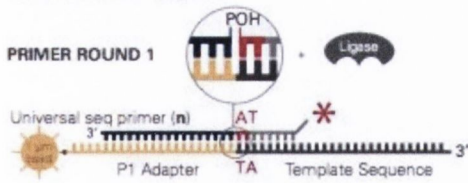
under comparison, with a threshold of \log_2 -fold change of 1.0. (p-value was not used as criteria due to lack of replicates).

Analysis of miRNA-targeted Genes

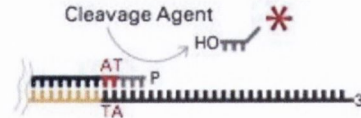
MiRNA species were extracted from the list of deregulated genes and submitted to miRWalk (<http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/index.html>), an online database of known and predicted miRNA targets. Analysis was restricted to the database of experimentally validated miRNA-target interactions.

Functional categorisation and pathways analysis was performed on the returned lists of validated miRNA-targeted genes. This process was carried out using the Bioconductor library GeneAnswers. GeneAnswers was used to generate lists of enriched GO terms for each gene set.

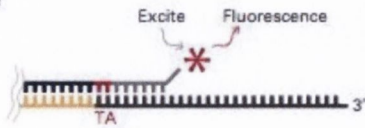
1. Prime and Ligate



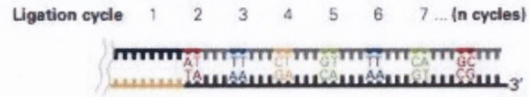
4. Cleave off Fluor



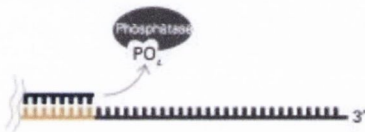
2. Image



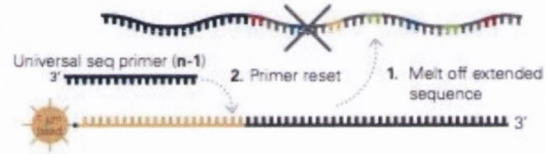
5. Repeat steps 1-4 to Extend Sequence



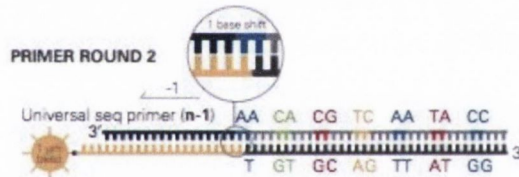
3. Cap Unextended Strands



6. Primer Reset



7. Repeat steps 1-5 with new primer



8. Repeat Reset with , n-2, n-3, n-4 primers

	Read Position	0	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	26	27	28	29	30	31	32	33	34	35			
Primer Round	1	Universal seq primer (n)	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
	2	Universal seq primer (n-1)	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	3	Universal seq primer (n-2)	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	4	Universal seq primer (n-3)	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	5	Universal seq primer (n-4)	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	

• Indicates positions of interrogation Ligation Cycle: 1 (blue), 2 (orange), 3 (green), 4 (red)

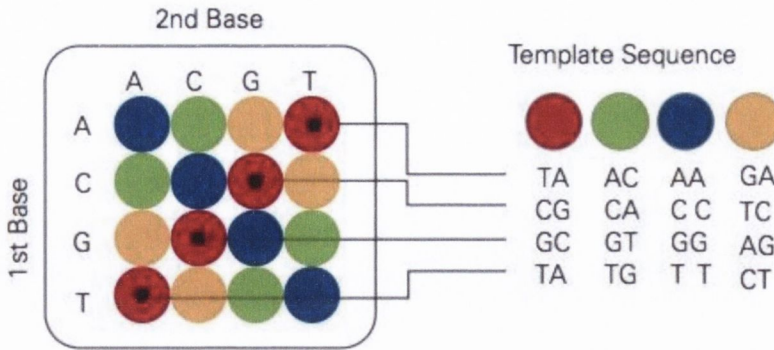
Figure 2.12.2(a): SOLiD™ Sequencing Chemistry

SOLiD uses a mixture of labeled oligonucleotides and queries the input strand with ligase. Each oligo has degenerate positions at bases 3-5 (N's), and one of 16 specific dinucleotides at positions 1-2 (numbered from the 3' end). Positions 6 through the 5' are also degenerate, and hold one of four fluorescent dyes. The sequencing involves:

1. Anneal a primer, then hybridize and ligate a mixture of fluorescent oligos (8-mers) whose 1st & 2nd 3' bases match that of the template
2. Capping unextended fragments with the same mixture of non-fluorescent probes
3. Phosphatase treatment to prevent any remaining unextended strands from contributing to out of phase ligation events
4. Detection of the specific fluor
5. Removal of fluor via two step chemical cleavage of the three 5' bases. This leaves behind a 5 base ligated probe, with a 5' phosphate
6. Repeat, this time querying the 6th & 7th bases
7. After 5-7 cycles of this, perform a "reset", in which the initial primer and all ligated portions are melted from the template and discarded.
8. Next a new initial primer is used that is N-1 in length. Repeating the initial cycling (steps 1-5) now generates an overlapping data set (bases 1/2, 6/7, etc, see Figure 2.12.2(a), Step 8 above).

Thus, 5-7 ligation reactions followed by 5 primer reset cycles are repeated generating sequence data for ~35 contiguous bases, in which each base has been queried by two different oligonucleotides. There are 16 possible dinucleotides (4²) and only 4 dyes so data from a single colour call does not tell what base is at a given position. There are 4 oligos for every dye, meaning there are four dinucleotides that are encoded by each dye.

Possible Dinucleotides Encoded By Each Color



Double Interrogation

With 2 base encoding each base is defined twice

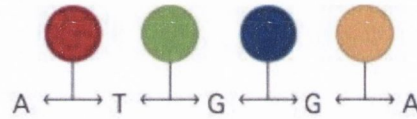


Figure 2.13.2(b): Dibase encoding and how it relates to calling the actual template sequence. In this example, dinucleotides CA, AC, TG and GT, are all encoded by the green dye. Because each base is queried twice, it is possible using the two colours, to determine which bases were at which positions. One of the side effects of this dual encoding is that when aligning to a reference and attempting to determine variants, true variants will follow specific colour change “rules”

Chapter 3

Gene Expression in Holoclones Derived From Cell Lines +/- BRAF V600E Mutation

3.1 Introduction

BRAF V600E mutation has been shown to have many downstream effects on cell signalling and its prevalence varies across cancer types.

Cancer stem cells present an attractive target for developing therapies to combat tumour (re)population and so by isolating cancer stem cells in wild type and BRAF V600E cell lines, we will also identify dysregulated genes and pathways central to the attainment and maintenance of stemness characteristics. This work proposes to generate greater insights into BRAF mediated biological pathways and molecular mechanisms that regulate cell fate, development, and disease progression.

Aims & Objectives

- Establish a cell line panel comprised of a wild type and V600E mutated BRAF Melanoma, Thyroid, Ovarian and Colorectal cell lines confirmed by BRAF SNP Genotyping
- Generate cancer stem cell progenies in the form of holoclones, from parental cells using a positive selection high salt agar assay
- Investigate expression of a panel of genes involved in various different biological processes involved in carcinogenesis such as stemness, Epithelial to Mesenchymal Transition (EMT) and cell signalling
- Investigate protein expression of the most consistently up regulated gene in parental cells and holoclones via confocal microscopy

3.2 TaqMan® SNP Genotyping of Cell Line Panel

Genomic DNA from all cell lines was SNP genotyped using allelic discrimination to confirm the status of the BRAF gene at nucleotide position 1799. The result was a wild type or SNP and genotype call accompanied by an associated measure of certainty (described by “call quality percentage”), in which a 100% score indicated 100% confidence in a base call.

3.2.1 Melanoma

Melanoma cell lines COLO794 and SK-Mel 28 underwent SNP genotyping to confirm BRAF status along with non-template controls (NTC) and known samples for other possible outcomes (homozygote allele 1 and 2) were included for control purposes.

Sample Name	SNP Assay Name	Call Quality (%)	Call
COLO794	BRAF 1799	100	Homozygous 1/1
COLO794	BRAF 1799	100	Homozygous 1/1
SK-Mel 28	BRAF 1799	100	Homozygous 2/2
SK-Mel 28	BRAF 1799	100	Homozygous 2/2
Control Name	SNP Assay Name	Call Quality (%)	Call
8505C	BRAF 1799	100	Homozygous 2/2
8505C	BRAF 1799	100	Homozygous 2/2
N-Thyori	BRAF 1799	100	Homozygous 1/1
N-Thyori	BRAF 1799	100	Homozygous 1/1
NTC	BRAF 1799	100	NTC
NTC	BRAF 1799	100	NTC

Table 3.2.1(a): SNP genotyping data for melanoma cell lines COLO794 and SK-Mel 28

The SNP genotyping assay confirmed that:

- Cell line COLO794 harboured a T nucleotide at position 1799 on both alleles and was BRAF homozygous wild type.
- Cell line SK-Mel 28 harboured an A on both alleles and was BRAF homozygous mutant.
- NTCs and other known samples all had expected observations with high quality scores of 100%.

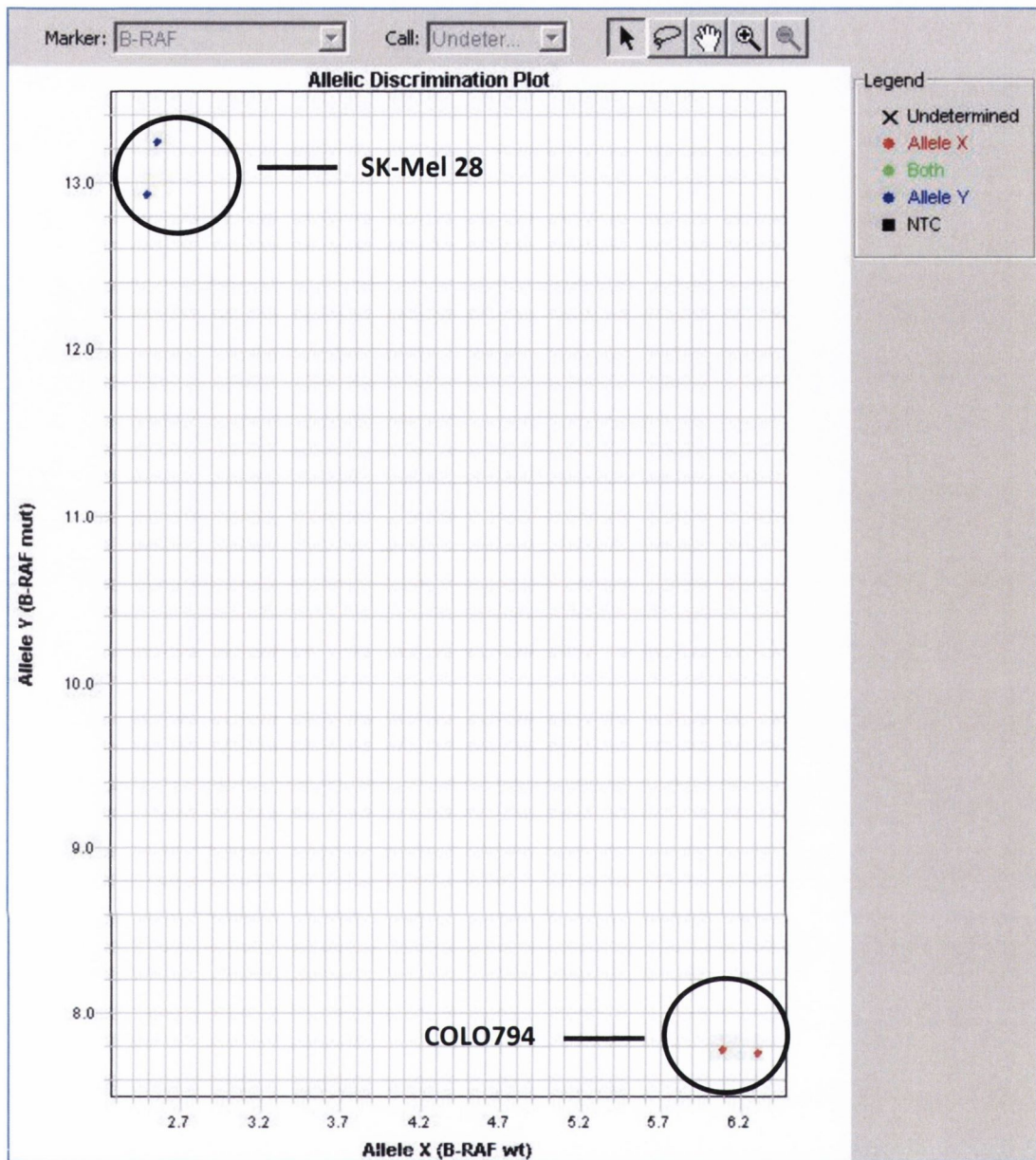


Figure 3.2.1(b): Allelic Discrimination plot illustrating the BRAF status of melanoma cell lines COLO794 and SK-Mel 28.

The allelic discrimination software interpreted the COLO794 cell line as having wild type BRAF status or a T nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 1/1 or **Allele X**.

The allelic discrimination software called the SK-Mel 28 cell line as having mutant BRAF or an A nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 2/2 or **Allele Y**.

3.2.2 Thyroid

Thyroid cell lines TPC-1 and 8505C underwent SNP genotyping to confirm BRAF status along with non-template controls (NTC) and known samples for other possible outcomes (heterozygote and homozygote allele 1) were included for control purposes.

Sample Name	SNP Assay Name	Call Quality (%)	Call
TPC-1	BRAF 1799	100	Homozygous 1/1
TPC-1	BRAF 1799	99.98	Homozygous 1/1
8505C	BRAF 1799	99.62	Homozygous 2/2
8505C	BRAF 1799	99.82	Homozygous 2/2
Control Name	SNP Assay Name	Call Quality (%)	Call
HT-29	BRAF 1799	99.26	Heterozygous 1/2
HT-29	BRAF 1799	99.47	Heterozygous 1/2
N-Thyori	BRAF 1799	99.26	Homozygous 1/1
N-Thyori	BRAF 1799	99.57	Homozygous 1/1
NTC	BRAF 1799	100	NTC
NTC	BRAF 1799	100	NTC

Table 3.2.2(a): SNP genotyping data for thyroid cell lines TPC-1 and 8505C

The SNP genotyping assay confirmed that:

- Cell line TPC-1 harboured a T nucleotide at position 1799 on both alleles and was consequently BRAF homozygous wild type.
- Cell line 8505C harboured an A on both alleles and was BRAF homozygous mutant.
- No template controls (NTC) and known samples for other possible outcomes (heterozygote, homozygote allele 1) all had expected observations with high quality scores of 99% or greater.

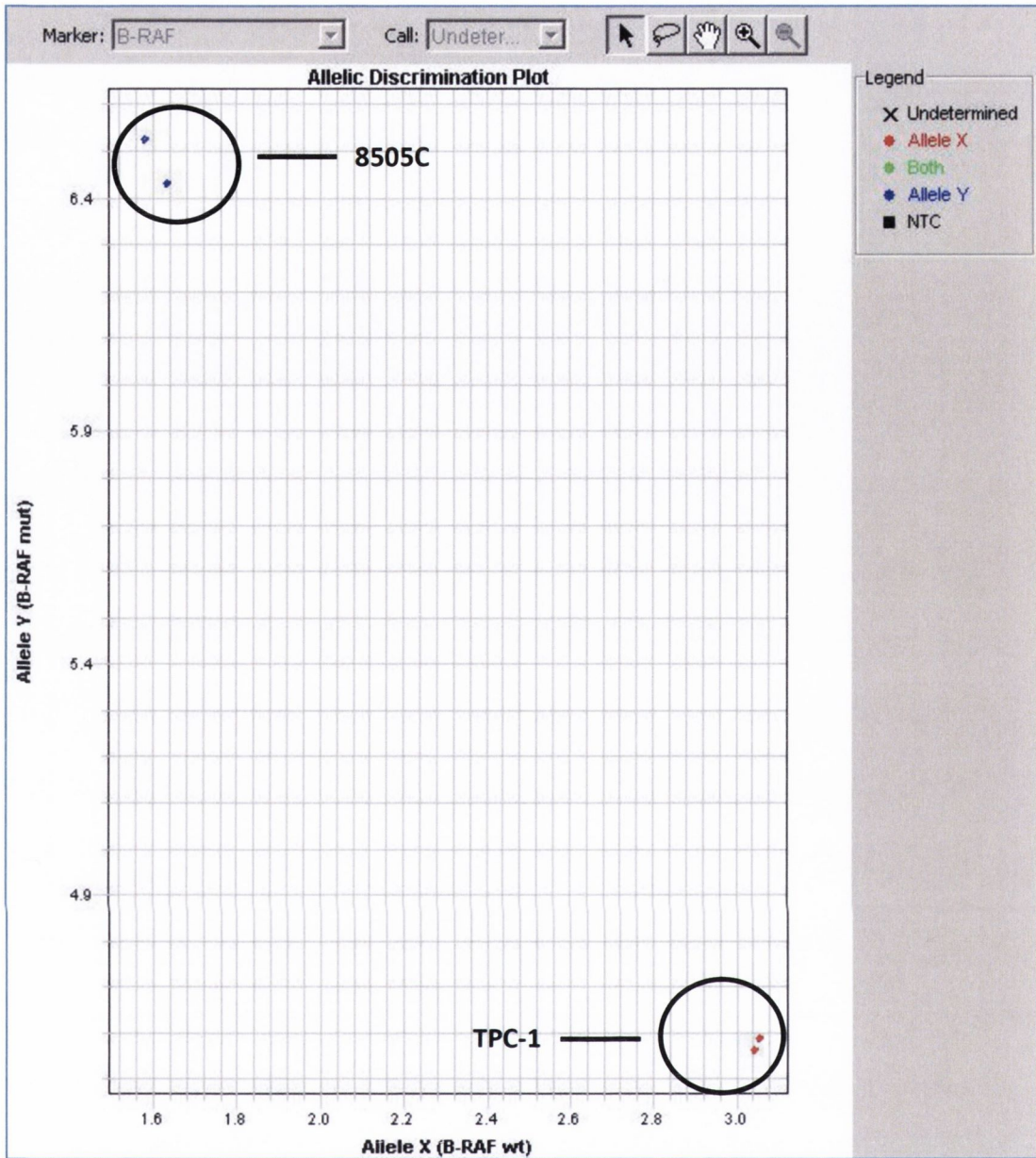


Figure 3.2.2(b): Allelic Discrimination plot illustrating the BRAF status of thyroid cell lines TPC-1 and 8505C.

The allelic discrimination software called the TPC-1 cell line as having wild type BRAF status or a T nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 1/1 or **Allele X**.

The allelic discrimination software called the 8505C cell line as having mutant BRAF or an A nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 2/2 or **Allele Y**.

3.2.3 Ovarian

Ovarian cell lines A2780 and ES-2 underwent SNP genotyping to confirm BRAF status along with non-template controls (NTC) and known samples for other possible outcomes (heterozygote, homozygote allele 1 and 2) were included for control purposes.

Sample Name	SNP Assay Name	Quality (%)	Call
A2780	BRAF 1799	99.84	Homozygous 1/1
A2780	BRAF 1799	99.99	Homozygous 1/1
ES-2	BRAF 1799	100	Homozygous 2/2
ES-2	BRAF 1799	100	Homozygous 2/2
Control Name	SNP Assay Name	Quality (%)	Call
HT-29	BRAF 1799	99.86	Heterozygous 1/2
HT-29	BRAF 1799	99.85	Heterozygous 1/2
8505C	BRAF 1799	99.71	Homozygous 2/2
8505C	BRAF 1799	100	Homozygous 2/2
N-Thyori	BRAF 1799	98.06	Homozygous 1/1
N-Thyori	BRAF 1799	98.55	Homozygous 1/1
NTC	BRAF 1799	100	NTC
NTC	BRAF 1799	100	NTC

Table 3.2.3(a): SNP genotyping data for ovarian cell lines A2780 and ES-2

The SNP genotyping assay confirmed that:

- Cell line A2780 harboured a T nucleotide at position 1799 on both alleles and was consequently BRAF homozygous wild type.
- Cell line ES-2 harboured an A on both alleles and was BRAF homozygous mutant.
- No template controls (NTC) and known samples for other possible outcomes (heterozygote, homozygote allele 1) all had expected observations with high quality scores of 98% or greater.

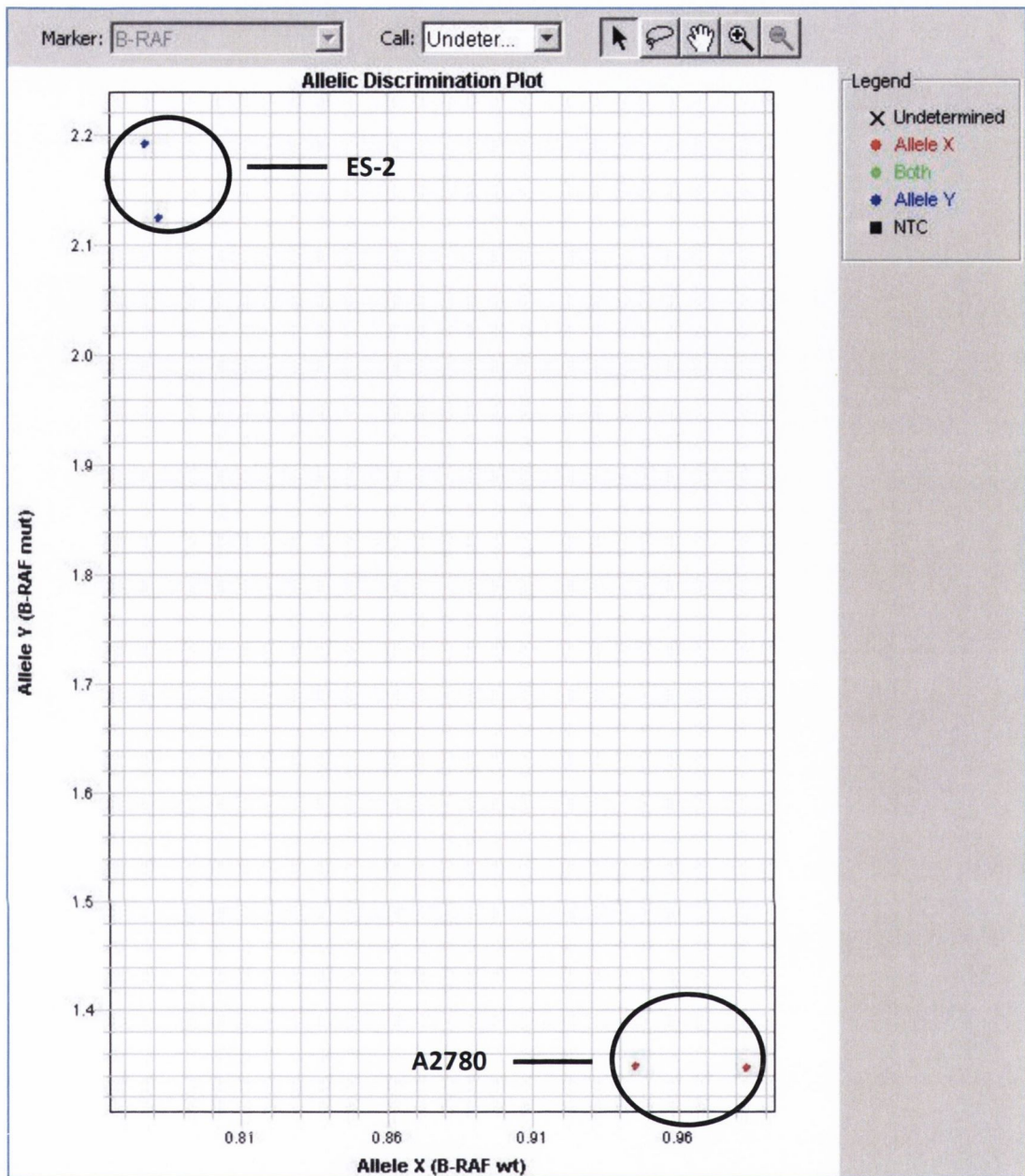


Figure 3.2.3(b): Allelic Discrimination plot illustrating the BRAF status of ovarian cell lines A2780 and ES-2.

The allelic discrimination software called the A2780 cell line as having wild type BRAF status or a T nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 1/1 or **Allele X**.

The allelic discrimination software called the ES-2 cell line as having mutant BRAF or an A nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 2/2 or **Allele Y**.

3.2.4 Colorectal

Colorectal cell lines COLO320 and HT-29 underwent SNP genotyping to confirm BRAF status along with non-template controls (NTC) and known samples for other possible outcomes (homozygote allele 1 and 2) were included for control purposes.

Sample Name	SNP Assay Name	Quality (%)	Call
COLO320	BRAF 1799	100	Homozygous 1/1
COLO320	BRAF 1799	100	Homozygous 1/1
HT-29	BRAF 1799	99	Homozygous 2/2
HT-29	BRAF 1799	100	Homozygous 2/2
Control Name	SNP Assay Name	Quality (%)	Call
N-Thy	BRAF 1799	100	Homozygous 1/1
N-Thy	BRAF 1799	100	Homozygous 1/1
8505C	BRAF 1799	100	Homozygous 2/2
8505C	BRAF 1799	99	Homozygous 2/2
NTC	BRAF 1799	100	NTC
NTC	BRAF 1799	100	NTC

Table 3.2.4(a): SNP genotyping data for colon cell lines COLO320 and HT-29

The SNP genotyping assay confirmed that:

- Cell line COLO320 harboured a T nucleotide at position 1799 on both alleles and was consequently BRAF homozygous wild type.
- Cell line HT-29 harboured an A on both alleles and was BRAF homozygous mutant.
- No template controls (NTC) and known samples for other possible outcomes (heterozygote, homozygote allele 1) all had expected observations with high quality scores of 99% or greater.

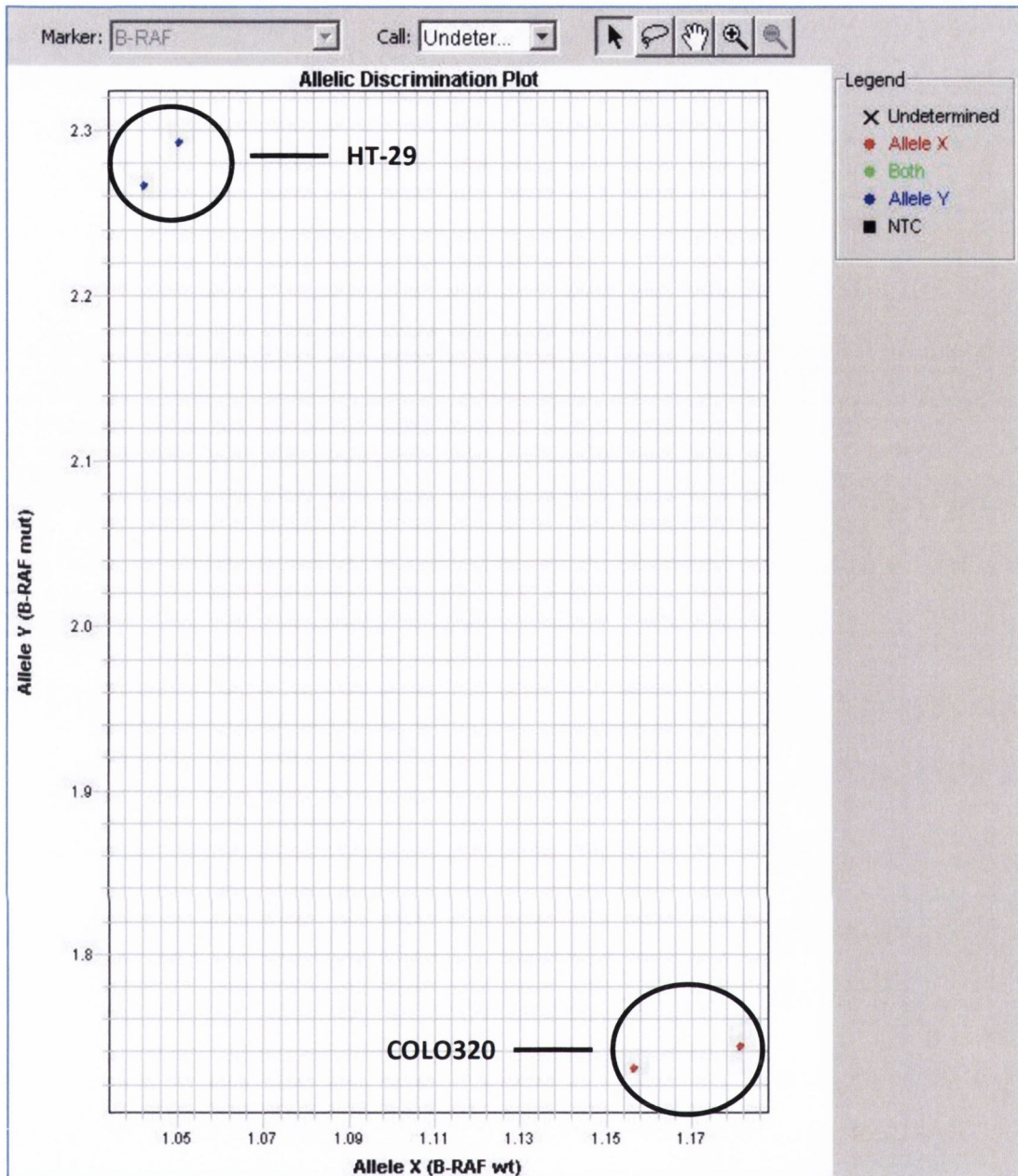


Figure 3.2.4(b): Allelic Discrimination plot illustrating the BRAF status of colorectal cell lines COLO320 and HT-29.

The allelic discrimination software called the COLO320 cell line as having wild type BRAF status or a T nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 1/1 or **Allele X**.

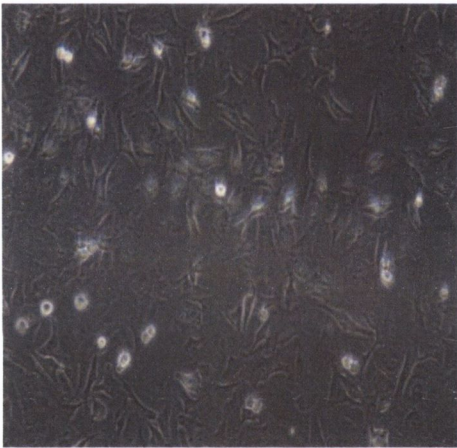
The allelic discrimination software called the HT-29 cell line as having mutant BRAF or an A nucleotide at position 1799. It is therefore illustrated on the above plot as being Homozygous 2/2 or **Allele Y**.

3.3 High Salt Agar Assay

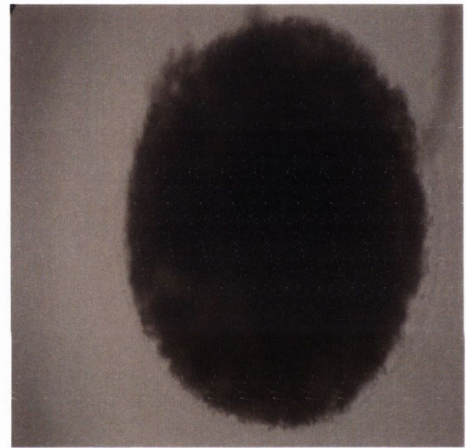
The following images depict the morphology of holoclones that were derived from parental cells using a high salt agar assay. All images of parental cells in culture and holoclones embedded in agarose were captured using 10x objective under a brightfield microscope.

3.3.1 Melanoma

In the wild type BRAF melanoma cell line COLO794, holoclone formation was recorded at 4 weeks. V600E mutated BRAF SK-Mel 28 cell line holoclones were observed before the end of the third week in culture. The SK-Mel 28 cell line has an extremely aggressive growth pattern which may be reflected in its ability to yield holoclones before its wild type BRAF counterpart.



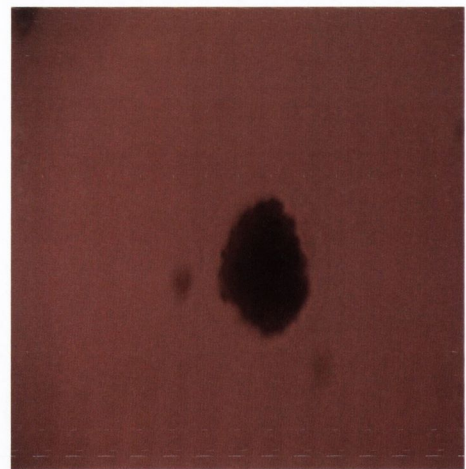
COLO794 Parental Cells in culture



COLO794 Holoclone embedded in agar



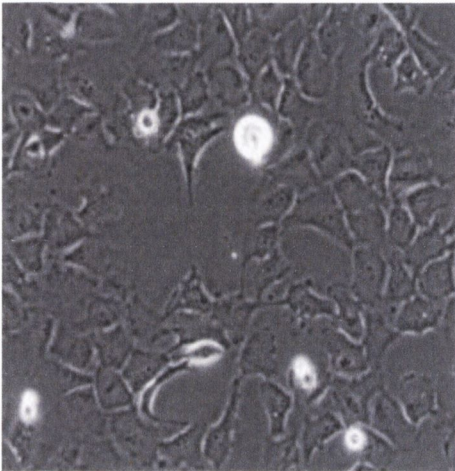
SK-Mel 28 Parental Cells in culture



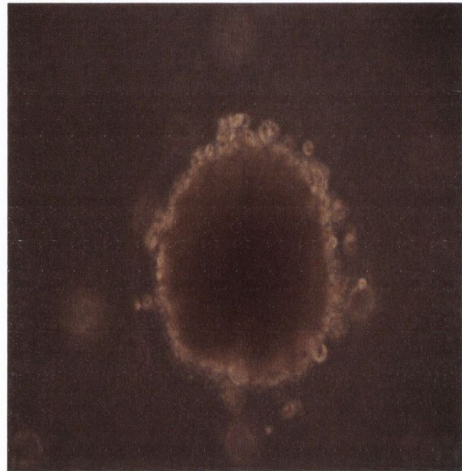
SK-Mel 28 Holoclone embedded in agar

3.3.2 Thyroid

In both the wild type BRAF Papillary Thyroid Carcinoma cell line TPC-1 and V600E mutated BRAF anaplastic thyroid cell line 8505C, holoclone formation was observed at 4 weeks. It is noteworthy that plates were maintained after holoclone extraction for a period of 1-2 weeks and in each case holoclones reformed at the precise site of previous extractions, highlighting their ability for self-renewal.



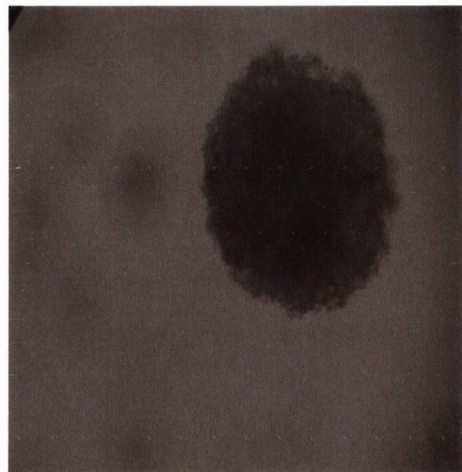
8505C Parental Cells in culture



8505C Holoclone embedded in agar



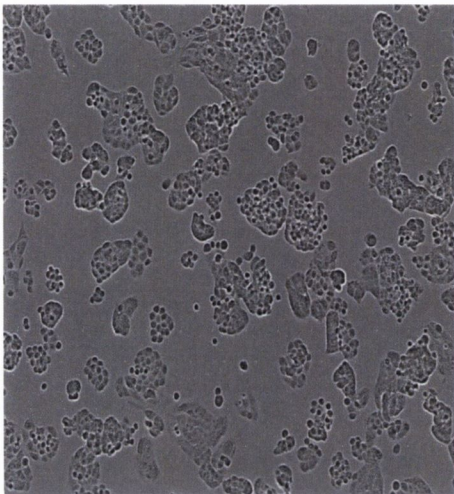
TPC-1 Parental Cells in culture



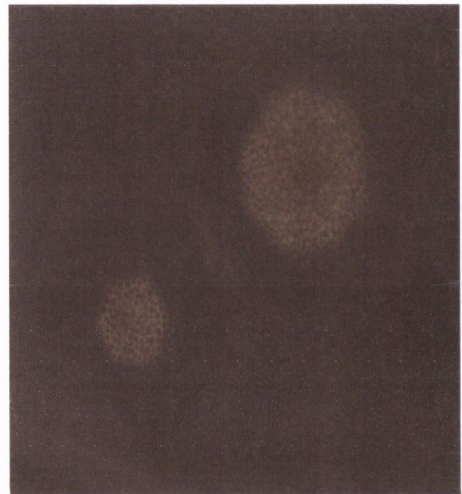
TPC-1 Holoclone embedded in agar

3.3.3 Ovarian

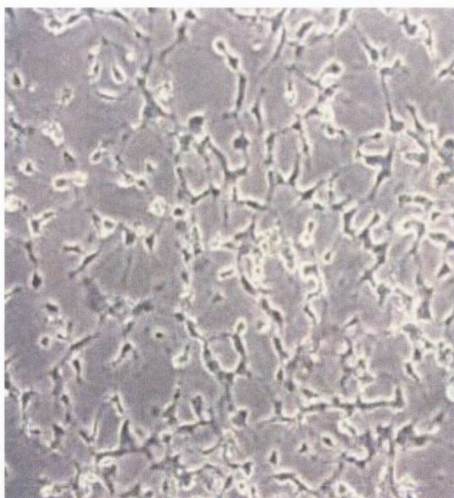
In the wild type BRAF ovarian cell line A2780, holoclone formation was observed after 2 weeks in culture whereas in the V600E mutated BRAF ES-2 cell line, holoclones were seen after 4-5 weeks. The A2780 cell line was derived from ascitic fluid from a patient with metastatic advanced ovarian carcinoma, and this might be a factor in the cell line's rapid rate of holoclone generation compared with the ES-2 cell line derived from a clear cell ovarian carcinoma.



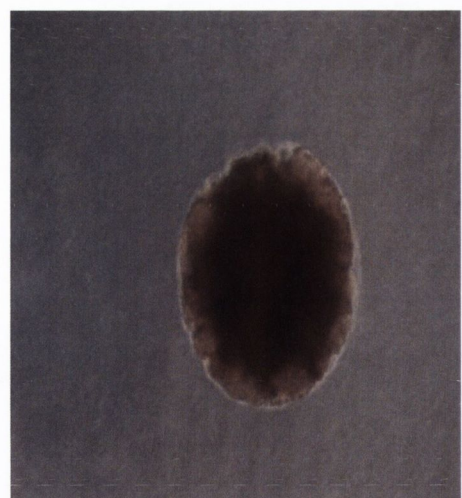
A2780 Parental Cells in culture



A2780 Holoclones embedded in agar



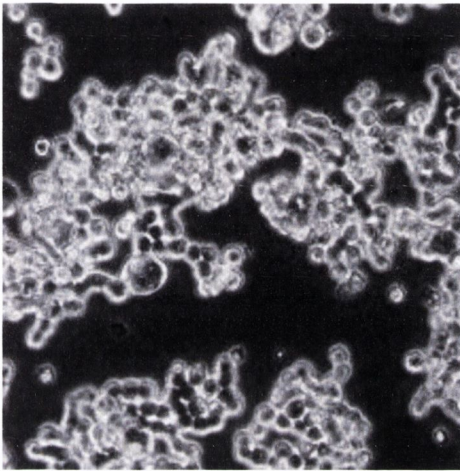
ES-2 Parental Cells in culture



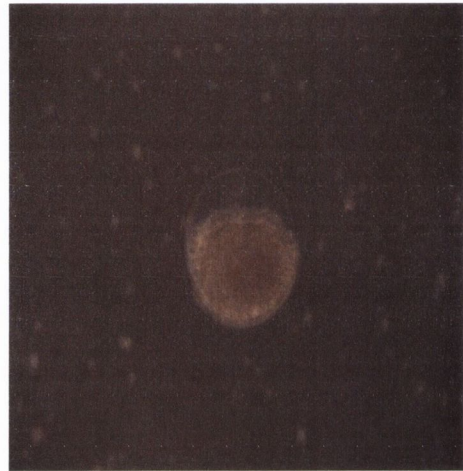
ES-2 Holoclone embedded in agar

3.3.4 Colorectal

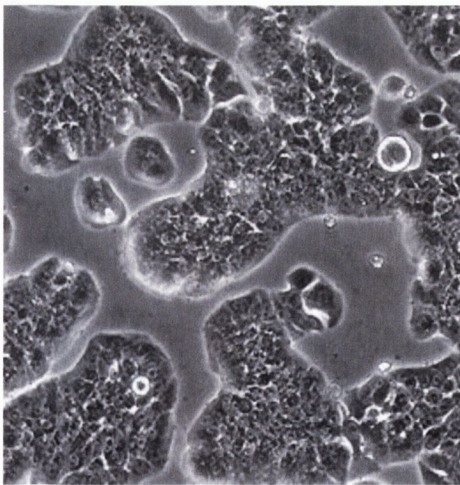
In the wild type BRAF colorectal cell line COLO320, holoclone formation was observed after 2 weeks in culture whereas in the V600E mutated BRAF HT-29 cell line, holoclones were observed at week 4. The COLO320 cell line has previously been shown to have strong colony forming ability due to its high expression levels of c-myc (Collins et al., 1992). Thus it seems plausible that COLO320 parental cells might generate holoclones faster than their HT-29 counterparts despite both cell lines being derived from colon adenocarcinomas.



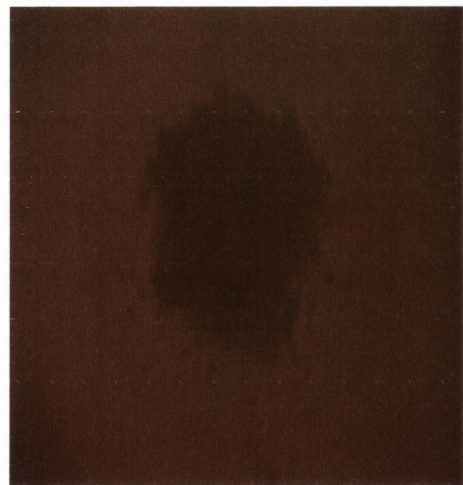
COLO320 Parental Cells in culture



COLO320 Holoclone embedded in agar



HT-29 Parental Cells in culture



HT-29 Holoclone embedded in agar

It is noteworthy that a halo-like structure surrounding holoclones embedded in high salt agar was observed in all 8 cell lines, regardless of BRAF status. Its function is not yet known. Perhaps it is the result of a secretion by the holoclone cells into the surrounding agar and may provide a protective barrier from the high salt conditions which surrounds it. Alternatively, it may represent a reservoir of nutrients absorbed from the media. It may provide a structural support for sustained holoclone growth. These suggestions while each is biologically plausible remain to be verified or excluded.

Given the consistency with which we observed the 'halo', it should be examined in greater detail. Future studies might employ Electron Microscopy to interrogate the structure and composition of the halo and this will undoubtedly provide further insight into CSC idiosyncrasies.

3.4 TaqMan® Gene Expression in Parental Cells & Holoclones

A panel of genes involved in different biological processes were examined in all parent cells and holoclonal cells by quantitative TaqMan RT-PCR. For display and discussion purposes, these genes were divided into 3 groups by function; Stemness, Epithelial to Mesenchymal Transition and Signalling.

Stemness – NANOG, Oct4, ALDH1

EMT – E-Cadherin, SNAI2

Signalling – SHH, TGF- β , β -Catenin

3.4.1 Melanoma

SK-Mel 28 BRAF V600E mutant

Gene Expression in SK-Mel 28 Holoclones v Parent Cells

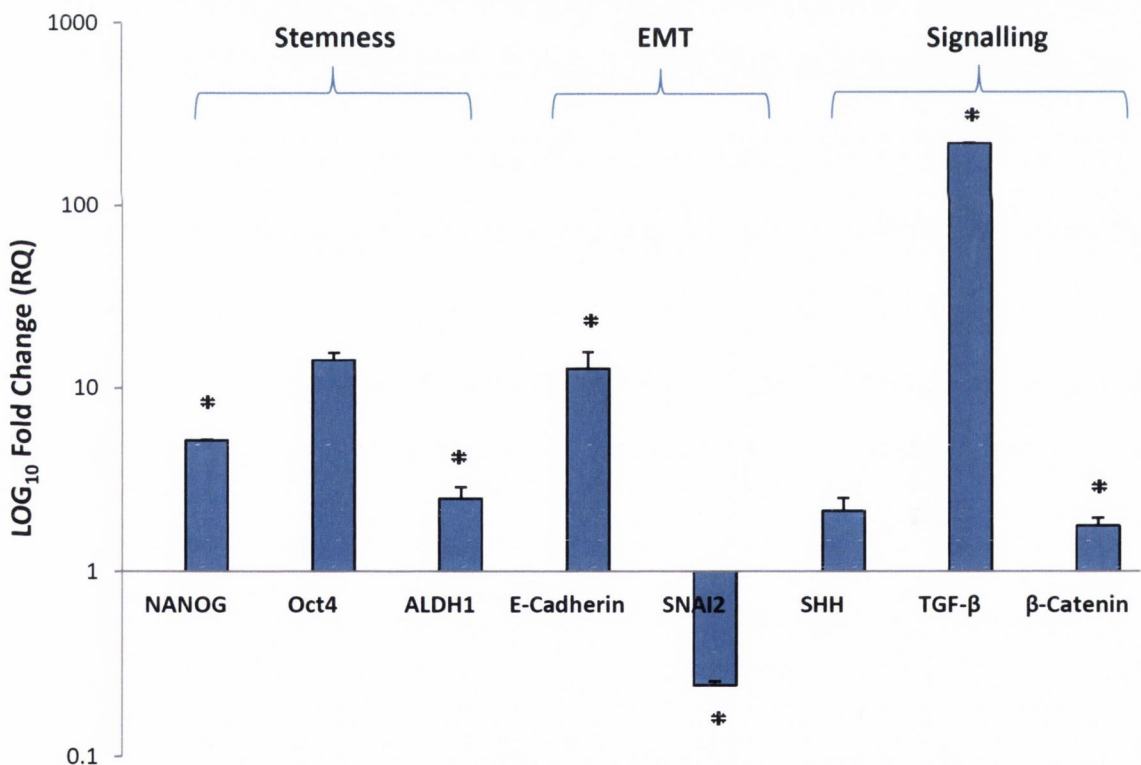


Figure 3.4.1(a): Grouped differential gene expression in SK-Mel 28 BRAF V600E mutated holoclonal cells relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p-value < 0.05 *). Error bars: SEM.

X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were

assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in SK-Mel 28 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Increased expression of pluripotency transcription factors NANOG and Oct4 was recorded in SK-Mel 28 holoclones compared to parent cells. We observed a statistically significant 5-fold increase (p-value 0.01) in NANOG expression and a 14-fold increase in Oct4 levels in the same holoclone extracts. A statistically significant 2-fold increase (p-value 0.01) in expression of stem cell marker ALDH1 gene was detected. EMT genes showed significant differential expression levels with E-Cadherin increasing 12-fold (p-value 0.0002) while SNAI2 was significantly down regulated 4-fold (p-value 0.02) in SK-Mel 28 holoclones.

All 3 cell signalling genes investigated in SK-Mel 28 holoclones showed increased expression in holoclone populations. TGF- β showed the greatest statistically significant gene up regulation with a 218-fold increase (p-value 0.009) relative to its founder cells and a 2-fold increase in SHH expression was recorded in SK-Mel 28 holoclones. A significant 2-fold increase (p-value 0.01) in expression was recorded in β -Catenin, a member of the Wnt signalling pathway.

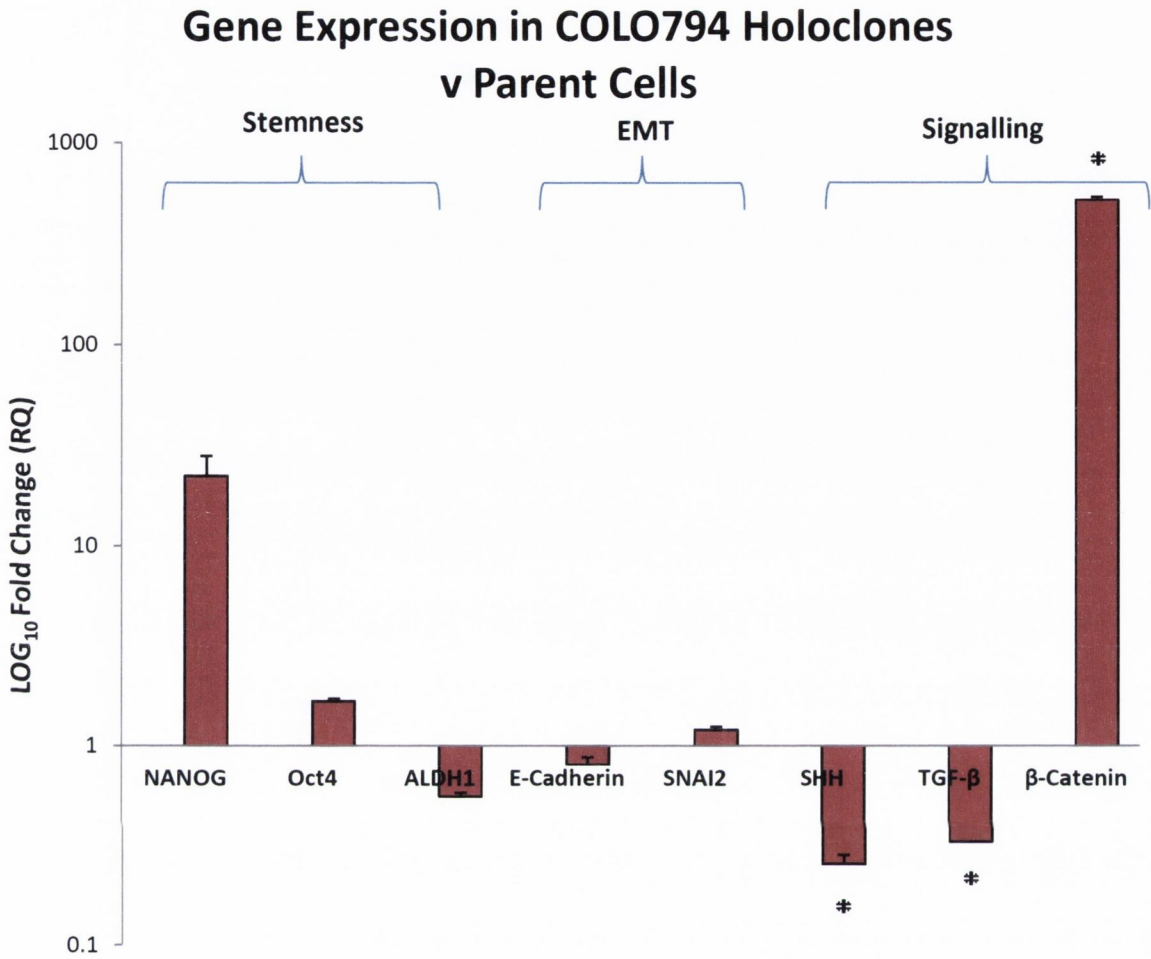


Figure 3.4.1(b): Grouped differential gene expression in COLO794 BRAF wild type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p-value <0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in COLO794 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Similar to its BRAF V600E mutated counterpart, increases in both NANOG and Oct4 were recorded in COLO794 holoclones. NANOG had a 22-fold increase in expression relative to parent cells, while Oct4 was 2-fold increased. In contrast, ALDH1 showed a 2-fold decrease in expression. This observation highlights a marked difference with the BRAF V600E mutated counterpart (SK-Mel 28).

There was distinct difference in expression among the EMT gene group in wild type holoclones compared to their BRAF mutated counterparts. E-Cadherin expression was marginally decreased (0.8 fold) while SNAI2 was 1.2-fold increased. Statistically significant changes in cell signalling gene expression was recorded between COLO794 parent calls and derived holoclones. SHH showed a significant 4-fold decrease (p-value 0.04) in expression in COLO794 holoclones. The biggest change between BRAF mutant and wild type holoclones was recorded in in the Wnt signalling pathway gene β -Catenin with a significant increase (p-value 0.04) of 520-fold in expression in the COLO794 holoclone population. A significant 3-fold decrease (p-value 0.002) in TGF- β expression was recorded in COLO794 holoclones.

These results illustrate differential expression across the panel of genes between SK-Mel 28 BRAF V600E mutated parental melanoma cells and derived holoclones.

3.4.2 Thyroid

8505C BRAF V600E mutant

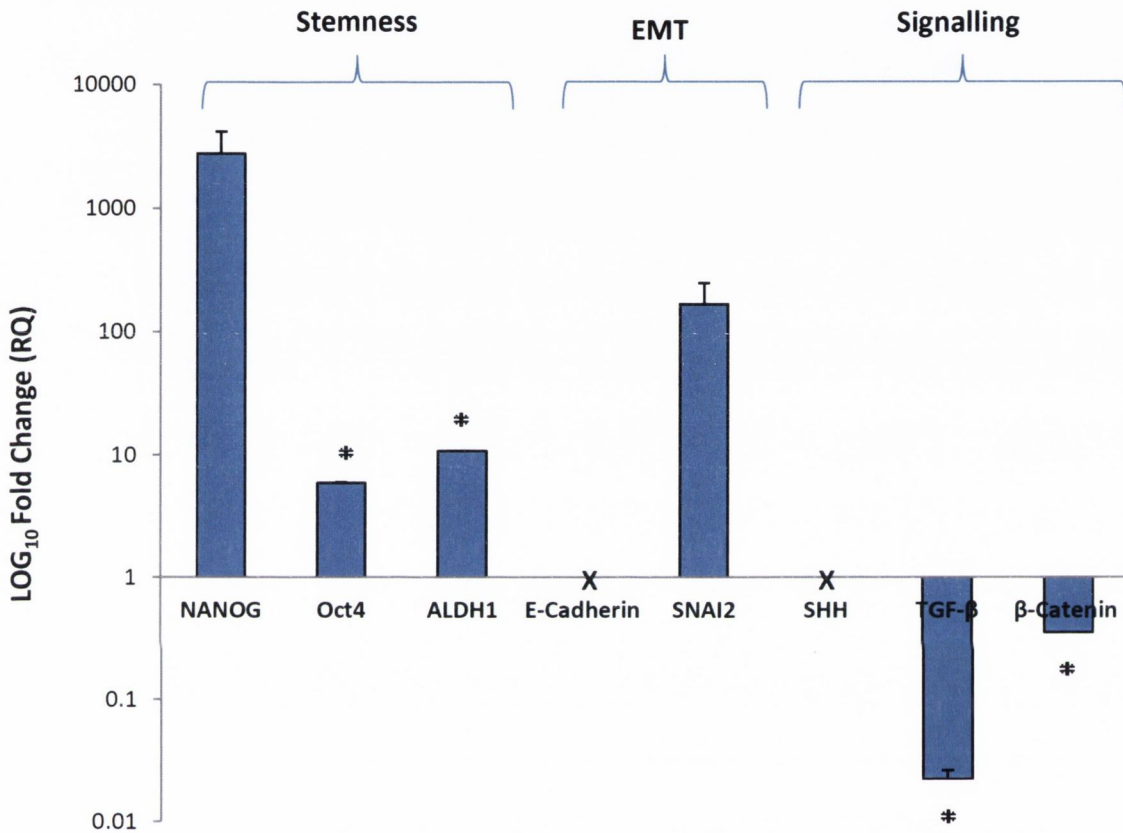


Figure 3.4.2(a): Grouped differential gene expression in 8505C BRAF V600E mutated holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log_{10} Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in 8505C holoclones were plotted using a LOG_{10} scale to illustrate up and down-fold changes

In 8505C BRAF V600E mutated thyroid holoclones; expression of all 3 stemness genes was increased relative to parental cell expression. A substantial 2,774-fold increase in expression was recorded for the NANOG gene and a statistically significant 6-fold increase (p-value 0.03) of Oct4 expression was observed. ALDH1 expression was also significantly up regulated by 10-fold (p-value 0.01) in 8505C holoclones.

Expression of the cellular adhesion gene E-cadherin was not expressed in the 8505C holoclone population but showed consistent expression levels in parental cells and expression of SNAI2 increased by 167-fold in holoclones.

2 of the 3 cell signalling genes assayed were down regulated in 8505C holoclones compared to parent cells. TGF- β expression was decreased by a statistically significant 45-fold (p-value 0.004) and a significant 3-fold decrease (p-value 0.01) in β -Catenin was recorded. SHH was not expressed in 8505C holoclone population but showed consistent expression levels in 8505C parental cells.

Gene Expression in TPC-1 Holoclones v Parent Cells

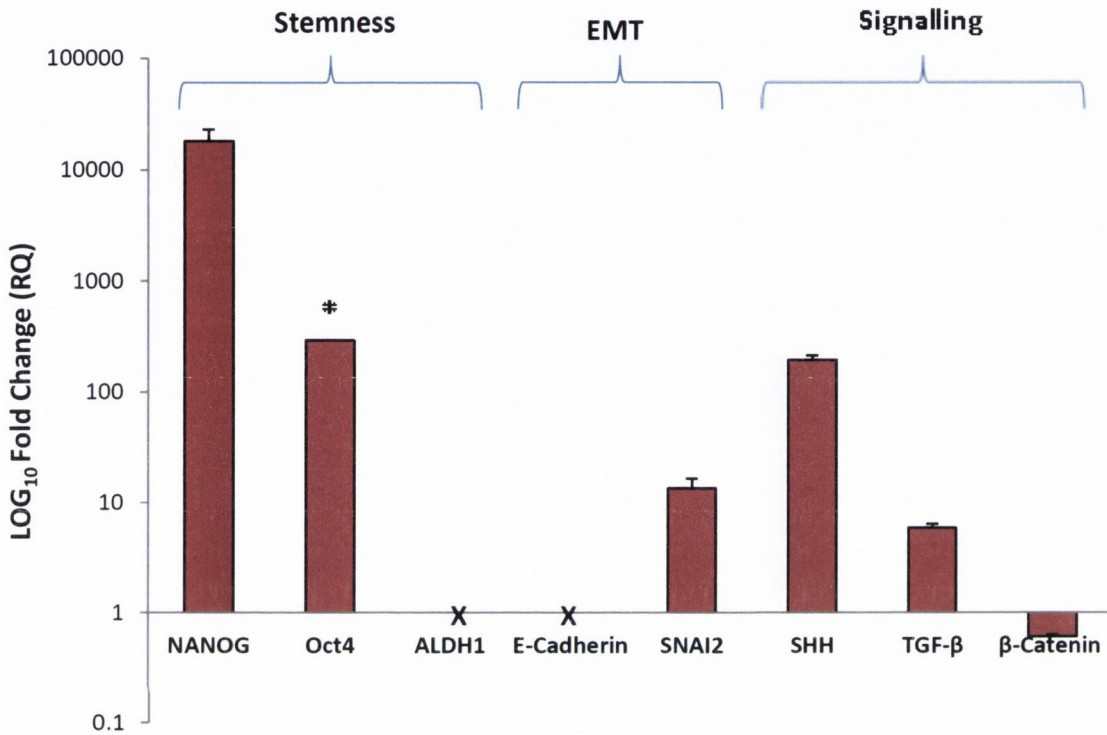


Figure 3.4.2(b): Grouped differential gene expression in TPC-1 BRAF wild type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in TPC-1 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

In TPC-1 BRAF wild type thyroid holoclones, a differential stemness gene expression profile was observed. A substantial increase in NANOG expression of over 18,000-fold was recorded and a statistically significant 293-fold increase (p-value 0.03) of Oct4 expression was observed in the TPC-1 holoclone population relative to TPC-1 parent cells. Unlike its BRAF mutated counterpart, ALDH1 was not expressed in TPC-1 holoclones. Loss of E-cadherin expression was recorded in the TPC-1 holoclone population but showed consistent expression levels in parental cells and expression of SNAI2 increased by 13-fold in holoclones. Expression of the SHH gene increased 194-fold in TPC-1 holoclones and TGF- β expression was also increased, by 6-fold in relative to TPC-1 founder cells. A 2-fold decrease in TPC-1 holoclone β -Catenin levels was recorded.

3.4.3 Ovarian

ES-2 BRAF V600E mutant

Gene Expression in ES-2 Holoclones v Parent Cells

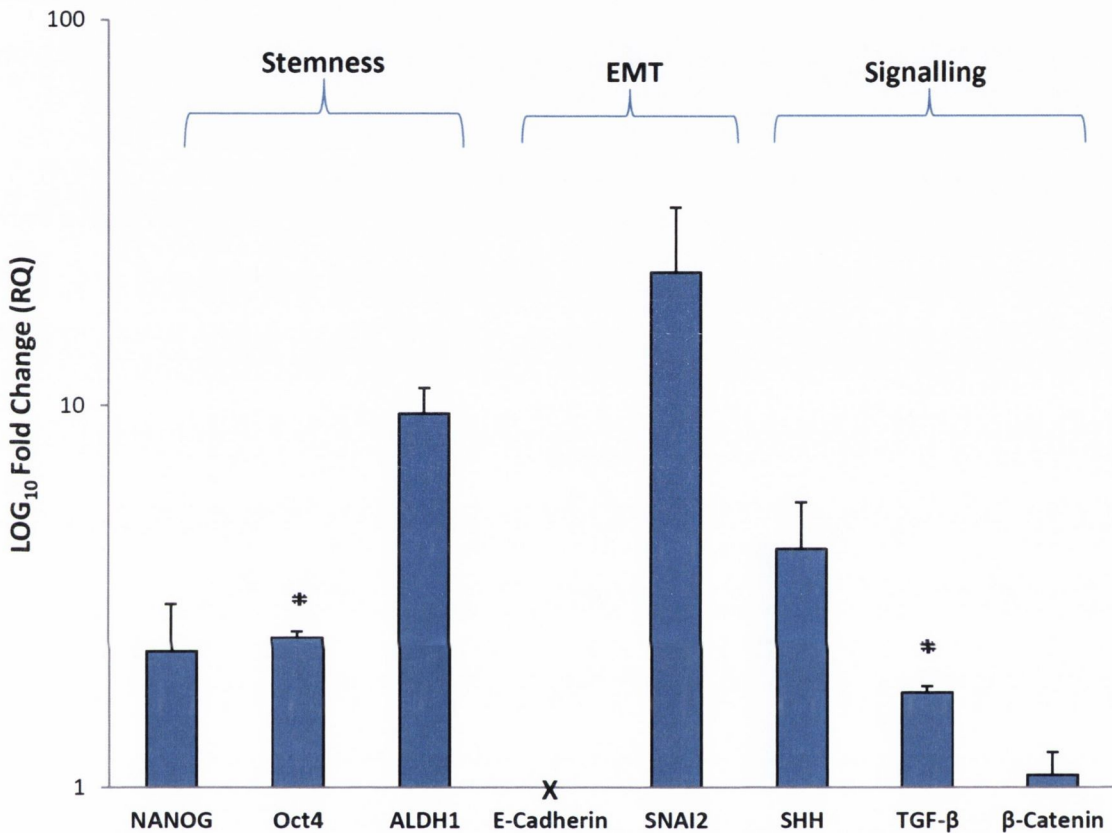


Figure 3.4.3(a): Grouped differential gene expression in ES-2 BRAF V600E mutated holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in ES-2 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

In ES-2 BRAF V600E mutated ovarian holoclones all 3 stemness genes showed increased expression. A 2-fold increase in NANOG expression was observed and Oct4 expression was increased by a statistically significant 2-fold (p-value 0.01). The biggest increase in stemness gene expression was recorded in ALDH1, which was up regulated by 10-fold. Expression of the cellular adhesion gene E-cadherin was not expressed in the holoclone population but showed consistent expression levels in parental cells. The greatest increase in expression across the panel of genes was observed in the EMT gene SNAI2 with a 22-fold increase over its founder cells.

All cell signalling genes in ES-2 holoclones either remained at baseline expression or showed a slight increase relative to parent cells. The greatest differential expression levels were recorded for SHH with a 4-fold increase whereas TGF- β showed an increase of just under 2-fold. β -Catenin remained at baseline expression, similar to the expression levels recorded in ES-2 parental cells.

Gene Expression in A2780 Holoclones v Parent Cells

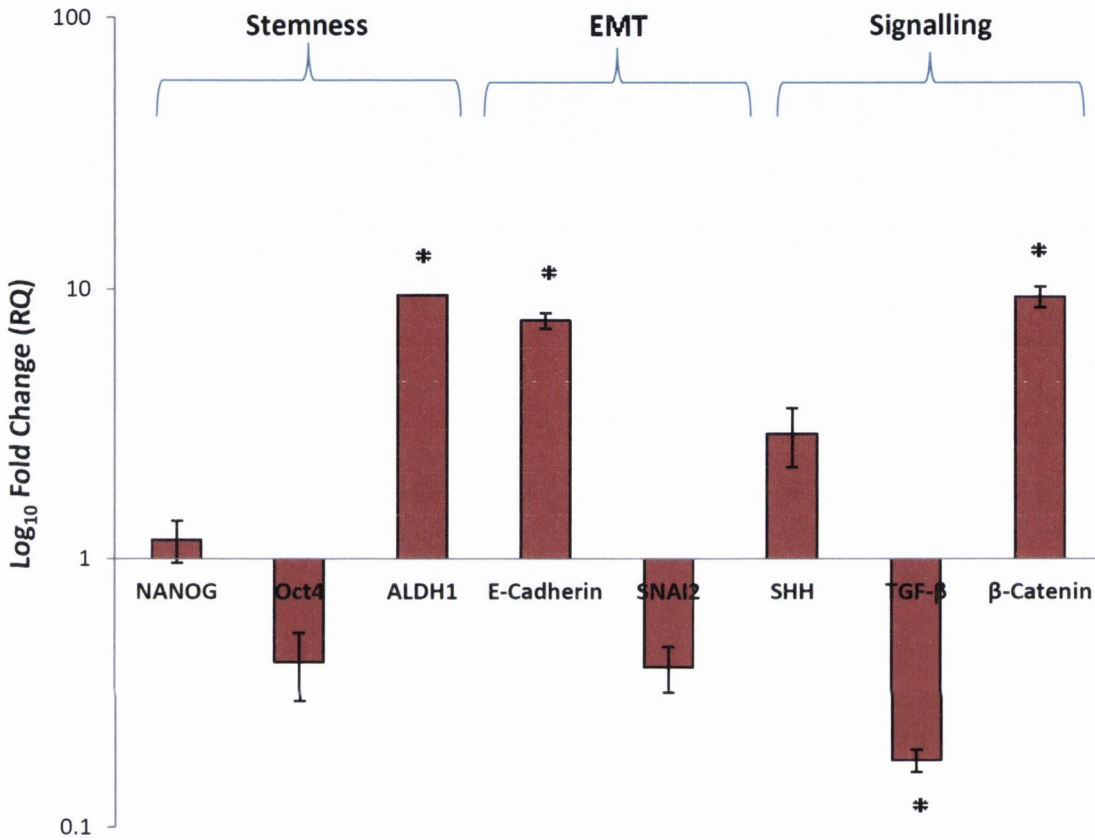


Figure 3.4.3(b): Grouped differential gene expression in A2780 BRAF wild type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p-value <0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in A2780 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Differences in expression of stemness genes were recorded between A2780 holoclones and their BRAF mutated counterparts. A small increase of 1.2-fold was observed in NANOG whereas both Oct4 and ALDH1 were down regulated. A statistically significant 2-fold decrease (p-value 0.01) was recorded in Oct4 and a further 10-fold decrease was seen in ALDH1.

Cellular adhesion gene E-Cadherin expression was significantly increased 7-fold (p-value 0.05) in A2780 holoclones whereas a 2.5-fold decrease was recorded in EMT gene SNAI2.

Statistically significant differential cell signalling gene expression was observed between in A2780 parental cells and holoclone populations. A 3-fold increase in SHH expression was recorded, similar to its BRAF mutated ovarian counterparts. TGF- β expression was significantly decreased by 5-fold (p-value 0.03). Only a slight increase above baseline expression in β -Catenin was seen in the BRAF V600E mutated cell line ES-2 but the increase was much greater in A2780 holoclones as we observed a significant 9-fold increase (p-value 0.01), the greatest differential expression level of all the genes assayed between A2780 parental cells and holoclones.

3.4.4 Colorectal

HT-29 BRAF V600E mutant

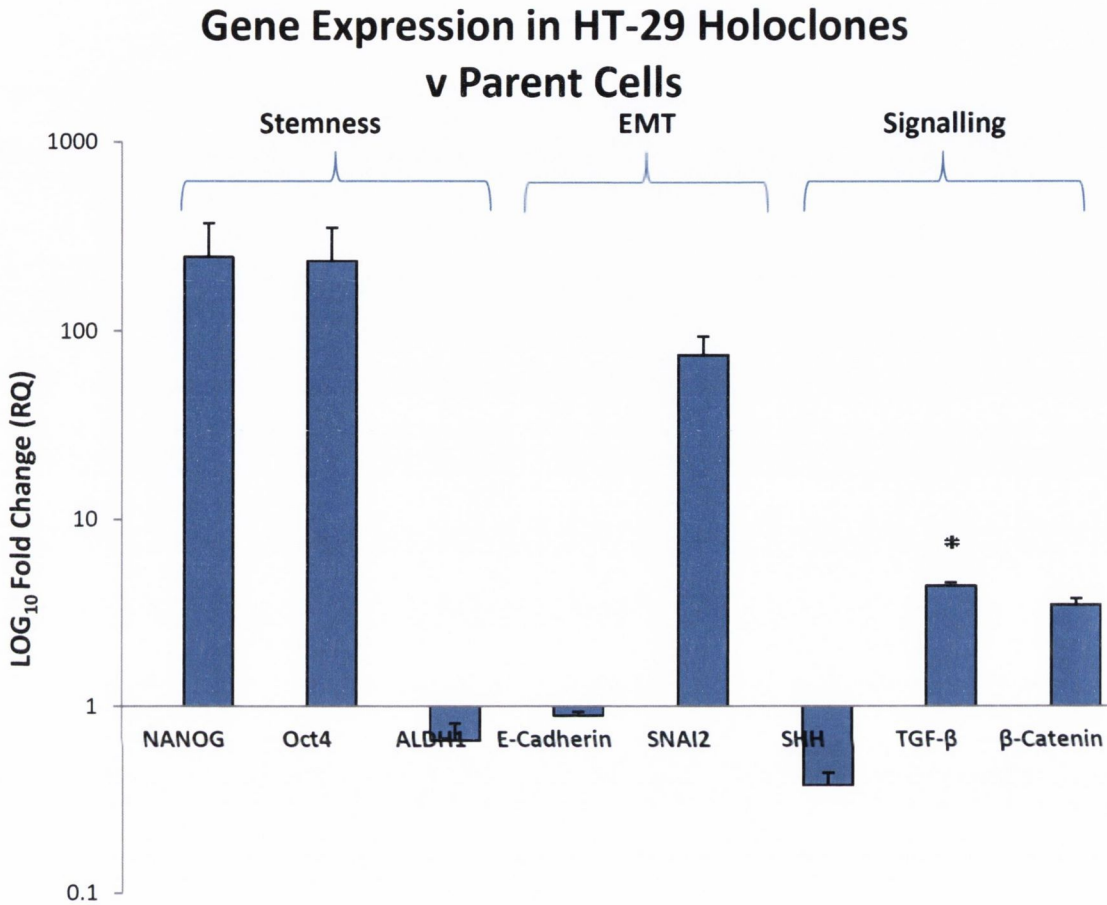


Figure 3.4.4(a): Grouped differential gene expression in HT-29 BRAF V600E mutated holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in HT-29 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Substantial increases in stemness gene expression was recorded between HT-29 parent cells and derived holoclones. Pluripotency gene NANOG increased by over 246-fold in HT-29 holoclones when compared to parental cell expression and similarly, expression of Oct4 increased by 234-fold in the holoclone population. Expression of the third stemness marker analysed, ALDH1, decreased by just over 1.5 fold.

Despite only a minor change below baseline expression being recorded in the cell adhesion gene E-Cadherin, a 75-fold increase in expression of EMT gene SNAI2 was recorded in HT-29 holoclones.

A diverse cell signalling gene expression profile was observed in HT-29 holoclones with a 2-fold decrease being recorded in SHH. A significant 3.5-fold (p-value 0.01) increase was recorded in TGF- β expression in HT-29 holoclones. Expression of the Wnt signalling gene β -Catenin, was up regulated 4-fold in HT-29 holoclone populations.

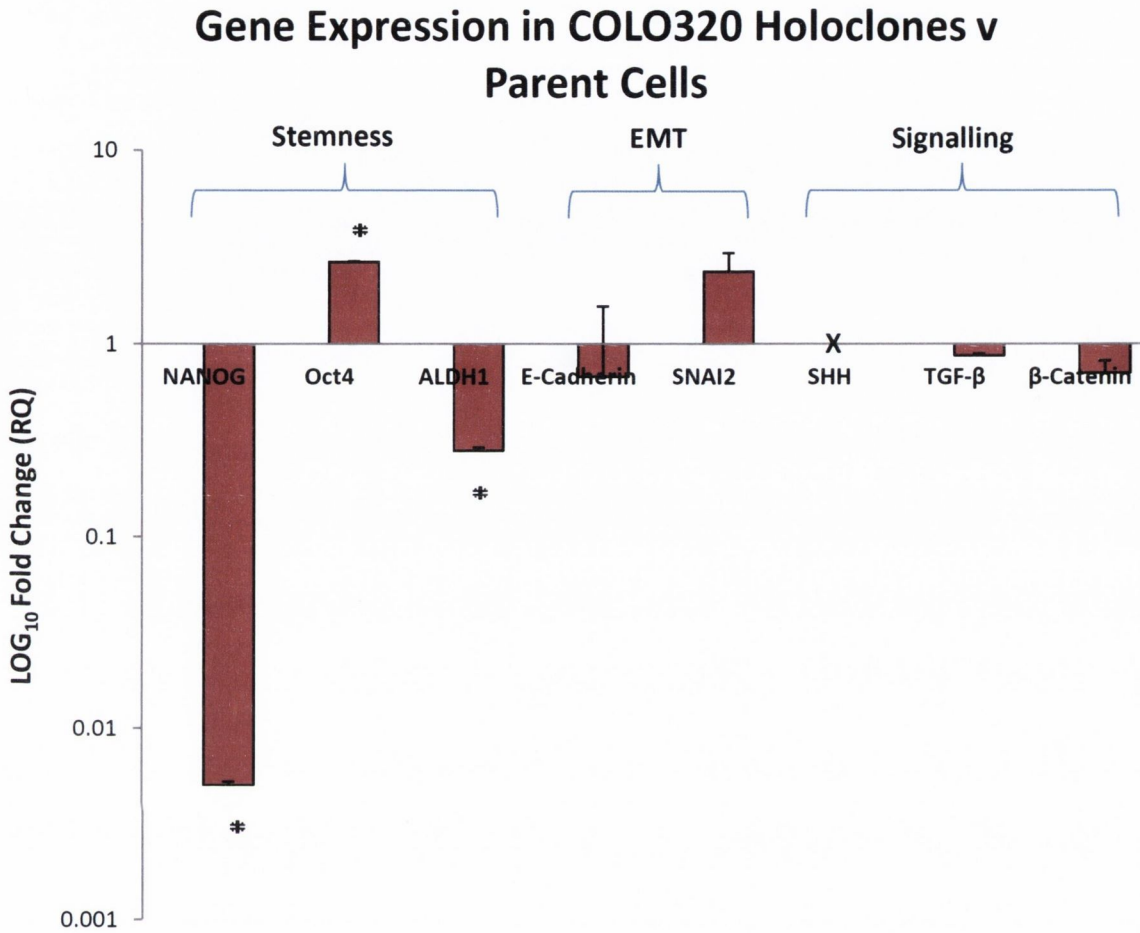


Figure 3.4.4(b): Grouped differential gene expression in COLO320 BRAF V600E mutated holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Assayed genes; Y-axis: Log₁₀ Avg Fold Change (RQ). A panel of genes were assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in COLO320 holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Statistically significant changes in the expression levels of all 3 stemness genes were recorded in COLO320 holoclones when compared to their founder cells. NANOG showed a significant 200-fold (p-value 0.0002) decrease in expression compared with that of the COLO320 parental cells. Expression of the other pluripotency marker analysed, Oct4, was increased significantly (p-value 0.01) by almost 3 fold. A significant 4-fold decrease (p-value 0.02) in ALDH1 expression was also observed in COLO320 holoclones

Minor changes in the expression of cellular adhesion gene E-Cadherin and EMT gene SNAI2, were recorded in COLO320 holoclones. E-Cadherin expression decreased by 1.5 fold relative to COLO320 parent cells and expression of SNAI2 increased by 2.3 fold in the holoclone population.

Similarly, changes of gene expression in cell signalling genes were minor. SHH was not expressed in COLO320 holoclones despite showing consistent expression in parental cells. TGF- β expression was slightly decreased at 0.8-fold below baseline expression and β -Catenin expression levels decreased by 0.7-fold relative to COLO320 parental cell expression.

3.4.5 Grouped Differential Gene Expression

Individual gene expression results from each cell line were combined to illustrate expression patterns between BRAF V600E mutated and BRAF wild type holoclones derived from melanoma, thyroid, ovarian and colorectal cancers.

As a result, differences and similarities in gene expression patterns of genes involved in stemness, EMT and cell signalling were identified.

Cell Line Matrix Holoclone NANOG Gene Expression

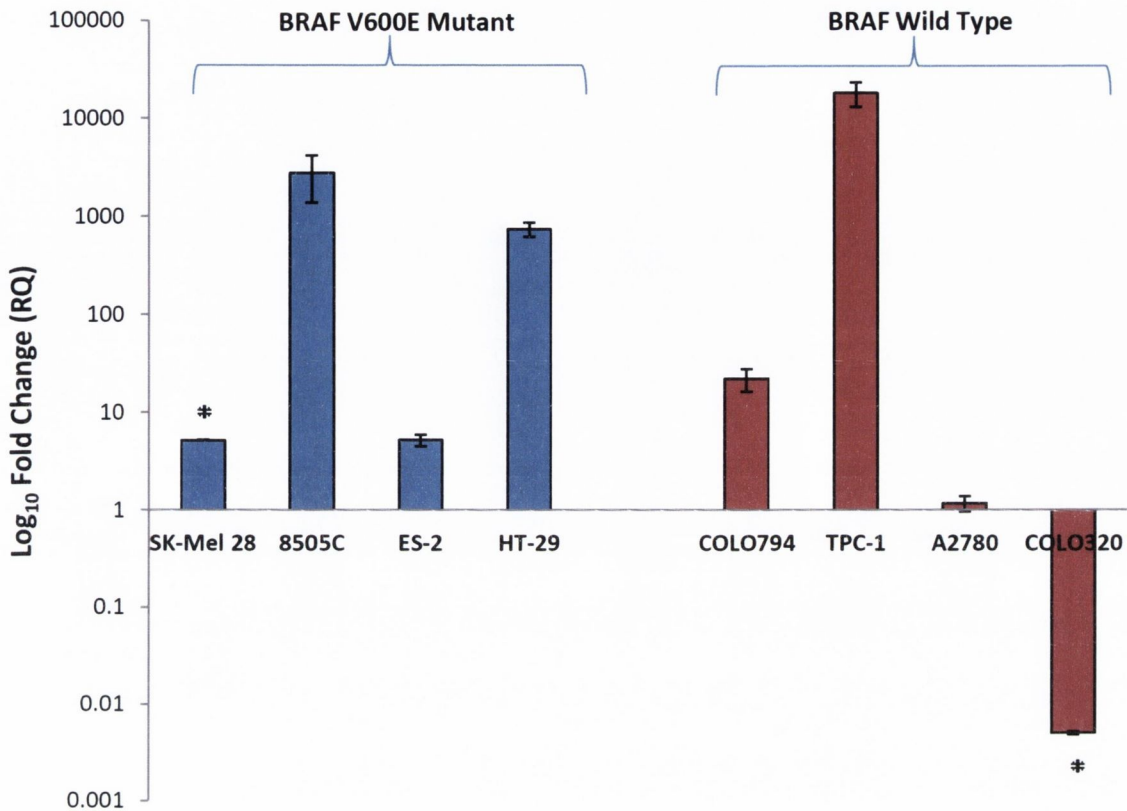


Figure 3.4.5(a): Differential gene expression of NANOG in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). NANOG was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Gene expression of pluripotency marker NANOG was increased in BRAF V600E mutated holoclones derived from all four cancer types with a statistically significant increase recorded in SK-Mel 28 holoclones.

NANOG expression was also increased in BRAF wild type holoclones with the exception of COLO320 in which the gene was significantly down regulated.

Cell Line Matrix Holoclone Oct4 Gene Expression

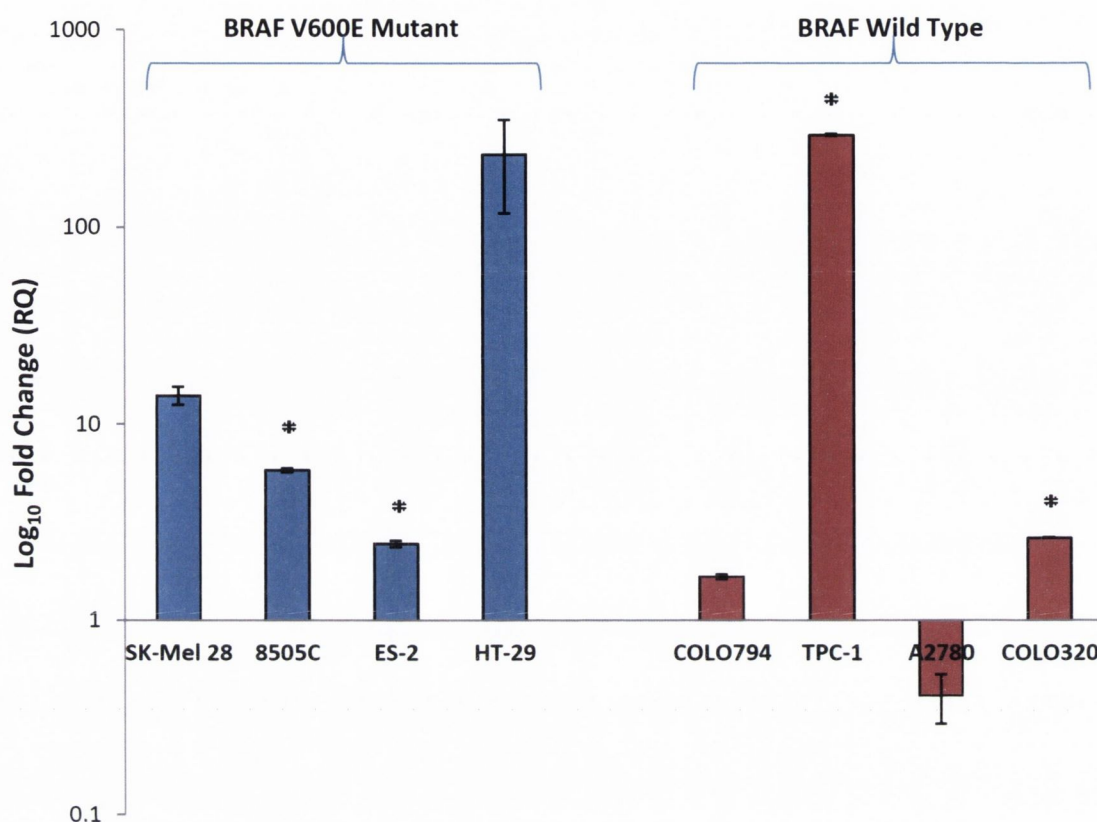


Figure 3.4.5(b): Differential gene expression of Oct4 in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p-value <0.05 *). Error bars: SEM.

X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). Oct4 was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Gene expression of pluripotency marker Oct4 was increased in BRAF V600E mutated holoclones derived from all four cancer types with statistically significant increases recorded in 8505C and ES-2 holoclones. Oct4 expression was also increased in BRAF wild type holoclones with the exception of A2780. Statistically significant increases were recorded in TPC-1 and COLO320 holoclones.

Cell Line Matrix Holoclone ALDH1 Gene Expression

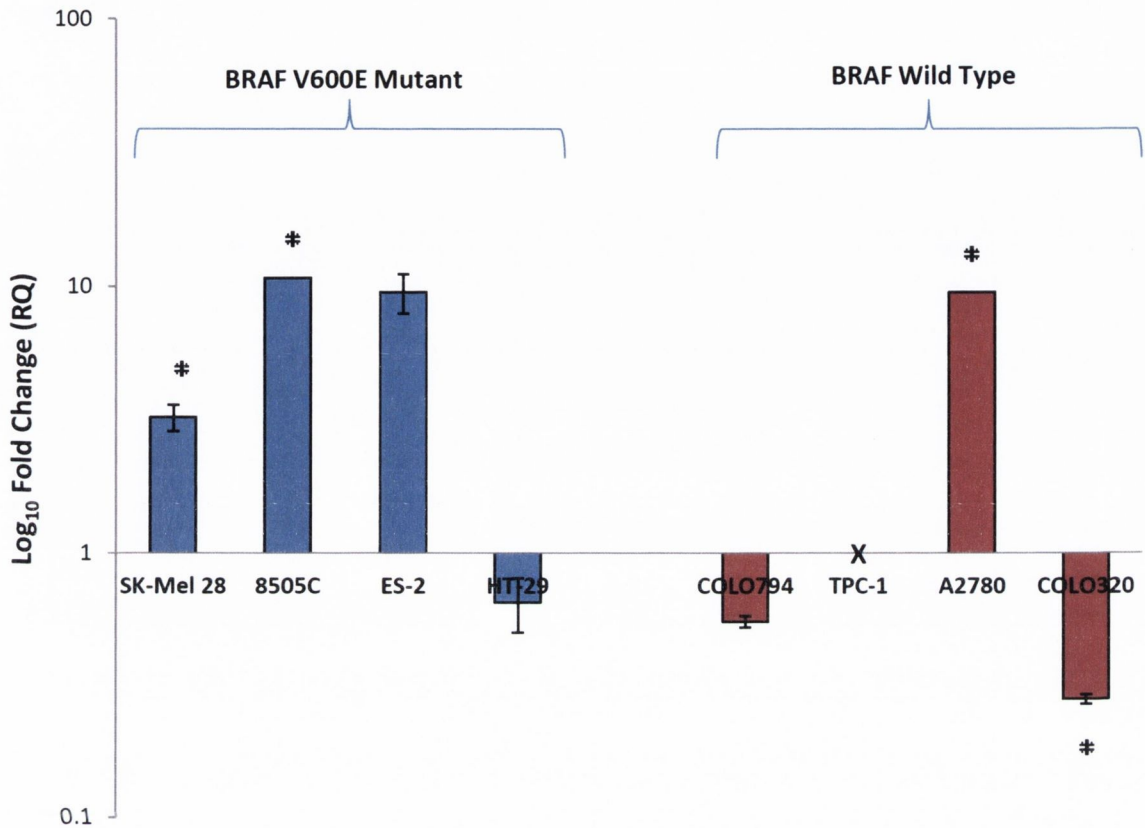


Figure 3.4.5(c): Differential gene expression of ALDH1 in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). ALDH1 was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Gene expression of stem cell marker ALDH1 was increased in BRAF V600E mutated holoclones with the exception of HT-29 and statistically significant increases were recorded in SK-Mel 28 and 8505C holoclones. ALDH1 expression was lost or decreased in BRAF wild type holoclones with the exception of A2780 in which a statistically significant increase was recorded.

Cell Line Matrix Holoclone E-Cadherin Gene Expression

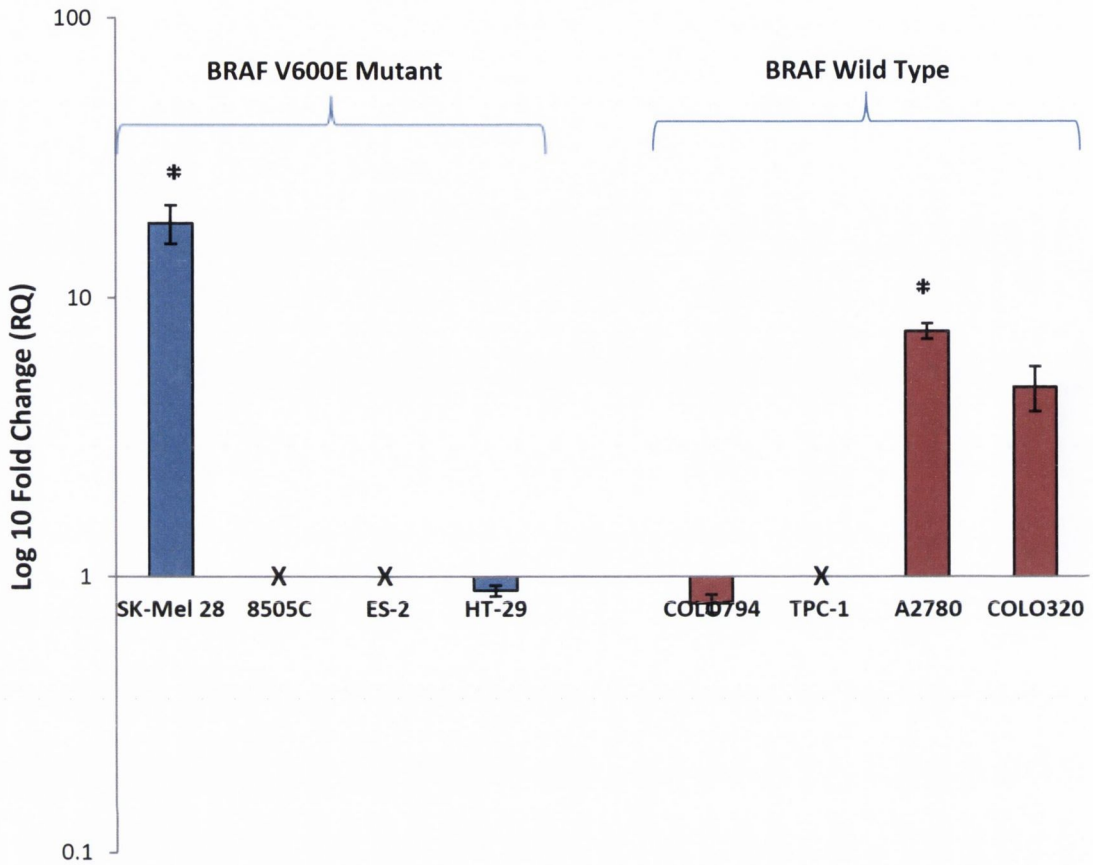


Figure 3.4.5(d): Differential gene expression of E-Cadherin in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Cell Lines; Y-axis: Log_{10} Avg Fold Change (RQ). E-Cadherin was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG_{10} scale to illustrate up and down-fold changes

No clear pattern of E-Cadherin expression was observed in BRAF V600E mutated holoclones as expression was either lost, decreased or increased significantly in the case of SK-Mel 28. Increases in E-Cadherin expression were also observed in BRAF wild type ovarian and colorectal holoclones but decreased in melanoma and thyroid holoclones.

Cell Line Matrix Holoclone SNAI2 Gene Expression

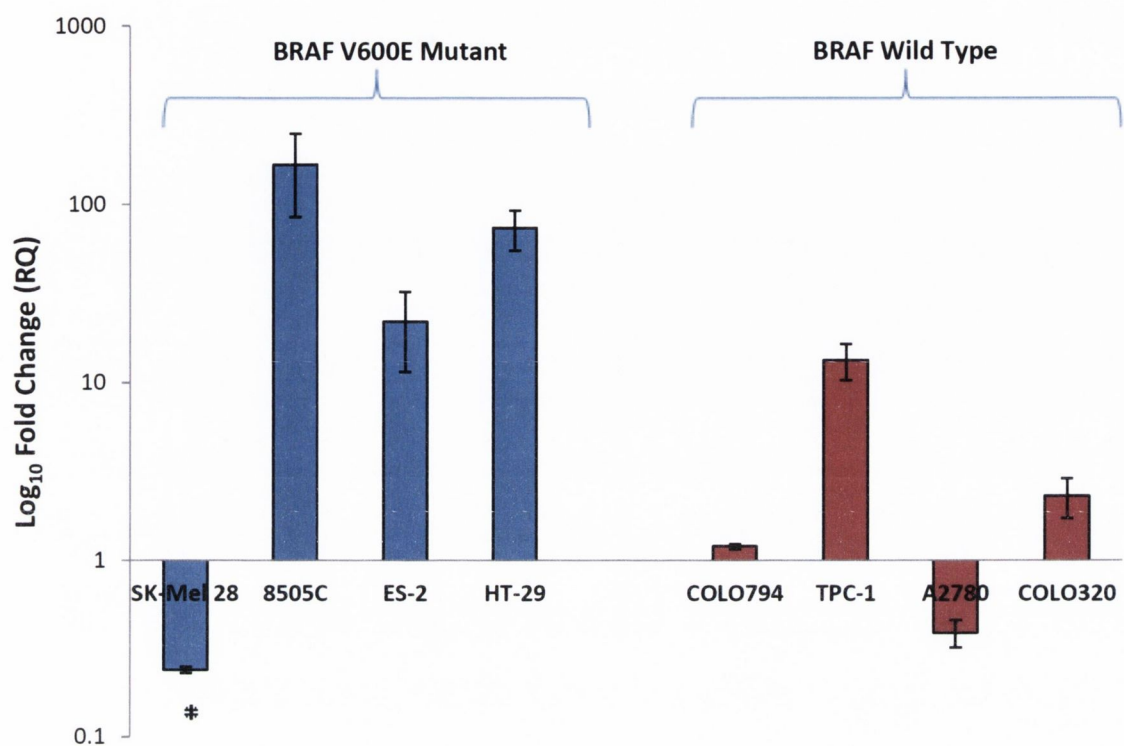


Figure 3.4.5(e): Differential gene expression of SNAI2 in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). SNAI2 was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Increases in expression of EMT gene SNAI2 were recorded in BRAF V600E mutated holoclones with the exception of SK-Mel 28 in which a significant decrease was observed. SNAI2 expression was also increased in BRAF V600E wild type holoclones with the exception of A2780 in which a decrease was recorded.

Cell Line Matrix Holoclone SHH Gene Expression

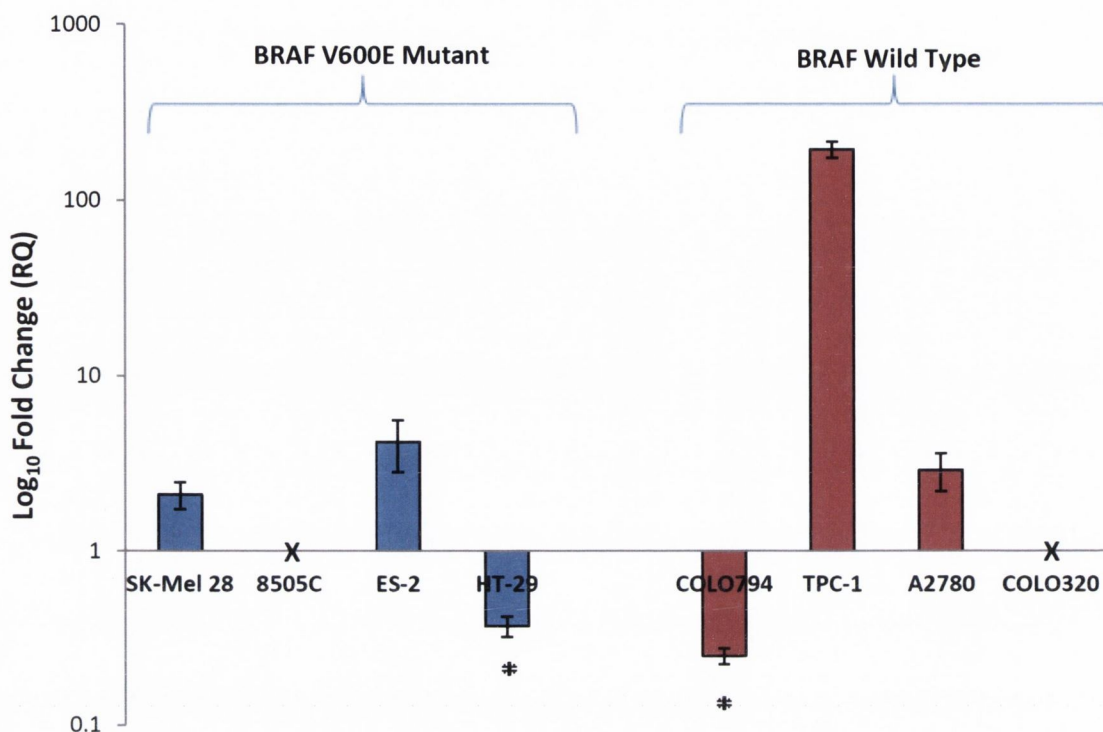


Figure 3.4.5(f): Differential gene expression of SHH in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM.

X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). SHH was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

No clear pattern of expression of the cell signalling gene SHH was observed in BRAF V600E mutated holoclones as expression was increased, lost across the range of cancer types or decreased in the case of HT-29 colorectal holoclones.

Similarly in BRAF wild type holoclones, SHH expression was increased, lost or significantly decreased in the case of COLO794 holoclones.

Cell Line Matrix Holoclone TGF- β Gene Expression

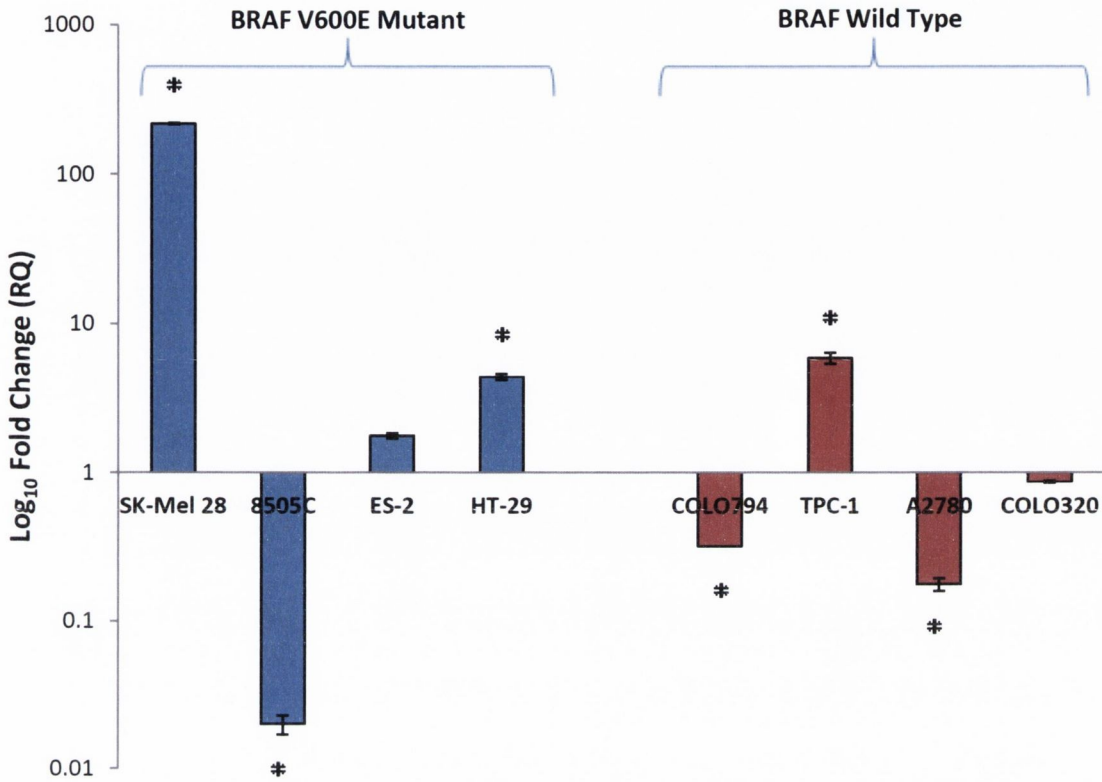


Figure 3.4.5(g): Differential gene expression of TGF- β in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). TGF- β was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Expression of TGF- β was increased in BRAF V600E mutated holoclones, including statistically significant increases in SK-Mel 28 and HT-29 whereas a significant decrease was recorded in 8505C holoclones. Expression of TGF- β was decreased in BRAF wild type holoclones, including statistically significant decreases in COLO794 and A2780 whereas a significant decrease was recorded in TPC-1 holoclones.

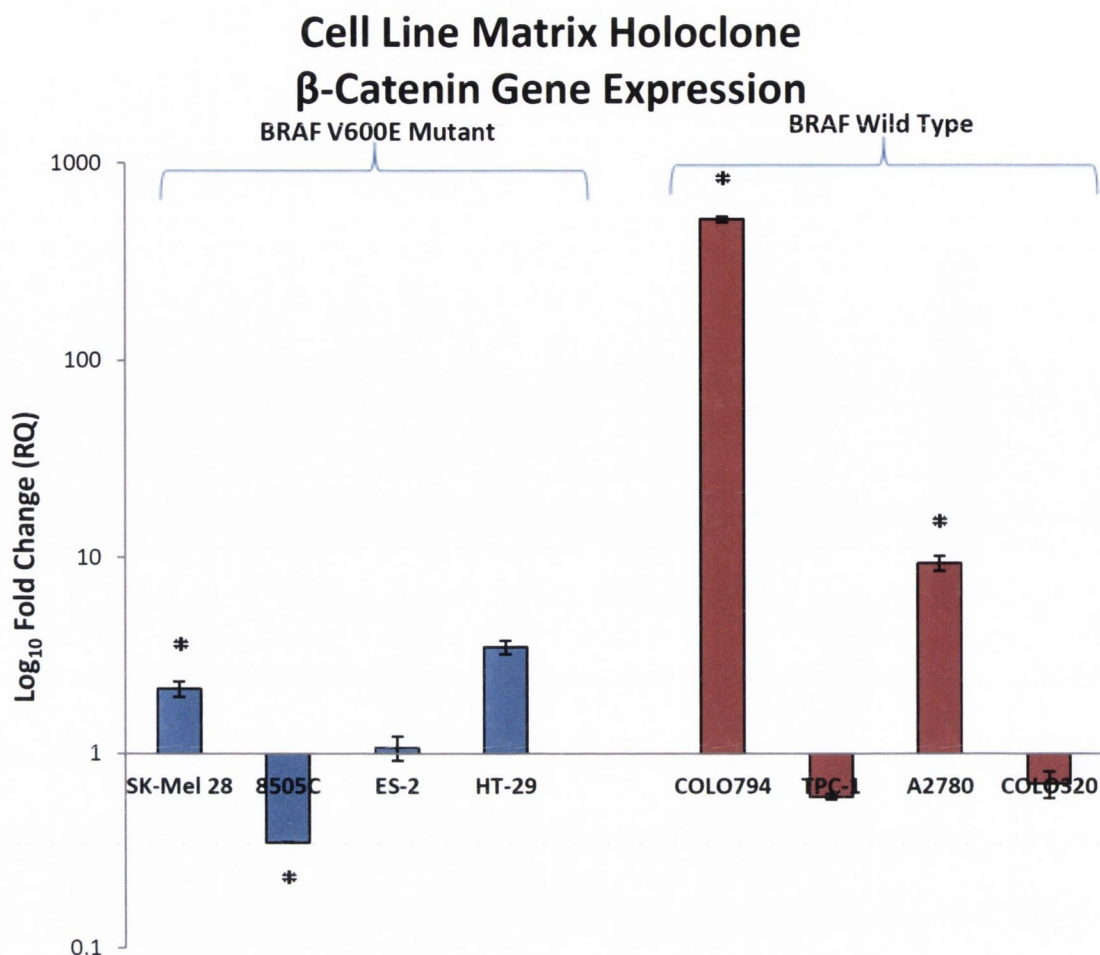


Figure 3.4.5(h): Differential gene expression of β -Catenin in BRAF V600E and BRAF Wild Type holoclones relative to parental cells.

Statistical significance: Student 2-tail t test +/- 2-fold (p -value < 0.05 *). Error bars: SEM. X-axis: Cell Lines; Y-axis: Log₁₀ Avg Fold Change (RQ). TGF- β was assayed by quantitative TaqMan PCR and results were analysed by using the DataAssist V3.0 software package to generate fold change results based on relative quantification values. RQ values of gene expression in holoclones were plotted using a LOG₁₀ scale to illustrate up and down-fold changes

Expression of β -Catenin was increased in BRAF V600E mutated holoclones, including a statistically significant increase in SK-Mel 28 and HT-29 whereas a significant decrease was recorded in 8505C holoclones. Expression of β -Catenin was also significantly increased in COLO794 and A2780 BRAF wild type holoclones, decreases were recorded in TPC-1 and COLO320 holoclones.

3.4.6 Discussion

Melanoma

Increased expression of pluripotency transcription factors NANOG and Oct4 in SK-Mel 28 holoclones suggests that cancer stem cells derived from V600E BRAF mutated melanoma cells display high levels of pluripotency and self-renewal potential. SK-Mel 28 parental cells are an aggressive malignant melanoma displaying a fast growth pattern. Coupled with the increased expression of stem cell marker ALDH1 gene, these results propose that SK-Mel 28 parental cells may have an enhanced ability to produce holoclones with strong stem identity and a strong resistance to high salt agar conditions.

Many studies have shown loss of E-cadherin expression contributing to melanoma spread (Bonitsis et al., 2006). SK-Mel 28 holoclones displayed lower levels of the cellular adhesion gene E-Cadherin, suggesting they may have superior invasive potential than the founder cells they were derived from. Several studies have investigated the role of TGF- β in driving the oncogenic process in melanoma but a better understanding of its role in melanoma cancer stem cell biology is needed (Nummela et al., 2012; Sun et al., 2011). Despite recording a significant increase in TGF- β expression levels in SK-Mel 28 holoclones, some studies have established that in melanoma, β -catenin is a suppressor of invasion as increased cytoplasmic and nuclear β -Catenin is currently postulated as a biomarker of good prognosis (Arozarena et al., 2011). An increase in SHH expression in SK-Mel 28 holoclones, suggests that activity of Hedgehog pathway components are increased in BRAF V600E mutated CSCs. In summary, the majority of genes investigated in SK-Mel 28 holoclones showed an increase in expression which may be contributing towards a strong stem, hyperactive signalling CSC population.

As melanoma shows the highest prevalence of BRAF V600E mutation in solid tumours, it could be suggested that BRAF mutation has widespread effects on many aspects of cancer biology (Cantwell-Dorris et al., 2011).

Similar to its BRAF V600E mutated counterpart, increases in both NANOG and Oct4 were recorded in BRAF wild type COLO794 holoclones. Despite a decrease in ALDH1 expression, these gene expression results suggest that CSC progenies derived from a wild type BRAF malignant melanoma have pluripotent characteristics. A distinct difference in expression of

E-cadherin and SNAI2 in wild type holoclones compared to their BRAF mutated counterparts suggests a different migratory potential however more research is needed to investigate invasive abilities of both melanoma cell lines.

Significant increases in both SHH and β -Catenin suggest increased cell signalling via the Hedgehog and WNT pathways although other components of these pathways must be investigated to get a greater understanding of their role in BRAF wild type melanoma CSCs.

SHH showed a significant 4-fold decrease (p-value 0.04) in expression in COLO794 holoclones. A significant decrease in TGF- β expression in COLO794 holoclones suggests that signalling may in fact be taking place through these pathways in place of TGF- β pathway activity.

It is evident from these results that CSC progenies derived from both wild type and BRAF V600E mutated melanomas have different gene expression profiles. Similar stemness profiles were to be expected if both progenies were truly “stem” but differences in both EMT and cell signalling processes in these holoclones may be under the influence of BRAF mutation directly or indirectly via interaction with other molecules.

Thyroid

Increased expression in of stemness genes in 8505C holoclones suggest that BRAF V600E mutated thyroid CSC progenies have a strong stem identity evident from elevated levels of stem marker expression. As the 8505C cell line was derived from an undifferentiated anaplastic thyroid carcinoma, it is possible that 8505C cells express high levels of stemness genes such as NANOG and ALDH1. Higher expression levels of these genes in 8505C holoclone populations suggest that these CSC progenies have a more undifferentiated, stem phenotype than their founder cells. Observations of self-renewal in high salt agar assays further strengthens the suggestion that V600E BRAF mutated CSC progenies have enhanced levels of pluripotency.

Increased expression of EMT gene SNAI2 and loss of E-Cadherin expression in 8505C holoclones suggest these CSCs have increased migratory potential over an already anaplastic aggressive parental cell line. However, further experiments to investigate if 8505C holoclones have invasive characteristics must be carried out to learn more about the potential for these CSC to metastasis in an animal model.

Decreases in SHH and TGF- β expression in 8505C suggest that cell signalling via the Hedgehog and TGF- β cell signalling pathways may be repressed in 8505C holoclones but further research into the other pathway components is needed to get a greater understanding of cell signalling in BRAF V600E mutated thyroid CSC progenies.

The high expression levels of the pluripotency markers NANOG and Oct4 in TPC-1 holoclones suggest that BRAF wild type thyroid CSC progenies possesses stem characteristics and are pluripotent. Similar to its BRAF V600E mutated cell line, observations of holoclone self-renewal made in TPC-1 cell lines in high salt agar assays, may supplement gene expression results and suggest that TPC-1 CSC progenies have self-renewal capabilities. Despite showing consistent expression in parental cells, TPC-1 holoclones exhibited loss of E-cadherin expression and coupled with an increase in SNAI2 expression suggests TPC-1 holoclones have reduced cellular adhesion and increased EMT potential based on expression results of these two genes.

Increased SHH and TGF- β gene expression results suggests that CSC progenies derived from PTC cells have increased Hedgehog cell signalling and unlike 8505C holoclones, may also have increased cell signalling via the TGF- β pathway. However more research into the cell signalling mechanisms is needed to get a better insight into the roles played by signalling genes in thyroid holoclones that harbour wild type BRAF.

It is clear from these results that CSC progenies derived from both wild type and BRAF V600E mutated thyroid carcinomas have different gene expression profiles. Similar expression of pluripotency markers were recorded suggesting both progenies have “stem” characteristics but differences in both EMT and cell signalling processes in these holoclones may be affected by the BRAF V600E mutation.

Ovarian

Increased expression in ES-2 holoclones of all 3 stemness markers investigated in this study; NANOG, Oct4 and ALDH1 BRAF V600E mutated ovarian CSC progenies have a strong stem identity. Despite consistent expression levels of E-Cadherin in parental cells, expression of the cellular adhesion gene E-cadherin was not expressed in the holoclone population suggesting a loss of cellular adhesion mechanisms. In conjunction with increased expression

of EMT gene SNAI2, loss of E-Cadherin expression in ES-2 holoclones suggest these CSCs may have increased migratory potential over their parental cells. Further studies, in mouse models may give a greater insight into the tumour forming capabilities of these BRAF V600E mutated ovarian CSCs and if they have metastatic potential based on results of aberrant EMT marker expression.

All 3 cell signalling genes; SHH, TGF- β , β -Catenin showed relatively minor changes in expression in comparison to parent cells. In other cancers investigated in this study in which BRAF mutation has a profound effect on cell signalling, expression of these genes showed widespread changes in expression profiles. However prevalence of the V600E mutation in ovarian cancers is much lower and results from this study may suggest that its effects may not be as widespread in this type of cancer (Estep et al., 2007). In BRAF V600E mutated ovarian holoclones only minor changes were recorded suggesting BRAF mutation in ovarian CSC progenies may not have such a profound effect on signalling pathways.

A2780 parent cells are an aggressive, invasive ovarian cancer cell line in culture which may be reflected in the high salt agar assay results observed in which holoclone morphology was recorded after only 2 weeks, in contrast with the cell line panel average of 4 weeks. Despite this evidence of a strong CSC progeny production capability, stemness genes NANOG, Oct4 and ALDH1 showed baseline or decreased expression. Expression of NANOG remained similar to that of its founder cells whereas both Oct4 and ALDH1 were down regulated.

As the A2780 cell line was derived from ascitic fluid from a patient with metastatic advanced ovarian carcinoma, as is the case with many advanced cancers, loss of cellular adhesion gene E-Cadherin might be expected. However E-Cadherin expression was significantly increased in A2780 holoclones expression of EMT gene SNAI2 was decreased suggesting the holoclone population may not have migratory potential like the parental cells they were derived from.

Statistically significant differential cell signalling gene expression was observed between in A2780 parental cells and holoclone populations. Increases in SHH expression suggest A2780 holoclones have increased signalling via the Hedgehog pathway although the expression of more genes in this pathway must be examined to get a better understanding of the role played by hyperactive cell signalling pathways in ovarian CSCs.

In contrast, TGF- β expression was significantly decreased suggesting signalling via the TGF- β pathway is repressed in CSC progenies derived from wild type BRAF ovarian carcinoma.

Based on our gene expression results, hyperactive BRAF kinase activity may play a role in increasing activity of signalling pathways involved in cell survival and proliferation such as the Wnt and TGF- β pathways. Components of both pathways showed increased expression in BRAF mutated holoclones warranting further investigation into the several other genes involved in these pathways to establish if their activity is increased in a similar fashion as a result of V600E mutation.

Colorectal

Statistically significant changes in the expression levels of all 3 stemness genes were recorded in HT-29 holoclones when compared to their founder cells indicating that CSC progenies derived from BRAF V600E mutated colorectal carcinomas may have a strong stem and pluripotent identity.

Despite only a minor change below baseline expression being recorded in the cell adhesion gene E-Cadherin, a 75-fold increase in expression of EMT gene SNAI2 was recorded in HT-29 holoclones. Similar profiles E-Cadherin loss and increased SNAI2 expression have been observed in other BRAF V600E mutated holoclones investigated in this study suggesting BRAF mutation may play a role in increasing EMT marker expression in holoclones derived from parental cells carrying the V600E mutation.

In this study we have shown like BRAF V600E mutated melanoma holoclones, that TGF- β levels are increased in CSC progenies with enhanced BRAF activity. Holoclones derived from colorectal adenocarcinoma cells possessing the BRAF V600E mutation have stem-like expression profiles evident from increased levels of pluripotency markers NANOG and Oct4. Activation of cell signalling pathways may also be increased based on the genes analysed in this study and may be involved in conferring a pro-survival cancer stem cell population with increased self-renewal and proliferation capabilities. Further investigation into cell signalling pathways in BRAF mutated colorectal CSC progenies are warranted based on these results.

Despite evidence of holoclone morphology appearing in week 2 of the high salt agar assay, expected increases in stemness gene expression in COLO320 holoclones were not recorded

with only 1 of the 3 stemness genes assayed showing an increase in expression. Previous studies (Collins et al., 1992) have highlighted the role in c-myc in COLO320 colony formation that may also play a role in the development of COLO320 holoclones but was not investigated as part of this study. A statistically significant decrease in NANOG expression coupled with a significant increase in Oct4 expression suggests that COLO320 holoclones may not possess the increased pluripotency levels like their BRAF V600E mutated counterparts. Investigating expression of c-myc in COLO320 holoclones may highlight a differential mechanism of holoclone development in BRAF wild type colorectal adenocarcinomas.

The minor changes observed in the expression of cellular adhesion gene E-Cadherin and EMT gene SNAI2 recorded in COLO320 holoclones suggest that colorectal holoclones lacking the BRAF V600E mutation may have low EMT potential based on these results.

Similarly, minor changes in cell signalling gene expression may indicate a CSC progeny with decreased cell signalling via the TGF- β and Wnt signalling pathways as evident from the decrease in TGF- β and β -Catenin expression. SHH was not expressed in COLO320 holoclones despite showing consistent expression in parental cells indicating possible loss of stem cell proliferation which has been implicated in studies outlined in chapter 1 (Song et al., 2011; Lee et al., 2008). These gene expression results suggest BRAF V600E mutation has many downstream effects in both wild BRAF colorectal holoclones and holoclones harbouring the mutation. Differential gene expression profiles were observed across the different gene groups in which the BRAF gene may play a role but needs further investigation to get a greater understanding of BRAF's involvement in colorectal CSC molecular biology.

Through gene expression analysis, we identified differential expression patterns of genes involved in EMT and cell signalling. We did however observe a consistent expression profile in stemness genes across the panel of BRAF V600E mutated holoclones suggesting that the MAPK-ERK signalling pathway may play a pivotal role in cancer stem cells. As TPC-1 cells carry the RET-PTC mutation upstream of BRAF resulting in a similar hyperactivation of the MAPK-ERK pathway (Smyth et al, 2005), it is possible that mutations in this cell signalling

mechanism confer stem-like characteristics in holoclones derived from a variety of different tumour types.

3.5 Immunofluorescence of NANOG Protein in Parental Cells & Holoclones

3.5.1 HT-29 BRAF V600E mutant

Cells were stained by immunofluorescence and visualized with confocal microscopy. Images were captured at 10X objective magnification and 1X zoom. Scans were performed at 1µm interval depths through the fixed cells, and single or merged images are presented as XY single planes through the mid-section of the cells. DAPI (4', 6-diamidino-2-phenylindole) was used to stain the cell nuclei (**Blue**), Alexa488 was used to stain NANOG (**Green**) and Rhodamine Phalloidin Tetramethylrhodamine B isothiocyanate (TRITC) was used to stain actin filaments (**Red**). The fourth panel displays the merged images of all three channels.

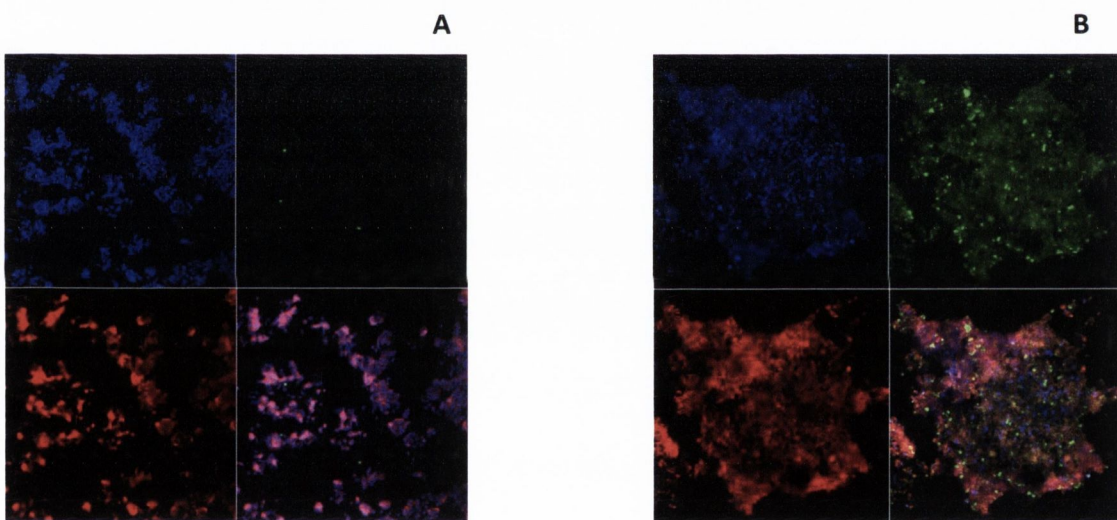


Figure 3.5.1(a): Confocal images of HT-29 parental cells **(A)** and HT-29 holoclone **(B)** stained for NANOG protein

Through immunofluorescence, a change in NANOG protein expression in HT-29 holoclones was observed. Only a small percentage of parental cells showed visible signs of NANOG protein expression, perhaps representative of the small percentage of cells, reported as being as low as 0.1-0.3%, that have cancer stem cell potential in a tumour population (Chan et al., 2011). HT-29 holoclone cells showed widespread fluorescence for NANOG protein

expression compared to parental cells. These results correlate with gene expression results in section 3.4.4 in which a substantial 246-fold increase in NANOG expression was recorded.

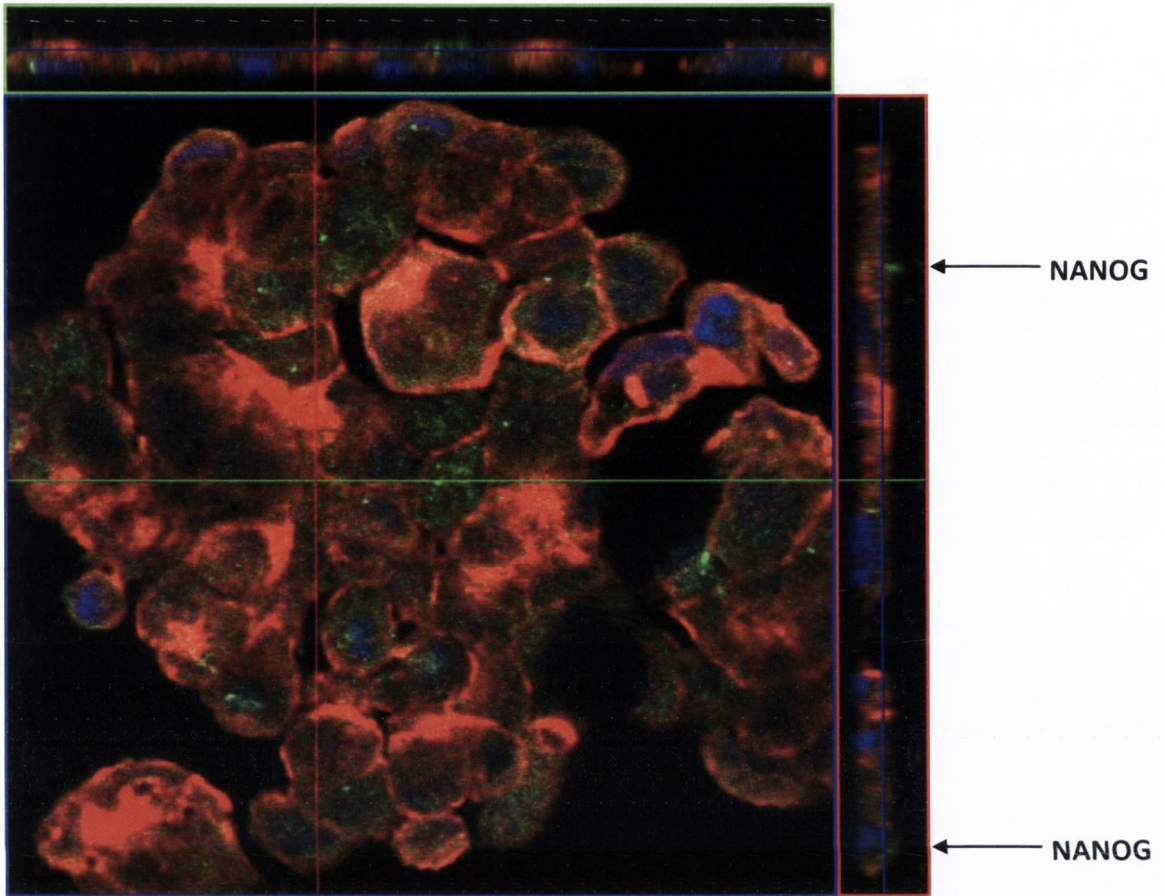


Figure 3.5.1(b): XY and ZY views illustrating the localisation of NANOG protein in a HT-29 holoclone.

In the HT-29 holoclone shown in Figure 3.5.1(b), NANOG appeared to be predominantly localized within the cell nucleus evident from its overlap with cell nuclei staining. However, there was evidence of NANOG protein localization to the apical side of the cell and on the cell surface as evidenced by orthogonal view of the cells and location of NANOG above the blue lines.

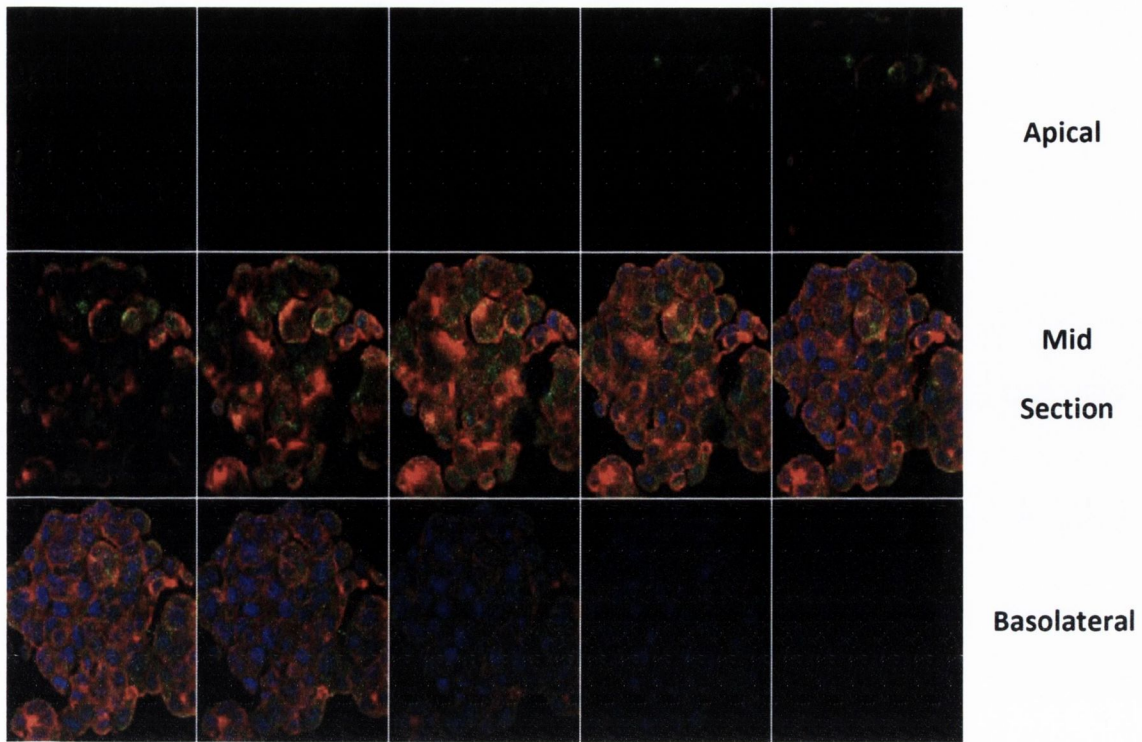


Figure 3.5.1(c): Z-Stacks illustrating NANOG expression in a HT-29 holoclone.

As in Figure 3.5.1(b), in HT-29 holoclones, NANOG appeared to be predominantly localized within the cell nucleus of a HT-29 holoclone.

Similarly, in an XYZ Stack view of the cells shown in Figure 3.5.1(c) there was evidence of NANOG localization to the apical side of the cell on the cell surface. NANOG was observed on the apical and mid-section of the cells with little NANOG evident in the basolateral sections.

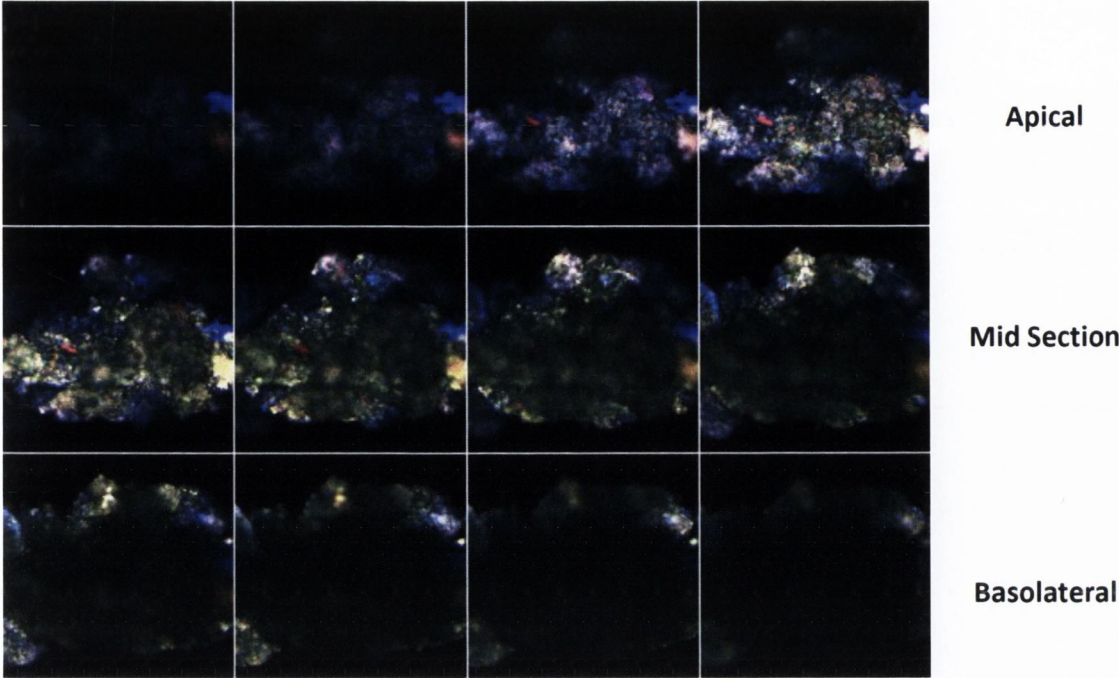


Figure 3.5.1(d): Confocal images illustrating NANOG protein localisation in HT-29 holoclone

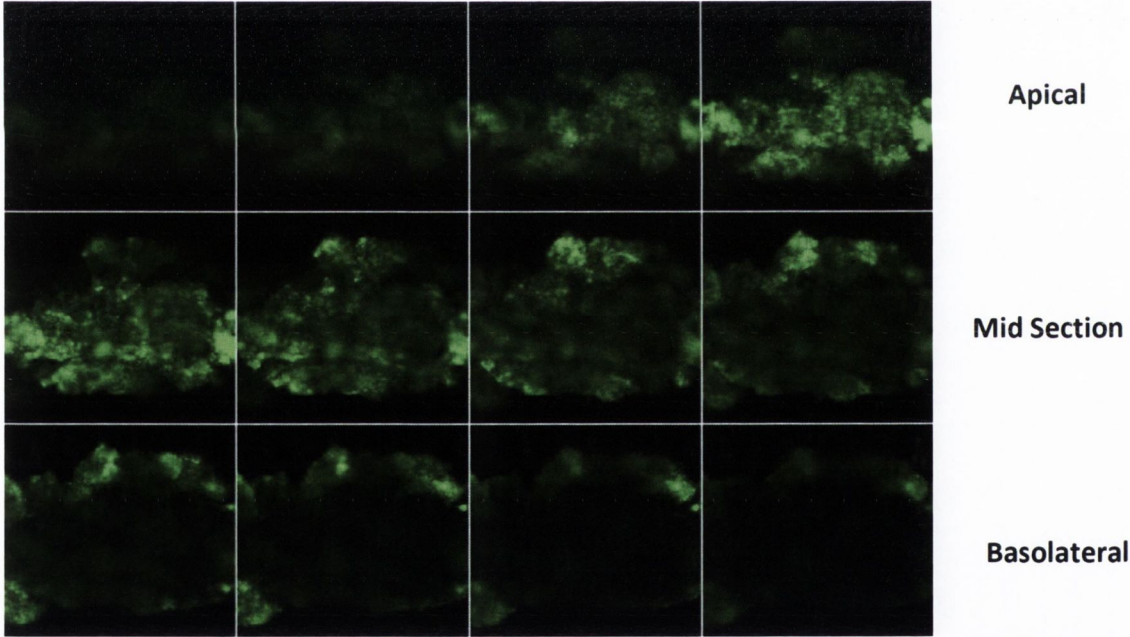


Figure 3.5.1(e): Confocal images illustrating NANOG protein localisation in HT-29 holoclone (no DAPI and no TRITC staining)

In the HT-29 holoclone in Figure 3.5.1(d), NANOG appears to be localized to the apical and mid-section of the scans and is largely absent on the basolateral sides of the scans.

In the same COLO320 holoclone in Figure 3.5.1(e), NANOG appears to be localized to the apical side of the cell on the cell surface as evident from the location of NANOG on apical and mid-section of the cells whereas there is little evidence on the basolateral sides.

3.5.2 COLO320

DAPI (4', 6-diamidino-2-phenylindole) was used to stain the cell nuclei (**Blue**), Alexa488 was used to stain NANOG (**Green**) and Rhodamine Phalloidin Tetramethylrhodamine B isothiocyanate (TRITC) was used to stain actin filaments (**Red**). The fourth panel displays the merged images of all three channels.

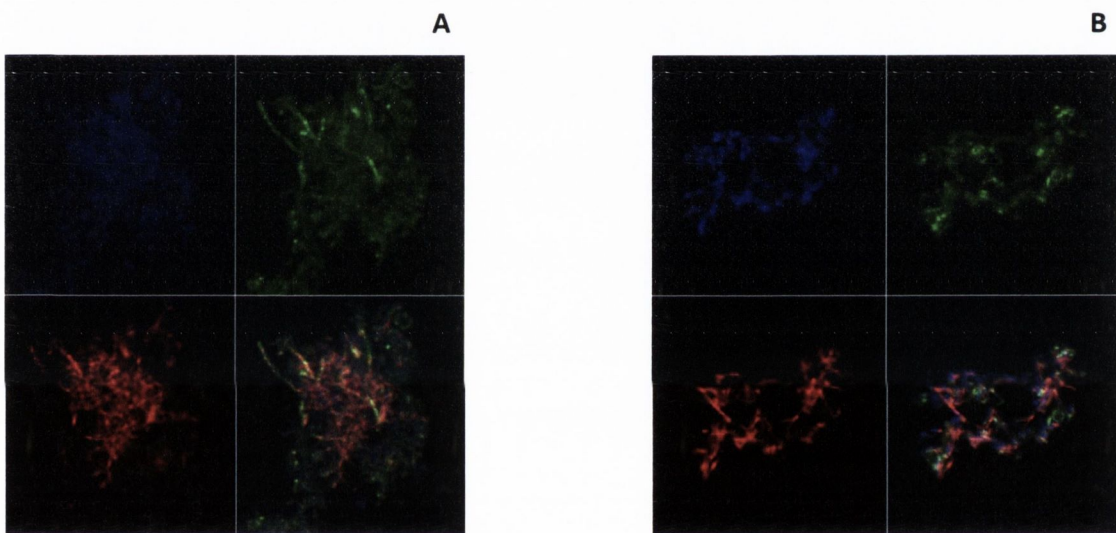


Figure 3.5.2(a): Confocal images of COLO320 parental cells **(A)** and COLO320 holoclone **(B)** stained for NANOG protein

Through immunofluorescence, a change in NANOG protein expression in COLO320 holoclones was observed. In contrast to BRAF V600E mutated HT-29 holoclones, COLO320 holoclones cells exhibited decreased NANOG protein expression compared to parental cells. Fluorescence intensity was greater in parental cells and these observations correlate with gene expression results in section 3.4.4 in which a significant decrease of 200-fold in NANOG expression was recorded in COLO320 holoclones.

3.5.3 Discussion

Based on confocal microscopy results, a correlation between gene and protein expression of the stem cell marker NANOG is evident. In HT-29 holoclones, a substantial increase in NANOG expression was recorded in the holoclone population relative to their founder cells. A similar increase was observed in NANOG protein expression in HT-29 holoclones compared to parental cells in which very low levels of the protein were visible by immunofluorescence indicating a strong correlation between NANOG transcription and translation levels. In contrast, in BRAF wild type COLO320 holoclones, a significant increase in NANOG expression was recorded in the holoclone population relative to their founder cells. Evident from confocal microscopy images, a decrease was observed in NANOG protein expression in COLO320 holoclones compared to parental cells. Parental cells had higher expression levels of NANOG protein, which like its BRAF V600E mutated counterpart, shows a strong correlation between decreased levels of the NANOG protein at both the message and protein levels.

NANOG expression and protein localisation has previously been described as being restricted to the inner cell mass (ICM) in mouse and human blastocysts, and are used to characterize undifferentiated ES cells in vitro (He et al., 2006). In HT-29 holoclones, we have shown that NANOG protein appeared to be predominantly localized within the cell nucleus of the inner cell mass that makes up a holoclone suggesting that CSCs derived from a BRAF V600E mutated colorectal adenocarcinoma possess similar expression patterns of the stem cell marker NANOG. In contrast, expression of NANOG protein in COLO320 holoclones decreased but in parent cells where NANOG expression was abundant, localization was predominantly within the cell nucleus of the inner cell mass. In conclusion, from correlated gene expression and protein expression results, our findings suggest that in CSC progenies that show overexpression of stemness marker NANOG, there is a correlation with increased NANOG protein expression which shows localization patterns similar to ES cells.

Chapter 4

Non-Coding RNA Expression in Holoclones Derived From Cell Lines +/- BRAF V600E Mutation

4.1 Introduction

Cancer stem cells have been reported in many human tumours and are thought to be responsible for tumour initiation, therapy resistance and metastasis. Currently, the regulation of cancer stem cells is not well understood at a molecular level despite recent research into the involvement of miRNAs in CSCs. We seek to identify dysregulated miRNAs in cancer stem cell populations derived from a range of tumours with and without BRAF V600E mutation. Similarly, by identifying validated target genes and biological processes dysregulated in holoclone populations we may highlight in greater detail the role of non-coding RNA in cancer stem biology.

Aims & Objectives

- Interrogate the miRNA profile of BRAF wild type and V600E mutant cell lines and derived holoclone populations using the Applied Biosystems SOLiD™ 4 Next Generation sequencing platform
- Investigate overrepresented Gene Ontology terms in holoclone populations to identify dysregulated biological processes unique to cancer stem cell progenies
- Interrogate the snoRNA profile of BRAF wild type and BRAF V600E mutant cell lines and derived holoclone populations using the Applied Biosystems SOLiD™ 4 Next Generation sequencing platform

4.2 SOLiD™ 4 Next generation sequencing and annotation of miRNAs

Global miRNA expression was analysed among parent and holoclone samples for SK-Mel 28, COLO794, 8505C, TPC-1, ES-2, A2780, COLO320 and HT-29 cell lines on the SOLiD™ 4 next generation sequencing platform. Workflow analysis (WFA) outlined in section 2.13.1, indicated a 0.4pM barcoded library was optimal (Figure 4.2(b)), however, as a single ePCR reaction failed during template bead preparation there was insufficient quantity of 0.4pM library and as such, the deposited library consisted of a mix of 0.4pM and 0.8pM libraries. We calculated that 708 million P2+ (enriched) beads were required for deposition. Assuming a loss of 15% during the deposition procedure, 814 million P2+ beads were required. Using data from the WFA, 1.085×10^9 beads with an average 75% P2+ beads, were prepared and deposited. Following the sequencing run, the data was assessed and analysed bioinformatically. The sequencing run passed all primary analyses and was deemed successful before colour space data was converted to 35bp reads and mapped to the reference.

A series of software programmes were used to bioinformatically identify and remove contaminants, and sequences were annotated based on their overlap with publicly available genome annotations of miRNAs, rRNAs, tRNAs, other small RNAs and genomic repeats. The total reads and proportion of miRNAs per mapped sample is found in Table 4.2(a). It should be noted that these merely represent the annotated reads and do not include novel transcripts. MiRNAs were the most abundant class of annotated small RNA on average. As there was variability in the number of total reads between samples, annotated reads were normalised to allow analysis of differential expression across samples (Li et al, 2011), using the following formula:

$$\text{Coverage target miRNA} = \frac{\text{count miRNA target}}{\text{Total count of miRNAs in sample}}$$

Differentially expressed miRNAs, based on normalised sequence number variations, were identified in holoclone samples and to increase stringency, any miRNA with ≤ 10 reads was excluded and the relative fold change of holoclone samples was calculated as follows:

$$\text{Fold Change} = \frac{\text{Coverage miRNA Holoclone}}{\text{Coverage miRNA Parent}}$$

A Log₂fold change of (+/-) 2-fold was considered biologically significant and analysed further. MiRNAs uniquely identified in either parent or holoclone samples were considered significant if >100 reads were identified.

Sample	Barcode	Barcode Sequence	Total miRNA reads
SK-Mel 28 Parent	24	GGGTCGGTAT	22,314,094
SK-Mel 28 Holoclone	45	GTA CTTGGCTC	4,511,702
COLO794 Parent	47	GAGGGATGGC	2,622,105
COLO794 Holoclone	5	GTGGTGTAAG	5,720,083
8505C Parent	7	GGGTTATGCC	9,539,315
8505C Holoclone	46	GGTCGTCGAA	6,733,553
TPC-1 Parent	9	AGGTTGCGAC	1,067,223
TPC-1 Holoclone	12	AAGAGGAAAA	7,802,720
ES-2 Parent	3	ATAGGTTATA	5,261,444
ES-2 Holoclone	2	AGGGAGTGGT	2,970,029
A2780 Parent	6	GCGAGGGACA	2,595,122
A2780 Holoclone	11	GTGCGACACG	3,714,542
HT-29 Parent	4	GGATGCGGTC	7,915,241
HT-29 Holoclone	10	GCGGTAAGCT	8,534,508
COLO320 Parent	1	GTGTAAGAGG	3,940,032
COLO320 Holoclone	48	GCCGTAAGTG	2,760,209

Table 4.2(a): Uniquely barcoded samples and total number of sequence reads generated for each library. All parent and holoclone small RNA libraries were assigned with unique nucleotide barcodes for bioinformatical identification purposes. Based on 120-130bp percentages determined through smear analysis of amplified libraries outlined in section 2.9.4, percentages were used to establish total miRNA reads generated for each individual sample.

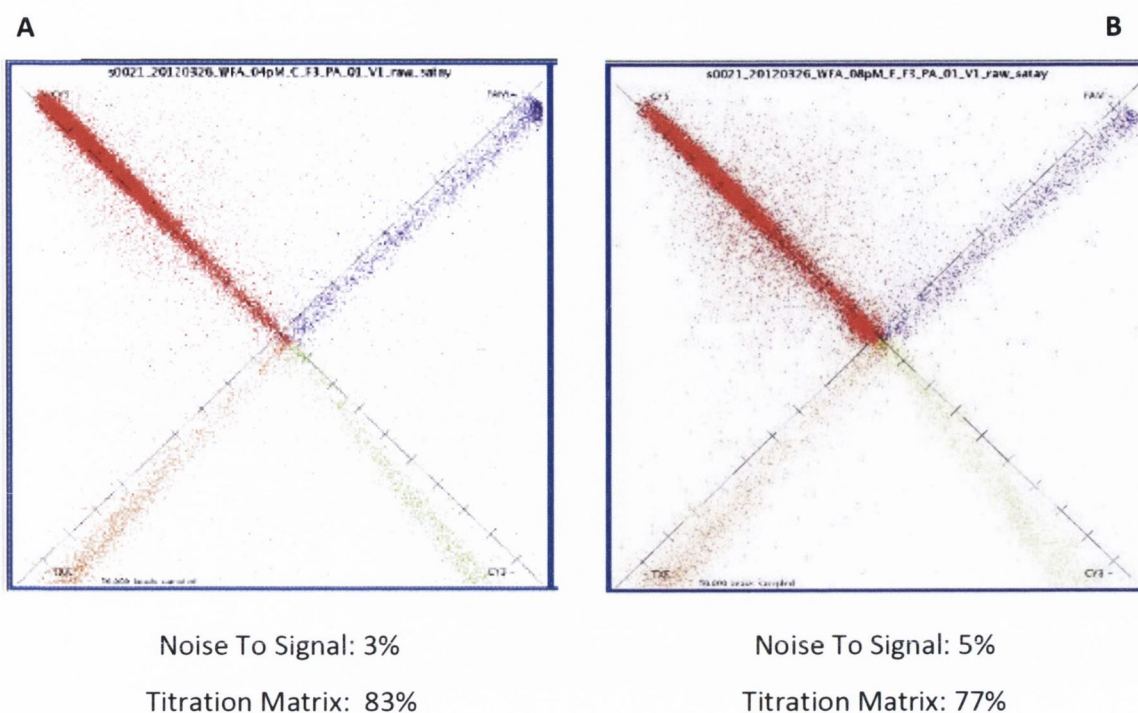


Figure 4.2(b): Workflow Analysis (WFA) for SOLiD next generation sequencing.

A WFA was performed for each successful ePCR for both the (A) 0.4pM and (B) 0.8pM barcoded sequencing library. Satay plots from a representative ePCR WFA for each library concentration are shown here. Satay plots demonstrate both bead spectral purity and signal intensity. Desired beads are spectrally pure and thus monochromatic, found within a 9° angle window of the axis. The noise-to-signal ratio (N2S) is the ratio between the brightest and second brightest dye and is an index of polyclonal beads, but deposition can also influence N2S: if beads are not properly declumped prior to deposition they can increase noise.

The titration matrix = (% p2⁺ beads) × (% On axis (within 10°) beads).

Both libraries met the criteria required for sequencing: >75% on-axisbeads and <25% N2S. A higher titration matrix is desired for optimal sequencing and therefore the 0.4pM library with a titration matrix of 83%, was preferred for this sequencing run.

4.3 Differential miRNA expression in holoclone populations

Total sequencing read numbers were generated and broken down for each uniquely barcoded sample library. Using Bioanalyser smear analysis, miRNA library percentages were recorded for each library and these were used to establish the total small RNA sequences generated for each sample. For example, a sample that generated 10 million sequencing reads with a miRNA library percentage of 50% established through smear analysis had 5 million total miRNA reads. These reads were then mapped back to different chromosomes on the human genome and parent reads were bioinformatically compared to holoclone reads to identify differentially expressed miRNAs. For each differentially expressed miRNA, a comprehensive array of data was generated including:

- Gene Symbol
- Parental and Holoclone Counts
- Fold Change (Log_2)
- Chromosomal Location
- Biotype (miRNA)

4.3.1 Melanoma

SK-Mel 28 BRAF V600E mutant

Sequencing of the SK-Mel 28 Parent Small RNA library yielded 22,314,094 unique 120-130bp sequencing reads and 4,511,702 reads for the SK-Mel 28 Holoclone Small RNA library. A total of 89 differentially expressed miRNAs (Appendix III, Table A.3.1.1(a)) was recorded between SK-Mel 28 parent and holoclone samples with 53 up regulated miRNAs being observed and 36 down regulated miRNAs.

In differentially expressed miRNAs in SK-Mel 28 holoclones, we identified overexpression of miR-103 transcripts. Expression of miR-103 in holoclone populations was decreased 2-fold in comparison to SK-Mel 28 parental cells with showed a substantial level of expression with more than 4,000 reads being recorded.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-103a2	20	22350.6	4848.9	-2.21

Table 4.3.1(a): Fold change of miR-103a in SK-Mel 28 holoclones

A novel miRNA transcript AC118345.2 located on chromosome 2 was identified as being substantially expressed in SK-Mel 28 holoclone populations with over 1,000 reads being recorded with none identified in parental cells, however its function is not yet known.

COLO794 BRAF Wild Type

Sequencing of the COLO794 Parent Small RNA library yielded 2,622,105 unique 120-130bp sequencing reads and 5,720,083 reads for the COLO794 Holoclone Small RNA library. A total of 112 differentially expressed miRNAs (Appendix III, Table A.3.1.1)) was recorded between COLO794 parent and holoclone samples with 50 up regulated miRNAs identified and 52 down regulated miRNAs.

In our study of differentially expressed miRNAs in COLO794 holoclones, we identified overexpression of miR-26 transcripts. Expression of miR-26a1 was increased 4-fold in COLO794 holoclones and a 2-fold up regulation of miR-26a2 was recorded.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-26a1	3	52.1	777.8	3.90
miR-26a2	12	973.3	4768.9	2.29

Table 4.3.1(b): Fold change of miR-26 transcripts in COLO794 holoclones

4.3.2 Thyroid

8505C BRAF V600E Mutant

Sequencing of the 8505C Parent Small RNA library yielded 9,539,315 unique 120-130bp sequencing reads and 6,733,553 reads for the 8505C Holoclone Small RNA library. A total of

93 differentially expressed miRNAs (Appendix III, Table A.3.1.2(a)) was recorded between BRAF V600E mutated 8505C parent and holoclone samples with 39 up regulated miRNAs being observed and 54 down regulated miRNAs.

Similar to observations made in BRAF V600E mutated SK-Mel 28 holoclones, miR-103a was also down regulated in BRAF V600E mutated 8505C thyroid holoclones. We identified over a 2-fold underexpression of miR-103a in holoclone populations which showed substantial expression in parental cells with over 15,000 reads recorded.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-103a2	20	15330.5	3203.5	-2.26

Table 4.3.2(a): Fold change of miR-103 in 8505C holoclones

A noteworthy observation in differential miRNA expression between BRAF V600E mutated 8505C holoclones and parent cells was expression of the miR-302 family. The 4 family members of the miR-302 cluster located on chromosome 4 showed increased expression by more than 4-fold with the greatest increase in 8505C holoclones being recorded in miR-302a of more than 5-fold (Table 4.3.2(b)).

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change(Log2)
miR-302a	4	4.8	234.4	5.61
miR-302b	4	23.2	703.2	4.92
miR-302d	4	15.3	429.7	4.81
miR-302c	4	3.7	78.1	4.38

Table 4.3.2(b): Fold change of the miR-302 family members in 8505C holoclones

TPC-1 BRAF Wild Type

These miRNAs are part of a miRNA cluster classified as an embryonic stem cell-specific cell cycle-regulating (ESCC) miRNA family (Subramanyam et al, 2011; Barroso-del Jesus et al.,

2008) along with miR-373, miR-374 and are often referred to as the miR-302/373/374/520 family (Mazda et al, 2011) despite being located on different chromosomes.

Sequencing of the TPC-1 Parent Small RNA library yielded 1,067,223 unique 120-130bp sequencing reads and 7,802,720 reads for the TPC-1 Holoclone Small RNA library. A total of 99 differentially expressed miRNAs (Appendix III, Table A.3.1.2(b)) was recorded between BRAF wild type TPC-1 parent and holoclone samples with 38 up regulated miRNAs being observed and 61 down regulated miRNAs.

In our study of differentially expressed miRNAs in TPC-1 holoclones, we identified overexpression of miR-26 family members. Expression of miR-26b was increased 4-fold in TPC-1 holoclones and a 1.4-fold up regulation of miR-26a1 was recorded.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-26b	2	19.0	382.6	4.34
miR-26a1	3	151.0	403.8	1.42

Table 4.3.2(c): Fold change of miR-26 family members in TPC-1 holoclones

We observed a correlation in the differential expression of the miR-302 family between TPC-1 holoclone populations and their BRAF V600E mutated 8505C holoclone counterparts. All 4 miR-302 family members were up regulated more than 5-fold in TPC-1 holoclones with the greatest increase being recorded in miR-302c of more than 8-fold.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-302c	4	4.	920	8.24
miR-302d	4	4	871	7.85
miR-302b	4	3	701	7.79
miR-302a	4	3	127	5.34

Table 4.3.2(d): Fold change of the miR-302 family members in TPC-1 holoclones

4.3.3 Ovarian

ES-2 BRAF V600E Mutant

Sequencing of the BRAF V600E mutated ES-2 Parent Small RNA library yielded 5,261,444 unique 120-130bp sequencing reads and 2,970,029 reads for the ES-2 Holoclone Small RNA library. Differential expression of 94 miRNAs (Appendix III, Table A.3.1.3(a)) was recorded between BRAF V600E mutated ES-2 parent and holoclone samples with 36 up regulated miRNAs being observed and 58 down regulated miRNAs.

In holoclones derived from ES-2 ovarian cancer cells, we identified an underexpression of 2 miR-103a transcripts. The biggest differential expression change was recorded in miR-103a1 located on chromosome 5 which was down regulated by more than 3-fold in ES-2 holoclones. A down regulation of almost 2-fold was recorded for miR-103a2 in holoclones populations which showed substantial expression in ES-2 parental cells with more than 20,000 reads being recorded.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-103a1	5	200.4	19.9	-3.33
miR-103a2	20	20958.4	6309.9	-1.73

Table 4.3.3(a): Fold change of miR-103a transcripts in ES-2 holoclones

A novel miRNA transcript AC103686.1 located on chromosome 8 was identified as being expressed in ES-2 holoclone populations with over 30 reads being recorded with none identified in parental cells, however its function is not yet known.

A2780 BRAF Wild Type

Sequencing of the BRAF wild type A2780 Parent Small RNA library yielded 2,595,122 unique 120-130bp sequencing reads and 3,714,542 reads for the A2780 Holoclone Small RNA library. Differential expression of 89 miRNAs (Appendix III, Table A.3.1.3(b)) was recorded between BRAF wild type A2780 parent and holoclone samples with 54 up regulated miRNAs being observed and 35 down regulated miRNAs.

Of the differentially expressed miRNAs recorded in A2780 holoclones, overexpressed miR-26 family members were identified. Expression of miR-26a1 was increased almost 2-fold and miR-26b was 1.5-fold up regulated in A2780 holoclone populations.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log ₂)
miR-26a1	3	122.2	400.1	1.71
miR-26b	2	249.3	721.3	1.53

Table 4.3.3(b): Fold change of miR-26 family members in A2780 holoclone samples

4.3.4 Colorectal

HT-29 BRAF V600E Mutant

Sequencing of the BRAF V600E mutated HT-29 Parent Small RNA library yielded 7,915,241 unique 120-130bp sequencing reads and 8,534,508 reads for the HT-29 Holoclone Small RNA library. Differential expression of 62 miRNAs (Appendix III, Table A.3.1.4(a)) was recorded between BRAF V600E mutated HT-29 parent and holoclone samples with 42 up regulated miRNAs being observed and 20 down regulated miRNAs.

As observed in holoclones derived from BRAF V600E mutated melanoma, thyroid and ovarian holoclones, we identified underexpression of miR-103a in BRAF V600E mutated HT-29 holoclones. Expression of miR-103a was down regulated by 2-fold in holoclone populations.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log ₂)
miR-103a1	5	245.6	64.2	-1.94

Table 4.3.4(a): Fold change of miR-103a transcripts in HT-29 holoclones

COLO320 BRAF Wild Type

Sequencing of the wild type COLO320 Parent Small RNA library yielded 3,940,032 unique 120-130bp sequencing reads and 2,760,209 reads for the COLO320 Holoclone Small RNA library. Differential expression of 93 miRNAs (Appendix III, Table A.3.1.4(b)) was recorded between BRAF wild type COLO320 parent and holoclone samples with 34 up regulated miRNAs being observed and 59 down regulated miRNAs.

Overexpression of one transcript of the miR-26 family was recorded in COLO320 holoclones. MiR-26a2 was up regulated by 3-fold and showed substantial expression levels in holoclone populations with more than 2,500 sequencing reads being recorded.

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
miR-26a2	12	322.5	2522.2	2.97

Table 4.3.4(b): Fold change of a miR-26a transcript in COLO320 holoclones

4.3.5 Discussion

Through differential miRNA expression analysis via the SOLiD™ next generation sequencing platform, we have identified different miRNA profiles between BRAF V600E mutant and BRAF wild type cancer stem cell progenies derived from a range of human cancers. The roles of some of these miRNAs have been previously investigated in the oncogenic processes involved in many different tumour types. However, some miRNAs, including dysregulated miRNA families have been identified here for the first time in CSC subpopulations with and without the BRAF V600E mutation. We observed differential miRNA expression profiles across the range of tumour types involved in this study.

Through global miRNA expression analysis in holoclones derived from BRAF V600E mutated melanoma, thyroid, colorectal and ovarian cancer cell lines, we identified a consistently underexpressed subset of miRNA transcripts. Down regulation of miR-103 was recorded across our BRAF V600E mutated holoclone populations. Studies have attempted to establish

the role played by miR-103 in human cancers such as endometrial and colorectal carcinomas (Yu et al., 2012; Chen et al., 2012). However, its role in cancer stem cells is not well understood and through next generation sequencing we have identified its dysregulation in BRAF V600E mutated cancer stem cell progenies. Using the MirWalk validated target identification software, 5 genes within the MAPK-ERK signalling pathway were identified as being targets of miR-103. These gene targets include Fibroblast Growth Factor which acts as a mitogen which can initiate the MAPK-ERK signalling pathway. MiR-103 also targets MAPK14 which functions as one of the kinase proteins within the signalling pathway. Our findings suggest that underexpression of miR-103 in response to BRAF V600E mutation contributes to the enhancement of MAPK-ERK signalling in cancer stem cell progenies harbouring the V600E mutation.

Through differential miRNA expression analysis we have identified consistent overexpression of members of the miR-26 family in holoclones derived from cancer cell lines with wild type BRAF. In holoclones derived from melanoma, thyroid, ovarian and colorectal members of the miR-26 family were up regulated in holoclone populations and represent a novel observation in a cancer stem cell setting. Using the MirWalk validated target identification software, 15 genes within the MAPK-ERK signalling pathway were identified as being targets of miR-26a and miR-26b. These gene targets include transcription factors c-myc and c-Jun targeted by miR-26b and kinase MAPK8 targeted by miR-26a. Despite recent research implicating miR-26 in cancers such as nasopharyngeal carcinoma and metastatic colorectal cancer (Lu et al., 2011; Ma et al., 2011), its role in carcinogenesis is not well known. Here we have described consistent miR-26a overexpression in cancer stem cell progenies derived from tumour types without BRAF V600E mutation. Our findings suggest that miR-26a plays a role in the regulation of many components of the MAPK-ERK signalling pathway in the absence of BRAF V600E mutation. Its conservation across BRAF wild type holoclones suggests that it is important for the maintenance of growth and proliferation regardless of tissue type. Future miR-26 knockdown experiments may investigate if BRAF wild type cancer stem cell progenies depend on miR-26 overexpression in order to maintain survival in a harsh environment simulated by a high salt agar assay.

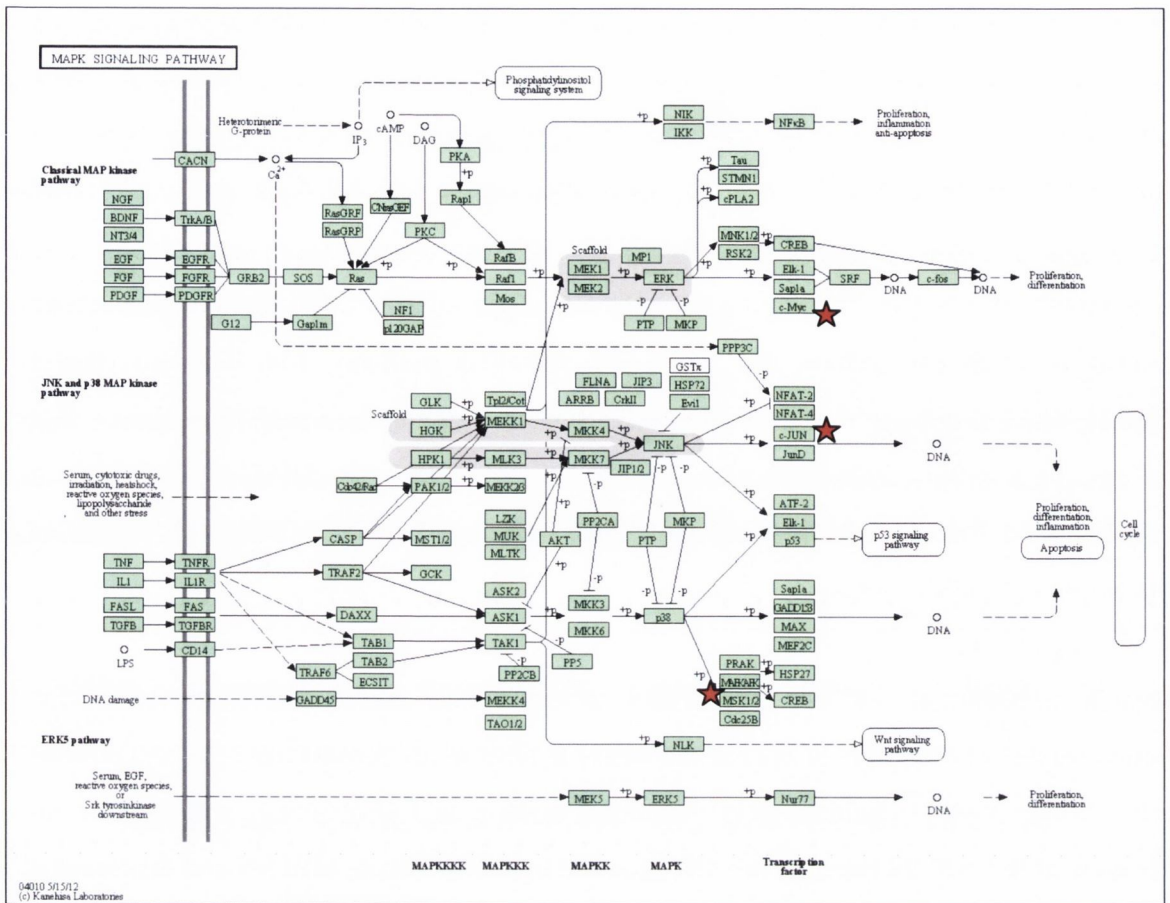


Figure 4.3.5(a): MAPK-ERK signalling pathway in which several genes are targeted by the miR-26 family members miR-26a and miR-26b (validated target genes indicated by the symbol ★ are c-Jun, c-myc and MEK)

A similar pattern of differential expression of the miR-302 family between TPC-1 holoclone populations and their BRAF V600E mutated 8505C holoclone counterparts was recorded as all 4 miR-302 family members were up regulated. In recent years, several studies into the role of the miR-302 family of miRNAs, located on chromosome 4, have identified it as a key regulator of stemness. Having been previously described by Subramanyam et al, 2011; Barroso-del Jesus et al., 2008, as an embryonic stem cell-specific cell cycle-regulating miRNA family, its importance in the maintenance of a stem state in human ES cells has been established. In our differential miRNA expression analysis in cancer stem cell progenies derived from thyroid cancers, we observed an overexpression of the miR-302 family in both BRAF V600E mutated and BRAF wild type holoclones. Regardless of BRAF status, thyroid

holoclonal cells showed a biologically significant increase of all 4 members of the miR-302 family in comparison to their parental cells.

miRNA	8505C Fold Change	TPC-1 Fold Change
miR-302a	5.61	5.34
miR-302b	4.92	7.79
miR-302d	4.81	7.85
miR-302c	4.38	8.24

Table 4.3.5(b): Fold change of the miR-302 family 8505C and TPC-1 thyroid holoclonal cells

Despite showing different expression levels, it is evident that both 8505C and TPC-1 holoclonal cells exhibit a similar pattern of miR-302 family expression. The identification of ES cell transcription factors Oct4 and NANOG as having a regulatory interaction with the miR-302 cluster (Ren et al, 2009), suggests that they may work in concert to establish a highly pluripotent stem-like state in CSCs derived from different types of thyroid cancer. As TPC-1 harbours a RET-PTC mutation upstream of BRAF (Smyth et al., 2005), the MAPK-ERK signalling pathway is hyperactive independent of BRAF mutation. These findings suggest that overexpression of the miR-302 family in holoclonal cells with dysregulated MAPK-ERK signalling plays a role in holoclone stemness. It is possible that this stem profile in both 8505C and TPC-1 holoclonal cells enables the self-renewal potential observed in the high salt agar assay in which holoclone reformation was identified in both thyroid cancer cell lines.

Using global miRNA analysis via SOLiD™ next generation sequencing, we have identified unique dysregulated miRNAs in cancer stem cells derived from BRAF V600E mutated and BRAF wild type cancer cell lines. The importance of these miRNAs in holoclonal cells may be further explored through functional miRNA work to get a greater understanding of their function in cancer stem cells.

4.4 Overrepresented Gene Ontology (GO) terms in holoclone populations

The computationally derived gene targets of the differentially expressed miRNAs with fold changes of greater than +/- 2 were examined by Gene Ontology (GO) "biological processes" classifications. Analytical software was used to identify GO terms under and overexpressed in holoclones for each cell line with significance set to $P < 0.01$. For each cell line, GO terms shared by both up and down regulated groups in holoclones were filtered out and the data generated was made up of uniquely up regulated and down regulated GO terms which contained the following information:

- Gene Ontology (GO) Term
- Number of genes in category
- Overrepresented fold change (Log_2)
- P-value

These lists were then plotted on heat maps, based on generated p-values, to illustrate differential GO term patterns between holoclones derived from a BRAF V600E mutated and BRAF wild type tumour type. Within these lists, searches were carried out to identify overrepresented GO terms involved in several biological processes that may be dysregulated according to the Rosen & Jordan, 2006, CSC hypothesis such as stem cell proliferation, differentiation, DNA damage repair and EMT. The heat maps were designed to illustrate differential patterns of under and overexpressed GO terms in holoclone populations derived from tumour types with and without BRAF V600E mutation.

4.4.1 Melanoma

SK-Mel 28 BRAF V600E Mutated Holoclone Overrepresented GO Terms

A total of 872 overrepresented GO terms in holoclones derived from BRAF V600E mutated SK-Mel 28 melanoma cells were identified (Appendix CD). This total was then broken down into 498 uniquely up regulated and 374 uniquely down regulated GO terms in holoclone populations.

The GO term “negative regulation of stem cell differentiation” was up regulated in SK-Mel 28 holoclones with a 10-fold increase (p-value <0.01) recorded. An interesting observation in up regulated GO terms in SK-Mel 28 holoclones was the overexpression of “positive regulation of epidermal cell differentiation” which increased 7-fold (p-value <0.01) over parental SK-Mel 28 cells. As these holoclones are derived from a malignant melanoma, it is possible that dysregulated differentiation within the epidermis may play a role in the acquisition of stem-like characteristics in malignant melanoma subpopulations harbouring the BRAF V600E mutation. Overexpression of “epithelial-mesenchymal cell signalling” was identified in SK-Mel 28 holoclones with a 14-fold (p-value <0.01) increase being recorded.

Of the 374 uniquely down regulated GO terms, “stem cell proliferation” decreased 8-fold (p-value <0.01) suggesting dysregulated miRNAs may be playing a role in loss of proliferation control in holoclones derived from SK-Mel 28 malignant melanoma cells. A 20-fold decrease (p-value <0.01) in genes involved in the canonical Wnt receptor signalling pathway involved in “positive regulation of epithelial to mesenchymal transition” was recorded in SK-Mel 28 holoclone populations.

The observation of overrepresented GO terms in SK-Mel 28 holoclones may highlight in greater details the downstream effect of differentially expressed miRNAs in CSC subpopulations in a BRAF V600E mutated malignant melanoma

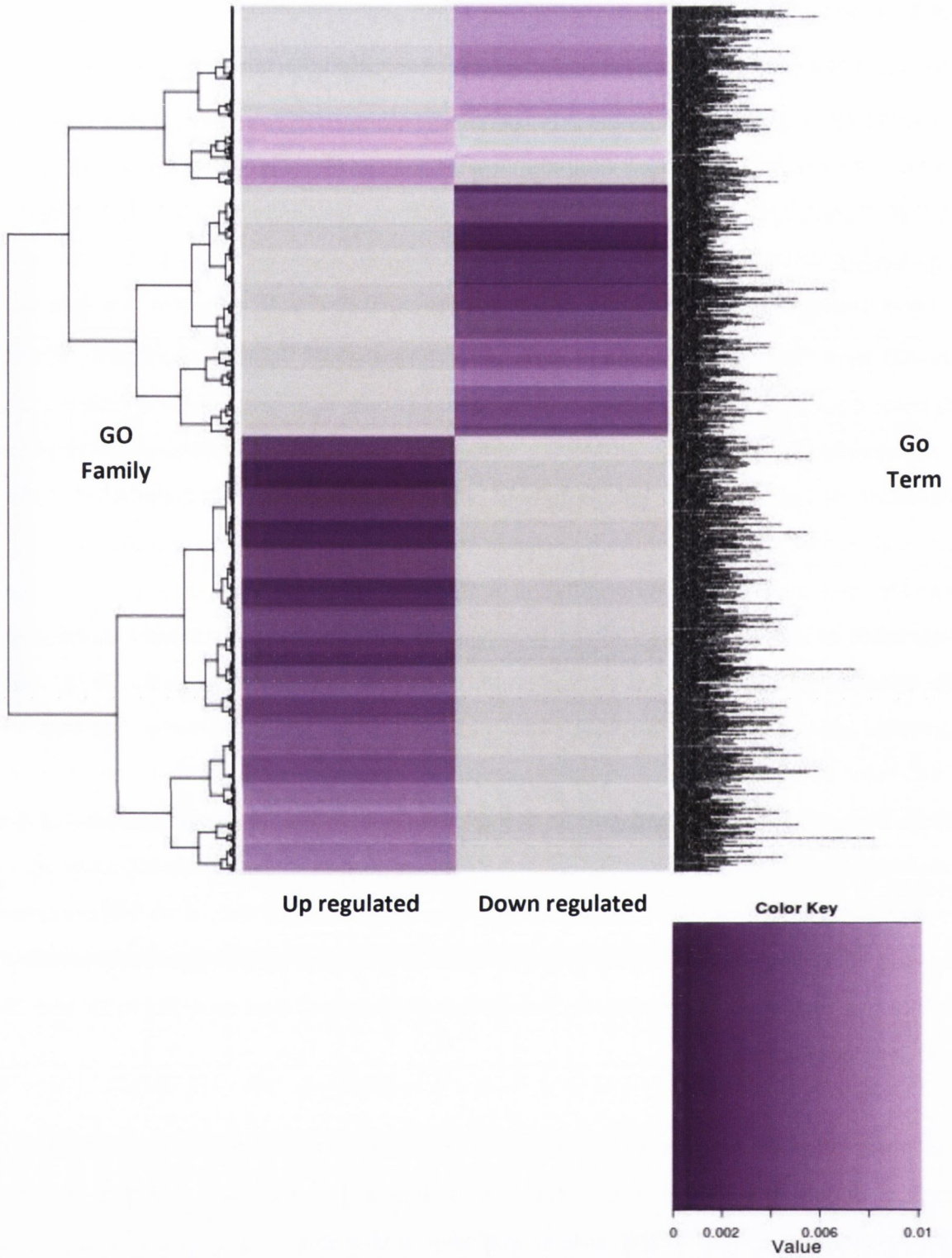


Figure 4.4.1(a): Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated SK-Mel 28 melanoma cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

COLO794 BRAF Wild Type Holoclone Overrepresented GO Terms

A total of 780 overrepresented GO terms in holoclones derived from BRAF wild type COLO794 melanoma cells were identified (Appendix CD). This total was then broken down into 478 uniquely up regulated and 312 uniquely down regulated GO terms in holoclone populations.

Of the 478 uniquely up regulated GO terms, “negative regulation of mesenchymal cell proliferation” was increased by 10-fold (p-value <0.01) in COLO794 holoclones. 4 genes involved in this process were identified as being under the regulation of differentially expressed miRNAs in holoclone populations. A 6-fold overexpression (p-value <0.01) in “regulation of endothelial cell differentiation” as a result of miRNA regulation was recorded in COLO794 holoclones.

The GO term “mesenchymal stem cell differentiation” was underexpressed by 14-fold (p-value <0.01) in COLO794 holoclones. One of the most down regulated GO terms recorded in COLO794 holoclones was the canonical Wnt receptor signaling pathway involved in “positive regulation of epithelial to mesenchymal transition”. A 14-fold decrease (p-value <0.01) in expression of genes in this category through miRNA regulation, was recorded in holoclone populations and represented one of the most underexpressed biological processes observed in COLO794 holoclone populations.

Taken together, these findings suggest that COLO794 holoclones express dysregulated miRNAs that may contribute to the underexpression of biological events such as stem cell differentiation and EMT regulation. Identifying differentially expressed miRNAs and highlighting their effects on biological processes can begin to give a greater insight into the molecular characteristics of CSC progenies derived from a BRAF wild type malignant melanoma.

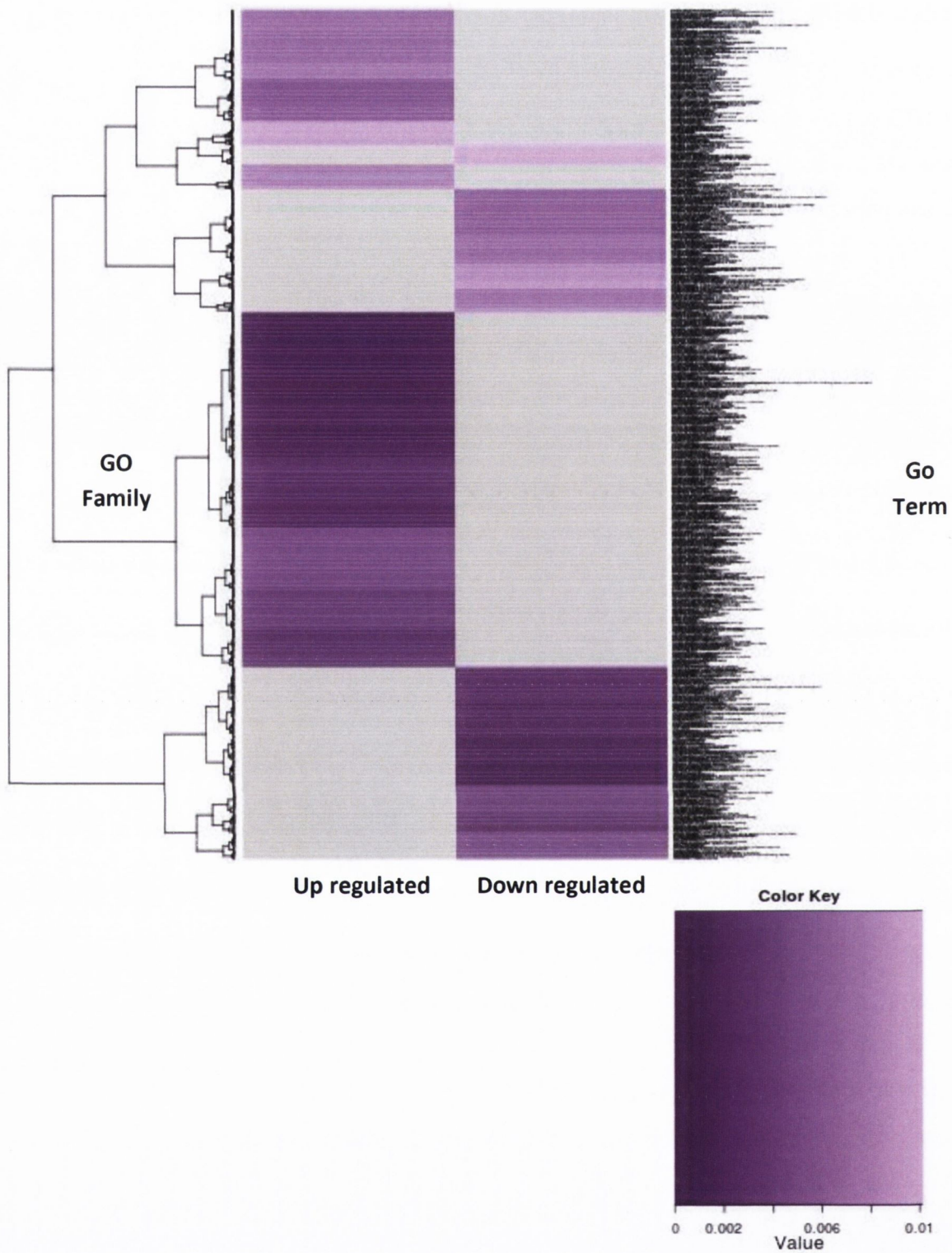


Figure 4.4.1(b): Heat map illustrating up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type COLO794 melanoma cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

4.4.2 Thyroid

8505C BRAF V600E Mutated Holoclone Overrepresented GO Terms

A total of 1,123 overrepresented GO terms in holoclones derived from BRAF V600E mutated 8505C thyroid cells were identified (Appendix CD). This total was then broken down into 503 uniquely up regulated and 620 uniquely down regulated GO terms in holoclone populations.

Of the 503 uniquely up regulated GO terms, 5 terms directly involved in stem cell biological processes were identified in 8505C holoclones. A 7-fold increase (p-value <0.01) in “regulation of stem cell proliferation” was recorded and a similar observation was made in the “regulation of stem cell maintenance” which was overexpressed by 6-fold (p-value <0.01) in the 8505C holoclone population. The “stem cell proliferation” GO term was also overexpressed by 6-fold (p-value <0.01) in 8505C holoclones based on validated gene targets of differentially expressed miRNAs. “Somatic stem cell division” was overexpressed by 5-fold (p-value <0.01) and the “stem cell division” GO term increased by 4-fold (p-value <0.01), suggesting increased division capabilities in CSC progenies derived from 8505C parent cells. Coupled with evidence of self renewal in the high salt agar assay, these results indicate that dysregulated miRNAs may contribute to 8505C holoclones possessing a highly proliferative stem identity.

“Negative regulation of mesenchymal cell proliferation” was 9-fold (p-value <0.01) underexpressed in 8505C holoclones. Of particular relevance in underexpressed GO terms was the “negative regulation of epithelial to mesenchymal transition”. A 6-fold underexpression (p-value <0.01) of this biological process was recorded and taken together these results suggest that regulation of EMT may be lost in 8505C holoclones and subsequent proliferation of mesenchymal-like cell populations may be increased. It is possible that as a result of miRNA dysregulation, CSC progenies derived from a BRAF V600E mutated cancer have a strong stem identity with the potential to undergo EMT and proliferate.

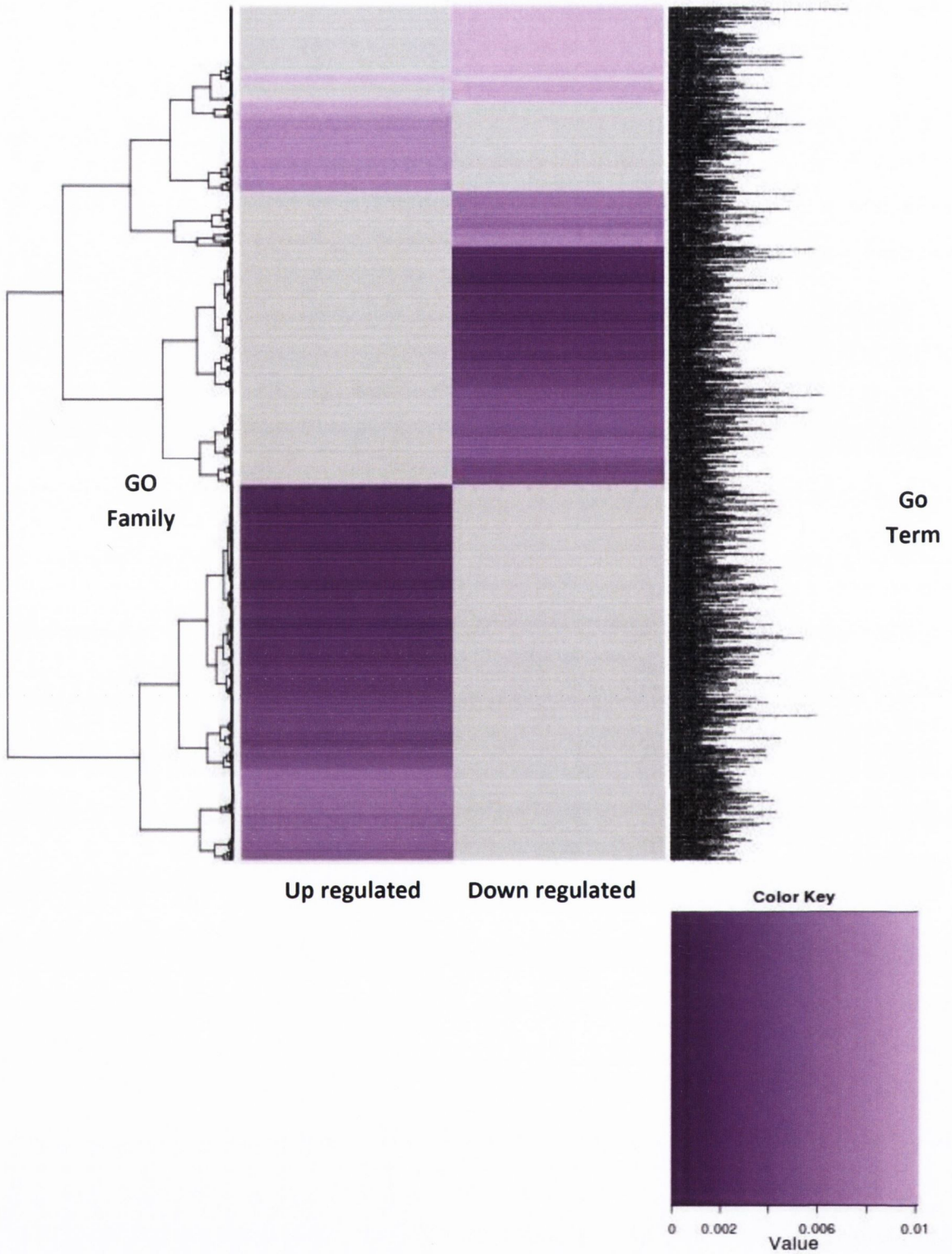


Figure 4.4.2(a): Heat map of up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated 8505C thyroid cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

TPC-1 BRAF Wild Type Holoclone Overrepresented GO Terms

A total of 993 overrepresented GO terms in holoclones derived from BRAF wild type TPC-1 thyroid cells were identified (Appendix CD). This total was then broken down into 428 uniquely up regulated and 505 uniquely down regulated GO terms in holoclone populations.

Of the 428 uniquely up regulated GO terms, similar to their BRAF V600E mutated counterparts, many overexpressed processes were stem cell biological processes. The “regulation of stem cell proliferation” GO term was overexpressed in holoclone populations by 9-fold (p-value <0.01). Similarly, gene targets involved in the “regulation of stem cell maintenance” were also overexpressed by 8-fold (p-value <0.01), suggesting that TPC-1 holoclones have stronger proliferative and maintained stem capabilities over their parental cells. Enhanced divisional capabilities of TPC-1 holoclone cells is suggested by the observation of 5-fold overexpressed (p-value <0.01) genes involved in “somatic stem cell division”. An interesting observation was in the biological process of “stem cell fate commitment” in TPC-1 holoclones. We recorded a 14-fold increase (p-value <0.01) in this GO term in which 2 target genes of differentially expressed miRNAs were overexpressed.

Many of the 505 uniquely down regulated GO terms recorded in TPC-1 holoclones showed similarity to GO terms underexpressed in 8505C holoclones. The most comparable of these were biological processes involved in EMT in which several target genes of differentially expressed miRNAs were underexpressed in holoclone populations. The most biologically relevant of these findings was the “negative regulation of mesenchymal cell proliferation” which was underexpressed by 7-fold (p-value <0.01) in TPC-1 holoclones.

The identification of several overexpressed GO terms involved in stemness suggests that like their BRAF V600E mutated counterparts, TPC-1 holoclones have a strong identity. It is possible that the combination of increased stem cell division and maintenance recorded through up regulated GO terms plays a role in establishing a stem-like state in holoclones derived from a BRAF wild type thyroid cancer.

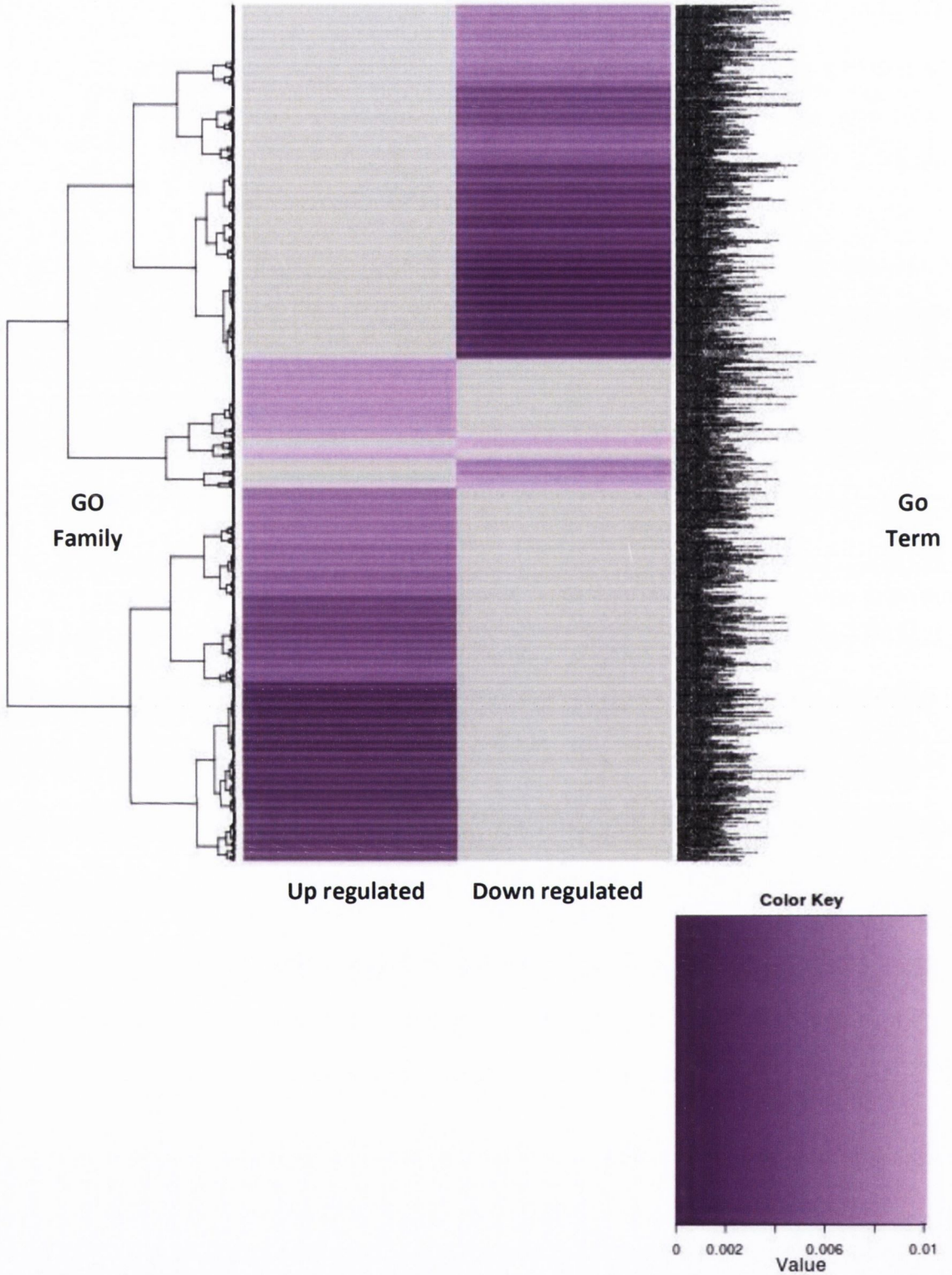


Figure 4.4.2(b): Heat map illustrating up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type TPC-1 thyroid cancer cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

4.4.3 Ovarian

ES-2 BRAF V600E Mutated Holoclone Overrepresented GO Terms

A total of 987 overrepresented GO terms in holoclones derived from BRAF V600E mutated ES-2 ovarian cells were identified. This total was then broken down into 606 uniquely up regulated and 381 uniquely down regulated GO terms in holoclone populations.

Several of the 606 uniquely up regulated GO terms identified in ES-2 holoclones were overexpressed stem cell biological processes. An 11-fold overexpression (p-value <0.01) of 3 genes involved in the “positive regulation of mesenchymal stem cell differentiation” was recorded. 3 genes positive regulation of “stem cell differentiation” were overexpressed by 11-fold (p-value <0.01) indicating a possible suppression of differentiation in cells with mesenchymal-like properties in ES-2 holoclone populations. Similarly, a 5-fold overexpression (p-value <0.01) of genes involved in “stem cell proliferation” was recorded indicating enhanced proliferative capabilities in CSC progenies derived from ES-2 ovarian cancer cells.

Of the 381 uniquely down regulated GO terms identified in ES-2 holoclones, only one stem cell GO term was identified as being underexpressed. A 10-fold underexpression (p-value <0.01) of 3 genes involved in “positive regulation of stem cell proliferation” was recorded in ES-2 holoclone populations. An interesting observation in down regulated GO terms was the “regulation of timing of cell differentiation” which was 6-fold underexpressed (p-value <0.01) indicating a possible inability in ES-2 holoclones to control differentiation and may play a role in keeping cells in a stem-like state.

The results of over and underexpressed GO terms as a result of differentially expressed miRNAs, may give a greater insight into the aberrant biological processes taking place within ES-2 holoclones. Increased stem cell proliferation and repression of differentiation suggests that CSC progenies derived from a BRAF V600E mutated ovarian carcinoma may have enhanced stem characteristics and display reduced differentiation mechanisms.

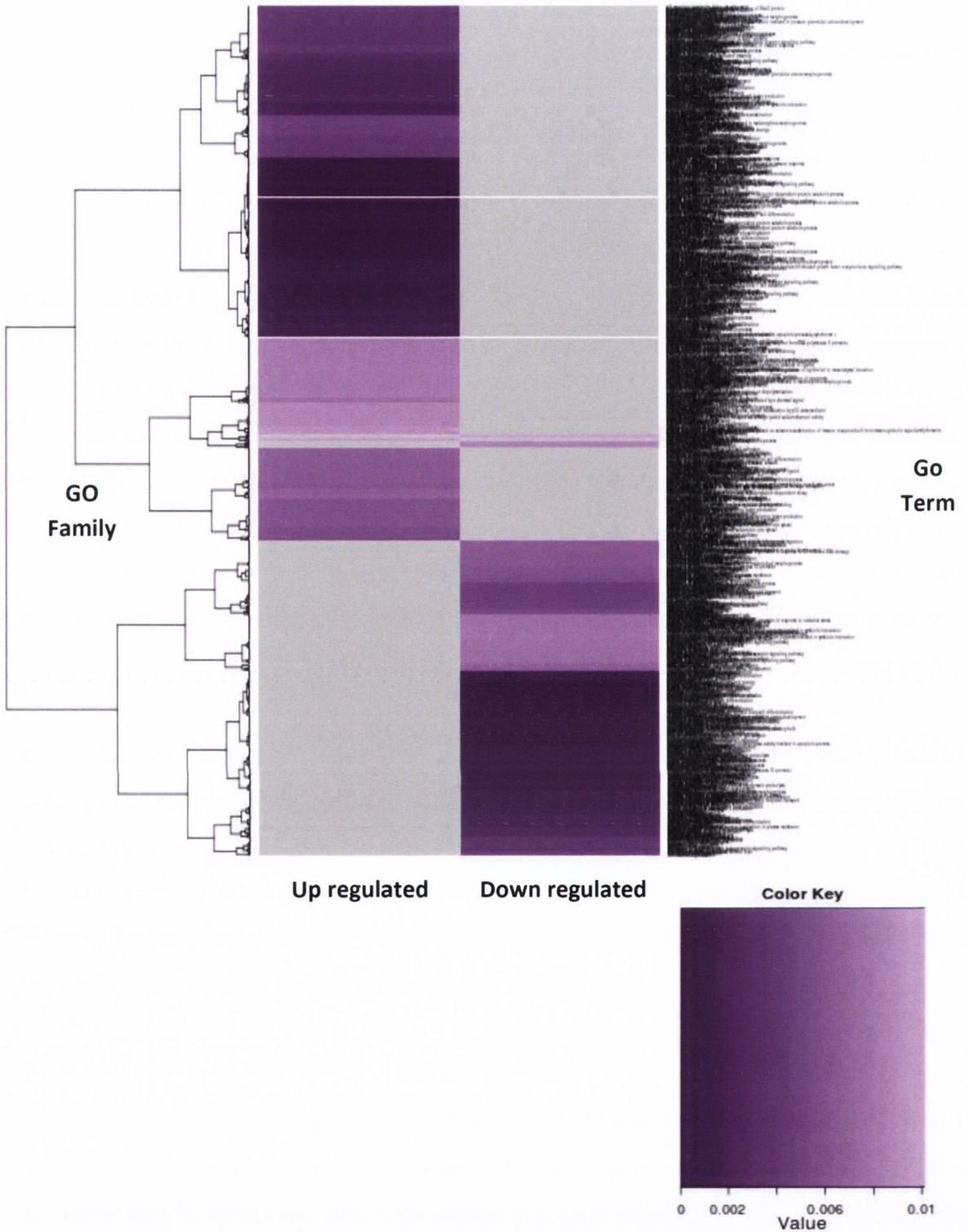


Figure 4.4.3(a): Heat map illustrating up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated ES-2 ovarian cancer cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

A2780 BRAF Wild Type Holoclone Overrepresented GO Terms

A total of 1,053 overrepresented GO terms in holoclones derived from BRAF wild type A2780 ovarian cells were identified. This total was then broken down into 696 uniquely up regulated and 357 uniquely down regulated GO terms in holoclone populations.

2 of the most highly overexpressed terms in BRAF wild type A2780 holoclones were involved in stem cell fate. An overexpression of 18-fold (p-value <0.01) was recorded in 3 genes involved in “stem cell fate commitment” and similarly an 18-fold overexpression (p-value <0.01) of “stem cell fate specification” was observed in A2780 holoclone populations. Correlating with observations made in BRAF V600E mutated ES-2 holoclones, many of the overexpressed terms recorded in A2780 holoclone populations were involved in stem cell regulation and proliferation. These overrepresented GO terms included “stem cell proliferation” which was overexpressed by 7-fold (p-value <0.01) in A2780 holoclones and a 4-fold overexpression of 5 genes involved in “stem cell division” (p-value <0.01).

3 of the top 20 underexpressed GO terms identified in A2780 holoclones were involved in stem cell differentiation. These highly down regulated GO terms were: “positive regulation of mesenchymal stem cell differentiation” which was underexpressed by 32-fold (p-value <0.01), “positive regulation of stem cell differentiation” also underexpressed by 32-fold with a p-value of <0.01 and “regulation of mesenchymal stem cell differentiation” which was underexpressed by 25-fold (p-value <0.01). The scale of the fold change in these underexpressed GO terms suggests that A2780 holoclones may lack the biological mechanisms that promote differentiation, thus keeping them in a pluripotent stem-state.

The results recorded in A2780 holoclones differ from their BRAF V600E mutated ES-2 holoclone counterparts in that mechanisms that govern stem cell fate were overexpressed. This observation that was unique to holoclones derived from a BRAF wild type ovarian carcinoma may highlight the effects that dysregulated miRNA expression has on stem cell maintenance and differentiation.

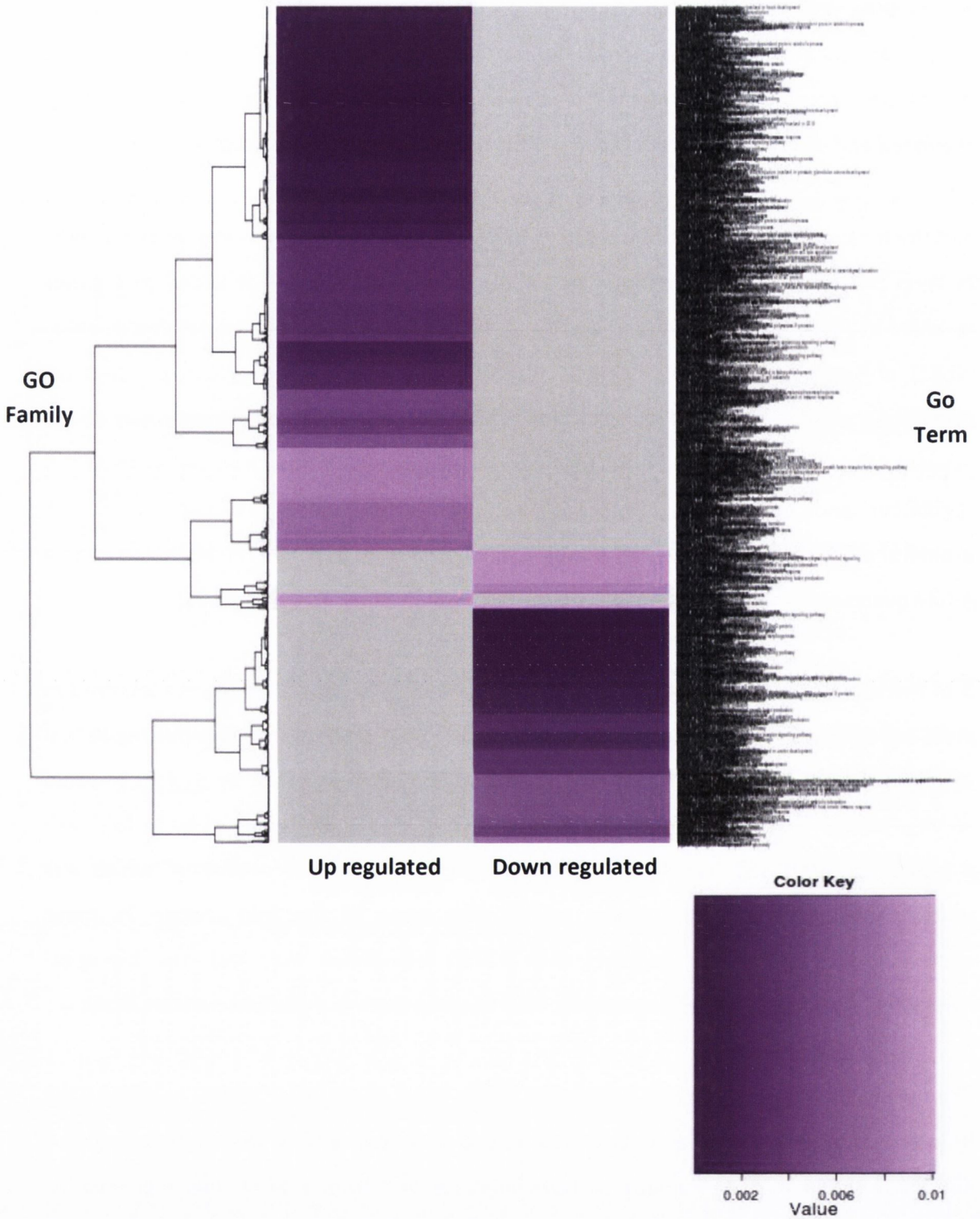


Figure 4.4.3(b): Heat map illustrating up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type A2780 ovarian cancer cell line. GO terms on the right are mapped according to p-value in which p value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

4.4.4 Colorectal

HT-29 BRAF V600E Mutated Holoclone Overrepresented GO Terms

A total of overrepresented 1,457 GO terms in holoclones derived from BRAF V600E mutated HT-29 colorectal cells were identified. This total was then broken down into 1,243 uniquely up regulated and 214 uniquely down regulated GO terms in holoclone populations. This pattern was in contrast to most other holoclone populations in which there was a more even spread between over and underexpressed GO terms however the majority of differentially represented GO terms in HT-29 holoclones were overrepresented.

As expected based on results observed in other cancer types in this study, of the 1,243 uniquely up regulated GO terms in BRAF V600E mutated HT-29 holoclones, many were involved in stem cell biological processes. However many differed in the number of genes overexpressed in each GO term category. For example, 27 genes involved in “stem cell development” with a fold change of 6 (p-value <0.01) were recorded in HT-29 holoclones. Similarly, 24 genes involved in “stem cell maintenance” were overexpressed by 6-fold (p-value <0.01) in HT-29 holoclone populations. The “regulation of stem cell differentiation” GO term was also overexpressed by 6-fold (p-value <0.01).

Fold changes (up to 184-fold) within underexpressed GO terms in HT-29 holoclones were much higher than those observed in other holoclone types indicating a significant regulatory effect of miRNAs on aberrant biological processes. Of the 214 uniquely down regulated GO terms identified in HT-29 holoclones, one of the most interesting observations was underexpression of “mesenchymal cell apoptosis”. This GO term was underexpressed by 123-fold (p-value <0.01) suggesting that if HT-29 holoclones possess true EMT characteristics, they may have dysfunctional apoptotic mechanisms allowing the persistence of CSC progenies with increased EMT potential.

In overexpressed GO terms in holoclones derived from other cancer types, overexpressed gene numbers were 6 or below. However the results recorded in HT-29 holoclones suggest that differentially expressed miRNAs have a dynamic and wide ranging effect on biological processes involved in stem cell populations.

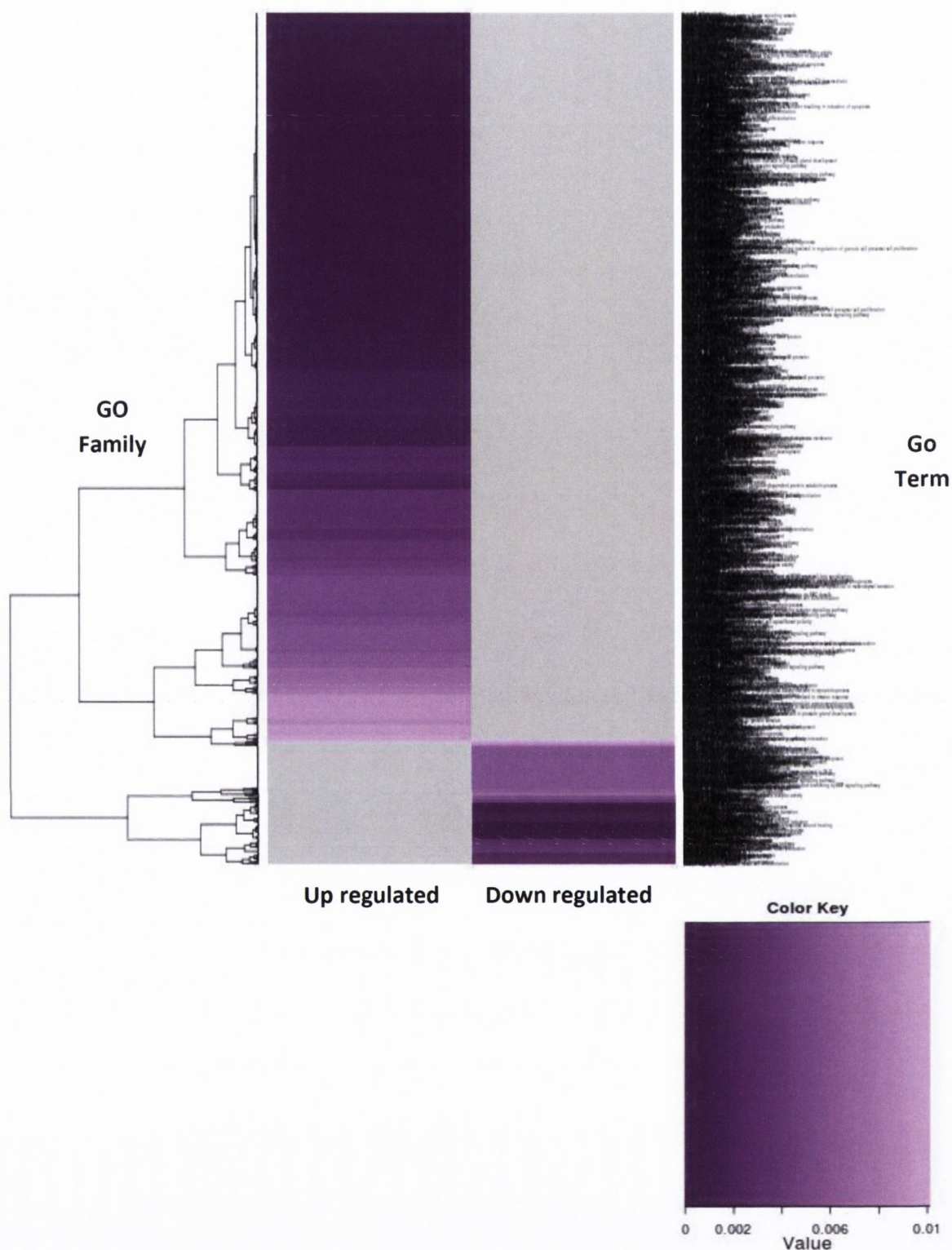


Figure 4.4.4(a): Heat map illustrating up regulated and down regulated GO terms in holoclones derived from BRAF V600E mutated HT-29 colorectal cancer cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped by biological function family (e.g. cell signalling or DNA repair)

COLO320 BRAF Wild Type Holoclone Overrepresented GO Terms

A total of 1,020 overrepresented GO terms in holoclones derived from BRAF wild type COLO320 colorectal cells were identified. This total was then broken down into 417 uniquely up regulated and 603 uniquely down regulated GO terms in holoclone populations.

Of the 417 uniquely up regulated GO terms identified in COLO320 holoclones, “regulation of mesenchymal stem cell differentiation” was the most overexpressed stem GO term. An 8-fold overexpression (p-value<0.01) of this biological process was recorded in COLO320 holoclone populations. The process of “stem cell division” was overexpressed by 4-fold (p-value <0.01). These observations were in contrast to results seen in HT-29 GO terms in which several stem cell biological processes were overexpressed in holoclone populations.

Of the 603 uniquely down regulated GO terms identified in COLO320 holoclones, “stem cell fate commitment” which was underexpressed by 13-fold (p-value <0.01) and “regulation of stem cell maintenance” which was underexpressed by 7-fold (p-value <0.01), were of particular relevance. This pattern of underexpressed GO terms is in contrast to HT-29 holoclones in which both biological processes were overexpressed in holoclone populations. However, in correlation with HT-29 holoclones, genes involved in the “mesenchymal cell apoptosis” GO term were underexpressed by 12-fold (p-value <0.01). Of particular relevance was a similar 12-fold (p-value <0.01) underexpression of genes involved in “regulation of mesenchymal cell migration”. Taken together these results suggest that holoclones derived from COLO320 colorectal cells may have mesenchymal properties with defective controlled cell death mechanisms.

The differentially represented GO term results recorded in COLO320 holoclones have both comparisons and contrasts to their counterparts. Opposite expression profiles of stem cell fate and maintenance processes were observed between CSC progenies derived from BRAF V600E mutant and BRAF wild type colorectal adenocarcinomas. However similar observations were made in mesenchymal GO terms suggesting that holoclones from both cell lines share certain aberrant biological processes as a result of differentially expressed miRNAs

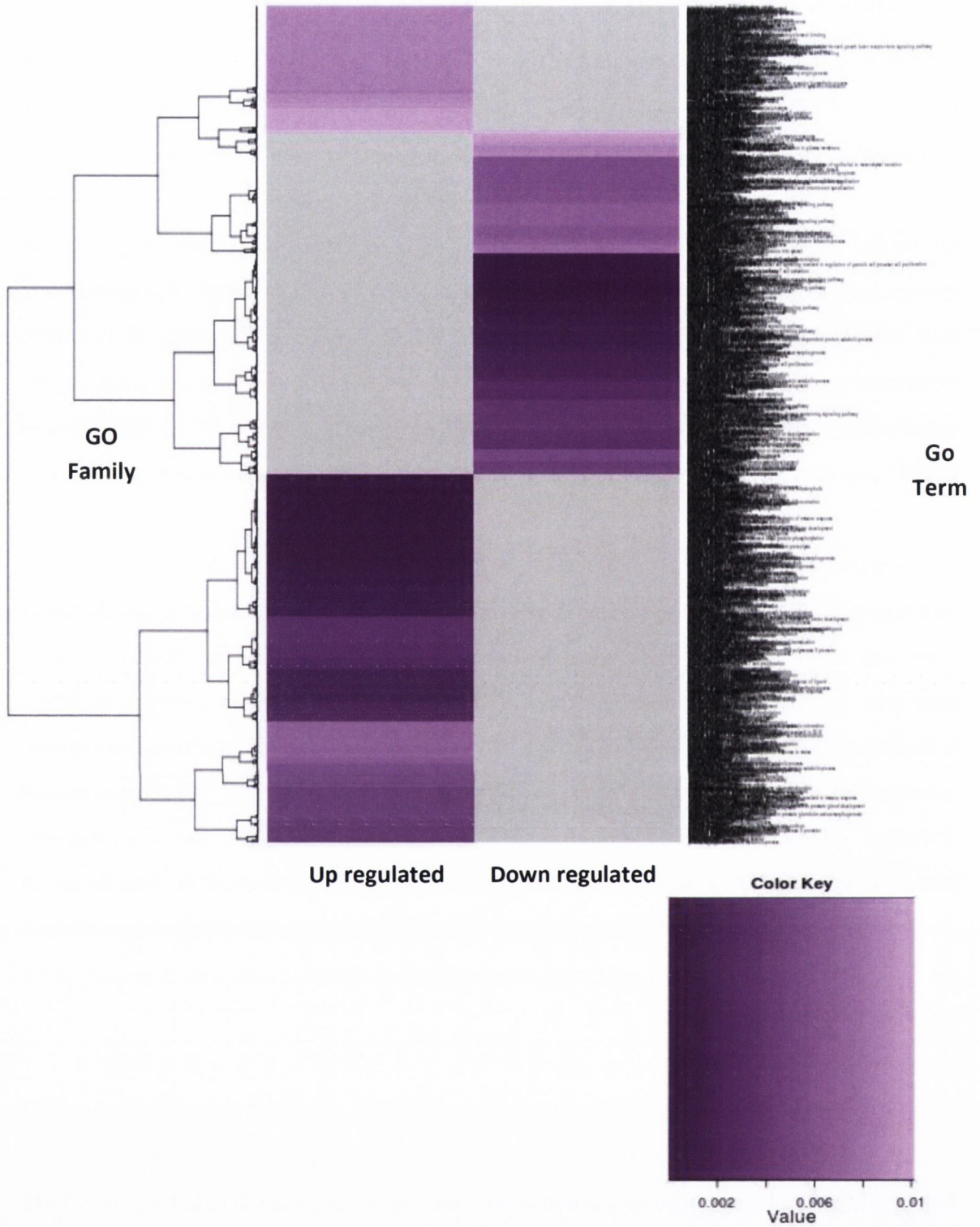


Figure 4.4.4(b): Heat map illustrating up regulated and down regulated GO terms in holoclones derived from BRAF Wild Type COLO320 colorectal cancer cell line. GO terms on the right are mapped according to p-value in which p-value strength is illustrated by the colour key and grouped according to biological function family (e.g. cell signalling or DNA repair)

4.4.5 Discussion

In this study we identified differential gene ontology term patterns as result of dysregulated miRNA expression between holoclones derived from BRAF V600E mutated and BRAF wild type tumours. Many of the overrepresented GO terms involving stem cell biological processes such as proliferation, division and regulation of differentiation were recorded in both BRAF V600E mutant and wild type holoclone populations. Gene expression analysis results (Section 3.4) demonstrated that subpopulations derived from a range of cancer cell lines express stem cell markers such as NANOG and Oct4. The identification of overexpressed stem cell processes such as loss of proliferation control and stem cell maintenance suggest that holoclones derived from these cell lines have strong stem identities and have aberrant biological processes that confer them such characteristics.

Melanoma

In BRAF V600E mutated SK-Mel 28 holoclones, we identified overexpressed EMT processes and loss of stem cell proliferation control which may play a role in these malignant melanoma CSC progenies possessing a greater stem and migratory potential than their founder cells. Similarly in BRAF wild type COLO794 holoclones, we recorded aberrantly expressed EMT processes including involvement of the canonical Wnt receptor signalling pathway. These findings suggest that despite showing differentially expressed miRNAs, similar processes are effected regardless of BRAF status in melanoma CSC progenies.

Thyroid

In BRAF V600E mutated 8505C holoclones and BRAF wild type TPC-1 holoclones, overexpressed stem cell proliferation and maintenance GO terms were recorded. We identified evidence of self-renewal in 8505C and TPC-1 holoclones in a high salt agar assay setting along with differentially expressed miRNAs implicated in stem cells and cancer stem cells alike. Taken together these findings suggest that CSC progenies derived from both BRAF V600E mutated and BRAF wild type thyroid carcinomas have enhanced proliferative capabilities and exhibit highly dysregulated biological processes that contribute to their stem identity.

Ovarian

In BRAF V600E mutated ES-2 holoclones, overrepresented genes involved in loss of stem cell differentiation regulation and enhanced stem cell proliferation were recorded in overexpressed GO terms. An observation exclusive to results recorded in ES-2 holoclones was an overexpression in the regulation of timing of cell differentiation. This finding suggests that in CSC progenies derived from a BRAF V600E mutated ovarian cancer, miRNAs may play a role indicating a possible differentiation mechanisms thus keeping ES-2 holoclones in a stem-like state. The GO term results recorded in A2780 holoclones were exclusive in that processes involved in stem cell fate were overexpressed. The unique observations made in these holoclones suggest that specific miRNAs may be working in concert to exert regulatory effects on these biological processes. Comprehensive downward fold changes were recorded in underexpressed GO terms in A2780 holoclones including loss of stem cell differentiation regulation. Dysregulated miRNAs in holoclones derived from a BRAF wild type ovarian carcinoma may contribute to pluripotency and an undifferentiated state in these CSC subpopulations.

Colorectal

Overexpressed GO terms recorded in BRAF V600E mutated HT-29 holoclones were of particular interest in that, comprehensive gene numbers within the same GO category appeared to be regulated by differentially expressed miRNAs. In some GO terms such a stem cell development, up to 27 genes involved in the same biological process were overexpressed, illustrating a substantial increase over holoclones derived from other cancer types. This observation suggests that dysregulated miRNAs have a greater influence on biological functions in stem cells derived from a BRAF V600E mutated colorectal cancer.

As a result of c-myc overexpression, COLO320 cells have been shown to have strong colony forming abilities (Collins et al., 1993). Enhanced holoclone formation in a high salt agar assay would appear to correlate with these findings and perhaps a loss in stem cell proliferation control would be expected in GO term analysis. We recorded an overexpression of stem cell division in COLO320 holoclones and an up regulation in stem cell maintenance suggesting that CSC progenies derived from a BRAF wild type colorectal cancer may have a more proliferative capability than their V600E mutated counterparts.

4.5 Differential snoRNA expression in holoclone populations

Sequencing of small RNAs via the SOLiD™ 4 next generation sequencing platform generated data detailing differentially expressed snoRNAs in holoclone samples. Global snoRNA expression was analysed among parent and holoclone samples for SK-Mel 28, COLO794, 8505C, TPC-1, ES-2, A2780, HT-29 and COLO320 cell lines. Using the same methods outlined in Section 4.2, annotated reads were normalised to allow analysis of differential expression across samples (Li et al, 2011), using the same formula:

$$\text{Coverage target snoRNA} = \frac{\text{count snoRNA target}}{\text{Total count of snoRNAs in sample}}$$

Differentially expressed snoRNAs, based on normalised sequence number variations, were identified in holoclone samples and the relative fold change of holoclone samples was calculated as follows:

$$\text{Fold Change} = \frac{\text{Coverage snoRNA Holoclone}}{\text{Coverage snoRNA Parent}}$$

Reads were mapped back to different chromosomes on the human genome and parent reads were bioinformatically compared to holoclone reads to identify differentially expressed snoRNAs. For each differentially expressed miRNA, a comprehensive array of data was generated including:

- Gene Symbol
- Parental and Holoclone Counts
- Fold Change (Log_2)
- Chromosomal Location
- Biotype (snoRNA)

We sought to identify differentially expressed snoRNAs that play in holoclones derived from cell lines with and without BRAF V600E mutation. At present, very little is known about snoRNAs in cancer stem cell populations and so by interrogating the snoRNA profiles of

holoclonal clones derived from different cancer types, we may highlight potential markers for future investigation in cancer stem cells.

4.5.1 Melanoma

SK-Mel 28 BRAF V600E Mutant

A total of 64 differentially expressed snoRNAs (Appendix III, Table A.3.2.1(a) and (b)) was recorded between SK-Mel 28 parent and holoclone samples with 8 up regulated snoRNAs being observed and 56 down regulated snoRNAs.

In SK-Mel 28 holoclonal clones we identified C/D box snoRNAs that were exclusively expressed. For one of these, SNORD114-3 located on chromosome 14, we recorded 27 sequencing reads in SK-Mel 28 holoclonal clones. We also identified another snoRNA, SNORD113-9 located at the same genomic region on chromosome 14 that was exclusively expressed in SK-Mel 28 holoclonal clones.

snoRNA	Chromosome	Parent Count	Holoclonal Count	Fold Change (Log ₂)
SNORD114-3	14	0	27.201	inf
SNORD113-9	14	0	25.201	inf

Table 4.5.1(a): Fold change of SNORD114 and SNORD113 in SK-Mel 28 holoclonal clones

COLO794 BRAF Wild Type

A total of 50 differentially expressed snoRNAs (Appendix III, Tables A.3.2.1(c) and (d)) was recorded between COLO794 parent and holoclone samples with 27 up regulated snoRNAs being observed and 23 down regulated snoRNAs.

We identified 2 C/D box snoRNAs that were overexpressed in COLO794 holoclonal clones and which also showed up regulation in BRAF wild type holoclonal clones derived from thyroid, ovarian and colorectal cancer cell lines. Expression of SNORD85, located on chromosome 1,

increased more than 2-fold in COLO794 holoclones and SNORD71 located on chromosome 16, was up regulated almost 2-fold.

snoRNA	Chromosome	Parent Count	Holoclonal Count	Fold Change (Log ₂)
SNORD85	1	5.326	26.524	2.316
SNORD71	16	13.456	47.011	1.805

Table 4.5.1(b): Fold change of SNORD85 and SNORD71 in COLO794 holoclones

4.5.2 Thyroid

8505C BRAF V600E Mutant

A total of 76 differentially expressed snoRNAs (Appendix III, Tables A.3.2.1(c) and (d)) was recorded between 8505C parent and holoclone samples with 75 up regulated snoRNAs being observed and 1 down regulated snoRNA.

Of the 75 up regulated snoRNAs identified in 8505C holoclones, many of these were C/D box snoRNAs located on chromosome 14. Similar to SK-Mel 28 holoclones we identified SNORD113 transcripts, also on chromosome 14 indicating a potential hotspot for snoRNA dysregulation at this genomic location. These transcripts showed up regulation of greater than 10-fold and 3 of these showed exclusive expression in holoclone populations with more than 400 reads being recorded for each.

snoRNA	Chromosome	Parent Count	Holoclonal Count	Fold Change (Log ₂)
SNORD114-14	14	0	586	inf
SNORD113-6	14	0	430	inf
SNORD113-7	14	0	820	inf
SNORD114-22	14	0.087	625	12.81
SNORD114-25	14	0.174	1133	12.67
SNORD113-9	14	0.609	2227	11.84
SNORD114-1	14	0.348	1094	11.62
SNORD114-3	14	0.696	1875	11.40
SNORD114-26	14	0.174	430	11.27
SNORD114-12	14	0.696	1406	10.98

SNORD114-9	14	0.435	859	10.95
------------	----	-------	-----	-------

Table 4.5.2(a): Fold changes of SNORD114 transcripts in 8505C holoclones

TPC-1 BRAF V600E Wild Type

A total of 57 differentially expressed snoRNAs (Appendix III, Tables A.3.2.2(c) and (d)) was recorded between TPC-1 parent and holoclone samples with 55 up regulated snoRNAs being observed and 2 down regulated snoRNAs. SNORD85 was up regulated almost 4-fold in TPC-1 holoclones showing similar expression patterns to other BRAF wild type holoclones.

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log ₂)
SNORD85	1	7.582	106.266	3.809

Table 4.5.2(b): Fold change of SNORD85 in TPC-1 holoclones

4.5.3 Ovarian

ES-2 BRAF V600E Mutant

A total of 49 differentially expressed snoRNAs (Appendix III, Tables A.3.2.3(a) and (b)) was recorded between ES-2 parent and holoclone samples with 43 up regulated snoRNAs being observed and 6 down regulated snoRNAs.

The majority of differentially expressed snoRNAs in ES-2 holoclones were overexpressed and were of the C/D box snoRNA family. 3 snoRNAs were identified that were uniquely expressed in holoclone populations and one of these was a SNORD114 transcript, similar to observations made in BRAF V600E mutated SK-Mel 28 and 8505C holoclones. 28 reads were identified in ES-2 holoclones for SNORD114-13 located on the same chromosome as other transcripts identified previously in this study.

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log ₂)
SNORD114-13	14	0	28.083	inf

Table 4.5.3(a): Fold changes of a SNORD114 transcript in ES-2 holoclones

A2780 BRAF V600E Wild Type

A total of 49 differentially expressed snoRNAs (Appendix III, Tables A.3.2.3(c) and (d)) was recorded between A2780 parent and holoclone samples with 45 up regulated snoRNAs being observed and 4 down regulated snoRNAs.

We identified overexpression of SNORD85 and SNORD71 in A2780 holoclone populations. SNORD85 was the most overexpressed snoRNA in A2780 holoclones and we recorded a 6-fold increase in its expression over parental cells. SNORD71 was up regulated by more than 3-fold in A2780 holoclones.

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log₂)
SNORD85	1	0.757	50.205	6.052
SNORD71	16	3.027	32.728	3.435

Table 4.5.3(b): Fold change of SNORD85 and SNORD71 in A2780 holoclones

4.5.4 Colorectal

HT-29 BRAF V600E Mutant

A total of 27 differentially expressed snoRNAs (Appendix III, Tables A.3.2.4(a) and (b)) was recorded between HT-29 parent and holoclone samples with 27 up regulated snoRNAs being observed and no down regulated snoRNAs.

Despite all of the differentially expressed snoRNAs identified in HT-29 holoclones being up regulated up we did not record an increase in expression of any SNORD114 transcripts. It is possible that expression of SNORD114 is similar to that of parent cells and was not identified as differentially expressed through bioinformatical tools. We did record several novel overexpressed C/D box family members however, as of yet their function is not yet known.

COLO320 BRAF V600E Wild Type

A total of 41 differentially expressed snoRNAs (Appendix III, Tables A.3.2.4(c) and (d)) was recorded between COLO320 parent and holoclone samples with 40 up regulated snoRNAs being observed and only 1 down regulated snoRNA.

Similar expression levels of SNORD85 and SNORD71 were recorded in COLO320 holoclones. SNORD85 expression increased by more than 2.5-fold and SNORD71 was up regulated by more than 2.5-fold in COLO320 holoclone populations.

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change (Log2)
SNORD85	1	3.139	19.144	2.608
SNORD71	16	6.153	36.373	2.563

Table 4.5.4(a): Fold change of SNORD85 and SNORD71 in COLO320 holoclones

4.5.5 Discussion

Through deep sequencing of non-coding RNA species using SOLiD™ next generation sequencing, differentially expressed snoRNAs were identified in holoclones derived from BRAF wild type and BRAF V600E mutated cancer types. Many of these snoRNAs and their expression patterns have been highlighted here for the first time in cancer stem cell subpopulations. We have identified expression of differential snoRNAs that are consistent across BRAF V600E mutated and also BRAF wild type cancer stem cell progenies. Their conservation of expression across holoclones derived from different cancer types suggests they may play an important role in the stemness of these subpopulations in association with BRAF status.

Small nucleolar RNA SNORD113 14q(I) is a C/D box snoRNA and may play a role in epigenetic imprinting (Cavaille et al., 2002). At the 14q32 genomic domain there are two clusters of tandemly repeated snoRNAs named 14q(I) and 14q(II). These clusters contain 9 and 31 highly related snoRNAs respectively, known as SNORD113 and SNORD114 respectively. The snoRNAs found in each cluster are related and are simply referred to with a 1-9 or 1-31 suffix. SNORD114 belongs to the C/D box class of snoRNAs which contain the

the human orthologue of the mouse MBII-239 and is predicted to guide 2'O-ribosemethylation of 5.8S rRNA on residue U1 (Hüttenhofer et al., 2001). SNORD71 has previously been identified as a sno-microRNA through Blast analysis in the same study (Hüttenhofer et al., 2001). The study identified 100% alignment between SNORD71 and the human miRNA miR-768 and the results showed that mature miR-768 was present in the SNORD71 gene sequence.

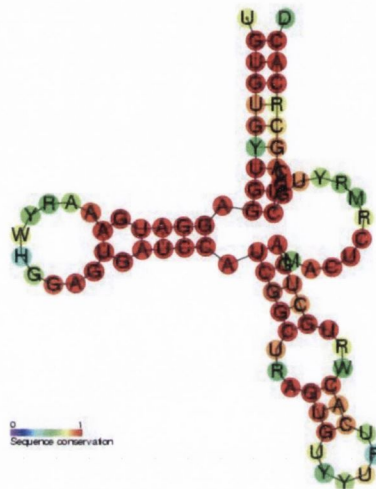


Figure 4.5.5(b): Predicted secondary structure of SNORD71 (Rfam, Gardener et al., 2010)

Despite the identification of SNORD71 and miR-768 alignment, we did not record an overexpression of miR-768 in holoclones derived from BRAF wild type holoclones. However, its consistent up regulation across cancer types suggests it may play a role in cancer stem cells which do not carry the V600E mutation. It is possible that defective miR-768 processing inhibits its expression as a mature miRNA sequence in holoclone and parent populations.

SNORD85 also showed consistent up regulation across the BRAF wild type holoclone panel and this observation represents a novel finding in cancer biology. Using the snoRNABase bioinformatics tool (Lestrade & Weber, 2006), PUM1 (pumilio homolog 1 (Drosophila) was identified as the host gene of SNORD85, and functions as an RNA-binding protein that regulates translation and mRNA stability by binding the 3'UTR of mRNAs (Van Etten et al.,

2012). PUM1 is one of two PUF RNA-binding proteins present in humans that regulate specific mRNAs and is thought to be required to support proliferation and self-renewal of stem cells (Spasov & Jurecic, 2003).

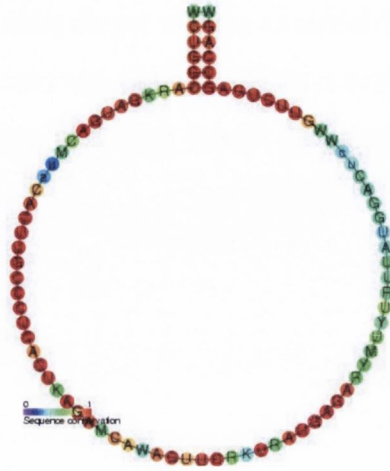


Figure 4.5.5(c): Predicted secondary structure of SNOORD85 (Rfam, Gardener et al., 2010)

Our findings of increased SNORD85 expression within its host gene PUM1, suggest that it may play a role in the maintaining proliferation of cancer stem cell progenies derived from melanoma, thyroid, ovarian and colorectal cancers that lack the BRAF V600E mutation. Future experiments investigating the expression of PUM1 in BRAF wild type holoclones may highlight a potential association between PUM1 and SNORD85 in cancer stem cell progenies derived from different tumour types lacking BRAF V600E mutation.

Through global snoRNA expression in holoclones derived from BRAF wild type and BRAF V600E mutated holoclones, we have made novel observations of differentially expressed snoRNAs. In BRAF V600E mutated cancer stem cell progenies we identified consistent overexpression of SNORD114 across 4 cancer types and following on from previous work by Valleron et al., 2012, we propose that up regulation of this snoRNA is important for proliferation within BRAF mutated holoclones.

In BRAF wild type holoclones we identified 2 snoRNAs that were consistently expressed across all 4 cancer types which represent novel findings in the field of cancer stem cell biology. Of particular interest is the observation of overexpressed SNORD85, located within

the PUM1 gene which is believed to play a role in stem cell self-renewal. These findings suggest that SNORD85 and PUM1 expression may be required for proliferation in BRAF wild type cancer stem cells.

These novel findings represent attractive targets for future work investigating the molecular biology of holoclones with and without BRAF mutation. Disruption of these non-coding RNA through sequence specific targeting may reveal their importance in BRAF wild type and BRAF V600E mutated cancer stem cell progenies.

Chapter 5

Methylation of MLH1 and RAR β 2 in a Cell Line Panel +/- BRAF V600E Mutation

5.1 Introduction

Through global snoRNA analysis of holoclones derived from a cell line panel with and without BRAF mutation, we identified several overexpressed C/D box snoRNAs. At present, little is known about the role of snoRNAs in cancer but research has established that C/D box snoRNAs play a role in methylation (Maden & Hughes, 1997). Having identified C/D box snoRNA overexpression in holoclones with and without BRAF mutation, we sought to interrogate the methylation profiles of cancer related genes. Our aim was to investigate if BRAF V600E mutation is associated with hypermethylation in the promoter regions of these genes in cancer cell lines derived from melanoma, thyroid, ovarian and colorectal tumours. Previous work by our group investigated the methylation profile of genes involved in the Wnt signalling pathway and other cancer related genes in cancer cell lines with and without BRAF V600E mutation. The hypothesis of the study was that BRAF V600E mutation plays a role in epigenetic modifications of these genes in a thyroid cancer cell line model. The study interrogated methylation profiles of CpG islands within the promoter region of these genes.

Gene	TPC-1		Kat10	
	CpGs Analysed	CpGs Methylated	CpGs Analysed	CpGs Methylated
APC	32	6	13	0
CD44	21	0	7	0
MLH1	92	5	23	23
E-Cadherin	16	15	25	2
p15	43	4	16	0
GSTP	7	0	0	0
RAR β 2	30	28	10	10

Table 5.1(a): Methylation results of gene panel in BRAF wild type and BRAF V600E mutant thyroid cancer cell lines. ‘CpGs Analysed’ is the total number of CpGs that were able to be called and ‘CpGs Methylated’ are the number of CpGs that showed methylation.

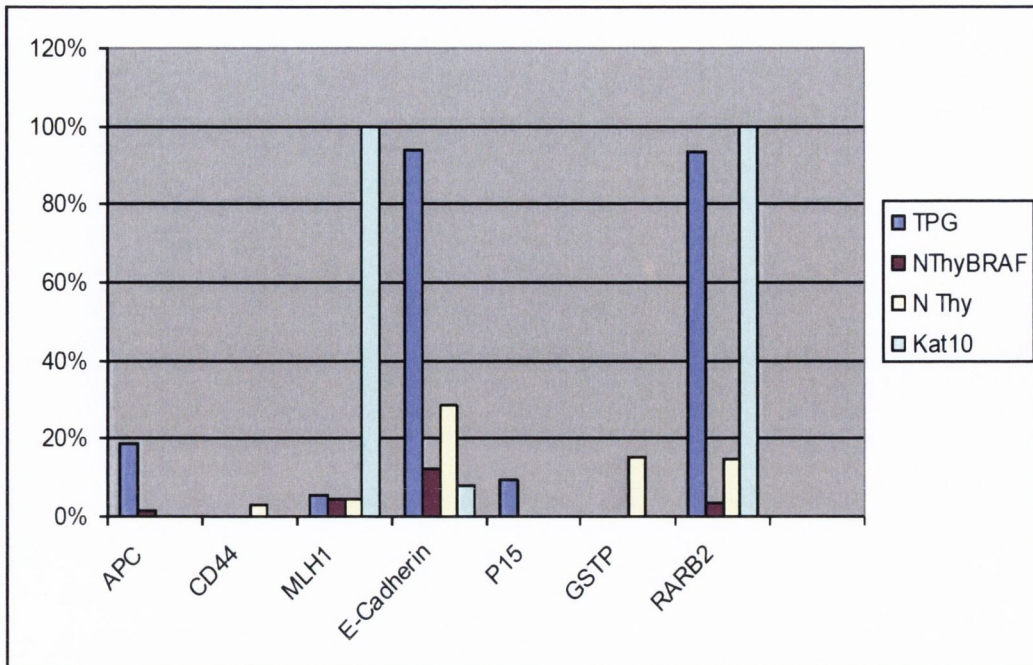


Figure 5.1(b): Pilot study illustrating the percentage of methylation in cancer related genes across a panel of cell lines. The degree of methylation shown by each sample is shown for each gene tested and the methylation percentage is expressed as the number of CpGs showing any methylation as a function of the total CpGs analysed.

The study identified methylation in 23 out of 23 CpG islands in MLH1 and 10 out of 10 CpG islands in RAR β 2 in the thyroid cancer cell line Kat10 (highlighted in light green) compared to wild type BRAF thyroid cancer cell lines TPC-1 and N-Thy in which less than 20% of CpG islands were methylated. The Kat10 thyroid cancer cell line harbours a BRAF V600E mutation suggesting that BRAF mutation plays a role in epigenetic silencing of genes involved in carcinogenesis. To expand this study we interrogated the methylation profiles of MLH1 and RAR β 2 across a panel of cell lines derived from different tumour types with and without BRAF V600E mutation.

By examining the methylation profiles of MLH1 and RAR β 2 IN BRAF V600E and BRAF wild type mutated cell lines derived from melanomas, thyroid, ovarian and colorectal cancers; we seek to establish if BRAF V600E mutation facilitates methylation of cancer related genes in different cancer types. We hypothesise that holoclones derived from these cell lines inherit these epigenetic modification which contribute to the genomic instability of cancer

stem cell progenies. The epigenetic modification of genes such as MLH1 and RAR β 2 in holoclone populations may contribute towards chemoresistant and tumourigenic characteristics features of a true cancer stem cell as outlined by Rosen & Jordan, 2006.

Aims & Objectives

- Interrogate methylation status of the mismatch repair gene MLH1 in a panel of cancer cell lines +/- BRAF V600E mutation
- Interrogate methylation status of the Retinoic acid receptor B2 gene RAR β 2 in the same cell line panel
- Investigate if increased methylation of MLH1 and RAR β 2 is associated with BRAF V600E mutation

5.2 PCR of Bisulfite Converted DNA from Cell Line Panel

Primers for β -Actin were used as a control to establish if bisulfite conversion of genomic DNA isolated from the cell line panel was successful. These primers were used to amplify converted DNA only and so the appearance of bands in the resulting PCR gel indicated that bisulfite conversion was successful.

Sample Legend:

Lad – Ladder

SK – SK-Mel 28 (BRAF V600E Mutant)

C7 – COLO794 (BRAF Wild Type)

85 – 8505C (BRAF V600E Mutant)

TP – TPC-1 (BRAF Wild Type)

ES – ES-2 (BRAF V600E Mutant)

A2 – A2780 (BRAF Wild Type)

HT- HT-29 (BRAF V600E Mutant)

C3 – COLO320 (BRAF Wild Type)

Lad – Ladder

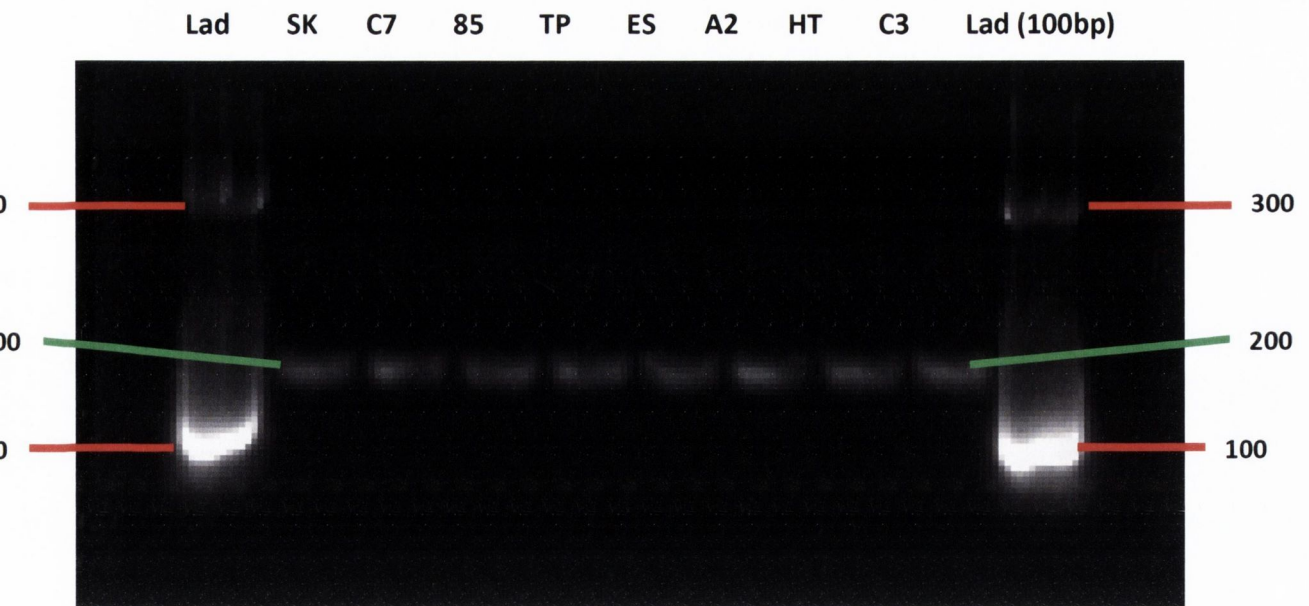


Figure 5.2(a): 2% agarose gel of successfully amplified control β -Actin in bisulfite converted DNA isolated from cancer cell line panel. Expected PCR product size of 200bp was observed for each sample.

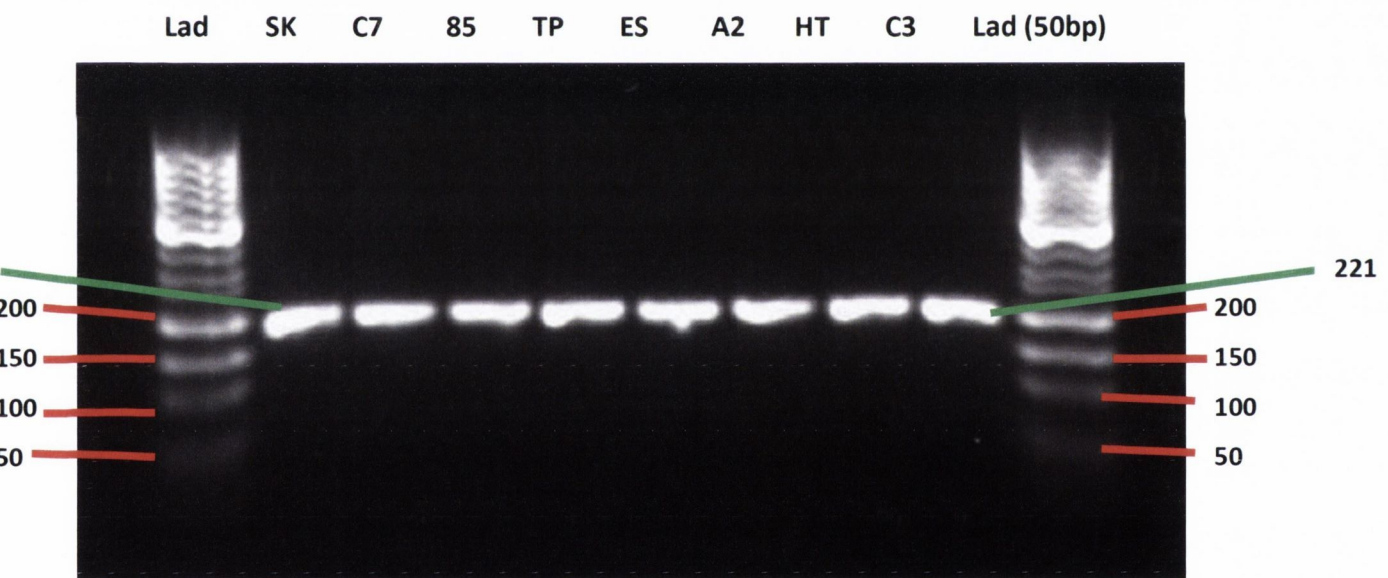


Figure 5.2(b): 2% agarose gel of successfully amplified MLH1 in bisulfite converted DNA isolated from cancer cell line panel. Expected PCR product size of 221bp was observed for each sample.

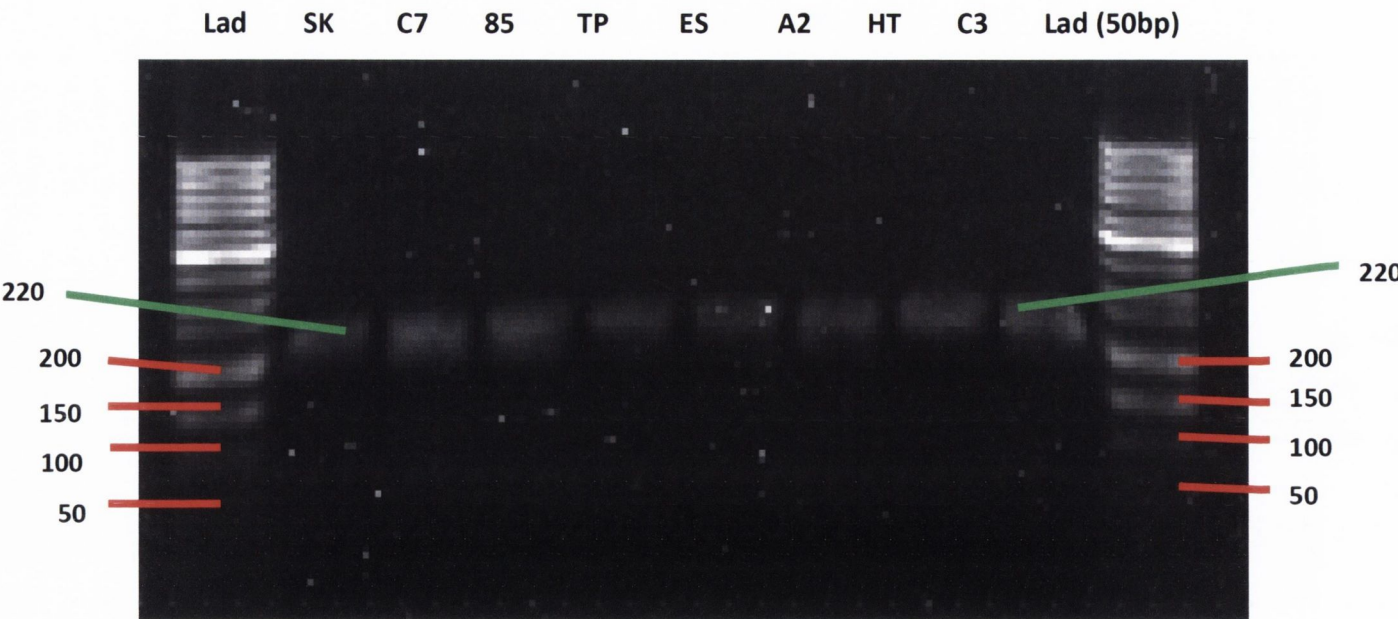


Figure 5.2(c): 2% agarose gel of successfully amplified *RARβ2* in bisulfite converted DNA isolated from cancer cell line panel. Expected PCR product size of 220bp was observed for each sample.

5.3 Methylation Sequencing Results

After PCR, the sequence of the bisulfite-converted DNA had C residues only if the C residues were already methylated. All other Cs were detected as thymines (T). Thus, an unmethylated gDNA sample produced no Cs in the sequencing data. Input reads obtained from methylation sequencing results (Appendix CD), were graphically aligned to a reference sequence using CLC Sequence Viewer.

5.3.1 MLH1

We identified 67 cytosine residues in the reference MLH1 sequence and the following results illustrate the number of cytosine residues (and percentage) in the resulting sequence, indicating a methylated residue:

SK-Mel 28 -	2 cytosine residues methylated (3%)
COLO794 -	0 cytosine residues methylated (0%)
8505C -	0 cytosine residues methylated (0%)
TPC-1 -	8 cytosine residues methylated (12%)

ES-2 -	1 cytosine residue methylated (1%)
A2780 -	0 cytosine residues methylated (0%)
HT-29 -	3 cytosine residues methylated (5%)
COLO320 -	2 cytosine residues methylated (3%)

5.3.2 RAR β 2

We identified 45 cytosine residues in the reference RAR β 2 sequence and the following results illustrate the number of cytosine residues (and percentage) in the resulting sequence, indicating a methylated residue:

SK-Mel 28 -	21 cytosine residues methylated (47%)
COLO794 -	0 cytosine residues methylated (0%)
8505C -	21 cytosine residues methylated (47%)
TPC-1 -	21 cytosine residues methylated (47%)
ES-2 -	0 cytosine residue methylated (0%)
A2780 -	1 cytosine residues methylated (2%)
HT-29 -	0 cytosine residues methylated (0%)
COLO320 -	4 cytosine residues methylated (9%)

5.3.3 Discussion

Across the panel of cell lines with and without BRAF V600E mutation, we recorded a varying degree of methylation within the promoter region of the MLH1 gene. In 3 out of the 4 cancer types; melanoma, ovarian and colorectal, subtle increases in percentage of methylation were recorded indicating a trend towards greater methylation in BRAF V600E mutated cells:

- Melanoma – 3% in mutant v 0% in wild type
- Ovarian – 1% in mutant v 0% in wild type
- Colorectal – 5% in mutant v 3% in wild type

The exception was the thyroid cell lines in which a substantially higher percentage (12%) of methylation was observed in the BRAF wild type TPC-1 cell line over its BRAF V600E mutated 8505C counterpart in which no methylation was recorded. These findings represented a

small increase over results recorded for MLH1 methylation in TPC-1 in previous work, in which 5% of CpGs within the promoter region of MLH1 were methylated (Figure 5.1(a)). In previous work, high levels of methylation were observed in the BRAF V600E mutated Kat10 cell line however we recorded 0% MLH1 methylation the BRAF V600E mutated 8505C cell line used in this study. The 8505C cell line is derived from an anaplastic thyroid carcinoma and our results showing 0% methylation in the MLH1 gene suggests that MLH1 function may be required for the increased cell proliferation that is characteristic of anaplastic cancer. Strong expression of MLH1 has been previously been shown through immunohistochemistry in anaplastic thyroid carcinoma (Broaddus et al., 2004) and may suggest why the 8505C cell line exhibited no MLH1 methylation. However, in cell lines derived from melanoma, ovarian and colorectal cancers, BRAF V600E mutation is associated with higher levels of MLH1 methylation and may play a role in methylation within the promoter region of MLH1 across a range of cancer types.

No differential pattern of expression linking BRAF status was observed in RAR β 2. However, the pattern of methylation of RAR β 2 in melanoma and thyroid cancer cell lines is noteworthy. Methylated sequences in the melanoma cell line SK-Mel 28 and thyroid cancer cell lines 8505C were identical and had the same methylated cytosine residues. These results correlate with observations made in previous work in which the Kat10 BRAF V600E mutated thyroid cancer cell line exhibited hypermethylation of RAR β 2 MLH1. However we also recorded a substantial level of methylation in the thyroid TPC-1 cell line which had a similar methylation pattern to the SK-Mel 28 and 8505C cell lines. These findings suggest that mutations in components of the MAPK-ERK signalling pathway such as BRAF or RET-PTC in thyroid cancers and melanoma, facilitates methylation of the RAR β 2 gene.

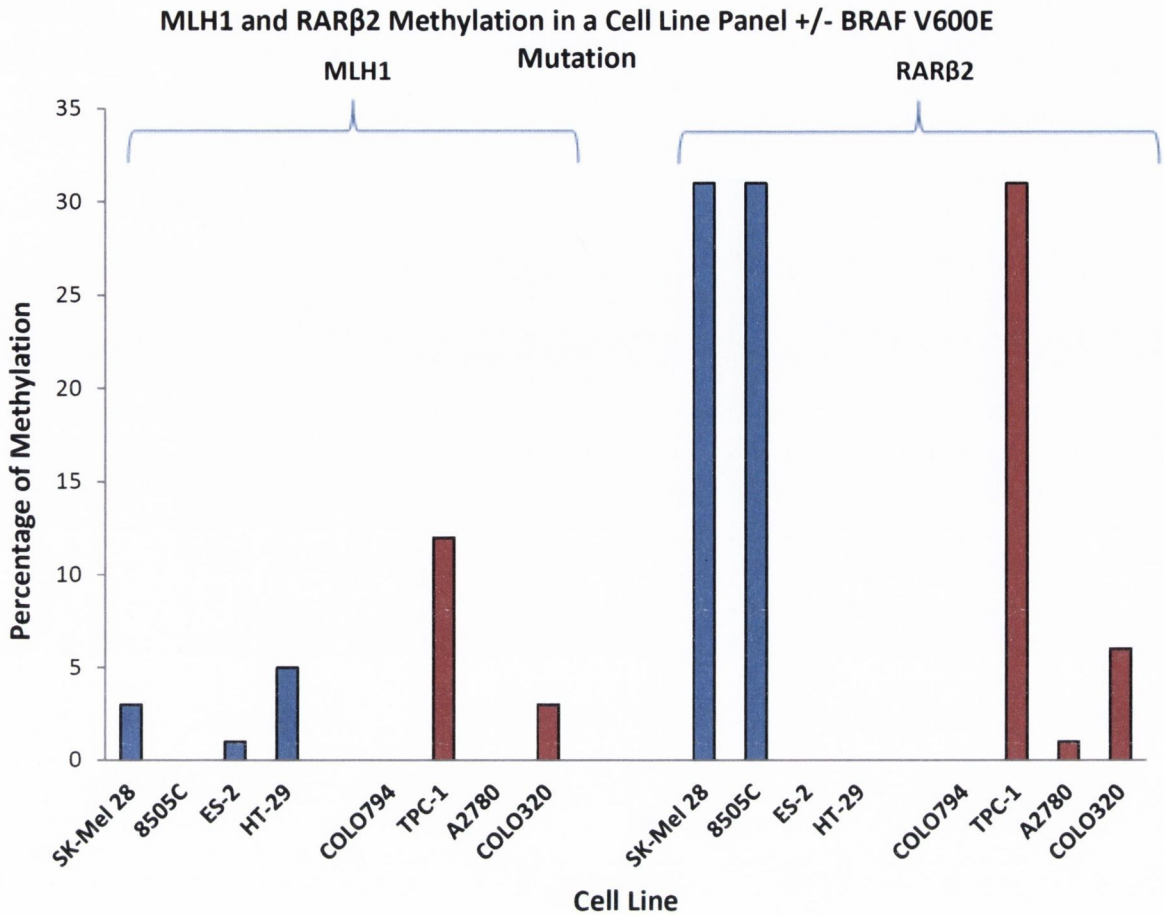


Figure 5.3.3(a): MLH1 and RAR β 2 methylation percentage in a cell line panel with (indicated in blue) and without (indicated in red) BRAF V600E mutation

The encoded RAR β 2 protein binds retinoic acid, and mediates cellular signalling in embryonic morphogenesis, cell growth and differentiation (Guan et al., 2001). It is thought that this protein limits growth of many cell types by regulating gene expression. However, silencing of the RAR β 2 gene as a result of hypermethylation may result in the loss of stem cell growth regulation and differentiation mechanisms. The highest levels of RAR β 2 methylation were observed in BRAF V600E mutated SK-Mel 28 holoclones and thyroid 8505C and TPC-1 holoclones. Both 8505C and TPC-1 holoclones exhibited overexpression of several C/D box snoRNAs. We investigated if any of these snoRNAs were consistently overexpressed across all 3 holoclone populations. Overexpression of only 5 C/D box snoRNAs were recorded in SK-Mel 28 holoclones however 3 of these snoRNAs were also uniquely expressed or up regulated in thyroid holoclones. SNORD85, SNORD20 and SNORD91A were up regulated in

SK-Mel 28, 8505C and TPC-1 holoclones suggesting a potential association between dysregulated MAPK-ERK signalling in the methylation of RAR β 2 and subsequent loss of stem cell growth control.

snoRNA	Chromosome	SK-Mel 28	8505C	TPC-1
SNORD85	1	Unique	4.0	3.8
SNORD20	2	Unique	1.6	2.3
SNORD91A	17	Unique	1.8	3.4

Figure 5.3.3(b): Log₂ fold changes of consistently overexpressed snoRNAs in SK-Mel 28, 8505C and TPC-1 holoclones

A clear differential methylation profile in RAR β 2 between BRAF V600E mutated and BRAF wild type cancer types was not recorded in this study. However a novel link was identified between commonly overexpressed C/D box snoRNAs and high levels of RAR β 2 promoter methylation. These findings suggest that hypermethylation of RAR β 2 in melanoma and thyroid cancer is associated with up regulation or unique expression of SNORD85, SNORD20 and SNORD91A in these cancer types. Of particular relevance is SNORD85 which has previously been shown to be located within a host gene PUM1 (pumilio homolog 1 (Drosophila), believed to be required to support proliferation and self-renewal of stem cells (Spassov & Jurecic, 2003). We propose that up regulation of SNORD85 in holoclones derived from cell lines exhibiting high levels of RAR β 2 methylation may play a role in the self-renewal of cancer stem cell progenies as evident from 8505C and TPC-1 holoclone repopulation in high salt agar assays.

By interrogating the methylation profile of MLH1 and RAR β 2 genes in both BRAF V600E mutated and BRAF wild type cell lines we have identified associations between BRAF mutation and MLH1 promoter methylation. Additionally, we have identified a novel biological link between methylation of cancer-related genes and snoRNA expression which may play a role in the stemness characteristics of holoclones derived from cancer types with

and without BRAF V600E mutation. These findings suggest BRAF mutation plays a pivotal role in many aspects of cancer biology including methylation and the acquisition of stem-like properties in cancer cell subpopulations.

Chapter 6

Discussion

The main emphasis of this thesis was BRAF status and its association with dysregulated non-coding RNAs in cancer stem cell subpopulations and also its possible role in the methylation of cancer-related genes. We have shown that holoclones can be successfully isolated from BRAF V600E mutated and BRAF wild type melanoma, thyroid, ovarian and colorectal cancer cell lines. Our primary objective was to study these cancer stem cells at a molecular level and identify dysregulated non-coding RNAs that may play a role in the stemness characteristics that holoclones possess.

The MAPK-ERK signalling pathway has been shown to be involved in up to 30% of cancers and BRAF mutation is believed to be implicated in approximately 8% of solid tumours (Cantwell-Dorris et al., 2011), and so components of this pathway represents an attractive target for molecular oncology research. By establishing a cell line panel of different tumour types, we represented the varying prevalence rates of BRAF V600E mutation which ranges from up to 70% in melanomas to 10-20% in colorectal cancer (Kumar et al., 2004; Yokota, 2012). The more recently formed hierarchical hypothesis suggests the bulk of cancer cells have limited regenerative capability, but a subpopulation of cancer stem cells can consistently proliferate and are tumorigenic (Sakashita et al., 2007). The aggressive nature of cancer stem cells and their resilience against conventional anti-cancer agents make them a potentially novel avenue in which cancer treatment can be aimed (Rosen & Jordan, 2009). Three cancer stem cell morphologies have been identified to date, and of these, holoclones have been shown to have the greatest self-renewal capability (Pfeiffer & Schalken, 2009) and using a high salt agar assay, positive selection of holoclones could be achieved. Once isolated from their parental cells, gene expression analysis could be employed to investigate if holoclones derived from BRAF V600E mutated holoclones possessed different expression profiles to their parent cells and to holoclones without BRAF mutation.

Gene expression analysis of holoclones and parent cell populations yielded noteworthy results in relation to stemness gene expression in both BRAF V600E mutated holoclones. Our study concluded that BRAF V600E mutation was associated with higher expression levels of stemness genes NANOG, Oct4 and ALDH1 which have been shown to be highly expressed by human ES cells (Chan et al., 2011). Consistent overexpression of these genes in

BRAF V600E mutated holoclones suggests that like ES cells, holoclones have enhanced pluripotency and self-renewal capabilities. Despite this conclusion, BRAF mutation was not associated with a clear enhanced EMT and cell signalling potential in V600E mutated holoclone populations derived from different cancer types. An inconsistent expression profile of EMT genes E-Cadherin and SNAI2, and cell signalling genes SHH, TGF- β and β -Catenin across a BRAF V600E mutated holoclone panel indicates that BRAF mutation may not have such widespread effects as those seen in the expression of stemness genes. Although expression of stemness genes may represent novel markers for identifying underlying subpopulations of cancer stem cells in the future, it is important that these changes in gene expression are reflected on the protein level. The most consistently up regulated stemness gene identified in our gene expression studies served as a candidate marker to take forward for protein expression analysis in holoclones and parent cells. Through immunofluorescence, we established a correlation between NANOG gene expression and protein expression in colorectal holoclones. Confocal microscopy identified increased expression of NANOG protein in BRAF V600E mutated HT-29 holoclones in which expression at the mRNA level was also observed. Conversely, NANOG protein expression was decreased in COLO320 BRAF wild type holoclones in which a significant decrease was observed at the mRNA level. In addition, NANOG protein localisation was observed to be similar to previous studies which identified that NANOG expression was restricted to the inner cell mass (ICM) in mouse and human blastocysts, and are used to characterize undifferentiated ES cells in vitro (He et al., 2006). Taken together, our gene and protein expression results suggest that BRAF V600E mutation has a significant impact on the expression of stem genes in holoclone populations that can be identified at both the message and protein level.

The SOLiD™ 4 next generation sequencing platform offered this project a unique opportunity to apply state-of-the-art technology to address fundamental questions of carcinogenesis. By carrying out deep sequencing of non-coding RNA species in holoclone populations we could get a greater insight into how BRAF V600E mutation affects the regulatory mechanisms that govern cancer stem cells subpopulations. With the current knowledge of miRNA involvement in human carcinogenesis by acting as tumour suppressors

or oncogenes, interrogating miRNA profiles in holoclones offers a novel approach to investigating the molecular biology of human cancer (Lynam-Lennon et al., 2009).

In BRAF V600E mutated holoclones we identified down regulation of miR-103 across BRAF V600E mutated holoclone populations. MirWalk validated target identification software identified several target genes of miR-103 within the MAPK-ERK signalling pathway suggesting that its underexpression may play a pivotal role in hyperactive signalling via this pathway in cancer stem cells harbouring the BRAF V600E mutation. The importance of the MAPK-ERK signalling pathway in cell growth, proliferation and apoptosis indicates that cancer stem cells exhibiting underexpressed miR-103 may have a proliferative advantage over their parental cells and may confer them enhanced tumorigenic potential (Brzezińska & Pastuszak-Lewandoska, 2011).

Similarly, we have identified highly dysregulated miRNAs common to BRAF wild type holoclones derived from different cancer types. Through global differential miRNA expression analysis, miR-26 family members were observed to be up regulated in melanoma, thyroid, ovarian and colorectal holoclones. As seen in miR-103, several gene targets of miR-26 family members were involved in MAPK-ERK signalling including kinases MAPK8 and transcription factors such as c-Jun. This observation represents a novel finding in cancer stem cell subpopulations and suggests that miR-26a plays a role in growth and proliferation of cancer stem cells by interacting with MAPK-ERK signalling components in the absence of BRAF V600E mutation.

Next generation sequencing has successfully identified consistently dysregulated miRNAs in both BRAF V600E mutated and BRAF wild type cancer stem cell populations across a range of tumour types. These results offer greater insights into the aberrant regulatory mechanisms associated with BRAF status in cancer stem cells. The ability to carry out global miRNA expression analysis on holoclone populations has enabled us to highlight the dysregulated biological processes that cancer stem cells possess. Gene ontology analysis has assisted in the identification of aberrant stem cell biological processes in holoclones with and without BRAF mutation. Identification of gene targets involved in processes such as stem cell proliferation control and differentiation regulation targeted by differentially expressed miRNAs. The results indicate that holoclones derived from both BRAF wild type

and BRAF V600E mutated cancer types, possess many hyperactive stem cell processes similar to those seen in human ES cells.

Recent studies have shown that snoRNAs are another class of non-coding RNAs that may play a role in human cancer (Mei et al., 2012; Askarian-Amiriet al., 2011). Despite what is known about the 2 classes of snoRNAs; C/D box snoRNAs and the H/ACA box snoRNAs, little is known about their role in cancer and their identification in cancer stem cell populations has not yet been described (Khanna & Stamm, 2010). Next generation sequencing offered this project the opportunity to identify differentially expressed snoRNAs in cancer stem cells for the first time. Our study identified SNORD114 as being consistently overexpressed in holoclones harbouring the BRAF V600E mutation and its role in cell growth promotion has previously been described (Valleron et al., 2012). Our findings propose that SNORD114 is involved in promoting the growth of cancer stem cells derived from BRAF V600E mutated cancer types and suggests that the mutation facilitates holoclone proliferation.

In holoclones derived from BRAF wild type cancer cell lines, we identified consistent overexpression of two C/D box snoRNAs SNORD85 and SNORD71. We propose that increased SNORD85 expression within its host gene PUM1 which has been shown to be involved in proliferation and self-renewal of stem cells (Spassov & Jurecic, 2003), plays a role in maintaining proliferation of cancer stem cell progenies derived from melanoma, thyroid, ovarian and colorectal cancers that lack the BRAF V600E mutation. Our identification of differentially expressed snoRNAs that are consistent across either BRAF wild type or BRAF V600E mutated holoclones represent novel observations in cancer stem cell subpopulations. Future work may yield a better understanding about their precise function in molecular biology and their role in tumourigenesis

Research into epigenetics has established that methylation plays a pivotal role in cancer development through genomic instability in tumour cells (Chen et al., 1998). Building on previous work we sought to address the biological question of what effect BRAF V600E mutation has in different cancer types. Increased methylation of the mismatch repair gene MLH1 was observed in BRAF V600E mutated cancer cell lines with the exception of thyroid

cell line 8505C. Previous work has shown that anaplastic thyroid cancers express MLH1 indicating that it may be required to assist in rapid cell turnover, a characteristic of anaplastic cancer (Broaddus et al., 2004). The results recorded for MLH1 methylation proposes that BRAF V600E mutation plays a role in increasing methylation within the promoter region of MLH1 across a range of cancer types.

Commonly overexpressed C/D box snoRNAs in holoclones derived from cell lines that exhibit high levels of RAR β 2 promoter methylation were identified through next generation sequencing. Up regulation of SNORD85, SNORD20 and SNORD91A was recorded across BRAF wild type SK-Mel 28, 8505C and TPC-1 holoclones. These results indicate that hypermethylation of RAR β 2 in melanoma and thyroid cancer is associated with up regulation or unique expression of SNORD85, SNORD20 and SNORD91A in the cancer stem cell subpopulations derived from these cancer types. A noteworthy observation was the up regulation of SNORD85 to which we propose that its overexpression in holoclones derived from cell lines exhibiting high levels of RAR β 2 methylation may play a role in the self-renewal of cancer stem cell progenies. Methylation analysis of cancer related genes across a cell line panel and snoRNA expression analysis in their derived holoclones has forged novel links between epigenetics and non-coding RNAs in cancer stem cells with and without BRAF mutation.

This study sought to understand the dysregulation of different cellular processes in tumourigenesis and to identify potential targets in the treatment of various human malignancies. BRAF mutation plays a pivotal role in many aspects of cancer biology including methylation and the acquisition of stem-like properties in cancer cell subpopulations. The work carried out in this thesis has identified areas for future directions which may advance our understanding of how cancer stem cells work at a molecular level and the role played by BRAF. There are several avenues which may be explored to investigate how specific genes and non-coding RNAs influence holoclone populations derived from both BRAF V600E mutated and BRAF wild type cancer types. Knockdown experiments of consistently up regulated stemness genes NANOG, Oct4 and ALDH1 in BRAF V600E holoclones may identify the necessity of these markers for holoclones to proliferate in high salt agar assays. Similar

to observations made in NANOG protein expression in this study, investigating the protein expression of other markers such as Oct4 in BRAF V600E mutated holoclones may establish if increases in stemness genes are consistent across mRNA and protein levels in holoclone populations.

Delivering holoclones into animal models is an important step in investigating if holoclones are truly tumorigenic and possess the ability to initiate tumour development. By interrogating the potential tumour forming capability of holoclones *in vivo*, a greater understanding of cancer stem cells may be achieved and may represent a novel capacity to which cancer treatments may be targeted in the future. The identification of aberrant non-coding RNA expression in holoclone populations illustrates the dysregulation of regulatory mechanisms in cancer stem cells. Functional analysis using pre-miRs and anti-miRs directed at differentially expressed miRNAs may elucidate the mechanisms critical for pluripotency and self-renewal in cancer stem cells. Future work is needed to gain a greater understanding of novel snoRNA analysis identified in cancer stem cells in this study. At present, researchers are only beginning to establish the roles of some of these non-coding RNAs in cancer. The results of this study highlight the need of a better understanding of how snoRNAs are involved in human carcinogenesis. We have shown that BRAF V600E mutation is associated with elevated methylation levels of cancer genes. Future directions may interrogate the methylation status of other cancer genes in different tumour types harbouring BRAF V600E mutation.

This thesis has added to the existing knowledge that BRAF mutation is pivotal in human cancer however it has also shed light on its involvement in cancer stem cell subpopulations that are believed to have resilience against conventional anti-cancer agents (Rosen & Jordan, 2009). Taken together, future work may explore the association between BRAF V600E mutation and its effects on cancer stem cells *in vivo* and facilitate the implementation of novel targeted strategies against tumours harbouring this specific mutation.

viii. Bibliography

- Adams M.S., Bronner-Fraser M. (2009). Review: the role of neural crest cells in the endocrine system. *Endocrine Pathology*, 20(2):92-100.
- Al-Sohaily S., Biankin A., Leong R., Kohonen-Corish M., Warusavitarne J. (2012). Molecular pathways in colorectal cancer. *Journal of Gastroenterology Hepatology*, 27(9):1423-31.
- Anders S., Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, 11:R106.
- Arozarena I., Bischof H., Gilby D., Belloni B., Dummer R., Wellbrock C. (2011). In melanoma, beta-catenin is a suppressor of invasion. *Oncogene*, 30(45):4531-43.
- Askarian-Amiri M.E., Crawford J., French J.D., Smart C.E., Smith M.A., Clark M.B., Ru K., Mercer T.R., Thompson E.R., Lakhani S.R., Vargas A.C., Campbell I.G., Brown M.A., Dinger M.E., Mattick J.S. (2011). SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA*, 17(5):878-91.
- Baak J.P., Wisse-Brekelmans E.C., Langley F.A., Talerman A., Delemarre J.F. (1986). Morphometric data to FIGO stage and histological type and grade for prognosis of ovarian tumours. *Journal of Clinical Pathology*, 39(12):1340-6.
- Barroso-del Jesus A., Romero-López C., Lucena-Aguilar G., Melen G.J., Sanchez L., Ligeró G., Berzal-Herranz A., Menendez P. (2008). Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. *Molecular and Cellular Biology*, 28(21):6609-19.
- Bartel D.P. (2004). MicroRNAs: Genomics biogenesis, mechanism, and function. *Cell*, 116, 281-297.

- Baum B., Georgiou M. (2011). Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. *The Journal of Cell Biology*, 192(6):907-17.
- Becker K.F., Rosivatz E., Blechschmidt K., Kremmer E., Sarbia M., Höfler H. (2002). Analysis of the E-cadherin repressor Snail in primary human cancers. *Cells, Tissues, Organs*, 185(1-3):204-12.
- Bomken S., Fiser K., Heidenreich O., Vormoor J. (2010). Understanding the cancer stem cell. *British Journal of Cancer*, 103(4):439-45
- Bonitsis N., Batistatou A., Karantima S., Charalabopoulos K. (2006). The role of cadherin/catenin complex in malignant melanoma. *Experimental Oncology*, 28(3):187-93.
- Brand K., Sandig V., Strauss M. (2000). Combined adenoviral transfer of tumor suppressor and cell-cycle genes for tumor-cell apoptosis. *Methods in Molecular Medicine*, 35:151-63.
- Broadus R.R., Lynch P.M., Lu K.H., Luthra R., Michelson S.J. (2004). Unusual tumors associated with the hereditary non polyposis colorectal cancer syndrome. *Modern Pathology: an Official Journal of the United States and Canadian Academy of Pathology, Inc*, 17(8):981-9.
- Bronner C.E., Baker S.M., Morrison P.T., Warren G., Smith L.G., Lescoe M.K., Kane M., Earabino C., Lipford J., Lindblom A., *et al.* (1994). Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature*, 368(6468):258-61.
- Brown J.W., Marshall D.F., Echeverria M. (2008). Intronic noncoding RNAs and splicing. *Trends in Plant Science*, 13(7):335-42.

Brzezianska E., Pastuszak-Lewandoska D. (2011). A mini review: the role of MAPK/ERK and PI3K/Akt pathways in thyroid follicular cell-derived neoplasm. *Frontiers in Bioscience: a Journal and Virtual Library*, 16:422-39.

Buda A., Pignatelli M. (2011). E-cadherin and the cytoskeletal network in colorectal cancer development and metastasis. *Cell Communication & Adhesion*, 18(6):133-43.

Burges A., Schmalfeldt B. (2011). Ovarian cancer: diagnosis and treatment. *Deutsches Ärzteblatt International*, 108(38):635-41.

Cahill S., Smyth P., Denning K., Flavin R., Li J., Potratz A., Guenther S.M., Henfrey R., O'Leary J.J., Sheils O. (2007). Effect of BRAF^{V600E} mutation on transcription and post-transcriptional regulation in a papillary thyroid carcinoma model. *Molecular Cancer*, 6:21

Cantwell-Dorris E.R., O'Leary J.J., Sheils O.M. (2011). BRAFV600E: Implications for Carcinogenesis and Molecular Therapy. *Molecular Cancer Therapeutics*, 10(3):385-94.

Cavaillé J., Seitz H., Paulsen M., Ferguson-Smith A.C., Bachellerie J.P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Human Molecular Genetics* 11(13):1527-38.

Chan Y.S., Yang L., Ng H.H. (2011). Transcriptional regulatory networks in embryonic stem cells. *Progress in Drug Research*, 67:239-52.

Chandra S.H., Wacker I., Appelt U.K., Behrens J., Schneikert J. (2012). A common role for various human truncated adenomatous polyposis coli isoforms in the control of beta-catenin activity and cell proliferation. *PLoS One*, 7(4):e34479.

Chen H.Y., Lin Y.M., Chung H.C., Lang Y.D., Lin C.J., Huang J., Wang W.C., Lin F.M., Chen Z., Huang H.D., Shyy J.Y., Liang J.T., Chen R.H. (2012). miR-103/107 promote metastasis of colorectal_cancer_by targeting the metastasis suppressors DAPK and KLF4. *Cancer Research*, 72(14):3631-41.

Chen L., Yabuuchi A., Eminli S., Takeuchi A., Lu C.W., Hochedlinger K., Daley G.Q. (2009). Cross-regulation of the Nanog and Cdx2 promoters. *Cell Research*, 19(9):1052-61.

Chen R.Z., Pettersson U., Beard C., Jackson-Grusby L., Jaenisch R. (1998). DNA hypomethylation leads to elevated mutation rates. *Nature*, 395(6697):89-93.

Choy S.W., Cheng S.H. (2012). Hedgehog signaling. *Vitamins and Hormones*, 88:1-23.

Cobaleda C., Pérez-Caro M., Vicente-Dueñas C., Sánchez-García I. (2007). Function of the zinc-finger transcription factor SNAI2 in cancer and development. *Annual Review of Genetics*, 41:41-61.

Collins J.F., Herman P., Schuch C., Bagby G.C. (1992). c-myc antisense oligonucleotides inhibit the colony forming capacity of Colo 320 colonic carcinoma cells. *The Journal of Clinical Investigation*, 89(5):1523-7.

Cornett W.R., Sharma A.K., Day T.A., Richardson M.S., Hoda R.S., van Heerden J.A., Fernandes J.K. (2007). Anaplastic thyroid carcinoma: an overview. *Current Oncology Reports*, 9(2):152-8.

Cunningham D., Atkin W., Lenz H.J., Lynch H.T., Minsky B., Nordlinger B., Starling N. (2010). Colorectal cancer. *Lancet*, 375(9719):1030-47.

Danzi S., Klein I. (2012). Thyroid hormone and the cardiovascular system. *The Medical Clinics of North America*, 96(2):257-68.

Davidson B., Tropé C.G., Reich R. (2012). Epithelial-mesenchymal transition in ovarian carcinoma. *Frontiers in Oncology*, 2:33.

De Cuevas M., Matunis E.L. (2011). The stem cell niche: lessons from the *Drosophila* testis. *Development*, 138(14):2861-9.

De Gruijl, F.R., van Kranen, H.J., Mullenders, L.H. (2012). UV-induced DNA damage, repair, mutations and oncogenic pathways in skin cancer. *Journal of Photochemistry and Photobiology. B, Biology*, 63(1-3):19-27.

De Santa Barbara P., van den Brink G.R., Roberts D.J. (2003). Development and differentiation of the intestinal epithelium. *Cellular and Molecular Life Sciences: CMLS*, 60(7):1322-32.

Derbel O., Limem S., Ségura-Ferlay C., Lifante J.C., Carrie C., Peix J.L., Borson-Chazot F., Bournaud C., Droz J.P., de la Fouchardière C. (2011). Results of combined treatment of anaplastic thyroid carcinoma (ATC). *BMC Cancer*, 11:469.

Douville J., Beaulieu R., Balicki D. (2009). ALDH1 as a functional marker of cancer stem and progenitor cells. *Stem Cells Development*, 18(1):17-25.

Downward J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nature Reviews. Cancer*, 3(1):11-22.

Duntas L.H., Tsakalacos N., Grab-Duntas B., Kalarritou M., Papadodima E. (2003). The use of recombinant human thyrotropin (Thyrogen) in the diagnosis and treatment of thyroid cancer. *Hormones (Athens)*, 2(3):169-74.

Dweep, H., Sticht, C., Pandey, P., Gretz, N., (2011). miRWalk - database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. *Journal of Biomedical Informatics*. 44(5):839-47.

Estep A.L., Palmer C., McCormick F., Rauen K.A. (2007). Mutation analysis of BRAF, MEK1 and MEK2 in 15 ovarian cancer cell lines: implications for therapy. *PLoS One*, 2(12):e1279.

Feng G. Du, P. Xia T., Kibbe W., Lin S. (2011). GeneAnswers: Integrated Interpretation of Genes. R package version 1.12.2.

Fire A., Xu S., Montgomery M.K., Kostas S.A., Driver S.E., Mello C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391(6669):806-11.

Förster T. (1948). Intermolecular Energy Migration and Fluorescence. *Annalen der Physik*, 2:55-75.

Fortier L.A. (2005). Stem cells: classifications, controversies, and clinical applications. *Veterinary Surgery: VS*, 34(5):415-23.

Galardi S., Fatica A., Bachi A., Scaloni A., Presutti C., Bozzoni I. (2002). Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Molecular and Cellular Biology*, 22(19):6663-8.

Galvin-Burgess K.E., Vivian J.L. (2011). Transforming growth factor-beta superfamily in mouse embryonic stem cell self-renewal. *Vitamins and Hormones*, 87:341-65.

Gardner P.P., Daub J., Tate J., Moore B.L., Osuch I.H., Griffiths-Jones S., Finn R.D., Nawrocki E.P., Kolbe D.L., Eddy S.R., Bateman A. (2011). Rfam: Wikipedia, clans and the "decimal" release. *Nucleic Acids Research*, 39(Database issue):D141-5.

Goodsell D.S. (1999). The molecular perspective: the ras oncogene. *Oncologist*, 4(3):263-4.

- Grover S., Kastrinos F., Steyerberg E.W., Cook E.F., Dewanwala A., Burbidge L.A., Wenstrup R.J., Syngal S. (2012). Prevalence and phenotypes of APC and MUTYH mutations in patients with multiple colorectal adenomas. *JAMA: the journal of the American Medical Association*, 308(5):485-92.
- Guan K., Chang H., Rolletschek A., Wobus A.M. (2001). Embryonic stem cell-derived neurogenesis. Retinoic acid induction and lineage selection of neuronal cells. *Cell and Tissue Research*, 305(2):171-6.
- Guo D., Xu B.L., Zhang X.H., Dong M.M. (2012). Cancer stem-like side population cells in the human nasopharyngeal carcinoma cell line cne-2 possess epithelial mesenchymal transition properties in association with metastasis. *Oncology Reports*, 28(1):241-7.
- Hammachi F., Morrison G.M., Sharov A.A., Livigni A., Narayan S., Papapetrou E.P., O'Malley J., Kaji K., Ko M.S., Ptashne M., Brickman J.M. (2012). Transcriptional activation by oct4 is sufficient for the maintenance and induction of pluripotency. *Cell Reports*, 1(2):99-109.
- Han M., Liu M., Wang Y., Chen X., Xu J., Sun Y., Zhao L., Qu H., Fan Y., Wu C. (2012). Antagonism of miR-21 Reverses Epithelial-Mesenchymal Transition and Cancer Stem Cell Phenotype through AKT/ERK1/2 Inactivation by Targeting PTEN. *PLoS One*, 7(6):e39520.
- He L., Hannon G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nature Reviews. Genetics*, 5(7):522-31
- He S., Pant D., Schiffmacher A., Bischoff S., Melican D., Gavin W., Keefer C. (2006). Developmental expression of pluripotency determining factors in caprine embryos: novel pattern of NANOG protein localization in the nucleolus. *Molecular Reproduction and Development*, 73(12):1512-22.
- Heinen F.L., Pérez, G. (2012). Ovarian mucinous borderline cystadenoma, in a premenarchal girl. *Archivos Argentinos de Pediatría*, 110(1):e4-8.

Hüttenhofer A., Kiefmann M., Meier-Ewert S., O'Brien J., Lehrach H., Bachellerie J.P., Brosius J. (2001). RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *The EMBO Journal*, 20(11):2943-53.

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011):931-45.

Iyer R.R., Pluciennik A., Burdett V., Modrich P.L. (2006). DNA mismatch repair: functions and mechanisms. *Chemical Reviews*, 106(2):302-23.

Jamieson C., Sharma M., Henderson B.R. (2012). Wnt signaling from membrane to nucleus: β -catenin caught in a loop. *The International Journal of Biochemistry and Cellular Biology*, 44(6):847-50.

Jaas J.R. (2006). Colorectal cancer: a multipathway disease. *Critical Reviews in Oncogenesis*, 12(3-4):273-87.

Jhanwar-Uniyal M. (2003). BRCA1 in cancer, cell cycle and genomic stability. *Frontiers in Bioscience: a Journal and Virtual Library*, 8:s1107-17.

Jones P.A., Baylin S.B. (2002). The fundamental role of epigenetic events in cancer. *Nature Reviews. Genetics*, 3(6):415-28.

Kalathur R.K., Hernández-Prieto M.A., Futschik M.E. (2012). Huntington's Disease and its therapeutic target genes: A global functional profile based on the HD Research Crossroads database. *BMC Neurology*, 12(1):47.

Kalirai H, Damato, B.E., Coupland SE. (2011). Uveal melanoma cell lines contain stem-like cells that self-renew, produce differentiated progeny, and survive chemotherapy. *Investigative Ophthalmology & Visual Science*, 52(11):8458-66

- Kanavy H.E., Gerstenblith M.R. (2011). Ultraviolet radiation and melanoma. *Seminars in Cutaneous Medicine and Surgery*, 30(4):222-8.
- Kelly R.G. (2012). The second heart field. *Current Topics in Developmental Biology*, 100:33-65.
- Khanna A., Stamm, S. (2010). Regulation of alternative splicing by short non-coding nuclear RNAs. *RNA Biology*, 7(4):480-5.
- Kim A., Ueda Y., Naka T., Enomoto T. (2012). Therapeutic strategies in epithelial ovarian cancer. *Journal of Experimental & Clinical Cancer Research: CR*, 31:14.
- Kim R.J., Nam J.S. (2011). OCT4 Expression Enhances Features of Cancer Stem Cells in a Mouse Model of Breast Cancer. *Laboratory Animal Research*, 27(2):147-52.
- Kondo T., Nakazawa T., Murata S., Kurebayashi J., Ezzat S., Asa S.L., Katoh R. (2007). Enhanced B-Raf protein expression is independent of V600E mutant status in thyroid carcinomas. *Human Pathology*, 38 1810–1818.
- Kooreman N.G., Wu J.C. (2010). Tumorigenicity of pluripotent stem cells: biological insights from molecular imaging. *Journal of the Royal Society, Interface / the Royal Society*, 7 Suppl 6:S753-63.
- Kumar R., Angelini S., Snellman E., Hemminki K. (2004). BRAF mutations are common somatic events in melanocytic nevi. *The Journal of Investigative Dermatology*, 122(2):342-8
- Kwak P.B., Iwasaki S., Tomari Y. (2010). The microRNA pathway and cancer. *Cancer Science*, 101(11):2309-15.

- Lakowicz J.R., Keating S. (1983). Binding of an indole derivative to micelles as quantified by phase-sensitive detection of fluorescence. *The Journal of Biological Chemistry*, 10;258 (9):5519-24.
- Larcher F. Dellambra E. Rico L. Bondanza S. Murillas R. Cattiglio C. Mavilio F. Jorcano J.L. Giovanna Z., Del Rio, M. (2007). Long-term engraftment of single genetically modified human epidermal holoclones enables safety pre-assessment of cutaneous gene therapy. *The American Society of Gene Therapy*, 15:9:1670-1676.
- Larrea A.A., Lujan S.A., Kunkel T.A. (2010). SnapShot: DNA mismatch repair. *Cell*, 141(4):730.e1
- Lauth M., Toftgård R. (2011). Hedgehog signaling and pancreatic tumor development. *Advances in Cancer Research*, 110:1-17.
- Lee C.J., Dosch J., Simeone D.M. (2008). Pancreatic cancer stem cells. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 26(17):2806-12.
- Lee R.C., Feinnbaum R.L., Ambros V. (1993). The *C.elegans* heterochronic gene *lin-4* encodes small RNA's with antisense complementary to *lin-14*. *Cell*, 75(5), 843-854
- Lee S.J., Jiko C., Yamashita E., Tsukihara T. (2011). Selective nuclear export mechanism of small RNAs. *Current Opinion in Structural Biology*, 21(1):101-8
- Lee Y., Ahn, C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Radmark O., Kim S., Kim V.N. (2003). The nuclear RNase III Drosha initiates miRNA processing. *Nature*, 425, 415-419
- Lestrade L., Weber M.J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Research*, 34(Database issue):D158-62.

Li C.Y., Cha W., Luder H.U., Charles R.P., McMahon M., Mitsiadis T.A., Klein O.D. (2012). E-cadherin regulates the behavior and fate of epithelial stem cells and their progeny in the mouse incisor. *Developmental Biology*, 366(2):357-66.

Li H., Chen X., Calhoun-Davis T., Claypool K., Tang D.G. (2008). PC3 human prostate carcinoma cell holoclones contain self-renewing tumor-initiating cells. *Cancer Research*, 68(6):1820-5.

Liede A., Karlan B.Y., Narod S.A. (2004). Cancer risks for male carriers of germline mutations in BRCA1 or BRCA2: a review of the literature. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 22(4):735-42.

Lim S.M., Shin S.J., Chung W.Y., Park C.S., Nam K.H., Kang S.W., Keum K.C., Kim J.H., Cho J.Y., Hong Y.K., Cho B.C. (2012). Treatment outcome of patients with anaplastic thyroid cancer: a single center experience. *Yonsei Medical Journal*, 53(2):352-7.

Livak K.J., Schmittgen T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ Method. *Methods: A Companion to Methods in Enzymology*, 25(4):402-8.

Lorenowicz M.J., Korswagen H.C. (2009). Sailing with the Wnt: charting the Wnt processing and secretion route. *Experimental Cell Research*, 315(16):2683-9.

Lu J., He M.L., Wang L., Chen Y., Liu X., Dong Q., Chen Y.C., Peng Y., Yao K.T., Kung H.F., Li X.P. (2011). MiR-26a inhibits cell growth and tumorigenesis of nasopharyngeal carcinoma through repression of EZH2. *Cancer Research*, 71(1):225-33.

Luke J.J., Hodi F.S. (2011). Vemurafenib and BRAF Inhibition: A New Class of Treatment for Metastatic Melanoma. *Clinical Cancer Research*, 18(1):9-14.

Lyamichev V, Brow M.A., Dahlberg J.E. (1993). Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science*, 260(5109):778-83.

Lynam-Lennon N., Maher S.G., Reynolds J.V. (2009). The roles of microRNA in cancer and apoptosis. *Biological reviews of the Cambridge Philosophical Society*, 84(1):55-71.

Ma I., Allan A.L. (2011). The role of human aldehyde dehydrogenase in normal and cancer stem cells. *Stem Cell Reviews*, 7(2):292-306.

Ma Y.L., Zhang P., Wang F., Moyer M.P., Yang J.J., Liu Z.H., Peng J.Y., Chen H.Q., Zhou Y.K., Liu W.J., Qin H.L. (2011). Human embryonic stem cells and metastatic colorectal cancer cells shared the common endogenous human microRNA-26b. *Journal of Cellular and Molecular Medicine*, 15(9):1941-54.

Macara I.G., Lounsbury K.M., Richards S.A., McKiernan C., Bar-Sagi D. (1996). The Ras superfamily of GTPases. *FASEB journal: Official Publication of the Federation of American Societies for Experimental Biology*, 10(5):625-30.

Maden B.E., Hughes J.M. (1997). Eukaryotic ribosomal RNA: the recent excitement in the nucleotide modification problem. *Chromosoma*, 105(7-8):391-400.

Mani S.A., Guo W., Liao M.J., Eaton E.N., Ayyanan A., Zhou A.Y., Brooks M., Reinhard F., Zhang C.C., Shipitsin M., Campbell L.L., Polyak K., Brisken C., Yang J., Weinberg R.A. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, 133(4):704-15

Mazda M., Nishi K., Naito Y., Ui-Tei K. (2011). E-cadherin is transcriptionally activated via suppression of ZEB1 transcriptional repressor by small RNA-mediated gene silencing. *PLoS One*, 6(12):e28688.

- Mei Y.P., Liao J.P., Shen J., Yu L., Liu B.L., Liu L., Li R.Y., Ji L., Dorsey S.G., Jiang Z.R., Katz R.L., Wang J.Y., Jiang F. (2012). Small nucleolar RNA 42 acts as an oncogene in lung tumorigenesis. *Oncogene*, 31(22):2794-804.
- Menczer J., Usviatov I., Ben-Shem E., Golan A., Levy T. (2011). Neoadjuvant chemotherapy in ovarian, primary peritoneal and tubal carcinoma: can imaging results prior to interval debulking predict survival? *Journal of Gynecologic Oncology*, 22(3):183-7.
- Meng W., Jiang L., Lu L., Hu H., Yu H., Ding D., Xiao K., Zheng W., Guo H., Ma W. (2012). Anti-miR-155 oligonucleotide enhances chemosensitivity of U251 cell to taxol by inducing apoptosis. *Cell Biology International*, 36(7):653-9.
- Miyanari Y., Torres-Padilla M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. *Nature*, 483(7390):470-3.
- Moncharmont C., Levy A., Gilormini M., Bertrand G., Chargari C., Alphonse G., Ardail D., Rodriguez-Lafrasse C., Magné N. (2012). Targeting a cornerstone of radiation resistance: cancer stem cell. *Cancer Letters*, 322(2):139-47.
- Morandell S., Yaffe M.B. (2012). Exploiting Synthetic Lethal Interactions Between DNA Damage Signaling, Checkpoint Control, and p53 for Targeted Cancer Therapy. *Progress in Molecular Biology and Translational Science*, 110:289-314.
- Moreb J.S. (2008). Aldehyde dehydrogenase as a marker for stem cells. *Current Stem Cell Research & Therapy*, 3(4):237-46.
- Morigi M., Imberti B., Zoja C., Corna D., Tomasoni S., Abbate M., Rottoli D., Angioletti S., Benigni A., Perico N., Alison M., Remuzzi G. (2004). Mesenchymal stem cells are renotropic, helping to repair the kidney and improve function in acute renal failure. *Journal of the American Society of Nephrology: JASN*, 15(7):1794-804.

Nakayama M., Gonzalzo M.L., Yegnasubramanian S., Lin X., De Marzo A.M., Nelson W.G. (2004). GSTP1 CpG island hypermethylation as a molecular biomarker for prostate cancer. *Journal of Cellular Biochemistry*, 91(3):540-52.

Nayak R.P. (2012). Latest in cystic fibrosis. *Molecular Medicine*, 109(2):127-32.

Nilsen T.W. (2007). Mechanisms of microRNA-mediated gene regulation in animal cells. *Trends in Genetics : TIG*, 23(5):243-9.

Nissan M.H., Solit D.B. (2011). The "SWOT" of BRAF inhibition in melanoma: RAF inhibitors, MEK inhibitors or both? *Current Oncology Reports*, 13(6):479-87.

Nummela P., Lammi J., Soikkeli J., Saksela O., Laakkonen P., Hölttä E. (2012). Transforming growth factor beta-induced (TGFBI) is an anti-adhesive protein regulating the invasive growth of melanoma cells. *The American Journal of Pathology*, 180(4):1663-74.

Oren M., Rotter V. (2010). Mutant p53 gain-of-function in cancer. *Cold Spring Harbour Perspectives in Biology*, 2(2):a001107.

Parkin D.M., Bray F., Ferlay J., Pisani P. (2002). Global cancer statistics, 2002. *CA: a Cancer Journal for Clinicians*, 55(2):74-108.

Parsons M.T., Buchanan D.D., Thompson B., Young J.P., Spurdle A.B. (2012). Correlation of tumour BRAF mutations and MLH1 methylation with germline mismatch repair (MMR) gene mutation status: a literature review assessing utility of tumour features for MMR variant classification. *Journal of Medical Genetics*, 49(3):151-7.

Pasieka J.L. (2003). Anaplastic thyroid cancer. *Current Opinion in Oncology*, 15(1):78-83.

Pei D., Zhang Y., Zheng J. (2012). Regulation of p53: a collaboration between Mdm2 and Mdmx. *Oncotarget*, 3(3):228-35.

Pérez-Mancera P.A., González-Herrero I., Pérez-Caro M., Gutiérrez-Cianca N., Flores T., Gutiérrez-Adán A., Pintado B., Sánchez-Martín M., Sánchez-García I. (2005). SLUG in cancer development. *Oncogene*, 24(19):3073-82.

Pfeiffer M.J., Schalken J.A. (2009). Stem cell characteristics in prostate cancer cell lines. *European Urology*, 57(2):246-54

Phipps K.R., Charette J.M., Baserga S.J. (2011). The small subunit processome in ribosome biogenesis—progress and prospects. *Wiley Interdisciplinary Reviews. RNA*, 2(1):1-21.

Piao L., Zhang M., Datta J., Xie X., Su T., Li H., Teknos T.N., Pan Q. (2012). Lipid-based nanoparticle delivery of Pre-miR-107 inhibits the tumorigenicity of head and neck squamous cell carcinoma. *Molecular Therapy: The Journal of The American Society of Gene Therapy*, 20(6):1261-9.

Piek J.M., van Diest P.J., Verheijen R.H. (2008). Ovarian carcinogenesis: an alternative hypothesis. *Advances in Experimental Medicine and Biology*, 622:79-87.

Poulogiannis G., Luo F., Arends M.J. (2012). RAS signalling in the colorectum in health and disease. *Cell Communication & Adhesion*, 19(1):1-9.

Pytliak M., Vargová V., Mechírová V. (2012). Matrix metalloproteinases and their role in oncogenesis: a review. *Onkologie*, 35(1-2):49-53.

Rasool S., A Ganai B., Syed Sameer A., Masood A. (2012). Esophageal cancer: associated factors with special reference to the Kashmir Valley. *Tumori*, 98(2):191-203.

Ratajczak M.Z., Zuba-Surma E.K., Wysoczynski M., Wan W., Ratajczak J., Wojakowski W., Kucia M. (2008). Hunt for pluripotent stem cell -- regenerative medicine search for almighty cell. *Journal of Autoimmunity*, 30(3):151-62.

- Ravnan M.C., Matalka M.S. (2012). Vemurafenib in Patients With BRAF V600E Mutation-Positive Advanced Melanoma. *Clinical Therapeutics*, 34(7):1474-86.
- Rederstorff M., Hüttenhofer A. (2010). Small non-coding RNAs in disease development and host-pathogen interactions. *Current opinion in Molecular Therapeutics*, 12(6):684-94.
- Reinhart B.J., Slack F.J., Basson M., Pasquinelli A.E., Bettinger J.C., Rougvié A.E., Horvitz H.R., Ruvkun G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 403, 901–906.
- Ren J., Jin P., Wang E., Marincola F.M., Stroncek D.F. (2009). MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells. *Journal of Translational Medicine*, 7:20.
- Rijsewijk F., Schuermann M., Wagenaar E., Parren P., Weigel D., Nusse R. (1987). The *Drosophila* homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell*, 50(4), 649–57.
- Rodrigues N.R., Rowan A., Smith M.E., Kerr I.B., Bodmer W.F., Gannon J.V., Lane D.P. (1990). p53 mutations in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 87(19):7555-9.
- Rosen J.M., Jordan C.T. (2009). The increasing complexity of the cancer stem cell paradigm. *Science*, 324(5935):1670-3
- Rumble S.M., Lacroute P., Dalca A.V., Fiume M., Sidow A., Brudno M. (2009). SHRiMP: accurate mapping of short color-space reads. *PLoS Computational Biology*, 5(5):e1000386.
- Rusconi P., Caiola E., Brogгинi M. (2012). RAS/RAF/MEK inhibitors in oncology. *Current Medicinal Chemistry*, 19(8):1164-76.

Ruvkun G. (2001). Molecular Biology; Glimpses of a tiny RNA world. *Science*, 294 (5543), 797-799.

Sakashita H., Ieta K., Haraguchi N., Inoue Y., Yoshizawa Y., Mori M. (2007). Cancer stem cell. *Gan To Kagaku Ryoho. Cancer & chemotherapy*, 34(11):1721-9.

Sala E., Mologni L., Truffa S., Gaetano C., Bollag G.E., Gambacorti-Passerini C. (2008). BRAF silencing by short hairpin RNA or chemical blockade by PLX4032 leads to different responses in melanoma and thyroid carcinoma cells. *Molecular Cancer Research*, 6(5):751-9.

Saldanha G., Potter L., Daforno P., Pringle J.H. (2006). Cutaneous melanoma subtypes show different BRAF and NRAS mutation frequencies. *Clinical Cancer Research*, (15):4499-505.

Schwarz D.S., Hutvagner G., Du T., Xu Z., Aronin N., Zamore P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2), 199-208.

Schwarz-Cruz-y-Celis A., Meléndez-Zajgla J. (2011). Cancer stem cells. *Revista de Investigación Clínica; Organo del Hospital de Enfermedades de la Nutrición*, 63(2):179-86.

Schoof C.R., Botelho E.L., Izzotti A., Vasques Ldos R. (2012). MicroRNAs in cancer treatment and prognosis. *American Journal of Cancer Research*, 2(4):414-33.

Shih S.R., Chiu W.Y., Chang T.C., Tseng C.H. (2012). Diabetes and thyroid cancer risk: literature review. *Experimental Diabetes Research*, 2012:578285.

Shirley S.H., Hudson L.G., He J., Kusewitt D.F. (2010). The skinny on Slug. *Molecular Carcinogenesis*, 49(10):851-61.

Silberstein E.B. (2012). Radioiodine: the classic theranostic agent. *Seminars in Nuclear Medicine*, 42(3):164-70.

Singh A., Settleman J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene*, 29(34):4741-51.

Sinicrope F.A., Sargent D.J. (2012). Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. *Clinical Cancer Research*, 18(6):1506-12.

Smyth P., Finn S., Cahill S., O'Regan E., Flavin R., O'Leary J.J., Sheils O. (2005). ret/PTC and BRAF act as distinct molecular, time dependant triggers in a sporadic Irish cohort of papillary thyroid carcinoma. *International Journal of Surgical Pathology*, 13(1):1-8.

Song Z., Yue W., Wei B., Wang N., Li T., Guan L., Shi S., Zeng Q., Pei X., Chen L. (2011). Sonic hedgehog pathway is essential for maintenance of cancer stem-like cells in human gastric cancer. *PLoS One*, 6(3):e17687.

Spassov D.S., Jurecic R. (2003). The PUF family of RNA-binding proteins: does evolutionarily conserved structure equal conserved function? *IUBMB Life*, 55(7):359-66.

Sterneckert J., Höing S., Schöler H.R. (2012). Concise review: Oct4 and more: the reprogramming expressway. *Stem Cells*, 30(1):15-21.

Su P.J., Chen J.S., Liaw C.C., Chang H.K., Wang H.M., Yang T.S., Lin Y.C., Liao C.T., Yang H.Y., Yeh K.Y., Ho M.M., Chang N.J., Wang C.H., Chang J.W. (2011). Biochemotherapy with carmustine, cisplatin, dacarbazine, tamoxifen and low-dose interleukin-2 for patients with metastatic malignant melanoma. *Chang Gung Medical Journal*, 34(5):478-86.

Subramanyam D. & Blelloch R. (2011). From microRNAs to targets: pathway discovery in cell fate transitions. *Current Opinion in Genetics & Development*, 21(4):498-503.

Sun H., Hu K., Wu M., Xiong J., Yuan L., Tang Y., Yang Y., Liu H. (2011). Contact by melanoma cells causes malignant transformation of human epithelial-like stem cells via alpha V integrin activation of transforming growth factor β 1 signaling. *Experimental Biology and Medicine (Maywood)*, 236(3):352-65.

Sun L. (2004). Tumor-suppressive and promoting function of transforming growth factor beta. *Frontiers in Bioscience: a Journal and Virtual Library*, 9:1925-35.

Tao Y., Mao J., Zhang Q., Li L. (2011). Overexpression of Hedgehog signaling molecules and its involvement in triple-negative breast cancer. *Oncology Letters*, 2(5):995-1001.

Tarhini A.A., Iqbal F. (2010). CTLA-4 blockade: therapeutic potential in cancer treatments. *OncoTargets and Therapy*, 3:15-25.

Terhaar sive Droste J.S., Oort F.A., van der Hulst R.W., Coupé V.M., Craanen M.E., Meijer G.A., Morsink L.M., Visser O., van Wanrooij R.L., Mulder C.J. (2010). Does delay in diagnosing colorectal cancer in symptomatic patients affect tumor stage and survival? A population-based observational study. *BMC Cancer*, 10:332.

Tian X., Liu Z., Niu B., Zhang J., Tan T.K., Lee S.R., Zhao Y., Harris D.C., Zheng G. (2011). E-cadherin/ β -catenin complex and the epithelial barrier. *Journal of Biomedicine and Biotechnology*, 2011:567305.

Tie J., Fan D. (2011). Big roles of microRNAs in tumorigenesis and tumor development. *Histology and Histopathology*, 26(10):1353-61.

Tinelli R., Tinelli A., Tinelli F.G., Cicinelli E., Malvasi A. (2006). Conservative surgery for borderline ovarian tumors: a review. *Gynecologic Oncology*, 100(1):185-91.

Trinh V.A., Hwu W.J. (2012). Ipilimumab in the treatment of melanoma. *Expert Opinion on Biological Therapy*, 12(6):773-82.

Tsao H., Chin L., Garraway L.A., Fisher D.E. (2012). Melanoma: from mutations to medicine. *Genes & Development*, 26(11):1131-55.

Unno K., Jain M., Liao R. (2012). Cardiac side population cells: moving toward the center stage in cardiac regeneration. *Circulation Research*, 110(10):1355-63.

Vakiani E., Solit D.B. (2011). KRAS and BRAF: drug targets and predictive biomarkers. *The Journal of Pathology*, 223(2):219-29.

Valenta T, Hausmann G, Basler K. (2012). The many faces and functions of β -catenin. *The EMBO Journal*, 31(12):2714-36.

Valleron W., Laprevotte E., Gautier E.F., Quelen C., Demur C., Delabesse E., Agirre X., Prósper F., Kiss T., Brousset P. (2012). Specific small nucleolar RNA expression profiles in acute leukemia. *Leukemia*, 26(9):2052-60.

Van Etten J., Schagat T.L., Hrit J., Weidmann C.A., Brumbaugh J., Coon J.J., Goldstrohm A.C. (2012). Human Pumilio Proteins Recruit Multiple Deadenylation Complexes to Efficiently Repress Messenger RNAs. *The Journal of Biological Chemistry*, 287(43):36370-83.

Verrecchia F., Mauviel A. (2002). Transforming growth factor-beta signaling through the Smad pathway: role in extracellular matrix gene expression and regulation. *The Journal of Investigative Dermatology*, 118(2):211-5.

Volkovova K., Bilanicova D., Bartonova A., Letašiová S., Dusinska M. (2012). Associations between environmental factors and incidence of cutaneous melanoma. *Environmental health: a Global Access Science Source*, 11Suppl 1:S12.

Wagers A.J., Weissman I.L. (2004). Plasticity of adult stem cells. *Cell*, 116(5):639-48.

Wang Z., Oron E., Nelson B., Razis S., Ivanova N. (2012). Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell*, 10(4):440-54.

Williams A.R., Hare, J.M. (2011). Mesenchymal stem cells: biology, pathophysiology, translational findings, and therapeutic implications for cardiac disease. *Circulatory Research*, 109(8):923-40

Wobus A.M., Boheler K.R. (2005). Embryonic stem cells: prospects for developmental biology and cell therapy. *Physiological Reviews*, 85(2):635-78

Yokota T. (2012). Are KRAS/BRAF mutations potent prognostic and/or predictive biomarkers in colorectal cancers? *Anticancer Agents in Medicinal Chemistry*, 12(2):163-71.

Yoshiki R., Nakamura M., Tokura Y. (2012). The biological role of UVB-induced cutaneous immunosuppression. *Journal of UOEH*, 34(1):77-83.

You J.S., Jones P. A. (2012). Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*, 22(1):9-20.

Yu D., Zhou H., Xun Q., Xu X., Ling J., Hu Y. (2012). microRNA-103 regulates the growth and invasion of endometrial cancer cells through the downregulation of tissue inhibitor of metalloproteinase 3. *Oncology Letters*, 3(6):1221-1226.

Zafarana G., Avery S.R., Avery K., Moore H.D., Andrews P.W. (2012). Specific knockdown of OCT4 in human embryonic stem cells by inducible short hairpin RNA interference. *Stem Cells*, 27(4):776-82.

Zafon C., Obiols G. (2009). The mitogen-activated protein kinase (MAPK) signaling pathway in papillary thyroid cancer. From the molecular bases to clinical practice. *Endocrinología y Nutrición: órgano de la Sociedad Española de Endocrinología y Nutrición*, 56(4):176-86.

Zhang B., Wang Q., Pan X. (2007). MicroRNAs and their regulatory roles in animals and plants. *Journal of Cell Physiology*, 210(2):279-89.

Zhang Z., Li Z., Gao C., Chen P., Chen J., Liu W., Xiao S., Lu H. (2008). miR-21 plays a pivotal role in gastric cancer pathogenesis and progression. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, 88(12):1358-66.

Zlobec I, Lugli A. Invasive front of colorectal cancer: dynamic interface of pro-/anti-tumor factors. *World Journal of Gastroenterology*, 15(47):5898-906.

Anders S. (2011). HTSeq: Analysing high-throughput sequencing data with Python.
<http://www-huber.embl.de/users/anders/HTSeq/>

Genome Reference Consortium. (2012).
<http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/data/index.shtml>

The Irish Cancer Society. (2012).
<http://www.cancer.ie/cancer-information/ovarian-cancer/about>

The National Cancer Registry Ireland. (2012).
http://www.ncri.ie/data.cgi/client/generate_stats.php

Pathology Outlines. (2012). Colon tumor Polyps Adenoma-carcinoma sequence of colon.
<http://www.pathologyoutlines.com/topic/colontumoradenomacarcinoma.html>

Appendix I
Supplementary Methods

A.1.1 DNA Extraction from FFPE Tissue

Overview

The ability to isolate nucleic acid that is suitable for molecular analysis from archived tissue samples provides a powerful tool in retrospective studies of diseased tissue at both the genomic and gene expression level. Standard preservation techniques for storage of biological tissue samples use formaldehyde or, less frequently, paraformaldehyde. The high reactivity of these chemicals makes them ideal for maintaining tissue structure and preventing putrefaction, however, tissues preserved in this fashion have historically been thought to be unusable for molecular analysis. Nucleic acid is both trapped and modified by extensive protein-protein and protein-nucleic acid crosslinks. RNA (and to a lesser extent DNA) in histological samples typically is fragmented and chemically modified to a degree that renders it incompatible with many molecular analysis techniques. DNA tends not to fragment as easily as RNA however; the nucleo-histone matrix is quite dense and also appears to be much more reactive to formaldehyde.

An overnight protease digestion is required to release substantial amounts of DNA from the tissue extract. After purification on the glass-fibre filter, DNA recovered with the RecoverAll Total Nucleic Acid Isolation Kit can typically be used for PCR and other downstream applications. The RecoverAll™ Total Nucleic Acid Isolation Kit is designed to extract total nucleic acids (RNA, miRNA, and DNA) from formaldehyde- or paraformaldehyde-fixed, paraffin-embedded (FFPE) tissues. Four 20µm sections were processed per reaction.

DNA isolation from FFPE material

Solution Preparation

Ethanol was added to Wash 1 concentrate & Wash 2/3 concentrate as per manufacturer's protocol.

Protease Digestion

Deparaffinisation of samples was performed as part of the H&E staining process. 400µl of Digestion Buffer and 4µl of Protease was added to each sample. Tubes were swirled gently to mix and to immerse the tissue. Samples were incubated in a heat block for 48 hours at 50°C.

Nucleic Acid Isolation

480µl of Isolation Additive was added to each sample and vortexed to mix. A white and cloudy solution indicated adequate mixing. 1.1ml 100% ethanol was added to each sample and mixed by pipetting until the solution became clear. A Filter Cartridge was placed into each one of the Collection Tubes supplied and 700µl of the sample/ethanol mixture was pipetted onto the filter. Samples were centrifuged at 10,000 x g (typically 10,000rpm) for 30–60 sec to pass the mixture through the filter. Flow-through was discarded and each Filter Cartridge was re-inserted into the same Collection Tube. This was repeated until all the sample mixture had passed through the filter. 700µl of Wash 1 was added to the Filter Cartridge and centrifuged for 30 sec at 10,000 x g to pass the mixture through the filter. Flow-through was discarded and the Filter Cartridge was re-inserted into the same Collection Tube. 500µl of Wash 2/3 was added to the Filter Cartridge and centrifuged for 30 sec at 10,000 x g to pass the mixture through the filter. Flow-through was discarded before re-inserting the Filter Cartridge into the same Collection Tube. Each column was centrifuged for an additional 30 sec to remove residual fluid from the filter.

Nuclease Digestion and Final Nucleic Acid Purification

For DNA isolation, 60µl RNase mix was added (containing 10µl RNase A and 50µl Nuclease-free Water) to the centre of each Filter Cartridge. Each tube was closed incubated for 30 min at 37°C. 700µl of Wash 1 was added to the Filter Cartridge and incubated for 30–60 sec at room temperature before being centrifuged for 30 sec at 10,000 x g. Flow-through was discarded and Filter Cartridge was re-inserted into the same Collection Tube. 500µl of Wash 2/3 was added to the Filter Cartridge and centrifuge for 30 sec at 10,000 x g. Flow-through was discarded, Filter Cartridge was re-inserted into the same Collection Tube. Wash was repeated a second time with 500µl of Wash 2/3 before being centrifuged for 1 min at 10,000 x g to remove residual fluid from the filter. The Filter Cartridge was transferred to a fresh Collection Tube. 30µl of Elution Solution was added to the centre of the filter and heated to 95°C. Samples were allowed to sit at room temperature for 1 min before being centrifuged for 1 min at maximum speed to pass the mixture through the filter. The Elution step was repeated with a second 30µl aliquot of solution, using the same Collection Tube and DNA was stored at -20°C.

A.1.2 Genomic DNA Purification

Genomic DNA was extracted from cultured cells using the Qiagen DNeasy Blood and Tissue Kit. All centrifugation steps were carried out at room temperature (15–25°C) in a microcentrifuge and vortexing was performed by pulse-vortexing for 5–10 seconds. Buffer AW1 and Buffer AW2 were supplied as concentrates. Before using for the first time, the appropriate amount of ethanol (96–100%) as indicated on the bottle was added to obtain a working solution.

Protocol for Cultured Cells

An appropriate number of cells (maximum 5×10^6) were centrifuged for 5 min at 300 x g. The pellet was resuspended in 200µl PBS before adding 20µl of proteinase K and 4µl of RNase A (100 mg/ml), vortexing to mix and incubating for 2 min at room temperature. When using a frozen cell pellet, cells were allowed to thaw before adding PBS until the pellet could be dislodged by gently flicking the tube. 200µl Buffer AL was added to the cells, mixed thoroughly by vortexing, and incubated at 56°C for 10 min. It was essential that the sample and Buffer AL were mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution. 200µl ethanol (96–100%) was added to the sample, and mixed thoroughly by vortexing. The homogenous mixture was pipetted into a DNeasy Mini spin column placed in a 2ml collection tube. The spin column was centrifuge at 6000 x g (8000 rpm) for 1 min before discarding the flow-through and collection tube. The DNeasy Mini spin column was placed into a new 2 ml collection tube and 500µl Buffer AW1 was added. The spin column was centrifuged for 1 min at 6000 x g (8000 rpm) before discarding the flow-through and collection tube.

The DNeasy Mini spin column was placed into a new 2 ml collection tube and 500µl Buffer AW2 was added. The spin column was centrifuged for 3 min at 20,000 x g (14,000 rpm) to dry the DNeasy membrane before discarding the flow-through and collection tube. (It was important to dry the membrane of the DNeasy Mini spin column, since residual ethanol can interfere with subsequent reactions. This centrifugation step ensured that no residual ethanol was carried over during the subsequent elution). Following the centrifugation step, the DNeasy Mini spin column was removed carefully so that the column does not come into contact with the flow-through, since this can result in carryover of ethanol. (If carryover of

ethanol occurred, the collection tube was emptied before reusing it in another centrifugation for 1 min at 20,000 x g (14,000 rpm)). The DNeasy Mini spin column was placed into a clean 1.5ml microcentrifuge tube and 200µl Buffer AE was added directly onto the DNeasy membrane. The tube was incubated at room temperature for 1 min and then centrifuged for 1 min at 6000 x g (8000 rpm) to elute.

Elution with 100µl (instead of 200µl) increased the final DNA concentration in the eluate, but also decreased the overall DNA yield. For maximum DNA yield, the elution step was repeated once as described above. This step gave increased overall DNA yield. A new microcentrifuge tube was used for the second elution step to prevent dilution of the first eluate.

A.1.3 RNA Extraction

The miRNeasy Mini Kit combines phenol/guanidine-based lysis of samples and silica-membrane based purification of total RNA. The miRNeasy Kit provides a separate specialized protocol for enrichment of miRNA and other small RNAs (less than ~200 nt) separate fractions. Buffers RWT and RPE are supplied as concentrates so the required volume of ethanol (96%–100%), as indicated on the bottle, was added to obtain a working solution.

Protocol for Purification of Total RNA, Including Small RNAs

Cells grown in a monolayer (not more than 1×10^7 cells) in cell-culture vessels can be either lysed directly in the vessel (up to 10 cm diameter) or trypsinized and collected as a cell pellet prior to lysis. Cells grown in a monolayer in cell-culture flasks should always be trypsinized.

Spent medium was removed the cells were washed with 1X PBS. PBS was aspirated 0.1–0.25% trypsin EDTA was added to the cell monolayer. After the cells detached from the dish or flask, appropriate complete culture medium was added to inactivate the trypsin and the cells were transferred to an RNase-free polypropylene centrifuge tube and centrifuged at 300 x g for 5 min. The supernatant was aspirated before progressing to cell lysis.

(Incomplete removal of cell-culture medium can inhibit cell lysis and dilute the lysate, affecting the conditions for binding of RNA to the RNeasy Mini spin column membrane which can reduce RNA yield.

Cells were disrupted by adding 700µl QIAzol Lysis Reagent to the pellet before being vortexed for 1 min to mix. (Incomplete loosening of the cell pellet can lead to inefficient lysis and reduced RNA yields). The homogenate was allowed to incubate at room temperature (15–25°C) for 5 min on the bench top to promote the dissociation of nucleoprotein complexes. 140µl chloroform was added to the homogenate and vortexed vigorously for 15 seconds as thorough mixing is important for subsequent phase separation. The homogenate was allowed to incubate at room temperature (15–25°C) for 2-3 min on the bench top. The homogenate was centrifuged for 15 min at 12,000 x g at 4°C to separate the sample into 3 phases: an upper, colourless, aqueous phase containing RNA; a white interphase; and a lower, red, organic phase. The volume of the aqueous phase was approximately 350µl.

Preparation of miRNA-Enriched Fractions Separate from Larger RNAs (>200 nt)

The upper aqueous phase was transferred into a new reaction tube and 350µl 70% ethanol was added before mixing thoroughly by vortexing. The sample (approx. 700µl), including any precipitate that may have formed, was pipetted into an RNeasy Mini spin column placed in a 2 ml collection tube. The sample was then centrifuged at 8000 x g (10,000 rpm) for 15 seconds at room temperature (15–25°C). The flow-through (which contains miRNA) was pipetted into a 2ml reaction tube. As both the miRNA-enriched fraction and larger RNAs (>200nt) were being purified, the RNeasy Mini spin column was stored at 4°C for later 450µl of 100% ethanol (0.65 volumes) to the flow-through from step 2 and mix thoroughly by vortexing. 700µl of the sample was pipetted into an RNeasy MinElute spin column placed in a 2 ml collection tube and centrifuged for 15 seconds at 8000 x g (10,000rpm) at room temperature the flow-through was discarded. This step was repeated until the whole sample had been pipetted into the spin column with the flow-through discarded each time. 500µl Buffer RPE was pipetted into the RNeasy MinElute spin column and centrifuged for 15 seconds at 8000 x g (10,000 rpm) before discarding the flow-through. 500µl of 80% ethanol was added to the RNeasy MinElute spin column and centrifuged for 2 min at 8000 x g (10,000 rpm) to dry the spin column membrane before discarding the flow-through and the

collection tube. (After centrifugation, the RNeasy MinElute spin column was removed carefully from the collection tube so that the column did not contact the flow-through as carryover of ethanol can occur if there is contact). The RNeasy MinElute spin column was placed into a new 2 ml collection tube with the lid open was centrifuged for 5 min at 8000 x g (10,000 rpm). The RNeasy MinElute spin column was then placed into a 1.5ml collection tube and 14µl RNase-free water was pipetted onto the spin column membrane. The lid was closed gently and centrifuged for 1 min at 8000 x g (10,000 rpm) to elute the miRNA-enriched fraction.

Purifying total RNA (>200nt) using the RNeasy Mini spin column

700µl Buffer RWT was pipetted into the RNeasy Mini spin column from the earlier steps and centrifuged for 15 s at 8000 x g (10,000 rpm) to wash the spin column membrane. The flow-through was discarded before adding 500µl Buffer RPE to the RNeasy Mini spin column and centrifuged for 15 s at 8000 x g (10,000 rpm) to wash the spin column membrane. The flow-through was discarded before pipetting another 500µl Buffer RPE into the RNeasy Mini spin column and centrifuging for 15 s at 8000 x g (10,000 rpm) to wash the spin column membrane. Both the flow-through and the collection tube were discarded before placing the RNeasy Mini spin column in a new 2 ml collection tube and centrifuging at full speed for 1 min. The RNeasy Mini spin column was then placed into a new 1.5 ml collection tube and 40µl RNase-free water was pipette directly onto the spin column membrane. The lid was closed gently and centrifuged for 1 min at 8000 x g (10,000 rpm) to elute the total RNA. The elution step was then repeated to ensure RNA yield was maximised as some RNA can remain in the filter after only one elution.

A.1.4 Nucleic Acid Quality Control

A.1.4.1 Analysing RNA by the Agilent 2100 Bioanalyser

Large RNA (above 200nt) integrity and concentration were analysed using the 2100 Bioanalyser.

Preparing the Gel

All reagents were allowed to equilibrate to room temperature for 30 minutes before use. 550µl of Agilent RNA 6000 Nano gel matrix (**red** colour) was placed into the top receptacle of a spin filter which was inserted into a microcentrifuge and centrifuged for 10 minutes at 1500g ± 20%. 65µl filtered gel was aliquoted into 0.5ml RNase-free microfuge tubes that were included in the kit and stored at 4°C

Preparing the Gel Dye Mix

All reagents were allowed to equilibrate to room temperature for 30 minutes before use with the dye concentrate being protected from light while bringing it to room temperature. The RNA 6000 Nano dye concentrate (**blue** colour) was vortexed for 10 seconds and spun down. 1µl of RNA 6000 Nano dye concentrate was added to a 65µl aliquot of filtered gel (prepared as described in “Preparing the Gel” above), vortexed thoroughly and visually inspected to ensure proper mixing of gel and dye. The tube was centrifuged at 13000g for 10 minutes at room temperature and was used within one day.

Loading the Gel Dye Mix

Before loading the gel-dye mix, it was confirmed that the base plate of the chip priming station was in position (C) as below in Figure A.1.4.1(a) (A) and the adjustable clip was set to the top position as below in Figure A.1.4.1(a) (B).

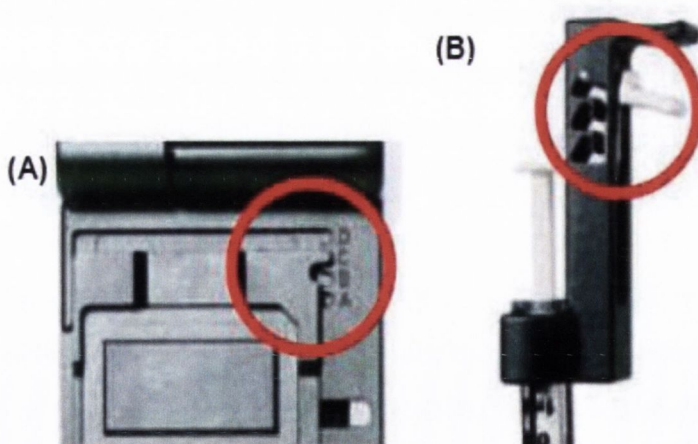


Figure A.1.4.1(a): Bioanalyzer (A) and (B) Priming Station

A new RNA Nano chip was placed on the chip priming station and 9.0µl of the gel-dye mix was pipetted into the bottom of the well marked **G**.

The plunger was positioned at 1ml and then the chip priming station was closed. After exactly 30 seconds the plunger was released via the clip release mechanism. After 5 seconds plunger was slowly pulled back up to the 1ml position. The chip priming station was opened and 9µl of the gel-dye mix was pipetted into each of the wells marked: **G**

Loading the RNA Nano Marker

5µl of the RNA 6000 Nano marker (**green** colour) was pipetted into the well marked with the ladder symbol and each of the 12 sample wells. Unused wells were filled with 6µl of the RNA 6000 Nano marker.

Loading Samples and Ladder

Before use, a ladder aliquot was thawed and kept on ice. To minimise secondary structure, the samples were heat-denatured at 70°C for 2 minutes before loading on the chip. 1µl of RNA ladder was pipetted into the ladder well and 1µl of each sample was pipetted into each of the 12 sample wells. The chip was vortexed for 60 seconds at 2400rpm and placed carefully into the instrument receptacle.

Starting the 2100 expert software

In the Instrument context, the appropriate assay was selected from the Assay menu. The Start button in the upper right of the window was clicked to start the chip run with the incoming raw signals displayed in the Instrument context. To check the results of your run, the Gel or Electropherogram tab was selected in the Data context. The electropherogram of the ladder well window resembled those shown below in Figure A.1.4.1 (b).

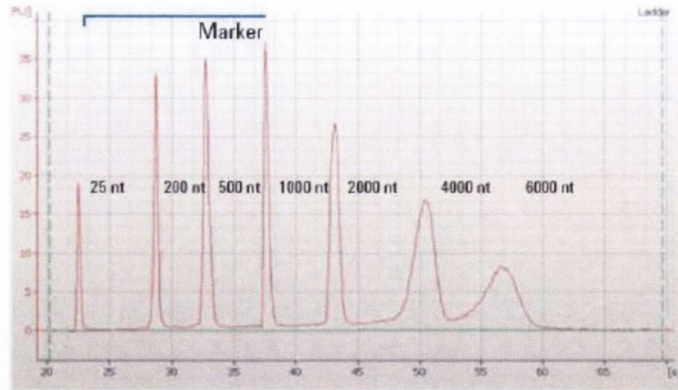


Figure A.1.4.1(b): RNA Nano Ladder Electropherogram

To review the results of a specific sample, the sample name in the tree view was selected and the Results sub-tab was highlighted. The electropherogram of the sample well window for total RNA (eukaryotic) resembled the one shown below in Figure A.1.4.1(c).

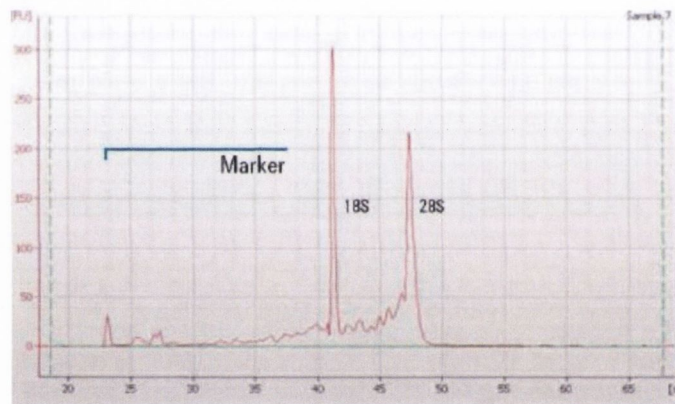


Figure A.1.4.1(c): RNA Nano Sample Electropherogram

By selecting the Results sub-tab, values for the calculated RNA concentration, the ribosomal ratio and the RNA Integrity Number (RIN) were displayed.

A.1.4.2 Analysing Small RNA by the Agilent 2100 Bioanalyser

All reagents were allowed to equilibrate to room temperature for 30 minutes before use. 650µl of Small RNA gel matrix (red colour) was placed into the top receptacle of a spin filter which was inserted into a microcentrifuge and centrifuged for 10 minutes at 10,000g ± 20%. The filter was removed to store the gel at 4°C and was used within one month of preparation.

Preparing the Gel-Dye Mix

All reagents were allowed to equilibrate to room temperature for 30 minutes before use with the dye concentrate being protected from light while bringing it to room temperature. Small RNA dye concentrate (**blue**) was vortexed for 10 seconds and spin down. 2 μ l of dye was aliquoted into 0.5 ml RNase free microtubes (provided with the kit). 40 μ l of filtered gel was added to an aliquot of dye and mixed by pipetting to obtain a homogeneous mixture before being centrifuged at 13000 x g for 10 minutes at room temperature.

Loading the Gel-Dye Mix

The gel-dye mix was allowed to equilibrate to room temperature for 30 minutes before use and protected from light during this time. A new Small RNA chip was placed on the chip priming station and 9 μ l of the gel-dye mix was pipetted into the bottom of the well marked **G**.

The plunger was positioned at 1ml before closing the chip priming station and pressed down until it was held by the clip. After exactly 60 seconds the plunger was released with the clip release mechanism and visually inspected to ensure that the plunger moved back at least to the 0.3ml mark. After 5 seconds plunger was pulled back slowly to the 1ml position and the chip priming station was opened. 9 μ l of the gel-dye mix was pipetted into each of the wells marked **G**.

Loading the Small RNA Conditioning Solution and Marker

9 μ l of the Small RNA conditioning solution (white) was pipette into the well marked CS.

5 μ l of the Small RNA marker (**green**) was pipette into the well marked with a ladder symbol and each of the 11 sample wells.

Loading Samples and Ladder

To minimize secondary structure, samples were heat denatured at 70 °C for 2 minutes before loading on the chip. 1 μ l of the Small RNA ladder was pipette into the well marked with the ladder symbol and 1 μ l of each sample was pipette into each of the 11 sample wells. The chip was then vortexed for 60 seconds at the indicated setting (2400 rpm) before being placed carefully into the instrument receptacle.

To check the results of the run, the Gel or Electropherogram tab was selected in the Data context. The electropherogram of the ladder well window resembled the one shown below in Figure A.1.4.2 (a).

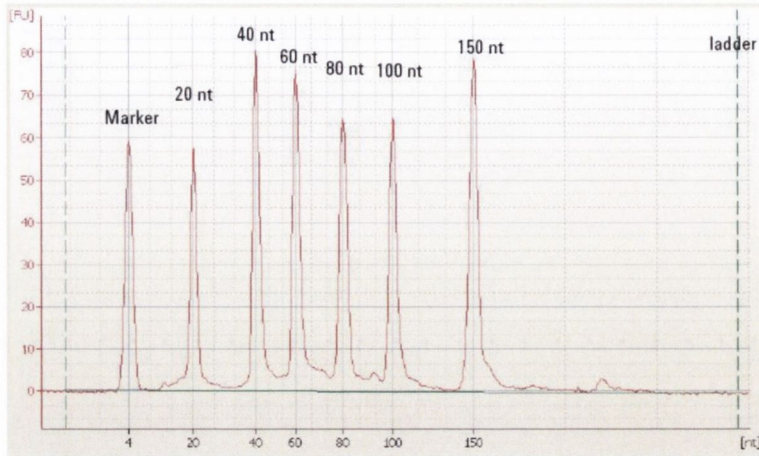


Figure A.1.4.3(a): Small RNA Ladder Electropherogram

Major features of a successful ladder run are:

- 1 marker peak
- 6 RNA peaks
- All 7 peaks are well resolved
- Correct peak size assignment in electropherogram

To review the results of a specific sample, each sample name was selected in the tree view and the Results sub-tab was highlighted. The electropherogram of the sample well window for total RNA (eukaryotic) resembled the one shown below in figure A.1.4.2(b).

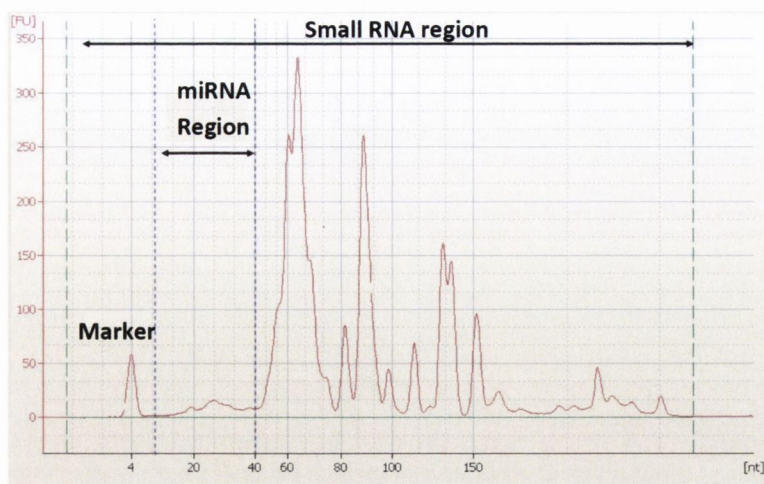


Figure A.1.4.2(b): Small RNA Sample Electropherogram

Major features for a successful total RNA run are:

- 1 marker peak

Two distinct regions could be defined arbitrarily with the Small RNA region from 0 to 150nt, and the micro RNA region (mi RNA) from the 10 to 40nt. These regions could be modified by selecting the Region table or by sliding borders in the electropherogram, implemented with the 2100 expert software version B02.04.

A.1.4.4 Analysing DNA by the Agilent 2100 Bioanalyser



Having been converted to cDNA and amplified, small RNA libraries were analyzed using to Agilent DNA 1000 Kit. Library integrity and concentration was analysed using the 2100 Bioanalyser and the software's smear analysis program allowed regions of interest to be investigated. To analyse DNA quality and concentration the protocol below was used.

Preparing the Gel-Dye Mix

DNA dye concentrate (**blue**) and DNA gel matrix (**red**) were allowed to equilibrate to room temperature for 30 min.

DNA dye concentrate (**blue**) was vortexed and 25µl was added to a DNA gel matrix vial (**red**). The solution was vortexed to mix and spun down before being transferred to a spin filter. The filter was then centrifuged at 2240 g ± 20 % for 15 min and stored at 4 °C protected from light.


Loading the Gel-Dye Mix

The gel-dye mix was allowed equilibrate to room temperature for 30 min before use and a new DNA chip was placed on the chip priming station. 9µl of gel-dye mix was pipetted into the well marked . The plunger was positioned at 1 ml, the chip priming station was closed and pressed down for exactly 60 seconds. After 5 seconds the plunger was pulled slowly back to the 1ml position. 9µl of gel-dye mix was then pipette into the wells marked .

Loading the Markers

5µl of marker (**green**) was pipetted into all 12 sample wells and ladder well.

Loading Samples and Ladder

1µl of DNA ladder (yellow) was pipetted into the well marked  and 1µl of sample (used wells) or 1µl of de-ionized water (unused wells) was pipetted into each of the 12 sample wells. The chip was then vortexed for 1 min at the indicated setting (2400 rpm) and run in the Agilent 2100 Bioanalyzer within 5 min.

Enabling Smear Analysis

The 2100 expert software allows you to perform a smear analysis for all electrophoresis assays. When the smear analysis is enabled, the software allows you to define regions of interest. These regions were used to define the area of broad peaks and determine their part of the total area. Smear analysis provide a means to analyse broad signals that can be hardly evaluated with the normal peak assignment.

We could then define regions of interest that contain the peaks (base pair size) that we were interested in. For these regions we could determine the covered area in relation to the total area.

To enable smear analysis:

- The Electropherogram tab in the Data context was selected:
- The set point explorer was highlighted the Local or Global tab, depending on which samples should be analysed was selected.
- Advanced mode was selected.

- Under Smear Analysis, the check box Perform Smear Analysis was selected.
- The Region Table sub-tab was then added to the Electropherogram tab.

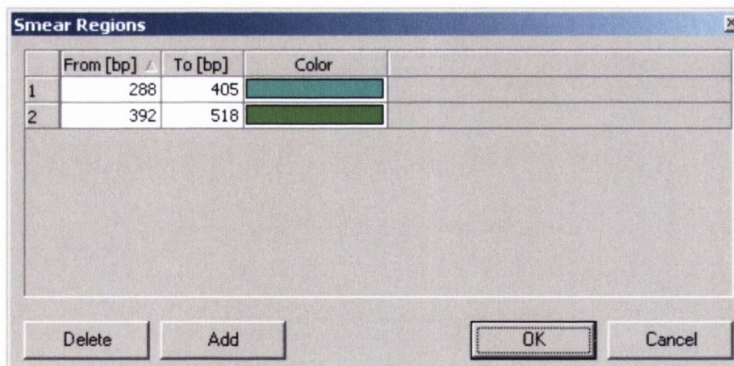
Performing Smear Analysis



After enabling the smear analysis in the set point explorer, we were able to insert regions of interest in the electropherogram.

To do so we:

- Selected the Region Table sub-tab in the Electropherogram tab.
- Right-clicked the electropherogram and selected Add region. A region was inserted into the electropherogram and the Region Table showed the values for the inserted region.
- Repeated the previous step until the number of required regions was inserted.
- Adjusted the regions by directly moving the dashed lines in the electropherogram.
- Removed a region if desired, by right-clicking the dashed line in the electropherogram and selecting Remove Region from the context menu.

In the smear analysis table, the Region Start Size and Region End Size could be edited, for example:



	From [bp]	To [bp]	Color
1	288	405	
2	392	518	

The screenshot shows a dialog box titled "Smear Regions" with a table containing two rows of data. The first row has a region starting at 288 bp and ending at 405 bp, with a green color swatch. The second row has a region starting at 392 bp and ending at 518 bp, also with a green color swatch. Below the table are buttons for "Delete", "Add", "OK", and "Cancel".

Figure A.1.4.3(a): Example of region start and end size in Smear Analysis tool

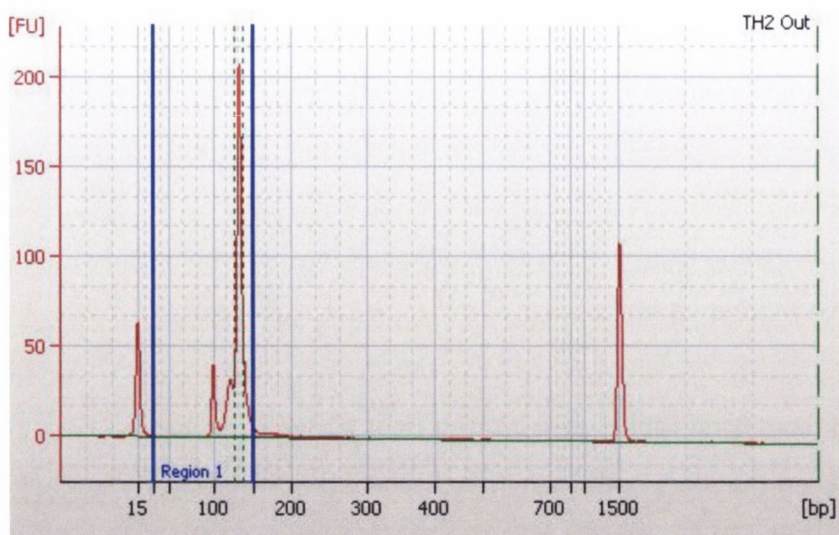


Figure A.1.4.3(b): Typical size profile of a successfully amplified library

A.1.5 Reverse Transcription of RNA to cDNA

The High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, CA, USA) was used to convert total RNA to single stranded cDNA. This kit uses the random primer scheme for initiating cDNA synthesis as it ensures that the first strand synthesis occurs efficiently with all species of RNA molecules present including messenger RNA (mRNA) and ribosomal RNA (rRNA). The cDNA synthesis protocol for reverse transcription is outlined below.

2X master mix preparation

The 2X reverse transcription (RT) master mix was prepared as follows: The volume of components needed to prepare cDNA for each RNA sample was calculated (with 10% overage), using the table below:

Component	Volume/Reaction (µL)	
	Kit with RNase Inhibitor	Kit without RNase Inhibitor
10X RT Buffer	2.0	2.0
25X dNTP Mix (100 mM)	0.8	0.8
10X RT Random Primers	2.0	2.0
MultiScribe™ Reverse Transcriptase	1.0	1.0
RNase Inhibitor	1.0	—
Nuclease-free H ₂ O	3.2	4.2
Total per Reaction	10.0	10.0

Table A.1.5: Reagents involved in the preparation of a cDNA synthesis reaction

cDNA RT reaction preparation

10µl of 2X RT master mix was pipetted into each individual PCR tube along with 10µl of RNA sample into each tube, pipetting up and down two times to mix. An equal volume of RNA sample was added per reaction to ensure the same amount of cDNA was being used in each downstream experiment. The tubes were sealed and centrifuged briefly to spin down the contents and to eliminate any air bubbles. The tubes were placed on ice until they were loaded into the thermal cycler.

Performing Reverse Transcription

The thermal cycler conditions for cDNA synthesis were as follows:

Incubation at 25°C for 10min, 37°C for 120min, 85°C for 5sec followed by indefinite hold at 4°C and the reaction volume to was set to 20µl. cDNA was then stored at a temperature of -20°C.

A.1.6 TaqMan® PCR

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.

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 (Förster, 1948, 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.





-  = Nonfluorescent quencher
-  = Minor groove binder
-  = Reporter
-  = Hot-start DNA polymerase

Figure A.1.6(a): Legend for figures b-e

During PCR, the TaqMan MGB probe anneals specifically to a complementary sequence between the forward and reverse primer sites. When the probe is intact (Figures A.1.6(b), A.1.6(c)), the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence, primarily by Förster-type energy transfer (Förster, 1948; Lakowicz, 1983).

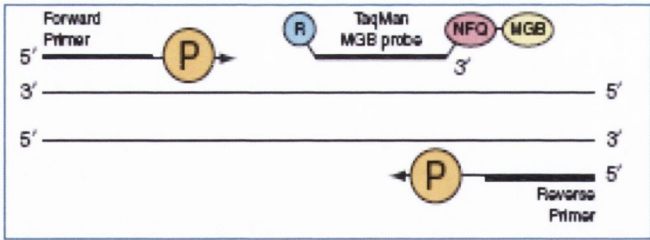


Figure A.1.6(b): Polymerization

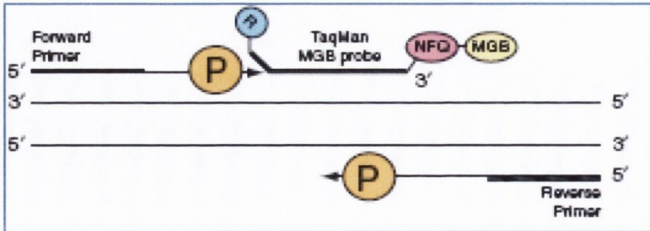


Figure A.1.6(c): Strand Displacement

The DNA polymerase cleaves only probes that are hybridized to the target (Figure A.1.6 (d)). Cleavage separates the reporter dye from the quencher dye; the separation of the reporter dye from the quencher dye results in increased fluorescence by the reporter. The increase in fluorescence occurs only if the target sequence is complementary to the probe and is amplified during PCR. Because of these requirements, non-specific amplification is not detected.

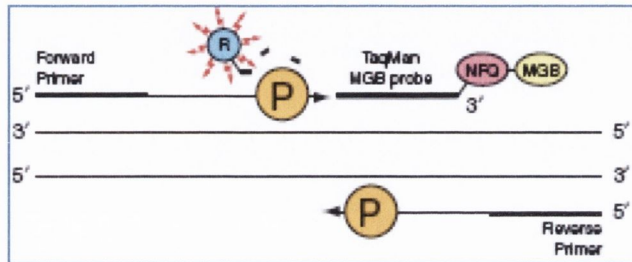


Figure A.1.6(d): Cleavage

Polymerization of the strand continues, but because the 3' end of the probe is blocked, no extension of the probe occurs during PCR.

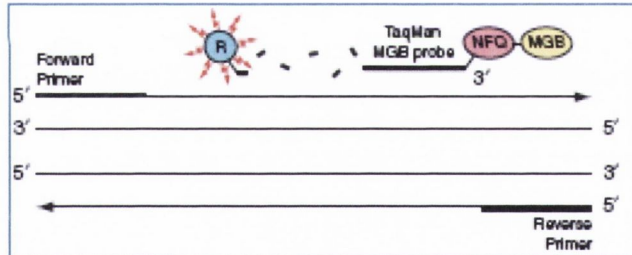


Figure A.1.6(e): Completion of Polymerization

A.1.6.1 TaqMan® SNP Genotyping and Allelic Discrimination Assay

TaqMan SNP Genotyping Assays are designed and optimized to work with TaqMan® Universal PCR Master Mix (with or without AmpErase® UNG) using the same thermal cycling conditions. SNP genotyping using this assay requires only three components:

1. 1 to 20 ng of purified gDNA sample per well
2. 20X, 40X, or 80X SNP Genotyping Assay (depending on product and assay scale)
3. TaqMan® Universal PCR Master Mix (with or without AmpErase® UNG)

It also requires only one PCR amplification step and an endpoint reading to obtain results. During the first step of a TaqMan SNP Genotyping Assay experiment, AmpliTaq Gold® DNA polymerase from the TaqMan Universal PCR Master Mix, No AmpErase UNG, amplifies target DNA using sequence-specific primers. TaqMan MGB probes from the SNP Genotyping Assay provide a fluorescence signal for the amplification of each allele.

The 20X, 40X, or 80X SNP Genotyping Assay contains sequence-specific forward and reverse primers to amplify the polymorphic sequence of interest and two TaqMan® MGB probes:

- One probe labelled with VIC® dye detects the Allele 1 sequence
- One probe labelled with FAM™ dye detects the Allele 2 sequence

Before use in a reaction it was recommended that SNP Genotyping Assay was diluted to a 20X working stock, then aliquoted for routine use. By having aliquots, it minimized freeze-thaw cycles and protected the SNP Genotyping Assay from exposure to light.

Diluting TaqMan® SNP Genotyping Assays:

40X SNP Genotyping Assay was diluted to a 20X working stock with 1X TE buffer. (The 1X TE buffer was 10 mM Tris-HCl, 1 Mm EDTA, pH 8.0 and made using DNase-free, sterile-filtered water). The assay was vortexed to mix and then centrifuged before storing multiple aliquots -20°C.

Preparing the reaction mix

The reaction mix was made from 20X SNP Genotyping Assay, TaqMan Universal PCR Master Mix, No AmpErase UNG (UMM) and DNase-free water. The recommended final reaction volume per well was 5µl for a 384-well plate and 25µl for a 96-well plate.

The number of reactions to be performed for each assay was calculated and it was recommended to include:

- At least two NTCs on each plate.
- If available, at least one known DNA control on each plate.

Each DNA sample was diluted with DNase-free water to deliver a final DNA mass in the range of 1 to 20ng per well as it was important that all wells belonging to the same assay contained the same amount of sample or control. Genomic DNA was delivered to the final

reaction mix. The 20XSNP Genotyping Assay briefly was vortexed and centrifuge briefly to mix. The required total volumes of UMM and 20XSNP Genotyping Assay were pipetted into a sterile microcentrifuge tube and was inverted several times to mix. The tube was then centrifuged briefly to spin down the contents and to eliminate any air bubbles from the reaction mix.

Adding the Reaction Mix

Into each well of the DNA reaction plate, the volume of reaction mix was pipette as indicated below: It was important to ensure that no cross-contamination occurred from well to well during pipetting. All the wells were inspected for uniformity of volume and wells that did not appear to contain the proper volume were noted. The plate was sealed with the appropriate cover and vortexed to mix the wells before being briefly spun down to eliminate any air bubbles.

Performing the PCR

The thermal cycling conditions were specified which were optimized for use only with TaqMan SNP Genotyping Assays on ABI instruments.

Standard Protocol		
AmpliTaq Gold Enzyme Activation	PCR (40 Cycles)	
HOLD	Denature	Anneal/Extend
10 min at 95 °C	15 sec at 92 °C	1 min at 60 °C

The reaction plate was loaded into the thermal cycler and the run was started.

Taqman® SNP genotyping/allelic discrimination assay

Detection of nucleotide mutations and polymorphisms is central to the modern science of molecular genetics. For example, allelic discrimination (AD) detects different forms of the same gene that differ by nucleotide substitution, insertion or deletion. Methods for mutation detection can be divided into two groups: scanning methods that can discover previously unknown nucleotide differences and diagnostic methods designed to detect specific, known mutations and polymorphisms. Large-scale scoring of known (single nucleotide polymorphisms) SNPs requires techniques with few steps and the ability to

automate each of these steps. In this regard, the 5' nuclease assay is ideal because it combines PCR amplification and detection into a single step. Figure A.1.6.1(a) demonstrates how fluorogenic probes and the 5' nuclease assay can be used for allelic discrimination. For a bi-allelic system, probes specific for each allele are included in the PCR assay. The probes can be distinguished because they are labelled with different fluorescent reporter dyes (typically FAM and VIC). A fully hybridised probe remains bound during strand displacement, resulting in efficient probe cleavage and release of the reporter dye. A mismatch between probe and target greatly reduces the efficiency of probe hybridisation and cleavage. Therefore, a substantial increase in either FAM or VIC fluorescence indicates homozygosity for the FAM- or VIC-specific allele. An increase in both signals indicates heterozygosity.

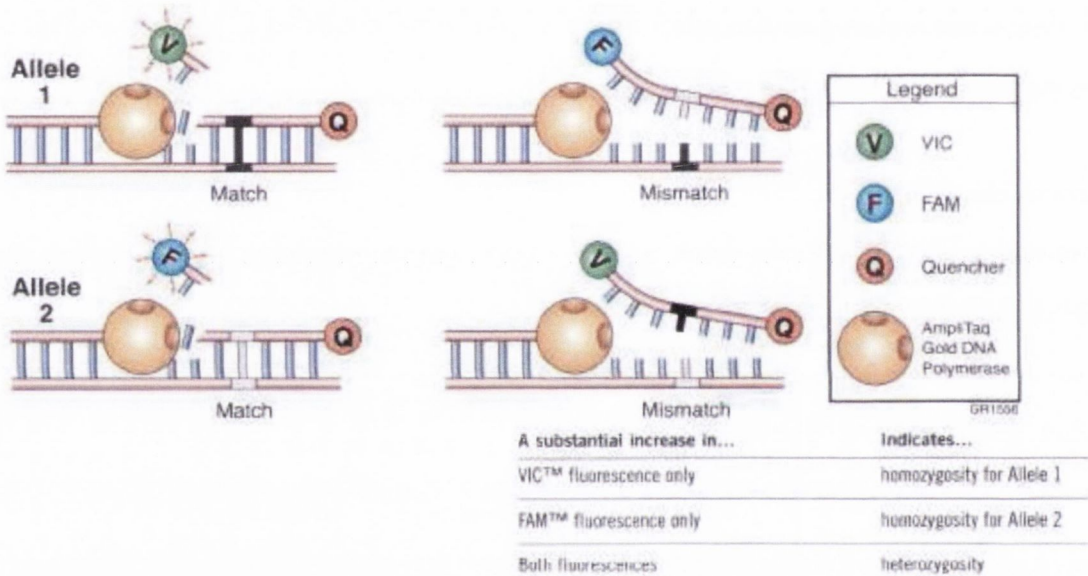


Figure A.1.6.1(a): Allelic discrimination using the 5' nuclease assay.

Three factors contribute to the discrimination based on a single mismatch:

1. The mismatch has a disruptive effect on hybridisation. A mismatched probe will have a lower T_m than a perfectly matched probe. Proper choice of annealing/extension temperature during PCR will favour hybridisation of an exact-match probe over a mismatched probe.
2. The assay is performed under competitive conditions with both probes present in the same reaction tube. Therefore, mismatched probes are prevented from binding due to stable binding of exact match probes.

3. The 5' end of the probe must start to be displaced before cleavage occurs. The 5' nuclease activity of Taq polymerase recognises a forked structure with a displaced 5' strand of at least 1 to 3 nucleotides (Lyamichev et al., 1993). Once a probe starts to be displaced, complete dissociation occurs faster with a mismatch than an exact match. This means there is less time for cleavage to occur with a mismatch probe. Thus, the presence of a mismatch promotes dissociation rather than cleavage of the probe. No template controls (NTC) and known samples for all 3 possible outcomes (heterozygote, homozygote allele 1 and 2) are included with each run for allele calling purposes. A generic graphical output from an AD run is shown in Figure 2.5.3(b). Primers/probes used for AD purposes were designed and purchased using the Custom TaqMan® SNP Genotyping Assays service.

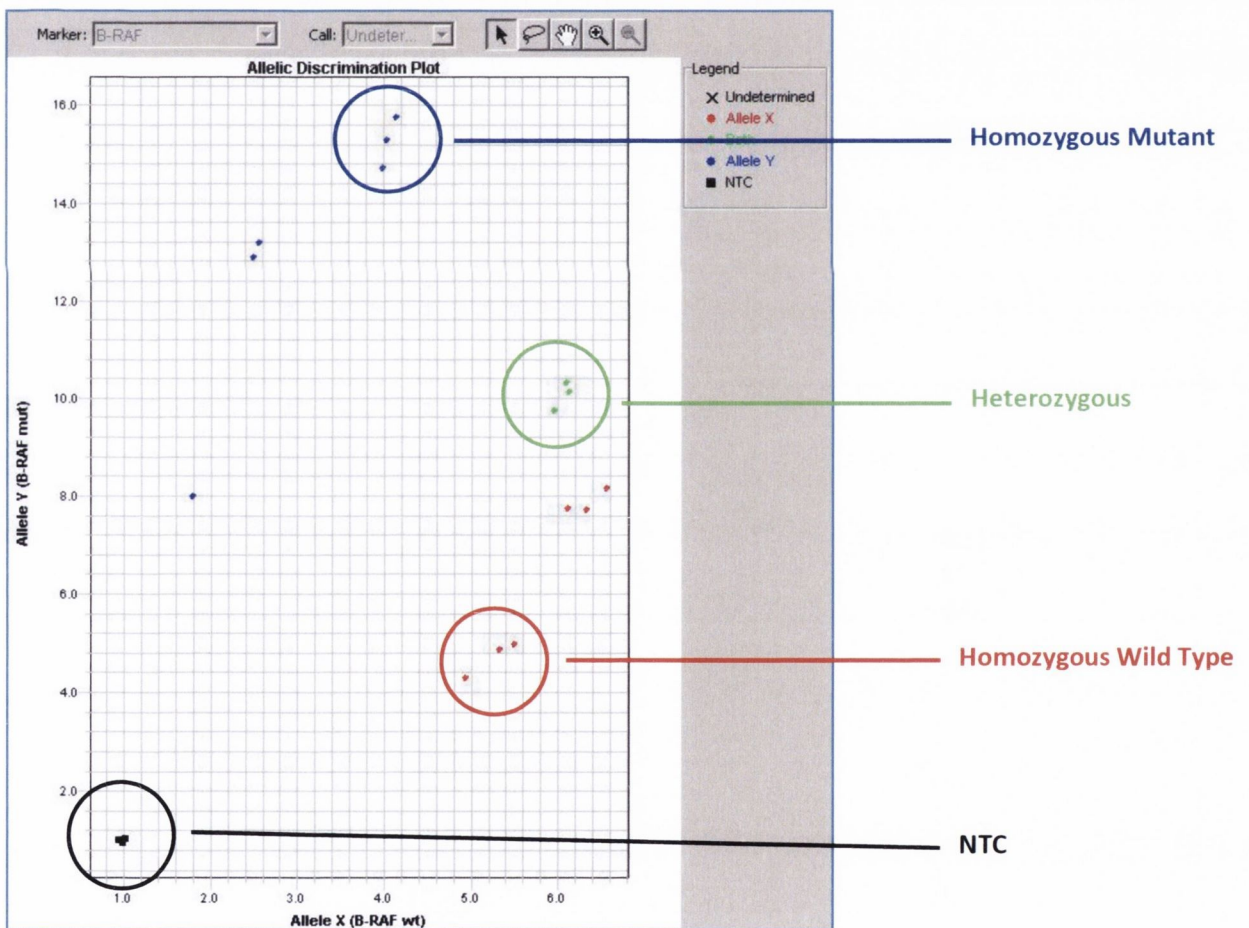


Figure A.1.6.1(b): Typical allelic discrimination output plot

The Custom TaqMan® SNP Genotyping Assay contained:

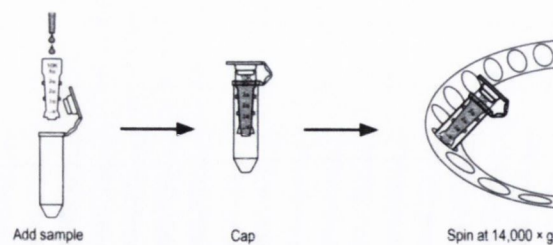
BRAF Forward Primer	5' CAT GAA GAC CTC ACA GTA AAA ATA GGT GAT 3'
BRAF Reverse Primer	5' GGA TCC AGA CAA CTG TTC AAA CTG A 3'
BRAF-P ^{WT}	VIC-5' CCA TCG AGA TTT CAC TGT AG 3'
BRAF-P ^{MUT}	FAM-5' CCA TCG AGA TTT CTC TGT AG 3'

A.1.7 Bisulfite Conversion

5µl methylSEQr™ Denaturation Buffer and 700ng of genomic DNA in 45µl deionized water were combined in a labelled microcentrifuge tube and mixed. The tube was incubated at 37 °C for at least 15 minutes while the conversion reagent was prepared. To make up the methylSEQr™ Conversion Reagent, 750µl deionized water and 210µl methylSEQr™ Denaturation Buffer was added to one tube of methylSEQr Conversion Reagent. The tube was vortexed for 1 minute and let rest for 2 minutes. This step was repeated 5 times to ensure the conversion reagent was now a homogenous mixture and suitable to add to genomic DNA samples. The prepared methylSEQr Conversion Reagent was protected from light and used within 1 hour of preparation. Denatured DNA was removed from 37 °C incubation and 100µl of the conversion reagent was added to each sample to give a final volume of 150µl. The samples were then incubated in the dark at 50 °C for 16 hours.

Purifying the sample

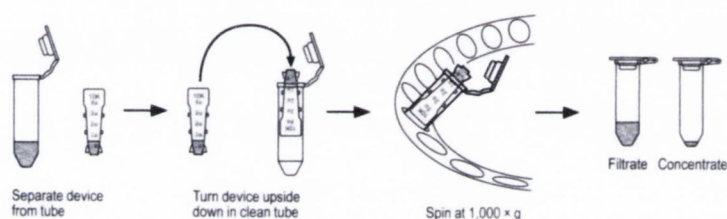
250µl of deionized water was added to an assembled methylSEQr™ Purification Column followed by the 150-µl incubated sample and was mixed carefully by pipetting up and down. The sample was centrifuged at a maximum of 14000 × g for 10 minutes and the filtrate was then discarded.



350µl of deionized water was added to the upper chamber of the purification column filter and centrifuge at $14000 \times g$ for 10 minutes, before discarding the filtrate. This step was then repeated to ensure adequate washing of the spin column membrane.

350µl of 0.1 M sodium hydroxide was then added to the upper chamber of the purification column filter and allowed to sit for 5 minutes before being centrifuged at $14000 \times g$ for 10 minutes and then discarding the filtrate. 350µl of deionized water was added to the upper chamber, and then centrifuged at a maximum of $14000 \times g$ for 5 minutes, until the membrane was just damp, or a small amount of liquid remained in the upper chamber. Finally 50µl of 1XTE buffer was added to the upper chamber and mixed gently by pipetting up and down and allowed to stand in the column for 5 minutes.

The column was then inverted as shown below and the bisulfite-treated genomic DNA was collected in a clean microcentrifuge tube by centrifuging at $1000 \times g$ for 30 seconds.



Bisulfite converted DNA sample were then quantified using the single stranded (ssDNA) parameters on a Nano-Drop® 1000 spectrophotometer to give a crude concentration measurement needed for downstream PCR experiments. Aliquots of converted DNA samples were then stored at -20°C to ensure one individual sample was not being subjected to freeze thaw cycles.

Appendix II
Supplementary Figures

A.2.1 Small RNA Sample Analysis

A.2.1.1 Melanoma

SK-Mel 28 BRAF V600E Mutant

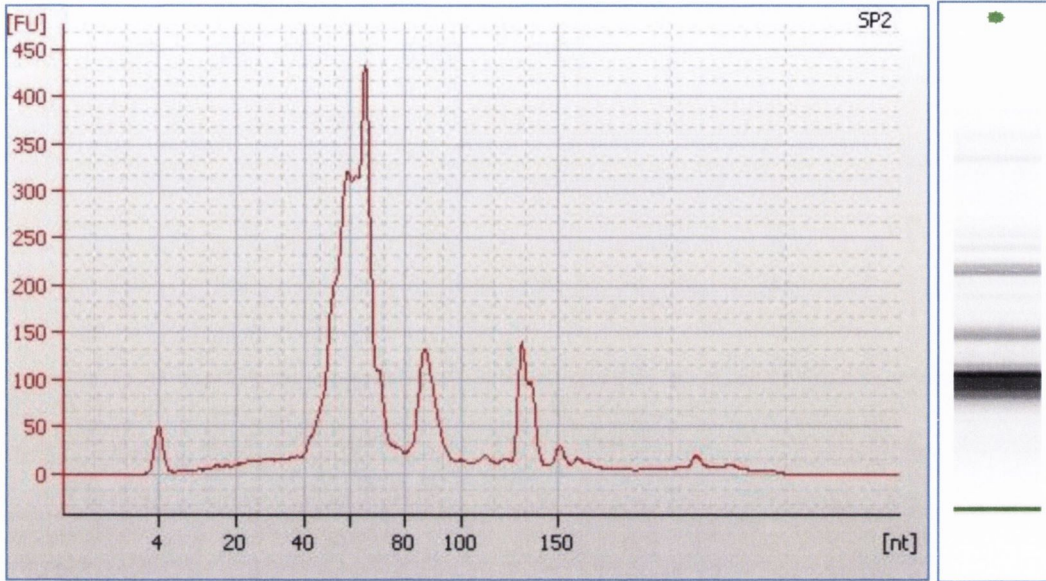


Figure A.2.1.1(a): Bioanalyser Analysis of SK-Mel 28 Parent Small RNA

miRNA Percentage – 15%

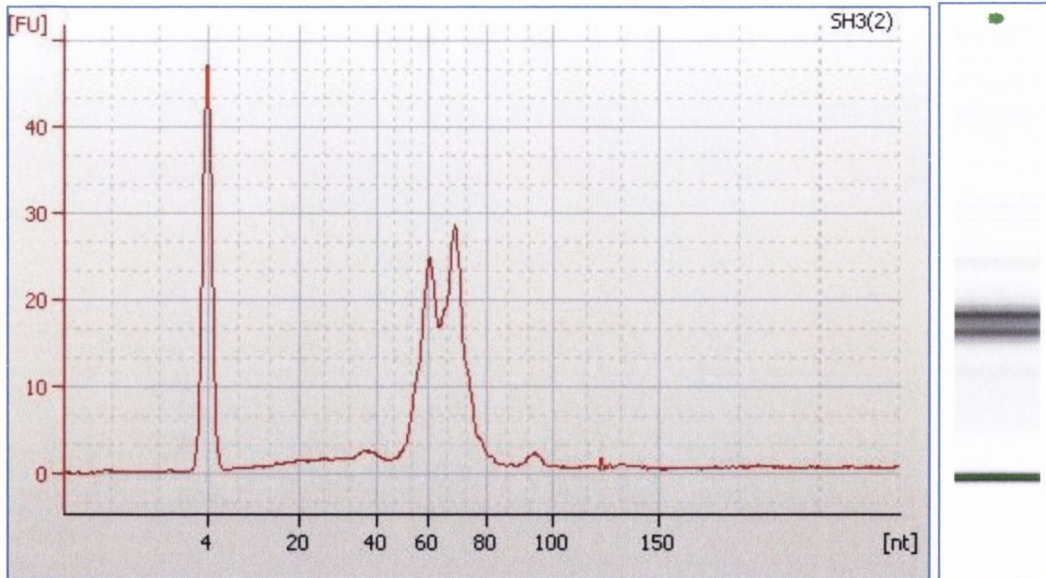


Figure A.2.1.1(b): Bioanalyser Analysis of SK-Mel 28 Holoclone Small RNA

miRNA Percentage – 29%

COLO794 BRAF Wild Type

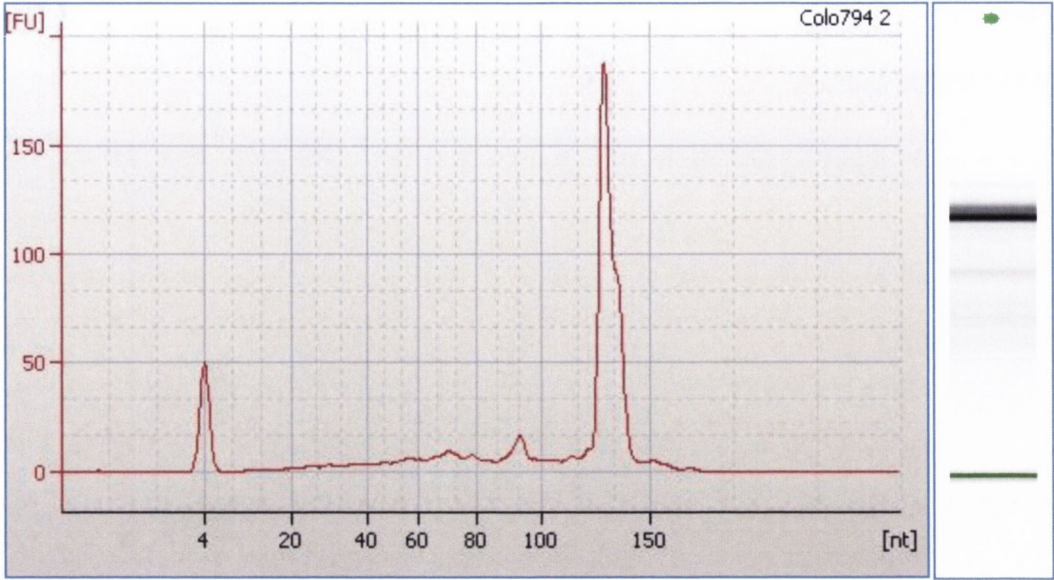


Figure A.2.1.1(c): Bioanalyser Analysis of COLO794 Parent Small RNA
miRNA Percentage – 10%

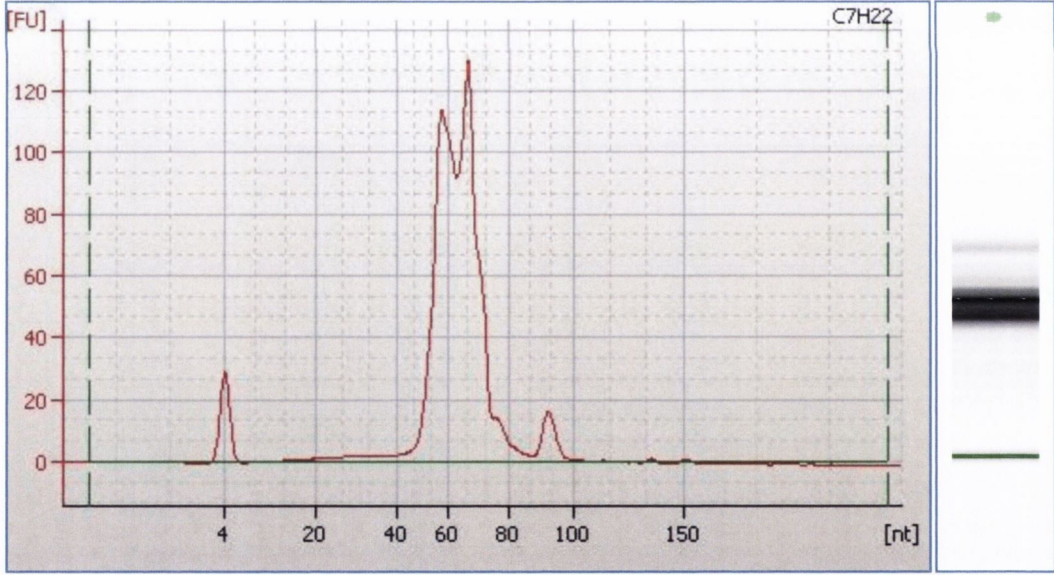


Figure A.2.1.1(d): Bioanalyser Analysis of COLO794 Holoclone Small RNA
miRNA Percentage – 5%

A.2.1.2 Thyroid

8505C BRAF V600E Mutant

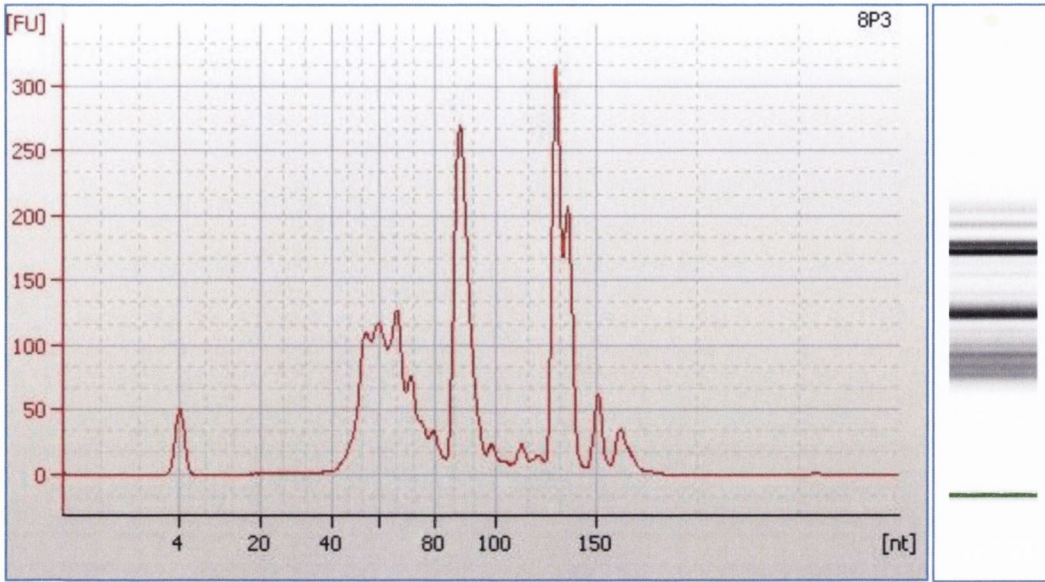


Figure A.2.1.2(a): Bioanalyser Analysis of 8505C Parent Small RNA

miRNA Percentage – 11%

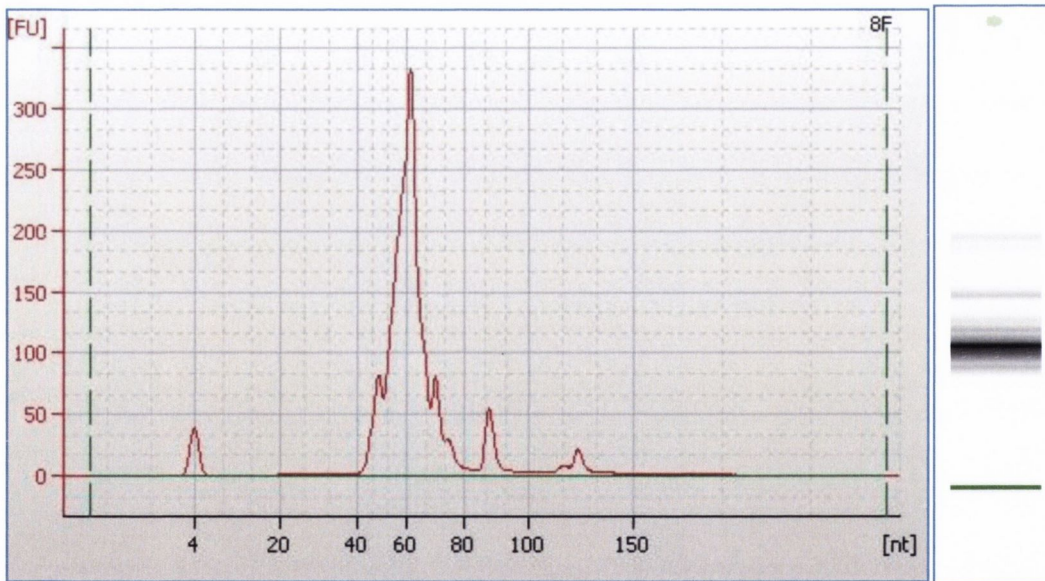


Figure A.2.1.2(b): Bioanalyser Analysis of 8505C Holoclone Small RNA

miRNA Percentage – 3%

TPC-1 BRAF Wild Type

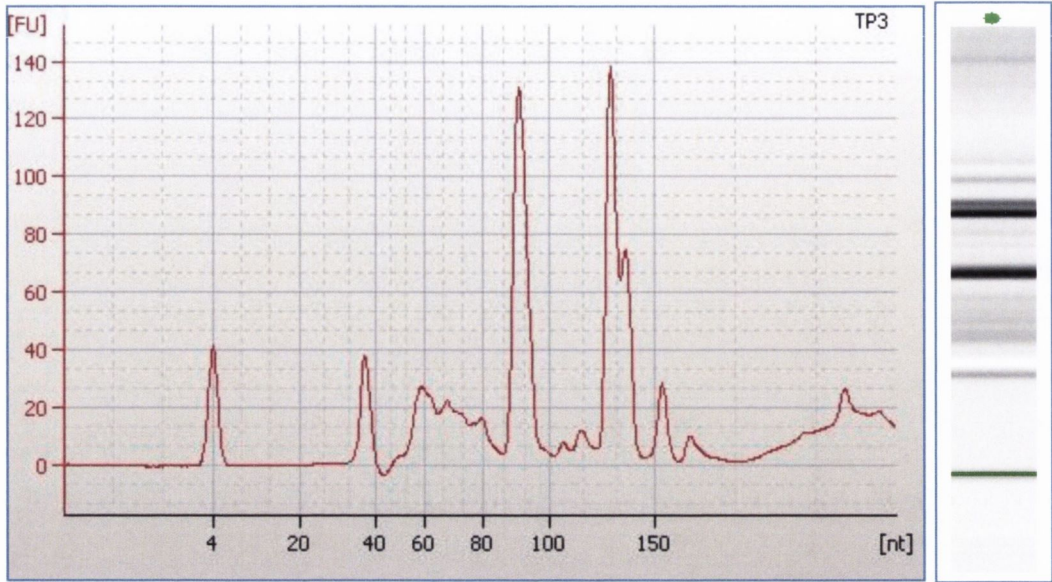


Figure A.2.1.2(c): Bioanalyser Analysis of TPC-1 Parent Small RNA

miRNA Percentage – 11%

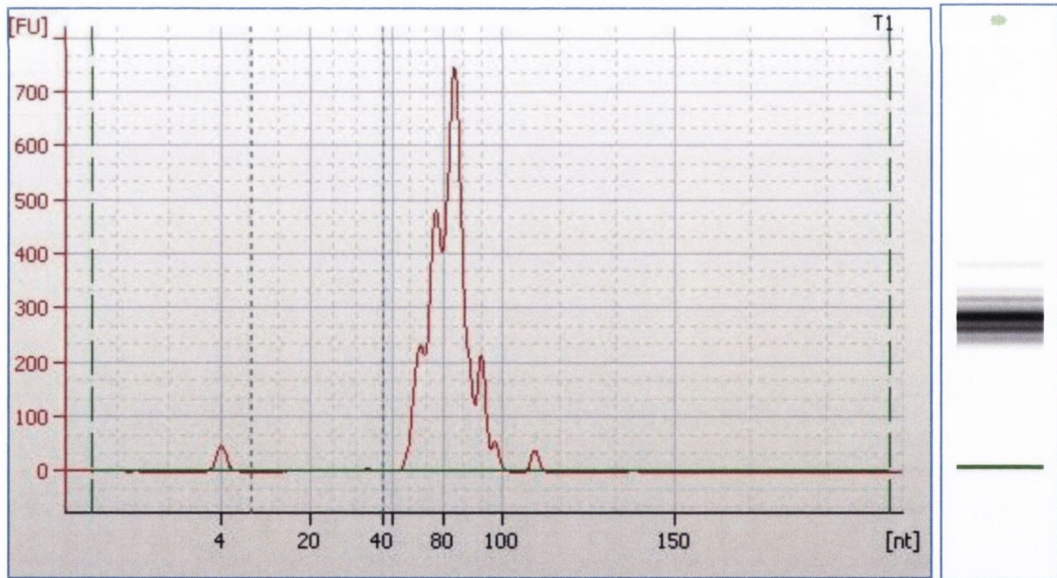


Figure A.2.1.2(d): Bioanalyser Analysis of TPC-1 Holoclone Small RNA

miRNA Percentage – 4%

A.2.1.3 Ovary

ES-2 BRAF V600E Mutant

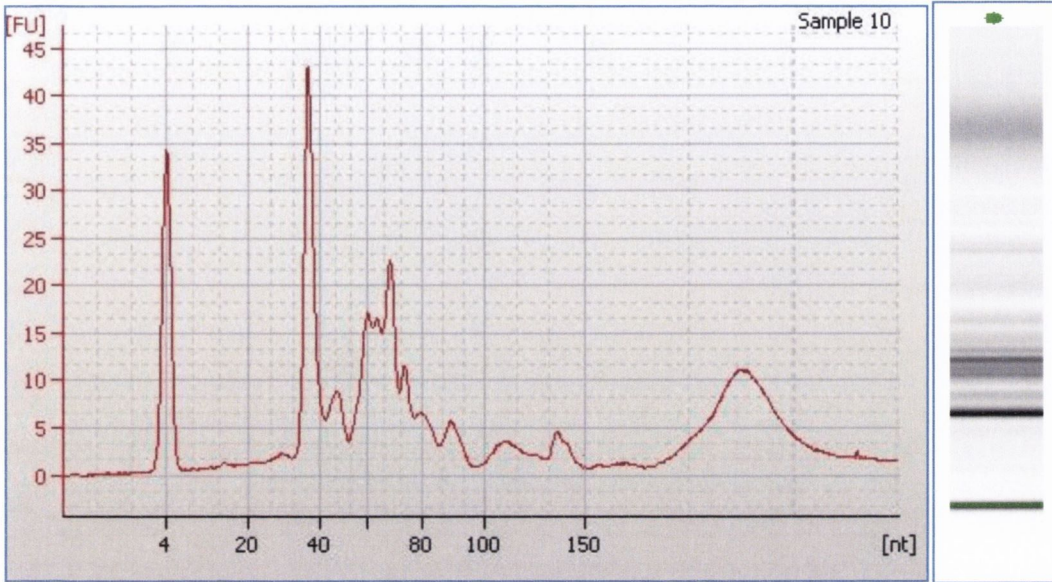


Figure A.2.1.3(a): Bioanalyser Analysis of ES-2 Parent Small RNA

miRNA Percentage – 10%

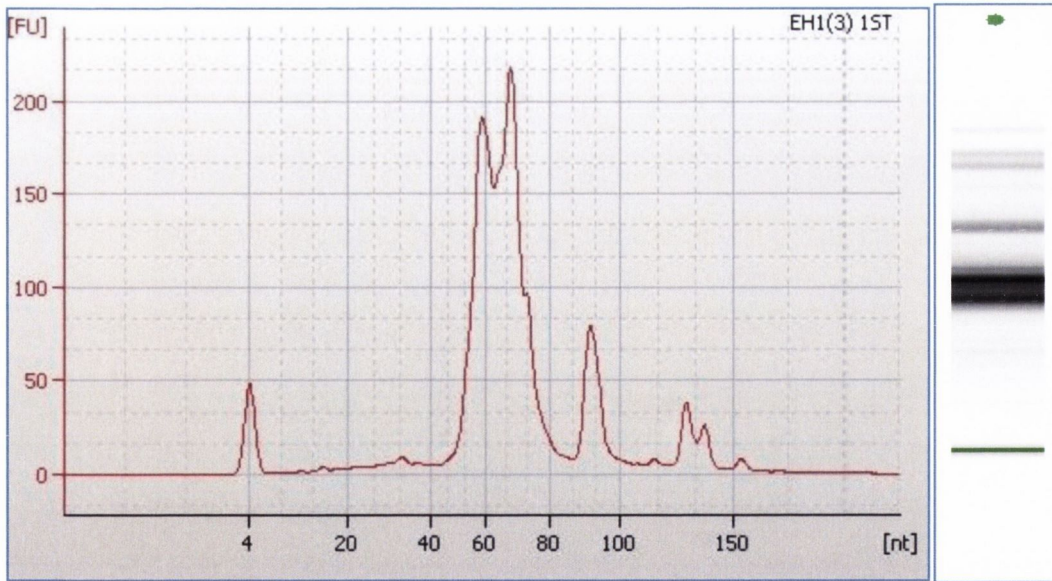


Figure A.2.1.3(b): Bioanalyser Analysis of ES-2 Holoclone Small RNA

miRNA Percentage – 12%

A2780 BRAF Wild Type

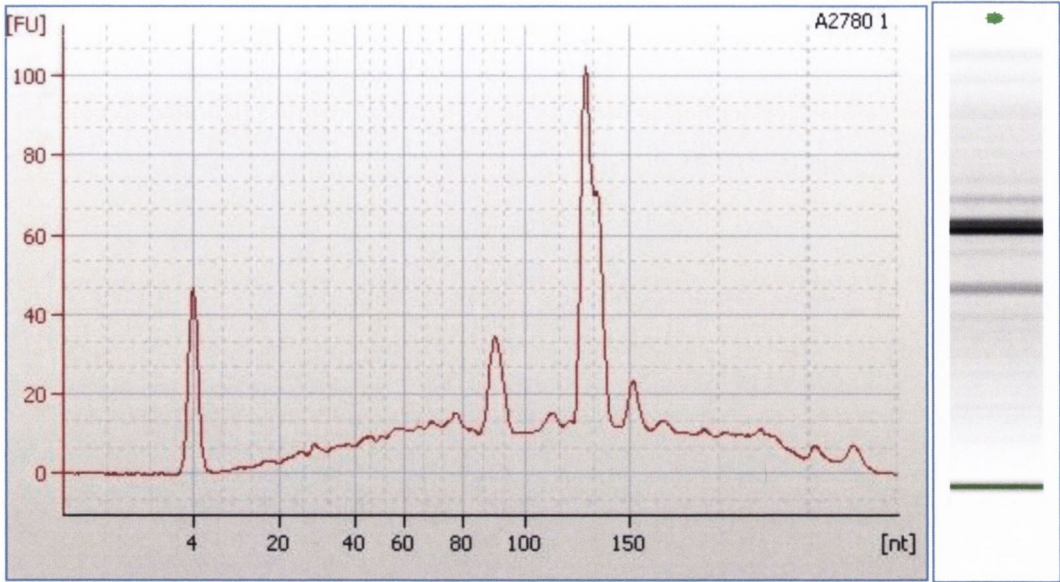


Figure A.2.1.3(c): Bioanalyser Analysis of A2780 Parent Small RNA

miRNA Percentage – 16%

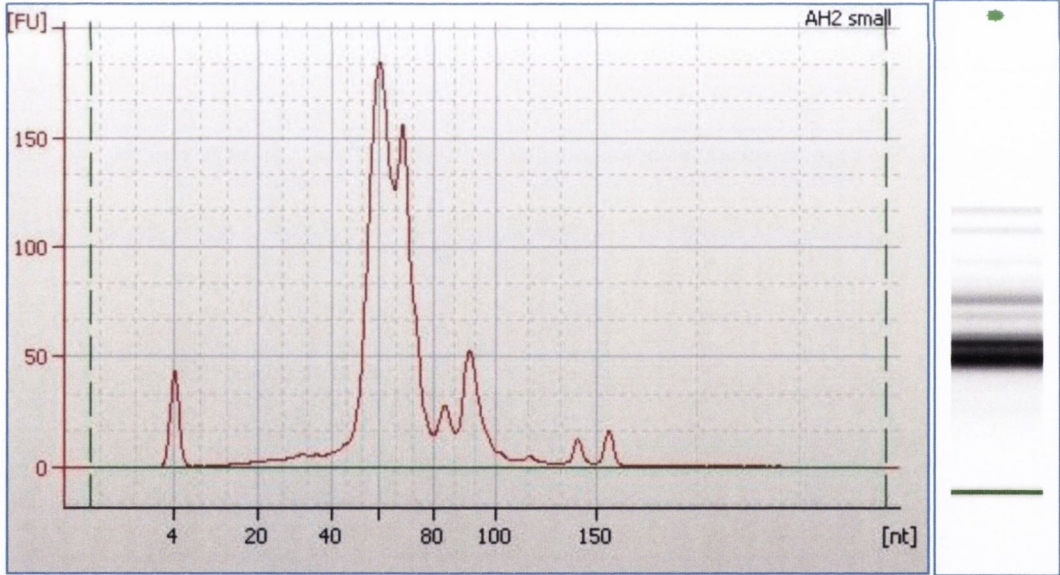


Figure A.2.1.3(d): Bioanalyser Analysis of A2780 Holoclone Small RNA

miRNA Percentage – 10%

A.2.1.4 Colorectal

HT-29 BRAF V600E Mutant

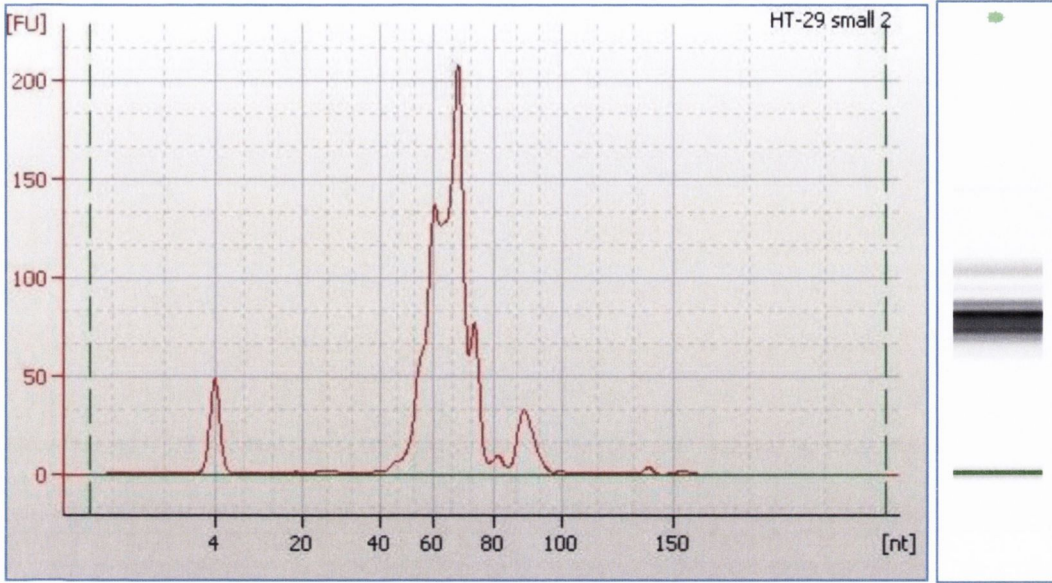


Figure A.2.1.4(a): Bioanalyser Analysis of HT-29 Parent Small RNA

miRNA Percentage – 6%

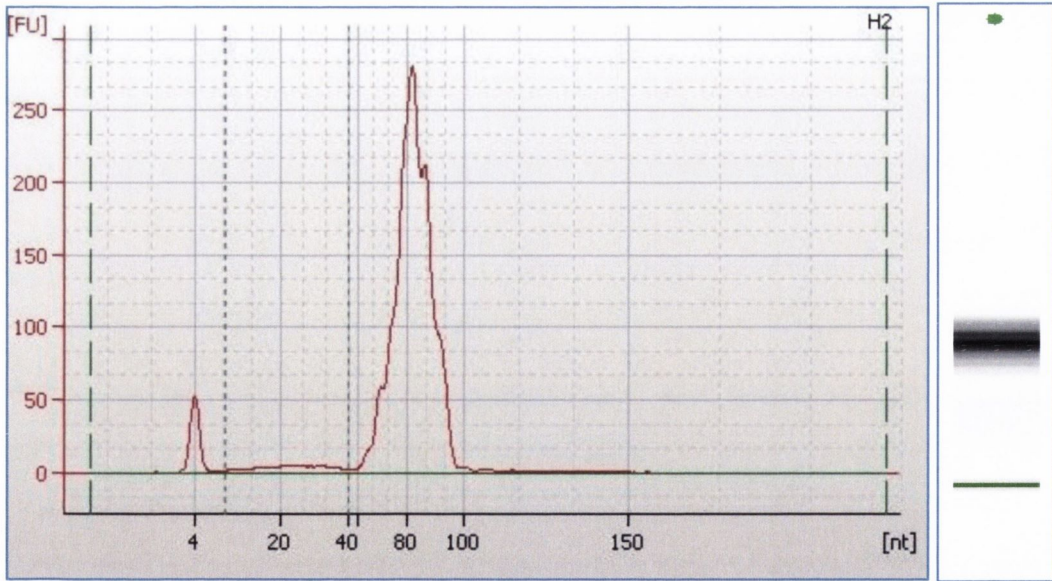


Figure A.2.1.4(b): Bioanalyser Analysis of HT-29 Holoclone Small RNA

miRNA Percentage – 17%

COLO320 BRAF Wild Type

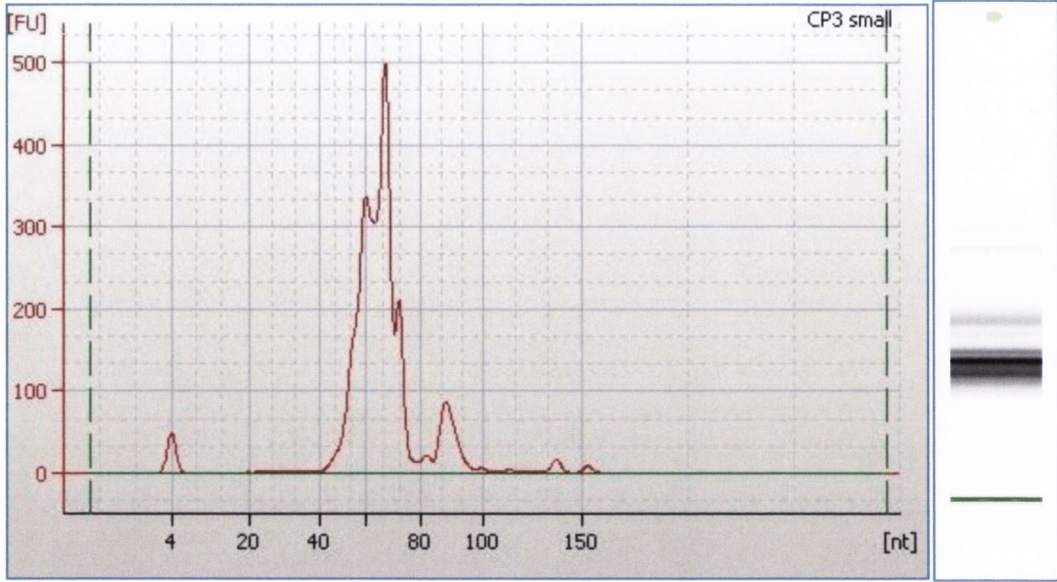


Figure A.2.1.4(c): Bioanalyser Analysis of COLO320 Parent Small RNA

miRNA Percentage – 3%

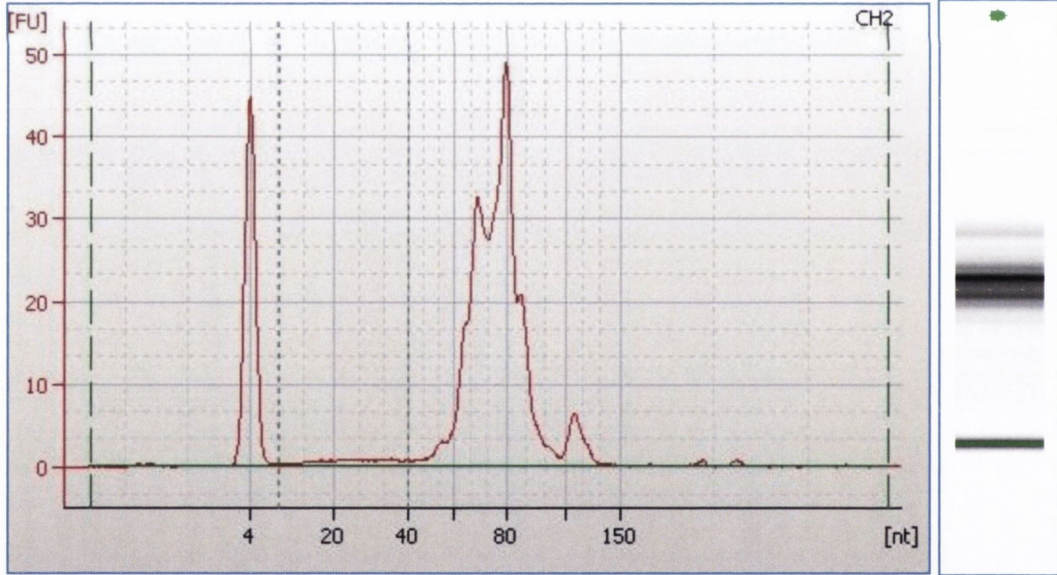


Figure A.2.1.4(d): Bioanalyser Analysis of COLO320 Holoclone Small RNA

miRNA Percentage – 12%

A.2.2 Bioanalyser Smear Analysis of Small RNA Sequencing Libraries

A.2.2.1 Melanoma

SK-Mel 28 BRAF V600E Mutant Parent

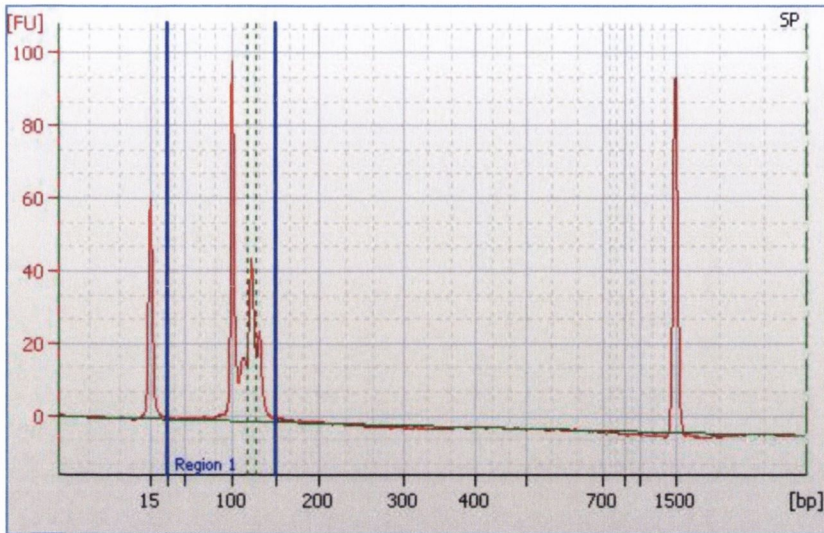


Figure A.2.2.1(a): Agilent Bioanalyser Smear Analysis of SK-Mel 28 Parent Small RNA Library. SK-Mel 28 parent small RNA library had a 117-127bp region area of 36.6 and a 25-150bp region area of 135.8 resulting in an amplified miRNA library percentage of 26%

SK-Mel 28 BRAF V600E Mutant Holoclone

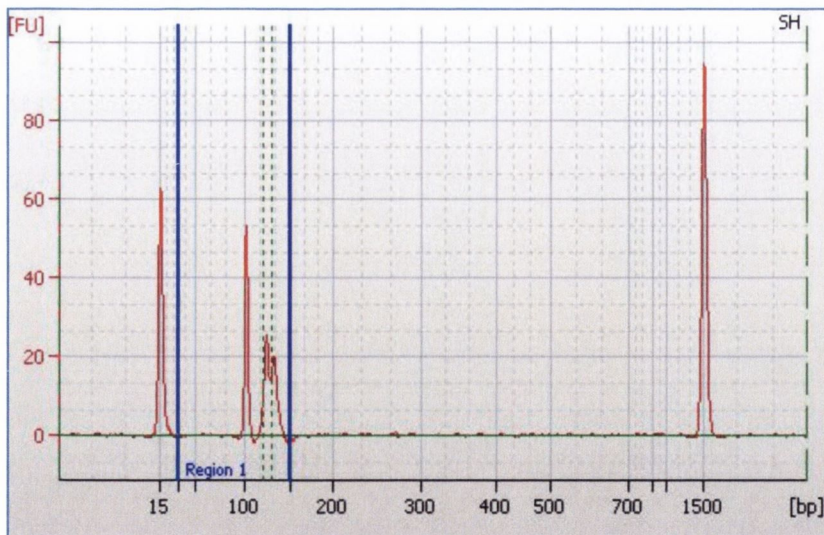


Figure A.2.2.1(b): Agilent Bioanalyser Smear Analysis of SK-Mel 28 Holoclone Small RNA Library. SK-Mel 28 holoclone small RNA library had a 122-132bp region area of 19.9 and a 25-150bp region area of 64.6 resulting in an amplified miRNA library percentage of 31%.

COLO794 BRAF Wild Type Parent

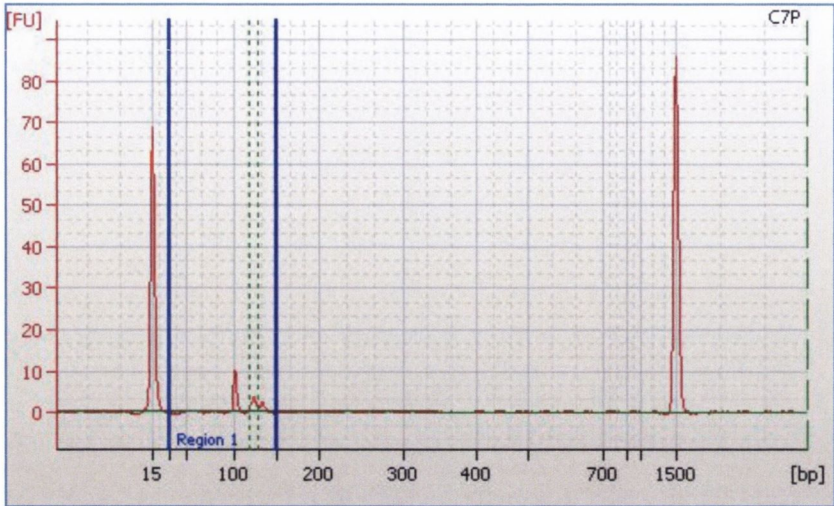


Figure A.2.2.1(c): Agilent Bioanalyser Smear Analysis of COLO794 Parent Small RNA Library. COLO794 parent small RNA library had a 118-128bp region area of 2.6 and a 25-150bp region area of 9.3 resulting in an amplified miRNA library percentage of 28%

COLO794 BRAF Wild Type Holoclone

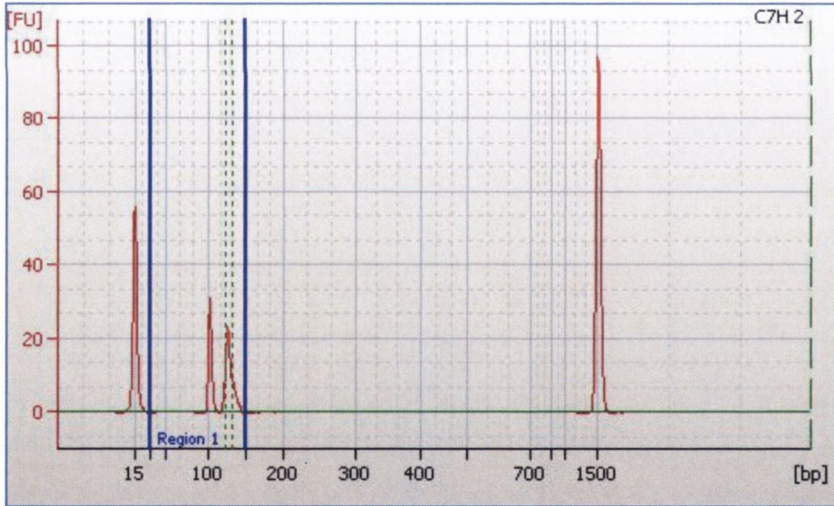


Figure A.2.2.1(d): Agilent Bioanalyser Smear Analysis of COLO794 Holoclone Small RNA Library. COLO794 holoclone small RNA library had a 122-132bp region of 16.1 and a 25-150bp region area of 36.8 resulting in an amplified miRNA library percentage of 43%

A.2.2.2 Thyroid

8505C BRAF V600E Mutant Parent

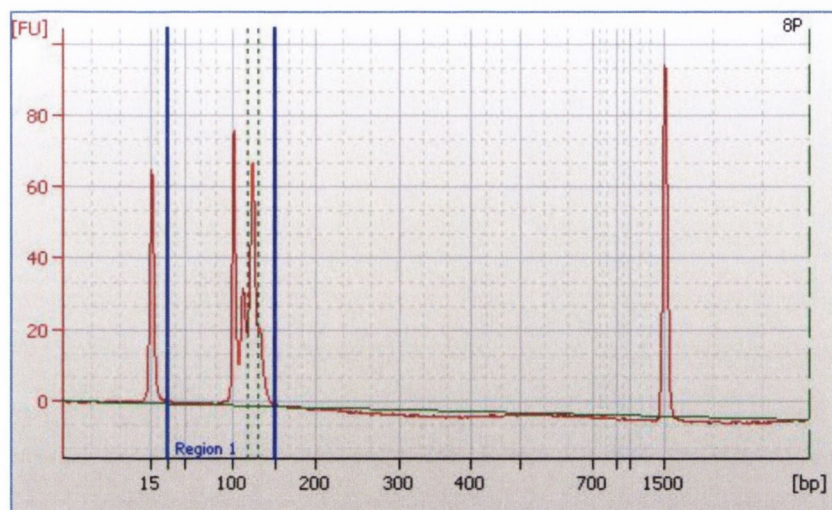


Figure A.2.2.2(a): Agilent Bioanalyser Smear Analysis of 8505C Parent Small RNA Library
8505C parent small RNA library had a 119-129bp region area of 53.6 and a 25-150bp region area of 151.2 resulting in an amplified miRNA library percentage of 35%

8505C BRAF V600E Mutant Holoclone

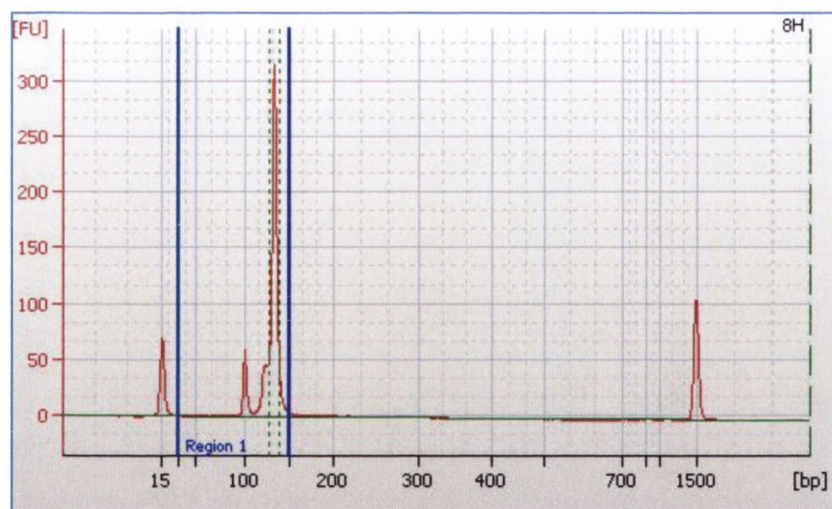


Figure A.2.2.2(b): Agilent Bioanalyser Smear Analysis of 8505C Holoclone Small RNA Library
8505C holoclone small RNA library had a 118-128bp region area of 214.8 and a 25-150bp region area of 324.3 resulting in an amplified miRNA library percentage of 66%

TPC-1 BRAF Wild Type Parent

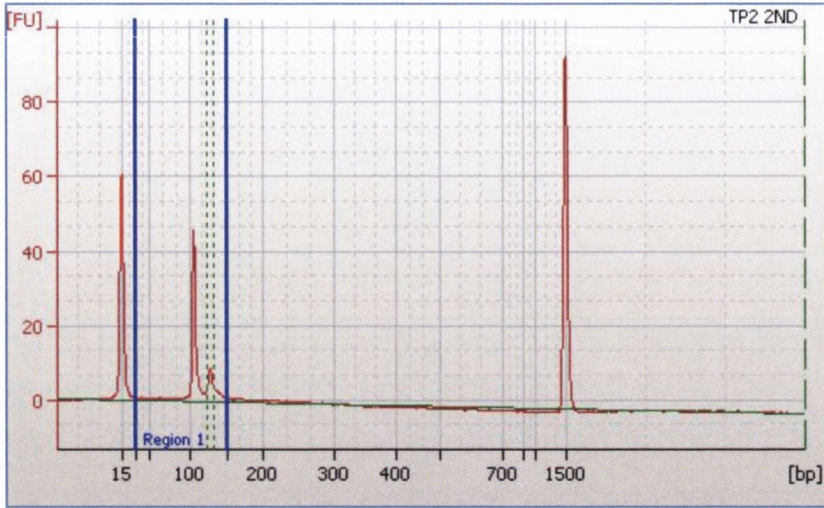


Figure A.2.2.2 (c): Agilent Bioanalyser Smear Analysis of TPC-1 Parent Small RNA Library
TPC-1 parent small RNA library had a 121-131bp region area of 6.0 and a 25-150bp region area of 41.9 resulting in an amplified miRNA library percentage of 14%

TPC-1 BRAF Wild Type Holoclone

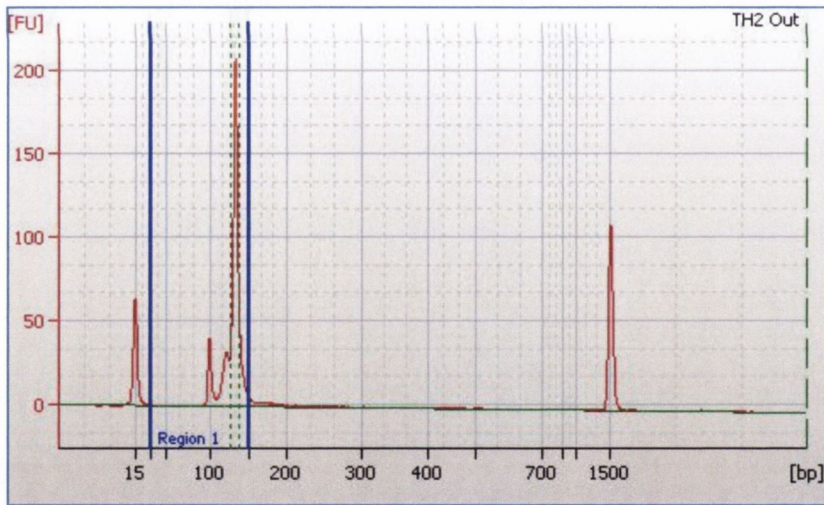


Figure A.2.2.2(d): Agilent Bioanalyser Smear Analysis of TPC-1 Holoclone Small RNA Library
TPC-1 holoclone small RNA library had a 128-138bp region area of 136.4 and a 25-150bp region area of 224.8 resulting in an amplified miRNA library percentage of 60%

A.2.2.3 Ovarian

ES-2 BRAF V600E Mutant Parent

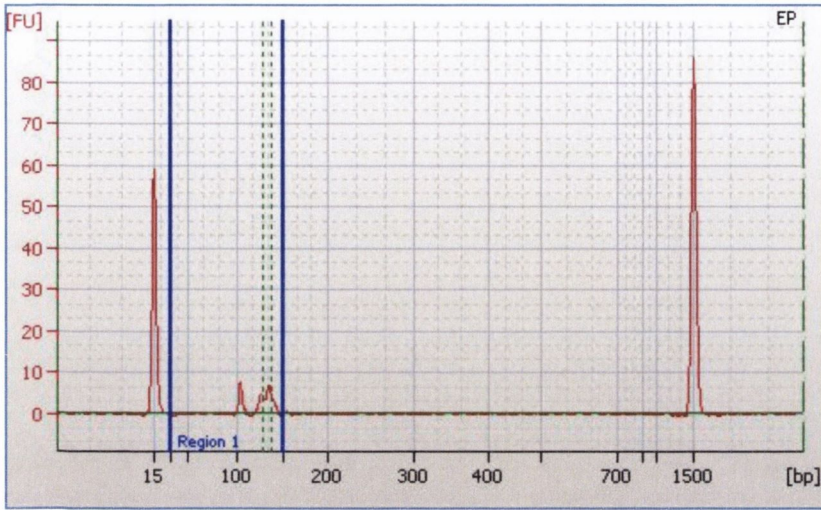


Figure A.2.2.3(a): Agilent Bioanalyser Smear Analysis of ES-2 Parent Small RNA Library

ES-2 parent small RNA library had a 127-137bp region area of 5.7 and a 25-150bp region area of 14.0 resulting in an amplified miRNA library percentage of 41%

ES-2 BRAF V600E Mutant Holoclone

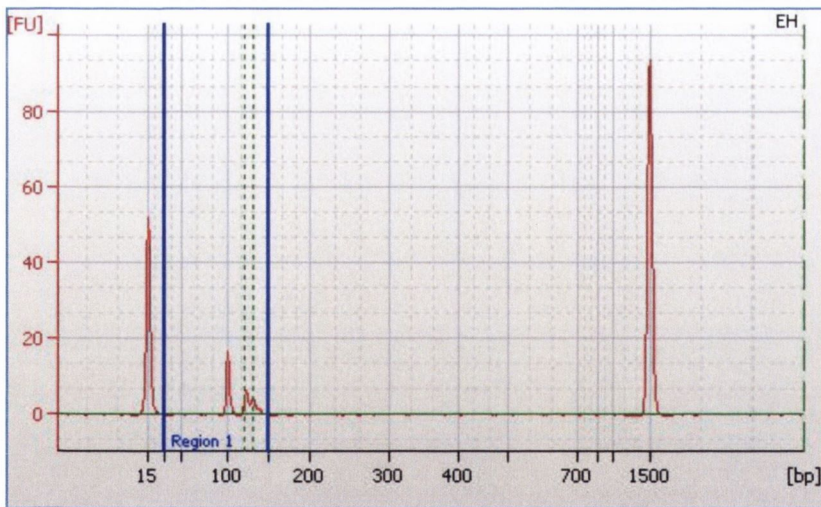


Figure A.2.2.3(b): Agilent Bioanalyser Smear Analysis of ES-2 Holoclone Small RNA Library

ES-2 holoclone small RNA library had a 120-130bp region area of 4.1 and a 25-150bp region area of 15 resulting in an amplified miRNA library percentage of 27%

A2780 BRAF Wild Type Parent

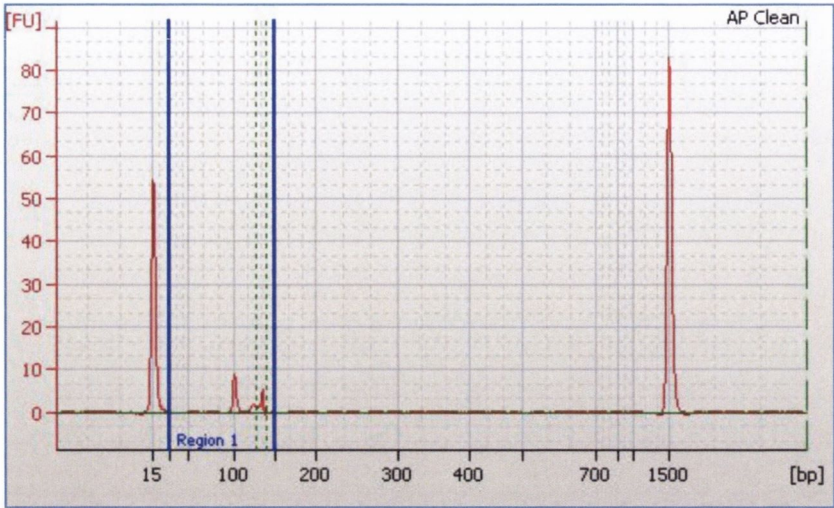


Figure A.2.2.3(c): Agilent Bioanalyser Smear Analysis of A2780 Parent Small RNA Library
A2780 parent small RNA library had a 128-138bp region area of 2.2 and a 25-150bp region area of 9.1 resulting in an amplified miRNA library percentage of 24%

A2780 BRAF Wild Type Holoclone

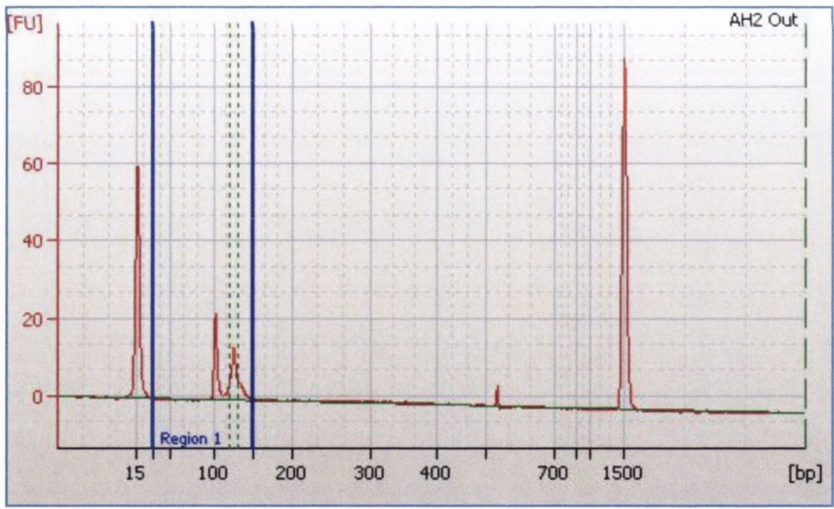


Figure A.2.2.3(d): Agilent Bioanalyser Smear Analysis of A2780 Holoclone Small RNA Library
A2780 holoclone small RNA library had a 128-138bp region area of 9.7 and a 25-150bp region area of 27.2 resulting in an amplified miRNA library percentage of 36%

A.2.2.4 Colorectal

HT-29 BRAF V600E Mutant Parent

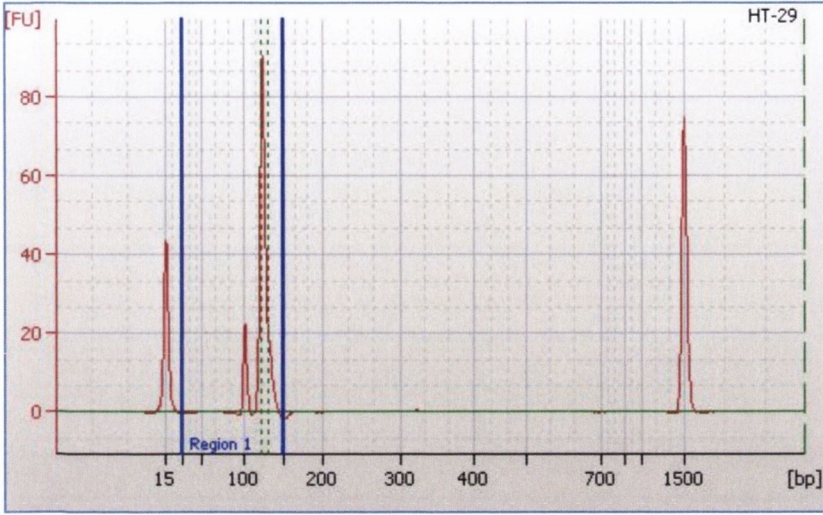


Figure A.2.2.4(a): Agilent Bioanalyser Smear Analysis of HT-29 Parent Small RNA Library

HT-29 parent small RNA library had a 120-130bp region area of 58.2 and a 25-150bp region area of 91.1 resulting in an amplified miRNA library percentage of 63%

HT-29 BRAF V600E Mutant Holoclone

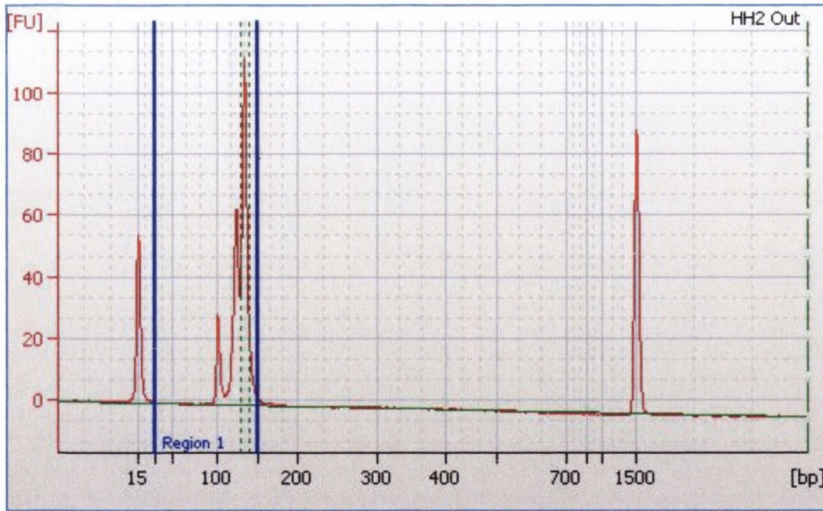


Figure A.2.2.4(b): Agilent Bioanalyser Smear Analysis of HT-29 Holoclone Small RNA Library

HT-29 holoclone small RNA library had a 128-130bp region area of 77.5 and a 25-150bp region area of 163.1 resulting in an amplified miRNA library percentage of 48%

COLO320 BRAF Wild Type Parent

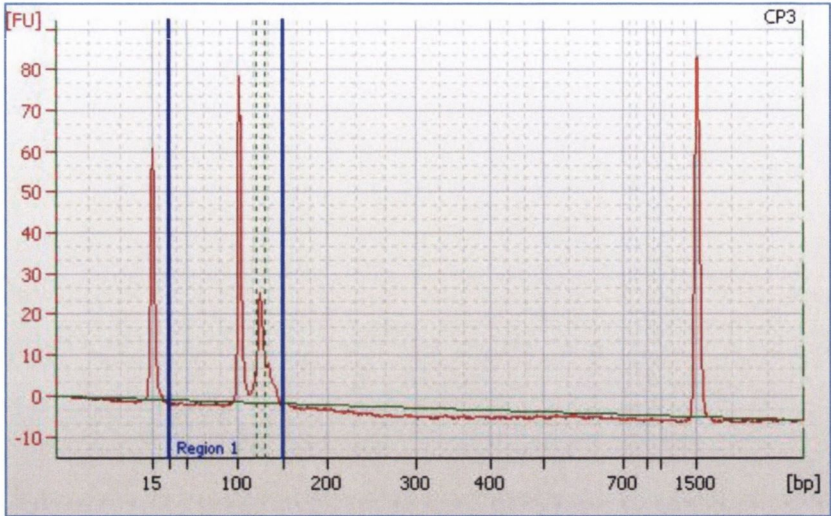


Figure A.2.2.4(c): Agilent Bioanalyser Smear Analysis of COLO320 Parent Small RNA Library. COLO320 parent small RNA library had a 128-138bp region area of 22.4 and a 25-150bp region area of 80.4 resulting in an amplified miRNA library percentage of 28%

COLO320 BRAF Wild Type Holoclone

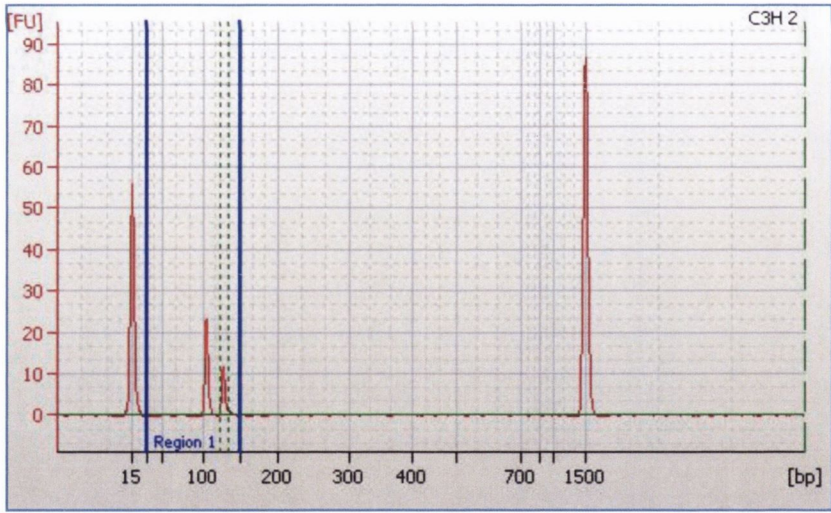


Figure A.2.2.4(d): Agilent Bioanalyser Smear Analysis of COLO320 Holoclone Small RNA Library. COLO320 holoclone small RNA library had a 122-132bp region area of 5.7 and a 25-150bp region area of 17.7 resulting in an amplified miRNA library percentage of 32%

Appendix III
Supplementary SOLiD™ Data

A.3.1 Holoclone Differentially Expressed miRNA Lists

All differentially expressed miRNAs in this section are displayed as Log₂ fold changes.

A.3.1.1 Melanoma

SK-Mel 28 BRAF V600E Mutant

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIR324	17	0.0	221.6	inf
MIR330	19	0.0	35.2	inf
MIR326	11	0.0	29.8	inf
MIR148B	12	0.0	97.2	inf
MIR365A	16	0.0	101.4	inf
MIR193A	17	0.0	37.2	inf
MIR512-1	19	0.0	22.6	inf
MIR183	7	0.0	45.4	inf
MIR32	9	0.0	39.6	inf
MIR451A	17	0.0	43.0	inf
MIR9-3	15	0.0	111.6	inf
MIR33A	22	0.0	30.2	inf
MIR296	20	0.0	115.6	inf
MIR628	15	0.0	116.2	inf
MIR320A	8	0.0	349.4	inf
MIR338	17	0.0	33.0	inf
MIR1180	17	0.0	27.8	inf
AC118345.2	2	0.0	1322.9	inf
MIRLET7G	3	52.2	1138.1	4.45
MIR211	15	339.4	6511.3	4.26
MIR29A	7	1331.6	14710.8	3.47
MIR26B	2	78.3	576.2	2.88
MIRLET7A2	11	313.3	2152.3	2.78
MIR148A	7	261.1	1686.7	2.69
MIRLET7E	19	26.1	156.8	2.59
MIRLET7I	12	26.1	148.8	2.51
MIR551B	3	391.7	2012.3	2.36
MIR423	17	52.2	253.6	2.28
MIR423	17	52.2	253.6	2.28
MIRLET7F1	9	208.9	987.7	2.24
MIR34A	1	574.4	2668.1	2.22
MIR30D	8	417.8	1757.7	2.07
MIR340	5	26.1	103.2	1.98
MIR15A	13	52.2	196.8	1.91

MIR26A2	12	391.7	1439.7	1.88
MIR27A	19	574.4	1875.1	1.71
MIR93	7	261.1	794.4	1.61
MIR197	1	130.6	386.6	1.57
MIR27B	9	104.4	303.2	1.54
MIR29B2	1	783.3	1929.1	1.30
MIR186	1	78.3	192.8	1.30
MIR101-1	1	156.7	378.8	1.27
MIRLET7A1	9	26.1	57.2	1.13
MIR125B2	21	26.1	57.2	1.13
MIR204	9	705.0	1472.7	1.06
MIR138-2	16	182.8	380.2	1.06
MIR16-2	3	182.8	378.0	1.05
MIR590	7	52.2	105.8	1.02
MIR21	17	15640.2	7211.4	-1.12
MIR145	5	26.1	11.6	-1.17
MIR1301	2	156.7	69.2	-1.18
MIR100	11	1175.0	507.2	-1.21
MIR125B1	11	3498.8	1478.9	-1.24
MIR598	8	26.1	10.2	-1.36
MIR769	19	26.1	8.8	-1.57
MIR218-1	4	208.9	68.4	-1.61
MIR192	11	339.4	110.4	-1.62
MIR181A2	9	208.9	64.4	-1.70
MIR455	9	52.2	14.4	-1.86
MIR181B2	9	130.6	33.0	-1.98
MIR103A2	20	22350.6	4848.9	-2.21
MIR1226	3	52.2	11.2	-2.22
MIR549	15	26.1	5.2	-2.33
MIR215	1	26.1	4.8	-2.44
MIR519D	19	26.1	3.8	-2.78
MIR139	11	26.1	3.0	-3.12
MIR30C2	6	208.9	21.0	-3.31
MIR199A1	19	130.6	11.4	-3.52
MIR149	2	78.3	6.4	-3.61
MIR520G	19	26.1	2.0	-3.71
MIR516B2	19	26.1	2.0	-3.71
MIR205	1	104.4	4.4	-4.57
MIR302A	4	208.9	4.6	-5.51
MIR371A	19	313.3	4.8	-6.03
MIR302B	4	2846.0	30.6	-6.54
MIR143	5	652.8	5.0	-7.03

MIR302C	4	626.7	4.6	-7.09
MIR302D	4	2193.3	8.0	-8.10

Table A.3.1.1(a): Differentially Expressed miRNAs in BRAF V600E mutated SK-Mel 28

Holoclones

COLO794 BRAF Wild Type

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIR296	20	0.0	180.8	-inf
MIR485	14	0.0	20.5	-inf
MIR2276	13	0.0	45.4	-inf
MIR489	7	20.1	0.0	inf
MIR214	1	67585.9	13.5	12.29
MIR335	7	2181.0	1.1	10.93
MIR135A2	12	96.0	0.3	8.42
MIR301A	17	2911.9	20.5	7.15
MIR218-2	5	181.6	2.2	6.34
MIR96	7	138.1	2.0	6.14
MIR182	7	1176.0	24.4	5.59
MIR218-1	4	2807.8	66.7	5.40
MIR199A2	1	778.1	21.9	5.15
MIR199B	9	4438.2	149.1	4.90
MIR301B	22	168.8	6.4	4.71
MIR340	5	157.1	6.4	4.61
MIR199A1	19	34109.9	1418.4	4.59
MIR99A	21	4641.1	212.2	4.45
MIR146B	10	110.3	5.3	4.37
MIR130A	11	10574.1	556.4	4.25
MIR26A1	3	777.8	52.1	3.90
MIR183	7	252.8	19.6	3.69
MIR20A	13	1957.3	162.0	3.60
MIR148A	7	513.6	43.2	3.57
MIR590	7	150.7	13.5	3.49
MIR181C	19	42.1	4.2	3.32
MIR192	11	220.4	22.1	3.32
MIR126	9	417.2	43.5	3.26
MIR454	17	34.9	4.8	2.87
MIR16-1	13	707.0	109.0	2.70
MIR598	8	34.9	5.9	2.57

MIR204	9	2227.6	388.5	2.52
MIR29C	1	107.7	19.6	2.46
MIR330	19	148.3	27.2	2.45
MIR26A2	12	4768.9	973.3	2.29
MIR30E	1	960.0	200.7	2.26
MIR425	3	773.2	176.0	2.14
MIR130B	22	511.8	136.5	1.91
MIR7-3	19	232.3	64.2	1.86
MIR338	17	146.7	42.3	1.79
MIR101-1	1	141.9	42.0	1.76
MIR30C1	1	1531.1	468.4	1.71
MIR132	17	21.2	6.7	1.66
MIR15A	13	346.6	114.9	1.59
MIR125A	19	27239.2	9235.5	1.56
MIR365A	16	115.6	40.9	1.50
MIR877	6	39.3	14.0	1.49
MIR874	5	160.2	57.5	1.48
MIR27B	9	303.1	118.0	1.36
MIR191	3	4125.1	1722.0	1.26
MIR34A	1	96.4	40.4	1.26
MIR128-1	2	119.1	52.4	1.18
MIR29B1	7	1540.7	687.9	1.16
MIR23B	9	6646.3	3087.2	1.11
MIR324	17	155.7	361.6	-1.22
MIR125B2	21	143.4	359.1	-1.32
MIR186	1	77.2	202.4	-1.39
MIR106B	7	191.2	538.2	-1.49
MIR23A	19	4573.0	13248.9	-1.54
MIR125B1	11	5630.9	17760.4	-1.66
MIR93	7	1162.5	3679.2	-1.66
MIR339	7	181.1	596.0	-1.72
MIRLET7C	21	62.2	220.3	-1.83
MIR197	1	290.7	1033.0	-1.83
MIR1229	5	5.1	18.5	-1.85
MIRLET7F1	9	99.7	375.1	-1.91
MIRLET7E	19	369.9	1775.0	-2.26
MIR140	16	880.8	4280.3	-2.28
MIR30D	8	324.0	1607.4	-2.31
MIR25	7	310.2	1584.4	-2.35
MIR1301	2	123.1	647.0	-2.39
MIR331	12	292.1	1595.3	-2.45
MIR15B	3	1953.2	10948.6	-2.49

MIR551B	3	28.7	187.3	-2.71
MIRLET7G	3	118.9	964.3	-3.02
MIR455	9	133.7	1190.8	-3.16
MIR320A	8	256.8	2421.7	-3.24
MIR589	7	8.0	77.7	-3.27
MIR152	17	20.3	207.2	-3.35
MIR550A3	7	6.6	67.3	-3.35
MIR302D	4	4.8	54.4	-3.52
MIR1180	17	25.1	292.1	-3.54
MIR302B	4	5.9	70.9	-3.60
MIR181A1	1	651.6	8115.6	-3.64
MIR139	11	3.3	43.5	-3.72
MIR100	11	315.4	5104.4	-4.02
MIR326	11	17.0	279.8	-4.04
MIR181B2	9	42.4	847.4	-4.32
MIR181A2	9	38.6	771.2	-4.32
MIR181B1	1	212.7	4275.2	-4.33
MIR760	1	2.9	62.0	-4.40
MIR211	15	1.5	52.4	-5.16
MIR629	15	3.7	136.0	-5.22
AC103686.1	8	2.2	85.2	-5.28
MIR451A	17	1.8	75.7	-5.37
MIRLET7A2	11	155.8	8831.0	-5.82
MIR135B	1	1.1	63.9	-5.86
MIR320B1	1	1.6	96.7	-5.88
MIR138-1	3	2.6	178.3	-6.12
MIR512-1	19	0.5	42.9	-6.29
MIR9-3	15	0.7	96.7	-7.05
MIRLET7I	12	2.6	345.4	-7.08
MIR138-2	16	2.6	363.9	-7.15
MIR320C1	18	0.5	102.0	-7.54
MIR31	9	2.4	700.2	-8.20
MIR30A	6	3.5	1052.9	-8.24
MIRLET7D	9	3.1	1719.8	-9.11

Table A.3.1.1(b): Differentially Expressed miRNAs in BRAF Wild Type COLO794 Holoclones

A.3.1.2 Thyroid

8505C BRAF V600E Mutant

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIRLET7C	21	39.7	0.0	-inf
MIR99B	19	169.6	0.0	-inf
MIR598	8	144.6	0.0	-inf
MIR181C	19	99.5	0.0	-inf
MIR32	9	99.5	0.0	-inf
MIR181B2	9	194.7	0.0	-inf
MIR33A	22	21.9	0.0	-inf
MIR296	20	42.7	0.0	-inf
MIR138-1	3	255.0	0.0	-inf
MIR744	17	340.8	0.0	-inf
MIR550A3	7	32.2	0.0	-inf
MIR301B	22	58.0	0.0	-inf
MIR143	5	0.1	468.8	12.40
MIR519A1	19	0.2	820.4	12.20
MIR1915	10	0.3	1211.1	12.18
MIR451A	17	1.0	2344.0	11.26
MIR517A	19	1.1	1836.2	10.67
MIR211	15	1.2	1406.4	10.17
MIR512-1	19	1.4	1250.2	9.81
MIR130B	22	26.0	19104.0	9.52
MIR373	19	0.7	312.5	8.81
MIR141	12	0.6	234.4	8.59
MIR145	5	1.2	312.5	8.00
MIR34A	1	2.3	429.7	7.57
MIR551B	3	7.6	1328.3	7.46
MIR205	1	3.0	351.6	6.85
MIR371A	19	3.8	390.7	6.67
MIR302A	4	4.8	234.4	5.61
MIR338	17	19.1	898.6	5.56
MIR449B	5	1.2	39.1	5.00
MIR302B	4	23.2	703.2	4.92
MIR302D	4	15.3	429.7	4.81
MIR302C	4	3.7	78.1	4.38
MIR204	9	13.5	234.4	4.12
MIR330	19	4.9	78.1	4.00
MIR214	1	16.2	234.4	3.86
MIR365A	16	3.0	39.1	3.68
MIR328	16	4.7	39.1	3.06

MIR335	7	10.0	78.1	2.97
MIR628	15	11.6	78.1	2.76
MIR199A1	19	23.4	117.2	2.32
MIR34C	11	8.1	39.1	2.27
MIR454	17	8.6	39.1	2.18
MIR197	1	19.9	78.1	1.97
MIR30D	8	359.9	1406.4	1.97
MIR326	11	22.6	78.1	1.79
MIR103A1	5	67.0	195.3	1.54
MIR148A	7	127.3	351.6	1.47
MIR1296	10	32.3	78.1	1.28
MIR125A	19	1432.1	2969.1	1.05
MIR193B	16	404.5	820.4	1.02
MIR126	9	556.3	273.5	-1.02
MIRLET7A2	11	1604.4	703.2	-1.19
MIR218-2	5	605.0	234.4	-1.37
MIR16-1	13	513.8	195.3	-1.40
MIR324	17	209.5	78.1	-1.42
MIR30B	8	2129.4	781.3	-1.45
MIR191	3	1924.7	703.2	-1.45
MIR101-1	1	226.3	78.1	-1.53
MIR425	3	233.6	78.1	-1.58
MIR186	1	124.1	39.1	-1.67
MIR29B2	1	5238.6	1640.8	-1.68
MIR181A1	1	4322.1	1211.1	-1.84
MIR339	7	931.7	234.4	-1.99
MIR181B1	1	468.4	117.2	-2.00
MIR23A	19	18915.7	4688.1	-2.01
MIR30C1	1	3078.2	742.3	-2.05
MIR106B	7	502.3	117.2	-2.10
MIR125B1	11	4266.5	976.7	-2.13
MIR193A	17	174.5	39.1	-2.16
MIR103A2	20	15330.5	3203.5	-2.26
MIR20A	13	187.6	39.1	-2.26
MIR182	7	188.9	39.1	-2.27
MIR192	11	621.6	117.2	-2.41
MIR301A	17	1461.0	273.5	-2.42
MIRLET7D	9	657.9	117.2	-2.49
MIR99A	21	2425.9	429.7	-2.50
MIR590	7	681.2	117.2	-2.54
MIR27A	19	5609.3	937.6	-2.58
MIR30A	6	2053.4	312.5	-2.72

MIR29B1	7	26912.8	3984.9	-2.76
MIR93	7	1106.3	156.3	-2.82
MIR24-2	19	5717.9	781.3	-2.87
MIRLET7G	3	2584.4	351.6	-2.88
MIR185	22	1263.6	156.3	-3.02
MIR29C	1	708.7	78.1	-3.18
MIR100	11	6131.2	625.1	-3.29
MIRLET7E	19	433.0	39.1	-3.47
MIRLET7F1	9	5704.7	507.9	-3.49
MIR130A	11	7250.1	351.6	-4.37
MIR21	17	81150.9	3359.8	-4.59
MIR218-1	4	5836.7	117.2	-5.64
MIR31	9	36234.8	507.9	-6.16

Table A.3.1.2(a): Differentially Expressed miRNAs in BRAF V600E mutated 8505C Holoclones

TPC-1 BRAF Wild Type

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIR429	1	72.0	0.0	-inf
MIRLET7A1	9	209.8	0.0	-inf
MIR598	8	67.0	0.0	-inf
MIR590	7	372.2	0.0	-inf
MIR139	11	173.1	0.0	-inf
MIR296	20	55.0	0.0	-inf
MIR877	6	22.7	0.0	-inf
MIR1226	3	58.8	0.0	-inf
MIR2355	2	26.5	0.0	-inf
MIR371A	19	0.0	340.1	inf
MIR515-1	19	0.0	318.8	inf
MIR199A1	19	0.0	11498.0	inf
MIR517C	19	0.0	637.6	inf
MIR519D	19	0.0	233.8	inf
MIR214	1	1.3	16981.3	13.71
MIR517A	19	1.3	2380.4	10.88
MIR451A	17	0.6	467.6	9.53
MIR145	5	1.3	786.4	9.28
MIR199B	9	1.9	701.4	8.53
MIR130B	22	91.0	32878.7	8.50
MIR302D	4	3.8	871.4	7.85
MIR302B	4	3.2	701.4	7.79

MIR375	2	1.3	191.3	7.24
MIR551B	3	4.4	382.6	6.44
MIR338	17	3.2	233.8	6.21
MIR218-1	4	9.5	531.3	5.81
MIR302A	4	3.2	127.5	5.34
MIR205	1	12.6	297.5	4.56
MIRLET7D	9	13.3	276.3	4.38
MIR26B	2	19.0	382.6	4.34
MIR7-2	15	5.1	85.0	4.07
MIR211	15	5.1	63.8	3.66
MIR769	19	5.1	63.8	3.66
MIR204	9	64.4	510.1	2.99
MIR192	11	686.2	4909.5	2.84
MIR215	1	8.2	42.5	2.37
MIR330	19	19.0	63.8	1.75
MIR194-2	11	6.3	21.3	1.75
MIR126	9	148.5	488.8	1.72
MIR550A3	7	19.6	63.8	1.70
MIR25	7	137.7	382.6	1.47
MIR26A1	3	151.0	403.8	1.42
MIR320A	8	41.1	106.3	1.37
MIR27B	9	103.6	233.8	1.17
MIR7-3	19	19.0	42.5	1.17
MIRLET7C	21	31.0	63.8	1.04
MIR103A2	20	19991.3	9563.9	-1.06
MIR185	22	314.0	148.8	-1.08
MIR331	12	830.9	340.1	-1.29
MIR191	3	2704.3	1105.2	-1.29
MIR16-2	3	323.5	127.5	-1.34
MIR328	16	109.3	42.5	-1.36
MIR23A	19	13043.6	4994.5	-1.39
MIR140	16	1501.2	573.8	-1.39
MIR135B	1	449.9	170.0	-1.40
MIR340	5	117.5	42.5	-1.47
MIR29B2	1	1816.5	637.6	-1.51
MIR148B	12	65.7	21.3	-1.63
MIR183	7	334.9	106.3	-1.66
MIR29B1	7	12520.5	3953.1	-1.66
MIR324	17	135.2	42.5	-1.67
MIR24-2	19	6417.6	1849.0	-1.80
MIR34A	1	702.0	191.3	-1.88
MIR125B1	11	2207.0	552.6	-2.00

MIR744	17	85.3	21.3	-2.01
MIR149	2	348.8	85.0	-2.04
MIR193A	17	365.2	85.0	-2.10
MIR200A	1	276.7	63.8	-2.12
MIR16-1	13	832.8	191.3	-2.12
MIR128-2	3	92.9	21.3	-2.13
MIR339	7	467.6	106.3	-2.14
MIR193B	16	2263.2	510.1	-2.15
MIRLET7G	3	1440.6	318.8	-2.18
MIR455	9	485.3	106.3	-2.19
MIR96	7	537.7	106.3	-2.34
MIR29A	7	5427.5	1062.7	-2.35
MIR181A2	9	117.5	21.3	-2.47
MIR182	7	1812.7	318.8	-2.51
MIR29C	1	385.4	63.8	-2.60
MIR181B2	9	262.2	42.5	-2.63
MIR301A	17	2686.6	425.1	-2.66
MIR152	17	154.2	21.3	-2.86
MIR301B	22	160.5	21.3	-2.92
MIR100	11	1879.7	233.8	-3.01
MIR30C1	1	4999.7	595.1	-3.07
MIR574	4	2713.7	318.8	-3.09
MIR93	7	418.3	42.5	-3.30
MIR181A1	1	4892.3	488.8	-3.32
MIR130A	11	3227.4	297.5	-3.44
MIRLET7I	12	257.2	21.3	-3.60
MIR197	1	551.0	42.5	-3.70
MIR138-1	3	1266.8	85.0	-3.90
MIR31	9	7549.8	488.8	-3.95
MIR181B1	1	1314.9	63.8	-4.37
MIR200B	1	457.5	21.3	-4.43
MIR30A	6	7023.5	212.5	-5.05
MIR21	17	176170.1	2762.9	-6.00
MIR138-2	16	2641.7	21.3	-6.96

Table A.3.1.2(b): Differentially Expressed miRNAs in BRAF Wild Type TPC-1 Holoclones

A.3.1.3 Ovarian

ES-2 BRAF V600E Mutant

miRNA	Chromosome	Parent Count	Holoclonal Count	Fold Change
MIR1301	2	285.0	0.0	-inf
AC103686.1	8	0.0	31.0	inf
MIR100	11	15.5	19665.6	10.31
MIR34A	1	1.2	1296.7	10.02
MIR199A2	1	0.2	77.1	8.53
MIR130A	11	0.6	196.6	8.30
MIR30A	6	2.6	732.8	8.14
MIR138-1	3	9.4	1232.7	7.04
MIR199B	9	0.5	58.1	6.80
MIR199A1	19	3.7	386.3	6.69
MIR125B1	11	407.6	25094.4	5.94
MIR139	11	3.2	192.7	5.90
MIR31	9	33.4	1894.6	5.83
MIR138-2	16	13.4	724.9	5.76
MIR125B2	21	36.6	1805.5	5.62
MIR874	5	0.7	29.7	5.35
MIR21	17	1162.7	46550.8	5.32
MIR214	1	3.3	103.2	4.95
MIR29B1	7	552.6	16887.7	4.93
MIR27A	19	29.1	756.6	4.70
MIR296	20	2.9	64.3	4.47
MIR142	17	1.9	30.7	4.04
MIR23A	19	274.1	4273.8	3.96
MIR29A	7	895.3	11226.3	3.65
MIR193A	17	10.0	114.9	3.53
MIR181A2	9	75.9	706.0	3.22
MIR574	4	253.5	2207.5	3.12
MIR29B2	1	97.8	775.2	2.99
MIR152	17	5.8	39.2	2.75
MIRLET7A2	11	254.7	1361.4	2.42
MIR125A	19	1194.9	6186.1	2.37
MIRLET7E	19	49.6	183.5	1.89
MIR320A	8	184.1	678.6	1.88
MIR99A	21	1189.8	4223.9	1.83
MIR181A1	1	735.5	2143.1	1.54
MIR365A	16	23.4	50.6	1.11
MIR26B	2	85.1	172.7	1.02
MIR193B	16	801.1	374.9	-1.10

MIR326	11	20.7	8.8	-1.23
MIR30C1	1	1011.3	366.7	-1.46
MIR128-1	2	70.2	25.1	-1.48
MIR30E	1	514.2	181.6	-1.50
MIR26A1	3	917.7	281.8	-1.70
MIR25	7	460.1	140.4	-1.71
MIR103A2	20	20958.4	6309.9	-1.73
MIR328	16	209.7	59.1	-1.83
MIR96	7	188.1	51.3	-1.88
MIR769	19	28.0	7.5	-1.90
MIRLET7F1	9	2786.3	725.3	-1.94
MIR629	15	26.5	6.2	-2.10
MIR30C2	6	37.9	8.8	-2.10
MIR183	7	272.1	61.1	-2.16
MIR106B	7	317.2	71.2	-2.16
MIR339	7	269.4	58.5	-2.20
MIR15B	3	4500.2	962.7	-2.23
MIR148B	12	127.7	26.5	-2.27
MIR182	7	2469.6	455.5	-2.44
MIR186	1	53.6	9.5	-2.50
MIR128-2	3	200.2	33.6	-2.57
MIR140	16	2747.1	441.2	-2.64
MIR101-1	1	189.7	29.1	-2.71
MIR589	7	24.0	3.3	-2.88
MIR218-2	5	218.2	27.4	-2.99
MIR218-1	4	2558.2	311.5	-3.04
MIR324	17	125.2	15.0	-3.06
MIR185	22	167.8	18.6	-3.17
MIR331	12	648.4	71.8	-3.17
MIR550A3	7	18.2	2.0	-3.22
MIR192	11	383.0	40.5	-3.24
MIR103A1	5	200.4	19.9	-3.33
MIR454	17	93.2	9.1	-3.35
MIR551B	3	404.6	31.0	-3.71
MIR590	7	524.4	38.9	-3.75
MIR23B	9	8042.2	538.1	-3.90
MIR340	5	98.6	6.5	-3.92
MIR27B	9	332.9	18.3	-4.19
MIR33B	17	27.1	1.3	-4.38
MIR1343	11	28.8	1.3	-4.46
MIR9-3	15	21.6	1.0	-4.47
MIR1296	10	24.4	1.0	-4.64

MIR744	17	230.6	9.1	-4.66
MIR149	2	34.4	1.3	-4.72
MIRLET7D	9	1064.2	32.3	-5.04
MIR301A	17	7287.3	213.6	-5.09
MIR7-2	15	491.5	12.1	-5.35
MIR1180	17	134.6	3.3	-5.37
MIR33A	22	95.2	2.3	-5.38
MIR32	9	29.4	0.7	-5.49
MIR15A	13	667.2	11.8	-5.83
MIR301B	22	513.8	7.2	-6.16
MIR7-3	19	901.2	11.1	-6.34
MIR126	9	360.1	2.9	-6.94
MIR338	17	246.0	0.7	-8.56
MIR148A	7	958.4	1.0	-9.93

Table A.3.1.3(a): Differentially Expressed miRNAs in BRAF V600E mutated ES-2 Holoclones

A2780 BRAF Wild Type

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIR375	2	0.0	46.7	inf
MIR200C	12	0.0	26.4	inf
MIR200B	1	0.0	115.7	inf
MIR489	7	2.3	302.8	7.06
MIR200A	1	1.1	99.5	6.45
MIRLET7D	9	0.8	64.8	6.42
MIR194-1	1	0.8	33.7	5.48
MIR181B1	1	4.9	197.6	5.33
MIR181D	19	1.1	21.9	4.27
MIRLET7A2	11	76.4	1137.6	3.90
MIR328	16	15.9	215.8	3.76
MIR31	9	7.9	97.6	3.62
MIR320C1	18	6.1	74.0	3.61
MIR365A	16	12.9	152.5	3.57
MIR150	19	1.9	22.2	3.56
MIRLET7C	21	29.9	327.0	3.45
MIR125B1	11	620.5	6124.0	3.30
MIRLET7A1	9	1.9	18.1	3.26
MIR517A	19	3.0	28.9	3.26
MIR181A1	1	51.5	476.3	3.21
MIR874	5	11.0	97.2	3.15

MIR181B2	9	4.5	37.8	3.06
MIR130B	22	174.8	1360.6	2.96
MIR181A2	9	3.0	22.9	2.92
MIR197	1	26.5	160.1	2.60
MIR215	1	4.2	21.0	2.33
MIR204	9	404.1	1927.2	2.25
MIR193A	17	64.3	287.6	2.16
MIR100	11	18.5	82.6	2.16
MIR330	19	20.4	86.7	2.09
MIR145	5	32.5	136.3	2.07
MIR1226	3	7.2	29.2	2.02
MIR1301	2	14.0	56.6	2.02
MIR125A	19	2867.7	11407.6	1.99
MIRLET7G	3	77.6	282.8	1.87
MIR146B	10	14.8	50.5	1.78
MIR26A1	3	122.2	400.1	1.71
MIR425	3	215.3	677.1	1.65
MIR451A	17	5.7	17.8	1.65
MIR574	4	222.1	689.8	1.64
MIRLET7E	19	334.4	990.4	1.57
MIR26B	2	249.3	721.3	1.53
MIR27B	9	296.2	840.1	1.50
MIR181C	19	10.6	29.9	1.50
MIR30D	8	362.4	990.4	1.45
MIRLET7F1	9	107.1	289.8	1.44
MIR26A2	12	1575.0	4241.7	1.43
MIR192	11	593.6	1574.1	1.41
MIR877	6	6.1	15.6	1.36
MIR93	7	669.6	1628.2	1.28
MIR338	17	79.8	191.6	1.26
MIR214	1	14790.2	34673.8	1.23
MIR199A1	19	22177.0	49158.2	1.15
MIR30E	1	877.0	431.5	-1.02
MIR29A	7	5121.0	2375.2	-1.11
MIR29C	1	302.3	137.0	-1.14
MIR182	7	1353.6	598.6	-1.18
MIR140	16	872.0	380.4	-1.20
MIR302B	4	15.9	6.0	-1.40
MIR335	7	2250.3	833.8	-1.43
MIR34A	1	244.4	90.2	-1.44
MIR148A	7	781.2	283.4	-1.46
MIR744	17	438.5	152.5	-1.52

MIR32	9	25.7	8.9	-1.53
MIR7-3	19	191.4	65.8	-1.54
MIR106B	7	127.5	41.9	-1.60
MIR186	1	104.8	33.4	-1.65
MIR30C1	1	2080.0	653.3	-1.67
MIR1296	10	135.1	40.7	-1.73
MIR30B	8	1074.4	322.2	-1.74
MIR629	15	21.6	6.4	-1.76
MIR339	7	441.9	128.7	-1.78
MIR301A	17	3207.4	925.3	-1.79
MIR29B2	1	501.7	142.4	-1.82
MIR30C2	6	67.0	17.8	-1.91
MIR21	17	12966.7	3227.7	-2.01
MIR454	17	79.1	19.4	-2.03
MIR301B	22	199.4	43.9	-2.19
MIR598	8	109.3	19.4	-2.50
MIR340	5	356.0	61.0	-2.55
MIR599	8	137.7	20.3	-2.76
MIR551B	3	62.8	8.9	-2.82
MIR29B1	7	6442.9	830.6	-2.96
MIR96	7	563.7	70.5	-3.00
MIR550A3	7	65.5	7.9	-3.04
MIR148B	12	231.9	27.6	-3.07
MIR590	7	1714.2	76.6	-4.48
MIR185	22	1876.1	58.5	-5.00

Table A.3.1.3(b): Differentially Expressed miRNAs in BRAF Wild Type A2780 Holoclones

A.3.1.4 Colorectal

HT-29 BRAF V600E Mutant

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIR517A	19	1.4	46.5	5.08
MIR199A1	19	1.2	31.1	4.69
MIR642A	19	4.0	41.5	3.39
MIR451A	17	2.3	22.0	3.26
MIR150	19	88.6	744.6	3.07
MIR30A	6	15.6	130.5	3.07
MIR328	16	19.4	149.8	2.95
MIR152	17	6.3	42.5	2.76
MIR375	2	244.5	1625.3	2.73
MIR200C	12	177.9	1141.1	2.68
MIR194-1	1	218.7	1369.2	2.65
MIR324	17	21.8	126.8	2.54
MIR130B	22	147.9	828.3	2.49
MIR26B	2	229.3	1239.8	2.44
MIR181A2	9	94.4	509.1	2.43
MIR26A1	3	151.4	751.7	2.31
MIR326	11	64.0	297.3	2.22
MIR574	4	230.3	1069.8	2.22
MIR125A	19	3153.1	14591.6	2.21
MIR31	9	4279.1	19687.6	2.20
MIR139	11	5.7	24.4	2.10
MIR126	9	14.9	62.6	2.07
MIR330	19	9.7	40.2	2.05
MIR197	1	83.2	330.1	1.99
MIR193B	16	335.0	1212.0	1.86
MIR186	1	92.4	306.4	1.73
MIR148A	7	11.2	35.9	1.69
MIR125B1	11	5.6	17.7	1.68
MIR365A	16	26.6	77.8	1.55
MIR215	1	298.9	873.0	1.55
MIR194-2	11	205.8	596.6	1.54
MIR320A	8	111.7	315.4	1.50
MIR577	4	72.0	185.4	1.36
MIR30D	8	1435.2	3537.9	1.30
MIR181C	19	8.1	19.0	1.23
MIR200B	1	1679.2	3881.6	1.21
MIR192	11	24499.4	54709.3	1.16
MIR141	12	350.5	762.4	1.12

MIR33A	22	60.3	130.0	1.11
MIR181A1	1	1931.0	3953.9	1.03
MIR26A2	12	1448.2	2944.0	1.02
MIR1343	11	12.3	24.9	1.02
MIRLET7C	21	17.9	8.6	-1.06
MIRLET7A1	9	181.0	84.0	-1.11
MIR21	17	44127.3	18457.2	-1.26
MIR1910	16	17.7	7.3	-1.27
MIR340	5	84.8	31.6	-1.42
MIR590	7	1813.6	645.8	-1.49
MIR16-2	3	598.9	205.4	-1.54
MIR7-3	19	1917.3	627.3	-1.61
MIR7-2	15	2053.4	623.3	-1.72
MIR185	22	1249.5	370.3	-1.75
MIR29B2	1	3889.3	1125.8	-1.79
MIR1180	17	31.3	8.9	-1.81
MIR103A1	5	245.6	64.2	-1.94
MIR301A	17	2004.9	507.7	-1.98
MIR301B	22	105.5	24.4	-2.11
MIR20A	13	3672.9	838.5	-2.13
MIR549	15	22.6	5.0	-2.19
MIR551A	1	44.9	7.5	-2.58
MIR454	17	55.9	6.5	-3.09

Table A.3.1.4(a): Differentially Expressed miRNAs in BRAF V600E mutated HT-29 Holoclones

COLO320 BRAF Wild Type

miRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
MIR518C	19	27.502	0	-inf
MIR604	10	0	5832.203	inf
MIR592	7	0	88.062	inf
MIR214	1	7.284	2232.184	8.26
MIR302D	4	2.888	501.571	7.44
MIR302B	4	4.27	572.404	7.067
MIR200A	1	0.251	20.101	6.323
MIR320A	8	11.176	715.026	5.999
MIR211	15	0.377	23.93	5.989
MIR302A	4	1.13	70.833	5.97
MIR302C	4	0.628	39.245	5.966
MIR130A	11	6.153	280.459	5.51

MIR371A	19	2.009	83.276	5.373
MIR31	9	2.26	75.619	5.064
MIR199A1	19	68.314	1695.196	4.633
MIR335	7	3.265	75.619	4.534
MIR34A	1	1.633	27.759	4.088
MIR204	9	18.586	255.572	3.781
MIR199B	9	21.599	227.813	3.399
MIR30C2	6	40.436	385.75	3.254
MIR205	1	19.088	175.167	3.198
MIRLET7E	19	5.149	44.031	3.096
MIR135B	1	10.549	88.062	3.061
MIR26A2	12	322.484	2522.215	2.967
MIR125A	19	824.168	4522.757	2.456
MIR365A	16	41.064	155.066	1.917
AC103686.1	8	5.023	17.23	1.778
MIR29B1	7	543.25	1813.889	1.739
MIR138-2	16	5.777	18.187	1.655
MIR877	6	8.79	25.844	1.556
MIR9-3	15	91.295	243.128	1.413
MIR197	1	117.666	301.517	1.358
MIRLET7D	9	368.32	911.252	1.307
MIR550A1	7	11.302	26.802	1.246
MIR30D	8	1015.799	2319.289	1.191
MIR100	11	68.691	140.708	1.035
MIR181B2	9	132.359	268.972	1.023
MIR551B	3	603.276	272.801	-1.145
MIR186	1	201.301	89.019	-1.177
MIR148B	12	249.9	107.206	-1.221
MIR101-1	1	497.916	211.541	-1.235
MIR589	7	18.46	7.658	-1.269
MIR16-2	3	235.961	93.805	-1.331
MIR301B	22	342.325	134.965	-1.343
MIR181C	19	14.693	5.743	-1.355
MIR15B	3	6229.793	2402.565	-1.375
MIRLET7C	21	68.691	25.844	-1.41
MIR218-1	4	2263.667	815.532	-1.473
MIR339	7	472.55	169.424	-1.48
MIR183	7	660.665	225.899	-1.548
MIR30B	8	798.047	272.801	-1.549
MIR550A3	7	47.845	16.272	-1.556
MIR1301	2	194.52	64.132	-1.601
MIR96	7	392.18	124.436	-1.656

MIR1323	19	18.209	5.743	-1.665
MIR331	12	1036.018	324.49	-1.675
MIR215	1	15.948	4.786	-1.737
MIR148A	7	1607.02	463.283	-1.794
MIR33A	22	161.242	45.945	-1.811
MIRLET7F1	9	7644.553	2155.608	-1.826
MIR145	5	190.627	47.86	-1.994
MIR517A	19	390.673	96.677	-2.015
MIRLET7G	3	4323.27	1057.703	-2.031
MIR128-1	2	176.939	43.074	-2.038
MIR326	11	39.306	9.572	-2.038
MIR24-2	19	2840.697	667.166	-2.09
MIR455	9	590.09	131.136	-2.17
MIR1180	17	240.858	49.774	-2.275
MIR126	9	1075.7	220.155	-2.289
MIR629	15	108.123	20.101	-2.427
MIR27B	9	609.429	107.206	-2.507
MIR143	5	38.929	6.7	-2.539
MIR512-1	19	96.695	16.272	-2.571
MIR93	7	1244.854	208.669	-2.577
MIR590	7	1063.645	172.296	-2.626
MIR1296	10	61.408	9.572	-2.682
MIR517C	19	76.854	11.486	-2.742
MIR744	17	559.198	81.362	-2.781
MIR29C	1	145.419	21.058	-2.788
MIR128-2	3	675.735	95.72	-2.82
MIR451A	17	276.899	37.331	-2.891
MIR20A	13	9475.856	1218.512	-2.959
MIR218-2	5	314.447	40.202	-2.967
MIR106B	7	777.327	95.72	-3.022
MIR32	9	48.096	5.743	-3.066
MIR33B	17	33.278	2.872	-3.535
MIR26A1	3	1367.794	107.206	-3.673
MIR454	17	136.252	10.529	-3.694
MIR515-1	19	68.817	3.829	-4.168
MIR519D	19	50.106	1.914	-4.71
MIR185	22	864.729	31.588	-4.775
MIR9-1	1	54.124	1.914	-4.821
MIRLET7A1	9	1207.682	39.245	-4.944
MIR769	19	35.413	0.957	-5.209

Table A.3.1.4(b): Differentially Expressed miRNAs in BRAF Wild Type COLO320 Holoclones

A.4.1 Holoclone Differentially Expressed snoRNA Lists

A.4.1.1 Melanoma

All differentially expressed snoRNAs in this section are displayed as Log₂ fold changes.

SK-Mel 28 BRAF V600E Mutant

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD85	1	0	21.001	inf
SNORA74A	5	0	24.001	inf
SNORD114-3	14	0	27.201	inf
SNORD113-9	14	0	25.201	inf
SNORD20	2	0	23.401	inf
SNORD91A	17	0	41.802	inf
SCARNA2	1	0	33.202	inf
SNORA36B	1	26.11	128.407	2.298

Table A.4.1.1(a): Up regulated snoRNAs in BRAF V600E mutated SK-Mel 28 Holoclones

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORA46	16	78.331	2	-5.291
SNORD19	3	52.221	2	-4.706
SNORD41	19	313.326	19.001	-4.044
SNORD30	11	234.994	16.201	-3.858
SNORA76	17	26.11	2	-3.706
SNORD42A	17	156.663	12.201	-3.683
SNORD94	2	52.221	5.4	-3.274
SNORD64	15	104.442	11.201	-3.221
SNORD12B	20	365.547	39.402	-3.214
SNORD101	6	208.884	23.001	-3.183
SCARNA4	1	26.11	3	-3.122
SNORD121A	9	52.221	6.8	-2.941
SNORD58C	18	182.773	24.001	-2.929
SNORD10	17	234.994	31.202	-2.913
SNORD88A	19	287.215	41.602	-2.787
SNORD114-22	14	26.11	3.8	-2.78
SCARNA18	18	26.11	4	-2.706
SNORD75	1	130.552	20.201	-2.692
SNORD38A	1	182.773	28.402	-2.686
SNORD12	20	365.547	59.603	-2.617

SNORD70	2	52.221	8.6	-2.602
SNORD60	16	130.552	22.601	-2.53
SNORD12C	20	287.215	53.403	-2.427
SNORD1B	17	704.983	132.807	-2.408
SNORA21	17	26.11	5	-2.385
SNORD9	14	26.11	5	-2.385
SNORA42	1	26.11	5	-2.385
SNORD117	6	156.663	30.802	-2.347
SNORD8	14	52.221	10.601	-2.3
SNORD69	3	469.989	97.205	-2.274
SNORD1A	17	443.878	92.205	-2.267
SNORD23	19	26.11	5.6	-2.221
SNORD83A	22	261.105	60.603	-2.107
SNORD82	2	208.884	48.603	-2.104
SNORD114-26	14	26.11	6.4	-2.028
SNORA54	11	26.11	6.4	-2.028
SNORD58A	18	182.773	46.002	-1.99
SNORD119	20	469.989	120.606	-1.962
SNORD90	9	26.11	6.8	-1.941
SNORD99	1	234.994	64.003	-1.876
SCARNA6	2	104.442	29.602	-1.819
SNORD35B	19	156.663	45.002	-1.8
SNORD6	11	130.552	37.802	-1.788
SNORD46	1	52.221	15.401	-1.762
SNORD125	22	26.11	7.8	-1.743
SCARNA10	12	26.11	8	-1.706
SNORD51	2	130.552	46.202	-1.499
SNORD63	5	104.442	37.602	-1.474
SNORD11B	2	26.11	10.001	-1.385
SNORD100	6	1305.524	507.227	-1.364
SNORD71	16	52.221	21.401	-1.287
SNORD17	20	78.331	33.802	-1.212
SNORD93	7	287.215	125.207	-1.198
SNORD59B	12	52.221	23.801	-1.134
SNORD78	1	1331.635	632.434	-1.074

Table A.4.1.1(b): Down regulated snoRNAs in BRAF V600E mutated SK-Mel 28 Holoclones

COLO794 BRAF Wild Type

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
snoR26	20	0.28	27.987	6.642
SNORD101	6	2.523	93.29	5.209
SNORA36B	1	12.895	110.668	3.101
SNORD20	2	10.372	83.595	3.011
SNORD88C	19	17.66	119.997	2.764
SNORD85	1	5.326	26.524	2.316
SNORD94	2	4.765	22.682	2.251
SNORD117	6	24.949	111.216	2.156
SNORD116-3	15	5.046	22.316	2.145
SNORA21	17	3.925	17.195	2.131
SNORD6	11	31.957	130.789	2.033
SNORD63	5	15.137	57.803	1.933
snR39B	2	11.213	41.157	1.876
SNORA14B	1	16.819	61.645	1.874
SNORA74A	5	7.849	28.17	1.844
SNORD51	2	16.539	59.084	1.837
SNORD71	16	13.456	47.011	1.805
SNORA65	9	5.046	17.56	1.799
SNORD105	19	7.008	23.963	1.774
SNORD78	1	487.763	1588.308	1.703
SNORD75	1	30.275	91.644	1.598
SNORD88A	19	35.321	102.436	1.536
SNORD91A	17	30.836	82.863	1.426
SNORD64	15	7.288	18.475	1.342
SNORD69	3	63.914	155.666	1.284
SNORD82	2	47.094	99.326	1.077
SNORD38A	1	19.062	38.596	1.018

Table A.4.1.1(c): Up regulated snoRNAs in BRAF Wild Type COLO794 Holoclones

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD114-10	14	42.048	0	-inf
SNORD114-1	14	32.237	0.183	-7.461
SNORD113-9	14	61.11	0.549	-6.799
SNORD60	16	2537.487	28.353	-6.484
SNORA31	13	42.609	0.549	-6.279
SNORD114-3	14	50.739	0.732	-6.116
SNORD114-12	14	46.534	0.915	-5.669

SNORD11	2	178.286	7.134	-4.643
SNORD70	2	71.202	4.39	-4.02
SNORD127	14	23.827	2.378	-3.325
SNORD93	7	402.264	46.462	-3.114
SNORD12C	20	100.075	18.292	-2.452
SNORD46	1	40.086	8.597	-2.221
SNORD42A	17	45.412	10.244	-2.148
SNORD125	22	21.865	5.122	-2.094
SNORD10	17	45.973	12.805	-1.844
SNORD1B	17	243.04	69.693	-1.802
SNORD12B	20	119.418	36.584	-1.707
SNORD58C	18	156.701	53.23	-1.558
SNORD4B	17	44.291	15.365	-1.527
SNORD58A	18	217.531	86.705	-1.327
SNORD15B	11	39.245	16.829	-1.222
SNORD99	1	179.968	85.424	-1.075

Table A.4.1.1(d): Down regulated snoRNAs in BRAF Wild Type COLO794 Holoclones

A.4.1.2 Thyroid

8505C BRAF V600E Mutant

snRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD114-14	14	0	586.011	inf
SNORD113-6	14	0	429.741	inf
SNORD113-7	14	0	820.415	inf
SNORA70	2	0	1054.82	inf
SNORA40	7	0.087	1601.763	14.168
SNORD111	16	0.174	2812.852	13.98
SNORD114-22	14	0.087	625.078	12.81
SNORD114-25	14	0.174	1132.954	12.668
SNORD113-9	14	0.609	2226.841	11.836
SNORD116-16	15	0.435	1562.696	11.81
SNORD114-1	14	0.348	1093.887	11.618
SNORD114-3	14	0.696	1875.235	11.395
SNORD114-26	14	0.174	429.741	11.27
SNORD114-12	14	0.696	1406.426	10.98
SNORD114-9	14	0.435	859.483	10.948
SNORD42A	17	0.261	390.674	10.547
SNORD19B	3	0.435	273.472	9.296
SNORD12B	20	1.653	820.415	8.955
SNORD64	15	0.783	273.472	8.448
SNORD46	1	1.044	273.472	8.033
SNORD30	11	2.175	429.741	7.626
SNORD75	1	5.569	976.685	7.454
SNORD9	14	0.696	117.202	7.395
SNORD121A	9	0.957	156.27	7.351
SNORD7	17	1.044	156.27	7.225
SNORD71	16	2.175	312.539	7.167
SNORD58C	18	4.09	546.944	7.063
SNORD100	6	59.868	7618.142	6.992
SNORD99	1	2.524	312.539	6.952
SNORD104	17	9.659	1132.954	6.874
SNORD37	19	2.436	273.472	6.81
SNORD112	2	1.74	195.337	6.81
SNORD4B	17	3.655	234.404	6.003
SNORD70	2	1.479	78.135	5.723
SNORA31	13	1.479	78.135	5.723
SNORD83A	22	6.787	351.607	5.695
SNORD69	3	17.317	859.483	5.633
SNORD58A	18	5.569	273.472	5.618
SNORD59B	12	10.616	507.876	5.58

SCARNA6	16	0.87	39.067	5.489
SNORD119	20	9.833	429.741	5.45
SNORA34	12	1.044	39.067	5.225
SNORD105B	19	16.62	586.011	5.14
SNORD12	20	2.262	78.135	5.11
SNORA5A	7	1.218	39.067	5.003
SNORA24	4	2.524	78.135	4.952
SNORD82	2	5.395	156.27	4.856
SNORA75	2	1.392	39.067	4.81
SNORA14B	1	2.872	78.135	4.766
SNORA33	6	3.307	78.135	4.563
SNORD8	14	1.914	39.067	4.351
SNORD63	5	7.658	156.27	4.351
SNORD1B	17	7.919	156.27	4.303
SNORD67	11	2.001	39.067	4.287
SNORA25	10	2.262	39.067	4.11
SNORD60	16	18.535	312.539	4.076
SNORD12C	20	31.239	507.876	4.023
SNORD38A	1	9.833	156.27	3.99
SNORD85	1	2.611	39.067	3.904
SNORD125	22	8.006	117.202	3.872
SNORD125	22	8.006	117.202	3.872
SNORD103A	1	5.395	78.135	3.856
SNORD6	11	14.358	195.337	3.766
SNORD93	7	20.449	273.472	3.741
SNORD35B	19	17.578	234.404	3.737
SNORD105	19	13.575	156.27	3.525
snR39B	2	52.472	468.809	3.159
SNORA71D	20	9.659	78.135	3.016
SNORD90	9	5.308	39.067	2.88
SNORD91B	17	49.861	312.539	2.648
SNORA65	9	12.879	78.135	2.601
SNORD117	6	27.411	156.27	2.511
SNORA36B	1	7.048	39.067	2.471
SNORA45	11	9.659	39.067	2.016
SNORD78	1	131.484	507.876	1.95
SNORD91A	17	22.538	78.135	1.794
SNORD20	2	12.792	39.067	1.611
SNORD17	20	13.575	39.067	1.525

Table A.4.1.2(a): Up regulated snoRNAs in BRAF V600E mutated 8505C Holoclones

snoRNA	Chromosome	Parent Count	Holoclonal Count	Fold Change
SNORD88A	19	80.143	0	-inf

Table A.4.1.2(b): Down regulated snoRNAs in BRAF V600E mutated 8505C Holoclones

TPC-1 BRAF Wild Type

snoRNA	Chromosome	Parent Count	Holoclonal Count	Fold Change
SNORA70	2	0	318.798	inf
SNORD7	17	0	340.051	inf
SNORD30	11	28.433	4718.212	7.375
SNORD100	6	338.665	50348.848	7.216
SNORD19B	3	6.318	573.837	6.505
SNORD11B	2	5.055	361.305	6.159
SNORD64	15	6.318	446.317	6.142
SNORD75	1	16.428	935.141	5.831
SNORD83A	22	20.219	1083.914	5.744
SNORD59B	12	29.696	1572.737	5.727
SNORD37	19	7.582	361.305	5.574
SNORD105B	19	8.214	361.305	5.459
SNORD12B	20	23.378	998.901	5.417
SNORD6	11	10.741	340.051	4.985
SNORD104	17	140.9	4101.869	4.864
SNORD82	2	10.741	233.785	4.444
SNORD69	3	40.438	828.875	4.357
SNORD71	16	4.423	85.013	4.265
SNORD58C	18	29.696	510.077	4.102
SNORD103A	1	6.318	106.266	4.072
SNORD90	9	10.109	170.026	4.072
SNORD99	1	41.701	595.09	3.835
SNORD85	1	7.582	106.266	3.809
SNORD125	22	11.373	148.772	3.709
SNORA76	17	5.055	63.76	3.657
SNORD88C	19	41.069	446.317	3.442
SNORD91A	17	28.433	297.545	3.387
SNORD126	14	6.95	63.76	3.198
SNORD42A	17	6.95	63.76	3.198
SNORD91B	17	28.433	255.038	3.165
SNORD17	20	10.109	85.013	3.072
SNORD35B	19	43.597	361.305	3.051

SNORD1A	17	24.642	191.279	2.957
SNORD66	3	89.721	680.103	2.922
SNORD11	2	5.687	42.506	2.902
SNORD8	14	8.846	63.76	2.85
SNORA14B	1	8.846	63.76	2.85
SNORD12C	20	19.587	127.519	2.703
SNORD12	20	13.9	85.013	2.613
SNORD1B	17	17.691	106.266	2.587
SNORD60	16	21.482	127.519	2.569
SNORD20	2	16.428	85.013	2.372
SNORD119	20	31.592	148.772	2.235
SNORD10	17	24.01	106.266	2.146
SNORA42	1	6.318	21.253	1.75
SNORD88A	19	19.587	63.76	1.703
SNORD101	6	27.801	85.013	1.613
SNORA74A	5	7.582	21.253	1.487
SNORA48	17	15.164	42.506	1.487
SNORD105	19	7.582	21.253	1.487
snR39B	2	15.164	42.506	1.487
SNORD117	6	40.438	106.266	1.394
SNORD70	2	8.214	21.253	1.372
SNORD78	1	366.465	892.635	1.284
SNORD58A	18	44.229	106.266	1.265
SNORA21	17	9.478	21.253	1.165
SNORA75	2	10.109	21.253	1.072

Table A.4.1.2(c): Up regulated snoRNAs in BRAF Wild Type TPC-1 Holoclones

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD93	7	152.273	42.506	-1.841
SNORA36B	1	66.343	21.253	-1.642

Table A.4.1.2(d): Down regulated snoRNAs in BRAF Wild Type TPC-1 Holoclones

A.4.1.3 Ovarian

ES-2 BRAF V600E Mutant

snRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD114-13	14	0	28.083	inf
SNORD64	15	0	20.246	inf
SNORD123	5	0	25.797	inf
SNORD42A	17	1.976	179.927	6.508
SNORD1A	17	1.976	77.391	5.291
SCARNA1	1	1.04	37.553	5.174
SNORD58A	18	5.929	162.947	4.78
SNORD75	1	13.419	180.907	3.753
SNORD46	1	6.865	82.943	3.595
SNORD58C	18	12.378	139.435	3.494
SNORD23	19	2.392	22.532	3.235
SNORD30	11	23.613	215.847	3.192
SCARNA7	3	2.913	20.899	2.843
SNORD6	11	42.024	260.91	2.634
SNORD17	20	13.419	80.983	2.593
SNORD7	17	3.433	19.266	2.489
SNORD99	1	37.343	209.316	2.487
SNORA33	6	9.57	51.268	2.421
SNORD63	5	12.69	66.942	2.399
SNORD15B	11	7.281	37.879	2.379
SCARNA10	12	6.449	32.981	2.354
SNORD12B	20	27.669	139.435	2.333
SNORD69	3	83.32	403.611	2.276
SNORD82	2	43.168	194.295	2.17
SNORD83A	22	31.934	138.782	2.12
SNORA42	1	4.577	19.266	2.074
SNORD67	11	6.241	26.124	2.065
SNORD104	17	225.619	921.84	2.031
SNORD119	20	101.836	383.692	1.914
SNORA65	9	14.875	54.86	1.883
SCARNA5	2	5.201	17.96	1.788
SNORD90	9	7.385	25.471	1.786
SNORD103A	1	6.137	20.899	1.768
SNORA74A	5	21.428	72.493	1.758
SNORD71	16	14.875	48.982	1.719
SNORD1B	17	58.251	174.702	1.585
SNORD41	19	14.043	40.165	1.516
SNORD12C	20	61.892	172.743	1.481
SNORD11	2	10.402	28.41	1.45

SNORD4B	17	16.331	42.124	1.367
SNORD12	20	59.395	146.293	1.3
SNORA45	11	10.818	26.45	1.29
SNORA36B	1	24.445	54.533	1.158
SNORD51	2	39.944	88.167	1.142

Table A.4.1.3 (a): Up regulated snoRNAs in BRAF V600E mutated ES-2 Holoclones

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD88B	19	71.878	6.204	-3.534
SNORD88C	19	516.043	132.251	-1.964
SNORD93	7	290.944	95.025	-1.614
SCARNA4	1	24.445	10.123	-1.272
snR39B	2	90.809	39.839	-1.189
SNORD88A	19	180.059	88.167	-1.03

Table A.4.1.3(b): Down regulated snoRNAs in BRAF V600E mutated ES-2 Holoclones

A2780 BRAF Wild Type

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD85	1	0.757	50.205	6.052
SNORD30	11	5.675	194.147	5.096
SNORD12B	20	3.027	86.746	4.841
SNORD100	6	93.446	2408.245	4.688
SNORD19B	3	1.513	33.682	4.476
SNORD104	17	18.16	403.863	4.475
SNORD99	1	6.432	140.447	4.449
SNORD94	2	1.135	20.972	4.208
SNORD89	2	1.513	27.327	4.175
SNORD82	2	7.945	139.176	4.131
SNORD116-3	15	3.027	48.934	4.015
SNORD58A	18	4.162	64.504	3.954
SNORD90	9	3.027	45.756	3.918
SNORD75	1	12.863	189.063	3.878
SNORD119	20	7.188	86.111	3.582
SNORD12	20	4.162	48.934	3.556
SNORD71	16	3.027	32.728	3.435
SNORD63	5	8.323	84.204	3.339
SNORD64	15	3.783	37.813	3.321

SNORD7	17	2.27	21.925	3.272
SNORD6	11	8.701	80.074	3.202
SNORD46	1	4.162	32.411	2.961
SNORD59B	12	13.241	101.999	2.945
SNORD67	11	2.648	19.383	2.872
SNORD8	14	3.027	21.607	2.836
SNORD83A	22	16.268	114.391	2.814
SCARNA5	2	3.405	22.243	2.708
SNORD37	19	6.053	36.224	2.581
SNORA36B	1	12.106	69.588	2.523
SNORD48	6	4.162	23.514	2.498
SNORD58C	18	8.323	46.71	2.489
SNORA74A	5	4.162	21.925	2.397
SNORD66	3	38.211	189.698	2.312
SNORD127	14	4.918	22.56	2.198
SNORD38A	1	5.297	23.514	2.15
SNORD91A	17	10.971	48.616	2.148
SNORD35B	19	16.646	73.719	2.147
SNORA33	6	6.053	25.42	2.07
SNORD20	2	29.131	114.391	1.973
SNORA14B	1	8.701	33.682	1.953
SNORD1A	17	15.133	56.56	1.902
SNORD51	2	15.133	55.607	1.878
SCARNA6	2	10.593	36.542	1.786
SNORD93	7	7.945	26.691	1.748
SNORD78	1	471.393	1563.659	1.73
SNORA45	11	7.945	26.056	1.714
SNORD4B	17	6.81	20.336	1.578
SNORD103A	1	9.836	27.644	1.491
SNORD103A	1	9.836	27.644	1.491
SNORD69	3	32.914	88.017	1.419
SNORD12C	20	7.566	19.065	1.333
SNORA21	17	8.323	19.701	1.243

Table A.4.1.3(c): Up regulated snoRNAs in BRAF Wild Type A2780 Holoclones

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORA48	17	73.773	18.43	-2.001
SNORD88B	19	20.43	5.72	-1.837
SNORD88A	19	88.528	28.28	-1.646
SNORD88C	19	191.432	87.382	-1.131

Table A.4.1.3(d): Down regulated snoRNAs in BRAF Wild Type A2780 Holoclones

A.4.1.4 Colorectal

HT-29 BRAF V600E Mutant

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORA21	17	5.333	101.563	4.251
SNORD58A	18	2.584	39.603	3.938
SNORD71	16	5.388	66.111	3.617
SNORD91A	17	5.168	62.119	3.587
SNORD23	19	2.364	24.752	3.388
SNORD103A	1	3.464	30.98	3.161
SNORD104	17	16.878	142.283	3.076
SNORD85	1	5.828	43.436	2.898
SNORD38A	1	3.354	18.045	2.428
SNORA14B	1	4.013	19.802	2.303
SNORD66	3	40.573	194.502	2.261
SNORD82	2	13.469	62.119	2.205
SNORA36B	1	22.376	98.369	2.136
SNORD83A	22	7.862	34.174	2.12
SNORD91B	17	18.307	73.936	2.014
SNORD125	22	4.178	15.969	1.934
SNORD35B	19	19.407	70.583	1.863
SNORD15B	11	4.948	16.927	1.774
SNORD63	5	6.157	20.76	1.753
SNORD6	11	35.185	118.49	1.752
SNORD17	20	20.341	64.195	1.658
SNORD75	1	6.432	20.121	1.645
SNORD12C	20	7.147	20.76	1.538
SNORD1B	17	15.668	45.352	1.533
SNORD12	20	17.977	45.192	1.33
SNORD100	6	162.622	392.836	1.272
SNORA74A	5	6.707	15.49	1.208

Table A.4.1.4(a): Up regulated snoRNAs in BRAF V600E mutated Holoclones

COLO320 BRAF Wild Type

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD95	9	0	30.63	inf
snoU109	15	0	44.031	inf

SNORD1A	17	0.753	24.887	5.046
SNORD58C	18	5.274	67.004	3.667
SNORD58A	18	4.018	44.988	3.485
SNORD30	11	7.409	65.089	3.135
SNORD69	3	24.99	203.883	3.028
SNORD85	1	3.139	19.144	2.608
SNORD71	16	6.153	36.373	2.563
SNORD104	17	88.156	506.357	2.522
SNORD37	19	4.521	23.93	2.404
SNORD83A	22	11.804	62.218	2.398
SNORD35B	19	19.59	102.42	2.386
SNORD82	2	24.613	126.35	2.36
SNORD66	3	64.547	307.26	2.251
SNORD6	11	14.19	65.089	2.198
SNORA21	17	6.279	24.887	1.987
SNORD90	9	7.158	27.759	1.955
SNORD38A	1	7.911	29.673	1.907
SNORA14B	1	13.06	48.817	1.902
SNORD100	6	219.51	813.618	1.89
SNORD63	5	8.163	28.716	1.815
SNORD41	19	4.646	16.272	1.808
SNORD51	2	15.321	52.646	1.781
SNORD119	20	39.557	134.965	1.771
SNORA74A	5	6.656	22.016	1.726
SNORD91A	17	23.734	77.533	1.708
SNORD75	1	14.818	44.988	1.602
SNORD46	1	5.4	16.272	1.591
SNORA45	11	5.149	15.315	1.573
SNORA65	9	6.153	17.23	1.485
SNORD93	7	124.95	343.634	1.46
SNORD12	20	16.325	42.117	1.367
SNORD12B	20	16.325	42.117	1.367
SNORD99	1	30.139	76.576	1.345
SNORD17	20	11.804	28.716	1.283
SNORD1B	17	36.543	79.447	1.12
SNORA36B	1	35.79	73.704	1.042
SNORD59B	12	35.413	71.79	1.02
SCARNA6	2	12.307	24.887	1.016

Table A.4.1.4(b): Up regulated snoRNAs in BRAF Wild Type COLO320 Holoclones

snoRNA	Chromosome	Parent Count	Holoclone Count	Fold Change
SNORD88B	19	17.958	5.743	-1.645

Table A.4.1.4(c): Down regulated snoRNAs in BRAF Wild Type COLO320 Holoclones