

https://dspace.org/

Mechanisms of cancer initiation/progression: an investigation of chromosomal "hot spots" in South African cancer patients

Willem-Belot, Pascale Sylvie

2009-10-14T10:46:36Z
http://hdl.handle.net/10539/7352

Downloaded from a DSpace repository. https://dspace.org/

MECHANISMS OF CANCER INITIATION/ PROGRESSION

An investigation of chromosomal "hot spots" in South African cancer patients

Pascale Willem

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, in fulfillment of the requirements for the degree of Doctor of Philosophy

Johannesburg, 2008

DECLARATION

I, Pascale Willem declare that this thesis is the result of my own work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.

……..

……..day of ……..200

I wish to acknowledge that the RT PCR presented in chapter 3, section 3.4.1, page 90 was performed by Jacky Brown under my supervision and as part of a paper published in BMC Cancer.

DEDICATION

I dedicate this thesis to Alain and Benjamin who have provided me with refreshing philosophical conversations and musical moments and with constant love and support.

PUBLICATIONS AND PRESENTATIONS ARISING FROM THIS STUDY

Publications

Chromosome 12p Rearrangement and DNA Amplification Mapped by Comparative Genomic Hybridization in a patient with Secondary Myeloid Leukaemia. **P. Willem**, B. Mendelow. Cancer Genetics and Cytogenetics. 1997, 99(1): 30-37

A novel approach to simultaneously scan genes at fragile sites. **Willem P**, Brown J, Schouten J. BMC Cancer, 2006, **6**:205.

Research papers in preparation

A common cluster of stem cell genes commonly amplified in testicular germ cell tumor is involved in a secondary leukemia. Willem P, Kerkeart P.

High frequency of FHIT and WWOX intragenic deletions in esophageal squamous cell carcinoma from a high endemic region in South Africa. Willem P, Stepien A, Veale R, Brown J.

Published conference presentation

A translocation t(3;11) in an ovarian carcinoma cell-line inactivates the FHIT gene. **P Willem** and N Reckviassvhili. Chromosome Research, 2004, Vol 12, Supl 1, PO:08:64

Conference presentations

Fourth International Chromosome 12 Workshop, April 1997. Nice, France. Human Genome Project, Organized by R Kucherlapati, P Marynen and C Turc-Carel. Invited speaker " Amplification at 12p11-p12 in a secondary leukemia **" P Willem.**

10th Human Genome Meeting April, 2005, Kyoto, Japan "Scanning genes at fragile sites in South African esophageal cancer." **P Willem**, J Brown, J Schouten.
15th International Chromosome Conference (ICC XV). Brunel University, UK, 5-10 September 2004. A translocation t(3;11) in an ovarian carcinoma cell line inactivates the FHIT gene. **P Willem**, N Rechviasvhili.

AORTIC 2007, 6th International Cancer Conference. Cape Town, South Africa, 24-28 October 2007. FRA3B, FRA16D, FHIT and WWOX in South African esophageal cancer. **P Willem**, J Brown, R Veale, R Ally, M Hale, A Stepien.

ABSTRACT

The human chromosome complement contains nonrandom genomic regions that are prone to breaks and recurrently altered in tumors. These chromosomal “hot spots” are preferentially involved in early events of genomic instability and point to genes in their vicinity that may participate to carcinogenesis. Amplicons and deletions are frequently generated at “hot spots”, including common fragile sites (CFS), and are thought to host cancer genes whose rearrangements drive cell proliferation and promote the initiation and progression of cancer. Chromosomal “hot spots” in South African cancer patients, were investigated in this study with a view to characterizing underlying gene alterations.

A chromosome 12p amplicon was mapped and array comparative genomic hybridization (aCGH) was pioneered to identify candidate genes in the amplicon. Genes of the 12p stem cell gene cluster (NANOG, STELLAR and GDF3) were involved in striking similarity to what has been reported in testicular germ cell tumors and suggest that they may be more commonly involved in different types of cancer.

Two tumor suppressor genes, FHIT and WWOX, are located at the two most commonly expressed fragile sites, FRA3B and FRA16D respectively. Alterations in these fragile site associated genes have been reported in a variety of tumors including lung, esophageal, gastric, breast and cervical cancers most frequently as a result of submicroscopic deletions. Genomic deletions at CFS have been mostly investigated using loss of heterozygosity assays that do not necessarily inform on gene exon deletions. A new method was developed based on multiplex ligation-dependent probe

amplification (MLPA) that screens for exon deletions/amplifications of genes at CFS. The assay was validated on five esophageal squamous carcinoma cell lines and showed deletions in the FRA3B-associated gene FHIT in four of the cell lines.

Two geographically distinct South African cohorts of esophageal squamous cell carcinoma were then screened for FHIT/WWOX exon deletions and a visual basic (VBA) encoded program was written to automate MPLA products analysis. A high frequency of intragenic deletions in FHIT and/or WWOX (73%) was observed in the Eastern Cape cohort. FHIT deletions were seen in 27% of specimens from the Gauteng cohort, which by contrast did not show WWOX deletions. This difference may however reflect a difference in sampling collection.

The breakpoints of a translocation t(3;11)(p14;p15.1) present in an ovarian carcinoma cell line was characterized using the above MLPA assay, aCGH and the polymerase chain reaction. The translocation was found to interrupt the FHIT gene making it the 5th cancer associated translocation involving FHIT. The evaluation of gene relative copy number by aCGH and MLPA were highly correlated further validating the power of the MLPA assay in fresh tissue.

The involvement of critical genes at “hot spots” in SA cancer patients was high in the context of this study raising questions about the possible role of environmental exposure. The new MLPA assay may assist to expand the screening of critical genes at fragile sites in the future.

ACKNOWLEDGMENTS

I would like to thank all the people, who have contributed to make this work possible.

First to all the external collaborators:

Professor Andrej Stepien from Umtata General hospital for his warm and always supportive collaboration. Andrej has provided samples from Umtata and reviewed the histology of each case with me. **Doctor R Ally** and **Mr H Hassan** from the Gastroenterology unit at Chris Hani Baraguanath Hospital who have assisted with the collection of patients' samples. To **Professors Martin Hale** and **Wayne Grayson** from the Histopathology Department, National Health Laboratories Services (NHLS) who have reviewed the histopathology of blocks and touch preparations from the Gauteng cohort. I am grateful to **Professors Peter Lichter**, from Heidelberg Germany and **Philippe Kerkeart** from Lille France who have taught me CGH and CGH array respectively and accommodated me in their facilities, **Joe Gray** who donated the plasmid chromosome libraries, **Dr Jan Schouten** from MRC Holland for his very warm and friendly collaboration in the MLPA FHIT/WWOX kit and **Harry Staines** who so kindly assisted with statistics.

To the molecular cytogenetics unit:

I thank every member of this great team for handling the routine diagnostic activities with professionalism. With a particular thank you to **Hannelie Bothma**, who, through her excellent quality of work, has relieved me from many diagnostic and management activities making this work possible. To my students and scientists in the unit, who have left or are present. **Adrene Laubsher** and **Natela Reckviashvhili** whose work constituted the preamble to chapter 4. **Nienke Herbst** who assisted with some DNA

extraction and **Jacqueline Brown** for her dedication, refreshing joyful spirit, and for being a friend. Jacqueline actively works on esophageal cancer, which provided the opportunity of delightful times for scientific and non scientific discussions.

To the Department of Haematology and Molecular Medicine and in particular to the head of department **Professor Wendy Stevens** for her warm and infallible support. To **Professor Barry Mendelow** for his initial support and inspiring enthusiasm.

To **Vuyo Ranke** who has initially assisted with VBA programming.

Lastly, but mostly I wish to thank **Dr Adrian Puren** for undertaking the task of enrolling as my supervisor 'en route', I will always be grateful for your commitment, the time you took for this task and your gentle yet firm support.

ETHICS CLEARANCE

Number M02-05-47, Dr Pascale Willem, Department of Molecular Medicine and Haematology, University of the Witwatersrand, Johannesburg, South Africa.

Number 28-06-90, Dr Pascale Willem, Department of Molecular Medicine and Haematology, University of the Witwatersrand, Johannesburg, South Africa.

GRANT FUNDINGS

I wish to thank the NHLS Research Trust and the National Cancer Association of South Africa, CANSA, for funding this project.

Table of Contents

DECLARATION ii

DEDICATION iii

PUBLICATIONS AND PRESENTATIONS ARISING FROM THIS STUDY iv

ABSTRACT vi

ACKNOWLEDGMENTS viii

Table of Contents x

Table of figures xiii

List of tables xv

List of Abbreviations xvi

1 INTRODUCTION 1

1.1 A brief overview of cancer 5
1.1.1 Cancer is a genetic disease 5
1.1.2 Inherited genetic predispositions 5
1.1.3 Interaction between the host genome constitution and environmental factors 7
1.1.4 Acquired somatic mutations and cancer gene activation 8
1.1.5 Current molecular strategies aimed at characterizing cancer gene mutations 10
1.1.6 How the study of chromosomal "hot spots" can contribute to knowledge and management of cancer 12
1.2 Aims of this Thesis 15

2 CHARACTERIZATION OF A 12p AMPLICON BY COMPARATIVE GENOMICHYBRIDIZATION IN A PATIENT WITH SECONDARY MYELOID LEUKEMIA INVOLVEMENT OF THE STEM CELL GENE CLUSTER 17

2.1 Introduction 17
2.1.1 Amplicons in cancer 17
2.1.2 Chromosome aberrations and 12p in secondary leukemia 21
2.1.3 CGH arrays 22
2.2 Aims 23
2.3 Case History 23
2.4 Materials and Methods 24
2.4.1 Cytogenetic Analysis 24
2.4.2 DNA preparation 25
2.4.3 Chromosome suppression hybridization 25
2.4.4 Comparative genomic hybridization 26
2.4.5 Slot blot analysis 27
2.4.6 CGH array 28
2.5 Results 29
2.5.1 Cytogenetic analysis 29
2.5.2 Chromosome suppression hybridization 30
2.5.3 Comparative genomic hybridization 32
2.5.4 Slot blot analysis 34
2.5.5 CGH array 34
2.6 Discussion 39
2.6.1 Chromosome 12p breaks on both homologs and complex rearrangements 39
2.6.2 Chromosome 12p amplicons 40

2.6.3 Possible mechanisms of amplification 43
2.6.4 The 12p stem cell gene cluster as candidate target 45

3 DEVELOPMENT OF A NOVEL APPROACH TO SCANNING GENES AT FRAGILE SITES 46

3.1 Introduction 46
3.1.1 Fragile sites 46
3.1.2 Fragile sites and replication 48
3.1.3 Fragile sites are sensitive to external carcinogens 49
3.1.4 Pathways and genes that control DNA repair at fragile sites 49
3.1.5 Fragile sites and cancer 51
3.1.6 Tumor suppressor genes at fragile sites 52
3.1.8 Methods of investigation of genes at fragile sites 55
3.2 Aims 56
3.3 Materials and Methods 58
3.3.1 Cell lines 58
3.3.2 DNA preparation 58
3.3.3 MLPA 59
3.3.4 Experimental data analysis 62
3.3.5 Fluorescence *in situ* hybridization (FISH) 63
3.3.6 RT-PCR 64
3.4 Results 65
3.4.1 Amplification pattern in controls 65
3.4.2 Comparison of FHIT and WWOX DNA exon copy number in cell lines and controls by MLPA 67
3.4.3 RT-PCR 78
3.5 Discussion 80

4 ASSESSMENT OF GENES AT FRAGILE SITES, FHIT AND WWOX, IN SOUTH AFRICAN PATIENTS WITH ESOPHAGEAL SQUAMOUS CELL CARCINOMA 84

4-1 Introduction 84
4.1.1 World epidemiology of Esophageal Squamous Cell Carcinoma 84
4.1.2 Genes previously investigated in South African ESCC tumor tissues 90
4.1.3 Fragile sites and ESCC 91
4.2 Aims 93
4.3 Materials and Methods 94
4.3.1 Patient samples 94
4.3.2 DNA Extraction 95
4.3.3 MLPA 97
4.3.4 Data Analysis 97
4.3.5 FISH 100
4.4 Results 101
4.4.1 Patient samples 101
4.4.2 Evaluation of FHIT /WWOX deletions in fresh tumor biopsies 102
Gauteng cohort 102
4.4.3 MLPA amplification profile consistency in paraffin-embedded specimens 103
4.4.4 FHIT/WWOX deletions in FFPE samples collected in the Eastern Cape 107
4.5 Discussion 125
4.5.1 Specimens collection 125
4.5.2 Strength and weakness of the FHIT/WWOX MLPA assay 126
4.5.3 Deletions of FHIT and WWOX in each respective cohort 128
4.5.4 Additional chromosome 16 deletions 129

5 CHARACTERIZATION OF A TRANSLOCATION INVOLVING THE FRA3B IN AN OVARIAN CARCINOMA CELL LINE 132

5.1 Introduction 132
5.2 Aims 138
5.3 Materials and Methods 138

5.3.1 Cell cultures 138
5.3.2 DNA extraction 138
5.3.3 MLPA 139
5.3.4 RNA extraction 139
5.3.5 RT-PCR for FHIT transcripts 139
5.3.6 3' Rapid amplification of cDNA ends (RACE PCR) 140
5.3.7 CGH array 141
5.4 Results 142
5.4.1 MLPA 142
5.4.2 RT-PCR 144
5.4.3 3' RACE PCR 145
5.4.4 CGH arrays 146
5.4.5 MLPA validation by CGH in cell line UWOV2 148
5.4.6 Status of WWOX at FRA16D in the UWOV2 cell line 151
5.6 Discussion 154

6 GENERAL CONCLUSIONS AND FUTURE WORK 157

6.1 Summary 157
6.1.1 Characterization of a 12p amplicon in a secondary leukemia patient 157
6.1.2 Characterization of fragile sites deletions affecting tumor suppressor genes 158
6.1.2 A cancer initiation/progression model involving fragile sites 159
6.2 Future work 161
6.2.1 Development of a FISH probe to detect amplification of the 12p stem cell gene cluster 161
6.2.2 Improvement of the MLPA assay 161
6.2.3 Furthering the investigation of the role of genes at fragile sites in SA ESSC 162
REFERENCES 165
APPENDICES 186

Appendix 1: list of fragile sites in the human genome 186

Appendix 2: MLPA Probe binding sequences: 188

Appendix 3: FHIT coding sequence and set of primers 190

Appendix 4: Heat maps representing MLPA probes ratios 192

Appendix 5: composite karyotype of cell line UWOV2 195

Appendix 6: FHIT and WWOX sequences corresponding to aCGH array 196

List of figures

Figure 1.1: Chromosomal "hot spots": chromatid breaks at fragile sites are shown here. Breaks at these sites can promote both the formation of amplicons or critical gene deletions**.** 4
Figure 1.2: Model showing the respective contribution of environmental (in green) and genetic in (blue) factors in cancer development. 13
Figure 2.1: Schematic representation of breakage-fusion bridge cycle. 19
Figure 2.2: G-Banded karyotype of a metaphase from the patient with secondary leukemia. Ring chromosomes are marked with arrows. The two derivatives of the t(3;12) are marked with arrowheads**.** 30
Figure 2.3: (A) Partial metaphase spread from a tumor cell counterstained with propidium iodine after chromosome 12 painting (FITC detection). 31
Figure 2.4: A control metaphase subjected to CGH shows the FITC fluorescence of hybridized tumor DNA sequences, left, and the rhodamine fluorescence of control reference female DNA hybridized to the same cell showing a homogenous staining of chromosomes, right. 33
Figure 2.5: Slot blot obtained with YAC 952a6 hybridized to an increasing amount of patient and reference DNA and showing homogenous signal intensity on both samples. 34
Figure 2.6: Ideogram of chromosomes 5, 7 and 12 with the corresponding hybridization signal intensity mean log ratio (blue line) of patient DNA versus reference DNA. 35
Figure 2.7: Genomic composition of the 12p13.31 amplicon. 38
Figure 2.8: Ideogram of chromosome 12, depicting on the left the number of genes corresponding to each cytogenetic sub-band. 42
Figure 3.1: Replication fork. While helicase unwinds the DNA, Polymerase α (grey-blue circle) leads the replication process on the lagging strand. 48
Figure 3.2: ATR-dependent cell cycle checkpoint that controls fragile site stability. 50
Figure 3.3: Principle of MLPA. The probe construct at the top shows target-specific hemiprobes in green, the stuffer sequence of variable length in blue and the universal MLPA primers in red. 60
Figure 3.4: (A) Mean "reference control" pattern of probe amplification established from 10 normal individual samples. Each probe is represented on the horizontal axis in order of its size.. 66
Figure 3.5: (A) The probes' fluorescent intensity is plotted on the Y-axis and the probes are ordered by size on the X-axis in the same order as on the Excel graph shown in Figure 3.4. Cell line WHCO1 raw data are represented with black arrows to indicate each FHIT-specific probe. The pink arrows indicate the WWOX probes. 68
Figure 3.5: (B) Data for cell line WHCO3 on the opposite did not show any major probe amplification differences (only probes for FHIT exons 4 and 5 have been indicated by black arrows for simplicity). 69
Figure 3.5: (C) Cell line WHCO6 showed significantly decreased signals for FHIT exons 4 and 5, indicated by black arrows. Cell line SNO had a homozygous deletion for FHIT exon 5 (orange arrows). 70

Figure 3.6: Comparisons of the probe peak heights between the cancer cell line WHCO5 (in dark grey) and the reference control (in white). Hemizygous deletions for FHIT exons 4 and 5 can clearly be seen. 71
Figure 3.7: Comparisons of the relative probe peak heights between cancer cell lines (in dark grey) and control (in white). 73
Figure 3.8: Graphic representation of FHIT and WWOX exon copy number in each cell line 76
Figure 3.9: (A) FISH with a 3p specific probe (red and arrowed) in each cell line… 77
Figure 3.10: (A) A 2% agarose gel, showing RT-PCR expression of FHIT (full transcript) in the 5 ESCC cell lines and a control (C1) 79
Figure 4.1: Potential combined effect of exogenous carcinogens in targeting fragile sites. 93
Figure 4.2: Positions of MLPA internal control probe CREM used for normalization and location of the validation FISH probe at 10p11 100
Figure 4.3: Graphic representation of FHIT and WWOX exon copy number in four primary ESCC tumors. 103
Figure 4.4: Comparison of pattern of amplification between fresh and FFPE control specimens. 104
Figure 4.5: Consistency of MLPA across varying extraction procedures. 105
Figure 4.6: Differing concentrations of starting DNA from a FFPE sample that had a high number of aberrations. 106
Figure 4.7: MLPA GeneScan raw data showing deletions of WWOX and chromosome 16 probes in the Eastern Cape FFPE specimens. 109
Figure 4.8: GeneScan raw data on MLPA results for A, a FFPE specimen that did not show any significant deletions and B, for a FFPE specimen with deletions in the FHIT locus. 110
Figure 4.9: Tumor sections on paraffin hybridized to chromosome 10p11 probe (red signals) on two FFPE specimens. The majority of cells have two signals, validating the choice of the internal control probe used for normalization. 111
Figure 4.10: Input page 113
Figure 4.11: Instruction page 114
Figure 4.12: Reference control 116
Figure 4.13: Test DNA result 117
Figure 4.14: Heat map representation of exons' copy number changes for genes at fragile sites in 17 specimens. 119
Figure 5.1: G banded karyotype of UWOV2 cell line at passage 50. 135
Figure 5.2: Summary of BACs and YACs mapping studies on the derivative (11) t(3;11). 136
Figure 5.3: FISH mapping data on der(11)t(3;11) in cell line UWOV2. Modified from Rechviasvhili, 2002. 137
Figure 5.4: MLPA results of cell line UWOV2 with the FHIT-WWOX probe mix. 143
Figure 5.5: 2% agarose gels showing the FHIT-specific RT-PCR results obtained with the UWOV2 cell line. 145
Figure 5.6: CGH array results for the UWOV2 cell line 147
Figure 5.7: Chromosome 10p control probes, correlation of MLPA and CGH array data. 150
Figure 5.8: Correspondence of MLPA and CGH array results for the control probe on chromosome16q, CDH13 gene. 151

Figure 5.9: Correlations of MLPA and CGH results in cell line UWOV2, showing deletions in the WWOX and FANCA loci on chromosome 16 q. 153
Figure 6.1: Model for the formation of malignant clones through deletions at fragile sites and DSB generated at the same time. 160
Figure 7.1: (A) aGCH array results obtained with the UWOV2 cell line showing the sequence information of oligonucleotide probes specific for the FHIT locus...197
Figure 7.1: (B) The FHIT gene exon/intron boundaries and sequence composition from the NCBI reference is shown……...198
Figure 7.2: (A) Hhybridization results on CGH analytics for the UWOV2 cell line. ……199
Figure 7.2: (B) The WWOX gene exon/intron boundaries and sequence composition from the NCBI reference sequence is shown…… 200

List of tables

Table 1-1: *Interactive events in cellular transformation*..... 8
Table 3.1: *Summary of FISH, MLPA and RT-PCR results for FHIT on the five ESCC cell lines***.** 80
Table 4.1: *bin sets attributed to MLPA signal amplitude ratios (tumor/reference control)* 100
Table 4.2: *Number of FFPE cases showing deletions in the respective fragile sites-associated genes, FHIT, WWOX, CAV1 and CAV2* 121

List of Abbreviations

AF B1	aflatoxin B1
AML	acute myeloid leukemia
APC	adenomatous polyposis coli
BAC	bacterial artificial chromosome
BFB	breakage-fusion bridge
CAV	caveolin gene family
CCD	cooled, charged couple device
CDH1	E-cadherin gene
CDH13	H-cadherin gene
CEPH	Centre d'Etudes du Polymorphisme Humain
CFS	common fragile sites
CGH	comparative genomic hybridization
CML	Chronic Myeloid Leukemia
DME	Dulbecco's Modified Eagle's Medium (DME)
EGFR	epidermal growth factor receptor
ESCC	esophageal squamous cell carcinoma
FB1	fumonisin
FBS	fetal bovine serum
FCS	fetal calf serum
FFPE	formalin-fixed, paraffin-embedded
FHIT	fragile histidine triad
FISH	fluorescence *in situ* hybridization
GDB	genome database
HBV	hepatitis B virus
HPV	human papilloma virus
IHC	immunohistochemical assay
IR	ionizing radiation
LOH	loss of heterozygosity
MLPA	multiplex ligation-dependent probe amplification
MLV	murine leukemia virus
MDS	myelodysplastic syndrome
OSCC	oral squamous cell carcinoma
PAH	polycyclic aromatic hydrocarbons
PAS	probe analysis software
RT-PCR	reverse trancriptase polymerase chain reaction
SCE	sister chromatid exchange
STD	Standard deviation
SULT1A1	sulfotransferase phase II detoxification enzyme
TGCT	testicular germ cell tumor
TF	transcription factor
TOC	Tylosis esophageal cancer
UV	ultra violet
VBA	Visual Basic Editor
WT	Wilms tumor
WWOX	WW domain containing oxidoreductase
YAC	yeast artificial chromosome

CHAPTER 1

INTRODUCTION

Within the last 20 years research in cancer has enabled considerable progress in disease management and early diagnosis. This has occurred primarily through the identification of predisposing gene mutations and molecular markers that also serve as therapeutic targets. Cancer as a disease, remains difficult to define as a single entity; most investigators view it as a collection of different diseases (Vogelstein and Kinzler 2004), a view which is supported by the discovery of specific molecular genetic signatures in different types of cancer (Herbst and Lippman 2007; Poulsen, et al. 2005). However, some common features are shared by all cancer cells, namely cellular functions such as cell renewal, cell proliferation and in particular, genetic instability are always altered. This implies the common participation of fewer biological pathways/functional systems than genes that are shared by all cancers, among which the DNA repair system plays a central role.

It has also become clear that cancer arises within the context of the interplay between the host genetic constitution and its environment. While inherited genetic predispositions that compromise a key gene (e.g. the adenomatous polyposis coli (APC) tumor suppressor gene in colon cancer) are known to lead to a sequence of molecular events that result in malignant transformation (Bertholon, et al. 2006), there is at the same time, mounting evidence to suggest that carcinogens target preferential regions of the genome, or chromosomal "hot spots" (Tsantoulis, et al. 2007). It is speculated that once altered, these "hot spots" may subsequently facilitate or participate in the steps that promote

carcinogenesis. These regions are common to all individuals and can be rearranged in a variety of cancers, raising the notion that they could represent a common feature in many cancers (Bartkova, et al. 2005). At the centre of the interplay between host and environment, these "hot spots" reflect both the exposure to external toxic factors and the level of efficiency of the DNA repair mechanism inherent to an individual genetic constitution.

From a molecular cytogenetic perspective, we understand how a chromosomal translocation activates (or, less frequently, deactivates) a "key" gene/oncogene through gene rearrangement. The Philadelphia chromosome translocation t(9;22)(q34;q22) and its resulting chimaeric gene product BCR-ABL, for example, results in a constitutive activation of the ABL oncogene tyrosine kinase activity (Randolph 2005) and epitomizes this mechanism. This translocation is strongly associated with Chronic Myeloid Leukemia (CML) (Randolph 2005). We also understand how chromosomal deletions can eliminate one copy of a tumor suppressor gene and facilitate steps for the inactivation of the other allele on the chromosome homolog. Chromosomal unstable regions, or "hot spots" on the other hand, have been more elusive with regard to their role in cancer initiation and progression since they could be byproducts of genetic instability without any down stream functional consequence. The presence of key genes located within these loci has changed this notion.

Two types of chromosomal "hot spots" are examined in this thesis: chromosomal amplification (amplicon) and chromosomal fragile sites.

Amplicons are genomic regions increased in copy number and undergoing cycles of amplification with subsequent cell divisions. They are common in solid tumors (Vogt, et

al. 2004). Their mechanism of formation is unclear, but recent evidence points to the involvement of neighboring unstable genomic regions, many of which (but not all) correspond to fragile sites. Importantly, it is thought that the mechanism of DNA amplification is driven by a key gene, or cluster of genes, that confers a proliferative cell advantage (Miller, et al. 2006). These key genes have both diagnostic and therapeutic values and they often constitute valid targets for therapy.

The HER2 amplicon for example, results in an increased HER2 gene copy number in ±30% of breast cancer patients. HER2 is an epidermal growth factor receptor that becomes constitutively activated as a result of amplification. HER2 amplification is associated with a more aggressive disease (Sawyers 2005). The specific antibody-derived therapy, Herceptin, targets the HER2 oncogene and is used successfully in patients who have HER2 amplification. The effectiveness of this drug underscores the importance of identifying new cancer target genes contained in amplicons.

Fragile sites (Figure 1.1) often delineate the boundaries of amplicons, but are also in themselves the site of frequent genomic deletions. They were originally identified as chromosomal regions that form gaps under culture conditions of replication stress and they are sensitive to external carcinogens (Yunis, et al. 1987). Several tumor suppressor genes map at the location of the most common fragile sites (Huebner and Croce 2001), but the nature of their role in carcinogenesis has not been completely elucidated. The main method of investigation of genomic aberrations at fragile sites has been loss of heterozygosity. However this method does not inform on the exon composition of genes localized at these loci and hence on the functional consequences of deletions observed in some tumors. While the most common fragile sites in cancer have been investigated in

other parts of the world, they have not been studied in South Africa (SA) to a great extent. In particular, they have not been investigated in esophageal squamous cell carcinoma (ESCC), known to have a high geographical prevalence in the Eastern Cape, the incidence of which is strongly associated with external carcinogens exposure.

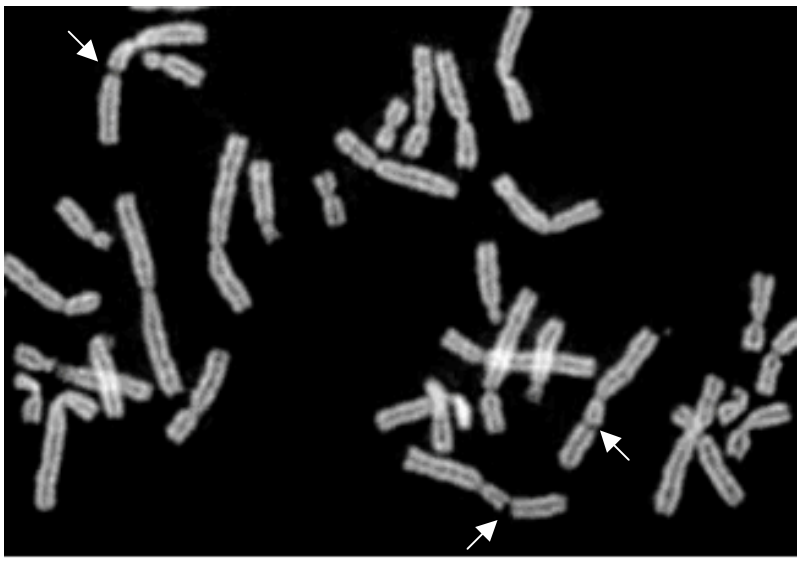


Figure 1.1: *Chromosomal "hot spots": chromatid breaks at fragile sites are shown here. Breaks at these sites can promote both the formation of amplicons or critical gene deletions.*

A brief and general overview of the main lines of cancer research follows to contextualize this work. Amplicons and fragile sites are reviewed in their respective chapters.

1.1 A brief overview of cancer

1.1.1 Cancer is a genetic disease

Cancer is a "genetic disease". This view is now generally accepted, as emphasized in a review (Vogelstein and Kinzler 2004). It signifies that every cancer cell carries some acquired or inherited genetic aberrations. Over the last 15 years, a great deal of research effort has gone into the identification of inherited predisposing gene mutations responsible for a host of familial cancers.

1.1.2 Inherited genetic predispositions

The study of familial cancer cases has promoted the discovery and characterization of several key genes, which, as single polymorphic alleles, confer a high risk of malignancy. These polymorphisms are rare functional alleles and have a high phenotypic penetrance. The most common genes involved include: BRCA1 and BRCA2 in breast and ovarian cancer; APC, MLH1, MLH6 and MSH1 in colon cancer; VHL in renal cancer; p53 in Lee Fraumeni syndrome; RB in retinoblastoma; and XP in Xeroderma Pigmentosum, to cite the main ones (Marsh and Zori 2002). Their detailed investigation has brought considerable new knowledge regarding the mechanisms, molecular pathways and stages of malignant transformation. However, the majority of cancers are not explained by familial mutations. While additional predisposing mutations are likely to be uncovered, they will probably not account for more than 30 percent of all cancers (Verkasalo, et al. 1999).

A different model for the involvement of genetic factors in "sporadic" cancers was proposed. It was thought that several polymorphic alleles may together influence cancer susceptibility, acting in synergy (Pharoah, et al. 2002; Ponder, et al. 2005); under this view cancer was coined a polygenic disease. This concept is supported by mouse models (Nagase, et al. 1999), where the co-participation of several loci together has been shown to confer resistance to malignant induction. There is no mechanistic explanation for most of these gene allele associations, and they do not provide any answer as to the cause of cancer. The co-participation of multiple alleles is therefore likely to remain difficult to assess with respect to their associated risk, especially if new mutations have recently arisen at different gene sites, as suggested by Balmain (Balmain, et al. 2003). These studies require numerous and complex analysis of haplotype associations in large population groups, backed up with mouse experimental models. While the results will clarify which genes may participate in carcinogenesis and may signal higher risk individuals, it will remain difficult to know if, how and when to intervene for these individuals.

Given these circumstances, and taking into consideration that the environment is an important induction factor, including in hereditary cancer, it will be important to clarify the part played by toxic exposures, (ionizing irradiations, viruses, tobacco, chemicals and other toxins) (Table 1.1), and evaluate the impact on the genome stability. This should significantly contribute to preventing cancer and developing adapted screening strategies.

1.1.3 Interaction between the host genome constitution and environmental factors

As stated earlier, the altered genetics in cancer cells results from the dynamic interactions between a host's genetic susceptibility and carcinogen exposure. Some external carcinogens are well known, such as smoking, ionizing radiations, toxic chemicals, (such as benzene) and viral infections, to cite the main ones. Smoking has been associated with an increased expression of fragile sites (Stein, et al. 2002). Other carcinogenic sources have been revealed in the cellular microenvironment. Hypoxia in the cellular microenvironment for example, affects fragile genomic regions and leads to gross genetic aberrations (Coquelle, et al. 1998). *In vitro* oncogene induced replication stress was also shown to preferentially target fragile sites (Tsantoulis, et al. 2007) and it would be interesting to assess if a similar stress induction may be involved in Neuroblastoma where neighboring non-cancerous cells can trigger cell transformation (Wu, et al. 2005; Zhu, et al. 2002). The table below illustrates the relationship between genotoxic factors, their impact on the genome, and the resulting biological consequences that lead the cell into a malignant phenotype**.**

Table 1-1: *Interactive events in cellular transformation*

Genotoxic factors	*Microcellular environment*	*Gene rearrangements*	*Biological cellular consequences*
External agents: *Ionizing radiation (IR), ultra violet (UV)* *Smoking, alcohol consumption, poor nutrition* *Toxic exposure (benzene)* *Viruses (HPV, HBV)* *Bacterial infections* *Internal agents:* *Genetic predispositions: Rare high penetrance allele or polygenic susceptibility* *Reactive oxygen species (ROS)* *Hormonal imbalances*	*Ageing* *Angiogenesis* *Hypoxia* *Stromal cells*	*Oncogene activation* *Tumor suppressor inactivation* *Alteration in genes monitoring genome stability* *Telomerase activation* *Epigenetic changes (methylation pattern)* *Loss or amplification of miRNA* *Mitochondrial alterations*	*Activation of cell proliferation/ survival mechanisms* *Block in differentiation* *Deficient cell-cycle check points* *Inhibition of apoptosis*

1.1.4 Acquired somatic mutations and cancer gene activation

Whether sporadically acquired or as a result of predisposing inherited mutations, a variety of differing nucleic acid aberrations can occur in cancer. The only meaningful genetic events in cancer cells are those activating key "gene(s)", which will provide the cell with a selective growth advantage and initiate clonal expansion. The nature of aberrations ranges from single point mutations and epigenetic modification to chromosomal deletions, amplifications, numerical aberrations and complex translocations. The term "somatic mutation" generally refers to any of these rearrangements, provided they occur during mitosis in somatic tissues and not in germ cells. Chromosomal aberrations in particular, provide powerful mechanisms of oncogene activation and tumor suppressor inactivation.

Oncogenes can be activated by simple transposition to another chromosome location that places it under the control of an active promoter and drives its over expression. This is the case for B cell lymphomas, where a variety of oncogenes are transposed and then fall under the control of the immunoglobulin gene heavy chain (IgH) promoter (in 14q23), which is actively transcribed in B cells (Butzler, et al. 1997; Martin-Subero, et al. 2005). Alternatively, a translocation can create a novel chimaeric gene made of two different gene regions and that acquire transforming properties; this is the case for the tyrosine kinase activity of the Philadelphia chromosome translocation mentioned earlier. Many translocations associated with acute leukemia work in this way and involve transcription factors (TF) that reciprocally swap their "activation" and "DNA binding" domains. This alters the expression of all responding downstream genes with major cellular consequences (Rowley 1998). Tumor suppressor genes can be inactivated by simple chromosomal deletion**,** point mutation, interstitial deletion, or translocation and subsequent inactivation of the other allele. This is the case for the Retinoblastoma gene for example (Kuchino 1994). Other genes involved in cancer normally function by maintaining the genome stability. Their role is to either repair DNA (due to replication errors, or after exposure to mutagens). They also ensure the control of the cell cycle checkpoints. Many inherited predisposing genes belong to this category and their alterations can also be acquired. BRCA1, for example, is involved in a DNA repair mechanism. Mutations that attenuate this function result in an increased risk of cancer.

The role of chromosomal "hot spots", on the other hand, has not been completely clarified. While it has long been recognized that specific chromosomal regions are more prone to breaks than others (Yunis 1988; Yunis 1990), their functional participation in malignancy has remained unclear until recently. Fragile sites that manifest as chromatin breaks and are

the site of frequent deletions, may occur as a mere consequence of genetic instability in cancer cells. The identification of several tumor suppressor genes at fragile sites (Smith, et al. 2007) and driver (onco)genes in amplicons (Myllykangas, et al. 2006) has recently highlighted their functional importance in either initiating or facilitating the process of malignancy. The recent report of non random involvement of genes at fragile sites in different cancers reinforces this idea (McAvoy, et al. 2007).

1.1.5 Current molecular strategies aimed at characterizing cancer gene mutations

Apart from linkage and familial studies conducted in inherited forms of cancer, the main genome-wide approaches to identifying molecular markers, signatures and potential target genes in tumors have expanded considerably, (Genome-Wide Association Studies (GWAS, http://grants.nih.gov/grants/gwas). Two general approaches make use of biochemical knowledge, while a third uses a whole genome approach for expression and copy number analysis.

a. Identification of biochemical and related molecular pathways

The molecular dissection of inherited cancer-predisposing mutations and the study of related tumor genetic aberrations have revealed several functional "molecular pathways" at work in the steps to cell transformation. In Vogelstein's words: "there are many fewer pathways than genes" (Vogelstein and Kinzler 2004). The alteration of any of several genes that belong to the same pathway leads to the same phenotypic consequence; thus, an increasing knowledge regarding the genes involved in these pathways and the nature of

their role, facilitates in turn the functional characterization of new candidate cancer genes involved in the same pathway.

b. Sequencing genes encoding protein kinases in cancers

Another large-scale approach to probing for "cancer driver genes" has recently been reported. It involves the systematic sequencing of all known protein kinases, screening for somatic point mutations present in a variety of tumors. Based on the assumption that only non-synonymous mutations (altering the gene protein products) would respond to a biological selection (mutation promoting cell growth), a higher than expected ratio of non-synonymous mutations versus synonymous mutations in 120 genes provided evidence for their contribution to carcinogenesis (Greenman, et al. 2007).

c. Molecular signatures

The use of micro-arrays in oncology has revealed several cancer-specific "signatures" in differing types of cancer. First expression signatures, with expression arrays (Valk, et al. 2005), and more recently copy number profiles, obtained with comparative genomic hybridization (CGH) arrays (Rubio-Moscardo, et al. 2005), have been characterized. The complexity of available information and the number of alterations observed in tumors poses a challenge to efforts to unravel both the sequence and meaning of specific genetic events involved in the process of carcinogenesis.

The characterization and screening of non-random chromosomal breakpoints

remains a powerful approach for the mapping and cloning of oncogenes and tumor suppressor genes (Rowley 1998), yet, very few studies have adopted a genome–wide approach with regards to chromosomal "hot spots" alterations in cancer. Three recent

studies stand out that confirm the early and nonrandom involvement of genes located in these regions (McAvoy, et al. 2007; Myllykangas, et al. 2006; Tsantoulis, et al. 2007).

1.1.6 How the study of chromosomal "hot spots" can contribute to knowledge and management of cancer

Environmental carcinogens and induced replication stress can directly affect "breakage-prone genomic regions" (Bartkova, et al. 2005; Gorgoulis, et al. 2005; Stein, et al. 2002; Tsantoulis, et al. 2007; Yunis, et al. 1987). An increased rate of breaks at these sites, may therefore either reflect significant exogenous exposure, exceeding the cell repair capacity, or a genetically deficient host repair pathway (particularly affecting the single strand DNA-break repair pathway). Both environmental and genetics factors are in essence co-participating in the process of carcinogenesis, with a respective contribution; the figure below illustrates this interplay (Figure 1.2).

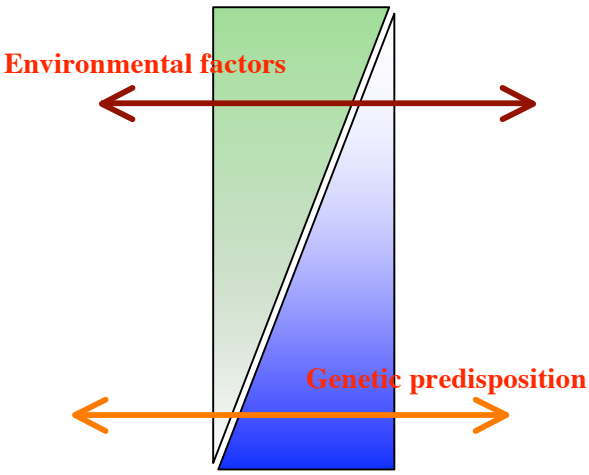


Figure 1.2: *Model showing the respective contribution of environmental (in green) and genetic in (blue) factors in cancer development.*
The green triangle represents a gradation in the contribution of environmental factors; all cancer at the top of the figure represent sporadic cancers caused by heavy mutagenic exposure (i.e. massive exposure to ionizing radiations, for example); at the bottom would be most of the rare inherited predisposing alleles associated with a high risk of malignancy. With respective gradation of both genetic (in blue) and environmental factors (in green) along the vertical axis of this figure, most cancers would fit somewhere in the middle, where the dynamic interaction between the host genetic susceptibility and environmental DNA damaging agents is critical.

The study of DNA damage at chromosomal "hot spots", interpreted together with epidemiological data, may signal the presence of environmental toxic exposure. Whether this occurs in a multifactorial predisposing genetic background in a subpopulation group shown to have a higher incidence for a particular cancer, would then have to be determined by population-based studies, but the detection of an excess rate of mutations at these sites justifies preventive efforts that explore and address the environment. In particular, the extent of deletions at fragile sites in endemic regions with a high prevalence of cancers clearly needs to be ascertained.

Deletions and amplifications may develop at these chromosomal regions and when they persist in cancer cells, they are likely to signal the presence of genes whose function has

been mechanically altered to support malignant cell growth (Ciullo, et al. 2002; Huebner, et al. 1998b; Miller, et al. 2006). There is a strong need at present to identify or confirm valid cancer targets for their use as both markers of disease and potential targets for therapy, and the study of recurrent breaks/amplicons is instrumental to their identification. A number of amplicons have not yet been characterized for the presence of oncogenes. Deletions of tumor suppressor genes at fragile sites may need to be clinically evidenced for potential future target therapy. The gene FHIT at FRA3B has already been shown to have effective anti-cancer activity in experimental mouse models (Dumon, et al. 2001a).

There is the intriguing notion that early cancer events are accompanied by an increased fragile site expression in a variety of cancers (Bartkova, et al. 2005; Gorgoulis, et al. 2005; Tsantoulis, et al. 2007); this would qualify them as useful predictive markers. The main method used up until recently to detect DNA copy number changes at fragile sites has been loss of heterozygosity (LOH), which may not always reflect true deletion but homodisomy, and does not inform on gene aberrations at these sites. This technique may have underestimated the number of rearrangements at these spots. While genomic arrays will contribute significantly to these investigations they are expensive and less suited to clinical screening. A cost effective screening method based on relative gene copy number would assist in detecting functional lesions in relevant cancer patients.

1.2 Aims of this Thesis

The main goal of this project was to investigate chromosomal "hot spots" in South African cancer patients, with a view to characterizing underlying gene alterations, with an emphasis on esophageal cancer of high prevalence in SA. Multiplex ligation-dependent probe amplification (MLPA) was used to design and develop a new screening approach for genes at fragile sites. The use of Comparative Genomic Hybridization (CGH) array was pioneered locally to characterize breakpoints at fragile sites.

The following specific objectives summarize the project and serve as a structure for this thesis:

1. To characterize a chromosome 12p amplicon in a secondary leukemia of iatrogenic origin and identify the potential driver genes contained in this amplicon. The rationale was that the high recurrence breakpoint at 12p and the overlap with an amplicon described in germ cell tumors suggested the presence of common driver genes (Chapter 2).

2. To develop and evaluate a new molecular screening test to detect DNA alterations affecting exons in genes at fragile sites. This was performed using the MLPA technique (Chapter 3).

3. To assess the exon copy number of two genes (FHIT and WWOX) at the two most commonly expressed fragile sites, FRA3B and FRA16D, in two geographically separate South African cohorts of esophageal squamous cell carcinoma (Chapter 4).

4. To automate the analysis of multiple probe fluorescence data for the MLPA test

developed and make suggestions for future improvements to the test (Chapter 4).

5. To characterize a translocation t(3;11)(p14; p15.1) mapping at the fragile site FRA3B and to assess if the breakpoint interrupted the tumor suppressor gene FHIT localized in this locus (Chapter 5).

Chapter 6 includes concluding remarks and outlines future work directions.

CHAPTER 2

CHARACTERIZATION OF A 12p AMPLICON BY COMPARATIVE GENOMIC HYBRIDIZATION IN A PATIENT WITH SECONDARY MYELOID LEUKEMIA INVOLVEMENT OF THE STEM CELL GENE CLUSTER

2.1 Introduction

2.1.1 Amplicons in cancer

Amplicons are discreet chromosomal regions that are dynamically amplified during successive somatic cell divisions to varying degrees of additional copy numbers. They result in the amplification of genes located in this region and usually correlate with an increased gene expression (Hyman, et al. 2002), although there are a few exceptions in which the amplified genes are not necessarily over-expressed (Platzer, et al. 2002). The mechanisms underlying DNA amplification are not fully understood, but at least two different modalities have been well characterized. First, the majority of intra-chromosomal amplicons are organized in large inverted repeats (Hellman, et al. 2002), and this structure suggests that they occurred through "breakage-fusion bridge" (BFB) cycles (Figure 2.1). This BFB mechanism, first described by Barbara McClintock (Mc Clintok, 1951), is probably the commonest in terms of amplification. It has since been observed and characterized by molecular cytogenetic techniques in early amplification events that occur in *in vitro* induced experimental models (Toledo, et al. 1992). On the other hand, extra chromosomal amplicons such as "double minutes" have been shown to originate from the excision of a single DNA molecule, followed by unequal segregation (Vogt, et al. 2004).

The molecular characterization of other amplicons suggests that there are other mechanisms more rarely involved in DNA amplification events; these have not yet been elucidated (Kuwahara, et al. 2004).

Amplicons are frequently observed in solid tumors and often correlate with an increased copy number and consequent over-expression of key oncogenes that map in the region. The in vitro experiments cited above, where the amplification of genes that confer drug resistance were induced (Toledo, et al. 1992), support the following scenario: a key gene, whose activation confers a cell growth advantage, is selected for during the successive cell divisions and drives the process, hence the amplicon. This is illustrated in Figure 2.1, which also depicts the BFB cycles mechanism.

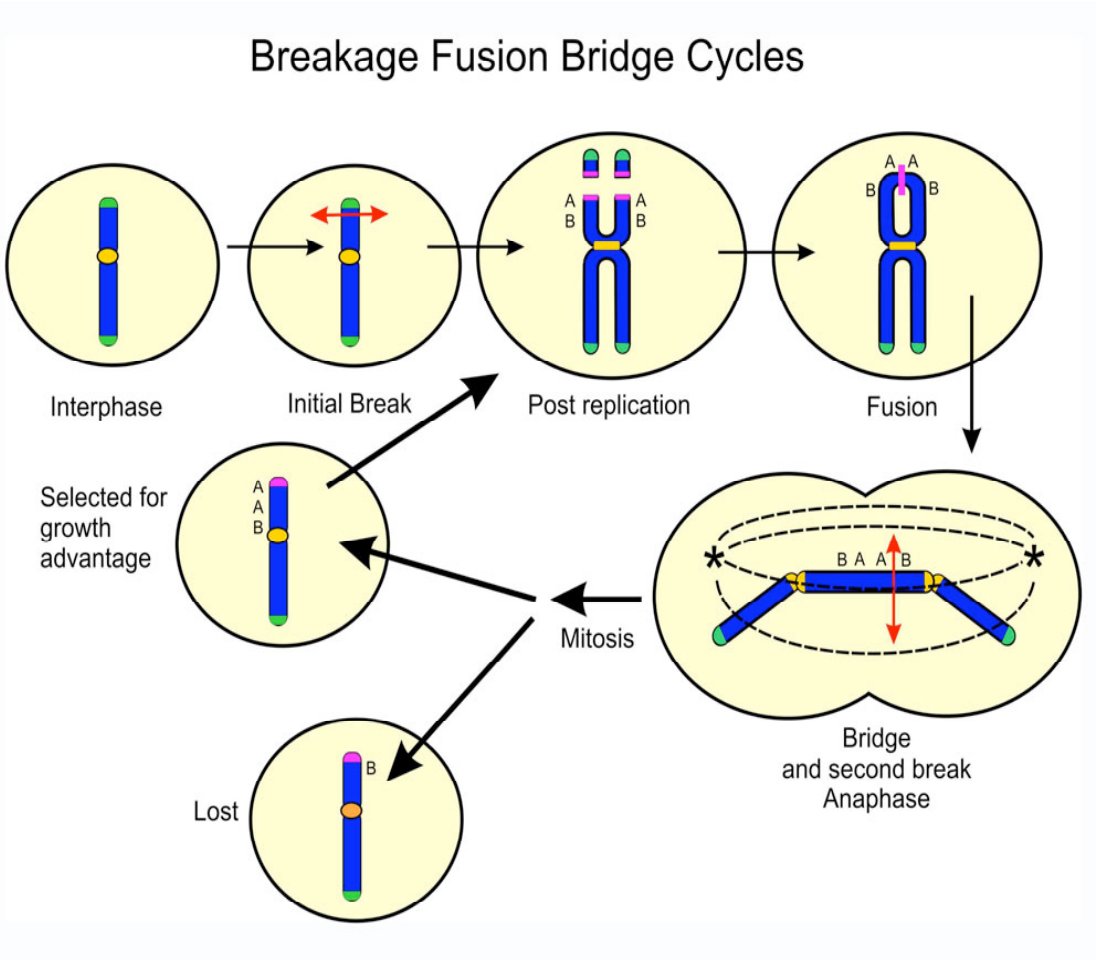


Figure 2.1: *Schematic representation of breakage-fusion bridge cycle. Chromosomes are shown in blue, centromeres in orange and telomeres in green. Breaks are shown with red double arrows and chromatid loose ends are pink. The amplified region carries a key driver gene A. Top left: initial break which, after DNA replication, results in two sister chromatids with loose ends, uncapped by telomeric sequences. Top right: Fusion of the uncapped sister chromatids forming a dicentric chromosome. Bottom right: the dicentric forms a bridge over the mitotic spindle, and is pulled in opposite directions by the two centromeres, each of which migrates to opposite poles of the cell. This leads to a second break at the anaphase stage of cell division and in two daughter cells with both uncapped chromosomes and an unequal copy number representation of the driver gene(s) A. The cell carrying more copies of the driver gene is selected for. Chomatids, with loose ends in pink, will fuse again, leading to another BFB cycle during the next cell division.*

This BFB scenario is at play in an amplicon commonly observed in gliomas where the epidermal growth factor receptor (EGFR) is the driver oncogene (Vogt, et al. 2004). This mechanism is also highly likely in the HER2 oncogene process of amplification which is observed in 30% of breast cancer patients (Mano, et al. 2007). Both EGFR and HER2 oncogenes are involved in a variety of cancers where their recurrent amplification status has led to the development of specific antibodies targeting their respective protein products for therapeutic intervention. Both "target therapies" are now successfully used in cancer patients who are positive for the respective intra-chromosomal 7q (EGFR) and 17q (HER2) amplicons (Mano, et al. 2007; Ramalingam and Belani 2007). These two examples underscore the importance of mapping tumor-associated recurrent amplicons in order to identify potential target genes.

While amplification is thought to be driven by a key gene(s) conferring a cell proliferative advantage, the boundaries of amplicons are clearly unstable regions of the genome that are prone to recurrent breaks under adverse external conditions. Several amplicon boundaries have now been mapped and found to lie within common fragile sites (CFS) in both tumor- and drug-resistant cells (Hellman, et al. 2002; Kuo, et al. 1998; Miller, et al. 2006). For example, the MET amplicon on 7q has its centromeric boundaries in the FRA7G at 7q31.2 (Hellman, et al. 2002; Miller, et al. 2006) and the duplicated region encompassing the PIP gene in breast cancer has boundaries in the FRA7I locus at 7q35 (Ciullo, et al. 2002). Fragile sites are therefore thought to be instrumental in the initiation of the amplification process, which is then driven by proliferative growth promoting genes. However, other amplicons have been described which do not localize near known CFS (Kuwahara, et al. 2004), inferring the existence of other types of unstable DNA domains involved in these breakpoints.

2.1.2 Chromosome aberrations and 12p in secondary leukemia

The biological events that trigger the evolution of a malignant phenotype in secondary leukemia remain virtually unknown. In most cases, therapy-induced myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML) are associated with a clonal monosomy or deletion of chromosomes 5 or 7 (Heller, et al. 2003; Weiss, et al. 2003). These abnormalities are usually followed by an increased chromosome instability and emergence of additional aberrations that parallel the progression of the disease. It is generally accepted that the secondary aberrations are pathognomonic in the biological steps leading to a full malignant phenotype (Pedersen-Bjergaard and Rowley 1994). Most secondary chromosome abnormalities observed in therapy-induced leukemia are unbalanced aberrations represented by the loss or gain of genetic material, and it has been suggested that genomic imbalance, rather than gene rearrangement, is at work in tumor progression evolution (Mertens, et al. 1994). Genomic imbalance in the form of DNA amplification has not been frequently described in leukemia, and it is worth observing that several reported cases had a prior history of anti-neoplastic therapy (Knutsen, et al. 2003; Mathew, et al. 2000; Viguie, et al. 2002). In solid tumors, on the other hand, gene amplification has been widely documented. This is particularly the case with the chromosome 12 short arm, the over-representation of which has been described as the most specific abnormality in testicular germ cell tumors (TGCT) (Looijenga, et al. 2003; Mostert, et al. 1998; Mostert, et al. 1996; Pienkowska-Grela, et al. 2002).

Although 12p amplified sequences are infrequent in hematological malignancies (Hoglund, et al. 1996; Mertens, et al. 1994), rearrangements of chromosome 12 short arm are a common finding in leukemias (Hoglund, et al. 1996; Kobayashi, et al. 1996; La Starza, et al. 1999; Raynaud, et al. 1999; Romana, et al. 1996) and have recently attracted

considerable interest due both to the discovery of the TEL gene at 12p13 and the observation of frequent interstitial deletions (Strefford, et al. 2007; Vieira, et al. 2005).

The molecular mapping of chromosomal breakpoints and amplicons is essential to the identification of key genes involved in carcinogenesis. This mapping has relied on conventional molecular cytogenetic techniques, namely fluorescence *in situ* hybridization (FISH) and conventional comparative genomic hybridization (CGH). More recently whole genome approaches such as CGH arrays (aCGH), tiling arrays and single nucleotide polymorphism (SNP) arrays, have considerably improved the resolution of mapping; the former *technique* is briefly described in the following section.

2.1.3 CGH arrays

Cytogenetic CGH was developed in 1992 (Kallioniemi, et al. 1992) to detect copy number changes in tumor cells as an alternative method to traditional cytogenetics, which is notoriously difficult in solid tumors. The method is based on the co-hybridization of differentially labeled normal and tumor DNA on normal metaphase chromosome spreads. Variations in fluorescent intensity ratios (tumor versus control) along each chromosome reflect DNA copy number gain or losses at these locations. The method has provided a vast amount of information in nearly all cancers, but has a limited resolution and can only detect changes larger than 5 to 10Mb. Based on the same conceptual approach, the microarray technology enabled the production of the first CGH arrays initially using bacterial artificial chromosomes (BACs) and cosmids clones immobilized on a slide instead of metaphases (Pinkel, et al. 1998; Solinas-Toldo, et al. 1997). The human cloned inserts ranged in size from 40 to 300Kb and arrays covering the whole genome reached

the unprecedented resolution of 1Mb (Chung, et al. 2004). The use of overlapping golden path sequencing clones then allowed for whole parts of the chromosomes to be covered, but their number limited each array to a specific chromosomal region. More recently, CGH arrays using oligonucleotides of ±60 mer-oligonucleotide probes in coding and non-coding regions of the genome have been developed. They allow a full coverage of the genome with, for the Agilent 44K Genomic array, an average resolution of ±35Kb (Agilent Technologies, http://www.chem.agilent.com/). Amplicons characterized by cytogenetics comparative genomic hybridization can therefore be redelineated at a much higher resolution with array CGH, facilitating the identification of potential target genes.

2.2 Aims

A post radiotherapy leukemia patient presented with a very complex karyotype that, upon further investigation, was found to be largely constituted of an increased copy number of chromosome 12 material. Due to the recurrent involvement of chromosome 12 amplification in other tumors, it was decided to determine the molecular identity of the 12p rearrangement in this patient with a view to identify genes commonly involved in cancer. Combined fluorescence *in situ* hybridization (FISH) and cytogenetic CGH strategies were used in the first phase. Array CGH was later applied.

2.3 Case History

A 44-year-old Caucasian male underwent lymph node irradiation prior to a renal transplant in 1986. In March 1993, he was found to have developed anemia (Hb 6.6g/dL) and marked neutropenia (WBC of 1.2 10^{9}/L, neutrophils 6%). A bone marrow aspirate

showed a mildly increased cellularity, reduced and disordered granulopoiesis, and the presence of 30% blasts. An early secondary myeloid leukemia was diagnosed and confirmed on a trephine biopsy. The patient was admitted five months later for a heterologous bone marrow transplant, but died shortly afterwards of septic complications.

2.4 Materials and Methods

The project received ethical clearance (WITS Ethics committee, protocol 28/6/90) and specimens were collected following voluntary informed consent from the patient and volunteer.

2.4.1 Cytogenetic Analysis

Cytogenetic analysis was performed on bone marrow aspirate as well as on peripheral blood specimens, on three different occasions. Short-term cultures were established in RPMI 1640 media supplemented with 10% fetal bovine serum (both Gibco-BRL, Life Sciences), 2mM L- glutamine (Sigma), 100 units/ml penicillin and 10mg/ml streptomycin (Sigma). Cells were incubated in a 5% CO_2 humidified atmosphere at 37°C and harvested the next day for karyotying and chromosome GTG-banding as per conventional methods (the Act Cytogenetics Laboratory Manual 2nd ed. New York: Raven Press; 1991). Clonal abnormalities were described according to ISCN (Heim 1996).

2.4.2 DNA preparation

High molecular-weight tumor DNA was extracted from the patient's marrow aspirate using modified standard procedures (Sambrook J 2001). In brief, the bone marrow cells were first resuspended in ice-cold amonium chloride (NH_4Cl) and left on ice for 15 minutes to lyse the red cells. The centrifuged cell pellet (1100rpm for 10 minutes) was then resuspended in a lysis buffer (40mM Tris pH 7.5, 20mM $MgCl_2$, 5% Tween 20) and incubated in the presence of 0.5μg/μl proteinase K at 56°C overnight. Phenol and chloroform/isoamyl alcohol (24:1) extraction was then carried out and the DNA was precipitated with absolute ethanol (+_93%) in the presence of 0.3M sodium acetate at -20°C for 20 minutes. After a 70% ethanol wash, the pellet was dried and re-dissolved in 10mM EDTA, 5mM Tris pH 8. The DNA concentration was estimated on a spectrophotometer. The control DNA was extracted using the same procedure, with a peripheral blood sample obtained from a healthy female volunteer.

2.4.3 Chromosome suppression hybridization

Human genomic DNA contains high amounts of repetitive sequences that can hybridize to many regions of the genome. In order to avoid this and retain each probe's unique specificity, these sequences are blocked by prehybridization to Cot1 DNA enriched in repetitive sequences, a process that is termed suppression hybridization.

Chromosomes 5 and 12 specific plasmid (pBS) libraries, kindly given by J Gray (Collins, et al. 1991), were grown in soft agar. A large number of colonies was pooled and the plasmid DNA was extracted by standard DNA plasmid preparations (Sambrook 2001). Labeling was performed by nick translation using biotin-16-dUTP (Boerhinger

Mannheim). In brief, 2µg of template DNA was mixed with nick-translation buffer (0.5 M Tris-HCl, pH8, 50mM MgCl, 0.5 mg/ml bovine serum albumin), 0.1M β-mercaptoethanol, deoxyribonuclease 1 (DNAse I) to a final concentration of 0.06ng/µl, nucleotide (dNTP) stocks (at final concentrations of 0.05mM of each dCTP, dATP and dGTP, 0.01mM dTTP) and 0.04mM Biotin-16-dUTP with 20 units of DNA Polymerase I (Promega, USA). The reaction was incubated at 15°C for 2 hours and stopped by heating the reaction product at 68°C for 15 minutes after the addition of 3µl 0.5M EDTA and 1µl of 1% SDS. The labeled product was run on a 2% agarose gel to control for the fragment size, which has to be between 200 and 500bp. After denaturation, the probes were preannealed for 30 minutes at 37°C with a 50-fold excess of Cot1 DNA (Boerhinger Mannheim). Meanwhile, the patient's slides with metaphase spreads were denatured in 70% formamide at 72°C for 3 minutes and dehydrated in ice-cold ethanol series (70%, 90%, 100%). The preannealed probes (from 37°C) were immediately applied to the slides and hybridized overnight at 37°C in a humidified chamber. Visualization of hybridized DNA was obtained using fluorescein conjugated avidin and biotinylated goat anti-avidin antibodies (both 4µg/ml) (Vector Laboratories). Metaphasic chromosome and interphase nuclei were counterstained with 4,6-diamidine-2-phenyllindole dihydrochloride (DAPI-Serva) (Vector Laboratories).

2.4.4 Comparative genomic hybridization

Tumor and control DNA were respectively nick translated as described above, using digoxigenin –11-dUTP (Boerhinger Mannheim) instead of Biotin for the control DNA. CGH was performed as described (du Manoir, et al. 1993; Joos, et al. 1993). In brief, 1ug of biotinylated tumor and digoxigenin-labeled control DNA, respectively, were co-

denatured and prehybridized for 30 minutes with 70μg heat-denatured Cot1 DNA (Boerhinger Mannheim). The probe mixture was applied to the control female normal metaphase, the tumor-hybridized biotinylated DNA was detected using FITC-conjugated avidin (Vector Laboratories) and the control DNA with sheep anti-digoxigenin conjugated to rhodamine (Boerhinger Mannheim). Slides were analyzed on an epifluorescence microscope with appropriate filters for the visualization of fluorescein, rhodamine, and DAPI fluorescence. Images were acquired and digitalized using a cooled, charged couple device (CCD) camera (Photometrics, Tucson, Arizona). Images of different fluorochromes from the same metaphase were overlaid using the Gene Join (Yale University) software package.

2.4.5 Slot blot analysis

Varying concentrations of total genomic DNA from the patient and female control were denatured in 15mM EDTA, 0.4M NaOH and applied to a nylon membrane (Hybond N, Amersham). After blotting, the filters were baked for 2 hours at 80°C. Gel purified inserts of the KRAS probe (ATCC Nb 41026), six YACs from the CEPH YAC library (770a11, 951h6, 753fl2, 884f4, 924c11 and 952a6) (kindly provided by Denis Le Paslier, Centre d' etudes du polymorphisme humain (CEPH), Paris), the N12C12 (12p13.2-p13.3) cosmid probe (kindly provided by Dr. R. Kucherlapati, New York) and the spectrin gene mapping on 14q22-23 and used as a control, were α^{32}P radiolabelled with the readyprime kit (Amersham) as per manufacturer recommendations. All YACs and the cosmid probe were preannealed for 2 hours with a 100-fold excess Cot1 DNA (Gibco). Baked filters were hybridized overnight at 62°C and washed under stringent conditions (0.1x SSC, 42°C). They were then autoradiographed for 2 to 5 days. All YACs were initially hybridized to

normal metaphases and their mapping to 12p was verified in a minimum of 15 prometaphases (results not shown); no cross-hybridization could be detected in any other chromosomes.

2.4.6 CGH array

An oligonucleotide-based array CGH was performed on a 44K CGH array slide as per manufacturer protocol (Human Genome microarray kit 44K, Agilent Technologies). Briefly, 3μg of tumor and reference high molecular weight DNA were first digested with a combination of *Alu1* and *Rsa1* (Promega). The patient was Caucasian, and control DNA was from a commercially available Caucasoid reference DNA (Promega) consisting of pooled normal male and female samples (n= 25 for each) (Promega) to correct for variations due to copy number polymorphisms. Random prime labeling was done with the exo-Klenow (Invitrogen); 21μl of digested DNA was added to 20μl of random primer solution (x2.5)(supplied in the kit) and heat denatured at 95°C for 3 minutes. Ten times dNTPs solution, exo Klenow fragment and cyanine 5-dUTP and cyanine 3-dUTP were respectively added to the tumor and control DNA as per instructions, and incubated at 37°C for 2 hours. The labeling reaction was stopped by heat at 65°C for 10 minutes. The labeled products were cleaned using the Microcon YM-30 filters (Millipore) and the labeling efficiency was assessed on a NanoDrop ND-1000 UV-VIS spetrophotometer at absorbances of A 550nm and A 650nm for cyanine 3 and 5 and respectively.

Both test and control-labeled DNA were at this stage combined and co-precipitated with 50μg of Cot1 DNA (Invitrogen). They were then resuspended in hybridization solution and preannealed for 30 minutes at 37°C. Hybridization was for 40 hours at 60°C in a gently rotating hybridization chamber. After hybridization, the slides were carefully

washed with stabilization and wash buffer solutions (Agilent Technologies) in a dust-free environment. The slides were air dried and immediately scanned on an Agilent scanner. The TIFF images were later loaded into Gene Spring CGH analytics software for analysis.

2.5 Results

2.5.1 Cytogenetic analysis

All the cells examined were abnormal, with a composite karyotype of 46-48, XY, t(3:12)(q21;p11)[24], -5[24], del(7)(q22)[24], + r1[24], + r2[15], +mar1[12], + mar2[4] [cp24] (Figure 2.2). Although the short arm of chromosome 12 homolog (not involved in the t(3;12)) was found to be abnormal in structure, it was not possible to ascertain and further characterize the anomaly cytogenetically. FISH allowed for the identification of chromosome 5-derived material within 12p, as described below.

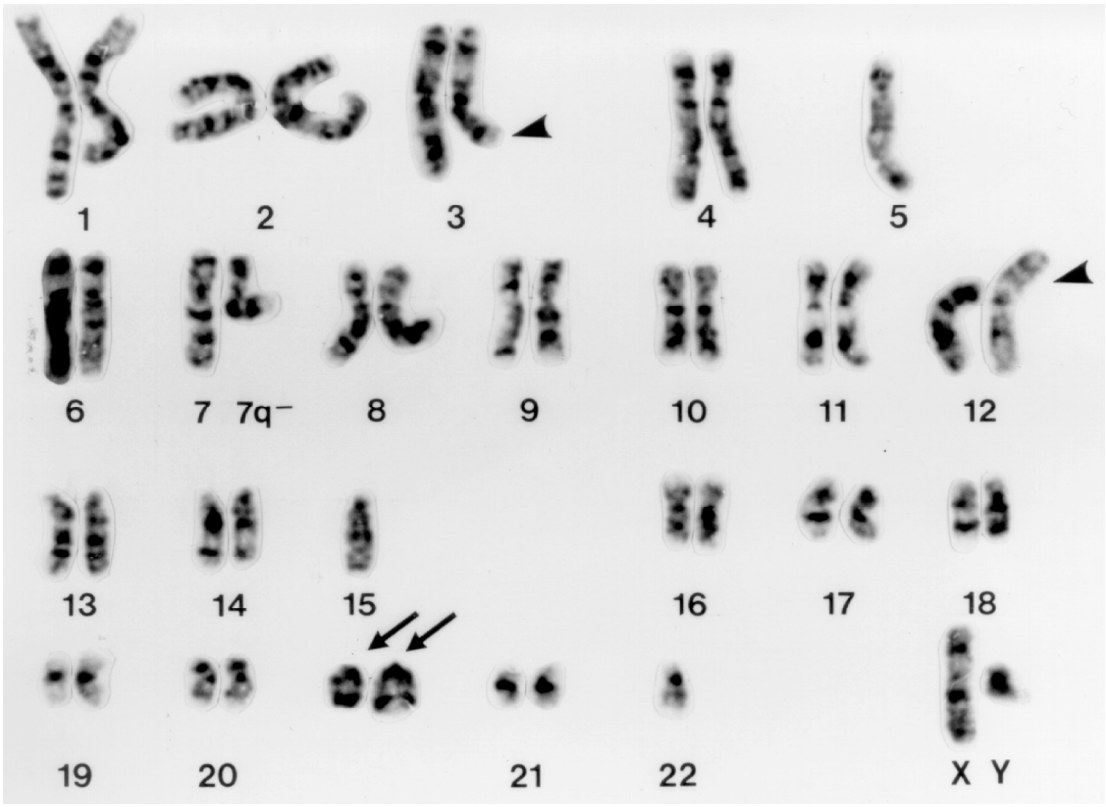


Figure 2.2: *G-Banded karyotype of a metaphase from the patient with secondary leukemia. Ring chromosomes are marked with arrows. The two derivatives of the t(3;12) are marked with arrowheads.*

2.5.2 Chromosome suppression hybridization

Chromosome 12-specific suppression hybridization was performed and confirmed the t(3;12) cytogenetics breakpoint. Results showed that the der(12)t(3;12) had retained the chromosome 12 centromeric region, and was consistent with 3q21 and 12p11 breakpoints (Figure 2.3 A). It was also observed that the ring chromosomes consistently exhibited chromosome 12 hybridization signals, although of weak intensity (Figure 2.3 B). It was considered that this low level of intensity might be due to the fact that 12 sequences involved in the rings represented several copies of a limited portion of chromosome 12

material not enriched in the library, and that they were mixed with sequences from another chromosome. With the aim of clarifying this possibility, other chromosome libraries were hybridized to the tumor cells. Chromosome 5 sequences were found to be present in the ring structures (Figure 2.3), and cohybridization of chromosome 5 and 12's libraries confirmed the participation of both in the rings. An interstitial deletion of 12p sequences was also detected on the other homolog (Figure 2.3 C), and it was seen that chromosome 5 material was inserted within 12p (results not shown). This resulted in a der(12)(12pter -> 12p13::5?pter -> 5?p11::12p11-> 12qter). It was not possible to determine whether chromosome 5 DNA originated from the short arm or long arm, but this rearrangement revealed a 12p11 breakpoint on the other homolog.

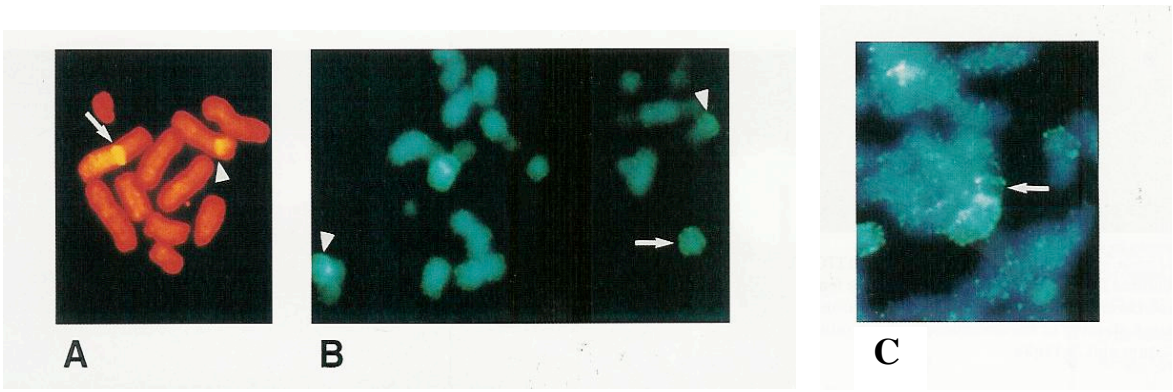


Figure 2.3: *(A) Partial metaphase spread from a tumor cell counterstained with propidium iodine after chromosome 12 painting (FITC detection).*
The 12p11 breakpoint on der(12)t(3;12) is marked with an arrow; the 12p material on the der(3)t(3;12) is marked with an arrowhead. (B) Partial metaphase spread from a tumor cell showing the ring maker (arrow) after chromosome 12 painting detected with FITC. Chromosomes are counterstained with DAPI; the derivatives of the t(3;12) are marked with arrowheads. (C) Enlargement of a partial tumor metaphase after chromosome 12 painting, counterstained with DAPI, showing the der(12) from ins(12;5)(p13p11;?) (arrow).

2.5.3 Comparative genomic hybridization

An analysis of 10 normal female metaphases, reverse painted with tumor DNA, revealed a consistent signal of very high intensity confined to a short region of chromosome 12 and mapping at 12p11p12. An increase in intensity was also noted on the chromosome 5 long arm, from q11 to q14, while the sequences q14 to q34 were clearly showing a reduced signal. The rest of chromosome 5 appeared otherwise normal. Taken together with FISH, these results suggest that, although not obvious from cytogenetic metaphases, the chromosome 5 material was involved in complex rearrangements which included an insertion, a ring formation, and a deletion. CGH showed that only five q14q34 sequences were truly lost (Figure 2.4). A decreased intensity on chromosome 7q confirmed the cytogenetic del (7) (q22).

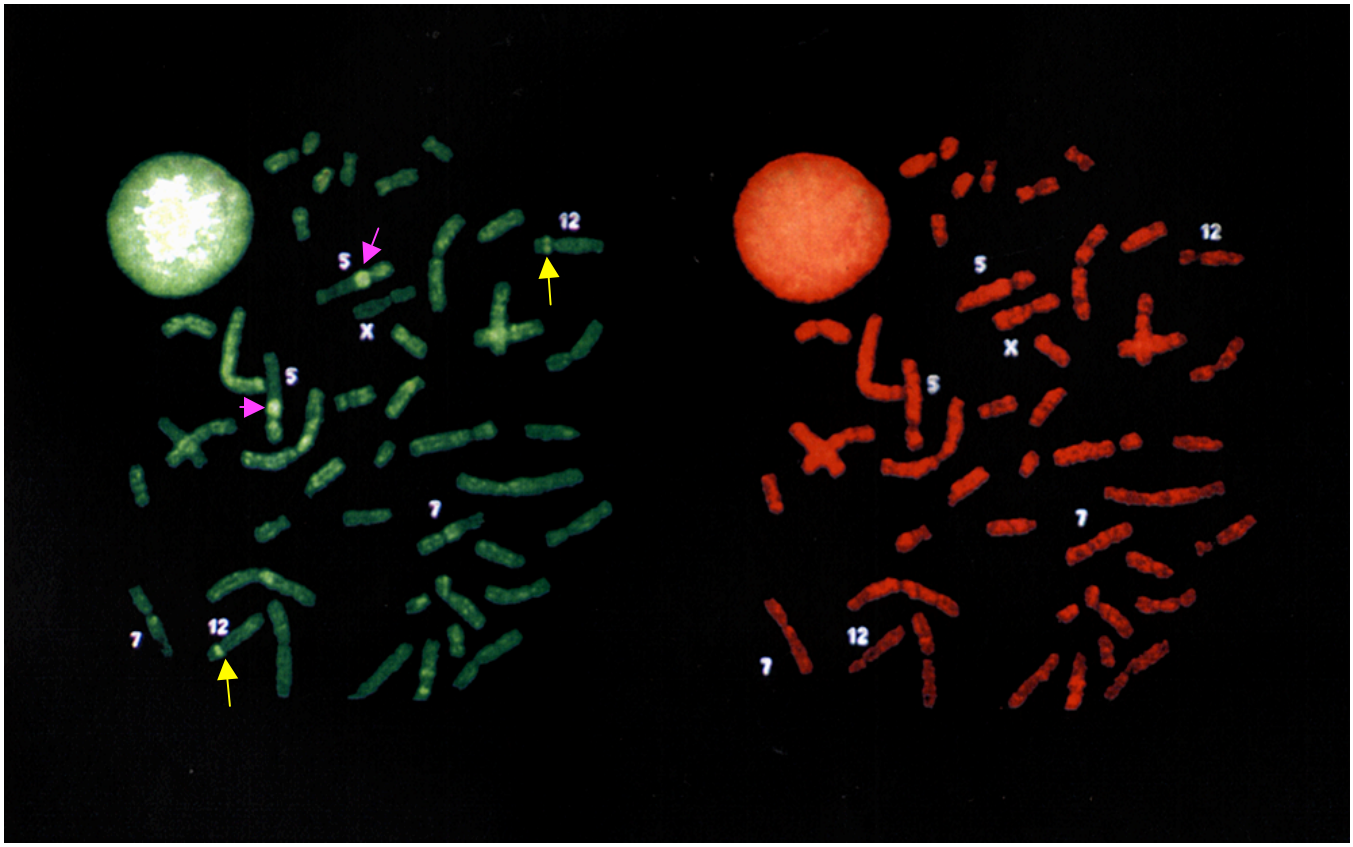


Figure 2.4: *A control metaphase subjected to CGH shows the FITC fluorescence of hybridized tumor DNA sequences, left, and the rhodamine fluorescence of control reference female DNA hybridized to the same cell showing a homogenous staining of chromosomes, right*. Two amplification sites at chromosome bands 12p11p12 (yellow arrows) and 5q11q14 (pink arrows) can be seen. The weak staining of chromosomes 5 and 7 reflect DNA losses corresponding to deletions mapping respectively to 7q22qter and 5q14q34.

2.5.4 Slot blot analysis

In order to map the 12p11p12 amplicon, we assessed the gene copy number of eight 12p-specific probes in the patient DNA, using slot blot analysis. No differences in signal intensity could be observed between the control DNA and the patient DNA in any of the eight 12p probes used, and these results were comparable when using the spectrin gene as a control (Figure 2.5)

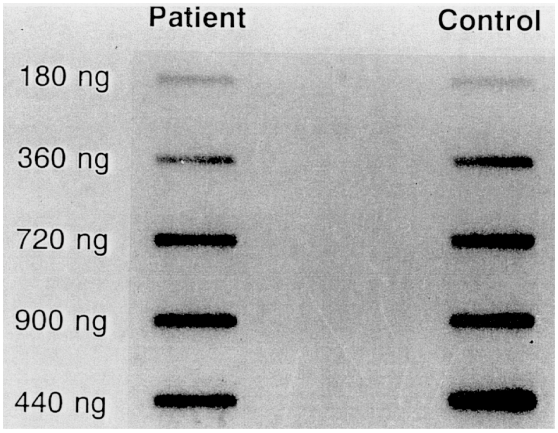


Figure 2.5: *Slot blot obtained with YAC 952a6 hybridized to an increasing amount of patient and reference DNA and showing homogenous signal intensity on both samples.*

2.5.5 CGH array

Overall, the oligonucleotide CGH array results highly correlated with those obtained through conventional CGH. Figure 2.6 shows ideograms of chromosomes 5, 7 and 12, confirming the aberrations described above.

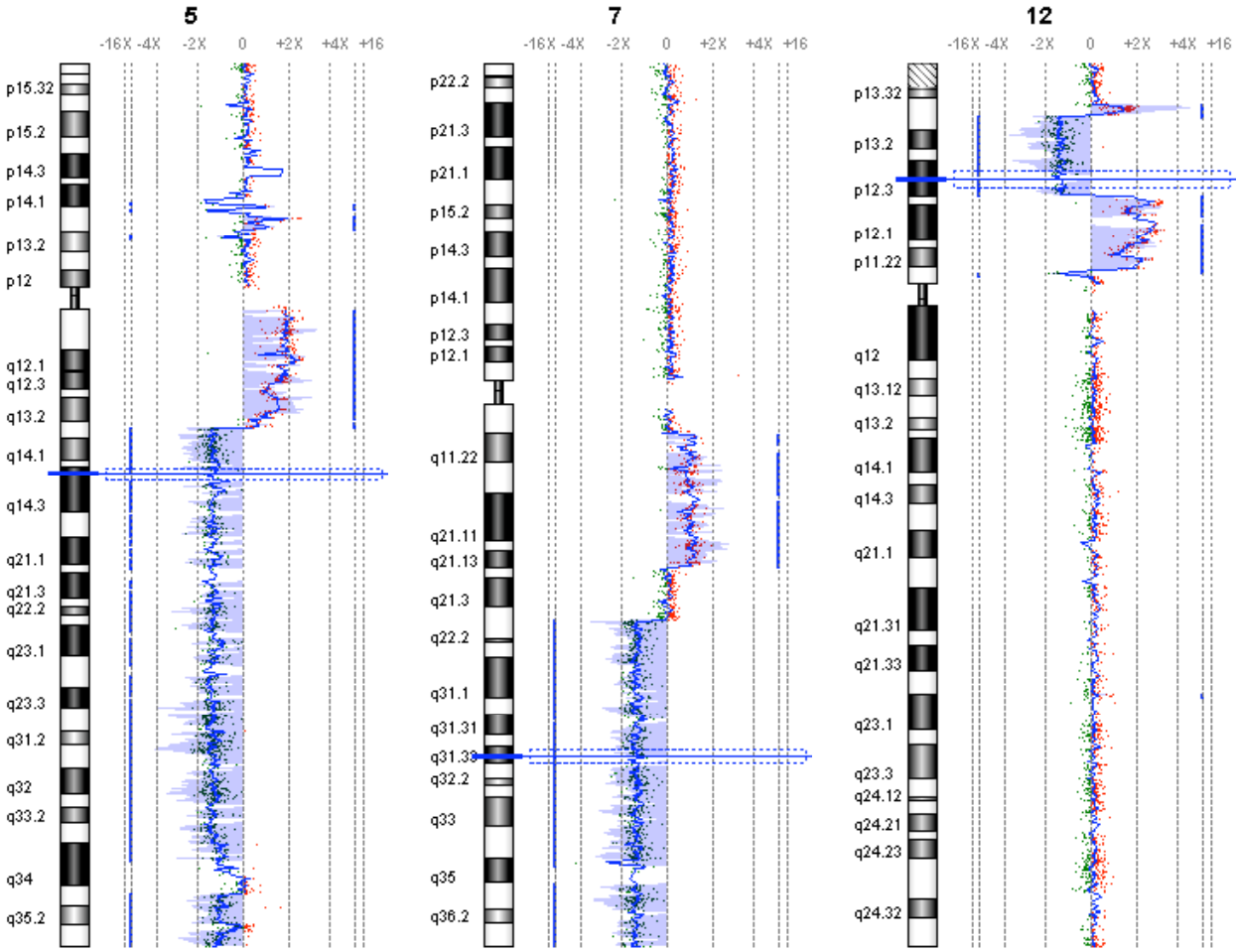


Figure 2.6: *Ideogram of chromosomes 5, 7 and 12 with the corresponding hybridization signal intensity mean log ratio (blue line) of patient DNA versus reference DNA.*
Images were analyzed using the CGH analytic software (Agilent Technologies). The zero line represents an equal copy number tumor/reference; deletions are represented by all dots on the left of this line, while amplifications are represented by dots on the right. The solid blue bars on the outside of the graph indicate global regions of deletions/amplifications respectively.

CGH array confirmed the regions of amplifications and deletions on chromosomes 5 and 7 respectively and brought a much higher level of resolution, allowing precise breakpoints to be defined as follows: chromosome 5: amp(5)(q11q13.3) and del(5) (q13.3 -> qter), (Figure 2.6).

On chromosome 5, the breakpoints of amplification and deletion events on the long arm of 5q13.3 coincided, suggesting that a common mechanism was involved, such as breakage fusion cycles, which selected part of the chromosome with overall selective positive and negative growth advantage. It may in fact be that the deluding 5q-gene deleted in secondary leukemia, could be a group of genes whose deletion sustains proliferation or other cancer phenotypic aberrations. Of note, small regions of both confirmed deletions and amplifications on 5p were detected by CGH array.

On chromosome 7, the region of amplification involved 7q11.1 -> 7q21.13 cytogenetic sub-bands and the deletion covered the 7q22.1 -> 7qter region almost entirely (Figure 2.6). Unlike what was seen for chromosome 5, for chromosome 7, regions of amplification and deletion were separated by a region of normal copy number implying that different breakpoints initiated these respective aberrations.

The amplicon initially detected by CGH on 12p11.2p12.1 was confirmed by CGH array. Given the extent of this amplicon, it is clear that YACs slot blot hybridization to tumor DNA with the collection of CEPH YACs described above did not reflect amplification in this region. This is probably due to a low sensitivity and the possibility that these mega YACs might have been unstable, losing parts of their human inserts.

Surprisingly, another, much smaller amplicon was revealed in 12p13.31, obviously beyond the detection resolution of conventional CGH. Because of its small size, and due to the fact that the same region had recently been reported to be amplified in TGCT, a full characterization of this amplicon was undertaken. As for our patient, this p13.31 amplicon had not been previously detected in TGCT, but was only later revealed with the high-resolution post-genomic tools of CGH arrays. The telomeric boundary of the 12p13.31 amplicon was distal to the Tumor Necrosis Factor Receptor Super Family, member 7 (TNRFSF7), and the centromeric boundary just telomeric to the SLC2a14 gene (solute carrier family 2, member 14, a glucose transporter). TNFRSF7 was included in the amplicon with 46 additional genes (Figure 2.7). The proximal boundary was particularly interesting, as it encompassed the Stem Cell Genes Cluster, including the homeobox gene, NANOG; the germ and embryonic stem cell enriched protein, STELLAR; and the growth differentiation factor 3, GDF3 (Figure 2.7).

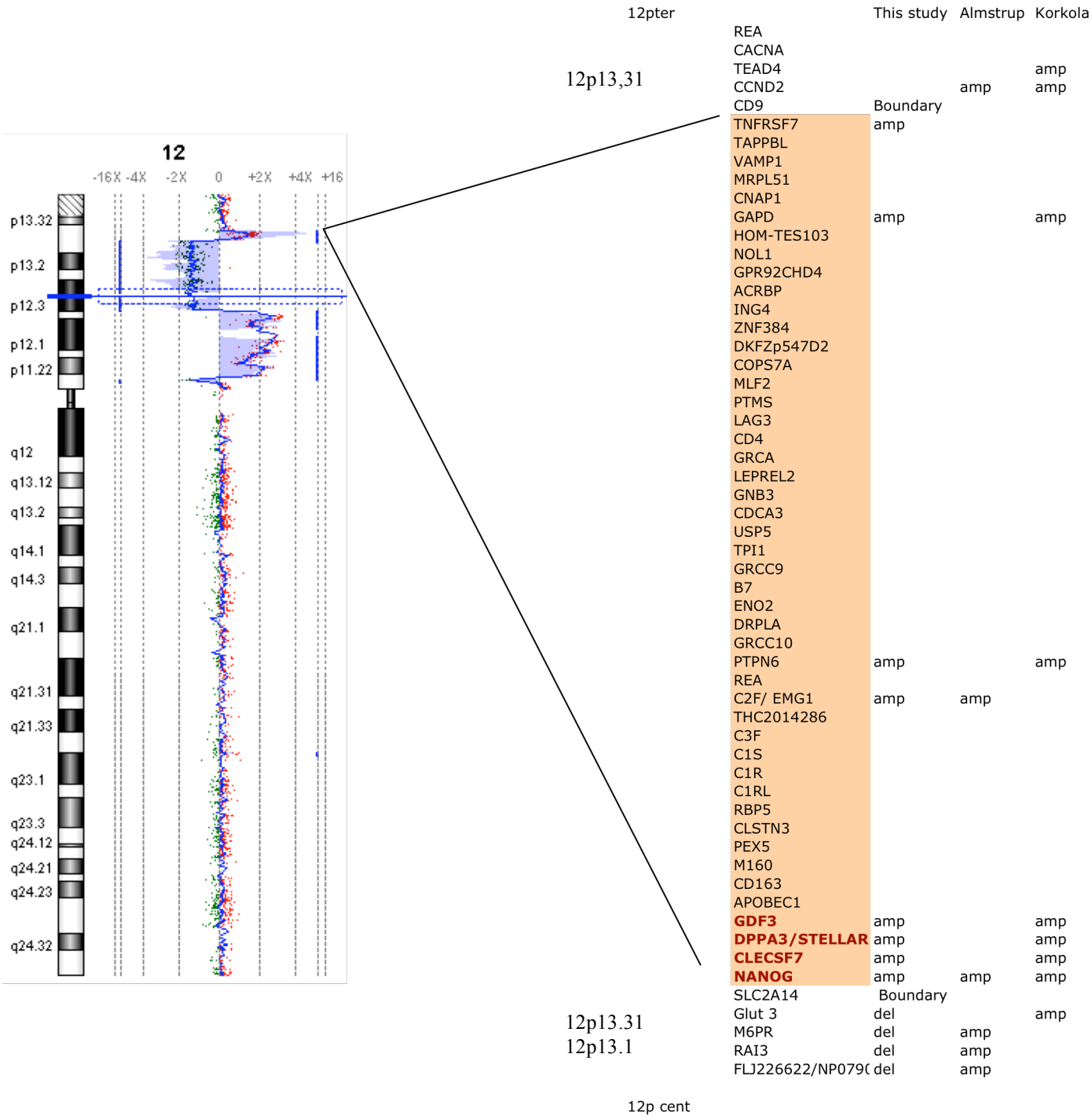


Figure 2.7: *Genomic composition of the 12p13.31 amplicon.*
On the right, genes included in the amplicon are highlighted in orange; the boundaries of the amplicon are indicated. In the three columns on the right are comparisons of genes that are found to be both amplified and over-expressed in carcinoma in situ CIS of TGCT (Almstrup, et al. 2005) and TGCT (Korkola, et al. 2006). The stem cell gene cluster is in dark red.

2.6 Discussion

Complex karyotypic aberrations observed in a patient with radiation-induced leukemia were molecularly characterized and revealed chromosome 12 short arm amplicons with strikingly overlapping amplicons described in TGCT.

2.6.1 Chromosome 12p breaks on both homologs and complex rearrangements

Chromosome 12 short arm anomalies were observed in addition to chromosomes 5 and 7's long arm partial monosomy, which are nonrandomly observed in secondary leukemia. The 12p11 region of both chromosome 12 homologs was involved in rearrangements, a t(3;12)(q21;p11) translocation, and a complex interstitial deletion and insertion. Interestingly, 12p11p12 sequences and 12p13.31 were also amplified. A 12p13, breakpoint was observed on one homolog.

Chromosome 12 abnormalities are a common finding in a wide variety of hematological malignancies. They have frequently been described in therapy-induced leukemia (Heller, et al. 2003; Penas, et al. 2003; Weiss, et al. 2003), and are mostly reported in the form of deletions involving various sub-bands of the short arm. Kobayashi et al. (Kobayashi and Rowley 1995) described a proximal interstitial deletion 12(12p13p12.2) in three myeloid leukemia patients, and suggested that cytogenetically different 12p rearrangements may in fact "represent 12p13 translocation accompanied by an interstitial deletion "; 12p13.2 is a critical band which harbors the ETS-like gene, TEL/ETV6, identified and fused to PDGFRB in t(5;12)(q33;p13) positive leukemia (Golub, et al. 1994). TEL is also commonly involved and fused to the AML1 gene in the t(12;21)(p12;q22) found in 15 to

35% of pediatric B-lineage ALL. TEL is further more frequently rearranged or deleted in hematological malignancies (Amor, et al. 1998; Raynaud, et al. 1996; Zelent, et al. 2004; Zuna, et al. 2004). On cytogenetic analysis there was a 12p13 breakpoint on one homolog, accompanied by an interstitial deletion and juxtaposed to chromosome 5 material. On array CGH, an interstitial deletion from 12p12.3 extending into the proximal part of the p13.31 sub-band was clear and spanned to the TEL gene.

Concomitant with the finding of a wide range of TEL rearrangements is the increasing number of reports describing 12p12p13 deletions (Hoglund, et al. 1996; Kobayashi and Rowley 1995). The former study showed that one copy of the KIP1/CDKN1 gene, whose product acts as a CDK-cyclin inhibitor, was lost in all their patients with a del (12p). The presence of these deletions implies the frequent existence of a proximal breakpoint in 12p11. Whether this breakpoint is of patho-physiological importance or is the mere consequence of a significant deletion remains to be clarified. However, the presence of a stable translocation t(3;12)(q21;p11) and the juxtaposition of chromosome 5 material to 12p11 sequences in our patient support the idea that this sub-band could harbor a gene involved in cancer phenotype. Indeed, 7 out of 10 patients with 12p12p13 deletions in the Hoglund et al. study had a stable translocation involving 12p11 (Hoglund, et al. 1996).

2.6.2 Chromosome 12p amplicons

The patient's tumor cells were all characterized by one or two ring chromosomes that were marked with a chromosome 12 library. Although ring chromosomes are not commonly described in leukemia, the presence of ring chromosome 12 is a highly specific abnormality described in lipomas and well-differentiated liposarcomas (Forus, et al. 2001;

Mandahl, et al. 1987; Micci, et al. 2002; Szymanska, et al. 1997). Two distinct amplification units, 12q14q15 and 12q21.3q22, have been mapped by CGH in these disorders (Micci, et al. 2002; Pedeutour, et al. 1999; Suijkerbuijk, et al. 1994a). We speculated that the same material could be involved in our patient. However, conventional CGH results revealed that the chromosome 12 short arm was amplified in 12p11p12. FISH slides obtained with a chromosome 12 library showed a weak but consistent signal on the ring chromosomes, which contrasted with the strong hybridization signal seen in 12p on the CGH slides. This observation strongly suggested that a relatively short-sized amplicon was amplified at a high copy number.

Supernumerary iso (12p) has been described in leukemia with half of the occurrences in secondary leukemia (Mertens, et al. 1994). Iso (12p) is furthermore a highly specific structural abnormality seen in 80% of testicular germ cell tumors (TGCT) (Parrington, et al. 1994; Skotheim, et al. 2006; Suijkerbuijk, et al. 1991) and was observed in mediastinal germ cell tumors (Aly, et al. 1994). FISH and CGH-based studies on these tumors have shown that the frequency of 12p sequences alterations is more frequent and extensive than seen on G-banded metaphases (Korkola, et al. 2006; Suijkerbuijk, et al. 1993) and multiple 12p copies have been detected in cytogenetically i(12p)-negative TGCT (Reilly, et al. 1993). More recently, in several studies, 12p amplicons have been further characterized with array technology (Korkola, et al. 2006; Skotheim, et al. 2006; Zafarana, et al. 2003).

Two main regions of 12p seem to significantly emerge from these studies. First, the 12p11.2p12.1 sub-region amplicon, which was also described in the metastasis of a TGCT (Suijkerbuijk, et al. 1994b), and the 12p13.31 region, more recently discovered (Almstrup, et al. 2005; Korkola, et al. 2006). This region has recently attracted attention since it is particularly rich in genes when compared to the rest of chromosome 12, as seen from the

Ensembl human genome database ideogram (Figure 2.8); it is also involved in a smaller size amplicon and contains a number of candidate cancer genes.

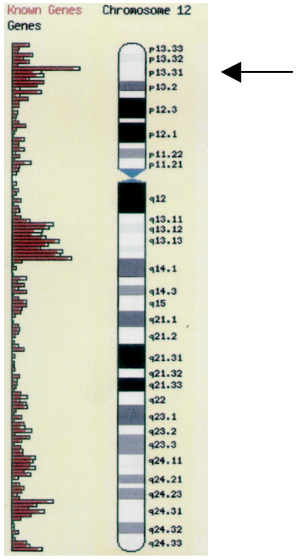


Figure 2.8: *Ideogram of chromosome 12, depicting on the left the number of genes corresponding to each cytogenetic sub-band. The sub-band 12p13.31 (arrowed) is one of the richest in genes content. (Diagram exported and adapted from the Ensembl database.)*

Most importantly, a cluster of stem cell-associated genes, NANOG, STELLAR and GDF3, maps to this region. These genes are known to be coordinately expressed in human embryonic stem cells, ESC and premeiotic germ cells; they are downregulated with cell differentiation (Clark, et al. 2004). The same genes were found to be amplified and over-expressed in both germ cell carcinoma in situ, CIS, (Almstrup, et al. 2005), as well as embryonal carcinoma and seminoma types of TGCT (Korkola, et al. 2006).

It is tempting to speculate that their inadequate amplification/activation confers a cell growth advantage and that they constitute valid candidates to be further evaluated as potential target genes in cancer. Indeed these genes were recently found to be over-expressed in breast cancer (Ezeh, et al. 2005).

It can also be speculated that these genes were activated in our patient, and although we did not have RNA source available to measure levels of expression, there has been a good correlation between DNA amplification and over-expression regarding this cluster (Almstrup, et al. 2005; Korkola, et al. 2006).

Interestingly, as already noted by others, spontaneous acquisition of karyotypic abnormalities occurs in *in vitro* cultures of human embryonic stem (hES) cells. The aberrations described have been gains from chromosomes 17 and 12 (Draper, et al. 2004; Pera 2004). One of these cell lines more specifically presented with an iso 12p, implying the presence of critical genes on chromosome 12 short arm, which support cell survival or proliferation.

2.6.3 Possible mechanisms of amplification

The presence of two separate amplicons located at the same chromosomal sub-regions in 12p in both our patient and many TGCTs is intriguing and suggests some common mechanism initially involving regions prone to breaks. Interestingly, in a CGH investigation of chronic B cell leukemia, Bentz et al. (Bentz, et al. 1995) described a patient with two amplifications on 12p13 and 12p11p12, respectively.

There are no known fragile sites in 12p; however, it is a likely possibility that BFB cycles were involved with the number of uncapped chromatid ends generated by initials breaks. It seems likely that ring formation occurred as a result of the complex chromosomes 5 and 12 rearrangements. The presence of several breakpoints on chromosomes 5 and 12 seems to have facilitated chromosome fusion, resulting in the juxtaposition of chromosome 12p11p12 sequences to chromosome 5 material, on both 12 homologs (Figure 2.3), and

possibly the region 5q11-q14, which was also amplified. This fusion derivative, having two sticky ends, would have circularized, and subsequent cycles of amplification may have been driven by "breakage fusion bridge cycles". After a succession of breakages and rearrangements, some ring chromosomes could have gained an increased copy number of chromosome 12p sequences by gradually shortening the amplicon in size and narrowing down to a sub-region of 12p11.2p12, in a mode similar to that observed in some *in vitro* DNA amplification models (Toledo, et al. 1992). Alternatively, amplification and size variation of the amplicon could have been caused by an odd number of sister chromatid exchanges. Unlike the boundaries of amplicons described on chromosomes 8 and 7, there are no known fragile sites on 12p. Bertoni et al. (Bertoni, et al. 1994) have shown that TTAGGG sequences of centromeric origin surrounded amplified regions in PALS-resistant hamster cell lines. They suggested that recombined events that accompany and possibly facilitate gene amplification frequently involve consensus centromeric and telomeric sequences. It would be interesting to test whether equivalent repetitive sequences surround the 12p amplification/deletion breakpoints in our patient. The near-centromeric 12p11 breakpoint, involving both homologs, could indeed involve repetitive sequences of a centromeric nature. The connection between 12p amplification and the 12p11 translocation breakpoint [t(3;12)(q21;p11)] is intriguing and suggests an initial common mechanism. Similar observations, where amplification clusters were accompanied by translocation breakpoints at 11q13 (Roelofs, et al. 1993), have been reported in breast cancer. Whether similar genes are involved in both amplification and translocation is unknown.

2.6.4 The 12p stem cell gene cluster as candidate target

The striking similarity of genes involved in liquid tumors, secondary leukemia and in germ cell tumors strongly suggests that the 12p13.31 cluster may be more commonly involved in cancer and in particular the stem cell genes NANOG, STELLAR and GDF3. The cause of TGCT is unknown, but thought to have a strong etiological environmental and possibly iatrogenic component, albeit of a hormonal nature (Almstrup, et al. 2005). Our patient was irradiated for renal transplantation but it will remain to be determined whether 12p is particularly sensitive to exogenous mutagens.

Forty-seven genes were involved in the 12p13.31 amplicon. Although the "stem cell cluster" is a likely candidate, other genes of interest also found to be amplified in several studies may participate in or even play the "driver role". These genes include TNFRSF7, GAPDH, PTPN6 and EMG1. Assessment of other cancers with 12p amplification with probes specific for these genes should help in establishing the extent of their involvement in malignancies.

CHAPTER 3

DEVELOPMENT OF A NOVEL APPROACH TO SCANNING GENES AT FRAGILE SITES

3.1 Introduction

Besides the formation of non-random amplicons hosting cancer "driver genes", chromosomal hot spots represented by common fragile sites have been linked to cancer initiation and progression (Arlt, et al. 2006; Iliopoulos, et al. 2007; Tsantoulis, et al. 2007). These sites are known to be sensitive to a variety of external carcinogens and are expressed as chromatid breaks or gaps under replication stress culture conditions. Genetic instability represented by fragile sites has been poorly investigated in South Africa, especially in cancers that are highly prevalent and whose occurrence is associated with exposure to external carcinogens as it is the case with esophageal squamous cell carcinoma (ESCC).

3.1.1 Fragile sites

Fragile sites are chromosomal regions non-randomly distributed in the genome, highly prone to breaks and manifesting in metaphase chromosomes as chromatid gaps or breaks under culture conditions of replication stress (e.g. folate deficiency or aphidicolin exposure). There are two types of fragile sites: rare fragile sites and common fragile sites. Rare fragile sites are sequence related and caused by the expansion of a repetitive stretch

of di or trinucleotides (Glover 1998). Rare fragile sites are present in less than 2.5% of the human population and are often associated with a congenital disorder caused by their interfering with the transcription of a downstream gene. The best known rare fragile site is the FRAXA, associated with the mental retardation syndrome,"fragile X syndrome" (Richards, et al. 1991). Rare fragile sites will not be discussed further in the context of this thesis.

The second type of fragile sites, common fragile sites (CFS), are present in every individual, and like the rare fragile sites, they can be induced *in vitro* under culture conditions of replication stress (Glover 1998; Glover, et al. 1984). In contrast to rare fragile sites, they do not contain stretches of repetitive di or tri nucleotide sequences to explain their sensitivity. Regions of high DNA flexibility as defined by sequence patterns, have been described (Mishmar, et al. 1998). Some of these sequences share partial analogy with rare fragile sites, but by and large, the reason for CFS fragility remains unknown. Intriguingly CFS are evolutionarily conserved across species, which implies that they have an important function (Glover, et al. 1998; Ruiz-Herrera, et al. 2004), and some studies have linked them to replication origins (Raveendranathan, et al. 2006; Toledo, et al. 2000). There are 88 common fragile sites in the human genome database (GDB) (http://www.gdb.org/gdb/queryPublic.html) (listed in Appendix 1), although this number can be increased by intensifying the conditions of replication stress in the culture medium. However, only 20 of these described fragile sites account for 80% of gaps and breaks observed in induced cultures (Glover, et al. 1984).

3.1.2 Fragile sites and replication

Rapidly dividing tissues replicate every 8 hours, and each human cell contains $6x10^9$ base pairs; and as a result, $2x10^5$ nucleotide base pairs are replicated every second in the human body.

Conditions of partial replication inhibition induce the replication fork to stall or collapse and produce single-strand stretches of DNA on the lagging strand that fail to complete replication (Figure 3.1). These are expressed as chromatid gaps or breaks in the subsequent metaphase at non-random fragile sites, CFS. Replication stress can be experimentally induced under culture conditions that partially inhibit DNA synthesis, such as folate deficiency or with the addition of low doses of aphidicolin, which inhibits the polymerase α (Glover 1998; Glover, et al. 1984).

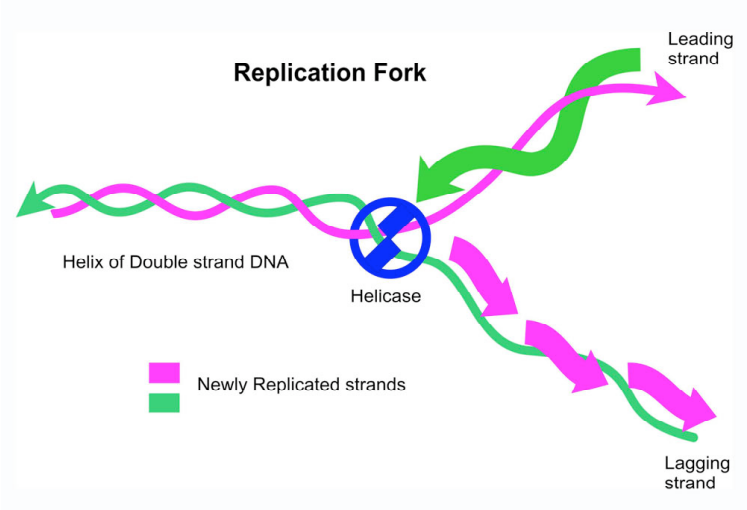


Figure 3.1: *Replication fork. While helicase unwinds the DNA, Polymerase α (grey-blue circle) leads the replication process on the lagging strand. The leading strand DNA is synthesized by Polymerases δ and ε. Uncompleted single-strand stretches of DNA (large pink arrows) on the lagging strand will produce gaps under conditions of replication inhibition.*

3.1.3 Fragile sites are sensitive to external carcinogens

Fragile sites are sensitive to external carcinogens and their expression as chromosomal gaps is enhanced in vivo by dietary and environmental factors (Fenech 2001; Stein, et al. 2002; Yunis 1988; Yunis, et al. 1987). Active tobacco smoking, in particular, directly affects CFS level of expression (Stein, et al. 2002). At the same time, a number of viruses, including the human papilloma virus (HPV), the hepatitis B virus (HBV) and the murine leukemia virus (MLV) (which has been used as a vector for gene therapy), all tend to preferentially integrate within fragile site loci (Bester, et al. 2006; Feitelson and Lee 2006; Thorland, et al. 2003; Thorland, et al. 2000; Wentzensen, et al. 2004). Ethanol, hypoxia and high doses of caffeine have also been found to trigger CFS expression, although evidence for caffeine comes from *in vitro* studies (Coquelle, et al. 2002; Coquelle, et al. 1998; Guven, et al. 1999; Kuwano and Kajii 1987; Richards 2001).

3.1.4 Pathways and genes that control DNA repair at fragile sites

Casper and Glover have recently unraveled the pathway that controls the stability and repair at fragile sites (Casper, et al. 2002). They showed that the ATR gene is the key player. Cells deficient in ATR were found to express a high level of both induced and spontaneous fragile sites. ATR is thought to work by activating both intra S and G2 cell cycle checkpoints under replication stress. Cell cycle arrest allows the tracking and repair of unreplicated strands by the protein complex ATRIP (Casper, et al. 2002). Other important genes in this pathway include BRCA1, FANCD2, and SMC1, all of which are involved in DNA repair/stability. In the fragile site repair pathway, they act as downstream targets of ATR (Glover 2006; Glover, et al. 2005) (Figure 3.2). Interestingly, deficiencies in any of these genes have been reported to be associated with an increased

instability of a fragile site (Arlt, et al. 2004; Focarelli, et al. 2006; Howlett, et al. 2005; Musio, et al. 2005).

Fragile site expression therefore occurs as a result of the combined effect of cellular damage due to external carcinogen exposure and of the "genetic fitness" of individuals (or subgroups of populations), to repair these breaks.

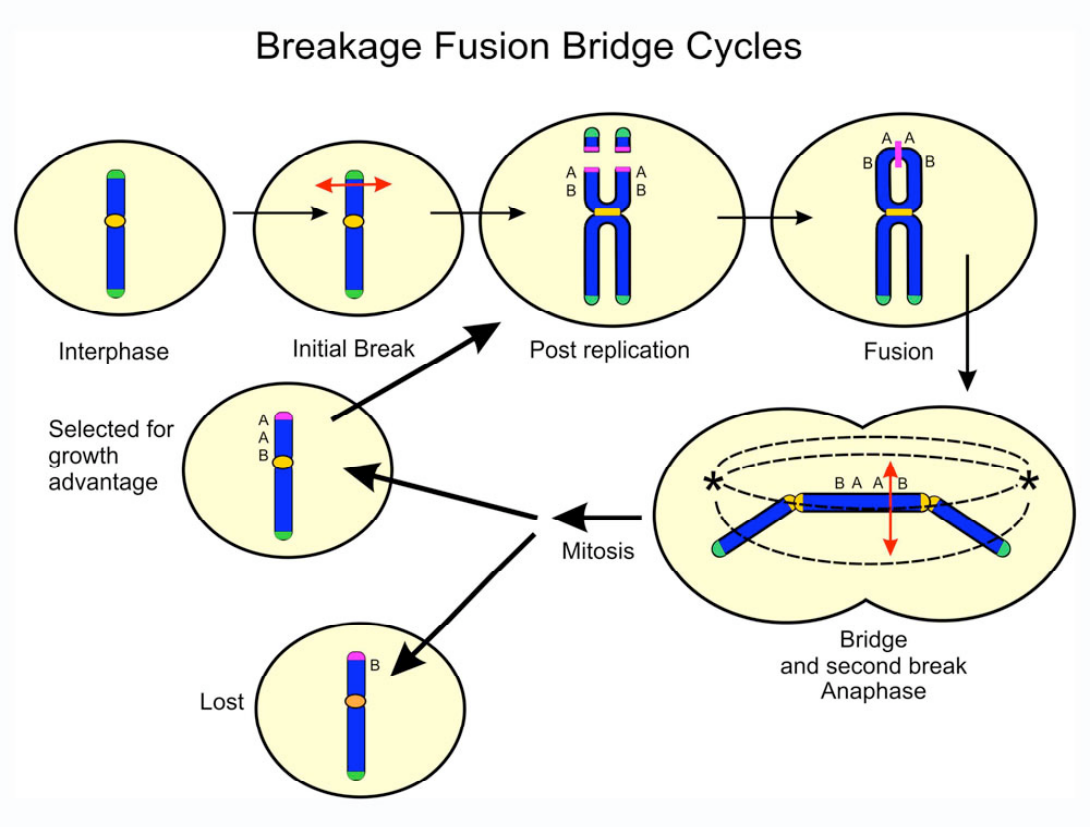


Figure 3.2: *ATR-dependent cell cycle checkpoint that controls fragile site stability. At the top is the replication fork; the lagging strand is replicated by polymerase α. Conditions of replication stress, shown here for specific culture conditions, will induce cell cycle arrest in S or G2. In a genetically "fit cell" (bottom left), the RPA protein will track the break and allow the repair complex ATRIP, made of RPA, ATR and the CHK1 proteins, to bind to the single-strand stretch of DNA. Other ATR dependant downstream genes involved in the repair pathway are indicated, namely: BRCA1, FANCD2 and SMC1. The bottom right represents events in a cell that fail to repair the single-strand break. The latter will either be expressed as a fragile site or can alternatively further evolve into a double-strand break (bottom).*

3.1.5 Fragile sites and cancer

In vitro experimental models have long shown that fragile sites are hot spots for sister chromatid exchange (SCE) (Glover and Stein 1987) and they exhibit a high rate of translocations and deletions in somatic cell hybrids (Wang, et al. 1993). Correspondence between fragile sites and several cancer-associated chromosomal breakpoints has been noted since the late 1980s (Yunis 1988); however, these observations did not attract a lot of attention in the last 15 years, apart from isolated groups.

The advance of post-genomic information and some recent discoveries have, however, started to spark a renewed interest in CFS. Indeed, a profiling of all DNA amplification events in cancers showed that many amplification sites occurred at CFS (Myllykangas, et al. 2006). Micro RNAs have recently emerged as functional regulators of gene expression/translation and it is estimated that a large proportion of micro RNAs (over 50%) are located within fragile sites (Calin, et al. 2004). This is likely to imply additional mechanisms by which CFS may functionally participate in cancer development. In addition, increased CFS expression was also recently found to be central in pre-malignant DNA damage response activation. Interestingly, two independent studies have shown that in both patients and *in vitro* induced pre-malignant lesions, activation of the DNA damage response is paralleled by an increased genetic instability and allelic loss at fragile sites, before any other genomic instability is observed (Bartkova, et al. 2005; Gorgoulis, et al. 2005). This observation was recently confirmed by a study where *in vitro* oncogene induced replication stress was shown to preferentially target CFS in a genome wide analysis (Tsantoulis, et al. 2007). The detection of genetic imbalances at fragile sites may therefore offer a powerful way to detect the "signature" of replication stress (Casper, et al. 2002) and identify pre-malignant lesions.

Most interest in CFS, however, has been linked to the presence of deletions at the most commonly expressed fragile sites that have been documented in a variety of cancers (for reviews see (Glover 2006; Huebner and Croce 2001)). Their role remained uncertain until recently, when several associated genes localized in these regions were clearly characterized as having tumor-suppressing activity.

3.1.6 Tumor suppressor genes at fragile sites

The two most commonly expressed human fragile sites, FRA3B and FRA16D, harbor the fragile histidine triad (FHIT) and the WW domain containing oxidoreductase (WWOX) genes respectively (Aqeilan, et al. 2007; Bednarek, et al. 2000; Corbin, et al. 2002; Ohta, et al. 1996; Trapasso, et al. 2003). Despite some controversy regarding the involvement of FHIT in oncogenesis, there is mounting evidence to suggest that it is more than a passive actor, although the nature of its role in differing cancer scenarios remains to be elucidated. FHIT is altered in a wide variety of tumors, mostly by genomic deletions and/or promoter hypermethylation, that results in inhibition of the FHIT product. Several reports have raised pertinent questions regarding the relevance of FHIT deletion to cancer development: exon skipping alternative transcripts have been found in normal tissues in addition to normal transcripts (Chen, et al. 1997; Chu, et al. 1998; Matthews, et al. 2001), the gene lies within an unstable genomic region which may be mechanistically deleted and the effect of hemizygous deletion in tumors is unclear. However FHIT was shown to function as a tumor suppressor gene (Dumon, et al. 2001a; Dumon, et al. 2001b) and FHIT was further found to induce apoptosis and modulate the expression of the DNA damage check-point proteins Chk1 and Hus1 (Deng, et al. 2007; Ishii, et al. 2006). These studies and many others show that FHIT has a definite tumor-suppressor function thus supporting the evidence for its role in carcinogenesis.

WWOX also participates in apoptosis (Chang, et al. 2003) and was functionally shown to act as a tumor suppressor in knock-out mice models and to suppress tumorigenicity in lung cancer cell-lines (Aqeilan, et al. 2007; Fabbri, et al. 2005). Furthermore, WWOX haploinsufficiency was found to be sufficient to increase tumor development (Aqeilan, et al. 2007). Both FHIT and WWOX can therefore participate in carcinogenesis when deleted or inactivated. Both genes have small exons distributed over their respective loci, and large intragenic deletions have been detected in a wide variety of malignant and pre-malignant tumors (for review see (Huebner, et al. 1998b)).

In addition to intragenic deletions, the aberrant hypermethylation of 5'CpG islands in the promoter region of both FHIT and WWOX genes has been shown to be closely associated with their transcriptional inactivation. This second mechanism of FHIT inactivation has been described in breast, lung, bladder and ESCC (Noguchi, et al. 2003; Iliopoulos, et al. 2005; Guo, et al. 2007). Interestingly, FHIT hypermethylation is also commonly associated with childhood hyperdiploid B cell acute lymphocytic leukemia (ALL) (Zheng, et al. 2004). Similarly, WWOX is also aberrantly methylated in breast, lung and hematological malignancies (Iliopoulos, et al. 2005; Ishii, et al. 2004) as well as in pancreatic cancer (Kuroki, et al.2004). The reversing of WWOX methylation was further shown to result in the inhibition of cancer cells growth in breast tumors (Iliopoulos, et al. 2007).

Genomic deletions and hypermethylation appear to be the main mechanism of FHIT and WWOX inactivation (Iliopoulos, et al. 2005; Kuroki, et al. 2003; Liu, et al. 2004; Noguchi, et al. 2003), which validates DNA based detection studies. A DNA approach can also make use of a significant amount of information that can be retrieved from the retrospective evaluation of archived paraffin-embedded specimens.

3.1.7 Fragile sites in epithelial tumors

Epithelial tumors affect tissues that are in contact with environmental exogenous factors and they may be prone to acquiring deletions at CFS. Esophageal squamous cell carcinoma (ESCC) is a common cancer in South Africa with a global prevalence higher than in the rest of the world, with geographical areas of very high occurrence. In the Eastern Cape for example, the age standardized incidence rate (ASIR) has been reported to be as high as 76.6/100 000 and 36.5/100 000 for males and females, respectively during the 1991- 1995 period (Somdyala, et al. 2003). ESCC ranks as the second most common malignancy in black males and third most common in black females (National Cancer Registry of South Africa 2004). The diagnosis is frequently delayed due to late presentation of patients at a stage of dysphagia and it has a very high mortality rate. Several studies (Hendricks and Parker 2002; Morita, et al. 2002; Pacella-Norman, et al. 2002) have shown that environmental factors and lifestyle habits are strongly associated with the risk of developing the disease. In South Africa, the main identified epidemiological risk factors include heavy smoking (Pacella-Norman, et al. 2002), exposure to fumonisin, a fungal toxin produced by fusarium fungi growing on local maize (Marasas, et al. 1979), and the consumption of alcohol and homemade beer fermented from infected maize (Pacella-Norman, et al. 2002) (Dandara, et al. 2005; Isaacson 2005). Other factors have been postulated as making a contribution, such as poor nutrition and Human Papillomavirus (HPV) infection (Pacella-Norman, et al. 2002) (Matsha, et al. 2002), but these factors are less significant. Most of the above risk factors and carcinogens (alcohol, tobacco consumption and HPV infection) have been shown in a variety of studies to potentially affect the integrity of fragile sites (Kuwano and Kajii 1987; Stein, et al. 2002; Thorland, et al. 2000; Wentzensen, et al. 2004).

In the rest of the world, several investigations have been conducted to assess the incidence of deletions, mainly in the FRA3B locus. These have revealed varying but significant degrees of deletions in both early cancerous lesions, as well as in association with progression. In Japan, loss of the Fhit protein has been documented in early stages of ESCC (Kuroki, et al. 2002), and FHIT gene alterations were also detected in association with severe malignant progression (Kitamura, et al. 2001; Mori, et al. 2000; Noguchi, et al. 2003; Stein, et al. 2002). While Fhit loss of expression has been documented in SA oral cavity squamous cell carcinoma (van Heerden, et al. 2001a; van Heerden, et al. 2001b) it was not investigated in ESCC. Only one investigation, also from Japan, reports on WWOX gene status in ESCC, and showed that it was altered in nearly half of tumors (Kuroki, et al. 2002). Neither of these genes has been investigated in South African esophageal carcinoma.

3.1.8 Methods of investigation of genes at fragile sites

Genomic deletions in FHIT and WWOX have mostly been investigated using loss of heterozygosity (LOH) analysis using polymorphic markers in separate assays. The markers used in these studies most often lie outside the coding sequence of CFS-associated genes and they do not inform on the functional consequences of deletions. LOH in the respective fragile locus does not always reflect the loss of coding exons or of protein product (Boldog, et al. 1997). In fact, they could reflect partial heterodisomy due to somatic recombination. Somatic recombination is a likely event in cancer cells and particularly at fragile sites that show increased sister chromatid exchanges in *in vitro* models (Glover and Stein 1987). In addition, deletions within fragile sites are known to be heterogeneous and can be heterogeneous within one cell line (Corbin, et al. 2002); in

effect, this means that differing regions within the fragile locus can be deleted in different cells. This could bias LOH results. Although the FHIT locus spans 1.4Mb and WWOX spans more than 750kb (Bednarek, et al. 2000; Druck, et al. 1997), both genes code for small transcripts of approximately 1kb.

Immunohistochemical assays (IHC) have been performed in a wide variety of tumors and provided large amounts of data regarding the loss of FHIT protein in solid tumors. This method remains subjective, results can vary widely with the antibody concentration used in the assay and there is a high inter-laboratory variation with this method in general (Vincent-Salomon, et al. 2003). Furthermore, IHC techniques do not provide information as to the nature of the genetic events that cause protein loss. This information may be useful regarding potential etiological factors of exposure. Reverse transcriptase polymerase chain reactions (RT-PCR) studies are dependent on the availability of fresh material, which is often difficult to obtain, as most biopsy tumor specimens are very small or fixed in paraffin.

For this reason a novel approach that specifically targets the exons of tumor suppressor genes at fragile sites may assist in evaluating the frequency of deletions likely to have a functional impact on fragile sites associated genes expression.

3.2 Aims

The aim of this chapter was to develop a new genetic assay that provides information on the exon composition of genes at fragile sites and that allows one to simultaneously screen for several genes associated with fragile sites in one assay.

The rationale was that under the exposure of common external carcinogens, one critical site hosting tumor-suppressor genes or another, or both may be affected. One gene may take over in driving tumorigenesis and deletions may therefore not necessarily affect all fragile sites equally. Yet, a high rate of deletions at fragile sites would be suggestive of replication stress. This assay could therefore provide information as to the respective pattern of deletions at fragile sites in cancers suspected to arise from exogenous carcinogen exposure.

In order to target several exons of several genes in one assay, the multiplex ligation dependant probe amplification (MLPA) technology was chosen. MLPA is a new PCR based method that evaluates copy number changes by comparative quantitation of PCR-amplified probes that are specific for the target (Schouten, et al. 2002). It uses universal primers that allow the amplification of numerous probes in one assay.

Five esophageal squamous cell carcinoma cell lines established in South Africa were ideally suited to evaluate the assay. These cell lines were simultaneously characterized by FISH and multicolor FISH by Jacqueline Brown as part of a Master's project conducted under my supervision. They were also cytogenetically characterized and there was therefore extensive information on each cell line for control purposes. They also constituted a pilot survey to predetermine the value of screening a larger number of ESCC in South African patients.

3.3 Materials and Methods

3.3.1 Cell lines

The five esophageal carcinoma cell lines investigated here were originally established from five black South African male patients from the Gauteng area, with known esophageal squamous cell carcinoma. These were previously described in the literature and are referred to as cell lines SNO, WHCO1, WHCO3, WHCO5 and WHCO6 (Bey, et al. 1976) (Veale, et al 1989). The cells were grown in Dulbecco's Modified Eagle's Medium (Bodmer, et al.): HAMS F12 (GIBCO) (3:1) supplemented with 10% fetal calf serum (FCS) in a humidified atmosphere at 37°C. (Ethics clearance reference M02-07-47).

3.3.2 DNA preparation

After trypsinisation and centrifugation of the cultured cell lines, DNA was extracted from each cell pellet by standard proteinase K digestion (performed at a concentration of 2mg/ml for 12 hours at 37°C) followed by phenol:chloroform extraction and ethanol precipitation (Sambrook J 2001). Samples were resuspended in TE buffer (pH 8.5). The DNA yield and sample protein purity was determined using a NanoDrop spectrophotometer at optical density (OD) of 260nm and 280nm. The integrity of DNA was further assessed on a 2% agarose gel.

Peripheral blood samples from 10 healthy volunteers were collected to act as controls. Red blood cells were first lysed twice in ice-cold amonium chloride buffer and the white blood

cell pellet was washed and collected by centrifugation. The DNA was then extracted from the pellet in the same way as the cell lines.

3.3.3 MLPA

Technical concept of MLPA

The MLPA method was initially developed and described by Schouten et al. in 2002.

The principle of the test is as follows: a series of MLPA probes, each consisting of two target-specific hemiprobes, are hybridized overnight at 60°C to the test DNA. Both hemiprobes are cloned and propagated in M13 vectors. One of the hemiprobes is linked to a stuffer sequence of variable length, which allows for the assignment of a different probe length specific for each probe set, irrespective of the human insert size (Figure 3.3). The hemiprobes are also linked to two end sequences recognized by universal the MLPA primer pair. Once hybridized to the target DNA, each set of hemiprobes, designed next to each other and separated by one or two nucleotides, is ligated at 54°C. The ligation products are then amplified by polymerase chain reaction using fluorescently labeled universal MLPA primers. The amplified products are separated based on size by capillary gel electrophoresis. For each probe, the amount of PCR product obtained, as measured by the fluorescence signal intensity, reflects the amount of target DNA in the sample, and each specific probe's relative target copy number can be measured.

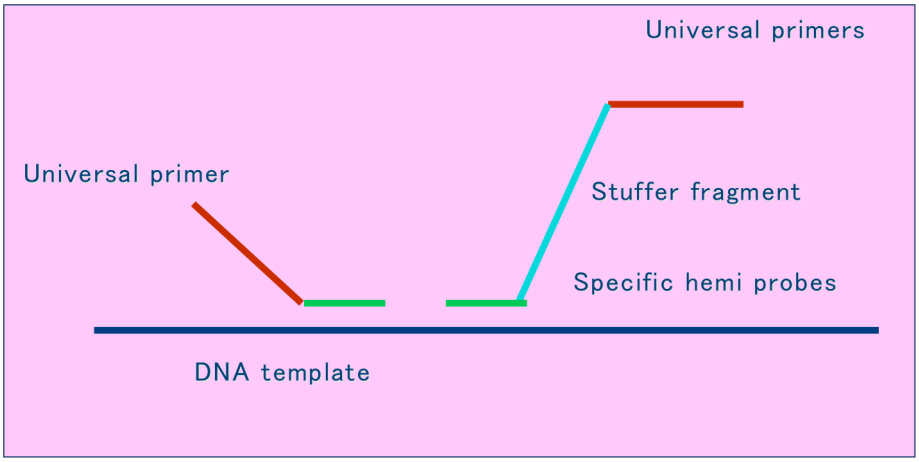


Figure 3.3: *Principle of MLPA. The probe construct at the top shows target-specific hemiprobes in green, the stuffer sequence of variable length in blue and the universal MLPA primers in red.*

MLPA FHIT/WWOX probe set design

In collaboration with Jan Schouten, a set of probes was designed to detect exons' specific copy number change in the FHIT and WWOX genes. FHIT and WWOX have 10 and 9 exons respectively. Although they span over a megabase, they code for small transcripts each of 1095bp (FHIT) and 1475bp (WWOX). Consequently, their exons are small; however, where possible, critical exons found to be preferentially deleted in previous studies were covered by two differing probe sets. Twenty-one out of 23 FHIT/WWOX target probes strictly mapped within an exon and could therefore be used to detect modifications of expression by MLPA. Sixteen probes for the FHIT gene covering all 10 exons were produced; two different sets of target probes covered exons 1 to 5 as well as exon 8. A single probe set covered exons 6, 7, 9 and 10. Seven probes for the WWOX gene covered 5 out of 9 exons, with two probe sets for exons 1 and 9, and a single probe set for exons 5, 6 and 7. Fifteen other probes for other human genes on chromosomes 3p

and 16q, as well as other chromosomes (8p, 9q, 10p, 12p, 16p) were included as internal control probes. The probes were designed and manufactured by Jan Schouten. Please refer to Appendix 2 for each probe's nucleotide sequence reference.

MLPA reaction

The MLPA reaction was performed as previously described (Schouten, et al. 2002) with some minor modifications. In brief, ±200ng of target DNA was resuspended in TE to a final volume of 5μl. Three μl of well-mixed probe cocktail was added and rapidly co-denatured with the target DNA (5 minutes at 98°C). Probe hybridization was allowed to take place overnight (18 hours) at 60°C. A ligase solution (MRC Holland) was prepared the next day, added to the hybridization mixture in appropriate buffer, and left at 54°C for 15 minutes before being heat inactivated. Thirteen μl of the ligation reaction was then pre-mixed with appropriate PCR buffers, MLPA universal primers and Taq polymerase (SALSA, MRC Holland, the Netherlands) for PCR amplification. PCR conditions were as follows: denaturation for 30 seconds at 95°C, annealing 30 seconds at 62°C, extension at 72°C for 60 seconds, repeated for 37 cycles.

After the first round PCR, the amplified products were run on a 2% agarose gel; this provided a good control step for both the ligation and PCR reactions. If a specimen was found not to have amplified, the PCR was repeated with an increased amount (up to 20μl) of starting ligation reaction. The PCR reaction was done in duplicate with the remainder of the ligation reaction otherwise stored at –20°C.

3.3.4 Experimental data analysis

The PCR amplified products were separated by capillary electrophoresis on an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The peak height values were obtained by the GeneScan analysis software (Applied Biosystems) and exported to an Excel spreadsheet for further processing.

Ten DNA samples extracted from healthy individuals were first analyzed to assess the pattern of probe amplification in normal tissue, exclude copy number polymorphism and create a mean "reference control" pattern. In brief: for each probe in each sample run, the relative peak height was calculated by dividing its absolute value by the sum of all probes' peak height for this run. A mean normalized peak height was calculated for each probe and a standard deviation established.

All five ESCC tumor cell lines were then processed and experiments were duplicated. DNA was extracted on different occasions from cell cultures grown at different times. Two to four control samples were included in each MLPA reaction. Their pattern of amplification was verified to be consistent with that of the "reference control" established on ten individuals. The cell lines' amplification patterns were consistent across experiments.

First, the ratios of the cell lines' probe peaks to control sample probe peaks were examined using relative peak height values (each probe's peak height being expressed relative to the sum of all probes' peak height for this run, and expressed as a percentage).

The analysis of data in cancer cells is complex, however, because a large number of chromosomal abnormalities may be present in a single sample. The normalization of probe peaks by relative values (percentage) may therefore distort the overall profile of the probe amplification pattern to the degree of copy number change of some of the probes. Other

groups, using differing MLPA probe sets in cancer cells (Buffart, et al. 2005; Postma, et al. 2005), have therefore normalized the ratio of peak heights using the absolute peak height of "control probes" in the tumor divided by the peak height of the same "control probes" in the reference specimen. Assigning this ratio to one, it can be used to normalize all peaks in that run. This approach avoids the artificial co-dependence of relative peak values; however, the choice of adequate control probes is critical due to the number of genetic alterations that may be present in cancer cells and affect the "control probes" themselves. Here, after first normalizing the runs with relative values, an absolute peak ratio, cancer versus reference control, was used, selecting internal control probes whose pattern of amplification did not deviate from that of the reference control in the same run and that were validated by other techniques. In this study, fluorescence *in situ* hybridization (FISH) and cytogenetics data were examined to validate the choice of internal control probes. Interpretation was as follows: a deletion was scored if the ratio was <0.7, and amplification was scored if the ratio was >1.3. Calculations were done in an Excel spreadsheet and graphs were established for visual interpretation. An automated spreadsheet was later developed in Visual Basic Editor (VBA), which took each analysis step into account and is described in Chapter 4.

3.3.5 Fluorescence *in situ* hybridization (FISH)

Metaphases from each cell line were prepared as per standard cytogenetic techniques (the AGT Cytogenetics Laboratory Manual, 2nd ed. New York: Raven Press; 1991), with minor modifications to adapt to permanent cell cultures. In brief, cells were exposed to 3–5µl of Colcemid (KaryoMAX® Colcemid solution, 10mg/ml, GibcoBRL) overnight; alternatively 20–25µl of Colcemid was added for 3 hours. Cells were then treated with

hypotonic solution 0.075M KCl for 20 minutes and fixed in methanol/acetic acid fixative (3:1 v/v). Slides were prepared per standard techniques.

FISH analysis using probes specific to the short arm of chromosome 3 and the long arm of chromosome 16 (Qbiogene) was performed as per manufacturer's protocols (Universal FISH protocol). Briefly, the probes were co-denatured with the target slides at 72°C in 70% formamide, 2xSSC (pH 7) and hybridized overnight in a humidified atmosphere, at 37°C. Slides were washed the next day at 65°C in 0.5 SSC/0.1% SDS for 5 minutes. They were then counterstained with DAPI (0.2mg/ml) (Vector Laboratories) and analyzed on an Olympus fluorescent microscope BX 61 equipped with appropriate fluorescent filters (Chroma). Images were acquired with the Applied Imaging Genus software.

3.3.6 RT-PCR

RNA was extracted from each cell line using the Qiagen RNeasy kit (Qiagen GmbH, Hilden, Germany), according to the manufacturer's instructions. RNA integrity was assessed on a 2% denaturing agarose gel and samples were quantified on a NanoDrop Spectrophotometer. Reverse transcription was performed with the OneStep RT-PCR kit (Qiagen). Five hundred ng of test RNA was mixed with the manufacturer's PCR reaction buffer, 0.4mM dNTPs, One Step enzyme mixture and target-specific forward and reverse primers. Cycling was done on an Eppendorf Mastercycler. Reverse transcription was performed at 50°C for 30 minutes, followed by a hot step at 95°C for 15 minutes. PCR conditions were as follows: 35 cycles with denaturation at 95°C for 1 minute, annealing at 62.5°C for 1 minute, and extension at 68°C for 3 minutes, with a final extension at the same temperature for 7 minutes. The full FHIT transcript was difficult to amplify and it

was found that longer extension times significantly improved the final yield. For WWOX amplification, the PCR conditions were: 95°C for 1 minute, 66°C for 1 minute, 72°C for 1 minute, repeated for 35 cycles with a final extension at 72°C for 7 minutes.

Reverse transcriptase polymerase chain reaction (RT-PCR) for the FHIT full transcript was performed using previously published primer sequences (Fang, et al. 2001) (Appendix 3). Primers were designed for WWOX full transcript analysis. They were: forward primer 5'-GAGTTCCTGAGCGAGTGG-3' mapping in exon 1, and reverse: 5'-GCTCGTTGGAGAAGAGGA-3' mapping in exon 9.

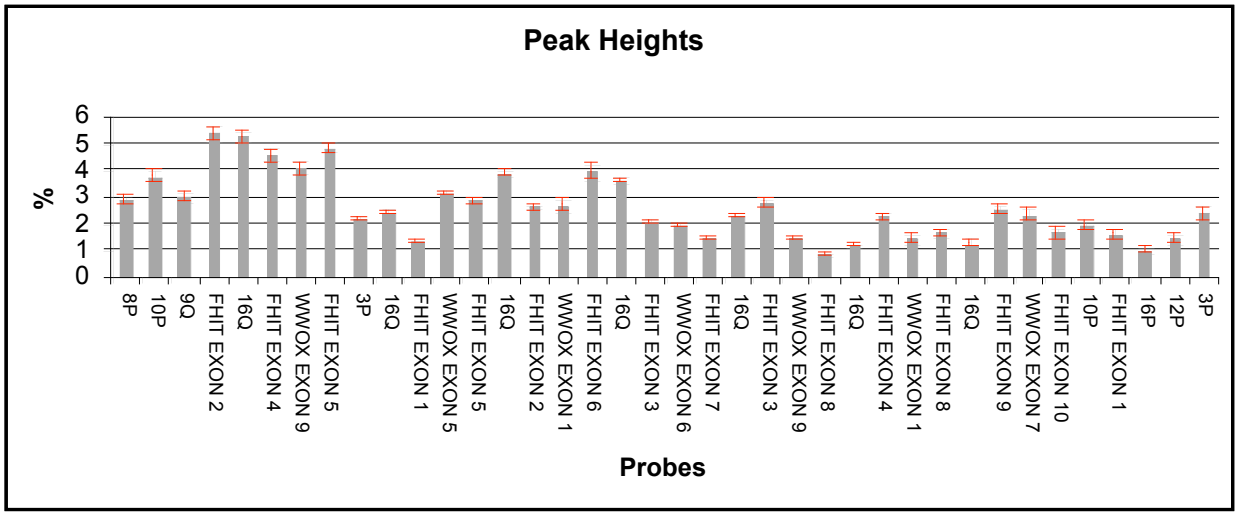
3.4 Results

3.4.1 Amplification pattern in controls

Five males and five female control DNA samples were used to establish the relative pattern of probe amplification in normal samples. A mean reference control pattern was derived (380 probe points) (Figure 3.4 A).

Although the sample size was small, standard deviations were used based on previous MLPA results and the consistency of amplification patterns (Schouten, et al. 2002; Vorstman, et al. 2006). This was done to verify the recommended confidence interval for interpretation of deletions and amplifications in this assay. In the mean reference control, when each probe peak value was normalized to one, the addition of two standard deviations above and below respectively remained within a range of 1.3 and 0.7 (Figure 3.4 B). A ratio of less than 0.7 was considered as a deletion, and a ratio greater than 1.3 was considered as a gain.

A



B

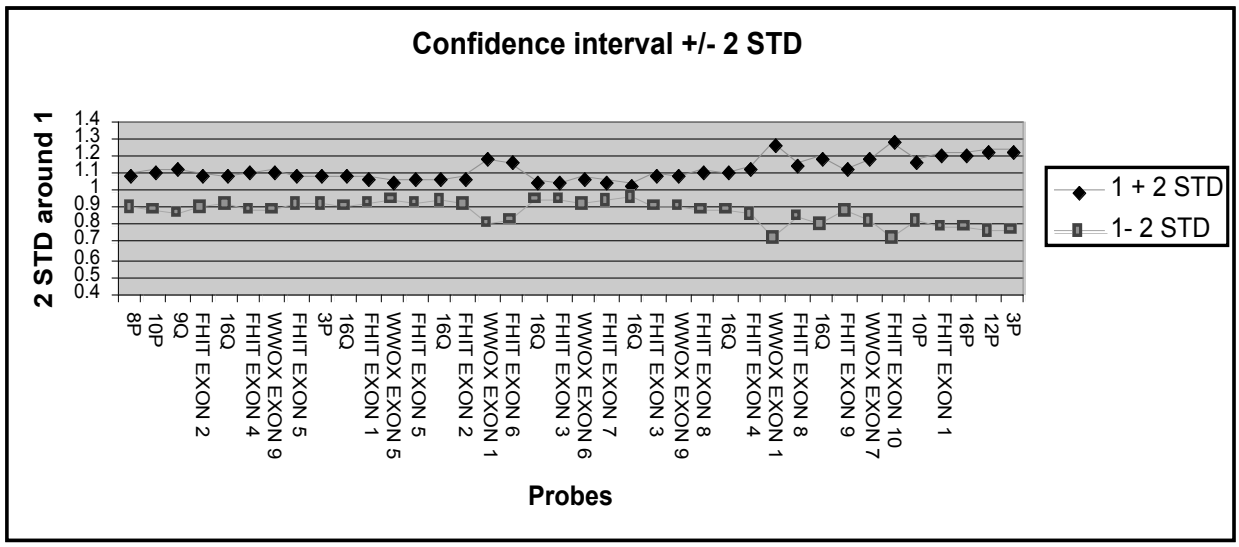


Figure 3.4: *(A) Mean "reference control" pattern of probe amplification established from 10 normal individual samples. Each probe is represented on the horizontal axis in order of its size. The normalized peak heights are expressed in percentages on the vertical axis.*

(B) Two standard deviations above and below the mean peak height (normalized to one) for each probe are shown and validate the scoring of deletion/amplification within the 0.7–1.3 ranges. Standard deviations are generally bigger for larger probes.

3.4.2 Comparison of FHIT and WWOX DNA exon copy number in cell lines and controls by MLPA

For subsequent experiments using the tumor cell lines, control specimens were included in each run and experiments were done in duplicate.

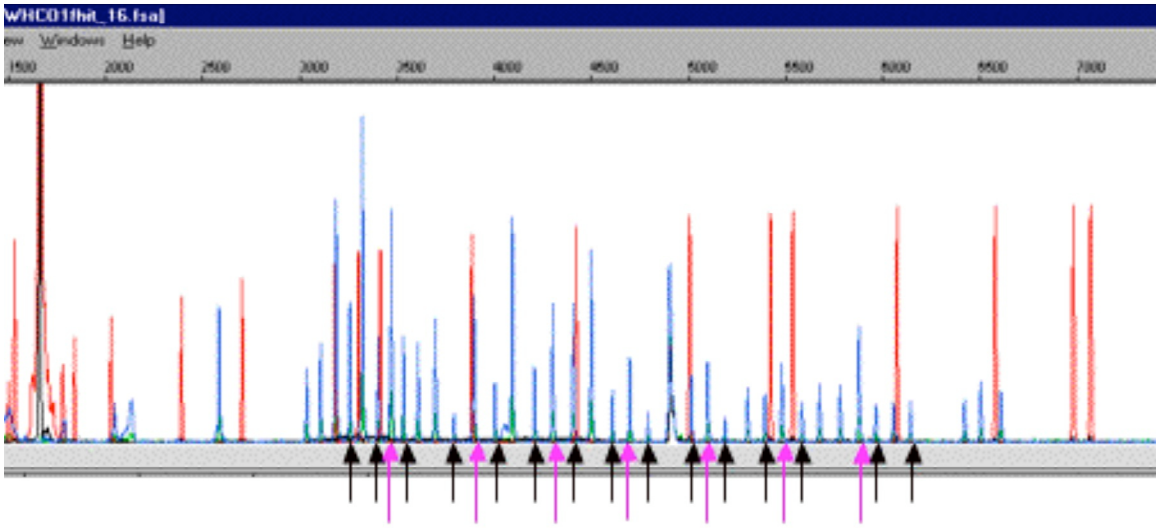
Visual examination of raw data

Upon visual examination of unprocessed MLPA data on GeneScan for each respective cell line, there was evidence of the presence of copy number aberrations (Figure 3.5, A, B and C).

When compared to the normal controls (Figure 3.5 A), FHIT heterozygous deletions were clearly seen in three cell lines: cell line WHCO1 (Figure 3.5 A), cell line WHCO5 (Figure 3.5, B) and cell line WHCO6 (Figure 3.5, C).

A homozygous deletion for FHIT exon 5 was also detected in cell line SNO (Figure 3.5, C).

WHCO1



CONTROL

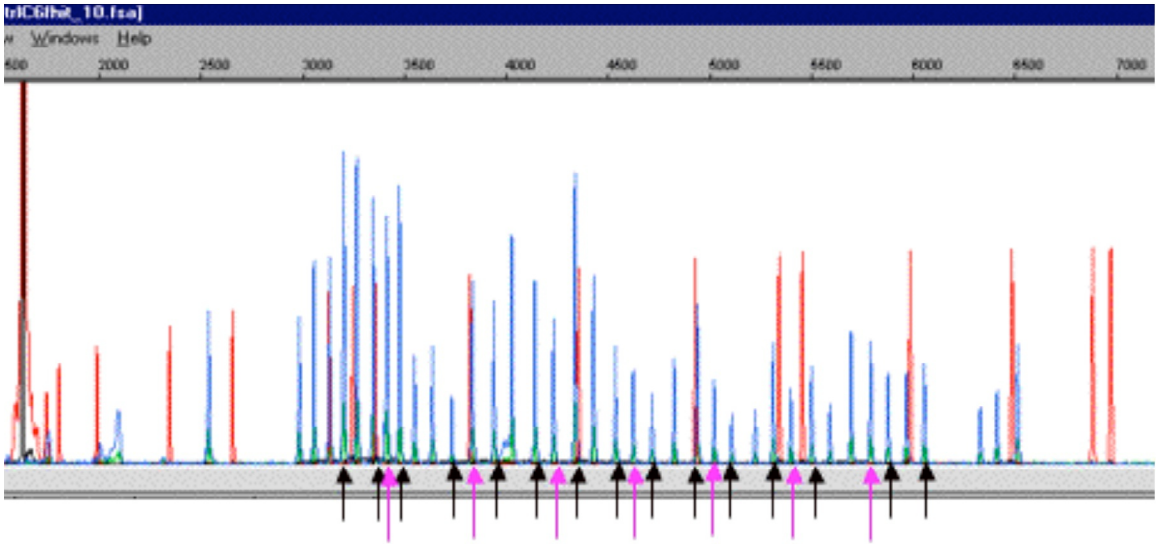


Figure 3.5: *(A) The probes' fluorescent intensity is plotted on the Y-axis and the probes are ordered by size on the X-axis in the same order as on the Excel graph shown in Figure 3.4. Cell line WHCO1 raw data are represented with black arrows to indicate each FHIT-specific probe. The pink arrows indicate the WWOX probes. A consistent decreased signal intensity for all FHIT probes could clearly be seen when compared to the same probes in the control specimen placed underneath. On the other hand, the WWOX probes shown in pink were not deleted when compared to control.*

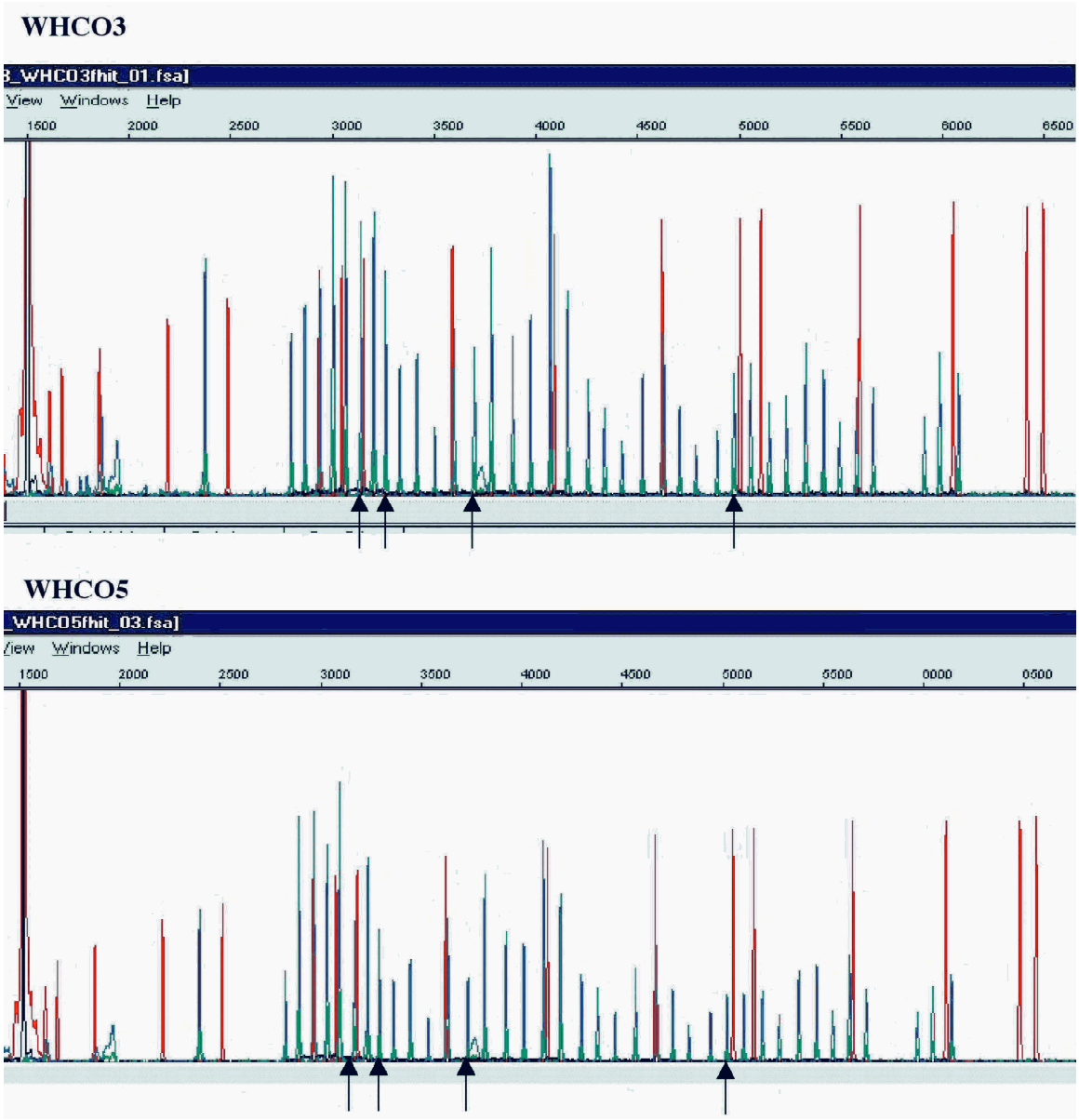
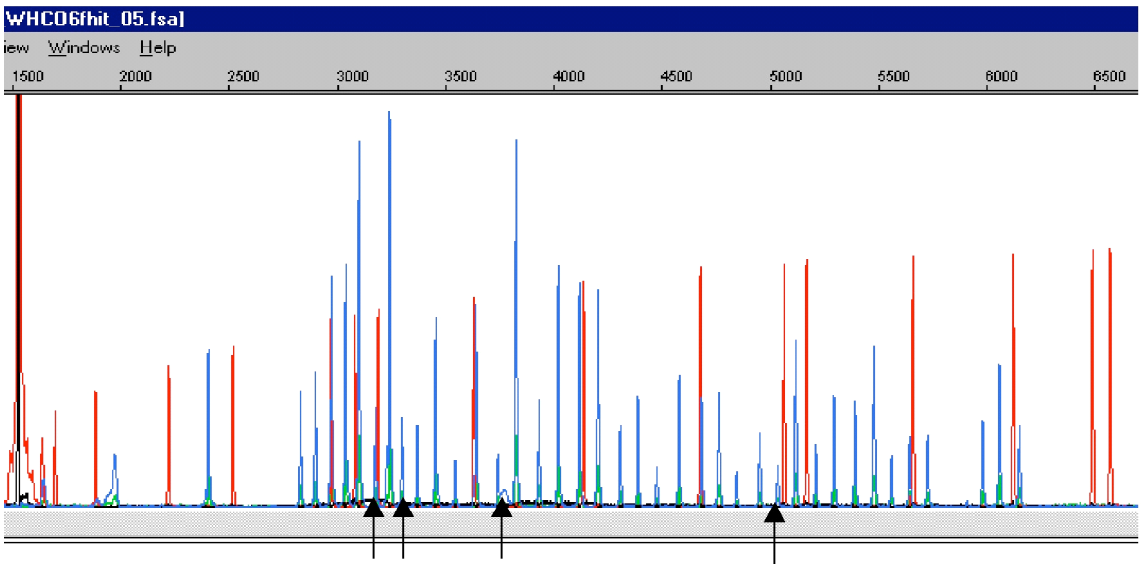


Figure 3.5: *(B) Data for cell line WHCO3 on the opposite did not show any major probe amplification differences (only probes for FHIT exons 4 and 5 have been indicated by black arrows for simplicity).*
Cell line WHCO5 showed significantly decreased signals for FHIT exons 4 and 5, indicated by black arrows when compared to the control (3.5 A).

WHCO6



SNO

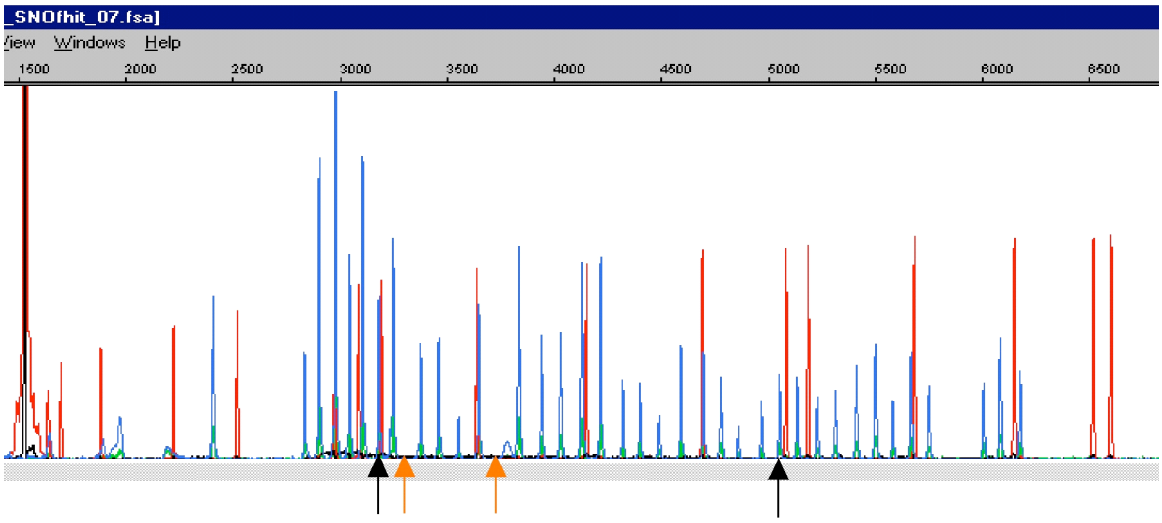


Figure 3.5: *(C) Cell line WHCO6 showed significantly decreased signals for FHIT exons 4 and 5, indicated by black arrows. Cell line SNO had a homozygous deletion for FHIT exon 5 (orange arrows).*

Figure 3.6 shows the same raw data exported in an Excel spreadsheet for WHCO5 and the reference control juxtaposed. A significant reduction (50%) in copy number of FHIT

exons 4 and 5 in WHCO5 was seen. The two control probes located on 3p were not significantly deleted, suggesting that *FHIT* deletion was intragenic.

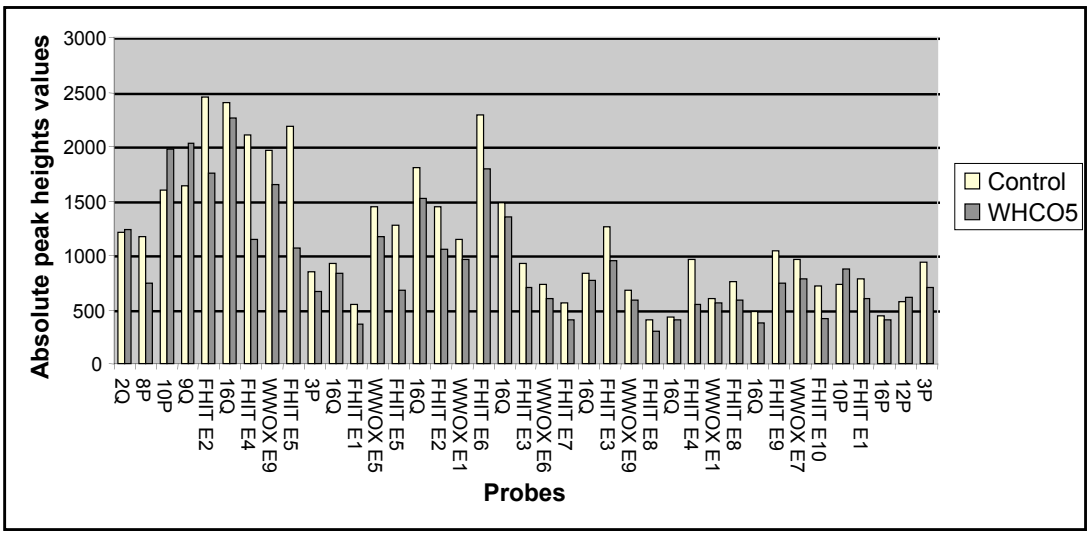


Figure 3.6: *Comparisons of the probe peak heights between the cancer cell line WHCO5 (in dark grey) and the reference control (in white). Hemizygous deletions for FHIT exons 4 and 5 can clearly be seen.*

Evaluation of the relative probe ratios: carcinoma cell lines versus controls

The analysis method mostly used and recommended for MLPA results interpretation is a ratio derived from each probe intensity, where each probe is expressed as a percentage of the total sum of all the probes' intensity for that run (relative peak value). This reflects the consistency of the pattern of probe amplification in normal specimens, as verified in the controls for this particular probe mix in Section 3.3.1. Any probe signal intensity deviation will be recorded in relation to the other probes.

The relative peak height value for each probe, in each cell line, was first compared to the same probe relative peak value in the external control. Cell line WHCO3 had a probe

amplification profile comparable to the control with no significant changes in FHIT and WWOX (Figure 3.7, A) and the homozygous deletion of FHIT exon 5 was evident in cell line SNO (Figure 3.7, B), both confirming visual examination.

However, except for frank homozygous deletions (as in cell line SNO) or for a probe amplification pattern normal or close to normal (as in WHCO3), this method is not suitable for analysis of MLPA data when too many probes are aberrant. It can be seen, for example, in cell line SNO (Figure 3.7, B) that the second and third internal control probes, mapping to 10p and 9q respectively, showed increase in copy number nearly double. This could be artificially created by the extent of probe deletions in this sample that could bias the use of percentage and relative values. Instead, an absolute ratio obtained by normalizing data on an internal control probe should more accurately reflect gene copy number deviation from a normal amplification pattern.

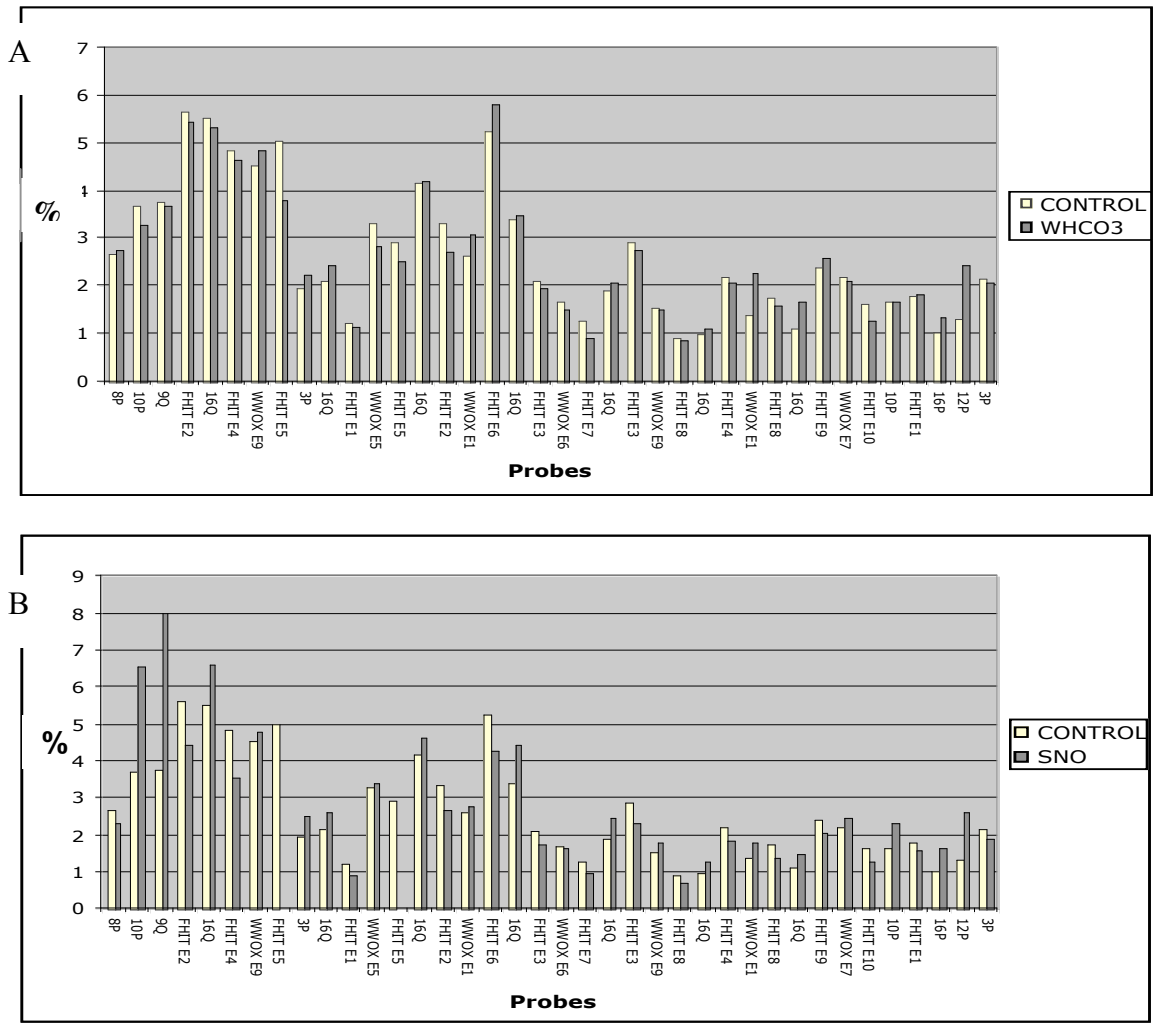


Figure 3.7: *Comparisons of the relative probe peak heights between cancer cell lines (in dark grey) and control (in white).*
The FHIT respective exon specific probes are marked FHIT E1, E2, etc., and the WWOX exon probes WWOX E1, E2, etc. Internal control probes are labeled according to their chromosomal location. A: Cell line WHCO3 showing no significant deletions or amplification in either FHIT or WWOX probes. B: Cell line SNO showed a homozygous deletion of FHIT exon 5 with both FHIT exon 5 probes.

Normalization of ratios on internal control probe

The peak height ratios were therefore normalized to a control probe whose amplification pattern did not deviate from its equivalent in the external control sample in the same run (original GeneScan data).

The results are summarized in Figure 3.8 which shows FHIT and WWOX exons' copy number in each cell line. The ratios derived from absolute values were comparable to ratios seen with relative values (Figure 3.7), but clarified the level of significance of deletions and corrected the level of some control probe over-representation.

Cell line WHCO3 had no significant exon deletions in both FHIT and WWOX. Cell lines WHCO5 and WHCO6 had preferential deletions of FHIT exons 4 and 5; in addition, all remaining FHIT exons in both cell lines had a ratio near or below 0.7 likely to reflect a hemizygous deletion of the entire FHIT gene. In cell line WHCO1, all FHIT exons had a ratio of +- 0.5 compatible with deletion of the gene on one chromosome and therefore hemizygosity. In cell line SNO, the internal control probes that mapped near the FRA3B locus, at 3p22 showed a borderline ratio (0.7) and the more distal probe at 3p25 was not deleted (Figure 3.8), although this cell line had a homozygous deletion of FHIT exon 5. Chromosome 3p deletions in SNO were therefore confined to the FRA3B region. In cell line WHCO1, the probe mapping at 3p22 was deleted, putting the telomeric boundary of the deletion between 3p25 and 3p22. Cell line WHCO5 had a borderline ratio for both 3p probes and in cell line WHCO6, which had a preferential deletion of FHIT exons 4 and 5, both 3p probes showed a ratio close to, but below 0.7. These ratios are likely to reflect hemizygous deletion of the whole 3p region 3p14 to 3p25 in cell lines WHCO5 and WHCO6 (Figure 3.8). These results were later confirmed by SNP array analysis (results not shown in this study). FISH analysis, using a chromosome 3 short arm library 3p, showed two 3p copies in cell line WHCO6 suggesting that the 3p deletion only involved the region 3p14.2p25 (Figure 3.8 B and Table 3.1). Cell line SNO also had 2 copies of 3p and cell line WHCO1 had 4 copies (2 large and 2 small) (Figure 3.9 A and Table 3.1). Cell line WHCO5 had a near tetraploid chromosome number of 117 with 4 to 6 copies of most

chromosomes and four copies of 3p. MLPA does not detect balanced copy number changes, such as tetraploidies. The unbalanced deletions of 3p were however clearly detected by MLPA in the cell lines. Overall the FISH results showed that 3p sequences were present in at least 2 copy number in all cell lines suggesting that the 3p deletions were therefore specific for the region surrounding FRA3B.

By contrast, in none of the cell lines was a WWOX deletion detected, with all ratios having values above 1. Cell line WHCO6 showed an increased copy number for exon 1 and possibly exon 9; chromosome 16 q was, however, present in three to four copies in this cell line (shown by FISH analysis with a chromosome 16q library, Figure 3.9 B). This partial tetrasomy 16 was paralleled by an increased MLPA signal for some of the control probes mapping to 16q and included in this assay. Duplicate probes for WWOX exon 1, each showed a difference in signal amplitude in cell lines WHCO6 and WHCO3. This may be due to the presence of an allelic point mutation in WWOX exon 1 in these cell lines, which would result in the lack of amplification of one allele. Exon 1 was not sequenced in these cell lines and no definite conclusion could be drawn regarding WWOX exon 1 duplication. The same applies to WWOX exon 9 in cell line WHCO6.

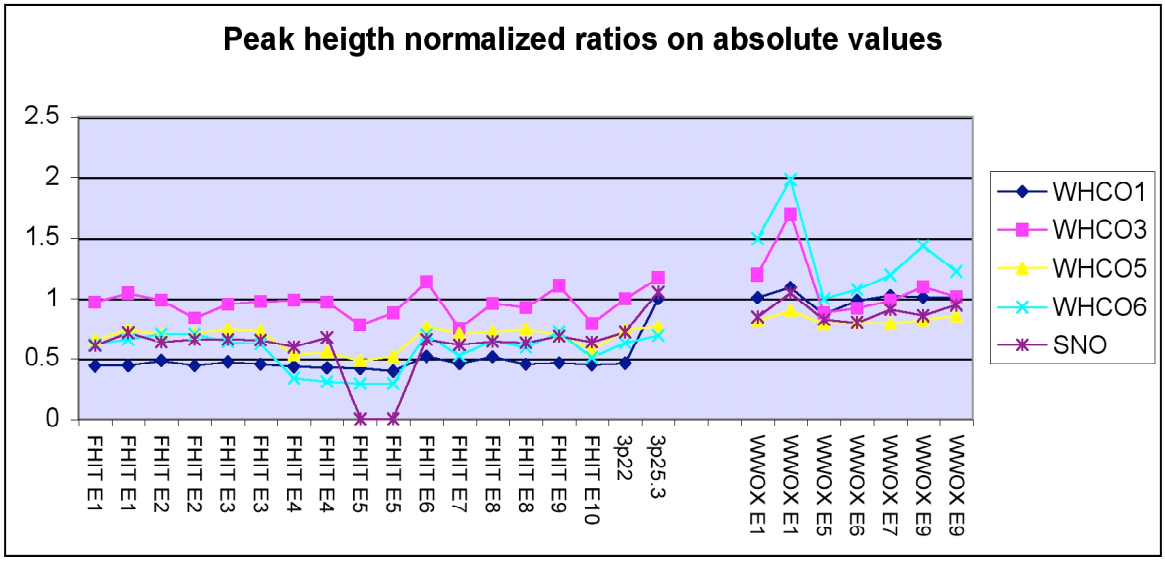


Figure 3.8: *Graphic representation of FHIT and WWOX exon copy number in each cell line. Absolute ratios were normalized to a control probe whose amplification was consistent with the control samples and are represented on the vertical axis. For clarity, exons have been represented in their coding order (horizontal axis).*

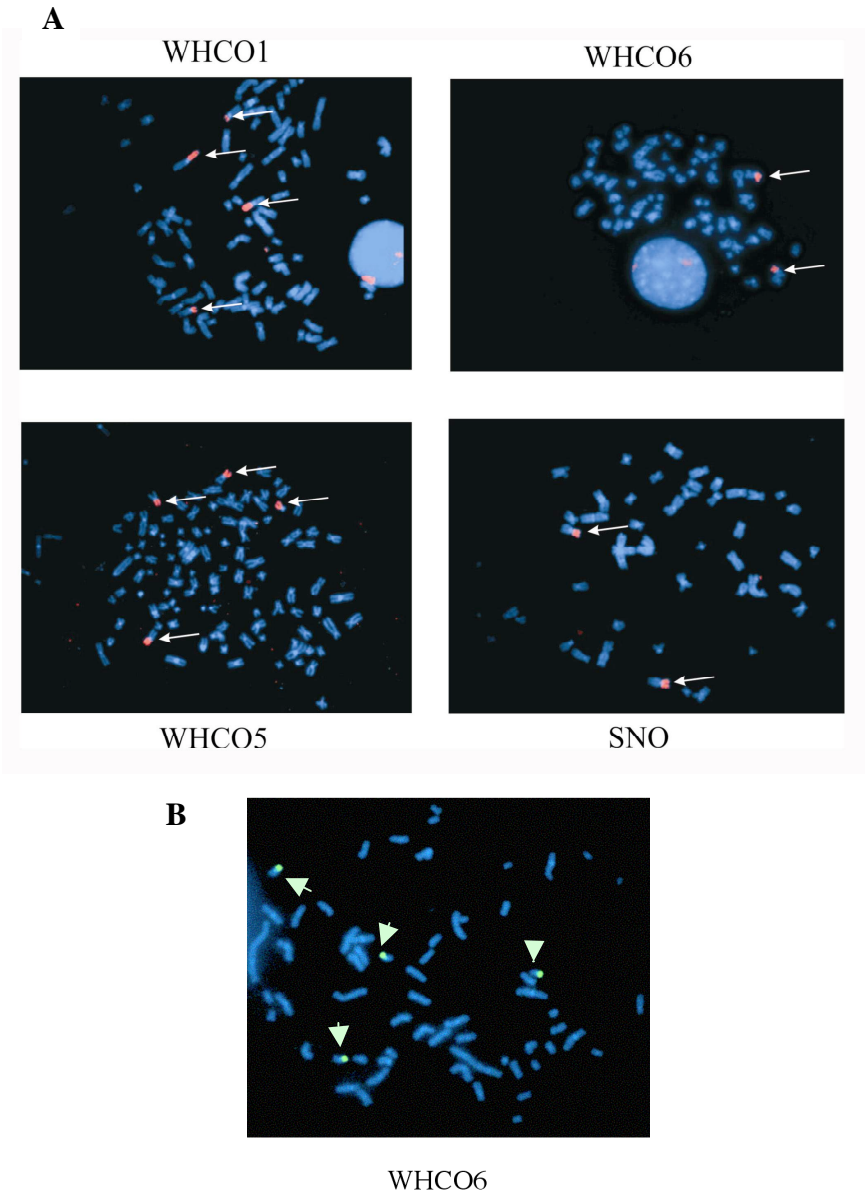


Figure 3.9: *(A) FISH with a 3p specific probe (red and arrowed) in each cell line. Cell-lines WHCO6 and SNO had 2 large 3p signals, WHCO1 showed 2 large and 2 smaller size signals, and cell-line WHCO5 had 4 large signals consistent with the near tetraploid karyotype of this cell-line.*

(B) FISH with a 16q specific probe (green signal) in WHCO6, showing four signals in an otherwise near diploid cell line.

3.4.3 RT-PCR

RT-PCR were performed on 4 occasions, all 4 experiments gave the consistent following results. In each cell line with a deletion in one or more exons of FHIT, there was an aberrant FHIT RT-PCR expression pattern (Figure 3.10 A). Primers that amplify the full FHIT transcript (1072bp) showed products of smaller size than expected (+_900bp in size) in both cell lines WHCO5 and WHCO6; these could represent alternative transcripts of the remaining allele. There were no normal size transcripts. Cell line WHCO1 showed a normal size full transcript (probably corresponding to the normal allele) as well as a small product of about 300bp. The identity of this smaller product is unclear. It is unlikely to be an artifact due to its consistent presence across repeated RT-PCR experiments and may represent an aberrant FHIT amplification product in this cell line. Cell line SNO exhibited two aberrantly sized transcripts (approximately 650 and 700bp). Exon 5, which was homozygously deleted in this cell line, is the first *FHIT* coding exon that ties in with the absence of a normal transcript. RT-PCR specific for exon 5 confirmed this deletion in cell line SNO (results not shown). Cell line WHCO3, which did not exhibit any deletion, had a normal FHIT transcript.

This was a qualitative RT-PCR assay and not a quantitative assay, however it was observed that FHIT PCR amplification was consistently difficult in the cell lines across repeated RT-PCR experiments done with different starting material from separate RNA extractions. On four occasions weak FHIT signals were consistently obtained in all the cell lines in contrast to the control specimens and strongly suggested a decreased FHIT expression.

RT- PCR on WWOX showed a normal amplification product size in all cell lines (Figure 3.10 B) and comparable to the control.

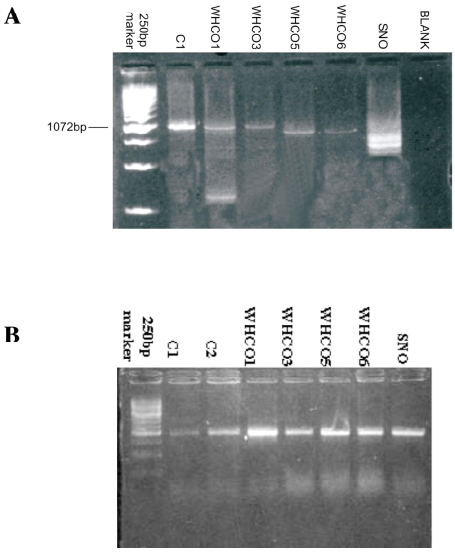


Figure 3.10: *(A) A 2% agarose gel, showing RT-PCR expression of FHIT (full transcript) in the 5 ESCC cell lines and a control (C1). (B) RT-PCR expression of the full WWOX transcript in the five cell lines and controls (C1; C2).*

The combined information regarding chromosome number, MLPA, FISH and RT-PCR is summarized in Table 3.1.

Table 3.1*: Summary of FISH, MLPA and RT-PCR results for FHIT on the five ESCC cell lines.*

Cell lines	**Modal chromosome number**	**FISH 3p copy number ***	**MLPA *FHIT* exon dosage deletion**	**RT-PCR**
Control	46	2	Normal	Normal transcript Strong expression
SNO	47	2+	Homozygous deletion exon 5	Aberrant transcripts only
WHCO1	52	2+ 2 small	Hemizygous deletion exons: 1 to 10	Normal transcript (Weak expression) Aberrant transcript
WHCO3	51	2+	Normal	Normal transcript (Weak expression)
WHCO5	117	4	Hemizygous deletion exons 4 and 5	Aberrant transcript only (Weak expression)
WHCO6	47	2	Hemizygous deletions exons 1, 3, 4, 5, 7, 10 (preferential 4 and 5)	Aberrant transcript only (Weak expression)

*: refers to an additional 3p copy number, full or partial.

3.5 Discussion

This study describes the value of an innovative approach using an MLPA assay specifically designed to scan the exon composition of the FHIT and WWOX genes involved at fragile sites, evaluated in the setting of ESCC cell lines established from South African patients. Many studies have used LOH assays to detect large intragenic deletion in genes at fragile sites (Cheung, et al. 2005; Lee, et al. 2004; Santos, et al. 2004). However, deletions occurring within fragile sites follow a complex pattern of deletions. They vary in size from 10 to hundred of kilobases.(Corbin, et al. 2002; Mimori, et al. 1999; Boldog, et al.1997), they can be discontinuous, affecting separate, discreet regions of the locus (Corbin, et al. 2002;Boldog,et al.1997), and they do not always affect coding exons (Druck, et al. 1997). Quantitative PCR may not accurately detect single genomic allelic loss. In the current approach used in this study the detection of deletions by MLPA is reflected by comparative ratios (Figure 3.8) and does not depend on the presence of informative markers. Hemi- and homozygous genomic deletions encompassing FHIT and WWOX exons were targeted to reflect the exon dosage composition of the genes.

The internal control probe used to normalize the peak heights should be selected in each cancer sample and should have an amplification pattern consistent with its equivalent in control samples (2 to 4 controls for 8 to 16 test samples) included in the same run. Fifteen internal control probes dispersed in the genome increased the possibilities. Only internal control probes that show a pattern of amplification consistent with those of the reference control specimen in the same run should be selected, which was done here.

In this study, there was a correlation between the presence of FHIT exon deletions and the aberrant pattern of RNA expression in the cell lines. Although only qualitative RT-PCR was performed, repeat experiments showed that FHIT expression was generally very low

in the established ESCC cell lines. Due to its exon-based design, this assay would miss intronic genomic deletions that do not affect exon integrity but may inhibit FHIT expression, as was observed in some studies (Druck, et al. 1997). In this regard, it is worth noting that since most of the probes were located within exons, the MLPA assay could be used for RNA studies.

Four out of five SA ESCC cell lines had a FHIT deletion, suggesting that FHIT might be of relevance in SA esophageal cancer, as in other parts of the world such as China and the Asian belt, where the incidence of this type of cancer is high (Kitamura, et al. 2001; Kuroki, et al. 2003; Liu, et al. 2004; Noguchi, et al. 2003). Exons 4 and 5 tended to be the preferential target in three cell lines, in accordance with previous observations (Corbin, et al. 2002; Tanimoto, et al. 2000). Cell line SNO had a homozygous deletion of exon 5 and no functional transcript. The other three cell lines had a hemizygous deletion of FHIT exons, with abnormal expression pattern. Experimental models have shown that either FHIT homozygous deletion or its haplo insufficiency confers an increased sensitivity to carcinogen exposure (Zanesi, et al. 2001;Zanesi, et al. 2005). It is possible that Fhit's decreased apoptotic function (Sard, et al. 1999) facilitates the occurrence of added deleterious genetic events. The evaluation of chromosome 3p copy number by FISH as well as the internal control probes on 3p in MLPA established that these deletions were specific to the FHIT locus.

FHIT has previously been investigated in South African oral squamous cell carcinoma (OSCC), which is known to share similar etiological risk factors with ESCC (van Heerden, et al. 2001a; van Heerden, et al. 2001b). In one study, the authors found reduced or absent Fhit protein in 12 of 17 tumors and aberrant RT-PCR products in a third of the

cases (van Heerden, et al. 2001a). In this study, FHIT exon deletions were associated with aberrant transcripts in the cell lines and observed in 4 out of 5 cell lines.

In contrast to FHIT, no deletion was observed of WWOX in these cell lines. Some studies have shown the simultaneous loss of both genes in several cancers (Guler, et al. 2005; (Ishii, et al. 2003). Since all common fragile sites are sensitive to external carcinogens, these findings raise questions as to the sequence of events involved in these cell lines. The FHIT locus could have been altered early in the patient's pre-malignant cells and a clone be driven/facilitated by its loss, FRA3B being the most sensitive fragile site (Glover and Stein 1987). Alternatively, deletions at fragile sites could represent a later event in the cell lines where FHIT haplo-insufficiency aggravates an already unstable genetic background.

FHIT loss, with or without WWOX loss, has been shown to correlate positively with tumor invasiveness and poor prognosis in several cancers (breast, pancreas, gastric cancers) (Aqeilan, et al. 2004; Arun, et al. 2005; Guler, et al. 2005). At the same time, Fhit protein loss can be detected early in some cancers, including ESCC, and appears to correlate with both early and advanced stages of malignant transformation (Huebner, et al. 1998a; Sozzi, et al. 1998; Kitamura, et al. 2001; Mori, et al. 2000). Whether involved early in the disease or deleted as a consequence of genomic instability and acting as a partner in the progression of the malignant phenotype, it will be important to establish the role of genes at fragile sites in relation to specific carcinogen exposure and disease behavior. FHIT and WWOX inactivation have been found to be coordinated in breast cancer (Guler, et al. 2004; Guler, et al. 2005).The FHIT/WWOX MLPA assay may offer a convenient method to assist in a deletion screening strategy in early tissue lesions associated with heavy carcinogen exposure (esophageal, head and neck, and lung

pathologies). It would 1. assist in the evaluation of the frequency of fragile sites associated gene deletions within coding regions in these early lesions, 2. clarify whether genes at fragile sites tend to be coordinately deleted (with a synergy of biological effects), or if they are independently affected, suggesting that the first deleted gene may take the main driver or facilitator role.

In conclusion, the FHIT/WWOX MLPA assay was suitable to detect fragile site associated genes deletions and showed that four out of five South African ESCC cell lines had FHIT exon deletions, which justified a larger screening in SA ESCC patients and is expanded in chapter 4.

CHAPTER 4

ASSESSMENT OF GENES AT FRAGILE SITES, FHIT AND WWOX, IN SOUTH AFRICAN PATIENTS WITH ESOPHAGEAL SQUAMOUS CELL CARCINOMA

4-1 Introduction

To clarify the extent of the involvement of fragile sites in South African cancer patients and in view of the previous findings in the five ESCC cell lines (Chapter 3), two cohorts of ESCC patients from two different geographical regions were organized. ESCC epidemiology worldwide is intriguing and a tumor genetic screening approach may facilitate the search for etiological factors when evaluated in the light of mutagen exposure and the indices of hereditary genetic components. Two cohorts were selected to assess potential tumor genetic differences.

The following review examines the epidemiology of ESCC world-wide and in South Africa as well as the genes that have been previously investigated in the South African setting.

4.1.1 World epidemiology of Esophageal Squamous Cell Carcinoma

World incidences

As stated before, the incidence of ESCC in South Africa is higher than in the Western part of the world, and it ranks second and third in black males and females respectively (National Cancer Registry of South Africa, 2004).

Worldwide, ESCC incidence is known to have a high geographical diversity, with pockets of higher incidence observed in both East and Central Africa (Chopra and Chopra 1977;

Madebo, et al. 1994), as well as in the "Central Asian belt", including the Caspian provinces, northern Iran (Golestan province), Turkmenistan, Uzbekistan, Kazakhstan (reviewed in (Kamangar, et al. 2007), and parts of northern China (particularly the Huaian region) (Wang, et al. 2006). Given the paucity of cancer registries and control case studies covering these areas, it is unclear whether particular geographical sub-regions (pockets) only are affected, but this seems probable. In South Africa, higher incidence rates have clearly been observed in the Eastern Cape (Pacella-Norman, et al. 2002), with an uneven distribution in different Eastern Cape provinces (Somdyala, et al. 2003).

It was recently observed in the USA that the incidence of ESCC was decreasing in the black American population (although remaining higher than in the white population), while the incidence of esophageal adenocarcinoma was rising (Holmes and Vaughan 2007). The pathology of each malignancy is different, and the epidemiological data support the view that they have differing etiological factors. Adenocarcinoma is associated with gastroesophageal reflux, obesity and good socioeconomic status, and ESCC with poor socioeconomic status and exogenous carcinogens exposure (Holmes and Vaughan 2007; Kamangar, et al. 2007; Pacella-Norman, et al. 2002).

Epidemiologic genetic component

High alcohol consumption, heavy smoking and poor dietary habits have clearly been associated with ESCC in differing parts of the world, including the USA, Africa and China (Brown, et al. 2001; Holmes and Vaughan 2007; Pacella-Norman, et al. 2002; Wang, et al. 2007). However, there are many areas of the world where population groups are exposed to a similarly high-risk profile without showing an increased incidence of ESCC. At the same time, in some areas showing some of the highest world incidence, e.g. Iran or northern regions of China, alcohol is not consumed for religious reasons and smoking has only been very weakly associated with ESCC incidence (Cook-Mozaffari, et

al. 1979; Kamangar, et al. 2007). This suggests that these populations may be exposed to different exogenous mutagens, combined or not with local practices (such as consumption of mold food). Alternatively an increased sensitivity linked to hereditary genetic factors may predispose certain ethnic groups. While population clustering's are clearly observed with ESCC, there are only very rare cases of familial clustering. Few familial cases are associated with the Tylosis locus (within which the causal gene has not yet been identified) (Ellis, et al. 1994). The paucity of familial cases suggests that there are probably, no other, or very few, identifiable high-penetrance alleles responsible for increased ESCC susceptibility. Instead, a role for polygenic factors has been inferred, and a large number of studies have examined the association of ESCC incidence with functional polymorphisms in several classes of genes (Tanaka, et al. 2000; Yang, et al. 2005a; Yokoyama, et al. 2006; Zhang, et al. 2006). Genes involved in the metabolism of environmental carcinogens, phase I/II enzymes (such as CYP1A1, CYP2E1 and SULT1A1) (Casson, et al. 2003) or enzymes controlling alcohol and acetaldehyde metabolism (ADH2, ALDH2) (Yokoyama, et al. 2006) have been preferentially investigated. In South Africa, three case-control studies stand out in this regard: first, a polymorphism in the phase enzyme CYP2E was found to be associated with an increased risk for ESCC (Kalnina, et al. 2005). This finding contrasts with a meta-analysis of literature studies, which reports no significant risk association with phase I/II enzymes except from a weak association with a polymorphism in the CYP1A1 gene (Yang, et al. 2005b). This difference may reflect other environmental selective pressures at work locally. A second positive association was shown by another SA group with the solute carrier gene SLC11A1 polymorphism (Zaahl, et al. 2005). This gene was investigated on the basis of its role in inflammation. Previous reports of mutations in the TP53 CpG island had also suggested the involvement of an inflammatory process (Sepehr, et al. 2001).

Lastly, an increased risk association was observed with a functional polymorphism in the sulfotransferase phase II detoxification enzyme, SULT1A1, among smokers. The risk was further enhanced when present in association with a CYP3A5 polymorphic genotype (Dandara, et al. 2006). SULT1A1 polymorphisms have been associated with an increased sensitivity to a variety of cancers, and in particular those with a history of tobacco exposure (Liang, et al. 2004; Tang, et al. 2003; Wu, et al. 2003). SULT1A1 may therefore represent an active player in the modulation of sensitivity to exogenous mutagens. Other genes, yet to be uncovered, are likely to participate in this variable sensitivity and to account for the discrepancy between the rate of critical functional polymorphisms and the high incidence of this cancer. The extent and strength of these genetic factors remain to be clarified.

Environmental factors

Environmental factors act in a complex interplay with individuals' genetic constitution, and the weight of their contribution in ESCC becomes clear in circumstances where geographical relocation of high-risk population groups correlates with a decrease in ESCC incidence. This was the case for immigrant Arabs who moved to African Zanzibar (Chopra and Chopra 1977) and this phenomenon underscores the importance of such exogenous factors, which may otherwise remain unsuspected.

In South Africa, in addition to tobacco and heavy alcohol consumption (Hendricks and Parker 2002; Pacella-Norman, et al. 2002), other environmental factors have been examined. In particular, Human Papillomavirus infections were thought to play a role, as reported in other parts of the world, such as China (Yao, et al. 2006). Investigations are, however, only partly convincing: although a statistically positive association with HPV16 sero-positivity and ESCC was noted (Sitas, et al. 2007), studies aimed at isolating HPV in

tumor tissues evidenced the prevalence of more indolent subtypes 11 and 39 (Matsha, et al. 2002) and not that of HPV16, or 18, which are the subtypes associated with carcinogenesis. An earlier study (Cooper, et al. 1995) reported a higher prevalence of HPV16 in retrospective tumor samples, but used in-situ hybridization techniques likely to be less specific with regard to subtype specificity than PCR methods used in the more recent studies.

Mycotoxins, and in particular fumonisin B1, a fungal toxin produced by fusarium fungi that grows freely on local corn, has long been identified by association studies as a candidate carcinogen in SA ESCC etiology; this was observed in theEastern Cape (Dutton 1996; Marasas 1995; Marasas 2001) and has also been reported in China (Zheng, et al. 2007). In SA, the incidence of ESCC increased at the same time as the rural population's change of diet from sorghum (not a good host to fusarium growth) to corn. High consumption of homemade beer, prepared with fermented maize, has further been proposed as increasing this risk (Isaacson 2005).

Fumonisin is thought to be carcinogenic and was shown to induce liver cancer in rats (Gelderblom, et al. 2001; Lemmer, et al. 2004; Ramljak, et al. 2000). One study suggests that one mechanism for fumonisin to induce hepatocarcinogenesis involves the stabilization of the cyclin D1 protein, increased phosphorylation of the retinoblastoma protein and disruption of the G1/S cell cycle control checkpoint (Ramljak, et al. 2000). A number of studies have also shown that exposure to fumonisin induces apoptosis, and increases the production of reactive oxygen species (ROS), in a dose-dependent manner (Stockmann-Juvala, et al. 2004; Stockmann-Juvala, et al. 2006) and increases the production of nitrosamines (Chu and Li 1994). Fumonisin was initially not thought to be genotoxic. However, it has since been found to be clastogenic by two separate groups. In

both studies, exposure in cell cultures increased the incidence of both chromosomal damage and sister chromatid exchange (SCE) (Knasmuller, et al. 997; Lerda, et al. 2005). Fumonisin is also known to block folate intake, and in this respect, has been associated with embryotoxicity and increased incidence of neural tube defects mimicking folate deficiency (Marasas, et al. 2004; Sadler, et al. 2002). There are therefore several ways in which this mycotoxin may actively participate in carcinogenesis.

Another mycotoxin, aflatoxin, is also present on corn in the Eastern Cape (van Rensburg, et al. 1990). It has not been strongly associated with increased incidence of ESCC compared to fumonisin, but it should be kept in mind that the combined exposure of both mycotoxins in experimental rat models strongly suggests an enhancement of carcinogenesis (Gelderblom, et al. 2002).

In addition to "passive" environmental exposure, cultural practices may co-facilitate carcinogenesis. In the Eastern Cape, a common religious practice consists of regular and frequent self-induced vomiting, using seawater or mechanical devices. This practice has been shown to positively correlate with chronic esophageal inflammation and could contribute to the high local incidence of ESCC (Matsha, et al. 2006).

Other cultural practices in other parts of the world may facilitate esophageal carcinogenesis, in particular, in regions of the world that have a high ESCC incidence and are strikingly devoid of most of the above risk factors (alcohol consumption and smoking), for example in Golestan in Iran, as well as some regions of China, differing cultural habits expose the population to other sources of exogenous carcinogens (Kamangar, et al. 2007). These include the common use of opium dross (less expensive than crude opium), the by-products of incomplete combustion of coal and tobacco, both of which produce high levels of carcinogenic compounds such as polycyclic aromatic hydrocarbons (PAH), and nitroso-compounds derived from food, tobacco and even water

(Kamangar, et al. 2007). Of relevance to the latter, high levels of salivary nitrites were found in populations in Iran (Eisenbrand, et al. 1980; Lin, et al. 2002).

Most of the above carcinogens can affect genomic integrity at fragile sites. Indeed, sodium nitrites have recently been found to increase damage at nucleotide stretches characteristically found at fragile sites, although this was demonstrated at a low pH (Smith and Schwarz 2006). Active smoking is well known to directly affect the stability of fragile sites (Stein, et al. 2002), and alcohol and fumonisin both affect folate intake, which may contribute to the expression of chromosomal fragile sites (Capaccio, et al. 2005; Stevens and Tang 1997; Yokoyama, et al. 2005). It would be of interest to clarify whether the clastogenic activity induced by fumonisin preferentially targets chromosomal fragile sites.

The above observations strengthen the hypothesis that the combinatorial effect of all or some of these toxic factors and carcinogens may play a role in the initiation of genetic instability early in the disease, and that fragile sites may be early targets (Figure 4.1).

4.1.2 Genes previously investigated in South African ESCC tumor tissues

Local investigations of gene mutations present in ESCC have clarified several facts.

Genes rarely involved

Point mutations in the tumor suppressor genes P53 and CDKN2 have been found to be present, but to lesser frequencies than in other parts of the world (Gamieldien, et al. 1998). Similarly, the presence of RAS mutations in the critical codons of the H, K and N-RAS genes were excluded (Victor, et al. 1990) and a role of mismatch repair genes was also found to be unlikely, given the lack of instability at several microsatellites loci (Naidoo, et

al. 2005). Loss of heterozygosity (LOH), tested at 2p, 3p, 5q and 18q, was also infrequent (Naidoo, et al. 1999).

Genomic regions suspected to have a role

A comparative genomic hybridization study in a cohort of 29 SA esophageal tumors has identified the most frequently amplified chromosomal regions as 3q, Xq, 8q and 2q, with highest amplifications observed at 2q and 8q. Deletions were most frequent on 19q and p, 4p and 18q (Du Plessis, et al. 1999).

Another recent study suggests that chromosome 17q may be the target of frequent deletions (Moodley, et al. 2006). The authors looked for loss of heterozygosity at six markers spread along the long arm of this chromosome, and found high frequencies of loss in particular at 17q21 and 17q25. They suggest a role for Tylosis esophageal cancer (TOC) locus associated with a rare familial skin disorder and early onset of ESCC (Ellis, et al. 1994), and deleted in some cases of sporadic esophageal cancer from differing geographical regions (Iwaya, et al. 1998; Shahabi, et al. 2004). While the marker at 17q21 investigated by Moodley, et al. 2006 is hundreds of megabases away from the putative gene locus, LOH at 17q25 is suggestive, and investigations of the TOC locus alterations are worth pursuing. The cytoglobin gene has recently been proposed as the candidate TOC gene and it was found to be downregulated in both familial and sporadic ESSC, through trans-allele inactivation and epigenetic regulation respectively (McRonald, et al. 2006).

4.1.3 Fragile sites and ESCC

LOH at FRA3B, and loss of the FHIT protein product or expression, has been documented previously in ESCC in several parts of the world (Kitamura, et al. 2001; Liu, et al. 2004;

Mori, et al. 2000; Noguchi, et al. 2003). The WWOX gene at FRA16D has only been investigated in one study on esophageal squamous carcinoma (Kuroki, et al. 2002). Genes at fragile sites have not been previously investigated in South African ESCC.

While numerous studies have established FHIT as a tumor-suppressor gene, it has also been suggested that its inactivation may be an ubiquitous event occurring in the course of carcinogenesis in many cancers (Bavi, et al. 2006). The relationship between chromosomes' "fragility" and exposure to exogenous carcinogens is however worth examining in the context of populations exposed to carcinogens and deletions at fragile sites need to be investigated. Indeed, nicotine was recently found to inactivate the FHIT gene in human esophageal squamous epithelial cells by hypermethylation (Soma, et al. 2006), and active smoking increases fragile site expression, which in turn may lead to double strand breaks at these loci (Stein, et al. 2002).

Of relevance, and as discussed briefly in Chapter 3 (Section 3.1), both pre-cancerous patient's tumor samples as well as in vitro-induced oncogene-driven cell divisions are characterized by the activation of the DNA damage response and fragile site expression. This was observed in early tumorigenesis of both retrospective paraffin-embedded specimens of a variety of cancers, as well as in in-vitro models by two separate groups (Bartkova, et al. 2005; Gorgoulis, et al. 2005). If fragile site expression therefore represents a "signature of replication stress", a phrase coined by Glover (Casper, et al. 2002; Glover 2006; Gorgoulis, et al. 2005), and specifically precedes cellular transformation, it is reasonable to speculate that genes at these sites may significantly contribute to carcinogenesis.

The figure below summarizes the potential combined effects of the exogenous factors discussed above. Smoking, alcohol, fumonisin exposure and DNA viruses are all

prevalent in SA regions of high ESCC incidence and are all likely to target genome "hot spots", particularly fragile sites.

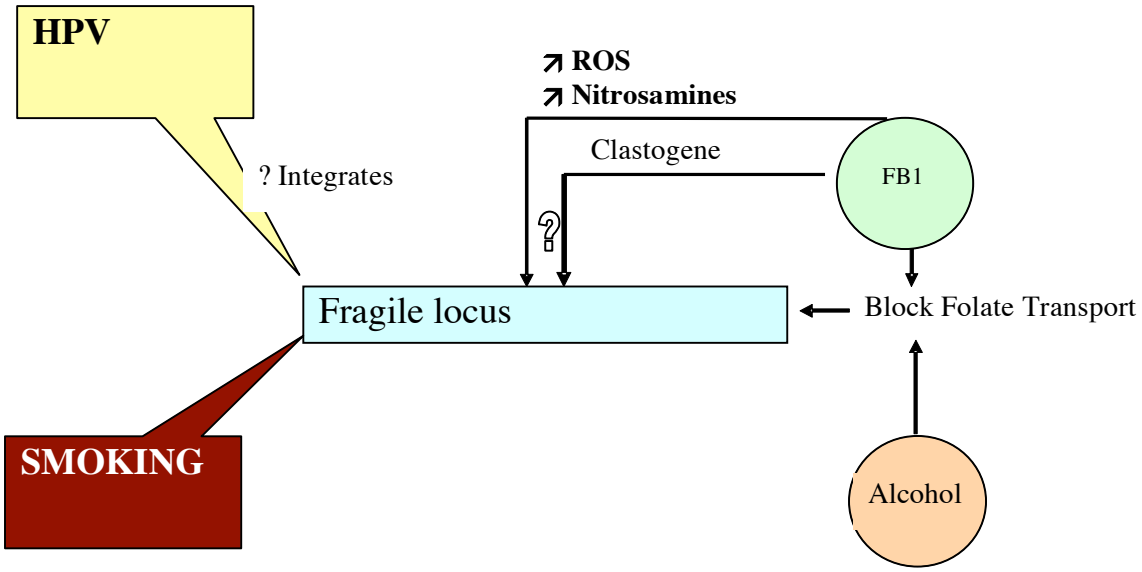


Figure 4.1: *Potential combined effect of exogenous carcinogens in targeting fragile sites.*

An investigation of FHIT and WWOX deletions in SA ESCC may give clues with respect to 1) the role of exogenous carcinogen exposure, 2) the extent of these genes' participation as tumor-suppressor genes in consecutive transforming pathways, and 3) could also provide information regarding a population weakness to repair genes at fragile sites.

4.2 Aims

The aim of this chapter was to investigate two geographically separate cohorts of SA ESCC patients for exon deletions of FHIT and WWOX that lie within the two most common fragile sites, FRA3B and FRA16D; it was speculated that a possible difference in

carcinogens exposure may be reflected by a different rate of alterations at fragile sites. One cohort was from the Eastern Cape and the other from Gauteng.

The MPLA kit described in Chapter 3 simultaneously scans for FHIT and WWOX exon deletions; since these two genes may be coordinately deleted in response to common carcinogens exposure, this kit was used to screen the above cohorts.

The manual analysis of MLPA raw data is tedious and while it can be handled in an excel spreadsheet for a small number of specimens (<15) it becomes difficult with a larger number of samples. A specific objective of this chapter was to develop a VBA encoded software to automate the analysis and interpretation of results.

4.3 Materials and Methods

4.3.1 Patient samples

Collaborators were contacted from two different geographical areas, Gauteng (Chris Hani Baragwanath Hospital), Johannesburg and the Eastern Cape (Umtata General Hospital), Umtata. The project was envisaged with a prospect beyond the scope of the present study (see Conclusions and Future Work); for this reason, we initiated a strategy for the collection of both prospective and retrospective specimens. In both centers, the strategy was to collect both fresh biopsy specimens (from the tumor site and adjacent to the tumor site) as well as left-over paraffin-embedded tissue collected for diagnostic purposes.

Ethical clearance was obtained and specimens were collected upon the signing of a patient informed consent form (Protocol number M02-05-47). The study was explained to the patients by a nurse, on site, in their home language.

4.3.1.1 Chris Hani Baragwanath cohort

Forty-three specimens were collected over a period of 18 months (2004–2005). Of these, 36 represented esophageal squamous cell carcinoma, six of which were carcinoma in situ and 30 invasive ESCC.

For each fresh biopsy, a touch preparation was done in order to assess the approximate percentage of tumor cells representation. Touch preparations were stained (Giemsa) and examined with the assistance of an experienced pathologist. All corresponding patient's diagnostic paraffin embedded specimens matching these biopsies were counter-examined with a pathologist to confirm the diagnosis of ESCC.

4.3.1.2 Umtata General Hospital cohort

Seventy-three retrospective specimens were collected, during the year 2002 and part of the year 2003. All specimens were invasive ESCC, they were stained and ringed by an experienced pathologist, and the marked area was manually micro-dissected.

4.3.2 DNA Extraction

Fresh biopsy specimens

DNA was extracted from prospective fresh specimens, as previously described in Chapter 3, Section 3.2.2.

Formalin-fixed, paraffin-embedded (FFPE) specimens

The formalin treatment of tissue prior to paraffin embedding, results in the formation of protein-protein and nucleic-acid-protein cross-links that can interfere with DNA extraction and MLPA probe hybridization. The MLPA probes hybridize to small regions of the

genome, smaller than 60 base pairs, and a moderate DNA degradation does not affect the results. A DNA extraction and thorough removal of cross-links is, however, necessary in order to have a homogeneous representation of all target sequences, and the emphasis was put on the removal of cross-links during DNA extraction. The following protocol was modified from the recommended method on the MRC Holland website (www.mrc-holland.com/mlpa) and was applied to FFPE samples.

Paraffin from micro-dissected fragments was extracted with three consecutive xylene washes and ethanol extraction (100% EtOH). The pellet was then washed overnight in 1M sodium thiocyanate solution at 37°C, to remove impurities. The following day, the specimens were spun and re-suspended in 200µl lysis buffer (consisting of 30mM Tris pH 8.5, 10mMEDTA, 1% SDS). Proteinase K was added at a final concentration of 2mg/ml and left overnight at 56°C. Heating and extensive proteinase K treatment ensured complete tissue dissolution, and removal of cross-links. If the specimen was found not to be lysed completely, proteinase K was added for an additional 2 hours. After digestion, proteinase K was inactivated at 80°C for 10 mns. This was followed by standard phenol/chloroform extractions and ethanol precipitation. Special care was taken not to leave any traces of phenol nor chloroform in the DNA phase, as they can inhibit subsequent PCR reactions. The specimens were resuspended in TE (pH 8) and evaluated by spectrophotometry (NanoDrop).

Where possible (sufficient amount of DNA), samples were run on a 1.8% agarose gel in order to assess the level of DNA degradation.

For controls, ten normal FFPE tissues were obtained from differing organs as well as differing individuals, and DNA was extracted using the same conditions under the same ethics clearance protocol (M02-05-47).

4.3.3 MLPA

MLPA reactions were performed as described in Chapter 3, Section 3.2.5.

The following additional control steps and reactions were added for FFPE tissues.

First, non-random DNA-protein cross-links may remain in formalin-treated specimens and could affect the amplification efficiency of some probes. In order to verify the completeness of the nucleic acid extraction and the consistency of the FHIT/WWOX MLPA pattern in normal FFPE specimens, two separate reactions were set up with 5 DNA samples from fresh specimens and 6 FFPE specimens extracted from normal tissues of differing origin. The respective amplification patterns were compared.

Second, to further control for possible differences arising from differing DNA extraction procedures, one normal FFPE specimen was extracted in 10 different reactions and the MLPA amplification patterns were compared across reactions.

Third, the presence of "impurities" in the samples extracted from FFPE specimens can decrease the signal of the stronger probes and increase that of the weaker probes (Schouten, www.mrc-holland.com/mlpa). The use of a lower concentration of sample decreases the quantity of "impurities" and is thus recommended. To exclude the possibility that highly aberrant patterns of amplification may be due to impurities, a range of varying amounts of starting DNA was assessed for amplification consistency in two tumor specimens that had highly abnormal MLPA patterns as well as in a control FFPE specimen.

4.3.4 Data Analysis

As in Chapter 3, the PCR-amplified products were separated on an ABI 3100 Genetic Analyser. The signal intensity of FAM-labeled probe peaks was plotted against their

respective size (automatically evaluated against a reference ROX labeled, lambda-size marker). Data were acquired with the GeneScan software (Applied Biosystem) and all the raw data were carefully visually examined prior to being exported into and processed in an Excel spreadsheet. The preliminary visual examination allowed the identification and exclusion of samples that had a failed or resulted from suboptimal PCR amplification; this was either obvious or evaluated against the DNA Quantity (DQ) control fragments that control for the starting amount of DNA sample. If the DQ control probes showed an amplification peak larger than half the amplification peak of the ligation dependant probe (92nt) the reaction was excluded and repeated (see MLPA DNA detection protocol: http://www.mrc-holland.com/pages/support_mlpa_protocolspag.html).

Excel spreadsheet probe analysis software, PAS

The manual export of data to an Excel spreadsheet, cleaning of unnecessary GeneScan information (fluorescent channels not used), elimination of primer-dimer peaks and background, followed by relative and absolute ratio analysis is very tedious. An automated spreadsheet was developed and written in Visual Basic Editor to process the raw data.

At the time of initiation of this project (2003), only a few MLPA kits had been developed and no analysis software was available. An Excel spreadsheet analysis program in VBA has since been developed at MRC Holland. This software has a sub-specific application for each MLPA kit commercially available at MRC Holland. There is no software available for the FHIT/WWOX, since it has not been released yet.

A program was designed and developed as an integral part of this thesis; it was conceptualized with a built-in flexibility that allows users to modify the probe mixture content and the respective probe identity and size. One spreadsheet can accommodate any kit and any kit variation/modification. This option is useful to adjust for probe migration

variations that occur during size separation from one batch to another, as well as from one platform to another, or for the addition of a probe in the probe mix.

Internal control probe

Comparison of probe amplification amplitudes in the control and tumor samples by visual examination of GeneScan raw data allowed the identification of internal control probes whose amplification did not depart from the normal control. FISH was used to validate the internal control probe used for normalization.

Consolidation of analyzed data and heat-map visual representation

While a graphic representation of MLPA results plotting each probe ratio was convenient to visualize a small number of samples it was not appropriate for large numbers of samples (n>15). In order to provide a more comprehensive overview of the complex patterns of deletions in the 73 FFPE specimens from the Eastern Cape cohort, a Visual Basic (VBA) program was developed that translates tumor/control probe ratios into a heat-map image. A color coding was attributed to respective values of the ratios to visualize meaningful deletions/gains along critical chromosomal loci covered by series of probes. The ratios were categorized into nine bin sets (Table 4.1). Ratios between 0.6 and 0.8 were further subdivided into two bins to allow for the flexibility of setting more stringent criteria for deletions (<0.6). Results were analyzed using both a ratio <0.7 and a ratio of 0.6 for comparison. Probes with a ratio of 0.8 to 1.2 were considered to be represented in normal copy number.

Table 4.1: *bin sets attributed to MLPA signal amplitude ratios (tumor/reference control)*

Deletion	Deletion	Deletion	Deletion	Border	Normal	Amp	Amp	Amp
0→0.2	0.2→0.4	0.4→0.6	0.6→0.7	0,7→0.8	0.8→1.2	1.2→1.4	1.6→1.8	>1.8

Amp = amplification

Each bin was assigned a specific color and the program automatically sorted input numerical values. Normal ratios were represented in yellow, deletions in orange turning to red with lowest ratio; amplifications were represented in green, the darker green, the higher the amplification.

4.3.5 FISH

Fluorescence in situ hybridization (Picchio, et al. 2004) was performed as described in Chapter 3 (Section 3.2.4). A centromeric probe for 10p11 (Createch) was chosen for its close proximity to the internal control probe CREM on 10p12.1 (Figure 4.2). The FISH probe was used to assess the number of chromosome 10 near centromeric signals and validate normalization**.**

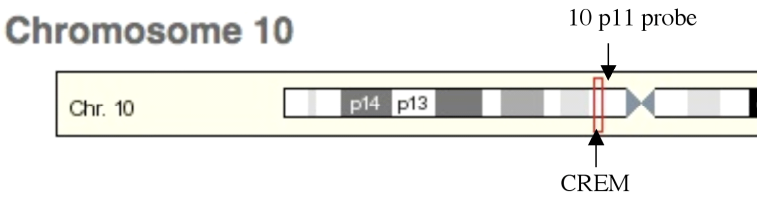


Figure 4.2: *Positions of MLPA internal control probe CREM used for normalization and location of the validation FISH probe at 10p11.*

Tissue sections were ringed by an experienced pathologist to indicate the tumor cell areas and the hybridized slides were analyzed on a fluorescent Olympus microscope with the Genus Software (Applied Imaging). Two hundred interphase nuclei were counted and the mean number of signals evaluated.

4.4 Results

4.4.1 Patient samples

Chris Hani Baragwanath cohort

Of the 43 specimens collected, 36 specimens were confirmed to be esophageal squamous cell carcinoma. Of these, 23 were males and 13 females. The age range was 38 to 95 years old and the mean age was 62.2 in males and 61.7 in females.

Despite the confidence of expert gastroenterologists who performed the biopsies, microscopic counter-examination proved disappointing regarding the percentage of tumor cells represented in many specimens: esophageal cancer cells typically invaginate under an apparently normal epithelium surface and many specimens consisted of a mixture of tumor and normal cells. If less than 70% of cells were cancer cells, the specimen was considered unreliable; only 15 specimens could be reliably assessed within the limits of sensitivity of the MLPA test.

Umtata General Hospital cohort

The age range was 40 to 79 years with a mean of 59 years. There was a gender imbalance, with 54 males versus 16 females. However, after discussion with the local surgeons and pathologists, it was discovered that, surprisingly, there had been no hospital beds for women during these years, which strongly biased this cohort; we recently received a second retrospective batch of specimens that were collected from the same hospital between 2005 and 2006 and found an equal representation of females and males. For this study, seventy three FFPE specimens were re-examined by an experienced histopathologist, confirmed to be invasive ESCC and suitable for this study. As opposed to

the fresh biopsies, it was easy to carefully ring each specimen on histopathology slides and select an area represented by at least (often well over) 70% cancer cells. Each specimen was then manually microdissected for subsequent DNA extraction.

4.4.2 Evaluation of FHIT /WWOX deletions in fresh tumor biopsies

Gauteng cohort

Chapter 3 describes how the MPLA FHIT/WWOX probe mix can be used to assess exon copy number variation in these two genes. Data analysis was performed as previously described (section 3.2.6).

Fifteen ESCC primary tumors were investigated with the assay. Six controls were included in the reactions, and PCR amplifications were performed in duplicate. The MPLA reaction was further repeated when an abnormal pattern was observed. Eleven tumor specimens showed an amplification profile comparable to that of the control specimens with no significant deletions (all ratios between 0.7 and 1.3) of FHIT or WWOX.

Four tumor specimens showed deletions covering or within the FHIT locus. By contrast, and as observed in the permanent ESCC cell lines, none of the WWOX exons were deleted in these primary tumors. Figure 4.3 shows the respective exon probe ratios (tumor/control), for FHIT and WWOX probes. ESCC samples designated CA1 and CA2 showed a pattern consistent with hemizygous deletion of the whole FHIT locus; a preferential deletion of exons 7 to 10 was noticeable in the sample CA1 (Figure 4.3). The tumor specimens CA3 and CA4 had preferential deletions of FHIT exon 5, and FHIT exons 3 and 4 respectively. Of note, 10 additional tumors showed a trend toward decreased signal intensity for one or more exons of the FHIT gene, but these were not significant (ratio above 0.7). Although suggestive, it could not be ascertained whether this

reflected an underlying true FHIT deletion in a small proportion of cancer cells whose significance could be questionable; these were therefore not included as deletions in the final results.

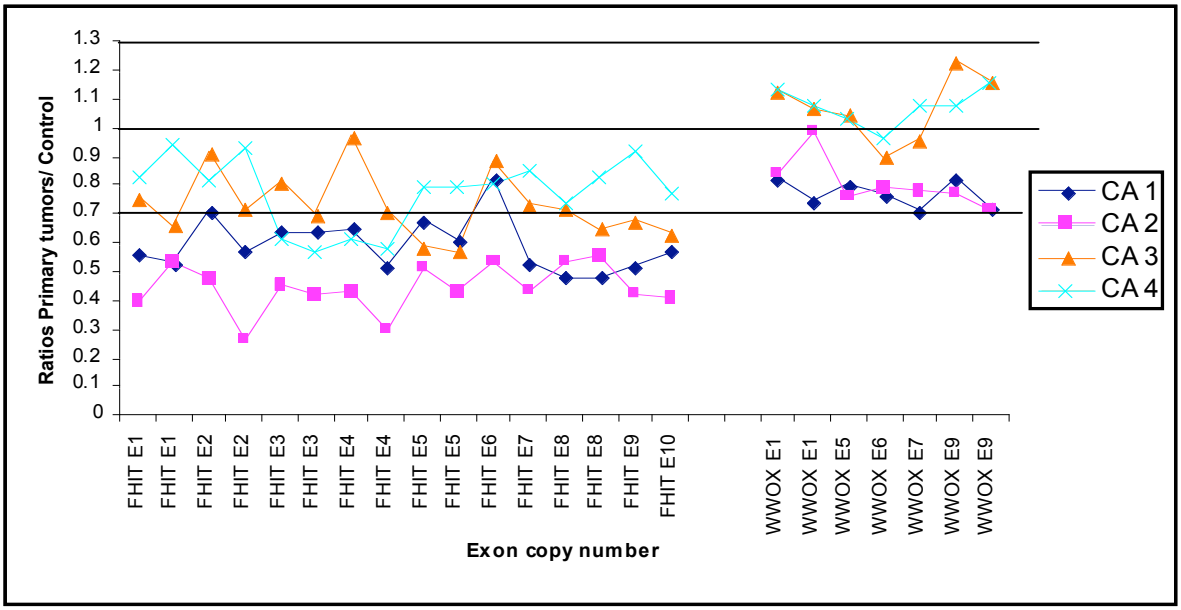


Figure 4.3: *Graphic representation of FHIT and WWOX exon copy number in four primary ESCC tumors.*
Exons in their coding order (horizontal axis) have been represented. Ratios are represented on the Y-axis. On the left, probes representing the FHIT gene show ratios below 1(equal copy number tumor/control), ratios with and partly (CA3, CA4) or wholly (CA1, CA2) significantly deleted exon ratios < 0.7). In contrast, all the WWOX exon probes displayed ratios above the 0.7 ratio line.

4.4.3 MLPA amplification profile consistency in paraffin-embedded specimens

The reproducibility and consistency of MLPA results obtained with FFPE samples when compared to fresh samples, was first evaluated.

FHIT/WWOX MLPA amplification profile from paraffin-embedded specimens compared to fresh tissue material.

Amplification patterns obtained from FFPE specimens were comparable to those of fresh samples within the same experiment (Figure 4.4). DNA extracted as described in Section

4.2.2 from paraffin-embedded specimens did not affect the yield of final MLPA amplification products, compared to those obtained from fresh samples, and was therefore a reliable source of material.

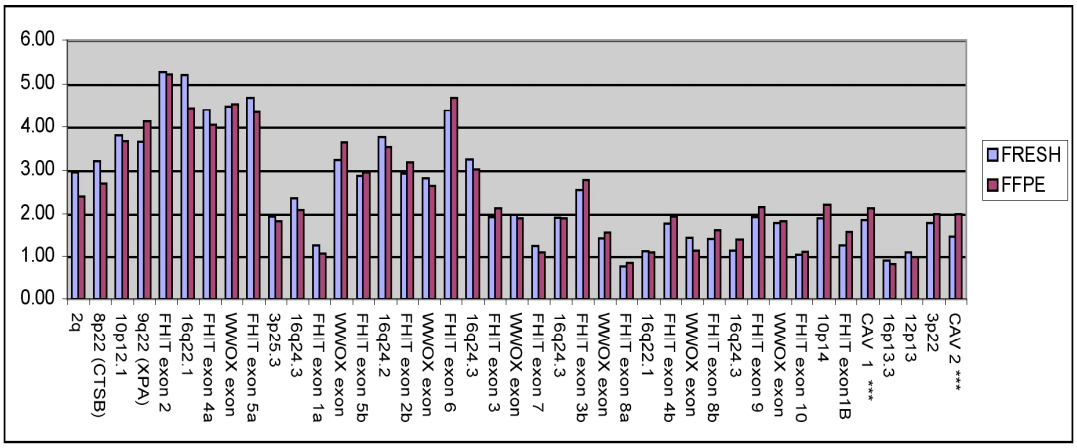


Figure 4.4: *Comparison of pattern of amplification between fresh and FFPE control specimens.*
Mean peak values of five fresh specimens in blue, and six FFPE specimens n=6, in purple.

Effect of differing DNA extractions on FFPE samples

DNA extracted at differing times from the same FFPE specimen obtained from a normal tissue (uterus) gave consistent results with the MLPA run in the same experiments (Figure 4.5).

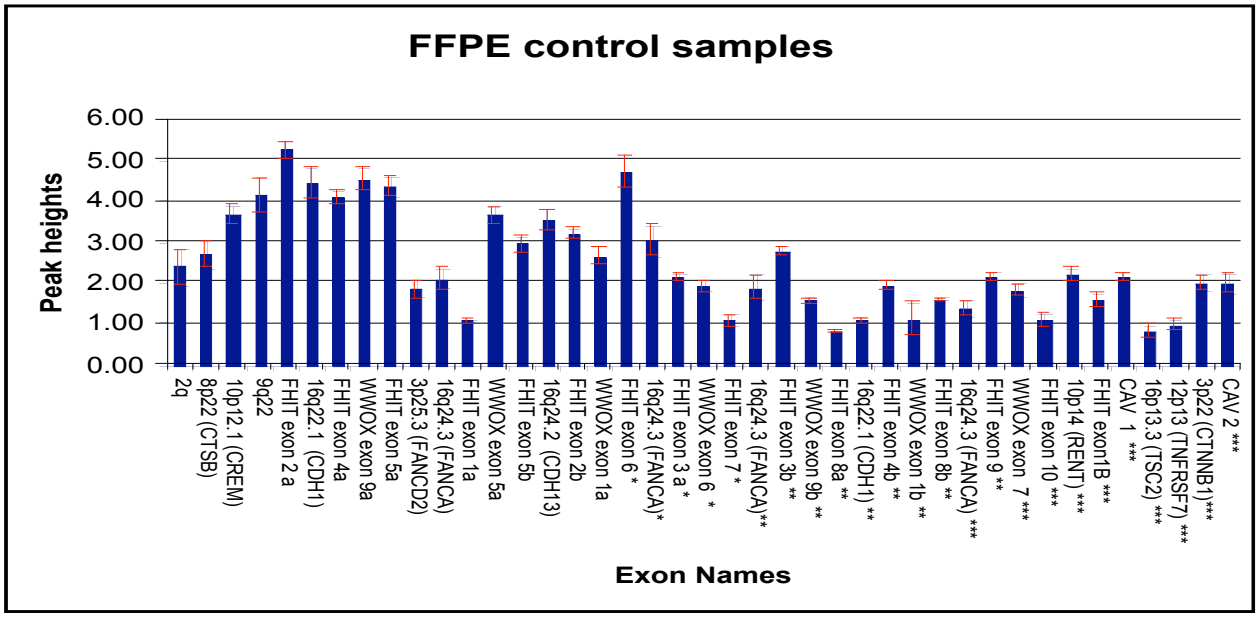


Figure 4.5: *Consistency of MLPA across varying extraction procedures.*
Mean peak height results from the amplification of 10 DNA specimens from the same original tissue material and extracted on different occasions. All samples were run in the same experiment.

The DNA extraction method used for FFPE samples gave consistent and reliable results on MLPA assays. Of note, it was crucial to obtain highly purified DNA to obtain good results. In particular, both A260/A280 and A260/A230 should be evaluated that inform on protein contamination and the presence of impurities respectively. Ideally, A260/A280 should be 1.8 to 2 and A260/A230 just above 2. As with fresh specimens, it is necessary to include normal FFPE controls in each run, as there is inter-run variability.

Varying amounts of starting DNA extracted from FFPE tissues

To further assess the possibility of interference due to impurities, as well as the influence of the starting amount of nucleic acid, varying starting amounts of DNA were loaded in the hybridization reaction. This was performed with a specimen that showed a high number of abnormal peaks, in order to ascertain that aberrations were genuine and not due to impurities or inadequate amounts of DNA (Figure 4.6).

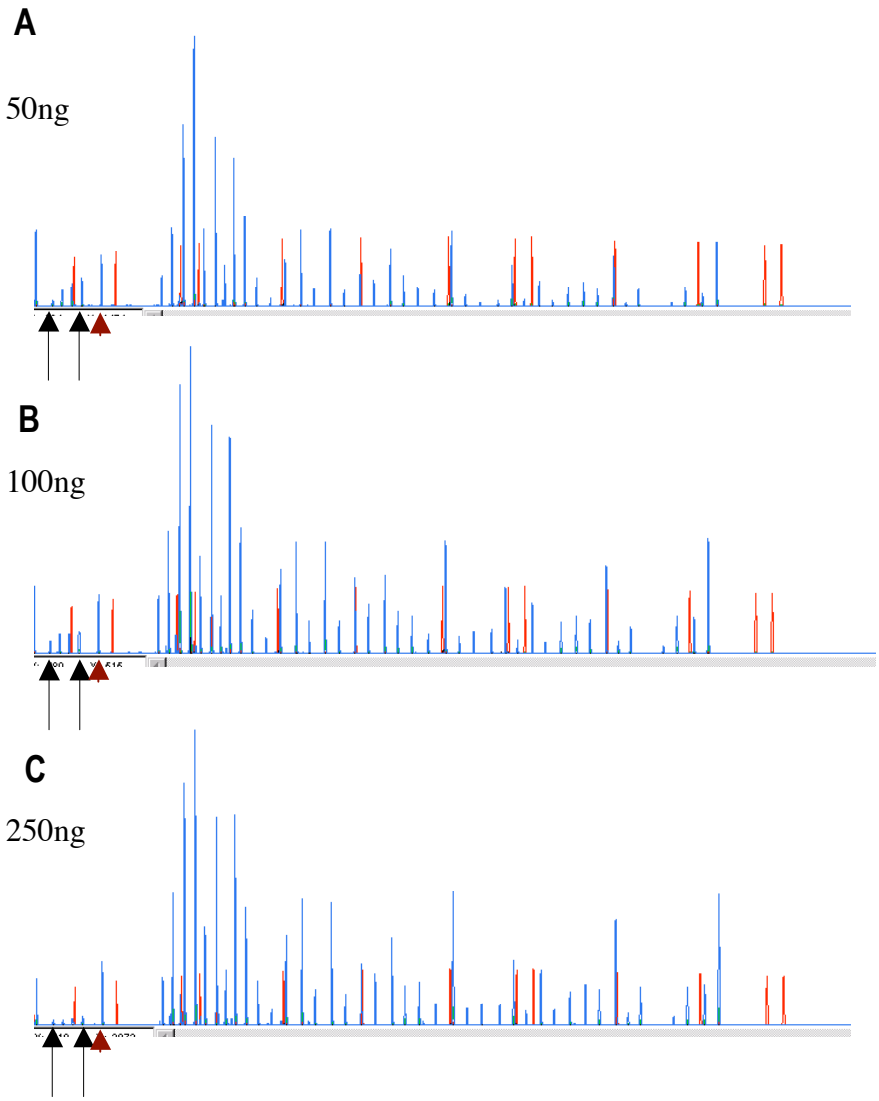


Figure 4.6: *Differing concentrations of starting DNA from a FFPE sample that had a high number of aberrations.*
The two arrows indicate the MLPA DQ control fragments: 64 bp (first arrow), 70 bp and 76bp in between and 82 bp (second arrow). These are standard MLPA control probes contained in all probe mixtures and serve to evaluate the adequacy of the starting amount of test DNA. These four fragments are ligation independent and their peak height is inversely proportional to the amount of DNA contained in the sample. A, with 50ng DNA, the DQ probes are bigger in size than in the C graph, where 250ng DNA was used. When DQ probes' peaks reach the size of the ligation-dependent probe 2q (red arrowhead), this indicates an insufficient amount of starting test DNA and the MLPA sample should be rejected. In A, with 50ng of sample DNA, the MLPA results are clearly not optimum. B and C, 100 and 250ng respectively; these two amplification patterns are comparable and the increased amount of DNA does not interfere with the probe's amplification pattern, as would happen in the presence of impurities.

There was no major variation in the probe amplification pattern between the FFPE and the fresh samples. The presence of impurities was also adequately removed from the FFPE samples using the above nucleic acid extraction method, and did not interfere with the MLPA results. Purity was confirmed by NanoDrop readings of both 260/280 and 230/260 ratios. Of note, it was generally found that a higher starting amount of material gave better results when using FFPE samples (200–300ng of DNA) with MLPA; this may be due to some degree of DNA degradation. A limited amount of DNA degradation, which is virtually inevitable in FFPE, did not affect the final results, which could be explained because the MPLA probes bind to small regions of the genome sizes (±60 mer).

4.4.4 FHIT/WWOX deletions in FFPE samples collected in the Eastern Cape

4.4.4.1 Visual examination of MLPA raw data

A total of 73 samples were processed as per Section 4.2 methods. Three specimens did not have good quality or sufficient amounts of DNA. A direct visual examination of the GeneScan MLPA raw data revealed a recurring pattern in the Eastern Cape specimens, with a trend towards the deletion of WWOX exons, as well as other probes on chromosome 16, including those mapping in the FANCA gene. This "trend" surfaced as a signature of the Eastern Cape specimens (Figure 4.7) and was in contrast to what had been observed in both the five ESCC cell lines (established from patients residing in the Gauteng area) (studied in Chapter 3) and the fresh specimens collected from patients in Gauteng, where no WWOX deletions had been found. Figure 4.7 shows the degree and extent of deletions in chromosome 16. While some specimens only had deletions in the WWOX locus, others involved the gene FANCA, and some extended into proximal regions of chromosome 16 (genes CDH1 and CDH13).

Furthermore, one or several exons in the FHIT gene were deleted in nearly all specimens (especially affecting exon 1, as well as exons 6 to 10) with or without WWOX deletion (Figure 4.8).

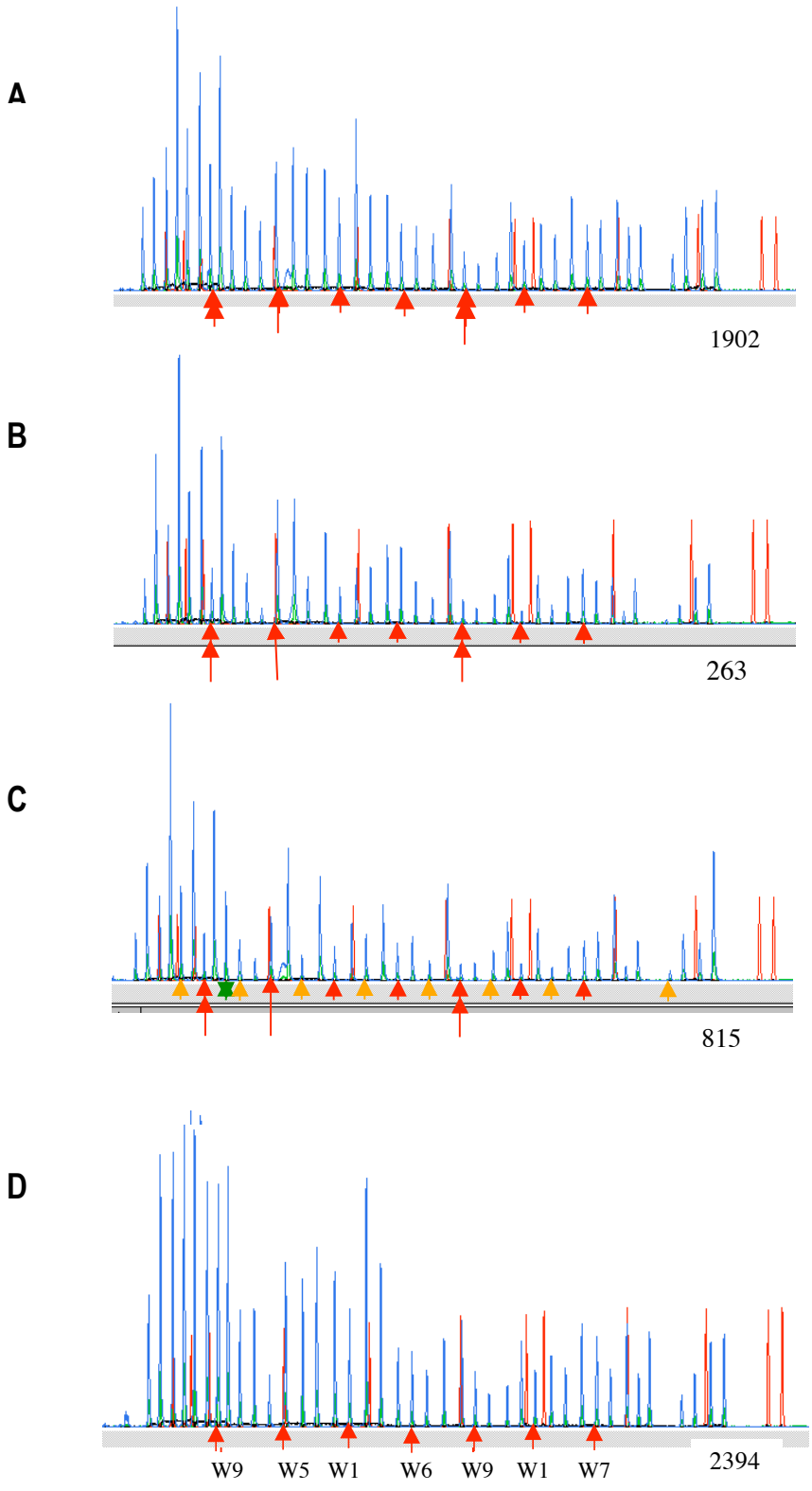


Figure 4.7: *MLPA GeneScan raw data showing deletions of WWOX and chromosome 16 probes in the Eastern Cape FFPE specimens.*
Three FFPE ESCC specimens, A to C, which show an increasing number of WWOX (red arrows) and chromosome 16 (orange arrows) probes with hemizygous deletion. The respective probes for WWOX exons are labeled for reference in D (W9 for WWOX exon 9; W5 for exon 5, etc.). A double red arrow points to WWOX exon 9 and the long arrow to WWOX exon 5 for reference in graphs A to C.

A) Sample showing hemizygous deletion for WWOX exon 9 only (double red arrow). B) Sample showing deletions for WWOX exons 9 and 1. C) Sample with extensive deletions in WWOX exons 9, 5,1 and 7, as well as in other genes on chromosome 16 (orange arrows), probes spanning the CDH1 and CDH13 genes as well as the FANCA gene on 16 q and including a probe on 16 p. The green inverted double arrow indicates the internal control probe that was used for normalization 3p25, FANCD2. D) A sample that showed no significant deletions.

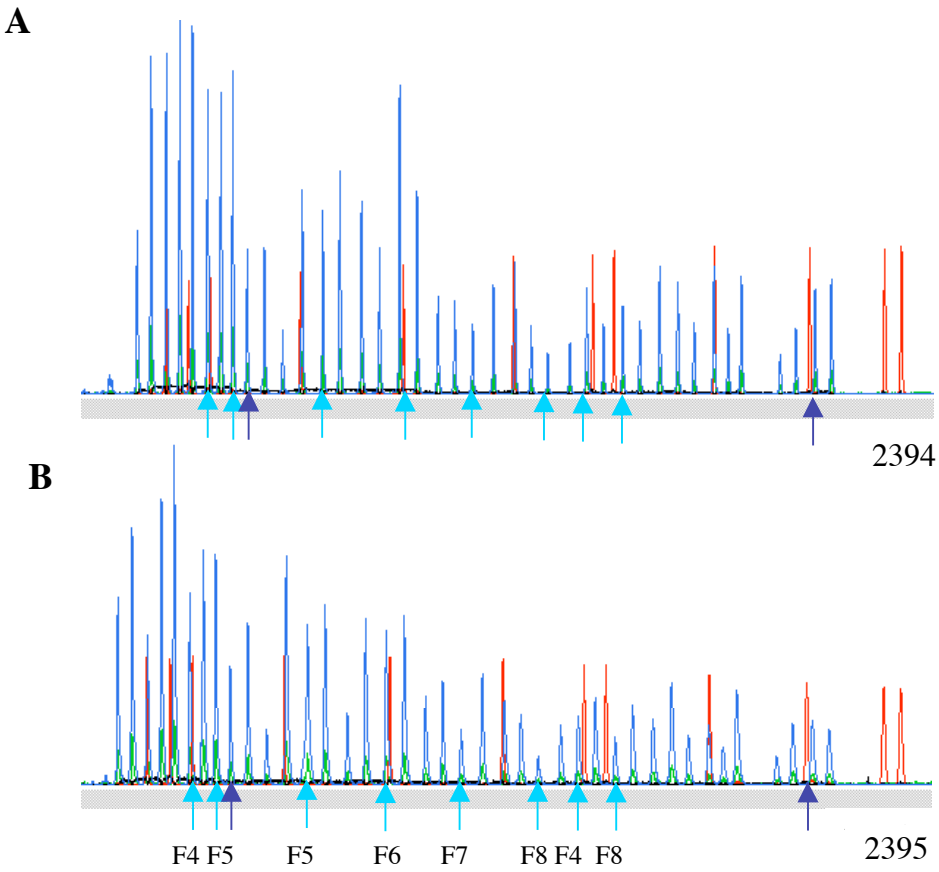


Figure 4.8: *GeneScan raw data on MLPA results for A, a FFPE specimen that did not show any significant deletions and B, for a FFPE specimen with deletions in the FHIT locus.*
FHIT exons are indicated by F, followed by the exon number (pale blue arrows). The two dark blue arrows point to other probes on chromosomes 3p, 3p25 on the left and 3p22 on the right. Deletions are preferentially affecting exons 4, 6 and 8. Both probes on chromosome 3p are also deleted.

4.4.4.2 Validation of internal control probe with FISH

In 85% of all samples from both cohorts, the internal control probes at 10p12.1 and/or 3p25 showed an amplification peak height of the same relative amplitude as that observed in normal FFPE control samples processed in the same experiment. These probes were used for normalization. For control, FISH studies with a 10p11 probe were performed to enumerate the number of signals at this chromosomal location. Eight ESCC specimens for which original FFPE material was still available were hybridized and showed the presence of, on average (on 200 cells), two copies at 10p11 in tumor cells, validating the choice of internal control probe for normalization (Figure 4.9).

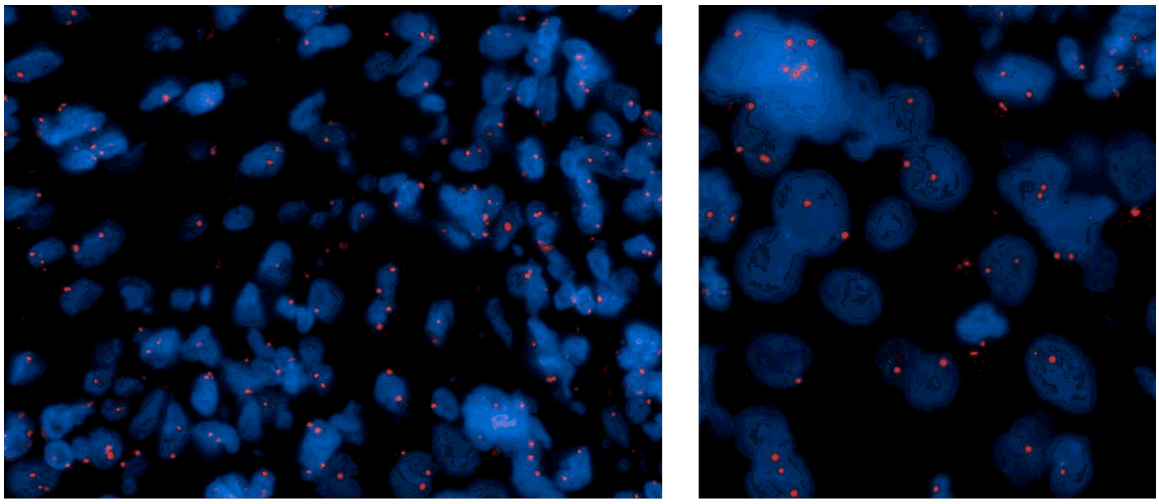


Figure 4.9: *Tumor sections on paraffin hybridized to chromosome 10p11 probe (red signals) on two FFPE specimens. The majority of cells have two signals, validating the choice of the internal control probe used for normalization.*

4.4.4.3 Probe analysis software PAS

PAS is a piece of software that was written in Visual Basic (VBA) as an integral part of this thesis. It allows for the automation of the analysis of MLPA products from the original GeneScan raw data. PAS allows for a customizable analysis for any probe mix.

The steps developed for the automated analysis are briefly described using subheadings for simplicity. Print screens of the actual program are shown to illustrate.

Pre-analysis considerations

MLPA experiments must include control DNA samples extracted in the same manner as the test DNA. At least 4 to 6 controls are included for 12 to 18 specimens.

MLPA reaction products are electrophoresed in capillaries on an ABI 3100 Genetic Analyser. The GeneScan raw data consist of the probes' fluorescent intensity plotted against their size along the X-axis. Sizes references are determined by a Rox lambda marker loaded with the PCR reaction product.

Visual examination of GeneScan raw data

A visual examination of each specimen graph gives immediate quality control information on each sample (see www.mrc-holland.com for details on MLPA DQ control probes). Ligation-independent controls inform on the adequacy of the starting amount of DNA; see Chapter 5, Section 5.2. The ligation-dependent probe controls for the ligation reaction.

A specimen that does not meet the quality criteria of PCR amplification can be excluded at this stage.

Steps involved in the PAS program

First, the GeneScan raw data are transferred into an Excel spreadsheet. The graphic representation of each specimen's original data is also kept in Word format.

Step 1: import controls specimens only in the appropriate VBA-encoded Excel workbook. The probes' sizes of the specific kit in use must match those pre-installed on the input page (Figure 4.10) or be changed accordingly.

Step 2: choose an arbitrary control sample for data filtering and signal-to-noise determination (Figure 4.11). Enter the corresponding sheet name in the input page (Figure 4.10). Determine the signal-to-noise determination by determining the smallest probe-specific probe peak height in the control run (enter manually in the input page; Figure 4.10). (Also see instruction page: Figure 4.11). During the analysis process, all peaks below the smallest specific peak for this run will be filtered out.

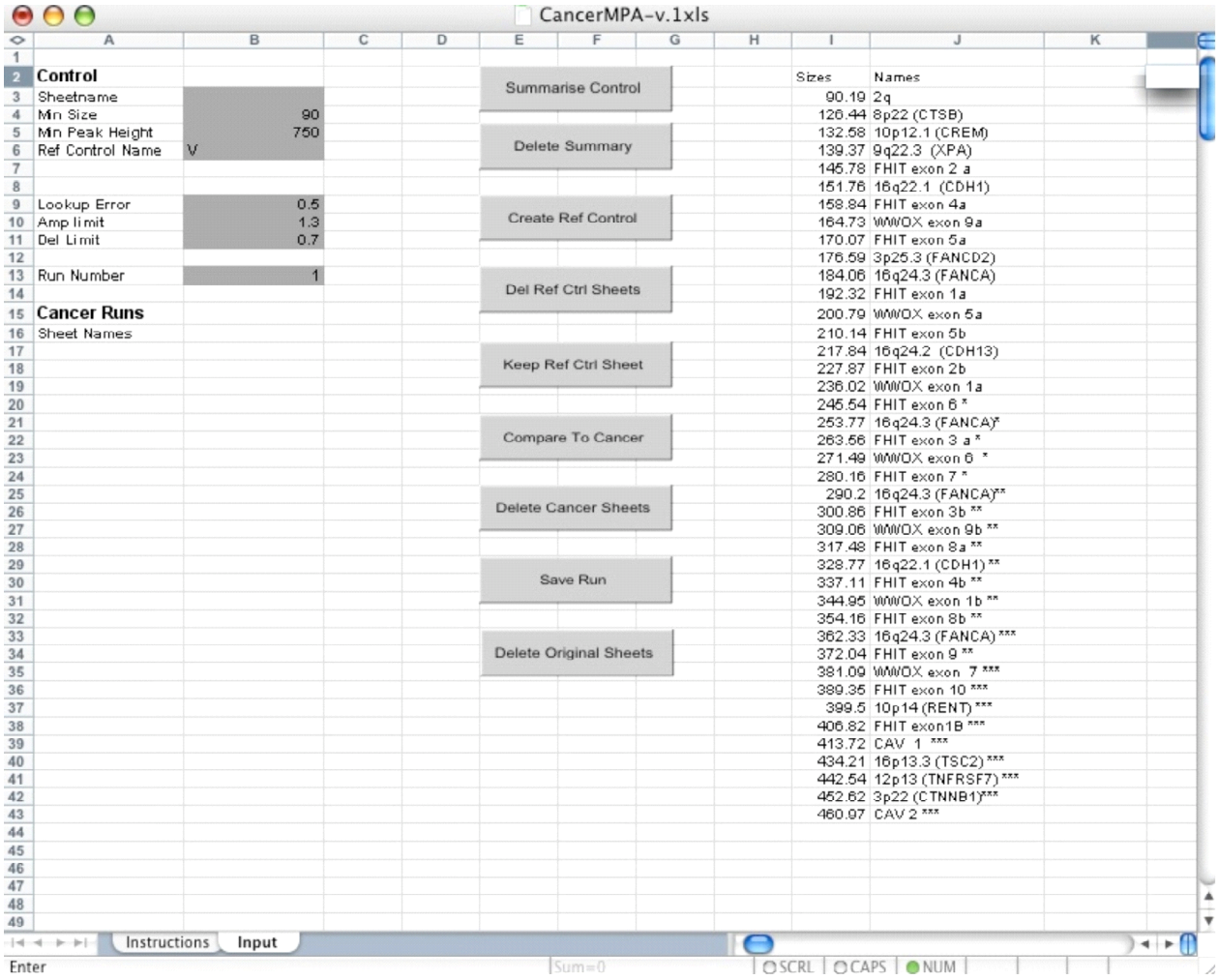


Figure 4.10: *Input page*

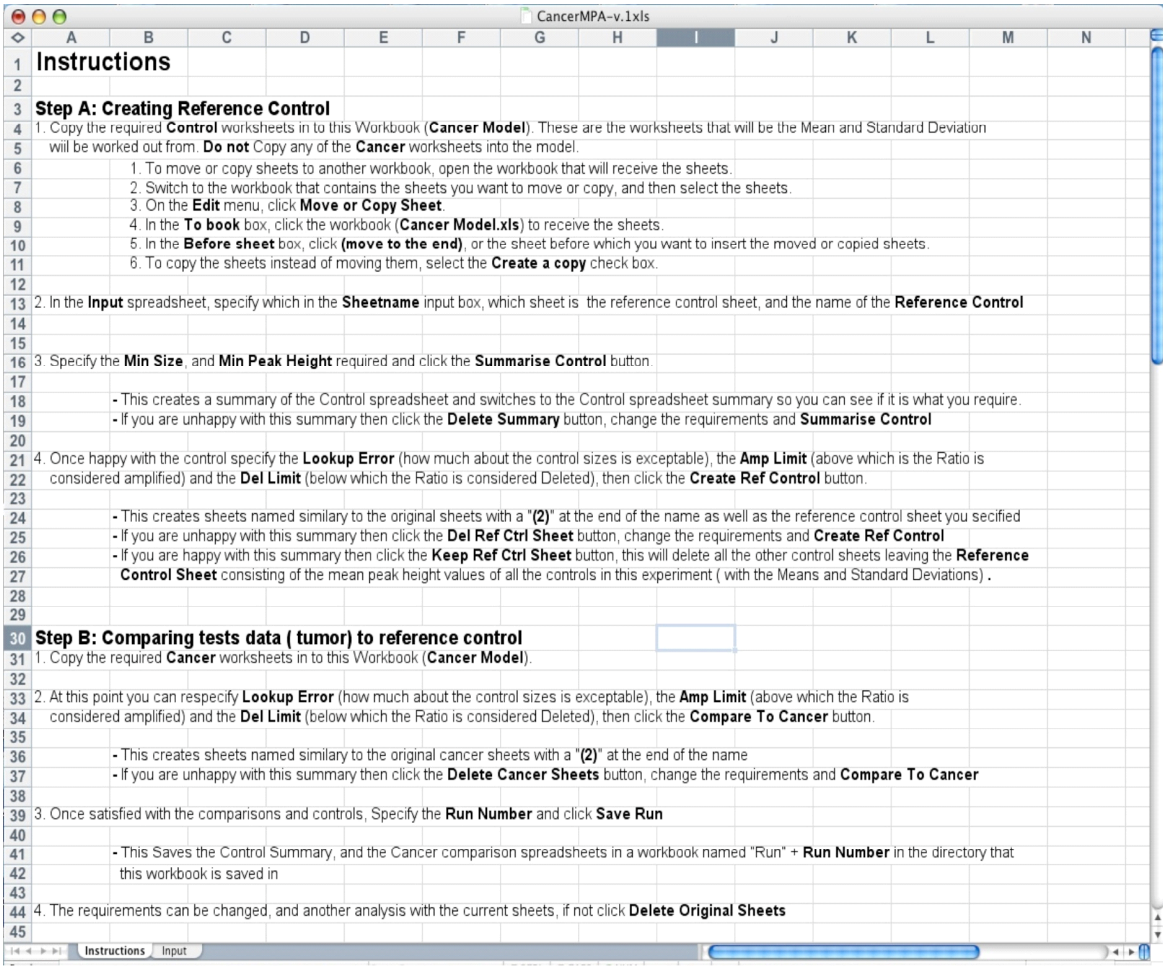


Figure 4.11*: Instruction page*

Step 3: Create a reference summary that can be examined, and either accepted or rejected. A bin set based on each respective probe size has now been established; if a bin set is found empty, it will be interpreted as a full deletion. Set a reasonable error variation around each bin set at this stage on the input page to accommodate inter-sample small differences in the probes' migration speed (Figure 4.11).

Step 4: Process all control samples to create a mean reference control. For each control sample, calculate the relative peak height values and a mean reference relative value for each probe, together with standard deviations (used as an indicator of each probe

amplification variation). This creates a mean reference control sample that can be used for the comparative analysis of tumor specimens run within the same experiment. The mean control is automatically plotted on a graph. Standard deviations (STDs) are graphically represented on the graph (Figure 4.12). If STDs on probes are too large, it may be necessary to adjust the choice of cut-off values set for interpreting a deletion or amplification, i.e. <0.7 for deletion and >1.3 for amplification may have to be changed to <0.6 and >1.4 respectively, and if STDs are too large the particular probe may be considered uninformative.

The analysis of controls informs both on the respective robustness of each probe as well as on the quality of the run in general.

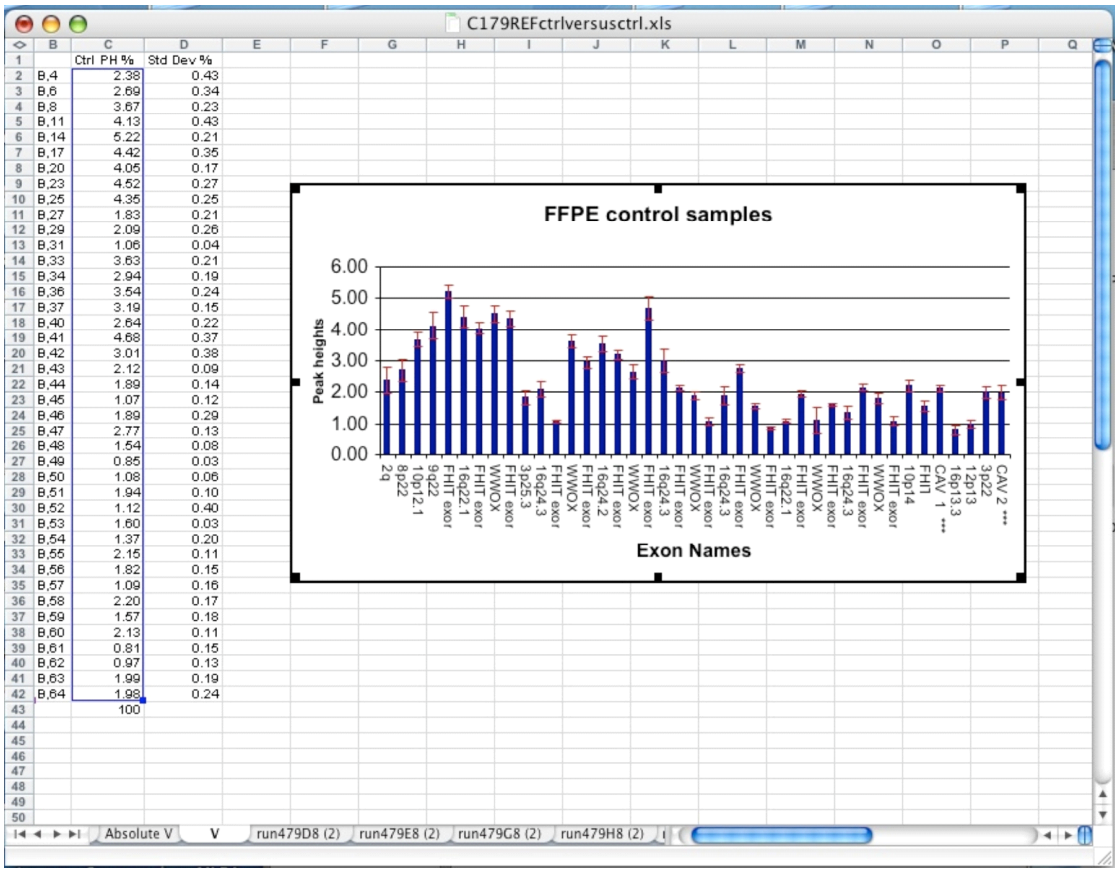


Figure 4.12: *Reference control*

Step 5: delete original control specimen sheets and validate/acceptance of the reference mean control.

Step 6: Import each test specimen spreadsheet into the workbook.

Step 7: Compare test samples to the reference control for this run.

Calculate automatically the relative pattern of amplification for each cancer specimen. Most importantly, calculate the absolute ratios of tumor to control specimens using an internal control probe. The results are expressed in ratios; the probes are ordered in a logical manner compatible with chromosomal representation and visualized graphically.

One sheet is dedicated to each test specimen. If the internal control probe is not valid, an alternative control can be chosen (Figure 4.13).

Step 8: Export the ratios to another spreadsheet to be translated into a heatmap color-coded format.

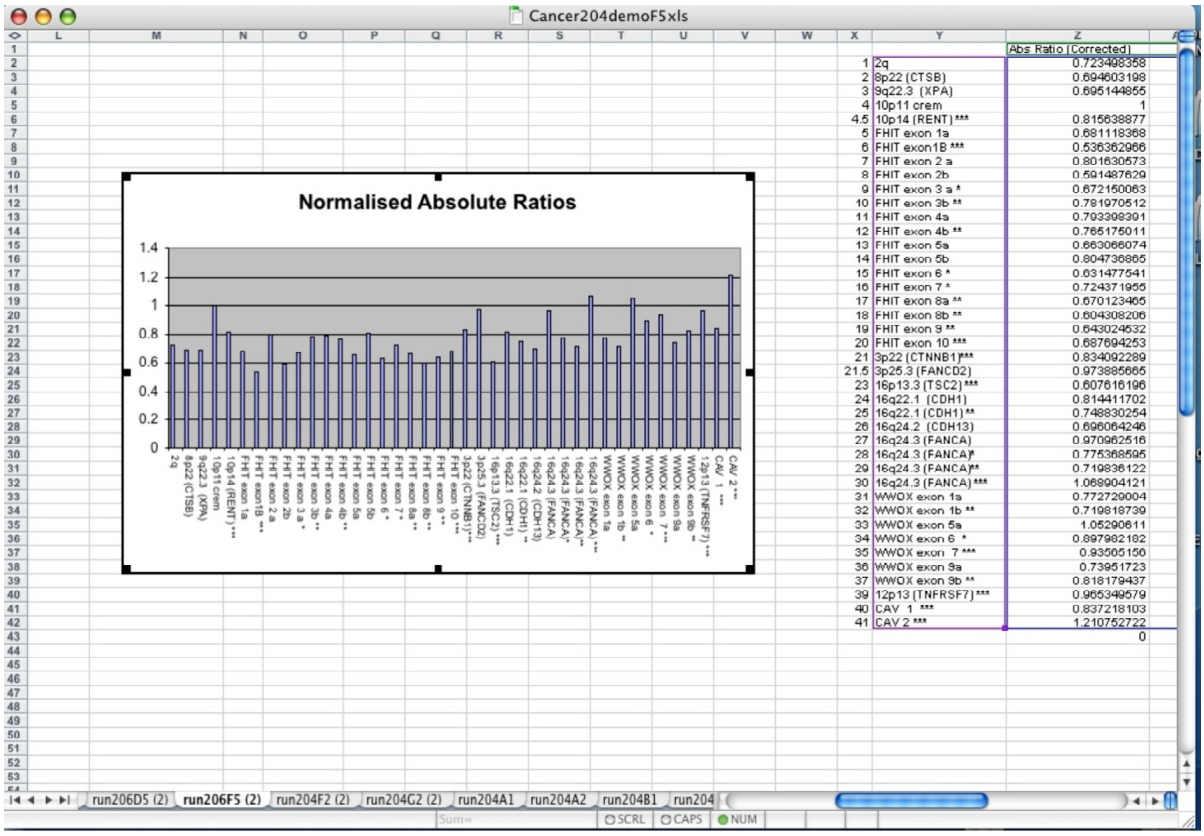


Figure 4.13*: Test DNA result*

4.4.4.4 Consolidated analysis

For each specimen, the original GeneScan raw data were exported to an Excel spreadsheet and analyzed with the VBA Excel program written for this project. Data for every sample were filtered and then normalized on the chosen internal control probe RENT and/or 3p25. The final results were expressed as ratios of tumor versus control for each MLPA probe.

In order to consolidate the normalized probe ratios, these were exported to a new spreadsheet and translated in a visual color codification to facilitate the analysis (Figure 4.14). Deletions were represented in red, amplifications in green and probes with normal amplification patterns were in yellow. Although not the focus of this work two probes for the caveolin genes 1 and 2, CAV1 and CAV2 were included during the course of this study. They both locate in a third fragile site on chromosome 7, FRA7G and there is mounting evidence that CAV1 may act as a tumor suppressor gene (Torres, et al. 2007). Both genes have however been found to be amplified in some tumor cells (including a few esophageal carcinoma cell lines (Torres, et al. 2007; Ando, et al. 2007) and two modes of action have been suggested in cancer cells that depend on caveolin interaction with other cell factors such as E-cadherin (Torres, et al. 2007).

The results were analyzed using the following criteria — counting a deletion in a gene if at least two probes had a ratio below 0.7 and excluding any data showing a discrepancy of ratios between two probes mapping in the same exon. Figure 4.14 shows results on 17 samples. Samples 1, 2, 3, 4, 5, 6, 8, 9, 10, 11 and 12 had FHIT heterozygous deletions (exons 6 to 10 as well as exons 4 and 5 in samples 1 and 10). No FHIT deletions were observed in samples 7 and 13 while the whole FHIT gene was represented in half copy number in samples 14 to 17. FHIT exon 1 deletions are commented upon in section 4.3.5 but were generally concordant with other FHIT exon deletions. Samples 1 to 9 all had WWOX exon deletions that extended on chromosome 16 to include CDH1, CDH13 and FANCA. Interestingly CAV1 was deleted in 10 out of 17 samples and CAV2 amplified in 2 samples.

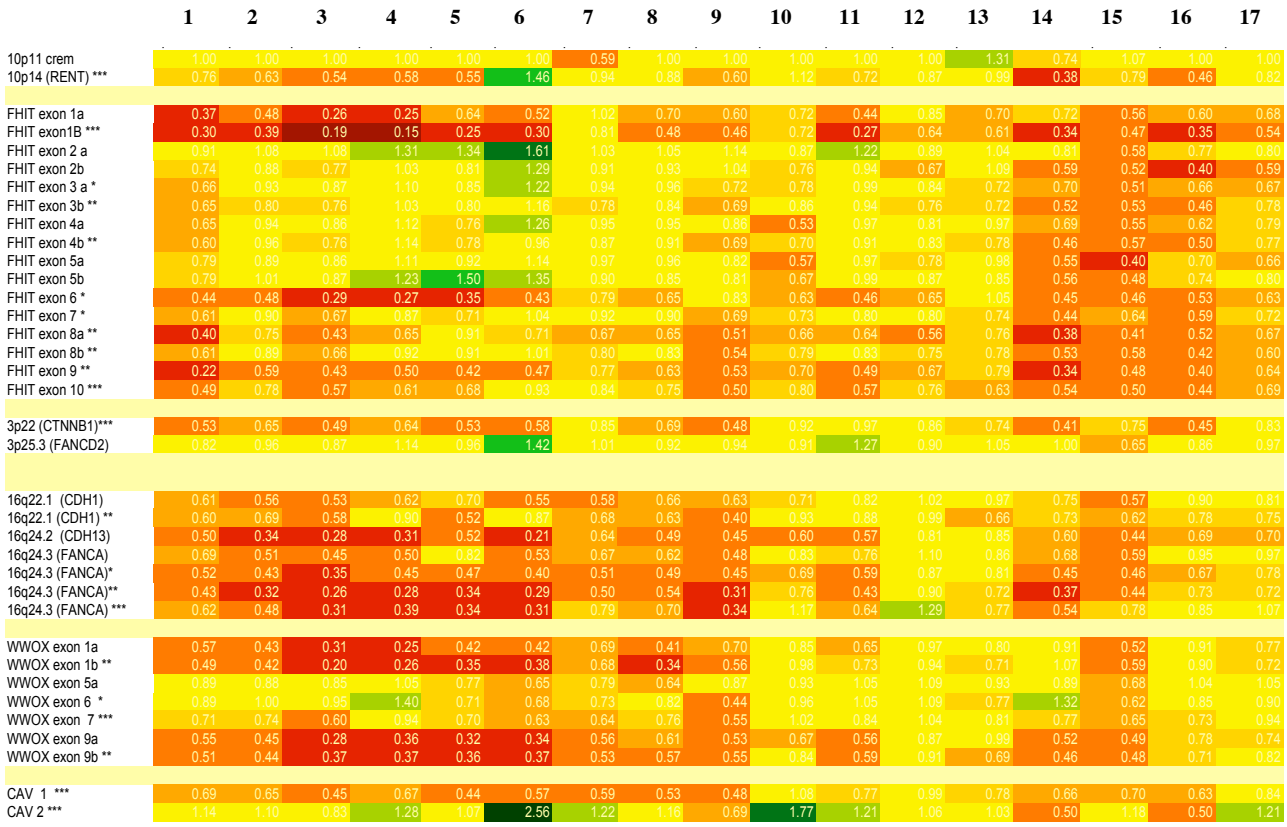


Figure 4.14: *Heat map representation of exons' copy number changes for genes at fragile sites in 17 specimens.*

Each exon probe is labeled accordingly; a and b are added where two different probes map within the same exon. The color coding corresponds to the value of normalized ratios "tumor /control", as discussed in section 4.2.4. Normal ratios are in yellow, deletions are in red and amplifications in green. The first two rows show control probes on chromosomes 10p. The probe for the gene CREM could be used in the majority of cases. Where an obvious aberrant copy number was seen for this probe the 3p25 probe was used as an alternative after assessment of the raw data. Both these control probes were usually consistent with each other. The second set of probes represented the FHIT locus on chromosome 3p, followed by two probes on 3p22 and 3p25 respectively. Beneath, are probes on chromosome 16q and those covering WWOX exons. The last set of probes represents the CAV1 and CAV2 genes respectively mapping within the FRA7G locus . On the left of the map, samples with a pattern of WWOX and chromosome 16 probe deletions have been clustered; on the right, samples with FHIT-predominant deletions are seen.

Heat maps corresponding to all samples analyzed can be seen in Appendix 4. Using the above criteria for counting deletions the compounded results could be summarized as follows: 62 out of 70 specimens had one or several exons that were deleted in the FHIT gene. When FHIT exon1 was excluded (see section 4.3.5) fifty-nine specimens had FHIT exon deletions. Forty-nine samples showed deletions in the WWOX gene, 39 in the CAV1 gene and 12 in the CAV2 gene. It should be noted that since only one probe was used for CAV1 and CAV2 respectively, results obtained on these 2 genes are bound to be less accurate.

Table 4.2 shows a summary of results for FFPE samples from the Eastern Cape cohort. Two criteria for the definitions of deletions are shown, first a ratio <0.7 and second a ratio <0.6. This allows checking the incidence of FHIT and WWOX deletions at higher levels of confidence. Unless specified otherwise in the text, results on deletions are always expressed for a ratio<0.7.

Table 4.2: *Number of FFPE cases showing deletions in the respective fragile sites-associated genes, FHIT, WWOX, CAV1 and CAV2.*

Gene	FHIT	WWOX	FHIT and WWOX	CAV1	CAV2
Deletion criteria: ratio < 0.7	59 (84%)	49 (70%)	48 (69%)	39 (56%)	12 (17%)
Deletion criteria: ratio < 0.6	50 (71.5%)	44 (63%)	43 (61%)	28 (40 %)	4 (6%)

Of note WWOX deletions were almost always accompanied by deletions in the FANCA gene on chromosome 16q , When WWOX was deleted in 44 samples, FANCA was deleted in 41 of these samples (under the deletion criteria set at a ratio <0.6).

4.4.4.5 Considerations regarding results obtained with FFPE samples

There were a high number of heterozygous intragenic deletions in one or several genes located at fragile sites. Surprisingly, after excluding FHIT exon1, 60 samples (86%) (59 with FHIT deletions, 48 with both FHIT and WWOX deletions and 1 sample with only WWOX deletions) had deletions in one or several exons of the FHIT and/or WWOX genes. No homozygous deletions were observed in either gene, possibly due to the admixture of few normal cells in those samples showing very low probe ratio.

The heterozygous deletions observed could not be attributed to the paraffin-embedded nature of the specimens, which was excluded by the assessed similarity of MLPA amplification profiles in FFPE and fresh control specimens (Section 4.3.3). The potential effect of differing DNA extraction procedures or the presence of impurities on the MLPA results was also excluded (Section 4.3 3). To further account for the possibility that some FFPE specimens might have had advanced DNA degradation, the results were also expressed using a higher level of stringency for the definition of deletion (ratio< 0.6), Table 4.2. Under these conditions, FHIT and /or WWOX still showed exon deletions in a high proportion 73% of samples (51 samples made of 50 with FHIT and /or WWOX deletions and 1 with WWOX deletion alone).

Discontiguous deletions

Deletions often appeared as non contiguous in both the FHIT and WWOX loci. Discontinuous deletions at fragile sites have been described in tumor cells (Corbin, et al. 2002) and would be expected. Indeed, *in vitro* model of oncogene induced fragile sites copy number alterations has shown that deletions at these sites can be genetically heterogeneous in early lesions (Tsantoulis, et al. 2007). In addition, it is likely that the presence of non-allelic deletions as well as a degree of cancer cells heterogeneity would

have added to the discontinuous appearance and complexity of deletion patterns observed at these two highly unstable regions of the genome.

Duplications/Amplifications

Two kind of duplication/amplification could be observed: intragenic duplications and duplications covering a large region of a chromosome.

Intragenic duplications

Intragenic duplications could be observed in samples 6 (Figure 4.14), as well as in samples S912, S972 and S1635 (Appendix6). Although the possibility of a biased PCR amplification cannot be excluded with certainty in these 4 samples, the duplication of experiments would weigh against it. Fragile sites are known to play a role in mechanisms of amplification (as discussed in section 3.1.5). Most of the reported events related to amplification at fragile sites have been characterized because they resulted in oncogene amplification (Hellman, et al. 2002). However, very recently, intragenic duplication neighboring deletions have been observed within FRA3B in 'in vitro' model of fragile site induction (Durkin, et al.2008), This *in vitro* model shows that duplication events can involve small portions of the fragile region. In a similar pattern, CAV2 was also found to be amplified/duplicated in 3 samples (samples 4 and 6 in Figure 4.14 and sample S1051 in Appendix 4), yet these same samples also had CAV1 deletions and both genes locate in FRA7G.

Duplications covering probes over a large portion of a chromosome

Three samples showed duplications of probes spanning chromosome 3p (S1984 in Appendix 4) chromosomes 3p and 16q (S1708) and chromosomes 16q and 7q (S220) . These are likely to represent chromosomal imbalances present in cancer cells in the form of partial or total trisomies.

FHIT deletions

The finding that FHIT was aberrant in 62 cases warranted a close examination; dissecting deletions in this gene, it was found that exon 1 was deleted in 57 cases. This exon was covered by two probes and both probes showed a small size probe peak on normal samples (see figures 3.4 in chapter 3 and 4.4 in this chapter). Despite the confidence interval verified by the mean controls probes' peak height and standard deviations, it is possible that small amplification variations for these peaks translate into a disproportional ratio variation and false deletions. These would have to be verified by other techniques, such as high resolution CGH arrays. FHIT exon 1 deletions, however, are probably functionally irrelevant, since the first coding exon is in FHIT exon 5. On the 3' part of the FHIT gene, probes for exons 6, 7, 9 and 10 were each covered by a single probe. Should a polymorphic mutation affect the ligating site of the two MLPA hemiprobes, this would translate into a lack of PCR amplification for this allele and could simulate a false deletion. While sequencing of the respective exons would inform on this possibility, this was beyond the capacity of this project. It is unlikely, however, that mutations in the ligation site would account for a high percentage of deletions, since the probes were designed in conserved regions with no known polymorphisms. In occasional instances, two different probes covering the same exon showed a discrepancy for their respective

ratios, which may reflect the presence of a polymorphic mutation; in these circumstances, the exon was excluded as being non-interpretable and not included in the final count.

FHIT exons 2, 3, 4 and 5 were each covered by a set of two probes widely spread over the probe size range, and which served to confirm each other. The two probes had sizes in both the large and small scales, which excludes the possibility of biased ratio, due to the fading away of the larger probes peaks (see section 4. 4, for discussion regarding this effect of larger probes). Deletions were also clearly seen on raw data (Figure 4.8).

WWOX deletions

WWOX deletions were seen in 70% of the samples, mostly targeting exons 9 and 1. WWOX exon 1 and 9 were each covered by a set of two probes each, with sizes in both the large and small scales (382 bp, 166bp for the larger and the smallest respectively); deletions in WWOX could clearly be seen in both probes sets (Figure 4.7).

Originally, the control probes were designed in the FANCA gene to assess the chromosome 16 locus specificity of deletions in the WWOX gene. FANCA is not known to be a target for deletions in sporadic cancers. Four probes covering this gene were, however, frequently deleted (at least 60% of cases), the significance of which is discussed below.

In summary, deletions that affect critical exons of genes at fragile sites were highly prevalent in ESCC specimens collected in the Eastern Cape, possibly in at least 73% of cases (this is when excluding isolated FHIT exon 1 deletions which may not affect transcription and when using a stringent definition of deletion for a ratio < 0.6).

4.5 Discussion

The aim of this study was to evaluate the extent of exon copy number changes in two genes located at specific fragile sites in ESCC specimens from South African patients, using a new MLPA-based screening method.

4.5.1 Specimens collection

Difficulties of differing natures were encountered at each centre. In Chris Hani Baragwanath Hospital, the biopsies were performed by gastroenterologists and consisted of small punch biopsies. The amount of material was limited and access to paraffin-embedded specimens left over from the pathologist's diagnostic need was not possible. Only fresh specimens could be used for this cohort. The histological assessment of these tissues depends upon both the site and the depth of the biopsy. Several biopsies are necessary for diagnosis purposes and one may differ from another. It follows that prospective specimens could not be accurately pre-assessed at a macroscopic level, which reduced the number of samples having at least 70% tumor cell representation, which was the criterion to qualify for this study. There were no resection specimens available, due to the fact that brachytherapy (radiotherapy) was the choice treatment, less invasive and probably the best option for ESCC patients in SA. Hence only 15 specimens were suitable for this study.

In Umtata General Hospital, the collection of fresh biopsies was difficult to organize and we could only reliably use paraffin-embedded specimens left over from diagnostic blocks. The paraffin tissue specimens were, however, of a larger size than those collected in

Gauteng and were performed by surgeons with the patient under heavy sedation. These specimens could easily be microdissected.

Although it would have interesting to isolate precancerous specimens this was not possible due to the fact that it is difficult to distinguish in situ carcinoma (CIS) from severe dysplasia. (Dawsey, et al. 1994).

4.5.2 Strength and weakness of the FHIT/WWOX MLPA assay

MLPA limitations

MLPA is a very robust assay in terms of the reproducibility of the probe amplification patterns and a number of deletions can clearly be seen on visual examination of raw data. The technique has been used widely for the screening of constitutional genetic aberrations. There have been, however, considerably fewer reports of MLPA applications in cancer compared to constitutional studies. This may be due to the following two limitations. The first practical concern is the choice of internal control probes for the normalization of ratios. It is critical and should ideally be validated by another method. FISH was chosen here within the limits of original specimens availability. The control gene RENT, on 10p is not in a fragile region (Appendix 1) and the FISH probe controlled for the average copy number of this chromosomal region. Alternatively several control probes could be used but once again the choice of probes located in genomic regions that are rarely involved in cancer is critical (see Chapter 6, future work).

The second limitation is due to the fact that large probes tend to fade away and have lower amplification peak heights in proportion to their size. If the control sample shows the same trend as the test samples, this doesn't interfere with the results that are expressed as ratios. Fading away can, however, be more pronounced in some abnormal samples. In the present

assay, critical exons were covered with two differing probes with different sizes correcting for this effect.

Some groups (MRC-Holland) have used log2 algorithm transformation of ratios values to first smooth inter-probe variations and assimilate all peaks to a value of 1. They then apply a regression coefficient calculated on the fading trend of all the internal control probes. In effect, this does not take into account the difference of regression slope between the control sample and the test samples, which is the only relevant factor. In addition, smoothing of probe ratios may introduce a larger distortion (control/test) for probes with a low peak height because they are corrected by a factor of 80% or more.

MLPA strengths

The MLPA assay allows the assessment of several genes at fragile sites in one reaction. The comparison of MLPA relative changes in copy number to CGH array results was strikingly consistent in the UWOV2 cell line (see Chapter 5). For this reason, MLPA strategies might be an excellent post-CGH array tool to screen for copy number-identified signatures in cancer samples. Preliminary results have been obtained on two FFPE specimens from the Eastern Cape that were also run on MLPA. Both samples were hybridized to a 244k oligonucleotide array (Agilent), in competition to control DNA extracted from normal FFPE samples. There was a good concordance between MLPA and CGH array for the WWOX probe in exon 9 (ratios of 0.66 and 0.67 in one specimen, 0.68 and 0.8 in the other specimen by MLPA and array CGH respectively).

4.5.3 Deletions of FHIT and WWOX in each respective cohort

After histological assessment, valid samples from Gauteng were limited in number, (15) but showed clear FHIT deletions in 27% of cases (4 out of 15). No WWOX deletions were detected.

By contrast in the Eastern Cape cohort, at least 73% of cases had intragenic deletions in FHIT and/or WWOX. Forty-eight cases (68%) had deletions in one or several exons of both genes.

The frequencies of deletions in the FHIT locus were consistent with reports of FHIT loss of expression, as evaluated by IHC and LOH in ESCC that were conducted for the most in Eastern countries such as Japan (Kitamura, et al. 2001; Morita, et al. 2006) and China (Liu, et al. 2004). They were also concordant in percentage with FHIT protein reduction, seen in association with promoter hypermethylation (Tzao, et al. 2006).

The only study that examined WWOX in ESCC was also conducted in Japan and showed LOH in nearly 40% of tumors (Kuroki, et al. 2002). As explained earlier, it is possible that LOH may underestimate the number of aberrations at fragile sites due to the possibility of partial or total heterodisomy.

Few studies have examined the occurrence of concordant aberrations in these two genes. Two reports by the same group show the concordant loss of FHIT and WWOX protein expression, as evaluated by IHC, in invasive breast carcinoma (Guler, et al. 2004; Guler, et al. 2005). The concordant loss at fragile sites is likely to point to a common DNA damage cause and reinforces the idea that external carcinogens are contributing to the high incidence of ESCC in the Eastern Cape area.

The difference of WWOX rate of deletions between the two cohorts is intriguing. Statistical evaluation using a Fisher's exact test shows that there is overwhelming evidence that the proportion of WWOX alterations in the Eastern Cape cohort is

significantly different from that observed in Gauteng (p value: 0.0000003) and this despite the small number of specimens from Gauteng.

Differences may reflect differences in toxic exposures. Although not confined to the Eastern Cape it is possible that the high consumption of tobacco and alcohol when combined to self-induced vomiting practices and exposure to fungus toxins, induces aberrations at fragile sites. However the difference in type of specimens (e.g. fresh versus FFPE) may be a confounder and this would need to be clarified by investigating FFPE specimens from the Gauteng area, although the high concordance of MLPA amplification patterns between fresh and FFPE samples weighs against this possibility.

4.5.4 Additional chromosome 16 deletions

Chromosome 16q showed a high rate of deletions overall; interestingly, loss of heterozygosity at the FRA16D in the 16q23 site has been reported in hepatocarcinoma that are associated to aflatoxin B1 (AF B1) exposure. Although the critical region deleted was found within WWOX intron 8, the WWOX gene itself was, not rearranged in 27 of the 157 cases that were specifically screened for WWOX mutations in this study (Yakicier, et al. 2001). AF B1 is present in high prevalence in SA and known to cause hepatocellular carcinoma, there are however no study that link AF B1 to ESCC in SA. On the other hand, fumonisin has been epidemiologically associated to ESCC in SA and it would be important to clarify whether the reported clastogenic effect of fumonisin preferentially targets fragile sites. The only studies regarding chromosomal breaks induced by fumonisin exposure in vitro unfortunately do not give any information related to the genomic location of the breaks, but mention an increased sister chromatid exchange rate, which is suggestive (Lerda, et al. 2005; Aranda, et al. 2000; Knasmuller, et al. 1997).

Probes spanning the FANCA gene were deleted in 42 of the 70 specimens from the Eastern Cape. FANCA gene germ line mutations are associated with one of the complementation groups of the congenital chromosomal instability syndrome, Fanconi anemia (Morgan, et al. 1999). This gene is not known to be frequently deleted in cancer; at the same time, acquired aberrations have rarely been assessed in sporadic tumors and some recent reports may signal a higher involvement than originally thought. Indeed, FANCA protein dysfunctions have been observed in sporadic cases of acute myeloid leukemia (AML) (Lensch, et al. 2003; Xie, et al. 2000), and heterozygous deletions were also detected in few sporadic AML cases (Lensch, et al. 2003; Tischkowitz, et al. 2004).

In squamous carcinoma, FANCA downregulation was reported in head and neck squamous cell carcinoma (Wreesmann, et al. 2007), and FANCA was underexpressed in young patients with sporadic oral squamous cell carcinoma (Tremblay, et al. 2006). It is possible that deletions in this gene have been underestimated until now.

Interestingly, tissue restricted hypersensitivity to cross-linking agents and chromosomal instability in ovarian epithelium have been linked to FANCD2 under expression in ovarian cancer patients and individuals at high risk of ovarian cancer (Pejovic, et al. 2006).

Like FANCD2, which is involved in the fragile site repair pathway (see Chapter 3, Section 3.1, Figure 3.2), FANCA is also strongly thought to be involved in the homologous recombination repair process, as shown by deficient knock-out mice models, and possibly via Fanca protein product interaction with Brca1(Yang, et al. 2005c). Given this role, deletions in FANCA would either facilitate or worsen deletions at fragile sites and it would be of interest to clarify the sequence of events in the Eastern Cape cohort.

Of note, in the ovarian carcinoma cell line studied in Chapter 5, the CGH array showed that WWOX and FANCA loci were very specifically deleted along the long arm of

chromosome 16; there were no other large deletions on this chromosome. It would be interesting to clarify whether these two sites constitute a preferential target of deletions.

Fifty percent of the microRNA (miRNA) molecules characterized thus far have been mapped at fragile sites (Calin, et al. 2004). Since deletions and to a lesser extent amplifications/duplications, occur at high frequency in fragile sites regions it will be interesting to see if a particular functional family of miRNA emerged from tumor-associated fragile sites aberrations.

In conclusion, deletions at both FHIT and WWOX loci were highly represented in the Eastern Cape cohort. It will be important to determine if the difference between the two cohorts observed in this study is confirmed, as it may indicate the presence of specific etiological factors in the Eastern Cape. It will be important to assess the effect of exogenous agents and carcinogens that are highly suspected to have a role in the Eastern Cape cohort in cell culture models. In addition, it will be valuable to screen for the presence of germ line mutations that could be associated with a population group "weakness" to repair genes at fragile sites.

CHAPTER 5

CHARACTERIZATION OF A TRANSLOCATION INVOLVING THE FRA3B IN AN OVARIAN CARCINOMA CELL LINE

5.1 Introduction

Ovarian cancer is the leading cause of fatality among gynecological malignancies (Fields and Chevlen 2006). Few cases are related to hereditary BRCA1 and BRCA2 predisposing mutations, but for most tumors, the molecular events driving the pathogenesis are not clearly elucidated (Goldberg and Borgen 2006).

The ovarian carcinoma cell line UWOV2 was established in the Johannesburg Department of Medicine from a papillary serous cystadenocarcinoma of the ovary (Golombick et al, 1990). A translocation derivative 11, der(11)t(3;11)(p14.2; p15) was one of the first, most stable and consistent abnormalities present in this cell line. This suggested that the translocation was likely to rearrange a critical gene, which would confer a cell growth advantage and play a role in the initiation or progression of malignancy. The translocation t(3;11) was of particular interest due to the non-random involvement of both chromosomes 3 and 11 in ovarian tumorigenesis (Borresen 1992; Huang, et al. 2002; Liu and Ganesan 2002). On the chromosome 11 short arm, the Wilms tumor gene(WT) locus maps at 11p13 (Call, et al. 1992). The WT gene is involved in gonad development (Logan, et al. 2003) and has previously been linked to serous ovarian carcinoma (Acs, et al. 2004; Hylander, et al. 2006).

On chromosome 3, several tumor suppressor functional regions have been characterized by research groups using two differing approaches. First by tumor-associated deletion

studies (Liu and Ganesan 2002; Manderson, et al. 2002), and secondly by *in vitro* experimental models, where parts of chromosome 3p genetic material were reintroduced into cancer cells. Suppression of tumorigenesis was achieved using the latter technology by independent groups (Cody, et al. 2007; Kost-Alimova and Imreh 2007). The FRA3B region was part of the critical regions identified that showed tumor suppressor activity in both strategies. On the other hand, the functional role of the FRA3B/FHIT in ovarian cancer has been a subject of controversy in the literature and the number of cases involving rearrangements at this locus remains small (Lounis, et al. 1998; Manderson, et al. 2002; Manning, et al. 1999; Ozaki, et al. 2001). Recent immuno-histochemistry (IHC) studies have reported the concomitant WWOX and FHIT proteins loss in breast tumors in association with a poorer clinical outcome (Guler, et al. 2005); this has not been investigated in ovarian tumors.

Although the FRA3B is a frequently deleted region in a variety of cancers, reports of translocation involving FHIT have only rarely been described. Only four tumor-associated FHIT translocations have been reported to date in the literature (Fang, et al. 2001; Gemmill, et al. 1998; Geurts, et al. 1997; Popovici, et al. 2002), of which two were present within the same cell line established from an esophageal adenocarcinoma (Fang, et al. 2001). The two other reports involve a breast cancer cell line (Popovici, et al. 2002) and a pleomorphic adenoma of the thyroid gland (Geurts, et al. 1997), respectively. A fifth translocation involving the FHIT locus had also been reported in one family as a constitutional translocation associated with hereditary renal cell cancer (Gemmill, et al. 1998). A chimeric transcript FHIT-TRC8 resulted from the latter and the thyroid-associated translocation fused FHIT to HMG A1. The other described cancer associated translocations either had non-coding chimeric transcripts or no transcripts at all,

suggesting that FHIT inactivation was the functional end point (Fang, et al. 2001; Popovici, et al. 2002).

Preamble : Previous work performed

The translocation t(3;11) present in the UWOV2 ovarian carcinoma cell-line was initially mapped by FISH techniques using yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs) on the short arm of each chromosome 3 and 11 respectively. The mapping of each derivative constituted a MSc project under my supervision, (Laubscher 2000; Rechviasvhili 2002). This previous work showed that the UWOV2 cells were nearly tetraploid. The composite karyotype was interpreted together with additional FISH data using YACs and chromosome-specific libraries, (Appendix 5). A chromosome complement of the UWOV2 cell line is shown in Figure 5.1. The two chromosome 11derivatives constituted the focus of this work and were described as: der(11)t(3;11)(p14.2;p14.3-15.1). Chromosome 3 homologs were rearranged with the two most consistent structural aberrations represented by deletion of chromosome 3 short arm in two copies (Figure 5.1, arrow heads), and a translocation with chromosome 8 long arm, t(3;8)(p13;q24) which did not involve chromosome 11 material; it was not possible to ascertain whether any of these structural aberrations represented reciprocal derivatives of the t(3;11) translocation at a cytogenetic level. Chromosome 11 homologs on the other hand, were only involved in a single structural rearrangement in the translocation t(3;11)(p13;q25).

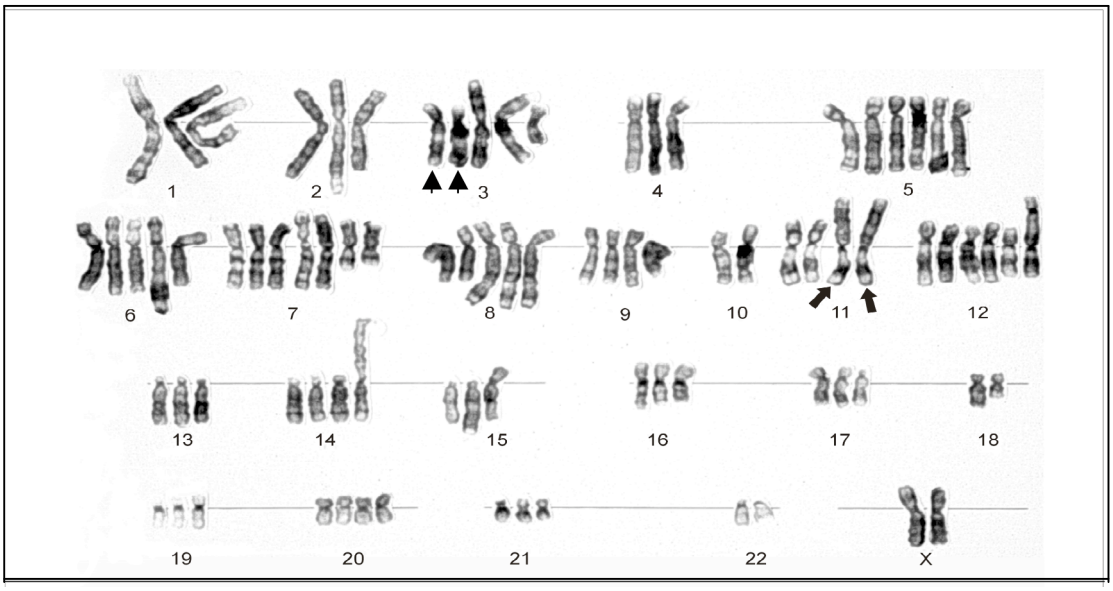


Figure 5.1: *G banded karyotype of UWOV2 cell line at passage 50. Arrows point to the derivative 11, der(11)t(3;11)(p14.2; p 15.1). Short arrows show del(3)(p13).*
(From A Lauscher, 2000).

The derivative 11 t(3;11) developed additional rearrangement (with added genetic material on the telomeric end of chromosome 3 sequences) during subsequent passages of the cell line; the chromosomes 3p14 and 11p15 breakpoints however, remained remarkably consistent, were present in every cell, from the first established cultures to all subsequent passages thus reflecting an early carcinogenic events with a likely growth promoting effect. The breakpoint on 11p was mapped to 11p15.1, centromeric to BAC 166e15 (BACPAC Resources Centre) and to the LDHC gene locus, and telomeric to BAC 13o19. No BAC nor YAC probes spanning the breakpoint could be identified (Laubscher 2000) (Figure 5.2). At the time of this initial study, the human genome mapping was incomplete and there was no BACs tiling path covering this region of chromosome 11.

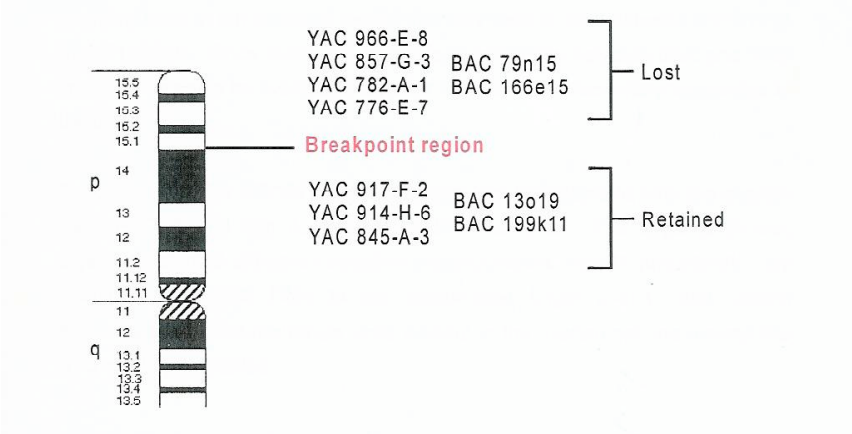


Figure 5.2: *Summary of BACs and YACs mapping studies on the derivative (11) t(3;11). Sequences telomeric to YAC 917F2 were lost; BAC 13o19 was retained.*
(from A Laubscher, 2000)

On the other partner, chromosome 3p, although no translocation spanning vector was identified, the breakpoint could be mapped with more precision and found to be flanked by YAC 850, retained on the derivative (11)t(3;11) while YAC 940H8 was lost (Figure 5.3). The YAC 850 was obtained from the Centre d'Etudes du Polymorphisme Humain (CEPH), Paris, and is known to span the FRA3B/FHIT locus, covering exons 1 to 5 of the gene.

Both FISH mapping studies, therefore showed that sequences 3p14.1 were juxtaposed to sequences 11p 15.1, and the translocation breakpoint could be corrected to: t (3;11)(q12 -> **p14.1:: p15.1**-> qter). These results also suggested that the breakpoint lay within the FRA3B locus within ±1Mb on chromosome 3p14.1, and probably between FHIT exon 7 and FHIT intron 4.

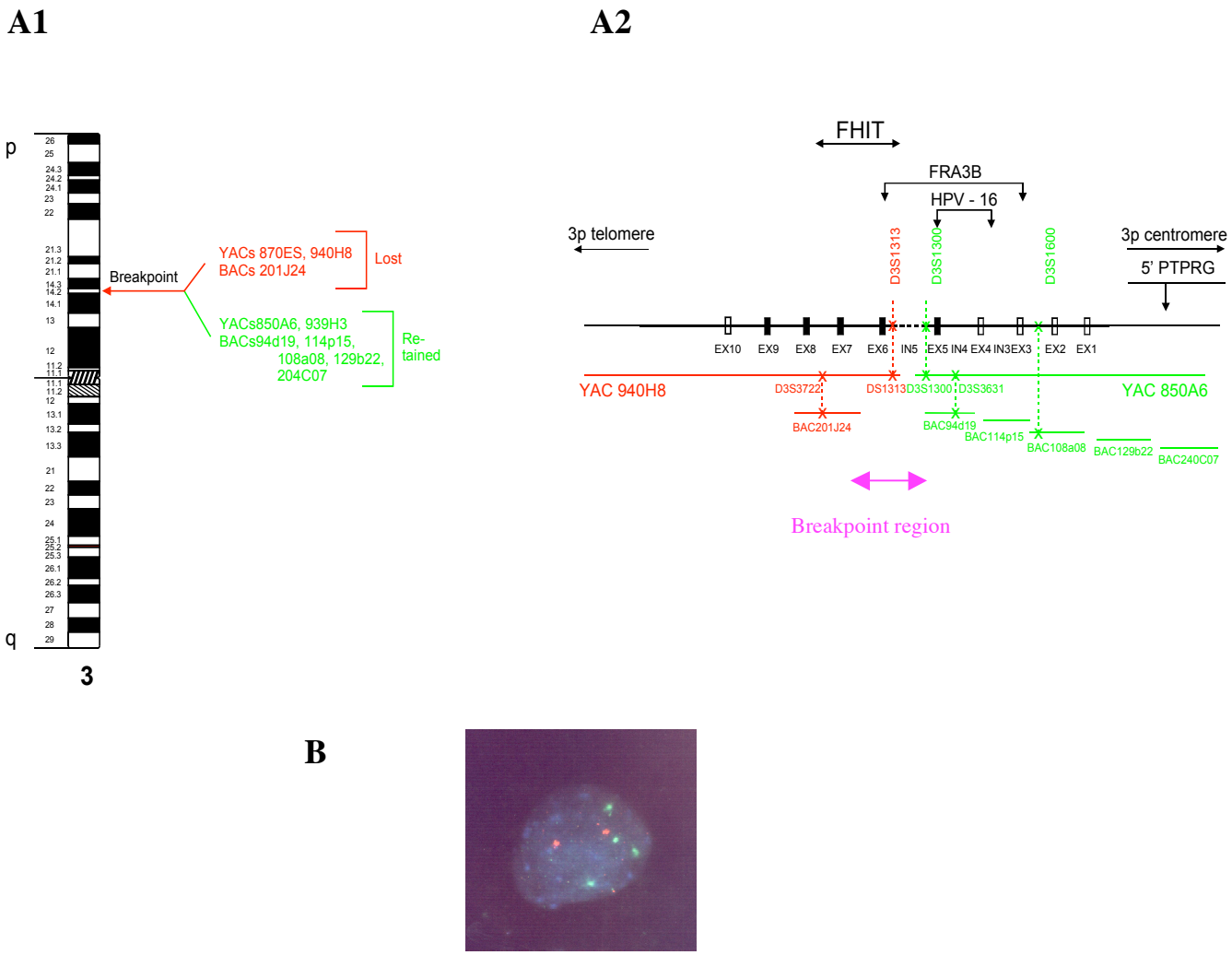


Figure 5.3: *FISH mapping data on der(11)t(3;11) in cell line UWOV2. Modified from Rechviasvhili, 2002.*
A1: chromosome 3 ideogram showing the respective BACs and YACs probes lost/or retained on the der(11)t(3;11). A2: corresponding map of the FRA3B locus, indicating the region of the breakpoint in purple.
B: Interphase cell showing the split between YAC 850A6 (red signals) retained on the der(11) and YAC 940H8 (green signal) lost from der(11). Of note, YAC 850A6 was present in two copies in all cells, while YAC 940H8 was present in three to four copies in 80% of cells.

5.2 Aims

The aims of this work were to:

Delineate further the translocation breakpoints on the derivative 11

Assess if the translocation interrupted the FHIT gene and evaluate the expression of FHIT in the cell line and

Characterize the translocation partner sequences on chromosome 11 p15.1.

5.3 Materials and Methods

5.3.1 Cell cultures

UWOV2 cells were grown in RPMI 1640 medium (with GlutaMax-1, 2mM HEPES, Gibco BRL), supplemented with 10% fetal bovine serum (FBS) (Gibco BRL) in a humidified atmosphere at 37°C.

5.3.2 DNA extraction

After trypsinisation and collection of the cell pellet by centrifugation, DNA was extracted by proteinase K (Roche) digestion followed by phenol:chloroform extraction, as in Chapter 3 (Section 3.2.2). The DNA yield and sample protein purity was determined using a NanoDrop spectrophotometer at an optical density (OD) of 260nm and 280nm. The integrity of the DNA was further assessed on a 2% agarose gel.

5.3.3 MLPA

MLPA was performed as described in Chapter 3 (Section 3.2.5), with a starting amount of 200 ng of high molecular weight DNA extracted from the UWOV2 cells in cultures. The reactions were performed in duplicate and three controls were included in each run. The raw data were first examined and compared to RT-PCR and CGH array results.
Normalization of the tumor to control ratios was done on the 16q probe for CDH13.

5.3.4 RNA extraction

Total RNA was extracted from freshly harvested cells using the Qiagen RNeasy midi kit, as per manufacturer's instructions. Briefly, UWOV2 cells were lysed in 2ml RLT buffer (Qiagen) containing 20μl of β-mercaptoethanol (1% final concentration). The RNA was then extracted using the RNeasy midi kit (Qiagen), according to the manufacturer protocol, and eluted into 150μl of RNase-free water. The yield and RNA quality were assessed by spectrophotometry and denaturing agarose gel electrophoresis respectively. Control (c1) RNA was extracted in the same manner from fresh ovarian cells that had been DMSO frozen (the cell line was obtained as part of a collaborative project on establishing ovarian cell cultures with the University of the Witwatersrand Department of Obstetrics and Gynecology). Control RNA (c2) was extracted from the peripheral blood cells of a healthy volunteer.

5.3.5 RT-PCR for FHIT transcripts

Reverse transcription and PCR reactions were performed in one reaction using the OneStep RT-PCR kit (Qiagen), according to the manufacturer's protocol and starting with an amount of 500ng to 1μg of total RNA. Both forward and reverse primers, 0.5mM of

each, (published primers by (Fang, et al. 2001) and detailed in Appendix 3) were included in the reaction. The reactions were performed on an Eppendorf Mastercycler. All reactions started with a RT reaction at 50°C for 30 minutes, followed by a hot step (95°C for 15 minutes). The full FHIT transcript was amplified with the following PCR steps: 95°C for 1 minute, 62.5°C for 90 seconds, and 68°C for 3 minutes for 35 cycles, followed by a final extension of 20 minutes at 68°C. A prolonged extension time was found to improve the yield of the FHIT full transcript with consistency.

PCR conditions for individual FHIT exons one , five and ten were 95°C for 1minute, 60°C for 1 minute, and 72°C for 1 minute, repeated for 35 cycles, with a final extension of 7 minutes at 72°C. The exon specific primers were previously published (Fang, et al. 2001) (Appendix 3.)

5.3.6 3' Rapid amplification of cDNA ends (RACE PCR)

Six μg of total cell line RNA was treated with 1μl DNAse I (1 unit/μl) (Invitrogen) in a total volume of 20μl for 15 minutes at room temperature to eliminate residual genomic DNA. DNAse I was then heat inactivated at 65°C for 15 minutes with EDTA (2.5mM final concentration) and the sample was placed on ice.

RACE PCR was done using the First Choice RLM-RACE kit (Ambion), according to the manufacturer's protocol. In brief, the first strand cDNA synthesis was performed with M-MLV reverse transcriptase at 42°C for 1 hour in the presence of the 3'RACE adapter included in the kit (enriched in polyT). First round PCR was performed with the 3' RACE outer primer and the FHIT exon 1, F1 primer (Appendix 3), using 0.5μl Platinum Taq (2.5 units/μl Invitrogen). The hot-start PCR conditions were 94°C for 3 minutes; 35 cycles of 94°C 30 seconds, 60°C 30 seconds, 72 °C 30 seconds, 72°C 10 minutes. When a second round PCR was performed to enrich for potential specific transcripts, this was done with

1µl of the first round product using the 3' RACE Inner primer and the nested FHIT primer F2 (Appendix 3). Products were run on a 2% low melting agarose gel containing 1µl/ml EtBr (10mg/ml) and visualized on a UV transilluminator. Bands were cut out from the gel and sent for sequencing (Inquaba Biotech).

5.3.7 CGH array

The Agilent Protocol "oligonucleotide array-based CGH for genomic analysis" was followed as in Chapter 2, Section 2.3.6. In brief, 3µg of cell line and control high molecular weight DNA were separately co-digested with Alu1 and Rsa1. The control DNA was a commercially available pooled Human Genomic DNA (Promega). Both control and tumor-digested DNA were then separately random prime labeled using the Exo-Klenow fragment (InVitrogen) in the presence of cyanine 3-dUTP (control) and cyanine 5-dUTP (tumor), respectively (both cyanines dUTP were from Ammersham). The reaction products were cleaned up on Microcon YM- 30 filters (Millipore) and the labeling efficiency was evaluated on a NanoDrop. The labelled samples were pre-annealed with 50µg of Cot1 DNA and hybridized to a 44K oligonucleotide Agilent CGH array for 40 hours at 65°C. The slides were then manually washed following the Agilent protocol, procedure B, which includes the stabilization and drying solution. The slides were scanned immediately on an Agilent scanner (Institut de Recherches sur le Cancer de Lille, Lille, France).

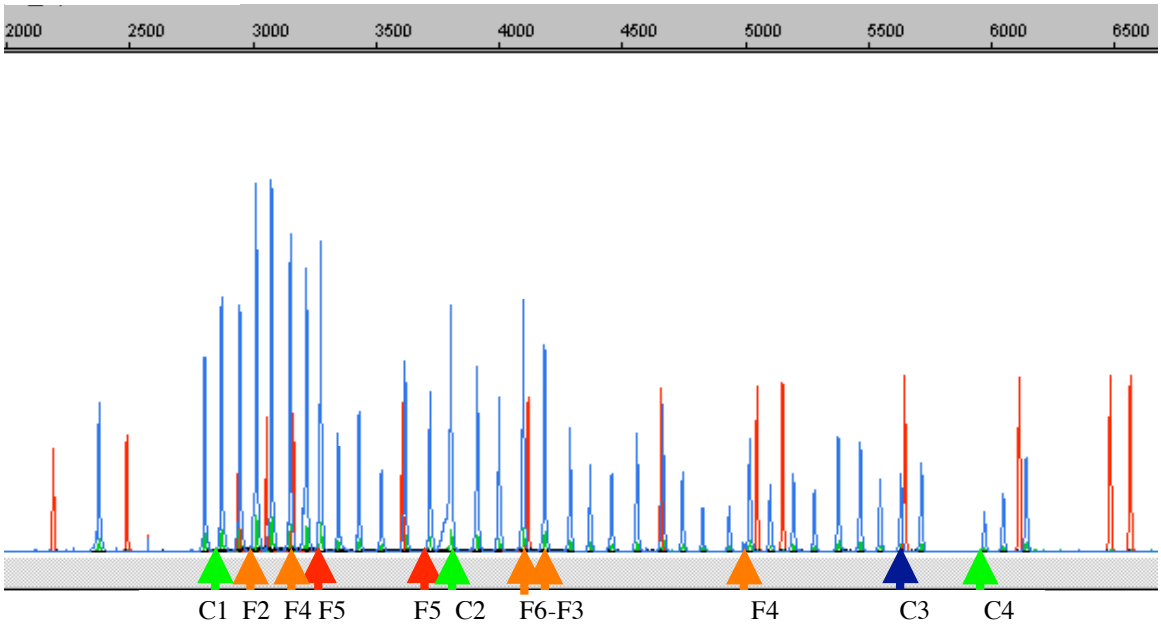
5.4 Results

5.4.1 MLPA

Results for the MLPA products of amplification obtained from the UWOV2 cell line were interpreted as in chapter 3, section 3.2.6. In brief, each probe fluorescent peak height was evaluated as a ratio against the same probe peak height observed in the reference control specimen, obtained from the 3 controls that were run in the same experiment. This ratio was then normalized to control probes whose amplification did not depart from that observed in the control specimen. For the UWOV2 cells, the control probes for the genes CREM and CDH13 were used for normalization. The choice of these control probes was further validated by CGH array data, see section 5.3.6, and the overall results were represented in a heat map format, further described in section 5.3.6.

Before normalization, the simple visual examination of the GeneScan raw data of MLPA products already evidenced homozygous deletion of FHIT exon 5, while FHIT exons 1 to 4 had a relative signal intensity reduced by half when compared to the control sample (Figure 5.4).

CONTROL



UWOV2

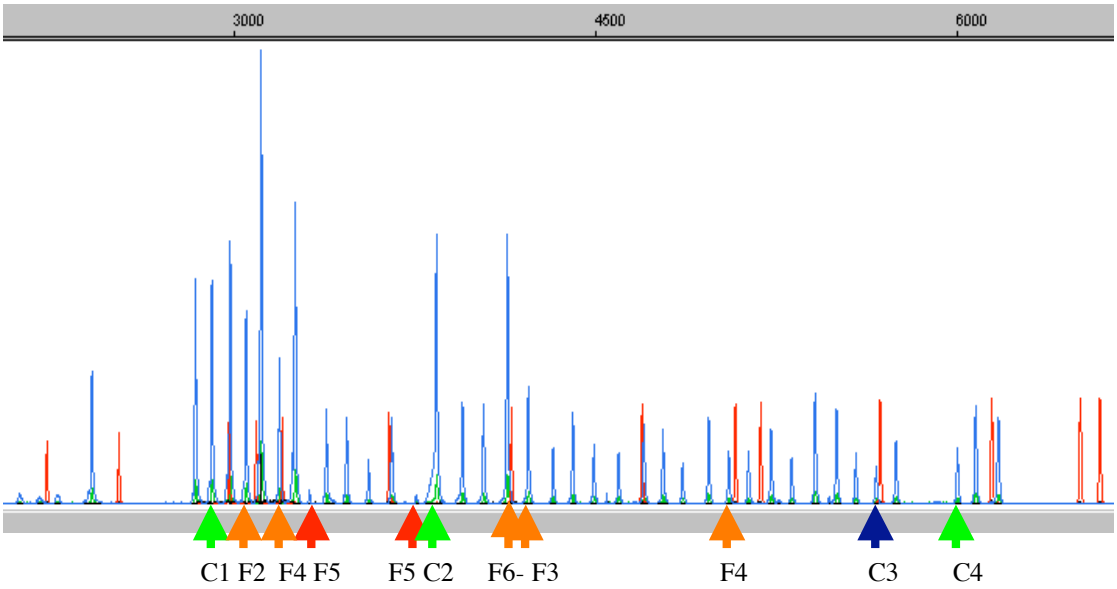


Figure 5.4: *MLPA results of cell line UWOV2 with the FHIT-WWOX probe mix. Raw data from GeneScan analysis software showing MLPA results for a control specimen run in the same experiment and the cell line UWOV 2 below. FHIT exons, highlighted here with orange arrows, are from left to right: FHIT exon 2: F2, FHIT exon 4: F4, FHIT exon 6: F6, FHIT exon 3: F3 and the second probe for FHIT exon 4. Homozygous deletions of FHIT exon 5: F5 were detected, indicated by red arrows. A hemi representation of FHIT exons 2, 3 and 4 was observed. Only the relevant probes are labeled for illustration purposes; (for the identity of all probes peaks please refer to Chapter 3, section 3.3.1). The green arrows show the internal control probes, C1: CREM*

on 10p1.2, C2: CDH13 on 16q16 q24.3, and C4: TSC2 on chromosome 16p13.3. The control probe C3: RENT on 10p, indicated with a blue arrow, also had a half representation copy number compared to the control and was not used for normalization. The C1 (CREM) and C2 probes (CDH13), which showed an amplification pattern similar to the normal control sample, was chosen for normalization and validated by CGH array.

These results suggested that the translocation t(3:11) therefore interrupted the FHIT gene and resulted in the loss of FHIT exon 5. FHIT exons 1 to 4 were present in half copy number. Taken together with previous FISH data showing that YAC 850A6 was retained on the two derivative (11)t(3;11), representing the 5' region of FHIT (exons 1 to 4) (Figure 5.3) and given the fact that neither exons 4 nor 6 were fully deleted, these results placed the translocation breakpoint in FHIT intron 4. FHIT exons 6 to 10 were deleted, but not significantly (ratios >0.8,0)(Figure 5.9, A). This was consistent with the presence of three copies of YAC 940H8 that covered the 3' part of FHIT (Figure 5.3, B).

5.4.2 RT-PCR

RT-PCR with primers in the first and tenth exon (Appendix 3) confirmed that there was no full FHIT transcript in the UWOV2 cell line (Figure 5.5, A). Exons 1 and 10 were, however, retained at both the DNA and RNA levels (Figure 5.5, B) and confirmed the MLPA results.

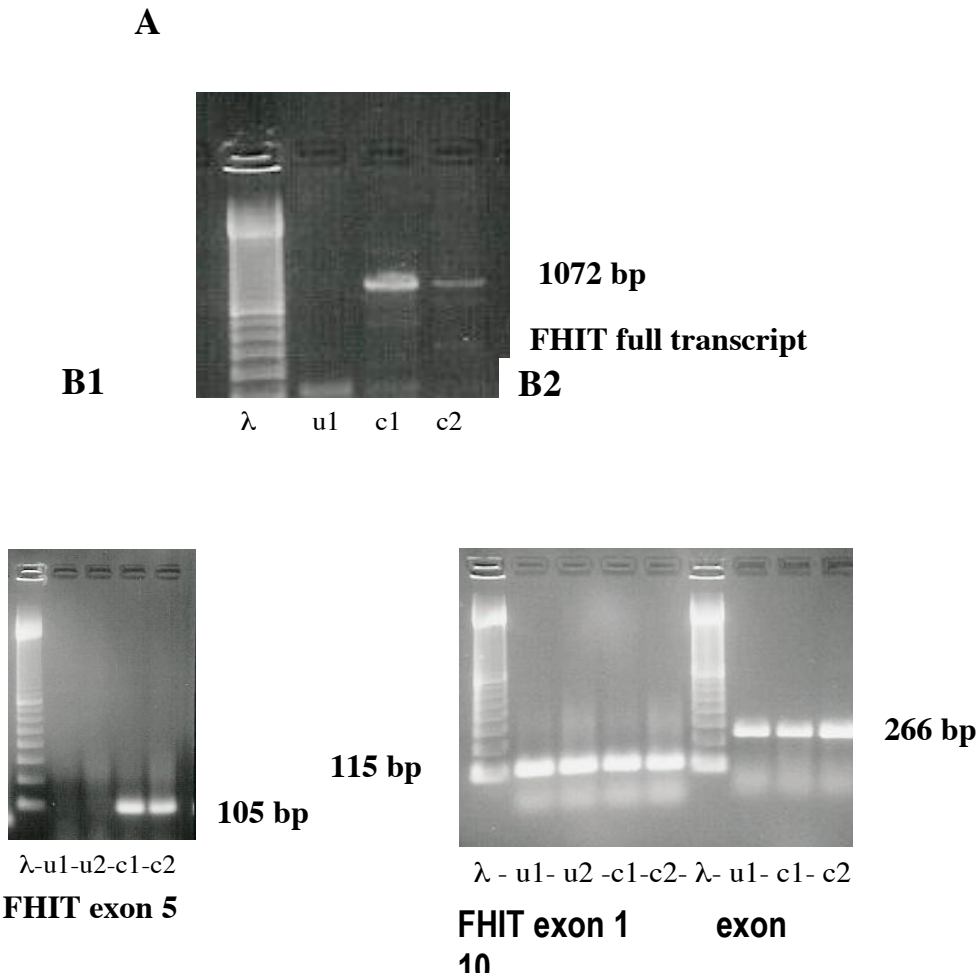


Figure 5.5: *2% agarose gels showing the FHIT-specific RT-PCR results obtained with the UWOV2 cell line. The λ marker used in all agarose gels was a 100bp marker (Promega).*
A: RT-PCR results on FHIT full transcript expression in the UWOV2 cell line from one culture (u1) as compared to controls (c1 and c2).
B: PCR results of the respective FHIT exon 5 (B1) and exon 1 and 10 gel (B2) in the UWOV2 cell line from two differing cultures (u1 and u2) as compared to controls (c1 and c2). FHIT exon 5, as well as the full FHIT transcript, was lost in the UWOV2 cell line.

5.4.3 3' RACE PCR

Several attempts at 3' RACE PCR failed to yield any chimeric transcripts, suggesting that the result of the translocation is the inactivation of the FHIT gene. Although FHIT exon 1

was present as established by RT-PCR, the RACE kit is designed based on the assumption that poly A sequences are present and it would not detect non-coding chimeric transcripts.

5.4.4 CGH arrays

Image analysis of the array raw data was performed using CHG Analytics software 3.4 (Agilent Technologies). Losses of chromosome 3 short arm, sequences p12.1 to p14.2 were detected in half the copy number (Figure 5.6, A). This is interpreted against a near tetraploid karytotype, evidenced by cytogenetics data, and is consistent with the presence of two copies of the derivative 11 t(3;11) which carry the sequences 3p12.1-3p14.2. This also confirmed the FISH results obtained with the YAC 850A6 probe (Figure 5.3, B)

The oligonucleotide probe that maps within FHIT exon 5 (as identified through the sequence accession number provided by the CGH analytics software oligonucleotide probe list) (Figure 5.6, B1, and Appendix 6) was completely deleted. This confirmed both the MLPA and PCR results.

The interpretation of CGH array data together with FISH results allowed for precise mapping of the translocation breakpoint on chromosome 11. The loss of sequences 11p, p15.1→pter (Figure 5.6, A) clearly resulted from the translocation and was in agreement with the FISH data, which showed the retention of the BAC 13o19 that spans the marker D11S2034 (Laubscher 2000).

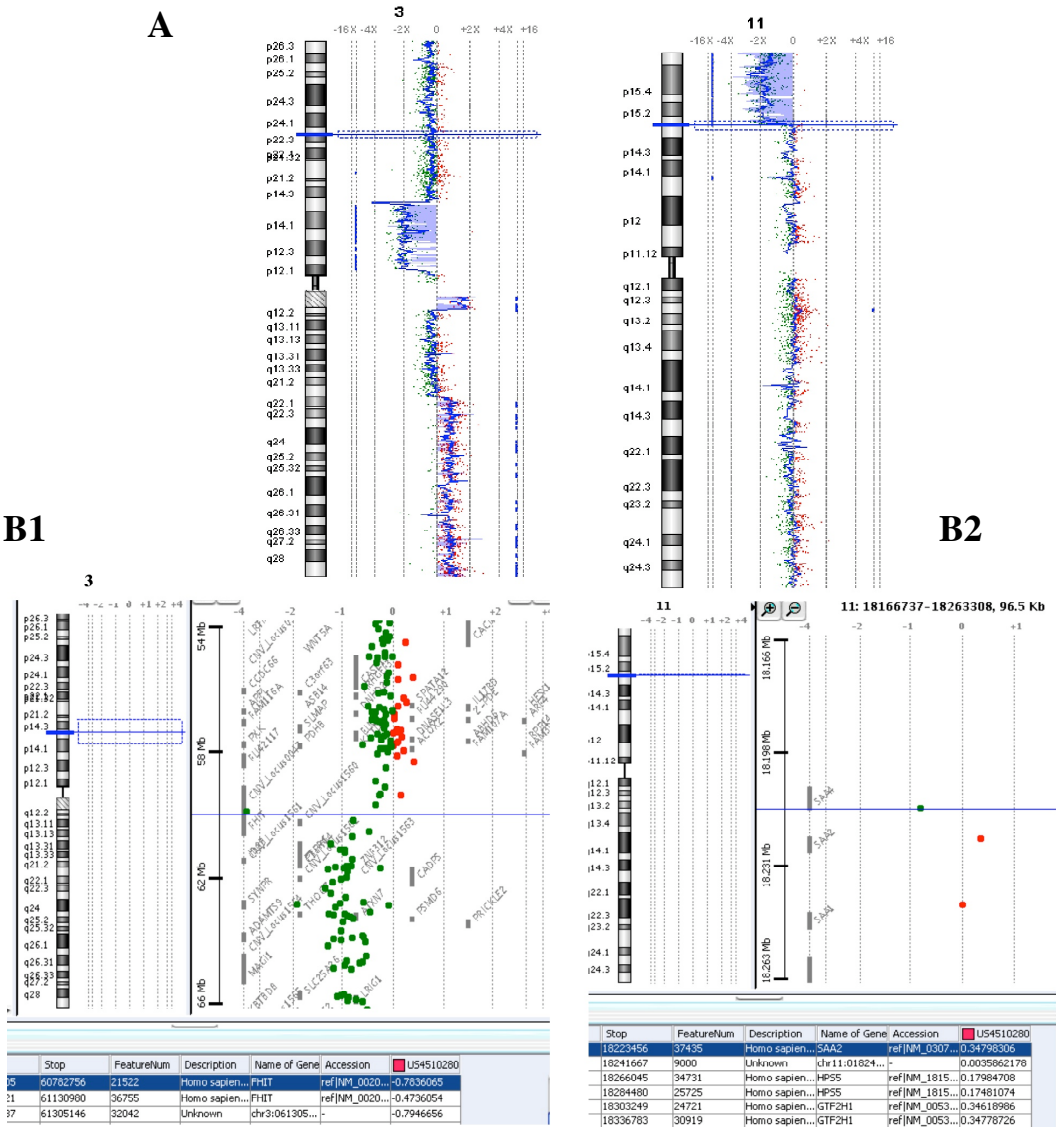


Figure 5.6: *CGH array results for the UWOV2 cell line.*
(A): chromosomes 3 and 11 ideograms; the mean average log2 ratio, blue line, reflects the copy number of probes mapping in the corresponding chromosomal regions. Deletions are shown on the left of the zero line and amplification on the right, respectively. CGH array like MLPA does not detect tetraploidy; copy number changes are therefore relative. (B1): detail of breakpoint mapping for chromosome 3p14 and (B2): for chromosome 11p15.1. The position of the horizontal blue line on the chromosome ideogram matches a green dot, which represents a specific oligonucleotide probe hybridization signal (right of the ideogram). The corresponding probe name and reference sequence accession number from the NCBI gene list is highlighted in blue at the bottom of the figure in the preloaded gene annotation list (CGH Analytics Software, Agilent). This allows for the verification of the exact gene location and composition of the selected probe. The FHIT probe closest to exon 5 was completely deleted. (B2): the breakpoint on 11 p15 can be inferred to lie within SAA2 and 4 genes.

Both chromosomes 3 and 11 short arms showed deletions of sequences: 3p11→p14.2 and 1p15.1→pter respectively. The boundaries of these deletions 3p14.2 and 11p15.1, corresponded to the translocation breakpoint, as seen by the combined interpretation of cytogenetics and FISH mapping results.

On the chromosome 11 ideogram (Figure 5.6, B2), the breakpoint on 11p15.1, as defined by the boundaries of deleted sequences p15 to pter, lied between the SAA2 and SAA4 genes respectively. The oligonucleotide probe specific for SAA4 was deleted indicating that it was not retained on the der(11)t(3;11). On chromosome 3 the integration of MLPA and CGH results suggested a breakpoint just distal to exon 4, in FHIT intron 4.

5.4.5 MLPA validation by CGH in cell line UWOV2

a) Concordance with MLPA

There was a very good correlation of results obtained with the MLPA and CGH arrays regarding relative copy number variation in the UWOV2 cell line. Apart from the homozygous deletion of FHIT exon 5, which was obvious, the MLPA internal control probes, which showed a normal pattern of amplification in MLPA, also had an even copy number representation on the CGH array on visual examination. Similarly, probes that showed decreased MLPA products compatible with hemizygous deletion also showed hemizygous deletion on the CGH array. Two control probes, localized on the chromosome 10 short arm, clearly illustrated this (Figure 5.7) and validated the choice of control. The probe for the CREM gene localized in 10p11.2 had a relative peak height comparable to that of the control specimen (Figure 5.4), while another probe for the RENT gene on 10p14 showed a MLPA relative peak height compatible with hemizygous deletion when compared to the control. These results were matched by the CGH array hybridization results and are shown in Figure 5.7. The larger view of chromosome 10 short arm clearly

shows a deletion from sub-band 10p12.31 to 10pter, explaining the copy number variation of these two probes.

Similarly, the probe for CDH13 on 16q24, which had an amplification pattern comparable to that of the control and was used to normalize the MLPA ratios, was found to be represented in normal copy number by both the MLPA and CGH arrays (Figure 5.8).

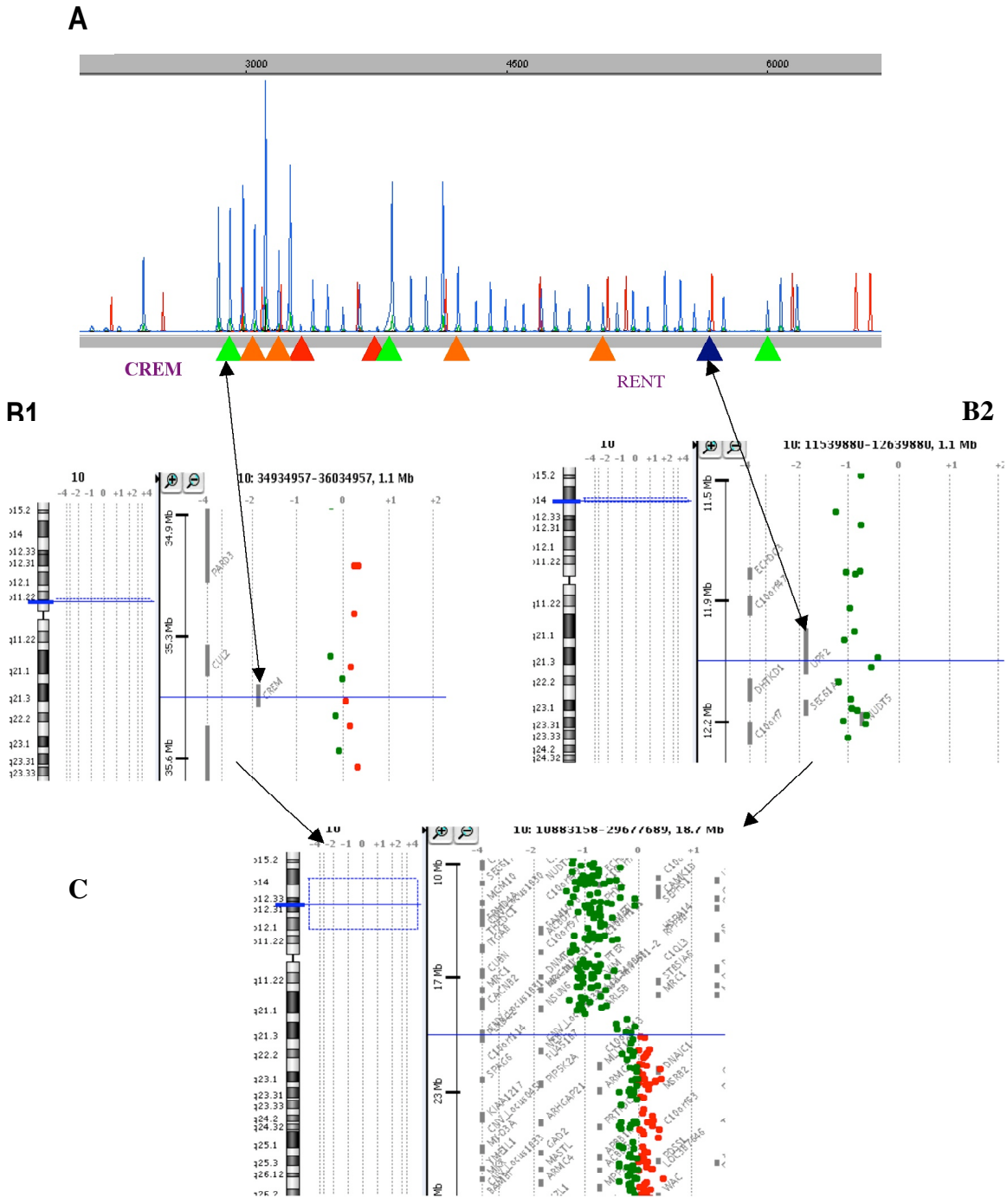


Figure 5.7: *Chromosome 10p control probes, correlation of MLPA and CGH array data. A: MLPA amplification results for the UWOV2 cell line. The black vertical arrows point respectively to the probe for the gene CREM, left (green arrowhead) and to the probe for the gene RENT, right (blue arrowhead), showing respective peak heights. B: the black arrows link to the probes' respective CGH hybridization data. B1: the oligonucleotide probe for the gene CREM is on the zero line, showing the normal copy number, while the probes for the gene RENT (B2) are near the "-1" line, indicating hemizygous deletion. C: a general view of chromosome 10p; a blue box on the ideogram show the magnified region represented on the right. Deletion of 10p is clearly visible at 10 p12.31.*

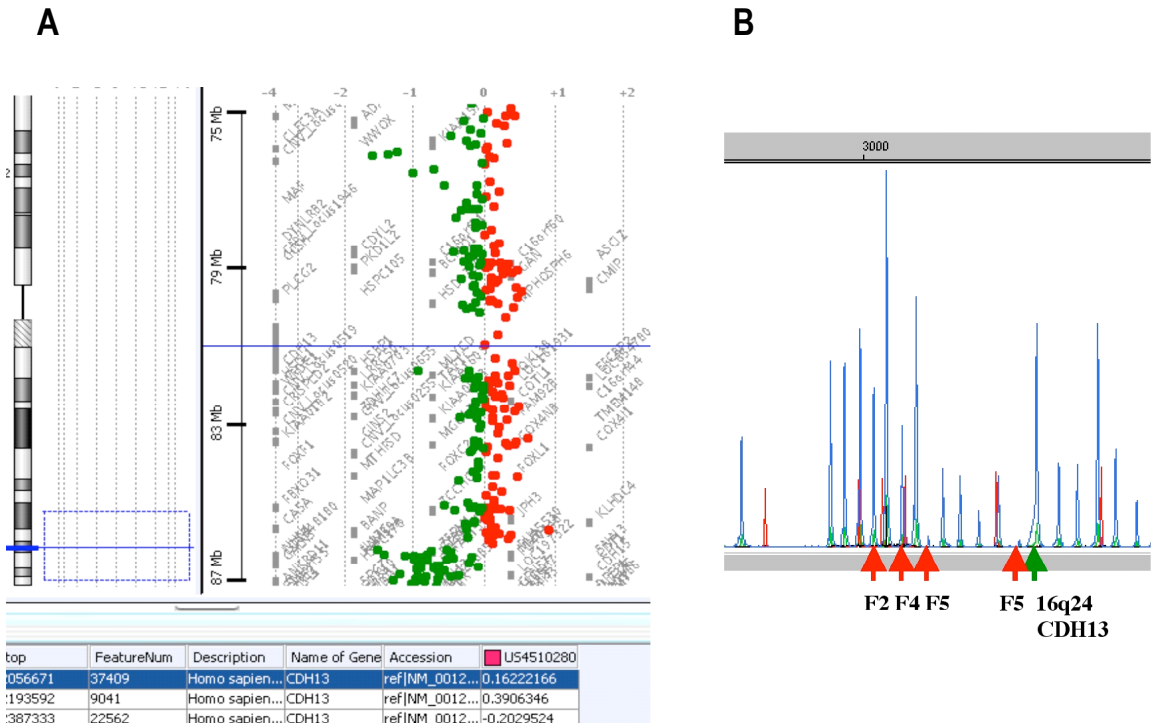


Figure 5.8: *Correspondence of MLPA and CGH array results for the control probe on chromosome16q, CDH13 gene.*
A: CGH data: the blue horizontal line on chromosome 16 ideogram is on the CDH13 probe, as indicated by the probe list below. B: UWOV2 MLPA GeneScan raw data showing the CDH13 probe peak signal (green arrow), comparable to that of the control (also see Figure 5.4).

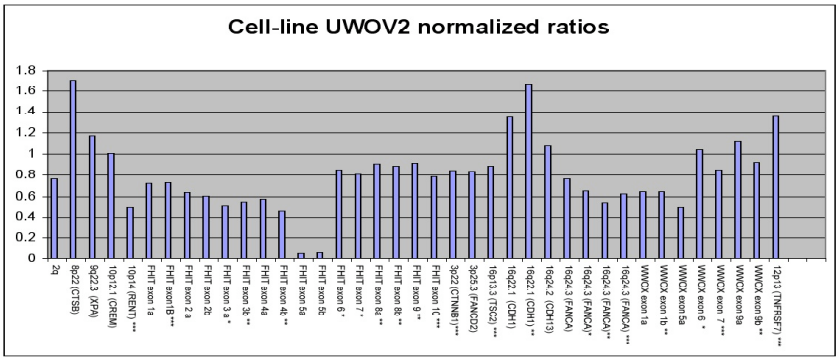
5.4.6 Status of WWOX at FRA16D in the UWOV2 cell line

Interestingly and in contrast to the five ESCC cell lines examined in Chapter 3, probes specific to WWOX exons in MLPA showed deletions of WWOX exons 1 and 5 in the cell line UWOV2 (Figure 5.9, A). With closer examination of the Agilent array oligonucleotide probes, this was confirmed by CGH array results with specific WWOX probes mapping near WWOX exon 5 (see Appendix 6) (Figure 5.9, B).

A set of four probes in the Fanconi Anemia A (FANCA) gene locus on 16q24 was built into the FHIT/WWOX MLPA assay to control for locus specificity of chromosome 16 deletions. Deletions in FANCA were detected in the UWOV2 cell line by both the MLPA and CGH arrays (Figure 5.9). Interestingly, a global view of the chromosome 16 long arm clearly demonstrated that the above deletions were site specific for both the WWOX and FANCA loci (Figure 5.9).

A . MLPA

10p12.1 (CREM)	1.00
10p14 (RENT)	0.49
FHIT exon 1a	0.72
FHIT exon1B	0.73
FHIT exon 2 a	0.63
FHIT exon 2b	0.60
FHIT exon 3 a	0.50
FHIT exon 3b	0.54
FHIT exon 4a	0.58
FHIT exon 4b	0.45
FHIT exon 5a	0.03
FHIT exon 5b	0.06
FHIT exon 6	0.85
FHIT exon 7	0.80
FHIT exon 8a	0.90
FHIT exon 8b	0.88
FHIT exon 9	0.91
FHIT exon 10	0.79
3p22 (CTNNB1)	0.84
3p25.3 (FANCD2)	0.83
16q22.1 (CDH1)	1.35
16q22.1 (CDH1)	1.67
16q24.2 (CDH13)	1.07
16q24.3 (FANCA)	0.76
16q24.3 (FANCA)	0.65
16q24.3 (FANCA)	0.53
16q24.3 (FANCA)	0.62
WWOX exon 1a	0.65
WWOX exon 1b	0.64
WWOX exon 5a	0.49
WWOX exon 6	1.04
WWOX exon 7	0.85
WWOX exon 9a	1.12
WWOX exon 9b	0.92



B. CGH array

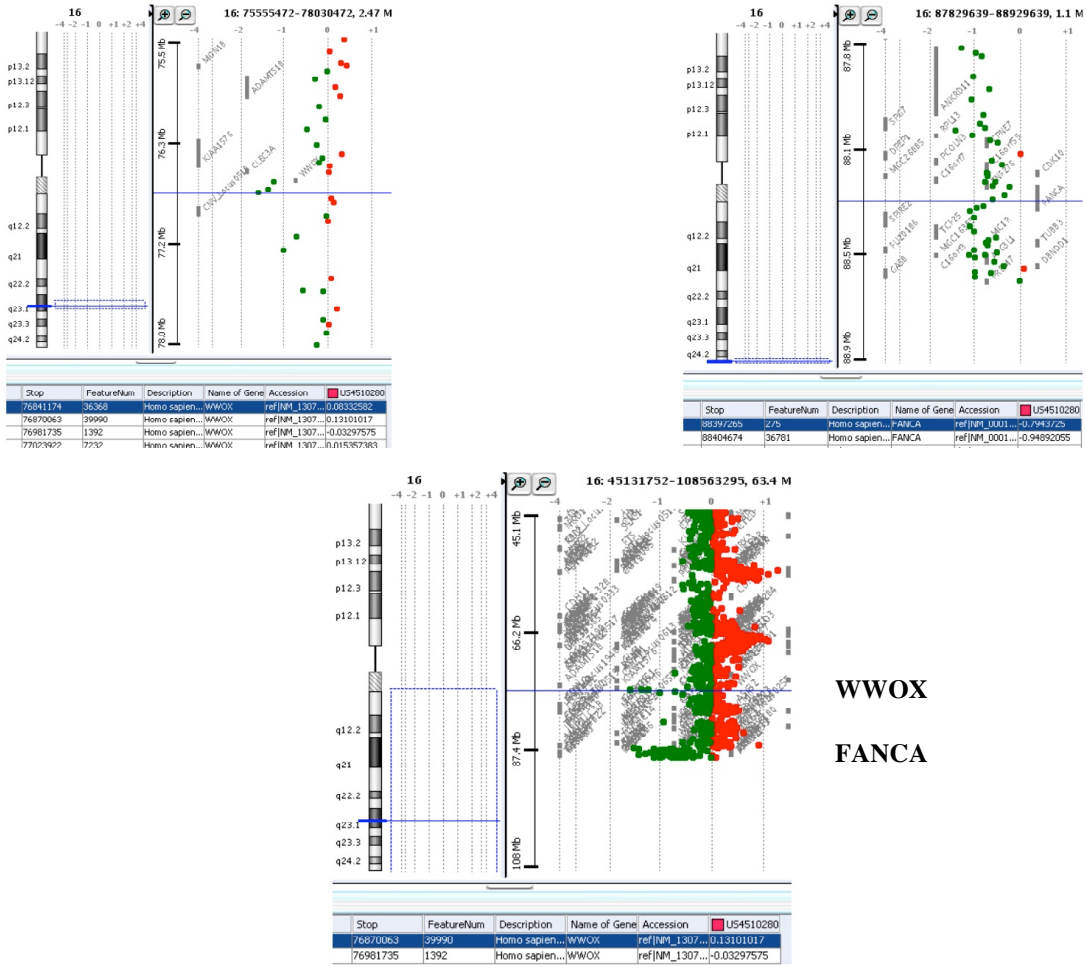


Figure 5.9: *Correlations of MLPA and CGH results in cell line UWOV2, showing deletions in the WWOX and FANCA loci on chromosome 16 q.*

(A) Left: normalized MLPA ratios visualized in a heat map (described in Chapter 4, Section 4.2.4); yellow blocks have normal ratios, orange and red blocks indicate deletions and green blocks amplifications. Probes in the FANCA gene and WWOX exons 1 and 5 showed significantly decreased ratios <0.7. The corresponding MLPA normalized ratio chart is on the right. (B) CGH array data showing, on the left, deletion in the WWOX locus (see Appendix 6) and in the FANCA gene (right). Bottom: overall representation of copy number variations along chromosome 16 showing locus-specific deletions for WWOX and FANCA loci.

5.6 Discussion

This is the first translocation t(3;11)(p14.2;p15.1) described and the first molecular characterization of these chromosomal regions in ovarian carcinoma.

Abnormalities of both chromosomes 3 and 11 have been recurrently described in ovarian cancer. On chromosome 11, aberrations usually result in the loss of genetic material 11p13 to 11pter (Struski, et al. 2003; Taetle, et al. 1999; Wilson, et al. 1996). Chromosome 3 has been less frequently implicated and is usually involved in structural rearrangement; a couple of reports have documented the involvement of both 11p15 and 3p14 (Panani and Roussos 2006; Rao, et al. 2002), although as separate aberrations.

In the UWOV2 cell line, the translocation breakpoint on 3p14.1 was localized in the FHIT gene. It was mapped by MPLA and CGH array within intron 4 and resulted in the homozygous loss of FHIT exon 5, which is the first coding exon. The loss of a FHIT transcript was confirmed by RT-PCR. The breakpoint on 11p15.1 was located between the SAA1 and SAA2 genes. These genes are not typically involved in malignancies, although SAA1 increased expression has recently been observed in colon tissues and correlated with neoplastic development. However, because the gene has no obvious functional contribution to the malignant process, it is thought to be activated by other mediators

present in cancer cells and may only act as a marker (Gutfeld, et al. 2006). Rapid amplification of cDNA ends (3' RACE PCR) did not identify any coding chimeric transcript in the UWOV2 cell line and the functional involvement of SAA1 or 2 is unlikely in this cell line. 5' RACE was not attempted in UWOV2, since there were no obvious reciprocal derivatives of the translocation t(3;11). Similarly to the UWOV2 cell line, no coding chimeric transcripts were detected in three of the five previously described tumor-associated translocations involving FHIT (Fang, et al. 2001; Popovici, et al. 2002). Rather it seems that inactivation of the FHIT gene is the key consequential event associated with malignancy. The fact that FHIT is found to accept a variety of partner chromosome sequences in translocations (chromosomes 4, 16 and 20 have been reported, (Fang, et al. 2001; Popovici, et al. 2002)) strongly supports the idea that the end functional result of structural rearrangements is FHIT's inactivation.

Interestingly, one of the rare FHIT translocations that have been described was present in a breast cancer cell line (Popovici, et al. 2002). Predisposing hereditary mutations in BRCA1 or BRCA2 are responsible for both breast and ovarian tumors in a subset of patients. It is tempting to speculate that a familial predisposing mutation affecting these genes may have been present in both patients' constitutional genome, patients from whom both the ovarian cell line presented in this study and the breast cell line from Popovici study were derived. BRCA1 and BRCA2 interfere with the repair pathway mechanism of genes at fragile sites and are downstream targets of the ATR gene (see Chapter 3, Section 3.1) and preliminary studies suggest that BRCA1 deletions correlate with FHIT inactivation in sporadic breast tumor cells. It would be interesting to assess FHIT and WWOX status in tumor cells of patients who have an inherited BRCA1 –BRCA2 predisposing mutation. It could be speculated that BRCA1 repair deficiency, which is

involved in the ATR dependant repair pathway, would predispose patients to having increased breaks at fragile sites.

In the UWOV2 cell line, the observation of a site-specific deletion within the WWOX locus on 16q using CGH array is reminiscent of the concomitant alteration of FHIT and WWOX proteins previously observed in breast cancer tissues (Guler, et al. 2005).

The Fanconi Anemia gene (FANCA) was found to be deleted in half copy number; since the related gene FANCD2 participates in the fragile sites repair pathway (Chapter 3, Section 3.1), it raises questions regarding the time chronology for this deletion. If FANCA deletions preceded alterations at fragile sites, this event may have contributed to FHIT and WWOX intragenic deletions. Indeed, a recent study showed the association of reduced FANCD2 expression with increased chromosomal instability and increased risk of developing ovarian cancer (Pejovic, et al. 2006). However, the extent of FANC gene deletions in ovarian cancer would have to be further evaluated to clarify the frequency of its participation.

In summary, a stable translocation that interrupts the FHIT gene was characterized; this is the sixth cancer associated translocation involving FHIT and the first described in ovarian carcinoma. The characterization of a translocation that targets the FHIT gene reinforces the notion that chromosomal instability and genes involved in these sites and pathways may have an important role in carcinogenesis.

CHAPTER 6

GENERAL CONCLUSIONS AND FUTURE WORK

6.1 Summary

This project was initiated to explore the extent of the involvement of chromosomal "hot spots" (as defined by regions of genomic instability preferentially targeted in cancer) in SA cancer patients, on the basis that they may host critical genes involved in the initiation and or progression of cancer and that they are sensitive to external carcinogens. Two types of chromosomal "hot spots" were examined in this thesis: chromosomal amplification (amplicon) and chromosomal fragile sites.

6.1.1 Characterization of a 12p amplicon in a secondary leukemia patient

A minimal region of amplification on 12p was characterized in a secondary leukemia patient and corresponded to a similar region described in testicular germ cell tumors. This region pointed to stem cell genes as potential key oncogenes. These findings suggested that the 12p stem cell cluster and in particular NANOG, STELLAR and GDF3 may be more commonly involved in a variety of cancers. GDF3 and NANOG are essential for embryonic stem cell self –renewal (Levasseur, et al. 2008) and their amplification is likely to promote stem cell qualities and enhanced cell proliferation. They may constitute potential target genes for therapy worth considering as proposed for Notch another stem cell self-renewal gene implicated in cancer (Chang, et al. 2008).

6.1.2 Characterization of fragile sites deletions affecting tumor suppressor genes

A novel approach to scanning genes at fragile sites was developed and evaluated in this work and will facilitate screening of FHIT and WWOX deletions as well as CAV1 and 2 copy number changes for research purposes. This MLPA assay may also be useful for clinical screening in cancers for which deletions of FHIT/ WWOX have been associated with an adverse prognosis such as breast and colorectal cancers. The PAS data analysis software and a VBA heat-map program were both developed, in a flexible format and can be used in a number of MLPA applications.

A novel translocation within the FHIT gene was characterized in an ovarian carcinoma cell line. This is the fifth tumor-associated translocation described that involved the FHIT gene, reinforcing the idea that this gene is indeed a critical anti-oncogene, involved in either initiation or progression of cancer. Array CGH experiments used in conjunction with the FHIT/WWOX MLPA assay showed a high concordance of results between the two approaches further validating the MLPA assay.

Exon deletions have been found to be frequent in both the FHIT and WWOX loci in ESCC cell lines and in patients' tumor biopsies. The difference of incidence between WWOX deletions in cohorts from Gauteng and the Eastern Cape needs to be verified by screening additional specimens from the Gauteng area, but already raises questions as to the possibility of carcinogen exposure in the Eastern Cape, preferentially targeting chromosome 16. In this regard, fumonisin could be a potential toxic agent with or without co-exposure with aflatoxin. Alternatively, the Eastern Cape population may carry a genetic predisposing weakness in a DNA repair pathway that participates in single-strand break pathway.

6.1.2 A cancer initiation/progression model involving fragile sites

While the mechanistic explanation for the fragility of chromosomal "hot spots" remain for the most unexplained, their involvement in cancer promoting events is obvious (Glover 2006). Not only are they sensitive to at least some external carcinogen but their rearrangement or deletion, which could be due to mechanistic events, may result in the rearrangement of a key tumor promoting gene(s) either through amplification of an oncogene or through deletion of a tumor suppressor gene.

There is a variation of individual response to conditions of fragile site induction (Lukusa and Fryns 2007) suggesting that there is a variable efficiency in DNA repair genes monitoring chromatin repair at fragile sites. These include ATR, BRCA1, FANCD2 and SMC1 (refer to Chapter 3, section 3.1.4). There is recent evidence that the ataxia-telangiectasisa mutated (ATM) pathway, which controls double strand breaks repair (DSB) is also playing a role in maintaining fragile site stability and that double strand breaks and DNA fragmentation do occur under condition of replication stress, contrasting with original notions (Ozeri-Galai, et al. 2007).

A model is proposed regarding the Eastern Cape population exposed to carcinogens that may induce breaks at fragile sites. The figure below (Figure 6.1) illustrates this model with the FHIT gene and would apply for WWOX as well since both genes are involved in apoptosis.

In brief, should breaks occur in high frequency at fragile sites this may inactivate FHIT/WWOX through deletions affecting coding exons. Cells with defective FHIT/WWOX, and therefore defective apoptosis may then progress through cell division and accumulate genetic aberrations. Similarly DSBs may go though cell division and generate interstitial deletions, translocations or inverted duplications. A genetic tendency

leading to ineffective repair mechanisms in either the ATR or ATM pathway is likely to favor an increased frequency of unrepaired breaks at CFS.

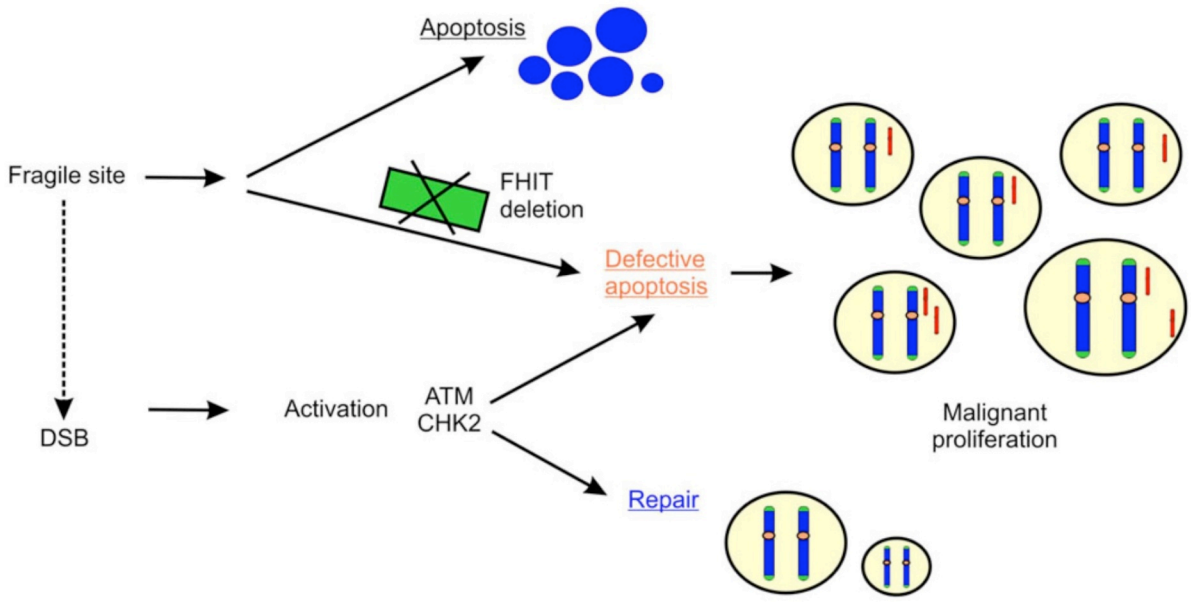


Figure 6.1: *Model for the formation of malignant clones through deletions at fragile sites and DSB generated at the same time.*

A recent study shows that the inactivation of genes at common fragile sites is non random (McAvoy, et al. 2007). One can speculate that the active gene machinery at work in a particular cell type at a particular stage of development is influencing the occurrence of deletions at some sites in preference to others. The chromatid would have to be loose and accessible at points of high gene transcription activity.

It will be interesting to investigate if a preferred subsequent molecular pathway follows on the inactivation of specific genes at fragile sites. WWOX and FHIT are coordinately deleted and associated with a poorer prognosis in breast cancer (Guler, et al. 2005) ; at the same time BRCA1 positive breast cancer tumors tend to have FHIT deletions (Cavalli, et al. 2004; Turner, et al. 2002). It will be important to assess the state of other repair genes

in breast cancer patients that have FHIT and WWOX deletions in tumor cells and who are otherwise BRCA1 and BRCA2 negative.

6.2 Future work

6.2.1 Development of a FISH probe to detect amplification of the 12p stem cell gene cluster

A FISH probe will be designed using BACs and or YACs mapping at 12p31 and which cover the stem cell gene cluster represented by NANOG, STELLAR and GDF3. This probe or probe cocktail will be useful in screening for amplification of this cluster in differing types of cancers. Interestingly 12p amplification has also been observed in some ESCC tumors.

6.2.2 Improvement of the MLPA assay

It is envisaged to improve the MLPA assay and to include new internal control probes. Control probes will be designed in order to span a conserved exon sequence of genes that are located in genomic regions rarely affected by copy number variations in cancer and that do not locate near a known fragile site.

A new algorithm for the correction of regression in appropriate samples is currently being designed that will more accurately correct for the signal “fading away” effect observed with the largest MLPA probes.

6.2.3 Furthering the investigation of the role of genes at fragile sites in SA ESSC

First, in order to verify if the surprising observed difference of WWOX rate of deletions between the two cohorts from Gauteng and the Eastern Cape is real, we will test FFPE samples from a Gauteng cohort using the same FHIT/ WWOX MLPA assay.

Second, we will assess the Eastern Cape FFPE specimens using CGH array for copy number analysis. This will be performed on forty patients and forty controls FFPE specimens represented by normal histo-pathological tissue sections collected from the same population group (Eastern Cape). The Affymetrix 500K SNP array will be used for this study. The SNP array will allow for the verification of the above data especially regarding FHIT and WWOX deletions and the extent of these deletions on chromosome 3 and 16 respectively. This will be achieved by looking at regions of loss of heterozygosity at these specific sites. The array work will also shed light on whether these deletions are locus specific and if they override in frequency other chromosomal aberrations. The latter proving positive would strongly suggest an early role for alterations at fragile sites.

Another aspect of interest derived from copy umber analysis will be to test the hypothesis that there may be nonrandom copy number changes associated with FHIT and /or WWOX exon deletions. Should this be the case, the nonrandom alterations may reflect a preferred pathway of molecular events associated with deletions of specific genes at fragile sites. The recent report of nonrandom involvement of genes at fragiles sites in a variety of cancer support this idea (McAvoy, et al. 2007).

Lastly but importantly, the choice of the Affymetrix SNP array platform will bring information regarding the possibility of associated polymorphisms in fragile sites associated repair genes and/ or others.

Third, to explore the effect of the putative local carcinogenic exposure and should a collaboration be feasible, an in-vitro model of fumonisin exposure in cell cultures should ideally be tested for fragile sites expression.

The findings by independent groups (Bartkova, et al. 2005; Gorgoulis, et al. 2005) that oncogene induction of replication stress and pre cancerous lesions are marked by an increased expression of fragile sites, point to the value of an appropriate screening test. Should fragile sites be over expressed in *in vivo* situation of carcinogen exposure, the early detection of deletions may be predictive of cancer and an assay could be very useful.

In the above work it has been shown that the exon based FHIT/WWOX MPLA assay can effectively detect those deletions at fragile sites likely to have significant functional consequences and potentially lead to either cancer initiation and/or cancer progression. However the MLPA assay will not detect all the deletions/ duplications occurring in noncoding regions of the fragile sites loci. If such lesions act as an "alarm" signaling an array based method with a high coverage in these sites (such as the 6K Affymetrix array or the 244K Agilent oligonucleotide array) would be ideal .

These assay are however prohibitively expensive in a clinical situation and more suited for research purposes. Indeed should a CGH array study be applied in screening for early fragile sites deletions in cancer induction model and early cancer lesions it may also lay the ground to develop a new MLPA test with probes specific for the most commonly deleted fragile sites regions.

The MPLA assay is sensitive in that it covers all critical exons in both genes with duplicate probes. The sensitivity of both MLPA and array will be of course dependant on the quality of the samples tested; a dilution of cancer cells with normal cells will decrease the sensitivity of both assays and give false negative results.

The MLPA assay is also robust as the probes are gene specific and point mutations that would result in non amplification (false positive results) are unlikely to occur in coding exon sequences. The FHIT/WWOX MLPA assay is also considerably a cheaper than a whole genome array analysis and would be ideally suited for prognosis evaluation in breast and colorectal cancer. It would also be of great value to screen early lesions in order to clarify the clinical role of the FHIT/WWOX assay in this context. It may prove very useful in detecting early events of cancer initiation and would thus lead to early treatment in the course of disease development.

REFERENCES

Acs G, Pasha T, Zhang PJ. 2004. WT1 is differentially expressed in serous, endometrioid, clear cell, and mucinous carcinomas of the peritoneum, fallopian tube, ovary, and endometrium. Int J Gynecol Pathol 23(2):110-8.

Ando T, Ishiguro H, Kimura M, Mitsui A, Mori Y, Sugito N, Tomoda K, Mori R, Harada K, Katada T and others. 2007. The overexpression of caveolin-1 and caveolin-2 correlates with a poor prognosis and tumor progression in esophageal squamous cell carcinoma. Oncol Rep 18(3):601-9.

Almstrup K, Ottesen AM, Sonne SB, Hoei-Hansen CE, Leffers H, Rajpert-De Meyts E, Skakkebaek NE. 2005. Genomic and gene expression signature of the pre-invasive testicular carcinoma in situ. Cell Tissue Res 322(1):159-65.

Aly MS, Dal Cin P, Jiskoot P, Deneffe G, Marynen P, Van den Berghe H. 1994. Competitive in situ hybridization in a mediastinal germ cell tumor. Cancer Genet Cytogenet 73(1):53-6.

Amor DJ, Algar EM, Slater HR, Smith PJ. 1998. High frequency of t(12;21) in childhood acute lymphoblastic leukemia detected by RT-PCR. Pathology 30(4):381-5.

Aqeilan RI, Kuroki T, Pekarsky Y, Albagha O, Trapasso F, Baffa R, Huebner K, Edmonds P, Croce CM. 2004. Loss of WWOX expression in gastric carcinoma. Clin Cancer Res 10(9):3053-8.

Aqeilan RI, Trapasso F, Hussain S, Costinean S, Marshall D, Pekarsky Y, Hagan JP, Zanesi N, Kaou M, Stein GS and others. 2007. Targeted deletion of Wwox reveals a tumor suppressor function. Proc Natl Acad Sci U S A 104(10):3949-54.

Aranda M, Perez-Alzola LP, Ellahuene MF, Sepulveda C. 2000. Assessment of in vitro mutagenicity in Salmonella and in vivo genotoxicity in mice of the mycotoxin fumonisin B(1). Mutagenesis 15(6):469-71.

Arlt MF, Durkin SG, Ragland RL, Glover TW. 2006. Common fragile sites as targets for chromosome rearrangements. DNA Repair (Amst) 5(9-10):1126-35.

Arlt MF, Xu B, Durkin SG, Casper AM, Kastan MB, Glover TW. 2004. BRCA1 is required for common-fragile-site stability via its G2/M checkpoint function. Mol Cell Biol 24(15):6701-9.

Arun B, Kilic G, Yen C, Foster B, Yardley DA, Gaynor R, Ashfaq R. 2005. Loss of FHIT expression in breast cancer is correlated with poor prognostic markers. Cancer Epidemiol Biomarkers Prev 14(7):1681-5.

Balmain A, Gray J, Ponder B. 2003. The genetics and genomics of cancer. Nat Genet 33 Suppl:238-44.

Bartkova J, Horejsi Z, Koed K, Kramer A, Tort F, Zieger K, Guldberg P, Sehested M, Nesland JM, Lukas C and others. 2005. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 434(7035):864-70.

Bavi P, Jehan Z, Atizado V, Al-Dossari H, Al-Dayel F, Tulbah A, Amr SS, Sheikh SS, Ezzat A, El-Solh H and others. 2006. Prevalence of fragile histidine triad expression in tumors from saudi arabia: a tissue microarray analysis. Cancer Epidemiol Biomarkers Prev 15(9):1708-18.

Bednarek AK, Laflin KJ, Daniel RL, Liao Q, Hawkins KA, Aldaz CM. 2000. WWOX, a novel WW domain-containing protein mapping to human chromosome 16q23.3-24.1, a region frequently affected in breast cancer. Cancer Res 60(8):2140-5.

Bentz M, Huck K, du Manoir S, Joos S, Werner CA, Fischer K, Dohner H, Lichter P. 1995. Comparative genomic hybridization in chronic B-cell leukemias shows a high incidence of chromosomal gains and losses. Blood 85(12):3610-8.
Bertholon J, Wang Q, Galmarini CM, Puisieux A. 2006. Mutational targets in colorectal cancer cells with microsatellite instability. Fam Cancer 5(1):29-34.
Bertoni L, Attolini C, Tessera L, Mucciolo E, Giulotto E. 1994. Telomeric and nontelomeric (TTAGGG)n sequences in gene amplification and chromosome stability. Genomics 24(1):53-62.
Bester AC, Schwartz M, Schmidt M, Garrigue A, Hacein-Bey-Abina S, Cavazzana-Calvo M, Ben-Porat N, Von Kalle C, Fischer A, Kerem B. 2006. Fragile sites are preferential targets for integrations of MLV vectors in gene therapy. Gene Ther 13(13):1057-9.
Bey E, Alexander J, Whitcutt JM, Hunt JA, Gear JH. 1976. Carcinoma of the esophagus in Africans: establishment of a continuously growing cell line from a tumor specimen. In Vitro 12(2):107-14.
Bodmer WF, Cottrell S, Frischauf AM, Kerr IB, Murday VA, Rowan AJ, Smith MF, Solomon E, Thomas H, Varesco L. 1989. Genetic analysis of colorectal cancer. Princess Takamatsu Symp 20:49-59.
Boldog F, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R and others. 1997. Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6(2):193-203.
Borresen AL. 1992. Oncogenesis in ovarian cancer. Acta Obstet Gynecol Scand Suppl 155:25-30.
Brown LM, Hoover R, Silverman D, Baris D, Hayes R, Swanson GM, Schoenberg J, Greenberg R, Liff J, Schwartz A and others. 2001. Excess incidence of squamous cell esophageal cancer among US Black men: role of social class and other risk factors. Am J Epidemiol 153(2):114-22.
Buffart TE, Coffa J, Hermsen MA, Carvalho B, van der Sijp JR, Ylstra B, Pals G, Schouten JP, Meijer GA. 2005. DNA copy number changes at 8q11-24 in metastasized colorectal cancer. Cell Oncol 27(1):57-65.
Butzler C, Zou X, Popov AV, Bruggemann M. 1997. Rapid induction of B-cell lymphomas in mice carrying a human IgH/c-mycYAC. Oncogene 14(11):1383-8.
Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M and others. 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 101(9):2999-3004.
Call KM, Ito CY, Lindberg C, Memisoglu A, Petrou C, Glaser T, Jones C, Housman DE. 1992. Mapping and characterization of 129 cosmids on human chromosome 11p. Somat Cell Mol Genet 18(5):463-75.
Cantor JP, Iliopoulos D, Rao AS, Druck T, Semba S, Han SY, McCorkell KA, Lakshman TV, Collins JE, Wachsberger P and others. 2007. Epigenetic modulation of endogenous tumor suppressor expression in lung cancer xenografts suppresses tumorigenicity. Int J Cancer 120(1):24-31.
Capaccio P, Ottaviani F, Cuccarini V, Cenzuales S, Cesana BM, Pignataro L. 2005. Association between methylenetetrahydrofolate reductase polymorphisms, alcohol intake and oropharyngolaryngeal carcinoma in northern Italy. J Laryngol Otol 119(5):371-6.
Casper AM, Nghiem P, Arlt MF, Glover TW. 2002. ATR regulates fragile site stability. Cell 111(6):779-89.

Casson AG, Zheng Z, Chiasson D, MacDonald K, Riddell DC, Guernsey JR, Guernsey DL, McLaughlin J. 2003. Associations between genetic polymorphisms of Phase I and II metabolizing enzymes, p53 and susceptibility to esophageal adenocarcinoma. Cancer Detect Prev 27(2):139-46.

Cavalli LR, Singh B, Isaacs C, Dickson RB, Haddad BR. 2004. Loss of heterozygosity in normal breast epithelial tissue and benign breast lesions in BRCA1/2 carriers with breast cancer. Cancer Genet Cytogenet 149(1):38-43.

Chang NS, Doherty J, Ensign A, Lewis J, Heath J, Schultz L, Chen ST, Oppermann U. 2003. Molecular mechanisms underlying WOX1 activation during apoptotic and stress responses. Biochem Pharmacol 66(8):1347-54.

Chen YJ, Chen PH, Lee MD, Chang JG. 1997. Aberrant FHIT transcripts in cancerous and corresponding non-cancerous lesions of the digestive tract. Int J Cancer 72(6):955-8.

Cheung AL, Si HX, Wang LD, An JY, Tsao SW. 2005. Loss of heterozygosity analyses of esophageal squamous cell carcinoma and precursor lesions from a high incidence area in China. Cancer Lett.

Chopra SA, Chopra FS. 1977. Cancer in the Africans and Arabs of Zanzibar. Int J Cancer 19(3):298-304.

Chu FS, Li GY. 1994. Simultaneous occurrence of fumonisin B1 and other mycotoxins in moldy corn collected from the People's Republic of China in regions with high incidences of esophageal cancer. Appl Environ Microbiol 60(3):847-52.

Chu TY, Shen CY, Chiou YS, Lu JJ, Perng CL, Yu MS, Liu HS. 1998. HPV-associated cervical cancers show frequent allelic loss at 3p14 but no apparent aberration of FHIT mRNA. Int J Cancer 75(2):199-204.

Chung YJ, Jonkers J, Kitson H, Fiegler H, Humphray S, Scott C, Hunt S, Yu Y, Nishijima I, Velds A and others. 2004. A whole-genome mouse BAC microarray with 1-Mb resolution for analysis of DNA copy number changes by array comparative genomic hybridization. Genome Res 14(1):188-96.

Ciullo M, Debily MA, Rozier L, Autiero M, Billault A, Mayau V, El Marhomy S, Guardiola J, Bernheim A, Coullin P and others. 2002. Initiation of the breakage-fusion-bridge mechanism through common fragile site activation in human breast cancer cells: the model of PIP gene duplication from a break at FRA7I. Hum Mol Genet 11(23):2887-94.

Clark AT, Rodriguez RT, Bodnar MS, Abeyta MJ, Cedars MI, Turek PJ, Firpo MT, Reijo Pera RA. 2004. Human STELLAR, NANOG, and GDF3 genes are expressed in pluripotent cells and map to chromosome 12p13, a hotspot for teratocarcinoma. Stem Cells 22(2):169-79.

Cody NA, Ouellet V, Manderson EN, Quinn MC, Filali-Mouhim A, Tellis P, Zietarska M, Provencher DM, Mes-Masson AM, Chevrette M and others. 2007. Transfer of chromosome 3 fragments suppresses tumorigenicity of an ovarian cancer cell line monoallelic for chromosome 3p. Oncogene 26(4):618-32.

Collins C, Kuo WL, Segraves R, Fuscoe J, Pinkel D, Gray JW. 1991. Construction and characterization of plasmid libraries enriched in sequences from single human chromosomes. Genomics 11(4):997-1006.

Cook-Mozaffari PJ, Azordegan F, Day NE, Ressicaud A, Sabai C, Aramesh B. 1979. Oesophageal cancer studies in the Caspian Littoral of Iran: results of a case-control study. Br J Cancer 39(3):293-309.

Cooper K, Taylor L, Govind S. 1995. Human papillomavirus DNA in oesophageal carcinomas in South Africa. J Pathol 175(3):273-7.

Coquelle A, Rozier L, Dutrillaux B, Debatisse M. 2002. Induction of multiple double-strand breaks within an hsr by meganucleaseI-SceI expression or fragile site activation leads to formation of double minutes and other chromosomal rearrangements. Oncogene 21(50):7671-9.

Coquelle A, Toledo F, Stern S, Bieth A, Debatisse M. 1998. A new role for hypoxia in tumor progression: induction of fragile site triggering genomic rearrangements and formation of complex DMs and HSRs. Mol Cell 2(2):259-65.

Corbin S, Neilly ME, Espinosa R, 3rd, Davis EM, McKeithan TW, Le Beau MM. 2002. Identification of unstable sequences within the common fragile site at 3p14.2: implications for the mechanism of deletions within fragile histidine triad gene/common fragile site at 3p14.2 in tumors. Cancer Res 62(12):3477-84.

Dandara C, Ballo R, Parker MI. 2005. CYP3A5 genotypes and risk of oesophageal cancer in two South African populations. Cancer Lett 225(2):275-82.

Dandara C, Li DP, Walther G, Parker MI. 2006. Gene-environment interaction: the role of SULT1A1 and CYP3A5 polymorphisms as risk modifiers for squamous cell carcinoma of the oesophagus. Carcinogenesis 27(4):791-7.

Dawsey SM, Lewin KJ, Wang GQ, Liu FS, Nieberg RK, Yu Y, Li JY, Blot WJ, Li B, Taylor PR. 1994. Squamous esophageal histology and subsequent risk of squamous cell carcinoma of the esophagus. A prospective follow-up study from Linxian, China. Cancer 74(6):1686-92.

Deng WG, Nishizaki M, Fang B, Roth JA, Ji L. 2007. Induction of apoptosis by tumor suppressor FHIT via death receptor signaling pathway in human lung cancer cells. Biochem Biophys Res Commun 355(4):993-9.

Draper JS, Smith K, Gokhale P, Moore HD, Maltby E, Johnson J, Meisner L, Zwaka TP, Thomson JA, Andrews PW. 2004. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat Biotechnol 22(1):53-4.

Druck T, Hadaczek P, Fu TB, Ohta M, Siprashvili Z, Baffa R, Negrini M, Kastury K, Veronese ML, Rosen D and others. 1997. Structure and expression of the human FHIT gene in normal and tumor cells. Cancer Res 57(3):504-12.

du Manoir S, Speicher MR, Joos S, Schrock E, Popp S, Dohner H, Kovacs G, Robert-Nicoud M, Lichter P, Cremer T. 1993. Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. Hum Genet 90(6):590-610.

Du Plessis L, Dietzsch E, Van Gele M, Van Roy N, Van Helden P, Parker MI, Mugwanya DK, De Groot M, Marx MP, Kotze MJ and others. 1999. Mapping of novel regions of DNA gain and loss by comparative genomic hybridization in esophageal carcinoma in the Black and Colored populations of South Africa. Cancer Res 59(8):1877-83.

Dumon KR, Ishii H, Fong LY, Zanesi N, Fidanza V, Mancini R, Vecchione A, Baffa R, Trapasso F, During MJ and others. 2001a. FHIT gene therapy prevents tumor development in Fhit-deficient mice. Proc Natl Acad Sci U S A 98(6):3346-51.

Dumon KR, Ishii H, Vecchione A, Trapasso F, Baldassarre G, Chakrani F, Druck T, Rosato EF, Williams NN, Baffa R and others. 2001b. Fragile histidine triad expression delays tumor development and induces apoptosis in human pancreatic cancer. Cancer Res 61(12):4827-36.

Durkin SG, Ragland RL, Arlt MF, Mulle JG, Warren ST, Glover TW. 2008. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. Proc Natl Acad Sci U S A 105(1):246-51.

Dutton MF. 1996. Fumonisins, mycotoxins of increasing importance: their nature and their effects. Pharmacol Ther 70(2):137-61.

Eisenbrand G, Spiegelhalder B, Preussmann R. 1980. Nitrate and nitrite in saliva. Oncology 37(4):227-31.
Ellis A, Field JK, Field EA, Friedmann PS, Fryer A, Howard P, Leigh IM, Risk J, Shaw JM, Whittaker J. 1994. Tylosis associated with carcinoma of the oesophagus and oral leukoplakia in a large Liverpool family--a review of six generations. Eur J Cancer B Oral Oncol 30B(2):102-12.
Ezeh UI, Turek PJ, Reijo RA, Clark AT. 2005. Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma. Cancer 104(10):2255-65.
Fabbri M, Iliopoulos D, Trapasso F, Aqeilan RI, Cimmino A, Zanesi N, Yendamuri S, Han SY, Amadori D, Huebner K and others. 2005. WWOX gene restoration prevents lung cancer growth in vitro and in vivo. Proc Natl Acad Sci U S A 102(43):15611-6.
Fang JM, Arlt MF, Burgess AC, Dagenais SL, Beer DG, Glover TW. 2001. Translocation breakpoints in FHIT and FRA3B in both homologs of chromosome 3 in an esophageal adenocarcinoma. Genes Chromosomes Cancer 30(3):292-8.
Feitelson MA, Lee J. 2006. Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis. Cancer Lett.
Fenech M. 2001. The role of folic acid and Vitamin B12 in genomic stability of human cells. Mutat Res 475(1-2):57-67.
Fields MM, Chevlen E. 2006. Ovarian cancer screening: a look at the evidence. Clin J Oncol Nurs 10(1):77-81.
Focarelli ML, Montagna C, Colombo R, Ried T, Vezzoni P, Musio A. 2006. SMC1 inhibition results in FRA3B expression but has no effect on its delayed replication. Mutat Res 595(1-2):23-8.
Forus A, Bjerkehagen B, Sirvent N, Meza-Zepeda LA, Coindre JM, Berner JM, Myklebost O, Pedeutour F. 2001. A well-differentiated liposarcoma with a new type of chromosome 12-derived markers. Cancer Genet Cytogenet 131(1):13-8.
Gamieldien W, Victor TC, Mugwanya D, Stepien A, Gelderblom WC, Marasas WF, Geiger DH, van Helden PD. 1998. p53 and p16/CDKN2 gene mutations in esophageal tumors from a high-incidence area in South Africa. Int J Cancer 78(5):544-9.
Gelderblom WC, Abel S, Smuts CM, Marnewick J, Marasas WF, Lemmer ER, Ramljak D. 2001. Fumonisin-induced hepatocarcinogenesis: mechanisms related to cancer initiation and promotion. Environ Health Perspect 109 Suppl 2:291-300.
Gelderblom WC, Marasas WF, Lebepe-Mazur S, Swanevelder S, Vessey CJ, Hall Pde L. 2002. Interaction of fumonisin B(1) and aflatoxin B(1) in a short-term carcinogenesis model in rat liver. Toxicology 171(2-3):161-73.
Gemmill RM, West JD, Boldog F, Tanaka N, Robinson LJ, Smith DI, Li F, Drabkin HA. 1998. The hereditary renal cell carcinoma 3;8 translocation fuses FHIT to a patched-related gene, TRC8. Proc Natl Acad Sci U S A 95(16):9572-7.
Geurts JM, Schoenmakers EF, Roijer E, Stenman G, Van de Ven WJ. 1997. Expression of reciprocal hybrid transcripts of HMGIC and FHIT in a pleomorphic adenoma of the parotid gland. Cancer Res 57(1):13-7.
Glover TW. 1998. Instability at chromosomal fragile sites. Recent Results Cancer Res 154:185-99.
Glover TW. 2006. Common fragile sites. Cancer Lett 232(1):4-12.
Glover TW, Arlt MF, Casper AM, Durkin SG. 2005. Mechanisms of common fragile site instability. Hum Mol Genet 14 Spec No. 2:R197-205.

Glover TW, Berger C, Coyle J, Echo B. 1984. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67(2):136-42.
Glover TW, Hoge AW, Miller DE, Ascara-Wilke JE, Adam AN, Dagenais SL, Wilke CM, Dierick HA, Beer DG. 1998. The murine Fhit gene is highly similar to its human orthologue and maps to a common fragile site region. Cancer Res 58(15):3409-14.
Glover TW, Stein CK. 1987. Induction of sister chromatid exchanges at common fragile sites. Am J Hum Genet 41(5):882-90.
Goldberg JI, Borgen PI. 2006. Breast cancer susceptibility testing: past, present and future. Expert Rev Anticancer Ther 6(8):1205-14.
Golub TR, Barker GF, Lovett M, Gilliland DG. 1994. Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. Cell 77(2):307-16.
Gorgoulis VG, Vassiliou LV, Karakaidos P, Zacharatos P, Kotsinas A, Liloglou T, Venere M, Ditullio RA, Jr., Kastrinakis NG, Levy B and others. 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. Nature 434(7035):907-13.
Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C and others. 2007. Patterns of somatic mutation in human cancer genomes. Nature 446(7132):153-8.
Guler G, Uner A, Guler N, Han SY, Iliopoulos D, Hauck WW, McCue P, Huebner K. 2004. The fragile genes FHIT and WWOX are inactivated coordinately in invasive breast carcinoma. Cancer 100(8):1605-14.
Guler G, Uner A, Guler N, Han SY, Iliopoulos D, McCue P, Huebner K. 2005. Concordant loss of fragile gene expression early in breast cancer development. Pathol Int 55(8):471-8.
Gutfeld O, Prus D, Ackerman Z, Dishon S, Linke RP, Levin M, Urieli-Shoval S. 2006. Expression of serum amyloid A, in normal, dysplastic, and neoplastic human colonic mucosa: implication for a role in colonic tumorigenesis. J Histochem Cytochem 54(1):63-73.
Guven GS, Hacihanefioglu S, Cenani A. 1999. Expression of aphidicolin, FUdR and caffeine-induced fragile sites in lymphocytes of healthy Turkish individuals. Genetica 105(1):109-16.
Heim S. 1996. Genetic nomenclature: ISCN and ISGN. Pediatr Hematol Oncol 13(2):iii.
Heller A, Trifonov V, Rubtsov N, Sauerbrey A, Starke H, Loncarevic IF, Claussen U, Liehr T. 2003. Complex chromosomal rearrangements in a secondary acute myeloblastic leukemia after chemotherapy in TRAPS. Oncol Rep 10(6):1789-92.
Hellman A, Zlotorynski E, Scherer SW, Cheung J, Vincent JB, Smith DI, Trakhtenbrot L, Kerem B. 2002. A role for common fragile site induction in amplification of human oncogenes. Cancer Cell 1(1):89-97.
Hendricks D, Parker MI. 2002. Oesophageal cancer in Africa. IUBMB Life 53(4-5):263-8.
Herbst RS, Lippman SM. 2007. Molecular signatures of lung cancer--toward personalized therapy. N Engl J Med 356(1):76-8.
Hoglund M, Johansson B, Pedersen-Bjergaard J, Marynen P, Mitelman F. 1996. Molecular characterization of 12p abnormalities in hematologic malignancies: deletion of KIP1, rearrangement of TEL, and amplification of CCND2. Blood 87(1):324-30.
Holmes RS, Vaughan TL. 2007. Epidemiology and pathogenesis of esophageal cancer. Semin Radiat Oncol 17(1):2-9.

Howlett NG, Taniguchi T, Durkin SG, D'Andrea AD, Glover TW. 2005. The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. Hum Mol Genet 14(5):693-701.
Huang NF, Gupta M, Varghese S, Rao S, Luke S. 2002. Detection of numerical chromosomal abnormalities in epithelial ovarian neoplasms by fluorescence in situ hybridization (FISH) and a review of the current literature. Appl Immunohistochem Mol Morphol 10(2):187-93.
Huebner K, Croce CM. 2001. FRA3B and other common fragile sites: the weakest links. Nat Rev Cancer 1(3):214-21.
Huebner K, Druck T, Siprashvili Z, Croce CM, Kovatich A, McCue PA. 1998a. The role of deletions at the FRA3B/FHIT locus in carcinogenesis. Recent Results Cancer Res 154:200-15.
Huebner K, Garrison PN, Barnes LD, Croce CM. 1998b. The role of the FHIT/FRA3B locus in cancer. Annu Rev Genet 32:7-31.
Hylander B, Repasky E, Shrikant P, Intengan M, Beck A, Driscoll D, Singhal P, Lele S, Odunsi K. 2006. Expression of Wilms tumor gene (WT1) in epithelial ovarian cancer. Gynecol Oncol 101(1):12-7.
Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, Ringner M, Sauter G, Monni O, Elkahloun A and others. 2002. Impact of DNA amplification on gene expression patterns in breast cancer. Cancer Res 62(21):6240-5.
Iliopoulos D, Fabbri M, Druck T, Qin HR, Han SY, Huebner K. 2007. Inhibition of breast cancer cell growth in vitro and in vivo: effect of restoration of Wwox expression. Clin Cancer Res 13(1):268-74.
Iliopoulos D, Guler G, Han SY, Johnston D, Druck T, McCorkell KA, Palazzo J, McCue PA, Baffa R, Huebner K. 2005. Fragile genes as biomarkers: epigenetic control of WWOX and FHIT in lung, breast and bladder cancer. Oncogene 24(9):1625-33.
Isaacson C. 2005. The change of the staple diet of black South Africans from sorghum to maize (corn) is the cause of the epidemic of squamous carcinoma of the oesophagus. Med Hypotheses 64(3):658-60.
Ishii H, Mimori K, Inoue H, Inageta T, Ishikawa K, Semba S, Druck T, Trapasso F, Tani K, Vecchione A and others. 2006. Fhit modulates the DNA damage checkpoint response. Cancer Res 66(23):11287-92.
Ishii H, Furukawa Y. 2004. Alterations of common chromosome fragile sites in hematopoietic malignancies. Int J Hematol 79(3):238-42.
Ishii H, Vecchione A, Furukawa Y, Sutheesophon K, Han SY, Druck T, Kuroki T, Trapasso F, Nishimura M, Saito Y and others. 2003. Expression of FRA16D/WWOX and FRA3B/FHIT genes in hematopoietic malignancies. Mol Cancer Res 1(13):940-7.
Iwaya T, Maesawa C, Ogasawara S, Tamura G. 1998. Tylosis esophageal cancer locus on chromosome 17q25.1 is commonly deleted in sporadic human esophageal cancer. Gastroenterology 114(6):1206-10.
Joos S, Scherthan H, Speicher MR, Schlegel J, Cremer T, Lichter P. 1993. Detection of amplified DNA sequences by reverse chromosome painting using genomic tumor DNA as probe. Hum Genet 90(6):584-9.
Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258(5083):818-21.
Kalnina Z, Zayakin P, Silina K, Line A. 2005. Alterations of pre-mRNA splicing in cancer. Genes Chromosomes Cancer 42(4):342-57.

Kamangar F, Malekzadeh R, Dawsey SM, Saidi F. 2007. Esophageal cancer in Northeastern Iran: a review. Arch Iran Med 10(1):70-82.
Kitamura A, Yashima K, Okamoto E, Andachi H, Hosoda A, Kishimoto Y, Shiota G, Ito H, Kaibara N, Kawasaki H. 2001. Reduced Fhit expression occurs in the early stage of esophageal tumorigenesis: no correlation with p53 expression and apoptosis. Oncology 61(3):205-11.
Knasmuller S, Bresgen N, Kassie F, Mersch-Sundermann V, Gelderblom W, Zohrer E, Eckl PM. 1997. Genotoxic effects of three Fusarium mycotoxins, fumonisin B1, moniliformin and vomitoxin in bacteria and in primary cultures of rat hepatocytes. Mutat Res 391(1-2):39-48.
Knutsen T, Pack S, Petropavlovskaja M, Padilla-Nash H, Knight C, Mickley LA, Ried T, Elwood PC, Roberts SJ. 2003. Cytogenetic, spectral karyotyping, fluorescence in situ hybridization, and comparative genomic hybridization characterization of two new secondary leukemia cell lines with 5q deletions, and MYC and MLL amplification. Genes Chromosomes Cancer 37(3):270-81.
Kobayashi H, Rowley JD. 1995. Identification of cytogenetically undetected 12p13 translocations and associated deletions with fluorescence in situ hybridization. Genes Chromosomes Cancer 12(1):66-9.
Kobayashi H, Satake N, Maseki N, Sakashita A, Kaneko Y. 1996. The der(21)t(12;21) chromosome is always formed in a 12;21 translocation associated with childhood acute lymphoblastic leukaemia. Br J Haematol 94(1):105-11.
Korkola JE, Houldsworth J, Chadalavada RS, Olshen AB, Dobrzynski D, Reuter VE, Bosl GJ, Chaganti RS. 2006. Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors. Cancer Res 66(2):820-7.
Kost-Alimova M, Imreh S. 2007. Modeling non-random deletions in cancer. Semin Cancer Biol 17(1):19-30.
Kuchino Y. 1994. [Tumor suppressor genes]. Nippon Rinsho 52(12):3319-29.
Kuo MT, Sen S, Hittelman WN, Hsu TC. 1998. Chromosomal fragile sites and DNA amplification in drug-resistant cells. Biochem Pharmacol 56(1):7-13.
Kuroki T, Tajima Y, Kanematsu T. 2004. Role of hypermethylation on carcinogenesis in the pancreas. Surg Today 34(12):981-6.
Kuroki T, Trapasso F, Shiraishi T, Alder H, Mimori K, Mori M, Croce CM. 2002. Genetic alterations of the tumor suppressor gene WWOX in esophageal squamous cell carcinoma. Cancer Res 62(8):2258-60.
Kuroki T, Trapasso F, Yendamuri S, Matsuyama A, Alder H, Mori M, Croce CM. 2003. Allele loss and promoter hypermethylation of VHL, RAR-beta, RASSF1A, and FHIT tumor suppressor genes on chromosome 3p in esophageal squamous cell carcinoma. Cancer Res 63(13):3724-8.
Kuwahara Y, Tanabe C, Ikeuchi T, Aoyagi K, Nishigaki M, Sakamoto H, Hoshinaga K, Yoshida T, Sasaki H, Terada M. 2004. Alternative mechanisms of gene amplification in human cancers. Genes Chromosomes Cancer 41(2):125-32.
Kuwano A, Kajii T. 1987. Synergistic effect of aphidicolin and ethanol on the induction of common fragile sites. Hum Genet 75(1):75-8.
La Starza R, Stella M, Testoni N, Di Bona E, Ciolli S, Marynen P, Martelli MF, Mandelli F, Mecucci C. 1999. Characterization of 12p molecular events outside ETV6 in complex karyotypes of acute myeloid malignancies. Br J Haematol 107(2):340-6.
Laubscher ALA. 2000. Characterization of the translocation t(3;11)(p14.2;p14.3-p15.1) present in an ovarian carcinoma cell-line. Mapping derivative 11.

Lee YC, Wu CT, Shih JY, Jou YS, Chang YL. 2004. Frequent allelic deletion at the FHIT locus associated with p53 overexpression in squamous cell carcinoma subtype of Taiwanese non-small-cell lung cancers. Br J Cancer 90(12):2378-83.
Lemmer ER, Vessey CJ, Gelderblom WC, Shephard EG, Van Schalkwyk DJ, Van Wijk RA, Marasas WF, Kirsch RE, Hall Pde L. 2004. Fumonisin B1-induced hepatocellular and cholangiocellular tumors in male Fischer 344 rats: potentiating effects of 2-acetylaminofluorene on oval cell proliferation and neoplastic development in a discontinued feeding study. Carcinogenesis 25(7):1257-64.
Lensch MW, Tischkowitz M, Christianson TA, Reifsteck CA, Speckhart SA, Jakobs PM, O'Dwyer ME, Olson SB, Le Beau MM, Hodgson SV and others. 2003. Acquired FANCA dysfunction and cytogenetic instability in adult acute myelogenous leukemia. Blood 102(1):7-16.
Lerda D, Biaggi Bistoni M, Peralta N, Ychari S, Vazquez M, Bosio G. 2005. Fumonisins in foods from Cordoba (Argentina), presence and genotoxicity. Food Chem Toxicol 43(5):691-8.
Li D, Dandara C, Parker MI. 2005. Association of cytochrome P450 2E1 genetic polymorphisms with squamous cell carcinoma of the oesophagus. Clin Chem Lab Med 43(4):370-5.
Liang G, Miao X, Zhou Y, Tan W, Lin D. 2004. A functional polymorphism in the SULT1A1 gene (G638A) is associated with risk of lung cancer in relation to tobacco smoking. Carcinogenesis 25(5):773-8.
Lin K, Shen W, Shen Z, Wu Y, Lu S. 2002. Dietary exposure and urinary excretion of total N-nitroso compounds, nitrosamino acids and volatile nitrosamine in inhabitants of high- and low-risk areas for esophageal cancer in southern China. Int J Cancer 102(3):207-11.
Liu FX, Huang XP, Zhao CX, Xu X, Han YL, Cai Y, Wu RL, Wu M, Zhan QM, Wang MR. 2004. [Allelic loss and down-regulation of FHIT gene expression in esophageal squamous cell carcinoma]. Ai Zheng 23(9):992-8.
Liu Y, Ganesan TS. 2002. Tumour suppressor genes in sporadic epithelial ovarian cancer. Reproduction 123(3):341-53.
Logan KA, McNatty KP, Juengel JL. 2003. Expression of wilms' tumor gene and protein localization during ovarian formation and follicular development in sheep. Biol Reprod 68(2):635-43.
Looijenga LH, Zafarana G, Grygalewicz B, Summersgill B, Debiec-Rychter M, Veltman J, Schoenmakers EF, Rodriguez S, Jafer O, Clark J and others. 2003. Role of gain of 12p in germ cell tumour development. Apmis 111(1):161-71; discussion 172-3.
Lounis H, Mes-Masson AM, Dion F, Bradley WE, Seymour RJ, Provencher D, Tonin PN. 1998. Mapping of chromosome 3p deletions in human epithelial ovarian tumors. Oncogene 17(18):2359-65.
Lukusa T, Fryns JP. 2007. Human chromosome fragility. Biochim Biophys Acta.
Madebo T, Lindtjorn B, Henriksen TH. 1994. High incidence of oesophagus and stomach cancers in the Bale highlands of south Ethiopia. Trans R Soc Trop Med Hyg 88(4):415.
Mandahl N, Heim S, Johansson B, Bennet K, Mertens F, Olsson G, Rooser B, Rydholm A, Willen H, Mitelman F. 1987. Lipomas have characteristic structural chromosomal rearrangements of 12q13-q14. Int J Cancer 39(6):685-8.
Manderson EN, Presneau N, Provencher D, Mes-Masson AM, Tonin PN. 2002. Comparative analysis of loss of heterozygosity of specific chromosome 3, 13, 17, and X loci and TP53 mutations in human epithelial ovarian cancer. Mol Carcinog 34(2):78-90.

Manning AP, Mes-Masson AM, Seymour RJ, Tetrault M, Provencher DM, Tonin PN. 1999. Expression of FHIT in primary cultures of human epithelial ovarian tumors and malignant ovarian ascites. Mol Carcinog 24(3):218-25.
Mano MS, Rosa DD, De Azambuja E, Ismael GF, Durbecq V. 2007. The 17q12-q21 amplicon: Her2 and topoisomerase-IIalpha and their importance to the biology of solid tumours. Cancer Treat Rev 33(1):64-77.
Marasas WF. 1995. Fumonisins: their implications for human and animal health. Nat Toxins 3(4):193-8; discussion 221.
Marasas WF. 2001. Discovery and occurrence of the fumonisins: a historical perspective. Environ Health Perspect 109 Suppl 2:239-43.
Marasas WF, Riley RT, Hendricks KA, Stevens VL, Sadler TW, Gelineau-van Waes J, Missmer SA, Cabrera J, Torres O, Gelderblom WC and others. 2004. Fumonisins disrupt sphingolipid metabolism, folate transport, and neural tube development in embryo culture and in vivo: a potential risk factor for human neural tube defects among populations consuming fumonisin-contaminated maize. J Nutr 134(4):711-6.
Marasas WF, van Rensburg SJ, Mirocha CJ. 1979. Incidence of Fusarium species and the mycotoxins, deoxynivalenol and zearalenone, in corn produced in esophageal cancer areas in Transkei. J Agric Food Chem 27(5):1108-12.
Marsh D, Zori R. 2002. Genetic insights into familial cancers-- update and recent discoveries. Cancer Lett 181(2):125-64.
Martin-Subero JI, Odero MD, Hernandez R, Cigudosa JC, Agirre X, Saez B, Sanz-Garcia E, Ardanaz MT, Novo FJ, Gascoyne RD and others. 2005. Amplification of IGH/MYC fusion in clinically aggressive IGH/BCL2-positive germinal center B-cell lymphomas. Genes Chromosomes Cancer 43(4):414-23.
Mathew S, Lorsbach RB, Shearer P, Sandlund JT, Raimondi SC. 2000. Double minute chromosomes and c-MYC amplification in a child with secondary myelodysplastic syndrome after treatment for acute lymphoblastic leukemia. Leukemia 14(7):1314-5.
Matsha T, Erasmus R, Kafuko AB, Mugwanya D, Stepien A, Parker MI. 2002. Human papillomavirus associated with oesophageal cancer. J Clin Pathol 55(8):587-90.
Matsha T, Stepien A, Blanco-Blanco E, Brink LT, Lombard CJ, Van Rensburg S, Erasmus RT. 2006. Self-induced vomiting -- risk for oesophageal cancer? S Afr Med J 96(3):209-12.
Matthews CP, Shera K, Kiviat N, McDougall JK. 2001. Expression of truncated FHIT transcripts in cervical cancers and in normal human cells. Oncogene 20(34):4665-75.
Mc CB. 1951. Chromosome organization and genic expression. Cold Spring Harb Symp Quant Biol 16:13-47.
McAvoy S, Ganapathiraju SC, Ducharme-Smith AL, Pritchett JR, Kosari F, Perez DS, Zhu Y, James CD, Smith DI. 2007. Non-random inactivation of large common fragile site genes in different cancers. Cytogenet Genome Res 118(2-4):260-9.
McRonald FE, Liloglou T, Xinarianos G, Hill L, Rowbottom L, Langan JE, Ellis A, Shaw JM, Field JK, Risk JM. 2006. Down-regulation of the cytoglobin gene, located on 17q25, in tylosis with oesophageal cancer (TOC): evidence for trans-allele repression. Hum Mol Genet 15(8):1271-7.
Mertens F, Johansson B, Mitelman F. 1994. Isochromosomes in neoplasia. Genes Chromosomes Cancer 10(4):221-30.

Micci F, Teixeira MR, Bjerkehagen B, Heim S. 2002. Characterization of supernumerary rings and giant marker chromosomes in well-differentiated lipomatous tumors by a combination of G-banding, CGH, M-FISH, and chromosome- and locus-specific FISH. Cytogenet Genome Res 97(1-2):13-9.

Miller CT, Lin L, Casper AM, Lim J, Thomas DG, Orringer MB, Chang AC, Chambers AF, Giordano TJ, Glover TW and others. 2006. Genomic amplification of MET with boundaries within fragile site FRA7G and upregulation of MET pathways in esophageal adenocarcinoma. Oncogene 25(3):409-18.

Mimori K, Druck T, Inoue H, Alder H, Berk L, Mori M, Huebner K, Croce CM. 1999. Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. Proc Natl Acad Sci U S A 96(13):7456-61.

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE and others. 1998. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. Proc Natl Acad Sci U S A 95(14):8141-6.

Moodley R, Reddi A, Chetty R, Naidoo R. 2006. Abnormalities of chromosome 17 in oesophageal cancer. J Clin Pathol.

Morgan NV, Tipping AJ, Joenje H, Mathew CG. 1999. High frequency of large intragenic deletions in the Fanconi anemia group A gene. Am J Hum Genet 65(5):1330-41.

Mori M, Mimori K, Shiraishi T, Alder H, Inoue H, Tanaka Y, Sugimachi K, Huebner K, Croce CM. 2000. Altered expression of Fhit in carcinoma and precarcinomatous lesions of the esophagus. Cancer Res 60(5):1177-82.

Morita M, Oyama T, Nakata S, Ono K, Sugaya M, Uramoto H, Yoshimatsu T, Hanagiri T, Sugio K, Yasumoto K. 2006. Expression of FHIT in esophageal epithelium and carcinoma: reference to drinking, smoking and multicentric carcinogenesis. Anticancer Res 26(3B):2243-8.

Morita M, Saeki H, Mori M, Kuwano H, Sugimachi K. 2002. Risk factors for esophageal cancer and the multiple occurrence of carcinoma in the upper aerodigestive tract. Surgery 131(1 Suppl):S1-6.

Mostert MC, Verkerk AJ, van de Pol M, Heighway J, Marynen P, Rosenberg C, van Kessel AG, van Echten J, de Jong B, Oosterhuis JW and others. 1998. Identification of the critical region of 12p over-representation in testicular germ cell tumors of adolescents and adults. Oncogene 16(20):2617-27.

Mostert MM, van de Pol M, Olde Weghuis D, Suijkerbuijk RF, Geurts van Kessel A, van Echten J, Oosterhuis JW, Looijenga LH. 1996. Comparative genomic hybridization of germ cell tumors of the adult testis: confirmation of karyotypic findings and identification of a 12p-amplicon. Cancer Genet Cytogenet 89(2):146-52.

Musio A, Montagna C, Mariani T, Tilenni M, Focarelli ML, Brait L, Indino E, Benedetti PA, Chessa L, Albertini A and others. 2005. SMC1 involvement in fragile site expression. Hum Mol Genet 14(4):525-33.

Myllykangas S, Himberg J, Bohling T, Nagy B, Hollmen J, Knuutila S. 2006. DNA copy number amplification profiling of human neoplasms. Oncogene 25(55):7324-32.

Nagase H, Mao JH, Balmain A. 1999. A subset of skin tumor modifier loci determines survival time of tumor-bearing mice. Proc Natl Acad Sci U S A 96(26):15032-7.

Naidoo R, Ramburan A, Reddi A, Chetty R. 2005. Aberrations in the mismatch repair genes and the clinical impact on oesophageal squamous carcinomas from a high incidence area in South Africa. J Clin Pathol 58(3):281-4.

Naidoo R, Tarin M, Reddi A, Chetty R. 1999. Allelic imbalance and microsatellite instability in chromosomes 2p, 3p, 5q, and 18q in esophageal squamous carcinoma in patients from South Africa. Diagn Mol Pathol 8(3):131-7.
Noguchi T, Takeno S, Kimura Y, Uchida Y, Daa T, Yokoyama S, Gabbert HE, Mueller W. 2003. FHIT expression and hypermethylation in esophageal squamous cell carcinoma. Int J Mol Med 11(4):441-7.
Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T and others. 1996. The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84(4):587-97.
Ozaki K, Enomoto T, Yoshino K, Fujita M, Buzard GS, Kawano K, Yamasaki M, Murata Y. 2001. Impaired FHIT expression characterizes serous ovarian carcinoma. Br J Cancer 85(2):247-54.
Ozeri-Galai E, Schwartz M, Rahat A, Kerem B. 2007. Interplay between ATM and ATR in the regulation of common fragile site stability. Oncogene.
Pacella-Norman R, Urban MI, Sitas F, Carrara H, Sur R, Hale M, Ruff P, Patel M, Newton R, Bull D and others. 2002. Risk factors for oesophageal, lung, oral and laryngeal cancers in black South Africans. Br J Cancer 86(11):1751-6.
Panani AD, Roussos C. 2006. Non-random structural chromosomal changes in ovarian cancer: i(5p) a novel recurrent abnormality. Cancer Lett 235(1):130-5.
Parrington JM, West LF, Heyderman E. 1994. Chromosome analysis of parallel short-term cultures from four testicular germ-cell tumors. Cancer Genet Cytogenet 75(2):90-102.
Pedersen-Bjergaard J, Rowley JD. 1994. The balanced and the unbalanced chromosome aberrations of acute myeloid leukemia may develop in different ways and may contribute differently to malignant transformation. Blood 83(10):2780-6.
Pedeutour F, Forus A, Coindre JM, Berner JM, Nicolo G, Michiels JF, Terrier P, Ranchere-Vince D, Collin F, Myklebost O and others. 1999. Structure of the supernumerary ring and giant rod chromosomes in adipose tissue tumors. Genes Chromosomes Cancer 24(1):30-41.
Pejovic T, Yates JE, Liu HY, Hays LE, Akkari Y, Torimaru Y, Keeble W, Rathbun RK, Rodgers WH, Bale AE and others. 2006. Cytogenetic instability in ovarian epithelial cells from women at risk of ovarian cancer. Cancer Res 66(18):9017-25.
Penas EM, Cools J, Algenstaedt P, Hinz K, Seeger D, Schafhausen P, Schilling G, Marynen P, Hossfeld DK, Dierlamm J. 2003. A novel cryptic translocation t(12;17)(p13;p12-p13) in a secondary acute myeloid leukemia results in a fusion of the ETV6 gene and the antisense strand of the PER1 gene. Genes Chromosomes Cancer 37(1):79-83.
Pera MF. 2004. Unnatural selection of cultured human ES cells? Nat Biotechnol 22(1):42-3.
Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. 2002. Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31(1):33-6.
Picchio MC, Martin ES, Cesari R, Calin GA, Yendamuri S, Kuroki T, Pentimalli F, Sarti M, Yoder K, Kaiser LR and others. 2004. Alterations of the tumor suppressor gene Parkin in non-small cell lung cancer. Clin Cancer Res 10(8):2720-4.
Pienkowska-Grela B, Grygalewicz B, Bregula U. 2002. Overrepresentation of the short arm of chromosome 12 in seminoma and nonseminoma groups of testicular germ cell tumors. Cancer Genet Cytogenet 134(2):102-8.

Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y and others. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20(2):207-11.
Platzer P, Upender MB, Wilson K, Willis J, Lutterbaugh J, Nosrati A, Willson JK, Mack D, Ried T, Markowitz S. 2002. Silence of chromosomal amplifications in colon cancer. Cancer Res 62(4):1134-8.
Ponder BA, Antoniou A, Dunning A, Easton DF, Pharoah PD. 2005. Polygenic inherited predisposition to breast cancer. Cold Spring Harb Symp Quant Biol 70:35-41.
Popovici C, Basset C, Bertucci F, Orsetti B, Adelaide J, Mozziconacci MJ, Conte N, Murati A, Ginestier C, Charafe-Jauffret E and others. 2002. Reciprocal translocations in breast tumor cell lines: cloning of a t(3;20) that targets the FHIT gene. Genes Chromosomes Cancer 35(3):204-18.
Postma C, Hermsen MA, Coffa J, Baak JP, Mueller JD, Mueller E, Bethke B, Schouten JP, Stolte M, Meijer GA. 2005. Chromosomal instability in flat adenomas and carcinomas of the colon. J Pathol 205(4):514-21.
Poulsen CB, Borup R, Nielsen FC, Borregaard N, Hansen M, Gronbaek K, Moller MB, Ralfkiaer E. 2005. Microarray-based classification of diffuse large B-cell lymphoma. Eur J Haematol 74(6):453-65.
Ramalingam S, Belani CP. 2007. Recent advances in targeted therapy for non-small cell lung cancer. Expert Opin Ther Targets 11(2):245-57.
Ramljak D, Calvert RJ, Wiesenfeld PW, Diwan BA, Catipovic B, Marasas WF, Victor TC, Anderson LM, Gelderblom WC. 2000. A potential mechanism for fumonisin B(1)-mediated hepatocarcinogenesis: cyclin D1 stabilization associated with activation of Akt and inhibition of GSK-3beta activity. Carcinogenesis 21(8):1537-46.
Randolph TR. 2005. Chronic myelocytic leukemia--Part I: History, clinical presentation, and molecular biology. Clin Lab Sci 18(1):38-48.
Rao PH, Harris CP, Yan Lu X, Li XN, Mok SC, Lau CC. 2002. Multicolor spectral karyotyping of serous ovarian adenocarcinoma. Genes Chromosomes Cancer 33(2):123-32.
Raveendranathan M, Chattopadhyay S, Bolon YT, Haworth J, Clarke DJ, Bielinsky AK. 2006. Genome-wide replication profiles of S-phase checkpoint mutants reveal fragile sites in yeast. Embo J 25(15):3627-39.
Raynaud SD, Baens M, Grosgeorge J, Rodgers K, Reid CD, Dainton M, Dyer M, Fuzibet JG, Gratecos N, Taillan B and others. 1996. Fluorescence in situ hybridization analysis of t(3; 12)(q26; p13): a recurring chromosomal abnormality involving the TEL gene (ETV6) in myelodysplastic syndromes. Blood 88(2):682-9.
Raynaud SD, Dastugue N, Zoccola D, Shurtleff SA, Mathew S, Raimondi SC. 1999. Cytogenetic abnormalities associated with the t(12;21): a collaborative study of 169 children with t(12;21)-positive acute lymphoblastic leukemia. Leukemia 13(9):1325-30.
Rechviasvhili NRN. 2002. Characterization of the translocation t(3;11)(p14.2;p14.3-p15.1) present in an ovarian carcinoma cell-line. .
Reilly PA, Heerema NA, Sledge GW, Jr., Palmer CG. 1993. Unusual distribution of chromosome 12 in a testicular germ-cell tumor cell line (833K) and its cisplatin-resistant derivative (64CP9). Cancer Genet Cytogenet 68(2):114-21.
Richards RI. 2001. Fragile and unstable chromosomes in cancer: causes and consequences. Trends Genet 17(6):339-45.

Richards RI, Holman K, Kozman H, Kremer E, Lynch M, Pritchard M, Yu S, Mulley J, Sutherland GR. 1991. Fragile X syndrome: genetic localisation by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. J Med Genet 28(12):818-23.

Roelofs H, Schuuring E, Wiegant J, Michalides R, Giphart-Gassler M. 1993. Amplification of the 11q13 region in human carcinoma cell lines: a mechanistic view. Genes Chromosomes Cancer 7(2):74-84.

Romana SP, Le Coniat M, Poirel H, Marynen P, Bernard O, Berger R. 1996. Deletion of the short arm of chromosome 12 is a secondary event in acute lymphoblastic leukemia with t(12;21). Leukemia 10(1):167-70.

Rowley JD. 1998. The critical role of chromosome translocations in human leukemias. Annu Rev Genet 32:495-519.

Rubio-Moscardo F, Climent J, Siebert R, Piris MA, Martin-Subero JI, Nielander I, Garcia-Conde J, Dyer MJ, Terol MJ, Pinkel D and others. 2005. Mantle-cell lymphoma genotypes identified with CGH to BAC microarrays define a leukemic subgroup of disease and predict patient outcome. Blood 105(11):4445-54.

Ruiz-Herrera A, Garcia F, Fronicke L, Ponsa M, Egozcue J, Caldes MG, Stanyon R. 2004. Conservation of aphidicolin-induced fragile sites in Papionini (Primates) species and humans. Chromosome Res 12(7):683-90.

Sadler TW, Merrill AH, Stevens VL, Sullards MC, Wang E, Wang P. 2002. Prevention of fumonisin B1-induced neural tube defects by folic acid. Teratology 66(4):169-76.

Sambrook J,Russell D.2001. Molecular Cloning Third edition, volume 2, Cold Spring Harbor Laboratory (CSHL) Press.

Santos SC, Cavalli LR, Cavalli IJ, Lima RS, Haddad BR, Ribeiro EM. 2004. Loss of heterozygosity of the BRCA1 and FHIT genes in patients with sporadic breast cancer from Southern Brazil. J Clin Pathol 57(4):374-7.

Sard L, Accornero P, Tornielli S, Delia D, Bunone G, Campiglio M, Colombo MP, Gramegna M, Croce CM, Pierotti MA and others. 1999. The tumor-suppressor gene FHIT is involved in the regulation of apoptosis and in cell cycle control. Proc Natl Acad Sci U S A 96(15):8489-92.

Sawyers CL. 2005. Making progress through molecular attacks on cancer. Cold Spring Harb Symp Quant Biol 70:479-82.

Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 30(12):e57.

Sepehr A, Kamangar F, Fahimi S, Saidi F, Abnet CC, Dawsey SM. 2005. Poor oral health as a risk factor for esophageal squamous dysplasia in northeastern Iran. Anticancer Res 25(1B):543-6.

Sepehr A, Taniere P, Martel-Planche G, Zia'ee AA, Rastgar-Jazii F, Yazdanbod M, Etemad-Moghadam G, Kamangar F, Saidi F, Hainaut P. 2001. Distinct pattern of TP53 mutations in squamous cell carcinoma of the esophagus in Iran. Oncogene 20(50):7368-74.

Shahabi M, Noori Daloii MR, Langan JE, Rowbottom L, Jahanzad E, Khoshbin E, Taghikhani M, Field JK, Risk JM. 2004. An investigation of the tylosis with oesophageal cancer (TOC) locus in Iranian patients with oesophageal squamous cell carcinoma. Int J Oncol 25(2):389-95.

Sitas F, Urban M, Stein L, Beral V, Ruff P, Hale M, Patel M, O'Connell D, Qin Yu X, Verzijden A and others. 2007. The relationship between anti-HPV-16 IgG seropositivity and cancer of the cervix, anogenital organs, oral cavity and pharynx,

oesophagus and prostate in a black South African population. Infect Agent Cancer 2:6.
Skotheim RI, Autio R, Lind GE, Kraggerud SM, Andrews PW, Monni O, Kallioniemi O, Lothe RA. 2006. Novel genomic aberrations in testicular germ cell tumors by array-CGH, and associated gene expression changes. Cell Oncol 28(5-6):315-26.
Smith DI, McAvoy S, Zhu Y, Perez DS. 2007. Large common fragile site genes and cancer. Semin Cancer Biol 17(1):31-41.
Smith SS, Schwarz RE. 2006. Gastric DNA damage through tobacco chewing: in vitro mechanistic studies of DNA nitrite attack. Cancer Lett 235(2):221-8.
Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. 1997. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer 20(4):399-407.
Soma T, Kaganoi J, Kawabe A, Kondo K, Imamura M, Shimada Y. 2006. Nicotine induces the fragile histidine triad methylation in human esophageal squamous epithelial cells. Int J Cancer 119(5):1023-7.
Somdyala NI, Marasas WF, Venter FS, Vismer HF, Gelderblom WC, Swanevelder SA. 2003. Cancer patterns in four districts of the Transkei region--1991-1995. S Afr Med J 93(2):144-8.
Sozzi G, Pastorino U, Moiraghi L, Tagliabue E, Pezzella F, Ghirelli C, Tornielli S, Sard L, Huebner K, Pierotti MA and others. 1998. Loss of FHIT function in lung cancer and preinvasive bronchial lesions. Cancer Res 58(22):5032-7.
Stein CK, Glover TW, Palmer JL, Glisson BS. 2002. Direct correlation between FRA3B expression and cigarette smoking. Genes Chromosomes Cancer 34(3):333-40.
Stevens VL, Tang J. 1997. Fumonisin B1-induced sphingolipid depletion inhibits vitamin uptake via the glycosylphosphatidylinositol-anchored folate receptor. J Biol Chem 272(29):18020-5.
Stockmann-Juvala H, Mikkola J, Naarala J, Loikkanen J, Elovaara E, Savolainen K. 2004. Fumonisin B1-induced toxicity and oxidative damage in U-118MG glioblastoma cells. Toxicology 202(3):173-83.
Stockmann-Juvala H, Naarala J, Loikkanen J, Vahakangas K, Savolainen K. 2006. Fumonisin B1-induced apoptosis in neuroblastoma, glioblastoma and hypothalamic cell lines. Toxicology 225(2-3):234-41.
Strefford JC, Worley H, Barber K, Wright S, Stewart AR, Robinson HM, Bettney G, van Delft FW, Atherton MG, Davies T and others. 2007. Genome complexity in acute lymphoblastic leukemia is revealed by array-based comparative genomic hybridization. Oncogene.
Struski S, Doco-Fenzy M, Koehler M, Chudoba I, Levy F, Masson L, Michel N, Ulrich E, Gruson N, Benard J and others. 2003. Cytogenetic evolution of human ovarian cell lines associated with chemoresistance and loss of tumorigenicity. Anal Cell Pathol 25(3):115-22.
Suijkerbuijk RF, Olde Weghuis DE, Van den Berg M, Pedeutour F, Forus A, Myklebost O, Glier C, Turc-Carel C, Geurts van Kessel A. 1994a. Comparative genomic hybridization as a tool to define two distinct chromosome 12-derived amplification units in well-differentiated liposarcomas. Genes Chromosomes Cancer 9(4):292-5.
Suijkerbuijk RF, Sinke RJ, Meloni AM, Parrington JM, van Echten J, de Jong B, Oosterhuis JW, Sandberg AA, Geurts van Kessel A. 1993. Overrepresentation of chromosome 12p sequences and karyotypic evolution in i(12p)-negative testicular germ-cell tumors revealed by fluorescence in situ hybridization. Cancer Genet Cytogenet 70(2):85-93.

Suijkerbuijk RF, Sinke RJ, Weghuis DE, Roque L, Forus A, Stellink F, Siepman A, van de Kaa C, Soares J, Geurts van Kessel A. 1994b. Amplification of chromosome subregion 12p11.2-p12.1 in a metastasis of an i(12p)-negative seminoma: relationship to tumor progression? Cancer Genet Cytogenet 78(2):145-52.

Suijkerbuijk RF, van de Veen AY, van Echten J, Buys CH, de Jong B, Oosterhuis JW, Warburton DA, Cassiman JJ, Schonk D, Geurts van Kessel A. 1991. Demonstration of the genuine iso-12p character of the standard marker chromosome of testicular germ cell tumors and identification of further chromosome 12 aberrations by competitive in situ hybridization. Am J Hum Genet 48(2):269-73.

Sun G, Wang S, Hu X, Su J, Huang T, Yu J, Tang L, Gao W, Wang JS. 2007. Fumonisin B1 contamination of home-grown corn in high-risk areas for esophageal and liver cancer in China. Food Addit Contam 24(2):181-5.

Szymanska J, Virolainen M, Tarkkanen M, Wiklund T, Asko-Seljavaara S, Tukiainen E, Elomaa I, Blomqvist C, Knuutila S. 1997. Overrepresentation of 1q21-23 and 12q13-21 in lipoma-like liposarcomas but not in benign lipomas: a comparative genomic hybridization study. Cancer Genet Cytogenet 99(1):14-8.

Taetle R, Aickin M, Yang JM, Panda L, Emerson J, Roe D, Adair L, Thompson F, Liu Y, Wisner L and others. 1999. Chromosome abnormalities in ovarian adenocarcinoma: I. Nonrandom chromosome abnormalities from 244 cases. Genes Chromosomes Cancer 25(3):290-300.

Tanaka H, Shimada Y, Harada H, Shinoda M, Hatooka S, Imamura M, Ishizaki K. 2000. Polymorphic variation of the ARP gene on 3p21 in Japanese esophageal cancer patients. Oncol Rep 7(3):591-3.

Tang D, Rundle A, Mooney L, Cho S, Schnabel F, Estabrook A, Kelly A, Levine R, Hibshoosh H, Perera F. 2003. Sulfotransferase 1A1 (SULT1A1) polymorphism, PAH-DNA adduct levels in breast tissue and breast cancer risk in a case-control study. Breast Cancer Res Treat 78(2):217-22.

Tanimoto K, Hayashi S, Tsuchiya E, Tokuchi Y, Kobayashi Y, Yoshiga K, Okui T, Kobayashi M, Ichikawa T. 2000. Abnormalities of the FHIT gene in human oral carcinogenesis. Br J Cancer 82(4):838-43.

Thorland EC, Myers SL, Gostout BS, Smith DI. 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. Oncogene 22(8):1225-37.

Thorland EC, Myers SL, Persing DH, Sarkar G, McGovern RM, Gostout BS, Smith DI. 2000. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. Cancer Res 60(21):5916-21.

Tischkowitz MD, Morgan NV, Grimwade D, Eddy C, Ball S, Vorechovsky I, Langabeer S, Stoger R, Hodgson SV, Mathew CG. 2004. Deletion and reduced expression of the Fanconi anemia FANCA gene in sporadic acute myeloid leukemia. Leukemia 18(3):420-5.

Toledo F, Coquelle A, Svetlova E, Debatisse M. 2000. Enhanced flexibility and aphidicolin-induced DNA breaks near mammalian replication origins: implications for replicon mapping and chromosome fragility. Nucleic Acids Res 28(23):4805-13.

Toledo F, Le Roscouet D, Buttin G, Debatisse M. 1992. Co-amplified markers alternate in megabase long chromosomal inverted repeats and cluster independently in interphase nuclei at early steps of mammalian gene amplification. Embo J 11(7):2665-73.

Torres VA, Tapia JC, Rodriguez DA, Lladser A, Arredondo C, Leyton L, Quest AF. 2007. E-cadherin is required for caveolin-1-mediated down-regulation of the inhibitor of apoptosis protein survivin via reduced beta-catenin-Tcf/Lef-dependent transcription. Mol Cell Biol 27(21):7703-17.

Trapasso F, Krakowiak A, Cesari R, Arkles J, Yendamuri S, Ishii H, Vecchione A, Kuroki T, Bieganowski P, Pace HC and others. 2003. Designed FHIT alleles establish that Fhit-induced apoptosis in cancer cells is limited by substrate binding. Proc Natl Acad Sci U S A 100(4):1592-7.

Tremblay S, Pintor Dos Reis P, Bradley G, Galloni NN, Perez-Ordonez B, Freeman J, Brown D, Gilbert R, Gullane P, Irish J and others. 2006. Young patients with oral squamous cell carcinoma: study of the involvement of GSTP1 and deregulation of the Fanconi anemia genes. Arch Otolaryngol Head Neck Surg 132(9):958-66.

Tsantoulis PK, Kotsinas A, Sfikakis PP, Evangelou K, Sideridou M, Levy B, Mo L, Kittas C, Wu XR, Papavassiliou AG and others. 2007. Oncogene-induced replication stress preferentially targets common fragile sites in preneoplastic lesions. A genome-wide study. Oncogene.

Turner BC, Ottey M, Zimonjic DB, Potoczek M, Hauck WW, Pequignot E, Keck-Waggoner CL, Sevignani C, Aldaz CM, McCue PA and others. 2002. The fragile histidine triad/common chromosome fragile site 3B locus and repair-deficient cancers. Cancer Res 62(14):4054-60.

Tyybakinoja A, Saarinen-Pihkala U, Elonen E, Knuutila S. 2006. Amplified, lost, and fused genes in 11q23-25 amplicon in acute myeloid leukemia, an array-CGH study. Genes Chromosomes Cancer 45(3):257-64.

Tzao C, Sun GH, Tung HJ, Hsu HS, Hsu WH, Wang YC, Cheng YL, Lee SC. 2006. Reduced acetylated histone H4 is associated with promoter methylation of the fragile histidine triad gene in resected esophageal squamous cell carcinoma. Ann Thorac Surg 82(2):396-401; discussion 401.

Valk PJ, Delwel R, Lowenberg B. 2005. Gene expression profiling in acute myeloid leukemia. Curr Opin Hematol 12(1):76-81.

van Heerden WF, Swart TJ, Robson B, Smith TL, Engelbrecht S, van Heerden MB, van Rensburg EJ, Huebner K. 2001a. FHIT RNA and protein expression in oral squamous cell carcinomas. Anticancer Res 21(4A):2425-8.

van Heerden WF, Swart TJ, van Heerden MB, Pekarsky Y, Sutherland R, Huebner K. 2001b. Fhit protein expression in oral epithelium: immunohistochemical evaluation of three antisera. Anticancer Res 21(4A):2419-23.

van Rensburg SJ, van Schalkwyk GC, van Schalkwyk DJ. 1990. Primary liver cancer and aflatoxin intake in Transkei. J Environ Pathol Toxicol Oncol 10(1-2):11-6.

Verkasalo PK, Kaprio J, Pukkala E, Koskenvuo M. 1999. Breast cancer risk in monozygotic and dizygotic female twins: a 20-year population-based cohort study in Finland from 1976 to 1995. Cancer Epidemiol Biomarkers Prev 8(3):271-4.

Victor T, Du Toit R, Jordaan AM, Bester AJ, van Helden PD. 1990. No evidence for point mutations in codons 12, 13, and 61 of the ras gene in a high-incidence area for esophageal and gastric cancers. Cancer Res 50(16):4911-4.

Vieira L, Marques B, Cavaleiro C, Ambrosio AP, Jorge M, Neto A, Costa JM, Junior EC, Boavida MG. 2005. Molecular cytogenetic characterization of rearrangements involving 12p in leukemia. Cancer Genet Cytogenet 157(2):134-9.

Viguie F, Aboura A, Bouscary D, Guesnu M, Baumelou E, Dreyfus F, Casadevall N, Tachdjian G. 2002. Isodicentric/pseudoisodicentric chromosome 21 amplification in four cases of acute myelocytic leukemia or myelodysplasia. Cancer Genet Cytogenet 138(1):80-4.

Vincent-Salomon A, MacGrogan G, Couturier J, Arnould L, Denoux Y, Fiche M, Jacquemier J, Mathieu MC, Penault-Llorca F, Rigaud C and others. 2003. Calibration of immunohistochemistry for assessment of HER2 in breast cancer: results of the French multicentre GEFPICS study. Histopathology 42(4):337-47.
Vogelstein B, Kinzler KW. 2004. Cancer genes and the pathways they control. Nat Med 10(8):789-99.

Vogt N, Lefevre SH, Apiou F, Dutrillaux AM, Cor A, Leuraud P, Poupon MF, Dutrillaux B, Debatisse M, Malfoy B. 2004. Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. Proc Natl Acad Sci U S A 101(31):11368-73.
Vorstman JA, Jalali GR, Rappaport EF, Hacker AM, Scott C, Emanuel BS. 2006. MLPA: a rapid, reliable, and sensitive method for detection and analysis of abnormalities of 22q. Hum Mutat.
Wang JM, Xu B, Rao JY, Shen HB, Xue HC, Jiang QW. 2007. Diet habits, alcohol drinking, tobacco smoking, green tea drinking, and the risk of esophageal squamous cell carcinoma in the Chinese population. Eur J Gastroenterol Hepatol 19(2):171-6.
Wang ND, Testa JR, Smith DI. 1993. Determination of the specificity of aphidicolin-induced breakage of the human 3p14.2 fragile site. Genomics 17(2):341-7.
Wang Z, Tang L, Sun G, Tang Y, Xie Y, Wang S, Hu X, Gao W, Cox SB, Wang JS. 2006. Etiological study of esophageal squamous cell carcinoma in an endemic region: a population-based case control study in Huaian, China. BMC Cancer 6:287.
Weiss B, Vora A, Huberty J, Hawkins RA, Matthay KK. 2003. Secondary myelodysplastic syndrome and leukemia following 131I-metaiodobenzylguanidine therapy for relapsed neuroblastoma. J Pediatr Hematol Oncol 25(7):543-7.
Wentzensen N, Vinokurova S, von Knebel Doeberitz M. 2004. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. Cancer Res 64(11):3878-84.
Wilson AP, Dent M, Pejovic T, Hubbold L, Radford H. 1996. Characterisation of seven human ovarian tumour cell lines. Br J Cancer 74(5):722-7.
Wreesmann VB, Estilo C, Eisele DW, Singh B, Wang SJ. 2007. Downregulation of fanconi anemia genes in sporadic head and neck squamous cell carcinoma. ORL J Otorhinolaryngol Relat Spec 69(4):218-25.
Wu M, Wallace MR, Muir D. 2005. Tumorigenic properties of neurofibromin-deficient Schwann cells in culture and as syngrafts in Nf1 knockout mice. J Neurosci Res 82(3):357-67.
Wu MT, Wang YT, Ho CK, Wu DC, Lee YC, Hsu HK, Kao EL, Lee JM. 2003. SULT1A1 polymorphism and esophageal cancer in males. Int J Cancer 103(1):101-4.
Xie Y, de Winter JP, Waisfisz Q, Nieuwint AW, Scheper RJ, Arwert F, Hoatlin ME, Ossenkoppele GJ, Schuurhuis GJ, Joenje H. 2000. Aberrant Fanconi anaemia protein profiles in acute myeloid leukaemia cells. Br J Haematol 111(4):1057-64.
Yakicier MC, Legoix P, Vaury C, Gressin L, Tubacher E, Capron F, Bayer J, Degott C, Balabaud C, Zucman-Rossi J. 2001. Identification of homozygous deletions at chromosome 16q23 in aflatoxin B1 exposed hepatocellular carcinoma. Oncogene 20(37):5232-8.
Yang CX, Matsuo K, Ito H, Shinoda M, Hatooka S, Hirose K, Wakai K, Saito T, Suzuki T, Maeda T and others. 2005a. Gene-environment interactions between alcohol

drinking and the MTHFR C677T polymorphism impact on esophageal cancer risk: results of a case-control study in Japan. Carcinogenesis 26(7):1285-90.
Yang CX, Matsuo K, Wang ZM, Tajima K. 2005b. Phase I/II enzyme gene polymorphisms and esophageal cancer risk: a meta-analysis of the literature. World J Gastroenterol 11(17):2531-8.
Yang YG, Herceg Z, Nakanishi K, Demuth I, Piccoli C, Michelon J, Hildebrand G, Jasin M, Digweed M, Wang ZQ. 2005c. The Fanconi anemia group A protein modulates homologous repair of DNA double-strand breaks in mammalian cells.Carcinogenesis 26(10):1731-40.
Yao PF, Li GC, Li J, Xia HS, Yang XL, Huang HY, Fu YG, Wang RQ, Wang XY, Sha JW. 2006. Evidence of human papilloma virus infection and its epidemiology in esophageal squamous cell carcinoma. World J Gastroenterol 12(9):1352-5.
Yokoyama A, Kato H, Yokoyama T, Igaki H, Tsujinaka T, Muto M, Omori T, Kumagai Y, Yokoyama M, Watanabe H. 2006. Esophageal squamous cell carcinoma and aldehyde dehydrogenase-2 genotypes in Japanese females. Alcohol Clin Exp Res 30(3):491-500.
Yokoyama T, Saito K, Lwin H, Yoshiike N, Yamamoto A, Matsushita Y, Date C, Tanaka H. 2005. Epidemiological evidence that acetaldehyde plays a significant role in the development of decreased serum folate concentration and elevated mean corpuscular volume in alcohol drinkers. Alcohol Clin Exp Res 29(4):622-30.
Yunis JJ. 1988. Fragile sites, mutagens and genomic rearrangements in cancer. Basic Life Sci 43:11-21.
Yunis JJ. 1990. Fragile sites, genomic rearrangements and cancer predisposition. Birth Defects Orig Artic Ser 26(1):67-83.
Yunis JJ, Soreng AL, Bowe AE. 1987. Fragile sites are targets of diverse mutagens and carcinogens. Oncogene 1(1):59-69.
Zaahl MG, Warnich L, Victor TC, Kotze MJ. 2005. Association of functional polymorphisms of SLC11A1 with risk of esophageal cancer in the South African Colored population. Cancer Genet Cytogenet 159(1):48-52.
Zafarana G, Grygalewicz B, Gillis AJ, Vissers LE, van de Vliet W, van Gurp RJ, Stoop H, Debiec-Rychter M, Oosterhuis JW, van Kessel AG and others. 2003. 12p-amplicon structure analysis in testicular germ cell tumors of adolescents and adults by array CGH. Oncogene 22(48):7695-701.
Zanesi N, Fidanza V, Fong LY, Mancini R, Druck T, Valtieri M, Rudiger T, McCue PA, Croce CM, Huebner K. 2001. The tumor spectrum in FHIT-deficient mice. Proc Natl Acad Sci U S A 98(18):10250-5.
Zanesi N, Mancini R, Sevignani C, Vecchione A, Kaou M, Valtieri M, Calin GA, Pekarsky Y, Gnarra JR, Croce CM and others. 2005. Lung cancer susceptibility in Fhit-deficient mice is increased by Vhl haploinsufficiency. Cancer Res 65(15):6576-82.
Zelent A, Greaves M, Enver T. 2004. Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia. Oncogene 23(24):4275-83.
Zhang XJ, Guo W, Wang N, Zhou RM, Dong XJ, Li Y. 2006. [The association of MMP-13 polymorphism with susceptibility to esophageal squamous cell carcinoma and gastric cardiac adeno-carcinoma]. Yi Chuan 28(12):1500-4.
Zheng H, Tsuneyama K, Takahashi H, Miwa S, Nomoto K, Saito H, Masuda S, Takano Y. 2007. Expression of PTEN and FHIT is involved in regulating the balance between apoptosis and proliferation in lung carcinomas. Anticancer Res 27(1B):575-81.

Zheng S, Ma X, Zhang L, Gunn L, Smith MT, Wiemels JL, Leung K, Buffler PA, Wiencke JK. 2004. Hypermethylation of the 5' CpG island of the FHIT gene is associated with hyperdiploid and translocation-negative subtypes of pediatric leukemia. Cancer Res 64(6):2000-6.

Zhu Y, Ghosh P, Charnay P, Burns DK, Parada LF. 2002. Neurofibromas in NF1Schwann cell origin and role of tumor environment. Science 296(5569):920-2.

Zuna J, Ford AM, Peham M, Patel N, Saha V, Eckert C, Kochling J, Panzer-Grumayer R, Trka J, Greaves M. 2004. TEL deletion analysis supports a novel view of relapse in childhood acute lymphoblastic leukemia. Clin Cancer Res 10(16):5355-60.

APPENDICES

Appendix 1: list of fragile sites in the human genome

Modified from From NCBI : http://www.ncbi.nlm.nih.gov/

Genome Data Base reference number	Chromosomal sub band		Name	Agent
GDB:119141	10q21		FRA10C	BrdU type
GDB:119142	10q22.1		FRA10D	Aphidicolin type
GDB:119143	10q25.2		FRA10E	Aphidicolin type
GDB:119923	10q26.1		FRA10F	Aphidicolin type
GDB:119145	11q23.3		FRA11B	
GDB:119146	11p15.1		FRA11C	Aphidicolin type
GDB:119147	11p14.2		FRA11D	Aphidicolin type
GDB:119148	11p13		FRA11E	Aphidicolin type
GDB:119149	11q14.2		FRA11F	Aphidicolin type
GDB:119150	11q23.3		FRA11G	Aphidicolin type
GDB:119924	11q13		FRA11H	Aphidicolin type
GDB:119152	12q21.3		FRA12B	Aphidicolin type
GDB:119154	12q24		FRA12E	Aphidicolin type
GDB:119155	13q13.2		FRA13A	Aphidicolin type
GDB:119156	13q21		FRA13B	BrdU type
GDB:119157	13q21.2		FRA13C	Aphidicolin type
GDB:119927	13q32		FRA13D	Aphidicolin type
GDB:119158	14q23		FRA14B	Aphidicolin type
GDB:119159	14q24.1		FRA14C	Aphidicolin type
GDB:119160	15q22		FRA15A	Aphidicolin type
GDB:119162	16q22.1		FRA16C	Aphidicolin type
GDB:119163	16q23.2		FRA16D	Aphidicolin type
GDB:119930	17q23.1		FRA17B	Aphidicolin type
GDB:119165	18q12.2		FRA18A	Aphidicolin type
GDB:119166	18q21.3		FRA18B	Aphidicolin type
GDB:119931	19q13		FRA19A	5-Azacytidine type
GDB:119168	1p36		FRA1A	Aphidicolin type
GDB:119169	1p32		FRA1B	Aphidicolin type
GDB:119170	1p31.2		FRA1C	Aphidicolin type
GDB:119171	1p22		FRA1D	Aphidicolin type
GDB:119172	1p21.2		FRA1E	Aphidicolin type
GDB:119932	1q21		FRA1F	Aphidicolin type
GDB:119173	1q25.1		FRA1G	Aphidicolin type
GDB:119174	1q42		FRA1H	5-Azacytidine type
GDB:119933	1q44		FRA1I	Aphidicolin type
GDB:119934	1q12		FRA1J	5-Azacytidine type
GDB:119935	1q31		FRA1K	Aphidicolin type
GDB:119936	1p31		FRA1L	Aphidicolin type
GDB:119176	20p12.2		FRA20B	Aphidicolin type
GDB:119178	22q12.2		FRA22B	Aphidicolin type
GDB:119181	2p24.2		FRA2C	Aphidicolin type

GDB:119182	2p16.2		FRA2D	Aphidicolin type
GDB:119183	2p13		FRA2E	Aphidicolin type
GDB:119184	2q21.3		FRA2F	Aphidicolin type
GDB:119185	2q31		FRA2G	Aphidicolin type
GDB:119186	2q32.1		FRA2H	Aphidicolin type
GDB:119187	2q33		FRA2I	Aphidicolin type
GDB:119188	2q37.3		FRA2J	Aphidicolin type
GDB:119190	3p24.2		FRA3A	Aphidicolin type
GDB:119191	3p14.2		FRA3B	Aphidicolin type
GDB:119192	3q27		FRA3C	Aphidicolin type
GDB:119937	3q25		FRA3D	Aphidicolin type
GDB:119193	4p16.1		FRA4A	Aphidicolin type
GDB:119194	4q12		FRA4B	BrdU type
GDB:119195	4q31.1		FRA4C	Aphidicolin type
GDB:119196	4p15		FRA4D	Aphidicolin type
GDB:119197	4q27		FRA4E	Unclassified
GDB:119198	5p13		FRA5A	BrdU type
GDB:119199	5q15		FRA5B	BrdU type
GDB:119200	5q31.1		FRA5C	Aphidicolin type
GDB:119938	5q15		FRA5D	Aphidicolin type
GDB:119939	5p14		FRA5E	Aphidicolin type
GDB:119940	5q21		FRA5F	Aphidicolin type
GDB:119202	6p25.1		FRA6B	Aphidicolin type
GDB:119203	6p22.2		FRA6C	Aphidicolin type
GDB:119204	6q13		FRA6D	BrdU type
GDB:119941	6q26		FRA6E	Aphidicolin type
GDB:119942	6q21		FRA6F	Aphidicolin type
GDB:119943	6q15		FRA6G	Aphidicolin type
GDB:119206	7p22		FRA7B	Aphidicolin type
GDB:119207	7p14.2		FRA7C	Aphidicolin type
GDB:119208	7p13		FRA7D	Aphidicolin type
GDB:119209	7q21.2		FRA7E	Aphidicolin type
GDB:119210	7q22		FRA7F	Aphidicolin type
GDB:119211	7q31.2		FRA7G	Aphidicolin type
GDB:119212	7q32.3		FRA7H	Aphidicolin type
GDB:119213	7q36		FRA7I	Aphidicolin type
GDB:119944	7q11		FRA7J	Aphidicolin type
GDB:119215	8q22.1		FRA8B	Aphidicolin type
GDB:119216	8q24.1		FRA8C	Aphidicolin type
GDB:119217	8q24.3		FRA8D	Aphidicolin type
GDB:119222	9p21		FRA9C	BrdU type
GDB:119945	9q22.1		FRA9D	Aphidicolin type
GDB:119223	9q32		FRA9E	Aphidicolin type
GDB:119946	9q12		FRA9F	5-Azacytidine type
GDB:119948	Xp22.31		FRAXB	Aphidicolin type
GDB:119949	Xq22.1		FRAXC	Aphidicolin type
GDB:119950	Xq27.2		FRAXD	Aphidicolin type

Appendix 2: MLPA Probe binding sequences:

FHIT exon 1
GAAGCTCAGGAAAGAAGAGAAATCCACTGA-– GAACAGTCTGTAAAG
GTAAGTGTATTTATTACATGGCCTTGAC

FHIT exon 2
CTGACAGGACAACTTCCTTAACCACCT–- GCTTGTCCTCCCTCTTAGGCA
AAAGTGCTGAGCAATAGTT

FHIT exon 3
GTCTCCCATCTTGTGTTTCTGCCTTTCCCT-– AGGTCCGTAGTGCTATCTA
CATCCAGACGGTGGAAGGGAG

FHIT exon 3
CAATCCTGTTGTCCTTTAACTTCCACTTT–-ACTGCCCTACAAGATGTCCC
AGGTAGGTTATATGCTTTTAA

FHIT exon 4
GTAGGTATCCTAGGAATACCTGCCTGCTTA–-GACCCTCTATAAAAGCTC
TGTGCATCCTGCCACTGAGGAC

FHIT exon 4
GCTCAGTATTCATGGGGAACGTTCCTATG –- AATTGTACCTAGAGCCATC
T GGATCACTGCTGAAATAATCAT

FHIT exon 5
GACTTCAACTGTGAGGACATGTCG-– TTCAGATTTCCGCACATCTCATCA
AGCCCTCTGTACTGTTTC

FHIT exon 5
CAGAACTGTCCTTCGCTCTTGTG-– AATAGGAAACCTGTGGTACCATGAC
GTATCCTTTCTTT

FHIT exon 6
CTGCGGCCAGTGGAGCGCTTCCA-– TGACCTGCGTCCTGATGAAGTGGCC
GATTTGTTTCAGA

FHIT exon 7
GGGTGATATTTCGAATAAGCAGGAGATACCATCTACCTAT–- TCTCACCTC
TG TTGGTTTTTTCCTACTTCTAGGATGGCCCCGAA

FHIT exon 8
GCCTCTGAAAAATAACCCAGCAGGAACCATCG-– TGGAGTAATTGGGCT
TCATGAGAGCATCACTGTCAAGTCTA

FHIT exon 8
CTATCCTGTGAAAGTCTCCAGCCTTC-– CTGGGAAGAACATGGACGTGAA
CGTGCTGAAATGTACAAG

FHIT exon 9
GGAGGACTTTCCTGCCTCTTGGA–- GATCAGAGGAGGAAATGGCAGCAG
AAGCCGCAGCTCTG

FHIT exon 10
GTATTCTTCAACCAGTGATCCCCCACCTC–- GGTCACTCCAACTCCCTTAA
AAT CCTAGACCTAAACGGCT

WWOX exon 1
CAGGAACTCACTCCCGCTCCAAACGA–- GACCCGCGCGCGTCGGGGC
CCGACCGCTCACGCCT

WWOX exon 1
CCATGGCAGCGCTGCGCTACG–- CGGGGCTGGACGACACGGACAGTGAG
GACGAGCTGCC

WWOX exon 4
CTTGGACCCAAGACTGGCGTTTACTGT-– GGATGATAATCCGACCAAGCC
AACCACCCGGCAAAGATACGAC

WWOX exon 5
CCAAGTCTTTTGCCCTCCATGGTGCACAT-– GTGATCTTGGCCTGCAGGAA
CATGGCAAGGGCGAGTGAGC

WWOX exon 6
GACCCTGGACCTCGCTCTGCTCCGTA-– CGCTGCAGCATTTTGCTGAAGC
ATTCAAGGCCAAGAATGTG

WWOX exon 7
GTCTCACCAAAGATGGCCTGGAGACC-– ACCTTTCAAGTGAATCATCTGG
GGCACTTCTACCCTGTCCAGCT

WWOX exon 9
GCAACAGGGAGCTGCCACCACCGTGT–- ACTGTGCTGCTGTCCCAGAACT
GGAGGGTCTGGGAGGGATG

WWOX exon 9
CGAGAGGCTGATCCAAGAACGGCT–- TGGCAGCCAGTCCGGCTAAGTGG
AGCTCAGAGCGGATGGG

Appendix 3: FHIT coding sequence and set of primers

Coding sequence:

CTTTTTgCCCTCTgTTCCCgggTCCCTCAggCggCCACCCAgTgggCACACTCCC
----------F1
AggCggCgCTCCggCCCCgCgCTCCCTCCCTCTgCCTTTCATTCCCAgCTgTCAA
R1
CATCCTggAAgCTTTgAAgCTCAggAAAgAAgAgAAATCCACTgAgAACAgTCT

GTAAAggTCCgTAgTgCTATCTACATCCAgACggTggAAgggAgAgAAAgAgAAA

GAAggTATCCTAggAATACCTgCCTgCTTAgACCCTCTATAAAAgCTCTgTgCA

TCCTgCCACTgAggACTCCgAAgAggTAgCAgTCTTCTgAAAgACTTCAACTgTg
*
AGgACATgTCgTTCAgATTTggCCAACATCTCATCTCATCAAgCCCTCTgTAgT
*
F5
gTTTCTCAAAACAgAACTgTCCTTCgCTCTTgTgAATAggAAACCTgTggTACC
R5
AgGACATgTCCTTgTgTgCCCgCTgCggCCAgTggAgCgCTTCCATgACCTgCgTC

CTgATgAAgTggCCgATTTgTTTCAgACgACCCAgAgAgTCgggACAGTggTggAA

AAACATTTCCATgggACCTCTCTCACCTTTTCCATgCAggATggCCCCgAAgCC

ggACAgACTgTgAAgCACgTTCACgTCCATgTTCTTCCCAggAAggCTggAgACT

TTCACAggAATgACAgCATCTATgAggAgCTCCAgAAACATgACAAggAggACT

TTCCTgCCTCTTggAgATCAgAggAggAAATggCAgCAgAAgCCgCAgCTCTgCgg

GTCTACTTTCAgTgACACAgATgTTTTTCAgATCCTgAATTCCAgCAAAAgAgC
F10
TATTgCCAACCAgTTTgAAgACCgCCCCCCgCCTCTCCCCAAgAggAAgTgAAT

CAgCATgAAAATgCAgTTTCTTCATTTTTTTCTCACCATCCTgTATTCTTCAA
R10c
CCAgTgATCCCCCACCTCggTCACTCCAACTCCCTTAAAATACCTAgACCTA

AACggCTCAgACAggCAgATTTgAggTTTCCCCCTgTCTCCTTATTCggCAgCCT

TATgATTAAACTTCCTTCTCTgCTgCAAAAAAAAAAAAAAA
R10a

Set of Primer (from Fang, et al. 2001)

FHIT full transcript

F 1 **5'-** CTTTTTGCCCTCTGTTCCCG **-3'**

R 10c **5'-** TCACTGGTTGAAGAATACAGG **-3'**

FHIT exon1

F 1 **5'-** CTTTTTGCCCTCTGTTCCCG **-3'**

R 1 **5'-** GGATGTTGACAGCTGGGAATG **-3'**

FHIT exon 5

F 5 **5'-** ATGTCGTTCAGATTTGGCC **-3'**

R 5 **5'-** CTGGTACCACAGGTTTCCTA **-3'**

FHIT exon10

F 10 **5'-** CCTCTTAAGGTCGTTTTCTC **-3'**

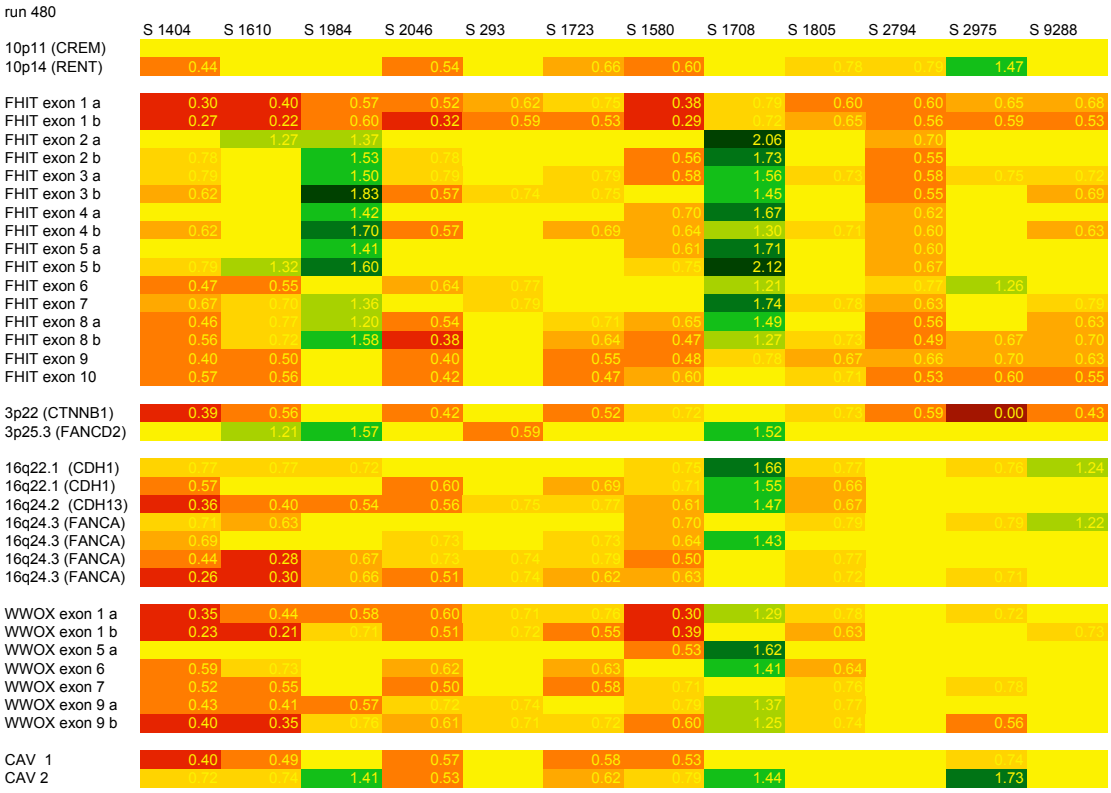
R 10a **5'-** GCAGCAGAGAAG GAAGTTTA **-3'**

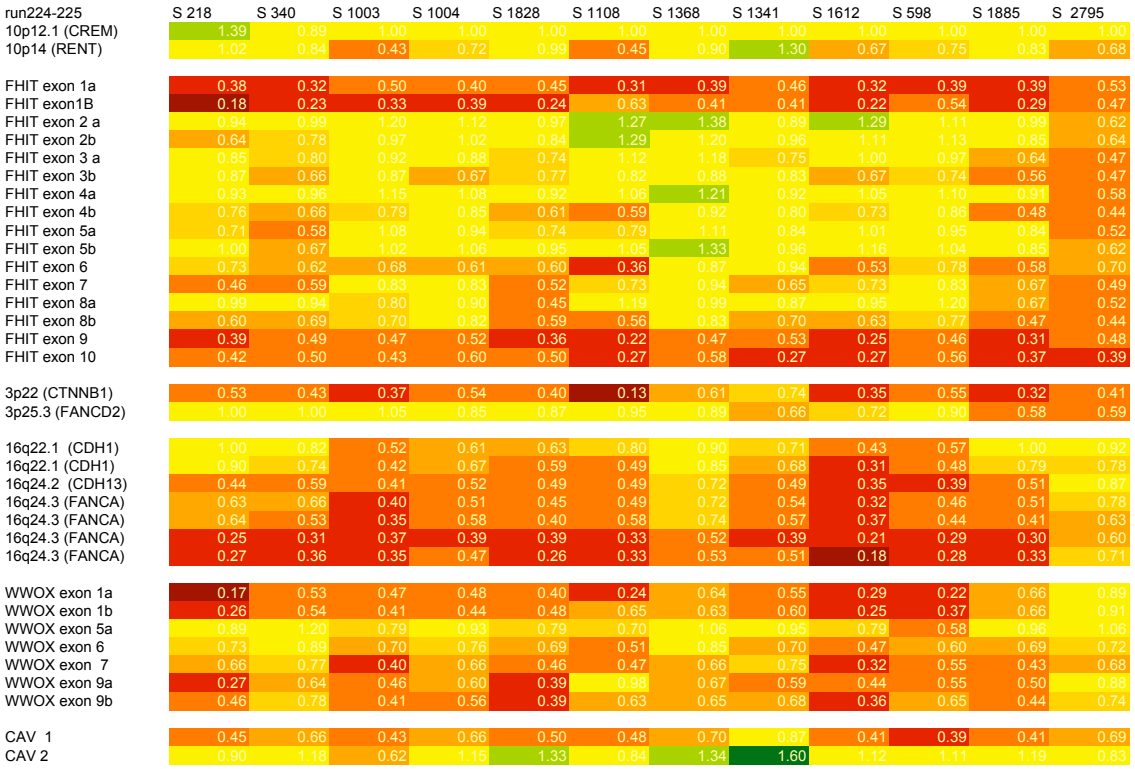
2. Other primers:

F2 (nested RACE PCR primer in exon 1):

5'- CATCCTGGAAGCTTTGAAGCTCA **-3'**

Appendix 4. Heat maps representing MLPA probes ratios for the Eastern Cape cohort





run 317	S 1051	S 220	S 1105	S 257	S 1635	S 912	S 1918
10p12.1 (CREM)	0.89	1.00	1.00	1.00	1.00	1.00	1.00
10p14 (RENT)	1.50	1.58	1.87	1.50	1.48	0.50	0.55
FHIT exon 1a	0.48	0.64	0.24	0.62	0.48	0.76	0.58
FHIT exon1B	0.53	0.95	0.38	0.54	0.51	0.38	0.56
FHIT exon 2 a	0.94	0.91	0.74	1.05	0.93	1.69	1.23
FHIT exon 2b	1.02	1.06	0.81	1.16	1.68	1.41	0.72
FHIT exon 3 a	0.85	0.83	0.88	0.81	1.33	1.42	0.73
FHIT exon 3b	0.84	0.83	0.84	0.96	1.07	1.02	0.67
FHIT exon 4a	0.75	0.87	0.74	0.89	1.10	1.95	1.12
FHIT exon 4b	1.04	0.93	0.97	0.89	1.44	1.07	1.02
FHIT exon 5a	0.78	0.81	0.67	0.83	1.05	0.95	0.81
FHIT exon 5b	0.86	0.94	0.79	1.05	1.55	0.98	0.84
FHIT exon 6	0.56	1.02	0.61	0.85	0.59	0.59	0.60
FHIT exon 7	0.67	0.82	0.53	0.88	1.18	0.70	0.49
FHIT exon 8a	0.85	1.18	0.52	0.89	1.38	0.52	0.65
FHIT exon 8b	0.70	1.10	0.81	0.97	1.18	0.60	0.56
FHIT exon 9	0.70	1.03	0.88	0.79	0.64	0.48	0.25
FHIT exon 10	0.82	1.00	0.63	0.88	1.02	0.58	0.32
3p22 (CTNNB1)	1.00	1.05	0.83	0.85	1.26	0.40	0.37
3p25.3 (FANCD2)	1.00	1.07	0.70	1.07	1.01	0.78	0.92
16q22.1 (CDH1)	0.64	1.05	0.95	1.35	0.62	0.95	0.90
16q22.1 (CDH1)	0.92	1.32	1.12	1.32	0.60	0.53	0.48
16q24.2 (CDH13)	0.63	1.02	0.51	1.13	0.56	0.56	0.71
16q24.3 (FANCA)	0.53	1.28	0.72	1.41	0.51	0.67	0.84
16q24.3 (FANCA)	0.48	1.30	0.98	1.32	0.51	0.51	0.67
16q24.3 (FANCA)	0.46	1.27	0.51	1.10	0.34	0.41	0.36
16q24.3 (FANCA)	0.63	1.56	0.56	1.27	0.49	0.37	0.31
WWOX exon 1a	0.64	1.09	0.49	1.04	0.57	0.51	0.57
WWOX exon 1b	0.57	1.16	0.70	1.07	0.57	0.31	0.29
WWOX E 5A	0.77	1.32	0.97	1.46	1.00	0.92	0.74
WWOX exon 6	0.82	1.24	0.97	1.39	1.10	0.80	0.56
WWOX exon 7	0.77	1.29	0.87	1.30	0.79	0.46	0.41
WWOX exon 9a	0.55	1.13	0.52	1.20	0.62	0.57	0.92
WWOX exon 9b	0.66	1.24	0.74	1.16	0.67	0.44	0.44
CAV 1	0.66	1.48	1.03	1.34	0.88	0.47	0.46
CAV 2	2.28	1.65	2.02	2.21	2.60	0.63	0.87

Run 361	S 2789	S 2790	S 2791	S 263	S 2793	S 2139	S 60	S 432
10p12.1 (CREM)	1.00	1.00	1.23	0.84	1.00	1.00	1.00	1.00
10p14 (RENT)	1.02	0.96	1.04	0.76	1.13	0.65	0.71	0.55
FHIT exon 1a	0.70	0.89	0.66	0.35	0.65	0.38	0.79	0.50
FHIT exon1B	0.54	0.51	0.48	0.18	0.43	0.32	0.47	0.30
FHIT exon 2 a	0.91	1.03	1.07	1.10	1.03	0.82	1.02	0.79
FHIT exon 2b	0.69	0.77	0.76	0.77	0.76	0.51	0.88	0.58
FHIT exon 3 a	0.67	0.87	0.82	0.95	0.86	0.58	0.80	0.61
FHIT exon 3b	0.66	0.81	0.89	0.74	0.78	0.64	0.71	0.51
FHIT exon 4a	0.85	1.07	0.93	1.07	1.01	0.71	0.95	0.77
FHIT exon 4b	0.79	0.77	0.91	0.94	0.70	0.58	0.78	0.61
FHIT exon 5a	0.95	1.06	0.97	1.00	0.96	0.61	1.05	0.79
FHIT exon 5b	0.70	0.86	0.85	1.06	0.91	0.70	0.97	0.64
FHIT exon 6	0.61	0.79	0.81	0.44	0.84	0.59	0.89	0.52
FHIT exon 7	0.85	1.10	0.89	1.08	1.04	0.53	0.98	0.74
FHIT exon 8a	0.66	0.68	0.68	0.74	0.71	0.48	0.59	0.52
FHIT exon 8b	0.74	0.68	0.85	0.62	0.60	0.54	0.65	0.54
FHIT exon 9	0.67	0.78	0.62	0.45	0.67	0.46	0.56	0.45
FHIT exon 10	0.79	0.64	0.72	0.81	0.66	0.52	0.55	0.18
3p22 (CTNNB1)	0.74	0.70	0.82	0.55	0.61	0.58	0.54	0.40
3p25.3 (FANCD2)	1.01	1.06	1.00	1.15	1.10	1.02	1.02	0.87
16q22.1 (CDH1)	0.92	1.16	0.84	0.68	0.88	0.64	0.79	0.74
16q22.1 (CDH1)	0.66	0.82	0.74	0.68	0.66	0.55	0.68	0.57
16q24.2 (CDH13)	0.73	1.01	0.68	0.38	0.69	0.47	0.67	0.59
16q24.3 (FANCA)	0.90	1.17	0.83	0.61	0.85	0.68	0.72	0.72
16q24.3 (FANCA)	0.99	1.10	0.93	0.65	0.80	0.61	0.70	0.33
16q24.3 (FANCA)	0.88	0.93	0.67	0.23	0.66	0.57	0.60	0.20
16q24.3 (FANCA)	0.82	1.00	0.65	0.26	0.59	0.55	0.48	0.53
WWOX exon 1a	0.85	1.30	0.95	0.37	1.01	0.34	0.76	0.62
WWOX exon 1b	0.69	1.05	0.63	0.31	0.70	0.30	0.51	0.46
WWOX E 5A	0.87	1.05	0.92	0.75	0.85	0.44	0.84	0.73
WWOX exon 6	0.66	0.84	0.79	0.83	0.61	0.59	0.69	0.64
WWOX exon 7	0.86	0.95	0.98	0.64	0.70	0.71	0.63	0.65
WWOX exon 9a	0.77	0.94	0.65	0.36	0.74	0.62	0.69	0.56
WWOX exon 9b	0.75	0.94	0.69	0.44	0.66	0.65	0.57	0.50
CAV 1	0.76	0.97	0.71	0.48	0.66	0.43	0.52	0.55
CAV 2	0.76	0.97	0.70	0.73	1.16	0.69	0.66	0.50

run 205	S 104	S 2989	S 408	S 1148	S 972	S 2133	S 1616
10p12.1 (CREM)	1.33	1.00	1.00	1.00	1.00	1.37	1.00
10p14 (RENT)	0.77	0.69	0.65	0.67	0.61	0.95	0.76
FHIT exon 1a	0.64	0.55	0.15	0.61	0.79	0.52	0.33
FHIT exon1B	0.56	0.39	0.00	0.45	0.46	0.55	0.26
FHIT exon 2 a	1.00	1.08	0.64	0.97	1.90	1.07	1.15
FHIT exon 2b	0.84	0.86	0.37	0.69	1.62	0.95	0.96
FHIT exon 3 a	0.90	0.89	0.60	0.78	1.38	0.85	0.81
FHIT exon 3b	0.80	0.74	0.45	0.75	0.97	0.64	0.57
FHIT exon 4a	0.88	0.81	0.56	0.72	1.96	0.86	1.10
FHIT exon 4b	0.84	0.79	0.37	0.68	1.33	0.71	0.65
FHIT exon 5a	0.90	0.89	0.51	0.76	1.06	0.88	0.81
FHIT exon 5b	0.97	0.92	0.66	0.90	0.99	0.67	0.77
FHIT exon 6	0.69	0.50	0.66	0.64	0.70	0.70	0.57
FHIT exon 7	0.89	0.87	0.32	0.75	0.59	0.68	0.64
FHIT exon 8a	0.77	0.59	0.52	0.70	0.62	0.57	0.55
FHIT exon 8b	0.81	0.66	0.45	0.67	0.57	0.67	0.63
FHIT exon 9	0.67	0.41	0.35	0.58	0.47	0.60	0.51
FHIT exon 10	0.23	0.52	0.36	0.53	0.47	0.71	0.44
3p22 (CTNNB1)	0.71	0.50	0.31	0.64	0.48	0.72	0.37
3p25.3 (FANCD2)	1.01	1.00	0.70	0.93	0.82	1.00	1.00
16q22.1 (CDH1)	0.86	0.66	0.82	0.61	0.92	0.80	0.75
16q22.1 (CDH1)	0.96	0.56	0.44	0.50	0.55	0.59	0.63
16q24.2 (CDH13)	0.66	0.51	0.52	0.40	0.60	0.58	0.42
16q24.3 (FANCA)	0.87	0.72	0.45	0.72	0.73	0.71	0.60
16q24.3 (FANCA)	0.34	0.54	0.48	0.52	0.55	0.71	0.46
16q24.3 (FANCA)	0.28	0.51	0.35	0.52	0.32	0.69	0.20
16q24.3 (FANCA)	0.99	0.58	0.19	0.59	0.34	0.65	0.33
WWOX exon 1a	0.79	0.69	0.31	0.58	0.49	0.73	0.38
WWOX exon 1b	0.74	0.69	0.35	0.58	0.36	0.59	0.36
WWOX E 5A	1.14	0.90	0.69	0.76	1.01	0.83	0.74
WWOX exon 6	1.14	0.80	0.38	0.73	0.77	0.63	0.76
WWOX exon 7	0.93	0.72	0.36	0.64	0.50	0.68	0.51
WWOX exon 9a	0.72	0.63	0.37	0.59	0.50	0.70	0.41
WWOX exon 9b	0.80	0.70	0.41	0.67	0.38	0.43	0.28
CAV 1	0.98	0.68	0.47	0.69	0.47	0.76	0.43
CAV 2	1.01	1.07	0.66	1.09	0.62	1.00	1.14

run 205 B	S 2134	S 2041	S 2135	S 2136	S 2391	S 2398	S 1978
10p12.1 (CREM)	1.44	0.54	1.11	1.00	1.23	0.78	1.00
10p14 (RENT)	0.94	0.45	0.90	1.02	0.95	1.00	0.45
FHIT exon 1a	0.79	0.47	0.45	0.50	0.82	0.59	0.35
FHIT exon1B	0.79	0.32	0.47	0.40	0.57	0.52	0.10
FHIT exon 2 a	0.82	1.08	0.85	0.73	1.26	0.68	1.29
FHIT exon 2b	0.78	0.75	0.86	0.47	1.02	0.52	0.88
FHIT exon 3 a	0.73	0.97	0.69	0.62	0.97	0.62	1.15
FHIT exon 3b	0.68	0.72	0.56	0.57	0.87	0.61	0.92
FHIT exon 4a	0.55	0.85	0.84	0.49	1.06	0.64	1.12
FHIT exon 4b	0.84	0.81	0.61	0.49	0.95	0.67	0.98
FHIT exon 5a	0.66	0.93	0.72	0.65	1.13	0.66	1.03
FHIT exon 5b	0.58	0.94	0.67	0.69	1.08	0.68	1.14
FHIT exon 6	0.83	0.42	0.75	0.55	0.81	0.53	0.39
FHIT exon 7	0.73	0.78	0.59	0.58	1.03	0.69	1.10
FHIT exon 8a	0.79	0.44	0.63	0.48	0.82	0.55	0.51
FHIT exon 8b	0.54	0.66	0.51	0.50	0.84	0.52	0.96
FHIT exon 9	0.67	0.42	0.62	0.50	0.68	0.61	0.40
FHIT exon 10	0.60	0.53	0.52	0.59	0.71	0.67	0.76
3p22 (CTNNB1)	0.55	0.48	0.61	0.62	0.77	0.72	0.46
3p25.3 (FANCD2)	0.48	0.97	0.81	0.89	0.97	0.88	1.16
16q22.1 (CDH1)	0.51	0.63	1.00	0.76	0.96	1.00	0.66
16q22.1 (CDH1)	0.44	0.65	0.75	0.84	0.90	1.15	0.86
16q24.2 (CDH13)	0.44	0.62	0.78	0.60	0.74	0.92	0.38
16q24.3 (FANCA)	0.61	0.61	0.99	0.88	0.99	1.03	0.64
16q24.3 (FANCA)	0.63	0.50	0.94	0.65	0.77	0.89	0.48
16q24.3 (FANCA)	0.42	0.45	0.68	0.61	0.78	0.94	0.26
16q24.3 (FANCA)	0.52	0.61	0.89	1.00	1.01	1.43	0.54
WWOX exon 1a	0.47	0.60	0.83	0.72	0.89	1.08	0.41
WWOX exon 1b	0.58	0.58	0.81	0.76	0.80	1.13	0.34
WWOX E 5A	0.44	0.93	1.03	0.88	1.22	1.16	0.81
WWOX exon 6	0.75	0.84	0.73	0.86	1.22	1.14	1.04
WWOX exon 7	0.40	0.68	0.78	0.77	0.99	1.29	0.77
WWOX exon 9a	0.65	0.53	0.87	0.69	0.93	0.86	0.30
WWOX exon 9b	0.31	0.50	0.58	0.75	1.03	1.03	0.39
CAV 1	0.54	0.64	0.75	0.76	0.86	1.15	0.69
CAV 2	0.94	0.67	1.04	1.25	1.20	1.29	0.95

Appendix 5: composite karyotype of cell line UWOV2

82~170, XXXX, -1, -1, add (1)(?q42), -2, add(2)(q37), del(3)(q12)x2, der(3)t(3;?)(p11;?), der(3)t(?;3;?)(?;q26-p11;?), - 4, + 5, +5, + add (6)(q24), - 7, +add(7)(p15)x2, +del(7)(q22) x2, -8, -8, +der(8)t(3;8)(p13;q24), -10, -10, -11, -11,+der(11) t(8;3;11)(?;p14; p15)x2, del(12)(p12), +add(12)(p13), +der(12)t(?;12;?), -13, add(14)(p11), -15, -15, +der(15)(?;15;?), -16, -17, -18, -18, del(18)(q 21), -19, -21, -22, -22, mar [cp20].

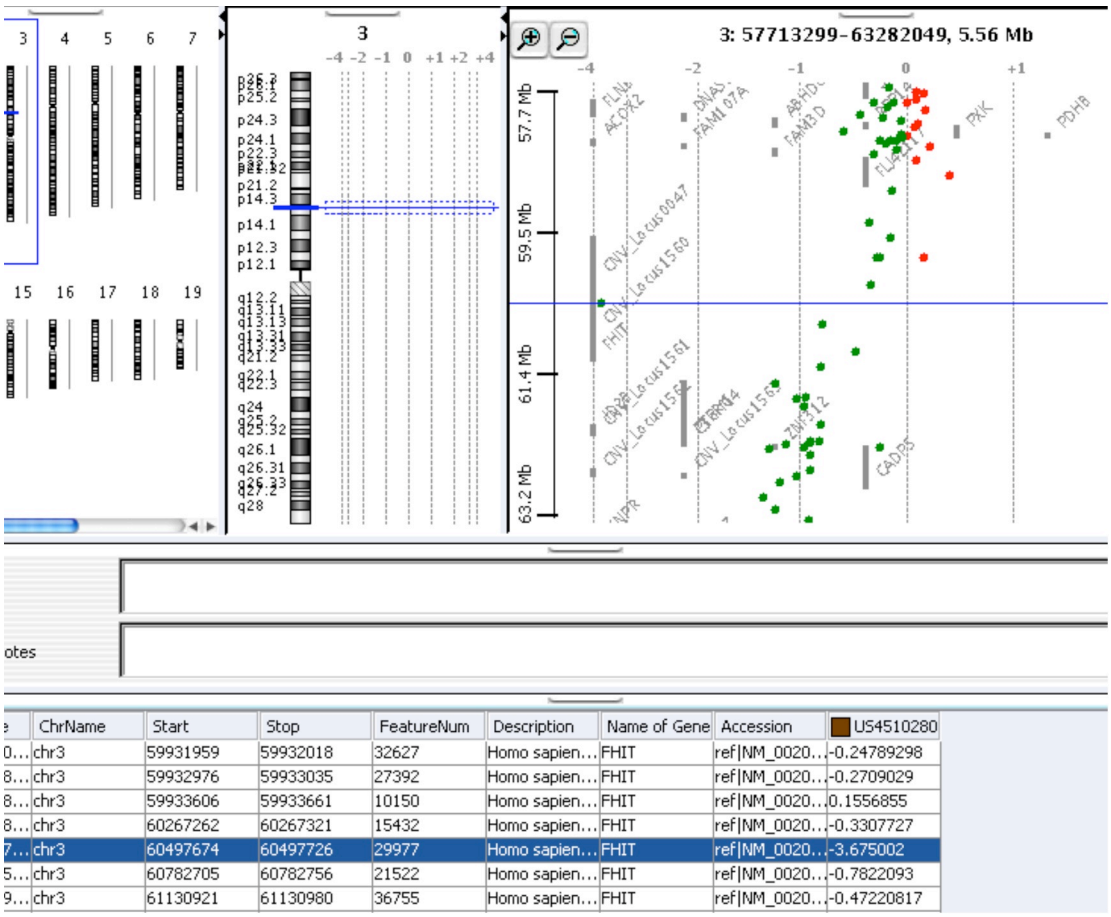
Appendix 6: FHIT and WWOX sequences corresponding to the CGH array 44k (Agilent) oligonucleotide probes.

A. FHIT sequences corresponding to the Agilent array Oligonucleotide probe and accessed through NCBI Accession number given on the CGH analytics software.

The figure on the next page (Figure 7.1) is an enlarged copy of part of Figure 5.6 (Chapter 5). It shows detailed information on the Agilent oligonucleotide probes used in the 44K array.

There are 7 oligonucleotide probes that span the FHIT locus on chromosome 3 in the Agilent CGH array (44k). One oligonucleotide probe showed a complete deletion in the near tetraploid UWOV2 cell line (- 4 copies) (Figure 7.1, A). The corresponding sequence of this probe starts at nucleotide 60497674 and end at nucleotide 60497726 on chromosome 3p. It is correlated here to FHIT transcript sequence information accessible via http://www.ensembl.org/Homo_sapiens/ and this probe is seen to overlap sequences within FHIT exon 5 (Figure 7.1, B). The results obtained by CGH array were therefore in agreement with those obtained by MPLA (see section 5.4.6 and Figure 5.9 in Chapter 5).

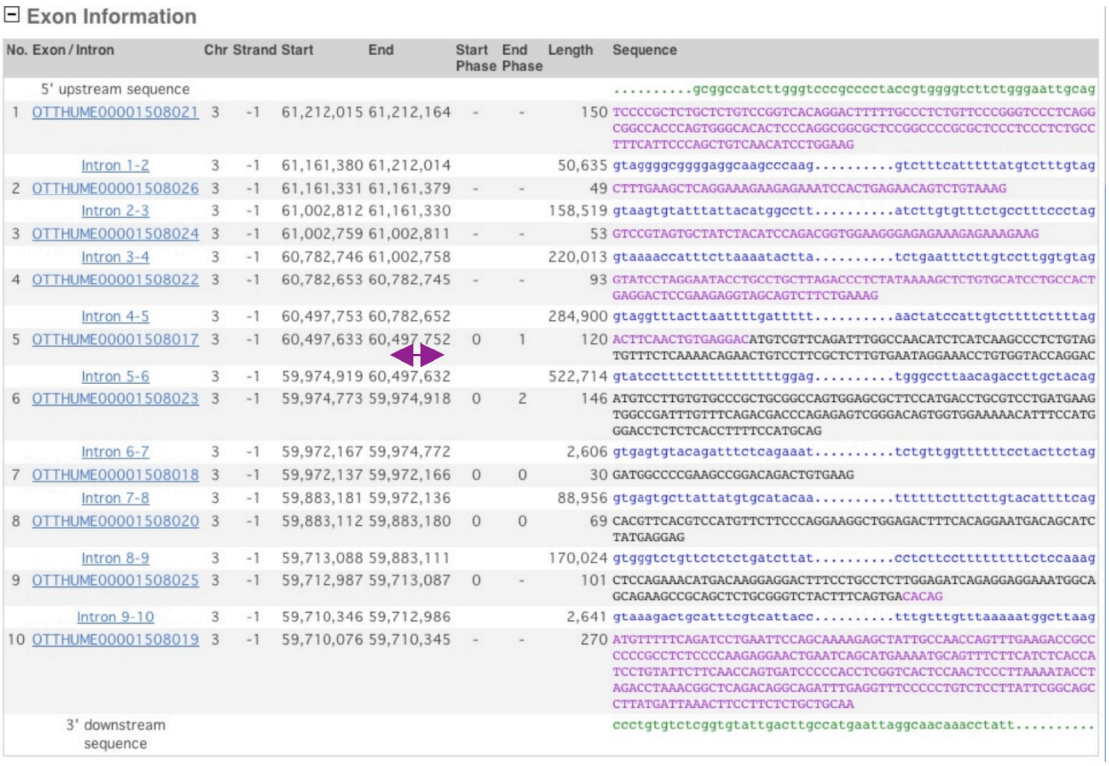
A



ChrName	Start	Stop	FeatureNum	Description	Name of Gene	Accession	US4510280
chr3	59931959	59932018	32627	Homo sapien...	FHIT	ref|NM_0020...	-0.24789298
chr3	59932976	59933035	27392	Homo sapien...	FHIT	ref|NM_0020...	-0.2709029
chr3	59933606	59933661	10150	Homo sapien...	FHIT	ref|NM_0020...	0.1556855
chr3	60267262	60267321	15432	Homo sapien...	FHIT	ref|NM_0020...	-0.3307727
chr3	60497674	60497726	29977	Homo sapien...	FHIT	ref|NM_0020...	-3.675002
chr3	60782705	60782756	21522	Homo sapien...	FHIT	ref|NM_0020...	-0.7822093
chr3	61130921	61130980	36755	Homo sapien...	FHIT	ref|NM_0020...	-0.47220817

Figure 7.1: *(A) aGCH array results obtained with the UWOV2 cell line showing the sequence information of oligonucleotide probes specific for the FHIT locus. The FHIT specific oligonucleotide probe selected by the blue cursor shows a homozygous deletion (-4 copies). It is highlighted below in blue with the corresponding nucleotide sequence information.*

B



Exon Information

No.	Exon / Intron	Chr	Strand	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence								.....gcggccatcttgggtcccgcccctaccgtggggtcttctgggaattgcag
1	OTTHUME00001508021	3	-1	61,212,015	61,212,164	-	-	150	TCCCCGCTCTGCTCTGTCCGGTCACAGGACTTTTTGCCCTCTGTTCCCGGGTCCCTCAGG CGGCCACCCAGTGGGCACACTCCCAGGCGGCGCTCCGGCCCCGCGCTCCCTCCCTCTGCC TTTCATTCCCAGCTGTCAACATCCTGGAAG
	Intron 1-2	3	-1	61,161,380	61,212,014			50,635	gtaggggcggggaggcaagcccaag.....gtctttcatttttatgtctttgtag
2	OTTHUME00001508026	3	-1	61,161,331	61,161,379	-	-	49	CTTTGAAGCTCAGGAAAGAAGAGAAATCCACTGAGAACAGTCTGTAAAG
	Intron 2-3	3	-1	61,002,812	61,161,330			158,519	gtaagtgtatttattacatggcctt.....atcttgtgtttctgcctttccctag
3	OTTHUME00001508024	3	-1	61,002,759	61,002,811	-	-	53	GTCCGTAGTGCTATCTACATCCAGACGGTGGAAGGGAGAGAAAGAGAAAGAAG
	Intron 3-4	3	-1	60,782,746	61,002,758			220,013	gtaaaaccatttcttaaaatactta.....tctgaatttcttgtccttggtgtag
4	OTTHUME00001508022	3	-1	60,782,653	60,782,745	-	-	93	GTATCCTAGGAATACCTGCCTGCTTAGACCCTCTATAAAAGCTCTGTGCATCCTGCCACT GAGGACTCCGAAGAGGTAGCAGTCTTCTGAAAG
	Intron 4-5	3	-1	60,497,753	60,782,652			284,900	gtaggtttacttaattttgattttt.....aactatccattgtcttttctttttag
5	OTTHUME00001508017	3	-1	60,497,633	60,497,752	0	1	120	ACTTCAACTGTGAGGACATGTCGTTCAGATTTGGCCAACATCTCATCAAGCCCTCTGTAG TGTTTCTCAAAACAGAACTGTCCTTCGCTCTTGTGAATAGGAAACCTGTGGTACCAGGAC
	Intron 5-6	3	-1	59,974,919	60,497,632			522,714	gtatcctttcttttttttttggag.....tgggccttaacagaccttgctacag
6	OTTHUME00001508023	3	-1	59,974,773	59,974,918	0	2	146	ATGTCCTTGTGTGCCCGCTGCGGCCAGTGGAGCGCTTCCATGACCTGCGTCCTGATGAAG TGGCCGATTTGTTTCAGACGACCCAGAGAGTCGGGACAGTGGTGGAAAAACATTTCCATG GGACCTCTCTCACCTTTTCCATGCAG
	Intron 6-7	3	-1	59,972,167	59,974,772			2,606	gtgagtgtacagatttctcagaaat.....tctgttggtttttcctacttctag
7	OTTHUME00001508018	3	-1	59,972,137	59,972,166	0	0	30	GATGGCCCCGAAGCCGGACAGACTGTGAAG
	Intron 7-8	3	-1	59,883,181	59,972,136			88,956	gtgagtgcttattatgtgcatacaa.....ttttttctttcttgtacattttcag
8	OTTHUME00001508020	3	-1	59,883,112	59,883,180	0	0	69	CACGTTCACGTCCATGTTCTTCCCAGGAAGGCTGGAGACTTTCACAGGAATGACAGCATC TATGAGGAG
	Intron 8-9	3	-1	59,713,088	59,883,111			170,024	gtgggtctgttctctctgatcttat.....cctcttcctttttttttctccaaag
9	OTTHUME00001508025	3	-1	59,712,987	59,713,087	0	-	101	CTCCAGAAACATGACAAGGAGGACTTTCCTGCCTCTTGGAGATCAGAGGAGGAAATGGCA GCAGAAGCCGCAGCTCTGCGGGTCTACTTTCAGTGACACAG
	Intron 9-10	3	-1	59,710,346	59,712,986			2,641	gtaaagactgcatttcgtcattacc.....tttgtttgtttaaaaatggcttaag
10	OTTHUME00001508019	3	-1	59,710,076	59,710,345	-	-	270	ATGTTTTTCAGATCCTGAATTCCAGCAAAAGAGCTATTGCCAACCAGTTTGAAGACCGCC CCCCGCCTCTCCCCAAGAGGAACTGAATCAGCATGAAAATGCAGTTTCTTCATCTCACCA TCCTGTATTCTTCAACCAGTGATCCCCCACCTCGGTCACTCCAACTCCCTTAAAATACCT AGACCTAAACGGCTCAGACAGGCAGATTTGAGGTTTCCCCCTGTCTCCTTATTCGGCAGC CTTATGATTAAACTTCCTTCTCTGCTGCAA
	3' downstream sequence								ccctgtgtctcggtgtattgacttgccatgaattaggcaacaaacctatt.....

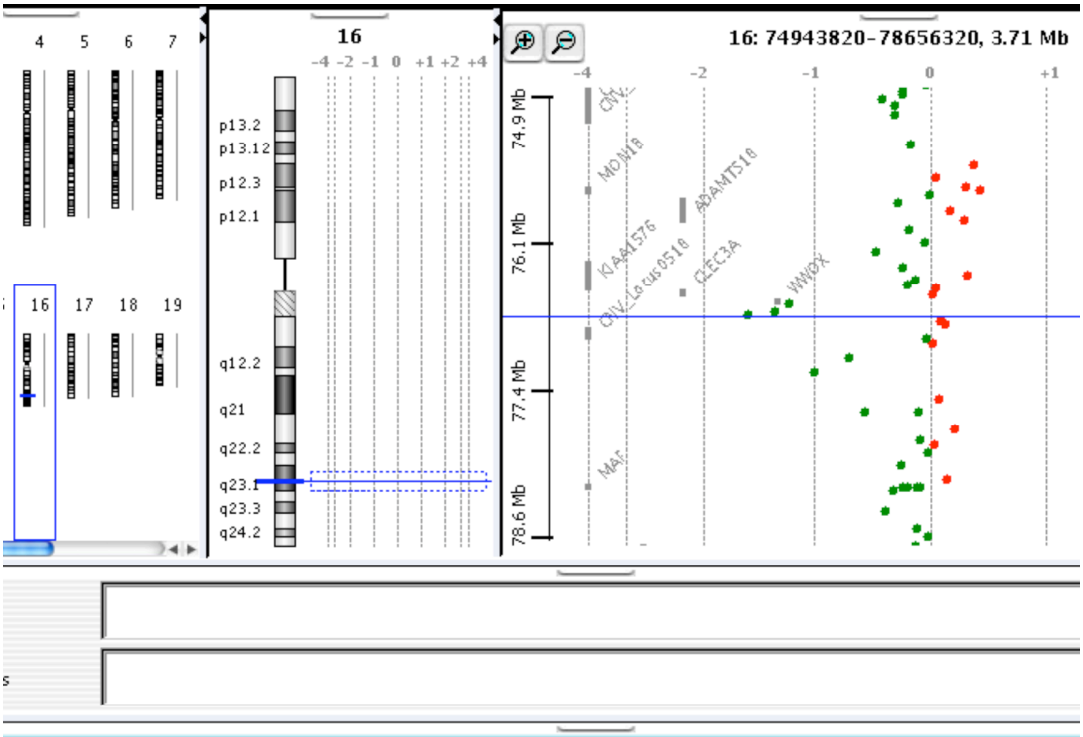
Figure 7.1*: (B) The FHIT gene exon/intron boundaries and sequence composition from the NCBI reference is shown. The purple double arrow indicates the position of the 44K Agilent CGH array oligonucleotide probe highlighted in figure 7.1 (A). This probe is seen to map within the exon 5 of the FHIT gene.*

B. WWOX sequences corresponding to the Agilent array Oligonucleotide probes and accessed through NCBI Accession number given on the CGH analytics software.

There are 14 oligonucleotide probes spanning the WWOX gene locus on chromosome 16 in the Agilent CGH array (44k). The WWOX oligonucleotide sequence close to WWOX exon 5 is shown below in Figure 7.2, A. This oligonucleotide probe, selected by the blue cursor in the CGH analytics image, starts at nucleotide: 76841115 and ends at 76841174. This probe maps within WWOX intron 5-6 close to exon 5 (Figure 7.2, B) and showed an hemizygous deletion in the cell line UWOV2. This corresponded closely to the MLPA

results obtained on the UWOV2 cell line that also showed hemizygous deletion for WWOX exon 5 (see Chapter 5, Figure 5.9, A). The probes 5' to exon 5 (intron 2-3) also showed hemizygous deletions in contrast to the probes located 3' to exon 5 (introns 6 to 10) that were not significantly deleted. These results were concordant to the ratios obtained by MLPA with the probes spanning WWOX exons 1 and 5 on the 5' region of the gene and exons 6, 7 and 9 on the 3' part of the WWOX locus.

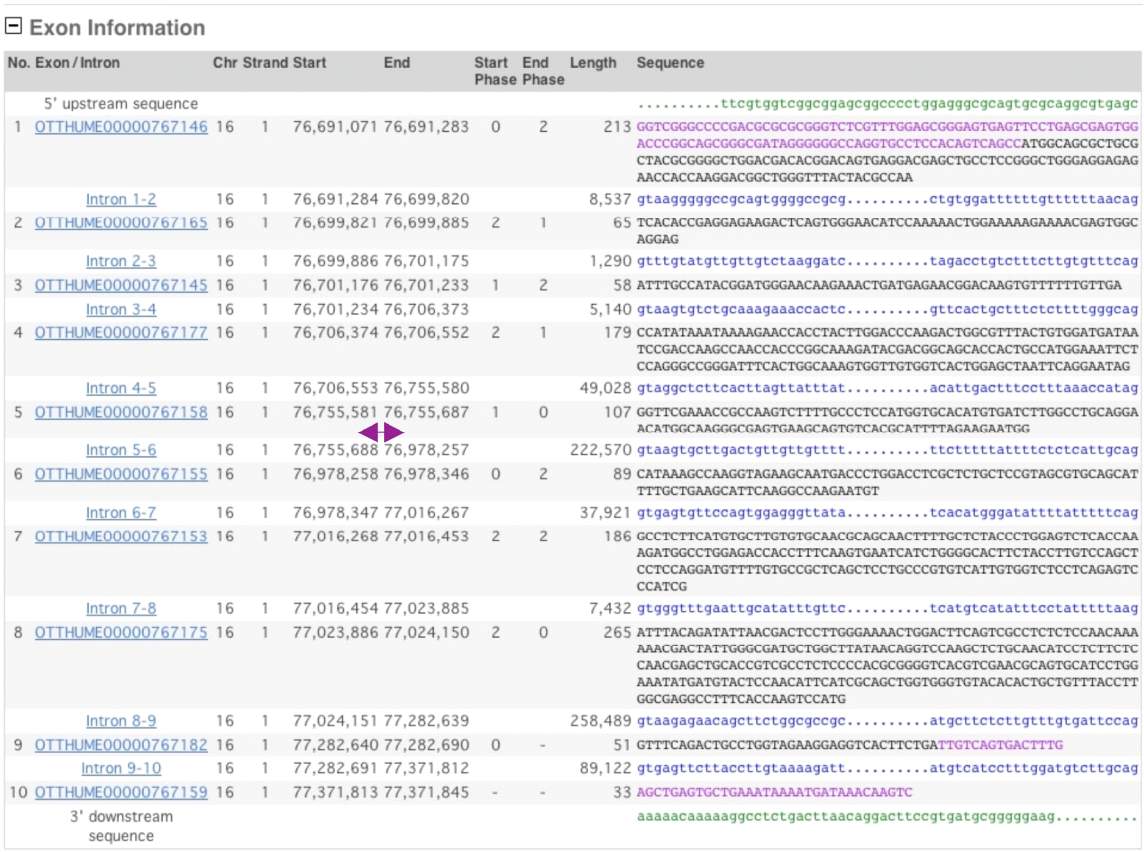
A



ChrName	Start	Stop	FeatureNum	Description	Name of Gene	Accession	US4510280
chr16	76701194	76701253	25648	Homo sapien...	WWOX	ref|NM_1307...	-1.2036757
chr16	76762418	76762477	30418	Homo sapien...	WWOX	ref|NM_1307...	-1.333209
chr16	76787484	76787543	25198	Homo sapien...	WWOX	ref|NM_1307...	-1.567372
chr16	76841115	76841174	36368	Homo sapien...	WWOX	ref|NM_1307...	0.08472306
chr16	76870007	76870063	39990	Homo sapien...	WWOX	ref|NM_1307...	0.13240741
chr16	76981676	76981735	1392	Homo sapien...	WWOX	ref|NM_1307...	-0.031578504
chr16	77023863	77023922	7232	Homo sapien...	WWOX	ref|NM_1307...	0.016754627
chr16	77143761	77143820	29694	Homo sapien...	WWOX	ref|NM_1307...	-0.692971
chr16	77251138	77251197	42062	Homo sapien...	WWOX	ref|NM_1307...	-0.997601
chr16	77472169	77472228	16360	Homo sapien...	WWOX	ref|NM_1307...	0.07618282
chr16	77573211	77573265	12485	Homo sapien...	WWOX	ref|NM_1307...	-0.56415784
chr16	77576005	77576064	33375	Homo sapien...	WWOX	ref|NM_1307...	-0.10472023
chr16	77715394	77715453	10006	Homo sapien...	WWOX	ref|NM_1307...	0.20729297
chr16	77803469	77803527	41654	Homo sapien...	WWOX	ref|NM_1307...	-0.09010079

Figure 7.2: *(A) Hybridization results on CGH analytics for the UWOV2 cell line. The blue cursor highlights a WWOX oligonucleotide probe that is deleted in half copy number. The corresponding sequence information is highlighted below in the probe list.*

B



Exon Information

No.	Exon / Intron	Chr	Strand	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence								.....ttcgtggtcggcggagcggcccctggagggcgcagtgcgcaggcgtgagc
1	OTTHUME00000767146	16	1	76,691,071	76,691,283	0	2	213	GGTCGGGCCCCGACGCGCGCGGGTCTCGTTTGGAGCGGGAGTGAGTTCCTGAGCGAGTGG ACCCGGCAGCGGGCGATAGGGGGGCCAGGTGCCTCCACAGTCAGCCATGGCAGCGCTGCG CTACGCGGGGCTGGACGACACGGACAGTGAGGACGAGCTGCCTCCGGGCTGGGAGGAGAG AACCACCAAGGACGGCTGGGTTTACTACGCCAA
	Intron 1-2	16	1	76,691,284	76,699,820			8,537	gtaaggggccgcagtggggccgcg.....ctgtggatttttgtttttaacag
2	OTTHUME00000767165	16	1	76,699,821	76,699,885	2	1	65	TCACACCGAGGAGAAGACTCAGTGGGAACATCCAAAAACTGGAAAAAGAAAACGAGTGGC AGGAG
	Intron 2-3	16	1	76,699,886	76,701,175			1,290	gtttgtatgttgttgtctaaggatc.....tagacctgtctttcttgtgtttcag
3	OTTHUME00000767145	16	1	76,701,176	76,701,233	1	2	58	ATTTGCCATACGGATGGGAACAAGAAACTGATGAGAACGGACAAGTGTTTTTTGTTGA
	Intron 3-4	16	1	76,701,234	76,706,373			5,140	gtaagtgtctgcaaagaaaccactc.....gttcactgctttctctttttgggcag
4	OTTHUME00000767177	16	1	76,706,374	76,706,552	2	1	179	CCATATAAATAAAAGAACCACCTACTTGGACCCAAGACTGGCGTTTACTGTGGATGATAA TCCGACCAAGCCAACCACCCGGCAAAGATACGACGGCAGCACCACTGCCATGGAAATTCT CCAGGGCCGGGATTTCACTGGCAAAGTGGTTGTGGTCACTGGAGCTAATTCAGGAATAG
	Intron 4-5	16	1	76,706,553	76,755,580			49,028	gtaggctcttcacttagttatttat.....acattgactttcctttaaaccatag
5	OTTHUME00000767158	16	1	76,755,581	76,755,687	1	0	107	GGTTCGAAACCGCCAAGTCTTTTGCCCTCCATGGTGCACATGTGATCTTGGCCTGCAGGA ACATGGCAAGGGCGAGTGAAGCAGTGTCACGCATTTTAGAAGAATGG
	Intron 5-6	16	1	76,755,688	76,978,257			222,570	gtaagtgcttgactgttgttgtttt.....ttctttttatttttctctcattgcag
6	OTTHUME00000767155	16	1	76,978,258	76,978,346	0	2	89	CATAAAGCCAAGGTAGAAGCAATGACCCTGGACCTCGCTCTGCTCCGTAGCGTGCAGCAT TTTGCTGAAGCATTCAAGGCCAAGAATGT
	Intron 6-7	16	1	76,978,347	77,016,267			37,921	gtgagtgttccagtggagggttata.....tcacatgggatattttatttttcag
7	OTTHUME00000767153	16	1	77,016,268	77,016,453	2	2	186	GCCTCTTCATGTGCTTGTGTGCAACGCAGCAACTTTTGCTCTACCCTGGAGTCTCACCAA AGATGGCCTGGAGACCACCTTTCAAGTGAATCATCTGGGGCACTTCTACCTTGTCCAGCT CCTCCAGGATGTTTTGTGCCGCTCAGCTCCTGCCCGTGTCATTGTGGTCTCCTCAGAGTC CCATCG
	Intron 7-8	16	1	77,016,454	77,023,885			7,432	gtgggtttgaattgcatatttgttc.....tcatgtcatatttcctatttttaag
8	OTTHUME00000767175	16	1	77,023,886	77,024,150	2	0	265	ATTTACAGATATTAACGACTCCTTGGGAAAACTGGACTTCAGTCGCCTCTCTCCAACAAA AAACGACTATTGGGCGATGCTGGCTTATAACAGGTCCAAGCTCTGCAACATCCTCTTCTC CAACGAGCTGCACCGTCGCCTCTCCCCACGCGGGGTCACGTCGAACGCAGTGCATCCTGG AAATATGATGTACTCCAACATTCATCGCAGCTGGTGGGTGTACACACTGCTGTTTACCTT GGCGAGGCCTTTCACCAAGTCCATG
	Intron 8-9	16	1	77,024,151	77,282,639			258,489	gtaagagaacagcttctggcgccgc.....atgcttctcttgtttgtgattccag
9	OTTHUME00000767182	16	1	77,282,640	77,282,690	0	-	51	GTTTCAGACTGCCTGGTAGAAGGAGGTCACTTCTGATTGTCAGTGACTTTG
	Intron 9-10	16	1	77,282,691	77,371,812			89,122	gtgagttcttaccttgtaaaagatt.....atgtcatcctttggatgtcttgcag
10	OTTHUME00000767159	16	1	77,371,813	77,371,845	-	-	33	AGCTGAGTGCTGAAATAAAATGATAAACAAGTC
	3' downstream sequence								aaaaacaaaaaggcctctgacttaacaggacttccgtgatgcgggggaag.....

Figure 7.2: ***(B)*** *The WWOX gene exon/intron boundaries and sequence composition from the NCBI reference sequence is shown. The purple double arrow indicates the position of the 44K Agilent CGH array oligonucleotide probe highlighted in figure 7.2 (A). This probe is seen to map within intron 5- 6 of the WWOX gene close to exon 5.*