

**The Use of Gene Expression Profiling to Identify
Novel Minimal Residual Disease Markers (MRD)
in Acute Myeloid Leukaemia (AML)**

Eleanor Woodward

PhD

2010

Cardiff University

UMI Number: U585569

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U585569

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

I would like to thank Dr Chris Pepper and Dr Richard Darley, without their help and supervision this project would not have been possible. I would also like to thank Professor Alan Burnett for allowing me to perform this research within his department.

I would also like to thank those within in the Department of Haematology, in particular to Amanda Gilkes, Dr Elisabeth Walsby, Michelle Lazenby and Val Walsh, and to Dr Peter Giles from the Department of Pathology for providing a considerable amount of support and guidance contributing towards my study.

Finally I would like to thank my family and friends for all the support they provided throughout my research, in particular to Tim for explaining the difficult bits, Joseph for reminding me all that really matters is physics and to Doug for providing a roof over my head.

Summary

Acute myeloid leukaemia (AML) is a heterogeneous disorder characterised by the accumulation of immature haematopoietic cells blocked at various stages of differentiation. Despite improved survival rates over the past decade, relapse occurs in approximately 70% patients undergoing chemotherapy. A potential reason for this is that current clinical protocols do not take account of the level of residual disease present at remission. Therefore, one strategy to reduce relapse rates is to monitor minimal residual disease and continue to treat until the patient is minimal residual disease negative. Current minimal residual disease markers are available for patients with characterised fusion genes but approximately 50% of patients have no detectable chromosomal aberration and therefore are without markers. Gene expression profiling is a powerful tool for disease classification, prognosis and therapeutic predictions. This study aimed to investigate the use gene expression profiling to identify novel minimal residual disease markers for specific AML sub-groups. Patient diagnostic samples were profiled to identify genes specific to AML patients with a favourable translocation in order to establish the "proof-of-principle". Several genes identified were followed in patient diagnostic and follow-up samples and compared to the markers currently used. Continuing with normal karyotype AML, genes were identified as specific to this sub-group. Several homeobox (*HOX*) genes and the Wilms' tumour (*WT1*) gene were identified and their MRD levels followed in diagnostic and follow-up samples. Only *WT1* identified as specific to normal karyotype AML met the necessary criteria to be an MRD marker. Although the majority of genes selected from the GEP in this study proved unsuitable as markers, the identification and validation of a marker already used for MRD monitoring, *WT1*, demonstrates the ability of gene expression profiling to identify potential minimal residual disease markers in normal karyotype AML.

Abbreviations

ABL	Abelson tyrosine kinase
ALCL	Anaplastic large cell lymphoma
ALL	Acute lymphoid leukaemia
AML	Acute myeloid leukaemia
AML1	Acute myeloid leukaemia 1
AML1-ETO	Acute myeloid leukaemia 1- Eight twenty-one Oncoproteins
AML-CFU	Proliferative leukaemic blasts
ANG-1	Angiopoietin 1
AP3S1	Adaptor-related protein complex 3, sigma 1 subunit
APL	Acute promyelocytic leukaemia
ARC	Activity-regulated cytoskeleton-associated protein
ASNase	L-asparaginase
ASNS	Asparagine synthetase
ASS1	Argininosuccinate
ATRA	All- <i>trans</i> retinoic acid
ANOVA	Analysis of variance
BAALC	Brain and acute leukemia, cytoplasmic
BIN	Biological interaction networks
BM	Bone marrow
BMEC	Bone marrow endothelial cells
BMP	Bone morphogenic protein
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
CALD1	Caldesmon 1
CBF	Core binding factor
CBF α	Core binding factor alpha
CBF β	Core binding factor beta
CBF β -MYH11	Core binding factor beta- smooth muscle myosin heavy chain
CDK	Cyclin-dependent kinase
CDKI	Cyclin-dependent kinase inhibitor
cDNA	Complementary DNA
CDX2	Caudal type homeobox 2
C/EBP	CCAAT/enhancer binding protein
C/EBP α	CCAAT/enhancer binding protein (C/EBP), alpha
CFU	Colony-forming unit
CFU-GEMM	Colony-forming units – granulocyte-erythroid- macrophage-megakaryocyte
CG	Control gene
CLIP3	CAP-GLY domain containing linker protein 3
CLL	Chronic lymphoid leukaemia
CLP	Common lymphoid progenitors
CML	Chronic myeloid leukaemia
CMP	Common myeloid progenitors
COL4A5	Collagen, type IV, alpha 5
Cp/Ct	Crossing point

CP	Chronic phase
CR	Complete remission
CSF	Colony stimulating factor
dATP	Deoxyadenosine triphosphate
DAVID	Database for annotation, visualisation and integrated discovery
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DMSO	Di-methyl sulphoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide triphosphate
dsDNA	Double stranded DNA
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
DX	Diagnostic
EAC	Europe Against Cancer
EB	Empirical Bayes
EDTA	Ethylenediaminetetracetic acid
ENU	<i>N</i> -ethyl- <i>N</i> -nitroso-urea
EPO	Erythropoietin
ET	Essential thrombocythemia
ETO	Eight twenty-one oncoproteins
EVI1	Ecotropic viral integration site 1
FAB	French-American-British
FBS	Foetal bovine serum
FCS	Foetal calf serum
FDR	False discovery rate
F-MuLV	Friend murine leukemia virus
FL	Fms-like tyrosine kinase 3 ligand
FGF-4	Fibroblast growth factor 4
FISH	Fluorescence <i>in situ</i> hybridisation
FLT3	Fms-like tyrosine kinase 3
FUP	Follow-up
G6PD	Glucose-6-phosphate dehydrogenase
GCT	Germ cell tumours
GEP	Gene expression profiling
GO	Gemtuzumab ozogamicin
G-CSF	Granulocyte - colony stimulating factor
GM-CSF	Granulocyte/macrophage- colony stimulating factor
GMP	Granulocyte/monocyte progenitor
GRIN1	Glutamate receptor, ionotropic, N-methyl D-aspartate 1
HAT	Histone acetyltransferase
HCL	Hairy cell leukaemia
HDAC	Histone deacetylase
HOX	Homeobox
HGF	Hepatocyte growth factor
HPC	Haematopoietic progenitor cell
HSC	Haematopoietic stem cell

HSCT	Haematopoietic stem cell transplant
HSPG2	Heparan sulphate proteoglycan 2
Ig	Immunoglobulin
IGF	Insulin-like growth factor
IGFBP2	Insulin-like growth factor binding protein 2
IL-3	Interleukin 3
IL5RA	Interleukin-5 receptor, alpha
IMDM	Iscoe's Modified Dulbecco's Medium
INK4	Inhibitors of CDK4
ITD	Internal tandem duplication
IVT	<i>In vitro</i> transcription
JAK2	Janus Kinase 2
LAIP	Leukaemia-associated aberrant immunophenotype
LIC	Leukaemia-initiating cells
LSC	Leukaemic stem cell
LT-HSC	Long-term haematopoietic stem cell
mAbs	Monoclonal antibodies
MAD	Median absolute deviation
MAP	Mitogen-activated protein
M-CSF	Macrophage colony-stimulating factor
M-CSFR	Macrophage colony-stimulating factor receptor
MDCK	Madin-Darby canine kidney
MDS	Myelodysplastic syndrome
MEM	Minimum essential media
MEP	Megakaryocyte/erythrocyte progenitor
MEIS1	Myeloid ecotropic viral integration site 1 homolog
MgCl ₂	Magnesium chloride
MFC	Multiparameter flow cytometry
MLL	Mixed lineage leukaemia
MLL-ENL	Mixed-lineage leukaemia - eleven nineteen leukaemia
MOZ-TIF2	Monocytic leukaemia zinc finger - TGF-beta induced factor-2
MNC	Mononuclear cell
MNI	Meningioma 1
MPD	Myeloproliferative disorder
MPN	Myeloproliferative neoplasms
MRC	Medical Research Council
MRD	Minimal residual disease
MRDv	Minimal residual disease value
MREC	Multicentre research ethics committee
mRNA	Messenger RNA
MSV-LTR	Moloney sarcoma virus terminal repeat
MuLV	Murine leukaemia virus
MYH11	Smooth muscle myosin heavy chain
NAB2	NGFI-A binding protein 2
NCoR	Nuclear receptor corepressor
NK	Natural killer
NMTS	Nuclear matrix attachment signal
NOD	Non-obese diabetic

NPM-ALK	Nucleophosmin-anaplastic lymphoma kinase 1
NPM-MLF1	Nucleophosmin-myelodysplasia/myeloid leukaemia factor 1
NPM-RAR α	Nucleophosmin-retinoic acid receptor α
NPM1	Nucleophosmin 1
N-ras	Neuroblastoma RAS viral (v-ras) oncogene homologue
NT5E	5'-nucleotidase, ecto (CD73)
NuMA	Nuclear mitotic apparatus protein
NUP98	Nucleoporin, 98kDa
OS	Overall survival
PAM	Predictive analysis for microarrays
PB	Peripheral blood
PBS	Phosphate buffered saline
PBSC	Peripheral blood stem cell
PBX3	Pre B-cell leukaemia transcription factor 3
PCA	Principle component analysis
PCR	Polymerase chain reaction
PDPN	Podoplanin
PGD2	Prostaglandin D2 synthase
Ph	Philadelphia
PHLDA1	Pleckstrin homology-like domain, family A, member 1
PI-3	Phosphoinositol-3
PKB	Protein kinase B
PLZF	Promyelocytic leukaemia zinc finger
PMF	Primary myelofibrosis
PML	Promyelocytic leukaemia
PML-RAR α	Promyelocytic leukaemia-retinoic acid receptor Alpha
POD	PML oncogenic domain
POU4F1	POU domain, class 4, transcription factor 1
PPR	Parathyroid hormone/parathyroid hormone-related peptide receptor
PRAME	Preferentially expressed antigen in melanoma
PTD	Partial tandem duplication
PU.1	Purine rich box-1
PV	Polycythemia vera
QN-RT-PCR	Quantitative nested RT-PCR
RAC3	Retinoic acid receptor interacting protein 3
RA	Retinoic acid
RAR	Retinoic acid receptor
RAR α -PML	Retinoic acid receptor alpha-promyelocytic leukaemia
RARE	Retinoic acid response elements
RAR-RXR	Retinoic acid receptor-retinoid X receptor
RAS	Rat sarcoma
RBCC	RING-finger domain, two B-box domains and an α -helical coiled-coil

RBPMs	RNA binding protein with multiple splicing
RFLP	Restriction fragment length polymorphism
RIN	Relevance interaction network
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
RQ-PCR	Real-time quantitative-polymerase chain reaction
RR	Relapse risk
RT	Reverse transcription
RT-PCR	Reverse transcriptase-PCR
RTK	Receptors with tyrosine kinase activity
RXR	Retinoid X receptor
RW1	Wash buffer 1
SAM	Significance analysis of microarrays
SAPE	Streptavidin-phycoerythrin dye
SCF	Stem cell factor
SCID	Severe combined immunodeficiency
SDF-1	Stromal-derived factor-1
SFT	Solitary fibrous tumours
siRNA	Small interfering RNA
SLC39A9	Solute carrier family 39, member 9
SL-IC	SCID leukaemia-initiating cell
SMRT	Silencing mediator for retinoid or thyroid-hormone receptors
SPARC	Secreted protein, acidic, cysteine-rich
ST18	Suppression of tumorigenicity 18
STAT	Signal transducer and activator of transcription
TALE	Three amino acid loop extension
t-AML	Therapy-related AML
TG	Test gene
TGF- α	Transforming growth factor alpha
TGT	Target Intensity
TIE-2	Tyrosine kinase with immunoglobulin-like and EGF-like domains
TKD	Tyrosine kinase domain
T _m	Melting temperature
t-MDS	Therapy-related MDS
TPO	Thyrotropin
TPPP3	Tubulin polymerisation-promoting protein family member 3
TPSAB1	Tryptase alpha 1 and beta 1
TPSB2	Tryptase beta 2
TRH	Thyrotropin-releasing hormone
WHO	World Health Organisation
WT1	Wilm's tumour 1
XCIP	X-chromosome inactivation pattern

Contents of Figures and Tables

Chapter 1

Figure number	Title	Page number
Figure 1	The production of multiple lineages from a single stem cell during normal haematopoiesis.	3
Figure 2	The FLT3 protein and the locations affected by the activating gene mutations.	36
Figure 3	The schematic structure of gemtuzumab ozogamicin.	51
Figure 4	The 50 "predictor" genes used to distinguish between 38 acute leukaemia samples.	69
Table 1	2008 WHO classification of myelodysplastic disorders	15
Table 2	The FAB classification of acute myeloid leukaemia	25
Table 3	The chromosomal aberrations and their resultant fusion products present in APL.	28
Table 4	The WHO classification of myeloid neoplasms focusing on the acute myeloid leukaemia and related neoplasms	44

Chapter 2

Figure number	Title	Page number
Figure 1	The amplification curves generated as double-stranded (ds) DNA is produced, incorporating more double strand-specific SYBR green dye, resulting in an increase in levels of fluorescence.	108
Figure 2	The melting curves and melting peaks generated in the final reaction step.	109
Table 1	The 5'-3' oligonucleotide primer sequences used for the study	78
Table 2	The source and culture conditions of the cell lines used	82

Figure number	Title	Page number
Table 3	The frequencies and percentages of the clinical and molecular characteristics of the 290 patients diagnosed with AML used in this study.	84
Table 4	The reaction mixture used for the 1 st strand cDNA synthesis.	87
Table 5	The reaction mixture used for the 2 nd strand cDNA synthesis.	88
Table 6	The reaction mixture used for the IVT reaction.	90
Table 7	The reaction mixture used for the cRNA fragmentation reaction.	92
Table 8	The reaction mixture used for the hybridisation reaction. *Fragmentation reaction mixture.	93
Table 9	The mixtures used for the stain solution and the antibody solution.	94
Table 10	The reaction mixture used for reverse-transcription of RNA.	106
Table 11	The cycling conditions used for reverse-transcription of RNA.	106
Table 12	The reaction mixture used for real-time quantitative-PCR.	110
Table 13	The cycling conditions used for real-time quantitative-PCR.	110
Table 14	The optimised magnesium chloride concentrations and annealing temperatures used for each primer set for Q-PCR.	111
Figure 15	The reaction mixture used for annealing the sense and antisense oligonucleotides used for siRNA transfection.	116
Figure 16	The reaction mixture used to label the annealed siRNA oligonucleotides with the FAM dye.	116

Chapter 3

Figure number	Title	Page number
Figure 1	The relative expression levels of the 18 probe sets identified, from probe sets that were classified as present in at least 1 sample, as being more significantly expressed in t(15;17) samples than non-t(15;17) samples and in donor samples by ≥ 3 fold.	134
Figure 2	The relative expression levels of the 17 probe sets identified, from probe sets that were classified as present in at least 2 samples and changing in expression, as being more significantly expressed in t(15;17) samples than non-t(15;17) samples and in donor samples by ≥ 3 fold.	135
Figure 3	A combination of the probe sets identified from SAM (FDR=0.05), ANOVA testing (FDR=0.05) and having a greater level of expression in t(15;17) samples than in non-t(15;17) samples and in donor samples by 3 fold change filtering. Seventeen probe sets were identified as being specifically expressed in t(15;17) samples.	138
Figure 4	The amplification curves of the primer sets used for <i>HGF</i> and <i>IGFBP2</i> and of the house-keeping gene, <i>S14</i> , to compare their amplification efficiencies.	140
Figure 5	Scatter plots of the spread of the normalised expression levels of the candidate genes measured in the t(15;17) AML diagnostic patient samples and the healthy normal donor samples	142
Figure 6	The fold changes in expression of <i>HGF</i> , <i>IGFBP2</i> , <i>PML-RARα</i> and <i>RARα-PML</i> in the diagnostic and follow-up samples of 3 t(15;17) patients, calculated as described in section 3.2.	146
Figure 7	The MRD values of <i>HGF</i> , <i>IGFBP2</i> , <i>PML-RARα</i> and <i>RARα-PML</i> in the diagnostic and follow-up samples from 3 t(15;17) patients, calculated as described in section 3.2.	151
Table 1	The number of probe sets identified as having a greater expression in t(15;17) AML patients than in non-t(15;17) patients or normal healthy donors.	131

Figure number	Title	Page number
Table 2	The 18 "present in at least 1 sample" probe sets identified from the 305 sample dataset as upregulated in t(15;17) AML patients samples when compared to non-t(15-17) patients and to healthy donor samples by fold change (FC>3), and the 17 "present and changing" probe sets identified from the 305 sample dataset as upregulated in t(15;17) AML patients samples when compared to non-t(15-17) patients and to healthy donor samples by fold change (FC>3).	133
Table 3	The number of probe sets identified as having a greater expression in t(15;17) AML patients than in non-t(15;17) patients or normal healthy donors through SAM and ANOVA analysis.	137
Table 4	The statistical significance of the mean normalised expression levels of the candidate genes measured in the t(15;17) AML diagnostic patient samples and the healthy normal donor samples.	141
Table 5	The probe set lists identified from the gene expression profiling analysis (Section 3.3.1 and 3.3.2).	148

Chapter 4

Figure number	Title	Page number
Figure 1	The relative expression levels of the 60 probe sets identified, from probe sets that were classified as present in at least 1 sample and as present and changing, as being more significantly expressed in t(8;21) samples than non-t(8;21) samples and in donor samples by ≥ 3 fold.	167
Figure 2	A combination of the probe sets identified from SAM (FDR=0.05), ANOVA testing (FDR=0.05) and having a greater level of expression in t(8;21) samples than in non-t(8;21) samples and in donor samples by 3 fold change filtering. Forty-one probe sets were identified as being specifically expressed in t(8;21) samples.	170
Figure 3	The relative expression levels of the 41 probe sets identified, from probe sets that were classified as present and changing, as being more significantly expressed in t(8;21) samples than non-t(8;21) samples and in donor samples by ≥ 3 fold and through SAM and ANOVA analysis with an FDR of 0.05.	172

Figure number	Title	Page number
Figure 4	The amplification curves of the primer sets used for <i>PRAME</i> , <i>POU4F</i> , <i>ETO</i> and <i>IL5RA</i> and of the house-keeping gene, <i>S14</i> , to compare their amplification efficiencies.	175
Figure 5	Scatter plots of the spread of the normalised expression levels of the candidate genes measured in the t(8;21) AML diagnostic patient samples and the healthy normal donor samples.	176
Figure 6	Longitudinal monitoring of the t(8;21) fusion gene in three t(8;21) AML patients, and of <i>WT1</i> in one t(8;21) AML patient .	181
Figure 7	Longitudinal monitoring of <i>PRAME</i> , <i>POU4F1</i> , <i>ETO</i> and <i>IL5RA</i> in three t(8;21) AML patients.	182
Figure 8	Longitudinal monitoring of <i>PRAME</i> , <i>POU4F1</i> , <i>ETO</i> and <i>IL5RA</i> in three t(8;21) AML patients.	183
Figure 9	The relative expression levels of the probe sets identified, from probe sets that were classified as present in at least 1 sample (n=100) and as present and changing (n=99), as being more significantly expressed in inv(16) samples than non-inv(16) samples and in donor samples by ≥ 3 fold.	190
Figure 10	A combination of the probe sets identified from SAM (FDR=0.05), ANOVA testing (FDR=0.05) and having a greater level of expression in t(8;21) samples than in non-t(8;21) samples and in donor samples by 3 fold change filtering.	193
Figure 11	The relative expression levels of the 66 probe sets identified, from probe sets that were classified as present and changing, as being more significantly expressed in inv(16) samples than non-inv(16) samples and in donor samples by ≥ 3 fold and through SAM and ANOVA analysis with an FDR of 0.05.	195
Figure 12	The amplification curves of the primer sets used for <i>MYH11</i> , <i>MN1</i> and <i>ST18</i> and of the house-keeping gene, <i>S14</i> , to compare their amplification efficiencies.	198

Figure number	Title	Page number
Figure 13	Scatter plots of the spread of the normalised expression levels of the candidate genes measured in the inv(16) AML diagnostic patient samples and the healthy normal donor samples.	199
Figure 14	Longitudinal monitoring of <i>MYH11</i> and <i>MN1</i> in five inv(16) AML patients.	204
Figure 15	Longitudinal monitoring of <i>MYH11</i> and <i>MN1</i> in five in(16) AML patients.	205
Figure 16	Longitudinal monitoring of the inv(16) fusion gene <i>CBFβ-MYH11</i> (a) and <i>WT1</i> (b) in five in(16) AML patients.	206
Table 1	The number of probe sets identified as having a greater expression in t(8;21) AML patients than in non-t(8;21) patients or normal healthy donors.	165
Table 2	The top 30 "present in at least 1 sample" and "present and changing" probe sets identified from the 305 sample dataset as upregulated in t(8;21) AML patients samples when compared to non-t(8;21) patients and to healthy donor samples by fold change (FC>3).	166
Table 3	The number of probe sets identified as having a greater expression in t(8;21) AML patients than in non-t(8;21) patients or normal healthy donors through SAM and ANOVA analysis.	169
Table 4	The 41 probe sets identified from the 305 sample dataset as upregulated in t(8;21) AML patients samples when compared to non-t(8;21) patients and to healthy donor samples by fold change (FC>3) and by SAM analysis and ANOVA analysis with a FDR of 0.05.	171
Table 5	The statistical significance of the mean normalised expression levels of the candidate genes measured in the t(8;21) AML diagnostic patient samples and the healthy normal donor samples.	177
Table 6	A summary of the correlations between the <i>PRAME</i> , <i>POU4F1</i> , <i>ETO</i> and <i>IL5RA</i> MRD data and the t(8;21) fusion gene and <i>WT1</i> normalised copy number data.	184

Figure number	Title	Page number
Table 7	The probe set lists identified from the gene expression profiling analysis (Section 4.3.1).	185
Table 8	The number of probe sets identified as having a greater expression in inv(16) AML patients than in non-inv(16) patients or normal healthy donors.	188
Table 9	The top 30 "present in at least 1 sample" and "present and changing "probe sets identified from the 305 sample dataset as upregulated in inv(16) AML patients samples when compared to non-inv(16) patients and to healthy donor samples by fold change (FC>3).	189
Table 10	The number of probe sets identified as having a greater expression in inv(16) AML patients than in non-inv(16) patients or normal healthy donors through SAM and ANOVA analysis.	192
Table 11	The 66 probe sets identified from the 305 sample dataset as upregulated in inv(16) AML patients samples when compared to non-inv(16) patients and to healthy donor samples by fold change (FC>3) and by SAM analysis and ANOVA analysis with a FDR of 0.05.	194
Table 12	The statistical significance of the mean normalised expression levels of the candidate genes measured in the inv(16) AML diagnostic patient samples and the healthy normal donor samples.	200
Table 13	A summary of the correlations between the <i>MYH11</i> and <i>MN1</i> MRD data and the inv(16) fusion gene and <i>WT1</i> normalised copy number data.	207
Table 14	The probe set lists identified from the gene expression profiling analysis (Section 4.3.3).	208

Chapter 5

Figure number	Title	Page number
Figure 1	Schematic of the method used by the outlier programme to identify and remove outliers from the data set.	234
Figure 2	Dot plot of the relationship between the gene expression values from the original core dataset and from the outlier “free” core dataset.	238
Figure 3	Heatmap of the 552 probe sets identified as being differentially expressed between the normal karyotype AML patient samples and the healthy donor samples (by Bayesian <i>t</i> -test ($p < 0.05$) and a fold change calculation ($FC > 4$) in the core dataset.	240
Figure 4	An unsupervised cluster analysis of the 1000 most variable genes identified from the expanded (with extra donors) dataset in the 54 healthy donor samples before batch correction and after a batch correction was applied.	242
Figure 5	Heatmap of the 552 probe sets previously identified from the core dataset as having differences in expression between the normal karyotype AML patient samples and the healthy donor samples by Bayesian <i>t</i> -test ($p < 0.05$) and a fold change calculation ($FC > 4$) applied to the batch-corrected expanded (with extra donors) dataset.	243
Figure 6	Heatmap of the 1577 probe sets identified as being differentially expressed between the normal karyotype AML patient samples and the healthy donor samples through a Bayesian <i>t</i> -test ($p < 0.05$) and a fold change calculation ($FC > 4$) in the expanded (with extra donors) dataset.	245
Figure 7	Heatmap of the 25 probe sets identified as being up regulated in the normal karyotype AML patient samples compared to the healthy donor samples through a Bayesian <i>t</i> -test ($p < 0.05$) and a fold change calculation ($FC > 4$) in the expanded (with extra donors) dataset.	247
Figure 8	A heatmap of the 25 probe sets identified as being up regulated in the normal karyotype AML patient applied to a new expanded dataset containing the 54 previously used healthy donor samples and 124 new AML normal karyotype samples from two independent studies.	249

Figure number	Title	Page number
Figure 9	A relevance interaction network (RIN) showing the proteins coded by the genes identified as up-regulated in the normal karyotype AML diagnostic samples compared to the healthy donor samples from the expanded (with donors) dataset and their involvement in binding interactions.	252
Figure 10	A relevance interaction network (RIN) showing the proteins encoded by the genes identified as up-regulated in the normal karyotype AML diagnostic samples compared to the healthy donor samples from the expanded (with donors) dataset and their interactions in the regulation of transcription and the network they form.	253
Figure 11	Schematics of the pathway maps (A) and diseases (B) identified by GeneGO MetaCore and the statistical significance of the association of the up-regulated genes with these pathways and diseases.	255
Table 1	The frequencies and percentages of the clinical and molecular characteristics of the 69 patients diagnosed with normal karyotype AML in the core dataset and the expanded (with extra donors) dataset.	230
Table 2	The frequencies and percentages of the clinical and molecular characteristics of the 124 patients diagnosed with normal karyotype AML in the expanded (with public data) dataset from the GEO database (GSE9476 and GSE1159).	231
Table 3	The 25 probe sets identified from the expanded (with extra donors) dataset as upregulated in normal karyotype AML patients samples when compared to healthy donor samples by fold change ($FC > 4$) and a Bayesian <i>t</i> -test ($P > 0.05$).	246
Table 4	The pathway maps that the genes from the 25 up-regulated probe set list were identified as having a statistically significant association with by GeneGO MetaCore.	254
Table 5	The statistically significant top 10 diseases that the genes from the 25 up-regulated probe set list were identified as having a statistically significant association with by GeneGO MetaCore.	254

Chapter 6

Figure number	Title	Page number
Figure 1	The amplification curves of the primer sets used for <i>HOXA5</i> , <i>HOXA9</i> , <i>HOXB6</i> , <i>WT1</i> and <i>COL4A5</i> and of the control gene, <i>ABL</i> , generated from serial dilutions of cell line cDNA to compare their amplification efficiencies.	277
Figure 2	Scatter plots of the spread of the normalised expression levels of the candidate genes measured in the normal karyotype AML diagnostic patient samples and the healthy normal donor samples.	280
Figure 3	Scatter plots of the spread of the normalised expression levels of the candidate genes measured in the <i>FLT3</i> -ITD positive normal karyotype AML diagnostic patient samples (<i>FLT3</i> -ITD +ve) and the <i>FLT3</i> -ITD negative normal karyotype AML diagnostic patient samples (<i>FLT3</i> -ITD –ve).	281
Figure 4	Scatter plots of the spread of the normalised expression levels of the candidate genes measured in the mutant <i>NPM1</i> normal karyotype AML diagnostic patient samples (Mutant) and the wild-type <i>NPM1</i> normal karyotype AML diagnostic patient samples (Wild-type).	282
Figure 5	Longitudinal monitoring of <i>HOXA5</i> , <i>HOXA9</i> , <i>HOXB6</i> , <i>WT1</i> and <i>COL4A5</i> in six normal karyotype AML patients.	288
Figure 6	Longitudinal monitoring of <i>HOXA5</i> , <i>HOXA9</i> , <i>HOXB6</i> , <i>WT1</i> and <i>COL4A5</i> in six normal karyotype AML patients.	289
Table 1	The frequencies and percentages of the clinical and molecular characteristics of the 52 patients diagnosed with normal karyotype AML in the sample set for RQ-PCR gene expression validation.	272
Table 2	The frequencies and percentages of the clinical and molecular characteristics of the 6 normal karyotype AML patients in the sample set for MRD detection.	273

Figure number	Title	Page number
Table 3	The statistical significance of the mean normalised expression levels of the candidate genes measured in the 52 normal karyotype (NK) AML diagnostic patient samples and the 15 healthy normal donor samples.	283
Table 4	The statistical significance of the mean normalised expression levels of the candidate genes measured in the 19 <i>FLT3</i> -ITD positive NK AML diagnostic patient samples (ITD +ve) and the 25 <i>FLT3</i> -ITD negative NK AML diagnostic patient samples (ITD -ve).	283
Table 5	The statistical significance of the mean normalised expression levels of the candidate genes measured in the 19 mutant <i>NPM1</i> NK AML diagnostic patient samples (Mutant) and the 20 wild-type <i>NPM1</i> NK AML diagnostic patient samples (Wild-type).	284

Chapter 7

Figure number	Title	Page number
Figure 1	The effect of ATRA on the growth (a) and viability (b) of NB4 cells over 72 hours when treated with 10^{-7} M ATRA at day 0 only and when treated at day 0, 1, 2 and 3 compared to untreated cells.	313
Figure 2	The effect of 10^{-7} M ATRA on the expression of <i>HOXA1</i> , <i>PRAME</i> , <i>MN1</i> and <i>HGF</i> in NB4, the t(15;17) cell line, cells over 72 hours when treated only at day 0 and when treated at day 0, 1, 2 and 3 compared to untreated cells.	315
Figure 3	The effect of ATRA on the expression of <i>HOXA1</i> , <i>PRAME</i> , <i>MN1</i> and <i>HGF</i> in HL-60, the promyelocytic cell line, cells over 72 hours when treated only at day 0 and when treated at day 0, 1, 2 and 3 compared to untreated cells.	315
Figure 4	The effect of ATRA on the expression of <i>HOXA1</i> , <i>PRAME</i> , <i>MN1</i> and <i>HGF</i> in Me-1, the inv(16) cell line, cells over 72 hours when treated only at day 0 and when treated at day 0, 1, 2 and 3 compared to untreated cells.	316

Figure number	Title	Page number
Figure 5	The effect of NK4 on growth (a) and viability (b) of NB4 cells over 72 hours when treated with 0.8µg/ml of NK4 at 0 hours and at 24 hours after initial treatment. Experiments were performed in triplicate and the error bars represent +/- one standard deviation from the mean.	317
Figure 6	The effect of 0.8µg/ml NK4 on the expression of cell surface markers of NB4 cells over 72 hours when treated at 0 hours and 24 hours after initial treatment. Experiments were performed once due to a lack of the NK4 inhibitor.	318

Contents

Chapter 1 Introduction.....	1
1.1 Haematopoiesis	2
1.1.1 Haematopoietic Hierarchy.....	2
1.1.2 Bone Marrow Microenvironment	5
1.1.3 Regulation of Haematopoiesis	7
1.2 Disorders of Haematopoiesis	13
1.2.1 Types of Leukaemia.....	13
1.2.2 Process of Myeloid Leukaemogenesis.....	17
1.3 Acute Myeloid Leukaemia.....	24
1.3.1 French-American-British (FAB) Cooperative Group	25
1.3.2 Cytogenetics of AML	26
1.3.3 Gene Mutations in AML.....	34
1.3.4 Aberrant Gene Expression in AML.....	39
1.3.5 WHO	42
1.3.6 Current and Potential AML Therapies.....	45
1.4 Minimal residual disease (MRD).....	54
1.4.1 Detection of MRD.....	55
1.4.2 Current MRD markers	64
1.5 Gene Expression Profiling.....	68
1.5.1 Early microarray studies.....	68
1.5.2 Identification of AML-specific sub-groups.....	69
1.5.3 Risk Classification and therapeutic targeting	75
1.6 Project aims and objectives	75
 Chapter 2 Materials and Methods	 77
2.1 Materials	78
2.1.1 Primer sequences	78
2.1.2 Suppliers	79
2.1.3 Enzymes and PCR reagents.....	79
2.1.4 Commercial Kits.....	80
2.1.5 Composition of Solutions	81
2.1.6 Antibodies	81
2.1.7 Laboratory Equipment.....	82
2.1.8 Cell Lines	82
2.2 Patient Database and Gene Expression Profiling.....	83
2.2.1 Patient and Donor Samples	83
2.2.2 Microarray Chips.....	85
2.2.3 From Patient to Chip	85
2.2.4 Data Normalisation	96

2.3 Tissue Culture.....	97
2.3.1 Cell counts and viability.....	97
2.3.2 Cryopreserving of lines	98
2.3.3 Extraction of Mononuclear Cells from Bone Marrow Aspirates and Peripheral Blood Samples.....	99
2.4 RNA: From Extraction to cDNA	100
2.4.1 RNA Extraction	100
2.4.2 RNA quantification and visualisation.....	104
2.4.3 Reverse Transcription (RT) of RNA	105
2.5 Real-Time Quantitative-PCR (RQ-PCR)	106
2.5.1 Primer design.....	109
2.5.2 Magnesium Chloride (MgCl ₂) Concentration and Annealing Temperature Optimisation	110
2.5.3 Optimised LightCycler PCR	112
2.6 Treatment of cell lines with NK4 and ATRA	113
2.6.1 Treatment of cells with ATRA.....	113
2.6.2 Determination of HGF protein concentration.....	113
2.6.3 Treatment of cells with NK4	115
2.7 siRNA.....	115
2.7.1 siRNA design and production.....	115
2.7.2 siRNA annealing	115
2.7.3 siRNA labelling.....	116
2.7.4 Transfection of siRNA	117
2.7.5 Determination of siRNA transfection efficiency	118
2.8 Immunophenotyping	118
2.8.1 Markers for NB4 cell lines	118
2.8.2 Flow analysis	119
 Chapter 3 Using gene expression profiling to identify novel minimal residual disease markers specific for t(15;17) acute myeloid leukaemia - a “proof of principle”	 120
3.1 Introduction.....	121
3.1.1 Acute Promyelocytic Leukaemia (APL).....	121
3.1.2 Minimal Residual Disease and APL	122
3.2 Methods.....	124
3.2.1 Patient and donor samples	124
3.2.2 Gene Expression Data Analysis.....	124
3.2.3 Diagnostic Screening of Candidate Genes	127
3.2.4 Validation of Candidate Genes as MRD Markers.....	127
3.2.5 Statistical Analysis of Diagnostic Screening and MRD Analyses	129

3.3 Results	129
3.3.1 Identification of genes expressed specifically in t(15;17) AML samples	129
3.3.2 Identification of Genes Differentially Expressed in t(15;17) AML Samples using Statistical Tools.....	136
3.3.3 Validation of Identified Genes by Real-Time Quantitative PCR	138
3.4 Discussion	149
 Chapter 4 Using gene expression profiling to identify novel minimal residual disease markers specific to CBF leukaemias ..	156
4.1 Introduction.....	157
4.1.1 Core Binding factors (CBF)	157
4.1.2 CBF Leukaemia	157
4.1.3 CBF Leukaemia and Gene Expression Profiling	158
4.1.4 CBF Leukaemia and MRD	159
4.1.5 Wilms' Tumour (WT1) as an MRD Marker	159
4.2 Methods.....	160
4.2.1 Patient and donor samples	160
4.2.2 Gene Expression Data Analysis.....	161
4.2.3 Diagnostic Screening of Candidate Genes	161
4.2.4 Validation of Candidate Genes as MRD Markers.....	162
4.3 Results	163
4.3.1 Identification of genes expressed specifically in t(8;21) AML samples	163
4.3.2 Validation of Candidate Genes as MRD Markers.....	174
B. inv(16) AML	186
4.3.3 Identification of genes expressed specifically in inv(16) AML samples	186
4.3.4 Validation of Candidate Genes as MRD Markers.....	196
4.3.5 Correlation between Candidate Markers and Current Markers	207
4.4 Discussion	209
4.4.1 t(8;21) AML	210
4.4.2 inv(16) AML.....	216
 Chapter 5 Gene Expression profiling of Normal Karyotype AML	222
5.1 Introduction.....	223
5.1.1 Cytogenetics	223
5.1.2 Minimal Residual Disease Markers.....	225
5.1.3 Gene Expression Profiling.....	226
5.2 Materials and Methods.....	227
5.2.1 Patient and donor samples	227
5.2.2 Data Analysis	232

5.3 Results	236
5.3.1 Expression profiling to identify MRD markers in normal karyotype AML patients.	236
5.3.2 Improving the estimate of fold change and its importance in MRD marker identification.....	237
5.3.3 Identification of MRD markers in the core dataset	238
5.3.4 Improved identification of MRD markers using an expanded donor dataset.....	241
5.3.5 Validation of identified MRD markers using publicly available datasets	247
5.3.5 Validation of identified MRD markers using publicly available datasets	248
5.3.6 Pathway analysis of identified genes	249
5.3.6 Pathway analysis of identified genes	250
5.4 Discussion	256
 Chapter 6 Monitoring MRD in Normal Karyotype AML	266
6.1 Introduction.....	267
6.1.1 MRD Detection.....	267
6.1.2 HOX Genes and MRD	268
6.1.3 WT1	269
6.1.4 COL4A5	270
6.2 Materials and Methods	271
6.2.1 Patient and Donor samples.....	271
6.2.2 Quantitative Real-Time PCR.....	274
6.3 Results	275
6.3.1 Validation of Identified Genes by RQ-PCR	275
6.4 Discussion	290
 Chapter 7 The treatment of AML cell lines with differentiation agents and gene-specific inhibitors.....	303
7.1 Introduction.....	304
7.1.1 Hepatocyte growth factor (HGF)	304
7.1.2 Preferentially expressed antigen in melanoma (PRAME) ..	306
7.1.3 Meningioma 1 (MN1)	306
7.1.4 HOXA1	307
7.1.5 All-trans retinoic acid (ATRA).....	307
7.1.6 NK4.....	308
7.1.7 siRNA.....	308
7.1.8 Cell surface markers	309
7.2 Materials and Methods	310
7.2.1 The effect of ATRA upon HGF expression in AML Cells....	310
7.2.2 The effect of NK4 upon HGF expression in AML cell lines	311
7.2.3 The effect of RNA interference upon the expression of genes identified as being specific to AML sub-groups in AML cell lines.	312

7.3 Results	312
7.3.1 The effect of ATRA on the growth of AML cells and the expression of selected genes.....	312
7.3.2 The effect of the HGF-antagonist NK4 on NB4 cells.....	316
7.3.3 Measurement of transfection efficiency of FAM-labelled siRNA against selected genes in NB4 cells	318
7.4 Discussion	319
 Chapter 8 Discussion	 323
 Appendix.....	 332
Appendix I: Supporting Ethical Permissions	333
Appendix II: Gene Lists.....	335
Appendix III: Cel Files	340
 References.....	 341

Chapter 1

Introduction

1.1 Haematopoiesis

Mature blood cells incapable of further growth or that are damaged are replaced constantly through the tightly controlled process of haematopoiesis. All mature blood cells are derived from haematopoietic stem cells (HSC). Early studies of lethally irradiated adult mice transplanted with bone marrow cells demonstrated the existence of a cell population able to sustain haematopoiesis (Ford *et al*, 1956). HSCs have the capacity to self-renew in order to generate more HSCs and to differentiate into progenitor cells leading to the development of cells of all lineages; B- and T-cells, Natural killer (NK) cells, erythrocytes, granulocytes and monocytes (Sorrentino, 2004;Bellantuono, 2004).

1.1.1 Haematopoietic Hierarchy

The process of haematopoiesis is hierarchically organised. HSCs give rise to a series of progenitors that become progressively more restricted in their differentiation capacity and lose the ability to self-renew as they progress along one of two lineages. The common lymphoid progenitors (CLP) give rise to only the B-, T- and NK cells (lymphocytes) whereas the common myeloid progenitors (CMP) give rise to the bipotent megakaryocyte/erythrocyte progenitors (MEP) and granulocyte/macrophage progenitors (GMP) (Kondo *et al*, 1997;Akashi *et al*, 2000). From *in vitro* culture assays of human bone marrow, multipotent progenitor cells termed colony-forming units (CFU) have been identified and characterised by their ability to form mixed

haematopoietic colonies that give rise to blood cells of multiple lineages. The early multipotent progenitors CFU-GEMM (colony-forming units – granulocyte-erythroid-macrophage-megakaryocyte) give rise to the granulocytes, erythrocytes, megakaryocytes and monocytes (Fauser & Messner, 1978; Fauser & Messner, 1979) (Figure 1).

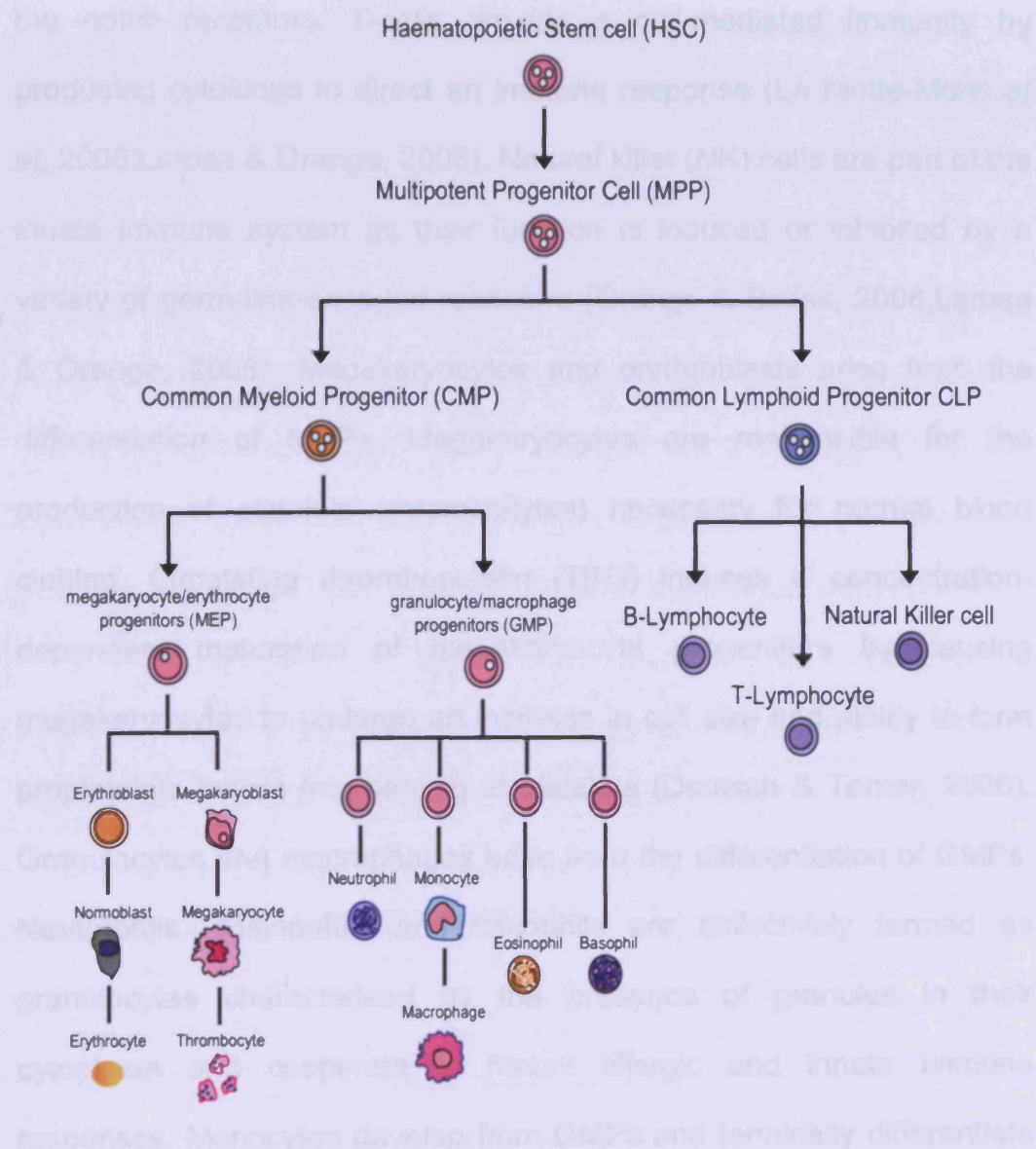


Figure 1. The production of multiple lineages from a single stem cell during normal haematopoiesis.

The terminally differentiated lymphocytes play major roles in the body's immune response system. B-cells develop in the bone marrow but reach their full maturity in peripheral blood organs. They provide a humoral immunity through the production of antibodies against extracellular pathogens (Larosa & Orange, 2008). T-cells develop from progenitors that have migrated to the thymus and received signals from the notch receptors. T-cells provide a cell-mediated immunity by producing cytokines to direct an immune response (La Motte-Mohs *et al*, 2005;Larosa & Orange, 2008). Natural killer (NK) cells are part of the innate immune system as their function is induced or inhibited by a variety of germ-line-encoded receptors (Orange & Ballas, 2006;Larosa & Orange, 2008). Megakaryocytes and erythroblasts arise from the differentiation of MEPs. Megakaryocytes are responsible for the production of platelets (thrombocytes) necessary for normal blood clotting. Circulating thrombopoietin (TPO) induces a concentration-dependent maturation of megakaryocyte progenitors by causing megakaryocytes to undergo an increase in cell size and ploidy to form proplatelets before fragmenting in platelets (Deutsch & Tomer, 2006). Granulocytes and macrophages arise from the differentiation of GMPs. Neutrophils, eosinophils and basophils are collectively termed as granulocytes characterised by the presence of granules in their cytoplasm and cooperate to mount allergic and innate immune responses. Monocytes develop from GMPs and terminally differentiate into macrophages and dendritic cells.

1.1.2 Bone Marrow Microenvironment

In the bone marrow, HSCs reside in a special microenvironment “niche” where they are supported by cell-cell interactions with non-haematological cells to maintain their cell functions of self-renewal and differentiation. Osteoblasts line the endosteal surface in the bone marrow and function as the endosteal niche to regulate and maintain HSC numbers. Several mechanisms of osteoblast control *in vivo* and *in vitro* have been identified. The activation of parathyroid hormone/parathyroid hormone-related peptide receptors (PPRs) stimulates the increase in the number of osteoblasts, which in turn produce high levels of the Notch ligand, jagged 1, that increases the number of HSCs (Calvi *et al*, 2003). Additionally, osteoblasts express bone morphogenic protein (BMP) receptor type IA through which the BMP signalling pathway acts and regulates the size of the endosteal niche to control the HSC numbers (Zhang *et al*, 2003). Angiopoietin 1 (Ang-1) in osteoblasts contribute to the quiescent and anti-apoptotic state of HSCs by interacting with a tyrosine kinase receptor, Tie-2, expressed in HSCs in the bone marrow. Tie-2/Ang-1 signalling enhances the adhesion of HSCs and osteoblasts resulting in the promotion and maintenance of HSC quiescence (Arai *et al*, 2004).

However, not all HSCs associate with osteoblasts and the mobilisation of HSCs into peripheral circulation suggests an association with blood vessels. Immunohistochemical analysis of primitive cells identified a number of HSCs associated with sinusoidal endothelial cells of

parafollicular areas of mobilised murine spleen and bone marrow (Kiel *et al*, 2005). The vascular endothelial cells of this “vascular niche” have been implicated in promoting the differentiation, proliferation and mobilisation of HSCs and their progenitor cells. Regions of vascular bone marrow endothelium cells (BMEC) are highlighted as areas of megakaryocyte regulation. *In vivo*, polyploid megakaryocytes localise to sinusoidal BMECs to form proplatelets and fragment into platelets that are released into the bone marrow vascular sinusoids (Zucker-Franklin & Philipp, 2000;Avecilla *et al*, 2004). The growth factor thrombopoietin (TPO) is suggested as the most effective cytokine supporting the proliferation and maturation of megakaryocyte progenitor cells. TPO is seen to stimulate the proliferation and differentiation of haematopoietic progenitors into megakaryocytes and platelets *in vitro* and *in vivo* (Kaushansky *et al*, 1994;Lok *et al*, 1994), but further *in vivo* evidence suggests TPO is not essential. An *in vivo* study of thrombopoietin (TPO)-deficient mice observed how the interaction of megakaryocytes with bone marrow endothelial cells (BMEC) mediated by megakaryocyte-active cytokines in TPO^{-/-} mice promoted TPO-independent maturation of megakaryocytes and the production of platelets. The cytokine fibroblast growth factor 4 (FGF-4) supported the adhesion of megakaryocytes to sinusoidal BMECs to increase their survival and maturation whilst stromal-derived factor-1 (SDF-1) promoted the affinity and migration of platelets across the BMECs (Avecilla *et al*, 2004).This data suggests that TPO is not essential for the regulation of megakaryocytes and BMECs of the vascular niche are

able to provide support for the differentiation and proliferation of HSCs through alternative factors.

1.1.3 Regulation of Haematopoiesis

a. Transcription Factors

Transcription factors contribute to the proliferation and differentiation signals required for haematopoiesis by regulating the expression of genes. The expression of each transcription factor involved in regulating haematopoiesis is not necessarily restricted to single cell lineage. For example, the development of erythrocytes and megakaryocytes are both under regulation of the zinc-finger transcription factor GATA-1. The expression of *GATA-1* mRNA coding for GATA-1 is detected in multipotent progenitors and the erythroid, megakaryocyte, eosinophil and mast cell lineages (Tsai *et al*, 1989;Romeo *et al*, 1990;Zon *et al*, 1993). Disruption of the *GATA-1* gene in mouse embryonic stem cells results in a block in the production of mature erythroid cells and whilst *GATA-1*⁻ embryonic stem cells can generate erythroid colonies containing proerythroblasts, they are unable to mature properly and apoptose (Pevny *et al*, 1991;Weiss & Orkin, 1995). Shivdasani *et al* (1997) also demonstrated a development of megakaryocytes with reduced ploidy and a failure to generate platelets in mice lacking expression of *GATA-1* (Shivdasani *et al*, 1997). Pu.1 (Purine rich box-1) belongs to the Ets (erythroblastosis)-family of transcription factors. Evidence of Pu.1 as a key regulator of haematopoiesis has been provided from mice knock-out studies. *Pu.1*^{-/-} mice lacked cells from

the myeloid and lymphoid lineages resulting in prenatal lethality (Scott *et al*, 1994). Transcription factors also contribute towards the self-renewal capacities of HSCs. For example, Homeobox (*HOX*) genes encode DNA-binding transcription factors. They are organised into four genomic clusters (A-D) and were first identified as regulators of positional identity along the anterior-posterior body axis in embryos (as reviewed by (Krumlauf, 1994). The majority of the A, B and C clusters are expressed in haematopoietic cells, and most of these *HOX* genes are preferentially expressed in HSCs and primitive progenitor cells, decreasing in expression as the cells differentiate and mature (Moretti *et al*, 1994; Sauvageau *et al*, 1994; Pineault *et al*, 2002). Southern blot analysis of cDNA identified *HOX* genes located at the 3' regions of the A and B clusters as enhanced in expression in primitive CD34⁺ cells whereas those located at the 5' end were expressed at relatively equal levels in all the primitive CD34⁺ sub-populations (Sauvageau *et al*, 1994). Mice knock-out studies have demonstrated the various effects of *HOX* genes on proliferation, differentiation and HSC-renewal. Knock-out studies revealed subtle reductions in the HSC and progenitor cell numbers without disrupting haematopoiesis in *HOXB4*-deficient mice and a more pronounced decrease in the proliferation capacity of these cells in *HOXB3*- and *HOXB4*-deficient mice (Bjornsson *et al*, 2003; Brun *et al*, 2004). *HOXB3* and *HOXB4* are both highly expressed in primitive CD34⁺ cells, but a decrease in their expression is observed as the cells differentiate and mature (Sauvageau *et al*, 1994). Despite their similar expression patterns and a decrease in HSC proliferation resulting from

their underexpression, the overexpression of these two genes generates very different results. An overexpression of *HOXB3* blocks the development of T- and B-cells whilst an overexpression of *HOXB4* results in the selective expansion of primitive haematopoietic cells (Sauvageau *et al*, 1995; Sauvageau *et al*, 1997).

b. Cytokines

Many of the regulatory pathways controlling the haematopoietic system are controlled by cytokines, a large family of extracellular ligands, binding to and activating a family of cytokine receptors. Cytokines are secreted from or presented on the surface of mesenchymal cells within the bone marrow, known as stromal cells, to promote the survival, proliferation and differentiation of haematopoietic stem cells and progenitors. Colony-stimulating factors (CSF) are synthesised locally by stromal cells and include macrophage-CSF (M-CSF), granulocyte-CSF (G-CSF) and granulocyte-macrophage-CSF (GM-CSF). M-CSF and G-CSF are relatively lineage-specific contributing to the differentiation and proliferation of macrophages and granulocytes. GM-CSF is produced by osteoblasts at relatively low levels and functions at earlier stages of lineage commitment, promoting the proliferation and maturation of granulocytes and macrophages and working with other cytokines as a growth factor for erythroid and megakaryocyte progenitors (Barreda *et al*, 2004). The multi-CSF is more commonly known as interleukin-3 (IL-3) contributing to the production of macrophages, neutrophils, eosinophils, basophils, mast cells, megakaryocytes, and erythrocytes

(as reviewed by (Barreda *et al*, 2004). Several cytokines work in synergy with other cytokines to promote haematopoiesis. Stem cell factor (SCF) functions with G-CSF, GM-CSF, IL-3 and EPO (erythropoietin) to support the growth of colony-forming units (CFU) in semisolid media (Martin *et al*, 1990).

c. Cytokine Receptors

Cytokines interact with specific membrane receptors that transmit a series of intracellular signals to the target cell following the cytokine-receptor interaction. The receptors for M-CSF and SCF, M-CSFR (also known as c-fms) and c-kit respectively, belong to the class III receptor tyrosine kinase (RTK) family (Yarden *et al*, 1987). Class III receptor RTKs are characterised by an extracellular ligand-binding domain containing five immunoglobulin-like domains, a single transmembrane domain, a juxtamembrane domain, two intracellular domains and a C-terminal domain (Ullrich & Schlessinger, 1990). Binding of a ligand to a class III RTK results in the activation of intrinsic tyrosine kinase activity (Ullrich & Schlessinger, 1990; Weiss & Schlessinger, 1998). The receptors mediating the effects of G-CSF, GM-CSF and IL-3 all belong to the type 1 cytokine receptor superfamily. Unlike M-CSF, G-CSFR, the receptor for G-CSF, has no intrinsic kinase activity. The GM-CSF receptor, GM-CSFR, and the IL-3 receptor are members of the gp140 family of type 1 cytokine receptors composed of two distinct chains; α and β . The α -chain is the primary binding chain of the ligands whilst the β -chain is necessary for signal transduction (Chiba *et al*, 1990; Barreda

et al, 2004). Binding of the ligands to their receptors results in the activation of kinases from the Janus kinase family (JAK) resulting in a cascade of phosphorylation events through the initiation of several pathways, for example the p21Ras (rat sarcoma)/MAP (mitogen-activated protein) kinase and PI-3 (phosphoinositol-3) kinase/PKB (protein kinase B) pathways, to promote cell survival, proliferation and differentiation (Ihle & Kerr, 1995).

d. Cell Cycle Regulation

An important characteristic of haematopoietic stem cells is their relative proliferative quiescent state. The clonal-succession model proposed by Kay *et al* (1965) hypothesised that only one or a few HSC clones from the large pool of HSCs gives rise to mature blood cells at any time whilst the remainder of HSC clones remain quiescent until they are needed due to the exhaustion of the proliferative capacity of the long-term HSC (LT-HSC) clone (Kay, 1965). Studies by Abkowitz *et al* (1990) and Gutterp *et al* (1990) supported the clonal-succession model. Safari cats heterozygous for the X chromosome linked enzyme glucose-6-phosphate dehydrogenase (G6PD) received autologous marrow transplantations with limited numbers of cells following lethal doses of radiation. The data suggests that haematopoiesis resumed to the state observed before irradiation but fewer stem-cell clones were contributing (Abkowitz *et al*, 1990;Gutterp *et al*, 1990). A later *in vivo* model studying the bromodeoxyuridine (BrdU) incorporation by replicating cells did not support the model. Cheshier *et al* (1999) proposed that long term-HSCs

enter the cell cycle in an asynchronous manner and that a few clones did not dominate proliferation. From their results they calculated that 99% of long term-HSCs divided, on average, every 57 days (Cheshier *et al*, 1999). The self-renewal capacity of HSCs requires maintaining the coordination of HSC progression through the cell cycle and HSC-fate choices. A direct control of the HSC cycle is by the activation of a group of enzymes known as cyclin-dependent kinases (CDK). CDKs are positive regulators inducing cell cycle progression; their activity partly controlled by their association with cyclins to form a cyclin-dependent kinase complex. The activities of the complexes are further regulated by cyclin-dependent kinase inhibitors (CDKIs) acting as negative regulators inhibiting the progression of the cell cycle. CDKIs belong to one of two families depending on their targets; the Cip/Kip family, including p21 and p27, inhibit a number of cyclin-dependent kinases whilst the Ink4 (inhibitors of CDK4) family, including p15 and p16, inhibits CDK4 and CDK6 (Harper *et al*, 1995; Sherr & Roberts, 1999; Attar & Scadden, 2004). The Cip/Kip family members inhibit the kinase activity of the cyclin/CDK2 complexes and over-expression of these inhibitors induces a cell cycle arrest in the G₁ phase (Harper *et al*, 1995; Lee & Yang, 2001). The p21 family member is under transcriptional control of p53 and appears to play a role in the inhibition of stem cell proliferation acting as a switch. Early studies with irradiated mice demonstrated the targeted disruption of the gene encoding p21 impaired the ability of cells to achieve cell cycle arrest (Deng *et al*, 1995). Study of mice deficient in p21 demonstrated the number of

primitive haematopoietic cells in the G₀ phase was reduced yet the number of primitive cells that gave rise to long-term multipotent colonies in culture was increased indicating the role p21 in controlling the quiescence of HSCs (Cheng *et al*, 2000).

1.2 Disorders of Haematopoiesis

A characteristic hallmark of leukaemia is the inhibition of cell differentiation. Unable to terminally differentiate, these cells instead retain their proliferative capacity (Andreeff, 1986). Some cells may only display the phenotype of hyperproliferation (myeloproliferative disorders) or defective differentiation (myelodysplasia) and therefore are not true leukaemias but these disruptions in the haematopoietic process can result in the progression to several forms of leukaemia (Sawyers *et al*, 1991). Depending on the symptoms patients present with, and the cells involved, leukaemias can be sub-grouped into acute and chronic and further classified depending on the cell of origin into lymphoid and myeloid leukaemia (Sawyers *et al*, 1991).

1.2.1 Types of Leukaemia

a. Myeloproliferative Disorders

Myeloproliferative disorders encompass clonal haematopoietic syndromes of increased proliferation and accumulation of erythroblasts, granulocytes and megakaryocytes in the bone marrow and peripheral blood (Dameshek, 1951). Myeloproliferative neoplasms (MPN) include

chronic myelogenous leukaemia (CML), polycythaemia vera (PV), essential thrombocythaemia (ET), primary myelofibrosis (PMF), mastocytosis, chronic eosinophilic leukaemia-not otherwise specified, chronic neutrophilic leukaemia, and “MPN, unclassifiable” according to the 2008 World Health Organization (WHO) classification system (Vannucchi *et al*, 2009;Vardiman *et al*, 2009). A single point mutation in the *JAK2* (Janus Kinase 2) gene, a valine to phenylalanine substitution at codon 617 (*JAK2V617F*), has been identified in a large proportion of patients with PV, ET and PMF (Baxter *et al*, 2005;Levine *et al*, 2005). The substitution results in the constitutive phosphorylation of *JAK2*. Mice transplanted with haematopoietic stem cells containing *JAK2V617F* rapidly develop an erythrocytosis which has identified *JAK2V617F* as a primary molecular event in the development of PV (James *et al*, 2005b;James *et al*, 2005a).

b. Myelodysplastic Syndromes

The myelodysplastic syndromes (MDS) are clonal haematopoietic cell disorders characterised by bone marrow failure leading to cytopenia with defects in erythroid, myeloid, and megakaryocytic maturation (Janssen *et al*, 1989). MDS are susceptible to undergoing transformation to a more acute phase with an increased risk of patients developing acute myeloid leukaemia (AML) (Bennett *et al*, 1982;Malcovati & Nimer, 2008). MDS is generally classified as “primary” MDS (*de novo*) or “therapy-related” MDS (t-MDS) in patients previously treated with chemotherapy and/or radiotherapy (Mijovic & Mufti, 1998).

Due to the challenges associated with accurate diagnosis, the World Health Organisation (WHO) has refined the previous MDS classification to include 10 sub-classifications of MDS (Table 1) (Vardiman *et al*, 2009). Each sub-classification is associated with an increasing risk in leukaemic transformation. MDS is thought to arise from abnormalities resulting in defective haematopoietic stem cell self-renewal and differentiation; chromosomal abnormalities have been detected in approximately 48% of patients presenting with MDS and in more than 90% of patients with t-MDS. The most common abnormalities involve interstitial deletions of the long arm of chromosomes 5 (Mufti, 1992; Malcovati & Nimer, 2008).

Myelodysplastic syndrome (MDS)
Refractory cytopenia with unilineage dysplasia Refractory anaemia (RA) Refractory neutropenia Refractory thrombocytopenia Refractory anaemia with ring sideroblasts (RARS) Refractory cytopenia with multilineage dysplasia Refractory anaemia with excess blasts (RAEB) Myelodysplastic syndrome with isolated del(5q) Myelodysplastic syndrome, unclassifiable Childhood myelodysplastic syndrome

Table 1. 2008 WHO classification of myelodysplastic disorders

c. Chronic Leukaemia

Patients with chronic leukaemias usually present with symptoms of fatigue, weight loss and an enlarged spleen. Chronic lymphocytic

leukaemia (CLL) is the most common form of adult leukaemia in western countries and mainly affects elderly individuals with a median age at diagnosis of 64 years (Dighiero & Binet, 2000;Montserrat & Moreno, 2008). CLL is characterised by the accumulation of mature monoclonal CD5⁺ B-lymphocytes in the blood, bone marrow and lymphoid tissues. These cells also express CD19 and CD23 whilst weakly expressing the surface immunoglobulins (Ig) IgM, IgD and CD79b (Dighiero, 2005;Chiorazzi *et al*, 2005). Hairy cell leukaemia (HCL) is an uncommon chronic mature B-lymphocyte disorder named for its characteristic “hairy” appearance of the leukaemic lymphoid cells due to the mixture of fine hair-like villi and membrane ruffles on the cell surface (Polliack, 2002;Tiacchi *et al*, 2006). Chronic myeloid leukaemia (CML) accounts for approximately 15-20% of adult leukaemias (Sessions, 2007). A clonal HSC disorder resulting in an increase in myeloid cells, erythroid cells and platelets in the peripheral blood, CML progresses from a chronic phase (CP) to a rapid blast crisis where cells fail to mature resembling cells found in acute leukaemia (Sawyers, 1999). CML is usually diagnosed from the detection of a translocation occurring between chromosomes 9 and 22, resulting in the Philadelphia (Ph) chromosome (Rowley, 1973).

d. Acute Leukaemia

Acute leukaemia is a heterogeneous group of malignant haematopoietic lymphoid and myeloid progenitor disorders. They are characterised by an increase in circulating cells that are unable to properly mature or

function (Sawyers *et al*, 1991). Patients present with fatigue, bone and joint pain, haemorrhages and infection (Jabbour *et al*, 2005). Acute lymphoblastic leukaemia (ALL) is the commonest leukaemia affecting children, peaking between the ages of 2 and 5 years with an incidence of 4 to 5 cases per 100,000 persons. Also occurring in adults, the incidence of ALL cases gradually rises at approximately 50 years of age but accounts for only 20% of all cases of adult leukaemia (Jabbour *et al*, 2005). Acute myeloid leukaemia (AML) is the commonest form of acute leukaemia affecting adults with an incidence of approximately 4 cases per 100,000 persons (Estey & Dohner, 2006; Teitell & Pandolfi, 2009). The median age of diagnosis of AML is approximately 65 years and above this age the incidence of cases increase to approximately 18 cases per 100,000 persons (Estey & Dohner, 2006). AML is characterised by an accumulation of immature blasts incapable of differentiation. According to the 2008 WHO classification, *de novo* AML is diagnosed when 20% or more blasts are detected in the bone marrow or peripheral blood (Vardiman *et al*, 2009).

1.2.2 Process of Myeloid Leukaemogenesis

a. Multistep Process

The progression of a pre-malignant clone through intermediate steps to eventually becoming a malignant neoplasm is promoted by an initiating event. The majority of malignancies are thought to be initiated by spontaneous or induced somatic mutations but latency periods between the development of a mutation and the transformation into a malignancy

suggests the need for more than one event (Knudson, 2001). Foulds (1954) described a step-wise development of neoplastic tumours of increasing autonomy due to changes in the cellular phenotype (Foulds, 1954) *In vitro models* and *in vivo* mice models of virally-infected leukaemia have helped to observe the multi-step process involved in the transformation of leukaemia. Following a long latency period from infection with a helper-independent Friend Murine Leukemia Virus (F-MuLV), bone marrow cultures and mice developed myeloblastic leukaemia displaying several steps in the progression to malignant transformation. During a “pre-leukaemic” stage the cells in culture displayed an abnormal response to GM-CSF, a promotion in their proliferation of colonies containing both mature granulo-macrophagic and immature myeloblastic cells, and acquired growth autonomy before finally achieving an *in vivo* tumourigenicity (Heard *et al*, 1984).

b. Clonality

Normal haematopoiesis is a process of polyclonal expansion; several stem cells proliferating and differentiating to generate a mixed pool of terminally-differentiated end-cells. However, the majority of haematological disorders result from monoclonal expansions. A genetic event resulting in a growth advantage of one cell over others can lead to the development of a clone. The multistep nature of AML can result in the generation of a range of sub-clones differing in their acquired secondary abnormality. One clone will usually dominate and direct the course of the disease but the application of pressure, for example

chemotherapy, can generate a clonal shift allowing a minor clone to dominate the disease. As relapsed leukaemia is often resistant to chemotherapy, it suggests that the event contributing to a leukaemic relapse is not always the original leukaemia present at diagnosis but instead a leukaemia resulting from a clonal shift and therefore relapse is accompanied by clonal evolution (Nakano *et al*, 1999). Monitoring of both the *FLT3* (Fms-like tyrosine kinase 3) gene and the *N-ras* (neuroblastoma RAS viral (v-ras) oncogene homologue) gene in AML have demonstrated the instability of such aberrations and their role as secondary events with the observation of the acquisition or loss of mutations over the course of the disease in AML patients (Nakano *et al*, 1999; Shih *et al*, 2004b; Chen *et al*, 2005).

Historically, methods used to identify clonality in haematological malignancies have adopted a system of genetic markers. Intrinsic markers, for example cellular markers or somatic cytogenetic aberrations, that have developed during normal cell processes or as part of the disease can act as useful intrinsic markers for clonality studies. The Philadelphia (Ph) chromosome is expressed in the majority of chronic myeloid leukaemic (CML) cells, but in almost no non-CML cells; the fusion product resulting from the translocation of chromosomes 9 and 22. This led to the suggestion that this leukaemia evolved from a single cell in which the aberrant chromosomal event occurred, therefore demonstrating the monoclonality of CML. However, the possibility of CML resulting from a polyclonal expansion can not be

ruled out based on this evidence alone and further analysis with extrinsic markers is required. Extrinsic markers take advantage of cellular mosaicism. The nature of mosaicism allows for markers completely independent of the disease of interest to be studied and to not be restricted to a particular cell lineage (Raskind *et al*, 1998). Mosaicism results from the inactivation of all but one X chromosome in cells containing two or more X chromosomes and occurs during the early stages of embryogenesis. The selection of either the maternal X chromosome or paternal X chromosome for inactivation is usually a random event and is stably transmitted during mitosis to daughter cells. Beutler *et al* (1962) demonstrated the mosaicism of cells in females using the X-linked enzyme G6PD as a marker (Beutler *et al*, 1962). Early studies of X-chromosome inactivation patterns (XCIPs), using isoforms of the G6PD enzyme resulting from a polymorphism, for determining the clonal origin of neoplastic cells were limited to small minority of the female population (Fialkow *et al*, 1967; Fialkow *et al*, 1981). Vogelstein *et al* (1987) were able to extend the use of XCIPs to a wider population using restriction fragment length polymorphisms (RFLP) and methylation patterns of the inactive and active X chromosomes.

c. Target Cells

As previously mentioned, an early study by Fialkow *et al* (1981) demonstrated the clonality of acute non-lymphocytic leukaemia through chromosomal studies of the x-linked enzyme G6PD. The data indicated

some cases of acute myeloid leukaemia (AML) originated from pluripotent stem cells whilst others originated from cells already restricted to the granulocyte/monocyte pathway (Fialkow *et al*, 1981). Later studies by Griffin *et al* (1986) of surface marker analysis further supported the heterogeneity of AML showing AML clonal cells arose from various points in the haematopoietic hierarchy in different patients (Griffin & Lowenberg, 1986; Griffin *et al*, 1986). These models predicted that differences in phenotype of the leukaemic stem cells would be observed between AML patients depending on the origin of their disease (Fialkow *et al*, 1981; Griffin & Lowenberg, 1986). Alternatively, McCulloch *et al* (1983) proposed leukaemias originate from stem cells with the ability to differentiate or acquire surface markers from the influence of transformation agents resulting in little variability of stem cell phenotype between patients (McCulloch, 1983). Fluorescence *in situ* hybridisation (FISH) and flow cytometric analysis studies of MDS and AML samples identified characteristic cytogenetic aberrations in the CD34⁺/CD38⁻ cell compartment providing evidence of primitive cell involvement in MDS and AML (Haase *et al*, 1995; Mehrotra *et al*, 1995). Lapidot *et al* (1994) identified SCID (severe combined immunodeficiency) leukaemia-initiating cells (SL-IC) as AML-initiating cells that could establish human leukaemia in SCID mice. The SL-ICs displayed an expression pattern of CD34⁺/CD38⁻ similar to that observed in normal haematopoietic stem cells and when transplanted into SCID mice they initiated leukaemia, unlike the transplanted CD34⁺/CD38⁺ cells, an expression pattern observed in more

differentiated cells, even when in the presence of AML-CFUs (Lapidot *et al*, 1994). Bonnet *et al* (1997) observed similar results when they transplanted SL-ICs into NOD/SCID (non-obese diabetic-severe combined immunodeficiency) mice. When transplanted, the CD34⁺/CD38⁻ SL-ICs generated large numbers of AML-CFUs and leukaemic blasts expressing the irregular combinations of surface antigens observed in the patient samples from which the cells originated (Bonnet & Dick, 1997). However, In favour of the hypothesis of Fialkow *et al* (1981), recent studies of mouse bone marrow transduced with oncogenic fusion genes have demonstrated the ability of MLL-ENL (mixed-lineage leukaemia - eleven nineteen leukaemia) and MOZ-TIF2 (monocytic leukaemia zinc finger - TGF-beta induced factor-2) to transform not only HSCs but committed myeloid progenitors (Cuzzio *et al*, 2003; Huntly *et al*, 2004). Kirstetter *et al* (2008) demonstrated the potential of leukaemic transformation to occur in progenitors with limited self-renewal capacity. Approximately 9% of *de novo* AML cases present with mutations of the C/EBP α (CCAAT/enhancer binding protein (C/EBP), alpha) gene. C/EBP α is produced of two polypeptides of 30 kDa (p30) and 42 kDa (p42), with most C/EBP α mutations in AML resulting in the loss of p42. Knock-in mice studies demonstrated that whilst GMPs were still generated, a loss of p42 resulted in myeloid progenitors with a vastly increased self-renewal capacity and progression to AML and ultimately death from liver and bone marrow failure. These mouse models demonstrated the role of C/EBP α mutations as AML-initiating events that result in the

generation of leukaemia-initiating cells (LIC) displaying an immunophenotype observed in myeloid-committed cells that generate only myeloid cells in irradiated recipients (Kirstetter *et al*, 2008). Therefore mutations occurring in HSCs or more committed progenitors may both give rise to the leukaemic stem cell.

d. Leukaemic Stem Cell

Early *in vitro* colony assays observed only a minority of proliferative leukaemic blasts (AML-CFU) were able to give rise to colonies. Several studies also identified similar properties of active proliferation, self-renewal and the ability to undergo differentiation of the normal haematopoietic progenitor cells in AML-CFUs *in vitro*. This led to the suggestion that a leukaemic clone exists in a hierarchy of proliferating progenitors differentiating into a population of non-cycling leukaemic blasts similar to that seen in normal haematopoiesis (Minden *et al*, 1978; Buick *et al*, 1979; Pessano *et al*, 1984; Griffin & Lowenberg, 1986; Bonnet, 2005). Evidence of a leukaemic stem cell (LSC) with the ability to initiate and sustain the growth of a leukaemic clone *in vivo* combined with their self-renewal capacity supports the concept of a hierarchical organisation of AML sustained by a small number of leukaemic stem cells transformed from normal haematopoietic cells. Their increased resistance to chemotherapeutic drugs have made it difficult to completely eradicate leukaemic stem cells and relapse often occurs. Similar to normal HSCs, a number of LSCs spend a large proportion of their time resting in the G₀ phase, unlike the AML-CFUs

they generate. Whilst cell cycle-specific chemotherapeutic drugs targeting the leukaemic blasts are able to eradicate the majority of blasts, the LSCs quiescent state protects them allowing relapse to occur due to their self-renewal capacity and generation of new progeny (Guan *et al*, 2003; Ravandi & Estrov, 2006; Misaghian *et al*, 2009). Unfortunately the concept of a hierarchical organisation in leukaemia similar to that in normal haematopoiesis is still a very controversial area. Data supporting the hierarchical organisation hypothesis was obtained through studies performing xenografts. It is argued that cells with a potential to be tumourigenic may require extrinsic factors from the surrounding microenvironment to engraft. Therefore when performing a xenograft between mouse and human, there may be a lack of appropriate microenvironment factors required for engraftment due to species-differences. When assessing the capacity of the grafted cells to initiate the growth of tumours, what may appear non-tumourigenic in the host may actually be tumourigenic under the correct microenvironment conditions in the donor (Rosen & Jordan, 2009).

1.3 Acute Myeloid Leukaemia

Acute myeloid leukaemia (AML) is an extremely heterogeneous disorder characterised by an accumulation of immature haematopoietic cells that have lost the ability to differentiate. The classification of myeloid neoplasms provides useful information regarding a patient's disease particularly with regard to prognosis and the approach to

treatment required. As our knowledge of AML increases it is important to change and add to these classifications to reflect this (Arber, 2001).

1.3.1 French-American-British (FAB) Cooperative Group

The French-American-British (FAB) cooperative group established objective criteria for the classification of AML in 1976 to allow for comparisons between therapeutic trials and reduce discrepancies. This classification was entirely based on the morphological appearance of bone marrow and peripheral blood (Bennett *et al*, 1976). The classification can only be applied to samples from patients who have not received cytotoxic drug treatment and patients fall into one of nine groups (Table 2).

M0	Minimally differentiated myeloblastic leukaemia
M1	Myeloblastic leukaemia without maturation
M2	Myeloblastic leukaemia with maturation
M3	Hypergranular promyelocytic leukaemia
M4	Myelomonocytic leukaemia
M4Eo	Myelomonocytic leukaemia with eosinophils
M5	Monocytic leukaemia
M6	Erythroleukaemia
M7	Megakaryocytic leukaemia

Table 2 The FAB classification of acute myeloid leukaemia (Bennett *et al*, 1976)

1.3.2 Cytogenetics of AML

More than 200 structural and numerical aberrations have been detected through cytogenetic studies. The cytogenetic characteristics of patients at diagnosis have been widely recognised as a powerful diagnostic and prognostic tool for AML. Cytogenetics at diagnosis can classify patients into one of three prognostic sub-groups; favourable, intermediate or poor. Patients presenting with a full or partial deletion of chromosome 5, deletion of chromosome 7 or with a complex karyotype (three or more chromosomal aberrations) are placed into the poor prognosis group. Patients presenting with a t(15;17) or a t(8;21) translocation or an inversion of chromosome 16 are placed into the favourable risk group. The intermediate risk group contains patients presenting with any other cytogenetic abnormalities including a translocation involving the mixed lineage leukaemia (*MLL*) gene located on chromosome 11q23 or with an apparently normal karyotype (Grimwade et al, 1998; Slovak et al, 2000).

a. t(15;17)

Based on the French-American-British (FAB) classification of AML, acute promyelocytic leukaemia (APL) is classified as FAB M3. Based on morphology, two subtypes of APL can be identified; hypergranular classical APL (M3) and hypogranular variant APL (M3v). In hypergranular APL hypergranular abnormal promyelocytes replace the bone marrow and patients often present with leukopenia. Less granular blasts are detected in hypogranular APL and patients typically present

with an elevated leukocyte count. Both forms of AML are characterised by a granulocytic differentiation block and the presence of chromosomal translocation involving chromosome 17q21 on which the retinoic acid receptor alpha (*RARα*) gene is located (Puccetti & Ruthardt, 2004). *RARα* encodes for the *RARα* protein that belongs to the retinoic acid receptor (RAR) family of nuclear receptors. RARs form heterodimers with members of the retinoid X receptor (RXR) family of nuclear receptors in order to transduce retinoid signalling. The RXR-RAR heterodimer binds to specific DNA sequences of target genes, known as retinoic acid response elements (RARE), and activates their transcription by allowing the recruitment of co-activators and histone acetyltransferases (HAT). In the absence of the ligands the RXR-RAR heterodimer binds to DNA and recruits the co-repressors NCoR (nuclear receptor corepressor), SMRT (silencing mediator for retinoid or thyroid-hormone receptors) and histone deacetylase (HDAC) resulting in a repression of transcription (Vitoux *et al*, 2007). Physiological levels of retinoic acid (RA) (1×10^{-9} M) can release the corepressor complex from the heterodimer and allow the recruitment of co-activators resulting in transcriptional activation of the *RARα*-target genes (Lo *et al*, 1999). More than 95% of cases of APL display the fusion of the promyelocytic leukaemia (*PML*) gene with *RARα* resulting from the t(15;17)(q22;q21) translocation. The PML protein is localised in discrete nuclear structures named PML nuclear bodies (Dyck *et al*, 1994). Cell line and mice knock-out studies have identified functions of PML in growth control and as a tumour suppressor (Mu *et al*, 1994; Wang *et al*, 1998; Ruggero *et al*,

2000). The PML-RAR α fusion protein can also bind to RXR to form a heterodimer with an increased affinity for corepressors resulting in constitutive transcriptional repression. Only pharmacological levels of RA (1×10^{-6} M) can release the PML-RAR α fusion protein to resume transcriptional activation (Grignani *et al*, 1998; Grignani *et al*, 2000; Vitoux *et al*, 2007). In the remaining cases of APL various chromosomal aberrations have resulted in the RAR α gene forming fusion gene products with other genes including promyelocytic leukaemia zinc finger (PLZF) gene, the nucleophosmin (NPM) gene, the nuclear mitotic apparatus (NuMA) gene and the signal transducer and activator of transcription 5b (STAT5b) gene (Mistry *et al*, 2003; Vitoux *et al*, 2007) (Table 3).

Chromosomal Aberration	Fusion product	Frequency
t(15;17)(q22;q21)	PML-RAR α RAR α -PML	>95%
t(11;17)(q23;q21)	PLZF-RAR α RAR α -PLZF	0.80%
t(11;17)(q13;q21)	NuMA-RAR α	Very rare
t(5;17)(q35;q21)	NPM-RAR α RAR α -NPM	<0.5%
der(17)	STAT5b-RAR α	Very rare

Table 3. The chromosomal aberrations and their resultant fusion products present in APL. RAR α , retinoic acid receptor alpha; PML, promyelocytic leukaemia; PLZF, promyelocytic leukaemia zinc finger; NuMA, nuclear mitotic apparatus protein; NPM, Nucleophosmin; STAT5b, signal transducer and activator of transcription 5B (Mistry *et al*, 2003)

b. t(8;21)

Rowley *et al* (1973) first reported the t(8;21) translocation in a leukaemia patient sample in the early 1970s but *AML1* and *ETO* were not identified as the two genes located at the breakpoints until the early 1990s (Gao *et al*, 1991; Miyoshi *et al*, 1991; Erickson *et al*, 1992; Peterson & Zhang, 2004). Commonly associated with FAB group M2, the translocation of chromosomes 8 and 21 results in the formation of the fusion gene *AML1-ETO*. *AML1* is located on the long arm of chromosome 21 and encodes for the AML1 protein. AML1, also known as CBF α 2 (core binding factor alpha-2) and RUNX1 (runt-related transcription factor 1), is a subunit of the core binding factor (CBF) complex. As a RUNX protein, AML1 contains a runt homology DNA binding domain, a transactivation domain, a nuclear matrix attachment signal (NMTS) and two inhibitory domains. The runt domain of CBF α interacts with the beta CBF subunit, CBF β , to form a heterodimeric complex in order to increase its DNA-binding affinity and protect it from ubiquitin-mediated proteolysis. AML1 activates gene transcription through facilitating the assembly of transcriptional activation complexes (Kitabayashi *et al*, 1998). Dependant upon the target gene, AML1 can also function as a repressor of transcription through the recruitment of co-repressors (Yagi *et al*, 1999; Perry *et al*, 2002). The core binding factor (CBF) complex regulates the transcription of several haematopoietic-specific genes, including *M-CSFR*, *IL-3* and *GM-CSF* (Cameron *et al*, 1994; Takahashi *et al*, 1995; Rhoades *et al*, 1996). The importance of AML1 to this complex and its regulation in

haematopoiesis was demonstrated through mouse knock-out studies. *AML1*-deficient embryos lacked definitive haematopoietic progenitors and failed to develop foetal liver haematopoiesis resulting in extensive haemorrhaging and mid-gestational lethality (Okuda *et al*, 1996;Wang *et al*, 1996a;Wang *et al*, 1996b). ETO is a nuclear protein containing regions characteristic to transcription factors; three proline/serine/threonine-rich regions, a leucine-rich region and two zinc-finger domains (Rhoades *et al*, 1996). ETO associates with the nuclear co-repressor NCoR, mSin3A and HDACs (Gelmetti *et al*, 1998;Lutterbach *et al*, 1998). In the resultant AML1-ETO fusion protein the C-terminal sequences for transactivation of AML1 are replaced by the ETO protein. ETO retains its capacity to form stable complexes with co-repressors allowing the fusion protein to block the transactivation of haematopoietic-specific genes by associating with the co-repressor complex containing NCoR, mSin3A and HDACs, resulting in a block of myeloid differentiation (Gelmetti *et al*, 1998). The runt-binding domain is also retained in the fusion gene allowing the AML1-ETO protein to interfere in the activation and expression of genes normally associated with the AML1-CBF β complex (Meyers *et al*, 1995;Perry *et al*, 2002). Patients with a t(8;21) translocation are found to have a relatively favourable prognosis (Grimwade *et al*, 1998;Slovak *et al*, 2000).

c. inv(16)

Arthur and Bloomfield (1983) first detected a del(16)(q22) in five AML patients with bone marrow eosinophilia ranging between 8% to 54%

(Arthur & Bloomfield, 1983). Further studies identified an association between an inversion of chromosome 16 and FAB M4 AML (Tantravahi *et al*, 1984) and the Fourth International Workshop on Chromosomes in Leukaemia: a Prospective Study of Acute Non-Lymphocytic Leukaemia confirmed an association between abnormal eosinophils, AML and structural rearrangements of chromosome 16 (Vardiman *et al*, 1984;Reilly, 2005). A later study by Larson *et al* (1986) identified what is now widely recognised as the breakpoints occurring at both 16p13 and 16q22 being most closely associated with M4Eo (Larson *et al*, 1986). An inversion of chromosome 16, inv(16)(p13q22), generates the fusion gene product core binding factor beta-smooth muscle myosin heavy chain (*CBFβ-MYH11*). The core binding factor beta (*CBFβ*) gene is situated at the chromosome locus 16q22 and encodes for the CBFβ protein; a subunit of the core binding factor (CBF). Widely expressed, CBFβ regulates transcription indirectly through interaction with the subunit CBFα2, otherwise known as AML1. Instead of binding directly to DNA, CBFβ increases the affinity of AML1 for the consensus DNA sequence TGT/cGGT. This is achieved through the interaction of the CBFβ core domain with the runt domain of AML1 (Wang *et al*, 1996b). *CBFβ-MYH11* knock-in mice studies have demonstrated the importance of the *CBFβ* gene in haematopoiesis and the pathogenic consequences of the *CBFβ-MYH11* fusion gene. Embryonic haematopoiesis was impaired at stem-progenitor cell level and mice were unable to achieve definitive haematopoiesis resulting in mid-gestational lethality as the function of the AML1-CBFβ heterodimer was suppressed (Wang *et al*,

1996a;Wang *et al*, 1996b;Kundu *et al*, 2002). Castilla *et al* (1999) demonstrated that whilst the *CBF β -MYH11* gene alone was not sufficient to induce AML in the mice, leukaemia characterised by the presence of myelomonocytic blasts and the occasional eosinophils developed approximately 4 months after treatment with *N*-ethyl-*N*-nitroso-urea (ENU) (Castilla *et al*, 1999). Patients presenting with an inversion of chromosome 16 are associated with a relatively favourable prognosis (Grimwade *et al*, 1998;Slovak *et al*, 2000).

d. 11q23

The mixed lineage leukaemia (*MLL*) gene is located at chromosome 11q23 and is widely expressed during development in most tissues. *MLL* is required for haematopoiesis with many clustered *HOX* genes as targets (Ernst *et al*, 2004;Harper & Aplan, 2008). Many chromosomal rearrangements involving the *MLL* gene are associated with myeloid, lymphoid and mixed lineage leukaemias (Ayton & Cleary, 2001). The commonest form of translocation involving the *MLL* gene results in the fusion of the C-terminus coding sequence of a partner gene and the N-terminus coding sequencing of the *MLL* gene. The resultant fusion genes encode for in-frame oncogenic proteins (Hess, 2004). Currently more than 51 fusion partners of *MLL* have been identified (Basecke *et al*, 2006). Translocations involving *MLL* occur in approximately <10% of AML cases and generate a similar prognosis to those AML cases with an intermediate prognosis. Common translocations detected in AML include the t(6;11)(q27;q23), t(9;11)(q22;q23) and t(10;11)(q12;q23)

translocations fusing *MLL* with *AF6*, *AF9* and *AF10* respectively (Martineau *et al*, 1998;Hess, 2004). In addition to the balanced chromosomal translocations, the partial tandem duplication (PTD) of exons encoding the N-terminal sequences of *MLL* has been recorded in approximately 10% of AML cases, predominantly in those cases with a normal karyotype or a trisomy 11, and confers a poor prognosis (Basecke *et al*, 2006).

e. Normal Karyotype

AML patients that are karyotypically normal following standard cytogenetic analysis at diagnosis constitute the largest cytogenetic sub-group in AML, accounting for approximately 45% of all *de novo* AML cases. These patients are placed in the intermediate risk group with only 20-42% being long-term survivors (Grimwade *et al*, 1998;Slovak *et al*, 2000;Byrd *et al*, 2002). Despite the lack of chromosomal aberrations, this sub-group of AML is extremely heterogenous displaying many gene mutations. The *FLT3* gene encodes for a class III receptor tyrosine kinase. An internal tandem duplication (ITD) in the juxtamembrane domain-coding sequence of the gene results in a constitutive autophosphorylation of the FLT3 protein. Normal karyotype AML patients with a FLT3 ITD differ clinically from those normal karyotype patients displaying a wild-type FLT3; an increased white blood count and increased percentages of peripheral blood and bone marrow blasts are observed in the FLT3 ITD patients. A poorer prognosis has been assigned to those normal karyotype patients displaying a FLT3 ITD

compared to those normal karyotype patients without the mutation (Frohling *et al*, 2002; Bienz *et al*, 2005; Mrozek *et al*, 2007b). Conversely mutations detected in the nucleophosmin 1 (*NPM1*) gene or the CCAAT/enhancer-binding protein α (*CEBPA*) gene confer a more favourable prognosis in normal karyotype AML patients (Mrozek *et al*, 2007b). As well as the various mutations detected, aberrant gene expressions are often detected in the normal karyotype AML patients. The over-expression of both the brain and acute leukemia, cytoplasmic (*BAALC*) gene and the meningioma 1 (*MN1*) gene are associated with an adverse prognosis in normal karyotype patients (Mrozek *et al*, 2007a). The numerous genetic factors displayed in normal karyotype AML generating different clinical outcomes may be able to provide useful information in order to improve the treatment stratification by providing molecular-based treatments (Radmacher *et al*, 2006; Baldus *et al*, 2007).

1.3.3 Gene Mutations in AML

a. FMS-like Tyrosine Kinase 3 (FLT3)

FLT3, FMS-like tyrosine kinase 3, is a class III receptor tyrosine kinase, composed of an extracellular ligand-binding domain containing five immunoglobulin (Ig) loops, a transmembrane domain, a juxtamembrane dimerisation domain and an intracellular kinase domain. FLT3 exhibits strong sequence and structural similarities with class III receptor tyrosine kinase receptor family members c-fms and c-kit; the cytokine

receptors for M-CSF and SCF (Yarden *et al*, 1987;Carow *et al*, 1996;Gilliland & Griffin, 2002). The ligand for the FLT3 protein (FL) is a type I transmembrane protein and a growth factor for stem cells and immature myeloid cells (Rusten *et al*, 1996). FL activates the tyrosine kinase activity of FLT3 and stimulates the proliferation of primitive haematopoietic progenitor cells and more committed early myeloid and lymphoid precursors through synergy with other growth factors and interleukins (IL) (Veiby *et al*, 1996;Ebihara *et al*, 1997;Lyman & Jacobsen, 1998;Meshinchi & Appelbaum, 2009). Preferentially expressed on haematopoietic CD34⁺ stem cells in normal bone marrow and decreasing in expression as differentiation progresses, high expression levels of both the *FLT3* gene and FLT3 protein have been observed in the majority of acute myeloid leukaemias (Birg *et al*, 1992;Small *et al*, 1994;Carow *et al*, 1996;Gilliland & Griffin, 2002;Kiyoi *et al*, 2005). As well as the high levels of expression observed, two types of activating mutations of the *FLT3* gene have been reported in haematological malignancies, and therefore the *FLT3* gene is considered to play an important role in leukaemogenesis (Gilliland & Griffin, 2002). The first reported *FLT3* gene mutation was internal tandem duplication (ITD) in the juxtamembrane domain-coding sequence; reported in approximately 20% of all cases of adult AML. The second type of FLT3 mutation is a point mutation in the activation loop of the tyrosine kinase domain (TKD)-coding sequence. The point mutation involving the aspartic acid 835 (D385) residue is the most frequently occurring point mutation and occurs in approximately 7% of

adult AML patients (Thiede et al, 2002; Yamamoto et al, 2001) (Figure 2). Both of these mutations result in constitutive autophosphorylation of the FLT3 protein and activation of downstream targets including the RAS/MAPK, STAT (signal transducer and activator of transcription) and the AKT/PI3 kinase pathways resulting in cytokine-dependent proliferation of cells (Mizuki et al, 2000; Hayakawa et al, 2000; Gilliland & Griffin, 2002). The presence of a *FLT3* ITD is associated with an unfavourable prognosis whilst the prognostic relevance of a *FLT3* TKD currently remains undefined (Abu-Duhier et al, 2000; Thiede et al, 2002; Moreno et al, 2003; Mead et al, 2008; Whitman et al, 2008).

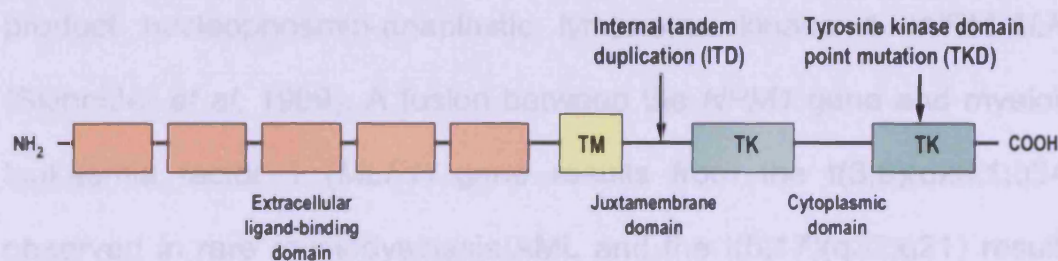


Figure 2. The FLT3 protein and the locations affected by the activating gene mutations (Gilliland & Griffin, 2002; Small, 2006) ECD – extracellular domain, TM – transmembrane domain, TK – tyrosine kinase domain.

b. Nucleophosmin 1 (NPM1)

Nucleophosmin (NPM1), also known as B23, belongs to the nucleoplasmin family of nuclear chaperones. NPM1 is predominantly located in the nucleolus and is ubiquitously expressed. Driven by a bipartite nuclear localization signal (NLS), the protein shuttles between the nucleus and the cytoplasm stimulating the import of proteins into the

nucleus (Szebeni *et al*, 1995;Hingorani *et al*, 2000;Falini *et al*, 2009). The *NPM1* gene is located on chromosome 5q35 and three isoforms exist as a result of alternative splicing; B23.1, B23.2 and B23.3 (Falini *et al*, 2007). The *NPM1* protein is seen to be aberrantly localised in the cytoplasm when a chromosomal translocation involving the *NPM1* gene occurs or a mutation in the *NPM1* gene is present (Falini *et al*, 2005). Chromosomal translocations disrupting the *NPM1* gene have been observed in various haematopoietic malignancies generating fusion proteins that retain the N-terminus of *NPM1*. Approximately 85% of cases of anaplastic large cell lymphoma (ALCL) display the t(2;5)(p23;q35) translocation at diagnosis resulting in the fusion gene product nucleophosmin-anaplastic lymphoma kinase 1 (*NPM-ALK*) (Skinnider *et al*, 1999). A fusion between the *NPM1* gene and myeloid leukaemia factor 1 (*MLF1*) gene results from the t(3;5)(q25.1;q34) observed in rare myelodysplasia/AML and the t(5;17)(q32;q21) results in the fusion of *NPM1* and the retinoic acid receptor α (*RAR α*) gene observed in APL patients (Redner *et al*, 1996;Yoneda-Kato *et al*, 1999;Falini & Mason, 2002). Mutations involving the C-terminus of exon 12 of the *NPM1* gene have been identified in approximately 25-35% of adult AML cases, and approximately 50% of AML patients with a normal karyotype (Schnittger *et al*, 2005;Falini *et al*, 2007). The most common mutation seen to occur in *NPM1* is called Type A, a TCTG insertion at position 596, and is observed in approximately 70-80% of normal karyotype AML cases with an *NPM1* mutation (Schnittger *et al*, 2005;Chen *et al*, 2006;Falini *et al*, 2007). The presence of an *NPM1*

gene mutation is seen to confer to a favourable outcome in AML patients with a normal karyotype when in the absence of a FLT3 ITD (Dohner *et al*, 2005;Falini *et al*, 2009).

c. CCAAT/enhancer binding protein, alpha (CEBPA)

CEBPA is a member of the leucine zipper transcription factor family with an N-terminal trans-activation domain and a leucine-rich dimerisation domain termed “leucine-zipper” at its C-terminus to mediate DNA-binding and dimerisation (Johnson, 2005;Fuchs, 2007). Expressed in many tissues, *CEBPA* plays an important role in the regulation of cell proliferation and differentiation. CEBPA mediates the actions of cyclin-dependent kinases (CDK) and the CDK2-inhibitor p21, amongst other targets, to up-regulate lineage-specific gene products and cause a proliferation arrest due to exit from the cell cycle (Fuchs, 2007). In haematopoiesis, *CEBPA* regulates myeloid differentiation; an up-regulation of the *CEBPA* gene results in the development of granulocytes and an inhibition in the development of monocytes (Radomska *et al*, 1998). A distinct pattern of *CEBPA* expression in the haematopoietic system has been generated; low levels of *CEBPA* expression are detected in HSCs, an up-regulation of expression in GMPs and a down-regulation through differentiation to megakaryocyte and erythroid progenitors (Mueller & Pabst, 2006). A mutation in the *CEBPA* gene is reported in approximately 15% of AML patients; approximately 70% of these cases also present with a normal karyotype (Pabst *et al*, 2001;Preudhomme *et al*, 2002;Frohling *et al*, 2004;Mueller

& Pabst, 2006). Mutations of the *CEBPA* gene result in disruptions to the leucine zipper affecting the binding of DNA and resulting in a loss of CEBPA function, or an induction of proliferation due to the premature termination of the normal 42 kDa form of CEBPA (Mueller & Pabst, 2006;Fuchs, 2007). The presence of a *CEBPA* mutation in AML is associated with a favourable prognosis (Preudhomme *et al*, 2002;Bienz *et al*, 2005).

1.3.4 Aberrant Gene Expression in AML

a. Brain and acute leukaemia, cytoplasmic (*BAALC*) gene

The brain and acute leukaemia cytoplasmic gene (*BAALC*) was identified from its over-expression in leukaemic blasts of AML with a trisomy 8 in comparison to normal karyotype AML. Located on chromosome 8q22.3, *BAALC* appears to have no sequence homology to other proteins or functional domains. *BAALC* is involved in neuroectodermal and haematopoietic development, expressed primarily in neuroectoderm-derived tissues and haematopoietic precursors (Tanner *et al*, 2001). High levels of *BAALC* expression have been detected in CD34⁺ progenitor cells but as the cells subsequently go through differentiation, the level of *BAALC* expression is seen to decrease (Langer *et al*, 2008;Baldus *et al*, 2003a;Tanner *et al*, 2001). High levels of *BAALC* have also been observed in cytogenetically normal AML in adults and is associated with an adverse clinical outcome (Baldus *et al*, 2003b;Bienz *et al*, 2005).

b. Meningioma 1 (MN1)

MN1 is a 136kDa nuclear protein functioning as a transcriptional co-activator. Transient transcription assays confirmed the transcriptional role of MN1 when the activation of transcription of the Moloney sarcoma virus long terminal repeat (MSV-LTR) in a murine fibroblast cell line was observed (Buijs *et al*, 2000). MN1 stimulates retinoic acid receptor/retinoid X receptor (RAR-RXR)-mediated transcription through the recruitment and co-expression of the protein co-activators p300 and RAC3 (retinoic acid receptor interacting protein 3) (Buijs *et al*, 2000;van Wely *et al*, 2003). Carella *et al* (2007) performed quantitative RT-PCR analysis of mouse bone marrow to identify which haematopoietic cells expressed the *MN1* gene. Low levels of *MN1* expression were detected in HSCs and high expression levels detected in GMPs but no expression in the CMP, CLP or MEP fractions suggesting a role of *MN1* in GMP-derived myelopoiesis (Carella *et al*, 2007). The balanced chromosome translocation t(21;22)(p13;q12) is associated with haematological malignancies including acute myeloid leukaemia. The translocation results in the fusion of the *MN1* gene with the ETS-transcription factor-encoding gene *TEL*. Studies of *MN1-TEL* knock-in mice displayed the oncogenic properties of the fusion gene in haematopoiesis; expression of the gene in myeloid and lymphoid compartments resulted in the development of T-cell lymphoma and AML (Kawagoe & Grosveld, 2005a;Kawagoe & Grosveld, 2005b). An over-expression of the *MN1* has also been associated with haematological malignancies with an inv(16) (Carella *et al*, 2007). An

over-expression of *MN1* has also been detected in some patients presenting with a normal karyotype where it is associated with a poor prognosis (Heuser *et al*, 2006).

c. Ecotropic Viral Integration Site-1 (EVI1)

The *EVI1* gene, located on chromosome 3q26 encodes for the nuclear transcription factor EVI1. Located in the nucleus, EVI1 consists of multiple zinc finger domains, a repression domain and an acidic region at the C-terminus. EVI1 plays important roles in both normal development and oncogenesis. Gene targeting mice studies have demonstrated a high expression of *EVI1* in embryonic and adult bone marrow HSCs and a role of EVI1 in the maintenance of HSCs. *EVI1*^{-/-} embryos displayed a markedly decreased number of HSCs with defective self-renewal, proliferation and repopulating capacity (Yuasa *et al*, 2005). The *EVI1*^{-/-} embryos also displayed a reduced expression of the *GATA-2* gene; the transcription factor encoded by *GATA-2* is essential for the proliferation of HSCs (Yuasa *et al*, 2005). EVI1 was originally identified as the integration site of the ecotropic retrovirus leading to myeloid leukaemia in mouse models (Goyama & Kurokawa, 2009). Chromosomal rearrangements involving the *EVI1* gene locus, 3q26, result in the formation of fusion genes involving *EVI1* and are commonly found in myeloid malignancies. The translocation of chromosome 3q26 is found in cases approximately 1-2% of cases of MDS, AML and CML in blast crisis; the translocation t(3;3)(q21;q26) or inversion inv(3)(q21q26) occurring most frequently (Nucifora, 1997).

The rearrangements involving 3q26 can result in the over-expression of *EVI1* but high levels of *EVI1* expression have also been detected in a group of AML cases without a 3q26 involvement. The cause of this over-expression of *EVI1* is unknown but it has been shown to be associated with a poor prognosis in *de novo* AML cases (Barjesteh van Waalwijk van Doorn-Khosrovani *et al*, 2003; Wieser, 2007).

1.3.5 WHO

As part of the 3rd edition of the series, *WHO Classification of Tumours*, a collaboration of the World Health Organization (WHO), the Society for Haematopathology and the European Association of Haematopathology published a "Classification of Tumours of the Hematopoietic and Lymphoid Tissues". Although the FAB classification of AML took into account the morphological heterogeneity, it did not reflect the genetic or clinical diversity of AML now recognised to be important (Vardiman *et al*, 2002). The WHO classification of tumours makes use of all available diagnostic information; genetic, immunophenotypic, biological, clinical **and** morphological, to define specific disease entities. In 2008, a revised classification was published as part of the 4th edition to incorporate the scientific and clinical information gathered over the 7 years between editions in order to define the diagnostic criteria of haematopoietic neoplasms (Vardiman, 2009). In the WHO classification, "Myeloid" includes all cells belonging to the granulocytic, monocytic/macrophage, erythroid, megakaryocytic and mast cell lineages. As with previous classifications, the blast percentage detected

in samples is used as a means of categorising the myeloid neoplasms and making decisions in regards to their progression; a myeloid neoplasm with more than 20% blasts detectable in the bone marrow or peripheral blood is considered to be AML when it occurs *de novo* (Vardiman *et al*, 2009). Based on the level of maturation, myeloid neoplasms are split into several groups. Those neoplasms comprised mainly of blasts with minimal, if any, maturation (AML) are placed together in one group. If any maturation of the blasts is detectable, the neoplasms are further split into one of three classes; those with effective maturation (MPN), those with ineffective maturation with dysplastic features (MDS) or those with both ineffective and effective maturation (MDS/MPN) (Vardiman, 2009). Since the 3rd edition, a number of gene mutations have been accepted as contributing to the leukaemic process and the morphologic and clinical characteristics of the disease; the mutational status of the *NPM1* and *CEBPA* genes have been incorporated as provisional entities (Vardiman *et al*, 2009). Despite incorporating the mutational status of several genes into the new classification, it is still necessary to use a “multiparameter” approach when classifying neoplasms; the blast count remains an important criteria factor in the classification process (Vardiman, 2009) (Table 4).

Myeloproliferative neoplasms (MPN)
Myeloid and lymphoid neoplasms associated with eosinophilia and abnormalities of PDGFRA, PDGFRB, or FGFR1
Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)
Myelodysplastic syndrome (MDS)
Acute myeloid leukaemia and related neoplasms
 Acute myeloid leukaemia with recurrent genetic abnormalities
 AML with t(8;21)(q22;q22); RUNX1-RUNX1T1
 AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); CBFB-MYH11
 APL with t(15;17)(q22;q12); PML-RARA
 AML with t(9;11)(p22;q23); MLLT3-MLL
 AML with t(6;9)(p23;q34); DEK-NUP214
 AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1
 AML (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1
 Provisional entity: AML with mutated NPM1
 Provisional entity: AML with mutated CEBPA
 Acute myeloid leukemia with myelodysplasia-related changes
 Therapy-related myeloid neoplasms
 Acute myeloid leukaemia, not otherwise specified
 AML with minimal differentiation
 AML without maturation
 AML with maturation
 Acute myelomonocytic leukaemia
 Acute monoblastic/monocytic leukaemia
 Acute erythroid leukaemia
 Pure erythroid leukaemia
 Erythroleukaemia, erythroid/myeloid
 Acute megakaryoblastic leukaemia
 Acute basophilic leukemia
 Acute panmyelosis with myelofibrosis
 Myeloid sarcoma
 Myeloid proliferations related to Down syndrome
 Transient abnormal myelopoiesis
 Myeloid leukaemia associated with Down syndrome
 Blastic plasmacytoid dendritic cell neoplasm
Acute leukemias of ambiguous lineage
B lymphoblastic leukaemia/lymphoma
T lymphoblastic leukaemia/lymphoma

Table 4. The WHO classification of myeloid neoplasms focusing on the acute myeloid leukaemia and related neoplasms (Vardiman *et al*, 2009).

1.3.6 Current and Potential AML Therapies

More than two decades ago the application of prolonged-duration multi-agent chemotherapy programmes generated disappointing results in AML despite the successful cure rates in childhood acute lymphoid leukaemia (ALL) (Vaughan & Karp, 2008). As the supportive care in transfusion and infectious disease management improved, investigators turned to researching more intensive remission induction regimes for the role of maintenance therapy and improving cure rates. Studies in timed-sequential chemotherapy to further intensify remission induction therapy or for consolidation therapy using a combination of two agents, anthracycline and cytarabine, were performed in clinical trials leading to a regime similar to that used today (Vaughan & Karp, 2008). As a consequence of increased knowledge about the molecular basis of leukaemogenesis more target-specific therapies are now being investigated to improve current regimes. Coordinated management of the treatment of AML has been the focus the Medical Research Council (MRC) trials throughout the UK; treatment improvements were made based on the results from previous trials. The current MRC AML 16 trial focuses on older patients (>60 years) diagnosed with AML or high-risk myelodysplastic syndrome (MDS) and not considered suitable for the MRC AML 15 trial for younger patients; patients are entered into one of two randomised treatment arms based on the consideration of their fitness for intensive treatment. The development of the next trial, MRC AML 17, has resulted from the recognition of the importance of cytogenetics and gene aberrations in AML; patients are randomised

based on t(15;17) translocations, core binding factor translocations or the presence of FLT3 mutations.

a. Chemotherapy

Currently patients diagnosed with AML aged 18-60 years go through several treatment phases. The initial therapy consists of induction chemotherapy; the removal of all haematopoietic elements from the bone marrow to allow re-population with normal cells. This often results in less than 5% leukaemic blasts in the bone marrow and therefore patients are deemed to be in remission. A regime commonly known as "3+7" is used; patients receive an anthracycline (daunorubicin or idarubicin) for 3 days followed by the cell cycle-specific cytarabine (ara-C) for 7 days (Stone *et al*, 2004; Estey, 2009). Approximately 50% to 75% of AML adults receiving this treatment regime will achieve complete remission (CR) (Tallman *et al*, 2005). To improve complete remission rates higher doses of the agents used have been tested in combination with other agents, for example etoposide; an inhibitor of microtubular assembly in dividing cells (Tallman *et al*, 2005). Following remission it is necessary that patients receive further therapy in an attempt to reduce the remaining undetectable leukaemic burden. Autologous and allogeneic bone marrow transplantation following remission provides a long-term disease-free survival but is accompanied by a high mortality and long-term morbidity rate. Post-remission therapy strategies include intensive consolidation therapies of either high-dose chemotherapy or chemo-radiotherapy with

allogeneic/autologous haematopoietic stem-cell transplantation (HSCT), or a low-dose maintenance therapy (Tallman *et al*, 2005). The vast heterogeneity of AML patients and their different response to standard treatments has provided a rationale for exploring more targeted therapies based on chromosomal aberrations and gene mutations and the expression of cell surface markers. Compared to the progress made in the treatment of younger patients, the treatment of AML in patients older than 60 years has not improved significantly. Following standard induction therapy, 43% to 57% of patients over the age of 60 years achieve a complete remission, long term survival ranges between 10% and 15% (Hiddemann *et al*, 1999;Goldstone *et al*, 2001). Elderly patients display differences in the biological and clinical behaviour of AML to their younger counterparts that often result in a poorer prognosis, for example chromosomal aberrations. The presence of co-morbid illnesses and poor performance status also affects their eligibility to receive standard induction therapy (Goldstone *et al*, 2001;Mulford, 2008).

b. Differentiation agents

The induction of terminal differentiation by retinoids in leukaemic cell lines and short-term culture APL cells was observed in the early 1980s. An *in vitro* study by Breitman *et al* (1980) demonstrated the differentiation of HL-60 cells induced by 10^{-6} M *all trans* retinoic acid (ATRA), resulting in terminal differentiation in 90% of the cells (Breitman *et al*, 1980). Further investigation into ATRA-induced differentiation in

APL cells exhibiting the t(15;17) translocation led to the development of ATRA as a therapy for APL (Breitman *et al*, 1981). As previously discussed in section 1.3.2, retinoic acid receptors (RARs) are a family of transcription factors that regulates granulocytic differentiation. The fusion product PML-RAR α is a negative regulator of transcription through the formation of homodimers with RARs, recruiting a co-repressor complex containing HDACs and inhibiting transcription of haematopoietic lineage-specific genes. Pharmacological concentrations of ATRA (10^{-6} M) induce a conformational change releasing the co-repressor complex and activating transcription (Grignani *et al*, 1998; Vitoux *et al*, 2007). ATRA was first used to treat APL patients in Shanghai, and later in France, with promising results (Huang *et al*, 1988; Castaigne *et al*, 1990). A progressive change in malignant cells in the bone marrow and signs of terminal differentiation was observed. Clonal cells exhibiting a t(15;17) were replaced by polyclonal haematopoietic cells and no primary resistance occurred. However, in some patients an increase in white blood count due to an activation of leukocytes during the first two weeks of treatment was observed in several patients, often accompanied by “retinoic acid syndrome” characterised by fever, weight gain, difficulty in breathing and pleural effusion, and occasionally resulting in death (Castaigne *et al*, 1990; De Botton S. *et al*, 1998; Fenaux *et al*, 2001). Following treatment with ATRA, a relapse within months is often observed when receiving ATRA alone or with mild chemotherapy due to a resistance to the drug. Following treatment with ATRA with intensive anthracycline-AraC

chemotherapy has been seen to successfully reduce relapse rates in APL cases (Fenaux *et al*, 1993;Fenaux *et al*, 1994).

c. Signal transduction inhibition

Now there is a better understanding into the molecular biology of leukaemogenesis, and in particular AML, treatment regimes can be improved with the development of targeted therapies. Receptor tyrosine kinase (RTK) signalling is essential for many cellular processes, including cell proliferation and survival and the control of the cell cycle. Many RTKs are constitutively activated in AML, for example FLT3 (Section 1.3.2a), as a result of molecular aberrations and therefore investigation into new targeted therapies has been directed towards the inhibition of tyrosine kinase activity. Various inhibitors have been used in clinical trials for the inhibition of the tyrosine kinase activity of FLT3. The indolinone compounds SU5416 and SU5614 designed to target the tyrosine kinase activity of FLT3 have been seen to decrease cell proliferation and increase cell apoptosis in AML. The compounds inhibit the activity of both wild-type FLT3 and FLT3 ITD isoforms resulting in the inhibition of downstream signalling via the MAPK and STAT pathway (Yee *et al*, 2002). In leukaemic cell lines and mouse models, G₁ cell cycle arrest and apoptosis in AML cells expressing mutant FLT3 was seen to be selectively induced through the small molecule FLT3 inhibitor PKC412 directly inhibiting the tyrosine kinase activity of FLT3 (Weisberg *et al*, 2002). Levis *et al* (2002) observed the inhibitory effects of the indolocarbazole derivative CEP-701 in AML cell lines and AML

primary blasts. CEP-701 was seen to be effective against cells expressing the wild-type, ITD and TKD isoforms of FLT3. By inhibiting the downstream targets ERK (extracellular signal-regulated kinases) and STAT5, a varying cytotoxic effect dependant upon the FLT3-expressing cell was observed (Levis *et al*, 2002).

d. Monoclonal antibodies

During the search for new therapies in AML attention has been focused on the use of monoclonal antibodies (mAbs) targeting antigens specifically expressed on myeloid cells. Monoclonal antibodies can provide a more specific and a less toxic means of treating AML compared to conventional chemotherapy (Mulford, 2008).

i. CD33

Currently the most successful mAb agent used in the battle against AML is the immunoconjugate known as gemtuzumab ozogamicin (GO), also known as Mylotarg, that binds to the CD33 antigen. CD33 is a transmembrane glycoprotein found specifically on myeloid cells but not expressed on normal CD34+ pluripotent haematopoietic and non-haematopoietic tissues (Andrews *et al*, 1983; Stasi *et al*, 2008). As the CD33 molecule is also expressed on more than 80% of myeloid leukaemia cells, an anti-CD33 immunoconjugate would allow for the selection and elimination of malignant myeloid cells whilst sparing immature haematopoietic progenitors (Mulford, 2008). GO is a humanised IgG4 anti-CD33 monoclonal antibody conjugated to a

hydrazide derivative of calicheamicin, a naturally occurring hydrophobic enediyne antibiotic (Figure 3). GO rapidly targets CD33+ cells and is internalised resulting in the release of the calicheamicin derivative by acid hydrolysis. Calicheamicin interferes with biological processes by cleaving free DNA and displacing a DNA-binding protein through competition or altering DNA structure leading to apoptosis (Stasi et al, 2008).

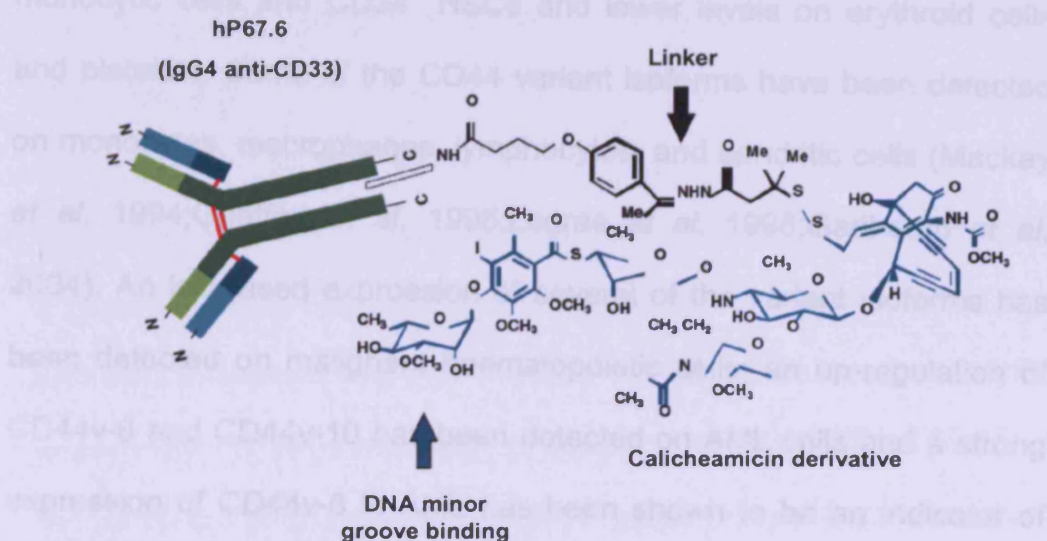


Figure 3. The schematic structure of gemtuzumab ozogamicin (Stasi *et al*, 2008)

ii. CD44

The cell surface antigen CD44 is a highly glycosylated type 1 transmembrane receptor involved in survival, migration, proliferation, differentiation and apoptosis. CD44 is encoded by a single gene consisting of 20 exons. Whilst 10 of the exons are constitutively expressed in all CD44⁺ cells generating a standard isoform of the CD44 molecule, the other 10 exons are alternatively spliced. Various

combinations of expression of the variant exons generates CD44 variant isoforms (CD44v) characterised by additional sequences inserted into the extracellular domain of the molecule. The expression of the standard CD44 isoform in haematopoiesis is dependant upon the cell lineage and stage of differentiation. Expressed on all types of mature blood cells and the majority of mononuclear bone marrow precursors, a high level of CD44 expression can be detected on monocytic cells and CD34⁺ HSCs and lower levels on erythroid cells and platelets. Some of the CD44 variant isoforms have been detected on monocytes, macrophages, lymphocytes, and dendritic cells (Mackay *et al*, 1994; Ghaffari *et al*, 1995; Legras *et al*, 1998; Gadhoun *et al*, 2004). An increased expression of several of the variant isoforms has been detected on malignant haematopoietic cells; an up-regulation of CD44v-6 and CD44v-10 has been detected on AML cells and a strong expression of CD44v-6 in AML has been shown to be an indicator of poor prognosis (Ghaffari *et al*, 1995; Legras *et al*, 1998). Charrad *et al* (1999) were the first group to demonstrate the potential use of anti-CD44 monoclonal antibodies (mAbs) as a therapy for AML. In primary AML blast cultures the addition of an anti-CD44 mAb could induce terminal differentiation and block proliferation of the AML blasts in FAB groups M1 to M5 (Charrad *et al*, 1999).

Treatment Failure

Treatment failures resulting in reduced complete remission rates and an increase in the number of relapse cases arise from a number of factors.

As previously discussed, age plays an important role in the success of AML treatment. Elderly patients may present with co-morbid illness or biological influences reducing the success of treatment they receive. The elderly are often unable to tolerate the intensive therapy required to eliminate AML, often developing life-threatening infections during the course of treatment (Estey *et al*, 1982). Secondary AML is associated with a resistance to standard chemotherapy and as a result, a poorer prognosis and treatment failure. Patients with a history of myelodysplastic syndrome (MDS) or a chronic myeloproliferative disorder often develop secondary AML (Larson, 2007). Alternatively, secondary AML is therapy-related (t-AML) arising from treatment with alkylating agents, topoisomerase II inhibitors, radiation therapy, or immunosuppressive agents (Leone *et al*, 1999;Larson, 2007). Importantly, a large number of AML cases do achieve complete remission but unfortunately relapse ultimately occurs. Induction therapy is designed to remove all the haematopoietic cells from the bone marrow to allow for repopulation of normal cells. As the cut-off for remission is less than 5% leukaemic blasts remaining, residual disease may confer a growth advantage and repopulate at a later date. The hypothesis of the leukaemic stem cell discussed previously (Section 1.2.2d) would suggest that only a few leukaemic stem cells remaining in the residual disease would be necessary for the development of a haematological relapse.

1.4 Minimal residual disease (MRD)

When a patient is diagnosed with AML, the number of leukaemic cells present may exceed 10^{12} . Defining a patient as being in complete remission (CR) depends on the identification of leukaemic cells based on their morphology and immunophenotype. A patient's bone marrow is classed as "normal" when the number of morphologically identifiable blasts calculated is less than 5%, which can equate to as many as 10^{10} leukaemic cells still present in the marrow (Campana & Pui, 1995). Following remission patients undergo post-remission treatment to further reduce the tumour load with the aim of eliminating the leukaemic clone. The patients receive treatment at the same intensity regardless of their individual residual tumour load and the decision to stop post-remission treatment is also decided without the knowledge of any residual disease present (Campana & Pui, 1995;Liu Yin, 2002f). Whilst survival rates have increased in adult patients, relapse occurs in up to 70% of the patients undergoing chemotherapy alone presumably due to residual leukaemic cells still present after treatment (Liu Yin, 2002e). Therefore, the study of MRD in patients with AML could produce information that could be used as a powerful tool for the prognosis of patients and therefore determining a treatment approach for patients (Grimwade & Lo Coco, 2002). Being aware of the level of leukaemic cells remaining following treatment would allow for therapy regimes to be altered and the risk of relapse reduced (Grimwade & Lo Coco, 2002;Liu Yin, 2002d;Chung *et al*, 2006). The various methods available

for the detection of MRD rely on the morphological, cytochemical, immunophenotypic and genetic properties of the leukaemic cells.

1.4.1 Detection of MRD

MRD detection methods require a sensitivity of detection of at least one leukaemic cell in 10^3 cells. To avoid false-positives techniques employed also require the ability to distinguish leukaemic cells from healthy normal cells. Over the years several techniques have been developed and evaluated that vary in sensitivity and specificity.

a. Conventional cytogenetics

Karyotyping is an essential tool routinely carried out at diagnosis. Clonal abnormalities including structural and numerical changes specific to the leukaemic cells can be detected and used to monitor the progression of the disease as they disappear during remission and reappear during relapse (Yin & Tobal, 1999; Liu Yin, 2002c). In a very early study by Testa *et al* (1979), chromosomal aberrations detected at diagnosis through karyotyping were seen to disappear when patients achieved complete remission and re-appear with the onset of relapse (Testa *et al*, 1979). MRD detection by cytogenetics also has the advantage of being able to detect new aberrations arising during the progression of the disease. However, karyotyping is quite insensitive and relies on the analysis of small numbers of metaphase cells (Liu Yin, 2002b).

b. Fluorescence *in situ* hybridisation (FISH)

Fluorescence *in situ* hybridisation (FISH) has the advantage over conventional karyotyping of being able to detect chromosomal abnormalities in both metaphase and interphase cells, increasing the chance of detecting abnormalities previously missed due to low proliferative rates. DNA probes specific to a region of a chromosomal breakpoint or to a whole chromosome are used to detect chromosomal abnormalities throughout a patient's disease. Structural rearrangements that would have previously been missed during traditional karyotyping can be detected by FISH (Yin & Tobal, 1999). To improve the accuracy of results, FISH analysis is often combined with morphologic or immunohistochemical analysis (Campana & Pui, 1995). Unfortunately FISH has a detection limit of only one abnormal cell per 100 cells (1%) due to aneuploid cells and other artefacts. The co-localisation of two signals can also generate false-positives further reducing the sensitivity of FISH as a means of MRD detection (Yin & Tobal, 1999;Liu Yin, 2002a).

c. Flow Cytometry

Despite a lack of leukaemia-specific antigens, the immunophenotype of some leukaemic cells differ enough from those of "normal" haematopoietic cells that immunophenotyping can be used for monitoring disease progression (Campana & Coustan-Smith, 1999;Liu Yin, 2002g). Staining with a number of fluorochrome-labelled antibodies can identify leukaemia-associated aberrant immunophenotypes (LAIP)

of leukaemic cells, which can alter between diagnosis and relapse (Kern et al, 2008). Multi-parameter flow cytometry can detect antigen co-expression in leukaemic cells and has been used for the detection of MRD and predicting relapse in AML patients in several studies (San Miguel et al, 1997; Sievers et al, 1996). Unique to this method, size and granularity of the cells can also be measured by flow cytometry (Campana & Coustan-Smith, 2002). A major advantage of the use of multi-parameter flow cytometry for the detection of MRD is its sensitivity (one leukaemic cell in 10^3 - 10^4 normal cells) and the speed of analysis over other methods (Liu Yin, 2002h). However, there are several problems existing with the use of flow cytometry for the detection of MRD, which contribute to reducing the sensitivity of this particular method. A lack of specificity or the use of inappropriate markers can result in false-positives due to background levels of "normal" cells with a similar immunophenotype to the leukaemic clones. False-negatives can also occur from a phenotypic-switch occurring at relapse (Campana & Coustan-Smith, 2002)

d. Reverse transcription-polymerase chain reaction and Polymerase chain reaction (RT-PCR/PCR)

DNA-amplification PCR methods are used to detect mutations, translocations, polymorphisms and aberrant changes in gene expression that can be used to follow the progression of a disease. RT-PCR and PCR methods are very sensitive capable of detecting one leukaemic cell in 10^3 - 10^6 normal cells, dependant upon many variables,

and so useful for the monitoring of MRD (Yin & Grimwade, 2002; Campana, 2003). Many studies have demonstrated the prognostic significance of the presence of disease-specific markers detected by PCR in AML (Trka *et al*, 2002; Krauter *et al*, 2003; Leroy *et al*, 2005; Perea *et al*, 2006).

Qualitative Reverse Transcriptase-PCR

Qualitative RT-PCR can be used to amplify target gene transcripts for the detection of gene rearrangements, for example fusion gene transcripts generated from chromosomal translocations including breakpoints spread over a large area (Yin & Tobal, 1999; Gabert *et al*, 2003). Complementary DNA (cDNA) generated in the first step of RT-PCR is amplified by PCR in the second step. The qualitative data generated, negative or positive expression, is dependant on the sensitivity of the assay. Inefficiency of the first step generating cDNA and degradation of the starting material can reduce the sensitivity of the process greatly. The sensitivity and the qualitative nature are major disadvantages for RT-PCR and its contribution to MRD detection. A negative result generated for a patient does not automatically imply there is no residual disease remaining; simply there is no residual disease that can be detected by the assay. In a similar manner, a positive result may not imply a potential relapse. Several studies have detected a persistence of cells bearing the t(8;21) translocation in AML patients in long-term remission after receiving chemotherapy or an autologous bone marrow transplant (Nucifora *et al*, 1993; Saunders *et*

al, 1994). As a result quantitative PCR methods for MRD detection have been developed.

Nested RT-PCR

Nested PCR is used to reduce contamination in the amplification process by using two sets of primers; an external primer set and an internal primer set. The amplified product generated from the first PCR run is amplified in a second PCR using internal oligonucleotide primers (Cross, 1995). A two-step nested RT-PCR can reach a minimal target sensitivity of 10^{-4} , and is reported to be more sensitive than real-time PCR techniques but as a qualitative method it can provide only limited information in regards to MRD kinetics (Takenokuchi *et al*, 2004;Schuler & Dolken, 2006). As a result, Takenokuchi *et al* (2004) developed a quantitative nested RT-PCR (QN-RT-PCR) technique. Compared to real-time PCR, QN-RT-PCR was 10 times more sensitive for the detection of plasmid DNA containing AML1-ETO fusion cDNA (Takenokuchi *et al*, 2004).

Competitive PCR

Competitive PCR has been shown to be useful in predicting relapse in CML and AML t(8;21) patients (Cross *et al*, 1993;Tobal & Yin, 1996). Serial dilutions of a known amount of internal standard, known as a “competitor”, are amplified using the same primers as the target gene of interest. The competitor produces different sized products to the gene target, which can be visualised by gel electrophoresis. As the amount of

competitor added initially to the PCR reaction is known, the amount of target gene transcript produced can be estimated (Cross, 1995; Yin & Tobal, 1999). A disadvantage of this form of quantitative PCR reaction is that it is time consuming and labour intensive (Gabert *et al*, 2003).

Quantitative Real-time PCR (RQ-PCR)

Quantitative real-time PCR (RQ-PCR) allows for a rapid and sensitive measurement of mRNA expression levels of AML-specific genes relative to a control/reference gene. The level of PCR product is measured as it accumulates throughout the amplification process rather than measuring end-point amount of product (Liu Yin, 2002i). RQ-PCR can be used to follow the progression of a patient's disease by detecting gene mutations, fusion-gene transcripts resulting from translocations and aberrantly expressed.

RQ-PCR Techniques. The hybridisation probe method utilises two juxtaposed sequence-specific probes; one labelled at the 3' end with a donor fluorochrome and one labelled at the 5' end with an acceptor fluorochrome. During the annealing phase, both probes hybridise to the target sequence. Light with a longer wavelength is emitted from the donor fluorochrome upon excitement, which in turn excites the acceptor fluorochrome. The SYBR green dye method is simpler than the hybridisation probe method. As DNA is generated, the SYBR green dye is incorporated into the molecule and its fluorescence increases; further increases are detected as the amount of DNA exponentially increases.

Compared to the hybridisation probe method, the SYBR green method has a reduced specificity and non-specific DNA products and primer dimers can be detected (Kern *et al*, 2005).

Sensitivity & Specificity. Despite studies demonstrating RQ-PCR as less sensitive by 10 times than nested PCR, it is a technique of choice. Unfortunately the high sensitivity of this method makes it vulnerable to the generation of false-positives from contamination (Schuler & Dolken, 2006). The definition of the sensitivity of a method in most MRD studies applies to lowest proportion of leukaemic cells that can be measured, although it is debated that sensitivity applies to the lowest proportion of leukaemic cells that can be distinguished from healthy or leukaemia-free, regenerating bone marrow (Steinbach & Debatin, 2008). The sensitivity of RQ-PCR assays can be affected by a number of variables. Similar to qualitative RT-PCR, factors occurring both pre- and during-PCR can greatly influence the sensitivity of detection (See *Quality Control and Standardisation* below). Important effectors of sensitivity to be aware of and control include the amount of RNA or DNA used, the quality of each sample and the number of replicates performed (Ross *et al*, 2009). The sensitivity of an assay can be determined through an experiment of serial dilutions of leukaemic cells in a healthy bone marrow sample; the lowest amount of leukaemic cells that can generate a positive result is deemed the sensitivity (Steinbach & Debatin, 2008). The specificity of an RQ-PCR assay can be influenced by the choice of SYBR green or probes. Hybridisation probes specifically targets the

sequence of interest whereas non-specific PCR products and primer-dimers can be amplified with SYBR green. When performing SYBR green assays, melting curves can be generated to assess the level of specificity; a single curve generated for each sample suggests the amplification of a single product (Kern *et al*, 2005).

Quality Control and Standardisation. The level of mRNA expression of an AML-specific gene is compared to the level of mRNA expression level of a control gene to account for and control the variability between samples that may arise. House-keeping genes are selected on the basis of them being expressed at similar levels in different cell types (Kern *et al*, 2008). Variability can arise during the pre-PCR stage due the choice of methods utilised by laboratories, for example the type of sample used, the method of mononuclear cell isolation, method of RNA extraction, method of cDNA synthesis. Even during the PCR process variability can arise. Results can be affected by the selection of reagents used, whether primers or probes are used and the PCR machine used, and even by the method of data analysis selected. A joint project between 25 university laboratories in ten European countries was designed via the “Europe Against Cancer” (EAC) programme to develop standardisation and quality control. The primary aim was to establish a standardised protocol to allow for comparison of MRD data for leukaemia bearing an appropriate molecular marker (Gabert *et al*, 2003).

Fusion Genes as Markers of MRD. The most common fusion genes used as MRD markers for detection by PCR are *PML-RAR α* , *AML1-ETO*, and *CBF β -MYH11*; the fusion genes generated from translocations associated with favourable AML. Using fusion gene products as markers allows for highly sensitive and specific MRD detection but due to their low incidence, detection of MRD in this manner only accounts for approximately 25% of adult *de novo* AML cases (Kern *et al*, 2008).

Gene Mutations as Markers of MRD. PCR detection of fusion genes limits the number of cases of AML that can be monitored; patients with a normal karyotype display no fusion gene products resulting from balanced chromosomal translocations. As a result, mutations detected in genes are being increasingly used for MRD monitoring. Assessing the behaviour of specific gene mutations allows for up to approximately 75% of normal karyotype AML cases to be monitored (Kern *et al*, 2008). Commonly monitored gene mutations include; *FLT3*, *NPM1* and *CEBPA*.

Aberrant Gene Expression as a Marker of MRD. As for gene mutations, monitoring the over-expression of genes provides a means of monitoring MRD in patients without a chromosomal translocation resulting in the generation of a fusion gene product. A disadvantage to the use of aberrant gene expression is its limited sensitivity compared to other markers; the genes monitored are rarely “undetected” in normal

healthy cells but instead display a low level background expression (Kern *et al*, 2008). The *WT1* gene has been demonstrated to be useful in monitoring the behaviour of MRD in normal karyotype AML patients.

1.4.2 Current MRD markers

As previously discussed, several sub-clones can exist in AML with a dominating clone directing the disease. Chemotherapy can result in the elimination of one sub-clone allowing complete remission to be achieved but may result in the domination of a minor sub-clone leading to the re-initiation of the disease and relapse through an alternative initiating event. Markers of MRD need to be stable events that are present at both diagnosis and remission, for example a primary event such as a chromosomal marker or a secondary event that is essential to the clone. Secondary abnormalities, for example *FLT3*, are not always present at relapse when previously observed at diagnosis and vice versa, and therefore are not as useful as MRD markers (Shih *et al*, 2004a).

a. PML-RAR α

With the advent of ATRA, APL is now one of most prognostically favourable subsets of AML (Grimwade *et al*, 1998; Grimwade, 2002). Despite the high cure rates, a small subgroup of patients with a poorer prognosis will ultimately relapse. With attempts to further improve cure rates, early studies suggested the monitoring of the *PML-RAR α* fusion gene product. An RT-PCR assay was developed to detect *PML-RAR α*

fusion gene transcripts in leukaemic samples with a greater detection sensitivity than conventional cytogenetics or Northern blot analysis (Castaigne *et al*, 1992; Miller, Jr. *et al*, 1992). Several groups observed a disappearance of the fusion gene following therapy but those patients remaining *PML-RAR α* positive following consolidation therapy or who became *PML-RAR α* positive after achieving a negative PCR result eventually relapsed (Biondi *et al*, 1992; Lo Coco *et al*, 1992; Miller, Jr. *et al*, 1992). Diverio *et al* (1993) went on to confirm an absence of the fusion gene in long-term remission APL patients (Diverio *et al*, 1993). Monitoring of MRD in APL patients is now mainly performed using quantitative RT-PCR, providing reliable prognostic indexes for APL following diagnosis and after consolidation therapy (Gallagher *et al*, 2003; Schnittger *et al*, 2003; Cassinat *et al*, 2009).

b. CBF fusion products

Patients exhibiting a t(8;21) translocation or an inv(16) at diagnosis are given a favourable prognosis and generally a high complete remission rate is achieved following standard induction chemotherapy (Byrd *et al*, 2002). Invariably though, some patients will relapse despite receiving intensive consolidation therapies or an autologous stem-cell transplantation. Studies into the usefulness of MRD monitoring in the CBF leukaemias by qualitative RT-PCR have generated conflicting data; *AML1-ETO* transcripts were detected in several long-term remission patients (Nucifora *et al*, 1993; Saunders *et al*, 1994; Miyamoto *et al*, 1996). As a result, quantitative competitor and real-time PCR

methods have been applied to CBF leukaemias for the monitoring of MRD. Several studies have demonstrated the prognostic value of monitoring MRD in these patients using quantitative methods (Tobal *et al*, 2000; Marcucci *et al*, 2001; Krauter *et al*, 2003)

c. Wilms' tumour (WT1)

Wilms' tumour 1 (WT1) was originally identified as a tumour-suppressor gene involved in the pathogenesis of Wilms' tumour, a kidney neoplasm (Haber *et al*, 1990). The *WT1* gene is found on chromosome 11p13 and spans approximately 50Kb of DNA (Coppes *et al*, 1993). *WT1* is expressed in a variety of tissues including the kidneys, gonads and those involved in haematopoiesis (Call *et al*, 1990) but an increase in expression is observed in Wilms' tumours (Bergmann *et al*, 1997). Many groups have detected low levels of *WT1* expression in normal bone marrow and peripheral blood samples yet high levels in AML samples and other leukaemic samples (Inoue *et al*, 1994) leading to the suggestion of *WT1* as a candidate for MRD monitoring in AML (Inoue *et al*, 1994). An advantage to *WT1* being used as a MRD marker is that it is expressed in 75-80% of patients compared other chromosomal aberrations that occur in only a small number of AML patients. In addition, PCR methods for the detection of *WT1* have a sensitivity of one cell of 10^4 - 10^5 cells (Bergmann *et al*, 1997). Despite these advantages, the use of *WT1* as an MRD marker and its clinical significance has been associated with some controversy. Schmid *et al* (1997) reported *WT1* expression in leukaemic blasts but found no

significant correlation between *WT1* expression and disease-free survival or overall survival. Their data was also too limited to come to a conclusion about the usefulness of *WT1* as a marker of MRD (Schmid *et al*, 1997). Using qualitative RT-PCR, Gaiger *et al* (1998) detected levels of *WT1* expression in AML patients following chemotherapy and marrow transplantation indicating *WT1* had no prognostic significance and could not reliably predict relapse (Gaiger *et al*, 1998). However, results from qualitative PCR methods need to be approached with caution as previously discussed. A low level of *WT1* expression preceding a relapse may not be detected whilst a positive result generated from low levels of *WT1* expression in a healthy normal cell will not always result in a relapse. More recent studies using quantitative PCR methods have demonstrated a positive prognostic value of *WT1* and its ability to predict relapse in AML following both conventional chemotherapy and allogeneic hematopoietic stem cell transplant (HSCT) (Cilloni *et al*, 2003;Ogawa *et al*, 2003;Candoni *et al*, 2009).

d. Preferentially Expressed Antigen of Melanoma (PRAME)

Overexpression of the *PRAME* gene in tumours encodes for a HLA-A24-restricted antigenic peptide (van Baren *et al*, 1998). *PRAME* is expressed at low levels in normal tissue cells but is predominantly expressed in acute leukaemia, in particular t(8;21) AML (Ikeda *et al*, 1997;van Baren *et al*, 1998). Using real-time RT-PCR, Tajeddine *et al* (2006) demonstrated that *PRAME* was a good marker of MRD in

PRAME⁺ AML displaying no other genetic markers. More recently Qin *et al* (2009) detected low levels of both *PRAME* and *WT1* in normal bone marrow and an increased expression of both genes in normal karyotype AML by RQ-PCR. They confirmed that the simultaneous monitoring of *PRAME* and *WT1* provides a more sensitive detection in MRD and may help to avoid false-positives (Qin *et al*, 2009).

1.5 Gene Expression Profiling

1.5.1 Early microarray studies

The identification of unique gene expression profiles has the potential to identify sub-groups within diseases with differing prognostic and therapeutic significance. Golub *et al* (1999) first approached the use of microarrays for disease classification. Using class discovery and a class prediction, previously unrecognisable subtypes were defined and “class predictors” were generated capable of assigning unknown samples to one of two classes. To validate the class prediction approach, Golub *et al* (1999) used 50 gene-predictors previously identified as most closely correlated with ALL-AML distinction in known samples to distinguish between 38 acute leukaemic samples (Figure 4) (Golub *et al*, 1999).

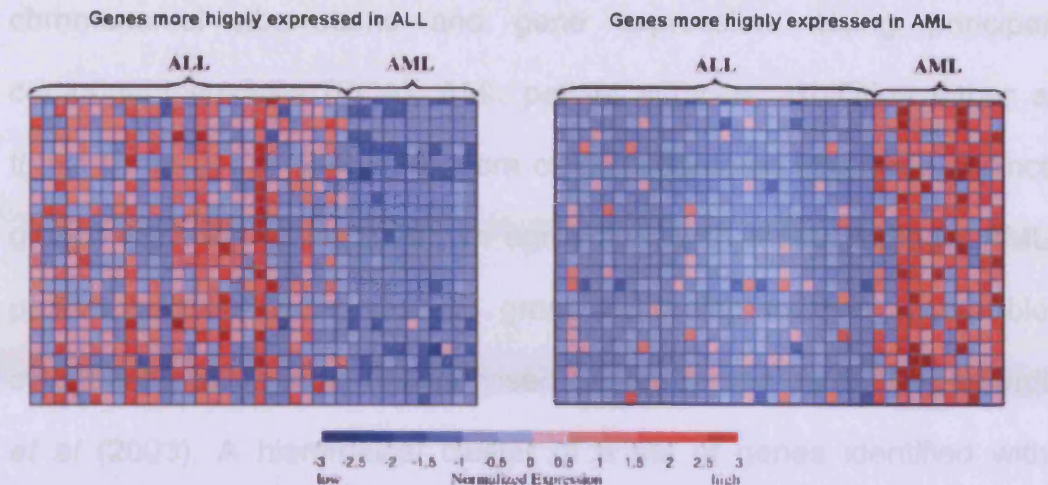


Figure 4. The 50 “predictor” genes used to distinguish between 38 acute leukaemia samples. Each row corresponds to a gene and each column corresponds to a sample. Expression levels greater than the mean are coloured red and those below are coloured blue. Adapted from Golub *et al* (1999) (Golub *et al*, 1999).

1.5.2 Identification of AML-specific sub-groups

Following on from the work by Golub *et al* (1999) many groups investigated the potential of gene expression profiling analysis to identify specific gene-expression signatures that could distinguish between the AML sub-groups and their cytogenetic aberrations, and identify new potential prognostic markers and targets for therapy.

a. Favourable translocations

To determine whether a favourable leukaemia-specific genotype is associated with a distinct gene expression profile, Schoch *et al* (2002) performed gene expression profiling analysis on 37 AML patient samples exhibiting one of the three favourable chromosomal translocations. They confirmed a correlation between AML-specific

chromosomal aberrations and gene expression. Using principal component analysis (PCA), AML patient samples exhibiting either a t(15;17), t(8;21) or an inv(16) were clearly separated into three distinct groups (Schoch *et al*, 2002). In agreement with these findings, AML patient samples were seen to group according to their favourable cytogenetic group in an unsupervised cluster carried out by Debernardi *et al* (2003). A hierarchical cluster of a set of genes identified with statistically significant differences in expression between the five groups analysed in this study (t(15;17), t(8;21), inv(16), 11q23 and normal karyotype) grouped the patient samples according to their cytogenetics (Debernardi *et al*, 2003). However, a more recent study by Bullinger *et al* (2007) demonstrated the molecular heterogeneity within the core binding factor (CBF) leukaemias. Rather than a distinct clustering of t(8;21) patient samples and inv(16) patient samples as previously seen, an unsupervised hierarchical cluster saw CBF AML samples segregate between two heterogeneous clusters. The larger of the two clusters displayed a separation of the t(8;21) patient samples and inv(16) patient samples into two distinct groups whilst the t(8;21) and inv(16) patient samples in the smaller cluster were mixed together (Bullinger *et al*, 2007). More recently, Verhaak *et al* (2009) used PAM (predictive analysis for microarrays) analysis to predict the cytogenetic status of AML patient samples with a favourable karyotype (t(15;17, t(8;21) or inv(16)) with 100% accuracy (Verhaak *et al*, 2009). These studies demonstrated a strong association between the three favourable chromosomal aberrations and their patterns of gene expression. The

distinct gene expression profiles observed within the CBF leukaemia may be a result of other clinical and genomic characteristics of the CBF patients that contribute to their disease and prognosis.

b. Normal Karyotype

Patients with a normal karyotype make up the largest sub-group in AML and are not homogenous in respect to their molecular genetics and therefore their clinical outcome (Marcucci et al, 2005). Gene expression profiling has been applied to normal karyotype AML patients to explore the heterogeneity amongst these patients and recognise distinct expression signatures (Radmacher et al, 2006). As previously discussed, in the study by Debernardi *et al* (2003), unsupervised clustering of 28 AML patient samples saw the clustering of samples based on their cytogenetics; with the exception of two samples clustering with the inv(16) patients, normal karyotype samples formed a distinct separate cluster (Debernardi *et al*, 2003). From the application of a statistical group comparison restriction filter to identify genes with statistically significant differences in expression among the cytogenetic groups, an increased expression of a number of homeobox (*HOX*) genes were seen in the normal karyotype group. Further unsupervised hierarchical clustering of 87 *HOX* genes saw the separation of samples into two distinct clusters; favourable AML with low *HOX* expression separating into one cluster and normal karyotype and 11q23 samples grouping together in the second cluster displaying a high expression of the *HOX* genes. Further analysis identified ten *HOX* genes specifically

over-expressed in normal karyotype AML; *HOXA4*, *HOXA5*, *HOXA9*, *HOXB2*, *HOXB3*, *HOXB5*, *HOXB6*, *HOXB7*, *MEIS1* (myeloid ecotropic viral integration site 1 homolog) and *PBX3* (pre-B-cell leukemia transcription factor 3) (Debernardi *et al*, 2003). Bullinger *et al* (2004) observed a grouping of normal karyotype patients into two distinct clusters; a prevalence of *FLT3* aberrations and FAB M1 and M2 samples were observed in cluster 1 whilst FAB M4 and M5 samples were common to cluster 2. Later studies into expression profiles of AML patients observed the clustering of normal karyotypes was also associated with other molecular aberrations. Valk *et al* (2004) observed a clustering of normal karyotype patients between 3 out of 16 clusters. The majority of the patients grouped into two of these clusters displayed a normal karyotype and an internal tandem duplication of the *FLT3* gene. The third cluster containing mostly patients with a normal karyotype presented with no consistent extra abnormalities (Valk *et al*, 2004). In determining whether or not gene expression profiling could improve outcome prediction and risk classification, Wilson *et al* (2006) identified one out of six clusters of AML patients with the best rate of overall survival; 25% at 5 years. This particular cluster contained a high number of *NPM1* mutations (78%), an unequal number of females (67%) and the majority of normal karyotype patients (75%) in the study (Wilson *et al*, 2006). These studies indicated that gene expression profiling might be useful for risk classification and prediction of outcome but further profiling work was required with the inclusion of other molecular aberrations and the influence they have on prognosis.

c. Gene mutations and aberrant gene expression

AML exhibiting mutations and/or aberrant expressions of *FLT3*, *NPM1*, *CEBPA* and *EVI1* have been associated with specific gene expression signatures and several investigations have been carried out to determine whether these signatures have the ability to predict the presence or absence of these mutations and determine risk and prognosis of the patients. In a study by Alcalay *et al* (2005) unsupervised hierarchical clustering of *de novo* AML displaying no AML-associated chromosomal translocations demonstrated the strongest clustering of patients based on their *NPM1* mutation status; *FLT3* mutations had no influence on the clustering (Alcalay *et al*, 2005). Within the cluster of samples displaying aberrant cytoplasmic *NPM1* localisation due to a mutation of the *NPM1* gene, a number of *HOX* genes were seen to be up-regulated in the gene expression signature (Alcalay *et al*, 2005). Similar to Alcalay *et al* (2005), Verhaak *et al* (2005) also identified a *HOX*-specific gene signature associated with mutant *NPM1* AML patients. Through PAM analysis, they were also able demonstrate the positive predictive power for *NPM1* mutations in both their training and independent data sets (Verhaak *et al*, 2005). Investigation into *FLT3*-specific gene expression profiles has generated some contradictory results. Several studies of normal karyotype samples have demonstrated the positive accuracy of specific gene signatures to separate AML samples based on the presence of a *FLT3*-TKD or a *FLT3*-ITD and of the presence of a *FLT3*-TKD or wild-type *FLT3* (Neben *et al*, 2005;Whitman *et al*, 2008). However, in an analysis

of paediatric AML, Lacayo *et al* (2004) observed a separation between wild-type *FLT3* samples and *FLT3*-ITD samples but several wild-type samples were seen to cluster with the *FLT3*-ITD samples, reducing the predictive power of *FLT3*-specific gene signatures (Lacayo *et al*, 2004).

In a gene expression analysis previously discussed, Valk *et al* (2004) identified two clusters of AML patients displaying a high frequency of *CEBPA* mutations, which are associated with a favourable outcome. The two clusters differed from each other by the genes that were up- or down-regulated in the samples (Valk *et al*, 2004). Combining the two clusters containing patients with a *CEBPA* mutation and using PAM analysis to identify genes capable of predicting karyotype, the presence of a *CEBPA* mutation was detected with 98% accuracy. The same group also observed an over-expression of the *EVI1* gene in 11 of 22 patients and in 5 of 14 patients in two clusters. The patients differed between the two clusters in the chromosomal abnormality they displayed; 60% of the *EVI1* over-expressing patients in the first cluster also displayed a chromosome 7 abnormality whilst the patients in the second cluster exhibited an 11q23 aberration (Valk *et al*, 2004). From their investigation, Verhaak *et al* (2009) concluded that gene expression profiling was capable of accurate prediction of the favourable AML subtypes but not for other mutations and molecular aberrations (Verhaak *et al*, 2009).

1.5.3 Risk Classification and therapeutic targeting

Microarray analysis began as a disease classification tool for identifying sub-group specific signatures and class prediction but gene expression profiling is now being directed towards the possibility of it being useful in risk classification and therapeutic targeting. Wilson *et al* (2006) identified six distinct groups based on gene expression with unsupervised hierarchical clustering on a cohort of 170 patients. A cluster dominated by patients with NPM1 mutations was identified with having the best overall and disease-free survival (Wilson *et al*, 2006). A gene expression profiling study was carried out to identify novel gene signatures and markers to predict patient response to the farnesyltransferase inhibitor tipifarnib (Zarnestra). The analysis identified eight genes as potential markers for response to treatment predictors and whilst the overall accuracies were only around 60%, the concept of gene expression profiling for identifying patients who respond to various therapies has been highlighted (Raponi *et al*, 2007).

1.6 Project aims and objectives

Complete remission rates of AML have greatly increased over the last 50 years but some patients still relapse within months of achieving complete remission. One approach to improving treatment outcomes is the development of sub-group specific therapies rather than the standardised management currently being employed. In addition, the effective and accurate monitoring of MRD within these sub-groups has the potential to improve patient outcomes. The aim of this study was to

utilise gene expression profiling as a tool to help provide a better clinical management for specific AML sub-groups.

Aims:

- Establish a “proof-of-principle” by using gene expression profiling to identify genes specific to t(15;17) AML.
- Use gene expression profiling to identify genes specific to CBF AML patients.
- Investigate the potential of the identified genes as markers of minimal residual disease within the favourable prognosis AML sub-groups and compare with markers currently used.
- Use gene expression profiling to identify genes specific to normal karyotype AML patients.
- Validate the expression of these genes and their ability to act as markers of minimal residual disease.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Primer sequences

	Gene	Direction of primer	Primer sequence (5' - 3')
Control genes	<i>S14</i>	Forward	ggcagaccgagatgaactct
		Reverse	ccagggtccaggggtcttgg
	<i>ABL</i>	Forward	cccaaccttttcgtgcactgt
		Reverse	cggctctcggaggagacgtaga
t(15;17)	<i>IGFBP2</i>	Forward	agcatggcctgtacaacctc
		Reverse	aagcctcctgctgctcatt
	<i>HGF</i>	Forward	acctgcaatggggagagagta
		Reverse	gggtgtgagggtcaagagta
t(8;21)	<i>PRAME</i>	Forward	tcagggagtgctgtgtgag
		Reverse	tgacatcctggcttagtg
	<i>POU4F1</i>	Forward	tgctgagcttctccttgg
		Reverse	caggacctggcccagataa
	<i>IL5RA</i>	Forward	ccctgaggacacgcagtatt
		Reverse	tgatcaaagggcctgatagc
	<i>ETO</i>	Forward	cccctccaccatagagaca
		Reverse	cgttactggccctctgtgtt
inv(16)	<i>ST18</i>	Forward	tttgagcaggatgtttgtga
		Reverse	gcatacatcattgccttttg
	<i>MYH11</i>	Forward	cagttcgaaagggatctcca
		Reverse	gtagctgcttgatggctcc
	<i>MN1</i>	Forward	acgaccccagaatgaaacag
		Reverse	ggattgcctaacacacagca
Normal	<i>HOXA5</i>	Forward	agatctaccctggatgcgc
		Reverse	ccttctccagctccagggtc
	<i>HOXA9</i>	Forward	gagtggagcgcgcataag
		Reverse	ggtgactgtcccacgctgac
	<i>HOXB6</i>	Forward	aggacaagagcgtgttcggc
		Reverse	ggcccaaaggaggaactgtg
	<i>WT1</i>	Forward	caggctgcaataagagatatttaagct
		Reverse	gaagtcacactggtatggttctca
	<i>COL4A5</i>	Forward	ttaagaccatatgcccctg
		Reverse	gaggccacaggaggattctt
cDNA synthesis	Affy_T7		ggccagtgaattgtaatacgactcact atagggaggcgggtttttttttttttttttttttt
siRNA	<i>PRAME</i>	Sense	gaggcuugccuauucugcau99
	<i>PRAME</i>	Antisense	augcagauaggcaagccuc99
	<i>MN1</i>	Sense	gcguugagucaucaacacu99
	<i>MN1</i>	Antisense	aguguugaugacucaacgc99
	<i>HOXA1</i>	Sense	gagcagaguuuacgggucu99
	<i>HOXA1</i>	Antisense	agacccguaaacucugcuc99

Table 1. The 5'-3' oligonucleotide sequences used for the study.

All oligonucleotide primer sets for RQ-PCR and siRNA transfection were supplied by Eurogentec Ltd, Southampton, Hampshire, UK.

2.1.2 Suppliers

Reagents used were supplied by Affymetrix UK Limited (High Wycombe, UK), Agilent Technologies UK Ltd (Stockport, Cheshire, UK), Alexis Biochemicals (Bingham, Nottingham, UK), Ambion Inc (Applied Biosystems, Warrington, UK), Applied Biosystems (Warrington, UK), Beckman Coulter UK Ltd (High Wycombe, Buckinghamshire, UK), Qiagen (Crawley, West Sussex, UK), DAKO Ltd, Ely, Cambridgeshire, UK, Invitrogen Ltd (Paisley, UK), R&D Systems Europe Ltd (Abingdon, Oxfordshire, UK), Roche Diagnostics Ltd (West Sussex, UK) and Sigma-Aldrich Company Ltd (Dorset, UK).

2.1.3 Enzymes and PCR reagents

The buffers (1st strand and 2nd strand), DNA ligase, DNA polymerase 1, RNase H and SuperScript 11 reverse transcriptase used for the gene chip arrays were supplied by Invitrogen Ltd (Paisley, UK),

Deoxynucleotide triphosphates (dNTPs), DNA polymerase, RNase inhibitor, MuLV (Murine Leukaemia Virus) reverse transcriptase, AmpliTaq gold DNA polymerase, MgCl₂ and buffers for PCR and reverse transcription PCR were supplied by Applied Biosystems Limited (Warrington, UK).

LightCycler FastStart enzyme, LightCycler FastStart reaction mix, MgCl₂ and PCR grade water used for the real-time quantitative PCR (RQ-PCR) were supplied by Roche Diagnostics Ltd (West Sussex, UK).

2.1.4 Commercial Kits

The precise compositions of some of the solutions provided with commercial kits were not supplied by the manufacturer.

RNeasy mini kit and RNase-free DNase Set were supplied by Qiagen, Crawley, West Sussex, UK. Buffers included in the RNeasy mini kit are as follows;

- RW1 (Wash buffer 1): used to wash away all components other than RNA bound to the column (i.e. DNA and proteins)
- RLT (Lysis buffer): contains a chaotropic guanidine salt to allow for the binding of RNA to the silica membrane of the column
- RPE: used to wash the RNA and removes the chaotropic guanidine salt left over from buffer RLT and RW1 and prepares the RNA for the final elution.

LightCycler FastStart DNA Master SYBR Green 1 was supplied by Roche Diagnostics Ltd, West Sussex, UK.

GeneChip® Sample Cleanup Module and Enzo BioArray HighYield RNA transcript Labelling kit were supplied by Affymetrix UK Limited, High Wycombe, UK.

Human hepatocyte growth factor (HGF) Quantikine enzyme-linked immunosorbent assay (ELISA) kit supplied by R&D Systems Europe Ltd, Abingdon, Oxfordshire, UK. Buffers included in the HGF ELISA kit are as follows;

- Calibrator diluent RD5P concentrate (5X): a concentrated buffered protein base with preservatives
- HGF standard: a recombinant human pro-HGF in a buffered protein base with preservatives
- Colour A reagent: a stabilised hydrogen peroxide
- Colour B reagent: a stabilised chromogen (tetramethylbenzidine)
- Assay diluent RD1W: a buffered protein base with preservatives
- HGF conjugate: a polyclonal antibody against HGF conjugated to horseradish peroxidase with preservatives
- Stop solution: 2 N sulphuric acid

Silencer siRNA Starter kit and Silencer siRNA Labelling kit supplied by Ambion Inc. (Applied Biosystems), Warrington, UK.

2.1.5 Composition of Solutions

1x siRNA annealing buffer; 50mM Tris, pH7.5-8.0, 100mM NaCl in DEPC (diethylpyrocarbonate)-treated water.

2.1.6 Antibodies

CD11b and CD38 mouse anti-human antibodies were supplied by DAKO Ltd, Ely, Cambridgeshire, UK.

2.1.7 Laboratory Equipment

Additional laboratory and tissue culture consumables were supplied by Bioquote Ltd (York, UK), Eppendorf UK limited (Histon, Cambridge, UK) Fisher Scientific UK Ltd (Loughborough, Leicestershire, UK), Immune Systems Ltd (Paignton, UK) and Thermo Life Sciences (Basingstoke, Hampshire, UK).

2.1.8 Cell Lines

The cell lines were originally obtained from their specific source and maintained within the department.

Cell Line	Cell type	Translocations	Source	Culture conditions
NB4	Human promyelocytic	t(15;17)	Beatson Institute for Cancer Research, Glasgow, UK	RPMI 1640, 10% FBS, 37°C, 5% CO ₂
HL-60	Human	None*	Dr Chris Bunce, Dept Medicine, University of Birmingham, UK	RPMI 1640, 10% FBS, 1% Pen/strep, 37°C, 5% CO ₂
Me-1	Human	Inv(16)	DSMZ, Braunschweig, Germany	RPMI 1640, 20% FBS, 1% Pen/strep, 37°C, 5% CO ₂
Kasumi-1	Human	t(8;21)	Dr. George Follows, Dept Molecular Medicine, Leeds University, UK	RPMI 1640, 20% FBS, 37°C, 5% CO ₂

Table 2. The source and culture conditions of the cell lines used. *Whilst HL-60 cells have no specific translocations, they do have a complex karyotype. (FBS - foetal bovine serum)

2.2 Patient Database and Gene Expression Profiling

2.2.1 Patient and Donor Samples

2.2 Patient Database and Gene Expression Profiling

2.2.1 Patient and Donor Samples

Gene expression microarrays were used to profile the diagnostic samples of 290 AML patients and 10 normal healthy donors. Duplicate microarray chips were performed for five of the AML patients, resulting in a total of 295 AML patient samples used in this study. The patients' clinical and molecular characteristics are presented in Table 3. The samples were obtained from the School of Medicine, Cardiff University, University Hospital Wales or from University Hospital, London, and processed by the School of Medicine, Cardiff University, University Hospital Wales. Patient confidentiality and permission to use excess biological material for research was been approved by the Multicentre Research Ethics Committee (MREC) (Appendix I).

Characteristic	Value
Sex - no. (%)	
Male	144 (49.7) (<i>*3 rep.</i>)
Female	132 (45.5) (<i>*2 rep.</i>)
Not determined	14 (4.2)
Age group – no. (%)	
<30 years	29 (10.0)
30-60 years	130 (44.8)
>60 years	119 (41.0)
Not determined	12 (4.2)
Age – years	
Median*	56
Range*	16-87
French-American British Classification – no. (%)	
M0	14 (4.8)
M1	53 (18.3)
M2	66 (22.8) (<i>*1 rep.</i>)
M3	30 (10.3) (<i>*2 rep.</i>)
M4	51 (17.6)
M5	33 (11.4) (<i>*2 rep.</i>)
M6	3 (1.0)
M7	2 (0.7)
Bi	1 (0.3)
Not determined	37 (12.8)
Cytogenetic Abnormalities – no. (%)	
t(15;17)	26 (9.0) (<i>*2 rep.</i>)
t(8;21)	15 (5.2)
inv(16)/t(16;16)	18 (6.2) (<i>*1 rep.</i>)
+8	6 (2.1)
+11	4 (1.4)
+21	0 (0)
-5	0 (0)
-5(q)	0 (0)
-7	5 (1.7)
3q	1 (0.3)
t(6;9)	3 (1.0)
t(9;22)	1 (0.3)
t(11q23)	4 (1.4)
Complex karyotype	20 (6.9)
Other abnormal karyotypes	21 (7.2)
Normal karyotype	119 (41.0) (<i>*2 rep.</i>)
Not determined	47 (16.3)
Molecular Abnormalities – no. (%)	
Mutation	
FLT3 Internal tandem duplication	63 (21.7) (<i>*2 rep.</i>)
FLT3 Tyrosine kinase domain	34 (11.7) (<i>*1 rep.</i>)

Table 3. The frequencies and percentages of the clinical and molecular characteristics of the 290 patients diagnosed with AML used in this study. *The characteristics of the 5 replicate patient samples are indicated in italics.

2.2.2 Microarray Chips

The microarray chips used to analyse the patient and donor samples were the Affymetrix human genome U133A microarray chips. On a single chip over 14,500 genes are represented by 22,283 oligonucleotides, referred to as probes, complementary to a corresponding region of a gene. The probe sets on each chip were selected from sequences in GenBank, dbEST and RefSeq; databases containing publicly available genomic DNA, transcripts and protein sequences. Each gene is represented by many probes evenly distributed to allow for an accurate reading of the hybridisation signals and to allow for the normalisation of any signals generated from mis-priming. Control probes are also present on the chip in the form of mis-matched probes, differing from the matched probes by one nucleotide. The mis-matched probes are able to hybridise to non-specific sequences in order to measure background 'noise' caused by non-specific interactions occurring on the chip. An intensity level is then generated based on the amount of gene transcript present in the sample taking into account the intensity levels obtained from the matched and mismatched probe sets.

2.2.3 From Patient to Chip

2.2.3a Isolation of Total RNA

The RNA used for the gene expression profiling analysis was extracted from patient mononuclear cells using TRIzol reagent as described in

section 2.4.1d. The extracted RNA was quantified using the Nanodrop 1000 spectrophotometer (Nanodrop products, Thermo Scientific, Wilmington, USA) and visualised on the Agilent Bioanalyser (Agilent Technologies UK Ltd, Stockport, Cheshire, UK) for quality control as described in section 2.4.2. If the RNA concentration was too low for efficient and reliable cDNA synthesis, the RNA was ethanol precipitated or two 1st strand synthesis reactions were run for the sample to be combined at the 2nd strand synthesis step. For ethanol precipitation of RNA, 1/10th volume of 3M Sodium Acetate (pH 5.2), 2.5 volumes of 100% ethanol and 1µl glycogen (5 mg/ml) were added to the RNA and allowed to precipitate at -20°C for at least 1 hour. The sample was then centrifuged at 12,000 x g for 20 minutes at 4°C. The RNA pellet was washed twice with 80% ethanol and centrifuged at 12,000 x g for 10 minutes at 4°C after each wash. The pellet was air-dried and resuspended in an appropriate volume of RNase-free water (10µl for 8-16µg total RNA/reaction).

2.2.3b cDNA Synthesis

The cDNA synthesis was carried out in microfuge tubes as an Applied Biosystems GeneAmp® PCR System 9700 thermal cycler was required for the 1st and 2nd strand synthesis step. Components for the cDNA synthesis were bought individually from Invitrogen Ltd (Paisley, UK).

i. First Strand Synthesis

A volume of 10 µl for 5-8µg total RNA was mixed with 2µl of 50pmol/µl T7 primer. The mix was incubated at 70°C for 10 minutes, put on ice and then briefly centrifuged down to collect the mix at the bottom of the tube. The following mixture was added to the RNA and T7 primer mixture, mixed well and briefly centrifuged down to collect the mix at the bottom of the tube (Table 4);

5 x 1 st Strand Buffer	4µl
DTT (0.1M)	2µl
dNTP mix (10mM)	1µl

Table 4. The reaction mixture used for the 1st strand cDNA synthesis. All reagents used were supplied by Invitrogen Ltd (Paisley, UK). DTT, Dithiothreitol.

The mix was then incubated at 42°C for 2 minutes before adding 1µl of SuperScript II reverse transcriptase enzyme, mixed well and incubated at 42°C for 1 hour.

ii. Second Strand Synthesis

The 1st strand synthesis reaction was placed on ice and briefly centrifuged down to collect the mix at the bottom of the tube. The following mixture was added to the 1st strand synthesis reaction, mixed well and briefly centrifuged down to collect the mix at the bottom of the tube (Table 5).

RNase-free water	91µl
5 x 2 nd Strand Buffer	30µl
dNTP mix (10mM)	3µl
DNA Ligase	1µl
DNA Polymerase I	4µl
RNase H	1µl

Table 5. The reaction mixture used for the 2nd strand cDNA synthesis. *71µl of RNase-free water was used if two separate 1st strand reactions had been carried out. All reagents used were supplied by Invitrogen Ltd (Paisley, UK).

The mix was incubated at 16°C for 2 hours until 10µl of 0.5 M EDTA (ethylenediaminetetraacetic acid) was added to stop the reaction.

The cDNA was cleaned up using the GeneChip Sample Cleanup Module (Affymetrix UK Limited, High Wycombe, UK). A volume of 24ml of ethanol (96-100%) was added to the cDNA wash buffer, as indicated on the bottle, to obtain a working solution. All steps were performed at room temperature. The cDNA reaction mixture was transferred into a 1.5ml microfuge tube. A volume of 600µl of cDNA Binding Buffer was added to the final double-stranded cDNA synthesis reaction mixture (162µl) and vortexed for 3 seconds. The colour of the reaction mixture was checked; the mixture should be yellow, similar to the cDNA Binding Buffer without the cDNA synthesis reaction. If the colour of the mixture was orange or violet, 10ml of 3M sodium acetate (pH 5.0) was added and mixed to turn the colour of the mixture yellow. 500µl of the sample was applied to the cDNA Cleanup Spin column sitting in a 2ml collection tube and centrifuged for one minute at 8000 x g; the flow-through was

discarded. The remaining mixture was added to the column and centrifuged as before, discarding the flow-through and collection tube. The spin column was transferred into a new 2ml collection tube and 750µl cDNA wash buffer was pipetted onto the column. The column was centrifuged for one minute at 8000 x g and the flow-through discarded. The cap of the spin column was opened and the column centrifuged for five minutes at maximum speed. The flow-through and collection tube were discarded. The spin column was transferred to a 1.5ml collection tube and 14µl of cDNA elution buffer was pipetted directly onto the spin column membrane. The column was incubated for one minute at room temperature and centrifuged for one minute at maximum speed to elute.

2.2.3c IVT Reaction

The IVT (*in vitro* transcription) reaction was carried out in an Eppendorf Mixer Comfort Mixer (Eppendorf UK limited, Histon, Cambridge, UK), programmed to incubate at 37°C for 6 hours with 15 seconds of mixing every 30 minutes. The IVT was performed using the cDNA prepared above and the Enzo BioArray High Yield RNA Transcript Labelling Kit, supplied by Affymetrix UK Limited (High Wycombe, UK).

The 12µl cDNA synthesis reaction mixture was added to a 1.5 ml microfuge tube. All solutions, except the enzymes, were brought to room temperature. The following mixture was added to the cDNA synthesis reaction, mixed well and briefly centrifuged down to collect the mix at the bottom of the tube to give a final volume of 40µl (Table 6).

RNase-free water	10µl
10 X HY buffer*	4µl
10 x Biotin-labelled Ribonucleotides	4µl
10 x DTT**	4µl
RNase Inhibitor mix	4µl
T7 RNA Polymerase	2µl

Table 6. The reaction mixture used for the IVT reaction. *HY – HighYield. **The DTT was mixed well to remove any precipitate. All reagents used were from the Enzo BioArray High Yield RNA Transcript Labelling Kit, supplied by Affymetrix UK Limited, High Wycombe, UK

The mix was incubated in the Eppendorf Mixer Comfort mixer for six hours. The cRNA was cleaned up using the GeneChip Sample Cleanup Module (Affymetrix UK Limited, High Wycombe, UK). The IVT reaction mixture was made up to 100µl by the addition of 60µl of RNase-free water and vortexed for 3 seconds. IVT cRNA Binding Buffer (350µl) was added to the sample and mixed by vortexing for 3 seconds. 250µl of ethanol (96-100%) was added to the lysate and mixed well by pipetting. The sample was applied to an IVT cRNA Cleanup spin column sitting in a 2ml collection tube and centrifuged for 15 seconds at 8000 x g; the flow-through and collection tube was discarded. The spin column was transferred to a new 2 ml collection tube and 500µl of IVT cRNA wash buffer was applied and centrifuged for 15 seconds at 8000 x g to wash; the flow-through was discarded. 500µl of 80% ethanol was pipetted onto the spin column and centrifuged for 15 seconds; the flow-through was discarded. The cap of the spin-column was opened and the column centrifuged for 5 minutes at maximum speed; the flow-through and the collection tube were discarded. The spin-column was transferred to a

1.5ml collection tube and 11µl of RNase-free water was pipetted directly onto the RNeasy membrane. After one minute, the column was centrifuged for one minute at 8000 x g to elute. A further 10µl of RNase-free water was pipetted onto the spin membrane and after one minute the column was centrifuged for one minute at 8000 x g. A further 15µl of RNase-free water was pipetted onto the spin membrane and after one minute the column was centrifuged for one minute at 8000 x g.

2.2.2d cRNA Fragmentation

i. cRNA Quantification

The adjusted cRNA yield was calculated using the following equation to reflect the carryover of unlabelled total RNA;

Adjusted Yield = (amount of cRNA measured after IVT, in µg) – (amount of total RNA used in cDNA synthesis reaction) x (fraction of cDNA synthesis reaction used in IVT)

The fragmentation and hybridisation of cRNA was only continued if there was a calculated adjusted cRNA volume of ~15µg.

ii. cRNA Fragmentation

The fragmentation reaction was performed in microfuge tubes and then these tubes were used directly to make up the hybridisation cocktail later. The following mix was used for the fragmentation mix (Table 7)

cRNA (15µg)	1-32µl
5 x Fragmentation buffer	8µl
RNase-free water	Make up to a volume of 40µl

Table 7. The reaction mixture used for the cRNA fragmentation reaction. All reagents were supplied by Affymetrix UK Limited, High Wycombe, UK.

The reaction mixture was incubated at 94°C for 35 minutes and then placed on ice and stored at -20°C until required for the hybridisation.

iii. Gel Analysis

The cRNA and fragmented cRNA were visualised on an Agilent RNA nano chip RNA for quality control as described in section 2.4.2.

2.2.3f Test3 Chip Hybridisation

Test3 chips were routinely run to check the integrity and quality of the cRNA before proceeding to the more expensive full chips. The protocol was set up in such a way that when the hybridisation cocktail was made for a full chip, there was automatically enough to run on a Test3 chip without needing to process more sample. A 1 x and 2 x hybridisation buffer was initially made up. The 20 x GeneChip Eukaryotic Control and B2 Oligo tubes were heated to 65°C for 5 minutes to ensure that the cRNA is completely resuspended. The array chips were removed from the bags and left to equilibrate to room temperature. The rotisserie oven (GeneChip® Hybridisation oven, Affymetrix UK Limited, High Wycombe,

UK) was switched on to warm up. The following mix was used for the hybridisation cocktail (Table 8).

Fragmented cRNA (15µg)	40µl*
Control Oligo B2	5µl
20 x Eukaryotic Hybridisation Controls	15µl
Herring Sperm DNA (10 mg/ml)	3µl
Acetylated BSA (50 mg/ml)	3µl
2 x Hybridisation Buffer	150µl
RNase-free water	Final volume of 300µl

Table 8. The reaction mixture used for the hybridisation reaction. *Fragmentation reaction mixture. BSA - Bovine Serum Albumin. All reagents used were supplied by Affymetrix UK Limited, High Wycombe, UK

The hybridisation cocktail was heated to 99°C for 5 minutes, followed by 45°C for 5 minutes in a PCR thermal cycler machine. Whilst the hybridisation cocktail was denaturing, the Test3 chip was filled with 100µl of 1 x Hybridisation Buffer. The chip was turned upside down and a venting tip was placed in the septum at the bottom of the chip and filled via the other one, to reduce leakage. The Test3 chip was pre-hybridised in the rotisserie oven for 10 minutes at 45°C, rotating at 60 rpm. The hybridisation cocktail was spun down at maximum speed in a microfuge for 5 minutes to remove any insoluble material from the hybridisation mixture. When the hybridisation cocktail was ready, and the chip was finished pre-hybridising, the chips were removed from the oven. The pre-hybridisation solution was removed from the Test3 chips and replaced with 80 µl of hybridisation cocktail. The chips were then

hybridised in the rotisserie oven for 16 hours at 45°C, rotating at 60 rpm. The remaining hybridisation cocktail was stored at -20°C.

2.2.3g Washing, Staining and Scanning of Test3 Chips

Buffer A and buffer B were filter sterilised before use and used to prime the fluidics station before use. The array chips were removed from the rotisserie oven and the hybridisation cocktail was removed from the chip and replaced in the tube with the remainder of the hybridisation cocktail and stored at -20°C. The chip was filled with 100µl of Wash Buffer A and placed in the fluidics station. A volume of 1200µl of stain solution was made up in an eppendorf, and split between two eppendorfs, each with 600µl, and 600µl of antibody solution was made up in a separate eppendorf (Table 9).

Stain Solution	
2 x Stain Buffer	160µl
Water	540µl
Acetylated BSA (50mg/ml)	48µl
SAPE (1mg/ml)	12µl
Antibody Solution	
2 x Stain Buffer	300µl
Water	266.4µl
Acetylated BSA (50mg/ml)	24µl
Normal Goat IgG (10mg/ml)	12µl
Biotinylated antibody (0.5 mg/ml)	3.6µl

Table 9. The mixtures used for the stain solution and the antibody solution. BSA - Bovine serum albumin. SAPE - Streptavidin-phycoerythrin dye. All reagents used were supplied by Affymetrix UK Limited, High Wycombe, UK

The required protocol was then run, adding the stain solution tubes and the antibody stain tube when required. When the protocol was finished, the Test3 chip was ejected and checked for bubbles. If bubbles were present the chip was replaced into the fluidics station and refilled with wash buffer A. When there were no bubbles visible on the chip, the chip was checked for leaks and then scanned in the scanner.

2.2.3h Full Chip Hybridisation

The array chips were removed from their bags and left to equilibrate to room temperature. The pooled hybridisation cocktail from the Test3 chips was thawed and heated to 99°C for 5 minutes, followed by 45°C for 5 minutes in a PCR thermal cycler. The hybridisation cocktail was then centrifuged at maximum speed in a microfuge for 5 minutes. Whilst the hybridisation cocktail was denaturing, the full chip was filled with 250µl of 1 x Hybridisation Buffer as described previously in section 2.2.3f. The full chips were pre-hybridised in the rotisserie oven for 10 minutes at 45°C, rotating at 60 rpm. When the hybridisation cocktail was ready, and the chip had finished pre-hybridising, the chip was removed from the oven. The pre-hybridisation solution was removed from the full chip and replaced with 200µl of hybridisation cocktail. The full chip was hybridised in the rotisserie oven for 16 hours at 45°C, rotating at 60 rpm, secured into the oven racks.

2.2.3i Washing, Staining and Scanning of Full Chips

The full chip was washed and stained as described previously in section 2.2.3g but with 250µl of Wash Buffer A pipetted into the full chip. The chip was placed in the fluidics station and the required protocol was selected. Following the completion of the protocol, the full chip was scanned in the scanner.

2.2.4 Data Normalisation

The gene expression data generated from the 305 Affymetrix human genome U133A microarray chips was analysed using GeneSpring GX 7.0 (Silicon Genetics, Agilent Technologies, Wokingham, Berkshire, UK). To allow for a comparison of gene expression changes between experiments and between experiments performed by other centres, the gene expression data was initially normalised by applying a target (TGT) normalisation value. The overall average intensity of every chip was scaled to 100 by multiplying the average intensity by a scaling factor. The data was also normalised using the GeneSpring software to standardise the data further with a per-chip normalisation and a per-gene normalisation. A per-chip normalisation is used to control variations in signal intensity across the chips due to experimental error. The expression value of each data point for a gene is normalised to the 50th percentile of all the data points. The per-gene normalisation allows for differences in detection efficiency between genes, enabling relative changes in gene expression to be compared. The normalisation is

achieved by dividing each signal value for a gene by the median of all the signals involved in the experiment (Silicon Genetics Tech notes).

Details of the probe sets identified, including name and nucleotide sequence was obtained from the Affymetrix NetAffx analysis centre (Affymetrix, <http://www.affymetrix.com/analysis/index.affx>).

2.3 Tissue Culture

2.3.1 Cell counts and viability

2.3.1a Fast Read Disposable Counting Chamber

Cells were counted using a Fast Read Disposable counting chamber. An aliquot of 10 μ l of cell culture was removed from the flask and mixed with 10 μ l 0.4% trypan blue. A 10 μ l aliquot of cells/trypan blue was placed on the counting chamber and analysed by a light microscope. The number of cells was calculated as:

$$\text{Cell count (16 squares)} \times \text{dilution factor} \times 1000 = \text{number cells/ml}$$

Cell viability was assessed using the trypan blue. Live cells excluded the dye whilst dead cells were stained blue. The number of live cells was expressed as a percentage of the total number of cells.

2.3.1b Vi-CELL™ XR Cell Viability Analyser

Cell growth and viability were assessed using a Vi-CELL™ XR cell viability analyser (Beckman Coulter, High Wycombe, UK), which uses the trypan blue exclusion method for calculating cell viability. An aliquot of 500µl of cell culture was used for growth and viability measurements.

2.3.2 Cryopreserving of lines

For long term storage, cultured cells were stored in liquid nitrogen. Cells (5×10^6) were centrifuged at 300 x g for 5 minutes and resuspended in 1 ml IMDM (Iscove's Modified Dulbecco's Medium) medium containing 20% foetal bovine serum (FBS) and 10% dimethyl sulphoxide (DMSO). The cells were aliquoted into 1.8ml cryopreservation vials and placed in a freezing box half filled with isopropan-1-ol to ensure slow freezing when placed at -80°C.

When required for use, cells were rapidly thawed at 37°C and 1ml of pre-warmed media was added drop-wise to the cryopreservation tube. The cells were transferred to a universal container and slowly made up to a volume of 5ml with pre-warmed media. The cells were then centrifuged at 300 x g for 5 minutes to harvest and resuspended in their appropriate culture medium.

2.3.3 Extraction of Mononuclear Cells from Bone Marrow Aspirates and Peripheral Blood Samples

An initial baseline count of the samples was performed using a Vi-CELL™ XR cell viability analyser (Beckman Coulter, High Wycombe, UK). A dilution medium was prepared using MEM (Minimum Essential Media) with 1% Pen/Strep (1ml/100ml) and 300U Heparin per 20ml. The sample was diluted with the MEM dilution medium to ensure there were no more than 100×10^6 cells per ficoll tube. If a significant dilution was required, the mixture included FCS (foetal calf serum) as a thickening agent. A volume of 6-7mls of Ficoll (histopaque) was aliquoted into a universal container. The same volume of diluted blood as Ficoll used was carefully run down the side of the universal container to form a layer on top of the Ficoll. The ficoll tube was centrifuged at $400 \times g$ for 30 minutes with the brake off. The mononuclear cell (MNC) interface layer was carefully removed from the ficoll tube using a syringe and quill. The quill tip was placed just above the interface layer and suction applied while gently moving the quill tip around horizontally. Care was taken not to disturb the red cell pellet. The interface layer was placed into a fresh universal container and diluted with MEM dilution medium to even out volumes if multiple ficoll tubes were used. The tube was centrifuged again at $400 \times g$ for 10 minutes. The supernatant was removed to leave a pellet of MNC's at the bottom of the tube. The cells were resuspended in an appropriate volume of MEM dilution medium and pooled into one container if more than one Ficoll tube was used. To calculate the

volume to add it was assumed there would be a 25% loss from the initial total cell count and enough was added to give 100×10^6 per ml. The number of MNC's was again counted using the coulter counter to give the number of cells/ml of cell preparation. The cells were either cryo-preserved for use later (Section 2.3.2) or continued onto RNA extraction (Section 2.4.1).

2.4 RNA: From Extraction to cDNA

2.4.1 RNA Extraction

2.4.1a Donated Patient and Normal Healthy Donor Material

The AML patient and healthy donor samples used for gene expression microarray analysis were obtained and processed by the School of Medicine, Cardiff University. The RNA was extracted from mononuclear cell of bone marrow aspirates using TRIzol reagent. The patient and normal healthy donor RNA and cell lysates used for real-time quantitative-PCR analysis were donated by the School of Medicine, Cardiff University and the Department of Clinical Haematology, Manchester Royal Infirmary. The centres extracted RNA from mononuclear cells of bone marrow aspirates or peripheral blood samples using either TRIzol reagent (Cardiff and Manchester) or the Qiagen RNeasy Mini kit (Cardiff). The cell lysates donated by the School of Medicine, Cardiff University, were patient or normal healthy donor mononuclear cells from bone marrow aspirates resuspended in

Qiagen RLT buffer and β -mercaptoethanol ready for RNA extraction using the Qiagen RNA extraction protocol.

2.4.1b Leukaemic Cell lines and Normal Healthy Donor Samples

Cells from leukaemic cell lines cultured within the department (1×10^7 cells) or from normal healthy donor mononuclear cells from bone marrow aspirates (5×10^6) were harvested by centrifugation at $300 \times g$ for 5 minutes. The supernatant was removed and the cell pellet was disrupted through the addition of 600 μ l or 350 μ l, respectively, of Qiagen RLT buffer and β -mercaptoethanol and passing the cell lysate through a 20-gauge needle fitted to an RNase-free syringe at least 5 times. The cell lysates were either stored at -80°C until required or continued through the extraction protocol.

2.4.1c Qiagen RNeasy Mini Protocol

A mononuclear cell pellet of 1×10^7 or of less than 5×10^6 cells were loosened and disrupted by the addition of 600 μ l or 350 μ l, respectively, of RLT buffer and the cell lysate passed through a 20-gauge needle fitted to an RNase-free syringe at least 5 times. A 1x volume of 70% ethanol was added to the homogenised sample and mixed by pipetting the sample up and down. Up to 700 μ l of the sample was applied to an RNeasy mini spin column and centrifuged for 15 seconds at $8000 \times g$. The flow-through from the column was discarded following centrifugation. Any remaining sample was added to the column and centrifuged for 15 seconds at $8000 \times g$. The flow-through from the

column was discarded following centrifugation. An on-column digestion of DNA was performed by adding a volume of 350µl of buffer RW1 to the column and centrifuging the column for 15 seconds at 8000 x g and discarding the flow-through. 10µl of DNase I stock solution was added to 70µl RDD buffer and mixed gently by inversion of the tube. The incubation mix was applied directly to the spin column membrane and left at 20-30°C for 15 minutes. A volume of 350µl of buffer RW1 was added to the column and the column was centrifuged for 15 seconds at 8000 x g and the flow-through was discarded. Following the DNA digestion, the RNeasy column was transferred to a clean collection tube and 500µl of RPE buffer was added to the column. The column was centrifuged at 8,000 x g for 15 seconds and the flow-through was discarded. 500µl RPE buffer was again added to the column and the column was centrifuged at 8000 x g for 2 minutes. The column was transferred to a new collection tube and centrifuged at a speed of 16000 x g for 1 minute to remove last traces of buffer. The column was then placed in a 1.5ml collection tube and 30µl of RNase-free water was added directly to the membrane at the base of the column. The column was spun at 8000 x g for 1 minute to elute the RNA. The RNA was stored at -80°C until required.

β-Mercaptoethanol was added to the RLT buffer and 100% ethanol was added to the RPE buffer before use.

2.4.1d TRIzol RNA Extraction Protocol

A cells pellet of $0.5-1 \times 10^7$ cells was homogenised by the addition of 1ml TRIzol reagent. The cell pellet was resuspended by pipetting to lyse the cells. A pre-dispensed Phase Lock Gel-Heavy tube was centrifuged at $1500 \times g$ for 30 seconds to collect the gel to the bottom of the tube. The cell lysate was then added to the tube and incubated at $15-30^\circ\text{C}$ for 5 minutes. A volume of 0.2ml of chloroform was added to the lysate in the Phase Lock Gel-Heavy tube and shaken vigorously for 15 seconds. The tube was centrifuged at no more than $12000 \times g$ for 10 minutes at $2-8^\circ\text{C}$. After centrifugation a clear, aqueous phase should be entirely atop the Phase Lock Gel and a red phenol chloroform phase and cloudy interphase should be below the Phase Lock Gel layer. The addition of chloroform and centrifugation was repeated if this was not the case. The aqueous phase containing the RNA was transferred to a new tube and the RNA was precipitated by the addition of 0.5ml isopropyl alcohol per 1ml TRIzol Reagent used initially. The sample was mixed by repeated inversion and allowed to incubate at $15-30^\circ\text{C}$ for 10 minutes and then centrifuged for 10 minutes at no more than $12000 \times g$ at $2-8^\circ\text{C}$. The RNA pellet was visible on the side and the bottom of the tube. The supernatant was decanted and 1ml of 75% ethanol per 1ml TRIzol reagent used initially was added to wash the RNA pellet. The sample was mixed to dislodge the pellet; a vortex was used if necessary, and centrifuged at no more than $7500 \times g$ for 5 minutes at $2-8^\circ\text{C}$. The supernatant was carefully decanted and the RNA pellet was left to air-dry for 5-10 minutes to remove any residual ethanol. The RNA pellet

was dissolved in 30µl RNase-free water and incubated at 55-60°C for 10 minutes to facilitate dissolution.

2.4.2 RNA quantification and visualisation

Extracted RNA was quantitated using the Nanodrop 1000 spectrophotometer (Nanodrop products, Thermo Scientific, Wilmington, USA). The instrument was blanked using 1.5µl of RNase-free water before applying 1.5µl of the extracted RNA to the spectrophotometer. The absorbance of the RNA was measured at 230nm, 260nm and 280nm to determine purity and concentration of the RNA.

The Agilent Bioanalyser (Agilent Technologies UK Ltd, Stockport, Cheshire, UK) produces profiles of RNA and DNA that can give an indication of their quality. Samples are mixed with a dye that specifically binds to either RNA or DNA and travels from the sample wells through microchannels. Each sample is injected into a separation channel where it is separated based on its size. At the detection point the fluorescence generated from the sample as it passes through a laser is detected before moving onto the waste/buffer well. Samples are visualised either as a graph of fluorescence against time or as a conventional electrophoresis gel showing bands of different sizes.

The extracted RNA was visualised on an Agilent RNA nano chip. All reagents required were left at room temperature for 30 minutes before use. A gel-dye mix was prepared by the addition of 550µl of RNA 6000

nano gel to a spin filter and centrifuged at 1500 x g for 10 minutes. RNA 6000 nano dye concentrate was vortexed and 1µl added to a 65µl aliquot of the filtered RNA 6000 nano gel. The mix was thoroughly vortexed for proper mixing before being centrifuged at 13000 x g for 10 minutes. 9µl of gel-dye mix was loaded into the gel well of a RNA nano chip and pressurised into the capillaries of the chip using a chip priming station. A further 9µl of gel-dye mix was added to the remaining 2 gel wells. A volume of 5µl of RNA 6000 nano marker was added to the 12 sample wells and to the ladder well. 1µl of RNA ladder previously heat denatured at 70°C for 2 minutes to minimise secondary structures was thawed and added to the ladder well. 1µl of RNA was added to the sample wells; the RNA sample added must be less than 500ng/µl. To any wells without RNA samples, 1µl of RNA nano 6000 marker was added. The RNA nano chip containing the RNA was vortexed for 1 minute at a speed of 2400 rpm using the IKA vortex mixer (IKA® Werke GmbH & Co. KG, Staufen, Germany) and inserted into the Agilent 2100 Bioanalyser.

2.4.3 Reverse Transcription (RT) of RNA

RNA from patient samples and cell lines was reverse transcribed into complementary DNA (cDNA) using the following reaction mixture and cycling conditions (Table 10 and Table 11);

RNA (3µl total volume made up with water)	1µg
MgCl ₂ (25mM)	4µl
10 x Buffer II	2µl
dATP (10mM)	2µl
dTTP (10mM)	2µl
dCTP (10mM)	2µl
dGTP (10mM)	2µl
Random hexamers (50µM)	1µl
RNase inhibitor (20U/µl)	1µl
MuLV reverse transcriptase (50U/µl)	1µl

Table 10. The reaction mixture used for reverse-transcription of RNA. All reagents used were supplied by Applied Biosystems Limited (Warrington, UK).

Temperature (°C)	Time (minutes)	Cycles
25	10	1
42	30	1
95	5	1
4	Soak	1

Table 11. The cycling conditions used for reverse-transcription of RNA. An initial priming step is performed at 25°C followed by the cDNA synthesis at 42°C. Finally a deactivation step is performed at 95°C. The reverse-transcriptase reaction was performed in an Applied Biosystems GeneAmp® PCR System 9700 thermal cycler.

A “no-RNA” control sample was included. The samples were stored at -20°C until required.

2.5 Real-Time Quantitative-PCR (RQ-PCR)

RQ-PCR was used to quantitate the level of gene expression in the AML patient samples and normal healthy donor samples. The LightCycler PCR system uses air to achieve temperature cycling

resulting in a faster process than conventional thermal blocks. The high surface-to-volume ratio of the reaction capillaries allows a rapid equilibration between the air and the reaction mixture. The optical unit of the LightCycler detects fluorescence from the tip of the capillary allowing a more sensitive method of detection. The fluorescence detected in the reaction corresponds to the amount of PCR product generated. As DNA is produced, more double strand-specific SYBR green dye is incorporated into the molecule and the levels of fluorescence increases, generating an amplification curve (Figure 1). At the end of the reaction the temperature of the thermal chamber is raised causing the double stranded (ds) DNA to denature and release the SYBR green dye. As each dsDNA has its own specific melting temperature (T_m) an analysis of the melting curve generated is useful for determining that the PCR products generated in each reaction are the same; a single peak should be generated for each gene (Figure 2). The T_m of DNA is the temperature at which half the DNA strands are present as double strands and half are present as single strands.

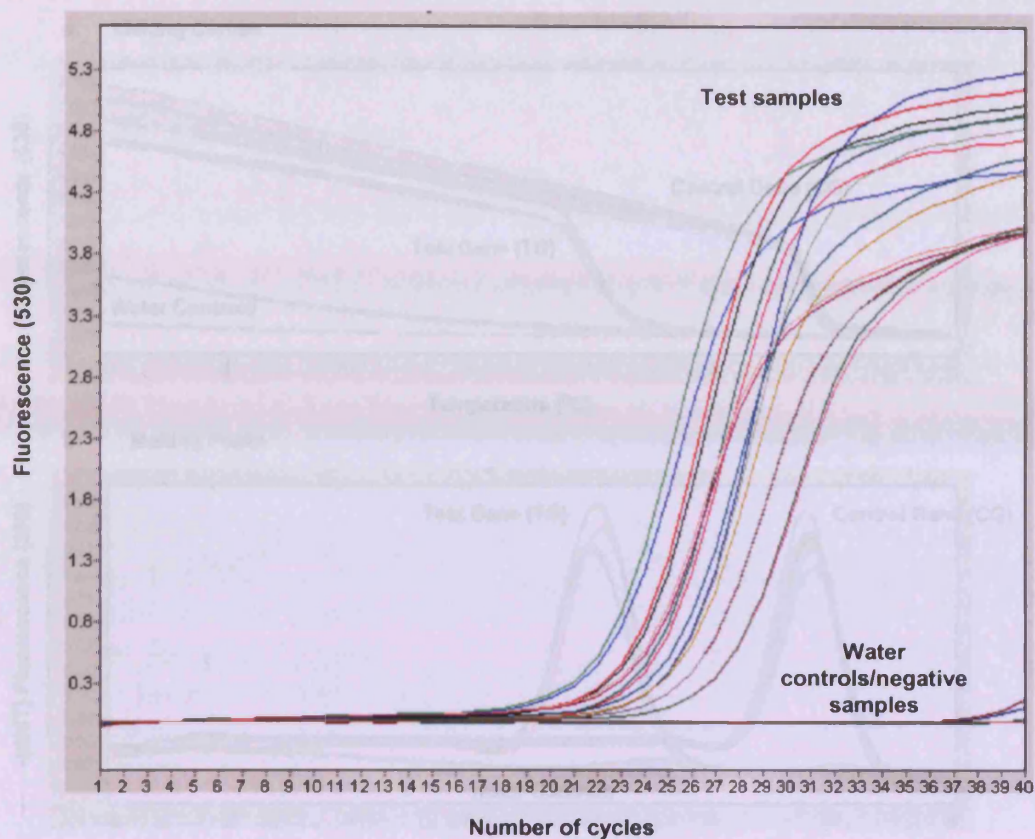


Figure 1. The amplification curves generated as double-stranded (ds) DNA is produced, incorporating more double strand-specific SYBR green dye, resulting in an increase in levels of fluorescence. As no dsDNA will be generated in water controls and RNA-negative samples, there should be no increase in the levels of fluorescence.

2.5.1 Primer design

Primers were designed to produce PCR products of approximately 150-200 base pairs for each gene of interest. The gene sequences used to design the primers were the sequences of the identified probe sets as supplied by the Affymetrix NetAffx analysis centre and primer sequences were generated using Primer 3 select software.

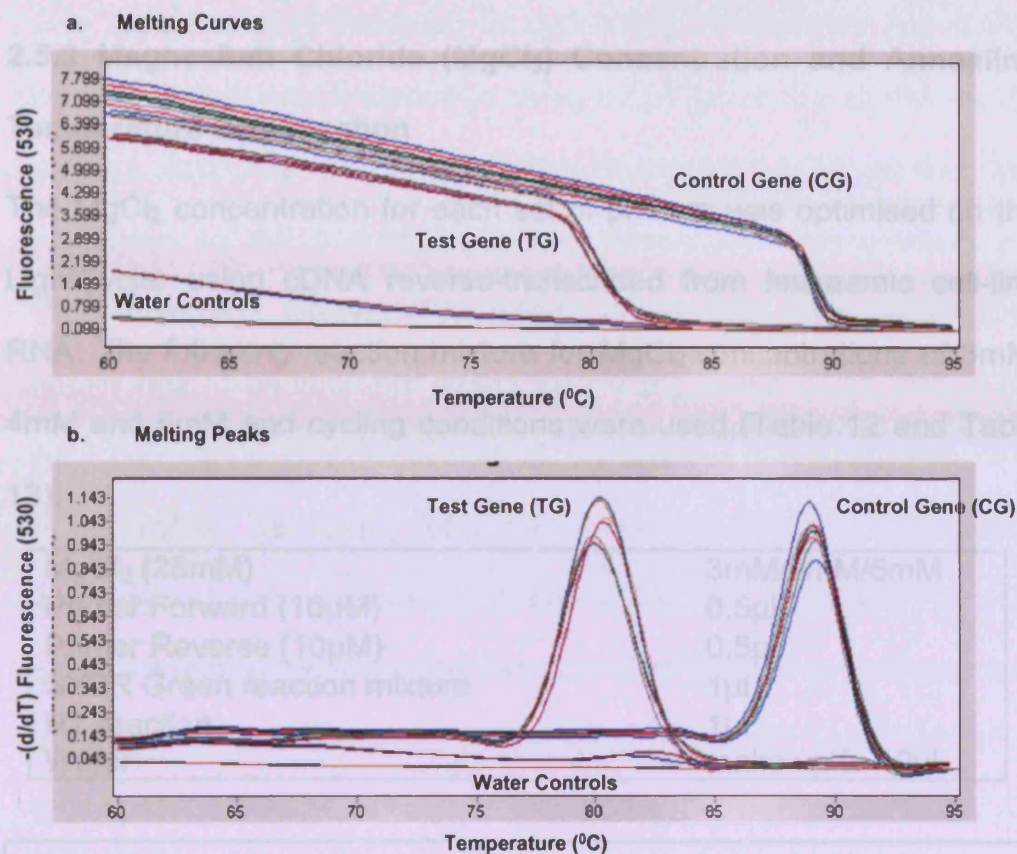


Figure 2. The melting curves (a) and melting peaks (b) generated in the final reaction step as the temperature is increased and the double-stranded (ds) DNA is degraded, releasing the double strand-specific SYBR green dye. A single peak for each gene indicates the PCR product generated in each reaction is the same. As no dsDNA will be generated in water controls and RNA-negative samples, there should be no increase in the levels of fluorescence.

2.5.1 Primer design

Primers were designed to produce PCR products of approximately 150-200 base pairs for each gene of interest. The gene sequences used to design the primers were the sequences of the identified probe sets as supplied by the Affymetrix NetAffx analysis centre and primer sequences were generated using Primer 3 select software.

2.5.2 Magnesium Chloride (MgCl₂) Concentration and Annealing Temperature Optimisation

The MgCl₂ concentration for each set of primers was optimised on the LightCycler using cDNA reverse-transcribed from leukaemic cell-line RNA. The following reaction mixture for MgCl₂ concentrations of 3mM, 4mM and 5mM and cycling conditions were used (Table 12 and Table 13).

MgCl ₂ (25mM)	3mM/4mM/5mM
Primer Forward (10µM)	0.5µl
Primer Reverse (10µM)	0.5µl
SYBR Green reaction mixture	1µl
RT reaction	1µl
Water	make up to 10µl

Table 12. The reaction mixture used for real-time quantitative-PCR. All reagents used were supplied by Roche Diagnostics Ltd (West Sussex, UK).

Programme	Temperature (°C)	Time (Seconds)	Cycles
Denaturation	95	600	1
Amplification	95	3	45
	60	5	
	72	15	
Melting	95	0	1
	70	15	
	95	0	
Cooling	40	30	1

Table 13. The cycling conditions used for real-time quantitative-PCR. The RQ-PCR reaction was performed in the Roche Applied Sciences LightCycler® Carousel-Based thermal cycler.

The samples were repeated in duplicates. A duplicate “no-cDNA” control sample was included in every LightCycler run to ensure the primers were RNA-specific. The accumulation of PCR product was measured by the acquisition of one fluorescent measurement per 72°C extension period. The melting curve data was obtained by continuous measurements during the melting stage at the end of the run. The MgCl₂ concentration that produced a melting curve most reproducible between replicates was used. The annealing temperature was increased or decreased as necessary to also generate reproducible results. The optimised MgCl₂ concentration and annealing temperature used for each set of primers were as follows (Table 14).

Primers	Temperature (°C)	MgCl ₂ Conc. (mM)
HGF2	60	4
ILGFBP2	60	4
ETO	60	5
IL5RA	60	5
PRAME	60	3
POU4F1	60	5
ST18	62	5
MYH11	60	5
MN1	60	4
HOXA5	60	4
HOXA9	60	4
HOXB6	60	4
COL45A	60	4
WT1	60	4

Table 14. The optimised magnesium chloride concentrations and annealing temperatures used for each primer set for Q-PCR.

2.5.3 Optimised LightCycler PCR

The optimised MgCl_2 concentration and annealing temperature for each set of primers was run on the LightCycler using the cycling conditions optimised as described above (Table 14). Per LightCycler run, every sample was run in duplicate for a test gene (TG) and for a control/housekeeping gene (CG). Duplicate “no-cDNA” samples were also included for both the TG and CG. Crossing point (Cp/Ct) values are generated using the second derivative maximum method of analysis; the point at which the rate of change of fluorescence is the fastest is usually where the sample fluorescence can be distinguished from the background fluorescence and the LightCycler software calculates the baseline value by subtracting the mean of the five lowest measured data points from each reading point in the sample. The Cp value generated is the number of cycles at the point where PCR amplification is clearly positive above the background phase. The crossing points are determined during the exponential phase of the PCR reaction (Technical note no. LC 10/update 2003, Roche Applied Science). The mean Cp value for the test genes of duplicate samples was normalised to the mean Cp value of the control genes of the corresponding duplicates. Samples with Cp values greater than 40 cycles were deemed as having an expression of a test gene as “not determined”; i.e. “negative”.

2.6 Treatment of cell lines with NK4 and ATRA

2.6.1 Treatment of cells with ATRA

An *All-trans* retinoic acid (ATRA) solution was made to a concentration of 3.3×10^{-3} M by dissolving 1 µg ATRA in 1 µl DMSO. 30 µl of the 1 µg/1 µl ATRA was added to 970 µl of RPMI + 10% FBS to give a final concentration of 10^{-4} M. The 10^{-4} M ATRA was used at 1 µl/1 ml of cell culture to treat the cells at a concentration of 10^{-7} M. The cells to be treated with ATRA were seeded to the required density and 1 µl of 10^{-4} M ATRA was added per 1 ml of cell culture. A set of the same type of cells were left untreated to act as a control. The cells' growth, viability and gene expression were assessed at time points throughout the experiment.

2.6.2 Determination of HGF protein concentration

The level of hepatocyte growth factor (HGF) protein present in the NB4 cell cultures was determined using the Quantikine® HGF ELISA kit. NB4 cells were grown overnight and volumes of cell culture at a density of 1×10^5 and 1×10^6 cells/ml were centrifuged at 300 x g for 5 minutes to separate the cells and the supernatant. The supernatant was removed and stored at -80°C until required.

A volume of 250 ml of wash buffer was initially prepared from 10 ml of wash buffer concentrate diluted with distilled water. A volume of 100 ml of calibrator diluent (1X) was prepared from 20 ml of calibrator diluent

RD5P concentrate (5X) diluted with distilled water. The HGF standard was reconstituted with 5ml of the calibrator diluent (1X) solution to produce a final concentration of 8000pg/ml. A volume of 500µl of calibrator diluent (1X) was pipetted into six eppendorfs (1.5ml) for a serial dilution of the reconstituted HGF standard; a concentration range of 125 pg/ml-8000pg/ml was generated and the calibrator diluent (1X) served as a 0pg/ml standard. The substrate solution was made up by the addition of 4ml of colour reagent A to 4ml of colour reagent B. The resultant substrate mixture was light sensitive and needed to be used within 15 minutes of its production. When required the cell culture supernatants were defrosted in a 37⁰C water bath. A volume of 150µl of assay diluent RD1W, followed by the addition of 50µl of standard, control or sample was added to each well and mixed by gentle tapping of the plate; the plate was covered with an adhesive strip. After an incubation of 2 hours at room temperature each well was aspirated and washed with 400µl with a multi-channel pipette for a total of 4 times. To ensure complete removal of wash buffer the plate was blotted against paper towels after the final wash. A volume of 200µl of HGF conjugate was then added to each well, the plate was covered with an adhesive strip and incubated at room temperature for 1.75 hours. After this time the process of washing each well with 400µl of wash buffer and aspirating was repeated 4 times. A volume of 200µl of substrate solution was added to each well and the plate was left to incubate at room temperature, protected from the light. After 30 minutes 50µl of stop solution was added to each well causing a colour change from blue

to yellow. Gently tapping the plate ensured mixing of the reagents. The optical density of each well was determined using a microplate reader set to 450nm and to 540nm for correction.

2.6.3 Treatment of cells with NK4

NK4 is an antagonist of HGF and contains the N-terminal domain and the 4 kringle domains of HGF. NB4 cells seeded to the required density were treated with 0.8µl of 10µl/ml NK4 per 1ml of cell culture over a period of 72 hours. Cells were either treated at day 0 only or at day 0 and day 1. A set of the same type of cells were left untreated to act as a control. The cells' growth, viability and cell surface marker expression were assessed at time points throughout the experiment.

2.7 siRNA

2.7.1 siRNA design and production

The siRNA oligonucleotides were designed and synthesised by Eurogentec Ltd, Southampton, Hampshire, UK. Eurogentec were supplied with the identified probe set sequences from the Affymetrix NetAffx analysis centre.

2.7.2 siRNA annealing

Oligonucleotides were centrifuged briefly, resuspended to 100µM using RNase-free water and stored at -70°C until required. The

oligonucleotides were diluted to 50 μ M aliquots for annealing. The following reaction mixture was used (Table 15);

Oligonucleotide sense	30 μ l
Oligonucleotide antisense	30 μ l
5x annealing buffer	10 μ l

Table 15. The reaction mixture used for annealing the sense and antisense oligonucleotides used for siRNA transfection.

The solution was incubated for 1 minute at 95°C and then allowed to cool to room temperature over 60 minutes. The tube containing the reaction mixture was briefly centrifuged to collect the mix to the bottom of the tube and stored at 4°C until required.

2.7.3 siRNA labelling

The dry fluorescein (FAM) labelling reagent was resuspended with 100 μ l reconstitution solution, vortexed and left at room temperature for 5 minutes before vortexing again to ensure proper resuspension. The following mixture for was used for the labelling reaction (Table 16);

RNase-free water	18.3 μ l
10x labelling buffer	5 μ l
siRNA duplex (5 μ g)	19.2 μ l
FAM labelling reagent	7.5 μ l

Table 16. The reaction mixture used to label the annealed siRNA oligonucleotides with the FAM dye.

The reagents were added in this order and mixed by vortexing whilst limiting the amount of time the reaction mixture was exposed to light. The reaction mix was incubated at 37°C for 1 hour in the dark. The labelled siRNA was retrieved by ethanol precipitation. For a 50µl reaction mix, 5µl of 5M NaCl and 125µl cold 100% ethanol was added and mixed well. The mixture was then placed at -20°C for 60 minutes. The labelled siRNA was pelleted by centrifugation at 8,000 x g for 20 minutes. The supernatant was aspirated off and 175µl of 70% ethanol was added to the pellet. The mix was centrifuged at 8,000 x g for 5 minutes and the supernatant was aspirated off. The siRNA was dried for 10 minutes at room temperature and resuspended in 19.2µl RNase-free water.

2.7.4 Transfection of siRNA

When warmed to room temperature, 10µl of siPORT NeoFX transfection agent was added to 90µl Opti-MEM (Gibco, Invitrogen Ltd) per reaction and incubated at room temperature for 10 minutes. Labelled siRNA was added to the transfection mixtures to achieve a final siRNA concentration of 50nM and 100nM and mixed by pipetting. The siRNA/transfection mixes were incubated at room temperature for 10 minutes. Cells set to the required density were added to the reaction mix and gently mixed by swirling. Following 24 hours, 3mls of fresh medium was added to the cells to reduce toxicity and after 48 hours following transfection the effect of siRNA transfection was determined by RQ-PCR (Section 2.5). A transfection reaction was carried out for

each gene at 50nM and 100nM, and for a scrambled siRNA at 50nM and 100nM and one without siRNA as negative controls.

2.7.5 Determination of siRNA transfection efficiency

The efficiency level of siRNA transfection into the cells was determined by flow cytometry on day 2 following transfection on a FACScan flow cytometer (Becton Dickinson, NJ, USA) using CellQuest software.

2.8 Immunophenotyping

To assess the expression of cell surface markers in the treated cell lines, fluorescently labelled antibodies were bound to the cell surface markers of interest. Cells were pelleted by centrifugation at 120 x g for 3 minutes. The pellet was washed 3 times with 1x PBS, 1% BSA and then resuspended in 10µl 1x PBS, 1% BSA. The resuspended cells were incubated in the dark at 4°C for 1 hour with 2.5µl of the labelled antibody. 50µl of FACSFlow was added to stop the binding reaction. The cells were pelleted by centrifugation and resuspended in 50µl FACSFlow before being transferred to a Luckham LP2 tube. The cells were analysed on a FACS Calibur flow cytometer (Becton Dickinson, NJ, USA) using CellQuest software.

2.8.1 Markers for NB4 cell lines

The antibodies used to assess the effect of NK4 on NB4 cells were mouse anti-human CD11b and mouse anti-human CD38. IgG was used

as a control to show the level of background or irrelevant binding of the antibody.

2.8.2 Flow analysis

A FACS Calibur was used to collect the flow cytometry data and the data was analysed using Summit 4.0 (DAKO, Ely, UK). The data was gated to exclude debris and background staining was determined using the data generated for the cells stained with the IgG control antibody. The level of antibody staining was expressed as the mean fluorescence intensity.

Chapter 3

Using gene expression profiling to identify novel minimal residual disease markers specific for t(15;17) acute myeloid leukaemia - a “proof of principle”

3.1 Introduction

3.1.1 Acute Promyelocytic Leukaemia (APL)

A reciprocal balanced translocation between chromosomes 15 and 17, resulting in the fusion of the promyelocytic (*PML*) gene on chromosome 15 and the retinoic acid receptor- α (*RAR α*) gene on chromosome 17, is the genetic hallmark of acute promyelocytic leukaemia (APL). The resulting fusion genes, *PML-RAR α* and *RAR α -PML*, are detected in more than 95% of APL patients (Lo-Coco & Ammatuna, 2006). The *PML* protein, encoded by the *PML* gene, co-localises with other proteins in discrete nuclear speckles to form complexes called nuclear *PML* bodies or PODs (*PML* oncogenic domains) (Jensen *et al*, 2001). The *PML* protein contains a **RING**-finger domain, two **B**-box domains and an α -helical coiled-coil domain, which form the RBCC domain; responsible for the homodimerisation, localisation and *PML*-protein interactions of *PML* (Zhong *et al*, 2000). Accumulating evidence suggests an important role for *PML* in the regulation of transcription; the conformation of the protein and its interactions with other proteins controlling the direction of its transcriptional regulation (Ahn *et al*, 1998; Zhong *et al*, 2000). The *RAR α* gene on chromosome 17 encodes for the *RAR α* protein; a member of a nuclear hormone receptor superfamily. *RAR α* acts as a ligand-inducible transcription factor by binding to specific response elements (RARE) of target genes. In the absence of ligand, *RAR α* acts as a transcriptional repressor by forming a heterodimeric complex with retinoid X receptor (*RXR*); the heterodimer recruits a repression complex consisting of histone deacetylases (HDACs) and nuclear

receptor co-repressors. In the presence of retinoic acid (RA), the repression complex is dissociated in favour of the recruitment of a co-activator complex consisting of histone acetylases (HAT) leading to the activation of gene transcription. The fusion protein PML-RAR α can bind to retinoic acid response elements (RAREs) as a homodimer unlike the wild-type RAR α (Perez *et al*, 1993). The fusion protein is also able to heterodimerise to retinoid receptors (RXR) with an increased affinity for the HDACs and nuclear receptor co-repressors compared to wild-type RAR α , resulting in an aberrant retinoid receptor leading to constitutive transcriptional repression and, thus, an increase in cell growth and a block in cell differentiation (Perez *et al*, 1993;Melnick & Licht, 1999;Pandolfi, 2001). The transcription repression resulting from the PML-RAR α protein can be released with pharmacological levels of retinoic acid (10^{-6} M) (Grignani *et al*, 1998;Mozziconacci *et al*, 1998).

3.1.2 Minimal Residual Disease and APL

Reverse-transcription polymerase chain reaction (RT-PCR) is an established method for the rapid and sensitive identification of the fusion gene product resulting from the rearrangement of the *PML* gene and the *RAR α* gene in APL cells. Three different breakpoints in the *PML* gene have been recorded, resulting in the generation of three possible *PML-RAR α* transcripts (Pandolfi *et al*, 1992). RT-PCR for the amplification of the *PML-RAR α* transcript has also become a valuable method for MRD detection in APL patients; RT-PCR positivity has been used to predict haematological relapse following consolidation therapy

whilst persistent RT-PCR negativity has been associated with a lower relapse rate (RR) and long-term survival (Chang *et al*, 1992;Huang *et al*, 1993;Miller, Jr. *et al*, 1992;Diverio *et al*, 1998;Jurcic *et al*, 2001). Unfortunately a number of studies have shown that despite achieving PCR negativity at the end of consolidation therapy, a number of patients go on to relapse suggesting more time-points and a longer follow-up are required for the monitoring of MRD. Furthermore, a method with greater sensitivity than RT-PCR (1 in 10^4) is required to correctly predict relapse or long-term survival (Devaraj *et al*, 1996;Mandelli *et al*, 1997;Burnett *et al*, 1999;Gameiro *et al*, 2001). An alternative to RT-PCR is real-time RT-PCR; the level of PCR product is quantified as it accumulates (Liu Yin, 2002). Subsequent studies have considered the use of real-time quantitative PCR (RQ-PCR) as a more sensitive method of MRD monitoring. RQ-PCR can achieve sensitivities of 1 leukaemic cell in 10^6 cells and serial MRD monitoring on a 3-monthly basis can prove informative in identifying APL patients at a high risk of relapse (Tobal *et al*, 2001;Gallagher *et al*, 2003;Lee *et al*, 2006).

The aim of this chapter was to use t(15;17) AML as a “proof of principle” to determine whether gene expression profiling can identify new MRD markers that are specific for leukaemic sub-groups. Genes specific to the t(15;17) sub-group were identified by gene expression profiling in order to identify potential candidate markers of minimal disease. Their levels of expression were tracked over time in t(15;17) AML patient samples using RQ-PCR and correlated with corresponding transcription

data for markers currently used for monitoring MRD in these patients (*PML-RAR α* and *RAR α -PML*).

3.2 Methods

3.2.1 Patient and donor samples

The sample dataset used for analysis in this section contained gene expression data from 295 patient samples (including 5 patient replicate samples) and 10 healthy normal donor samples processed by the School of Medicine, Cardiff University as described in section 2.2.1 (Chapter 2). The patient cohort used contained 26 AML patients presenting with a t(15;17) including two t(15;17) patient replicate samples (Chapter 2, Table 3).

The t(15;17) AML patient and normal healthy donor RNA was processed from bone marrow aspirates and supplied from within the Department of Haematology, School of Medicine, Cardiff University.

3.2.2 Gene Expression Data Analysis

3.2.2a Expression summarisation

Image files (CEL files) were analysed using the statistical algorithms in MAS 5.0 using target intensity (TGT) of 100. The resultant data matrix of expression levels for each probe set was taken forward for further analysis (see accompanying CD for CEL files).

3.2.2b Quality Control of probe Sets

Before starting the gene expression analysis process, probes representing genes that were not expressed in any of the samples were removed. Retaining such probe sets could affect the analysis by generating false-positives in a statistical analysis or grouping dissimilar expression profiles in a cluster analysis. A method of “filtering on flags” was used to remove such probe sets. Probe sets were assigned a “flag” of either present, marginal or absent based on the reliability of the intensity of certain features, for example level of intensity compared to the level of saturation or whether the feature is a population outlier (Agilent GeneSpring GX9 Tutorial). Further, more stringent, quality controls were applied by identifying probes sets flagged as present in at least 2 samples. From this set, probe sets that did not increase or decrease in expression level by more than 2 fold, between 0.5 and 2, (an accepted level of change) across the 305 samples were classed as non-changing and were removed to leave a list of probe sets called “present and changing”.

3.2.2c Fold Change Calculation

A fold change calculation is a simple way of identifying differentially expressed genes between two states. Using mean expression values, a gene is said to be differentially expressed if its mean expression value varies by more than a set factor between the control state and the test state. To identify genes differentially expressed between the t(15;17) AML patients, the non-t(15;17) AML patients and the healthy normal

donors, the probe sets were filtered on a change of 2-, 3-, 4- and 5-fold. The fold change filtering process was applied to both the list of probe sets classified as “present in at least 1 sample” and the set classified as “present and changing”.

3.2.2d Significance Analysis of Microarrays (SAM)

Significance Analysis of Microarrays (SAM) was used as another method for analysing the microarray data. SAM identifies changes in gene expression that are statistically significant by using gene-specific *t*-tests. A false discovery rate (FDR) threshold is applied to the data to control the expected proportion of false-positives identified. Adjusting the threshold alters the number of gene identified (Tusher *et al*, 2001). An FDR of 0.05 was used as a 5% probability of identifying false-positives is an established significance level; an FDR of 0.01 was used for more stringent testing.

3.2.2e Analysis of Variance (ANOVA)

Analysis of Variance (ANOVA) is a statistical test used to assess significant differences between the means of datasets. The variances of the datasets of interest are compared in order to test for the differences between the means. ANOVA was used to analyse the microarray data in this study by identifying genes with a statistically significant difference in their expression between the t(15;17) AML patients and the normal healthy donors.

3.2.3 Diagnostic Screening of Candidate Genes

The genes identified from the gene expression profiling analysis were screened in a group of diagnostic t(15;17) AML patient bone marrow samples (n=9) and normal healthy donor bone marrow samples (n=16). Real-time quantitative PCR (RQ-PCR) was used as described in section 2.5 (Chapter 2) to measure the expression levels of the candidate gene transcripts. Each sample was run in duplicate and a duplicate “no-RNA” cDNA control sample and “no-cDNA” control sample was included as negative-controls. The expression data generated from the real-time PCR was normalised to a house-keeping gene, *S14* (gene encoding for the 40S ribosomal subunit protein S14) (Rhoads *et al*, 1986), as recommended by Roche for relative gene expression studies on their LightCycler® Real-Time PCR System and optimised and published from in the Department to allow for comparisons with other on-going projects within the laboratory (Guinn *et al*, 2005;Walsby *et al*, 2008).

3.2.4 Validation of Candidate Genes as MRD Markers

The candidate genes screened in diagnostic patient samples were validated as markers of MRD using RQ-PCR as previously described. Duplicate samples and negative controls were used and all gene expressions were normalised to the transcript expression of *S14*. Diagnostic (DX) and follow-up (FUP) patient bone marrow samples were used for quantitative PCR analysis of the genes of interest. A normalised fold change in expression value was calculated for every

gene expression value for each patient sample using the following equation;

$$\text{Fold change} = 2^{(\text{Control gene Ct value} - \text{Test gene Ct value})}$$

A minimal residual disease value (MRDv) was calculated for every gene expression value and compared to MRDv data for the markers currently used. The MRDv was calculated using the $\Delta\Delta\text{Ct}$ method as described by Beillard *et al* (2003) in the “Europe Against Cancer” programme based on the assumption that the intercept and slope values between the control gene and the gene of interest are identical, using the following calculation

$$\text{MRD value} = 10^{((\Delta\text{Ct}(\text{FUP}) - \Delta\text{Ct}(\text{DX})) / -3.4)}$$

Where $\Delta\text{Ct}(\text{FUP})$ is the normalised crossing point for the gene of interest in a patient follow-up sample, $\Delta\text{Ct}(\text{DX})$ is the normalised crossing point for the gene of interest in the patient diagnostic sample and -3.4 is the average slope value from plasmid standard curves (Beillard *et al*, 2003; Gabert *et al*, 2003). Data for the markers currently used, *PML-RAR α* and *RAR α -PML*, was supplied by Professor D. Grimwade, Guy’s Hospital, London.

3.2.5 Statistical Analysis of Diagnostic Screening and MRD Analyses

The gene expression data generated from the diagnostic patient and normal healthy donor samples was analysed using an unpaired student *t*-test. All MRD data was analysed using a two-tailed Pearson's correlation. A Pearson's correlation was used to visualise the degree of linear relationship between the MRD data generated for the candidate genes and for markers currently used.

3.3 Results

3.3.1 Identification of genes expressed specifically in t(15;17) AML samples

Using filtering tools based on present/absent calls and fold change, 22,283 probe sets in 295 AML diagnostic patient samples (representing 290 patients) and 10 donor samples were analysed to identify genes specific to those patients with a t(15;17) translocation.

3.3.1a Quality Control of Probe sets

From 22,283 probe sets, 17,844 probe sets (probe set list 3.1) were flagged as "present", i.e. assigned a call of present, in at least one sample of the patient and donor cohort. Further filtering through flag assignment with an increased stringency identified 16,988 probes sets (probe set list 3.5) assigned a call of present in at least two samples in the cohort. From this group of probe sets, eleven probe sets did not

increase or decrease in expression by at least 2 fold (probe set list 3.6) and so were removed for being “non-changing”. This left a group of 16,977 probe sets classified as “present and changing” (probe set list 3.7).

Gene expression profiling was performed on both the group of “present in at least 1 sample” probe sets and the group of “present and changing” probe sets to compare the robustness and reliability of the analysis.

3.3.1b Identification of Differentially Expressed Genes by Fold change

The “present in at least 1 sample” probe sets were filtered on a 2-, 3-, 4- and 5-fold change in expression level between the t(15;17) patients and non-t(15;17) patients and between the t(15;17) patients and the normal healthy donor patients (Table 1). By using a ≥ 3 fold change as the threshold for significance, the group of 27 probe sets identified as having a greater expression in the t(15;17) patients than in the non-t(15;17) patients (probe set list 3.2) was combined with the group of 204 probe sets identified as having a greater expression in the t(15;17) patients than in the normal healthy donors (probe set list 3.3) to generate a group of 18 probe sets with a greater level of expression in t(15;17) AML patients when compared with other AML patient subgroups or normal healthy donors (Probe set list 3.3) (Figure 1; Table 2).

The “present and changing” probe sets were filtered on a 2-, 3-, 4- and 5-fold change in expression level between the t(15;17) patients and non-t(15;17) patients and between the t(15;17) patients and the normal healthy donor patients (Table 1). By using a ≥ 3 fold change as the threshold for significance, the group of 26 probe sets identified as having a greater expression in the t(15;17) patients than the non-t(15;17) patients (probe set list 3.8) was combined with the group of 202 probe sets having a greater expression in the t(15;17) patients than the normal healthy donors (probe set list 3.9) to generate a group of 17 probe sets with a greater level of expression in t(15;17) AML patients than in any other AML patient sub-group or the normal healthy donor group (Probe set list 3.10) (Figure 2; Table 2).

Probe set group	Fold Change	t(15;17) > non-t(15;17)	t(15;17) > Donors
Present in at least 1 sample	2	140	873
	3	27	204
	4	6	43
	5	5	26
Present and changing	2	139	833
	3	26	202
	4	6	43
	5	5	26

Table 1. The number of probe sets identified as having a greater expression in t(15;17) AML patients than in non-t(15;17) patients or normal healthy donors. Fold change filtering was applied to both probe set lists generated for quality control purposes as described in section 3.3.1a.

The group of 18 “present in at least 1 sample” probe sets and the group of 17 “present and changing” probe sets contained identical probe sets

bar one; probe set 204879_at relating to the podoplanin (*PDPN*) gene was not identified during the filtering of probe sets classified as “present and changing”. The 17 probe sets represented 12 genes. Furthermore, the 7 probe sets showing the highest fold change in expression in both groups of probe sets represented 3 genes; hepatocyte growth factor (*HGF*), insulin-like growth factor binding protein 2 (*IGFBP2*) and prostaglandin D2 synthase (*PGD2*) (Table 2).

Performing the filtering analysis on both probe set groups allowed for the reliability and robustness of the filtering process to be confirmed. The identification of the same 17 probe sets specific to both groups suggests filtering on fold change is a robust and reliable method of analysis and that probe sets identified really were specific to t(15;17) AML and were not false-positives. The identification of the 18th probe set, 204879_at, may be a false-positive due to the reduced stringency of the filtering on flags method originally applied to the probe set (present in only 1 sample rather than present in 2 and samples changing and changing in expression), or whilst it may be more specific to t(15;17) AML rather than non-t(15;17) AML it is not specific enough to be detected with more stringent analysis and therefore would not make a suitable MRD marker.

	Probe Set	Gene Symbol	Gene Title	Fold Change
1	210794_s_at	<i>MEG3</i>	Maternally expressed 3 (non-protein coding)	3.002
2	208581_x_at	<i>MT1X</i>	Metallothionein 1X	3.149
3	209686_at	<i>S100B</i>	S100 calcium binding protein B	3.184
4	203074_at	<i>ANXA8</i>	Annexin A8	3.387
5	221898_at*	<i>PDPN</i>	Podoplanin	3.689
6	205047_s_at	<i>ASNS</i>	Asparagine synthetase	3.706
7	*204879_at	<i>PDPN</i>	Podoplanin	3.739
8	219837_s_at	<i>CYTL1</i>	Cytokine-like 1	3.838
9	207076_s_at	<i>ASS1</i>	Argininosuccinate synthetase 1	3.874
10	211663_x_at	<i>PTGDS</i>	Prostaglandin D2 synthase 21kDa (brain)	3.961
11	203068_at	<i>KLHL21</i>	Kelch-like 21 (Drosophila)	3.970
12	209960_at	<i>HGF</i>	Hepatocyte growth factor	4.796
13	202718_at	<i>IGFBP2</i>	Insulin-like growth factor binding protein 2	4.800
14	210755_at	<i>HGF</i>	Hepatocyte growth factor	5.263
15	212187_x_at	<i>PTGDS</i>	Prostaglandin D2 synthase 21kDa (brain)	6.344
16	211748_x_at	<i>PTGDS</i>	Prostaglandin D2 synthase 21kDa (brain)	6.950
17	210998_s_at	<i>HGF</i>	Hepatocyte growth factor	7.367
18	210997_at	<i>HGF</i>	Hepatocyte growth factor	8.958

Table 2. The 18 “present in at least 1 sample” probe sets and the 17 “present and changing” probe sets identified from the 305 sample dataset as upregulated in t(15;17) AML patients samples when compared to non-t(15-17) patients and to healthy donor samples by fold change (FC>3). The fold change value displayed is the mean fold change value of the probe sets across the t(15;17) samples. The group of 17 “present and changing” probe sets were identical to the group of 18 “present in at least 1 sample” probe sets bar probe set 204879_at relating to the podoplanin gene*.

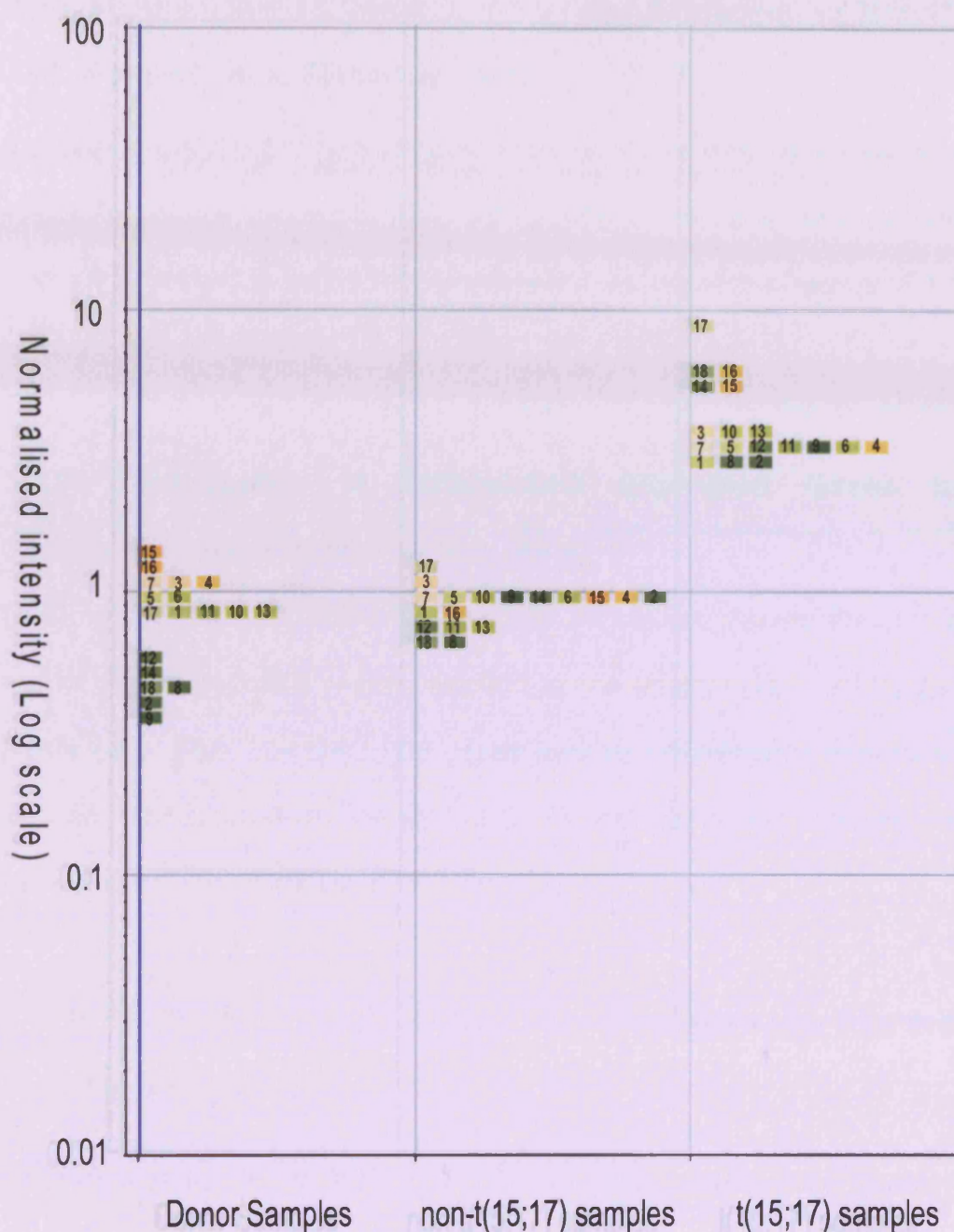


Figure 1. The relative mean expression levels of the 18 probe sets identified, from probe sets that were classified as present in at least 1 sample, as being more significantly expressed in t(15;17) samples than non-t(15;17) samples **and** in donor samples by ≥ 3 fold. The probe sets are coloured according to their expression levels in their respective sub-group relative to their expression level in the donor samples. The numbers on each block correspond to the probe set it represents - see Table 2.

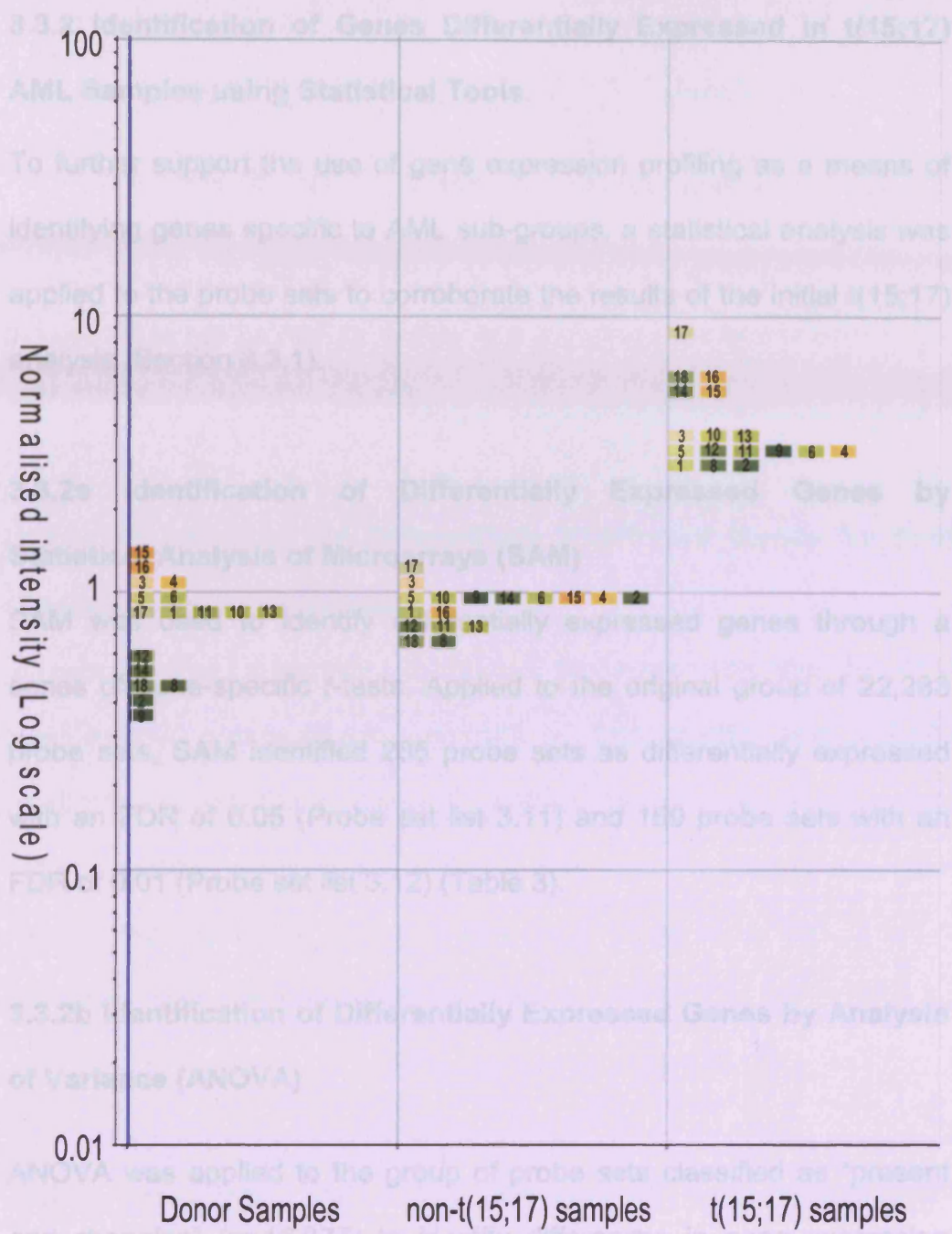


Figure 2. The relative mean expression levels of the 17 probe sets identified, from probe sets that were classified as present in at least 2 samples and changing in expression, as being more significantly expressed in t(15;17) samples than non-t(15;17) samples **and** in donor samples by ≥ 3 fold. The probe sets are coloured according to their expression levels in their respective sub-group relative to their expression level in the donor samples. The numbers on each block correspond to the probe set it represents - see Table 2.

3.3.2 Identification of Genes Differentially Expressed in t(15;17) AML Samples using Statistical Tools.

To further support the use of gene expression profiling as a means of identifying genes specific to AML sub-groups, a statistical analysis was applied to the probe sets to corroborate the results of the initial t(15;17) analysis (Section 3.3.1).

3.3.2a Identification of Differentially Expressed Genes by Statistical Analysis of Microarrays (SAM)

SAM was used to identify differentially expressed genes through a series of gene-specific *t*-tests. Applied to the original group of 22,283 probe sets, SAM identified 285 probe sets as differentially expressed with an FDR of 0.05 (Probe set list 3.11) and 159 probe sets with an FDR of 0.01 (Probe set list 3.12) (Table 3).

3.3.2b Identification of Differentially Expressed Genes by Analysis of Variance (ANOVA)

ANOVA was applied to the group of probe sets classified as “present and changing” (n=16,977) to identify differences in gene expression levels between the t(15;17) patients and the non-t(15;17) patients and normal healthy donors by comparing the variances of the test groups. Using an FDR of 0.05 7,680 probe sets (probe set list 3.13) were identified, while 5,501 probe sets (probe set 3.14) were identified with an FDR of 0.01 (Table 3).

FDR	SAM	ANOVA
0.05	285	7,680
0.01	159	5,501

Table 3. The number of probe sets identified as having a greater expression in t(15;17) AML patients than in non-t(15;17) patients or normal healthy donors through SAM and ANOVA analysis.

3.3.2c Identification of Differentially Expressed Genes by Fold Change and Statistical Tools

Combining the differentially expressed probe sets identified by SAM and ANOVA analysis with an FDR of 0.05 (n=285 and n=7,680 respectively) with the probe sets identified as “present and changing” with a greater expression in t(15;17) AML patients than in any other AML sub-group or the normal healthy donor group by ≥ 3 fold, a set of 17 probe sets was generated; this set was identical to the original list identified in section 3.3.1b (Figure 3; Figure 2; Table 2).

Combining statistical and filtering tools generates a far more stringent method of identification of probe sets specific to t(15;17) by increasing the thresholds the probe sets are required to achieve for selection. The combination of statistical tools and fold change filtering identifying the same probe sets as unique to t(15;17) AML as filtering alone indicates that, for t(15;17) AML, the initial analysis of filtering alone was robust

and reliable enough to identify potential markers of MRD for this subgroup and that the probe sets identified are not false-positives.

Identified in the gene expression analysis, HGF was selected for further investigation to assess its potential as an MRD marker in AML.

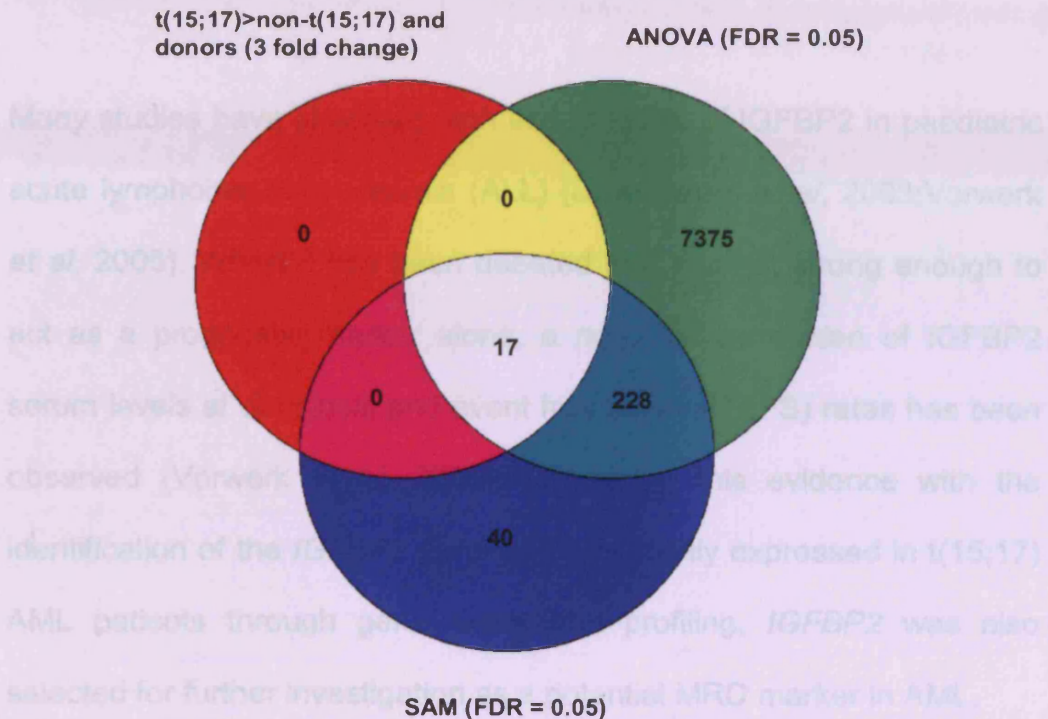


Figure 3. A combination of the probe sets identified from SAM (FDR=0.05), ANOVA testing (FDR=0.05) and having a greater level of expression in t(15;17) samples than in non-t(15;17) samples and in donor samples by ≥ 3 fold change filtering. Seventeen probe sets were identified as being specifically expressed in t(15;17) samples.

3.3.3 Validation of Identified Genes by Real-Time Quantitative PCR

Elevated levels of the HGF protein have been previously noted in leukaemic cell line populations and AML patient samples suggesting a role for HGF in AML (Nakamura *et al*, 1994; Hino *et al*, 1996). Weimar *et al* (1998) observed that HGF-induced colony formation and proliferation in leukaemic cells; an effect not observed in normal haematopoietic

cells. Given the potential importance of the HGF protein in AML and the high expression levels of the probes sets representing the *HGF* gene identified in the gene expression analysis, *HGF* was selected for further investigation to assess its potential as an MRD marker in AML.

Many studies have observed high serum levels of IGFBP2 in paediatric acute lymphoblastic leukaemia (ALL) (Dawczynski *et al*, 2003; Vorwerk *et al*, 2005). Whilst it has been debated that it is not strong enough to act as a prognostic marker alone, a negative correlation of IGFBP2 serum levels at diagnosis and event free survival (EFS) rates has been observed (Vorwerk *et al*, 2005). Combining this evidence with the identification of the *IGFBP2* gene as being highly expressed in t(15;17) AML patients through gene expression profiling, *IGFBP2* was also selected for further investigation as a potential MRD marker in AML.

3.3.3a Diagnostic Screening of Candidate Genes

Transcript expression levels of *HGF* and *IGFBP2* were measured in diagnostic t(15;17) AML patient samples and compared to levels measured in healthy normal donor samples. Initially primer sets were designed for the amplification of *HGF* and *IGFBP2* as described in the Materials and Methods section 2.5.1 (Chapter 2) and using serial dilutions of cDNA, their amplification efficiency was compared to that of the house-keeping gene, *S14*. Both *HGF* and *IGFBP2* produced similar amplification efficiencies to that produced by *S14*. Therefore *S14* was considered suitable as the normalising control gene (Figure 4).

The levels of expression of *HGF* and *IGFBP2* were measured in 6 and 9 diagnostic AML patient samples (respectively) presenting with a t(15;17) and were compared to the levels of expression in 18 normal healthy bone marrow donor samples (Figure 5). A mean normalised Ct value of *HGF* of 37.47 (range 35.19 to 39.75) was detected at a level of *IGFBP2* of 1.247 (range 0.834 to 1.659) was detected in the healthy normal donor samples. A mean normalised Ct value of *IGFBP2* of -5.19 (range -5.57 to -4.81) was detected in the healthy normal donor samples (Figure 5). Using *t*-test, a strong significant difference was observed between the mean normalised Ct values of the candidate genes in the diagnostic t(15;17) AML patient samples and the mean normalised Ct values of the candidate genes in the healthy normal donor samples (Table 4).

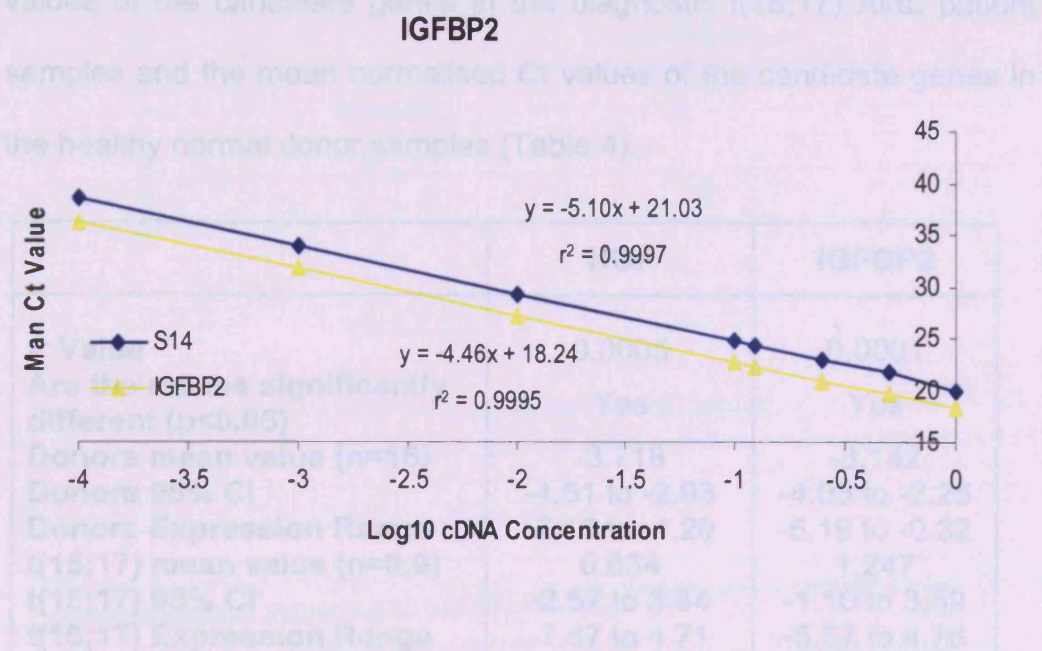
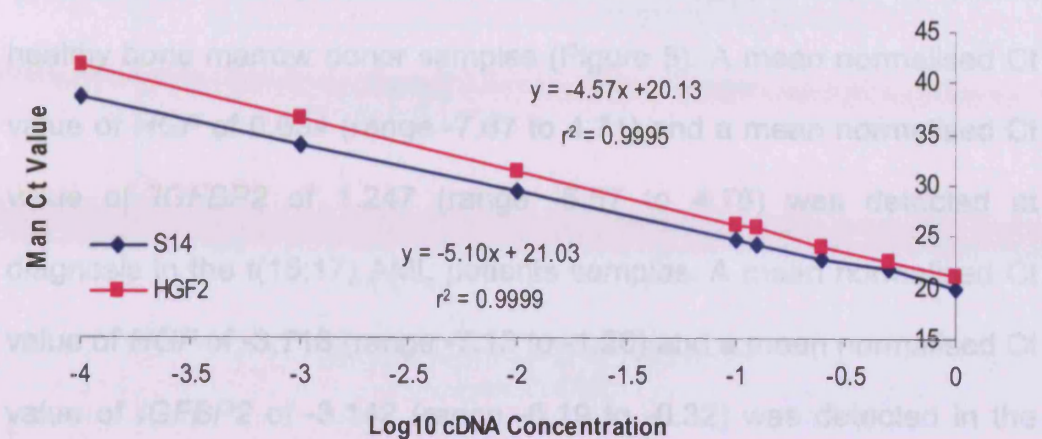


Figure 4. The amplification curves of the primer sets used for *HGF* and *IGFBP2* and of the house-keeping gene, *S14*, to compare their amplification efficiencies.

The levels of expression of *HGF* and *IGFBP2* were measured in 8 and 9 diagnostic AML patient samples (respectively) presenting with a t(15;17) and were compared to the levels of expression in 16 normal healthy bone marrow donor samples (Figure 5). A mean normalised Ct value of *HGF* of 0.634 (range -7.67 to 4.71) and a mean normalised Ct value of *IGFBP2* of 1.247 (range -5.57 to 4.76) was detected at diagnosis in the t(15;17) AML patients samples. A mean normalised Ct value of *HGF* of -3.718 (range -7.13 to -1.20) and a mean normalised Ct value of *IGFBP2* of -3.142 (range -6.19 to -0.32) was detected in the healthy normal donor samples (Figure 5). Using *t*-tests, a strong significant difference was observed between the mean normalised Ct values of the candidate genes in the diagnostic t(15;17) AML patient samples and the mean normalised Ct values of the candidate genes in the healthy normal donor samples (Table 4).

	HGF	IGFBP2
P Value	0.0005	0.0001
Are the means significantly different (p<0.05)	Yes	Yes
Donors mean value (n=16)	-3.718	-3.142
Donors 95% CI	-4.51 to -2.93	-4.03 to -2.25
Donors Expression Range	-7.13 to -1.20	-6.19 to -0.32
t(15;17) mean value (n=8,9)	0.634	1.247
t(15;17) 95% CI	-2.57 to 3.84	-1.10 to 3.59
t(15;17) Expression Range	-7.67 to 4.71	-5.57 to 4.76

Table 4. The statistical significance of the mean Ct values of the candidate genes measured in the t(15;17) AML diagnostic patient samples and the healthy normal donor samples normalised to the Ct values of *S14*. CI = confidence interval.

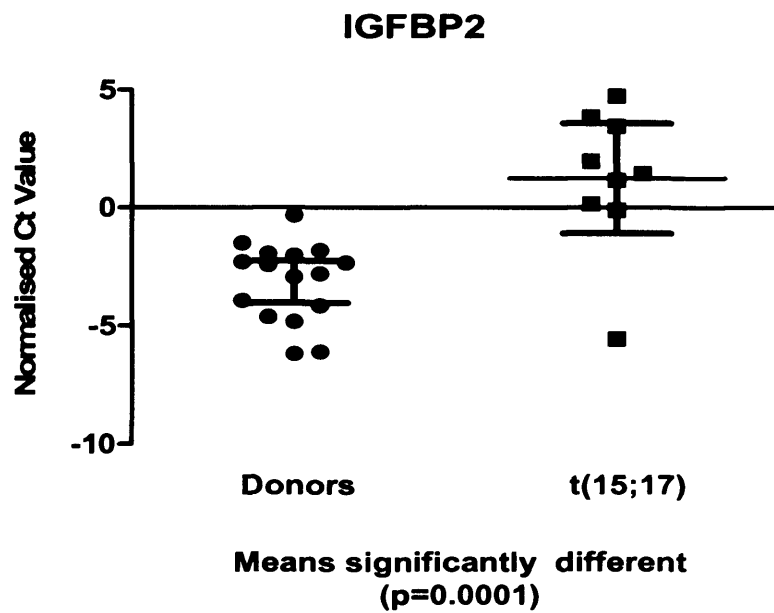
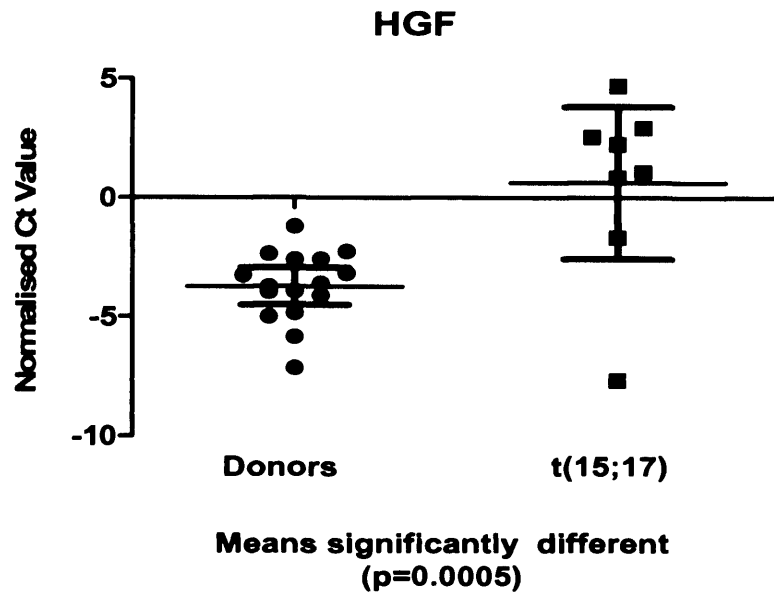


Figure 5. Scatter plots of the spread of the normalised Ct values of the candidate genes measured in the t(15;17) AML diagnostic patient samples and the healthy normal donor samples. Mean value $\pm$ 95% confidence interval plotted.

As both *HGF* and *IGFBP2* were more highly expressed in diagnostic samples from t(15;17) AML patients compared to normal healthy donors, both genes were followed through the progression of the disease of t(15;17) AML patients to validate their potential as markers of MRD

3.3.3b Hepatocyte Growth Factor (*HGF*) as an MRD Marker

The transcript levels and MRD values for the *HGF* gene were followed in the diagnostic and follow-up samples of three t(15;17) patients as described in section 2.5 and 3.2. The level of MRD in the first follow-up sample for the first two t(15;17) patients was two logs lower than the level of transcription observed at diagnosis (Figure 7a, 7b), and nearly three logs lower in the third t(15;17) patient (Figure 7c). The level of *HGF* MRD remained constant at these levels for the remainder of the follow-up period. Over the course of their disease, the levels of *HGF* MRD in the three t(15;17) patients displayed a similar pattern to the corresponding MRD values of the t(15;17) fusion gene products, *PML-RAR α* and *RAR α -PML*. A strong positive correlation was seen between the fusion gene product expression data of *PML-RAR α* and *RAR α -PML* and the expression data for *HGF* ($r^2 = 0.9947$, $p < 0.0001$ and $r^2 = 0.9982$, $p < 0.0001$ respectively).

The expression of *HGF* detected in the three t(15;17) patients was above the upper threshold of expression detected in the healthy normal donor samples at diagnosis by at least 1 log. Following the start of

induction therapy the expression of the candidate gene decreased below the upper threshold detected in the healthy normal donors by at least 1 log and remained so for the duration of the patients' disease (Figure 6).

3.3.3c Insulin-like Growth Factor Binding Protein 2 (*IGFBP2*) as an MRD Marker

The transcript levels and MRD values of *IGFBP2* were followed in diagnostic and follow-up samples from three t(15;17) patients. The level of *IGFBP2* MRD in the first follow-up sample of the first two t(15;17) patients was two logs lower than the normalised level of transcripts observed at diagnosis (Figure 7a, 7b), and nearly three logs lower in the third t(15;17) patient (Figure 7c). The level of *IGFBP2* MRD remained constant at these levels for the remainder of the follow-up period. The *IGFBP2* transcript levels and MRD values of the 3 t(15;17) patients through the course of the disease displayed a similar pattern to the transcript levels and MRD values of the fusion gene products, *PML-RAR α* and *RAR α -PML*. A strong positive correlation was seen between the fusion gene product expression data of *PML-RAR α* and *RAR α -PML* and the expression data for *IGFBP2* ($r^2 = 0.9993$, $p < 0.0001$ and $r^2 = 0.9990$, $p < 0.0001$ respectively).

The expression of *IGFBP2* detected in the three t(15;17) patients was above the upper threshold of expression detected in the healthy normal donor samples at diagnosis by at least 1 log. Following the start of

induction therapy the expression of the candidate gene decreased below the upper threshold detected in the healthy normal donors by at least 1 log and remained so for the duration of the patients' disease (Figure 6).

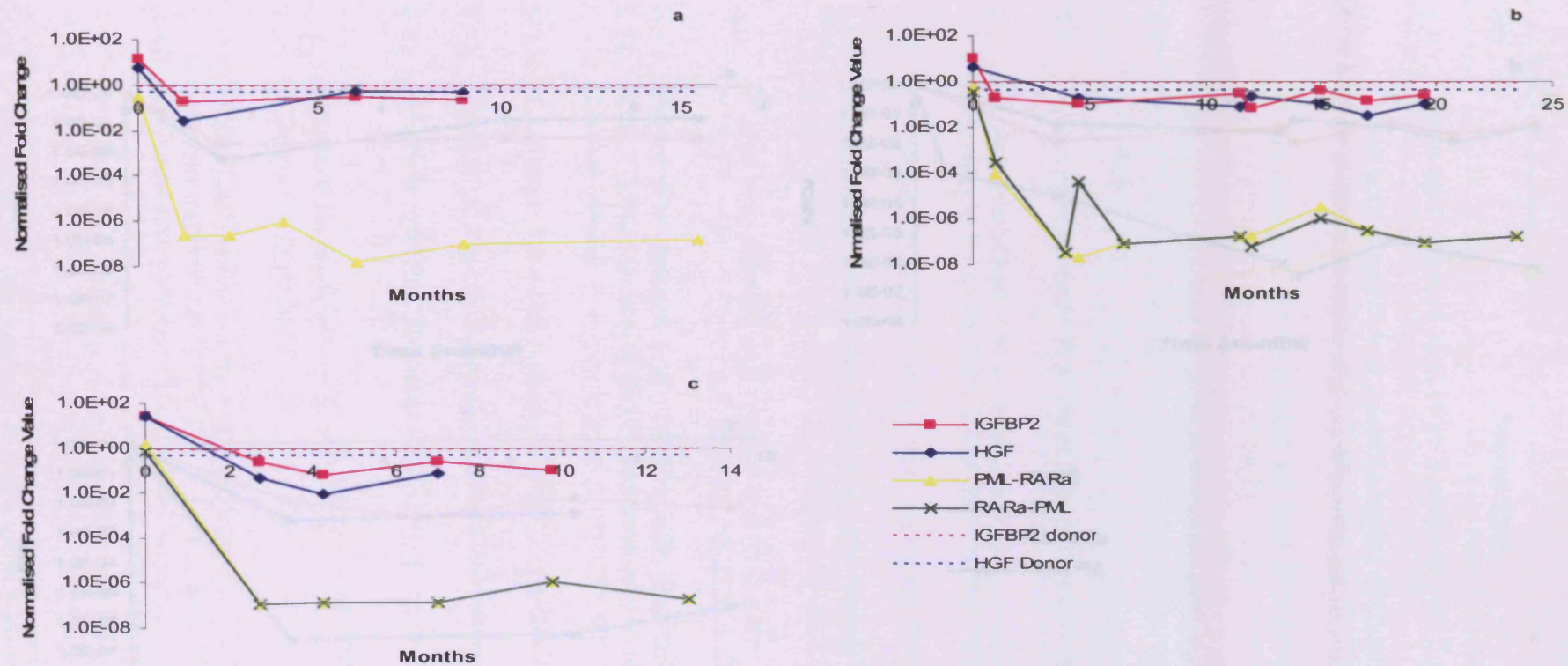


Figure 6. The fold changes in expression of *HGF*, *IGFBP2*, *PML-RARα* and *RARα-PML* in the diagnostic and follow-up samples of 3 t(15;17) patients (a-c), calculated as described in section 3.2. The dashed lines indicate the upper level of fold changes in expression of *HGF* and *IGFBP2* observed in the healthy normal donor samples.

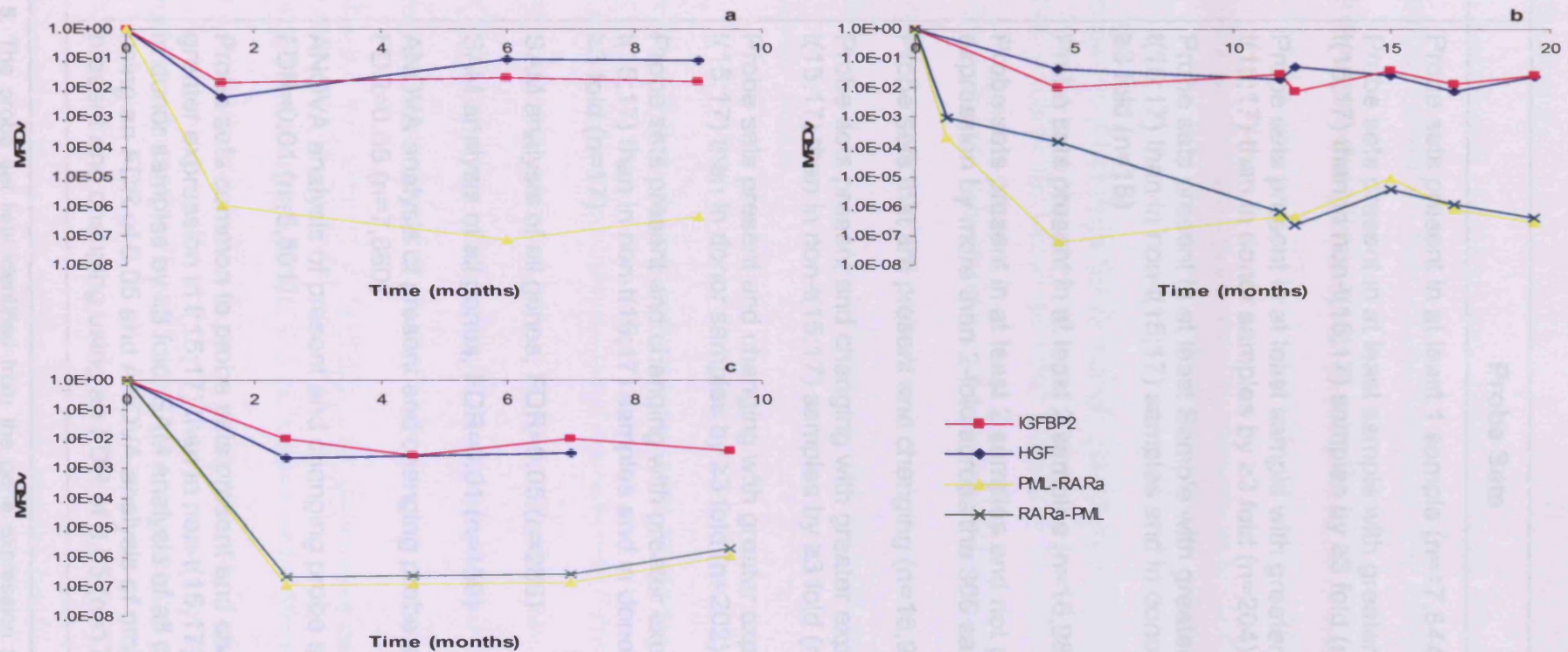


Figure 7. The MRD values of HGF, IGFBP2, PML-RAR α and RAR α -PML in the diagnostic and follow-up samples from 3 t(15;17) patients (a-c), calculated as described in section 3.2. A strong positive correlation was seen between the fusion gene product expression data of PML-RAR α and RAR α -PML and the expression data of HGF and of IGFBP2.

List	Probe Sets
3.1	Probe sets present in at least 1 sample (n=17,844)
3.2	Probe sets present in at least sample with greater expression in t(15;17) than in non-t(15;17) samples by ≥ 3 fold (n=27)
3.3	Probe sets present in at least sample with greater expression in t(15;17) than in donor samples by ≥ 3 fold (n=204)
3.4	Probe sets present in at least Sample with greater expression in t(15;17) than in non-t(15;17) samples and in donor samples by ≥ 3 fold (n=18)
3.5	Probe sets present in at least 2 samples (n=16,988)
3.6	Probe sets present in at least 2 samples and not changing in expression by more than 2-fold across the 305 samples (n=11)
3.7	Probe sets that are present and changing (n=16,977)
3.8	Probe sets present and changing with greater expression in t(15;17) than in non-t(15;17) samples by ≥ 3 fold (n=26)
3.9	Probe sets present and changing with greater expression in t(15;17) than in donor samples by ≥ 3 fold (n=202)
3.10	Probe sets present and changing with greater expression in t(15;17) than in non-t(15;17) samples and in donor samples by ≥ 3 fold (n=17)
3.11	SAM analysis of all genes, FDR=0.05 (n=285)
3.12	SAM analysis of all genes, FDR=0.01 (n=159)
3.13	ANOVA analysis of present and changing probe sets, FDR=0.05 (n=7,680)
3.14	ANOVA analysis of present and changing probe sets, FDR=0.01 (n=5,501)
3.15	Probe sets common to probe sets present and changing with greater expression in t(15;17) than in non-t(15;17) samples and in donor samples by ≥ 3 fold, SAM analysis of all probe sets using an FDR of 0.05 and ANOVA analysis of probe sets present and changing using an FDR of 0.05 (n=17)

Table 5. The probe set lists identified from the gene expression profiling analysis (Section 3.3.1 and 3.3.2).

3.4 Discussion

Gene expression profiling has been used to identify leukaemic populations based on their clinical subtypes and karyotypes and has identified genes specific for those groups (Andersson *et al*, 2005; Valk *et al*, 2004; Bullinger *et al*, 2004). Using this information, it may be possible to identify genes that can be used as markers of MRD for specific AML sub-groups. As a “proof of principle”, gene expression profiling was performed on AML patients to identify genes specific to t(15;17) AML patients; patients for which an MRD marker is already available. This approach allowed the selection of potential MRD marker candidates that were then compared with the markers currently used for MRD monitoring in t(15;17) AML patients i.e. the *PML-RAR α* and *RAR α – PML* fusion genes.

By profiling 295 diagnostic AML samples and 10 donor samples, a list of 18 probe sets, present in at least 1 sample, was identified as being specific for AML patients with a t(15;17). Further analysis identified 17/18 of these probe sets as present and changing in at least 1 sample and as being specific for those AML patients with a t(15;17). The probe set filtered out when the analysis was performed on present and changing probe sets related to the podoplanin (*PDPN*) gene. *PDPN* is a small mucin-like transmembrane protein. Whilst its physiological function is unknown, its expression is upregulated in many human cancers, including squamous cell carcinomas of the oral cavity and larynx (Martin-Villar *et al*, 2005). Various studies by Wicki *et al* (2006)

and Martin-Villar *et al* (2006) have shown the expression of *PDPN* in human cancers promotes the migration and invasion of cancer cells (Martin-Villar *et al*, 2006;Wicki *et al*, 2006). The identification of this gene from the less stringently filtered group of probe sets may be due to one of several reasons. The probe set may be false-positive identified as specific to t(15;17) AML because the original filtering to identify probe sets assigned a call of “present “ in at least 1 sample” was not stringent enough and did not take into account a lack of change in expression across the test sets, or the probe set may be specific to t(15;17) AML but not enough to act reliably as a marker of MRD for this sub-group and may be specific to AML as a whole rather than to an individual sub-group.

The 17 probe sets identified as specific for t(15;17) patients related to 12 genes. Hepatocyte growth factor (*HGF*) was represented by 4 of the identified probes. *HGF* was also identified by Gutierrez *et al* (2005) as being the second most significantly over expressed gene in APL patients compared to non-APL patients by gene expression profiling. This finding was supported by Valk *et al* (2004) identifying *HGF* as being specific to the AML t(15;17) sub-group (Gutierrez *et al*, 2005;Valk *et al*, 2004). *HGF* is a pleiotropic cytokine that promotes the growth and motility of epithelial cells in several organs (Borset *et al*, 1996) (Bussolino *et al*, 1992). The ability of *HGF* to promote cell proliferation, invasion and motility has encouraged the proposal that the *HGF* gene plays a role in cancer growth and metastasis (Bussolino *et al*, 1992).

However, a role in leukaemogenesis does not necessarily translate to a potential MRD marker; some genes displaying an increase in expression with the onset of malignancy may play no role in the pathogenesis. Many groups have reported increased concentrations of the HGF protein in AML patients' serum samples at diagnosis compared to normal controls in haematological malignancies, including AML; Verstovsek *et al* (2001) measured significantly higher concentrations of plasma HGF in diagnostic AML and MDS (Myelodysplastic Syndrome) samples compared to healthy normal samples ($p < 0.0001$) and Gwang *et al* (2005) found significantly higher HGF serum concentrations in patients with CML (Chronic Myeloid Leukaemia) and patients with AML than in healthy normal donors ($P < 0.0001$, $P < 0.001$ respectively) (Verstovsek *et al*, 2001; Gwang *et al*, 2005). In this study, the level of *HGF* expression was screened in 8 t(15;17) AML patient samples and 10 healthy normal donor samples. Whilst Verstovsek *et al* (2001) and Gwang *et al* (2005) screened the levels of the HGF protein; here *HGF* transcript levels were measured. Similar to the two protein studies, the levels of *HGF* transcript expression were significantly increased in the AML diagnostic samples compared to the healthy normal samples ($P < 0.0005$) (Verstovsek *et al*, 2001; Gwang *et al*, 2005). An obvious next step would be to compare the protein levels and transcript levels of *HGF* in the same samples. However, the circumstantial evidence from these unrelated studies suggests that there is probably a positive correlation between the level

of expression of *HGF* transcript in t(15;17) AML and the amount of HGF protein translated.

The levels of *HGF* mRNA were screened in three t(15;17) AML patient diagnostic and follow-up samples. In all three patients the normalised expression of *HGF* decreased in the follow-up samples when compared to the diagnostic samples. A very strong correlation between *HGF* and PML-RAR α and RAR α -PML ($p < 0.0001$) was also observed. The only comparable patient study of a haematological malignancy to relate these results to is a myeloma study by the Nordic Myeloma Study Group (1998). Seidel *et al* (1998) observed that HGF serum levels were raised in diagnostic samples of multiple myeloma patients compared to normal samples, decreased after treatment response and increased during relapse, and therefore suggested HGF as a candidate disease marker (Seidel *et al*, 1998).

Another probe set, related to insulin-like growth factor binding protein 2 (*IGFBP2*) gene, was identified as being specific for AML patients with t(15;17). In agreement with these findings, Gutierrez *et al* (2005) also identified *IGFBP2* as being highly expressed in APL patients (Gutierrez *et al*, 2005). A previous study by Vorwerk *et al* (2005) found highly elevated serum levels of IGFBP2 in acute lymphoblastic leukaemia (ALL) patients at diagnosis (Vorwerk *et al*, 2005). Insulin-like growth factor binding proteins (IGFBPs) bind to insulin-like growth factors (IGFs) and mediate the actions of IGFs through transport to specific

cells (Rodriguez *et al*, 2007;Hettmer *et al*, 2005;Rodriguez *et al*, 2007). IGFBP2 has also been found to act as a regulatory carrier for IGF in childhood leukaemia (Vorwerk *et al*, 2002). The level of *IGFBP2* expression was screened in 9 t(15;17) AML diagnostic patient samples and 10 healthy normal donor samples. Similar to the findings of Vorwerk *et al* (2005), the levels of *IGFBP2* expression were significantly higher in the t(15;17) AML diagnostic patients samples compared to the healthy normal donor samples ($P<0.0001$). The levels of *IGFBP2* MRD were screened in three t(15;17) AML patient diagnostic and follow-up samples. In all three patients the expression of *IGFBP2* decreased following diagnosis and the commencement of induction therapy and remained at these lower levels for the remainder of the disease. As with *HGF*, a very strong correlation between *IGFBP2* and *PML-RAR α* and *RAR α -PML* ($p<0.0001$) was observed. To date *IGFBP2* has not been investigated as a marker of MRD but another study by Vorwerk *et al* (2002) observed relapse of children with ALL was restricted to those with raised serum IGFBP2 concentrations at diagnosis and a difference was seen in the serum IGFBP2 concentration at diagnosis of those children who remained in continuous remission (Vorwerk *et al*, 2002).

Combining the gene expression data generated from this study with the findings of various studies by other groups supports the potential role of gene expression profiling for the identification of markers of MRD for AML sub-types, in particular those currently without markers. Further gene profiling analysis using statistical tools (ANOVA, SAM) combined

with fold change thresholds identified a set of probe sets identical to the previous list identified in section 3.3.1b. These findings validate the strength of the results generated from the gene expression profiling analysis of fold change calculations alone, and the potential of gene expression profiling for identifying MRD markers. The diagnostic screening of two of the identified genes, *HGF* and *IGFBP2*, compared to healthy normal donors and the longitudinal monitoring their levels of expression in patient follow-up samples indicated that both genes may have the potential to act as markers of MRD in t(15;17) AML. Further investigations are now clearly warranted in which transcript and protein levels of HGF and IGFBP2 are serially monitored in order to establish the potential utility of these genes in MRD monitoring and the role they play in leukaemogenesis.

Whilst the two candidate genes, *HGF* and *IGFBP*, displayed expression patterns that would be expected of an ideal MRD marker in patients that achieved and remained in complete remission, it was unfortunate that longitudinal monitoring was not performed in any t(15;17) APL patients that subsequently relapsed to observe an increase in expression with the onset of or prior to relapse to confirm their potential as MRD markers. However, due to the improved treatment strategies and remission rates achieved for APL a relapse in APL is not commonly seen.

A limitation to the study could have resulted from the sensitivity of the technique used. The sensitivity of the RQ-PCR technique used can be defined as the lowest dilution of cDNA still able to generate specific amplification (Papadaki *et al*, 2009). AML cell line cDNA was serially diluted to maximum of 1 in 10,000. At a sensitivity level of 10^{-4} specific amplification of the cDNA using the *HGF* and *IGFBP2* primers was observed for all samples; samples generated a Cp value of less than 40 cycles and were deemed as “positive” (Section 2.5.3). If a sample had generated a Cp value of greater than 40 it would have been deemed “negative” and all residual disease presumed to be absent. However, a more sensitive RQ-PCR (10^{-5} - 10^{-6}) might detect some residual disease proving its absence in this current study as a “false-negative”. Fortunately specific amplification of all samples in this section was detected with 40 cycles and issues of “false-negative” MRD presence were avoided.

To summarise, gene expression profiling was able to identify a list of genes specific to t(15;17) AML patients. Selecting two of these genes and following the levels of expression in three t(15;17) patients, a strong correlation was seen between the expression of these genes and the expression of the two markers currently used for MRD monitoring, *PML-RAR α* and *RAR α -PML*. This suggests that gene expression profiling can identify genes specific to t(15;17) AML and therefore has the potential to identify minimal residual disease markers for monitoring in other subgroups of AML.

Chapter 4

Using gene expression profiling to identify novel minimal residual disease markers specific to CBF leukaemias

4.1 Introduction

4.1.1 Core Binding factors (CBF)

Core binding factors (CBF) are a group of transcriptional regulators that contain a beta (CBF β) subunit and one of three alpha (CBF α) subunits. Both subunits are required for normal haematopoietic development (Speck & Gilliland, 2002). The CBF members are common targets for gene mutations and rearrangements, and CBF acute leukaemia results from translocations involving either a CBF α subunit or a CBF β subunit (Appelbaum *et al*, 2006).

4.1.2 CBF Leukaemia

CBF leukaemia account for approximately 15% of AML cases under the age of 55 years (Slovak *et al*, 2000; Grimwade *et al*, 1998).

4.1.2a t(8;21)

Acute myeloid leukaemia 1 (AML1), the CBF alpha subunit, is involved in many chromosomal translocations in several haematological malignancies (Burel *et al*, 2001). The t(8;21) chromosomal translocation is commonly associated with the AML FAB group M2. The t(8;21) involves a translocation between the *AML1* gene on chromosome 21 and the eight twenty-one oncoprotein (*ETO*) gene located on chromosome 8, resulting in the fusion gene product *AML1-ETO* (Peterson & Zhang, 2004). The role *AML1-ETO* plays in leukaemogenesis is to block cell differentiation and down-regulate the

expression of CCAAT/enhancer binding protein, alpha (CEBP α), a transcription factor involved in the commitment and differentiation of myeloid cells (Burel *et al*, 2001).

4.1.2b inv(16)

The beta subunit, CBF β , is involved in the inv(16) commonly associated with the AML FAB group M4Eo (Delaunay *et al*, 2003). An inversion of chromosome 16 results in an inversion of CBF β and the smooth muscle myosin heavy chain (MYH11) gene located on chromosome 16, resulting in the fusion gene product CBF β -MYH11. The CBF β unit is able to regulate transcription by interacting with the AML1 DNA-binding transcription factor. Disruption to CBF β through chromosome 16 inversion leads to a disruption of AML1 and transcription is negatively regulated (Lutterbach *et al*, 1999; Wang *et al*, 1993).

4.1.3 CBF Leukaemia and Gene Expression Profiling

Gene expression studies with microarrays have revealed distinct transcriptional differences between the two CBF leukaemias (Mrozek & Bloomfield, 2006) (Bullinger *et al*, 2007). Several groups have performed unsupervised hierarchical clustering analyses on sets of AML patient samples and found that despite both belonging to the collective group of CBF leukaemias, the t(8;21) patients and the inv(16) patients cluster into two distinct groups with different gene expression

patterns (Bullinger *et al*, 2004; Bullinger *et al*, 2007; Ichikawa *et al*, 2006).

4.1.4 CBF Leukaemia and MRD

Minimal residual disease (MRD) in patients with a CBF leukaemia can be monitored by quantitative real-time PCR (RQ-PCR) of their resultant fusion gene products, *AML1-ETO* and *CBF β -MYH11*. However, various studies have shown that the t(8;21) fusion transcript can still be detected in some t(8;21) following complete remission and throughout long-term remission (Kwong *et al*, 1996; Nucifora *et al*, 1993; Downing *et al*, 1993).

4.1.5 Wilms' Tumour (*WT1*) as an MRD Marker

Wilms' tumour 1 (*WT1*) was originally identified as a tumour suppressor gene involved in the pathogenesis of Wilms' tumour, a kidney neoplasm (Haber *et al*, 1990). Expression of *WT1* is found in a variety of tissues, including those involved in haematopoiesis and the kidney (Call *et al*, 1990). Several studies have detected low levels of *WT1* expression in normal bone marrow and peripheral blood samples yet high levels in AML samples and other leukaemic samples (Inoue *et al*, 1996; Inoue *et al*, 1994). Whilst it is unknown why *WT1* expression has been found to be increased, *WT1* has been suggested as a candidate for MRD monitoring in AML and has the potential to predict relapse (Inoue *et al*, 1994). However, some controversy surrounds the clinical significance of

WT1 expression and data generated by reverse transcriptase-PCR. Schmid *et al* (1997) found that while *WT1* was expressed in leukaemic blasts, the data was too limited to come to a conclusion about the usefulness of *WT1* as a marker of MRD (Schmid *et al*, 1997), whilst Gaiger *et al* (1998) still detected *WT1* expression in patients following complete remission (Gaiger *et al*, 1998). Elmaagacli *et al* (2000) also detected *WT1* expression in patients following transplantation (Elmaagacli *et al*, 2000).

Due to the persistence of t(8;21) gene transcripts following remission in t(8;21) patients and the need for further validation for the use of *WT1* as an MRD marker, other potential markers may provide more reliable information about CBF leukaemia. Building on previous gene expression profiling studies performed on CBF leukaemias and the methods used in the previous chapter for acute promyelocytic leukaemia, the aim of this chapter was to identify potential MRD markers for t(8;21) AML patients and inv(16) AML patients and compare their ability to monitor CBF AML with *AML1-ETO*, *CBF β -MYH11* and *WT1* transcript data from these patients.

4.2 Methods

4.2.1 Patient and donor samples

The sample dataset used for analysis in this section contained gene expression data for 295 patient samples (5 patient replicate samples)

and 10 healthy normal donor samples processed within the School of Medicine, Cardiff University as described in section 2.2.1 (Chapter 2). The patient cohort used contained 15 AML patients presenting with a t(8;21) and 18 AML patients presenting with an inv(16) (1 inv(16) patient replicate sample) (Chapter 2, Table 3).

The t(8;21) and inv(16) AML patient RNA for RQ-PCR was processed from bone marrow aspirates and supplied from the Department of Clinical Haematology, Manchester Royal Infirmary. The normal healthy donor RNA was processed from bone marrow aspirates and supplied from within the Department of Haematology, School of Medicine.

4.2.2 Gene Expression Data Analysis

The gene expression profiling analysis to determine genes differentially expressed in t(8;21) AML patients and in inv(16) AML patients was performed in an identical manner to the analysis of t(15;17) patients as described in chapter 3, section 3.2.2. As previously seen, a quality control test was applied to the probes sets before applying fold change and statistical analyses.

4.2.3 Diagnostic Screening of Candidate Genes

The genes identified from the t(8;21) gene expression profiling analysis were screened in a group of diagnostic t(8;21) AML patient bone marrow samples (n=10) and normal healthy donor bone marrow

samples (n=16). The genes identified from the inv(16) gene expression profiling analysis were screened in a group of diagnostic inv(16) AML patient bone marrow samples (n=13) and normal healthy donor bone marrow samples (n=16). Real-time quantitative PCR (RQ-PCR) was used as previously described in section 2.5 and section 3.2.3 (Chapter 2 & Chapter 3) to measure the expression levels of the candidate genes.

4.2.4 Validation of Candidate Genes as MRD Markers

The candidate genes screened in diagnostic patient samples were validated as markers of MRD using RQ-PCR in the same manner as described in chapter 3 (Section 3.2.4). Data for the markers currently used, *AML1-ETO*, *MYH11-CBF β* and *WT1*, was supplied by Professor J L Yin, Department of Clinical Haematology, Manchester Royal Infirmary. The data were displayed in the form of normalised copy number using the following calculation:

$$\text{Normalised copy} = (\text{number of copies/number of } ABL \text{ copies}) \times 100$$

4.2.5 Statistical Analysis of Diagnostic Screening and MRD Analyses

The gene expression data generated from the diagnostic patient and normal healthy donor samples was analysed using an unpaired student *t*-test. All MRD data was analysed using a two-tailed Pearson's correlation. A Pearson's correlation is used to visualise the degree of

linear relationship between the MRD data generated for the candidate genes and for markers currently used.

4.3 Results

A. t(8;21) AML

4.3.1 Identification of genes expressed specifically in t(8;21) AML samples

Using filtering tools based on present/absent calls and fold change and statistical analysis tools, 22,283 probe sets in 295 AML diagnostic patient samples (representing 290 patients) and 10 donor samples were analysed to identify genes specific to those patients with a t(8;21) translocation.

4.3.1a Identification of Differentially Expressed Genes by Fold change

The fold change calculation used was applied to the two probe set groups previously identified as “present in at least 1 sample” and “present and changing” through filtering on flags (Chapter 3, Section 3.3.1a) to remove any inconsistent data. Filtering was applied to both groups of probe sets to assess the robustness and reliability of the analysis.

The group of “present in at least 1 sample” probe sets were filtered on a 2-, 3-, 4- and 5-fold change in expression level between the t(8;21)

patients and non-t(8;21) patients and between the t(8;21) patients and the normal healthy donor patients (Table 1). By using a ≥ 3 fold change as the threshold for significance, the group of 94 probe sets identified as having a greater expression in the t(8;21) patients than in the non-t(8;21) patients (probe set list 4.1) was combined with the group of 272 probes sets identified as having a greater expression in the t(8;21) patients than in the normal healthy donors (probe set list 4.2) to generate a group of 60 probe sets with a greater level of expression in t(8;21) AML patients when compared to other AML patient sub-groups or normal healthy donors (Probe set list 4.3) (Figure 1; Table 2).

The “present and changing” probe sets were filtered on a 2-, 3-, 4- and 5-fold change in expression level between the t(8;21) patients and non-t(8;21) patients and between the t(8;21) patients and the normal healthy donors (Table 1). By using a ≥ 3 fold change as the threshold for significance the group of 94 probe sets identified as having greater expression in the t(8;21) patients than in the non-t(8;21) patients (probe set list 4.4) was combined with the group of 272 probe sets identified as having a greater expression in t(8;21) patients than in the normal healthy donors (probe set list 4.5) to generate a group of 60 probe sets with a greater level of expression in t(8;21) AML patients when compared to other AML patient sub-groups or normal healthy donor (Probe set list 4.6) (Figure 1; Table 2).

The group of 60 probe sets identified from the “present in at least 1 sample” probe sets (Probe set list 4.3) and the group of 60 probe sets identified from the “present and changing” probe sets (Probe set list 4.6) were identical confirming the robustness of the filtering analysis and the unlikelihood that the probe sets identified as specific to t(8;21) AML were false-positives.

Probe set group	Fold Change	t(8;21)>non-t(8;21)	t(8;21)>Donors
Present in at least 1 sample	2	318	930
	3	94	272
	4	31	95
	5	23	64
Present and changing	2	313	930
	3	94	272
	4	31	75
	5	23	23

Table 1. The number of probe sets identified as having a greater expression in t(8;21) AML patients than in non-t(8;21) patients or normal healthy donors. Fold change filtering was applied to both sets of probe sets generated for quality control purposes as described in section 3.3.1a (Chapter 3).

The 60 probe sets represented 46 genes. Furthermore amongst the 16 probe sets with the highest fold changes (greater than 8 fold) 7 probe sets identified the tryptase alpha 1 and beta 1 genes (*TPSAB1*) and tryptase beta 2 (*TPSB2*) gene involved in proteolysis.

	Probe Set	Gene Symbol	Gene Title	Fold Change
31	203913_s_at	<i>HPGD</i>	Tydoxyprostaglandin dehydrogenase 15-(NAD)	4.220
32	203065_s_at	<i>CAV1</i>	Caveolin 1, caveolae protein, 22kDa	4.439
33	204468_s_at	<i>TIE1</i>	Tyrosine kinase with immunoglobulin-like and EGF-like domains 1	4.551
34	219686_at	<i>STK32B</i>	Serine/threonine kinase 32B	4.716
35	204885_s_at	<i>MSLN</i>	Mesothelin	5.182
36	209170_s_at	<i>GPM6B</i>	Glycoprotein M6B	5.236
37	218876_at	<i>TPPP3</i>	Tubulin polymerization-promoting protein family member 3	5.312
38	208534_s_at	<i>RASA4</i>	RAS p21 protein activator 4	5.353
39	209010_s_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	5.453
40	209167_at	<i>GPM6B</i>	Glycoprotein M6B	6.328
41	206726_at	<i>PGDS</i>	Prostaglandin D2 synthase, hematopoietic	6.844
42	206940_s_at	<i>POU4F1</i>	POU class 4 homeobox 1	7.522
43	212097_at	<i>CAV1</i>	Caveolin 1, caveolae protein, 22kDa	8.016
44	213194_at	<i>ROBO1</i>	Roundabout, axon guidance receptor, homolog 1 (Drosophila)	8.209
45	207741_x_at	<i>TBSAB1</i>	Tryptase alpha/beta 1	8.355
46	204086_at	<i>PRAME</i>	Preferentially expressed antigen in melanoma	8.990
47	207134_x_at	<i>TPSB2</i>	Tryptase beta 2	9.150
48	201655_s_at	<i>HSPG2</i>	Heparan sulfate proteoglycan 2	9.649
49	205683_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	10.01
50	216474_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	12.06
51	216831_s_at	<i>RUNX1T1*</i>	Runt-related transcription factor 1	13.09
52	205529_s_at	<i>RUNX1T1*</i>	Runt-related transcription factor 1	13.25
53	215382_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	13.61
54	217023_x_at	<i>TPSB2</i>	Tryptase alpha/beta 1	14.81
55	210744_s_at	<i>IL5RA</i>	Interleukin 5 receptor, alpha	16.56
56	205528_s_at	<i>RUNX1T1*</i>	Runt-related transcription factor 1	18.12
57	210084_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	21.71
58	211517_s_at	<i>IL5RA</i>	Interleukin 5 receptor, alpha	25.35
59	211341_at	<i>POU4F1</i>	POU class 4 homeobox 1	39.08
60	206622_at	<i>TRH</i>	Thyrotropin-releasing hormone	39.27

Table 2. The top 30 “present in at least 1 sample” and “present and changing” probe sets identified from the 305 sample dataset as upregulated in t(8;21) AML patients samples when compared to non-t(8;21) patients and to healthy donor samples by fold change (FC>3). The fold change value displayed is the mean fold change value of the probe sets across the t(8;21) samples. The numbers assigned to the probe sets correspond to their location in figure 1. For a complete list of the 60 probe sets see appendix II.**RUNX1T1* is also known as *ETO*, the gene previously mentioned as involved in the t(8;21) (Section 4.2.1a).

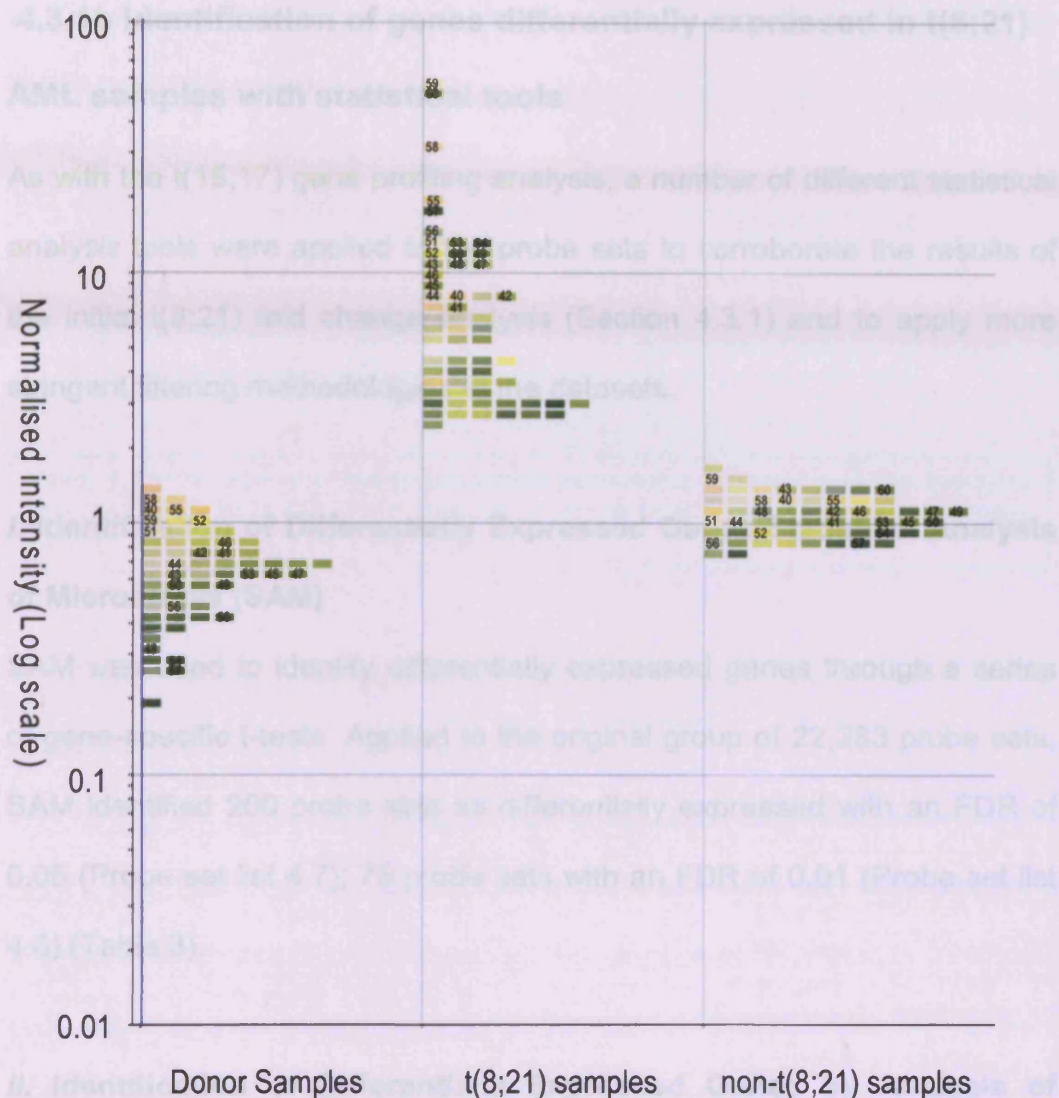


Figure 1. The relative mean expression levels of the 60 probe sets identified, from probe sets that were classified as present in at least 1 sample and as present and changing, as being more significantly expressed in t(8;21) samples than non-t(8;21) samples and in donor samples by ≥ 3 fold. The probe sets are coloured according to their mean expression levels in their respective sub-group relative to their mean expression level in the donor samples. The numbers on each block correspond to the probe set it is representing - see table 2 (only top 20 probe sets numbered).

4.3.1b Identification of genes differentially expressed in t(8;21)

AML samples with statistical tools

As with the t(15;17) gene profiling analysis, a number of different statistical analysis tools were applied to the probe sets to corroborate the results of the initial t(8;21) fold change analysis (Section 4.3.1) and to apply more stringent filtering methodologies to the datasets.

i. Identification of Differentially Expressed Genes Statistical Analysis of Microarrays (SAM)

SAM was used to identify differentially expressed genes through a series of gene-specific t-tests. Applied to the original group of 22,283 probe sets, SAM identified 200 probe sets as differentially expressed with an FDR of 0.05 (Probe set list 4.7); 75 probe sets with an FDR of 0.01 (Probe set list 4.8) (Table 3).

ii. Identification of Differentially Expressed Genes by Analysis of Variance (ANOVA)

ANOVA was applied to the group of probe sets classified as “present and changing” (n=16,977) to identify differences in gene expression levels between the t(8;21) patients and the non-t(8;21) patients and normal healthy donors by comparing the variances of the test groups. Using an FDR of 0.05, 7,271 probe sets (probe set list 4.9) were identified, while

4,957 probe sets (probe set 4.10) were identified with an FDR of 0.01 (Table 3).

FDR	SAM	ANOVA
0.05	200	7,271
0.01	75	4,957

Table 3. The number of probe sets identified as having a greater expression in t(8;21) AML patients than in non-t(8;21) patients or normal healthy donors through SAM and ANOVA analysis.

4.3.1c Identification of Differentially Expressed Genes by Fold change and Statistical tools

Combining the differentially expressed probe sets identified by SAM and ANOVA analysis with an FDR of 0.05 (n=200, n=7,271) with the probe sets identified as “present and changing” with a greater expression in t(8;21) AML patients than in any other AML sub-group or normal healthy donors (n=60), a set of 41 common probe sets was generated as being specific for t(8;21) samples (Probe set list 4.11; Figure 2; Table 4; Figure 3).

Unlike the analysis performed for the t(15;17) sub-group (Chapter 3), statistical tools used for the identification of probe sets specific to t(8;21) AML filtered out many probe sets identified as specific from the original fold change filtering. This could be due to a result of a greater variance in the expression levels of the probe sets amongst the t(8;21) samples compared

to the t(15;17) samples resulting in less probe sets originally identified achieving the required threshold expression levels of SAM and ANOVA analysis to be identified as specific to t(8;21) AML.

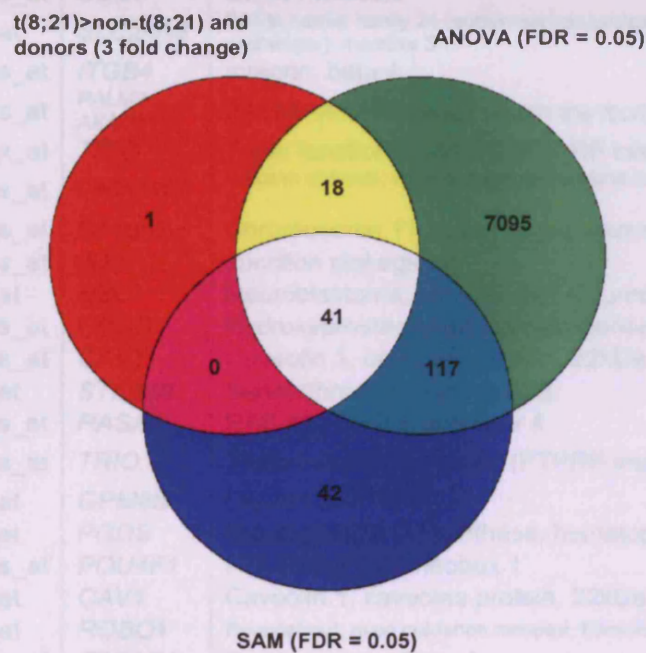


Figure 2. A combination of the probe sets identified from SAM (FDR=0.05), ANOVA testing (FDR=0.05) and having a greater level of expression in t(8;21) samples than in non-t(8;21) samples and in donor samples by 3 fold change filtering. Forty-one probe sets were identified as being specifically expressed in t(8;21) samples.

	Probe Set	Gene Symbol	Gene Title	Fold Change
3	202242_at	<i>TSPAN7</i>	Tetraspanin 7	3.010
5	208178_x_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	3.033
6	206187_at	<i>PTGIR</i>	Prostaglandin I2 (prostacyclin) receptor (IP)	3.082
9	207030_s_at	<i>CSRP2</i>	Cysteine and glycine-rich protein 2	3.229
12	211140_s_at	<i>CASP2</i>	Caspase 2, apoptosis-related cysteine peptidase	3.247
14	207496_at	<i>MS4A2</i>	Membrane-spanning 4-domains, subfamily A, member 2	3.293
16	211549_s_at	<i>HPGD</i>	Hydroxyprostaglandin dehydrogenase 15-(NAD)	3.439
19	209543_s_at	<i>CD34</i>	CD34 molecule	3.542
20	219090_at	<i>SLC24A3</i>	Solute carrier family 24 (sodium/potassium/calcium exchanger), member 3	3.590
21	204990_s_at	<i>ITGB4</i>	Integrin, beta 4	3.614
22	202760_s_at	<i>PALM2-AKAP2</i>	PALM2-AKAP2 readthrough transcript	3.653
23	209013_x_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	3.671
25	204811_s_at	<i>CACNA2D2</i>	Calcium channel, voltage-dependent, alpha 2/delta subunit 2	3.787
26	204073_s_at	<i>C11orf9</i>	Chromosome 11 open reading frame 9	3.843
27	201015_s_at	<i>JUP</i>	Junction plakoglobin	3.932
29	201621_at	<i>NBL1</i>	Neuroblastoma, suppression of tumorigenicity 1	3.986
31	203913_s_at	<i>HPGD</i>	Hydroxyprostaglandin dehydrogenase 15-(NAD)	4.220
32	203065_s_at	<i>CAV1</i>	Caveolin 1, caveolae protein, 22kDa	4.439
34	219686_at	<i>STK32B</i>	Serine/threonine kinase 32B	4.716
38	208534_s_at	<i>RASA4</i>	RAS p21 protein activator 4	5.353
39	209010_s_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	5.453
40	209167_at	<i>GPM6B</i>	Glycoprotein M6B	6.328
41	206726_at	<i>PGDS</i>	Prostaglandin D2 synthase, hematopoietic	6.844
42	206940_s_at	<i>POU4F1</i>	POU class 4 homeobox 1	7.522
43	212097_at	<i>CAV1</i>	Caveolin 1, caveolae protein, 22kDa	8.016
44	213194_at	<i>ROBO1</i>	Roundabout, axon guidance receptor, homolog 1	8.209
45	207741_x_at	<i>TBSAB1</i>	Tryptase alpha/beta 1	8.355
46	204086_at	<i>PRAME</i>	Preferentially expressed antigen in melanoma	8.99
47	207134_x_at	<i>TPSB2</i>	Tryptase beta 2	9.15
48	201655_s_at	<i>HSPG2</i>	Heparan sulfate proteoglycan 2	9.649
49	205683_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	10.01
50	216474_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	12.06
51	216831_s_at	<i>RUNX1T1</i>	Runt-related transcription factor 1	13.09
52	205529_s_at	<i>RUNX1T1</i>	Runt-related transcription factor 1	13.25
53	215382_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	13.61
54	217023_x_at	<i>TPSB2</i>	Tryptase alpha/beta 1	14.81
55	210744_s_at	<i>IL5RA</i>	Interleukin 5 receptor, alpha	16.56
56	205528_s_at	<i>RUNX1T1</i>	Runt-related transcription factor 1	18.12
57	210084_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	21.71
58	211517_s_at	<i>IL5RA</i>	Interleukin 5 receptor, alpha	25.35
59	211341_at	<i>POU4F1</i>	POU class 4 homeobox 1	39.08
60	206622_at	<i>TRH</i>	Thyrotropin-releasing hormone	39.27

Table 4. The 41 probe sets identified from the 305 sample dataset as upregulated in t(8;21) AML patients samples when compared to non-t(8;21) patients and to healthy donor samples by fold change (FC>3) and by SAM analysis and ANOVA analysis with a FDR of 0.05. The numbers assigned to the probe sets correspond to their location in figure 1 and 3.

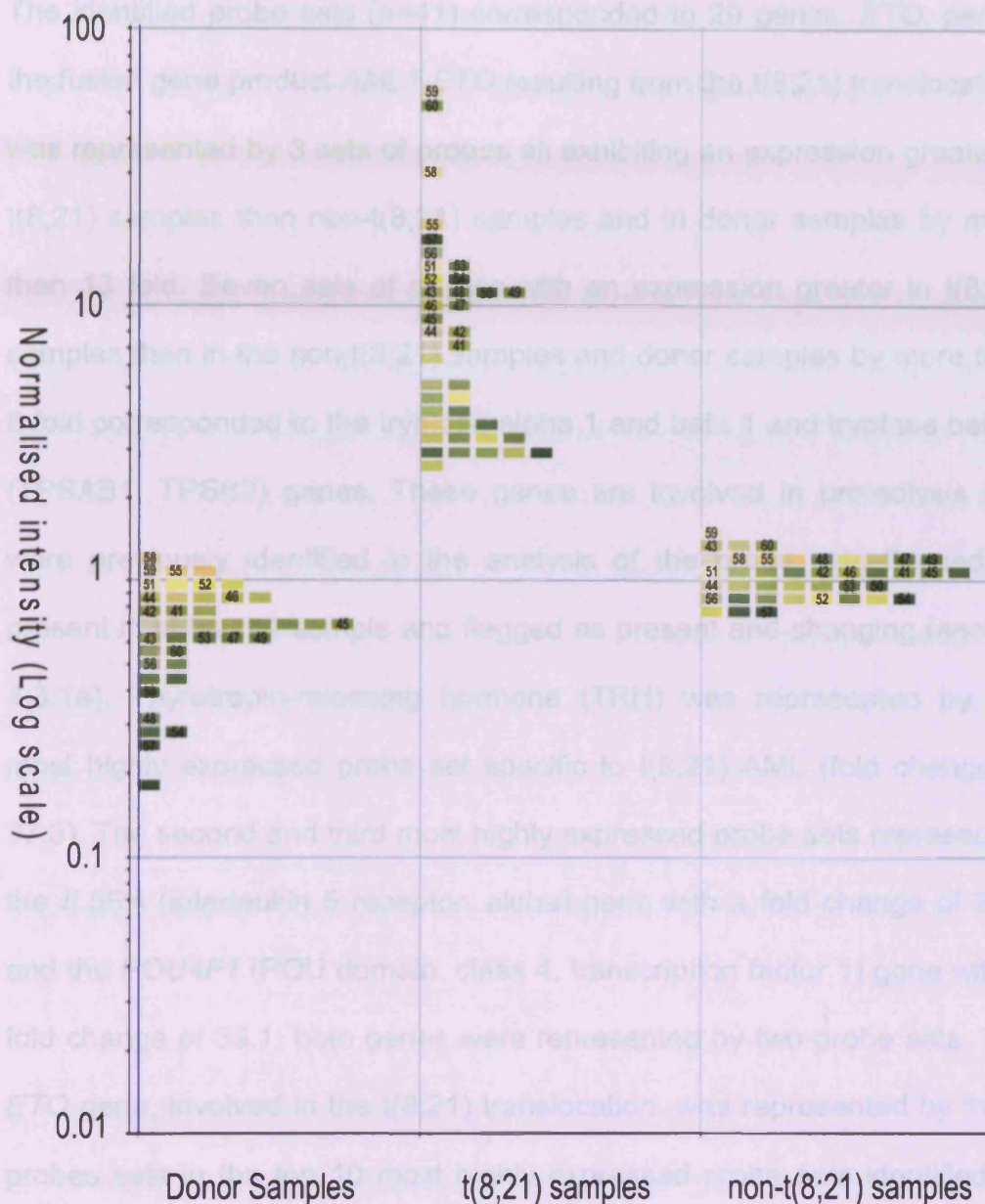


Figure 3. The relative mean expression levels of the 41 “present and changing” probe sets identified as being more significantly expressed in t(8;21) samples than non-t(8;21) samples and in donor samples by ≥ 3 fold and through SAM and ANOVA analysis with an FDR of 0.05. The probe sets are coloured according to their expression levels in their respective sub-group relative to their expression level in the donor samples. The numbers on each block correspond to the probe set it is representing - see Table 2 (only top 20 genes numbered).

The identified probe sets (n=41) corresponded to 29 genes. *ETO*, part of the fusion gene product *AML1-ETO* resulting from the t(8;21) translocation, was represented by 3 sets of probes all exhibiting an expression greater in t(8;21) samples than non-t(8;21) samples and in donor samples by more than 13 fold. Seven sets of probes with an expression greater in t(8;21) samples than in the non-t(8;21) samples and donor samples by more than 8 fold corresponded to the tryptase alpha 1 and beta 1 and tryptase beta 2 (TPSAB1, TPSB2) genes. These genes are involved in proteolysis and were previously identified in the analysis of the probe sets flagged as present in at least 1 sample and flagged as present and changing (section 4.3.1a). Thyrotropin-releasing hormone (TRH) was represented by the most highly expressed probe set specific to t(8;21) AML (fold change of 37.3). The second and third most highly expressed probe sets represented the *IL5RA* (interleukin 5 receptor, alpha) gene with a fold change of 26.4 and the *POU4F1* (POU domain, class 4, transcription factor 1) gene with a fold change of 39.1; both genes were represented by two probe sets. The *ETO* gene, involved in the t(8;21) translocation, was represented by three probes sets in the top 10 most highly expressed probe sets identified as having a greater expression in t(8;21) AML than non-t(8;21) AML or donors by more than 13 fold.

4.3.2 Validation of Candidate Genes as MRD Markers

The *ETO*, *POU4F1*, *IL5RA* and *PRAME* genes were selected for further investigation as potential markers of MRD in t(8;21) AML. The *ETO* gene was selected as it plays a major role in leukaemogenesis, being part of the *AML1-ETO* fusion gene product detected in t(8;21) patients (Nucifora *et al*, 1993). Several studies has detected high levels of *PRAME* in AML and it has already been suggested as a potential MRD marker (Paydas *et al*, 2005; Matsushita *et al*, 2001). The *IL5RA* gene encoders a cytokine receptor involved in the regulation of haematopoietic cells (Pierce *et al*, 1998). The *POU4F1* genes was selected for further investigation from the observation that over-expression of *POU4F1* is associated with a protection of cells from apoptosis (Latchman, 1998; Smith *et al*, 1998). All four genes were also represented by probe sets identified as more highly expressed in t(8;21) AML than non-t(8;21) AML and donors by more than 7.5 fold.

4.3.2a Diagnostic Screening of Candidate Genes

Expression levels of the four genes (*PRAME*, *POU4F*, *ETO* and *IL5RA*) were measured in diagnostic t(8;21) AML patient samples and compared to expression levels measured in healthy normal donor samples.

Initially primer sets were designed for the amplification of the four genes as described in the Materials and Methods section 2.5.1 (Chapter 2) and

using serial dilutions of cDNA, their amplification efficiencies were compared to that of the house-keeping gene, *S14*. All four genes produced similar amplification efficiencies to that produced by *S14*. Therefore *S14* was considered suitable as the normalising control gene (Figure 4).

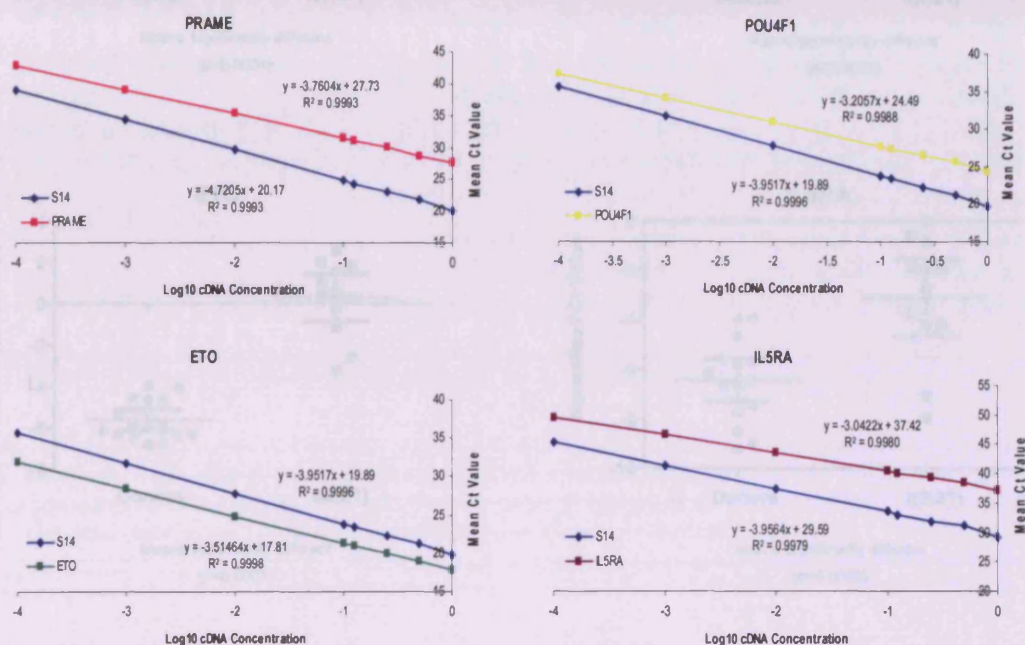


Figure 4. The amplification curves of the primer sets used for *PRAME*, *POU4F*, *ETO* and *IL5RA* and of the house-keeping gene, *S14*, to compare their amplification efficiencies.

The levels of expression of *PRAME*, *POU4F*, *ETO* and *IL5RA* were measured in AML diagnostic patient samples (n=11, 11 12 and 12, respectively) presenting with a t(8;21) and compared to the levels of expression in normal healthy bone marrow donor samples (n=15, 14, 16 and 16, respectively) (Figure 5).

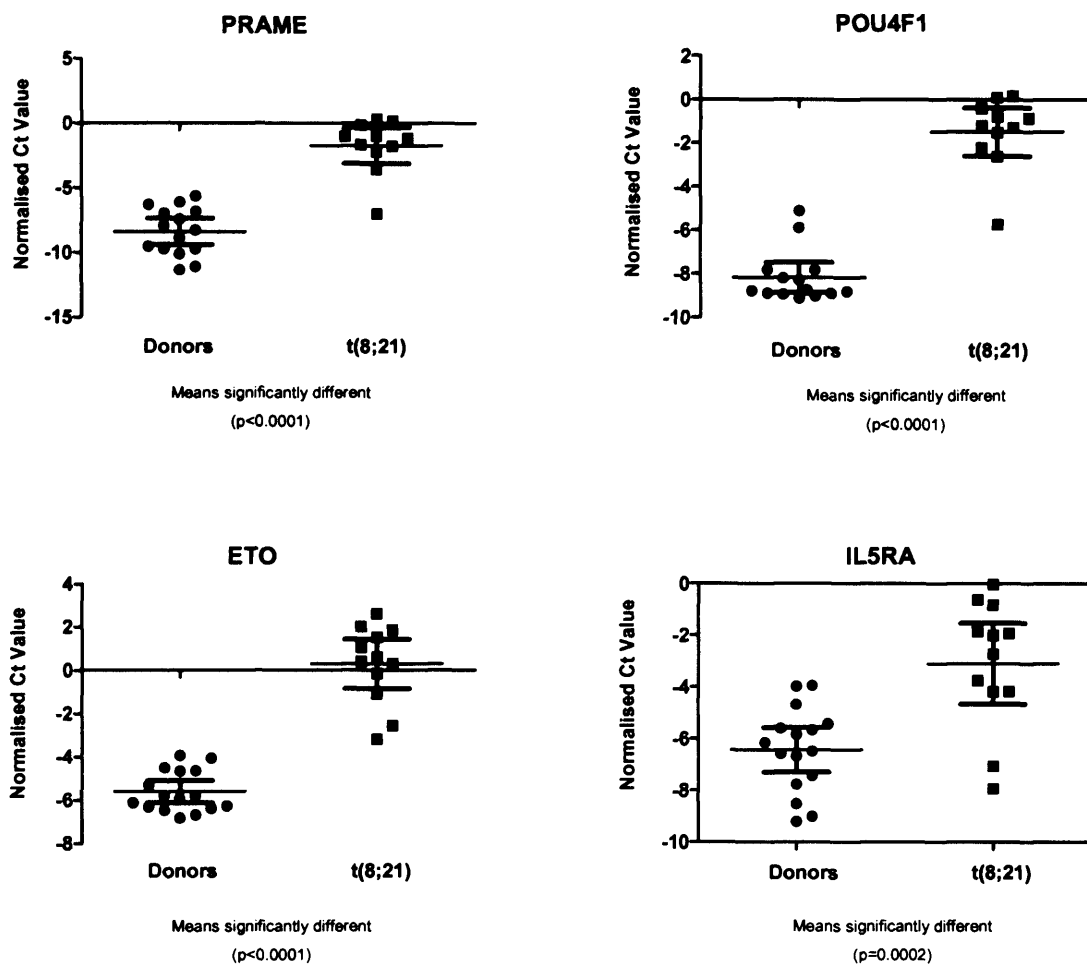


Figure 5. Scatter plots of the spread of the normalised Ct values of the candidate genes measured in the t(8;21) AML diagnostic patient samples and the healthy normal donor samples. Mean value $\pm$ 95% confidence interval plotted.

Using *t*-tests, highly significant differences were observed between the mean normalised Ct values of the candidate genes in the diagnostic t(8;21) AML patient samples and mean normalised Ct Values of the candidate genes in the healthy normal donor samples (Table 5).

	PRAME	POU4F1	ETO	IL5RA
P Value	<0.0001	<0.0001	<0.0001	0.0002
Are the means significantly different (p<0.05)	Yes	Yes	Yes	Yes
Number of donors	11	11	12	12
Donors mean value	-8.401	-8.165	-5.619	-6.442
Donors 95% CI	-9.41 to -7.39	-8.87 to -7.46	-6.13 to -5.11	-7.31 to -5.58
Donors expression range	-11.34 to -5.65	-9.12 to -5.89	-6.83 to -3.95	-9.22 to -3.95
Number of t(8;21)	15	14	16	16
t(8;21) mean value	-1.742	-1.470	0.285	-3.094
t(8;21) 95% CI	-3.13 to -0.35	-2.58 to -0.36	-0.86 to 1.43	-4.66 to -1.53
t(8;21) expression range	-7.02 to 0.30	-5.74 to 0.17	-3.18 to 2.02	-7.94 to -0.62

Table 5. The statistical significance of the mean Ct values of the candidate genes measured in the t(8;21) AML diagnostic patient samples and the healthy normal donor samples normalised to the Ct values of S14. CI = confidence interval.

Based on these findings it was agreed to continue following the expression levels of these genes in AML patient diagnostic and follow-up samples to establish whether any had the potential to act as a marker of MRD for t(8;21) AML.

4.3.2b Candidate Genes as MRD Markers

The transcript expression levels of the four genes were monitored in the diagnostic sample and follow-up samples of three AML patients with a t(8;21).

The fold change of the genes' transcript expression levels relative to *S14* were calculated as described in Chapter 3 and compared to the upper fold change in expression relative to *S14* observed in the healthy normal donor samples (Figure 5). The MRD values (MRDv) of the genes in the patient samples were calculated as described in Chapter 3 and compared to the level of transcript detected at diagnosis (Figure 6).

Normalised copy number data for the fusion gene *AML1-ETO* and the *WT1* gene was supplied from the Department of Clinical Haematology, Manchester Royal Infirmary for comparison with the fold change data generated for the candidate genes:

i. Patient 1

The fold change in expression of *PRAME* and *ETO* decreased by approximately 1 log over the first 9 months of the disease following induction therapy; increasing by approximately 1 log over the following 2.5 months. The level of expression detected of both *PRAME* and *ETO* remained above the upper level detected in healthy normal donor samples, dipping below this level by approximately half a log between 8 and 11 months following diagnosis. The fold change in expression of *POU4F1* and *IL5RA* increased by approximately 2 logs and 1 log, respectively, over the first 3 months of the disease following diagnosis. By 11.5 months post diagnosis, the fold change in expression of *POU4F1* and *IL5RA* had

decreased to a level similar to that detected at diagnosis. The fold change in expression of *IL5RA* remained below the upper level detected in healthy normal donors for the entirety of the monitoring period. This patient relapsed at 11.5 months following diagnosis (Figure 7).

ii. Patient 2

The fold change in expression of *PRAME*, *POU4F1*, *ETO* and *IL5RA* decreased by approximately 4 logs, 3 logs, 3 logs, and 1 log over the first 4.5 months of the disease following induction therapy. By 13.5 months following diagnosis, the fold change in expression of the four genes had increased by approximately 3 logs. Following the initial decrease after diagnosis and induction therapy, the level of expression of *PRAME* and *ETO* detected remained below the upper level detected in healthy normal donors until approximately 11 months following diagnosis. The level of expression of *POU4F1* and *IL5RA* increased to above the upper level detected in healthy normal donors at approximately 9 and 13 months, respectively, following diagnosis. This patient relapsed at 13.5 months following diagnosis (Figure 7).

iii. Patient 3

The fold change in expression of the four genes in patient 3 decreased by approximately 2 logs following diagnosis. The level of fold change in expression of the genes fluctuated for the remainder of the monitoring

period, remaining at least 1 log lower than that detected at diagnosis. Following an initial decrease after diagnosis and induction therapy, the levels of expression detected of the four genes consistently remained below the upper level detected in healthy normal donors. This patient achieved complete remission (CR) (Figure 7)

The largest decrease in MRD from the level of transcript detected at diagnosis was detected in patient 2. *PRAME*, *POU4F1* and *ETO* decreased by approximately 3 logs but the patient relapsed at 13.5 months. The corresponding normalised number of *AML1-ETO* transcripts detected for this patient also displayed a decrease following diagnosis and induction therapy but by approximately 4 logs. By 13.5 months when the patient relapsed, the MRD levels of *PRAME*, *POU4F1* and *ETO* had increased by at least 2 logs whilst the normalised number of *AML1-ETO* transcripts had increased by approximately 4 logs. The levels of MRD of *PRAME*, *POU4F1* and *IL5RA* for patient 1 fluctuated by a maximum of 1 log from the level of transcript detected at diagnosis for the duration of monitoring; the corresponding normalised number of *AML1-ETO* transcripts displayed a decrease by approximately 4 logs from diagnosis and increased by approximately 4 logs between 6 months and 13 months following diagnosis. The MRD levels of all four genes decreased following diagnosis and induction therapy in patient 3, the patient that achieved CR, reaching a minimum MRD value of approximately 3 logs lower than the

level of transcript detected at diagnosis; the normalised copy number of *AML1-ETO* transcripts displayed the same pattern of behaviour but decreased from the number detected at diagnosis by 6 logs. (Figures 6 and 8)

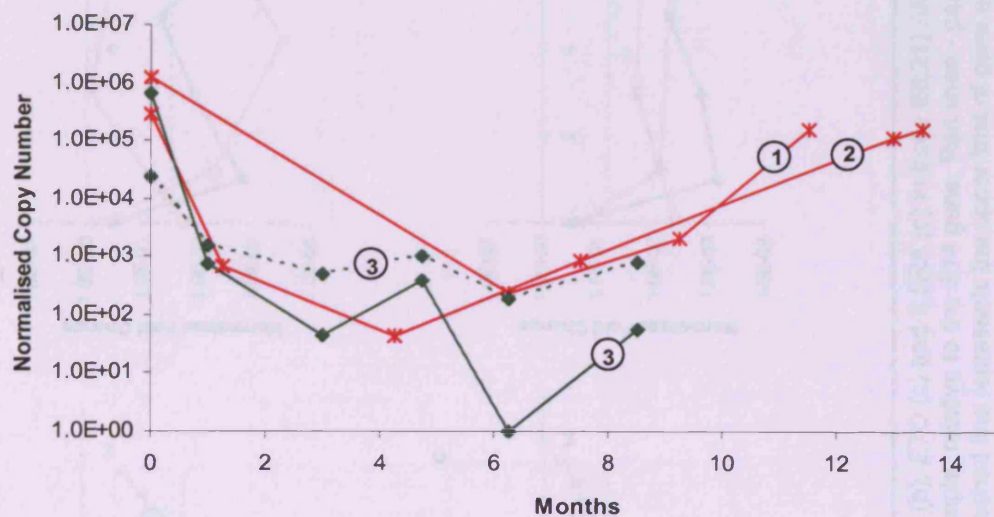


Figure 6. Longitudinal monitoring of the t(8;21) fusion gene in three t(8;21) AML patients, and of *WT1* in one t(8;21) AML patient (no data available for patients 1 and 2). The graphs show the normalised copy numbers of the genes in each patient, relative to the *ABL* gene. Red lines - normalised number of *AML1-ETO* transcripts in patients who relapsed. Green lines - normalised number of *AML1-ETO* transcripts in patient remaining in complete remission (CR). Green dashed line - *WT1* normalised copy numbers in patient remaining in complete remission (CR). The numbers in the circles indicate the patient number.

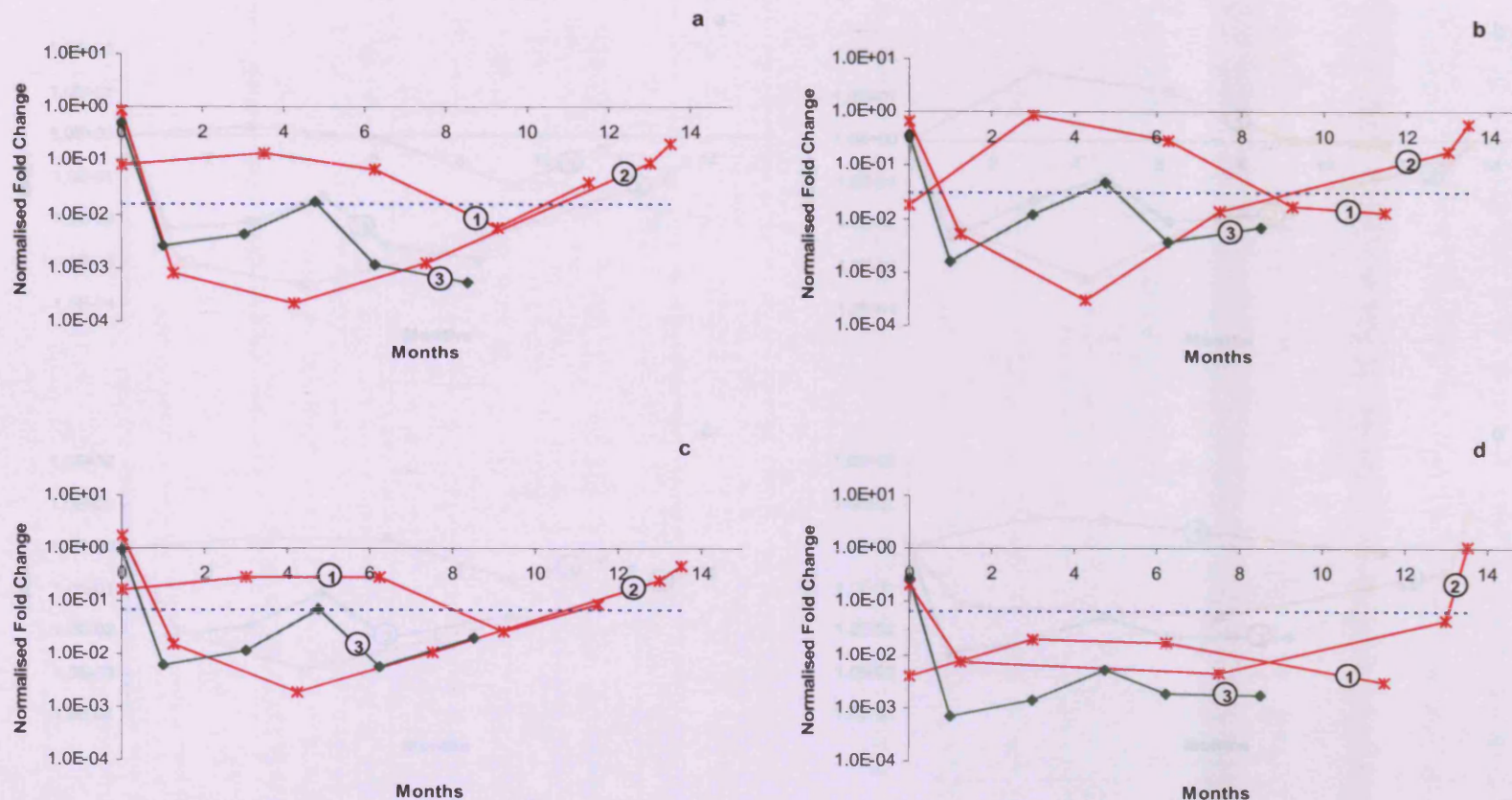


Figure 7. Longitudinal monitoring of *PRAME* (a), *POU4F1* (b), *ETO* (c) and *IL5RA* (d) in three t(8;21) AML patients. The graphs show the fold change values generated for respective genes in each sample, relative to the *S14* gene. Red lines - patients known to relapse. Green lines - patients remaining in complete remission (CR). The blue dashed line represents the upper limit of gene expression in the healthy normal donor samples. The numbers in the circles indicate the patient number.

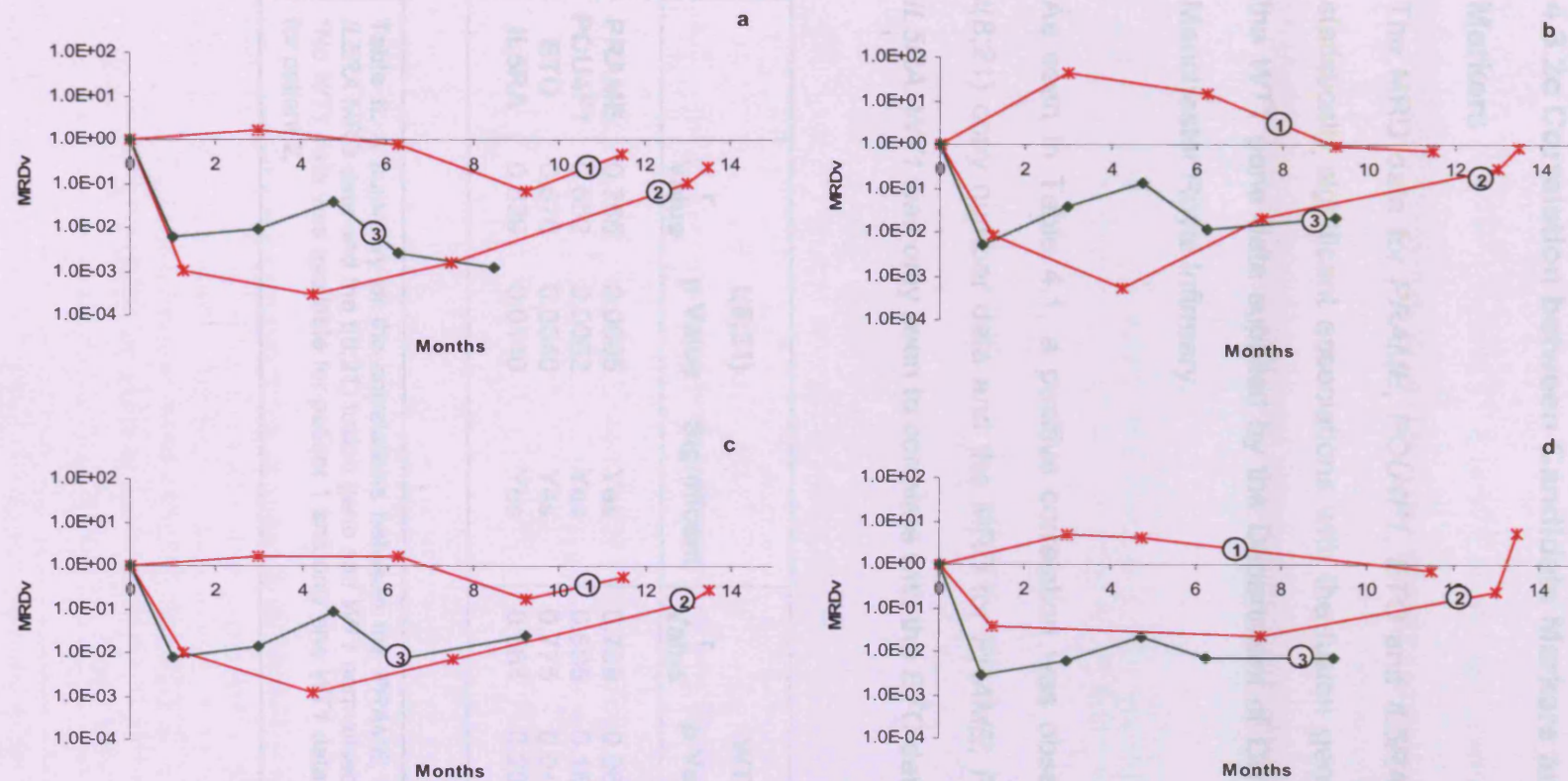


Figure 8. Longitudinal monitoring of *PRAME* (a), *POU4F1* (b), *ETO* (c) and *IL5RA* (d) in three t(8;21) AML patients. The graphs show the MRD values calculated for the genes in each sample. Red lines - patients known to relapse. Green lines - patients remaining in complete remission (CR). The numbers in the circles indicate the patient number.

4.3.2c Correlation between Candidate Markers and Current

Markers

The MRD data for *PRAME*, *POU4F1*, *ETO* and *IL5RA* was analysed for statistically significant associations with the fusion gene t(8;21) data and the *WT1* gene data supplied by the Department of Clinical Haematology, Manchester Royal Infirmary.

As seen in Table 4.1, a positive correlation was observed between the t(8;21) copy number data and the MRD for *PRAME*, *POU4F1*, *ETO* and *IL5RA*. *WT1* was only seen to correlate with the *ETO* data.

	t(8;21)			WT1*		
	r Value	p Value	Significant	r Value	p Value	Significant
PRAME	0.766	0.0005	Yes	0.739	0.0663	No
POU4F1	0.652	0.0062	Yes	0.595	0.1667	No
ETO	0.676	0.0040	Yes	0.775	0.0480	Yes
IL5RA	0.639	0.0140	Yes	0.564	0.2000	No

Table 6. A summary of the correlations between the *PRAME*, *POU4F1*, *ETO* and *IL5RA* MRD data and the t(8;21) fusion gene and *WT1* normalised copy number data.
*No *WT1* data was available for patient 1 and only one *WT1* data point was available for patient 2.

List	Probe Sets
4.1	Probe sets present in at least sample with greater expression in t(8;21) than in non-t(8;21) samples by ≥ 3 fold (n=94)
4.2	Probe sets present in at least sample with greater expression in t(8;21) than in donor samples by ≥ 3 fold (n=272)
4.3	Probe sets present in at least sample with greater expression in t(8;21) than in non-t(8;21) samples and in donor samples by ≥ 3 fold (n=60)
4.4	Probe sets present and changing with greater expression in t(8;21) than in non-t(8;21) samples by ≥ 3 fold (n=94)
4.5	Probe sets present and changing with greater expression in t(8;21) than in donor samples by ≥ 3 fold (n=272)
4.6	Probe sets present and changing with greater expression in t(8;21) than in non-t(8;21) samples and in donor samples by ≥ 3 fold (n=60)
4.7	SAM analysis of all genes, FDR=0.05 (n=200)
4.8	SAM analysis of all genes, FDR=0.01 (n=75)
4.9	ANOVA analysis of present and changing probe sets, FDR=0.05 (n=7,271)
4.10	ANOVA analysis of present and changing probe sets, FDR=0.01 (n=4,957)
4.11	Probe sets common to present and changing probe sets with greater expression in t(8;21) than in non-t(8;21) samples and in donor samples by ≥ 3 fold, SAM analysis of all probe sets using an FDR of 0.05 and ANOVA analysis of probe sets present and changing using an FDR of 0.05 (n=41)

Table 7. The probe set lists identified from the gene expression profiling analysis (Section 4.3.1).

B. inv(16) AML

4.3.3 Identification of genes expressed specifically in inv(16) AML samples

As previously shown with the t(8;21) AML gene expression profiling, filtering tools based on present/absent calls and fold change and statistical analysis tools were used to analysis 22,283 probe sets in 295 AML diagnostic patient samples (representing 290 patients) and 10 donor samples to identify genes specific to those patients with a inv(16).

4.3.3a Identification of Differentially Expressed Genes by Fold change

The fold change calculation used was applied to the two probe set groups previously identified as “present in at least 1 sample” and “present and changing” through filtering on flags (Chapter 3, Section 3.3.1a) to remove any inconsistent data.

The group of “present in at least 1 sample” probe sets were filtered on a 2-, 3-, 4- and 5-fold change in expression level between the inv(16) patients and non-Inv(16) patients and between the inv(16) patients and the normal healthy donor patients (Table 8). By using a ≥ 3 fold change as the threshold for significance, the group of 131 probe sets identified as having a greater expression in the inv(16) patients than in the non-inv(16) patients (probe set list 4.12) was combined with the group of 453 probes sets identified as having a greater expression in the inv(16) patients than in the

normal healthy donors (probe set list 4.13) to generate a group of 100 probe sets with a greater level of expression in inv(16) AML patients when compared to other AML patient sub-groups or normal healthy donor (Probe set list 4.14) (Figure 9; Table 9).

The “present and changing” probe sets were filtered on a 2-, 3-, 4- and 5-fold change in expression level between inv(16) patients and non-inv(16) patients and between the inv(16) patients and the normal healthy donor patients (Table 8). By using a ≥ 3 fold change as the threshold for significance, the group of 130 probe sets identified as having a greater expression in the inv(16) patients than in the non-inv(16) patients (probe set list 4.15) was combined with the group of 446 probes sets identified as having a greater expression in the inv(16) patients than in the normal healthy donors (probe set list 4.16) to generate a group of 99 probe sets with a greater level of expression in inv(16) AML patients when compared to other AML patient sub-groups or normal healthy donor (Probe set list 4.17) (Figure 9; Table 9).

The 99 probe sets identified from the “present and changing” group of probe sets were all identified in the list of 100 probe sets previously identified from the “present in at least 1 sample” group of probe sets. The extra probe set identified represented the activity-regulated cytoskeleton-associated protein (*ARC*) gene.

The identification of the same 99 probe sets as specific to inv(16) AML from both probe set groups re-affirms the reliability and robustness of fold change filtering as a mean of identifying disease-specific genes. The extra probe set identified from the “present in at least 1 sample” may be a false positive due to less stringent filtering methods or specific to AML but not enough to reach the threshold limits of the more stringent filtering methods and therefore, not a suitable marker for MRD.

Probe set group	Fold Change	Inv(16)>non-inv(16)	Inv(16)>Donors
Present in at least 1 sample	2	433	1352
	3	131	453
	4	32	158
	5	18	107
Present and changing	2	424	1283
	3	130	446
	4	32	158
	5	18	107

Table 8. The number of probe sets identified as having a greater expression in inv(16) AML patients than in non-inv(16) patients or normal healthy donors. Fold change filtering was applied to both sets of probe sets generated for quality control purposes as described in section 3.3.1a (Chapter 3).

The 100 probe sets represented 81 genes and the 99 probe sets identified represented 80 genes.

	Probe Set	Gene Symbol	Gene Title	Fold Change
71*	208450_at	<i>LGALS2</i>	Lectin, galactoside-binding, soluble, 2	5.059
72	204163_at	<i>EMILIN1</i>	Elastin microfibril interfacier 1	5.240
73	210139_s_at	<i>PMP22</i>	Peripheral myelin protein 22	5.296
74*	215382_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	5.322
75	214390_s_at	<i>BCAT1</i>	Branched chain aminotransferase 1, cytosolic	5.524
76*	215666_at	<i>HLA-DRB4</i>	Major histocompatibility complex, class II, DR beta 4	5.615
77*	206622_at	<i>TRH</i>	Thyrotropin-releasing hormone	5.722
78	201005_at	<i>CD9</i>	CD9 molecule	5.828
79	205239_at	<i>AREG</i>	Amphiregulin	5.858
80*	210495_x_at	<i>FN1</i>	Fibronectin 1	5.859
81	200665_s_at	<i>SPARC</i>	Secreted protein, acidic, cysteine-rich (osteonectin)	5.889
82*	210084_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	5.904
83	204661_at	<i>CD52</i>	CD52 molecule	5.964
84	221914_at	<i>SYN1</i>	Synapsin I	6.055
85	208978_at	<i>CRIP2</i>	Cysteine-rich protein 2	6.491
86	213479_at	<i>NPTX2</i>	Neuronal pentraxin II	6.530
87	203036_s_at	<i>MTSS1</i>	Metastasis suppressor 1	6.634
88	209387_s_at	<i>TM4SF1</i>	Transmembrane 4 L six family member 1	6.767
89	201324_at	<i>EMP1</i>	Epithelial membrane protein 1	6.792
90	205987_at	<i>CD1C</i>	CD1c molecule	7.030
91	205330_at	<i>MN1</i>	Meningioma (disrupted in balanced translocation) 1	7.947
92	203939_at	<i>NT5E</i>	5'-nucleotidase, ecto (CD73)	8.023
93	208789_at	<i>PTRF</i>	Polymerase I and transcript release factor	8.230
94	201325_s_at	<i>EMP1</i>	Epithelial membrane protein 1	8.544
95	209488_s_at	<i>RBPM5</i>	RNA binding protein with multiple splicing	8.791
96	212358_at	<i>CLIP3</i>	CAP-GLY domain containing linker protein 3	9.902
97	201497_x_at	<i>MYH11</i>	Myosin, heavy chain 11, smooth muscle	10.32
98	218876_at	<i>TPPP3</i>	Tubulin polymerization-promoting protein family member3	10.81
99*	204885_s_at	<i>MSLN</i>	Mesothelin	15.47
100	206135_at	<i>ST18</i>	Suppression of tumorigenicity 18 (breast carcinoma)	26.59

Table 9. The top 30 “present in at least 1 sample” and “present and changing” probe sets identified from the 305 sample dataset as upregulated in inv(16) AML patients samples when compared to non-inv(16) patients and to healthy donor samples by fold change (FC>3). The fold change value displayed is the mean fold change value of the probe sets across the inv(16) samples. The numbers assigned to the probe sets correspond to their location in figure 9. For the complete list of probe sets see appendix II. *Probe sets not identified from the “present and changing” probe sets.

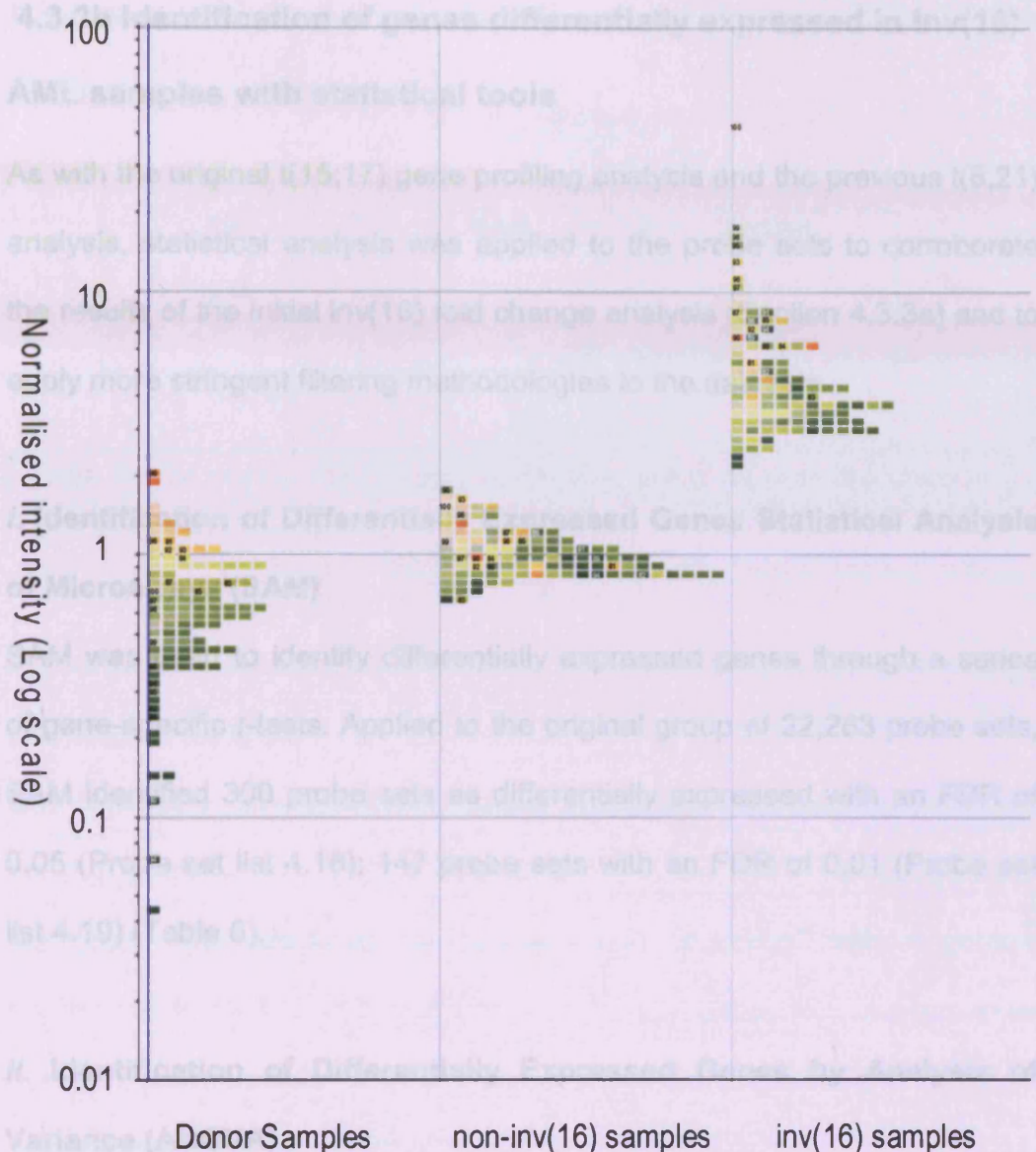


Figure 9. The relative mean expression levels of the probe sets identified, from probe sets that were classified as present in at least 1 sample (n=100) and as present and changing (n=99), as being more significantly expressed in inv(16) samples than non-inv(16) samples **and** in donor samples by ≥ 3 fold. The probe sets are coloured according to their expression levels in their respective sub-group relative to their expression level in the donor samples. The numbers on each block correspond to the probe set it is representing - see Table 5 (only top 20 probe sets numbered).

4.3.3b Identification of genes differentially expressed in inv(16)

AML samples with statistical tools

As with the original $t(15;17)$ gene profiling analysis and the previous $t(8;21)$ analysis, statistical analysis was applied to the probe sets to corroborate the results of the initial inv(16) fold change analysis (Section 4.3.3a) and to apply more stringent filtering methodologies to the datasets.

***i.* Identification of Differentially Expressed Genes Statistical Analysis of Microarrays (SAM)**

SAM was used to identify differentially expressed genes through a series of gene-specific t -tests. Applied to the original group of 22,283 probe sets, SAM identified 300 probe sets as differentially expressed with an FDR of 0.05 (Probe set list 4.18); 147 probe sets with an FDR of 0.01 (Probe set list 4.19) (Table 6).

***ii.* Identification of Differentially Expressed Genes by Analysis of Variance (ANOVA)**

ANOVA was applied to the group of probe sets classified as “present and changing” ($n=16,977$) to identify differences in gene expression levels between the inv(16) patients and the non-inv(16) patients and normal healthy donors by comparing the variances of the test groups. Using an FDR of 0.05, 7,654 probe sets (probe set list 4.20) were identified, while

5,405 probe sets (probe set 4.21) were identified with an FDR of 0.01 (Table 10).

FDR	SAM	ANOVA
0.05	300	7,654
0.01	147	5,405

Table 10. The number of probe sets identified as having a greater expression in inv(16) AML patients than in non-inv(16) patients or normal healthy donors through SAM and ANOVA analysis.

4.3.3c Identification of Differentially Expressed Genes by Fold change and Statistical tools

Combining the differentially expressed probe sets identified by SAM and ANOVA analysis with an FDR of 0.05 (n=300, n=7,654, respectively) with the probe sets identified as “present and changing” with a greater expression in t(8;21) AML patients than in any other AML sub-group or normal healthy donors (n=99), a set of 66 probe sets was generated as being specific to inv(16) samples (Probe set list 4.22; Figure 10; Table 11; Figure 11).

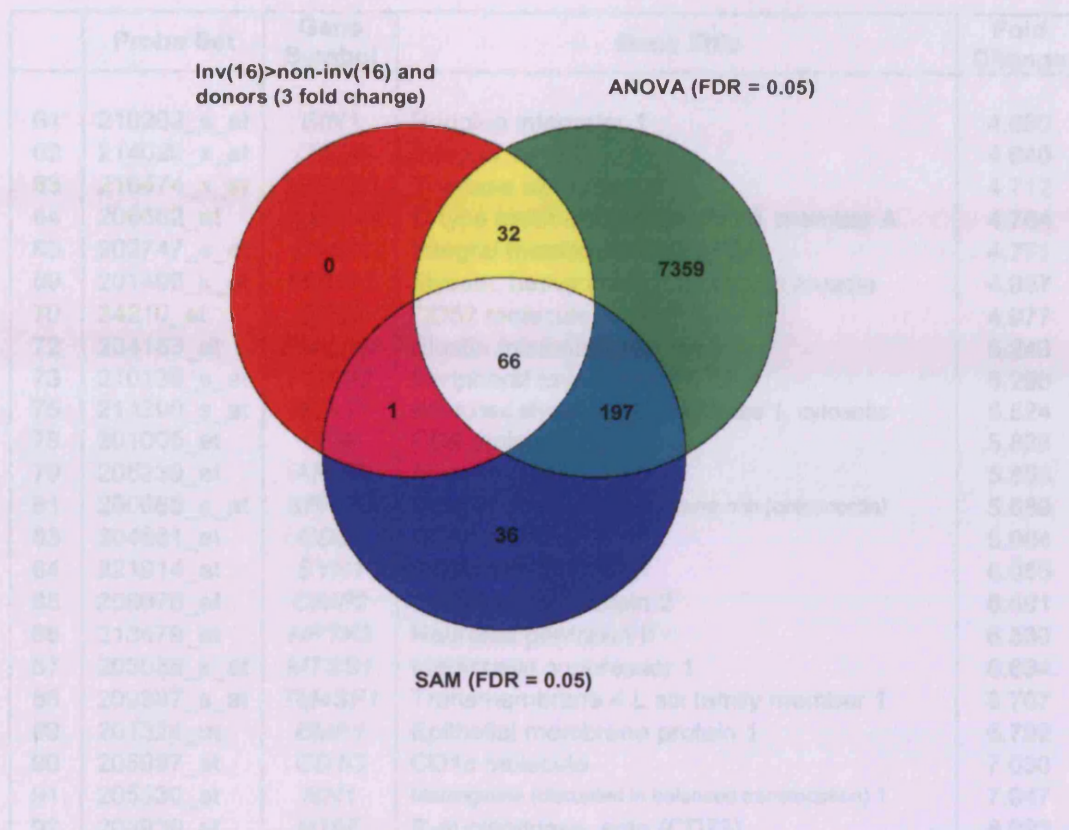


Figure 10. A combination of the probe sets identified from SAM (FDR=0.05), ANOVA testing (FDR=0.05) and having a greater level of expression in t(8;21) samples than in non-t(8;21) samples and in donor samples by 3 fold change filtering. Forty-one probe sets were identified as being specifically expressed in t(8;21) samples.

Similar to the t(8;21) AML analysis, statistical tools combined with filtering methods filtered out many probe sets identified as specific from the original fold change filtering. This could be due to a result of a greater variance in the expression levels of the probe sets amongst the inv(16) samples resulting in less probe sets originally identified achieving the required threshold expression levels of the SAM and ANOVA analysis to be identified as specific to inv(16) AML.

	Probe Set	Gene Symbol	Gene Title	Fold Change
61	210202_s_at	<i>BIN1</i>	Bridging integrator 1	4.630
62	214020_x_at	<i>ITGB5</i>	Integrin, beta 5	4.640
63	216474_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	4.712
64	206682_at	<i>CLEC10A</i>	C-type lectin domain family 10, member A	4.764
65	202747_s_at	<i>ITM2A</i>	Integral membrane protein 2A	4.771
69	201496_x_at	<i>MYH11</i>	Myosin, heavy chain 11, smooth muscle	4.967
70	34210_at	<i>CD52</i>	CD52 molecule	4.977
72	204163_at	<i>EMILIN1</i>	Elastin microfibril interfacier 1	5.240
73	210139_s_at	<i>PMP22</i>	Peripheral myelin protein 22	5.296
75	214390_s_at	<i>BCAT1</i>	Branched chain aminotransferase 1, cytosolic	5.524
78	201005_at	<i>CD9</i>	CD9 molecule	5.828
79	205239_at	<i>AREG</i>	Amphiregulin	5.858
81	200665_s_at	<i>SPARC</i>	Secreted protein, acidic, cysteine-rich (osteonectin)	5.889
83	204661_at	<i>CD52</i>	CD52 molecule	5.964
84	221914_at	<i>SYN1</i>	Synapsin I	6.055
85	208978_at	<i>CRIP2</i>	Cysteine-rich protein 2	6.491
86	213479_at	<i>NPTX2</i>	Neuronal pentraxin II	6.530
87	203036_s_at	<i>MTSS1</i>	Metastasis suppressor 1	6.634
88	209387_s_at	<i>TM4SF1</i>	Transmembrane 4 L six family member 1	6.767
89	201324_at	<i>EMP1</i>	Epithelial membrane protein 1	6.792
90	205987_at	<i>CD1C</i>	CD1c molecule	7.030
91	205330_at	<i>MN1</i>	Meningioma (disrupted in balanced translocation) 1	7.947
92	203939_at	<i>NT5E</i>	5'-nucleotidase, ecto (CD73)	8.023
93	208789_at	<i>PTRF</i>	Polymerase I and transcript release factor	8.230
94	201325_s_at	<i>EMP1</i>	Epithelial membrane protein 1	8.544
95	209488_s_at	<i>RBPM5</i>	RNA binding protein with multiple splicing	8.791
96	212358_at	<i>CLIP3</i>	CAP-GLY domain containing linker protein 3	9.902
97	201497_x_at	<i>MYH11</i>	Myosin, heavy chain 11, smooth muscle	10.32
98	218876_at	<i>TPPP3</i>	Tubulin polymerization-promoting protein family member 3	10.81
100	206135_at	<i>ST18</i>	Suppression of tumorigenicity 18 (breast carcinoma)	26.59

Table 11. The 66 probe sets identified from the 305 sample dataset as upregulated in inv(16) AML patients samples when compared to non-inv(16) patients and to healthy donor samples by fold change (FC>3) and by SAM analysis and ANOVA analysis with a FDR of 0.05. The fold change value displayed is the mean fold change value of the probe sets across the inv(16) samples. The numbers assigned to the probe sets correspond to their location in figure 11 and to their identification number in table 8.

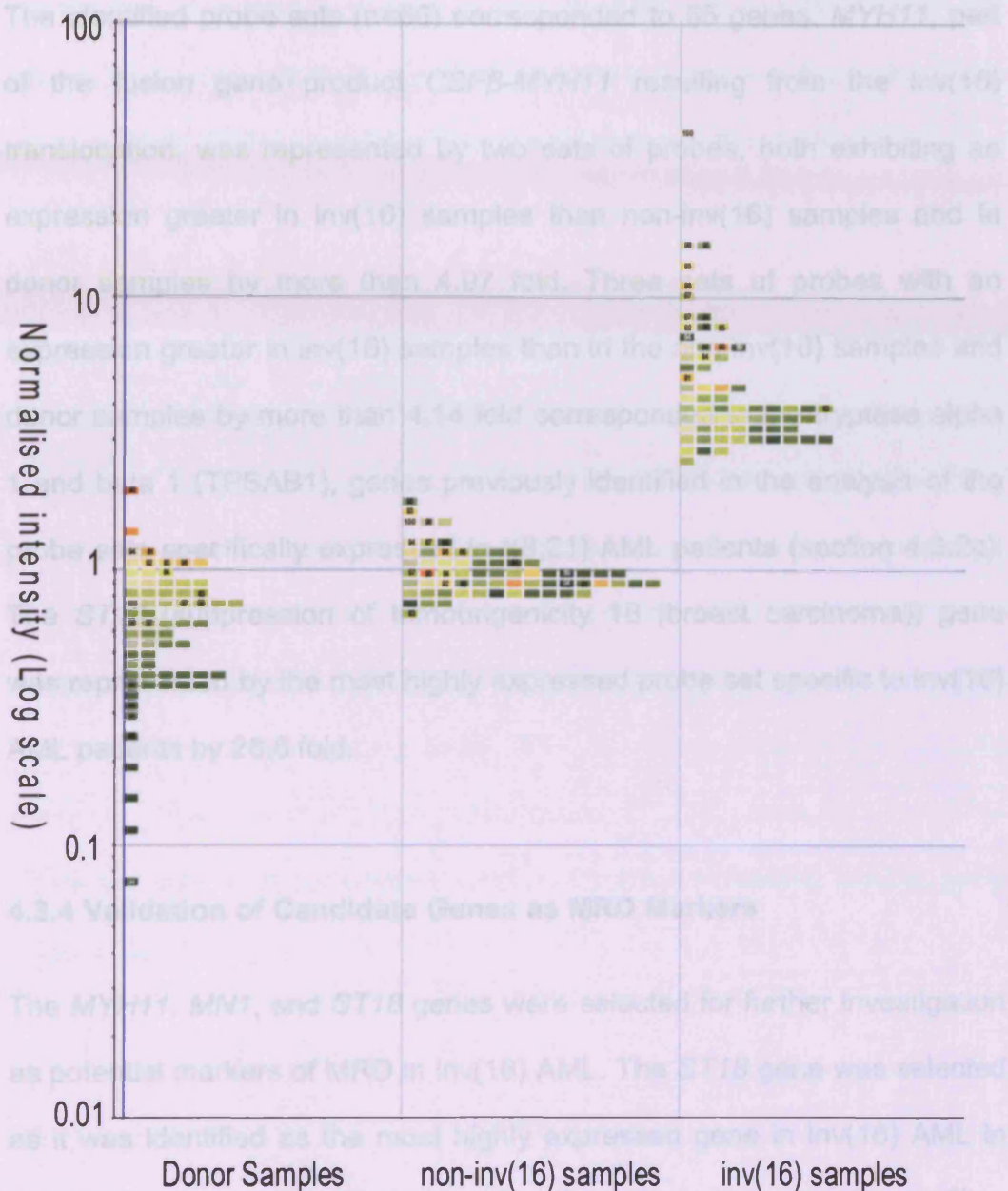


Figure 11. The relative mean expression levels of the 66 probe sets identified, from probe sets that were classified as present and changing, as being more significantly expressed in inv(16) samples than non-inv(16) samples and in donor samples by ≥ 3 fold and through SAM and ANOVA analysis with an FDR of 0.05. The probe sets are coloured according to their expression levels in their respective sub-group relative to their expression level in the donor samples. The numbers on each block correspond to the probe set it is representing - see table 7 (only top 20 genes numbered).

The identified probe sets (n=66) corresponded to 55 genes. *MYH11*, part of the fusion gene product *CBF β -MYH11* resulting from the inv(16) translocation, was represented by two sets of probes, both exhibiting an expression greater in inv(16) samples than non-inv(16) samples and in donor samples by more than 4.97 fold. Three sets of probes with an expression greater in inv(16) samples than in the non-inv(16) samples and donor samples by more than 4.14 fold corresponded to the tryptase alpha 1 and beta 1 (TPSAB1), genes previously identified in the analysis of the probe sets specifically expressed in t(8;21) AML patients (section 4.3.2c). The *ST18* (suppression of tumourigenicity 18 (breast carcinoma)) gene was represented by the most highly expressed probe set specific to inv(16) AML patients by 26.6 fold.

4.3.4 Validation of Candidate Genes as MRD Markers

The *MYH11*, *MN1*, and *ST18* genes were selected for further investigation as potential markers of MRD in inv(16) AML. The *ST18* gene was selected as it was identified as the most highly expressed gene in inv(16) AML in this study and was previously identified through gene expression profiling as having a greater expression by 50 times in a pool of 10 AML patient samples compared to pools of healthy bone marrow samples in an earlier study by Steinbach *et al* (2006) (Steinbach *et al*, 2006). The *MYH11* gene was selected for further analysis due to its representation by two probe sets identified as being more highly expressed in inv(16) AML than in non-

inv(16) AML and healthy normal donors (by 4.97 fold and 10.32 fold) and the important role it plays in inv(16) AML. *MYH11* is part of the fusion gene product *CBF β -MYH11* resulting from the inversion of chromosome 16.

4.3.4a Diagnostic Screening of Candidate Genes

Expression levels of the three genes (*MYH11*, *MN1*, and *ST18*) were measured in diagnostic inv(16) AML patient samples and compared to expression levels measured in healthy normal donor samples. Initially primer sets were designed for the amplification of the three genes as described in the Materials and Methods section 2.5.1 (Chapter 2) and using serial dilutions of cDNA, their amplification efficiency was compared to that of the house-keeping gene, *S14*. All three genes were seen to generate similar amplification efficiencies to that produced by *S14*. Therefore *S14* was considered suitable as the normalising control gene (Figure 12).

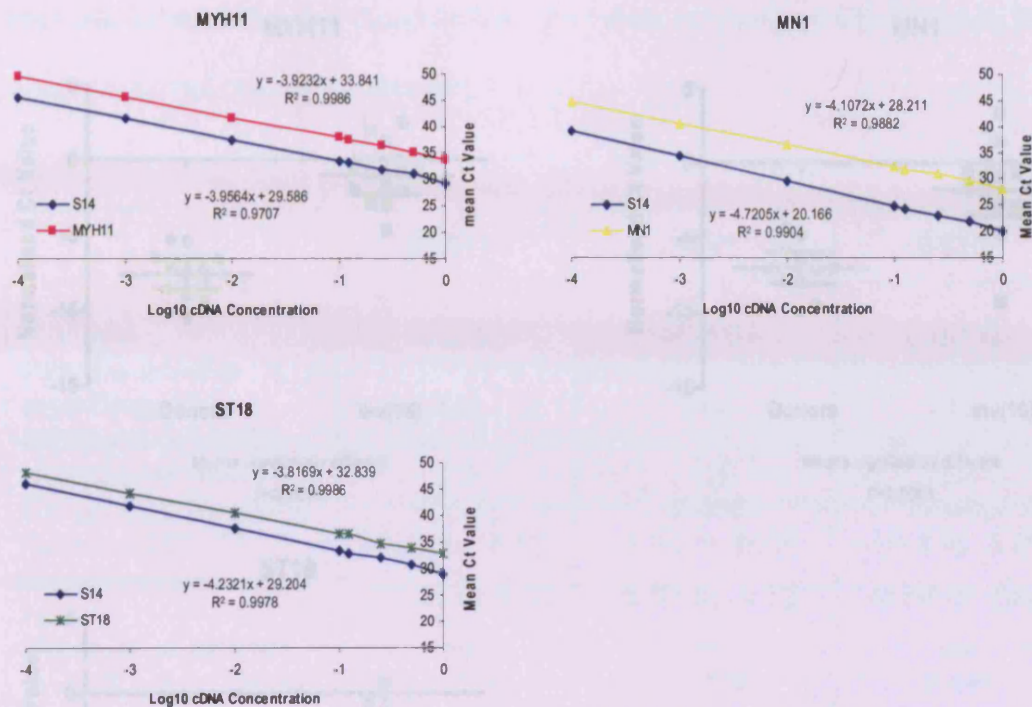


Figure 12. The amplification curves of the primer sets used for *MYH11*, *MN1* and *ST18* and of the house-keeping gene, *S14*, to compare their amplification efficiencies.

The levels of expression of *MYH11*, *MN1* and *ST18* were measured in AML diagnostic patient samples presenting with a *inv(16)* (n=13, 13 and 10, respectively) and compared to the levels of expression in normal healthy bone marrow donor samples (n=12, 12 and 5, respectively) (Figure13).

Using *t*-tests, highly significant differences were observed between the mean normalized Ct values of *MYH11* and *MN1* in the diagnostic *inv(16)* AML patient samples and mean normalized Ct values of the two genes in the healthy normal donor samples. The mean normalized Ct values of *ST18*

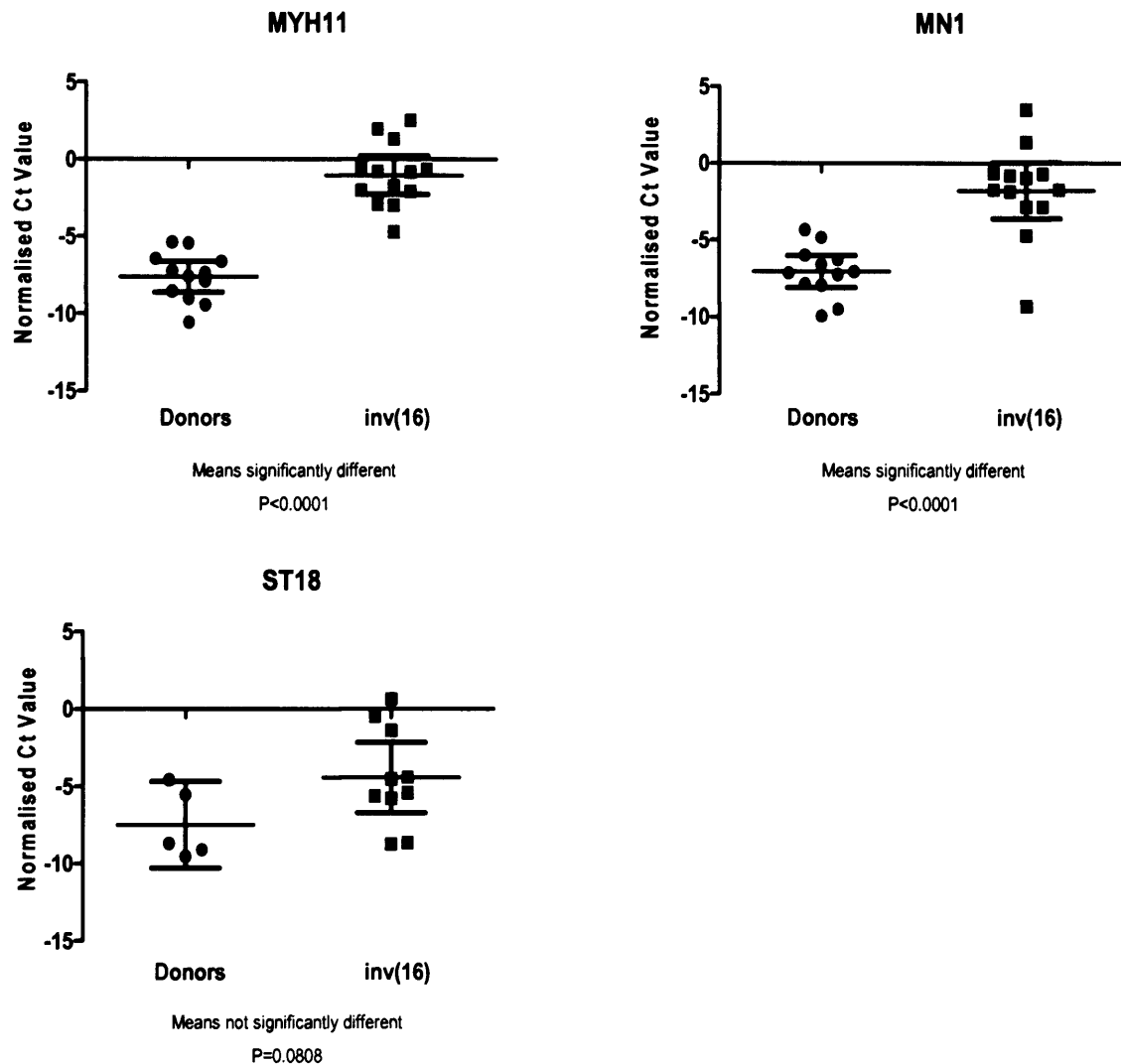


Figure 13. Scatter plots of the spread of the normalised Ct values of the candidate genes measured in the inv(16) AML diagnostic patient samples and the healthy normal donor samples. Mean value $\pm$ 95% confidence interval plotted.

Using *t*-tests, highly significant differences were observed between the mean normalised Ct values of *MYH11* and *MN1* in the diagnostic inv(16) AML patient samples and mean normalised Ct values of the two genes in the healthy normal donor samples. The mean normalised Ct values of *ST18*

was not found to be significantly from its mean normalised Ct values in the healthy normal donors (Table 12).

	MYH11	MN1	ST18
P Value	<0.0001	<0.0001	0.0808
Are the means significantly different (p<0.05)	Yes	Yes	No
Number of donors	12	12	5
Donors mean value	-7.628	-7.050	-7.490
Donors 95% CI	-8.63 to -6.63	-8.10 to -6.00	-10.3 to -4.68
Donors expression range	-10.56 to -5.37	-9.93 to -4.33	-9.54 to -548
Number of inv(16)	13	13	10
inv(16) mean value	-1.039	-1.769	-4.441
inv(16) 95% CI	-2.29 to 0.21	-3.59 to 0.06	-6.73 to -2.16
inv(16) expression range	-4.71 to 2.53	-4.68 to 3.51	-8.76 to 0.66

Table 12. The statistical significance of the mean Ct values of the candidate genes measured in the inv(16) AML diagnostic patient samples and the healthy normal donor samples normalised to the Ct values of S14. CI = confidence interval.

4.3.4b Candidate Genes as MRD Markers

Based on findings in Section 4.3B, the expression levels of *MYH11* and *MN1* were monitored in AML patient diagnostic and follow-up samples of five inv(16) AML patients to identify the potential for either gene to act as a marker of MRD for inv(16) AML.

The fold change of the genes' transcript expression levels relative to *S14* were calculated as described in Chapter 3 and compared to the upper fold

change in expression observed in the healthy normal donor samples (Figure 5). The MRD values (MRDv) of the genes in the patient samples were calculated as described in Chapter 3 and compared to the level of transcript detected at diagnosis (Figure 6).

Normalised copy number data for the fusion gene *CBFβ-MYH11* and the *WT1* gene was supplied from the Department of Clinical Haematology, Manchester Royal Infirmary for comparison with the fold change data generated for the candidate genes:

i. Patient 1

The fold change in expression of *MYH11* and *MN1* decreased by approximately 1.5 logs in patient 1 following diagnosis and induction therapy. The fold changes in expression remained at this lower level until 9.5 months following diagnosis when they began to increase, reaching a fold change in expression level of approximately 1 log higher by 12.75 months following diagnosis. Following the initial decrease after diagnosis, the levels of expression of both *MYH11* and *MN1* detected remained lower than the upper levels detected in healthy normal donors. This patient relapsed at 13 months following diagnosis (Figure 14).

ii. Patient 2

The fold change in expression of *MYH11* and *MN1* decreased by approximately 3 logs in patient 2 following diagnosis and induction therapy. The fold changes in expression of these genes remained at this lower level until 8 months following diagnosis when the fold change in expression of *MN1* began to increase, reaching a fold change in expression level of approximately 1.5 logs higher by 12 months following diagnosis. Following the initial decrease after diagnosis, the levels of expression of both *MYH11* and *MN1* detected remained lower than the upper levels detected in healthy normal donors until approximately 10 months following diagnosis when the level of *MN1* expression was higher than the upper level detected in healthy donors. This patient relapsed at 11 months following diagnosis (Figure 14).

iii. Patient 3

The fold change in expression of *MYH11* and *MN1* decreased by approximately 2 logs and 1 log, respectively, in patient 3 following diagnosis and remained between 0.5 logs and 1 log lower than the upper level detected in healthy normal donors for the duration of the disease. This patient achieved and remained in complete remission (Figure 14).

iv. Patient 4

The fold change in expression of *MYH11* and *MN1* decreased by approximately 5 logs and 2.5 logs, respectively, in patient 4 following

diagnosis. By 16.5 months following diagnosis the fold change in expression of both genes had increased to a level of approximately 1 log lower than that detected at diagnosis (Figure 14).

v. Patient 5

The fold change in expression of *MYH11* and *MN1* fluctuated throughout the disease of patient 5; both genes within 2 logs of the level of fold change detected at diagnosis. For the duration of the disease, the levels of expression of *MYH11* and *MN1* detected fluctuated within approximately 1.5 logs of the upper level detected in healthy normal donors (Figure 14).

The MRD levels of *MYH11* and *MN1* decreased between 1.5 logs and 3 logs from the level of transcript detected at diagnosis in the inv(16) patients that relapsed; the corresponding normalised number of *WT1* and *CBFβ-MYH11* transcripts decreased by between 2.5 logs and 5 logs following diagnosis. An increase in the MRD level of *MYH11* and *MN1* were detected at least 3 months prior to haematological relapse in the two patients; by the time of relapse the MRD levels of both genes had increased between 1 and 2 logs. The normalised number of *WT1* and *CBFβ-MYH11* transcripts also began to increase 3 months prior to haematological relapse in patient 2 but an increase was detected in patient 1 8 months prior to the haematological relapse; by the time of haematological relapse the normalised number of

WT1 and *CBF β -MYH11* transcripts had increased by between 2 and 5 logs (Figures 15 and 16).

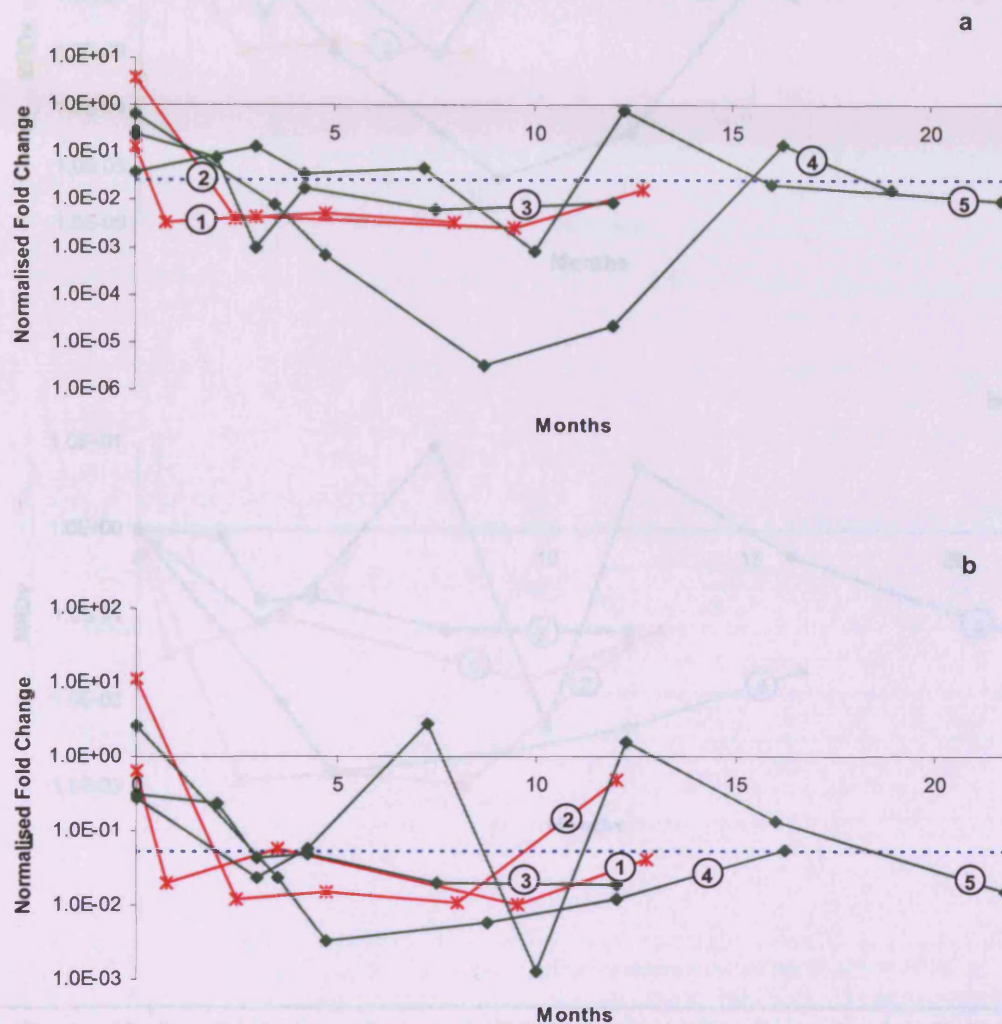


Figure 14. Longitudinal monitoring of *MYH11* (a) and *MN1* (b) in five *inv(16)* AML patients. The graphs show the fold change values generated for respective genes in each sample, relative to the *S14* gene. Red lines - patients who relapsed. Green lines - patients who remained in complete remission (CR). The blue dashed line represents the upper limit of gene expression in the healthy normal donor samples. The numbers in the circles indicate the patient number.

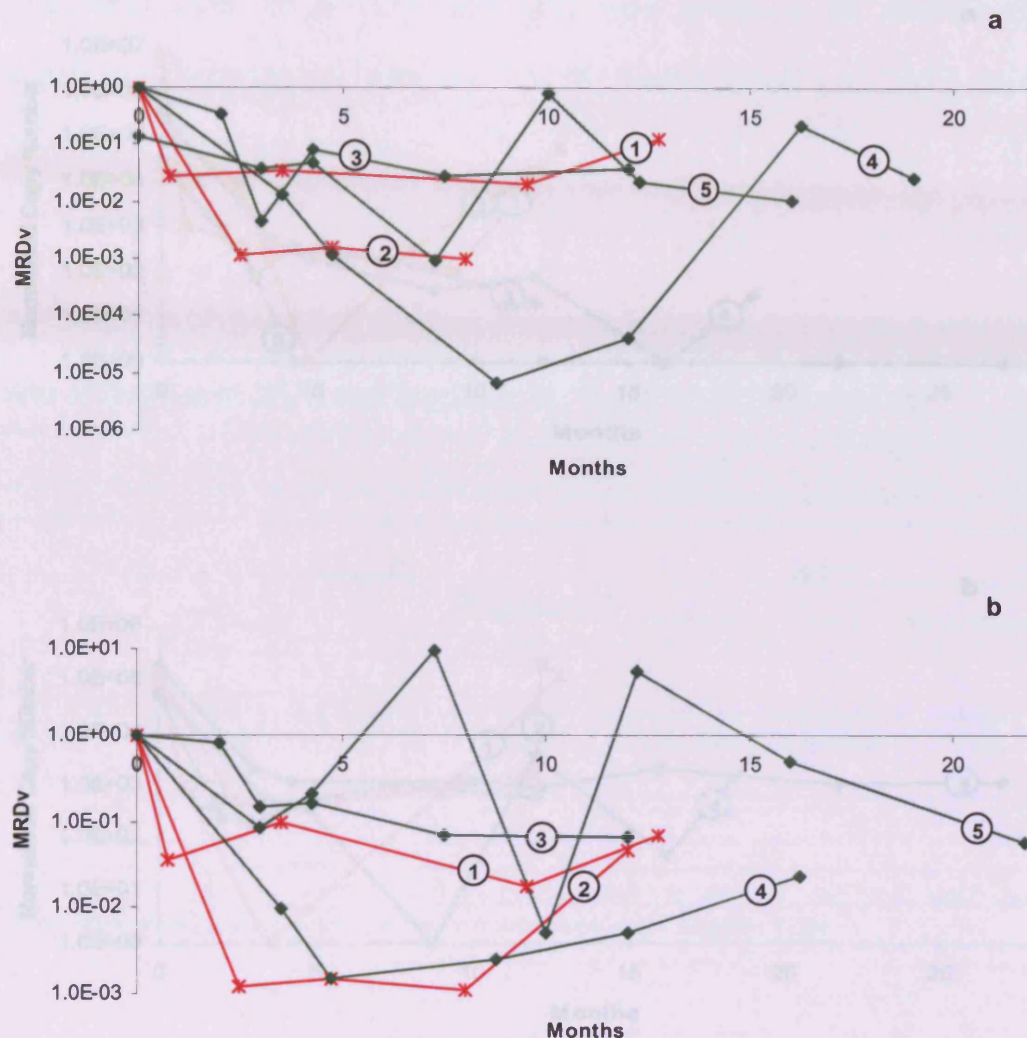


Figure 15. Longitudinal monitoring of *MYH11* (a) and *MN1* (b) in five in(16) AML patients. The graphs show the MRD values calculated for the genes in each sample. Red lines - patients who relapsed. Green lines - patients who remained in complete remission (CR). The numbers in the circles indicate the patient number.

4.3.5 Correlation between Candidate Markers and Current Markers

The MRQ data for *MYH11* and *WT1* was analysed for stability

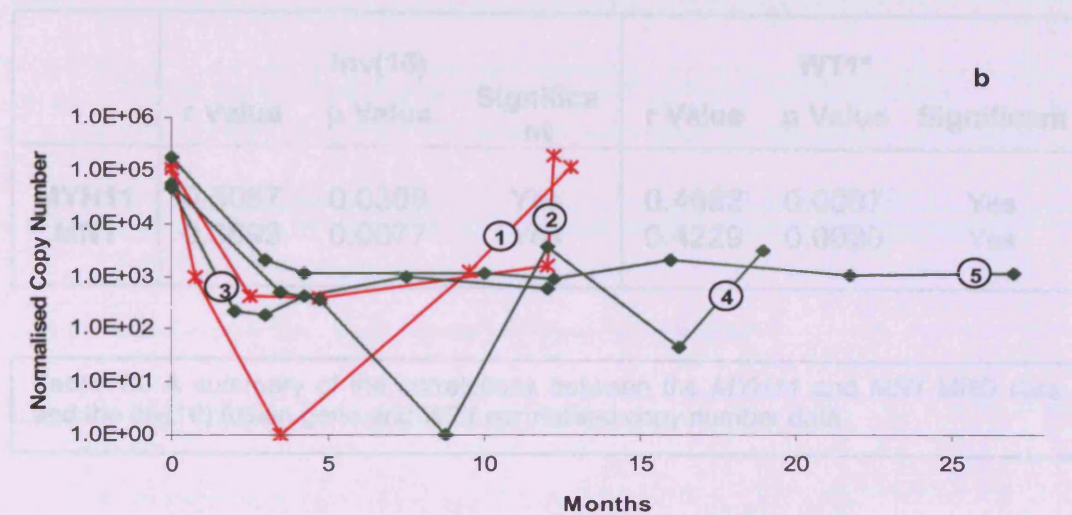
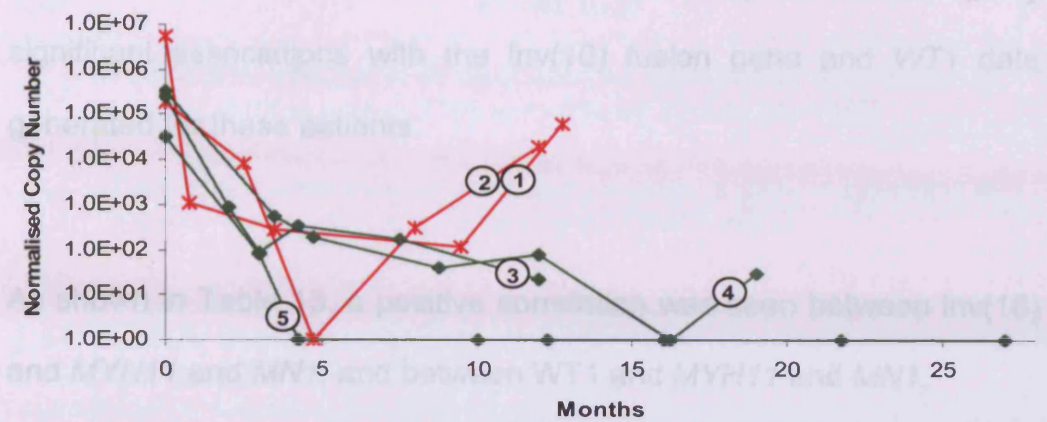


Figure 16. Longitudinal monitoring of the inv(16) fusion gene *CBFβ-MYH11* (a) and *WT1* (b) in five inv(16) AML patients. The graphs show the normalised transcript numbers calculated for the genes in each sample. Red lines - patients who relapsed. Green lines - patients who remained in complete remission (CR). The numbers in the circles indicate the patient number.

4.3.5 Correlation between Candidate Markers and Current Markers

The MRD data for *MYH11* and *MN1* was analysed for statistically significant associations with the *inv(16)* fusion gene and *WT1* data generated for these patients.

As shown in Table 13, a positive correlation was seen between *inv(16)* and *MYH11* and *MN1*, and between *WT1* and *MYH11* and *MN1*.

	Inv(16)			WT1*		
	r Value	p Value	Significant	r Value	p Value	Significant
MYH11	0.3087	0.0309	Yes	0.4692	0.0007	Yes
MN1	0.3693	0.0077	Yes	0.4229	0.0020	Yes

Table 13. A summary of the correlations between the *MYH11* and *MN1* MRD data and the *inv(16)* fusion gene and *WT1* normalised copy number data.

List	Probe Sets
4.12	Probe sets present in at least sample with greater expression in inv(16) than in non-inv(16) samples by ≥ 3 fold (n=131)
4.13	Probe sets present in at least sample with greater expression in inv(16) than in donor samples by ≥ 3 fold (n=453)
4.14	Probe sets present in at least sample with greater expression in inv(16) than in non-inv(16) samples and in donor samples by ≥ 3 fold (n=100)
4.15	Probe sets present and changing with greater expression in inv(16) than in non-inv(16) samples by ≥ 3 fold (n=94)
4.16	Probe sets present and changing with greater expression in inv(16) than in donor samples by ≥ 3 fold (n=446)
4.17	Probe sets present and changing with greater expression in inv(16) than in non-inv(16) samples and in donor samples by ≥ 3 fold (n=99)
4.18	SAM analysis of all genes, FDR=0.05 (n=300)
4.19	SAM analysis of all genes, FDR=0.01 (n=147)
4.20	ANOVA analysis of present and changing probe sets, FDR=0.05 (n=7,654)
4.21	ANOVA analysis of present and changing probe sets, FDR=0.01 (n=5,405)
4.22	Probe sets common to present and changing probe sets with greater expression in inv(16) than in non-inv(16) samples and in donor samples by ≥ 3 fold, SAM analysis of all probe sets using an FDR of 0.05 and ANOVA analysis of probe sets present and changing using an FDR of 0.05 (n=66)

Table 14. The probe set lists identified from the gene expression profiling analysis (Section 4.3.3).

4.4 Discussion

Core binding factor (CBF) leukaemia results from chromosomal aberrations involving either the CBF α subunit or the CBF β subunit (Appelbaum *et al*, 2006). Patients present with either a t(8;21) or an inv(16) translocation and are seen to have a relatively favourable prognosis (Grimwade *et al*, 1998). The resulting fusion gene products from these translocations, the *AML1-ETO* gene product generated from t(8;21) translocation and the *CBF β -MYH11* gene product generated from the inv(16), have been suggested as markers to monitor minimal residual disease (MRD) in the t(8;21) and inv(16) patients (Claxton *et al*, 1994). However, controversy has surrounded the use of these fusion genes products with conflicting results about the detection of these fusion products in some patients. Several groups were able to detect *AML1-ETO* and *CBF β -MYH11* transcripts using highly sensitive RT-PCR methods in some patients in long-term remission (Guerrasio *et al*, 1995; Tobal *et al*, 1995; Jurlander *et al*, 1996). The Wilms' tumour 1 (*WT1*) gene is often used as an MRD marker in AML patients but similar conflicting results regarding its expression in patients in remission has clouded the significance of its clinical use (Gaiger *et al*, 1998; Elmaagacli *et al*, 2000). Therefore gene expression profiling was applied to a cohort of diagnostic AML patient samples and donor samples to identify potential MRD markers specific to t(8;21) patients and to inv(16) AML patients.

4.4.1. t(8;21) AML

Gene expression profiling was applied to the expression data of 290 patients, 15 (5.2%) of which presented with a t(8;21) at diagnosis. By applying fold change calculations and statistical tools to the probe sets in the dataset identified as present in at least two samples and changing in expression by more than two fold, 41 probe sets were identified as being highly expressed in t(8;21) patients compared to healthy normal donors. The ten most highly expressed probe sets were more highly expressed in t(8;21) AML than in non-t(8;21) AML and healthy donors by a fold of between 13.25 and 39.27 and represented six genes; *TRH* (thyrotropin-releasing hormone), *POU4F1* (POU class 4 homeobox 1), *IL5RA* (interleukin 5 receptor, alpha), *TPSAB1* (tryptase alpha/beta 1), *ETO* (runt-related transcription factor 1) and *TPSB2* (tryptase alpha/beta 2). The identification of several of these genes as t(8;21)-specific is supported by the analysis of gene expression data from several other studies: Ichikawa *et al* (2006) analysed the gene expression profiles of leukaemic cells in 50 paediatric AML patients using microarrays and identified genes that were up- and down-regulated in the CBF sub-groups. As with the current study, the *ETO* gene and the *POU4F1* gene were identified as highly expressed t(8;21)-specific genes. The *TRH* gene and the *TPSAB1* gene were also identified as highly expressed genes but specific to both t(8;21) and inv(16) AML (Ichikawa *et al*, 2006). Using SAM analysis on 93 CBF AML samples, Bullinger *et al* (2007) identified more

than 1000 genes that significantly correlated with the t(8;21) and inv(16) CBF groups. Genes defining the t(8;21) sub-group included *POU4F1* and *TRH* (Bullinger *et al*, 2007). Steinbach *et al* (2006) also used microarrays to compare gene expressions profiles of healthy bone marrow CD34⁺ haematopoietic stem cells and 52 paediatric AML patients. Whilst this study was not limited to CBF AML, 23% presented with a CBF translocation at diagnosis. The analysis identified 158 genes as potential MRD markers, including *PRAME*, *POU4F1* and *TRH*; genes identified as t(8;21)-specific in this present study (Steinbach *et al*, 2006). Bullinger *et al* (2004) and Valk *et al* (2004) both identified *ETO* as discriminative to t(8;21) AML in studies using gene expression profiling to identify prognostically useful AML sub-groups (Bullinger *et al*, 2004;Valk *et al*, 2004). A study by Ross *et al* (2004) explored the concept of gene expression profiling to identify prognostically useful sub-groups specifically in paediatric AML as previously seen for adult AML. Using SAM analysis they also identified genes discriminative to t(8;21) AML. From their list of top 100 probe sets, 18 were identified in this current study as t(8;21)-specific. Several of these probe sets represented the genes *PRAME*, *POU4F1*, *ETO* and *IL5RA* (Ross *et al*, 2004).

Based on their level of expression from the microarray analysis and previous associations with AML, four genes were selected for validation as MRD markers; *PRAME*, *POU4F1*, *ETO* and *IL5RA*. Using RQ-PCR, the

normalised expression of the four genes relative to a house-keeping gene *S14* were measured in a group of diagnostic t(8;21) AML patients and a group of healthy normal donors. Student *t*-tests identified a significant difference in the expression of the genes between the two test groups. The work previously mentioned by Steinbach *et al* (2006) further supports these findings with their RQ-PCR analysis of 65 of the 200 genes they previously identified through microarray analysis. They identified 12 genes that were more highly expressed by more than 50 times in one of three diagnostic AML groups (myelocytic AML (FAB M1/2), myelomonocytic AML (FAB M4) or monocytic AML (FAB M5)) compared to in healthy bone marrow CD34⁺ haematopoietic stem cells. Included in this group of genes were *PRAME* and *POU4F1* (Steinbach *et al*, 2006). Similarly Paydas *et al* (2005) detected significant levels of *PRAME* expression in AML patient samples compared to the very low levels detected in healthy individuals (2 copies/10,000 cells) through RQ-PCR of peripheral blood samples (Paydas *et al*, 2005). Van Baren *et al* (1998) observed preferential expression of *PRAME* in AML-M2 and AML-M3 patients, the AML FAB groups associated with the t(8;21) and t(15;17), respectively, compared to healthy normal donors and other FAB groups (van Baren *et al*, 1998). Qin *et al* (2009) investigated the expression levels of both *WT1* and *PRAME* in 204 diagnostic AML samples. They detected the highest levels of *PRAME* expression in FAB M2 AML patients with *AML1-ETO* and the second highest levels in APL patients with *PML-RAR α* compared to other FAB

group patients. Minimal levels of *PRAME* expression were detected in normal bone marrow and peripheral blood samples (Qin *et al*, 2009). A very early study by Nishii *et al* (1995) used Northern-blot hybridisation and reverse transcriptase PCR (RT-PCR) to detect the expression of *IL5RA* in a cohort of 35 AML patients. Within this sample group, *IL5RA* was detected in 17 patients; all nine FAB M2 AML patients with a t(8;21) were positive for *IL5RA* expression, as were two patients with structural abnormalities of chromosome 16. Only one of five M2 patients without a t(8;21) and six of 24 patients lacking a t(8;21) or structural abnormalities of chromosome 16 were also positive for *IL5RA* (Nishii *et al*, 1995). Whilst no healthy normal samples were included in this study to compare the expression of *IL5RA* with, it does highlight the specificity of *IL5RA* expression to CBF AML, in particular t(8;21), and supports the significant expression of *IL5RA* detected in the diagnostic t(8;21) samples compared to the healthy normal donors in this current study.

The fold change in expression and the MRD values of the four genes were followed in the diagnostic and several follow-up samples of three t(8;21) AML patients. The MRD levels of all four genes decreased by at least one log from the level of transcript detected at diagnosed and remained at this lower level in the t(8;21) patient that achieved CR and one t(8;21) patient that relapsed. An increase in MRD level of more than 1 log over serial samples was observed for *PRAME*, *POU4F1* and *ETO* several months

prior to haematological relapse for this relapsing patient; an increase in *IL5RA* was not detected till a haematological relapse occurred. The MRD levels of the four genes provided little information for the other relapsing t(8;21) patient; an increase in MRD of *PRAME* and *ETO* of less than one log was observed one month prior to haematological relapse. Whilst a positive correlation was observed between this data and the normalised transcript data of the t(8;21) fusion gene, it was apparent the t(8;21) fusion gene provided more useful information regarding the progression of the patient's disease than the four candidate genes. A significant increase in the normalised number of *AML1-ETO* transcripts was observed in both relapsing patients; a three log increase from the level detected at least four months prior to when the haematological relapse occurred. Unlike the prognostic value of the candidate genes displayed in this current study, other groups have shown there to be more potential in the *PRAME* gene acting as an MRD marker. Paydas *et al* (2005) previously monitored *PRAME* in 15 cases of AML and detected decreased level of *PRAME* at remission and increased levels at relapses (Paydas *et al*, 2005). When monitoring *PRAME* in 84 serial samples collected from 14 t(8;21) M2 AML patients, Qin *et al* (2009) observed a good correlation between *PRAME* and *AML1-ETO* transcript levels in patients expressing *AML1-ETO* (Qin *et al*, 2009). The *POU4F1* gene was identified by gene expression microarrays and validated by RQ-PCR in a previous study by Steinbach *et al* (2006) but further analysis to identify candidate MRD markers removed

POU4F1 from the selection and the use of *POU4F1* as an MRD marker has not as yet been investigated by other groups. As the *ETO* gene forms one half of the fusion gene resulting from t(8;21), the potential of *ETO* as an MRD marker has not yet been investigated; instead the use of *AML1-ETO* as an MRD marker has been explored. The primers used in the current study for the amplification of *ETO* did not cover the breakpoint region of the *ETO* gene involved in the translocation of chromosomes 8 and 21. Therefore the *ETO* transcripts detected may have actually been from *AML1-ETO* transcripts and therefore do not properly represent the behaviour of the *ETO* gene in t(8;21) AML patients. This problem could be solved by designing primers to cover the region involved in the translocation so only *ETO* transcripts are detected rather than *AML1-ETO* transcripts.

Gene expression profiling was able to identify genes specifically expressed in t(8;21) AML that have previously been identified in other gene expression profiling studies. Further investigation of the four genes *PRAME*, *POU4F1*, *ETO* and *IL5RA* by RQ-PCR corroborated the microarray results by detecting significantly high levels of expression of these genes in diagnostic t(8;21) AML compared to healthy normal donors. Unfortunately the detection of these genes in patient follow-up samples displayed their limitations to act as markers of MRD. Whilst some degree of prognostic information was obtained by the monitoring of these genes,

the monitoring of the markers currently used, *WT1* and *AML1-ETO*, provided more sensitive and reliable information. However, this could be a result of the lack of standardisation and optimisation of the assays used in this current study compared to the assays used by the Department of Clinical Haematology, Manchester Royal Infirmary. Further standardisation and optimisation of assays for the detection of the genes may provide more insightful results into the true potential of the genes. Various studies have already shown the strong potential of *PRAME* gene to act as a marker for t(8;21) AML. The investigation in *ETO* as a marker was limited by the design of primers for also allowing the potential detection of *AML1-ETO* transcripts. It would be insightful to see where the sequences of the probe sets representing *ETO* identified in the microarray analysis were located in relation to the translocation breakpoint to know whether they were detecting the expression of *ETO* or *AML1-ETO* in the t(8;21) patients and if the monitoring of *ETO* alone could help foresee haematological relapses. If so, *ETO* may also act as a useful marker for AML patients not expressing *AML1-ETO*.

4.4.2. inv(16) AML

Gene expression profiling was applied to the expression data of 290 patients, 18 (6.2%) of which presented with an inversion of chromosome 16 at diagnosis. By applying the fold change calculations and statistical tools described for the t(8;21) patients, 66 probe sets were identified as

being highly expressed in inv(16) patients at diagnosis compared to healthy normal donors. The five most highly expressed probe sets were more highly expressed in inv(16) AML than in healthy donors by a fold of between 8.79 and 26.59 and represented five genes; *ST18* (Suppression of tumourigenicity 18), *TPPP3* (Tubulin polymerization-promoting protein family member 3), *MYH11* (Myosin, heavy chain 11, smooth muscle), *CLIP3* (CAP-GLY domain containing linker protein 3) and *RBPMS* (RNA binding protein with multiple splicing). Both Bullinger *et al* (2004) and Valk *et al* (2004) were able to identify specific gene signatures for the AML cytogenetic sub-classes, including the inv(16) sub-group, using gene expression microarrays. Within the gene signatures/clusters generated for AML patients exhibiting an inv(16), *MYH11*, a gene involved in the fusion gene created from the inversion of chromosome 16, was identified as discriminative to that particular sub-group (Bullinger *et al*, 2004;Valk *et al*, 2004). As mentioned previously, Ichikawa *et al* (2006) identified common gene expression signatures in paediatric CBF leukaemic samples. Ichikawa *et al* (2006) also identified *MYH11* as a highly expressed inv(16)-specific gene (Ichikawa *et al*, 2006). In agreement with a later study by Bullinger *et al* (2007), the *SPARC* (secreted protein, acidic, cysteine-rich) gene was among the genes identified as being more highly expressed in inv(16) AML than in healthy donors by 5.9 fold. Bullinger *et al* (2007) used SAM analysis to identify genes significantly correlated with CBF leukaemia; high-level expression of *SPARC* was found in cases with inv(16) (Bullinger

et al, 2007). In this current study the *MN1* (Meningioma 1) gene was identified as the 9th most highly expressed gene in inv(16) AML by a fold of 7.95. Ichikawa *et al* (2006) also found a high expression of this gene in their gene expression profiling study but they identified it as common to both t(8;21) and inv(16) AML (Ichikawa *et al*, 2006). Whilst not a study restricted to CBF leukaemia, Steinbach *et al* (2006) identified *ST18* as a potential MRD marker for AML patients using Affymetrix human U133A chips (Steinbach *et al*, 2006). In the study previously mentioned by Ross *et al* (2004), SAM analysis was used to identify genes discriminative to the paediatric AML sub-groups. From the list of probe sets identified for inv(16) AML, 12 were identified in this current study also as specific to inv(16) AML; *MYH11*, *MN1* and *ST18* were represented by several of the probe sets identified (Ross *et al*, 2004).

Based on their level of expression from the microarray analysis and previous associations with AML, three genes were selected for validation as MRD markers; *MYH11*, *MN1*, and *ST18*. As previously seen with the t(8;21) genes, RQ-PCR was used to detect the levels of these three genes in diagnostic inv(16) AML patients and normal healthy donors. A significant difference in the mean expressions of *MYH11* and *MN1* in the inv(16) AML patient group and the healthy normal donor group was observed ($p < 0.0001$) but not for *ST18* ($p = 0.0808$). In disagreement with the results for *ST18*, Steinbach *et al* (2006) followed up on the identification of *ST18*

through microarray by validating its expression in two pools of AML samples and two pools of healthy bone marrow samples by RQ-PCR. *ST18* expression was detected in the AML pools at a level 50 times greater than in the healthy bone marrow pools (Steinbach *et al*, 2006). However these findings would be influenced by the mix of cytogenetics within the AML pools tested whereas the AML samples study in this study were known to all exhibit an inv(16) and arrayed separately. Carella *et al* (2007) detected levels of *MN1* in inv(16) M4 AML samples at significantly higher level than in normal bone marrow samples by QRT-PCR. They also observed the levels of *MN1* mRNA detected were far greater than determined by microarray analysis (Carella *et al*, 2007).

The fold change in expression and the MRD values of *MYH11* and *MN1* were followed in the diagnostic and several follow-up samples of five inv(16) AML patients. Both genes displayed a decrease in expression fold change following diagnosis to levels lower than the upper limit detected in healthy normal donors in the five patients. The levels of MRD remained fairly consistent during the disease of the patients that achieved a complete remission but began to rise in the two patients that relapsed approximately three months prior to a haematological relapse and a positive correlation was seen between the MRD data for *MYH11* and *MN1* and the *WT1* and *CBFβ-MYH11* transcript data. In some of the patients the decrease and increase in *WT1* and *CBFβ-MYH11* transcripts was greater

by several logs than that detected for the *MYH11* and *MN1* MRD level suggesting the markers currently used are more sensitive and provide more reliable information in regards to the patient's disease. There are currently no studies into the use of *MYH11* as an MRD marker in inv(16) AML. As it forms one half of the resulting fusion gene *CBFβ-MYH11* studies have understandably concentrated on the inv(16) fusion gene as an MRD marker instead (Marcucci *et al*, 2001;Guerrasio *et al*, 2002;Perea *et al*, 2006;Lane *et al*, 2008). The inversion of chromosome 16 results in the fusion of the majority of the *CBFβ* gene and the tail domain of the *MYH11* gene; Liu *et al* (1993) identified a single breakpoint located close to the 3' end of the coding region and 3 breakpoints in the *MYH11* gene (Liu *et al*, 1993). The primers used in the current study for *MYH11* did not cover any of the three breakpoints previously identified. As a result the *MYH11* transcript supposedly detected by RQ-PCR had the potential to be a result of the detection of *CBFβ-MYH11* mRNA rather than *MYH11* mRNA alone. Therefore, as with the analysis of *ETO* in the t(8;21) patients, the analysis of *MYH11* may not be a true representation of the potential of *MYH11* to act as an MRD marker. To overcome this problem the primers designed for the amplification of *MYH11* need to stretch across the region disrupted through the inversion of chromosome 16; a potential problem due to the option of three breakpoints.

Gene expression profiling was able to identify genes specifically expressed in inv(16) AML that have previously been identified in other gene expression profiling studies, as for t(8;21) AML. Further investigation in the three genes *MYH11*, *MN1* and *ST18* by RQ-PCR identified *MYH11* and *MN1* displaying a significant level of expression in diagnostic inv(16) patients compared to healthy normal donors but *ST18* did not despite being represented by the probe set identified in the initial microarray analysis as the most highly expressed probe set in inv(16) AML compared to healthy normal donors by a fold of 26.59. As with the markers currently used for t(8;21) monitoring, the inv(16) fusion gene and *WT1* provided more sensitive and reliable information in some of the patients. Further standardisation and optimisation of assays for the detection of the genes may also prove to be more insightful results in this AML sub-group also.

As for *ETO*, the investigation in *MYH11* as a marker was limited by the design of primers allowing for the detection of *CBFβ-MYH11* transcripts also. Further investigation involving primers covering the *MYH11* breakpoints would help solve this problem but with multiple breakpoints it may be necessary to generate multiple primers for individual cases. Finally a much larger patient cohort is required to conclusively determine how well these genes monitor MRD and the clinical significance of their behaviour during the disease.

Chapter 5

Gene Expression profiling of

Normal Karyotype AML

5.1 Introduction

5.1.1 Cytogenetics

Cytogenetic aberration is recognised as a very important factor in the prognosis of acute myeloid leukaemia (AML) and is used routinely in clinical practice to predict a patient's outcome (Grimwade *et al*, 1998; Mrozek *et al*, 2004). Newly diagnosed AML patients exhibiting a translocation between chromosome 8 and chromosome 21, a translocation between chromosome 15 and chromosome 17 or an inversion of chromosome 16 are classified as having a favourable prognosis. Various studies have identified patients bearing these translocations to have low rates of primary drug resistance, greater complete remission (CR) rates compared to those with normal cytogenetics and greater overall survival (OS) associated with a reduced relapse risk (RR) (Grimwade *et al*, 1998; Byrd *et al*, 2002). AML patients with a normal karyotype belong to the largest cytogenetic group, accounting for approximately 45% of newly diagnosed *de novo* AML patients (Baldus *et al*, 2007). These patients are placed in the intermediate risk group because their CR rates and survival probabilities are lower than those with a t(8;21), t(15;17) or an inv(16) (Grimwade *et al*, 1998; Byrd *et al*, 2002). Despite having no detectable chromosomal abnormality following standard metaphase analysis, this subset of patients is extremely heterogeneous at the molecular level; several gene mutations and changes in gene expression have been identified in normal karyotype AML patients that have been shown to have an effect on the prognosis and clinical outcome of the patient's

disease (Mrozek *et al*, 2007b). An internal tandem duplication (ITD) of the fms-like tyrosine kinase 3 (*FLT3*) gene is a commonly seen genetic alteration in AML, occurring in ~30% of AML cases (Kiyoi *et al*, 1999;Thiede *et al*, 2002;Mrozek *et al*, 2007b). In addition to the ITD, point mutations in the activation loop of the tyrosine kinase domain (TKD) of the *FLT3* gene have been reported in ~7% of AML patients (Yamamoto *et al*, 2001;Thiede *et al*, 2002). Resulting in constitutive activation of the FLT3 kinase, an ITD of the *FLT3* gene is associated with a decreased overall survival (OS) in normal karyotype patients (Kottaridis *et al*, 2001;Whitman *et al*, 2001;Bienz *et al*, 2005). To summarise, AML patients under the age of 60 years presenting with *FLT3*-ITD confer a poorer prognosis (Abu-Duhier *et al*, 2000). A mutation in exon 12 of the nucleophosmin (*NPM1*) gene results in the abnormal localisation of the nucleophosmin protein in the cytoplasm (Falini *et al*, 2005;Suzuki *et al*, 2005). The incidence of *NPM1* mutations in normal karyotype patients is approximately 40-50% and they are associated with significantly higher complete remission (CR) rates (Falini *et al*, 2005;Suzuki *et al*, 2005;Schnittger *et al*, 2005;Mrozek *et al*, 2007a). Approximately 40% of normal karyotype AML patients harbouring an *NPM1* mutation also harbour a *FLT3*-ITD; in this context the favourable prognosis of the *NPM1* mutation is lost (Schnittger *et al*, 2005;Dohner *et al*, 2005;Thiede *et al*, 2006).

5.1.2 Minimal Residual Disease Markers

Many studies have assessed the usefulness of monitoring minimal residual disease (MRD) in AML patients to improve their treatment regimes but these are limited to the fusion gene products resulting from the favourable translocations detected through RQ-PCR in those “favourable risk” patients (Grimwade & Lo Coco, 2002; Guerrasio *et al*, 2002). Those patients with a “normal” karyotype currently lack a cytogenetic lesion that can be monitored for the purpose of MRD. Despite this, some genes are currently being investigated for and are currently used for the monitoring of MRD in patients with a favourable prognosis and those that are cytogenetically normal but controversy surrounds their use as markers. One such gene is Wilms’ Tumour 1 (*WT1*), a tumour-suppressor located on chromosome 11 that was originally identified for its role in Wilms’ tumour, a tumour of the kidneys (Call *et al*, 1990). In normal haematopoietic cells *WT1* has been shown to have a significantly lower level of expression compared to that observed in leukaemic cells at diagnosis. Furthermore, an increase in *WT1* expression following treatment has been associated with relapse leading *WT1* to be suggested as a good candidate for MRD monitoring (Inoue *et al*, 1997) (Cilloni *et al*, 2002). Despite many studies agreeing that *WT1* has the potential to act as an MRD marker, others disagree. Gaiger *et al* (1998) observed no difference in disease-free survival (DFS) and survival from remission in *WT1*-positive and *WT1*-negative AML patients in CR following treatment whilst Elmaagacli *et al* (2000) observed that following allogenic bone marrow (BM) or peripheral blood

stem cell (PBSC) transplant, *WT1* was not reliable as an MRD marker or as a predictor of relapse (Gaiger *et al*, 1998) (Elmaagacli *et al*, 2000). Mutations in exon 12 of the *NPM1* gene have been reported as the most common genetic abnormality occurring in *de novo* normal karyotype AML patients at diagnosis (Boissel *et al*, 2005;Falini *et al*, 2007). As a result of these findings many groups have investigated the potential of the *NPM1* gene as a marker of MRD for long-term monitoring. Sensitive RQ-PCR assays have been developed that have demonstrated the reliability of *NPM1* gene mutations as MRD markers during long-term follow-up periods in the majority of normal karyotype AML patients (Gorello *et al*, 2006;Chou *et al*, 2007). However, other studies have reported a loss of *NPM1* mutations in relapsed AML patients and an AML patient presenting with clonal evolution at relapse questioning the long-term stability and reliability of *NPM1* as a marker of MRD (Suzuki *et al*, 2005;Papadaki *et al*, 2009).

5.1.3 Gene Expression Profiling

Transcriptomics is a widely used molecular approach allowing the simultaneous analysis of thousands of genes and allowing for the identification of changes in mRNA level between experimental conditions. Various methods are used by researchers for the selection of differentially expressed genes to establish different behaviours between control and diseased/treatment situations (Baldi & Long, 2001;Mutch *et al*, 2002). Microarray gene expression profiling is commonplace in AML for identifying specific gene signatures that can

help provide important information regarding diagnosis, prognosis and relevant therapies of patients. An early study by Golub *et al* (1999) suggested roles of gene expression profiling for class discovery (the identification of new subgroups), class prediction (using genes to predict already defined subgroups) and finally for class comparison (the identification of deregulated genes that may be responsible for certain biological properties) (Golub *et al*, 1999). Later studies demonstrated the association of the disease-specific favourable gene translocations; t(8;21), t(15;17) and inv(16), with distinct gene expression profiles (Schoch *et al*, 2002; Valk *et al*, 2004). The uses of gene expression profiling has developed further as various studies have shown its role in risk classification and subset identification for therapeutic targeting (Wilson *et al*, 2006; Raponi *et al*, 2007).

The aim of this chapter was to use gene expression profiling on a cohort of normal karyotype diagnostic AML patients and normal healthy donors to identify genes that may have the potential to monitor MRD in AML patients due to their specific expression patterns and significant levels of expression compared to those in normal healthy donors.

5.2 Materials and Methods

5.2.1 Patient and donor samples

Three gene expression datasets were used for the gene expression analysis to identify differentially expressed genes between normal karyotype AML patients and healthy donors.

5.2.1a Core Dataset

The core dataset contained gene expression data from 72 normal karyotype AML diagnostic patient samples, obtained from 69 patients, and 10 healthy donor samples. The samples were obtained at the time of diagnosis and processed for microarray analysis by either the School of Medicine, Cardiff University or the University College Hospital, London (duplicate samples, processed on both sites, were available for three of the normal karyotype AML patients) (Table 1).

5.2.1b Expanded (with extra donors) dataset

The expanded (with extra donors) dataset contained data for the 82 samples in the core dataset (described in 5.2.1a) plus expression data for an extra 44 healthy donor samples generated by other research groups; 19 samples from a second study by the School of Medicine, Cardiff University, 5 healthy donor samples from a Dutch-Belgium study (GEO accession: GSE 1159) (Valk *et al*, 2004) and 20 healthy donor samples from an American study (GEO accession number: GSE 9476) (Stirewalt *et al*, 2008) (Table 1).

5.2.1c Expanded (with public data) dataset

The expanded (with public data) dataset included the gene expression data for the extra 54 healthy donor samples previously mentioned for the expanded (with extra donors) dataset (5.2.1b) plus gene expression data for 124 normal karyotype AML diagnostic patient samples generated by two other research groups; 115 normal karyotype AML

diagnostic samples from the Dutch-Belgium study (GEO accession: GSE 1159) (Valk *et al*, 2004) and 9 normal karyotype AML diagnostic samples from the American study (GEO accession number: GSE 9476) (Stirewalt *et al*, 2008) (Table 2).

The patient and donor RNA was extracted from bone marrow aspirates and peripheral blood using guanidinium thiocyanate (GSE 1159), TRIzol[®] reagent (GSE 9476, Cardiff samples) or the RNeasy[®] kit (Cardiff samples). The effect of total RNA isolation techniques on the precision and reproducibility of gene expression data has been investigated in a study by Campo Dell'Orto *et al* (2007). By comparing different total RNA isolation methods they concluded that the underlying biological characteristics of paediatric acute leukaemia classes largely surpassed any variations generated between the different protocols used and did not impair the overall distribution of the probe set signal intensity (Campo Dell'Orto *et al*, 2007). The samples were hybridised to Affymetrix human U133A or U133Plus2.0 Genechips.

Characteristic	Value	
Sex - no. (%)		
Male	32	(46.4)
Female	36	(52.2)
Not determined	1	(1.4)
Age group – no. (%)		
<30 years	4	(5.8)
30-60 years	22	(31.9)
>60 years	43	(62.3)
Age – years		
Median	64	
Range	17-87	
French-American British Classification – no. (%)		
M0	4	(5.8)
M1	18	(26.1)
M2	16	(23.2)
M3	0	(0)
M4	15	(21.7)
M5	7	(10.1)
M6	2	(2.9)
M7	1	(1.4)
Bi	1	(1.4)
Not determined	5	(7.4)
Molecular Abnormalities – no. (%)		
Mutation		
FLT3 Internal tandem duplication (ITD)	27	(39.1)
FLT3 Tyrosine kinase domain (TKD)	8	(11.6)
NPM1	24	(34.8)
ITD ⁺ /NPM1 ⁺	11	(15.9)
ITD ⁺ /NPM1 ⁻	6	(8.7)
ITD ⁻ /NPM1 ⁺	13	(18.8)
ITD ⁻ /NPM1 ⁻	13	(18.8)

Table 1. The frequencies and percentages of the clinical and molecular characteristics of the 69 patients diagnosed with normal karyotype AML in the core dataset and the expanded (with extra donors) dataset. Duplicate samples were available for 3 of the patients.

Characteristic	Value	
Sex - no. (%)		
Male	5	(4.0)
Female	4	(3.2)
Not determined	115	(92.8)
Age group – no. (%)		
<30 years	0	(0)
30-60 years	5	(4.0)
>60 years	4	(3.2)
Not determined	115	(92.8)
Age – years		
Median*	58	
Range*	38-79	
French-American British Classification – no. (%)		
M0	0	(0)
M1	40	(32.3)
M2	24	(19.4)
M3	0	(0)
M4	21	(16.9)
M5	32	(25.8)
M6	2	(1.6)
M7	0	(0)
MDS/AML	1	(0.8)
Not determined	4	(3.2)
Molecular Abnormalities – no. (%)		
Mutation		
FLT3 Internal tandem duplication	57	(46.0)
FLT3 Tyrosine kinase domain	9	(8.1)

Table 2. The frequencies and percentages of the clinical and molecular characteristics of the 124 patients diagnosed with normal karyotype AML in the expanded (with public data) dataset from the GEO database (GSE9476 and GSE1159) (*data not available for GSE1159). No NPM1 data was available for either dataset.

5.2.2 Data Analysis

5.2.2a Expression summarisation

Image files (CEL files) were analysed using the statistical algorithms in MAS 5.0 using target intensity (TGT) of 100. The resultant data matrix of expression levels for each probe set was taken forward for further analysis (see accompanying DVD for CEL files).

5.2.2b Fold Change Calculation

A simple method for identifying differentially expressed genes between test groups is that of fold change. A gene is said to have significantly changed in expression if its mean expression value varies by more than a set factor between the control group and the test group. As a fold change calculation only takes into account the mean of the expression values, outlier values can greatly affect the mean expression value such that genes may be identified as significantly under or over expressed when in fact they are not in the majority of samples.

5.2.2c Improved Fold Change Calculation

As described above, the process of identifying differentially expressed genes on a global scale can easily be marred by the odd erroneous gene expression level present in some samples that underplay (or overplay) the true difference in fold change of the genes between the sample groups. To overcome this problem, and calculate a more accurate fold change value for the genes, the process of outlier detection and removal can be applied. A programme designed to

identify and remove data outliers initially calculates the median and the median absolute deviation value (MAD) for the whole data set. The MAD is the median of the absolute values of the differences between values and is a more robust measurement of variance than standard deviation; the standard deviation is more sensitive to outliers whereas the MAD value is not altered by their presence. Data values of more than 3 MAD away from the median value are identified as outliers and removed from the data set. The median and the MAD are then recalculated for the remaining data and again data values of more than 3 MAD away from the median value are identified as outliers and removed from the data set. This process is continued until all the data values sit within 3 MAD of the current median value of the data set (Figure 1).

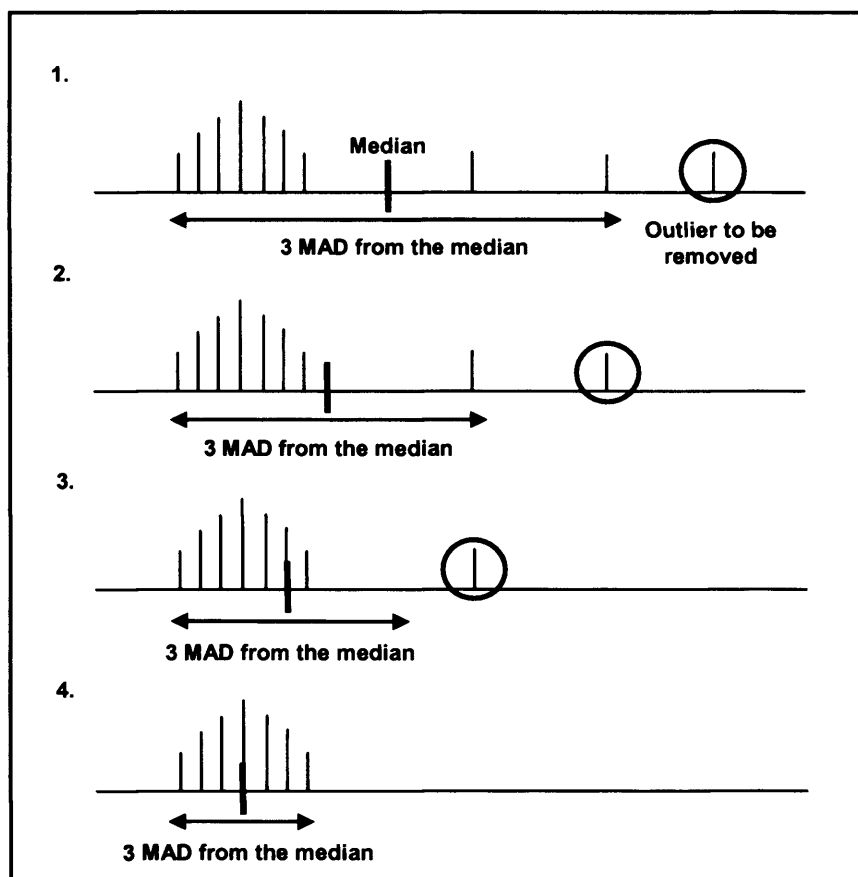


Figure 1. Schematic of the method used by the outlier programme to identify and remove outliers from the data set. The median value of the data set is calculated and values more than 3 MAD away are identified as outliers and removed (1). The median value is re-calculated and values more than 3 MAD away from the median value are identified as outliers and removed (2). The process is repeated (3) until all values lie within 3 MAD of the median value (4).

5.2.2d Detecting Differential Gene Expression

A simple fold-change approach for gene expression profiling is unlikely to yield optimal results as it only looks at the mean of the expression values. A *t*-test approach can provide information of greater significance by looking at the spread of the data (variance) as well as the location (mean) to give a statistical score of any “true” differences observed between two test sets. Disadvantages of a basic *t*-test include low repetition numbers leading to an underestimation of variance. A Bayesian *t*-test combines the use of prior distributions and likelihood to

provide all relevant information regarding mean and variance allowing for the detection of potentially significant changes in gene expression between groups that would otherwise be missed by fold-change or simple *t*-test calculations (Baldi & Long, 2001).

5.2.2e Batch Correction

Non-biological differences between samples can occur when combining gene expression data generated by several different groups or platforms. When data is initially analysed, samples are seen to cluster based on which experiment they belong to rather than any informative biological differences occurring between the samples. This “batch-effect” can be adjusted using several methods. Unfortunately many methods may remove real biological variation from the data and require large numbers of samples for optimum results. The Empirical Bayes (EB) method is a robust method suitable for small batch sizes commonly used to adjust batch effects seen in microarray data. This method assumes that phenomena resulting in a batch effect will affect many genes in a similar way and uses this idea to make better estimates and adjustments across genes in an experiment (Johnson *et al*, 2007).

5.2.2f Biological Pathways and Networks

PathwayArchitect® (Stratagene) uses a repository of over 2 million biological findings derived from interaction databases and current literature. Using manual curation and text mining algorithms, it identifies

the relationships, or *interactions*, between proteins and small molecules, also known as *entities*. A collection of *entities* and *interactions* are visualised through biological interaction networks (BIN). To identify more statistically relevant interactions, relevance interaction networks (RIN) are generated – advanced statistical models are used to maximise the biological relevance of the genes/proteins and their interactions (Stratagene Software Solutions). MetaCore™ (GeneGO) is a manually curated database of protein-protein and protein-DNA interactions, and metabolic and signalling pathways that can also analyse microarray data for the identification of statistically significant pathways and networks.

5.3 Results

5.3.1 Expression profiling to identify MRD markers in normal karyotype AML patients.

Gene expression profiling is a useful and powerful tool to help identify specific gene expression patterns across different disease states. To identify potential markers of minimal residual disease for AML patients with a normal karyotype, gene expression profiling was used to identify genes differentially expressed between normal karyotype AML patients and normal healthy donors. The initial analysis was performed on a core dataset consisting of 72 normal karyotype AML samples and 10 healthy bone marrow samples.

5.3.2 Improving the estimate of fold change and its importance in MRD marker identification

To act as a reliable marker of MRD, the expression levels of a particular gene need to be significantly differentially expressed between the disease state and the normal healthy state. Ideally a candidate gene should be highly expressed in the normal karyotype AML patients but have much lower, if not absent, levels of expression in healthy normal donors. To begin the analysis, a simple fold change calculation was applied to the core dataset to identify genes differentially expressed between the normal karyotype AML samples and the healthy donor samples by more than 4 fold and compared to genes similarly identified from the core dataset with any outliers removed. To remove the outliers, expression values more than 3 median absolute deviation values (MAD) away from the median expression value were identified as “outliers” and removed resulting in the median value being recalculated and the process repeated (see materials and methods) (Figure 2). With the removal of outliers and more accurate mean expression values calculated, a more accurate list of differentially expressed genes was generated.

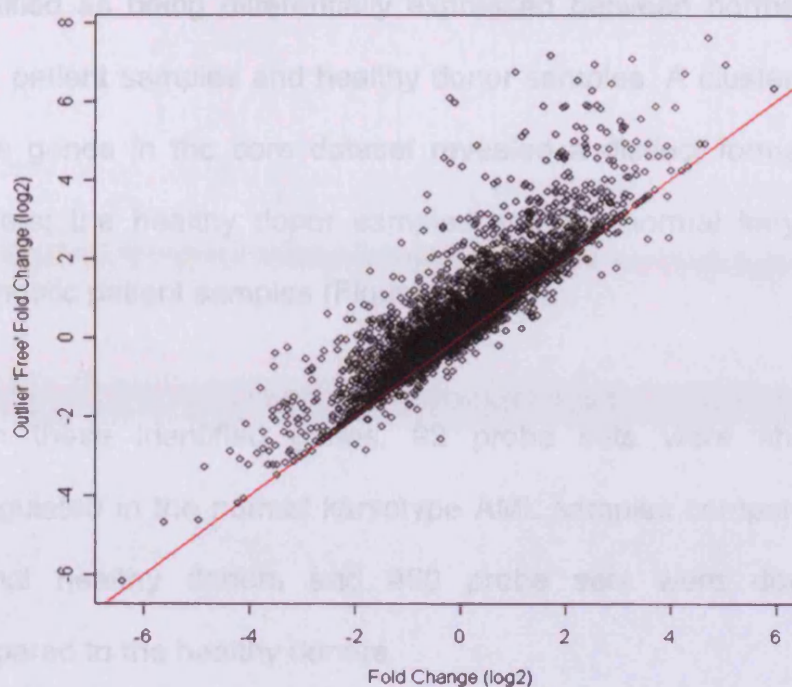


Figure 2. Dot plot of the relationship between the gene expression values from the original core dataset and from the “outlier free” core dataset.

5.3.3 Identification of MRD markers in the core dataset

Continuing with the “outlier-free” core dataset, a fold change calculation was applied to identify fold change differences between the normal karyotype AML diagnostic patient sample and the healthy donor samples. A Bayesian *t*-test was also applied to the core dataset, without the outliers removed, to improve the identification of genes with differences in expression between the two test groups by looking at the spread of the data (variance) as well as the mean of the data. Combining the genes identified through the Bayesian *t*-test ($p < 0.05$) with the “outlier-free” genes identified as being differently expressed between the two test groups by more than 4 fold, 552 probe sets were

identified as being differentially expressed between normal karyotype AML patient samples and healthy donor samples. A cluster analysis of these genes in the core dataset revealed a distinct formation of two clusters; the healthy donor samples and the normal karyotype AML diagnostic patient samples (Figure 3).

From these identified genes, 92 probe sets were shown to be upregulated in the normal karyotype AML samples compared to in the normal healthy donors and 460 probe sets were downregulated compared to the healthy donors.

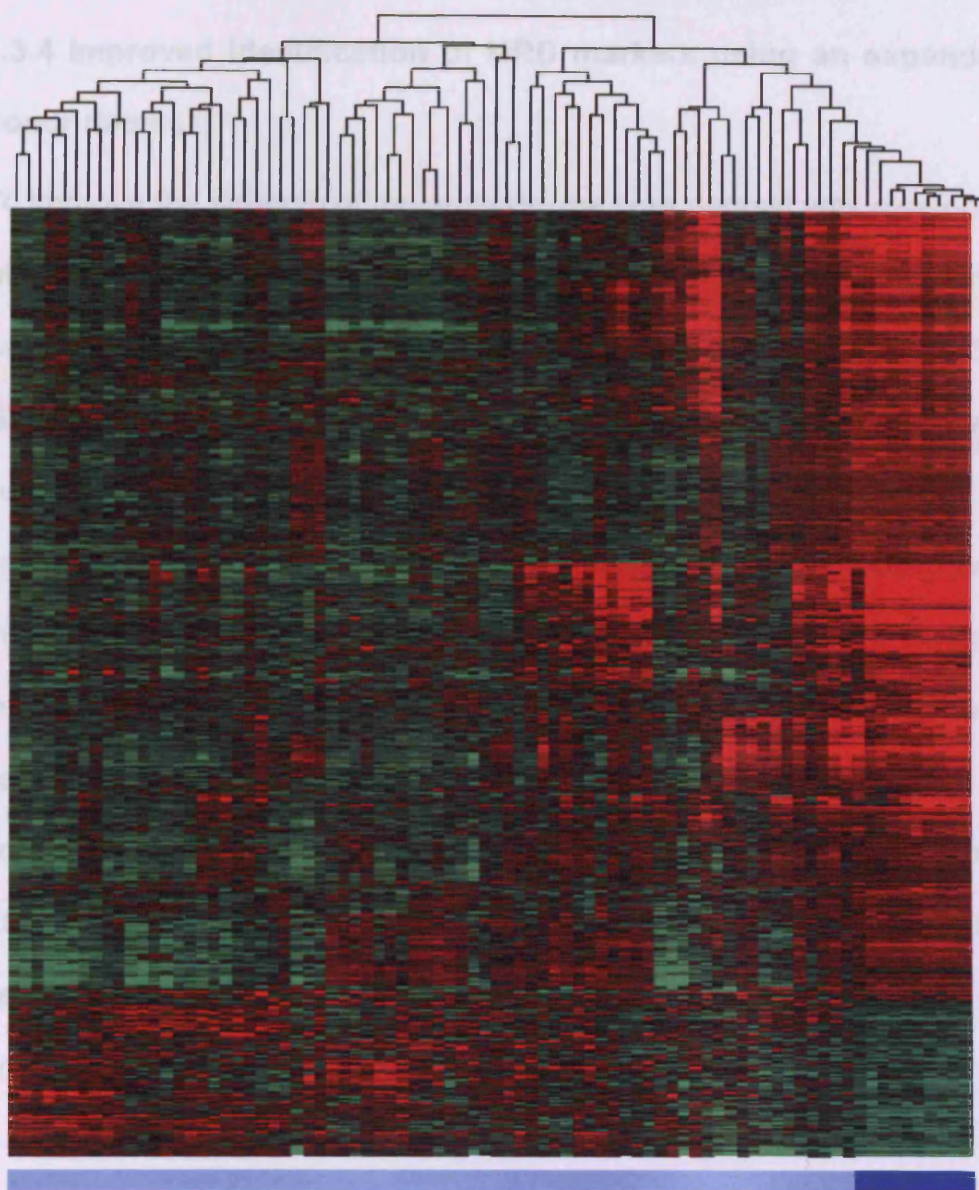


Figure 3. Heatmap of the 552 probe sets identified as being differentially expressed between the normal karyotype AML patient samples (light blue) and the healthy donor samples (dark blue) by Bayesian *t*-test ($p < 0.05$) and a fold change calculation ($FC > 4$) in the core dataset. The colours relate to the expression level of the genes to indicate either a high level of expression (red) or a low level of expression (green). The genes are displayed on the vertical axis and the samples are displayed on the horizontal axis.

5.3.4 Improved identification of MRD markers using an expanded donor dataset

To improve the strength of the analysis, the core dataset was expanded with the addition of the expression data for another 44 healthy donor samples derived from two other studies (GEO accession: GSE 1159, GSE 9476). An analysis to identify the 1000 most variable genes was performed on the newly expanded (with extra donors) dataset and an unsupervised cluster analysis of these genes identified a problem with the donor samples; the cluster analysis identified 3 distinct clusters defined by the specific experiment they came from rather than their gene expressions. An EB-method of batch correction was applied to adjust the data and remove any distinct experimental-source problems. After the batch correction had been applied to the data no specific gene-expression-sample origin pattern was observed with an unsupervised cluster analysis (Figure 4).

A further cluster analysis of the 552 probe sets previously identified from the core dataset in the expanded (with extra donors) dataset of 126 samples demonstrated the formation of two distinct clusters based on their levels of gene expression between the healthy donor samples and the normal karyotype AML patient samples, with the exception of 3 healthy donor samples originating from two Cardiff studies that clustered in with the normal karyotype AML patient samples (Figure 5).

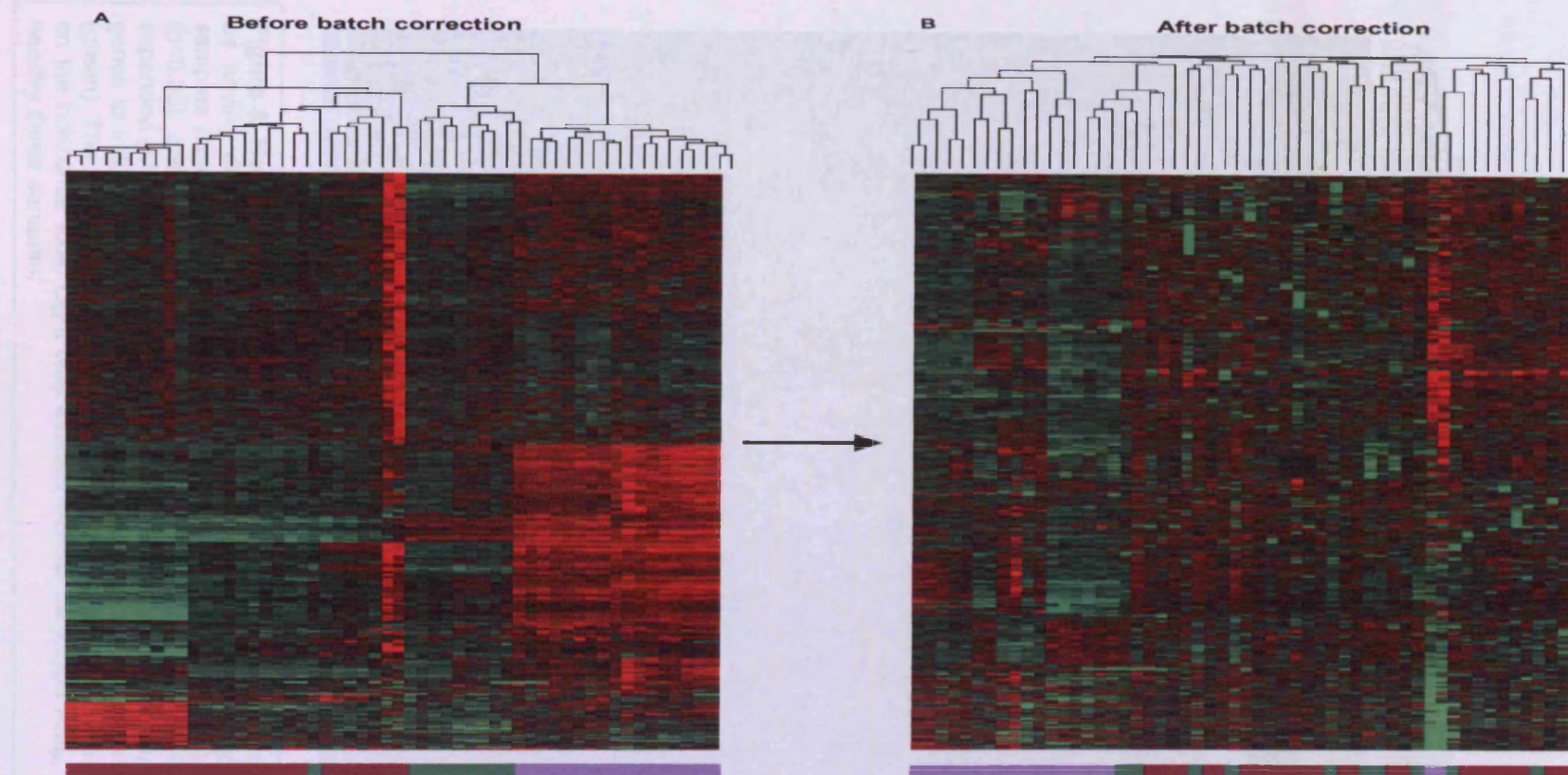


Figure 4. An unsupervised cluster analysis of the 1000 most variable genes identified from the expanded (with extra donors) dataset in the 54 healthy donor samples before batch correction (A) and after a batch correction (B) was applied. The colours of the heatmaps relate to the expression level of the genes to indicate either a high level of expression (red) or a low level of expression (green). The genes are displayed on the vertical axis and the samples are displayed on the horizontal axis. (Maroon = GEO samples, sea green = Cardiff samples HU133A, Lilac = Cardiff samples HU133Plus2.0)

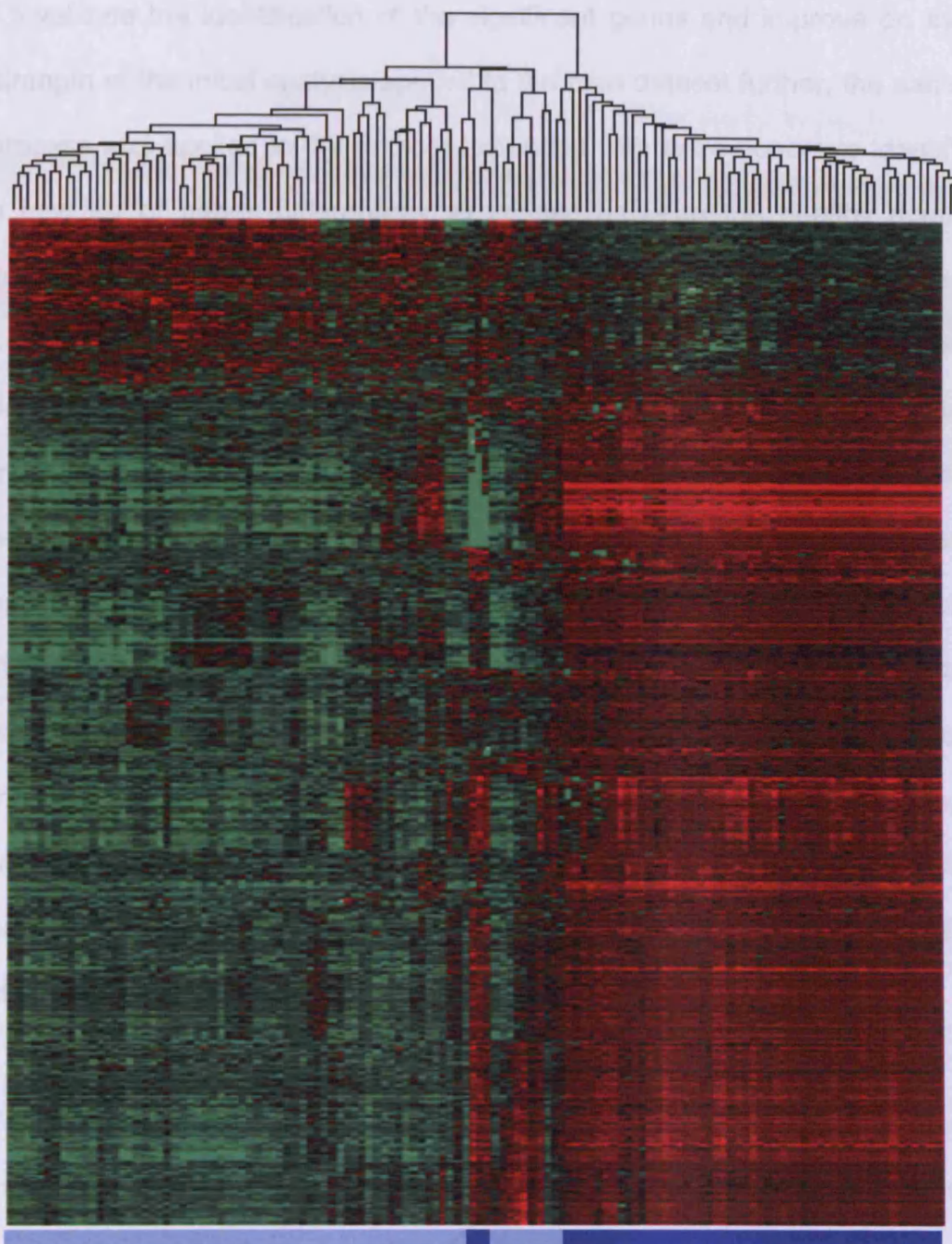


Figure 5. Heatmap of the 552 probe sets previously identified from the core dataset as having differences in expression between the normal karyotype AML patient samples (light blue) and the healthy donor samples (dark blue) by Bayesian t -test ($p < 0.05$) and a fold change calculation ($FC > 4$) applied to the batch-corrected expanded (with extra donors) dataset. The colours relate to the expression level of the genes to indicate either a high level of expression (red) or a low level of expression (green). The genes are displayed on the vertical axis and the samples are displayed on the horizontal axis. Light blue indicate normal karyotype AML, dark blue indicate healthy donor samples.

To validate the identification of the significant genes and improve on the strength of the initial analysis applied to the core dataset further, the same process was applied to the dataset expanded with extra donors to identify a new list of genes differentially expressed between the healthy donor samples and the normal karyotype patient AML diagnostic samples. The same outlier programme was applied to the expanded (with extra donors) dataset to identify and remove any outlier values as before and provide more “accurate” fold change values. The fold change calculation ($FC > 4$) was applied to the “outlier-free” data and the Bayesian *t*-test ($P < 0.05$) was applied to whole data set to identify genes with differences in expression. When the two sets of identified genes were combined 1577 probe sets were identified as being common to both analyses. Cluster analysis of these genes in the 126 sample dataset demonstrated a distinct clustering of the healthy donor samples and of the normal karyotype AML diagnostic patient samples, with the same 3 healthy donor samples seen previously clustered with normal karyotype AML samples (Figure 6).

From these identified genes, 25 probe sets were shown to be upregulated in the normal karyotype AML samples compared to in the donor healthy normal donors (Table 3, Figure 7) and 1552 probe sets were downregulated. From the new list of probe sets identified, 13/25 (52%) were previously identified in the initial analysis of the core dataset and 6/7

(85.7%) of the new probe sets with a fold change of more than 10 were also previously identified in the initial analysis.

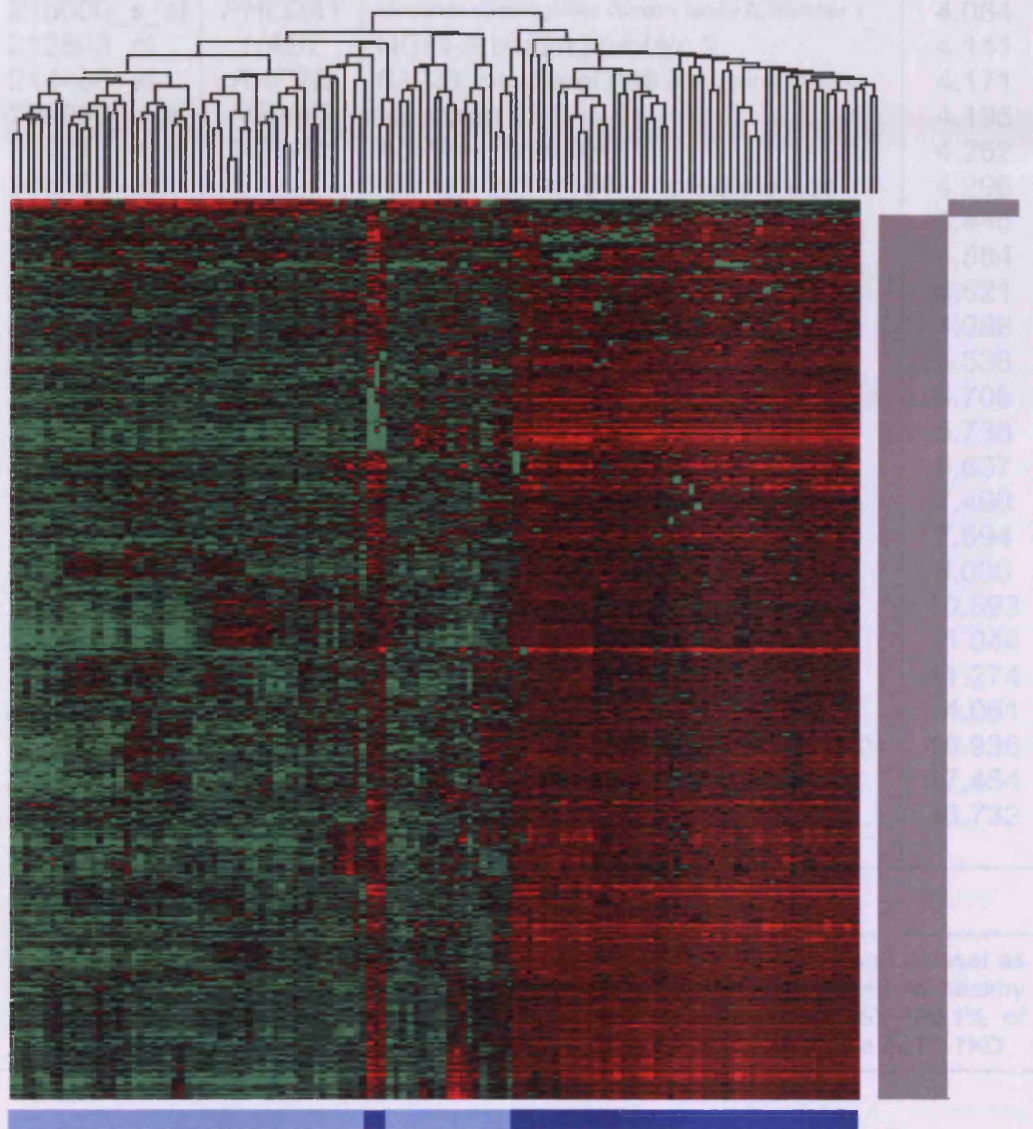


Figure 6. Heatmap of the 1577 probe sets identified as being differentially expressed between the normal karyotype AML patient samples (light blue) and the healthy donor samples (dark blue) through a Bayesian t -test ($p < 0.05$) and a fold change calculation ($FC > 4$) in the expanded (with extra donors) dataset. The colours relate to the expression level of the genes to indicate either a high level of expression (red) or a low level of expression (green). The genes are displayed on the vertical axis and the samples are displayed on the horizontal axis.

Probe set	Gene Symbol	Gene Title	Fold Change
213147_at	<i>HOXA10</i>	Homeobox A10	4.030
218000_s_at	<i>PHLDA1</i>	Pleckstrin homology-like domain, family A, member 1	4.084
212803_at	<i>NAB2</i>	NGFI-A binding protein 2	4.141
214488_at	<i>RAP2B</i>	RAP2B, member of RAS oncogene family	4.171
209983_s_at	<i>NRXN2</i>	Neurexin 2	4.195
201243_s_at	<i>ATP1B1</i>	ATPase, Na ⁺ /K ⁺	4.252
206532_at	-	Human Ini1 mRNA, complete cds	4.296
204069_at	<i>MEIS1</i>	Meis homeobox 1	4.446
206674_at	<i>FLT3*</i>	fms-related tyrosine kinase 3	4.584
205453_at	<i>HOXB2</i>	Homeobox B2	4.621
214880_x_at	<i>CALD1</i>	Caldesmon 1	4.988
211039_at	<i>CHRNA1</i>	Cholinergic receptor, nicotinic, alpha 1	5.538
219908_at	<i>DKK2</i>	Dickkopf homolog 2 (<i>Xenopus laevis</i>)	5.705
213150_at	<i>HOXA10</i>	Homeobox A10	5.736
211125_x_at	<i>GRIN1</i>	Glutamate receptor	6.637
215441_at	-	Human clone 23842 mRNA sequence	7.490
216011_at	<i>SLC39A9</i>	Solute carrier family 39	7.594
206310_at	<i>SPINK2</i>	Serine peptidase inhibitor	8.030
213110_s_at	<i>COL4A5</i>	Collagen, type IV, alpha 5	10.593
214651_s_at	<i>HOXA9</i>	Homeobox A9	11.044
213844_at	<i>HOXA5</i>	Homeobox A5	11.274
206067_s_at	<i>WT1</i>	Wilms' tumour 1	14.051
217698_at	-	/DB_XREF=AV651668/CLONE=GLCCSC04	16.936
209905_at	<i>HOXA9</i>	Homeobox A9	17.454
205366_s_at	<i>HOXB6</i>	Homeobox B6	36.732

Table 3. The 25 probe sets identified from the expanded (with extra donors) dataset as upregulated in normal karyotype AML patients samples when compared to healthy donor samples by fold change (FC>4) and a Bayesian *t*-test (P>0.05). *39.1% of patients presented with a *FLT3*-ITD and 11.3% of patients presented with a *FLT3*-TKD.

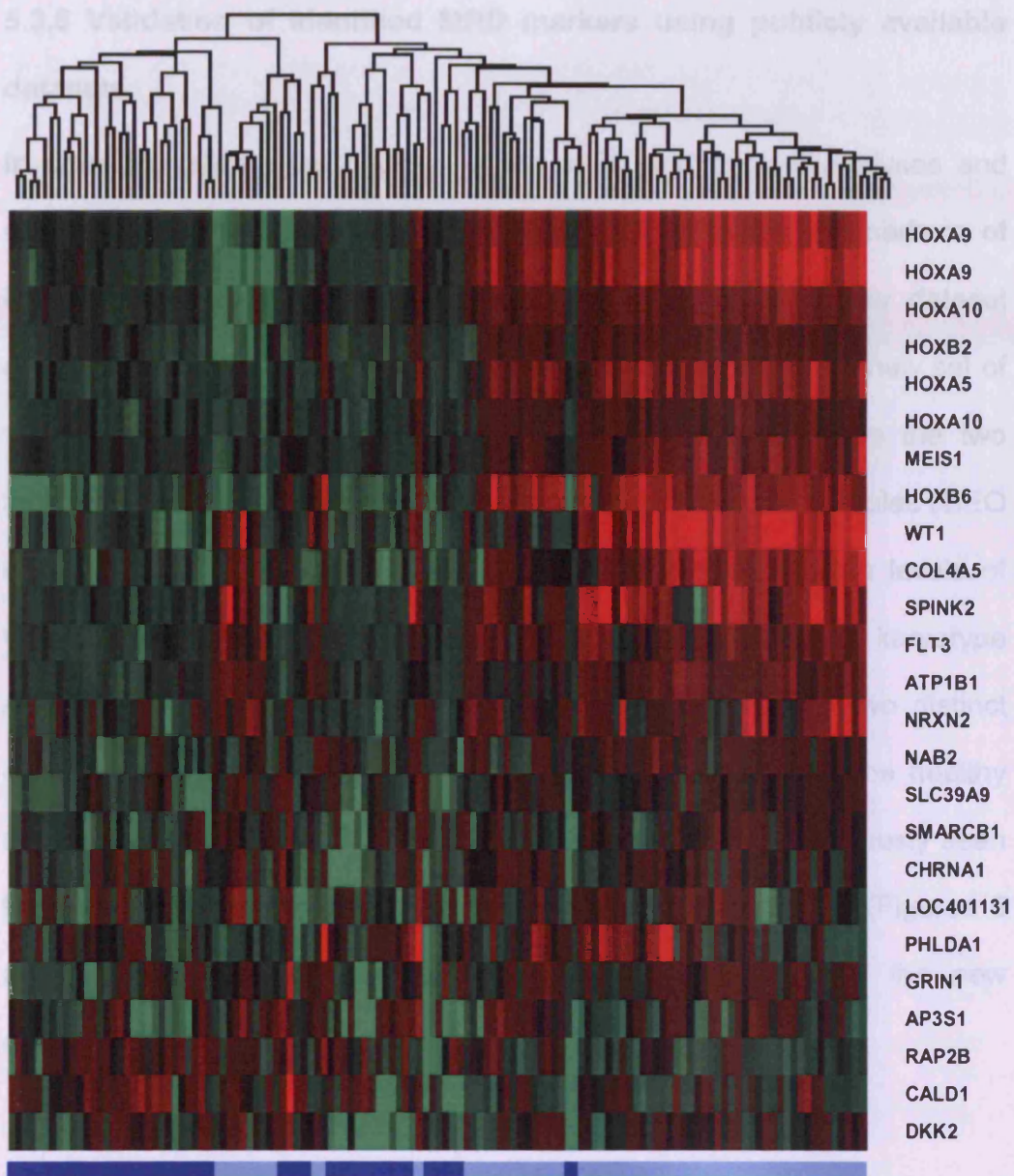


Figure 7. Heatmap of the 25 probe sets identified as being up regulated in the normal karyotype AML patient samples (light blue) compared to the healthy donor samples (dark blue) through a Bayesian t -test ($p < 0.05$) and a fold change calculation ($FC > 4$) in the expanded (with extra donors) dataset. The colours relate to the expression level of the genes to indicate either a high level of expression (red) or a low level of expression (green). The genes are displayed on the vertical axis and the samples are

5.3.5 Validation of identified MRD markers using publicly available datasets

In order to validate the findings of the initial and second analyses and confirm the potential application of the identified genes as markers of MRD, the 25 up-regulated probe sets were applied to a new dataset containing the 54 healthy donor samples previously used and a new set of 124 normal karyotype diagnostic AML samples obtained from the two independent studies used previously for their healthy donor samples (GEO accession: GSE 1159, GSE 9476). Looking at the expression levels of these 25 probe sets only in healthy normal samples and normal karyotype AML diagnostic samples, the two conditions separated into two distinct clusters; the normal karyotype AML diagnostic samples and the healthy normal samples (Figure 8). The 3 normal donor samples, previously seen to cluster with the normal karyotype diagnostic AML samples (Figures 5 and 6) clustered with the other healthy donor samples in the new expanded dataset.

5.3.5 Pathway analysis of enriched genes

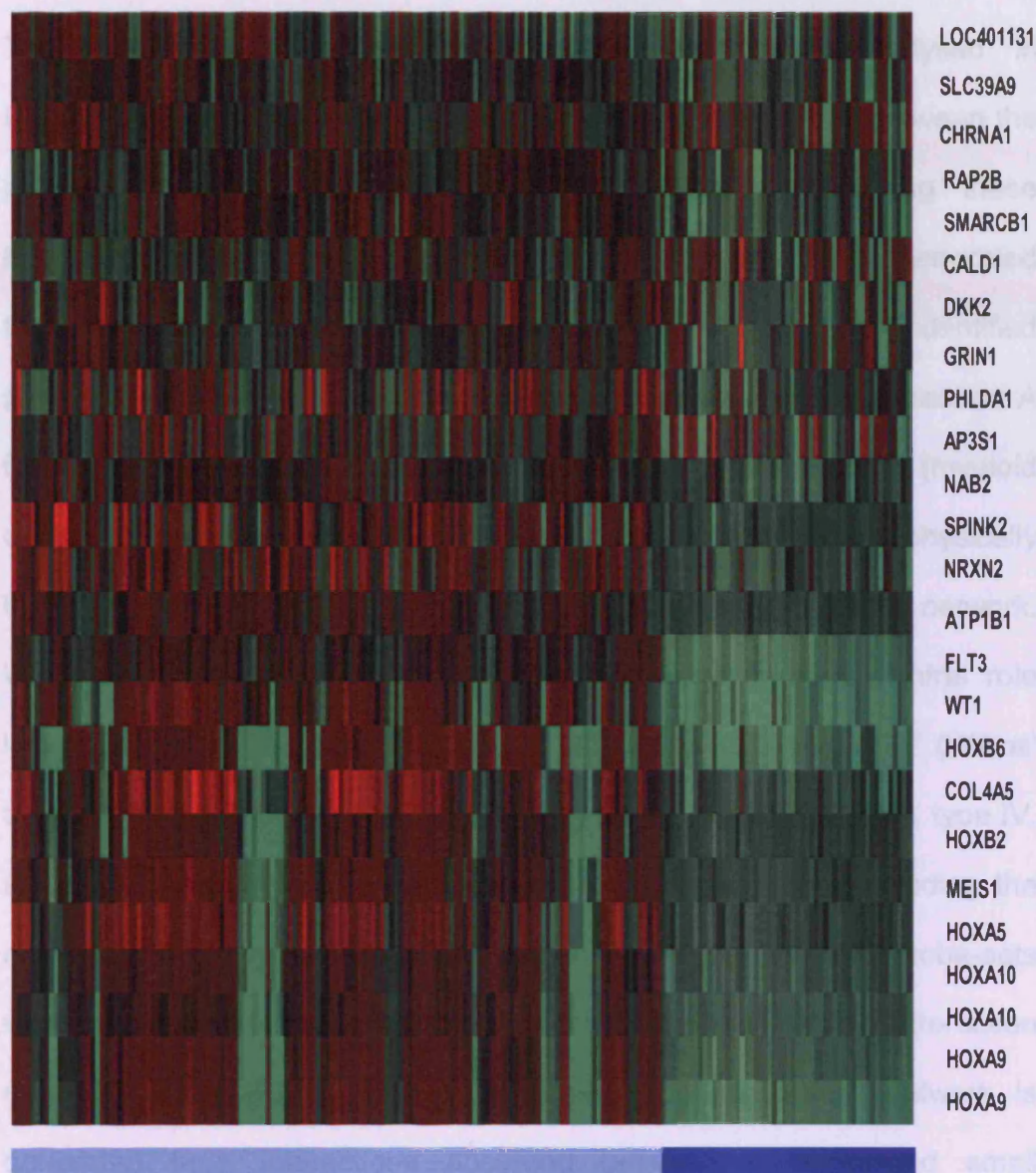


Figure 8. A heatmap of the 25 probe sets identified as being up regulated in the normal karyotype AML patient applied to a new expanded dataset containing the 54 previously used healthy donor samples (dark blue) and 124 new AML normal karyotype samples from two independent studies (light blue) (GEO accession: GSE 1159, GSE 9476). The colours relate to the expression level of the genes to indicate either a high level of expression (red) or a low level of expression (green). The genes are displayed on the vertical axis and the samples are displayed on the horizontal axis.

5.3.6 Pathway analysis of identified genes

The 25 probe sets identified as up-regulated were analysed in PathwayArchitect® to identify statistically relevant interactions between the proteins coded by the genes and the networks containing these interactions. A “relevance binding Interaction” network can be generated for proteins that physically interact with each other. From the 25 identified probe sets, proteins encoded by 13 genes, including the 3 homeobox A family members (*HOXA*) and the homeobox gene *MEIS1* (myeloid ecotropic viral integration site 1 homolog), were identified as physically interacting with each other and forming a biologically relevant network. Within the pathways the HOXA9 protein appears to have central role interacting with the proteins MEIS1, HOXA5, HOXA10, WT1 (Wilms’ tumour), NAB2 (NGFI-A binding protein 2) and COL4A5 (collagen, type IV, alpha 5) through other intermediate proteins. Nine genes, including the homeobox B family member HOXB6, from the original list of 25 probe-sets were not found to code for proteins involved in a binding interaction network (Figure 9). A “relevance transcription regulation” network is generated from interactions occurring between proteins and small molecules involved in the regulation of transcription. From the original list of 25 probe sets, proteins encoded by 4 genes were identified as interacting in the regulation of transcription and forming a biologically statistically relevant network for this regulation; WT1, glutamate receptor, ionotropic, N-methyl D-aspartate 1 (GRIN1), HOXA10 and NAB2 (Figure

10). All of these 4 genes were identified as coding for proteins with a statistically significant role in the relevance binding network also. The smallest binding network consisted of two identified proteins; SLC39A9 (solute carrier family 39 (zinc transporter), member 9) and PHLDA1 (pleckstrin homology-like domain, family A, member 1), connected by an intermediate protein, TCF (transcription factor). The third pathway contained 4 proteins encoded by the list of upregulated genes; GRIN1 (glutamate receptor, ionotropic, N-methyl D-aspartate 1), CALD1 (caldesmon 1), AP3S1 (adaptor-related protein complex 3, sigma 1 subunit) and FLT3 (fms-like tyrosine kinase 3). The *FLT3* gene is preferentially expressed in early haematopoietic progenitors and expressed at high levels in approximately 90-100% of AML cases (Birg *et al*, 1992;Carow *et al*, 1996)

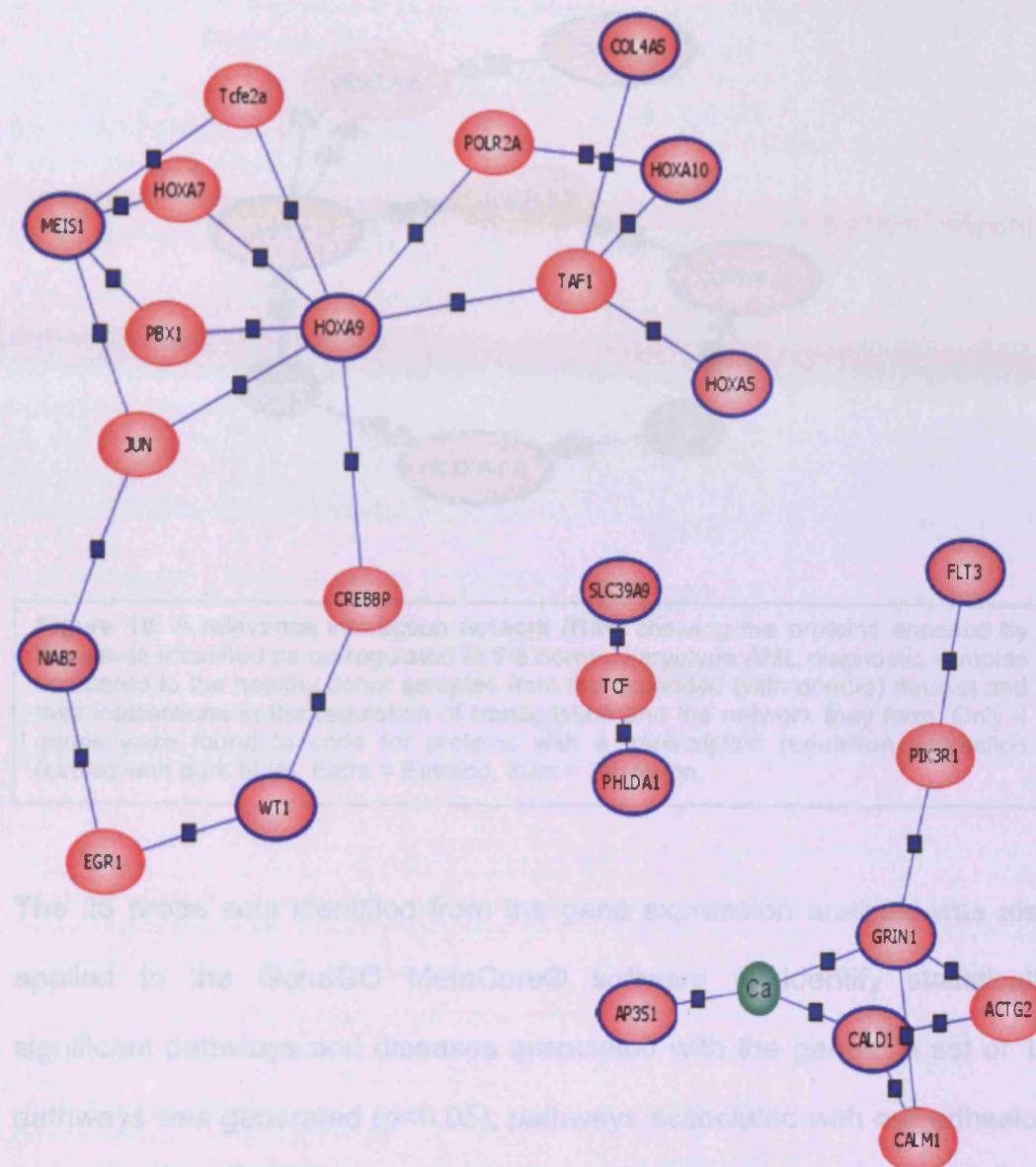
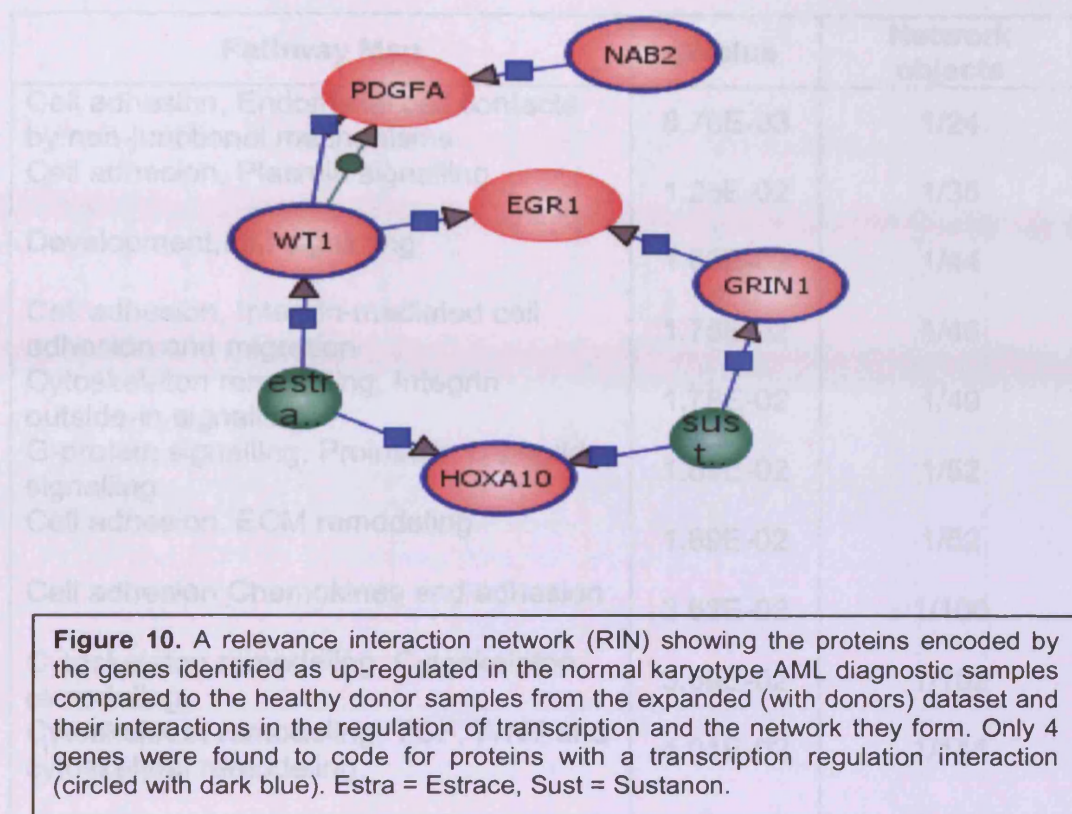


Figure 9. A relevance interaction network (RIN) showing the proteins coded by the genes identified as up-regulated in the normal karyotype AML diagnostic samples compared to the healthy donor samples from the expanded (with donors) dataset and their involvement in binding interactions i.e. they physically interact with each other. Only 13 genes were found to code for a protein with a binding interaction (circled with dark blue) and 3 networks were identified. Ca = Calcium



The 25 probe sets identified from the gene expression analysis was also applied to the GeneGO MetaCore® software to identify statistically significant pathways and diseases associated with the genes. A set of 10 pathways was generated ($p < 0.05$); pathways associated with cell adhesion ($n = 5$, 50%), cytoskeleton remodelling ($n = 3$, 30%), development ($n = 1$, 10%) and G-protein signalling ($n = 1$, 10%) (Table 4, Figure 11). The upregulated genes were also identified as having a statistically significant association with 50 diseases ($p < 0.005$); 19/50 (38%) were haematological diseases, 12/50 (24%) were development of kidney and gonad-related diseases (Table 5, Figure 11).

Pathway Map	pValue	Network objects
Cell adhesion, Endothelial cell contacts by non-junctional mechanisms	8.76E-03	1/24
Cell adhesion, Plasmin signalling	1.28E-02	1/35
Development, Flt3 signalling	1.60E-02	1/44
Cell adhesion, Integrin-mediated cell adhesion and migration	1.75E-02	1/48
Cytoskeleton remodeling, Integrin outside-in signalling	1.78E-02	1/49
G-protein signalling, Proinsulin C-peptide signalling	1.89E-02	1/52
Cell adhesion, ECM remodeling	1.89E-02	1/52
Cell adhesion Chemokines and adhesion	3.62E-02	1/100
Cytoskeleton remodeling, Cytoskeleton remodeling	3.69E-02	1/102
Cytoskeleton remodeling, TGF, WNT and cytoskeletal remodeling	4.01E-02	1/111

Table 4. The pathway maps that the genes from the 25 up-regulated probe set list were identified as having a statistically significant association with by GeneGO MetaCore.

Disease	pValue	Network objects
Leukemia, Myeloid	4.52E-09	10/535
Myelodysplastic Syndromes	9.73E-09	10/579
Bone Marrow Diseases	4.86E-08	10/684
Leukemia	3.29E-07	11/1098
Leukemia, Myelocytic, Acute	8.75E-07	7/325
Proteinuria	2.76E-06	4/56
Denys-Drash Syndrome	3.24E-06	2/2
Hematologic Neoplasms	5.04E-06	4/65
Leukemia, Non-lymphocytic, Acute	5.20E-06	7/424
Neoplasm, Residual	7.94E-06	3/22

Table 5. The statistically significant top 10 diseases that the genes from the 25 up-regulated probe set list were identified as having a statistically significant association with by GeneGO MetaCore.

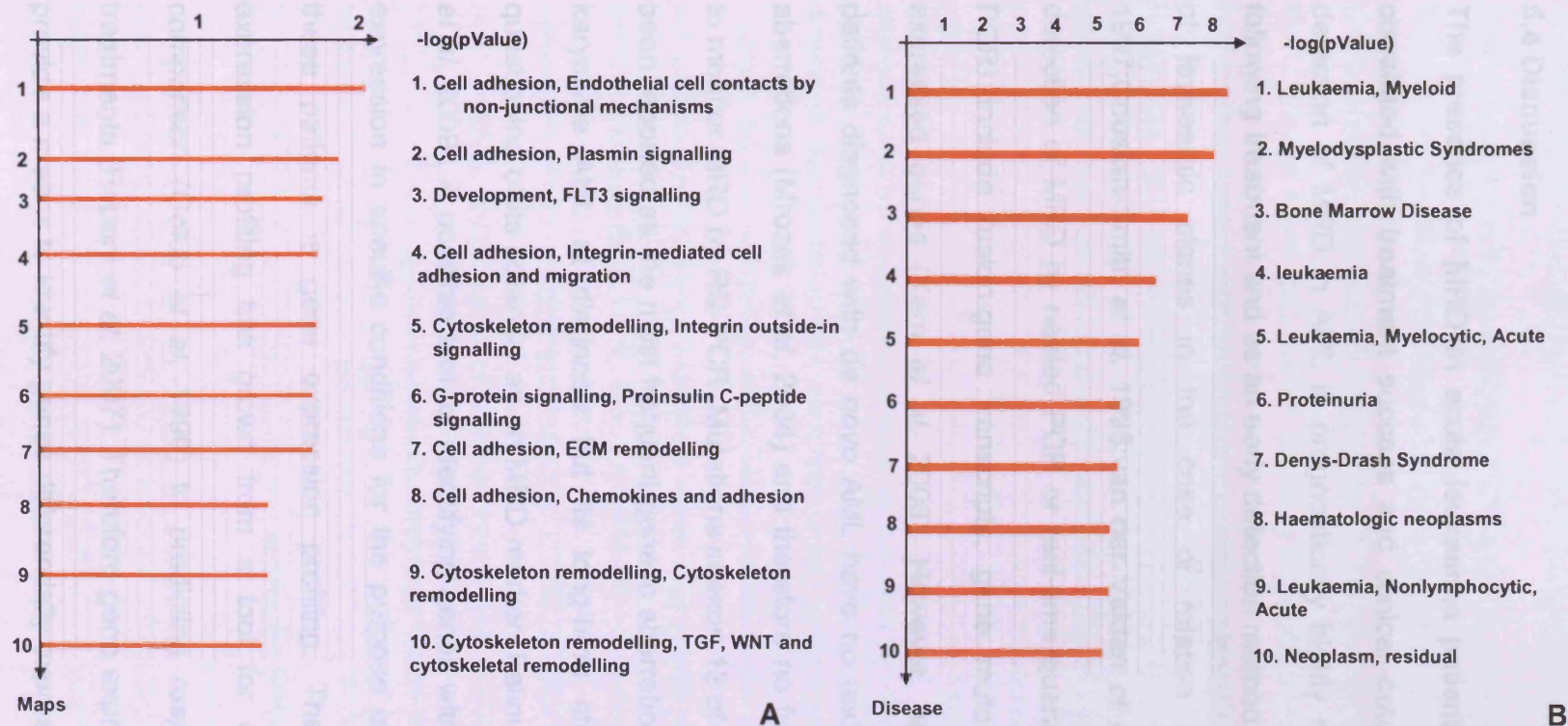


Figure 11. Schematics of the pathway maps (A) and diseases (B) identified by GeneGO MetaCore and the statistical significance of the association of the up-regulated genes with these pathways and diseases.

5.4 Discussion

The presence of MRD in acute leukaemia patients has been strongly correlated with treatment success and clinical outcome. Therefore, the detection of MRD in AML is prognostically highly relevant at diagnosis, following treatment and as an early detection method for the reappearance of leukaemic clones in the case of relapse (San Miguel *et al*, 1997; Coustan-Smith *et al*, 1998; van der Velden *et al*, 2003). Targets for detection of MRD by nested PCR or real-time quantitative RT-PCR (RQ-PCR) include fusion-gene transcripts, gene mutations and aberrantly expressed genes (Kern *et al*, 2008). However, approximately 45% of patients diagnosed with *de novo* AML have no underlying chromosomal aberrations (Mrozek *et al*, 2004) and therefore no fusion-gene transcripts to monitor MRD by RQ-PCR. Mutations in exon 12 of the *NPM1* gene have been reported as the most frequent genetic aberration occurring in normal karyotype AML at diagnosis but its long-term stability has led to a questioning of its potential as an MRD marker (Falini *et al*, 2007; Papadaki *et al*, 2009). A possible tool for identifying genes with an aberrant level of expression in specific conditions for the purpose of MRD monitoring in these patients is gene expression profiling. The potential of gene expression profiling has grown from a tool for class prediction and comparison (Golub *et al*, 1999) to predicting response to therapeutic treatments (Raponi *et al*, 2007). Therefore gene expression profiling could provide a means to identify genes differentially expressed between healthy

normal donors and normal karyotype AML patients, which can then be further monitored by quantitative PCR assays in patient follow-up samples to observe significant changes in gene expression levels, monitor MRD levels and thereby predict early relapse. Various groups have investigated the ability of gene expression profiling to identify potential MRD markers but studies have been limited to paediatric patients, particularly in acute lymphoblastic leukaemia (ALL). Chen *et al* (2001) used cDNA arrays to identify several genes over-expressed for MRD in a small cohort of ALL patients. They identified 64 genes that were over-expressed in three ALL patients and validated the expression of 7 proteins encoded by these genes by flow cytometry (Chen *et al*, 2001). Steinbach *et al* (2006) analysed five pools of 10 paediatric AML patient samples based on the hypothesis that for every useful MRD marker there would be at least 1 in 10 patients expressing the gene at a high level. Using oligonucleotide arrays and real-time PCR, they identified 5 potential markers of MRD for paediatric AML (Steinbach *et al*, 2006). Although these data are encouraging, it is clear that further investigation into the identification of MRD markers in acute leukaemias through gene expression profiling is required in much larger patient cohorts for the genes identified to have a stronger statistical potential to act as markers for MRD in AML.

Using gene expression profiling to identify potential MRD markers for AML in this study, 25 probe sets were identified as upregulated in diagnostic

normal karyotype AML samples compared to healthy donor samples, by a mean change of between 4.0 and 36.7 fold. Homeobox (*HOX*) genes and *HOX*-related genes were represented by 8/25 (32%) of the upregulated probe sets; *HOXA5*, *HOXA9*, *HOXA10*, *HOXB2*, *HOXB6* and *MEIS1*, a TALE (three amino acid loop extension) subfamily member of non-*HOX* homeobox genes. The identification of this *HOX* expression pattern through gene expression profiling is strongly supported by a previous study by Debernardi *et al* (2003) in which AML patients with a normal karyotype were found to have a distinctive up-regulation of several *HOXA* and *HOXB* genes and *HOX*-related genes compared to AML patients with a favourable risk translocation through microarray-based gene expression analysis. Comparing the expression profiles of patients with either favourable risk translocations, rearrangements of the *MLL* gene or a normal karyotype, Debernardi *et al* (2003) found the normal karyotype patient group were characterised predominantly by elevated levels of *HOXA5*, *HOXA9*, *HOXB2*, *HOXB6* and *MEIS1*; these findings strongly support the analysis presented in this chapter. They also observed elevated levels of *HOXA4*, *HOXB3*, *HOXB5*, *HOXB6* and *PBX3* (pre-B-cell leukaemia transcription factor 3), another member of the *TALE* (three amino acid loop extension) gene family, in the normal karyotype patients; these genes were not found to be upregulated in this present study (Debernardi *et al*, 2003). Debernardi *et al* (2003) identified a strong link between expression patterns and karyotypic status that implied a common

genetic lesion between these patients (Debernardi *et al*, 2003). Bullinger *et al* (2004) performed gene expression profiling on 116 adult samples to identify prognostic subgroups within AML and identified the characterisation of normal karyotype AML through the overexpression of *HOX* genes, including *HOXA4*, *HOXA9*, *HOXA10* and *MEIS1* (Bullinger *et al*, 2004). Genes containing a homeobox encode for transcription factors controlling pattern formation, differentiation and proliferation, particularly in haematopoiesis (Lawrence *et al*, 1996) and are expressed in a stage- and lineage-dependent manner. Various studies have seen higher expression levels of *HOXA9*, *HOXA10* and *MEIS1* in immature CD34⁺CD38⁻ normal cell populations compared to more mature, differentiated populations of CD38⁻ cells, where their expression levels have significantly decreased. In AML cells, the expression levels of these three genes do not decrease in the same manner in more differentiated cells, but have been reported to be significantly higher than in the CD34⁺CD38⁻ normal cell population (Kawagoe *et al*, 1999; Lawrence *et al*, 1999). Findings like these strongly support the concept of *HOX* genes playing an important role in the regulation normal haematopoietic development and a disruption to their regulation results in the development of myeloid leukaemia. This contributes to the notion that certain *HOX* genes, for example *HOXA9* and *HOXA10*, have the potential to act as markers of MRD in AML. If these *HOX* family members play a role in haematopoietic development in a stage-dependent manner, as patients achieve a “cure” their levels of

expression should decrease as the immature leukaemic cells are removed and normal cell differentiation is resumed. Conversely, relapse should result in an increase in expression of these genes as the differentiation block is resumed and the cells accumulate in an immature stage of development. Monitoring the changes in their expression would allow for the commencement of any necessary treatment to avoid a haematological relapse from occurring. It is of great importance that genes used as MRD markers in AML show a significant difference in transcription levels between diagnosis and when remission is achieved and should clearly distinguish between the disease state and normal health. In this study the *HOX* genes were over-expressed in the normal karyotype diagnostic AML samples between 4.0 and 37.7 fold compared to the healthy donor samples. As previously mentioned, oligonucleotide microarrays underestimate the fold changes of genes compared to real-time quantitative PCR assays (Yuen *et al*, 2002; Yao *et al*, 2004). Therefore, despite some of these lower fold changes measured by the microarray analysis, several of the *HOX* genes identified may indeed exhibit a larger, and more significant, fold change when investigated by real-time quantitative PCR assays making them more desirable as markers of MRD.

Two of the upregulated probe sets identified in this analysis are already used as markers of MRD in AML; the fms-related tyrosine kinase 3 (*FLT3*) gene and the Wilms tumour (*WT1*) gene. In normal bone marrow cells

FLT3 is preferentially expressed in early progenitors, including CD34⁺ cells and decreases in expression with increasing differentiation (Small *et al*, 1994;Weisel *et al*, 2009). Whilst *FLT3* plays an important role in haematopoiesis, mutations of the *FLT3* gene lead to a pathogenic role in AML. An internal tandem duplication (ITD) of the *FLT3* gene results in a constitutive activation of the FLT3 kinase protein and is associated with a poorer prognosis (Kiyoi *et al*, 1998;Kiyoi *et al*, 1999;Abu-Duhier *et al*, 2000). ITD mutations have been reported to occur in approximately 20-30% of all AML cases and 80-90% of *FLT3*-ITDs are found in normal karyotype AML (Kiyoi *et al*, 1999;Thiede *et al*, 2002;Beretta *et al*, 2004). The role *FLT3* plays in both normal haematopoiesis and in AML and the incidence of *FLT3*-ITD activating mutations occurring in AML have led to various studies investigating the role of a *FLT3*-ITD as a marker of MRD. Beretta *et al* (2004) found *FLT3*-ITD monitoring in eight paediatric AML patients agreed with the clinical course of each patient suggesting the *FLT3*-ITD mutations have the potential to act as MRD markers (Beretta *et al*, 2004). The reappearance of *FLT3*-ITDs with relapse and, so its ability to act as an MRD marker, has been investigated by many groups with varying results; both Nakano *et al* (1999) and Kottaridis *et al* (2002) identified patients presenting with a *FLT3*-ITD at diagnosis but no mutation at relapse and vice versa (Nakano *et al*, 1999;Kottaridis *et al*, 2002). In a much larger cohort of patients, Shih *et al* (2002) found a prevalence of *FLT3*-ITD at relapse than originally identified at diagnosis and a higher

mutational level at relapse than at diagnosis (Shih *et al*, 2002). Despite these results the studies were able to show significant differences between the diagnostic and relapse samples suggesting the potential of *FLT3*-ITD as an MRD marker. In this particular study the *FLT3* gene was identified as over-expressed in normal karyotype AML patients at diagnosis and the single study into the use of *FLT3* expression as an MRD marker is not as positive as the studies into the use of *FLT3* mutations. Kainz *et al* (2005) studied the expression levels of wild-type *FLT3* in 85 AML patients and concluded *FLT3* expression levels were of limited value for the monitoring of MRD. Higher levels of *FLT3* expression were observed in patients with a *FLT3*-ITD but this was not statistically significant (Kainz *et al*, 2005). These studies clearly highlight an involvement of the *FLT3* gene in AML and when identifying MRD markers for normal karyotype AML, it may be necessary to take into account the *FLT3* status of patients. A further step of gene expression profiling analysis between *FLT3* wild-type and *FLT3*-ITD patients may prove beneficial for identifying more specific and sensitive MRD markers for normal karyotype patients.

WT1 was identified as having a mean 14 fold higher expression in the normal karyotype diagnostic samples when compared with normal donor samples. *WT1* was initially identified as a tumour suppressor gene involved in Wilms' tumour, a childhood kidney malignancy (Coppes *et al*, 1993). However, recent studies have identified a potential role for *WT1* as an

oncogene due to its overexpression in a variety of cancers (Yang *et al*, 2007). *WT1* is preferentially expressed in developing genitourinary cells and haematopoietic cells (Ariyaratana & Loeb, 2007). Like the *FLT3* gene, *WT1* is expressed in haematopoietic CD34⁺ early progenitors and decreases in expression with increasing differentiation (Fraizer *et al*, 1995). Very low levels of *WT1* gene expression are observed by RT-PCR in normal haematopoietic cells. In contrast, the expression levels of *WT1* in leukaemias are often as much as ten times greater than those observed in normal bone marrow and progenitors (Inoue *et al*, 1994; Inoue *et al*, 1997). A recent gene expression profiling study by Stirewalt *et al* (2008) identified *WT1* as being over expressed in AML blasts compared to normal haematopoietic cells (Stirewalt *et al*, 2008). Due to the significance of this over expression of *WT1* in AML, it has been investigated as a potential marker of MRD in normal karyotype AML patients. In an early study by Gaiger *et al* (1998), AML patients with detectable *WT1* mRNA at diagnosis and who went on to achieve complete remission following chemotherapy were monitored by PCR. No difference was observed in disease-free survival (DFS) and survival from remission between *WT1*-negative patients and *WT1*-positive patients in complete remission following chemotherapy. The authors therefore concluded that *WT1* was of limited value in MRD monitoring (Gaiger *et al*, 1998). However, several other groups have more recently investigated *WT1* as an MRD marker using real-time quantitative PCR (RQ-PCR) in order to improve sensitivity. They all agree that *WT1*

does have the ability to act as an MRD marker in AML patients (Cilloni *et al*, 2002;Ostergaard *et al*, 2004;Weisser *et al*, 2005).

Several of the genes identified as upregulated in the normal karyotype patient samples were found to interact with each other in biological pathway networks. GeneGO analysis further highlighted the interactions of several of these genes pointing to their potential contribution to the pathology of AML; of the 50 diseases identified by GeneGO, 19 (38%) were haematological malignancies. The statistically significant associations observed between the genes identified in this study and several haematological malignancies pave the way for further promising avenues of research into the identification of key biological processes that underpin the aetiology of AML.

The aim of this chapter was to identify potential MRD markers for normal karyotype AML patients using gene expression profiling. A list of 25 probe sets, relating to 20 genes, were identified as upregulated in the normal karyotype diagnostic AML patient samples compared to healthy donor samples. Using an independent dataset, the strength of the group of upregulated probe sets' ability to classify between normal karyotype AML samples and healthy donor samples was tested and found to be very strong, indicating the robustness of gene expression profiling. Furthermore, previous studies have identified several of these genes, such as the

homeobox and homeobox-related genes, as having some role in leukaemic transformation indicating their importance in AML. The studies by various groups highlighting the roles of these genes in AML, their involvement in particular networks and diseases and their expression levels observed in this current study suggest several may have the potential to act as markers of MRD, which further supports gene expression profiling as a tool for identifying markers of MRD in AML.

Chapter 6

Monitoring MRD in Normal Karyotype AML

6.1 Introduction

6.1.1 MRD Detection

Complete remission (CR) is achieved by 60-80% of AML patients undergoing intensive induction chemotherapy regimens (Kern *et al*, 2005). Despite the relative success of these treatments and the addition of new “targeted” therapies, relapse remains a major obstacle in the fight against leukaemia, occurring in up to 70% of patients undergoing chemotherapy (Liu Yin, 2002b). The purpose of minimal residual disease (MRD) monitoring is to detect and quantify any persistent or relapsing leukaemic blasts, with the objective of eradicating them to prevent a haematological relapse (Chung *et al*, 2006). Knowing the level of MRD in a patient following induction treatment can provide useful information on the course of the disease and can also be used in decisions relating to treatment strategies following induction therapy (Kern *et al*, 2005). A patient’s disease can be monitored through conventional cytogenetics, fluorescence *in situ* hybridisation (FISH), multi-parameter flow cytometry (MFC) and PCR techniques. PCR approaches for MRD detection are very sensitive with methodologies capable of detecting one leukaemic cell in 10^4 - 10^6 normal cells (Campana & Pui, 1995;van Dongen *et al*, 1999;Liu Yin, 2002a). Early MRD studies used qualitative PCR methods. Assays generated either a positive or negative result that was dependent on the assay sensitivity; a negative result did not necessarily translate to an actual absence of residual disease but rather indicated that the threshold of detection of

the assay had been reached (Liu Yin, 2002c). Real-time quantitative PCR (RQ-PCR) is a sensitive method of MRD detection quantifying the levels of PCR product produced as it accumulates rather than an estimate at the end-point of the reaction. Due to the high levels of sensitivity of RQ-PCR, great care must be taken to avoid false-positive resulting from contamination. As with qualitative PCR methods the results generated from RQ-PCR are dependent upon the sensitivity of the protocol (Liu Yin, 2002d).

6.1.2 HOX Genes and MRD

The *HOX* family of Homeobox genes (Class I) encode for DNA-binding transcription factors that act during development. In mammals there are 39 *HOX* genes spread over four different chromosomes in four clusters; A, B, C and D (Argiropoulos & Humphries, 2007). The majority of *HOX* genes (clusters A-C) are preferentially expressed in haematopoietic stem cell-enriched sub-populations and in immature progenitor compartments. These genes are often down-regulated during differentiation and maturation, suggesting a key role in early haematopoietic cells (Sauvageau *et al*, 1994; Kawagoe *et al*, 1999). The observed down-regulation and up-regulation of some *HOX* genes in AML blasts compared to normal bone marrow cells has led to the suggestion that disruption of the normal *HOX* regulatory pathway is extremely important in leukaemogenesis (Lawrence *et al*, 1999; Kawagoe *et al*, 1999; Thompson *et al*, 2003; Argiropoulos &

Humphries, 2007). Expression of the *HOXA9* gene is regularly altered in myelodysplastic syndromes and acute myeloid leukaemia. The t(7;11)(p15;p15) chromosomal translocation results in the fusion of the *NUP98* (nucleoporin, 98kDa) gene on chromosome 11 and the *HOXA9* gene on chromosome 7 (Reviewed by (Romana *et al*, 2006;Moore *et al*, 2007). Kroon *et al* (2001) observed transplanting primary bone marrow cells transduced with a retrovirus expressing the *NUP98-HOXA9* fusion gene induced myeloproliferative disease (MPD) in mice, which developed into AML following a latency period of at least 4 months (Kroon *et al*, 2001). Takeda *et al* (2006) observed impaired myeloid and erythroid differentiation and the induction of proliferation of primitive cells capable of long-term growth following an initial suppression in a population of human CD34⁺ haematopoietic progenitors transduced with a *NUP98-HOXA9* retrovirus (Takeda *et al*, 2006). High levels of *HOX* gene expression have been observed intermediate risk AML patients and those with a normal karyotype leading to the suggestion of the use of *HOX* genes in the clinical management of AML patients (Drabkin *et al*, 2002;Debernardi *et al*, 2003;Roche *et al*, 2004). However, no further investigations into the use of *HOX* genes as markers of MRD have been carried out

6.1.3 WT1

As previously discussed in Chapter 4, *WT1* was originally identified as a tumour suppressor gene in patients with Wilms' tumours, a kidney neoplasm (Haber *et al*, 1990). Many groups have provided evidence

that strongly supports the use of *WT1* gene expression level as an alternative marker of MRD (Cilloni *et al*, 2002;Ostergaard *et al*, 2004;Candoni *et al*, 2009).

6.1.4 COL4A5

The *COL4A5* gene, encoding for the $\alpha 5$ (IV) chain of type IV collagen, is co-localised at chromosome Xq22 with the *COL4A6* gene (Sugimoto *et al*, 1994;Oohashi *et al*, 1994). Mutations in the *COL4A5* gene result in Alport Syndrome; patients present with a progressive hereditary nephritis often combined with a sensorineural deafness. To date, 38 large/medium-sized deletions and 138 small mutations have been identified in the *COL4A5* gene (as reviewed by (Lemmink *et al*, 1997). Previous gene expression profiling studies have identified an over-expression of *COL4A5* in the yolk sac tumour elements of germ cell tumours (GCT) and in solitary fibrous tumours (SFT) but to date an over-expression of *COL4A5* in a haematological malignancy has not been recorded (Juric *et al*, 2005;West *et al*, 2005).

The aim of this chapter was to use real-time quantitative PCR (RQ-PCR) to evaluate the candidate genes identified previously through gene expression profiling (GEP) (Chapter 5) in a group of normal karyotype diagnostic AML patients and normal healthy donors to validate the GEP analysis and monitor these genes in diagnostic and follow-up patient samples to assess their ability to act as markers of MRD.

6.2 Materials and Methods

6.2.1 Patient and Donor samples

6.2.1a Diagnostic Screening of Candidate Genes

Cell lysates and RNA from bone marrow aspirates of 15 healthy donors and 52 normal karyotype AML patients at diagnosis were obtained from the School of Medicine, Cardiff University (Table 1). Mononuclear cells were isolated from the bone marrow aspirates as described in section 2.3.3 (Chapter 2) and RNA was extracted using TRIzol reagent or the Qiagen RNeasy Mini Kit as described in section 2.4.1 (Chapter 2).

6.2.1b MRD Monitoring by RQ-PCR

Diagnostic and follow-up RNA from 6 normal karyotype AML patients entered into the AML15 trial were obtained from the School of Medicine, Cardiff University and processed for MRD monitoring by RQ-PCR (Table 2). The diagnostic and follow-up RNA samples were processed as described above.

Characteristic	Value	
Sex - no. (%)		
Male	27	(51.9)
Female	25	(48.1)
Not determined	0	(0)
Age group – no. (%)		
<30 years	4	(7.7)
30-60 years	19	(36.5)
>60 years	29	(55.8)
Age – years		
Median	54	
Range	12-78	
French-American British Classification – no. (%)		
M0	2	(3.8)
M1	6	(11.5)
M2	14	(26.9)
M3	0	(0)
M4	7	(13.5)
M5	9	(17.4)
M6	2	(3.8)
M7	0	(0)
Other	2	(3.8)
Not determined	10	(19.3)
Molecular Abnormalities – no. (%)		
Mutation		
FLT3 Internal tandem duplication	19	(36.5)
FLT3 Tyrosine kinase domain	3	(5.8)
NPM1	20	(38.5)
ITD ⁺ /NPM1 ⁺	10	(19.2)
ITD ⁺ /NPM1 ⁻	6	(11.5)
ITD ⁻ /NPM1 ⁺	9	(17.3)
ITD ⁻ /NPM1 ⁻	14	(26.9)

Table 1. The frequencies and percentages of the clinical and molecular characteristics of the 52 patients diagnosed with normal karyotype AML in the sample set for RQ-PCR gene expression validation.

Characteristic	Value	
Sex - no. (%)		
Male	3	(50.0)
Female	3	(50.0)
Not determined	0	(0)
Age group – no. (%)		
<30 years	1	(16.7)
30-60 years	5	(88.3)
>60 years	0	(0)
Age – years		
Median	42	
Range	17-47	
French-American British Classification – no. (%)		
M0	0	(0)
M1	3	(50.0)
M2	2	(33.3)
M3	0	(0)
M4	1	(16.7)
M5	0	(0)
M6	0	(0)
M7	0	(0)
Other	0	(0)
Not determined	0	(16.7)
Molecular Abnormalities – no. (%)		
Mutation		
FLT3 Internal tandem duplication	4	(66.7)
FLT3 Tyrosine kinase domain	0	(0)
FLT3 Not determined	0	(0)
NPM1 Mutant	4	(66.7)
NPM1 Not determined	1	(16.7)

Table 2. The frequencies and percentages of the clinical and molecular characteristics of the 6 normal karyotype AML patients in the sample set for MRD detection.

6.2.2 Quantitative Real-Time PCR

6.2.2a Gene Expression Validation

RQ-PCR was performed using the Roche LightCycler system as described in section 2.5 (Chapter 2). Each sample was run in duplicate and a duplicate “no-RNA” cDNA control sample and “no-cDNA” control sample was included. The Abelson (*ABL*) gene was used as the control house-keeping gene to evaluate the quality and quantity of the starting RNA and correct for any variations in expression arising from this between samples. *ABL* is a proto-oncogene that encodes for a nuclear and cytoplasmic tyrosine kinase protein. In the “A Europe Against Cancer” Programme, the *ABL* gene transcript was found to be a reliable control gene for the comparison of diagnostic and MRD samples; a similar expression of the *ABL* gene transcript was seen in both the normal and the diagnostic samples. The correlation between the *ABL* gene expression and the expression of the candidate genes were the highest observed amongst the fourteen control genes tested (Beillard *et al*, 2003; Gabert *et al*, 2003).

6.2.2b MRD Monitoring

RQ-PCR was performed on cDNA generated from AML diagnostic samples and healthy donor samples to monitor expression of the candidate genes in the patient samples. As with the diagnostic screening, the Ct values generated (see Chapter 2, Section 2.5.3) for the candidate genes were normalised to the Ct values of the control

gene *ABL*. Samples with Ct values greater than 40 cycles were deemed as “not determined”; i.e. “negative”.

6.3 Results

6.3.1 Validation of Identified Genes by RQ-PCR

Previously, gene expression analysis was applied to a dataset of normal karyotype AML diagnostic samples and healthy donor samples to identify differentially expressed genes between the two groups. From this analysis 25 probe sets relating to 20 genes were identified as up-regulated in the normal karyotype AML diagnostic samples compared to the normal healthy donor samples (Chapter 5). In order to establish if these genes have the potential to act as markers of MRD in normal karyotype AML, RQ-PCR was performed to screen selected genes across a panel of normal karyotype AML diagnostic patient samples and healthy donor samples. Probe sets with a greater level of expression (>10 fold) in the normal karyotype AML samples compared to the healthy donor samples were selected for gene expression screening in patient samples at diagnosis. The probe sets related to five genes; *HOXA5*, *HOXA9*, *HOXB6*, *WT1* and *COL4A5* (Chapter 5, Table 3).

6.3.1a Diagnostic Screening of Candidate Genes

Expression of the candidate gene transcripts was measured in diagnostic normal karyotype AML patient samples and compared to

levels measured in healthy normal donor samples. Initially primer sets were designed for the amplification of the five candidate genes as described in the Materials and Methods section 2.5.1 (Chapter 2) and using serial dilutions of cDNA, their amplification efficiencies were compared to that of the control gene, *ABL* (Figure 1). The amplification curves generated from the cell line cDNA for the five candidate genes were similar to those generated for the control gene *ABL* across the dynamic range of the cDNA template dilutions used. Therefore *ABL* was considered suitable as the normalising control gene (Figure 1).

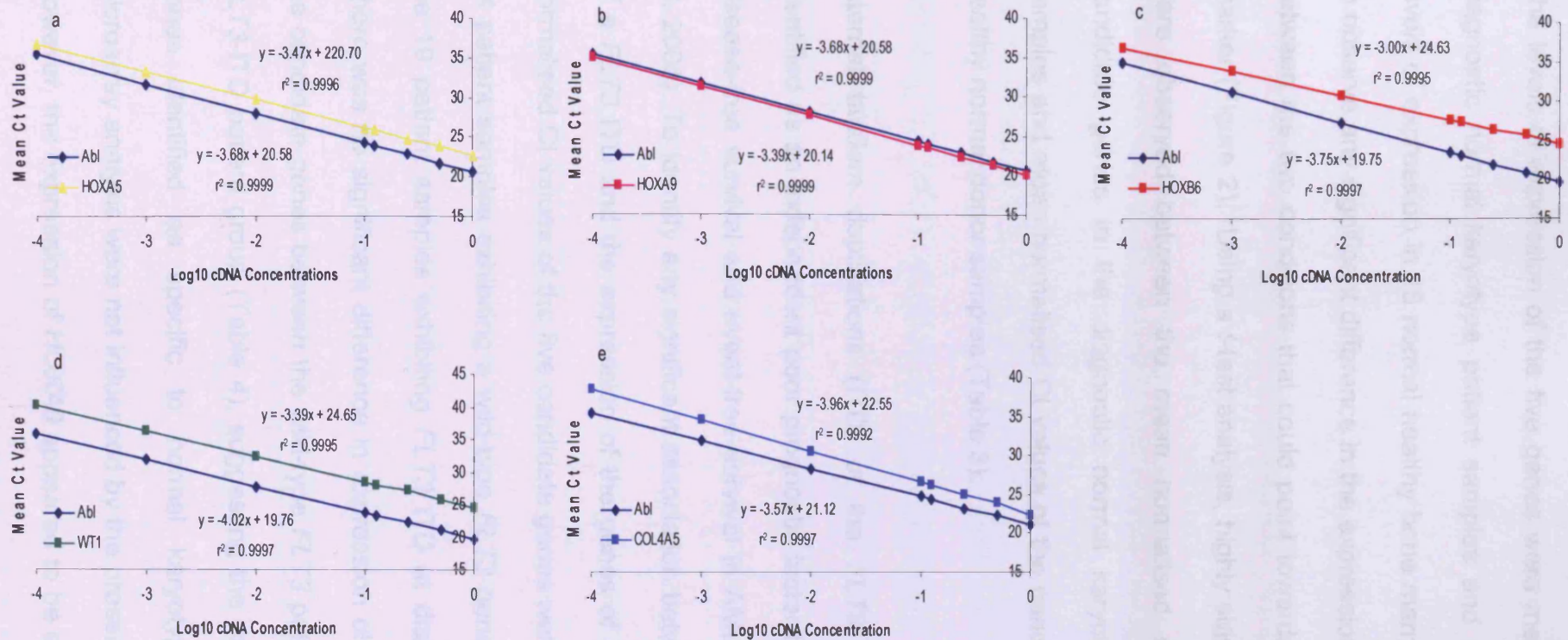


Figure 1. The amplification curves of the primer sets used for *HOXA5* (a), *HOXA9* (b), *HOXB6* (c), *WT1* (d) and *COL4A5* (e) and of the control gene, *ABL*, generated from serial dilutions of cell line cDNA to compare their amplification efficiencies.

The levels of expression of the five genes were measured in 52 AML diagnostic normal karyotype patient samples and compared to their levels of expression in 15 normal healthy bone marrow donor samples to observe any significant difference in the expression of the five genes between the two conditions that could point towards a potential MRD marker (Figure 2). Using a *t*-test analysis, highly significant differences were observed between the mean normalised Ct values of the candidate genes in the diagnostic normal karyotype AML patient samples and mean normalised Ct values of the candidate genes in the healthy normal donor samples (Table 3).

Internal tandem duplications (ITD) of the *FLT3* gene have been identified as an independent poor prognostic factor for overall survival, disease-free survival and event-free survival in AML patients (Kiyoi *et al*, 2005). To identify any significant association between the presence of a *FLT3* ITD and the expression of the genes of interest, the mean normalised Ct values of the five candidate genes were measured in the 24 patient samples exhibiting a wild-type *FLT3* gene at diagnosis and the 19 patient samples exhibiting *FLT3* ITD at diagnosis (Figure 3). There was no significant difference in expression observed for four of the candidate genes between the wild-type *FLT3* patient group and the *FLT3*-ITD patient group (Table 4), suggesting the expression of these genes identified as specific to normal karyotype AML through microarray analysis were not influenced by the presence of a *FLT3* ITD. However, the expression of *HOXA9* appeared to be associated with the

presence of *FLT3* ITD. This finding requires further investigation in a larger cohort of normal karyotype AML samples.

Mutation of the *NPM1* gene in normal karyotype AML patients is associated with a favourable outcome both in terms of event-free survival and overall survival (Chen *et al*, 2006). To identify any significant associations between the mutational status of *NPM1* and the expression of the candidate genes, the mean normalised Ct values of the five candidate genes were also measured in the 20 patient samples exhibiting a wild-type *NPM1* gene at diagnosis and 19 patient samples exhibiting an *NPM1* mutation at diagnosis (Figure 4). There was no significant difference observed between the wild-type *NPM1* patient group and the mutated *NPM1* patient group for *WT1* and *COL4A5*. However, a significant difference between the *NPM1* wild-type and mutant samples was observed for *HOXA5*, *HOXA9* and *HOXB6* (Table 5) implying that the expression of these genes may be modulated by *NPM1* status. Once again, this potentially interesting finding would need to be examined prospectively in a larger cohort of patients.

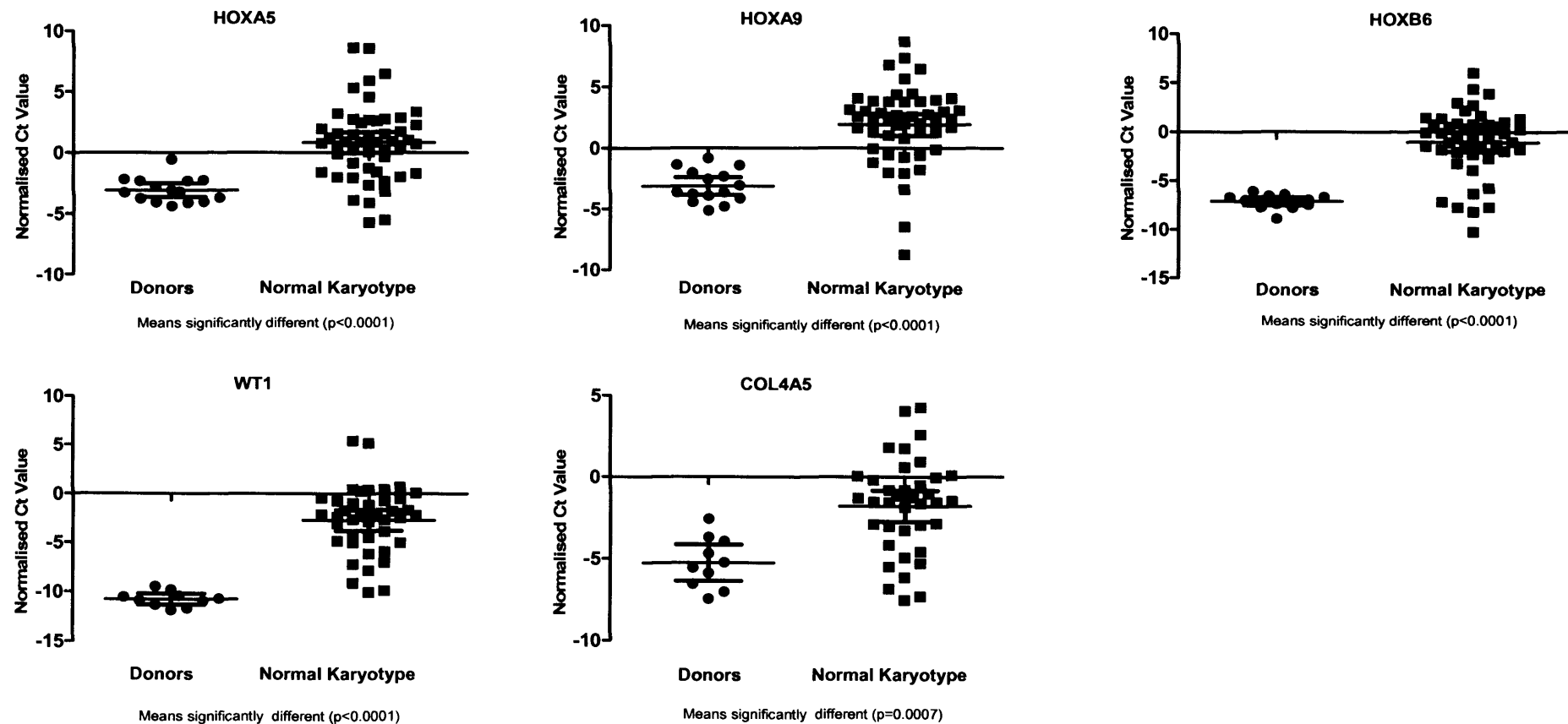


Figure 2. Scatter plots of the spread of the normalised Ct values of the candidate genes measured in the normal karyotype AML diagnostic patient samples and the healthy normal donor samples. Mean value $\pm$ 95% confidence interval plotted.

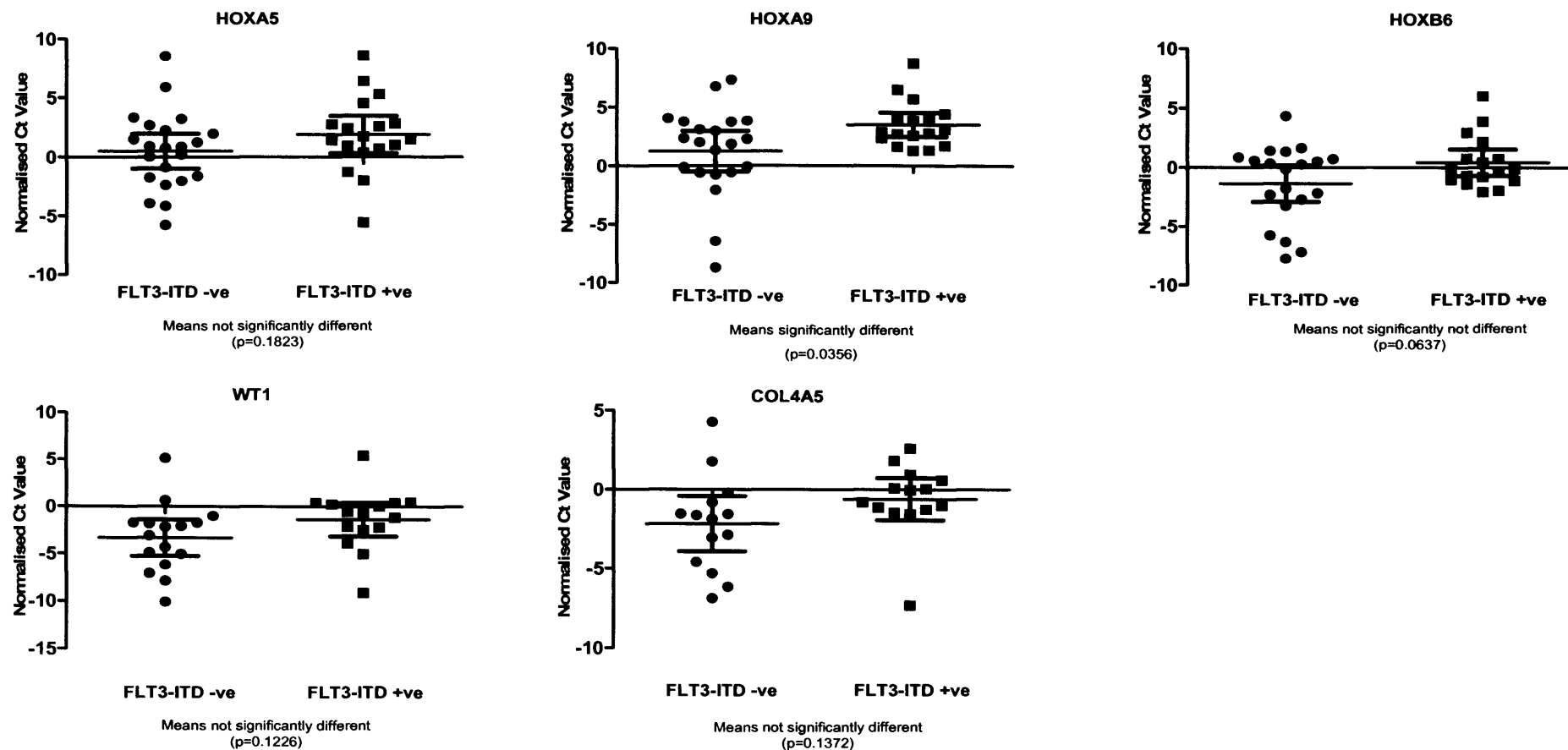


Figure 3. Scatter plots of the spread of the normalised Ct values of the candidate genes measured in the *FLT3*-ITD positive normal karyotype AML diagnostic patient samples (*FLT3*-ITD +ve) and the *FLT3*-ITD negative normal karyotype AML diagnostic patient samples (*FLT3*-ITD -ve). Mean value $\pm$ 95% confidence interval plotted.

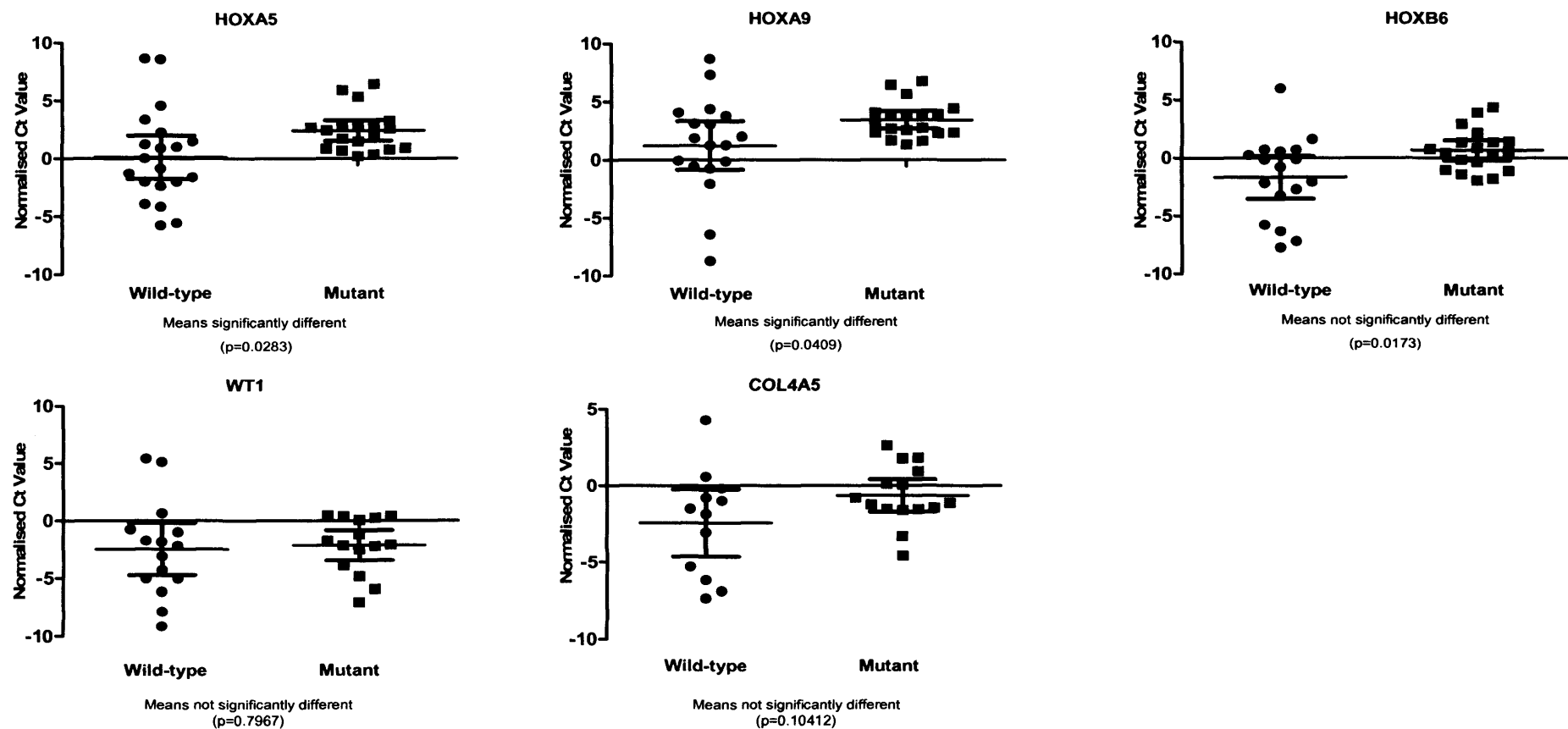


Figure 4. Scatter plots of the spread of the normalised Ct values of the candidate genes measured in the mutant *NPM1* normal karyotype AML diagnostic patient samples (Mutant) and the wild-type *NPM1* normal karyotype AML diagnostic patient samples (Wild-type). Mean value $\pm$ 95% confidence interval plotted.

	HOXA5	HOXA9	HOXB6	WT1	COL4A5
P Value	<0.0001	<0.0001	<0.0001	<0.0001	0.0007
Are the means significantly different (p<0.05)	Yes	Yes	Yes	Yes	Yes
Donors mean value	-3.070	-3.130	-7.104	-10.770	-5.266
Donors 95% CI	-3.64 to -2.50	-3.86 to -2.40	-7.51 to -6.69	-11.3 to -10.2	-6.38 to -4.15
Donors Expression Range	-4.35 to -0.56	-5.12 to -0.82	-8.89 to -6.05	-11.89 to -9.47	-7.46 to -2.57
NK mean value	0.820	1.907	-1.056	-2.749	-1.785
NK 95% CI	-0.05 to 1.69	0.99 to 2.82	-2.04 to -0.08	-3.83 to -1.67	-2.74 to -0.83
NK Expression Range	-5.77 to 8.64	-8.71 to 8.71	-10.24 to 6.01	-10.07 to 5.41	-7.55 to 4.28

Table 3. The statistical significance of the mean Ct values, normalised to the Ct values of *ABL*, of the candidate genes measured in the 52 normal karyotype (NK) AML diagnostic patient samples and the 15 healthy normal donor samples. CI = confidence interval.

	HOXA5	HOXA9	HOXB6	WT1	COL4A5
P Value	0.1823	0.0356	0.0637	0.1226	0.1372
Are the means significantly different (p<0.05)	No	Yes	No	No	No
ITD +ve mean value	1.644	3.226	0.196	-1.719	-1.068
ITD +ve 95% CI	0.32 to 3.51	2.213	-0.69 to 1.56	-3.13 to 0.42	-1.94 to 0.73
ITD +ve Expression Range	-5.58 to 8.64	-1.19 to 8.71	-3.89 to -6.01	-9.14 to 5.41	-7.55 to 2.61
ITD -ve mean value	0.6021	1.457	-1.045	-3.303	-1.846
ITD -ve 95% CI	-0.96 to 1.97	-0.47 to 2.99	-2.90 to 0.18	-5.25 to -1.38	-3.91 to -0.42
ITD -ve Expression Range	-5.77 to 8.58	-8.71 to 7.35	-7.72 to 4.34	-10.07 to 0.65	-6.87 to 5.13

Table 4. The statistical significance of the mean Ct values, normalised to the Ct values of *ABL*, of the candidate genes measured in the 19 *FLT3*-ITD positive NK AML diagnostic patient samples (ITD +ve) and the 25 *FLT3*-ITD negative NK AML diagnostic patient samples (ITD -ve). CI = confidence interval.

	HOXA5	HOXA9	HOXB6	WT1	COL4A5
P Value	0.0283	0.0409	0.0173	0.7967	0.1041
Are the means significantly different (p<0.05)	Yes	Yes	Yes	No	No
Mutant mean value	2.425	3.447	0.6705	-2.129	-0.664
Mutant 95% CI	1.55 to 3.30	2.69 to 4.20	-0.19 to 1.53	-3.46 to -0.80	-1.74 to 0.41
Mutant Expression Range	0.77 to 5.34	1.34 to 6.80	-1.77 to 4.34	-7.06 to 0.45	-4.58 to 2.61
Wild-type mean value	0.1270	1.243	-1.652	-2.450	-2.434
Wild-type 95% CI	-1.75 to 2.01	-0.86 to 3.03	-3.49 to 0.19	-4.74 to -0.16	-4.63 to -0.24
Wild-type Expression Range	-5.77 to 8.634	-8.71 to 8.71	-7.72 to 6.01	-9.14 to 5.41	-7.33 to 4.28

Table 5. The statistical significance of the mean Ct values, normalised to the Ct values of *ABL*, of the candidate genes measured in the 19 mutant *NPM1* NK AML diagnostic patient samples (Mutant) and the 20 wild-type *NPM1* NK AML diagnostic patient samples (Wild-type). CI = confidence interval.

Based on these findings, the expression levels of these five genes were monitored in AML patient diagnostic and follow-up samples to identify whether any had the potential to act as a marker of MRD for normal karyotype AML.

6.3.1b Candidate Genes as MRD Markers

The transcript expression levels of the five genes were monitored in the diagnostic sample and follow-up samples of six AML patients with a normal karyotype.

The fold change of gene transcript expression levels relative to *ABL* were calculated as described in Chapter 3 and compared to the upper fold change in expression observed in the healthy normal donor samples (Figure 5). The MRD values (MRDv) of the genes in the patient samples were calculated as described in Chapter 3 and compared to the level of transcript at diagnosis (Figure 6). Patients 1 to 4 achieved and remained in complete remission. Patient 5 relapsed at 252 days following diagnosis and patient 6 relapsed at 125 days following diagnosis.

i. HOX Genes

The fold changes in expression of the three *HOX* genes relative to *ABL* displayed inconsistent changes following induction therapy to what would be expected from the status of six patients of complete remission (CR) or relapse; both increases and decreases in fold changes in expression of the three genes were detected following induction therapy in the four patients achieving complete remission and in the two that relapsed. These fold changes in expression were within two logs of that detected diagnosis bar for *HOXA9* in one patient; an increase by approximately three logs following induction therapy in a patient that achieved complete remission. The level of expression of the three genes in the two patients that relapsed remained above the upper levels detected in healthy normal donors for the disease length. Again, inconsistency was observed for the four patients achieving CR in

regards to their gene levels relative to the upper levels detected in healthy normal donors (Figure 5a, 5b and 5c).

ii. WT1

Expression data was only available for three normal karyotype patients that achieved CR and one normal karyotype patient that relapsed.

The levels of *WT1* expression of all four patients at diagnosis was between one and two logs higher than the upper level of *WT1* expression detected in healthy normal donor samples. Following induction therapy the fold change in expression of *WT1* in the three CR patients decreased and continued so for the remainder of their disease. The fold change in expression of *WT1* in the patient that relapsed displayed an expected behaviour of decreasing following the start of induction therapy before increasing 90 days prior to a haematological relapse by more than 1 log (Figure 5d).

iii. COL4A5

The fold change in *COL4A5* expression relative to *ABL* of the four patients that achieved CR displayed similar expression patterns as the three *HOX* genes in these patients; inconsistent fold changes in the level of expression following the start of therapy by approximately one log from the levels detected at diagnosis. Again, inconsistent fold changes in expression following diagnosis and the start of therapy were detected in the two patients that relapsed (5 and 6). All four patients

that achieved CR displayed expression levels below the upper level measured in the healthy donor samples (Figure 5e).

WT1 was the only candidate gene to display a fall in its level of transcription in all patients monitored; increases **and** decreases in levels of *HOXA5*, *HOXA9*, *HOXB6* and *COL4A5* from their transcript level at diagnosis were observed in both the patients that achieved CR **and** the patients that relapsed. *WT1* also displayed the largest increase in transcript level following the initial decrease following induction therapy in the patient that relapsed; 2.5 logs compared to the one log increase observed for the three *HOX* genes (Figure 6).

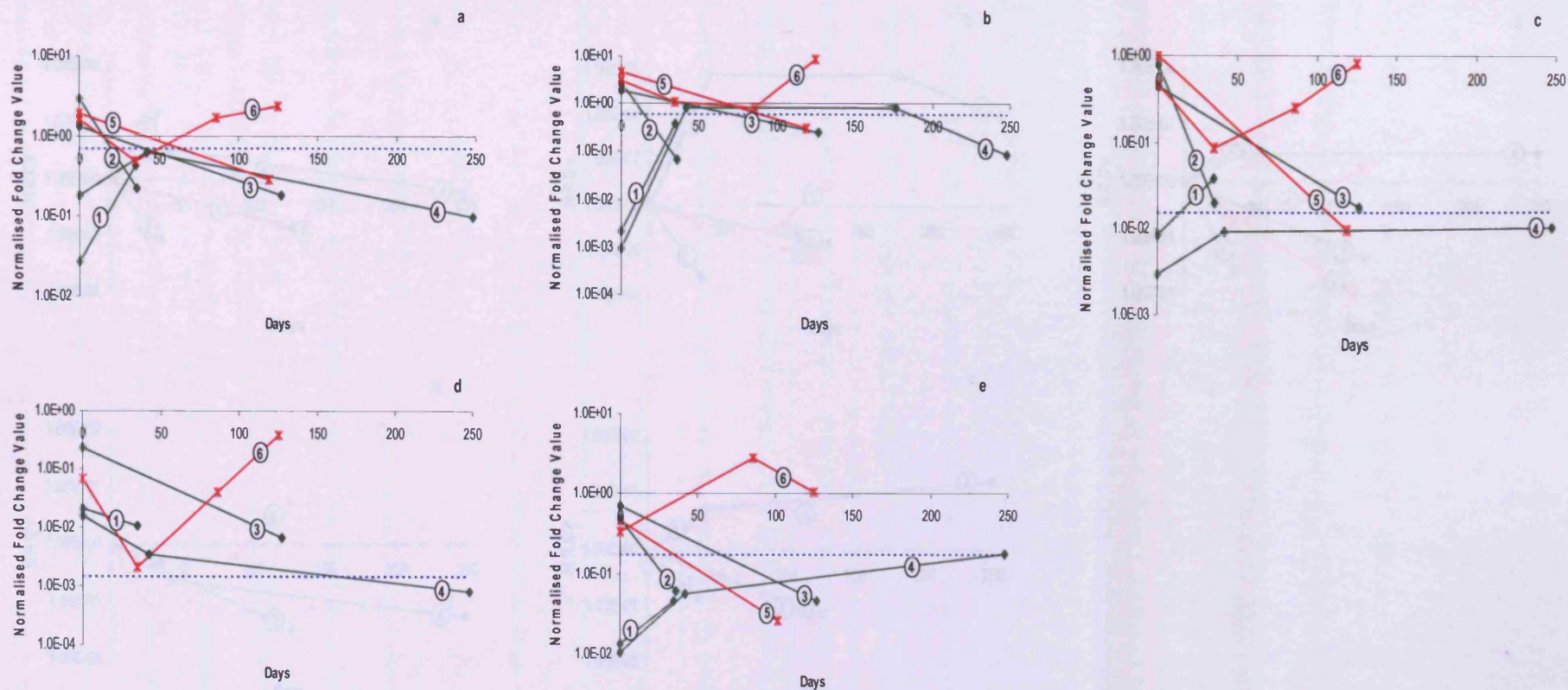


Figure 5. Longitudinal monitoring of *HOXA5* (a), *HOXA9* (b), *HOXB6* (c), *WT1* (d) and *COL4A5* (e) in six normal karyotype AML patients. The graphs show the fold change values generated for respective genes in each sample, relative to the *ABL* gene. Only four patient sample sets were available for monitoring of *WT1*. Red lines - patients who relapsed. Green lines - patients who remained in complete remission (CR). The blue dashed line represents the upper limit of gene expression in the healthy normal donor samples. Patient 5 relapsed at 252 days following diagnosis; patient 6 relapsed at 125 days following diagnosis.

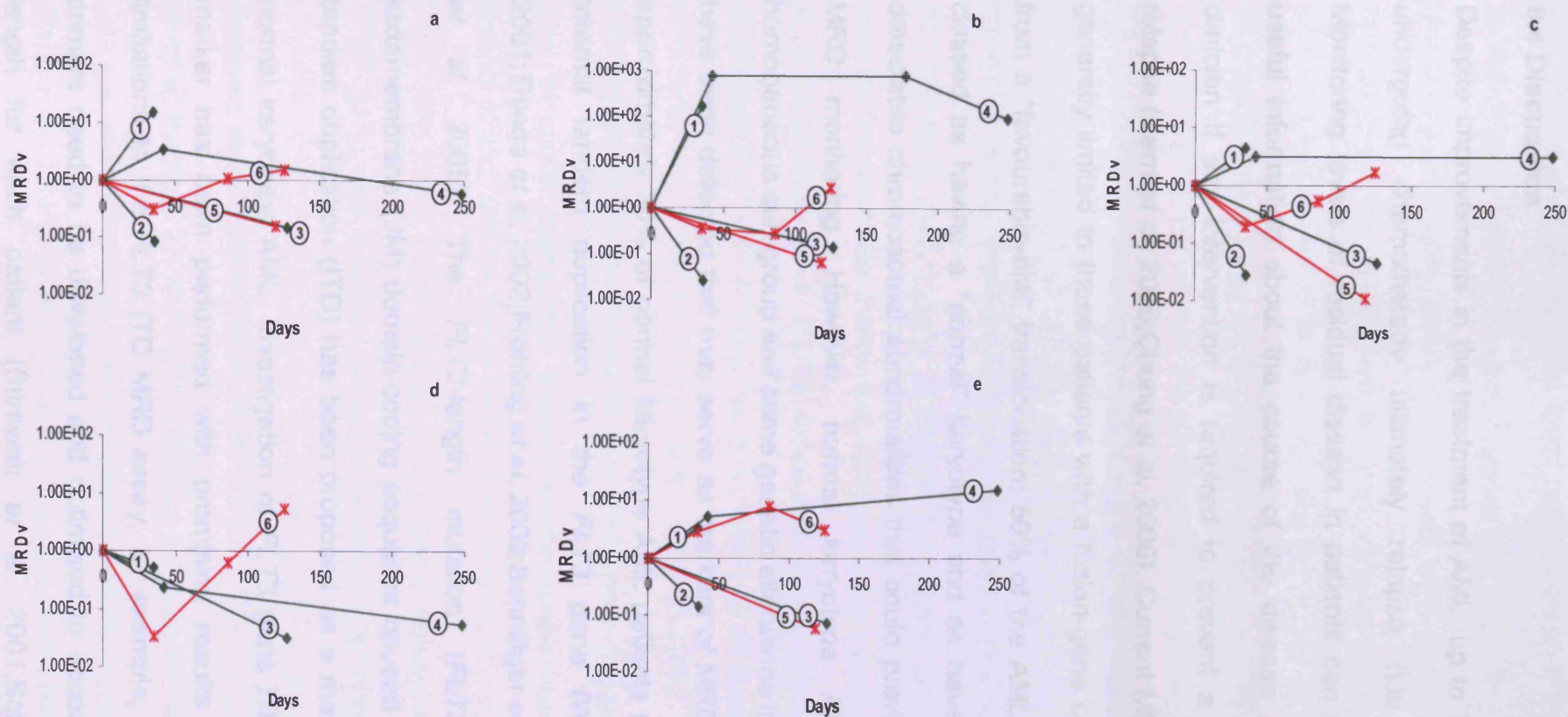


Figure 6. Longitudinal monitoring of *HOXA5* (a), *HOXA9* (b), *HOXB6* (c), *WT1* (d) and *COL4A5* (e) in six normal karyotype AML patients. The graphs show the MRD values calculated for the genes in each sample. Only four patient sample sets were available for monitoring of *WT1*. Red lines - patients who relapsed. Green lines - patients who remained in complete remission (CR). Patient 5 relapsed at 252 days following diagnosis; patient 6 relapsed at 125 days following diagnosis.

6.4 Discussion

Despite improvements in the treatment of AML, up to 70% of patients undergoing chemotherapy ultimately relapse (Liu Yin, 2002e). Monitoring levels of residual disease in patients can help to provide useful information about the course of the disease and inform the clinician if any intervention is required to prevent a haematological relapse (Kern *et al*, 2005; Chung *et al*, 2006). Current MRD markers are generally limited to those patients with a fusion-gene product resulting from a “favourable-risk” translocation; 50% of the AML population are classed as having a “normal” karyotype and so have no underlying detectable chromosomal abnormalities that could provide markers for MRD monitoring. However, normal karyotype AML is not a homogeneous sub-group and some genetic alterations in specific genes have been detected that may serve as markers of MRD. For example, approximately 30% of normal karyotype AML patients present with an internal tandem duplication in the *FLT3* gene (Whitman *et al*, 2001; Thiede *et al*, 2002; Frohling *et al*, 2002; Schnittger *et al*, 2005; Bienz *et al*, 2005). The *FLT3*-length mutation (*FLT3*-LM) in the juxtamembrane (JM) domain-coding sequence caused by the internal tandem duplication (ITD) has been proposed as a marker of MRD in normal karyotype AML. Investigation of *FLT3* gene ITDs as an MRD marker has been performed with promising results but there are limitations to the *FLT3* ITD MRD assay, for example, patient-specific primers need to be developed and optimised to account for the ITD length for each patient (Stirewalt *et al*, 2001; Schnittger *et al*,

2002;Scholl *et al*, 2005). The *NPM1* (nucleophosmin 1) gene is partner to many other genes in translocations occurring in leukaemias and lymphomas, including the *ALK* (anaplastic lymphoma receptor tyrosine kinase) gene, the *RAR α* (retinoic acid receptor alpha) gene and the *MLF1* (myeloid leukaemia factor 1) gene (Morris *et al*, 1994;Redner *et al*, 1996;Yoneda-Kato *et al*, 1999). Normally shuttling between the nucleus and the cytoplasm, the NPM1 protein aberrantly localises in the cytoplasm when a mutation occurs in the *NPM1* gene (Quentmeier *et al*, 2005;Wang *et al*, 2005). Many groups have revealed *NPM1* mutations in AML patients especially in those with a normal karyotype and therefore, *NPM1* gene mutations have been predicted to be an ideal marker for MRD (Falini *et al*, 2005;Schnittger *et al*, 2005;Boissel *et al*, 2005;Verhaak *et al*, 2005;Dohner *et al*, 2005). Gorello *et al* (2006) developed an RQ-PCR assay for both RNA and DNA to monitor *NPM1* mutations in normal karyotype AML adult patients. In 13 patients carrying mutations in exon 12 of the *NPM1* gene, the number of *NPM1* type A-mutated transcripts closely correlated with tumour burden at diagnosis and with response to induction therapy (Gorello *et al*, 2006). Due to the small patient cohort in this study further validation of the *NPM1* gene was indicated. Chou *et al* (2007) performed RQ-PCR on samples from 38 AML patients bearing an *NPM1* gene mutation and observed more mutant copies of the *NPM1* gene transcript in samples taken at the first morphological CR from the patients with subsequent clinical relapse than those from the patients without subsequent relapse (Chou *et al*, 2007). Whilst the use of the type A mutation in the *NPM1*

gene as an MRD marker for normal karyotype AML is promising, an *NPM1* mutation is detected in only approximately 50% of patients with a normal karyotype (Falini *et al*, 2005; Boissel *et al*, 2005; Schnittger *et al*, 2005; Dohner *et al*, 2005; Thiede *et al*, 2006). The association of an *NPM1* mutation with a favourable outcome may also influence its ability as a reliable marker of MRD (Chen *et al*, 2006).

Previously gene expression profiling was used to identify potential markers of MRD in AML patients with a normal karyotype (Chapter 5). Twenty genes were identified as being highly expressed in the normal karyotype patients compared to healthy normal donors; five genes, represented by probe sets, with a greater expression in the normal karyotype AML samples than the healthy donor samples by more than 10 fold. Initially the *HOXA5*, *HOXA9*, *HOXB6*, *WT1* and *COL4A5* genes were screened in 52 diagnostic normal karyotype AML patient samples and 15 healthy normal donor samples to identify their potential to act as markers of MRD. A statistically significant difference was observed for the increase in mean normalised expression levels for all five genes in the diagnostic normal karyotype AML samples compared to the healthy normal donor samples; the largest difference in mean expression level between the two groups was observed for the *WT1* gene. Whilst not solely restricted to the study of normal karyotype AML, many previous studies have shown an aberrant expression of the *WT1* gene and the *HOX* genes in AML patients lacking a specific fusion gene. A recent study by Stirewalt *et al* (2008), in which Affymetrix HG-U133A gene

chips were initially used to identify AML-specific gene expression changes, identified the *WT1* gene as AML-specific. Validating their data with quantitative RT-PCR, a significant increase in *WT1* expression was observed in AML patients compared to normal hematopoietic cell donors (Stirewalt *et al*, 2008) corroborating earlier studies by Inoue *et al* (1997) where it was also observed the *WT1* expression levels in leukaemic cells were at least one log higher than in normal hematopoietic progenitor cells (Inoue *et al*, 1997). Qin *et al* (2009) went a step further to investigate the expression levels of *WT1*, along with the *PRAME* gene (previously mentioned in chapter 4), in patients lacking a specific fusion gene. They found that *WT1* was detected at minimal levels in normal bone marrow samples but was detected at its highest levels in those patients lacking a specific fusion gene compared to patients with a specific fusion gene (Qin *et al*, 2009). In support of the significant increase in the expression of *HOXA5*, *HOXA9* and *HOXB6* in normal karyotype patients in this study, a study by Roche *et al* (2004) used Q-RT/PCR with SYBR-green to group AML patients into clusters marked by the expression of low, intermediate and high *HOX* expression. They identified 94% of the patients in the high *HOX* expression group as belonging to the intermediate cytogenetic group i.e. no specific fusion genes (Roche *et al*, 2004). In particular, Giampaolo *et al* (2002) observed absent *HOXB6* expression in almost all AML samples with fusion genes resulting to the commonest translocations compared to the AML samples lacking chromosomal translocations that were *HOXB6*⁺ (Giampaolo *et al*, 2002). Whilst these

studies were not specifically investigating normal karyotype AML, patients with a normal karyotype are placed in the intermediate cytogenetics/risk group and therefore the results of these studies do bear some significance with and support the findings in this particular section; the *WT1*, *HOXA5*, *HOXA9* and *HOXB6* genes are significantly increased in expression in diagnostic normal karyotype AML patients compared to their healthy normal counterparts. To date there is no literature on *COL4A5* expression in AML. In order to identify the true potential of *COL4A5* as a marker of MRD or just an over-expression with no clinical relevance, the expression of the *COL4A5* gene needs to be followed in longitudinal patient samples.

In agreement with previous studies, 36.5% of the 52 normal karyotype AML samples presented with a *FLT3*-ITD whilst the incidence of an *NPM1* mutation in these patients at diagnosis (38.5%) was slightly lower than the incidences of approximately 50% previously reported by other groups (Whitman *et al*, 2001;Frohling *et al*, 2002;Dohner *et al*, 2005;Falini *et al*, 2005;Boissel *et al*, 2005;Schnittger *et al*, 2005). Using *t*-tests, no statistically significant difference in the mean normalised expression levels of the genes between the *FLT3*-ITD normal karyotype patient samples and the wild-type *FLT3* normal karyotype patient samples was observed for 4 of the genes; only a small correlation was observed for *HOXA9* ($p=0.0356$). Unfortunately the *FLT3* results disagree with the study by Roche *et al* (2004) in which approximately twice the number of *FLT3* mutations occurred in high *HOX* expression

AML cases compared to low or intermediate *HOX* expression AML cases (Roche et al, 2004). However, these findings were not found to be statistically significant and were extrapolated from AML cases across the sub-group spectrum rather than just those with intermediate cytogenetics. However, a statistically significant difference in the normalised mean expression levels of *HOXA5*, *HOXA9* and *HOXB6* was observed between the mutant *NPM1* normal karyotype patients and the wild-type *NPM1* normal karyotype patients. Previous microarray studies have demonstrated distinct expression patterns of homeobox genes in normal karyotype AML displaying *NPM1* mutations. Verhaak et al (2005) identified a positive correlation between AML normal karyotype and mutations of *NPM1* from the analysis of 275 primary AML cases. From a cohort of 58 patients presenting with aberrant cytoplasmic NPM1 localization (NPMc⁺), Alcalay et al (2005) identified 52 patients exhibiting both a normal karyotype and an *NPM1* mutation. Microarray analysis of this patient group demonstrated a striking gene expression pattern highlighting the up-regulation of several homeodomain-containing genes, including the *HOXA9* and *HOXB6* genes selected for further investigation this current study (Alcalay et al, 2005). In a later study, Andreeff et al (2008) went on to use quantitative RT-PCR to analyse 42 genes, including 39 homeodomain genes, in 119 newly diagnosed AML samples. They detected low levels of *HOXA* and *HOXB* gene expression in favourable risk AML yet a marked over-expression of these genes was observed in AML cases with intermediate cytogenetics and *NPM1* mutations (Andreeff et al, 2008).

After identifying the significant difference in expression of the five genes in normal karyotype AML patients at diagnosis compared to the healthy normal bone marrow samples, the genes were monitored in six normal karyotype AML patients to observe their ability to monitor the disease. The expression levels of the *HOXA5*, *HOXA9*, *HOXB6* and *COL4A5* transcripts displayed no consistent correlation with the patients' disease status. Following diagnosis and the start of treatment the expression levels of the four gene transcripts decreased in two of the four of the patients remaining in complete remission (CR) but increased in the remaining two CR patients. An increase in gene transcript level of the three *HOX* genes was observed in one of the two patients that relapsed, but the increase seen was less than one log; not significant enough to predict a potential molecular relapse. From their studies in CBF leukaemias, Lane *et al* (2008) suggested that a $\geq 1 \log_{10}$ rise in RQ-PCR transcript levels observed in serial remission bone marrow samples is a clinically useful definition of and reliably predicts a morphologic relapse (Lane *et al*, 2008). Currently there has been little investigation into the monitoring of *HOX* gene expression as a marker of MRD in haematological malignancies. Riedt *et al* (2009) investigated the caudal (*CDX*) family of the homeobox genes specifically in acute lymphoblastic leukaemia (ALL). They observed an expression of the *CDX2* (caudal type homeobox 2) gene in the mononuclear cells of 9 out of 10 AML patients compared to no expression detected in healthy donor bone marrow and peripheral blood cells. They also identified low or negative *CDX2* expression in only 16% of paediatric ALL patient

samples comparable to levels in healthy donors. Measuring the expression of *CDX2* at diagnosis and at two subsequent time points in ALL patients led to the suggestion that the expression levels of *CDX2* at the manifestation of the disease may be used to foresee the response to treatment of patients that are MRD-positive at day 33 following diagnosis (Riedt *et al*, 2009). Scholl *et al* (2007) detected *CDX2* expression in 90% of their AML patient cohort yet no expression in haematopoietic stem and progenitor cells from healthy normal donors; the second highest levels of *CDX2* expression identified in patients with a normal karyotype. They found that the level of *CDX2* expression in 2 patients over the course of their disease correlated with the disease burden (Scholl *et al*, 2007a). *CDX2* regulates expression of *HOX* genes during embryogenesis and developmental haematopoiesis (Deschamps *et al*, 1999; Abramovich & Humphries, 2005; Scholl *et al*, 2007b). Whilst it is apparent that *HOX* genes and their regulators have a prominent role in leukaemogenesis, the *HOX* genes in this study do not appear to have the potential to act as MRD markers despite their increased expression levels at diagnosis. It may be the answer lies within other *HOX* family members. From the original analysis, probe sets representing *HOXA10*, *HOXB2* and *MEIS1* were also identified as being highly expressed in normal karyotype AML compared to healthy normal donors. Further investigation into these *HOX* genes and co-factors may identify a candidate for monitoring MRD. It may also prove more informative if the expression of several *HOX* genes and co-factors were monitored simultaneously to gain a more reliable insight into the

directional progression of a patient's disease rather than using just one *HOX* genes as a marker. Various studies have demonstrated the differential pattern of the *HOX* genes during the stages of haematopoiesis and that the *HOX* genes are switched on and off in blocks (Magli *et al*, 1991; Sauvageau *et al*, 1994; Magli *et al*, 1997).

In the four normal karyotype patients in which *WT1* was monitored a consistent pattern in *WT1* expression could be seen with the remission/relapse status of the patients. The expression levels of *WT1* in the patient that relapsed never decreased below the upper level of expression detected in the healthy donor samples. After the initial decrease in *WT1* MRD by 1.5 logs from the transcript levels detected at diagnosis, an increase in MRD by 2.5 logs over three serial bone marrow samples was observed. This strongly suggests a molecular relapse supported by the study by Lane *et al* (2008) in which a 1 log rise in transcript level in serial bone marrow remission samples is enough to reliably predict a morphological relapse. The levels of *WT1* MRD in the three patients that achieved complete remission decreased from the level of transcript detected at diagnosis. It was unfortunate further remission samples were unavailable for these patients as it would be insightful to know if the level of *WT1* expression continued to decrease to below the upper level detected in the healthy normal donor samples and remained below this level. *WT1* expression has been evaluated for MRD monitoring in AML in many previous studies (Inoue *et al*, 1996; Cilloni *et al*, 2002; Barragan *et al*, 2004; Ostergaard *et al*, 2004; Weisser *et al*, 2005; Ommen *et al*, 2008). These studies all agree

that *WT1* is useful as a marker of MRD, particularly in those patients without markers but others have reported background levels of *WT1* during remission and following bone marrow or peripheral blood stem cell transplantation and therefore argue that *WT1* is only useful to those patients without an alternative marker (Gaiger *et al*, 1998;Elmaagacli *et al*, 2000;Hamalainen *et al*, 2008).

The mean $\pm$ 95% confidence intervals generated for the three *HOX* genes, *WT1* and *COL4A5* in the diagnostic normal karyotype AML patient samples displayed no overlap with the mean $\pm$ 95% confidence intervals generated for the genes in the healthy normal donor samples supporting their selection for further longitudinal MRD studies in section 6.3.1b. However, from the longitudinal studies it was clear that only *WT1* had the ability to monitor the disease and predict relapse in normal karyotype patients. This could have possibly been predicted at the diagnostic screening stage. The largest difference observed in the $\pm$ 95% confidence interval between patients and the healthy normal donors was for the *WT1* gene suggesting that this particular gene would go on to act as a useful MRD marker. However, whilst still no overlap was observed, a significantly smaller gap was observed for the mean $\pm$ 95% confidence intervals between patients and healthy normal donors for the *COL4A5* gene highlighting an insufficient difference in transcription between the patients and the healthy normal donors to be confident that the gene could be used as MRD markers.

A limitation of this investigation lies within the number of follow-up samples and the relatively small number of monitoring time points available for this study. Patient 5 relapsed 252 days following diagnosis but in this study we were only able to follow their disease up to 159 days following diagnosis, approximately 3 months prior to relapse. Therefore an increase in transcript level would not necessarily be expected within this time-frame, instead a molecular relapse might have been observed after 159 days. Innate noise within the system also provides a problem for MRD monitoring and the reliability of increases and decreases in transcript levels detected. As previously mentioned Lane *et al* (2008) suggested that a $\geq 1 \log_{10}$ rise in RQ-PCR transcript levels observed in serial remission bone marrow samples is a clinically useful definition of and reliably predicts a morphologic relapse (Lane *et al*, 2008). Lane *et al* (2008) were monitoring fusion gene products generated from t(8;21) and inv(16) in CBF AML patients. In this current study fluctuations of more than 1 log were observed in some of the CR patients for the *HOX* genes over two follow-up samples. These fluctuations may have been due to the methodology, for example the normalisation approach used, or due to biological causes, for example a natural innate variation in the genes of interest. Events like this need to be taken into account when analysing transcript levels for relapse prediction. An important caveat to this study is the sample size. Further investigations of a larger normal karyotype patient cohort and serial bone marrow samples following diagnosis is required to support the findings of this particular study. However, these samples are incredibly

difficult to source and realistically would only be made available through a prospectively designed clinical trial.

HOXA5, *HOXA9*, *HOXB6*, *COL4A5* and *WT1* were identified through gene expression profiling as being more highly expressed in normal karyotype AML patients compared to healthy normal donors. Their over-expression was validated in diagnostic normal karyotype patient samples by RQ-PCR and compared to the expression levels detected in healthy normal donor samples. A statistically significant difference in mean expression was observed between the two sample groups indicating the potential of all five genes to act as markers of MRD in normal karyotype AML. Further longitudinal monitoring highlighted the ability of the *WT1* gene to act as a marker of MRD but the three *HOX* genes and *COL4A5* failed to fill such a role. Supported by previous studies, *WT1* appears to be a strong candidate for a marker of MRD. Whilst there is some contradicting evidence in regards to the expression of *WT1* in healthy bone marrow samples and its level of sensitivity in comparison to markers currently used, in terms of normal karyotype AML where there are currently no markers, *WT1* may provide useful and reliable clinical information in regards to prognosis and predicting a molecular relapse. *WT1* has effectively worked as a positive control in this study demonstrating the ability of gene expression profiling using the Affymetrix platform to identify potential markers of MRD in normal karyotype AML. As *WT1* has been previously identified as an MRD marker in several investigations (Cilloni *et al*, 2002; Ogawa *et al*,

2003;Weisser *et al*, 2005), its identification here suggests that the present study was sufficiently sensitive to reiterate what has been shown previously. In this context, the inability to identify any novel MRD markers suggests that a more sensitive microarray platform may be required in future studies. Alternatively, we may have to concede that global gene expression profiling may be incapable of identifying better MRD markers in the context of NK AML.

Chapter 7

The treatment of AML cell lines with differentiation agents and gene-specific inhibitors

7.1 Introduction

In the previous chapters, gene expression profiling was used to identify genes over-expressed in specific AML subgroups. Comparing minimal residual disease (MRD) data for the identified genes in patient diagnostic and follow-up samples revealed a strong correlation between these genes and the markers currently in use for MRD evaluation. Several of the genes identified have previously been shown to play important roles in the development of leukaemia and so represent potentially very interesting alternative biomarkers of response to therapy.

7.1.1 Hepatocyte growth factor (*HGF*)

HGF was identified as being over-expressed specifically in AML patients with a t(15;17) and showed a very strong correlation with the markers currently used, *PML-RAR α* and *RAR α -PML* (Chapter 3). *HGF* was originally identified as a mitogen in primary cultured adult rat hepatocytes and plays an important role in the regeneration of hepatic cells (Hino *et al*, 1996). Evarts *et al* (1993) observed increased levels of both *HGF* and transforming growth factor α (*TGF- α*) following a partial hepatectomy suggesting a role of *HGF* in proliferation and differentiation of hepatic cells (Evarts *et al*, 1993). Several studies have identified *HGF* as a multi-functional cytokine. The induction of tubule formation by MDCK (Madin-Darby canine kidney) cells showed that *HGF* was produced by fibroblasts (Rubin *et al*, 1991) whilst *HGF*, or a

molecule biologically and immunologically related to HGF, could induce tubule formation (Montesano *et al*, 1991). The motogenic properties of HGF were identified in a study by Stoker and Perryman (1985). A motility factor was identified that was able to affect the motility of epithelial cells; its presence caused the expansion of epithelial colonies and cells to be scattered, resulting in its name of "Scatter factor" (Stoker & Perryman, 1985;Stoker *et al*, 1987). Subsequent analysis of the biological activity and comparative immunology, partial protein analysis and cDNA sequences identified scatter factor and HGF as the same molecule (Weidner *et al*, 1991). HGF has been demonstrated to regulate haematopoiesis in mouse foetal liver and adult bone marrow (Nishino *et al*, 1995) and various studies have observed an increased level of plasma HGF in AML and CML patients compared to healthy samples, yet low levels in ALL patients, suggesting a role in myeloid leukaemias (Hino *et al*, 1996;Nakamura *et al*, 1994). HGF was also found to stimulate the proliferation or colony formation and affect the migration of myeloid leukaemia cells in bone marrow or peripheral blood samples (Weimar *et al*, 1998). The receptor for HGF is a tyrosine kinase protein encoded by the proto-oncogene c-Met (Park *et al*, 1986). The binding of the HGF ligand to its receptor c-Met results in autophosphorylation of c-Met, which results in constitutive downstream signalling. Simultaneous expression of HGF and c-Met has not been seen to occur in normal cells but concomitant expression has been observed in myeloma cells (Borset *et al*, 1996a;Borset *et al*, 1996b)

7.1.2 Preferentially expressed antigen in melanoma (*PRAME*)

PRAME (preferentially expressed antigen in melanoma) was identified as being over-expressed specifically in t(8;21) AML patients and a correlation was observed with *WT1* and t(8;21), the markers currently used in the assessment of MRD (Chapter 4). *PRAME* was originally identified as a tumour antigen in malignant melanoma (Ikeda *et al*, 1997). However, several studies have shown over-expression of *PRAME* in AML samples (van Baren *et al*, 1998) and have suggested that *PRAME* could be a potential MRD monitoring candidate (Matsushita *et al*, 2001; Paydas *et al*, 2005).

7.1.3 Meningioma 1 (*MN1*)

MN1 (meningioma 1) was identified as being over-expressed specifically in inv(16) AML patients and its expression correlated with *WT1* and *CBFβ-MYH11* MRD (Chapter 4). *MN1* was first identified as a target of a t(4;22) in a patient with meningioma (Lekanne Deprez *et al*, 1991). *MN1* acts as a transcriptional co-activator by stimulating retinoic acid receptor/retinoid X receptor (RAR-RXR)-mediated transcription through the recruitment of co-activators p300 and RAC3 (retinoic acid receptor interacting protein 3) (Buijs *et al*, 2000; van Wely *et al*, 2003). *MN1* is also a target in the chromosome translocation t(12;22) resulting in the fusion gene product *MN1-TEL*. *MN1-TEL* is only associated with myeloid diseases and acts as an oncogene repressing transcription (Buijs *et al*, 2000; Kawagoe & Grosveld, 2005). As well as forming part of an oncogene in myeloid diseases, an overexpression of *MN1* has

been observed in inv(16) AML patients in several gene expression profiling studies and in QRT-PCR studies (Carella *et al*, 2007;Ichikawa *et al*, 2006).

7.1.4 HOXA1

HOXA1 was identified as being over-expressed specifically in normal karyotype AML patients and its expression was strongly correlated with *WT1* MRD data. *HOXA1* is a member of the homeodomain containing transcription factor family involved in cell differentiation and growth. Low levels of *HOXA1* expression is seen in human mammary glands during normal growth and differentiation but higher levels of expression has been observed in a variety of breast cancer lesions (Chariot & Castronovo, 1996) and enhances the proliferation and survival of the cells observed in oncogenesis (Zhang *et al*, 2003).

7.1.5 All-trans retinoic acid (ATRA)

Several studies identified the ability of retinoids to induce differentiation in HL-60 cells and in promyelocytes from acute promyelocytic leukaemia (APL) patients (Breitman *et al*, 1980;Breitman *et al*, 1981). All-trans retinoic acid (ATRA) is a natural isoform of retinoic acid successfully used in for the treatment of APL (Huang *et al*, 1988). ATRA induces a terminal differentiation followed by apoptosis of malignant cells (Huang *et al*, 1988). When APL cells are treated with pharmacological doses of ATRA (>1µM) co-repressors that bind to the

PML-RAR α fusion product and result in a negative control of transcription are released, allowing the expression of genes involved in myeloid cell differentiation (Barber *et al*, 2008).

7.1.6 NK4

NK4 is a small molecule containing the N-terminal domain and the 4 kringle domains of HGF and is an antagonist of HGF. The kringle domains are required for receptor binding and therefore necessary for the biological function of HGF (Chirgadze *et al*, 1998). NK4 binds to c-Met without activating the autophosphorylation that normally occurs when c-Met is bound to HGF and inhibits any downstream mitogenic, motogenic and morphogenic activities (Date *et al*, 1997). NK4 has also been observed to inhibit angiogenesis in human lung-derived endothelial cells (Kuba *et al*, 2000). Due to the important role of HGF in cancer progression, NK4 has been suggested as a potential therapeutic and several studies have observed a suppression of tumour angiogenesis, growth and metastasis in colon cancer cells, cervical carcinoma cells and multiple myeloma cells (Date *et al*, 1998;Kuba *et al*, 2000;Wen *et al*, 2007;Du *et al*, 2007).

7.1.7 siRNA

RNA interference (RNAi) is a sequence-specific, post-transcriptional process for gene silencing. Double stranded RNA molecules are recognised by the cell as foreign and cleaved into short fragments

called small interfering RNA (siRNA) by a ribonuclease III-type protein, dicer RNase III (Elbashir *et al*, 2001). The resulting siRNA molecules are 21-23 nucleotides long and contain 3'-overhanging ends 2-3 nucleotides long (Zamore *et al*, 2000). A nuclease-containing complex called the RNA-induced silencing complex (RISC) incorporates the siRNAs, which are unwound by an RNA helicase. The RISC complex is activated and guided to the complementary mRNA strand by the siRNA (Elbashir *et al*, 2001). The siRNA hybridises to the mRNA causing the RISC complex to catalyse cleavage of the mRNA within the target site producing unprotected ends and resulting in degradation of the mRNA (Zamore *et al*, 2000). The RISC complex is recovered allowing further binding and cleavage to occur, resulting in a decreased expression of the target gene.

7.1.8 Cell surface markers

Currently it is standard practice to carry out immunophenotyping of blasts from bone marrow and peripheral blood samples of AML patients at diagnosis. A variety of differentiation markers characterise the leukaemic blasts and are associated with their commitment to the myeloid lineage and their level of maturation. The identification of markers that are not normally associated with the myeloid lineage can be used for diagnostic purposes and for the detection of minimal residual disease (Mason *et al*, 2006). CD11b is such a marker associated with the myeloid lineage. It is expressed as a heterodimer with CD18 and has been observed on mature monocytes and

macrophages. A study by Drayson *et al* (2001) identified monocytes and neutrophils as expressing CD11b whereas HL-60 did not express CD11b until treatment with the differentiating agent ATRA (Drayson *et al*, 2001). CD38 is a transmembrane glycoprotein and is expressed during the early stages of myeloid differentiation during normal haematopoiesis cells (Drach *et al*, 1994). Expression of CD38 is highly regulated by retinoids and in HL-60 cells ATRA was identified as a potent inducer of CD38 expression (Drach *et al*, 1993; Drach *et al*, 1994).

Many genes have previously been identified as being specific to various sub-groups of AML but their relevance to and how they function in the development leukaemogenesis is not fully known. Selecting a gene identified as specific from each sub-group, the aim of this chapter was to treat AML cell lines with differentiating agents and gene-specific inhibitors and observe their effects on the cells and the expression of the identified genes.

7.2 Materials and Methods

7.2.1 The effect of ATRA upon HGF expression in AML Cells

To determine the effect of ATRA on the expression of HGF in AML, three cell lines; NB4 (a t(15:17) AML cell line), HL-60 (an APL cell line) and Me-1 (an inv(16) AML cell line), were treated with ATRA at a concentration of 10^{-7} M over 4 days. The cells were treated at day 0 only

or treated every 24 hours from day 0. Cells were harvested at the start of treatment and at 1, 2, 3 and 4 days after treatment. The cells were snap-frozen and subsequently RNA was extracted as previously described (Chapter 2). The expression levels of the four genes identified as being specific to either one of the three favourable translocations or to a normal karyotype were monitored in three AML cell lines to observe the effects of ATRA upon their expression. Real-time PCR was performed using 1µg of RNA reverse transcribed into cDNA (Chapter 2).

7.2.2 The effect of NK4 upon HGF expression in AML cell lines

To observe the effects of the HGF-antagonist NK4 in leukaemic cells, NB4 cells were treated with NK4 (supplied by Dr Gaynor Davies and Dr Christian Parr, Department of Surgery, Cardiff University) over a period of 72 hours. To determine the concentration of NK4 required for the treatment of NB4 cells, the level of HGF protein was quantified using the Quantikine® HGF ELISA kit (Section 2.5.2). A standard curve was generated using HGF standards and the HGF concentration in NB4 cells was calculated. The amount of NK4 required for inhibitory effects was ten times the amount of HGF present. NB4 cells were set at a density of 1×10^5 cells/ml and initially treated with 0.8µg/ml of NK4 and left to grow for 72 hours. Cell counts and viability counts were performed at 0, 3, 6, 9, 15, 24, 48 and 72 hours. The experiment was performed in triplicate. The cells were pelleted and resuspended in 1x phosphate buffered saline (PBS) for immunophenotyping to observe the

effects of NK4 upon the expression of surface markers in these cells (Chapter 2).

7.2.3 The effect of RNA interference upon the expression of genes identified as being specific to AML sub-groups in AML cell lines.

To observe the effect of RNA interference on the expression levels of selected genes in NB4 cells, cells were transfected with siRNA designed to silence with the expression of *HGF*, *HOXA1*, *MN1* and *PRAME*. In parallel experiments a negative control scrambled siRNA was used. The protocol for the transfection of siRNA and the efficiency of transfection is described in chapter 2.

7.3 Results

7.3.1 The effect of ATRA on the growth of AML cells and the expression of selected genes

The AML cell lines NB4, HL-60 and Me-1 were treated with 10^{-7} ATRA over a period of 4 days to monitor its effects upon cell growth, viability and expression of the genes *HGF*, *PRAME*, *MN1* and *HOXA1*. These genes were selected for investigation to compare a potential ATRA responsive gene (*HGF*) to genes identified from ATRA non-responsive leukaemias; t(8;21), inv(16) and normal karyotype.

a. Cell growth and viability

The results shown in Figure 1 indicate an inhibition of growth by day 2 of the treatment when compared to untreated cells in all three cell lines cells treated with ATRA. The greatest effect was seen in the NB4 cells and the HL-60 cells. There was no significant difference in the rate of growth between those cells treated with ATRA on day 0 or those treated with ATRA on day 0, 1, 2 and 3. Further, despite the inhibition of growth, there was no significant change in cell viability between the ATRA-treated and untreated cells.

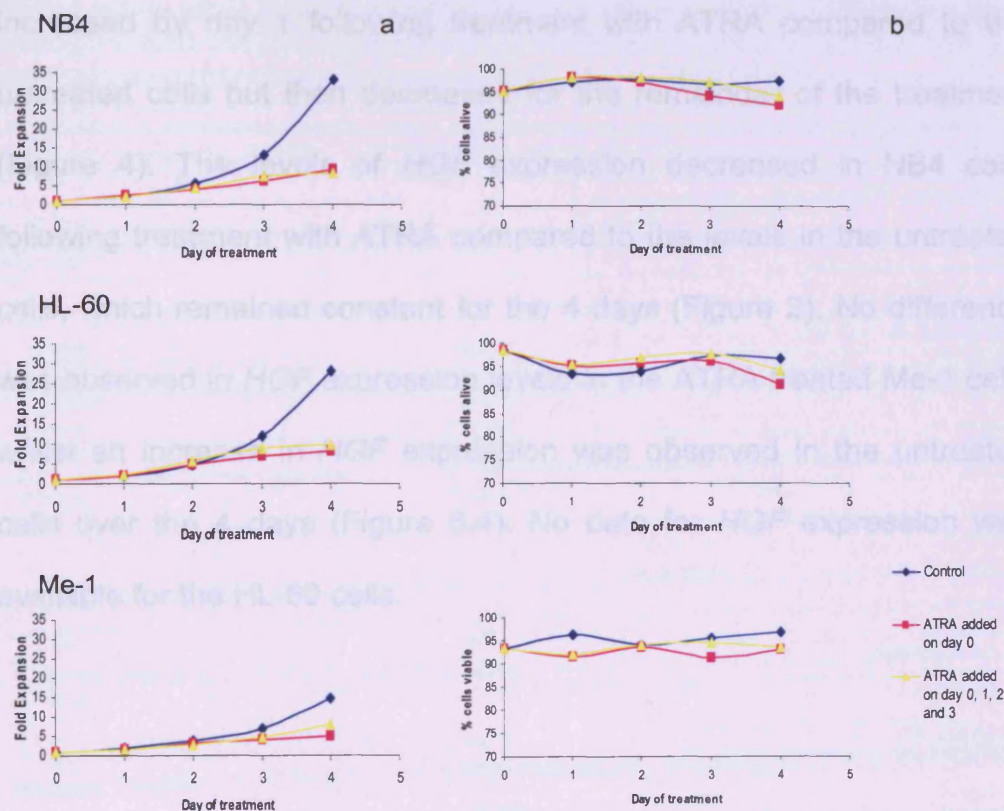


Figure 1. The effect of ATRA on the growth (a) and viability (b) of NB4 cells over 72 hours when treated with 10^{-7} M ATRA at day 0 only and when treated at day 0, 1, 2 and 3 compared to untreated cells.

b. Gene expression

The level of expression of the *HOXA1* gene appeared to increase in all three cell lines following treatment with ATRA over 4 days compared to the levels observed in the untreated cells, which remained constant over the 4 days (Figures 2, 3 and 4). No significance difference in the level of expression of *PRAME* was observed between the ATRA treated cells and the untreated cells in any of the 3 cell lines (Figures 2, 3 and 4). There appeared to be no difference in level of expression of *MN1* between the treated and untreated NB4 cells (Figures 2 and 3). In contrast, the level of *MN1* expression in the HL-60 and Me-1 cell lines increased by day 1 following treatment with ATRA compared to the untreated cells but then decreased for the remainder of the treatment (Figure 4). The levels of *HGF* expression decreased in NB4 cells following treatment with ATRA compared to the levels in the untreated cells, which remained constant for the 4 days (Figure 2). No difference was observed in *HGF* expression levels in the ATRA treated Me-1 cells whilst an increase in *HGF* expression was observed in the untreated cells over the 4 days (Figure 6.4). No data for *HGF* expression was available for the HL-60 cells.

NB4

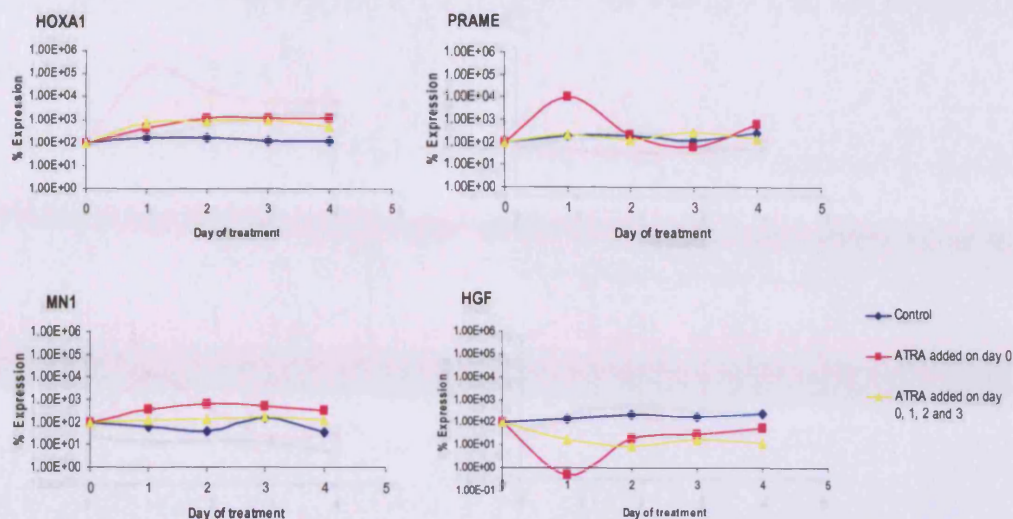


Figure 2. The effect of 10^{-7} M ATRA on the expression of *HOXA1*, *PRAME*, *MN1* and *HGF* in NB4, the t(15;17) cell line, cells over 72 hours when treated only at day 0 and when treated at day 0, 1, 2 and 3 compared to untreated cells.

HL-60

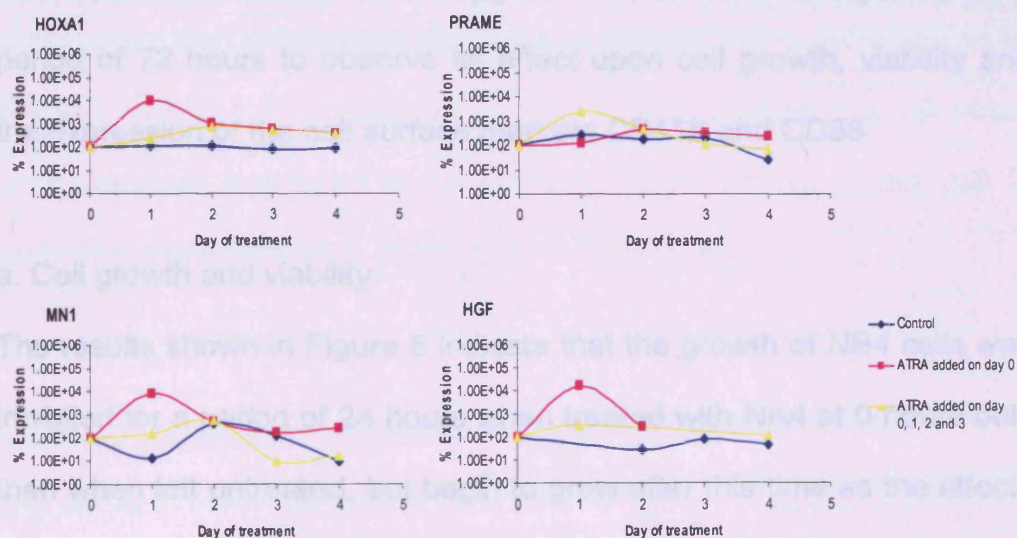


Figure 3. The effect of ATRA on the expression of *HOXA1*, *PRAME*, *MN1* and *HGF* in HL-60, the promyelocytic cell line, cells over 72 hours when treated only at day 0 and when treated at day 0, 1, 2 and 3 compared to untreated cells.

Me-1

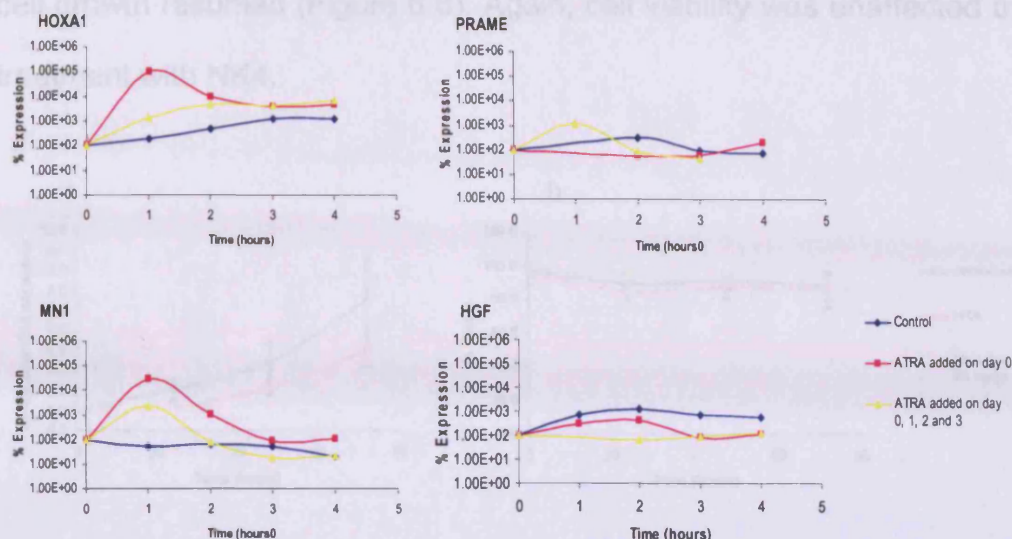


Figure 4. The effect of ATRA on the expression of *HOXA1*, *PRAME*, *MN1* and *HGF* in Me-1, the inv(16) cell line, cells over 72 hours when treated only at day 0 and when treated at day 0, 1, 2 and 3 compared to untreated cells.

7.3.2 The effect of the HGF-antagonist NK4 on NB4 cells

NB4 cells were treated with 0.8µg/ml NK4, an HGF antagonist, for a period of 72 hours to observe its effect upon cell growth, viability and the expression of the cell surface markers CD11b and CD38

a. Cell growth and viability

The results shown in Figure 5 indicate that the growth of NB4 cells was inhibited for a period of 24 hours when treated with NK4 at 0 hours only than when left untreated, but begin to grow after this time as the effects of NK4 ceased. No significant change in cell viability between treated and untreated cells was observed. When the NB4 cells were treated with NK4 at 0 hours and at 24 hours after the initial treatment, the

growth of cells was inhibited until 48 hours after initial treatment, when cell growth resumed (Figure 6.6). Again, cell viability was unaffected by treatment with NK4.

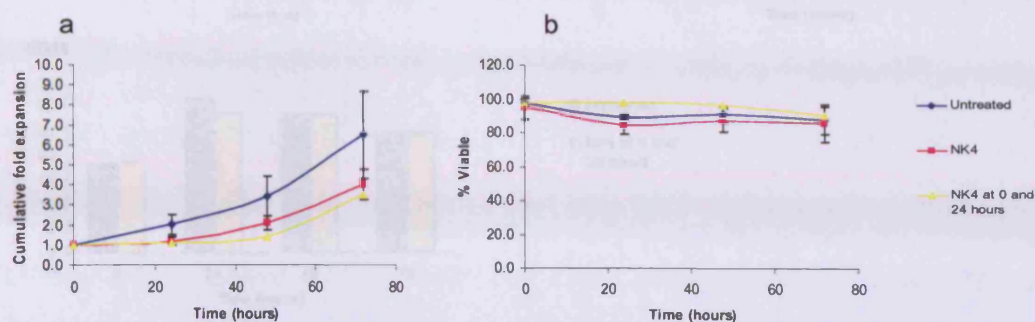


Figure 5. The effect of NK4 on growth (a) and viability (b) of NB4 cells over 72 hours when treated with 0.8 μ g/ml of NK4 at 0 hours and at 24 hours after initial treatment. Experiments were performed in triplicate and the error bars represent +/- one standard deviation from the mean.

b. Cell surface markers

As seen in Figure 6, no consistent change in expression of CD38 was observed in the NK4-treated NB4 cells compared to the untreated cells. A slight increase in expression of CD11b was observed at 48 hours after treatment compared to the untreated cells but due to the lack of replicates the significance of this data is inconclusive.

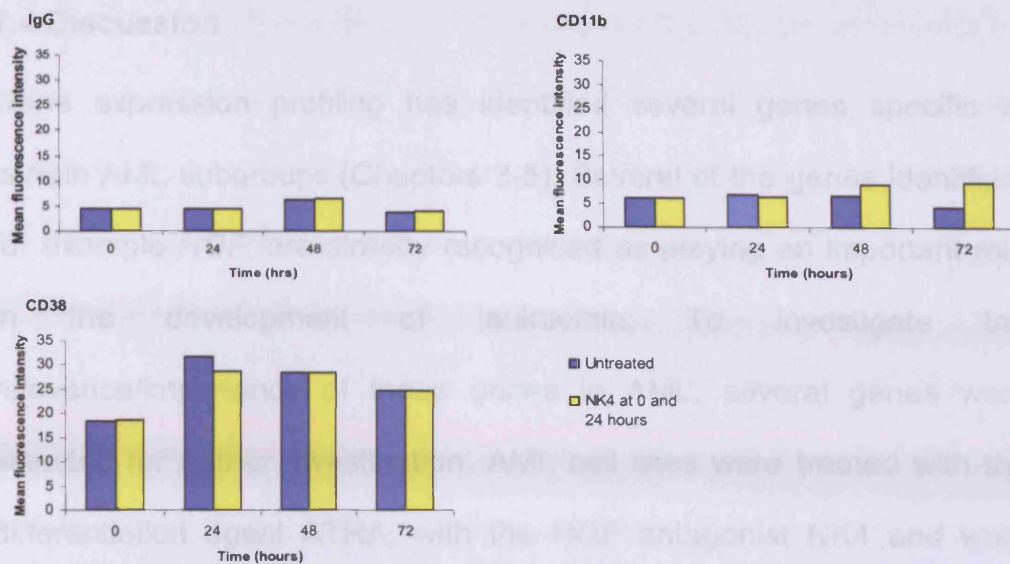


Figure 6. The effect of 0.8 μ g/ml NK4 on the expression of cell surface markers of NB4 cells over 72 hours when treated at 0 hours and 24 hours after initial treatment. Experiments were performed once due to a lack of the NK4 inhibitor.

7.3.3 Measurement of transfection efficiency of FAM-labelled siRNA against selected genes in NB4 cells

NB4 cells were transfected with 50nM and 100nM siRNA designed to silence *HGF*, *MN1*, *PRAME* and *HOXA1* and with a scrambled siRNA to act as a negative control. The efficiency of siRNA transfection into the NB4 cells was determined by flow cytometry on day 2 following transfection so that any changes in cell growth or gene expression in the cells could be correctly attributed to the actions of the siRNA. There was no significant difference in the level of fluorescence between the transfected and non-transfected cells suggesting inadequate transfection of the siRNA. It was decided due to time restraints not carry out any further investigation into the effect of siRNA on gene expression.

7.4 Discussion

Gene expression profiling has identified several genes specific to certain AML subgroups (Chapters 3-5). Several of the genes identified, for example *HGF*, are already recognised as playing an important role in the development of leukaemia. To investigate the relevance/importance of these genes in AML, several genes were selected for further investigation. AML cell lines were treated with the differentiation agent ATRA, with the HGF antagonist NK4 and were transfected with siRNA designed to silence the expression of *HGF*, *MN1*, *PRAME* and *HOXA1* in order to examine the effects upon cell growth, viability, gene expression and cell-surface marker expression.

When treated with 10^{-7} M ATRA, a decrease in growth was observed in the AML cell lines NB4, HL60 and Me-1 after approximately 2 days following treatment. No significant change in cell viability was observed in any of the cell lines. This data is supported by a study by Ozeki *et al* (2008) in which a growth arrest in HL-60 cells was observed 1-2 days following treatment with ATRA and <15% early apoptotic cells were observed throughout the 5-day treatment (Ozeki & Shively, 2008). The growth inhibition observed was much greater in the promyelocytic cell lines compared to the Me-1 cells, the cell line harbouring an inv(16) translocation. Previously a gene identified as specific to inv(16) patients was *MN1* (Chapter 4). The MN1 protein has been identified as functioning as a transcription co-activator for the retinoic acid receptor (van Wely *et al*, 2007). Heuser *et al* (2007) later observed that MN1

greatly increases the resistance of cells to ATRA-induced differentiation and cell cycle arrest (Heuser *et al*, 2007). Therefore the lower growth inhibition observed in the Me-1 cells could be due to the presence of an inv(16) and the subsequent over-expression of *MN1*. Treatment with ATRA saw the level of *HGF* expression to decrease in NB4 cells, increase in HL-60 cells and remained unchanged in Me-1 cells over the 4 days. The level of *HOXA1* increased in all 3 cell lines treated with ATRA over 4 days whereas the level of *PRAME* expression remained unchanged in the cell lines. The expression levels of *MN1* remained unchanged in the ATRA-treated NB4 cells, an increase in *MN1* expression was observed in the HL-60 cells and the Me-1 cells at 24 hours following treatment, but then decreased for the remainder of the treatment. NB4 is a promyelocytic cell line and HL-60 cells, derived from an AML FAB type M2 patient, display a myeloblastic/promyelocytic morphology but the t(15;17) translocation is only expressed in the NB4 cells (Birnie, 1988). *HGF* was previously identified as being specific to t(15;17) AML patients (Chapter 3). Whilst the effect of ATRA on the expression of *HGF* has not been investigated in haematological malignancies, a study using a human malignant astrocytoma cell line, U87, did observe a decrease in the expression upon treatment with ATRA in a time- and concentration-dependent manner (Chattopadhyay *et al*, 2001). Despite NB4 cells being promyelocytic and HL-60 displaying a promyelocytic morphology, the difference in effect on *HGF* expression between the HL-60 cells and the NB4 cells could be due to the absence of the t(15;17) translocation in HL-60 cells but this can only

be speculative due to the lack of confirmatory experimental data. The lack of change in *HGF* expression in the Me-1 cells could again be due to an over-expression of *MN1* resulting resistance to ATRA, or *HGF* not having as a significant role in inv(16) AML than in t(15;17) AML. *PRAME* was identified as being specific to t(8;21) patients and no change in its expression was observed in the 3 cell lines suggesting that *PRAME* is not regulated by retinoids in these cell lines. Unfortunately these results are not conclusive due to the lack of replicates. It would also be interesting to include a t(8;21) AML cell line, such as Kasumi, in the study to observe the effect of ATRA upon all the genes selected in their specific sub-group.

When treated with NK4, the NB4 cells exhibited an inhibition of growth for 24 hours following treatment. No significant change in cell viability was observed. No change in expression of CD38 was observed upon treatment with NK4 but a slight increase in CD11b was observed at 48 hours following treatment. The significance of the CD11b cannot be determined due to the lack of replicates. The expression of both CD11b and CD38 has been observed to increase in HL-60 cells upon treatment with the differentiating agent ATRA (Drach *et al*, 1993; Drayson *et al*, 2001). The lack of significant increase of both surface antigens in this study suggests that NK4 does inhibit the growth of cells but does not cause them to differentiate. *HGF* has been observed as a multifunctional protein that promotes tumour cell growth, which is inhibited by NK4 (Du *et al*, 2007). The cell surface marker assays need

to be repeated in order to determine their statistical significance. It would have also been interesting to follow the effect of NK4 in a wider panel of AML cell lines and look at the expression levels of HGF, MN1 and HOXA1 in the treated cells but unfortunately a problem with the supply of the NK4 meant this and further immunoassays were not possible.

Inhibition of the selected genes by siRNA was attempted but transfection of the AML cell lines was unsuccessful. Further investigation into the silencing of genes identified as specific to AML sub-groups could provide useful information, potentially leading to the validation of new therapeutic targets.

ATRA is currently used to induce differentiation in acute promyelocytic leukaemia patients and the ability of NK4 to down-regulate HGF and inhibit tumour growth has already been recognised in several cancers. However the observation of growth inhibition of AML cell lines would suggest that NK4 has the potential as a therapeutic treatment for AML and should be investigated further. This investigation into possible novel markers of MRD has also highlighted that the AML sub-group-specific genes may have the potential to identify new therapeutic targets.

Chapter 8

Discussion

Despite improved survival rates over the past decade, relapse still remains a major obstacle to the cure of leukaemia (Chung *et al*, 2006). One strategy to combat relapse is the early detection and treatment of minimal residual disease (MRD) in order to reduce the risk of a potential relapse. Therefore, MRD markers can be used to tailor treatment by identifying patients who need further therapy (MRD positive) and those who do not require further intervention (MRD negative). However, these MRD markers need to be tested prospectively and must meet exacting sensitivity and specificity criteria. Real-time PCR (RQ-PCR) is already routinely used to quantitate fusion gene products resulting from common translocations occurring in AML; *PML-RAR α* , *AML1-ETO* and *CBF β -MYH11*. Unfortunately patients presenting with these chromosomal abnormalities only account for approximately 20% of the adult AML cases and approximately 50% of AML cases have no detectable aberrations. The use of microarrays as a tool for disease classification was first explored by Golub *et al* (1999) using class discovery and class prediction to distinguish between ALL and AML samples (Golub *et al*, 1999). The predictive power of gene expression profiling was further utilised to identify distinct gene expression profiles for AML-specific genetic alterations and later with specific gene mutations and aberrant gene expressions within AML. Using these profiles, markers have been identified that can help diagnose individual patients and predict prognosis. Reports such as these have prompted investigations into the suitability of gene expression profiling to identify

new therapeutic targets and predict therapeutic outcome and patient responses.

With the need for novel MRD markers in AML clearly evident and the potential of gene expression profiling to identify such markers, the aim of this study was to use gene expression profiling to identify new markers of MRD specific to AML sub-groups.

As a “proof-of-principle”, gene expression profiling was initially applied to a cohort of AML patients to identify potential markers of MRD specific to AML bearing the t(15;17) translocation (Chapter 3). Statistical analysis and filtering tools identified a list of probe sets specific to this AML sub-group. Several probe sets identified as significantly over-expressed in t(15;17) patients in comparison to non-t(15;17) patients, including the two probe sets with the greatest fold change in expression, represented the hepatocyte growth factor (*HGF*) gene. This is consistent with previous gene expression profiling studies of APL. Valk *et al* (2004) identified the *HGF* gene as specific to a cluster of patients, all of whom were diagnosed with APL and as the best predictor for the t(15;17) aberration in AML whilst Gutierrez *et al* (2005) identified *HGF* as significantly over-expressed in APL patients and as the second most over-expressed gene in APL patients when compared to non-APL patients. The observation of a statistically significant difference in expression of *HGF* and *IGFBP2* between t(15;17) AML diagnostic patient samples and healthy normal bone marrow samples led to longitudinal monitoring of these genes in three t(15;17) patients

that achieved complete remission. The candidate genes displayed similar expression patterns to the genes currently used as MRD markers, *PML-RAR α* and *RAR α -PML*, and remained lower than the upper levels detected in healthy normal donor bone marrow samples. However, only a two log decrease in *HGF* and *IGFBP2* transcripts levels from diagnosis and induction therapy to complete remission was observed compared to the six log decrease observed for the *PML-RAR α* and *RAR α -PML* transcripts. This relatively small change in transcript levels probably limits the usefulness of these markers in the MRD setting but they may still be biologically important in the context of remission and disease relapse.

Using the same principle for the core binding factor leukaemias, gene expression profiling was applied to the cohort of AML patients to identify potential markers of MRD specific to AML bearing the t(8;21) translocation or inversion of chromosome 16 (Chapter 4). Many genes identified in this study as over-expressed in either t(8;21) or inv(16) AML have previously been identified in other investigations. Confirming the results of the current study, Valk *et al* (2004) identified the *ETO* gene as the most discriminating genes for a cluster containing all patients exhibiting a t(8;21), whilst Gutierrez *et al* (2005) observed an elevated *MYH11* expression in inv(16) AML patients. In further agreement with this study, Bullinger *et al* (2007) identified the *POU4F1* (POU domain, class 4, transcription factor 1), *CAV1* (caveolin 1), *HSPG2* (heparan sulphate proteoglycan 2) and *TRH* (Thyrotropin-releasing hormone)

genes as defining the t(8;21) patients and the *NT5E* (5'-nucleotidase, ecto (CD73)) and *SPARC* (secreted protein, acidic, cysteine-rich) genes defining the inv(16) patients. Diagnostic monitoring of several genes, selected on the basis of their level of expression in the gene expression profiling analysis and current literature, from each CBF groups in CBF patients and in healthy normal bone marrow demonstrated promising results but longitudinal studies demonstrated a similar problem seen for the t(15;17) monitoring – much smaller changes in transcript levels for the candidate compared to the markers currently used and therefore, the relevance of their change in expression can be questioned. However, in agreement with earlier studies of Matsushita *et al* (2001) and Tajeddine *et al* (2006), promising results were seen for the *PRAME* gene, identified as over-expressed in t(8;21) AML, and the monitoring of residual disease in several t(8;21) patients.

The potential of gene expression profiling to identify candidate MRD makers demonstrated in the favourable AML sub-groups led the investigation to identify potential markers for normal karyotype AML (Chapter 5). From a core dataset of normal karyotype AML patients and healthy normal donors, 25 probe sets were identified as up-regulated in the normal karyotype AML patients and correctly clustered samples as normal karyotype AML or healthy normal bone marrow in an independent dataset. Similar to the investigation by Debernardi *et al* (2003), a number of *HOX* gene and *HOX*-related genes were up-regulated in the normal karyotype AML samples. Selected on the basis

of their level in expression in the normal karyotype AML samples from the gene profiling analysis, several *HOX* genes, *COL4A5* and *WT1*, the gene currently used as an MRD marker, were monitored in diagnostic normal karyotype AML samples and healthy bone marrow (Chapter 6). Whilst all displaying statistically significant differences in mean expression levels between the normal karyotype AML samples and the healthy normal bone marrow samples, the lack of spread of expression levels of *WT1* between the two groups indicated *WT1* had the greatest potential from the group of genes to discriminate between residual disease indicating a relapse and background levels similar to those detected in healthy normal samples. This was confirmed from longitudinal monitoring of the genes in normal karyotype AML patients. A decrease in *WT1* expression levels following induction therapy was observed in all patients. In those patients achieving complete remission, *WT1* expression levels remained at this lower level whereas an increase in *WT1* expression was observed approximately three months prior to relapse in the patient that relapsed at 125 days following diagnosis. The three *HOX* genes and *COL4A5* displayed patterns of expression that would not be expected of a marker displaying either remission or relapse and of unknown relevance to their disease. *WT1* has a long history of investigation as a potential marker of MRD in AML (Inoue *et al*, 1996; Cilloni *et al*, 2002; Trka *et al*, 2002; Ogawa *et al*, 2003; Weisser *et al*, 2005; Candoni *et al*, 2009). The identification of *WT1* in the current study, and its validation as a marker, confirms the ability of gene expression profiling to identify new markers of MRD for

AML. The lack of other genes confirmed as suitable MRD markers may be due to a lack of sensitivity of the Affymetrix platform to identify further markers or that gene expression profiling is simply not a method suitable for identifying markers better than those already in use.

One possible limitation to this study may be the time-frame of patient follow-up samples that were available for analysis. The follow-up samples used for monitoring all the patients in the current study were not always available to the point of relapse. As a consequence, potentially important changes of gene expression relating to relapse were not observed as samples close enough to the point of relapse were not always available. Fluctuations in gene expression, possibly due to methodology and/or biological reasons, were regularly detected without a change in the patient's disease status. Following the rule of Lane *et al* (2008); the requirement for a greater than one log rise in RQ-PCR transcript levels detected in serial bone marrow samples, most of the fluctuations could be ignored as innate noise.

The use of hybridisation probes to detect the transcript levels of the candidate genes would improve the sensitivity of the detection during longitudinal monitoring and assess the potential relevance of any changes in expression detected. In the current study a SYBR green method with primers was utilised for the detection of transcripts by RQ-PCR. A disadvantage of this method is the reduced specificity of the primers compared to hybridisation probes resulting in the non-specific amplification of PCR products and primer dimers. To account for this

problem in the current study a melting curve analysis was performed to identify any non-specific amplification. However, this does not eliminate non-specific amplification and reduces the specificity and sensitivity of the amplification.

The small size of the sample populations monitored within this study precludes the drawing of definitive conclusions and further investigations of a larger normal karyotype patient cohort and serial bone marrow samples following diagnosis is required to corroborate the findings. However, it should be noted that these samples are very difficult to source and realistically would only be made available through a prospectively designed clinical trial.

This study identified a series of sub-group specific genes which were compared to existing MRD markers. A few of these genes correlated with the expression of existing markers and further validation in larger MRD studies could prove their usefulness in the clinical environment. Unfortunately, the majority of the candidate markers identified failed to meet all of the necessary suitability criteria for use as MRD markers. However, this does not invalidate gene expression profiling as a method of identifying potential MRD markers, it merely highlights the complexity of the problem and points to the need for a large-scale studies to minimise false discovery rates and positively identify important candidates. The identification and validation of a gene currently used as an MRD marker, *WT1*, demonstrates the ability of gene expression

profiling to identify potentially significant genes for AML sub-groups that may ultimately be used as markers of MRD.

Possible future work;

- Re-evaluate the current genes identified as specific to AML sub-groups to assess other potential MRD marker candidates previously over-looked.
- Apply gene expression profiling to identify genes specific to normal karyotype AML patients bearing significant prognostic markers, for example NPM1, FLT3, CEBPA, MN1
- Validate the potential MRD markers in a larger sample group and observe their behaviour in patients from other AML sub-groups.
- Use the standardised MRD protocols for monitoring the candidate markers.

Appendix

Appendix I: Supporting Ethical Permissions

AML 14

Multi-Centre Research
Ethics Committee for
Wales

Chairman/Cadeirydd:
Dr John Saunders

**MREC
for
WALES**

Pwyllgor
Ymchwil Eithgeau
Aml-Ganolfan
yng Nghymru

Administrator/Gweinyddes:
Corinne Scott

Temple of Peace and Health, Cathays Park, Cardiff CF10 3NW
Teml Heddwch ac Iechyd, Parc Cathays, Caerdydd CF10 3NW

WYTH 01800

Telephone enquiries to: 029 2049 2444

Fax No. 029 2049 2444

MREC website: <http://mrc.mrecc.org.uk>
e-mail: corinne.scott@tre-taf-lly-wales.nhs.uk

Professor Alan Burnett,
Department of Haematology,
University Hospital of Wales,
Heath Park,
Cardiff CF14 4XN

February 3rd 2004

Dear Professor Burnett

MREC 98/98
LFR AML 14 Trial

PLEASE QUOTE THIS IN ALL CORRESPONDENCE

The Executive Sub Committee of the MREC for Wales further reviewed your application at its meeting held on January 30th 2004

The documents reviewed were as follows:

- Protocol Amendment Form - Amendment Six - dated January 7th 2004
- Patient Information Sheet 1 (amendment 5) dated January 7th 2004
- Patient Consent Form 1, version 4 dated January 7th 2004
- Patient Information Sheet 2 (amendment 5) dated January 7th 2004
- Patient Consent Form 2, version 4 dated January 7th 2004
- Patient Information Sheet 3 (amendment 3) dated January 7th 2004
- Patient Consent Form 3, version 4 dated January 7th 2004
- GP letter 1, version 3 dated January 7th 2004
- GP letter 2, version 3 dated January 7th 2004
- GP letter 3, version 3 dated January 7th 2004
- Study protocol, version 3 dated January 7th 2004

The Committee members present agreed that they had no ethical objections to the proposed amendment and were pleased to approve it. A copy of the completed MREC Amendment Report Form is attached for your records

AML 15

Multi-Centre Research
Ethics Committee for
Wales

Chairman/Cadeirydd:
Dr John Saunders

**MREC
for
WALES**

Pwyllgor
Ymchwil Eithgeau
Aml-Ganolfan
yng Nghymru

Administrator/Gweinyddes:
Corinne Scott

Temple of Peace and Health, Cathays Park, Cardiff CF10 3NW
Teml Heddwch ac Iechyd, Parc Cathays, Caerdydd CF10 3NW

WYTH 01800

Telephone enquiries to: 029 2049 2444

Fax No. 029 2049 2444

MREC website: <http://mrc.mrecc.org.uk>
e-mail: corinne.scott@tre-taf-lly-wales.nhs.uk

Professor Alan Burnett
Department of Haematology
University of Wales College of Medicine
Heath Park,
Cardiff CF14 4XN

June 20th 2002

Dear Professor Burnett,

Research Protocol MREC 01/0/53 (Please quote this in all correspondence)
Medford Research Council Acute Myeloid Leukaemia 15 Trial

Further to my telephone conversation earlier today with Alison Jenkins, I am writing to confirm the documents approved by the MREC for Wales for use with this study.

The documents are as follows:

- Full application form including Annex A and Annex C
- Study Protocol: MRC AML 15 Trial, dated December 2001.
- ✓ Patient Information Sheet 1, version 1 dated November 16th 2001.
- ✓ Written Consent Form 1, version 1 dated November 16th 2001.
- ✓ Patient Information Sheet 2, version 1 dated November 16th 2001.
- ✓ Written Consent Form 2, version 1 dated November 16th 2001.
- ✓ Patient Information Sheet 3, version 1 dated November 16th 2001.
- ✓ Written Consent Form 3, version 1 dated November 16th 2001.
- ✓ Patient Information Sheet 4, version 1 dated November 16th 2001.
- ✓ Written Consent Form 4, version 1 dated November 16th 2001.
- ✓ Patient Information Sheet 5, version 1 dated November 16th 2001.
- ✓ Written Consent Form 5, version 1 dated November 16th 2001.
- ✓ Patient Information Sheet 6, version 1 dated November 16th 2001.
- ✓ Consent Form 6, version 1 dated November 16th 2001.
- ✓ Patient Information Sheet "Taking Part In Research", version dated March 22nd 1999.
- ✓ Written consent form, undated.
- ✓ Invitation Letter and Consent Form, dated March 22nd 1999
- GP Information Sheet, version 1 dated November 16th 2001.
- GP Information Sheet 2, version 1 dated November 16th 2001.
- GP Information Sheet 3, version 1 dated November 16th 2001.
- GP Information Sheet 4, version 1 dated November 16th 2001.
- Quality of Life Study.
- Updated list of Local Researchers dated January 16th 2002.

AML16

Multi-Centre Research
Ethics Committee for
Wales

Acting Chairman:
Cadeirydd Dros Dro:
Dr Gordon Taylor

MREC
for
WALES

Pwyllgor
Ymchwil Eithgeau
Aml-Ganolfan
yng Nghymru

Administrator/Gweinyddes:
Dr. Corinne Scott

Churchill House, Fourth Floor, 17 Churchill Way, Cardiff, CF10 2TW
17 Churchill, 17 Ffordd Churchill, Caerdydd, CF10 2TW

WELSH 01809

Telephone enquiries to: 02920 576829

Fax No: 02920 576824

MREC website: <http://www.mrec.org.uk>
e-mail: mrec@hsc.wales.nhs.uk

Please note new address and contact details for MREC for Wales.

22 November 2006

Professor Alan K Burnett
Professor of Haematology
Cardiff University
Department of Haematology
School of Medicine
Heath Park
Cardiff CF3 6XX

Dear Professor Burnett

Full title of study:	AML 16: A Programme of Treatment Development for Older Patients with Acute Myeloid Leukaemia and High Risk Myelodysplastic Syndrome
REC reference number:	06/MREC08/84
Protocol number:	2
EudraCT number:	2005-002846-14

The REC gave a favourable ethical opinion to this study on 06 December 2005.

Further notification(s) have been received from local site assessor(s) following site-specific assessment. On behalf of the Committee, I am pleased to confirm the extension of the favourable opinion to the new site(s). I attach an updated version of the site approval form, listing all sites with a favourable ethical opinion to conduct the research.

Research governance approval

The Chief Investigator or sponsor should inform the local Principal Investigator at each site of the favourable opinion by sending a copy of this letter and the attached form. The research should not commence at any NHS site until research governance approval from the relevant NHS care organisation has been confirmed.

Statement of compliance

This Committee is recognised by the United Kingdom Ethics Committee Authority under the Medicines for Human Use (Clinical Trials) Regulations 2004, and is authorised to carry out the ethical review of clinical trials of investigational medicinal products.

Appendix II: Gene Lists

t(15;17)-Specific Genes

Genes identified as over-expressed in t(15;17) AML through filtering on fold-change, ANOVA and SAM analysis of probe sets classified as “present in at least 1 sample” and as “present and changing” (Chapter 3).

	Probe Set	Gene Symbol	Gene Title	Fold Change
1	210794_s_at	<i>MEG3</i>	Maternally expressed 3 (non-protein coding)	3.002
2	208581_x_at	<i>MT1X</i>	Metallothionein 1X	3.149
3	209686_at	<i>S100B</i>	S100 calcium binding protein B	3.184
4	203074_at	<i>ANXA8</i>	Annexin A8	3.387
5	221898_at	<i>PDPN</i>	Podoplanin	3.689
6	205047_s_at	<i>ASNS</i>	Asparagine synthetase	3.706
7	*204879_at	<i>PDPN</i>	Podoplanin	3.739
8	219837_s_at	<i>CYTL1</i>	Cytokine-like 1	3.838
9	207076_s_at	<i>ASS1</i>	Argininosuccinate synthetase 1	3.874
10	211663_x_at	<i>PTGDS</i>	Prostaglandin D2 synthase 21kDa (brain)	3.961
11	203068_at	<i>KLHL21</i>	Kelch-like 21 (Drosophila)	3.970
12	209960_at	<i>HGF</i>	Hepatocyte growth factor	4.796
13	202718_at	<i>IGFBP2</i>	Insulin-like growth factor binding protein 2	4.800
14	210755_at	<i>HGF</i>	Hepatocyte growth factor	5.263
15	212187_x_at	<i>PTGDS</i>	Prostaglandin D2 synthase 21kDa (brain)	6.344
16	211748_x_at	<i>PTGDS</i>	Prostaglandin D2 synthase 21kDa (brain)	6.950
17	210998_s_at	<i>HGF</i>	Hepatocyte growth factor	7.367
18	210997_at	<i>HGF</i>	Hepatocyte growth factor	8.958

t(8;21)-Specific Genes

Genes identified as over-expressed in t(8;21) AML through filtering on fold-change, ANOVA and SAM analysis of probe sets classified as “present in at least 1 sample” and as “present and changing” (Chapter 4).

	Probe Set	Gene Symbol	Gene Title	Fold Change
1	218825_at	<i>EGFL7</i>	EGF-like-domain, multiple 7	3.004
2	209728_at	<i>HLA-DRB4</i>	Major histocompatibility complex, class II, DR beta 4	3.008
3	202242_at	<i>TSPAN7</i>	Tetraspanin 7	3.010
4	207550_at	<i>MPL</i>	Myeloproliferative leukemia virus oncogene	3.029
5	208178_x_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	3.033
6	206187_at	<i>PTGIR</i>	Prostaglandin I2 (prostacyclin) receptor (IP)	3.082
7	210997_at	<i>HGF</i>	Hepatocyte growth factor	3.102
8	222222_s_at	<i>HOMER3</i>	Homer homolog 3 (Drosophila)	3.189
9	207030_s_at	<i>CSRP2</i>	Cysteine and glycine-rich protein 2	3.229
10	209560_s_at	<i>DLK1</i>	delta-like 1 homolog (Drosophila)	3.242
11	204548_at	<i>STAR</i>	Steroidogenic acute regulatory protein	3.244
12	211140_s_at	<i>CASP2</i>	Caspase 2, apoptosis-related cysteine peptidase	3.247
13	212671_s_at	<i>HLA-DQA1</i>	Major histocompatibility complex, class II, DQ alpha 1	3.276
14	207496_at	<i>MS4A2</i>	Membrane-spanning 4-domains, subfamily A, member 2	3.293
15	219790_s_at	<i>NPR3</i>	Natriuretic peptide receptor C/guanylate cyclase C	3.424
16	211549_s_at	<i>HPGD</i>	Hydroxyprostaglandin dehydrogenase 15-(NAD)	3.439
17	212013_at	<i>PXDN</i>	Peroxidasin homolog (Drosophila)	3.533
18	210012_s_at	<i>EWSR1</i>	Ewing sarcoma breakpoint region 1	3.537
19	209543_s_at	<i>CD34</i>	CD34 molecule	3.542
20	219090_at	<i>SLC24A3</i>	Solute carrier family 24 (sodium/potassium /calcium exchanger), member 3	3.590
21	204990_s_at	<i>ITGB4</i>	Integrin, beta 4	3.614
22	202760_s_at	<i>PALM2-AKAP2</i>	PALM2-AKAP2 readthrough transcript	3.653
23	209013_x_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	3.671
24	219837_s_at	<i>CYTL1</i>	Cytokine-like 1	3.731
25	204811_s_at	<i>CACNA2D2</i>	Calcium channel, voltage-dependent, alpha 2/delta subunit 2	3.787
26	204073_s_at	<i>C11orf9</i>	Chromosome 11 open reading frame 9	3.843
27	201015_s_at	<i>JUP</i>	Junction plakoglobin	3.932
28	214183_s_at	<i>TKTL1</i>	Transketolase-like 1	3.938
29	201621_at	<i>NBL1</i>	Neuroblastoma, suppression of tumorigenicity 1	3.986
30	200671_s_at	<i>SPTBN1</i>	Spectrin, beta, non-erythrocytic 1	4.171
31	203913_s_at	<i>HPGD</i>	Hydroxyprostaglandin dehydrogenase 15-(NAD)	4.220
32	203065_s_at	<i>CAV1</i>	Caveolin 1, caveolae protein, 22kDa	4.439
33	204468_s_at	<i>TIE1</i>	Tyrosine kinase with immunoglobulin-like and EGF-like domains 1	4.551
34	219686_at	<i>STK32B</i>	Serine/threonine kinase 32B	4.716
35	204885_s_at	<i>MSLN</i>	Mesothelin	5.182
36	209170_s_at	<i>GPM6B</i>	Glycoprotein M6B	5.236
37	218876_at	<i>TPPP3</i>	Tubulin polymerization-promoting protein family member 3	5.312
38	208534_s_at	<i>RASA4</i>	RAS p21 protein activator 4	5.353
39	209010_s_at	<i>TRIO</i>	Triple functional domain (PTPRF interacting)	5.453
40	209167_at	<i>GPM6B</i>	Glycoprotein M6B	6.328
41	206726_at	<i>PGDS</i>	Prostaglandin D2 synthase, hematopoietic	6.844
42	206940_s_at	<i>POU4F1</i>	POU class 4 homeobox 1	7.522
43	212097_at	<i>CAV1</i>	Caveolin 1, caveolae protein, 22kDa	8.016
44	213194_at	<i>ROBO1</i>	Roundabout, axon guidance receptor, homolog 1	8.209
45	207741_x_at	<i>TBSAB1</i>	Tryptase alpha/beta 1	8.355
46	204086_at	<i>PRAME</i>	Preferentially expressed antigen in melanoma	8.990
47	207134_x_at	<i>TPSB2</i>	Tryptase beta 2	9.150

48	201655_s_at	<i>HSPG2</i>	Heparan sulfate proteoglycan 2	9.649
49	205683_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	10.01
50	216474_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	12.06
51	216831_s_at	<i>RUNX1T1</i>	Runt-related transcription factor 1	13.09
52	205529_s_at	<i>RUNX1T1</i>	Runt-related transcription factor 1	13.25
53	215382_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	13.61
54	217023_x_at	<i>TPSB2</i>	Tryptase alpha/beta 1	14.81
55	210744_s_at	<i>IL5RA</i>	Interleukin 5 receptor, alpha	16.56
56	205528_s_at	<i>RUNX1T1</i>	Runt-related transcription factor 1	18.12
57	210084_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	21.71
58	211517_s_at	<i>IL5RA</i>	Interleukin 5 receptor, alpha	25.35
59	211341_at	<i>POU4F1</i>	POU class 4 homeobox 1	39.08
60	206622_at	<i>TRH</i>	Thyrotropin-releasing hormone	39.27

inv(16)-Specific Genes

Genes identified as over-expressed in inv(16) AML through filtering on fold-change, ANOVA and SAM analysis of probe sets classified as “present in at least 1 sample” and as “present and changing” (Chapter 4).

	Probe Set	Gene Symbol	Gene Title	Fold Change
1	216953_s_at	<i>WT1</i>	Wilms tumor 1	3.007
2	202627_s_at	<i>SERPINE1</i>	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	3.042
3	209013_x_at	<i>TRIO</i>	Triple functional domain (PTPRF intera	3.068
4	219901_at	<i>FGD6</i>	FYVE, RhoGEF and PH domain conta	3.097
5	219051_x_at	<i>METRN</i>	Meteorin, glial cell differentiation regul	3.112
6	205081_at	<i>CRIP1</i>	Cysteine-rich protein 1 (intestinal)	3.114
7	201841_s_at	<i>HSPB1</i>	Heat shock 27kDa protein 1	3.116
8	208605_s_at	<i>NTRK1</i>	Neurotrophic tyrosine kinase, receptor, type 1	3.135
9	203329_at	<i>PTPRM</i>	Protein tyrosine phosphatase, receptor type, M	3.14
10	203407_at	<i>PPL</i>	Periplakin	3.173
11	215111_s_at	<i>TSC22D1</i>	TSC22 domain family, member 1	3.185
12	211030_s_at	<i>SLC6A6</i>	Solute carrier family 6 (neurotransmitter transporter, taurine), member 6	3.191
13	212998_x_at	<i>HLA-DRB4</i>	Major histocompatibility complex, class II, DQ beta 1	3.235
14	210090_at	<i>ARC</i>	Activity-regulated cytoskeleton-associated protein	3.238
15	202291_s_at	<i>MGP</i>	Matrix Gla protein	3.264
16	202827_s_at	<i>MMP14</i>	Matrix metalloproteinase 14 (membrane-inserted)	3.296
17	203973_s_at	<i>CEBPD</i>	CCAAT/enhancer binding protein (C/EBP), delta	3.343
18	204134_at	<i>PDE2A</i>	Phosphodiesterase 2A,	3.36
19	202145_at	<i>LY6E</i>	Lymphocyte antigen 6 complex, locus E	3.363
20	219308_s_at	<i>AK5</i>	Adenylate kinase 5	3.364
21	211182_x_at	<i>RUNX1</i>	Runt-related transcription factor 1	3.392

22	207075_at	<i>CIAS1</i>	NLR family, pyrin domain containing	3.414
23	47553_at	<i>DFNB31</i>	Deafness, autosomal recessive 31	3.435
24	213895_at	<i>EMP1</i>	Epithelial membrane protein 1	3.445
25	208456_s_at	<i>RRAS2</i>	Related RAS viral (r-ras) oncogene homolog 2	3.456
26	205780_at	<i>BIK</i>	BCL2-interacting killer (apoptosis-inducing)	3.474
27	206067_s_at	<i>WT1</i>	Wilms tumor 1	3.475
28	212803_at	<i>NAB2</i>	NGFI-A binding protein 2 (EGR1 binding protein 2)	3.487
29	201667_at	<i>GJA1</i>	Gap junction protein, alpha 1, 43kDa	3.518
30	201015_s_at	<i>JUP</i>	Junction plakoglobin	3.528
31	200951_s_at	<i>CCND2</i>	Cyclin D2	3.541
32	218825_at	<i>EGFL7</i>	EGF-like-domain, multiple 7	3.549
33	202746_at	<i>ITM2A</i>	Integral membrane protein 2A	3.575
34	209975_at	<i>CYP2E1</i>	Cytochrome P450, family 2, subfamily E, polypeptide	3.588
35	202016_at	<i>MEST</i>	Mesoderm specific transcript homolog (mouse)	3.593
36	201739_at	<i>SGK</i>	Serum/glucocorticoid regulated kinase 1	3.71
37	209386_at	<i>TM4SF1</i>	Transmembrane 4 L six family member 1	3.717
38	217388_s_at	<i>KYNU</i>	Kynureninase (L-kynurenine hydrolase)	3.722
39	202628_s_at	<i>SERPINE1</i>	Serpin peptidase inhibitor,	3.778
40	205157_s_at	<i>KRT17</i>	Keratin 17	3.8
41	215034_s_at	<i>TM4SF1</i>	Transmembrane 4 L six family member 1	3.847
42	204787_at	<i>VSIG4</i>	V-set and immunoglobulin domain containing 4	3.917
43	204011_at	<i>SPRY2</i>	Sprouty homolog 2 (Drosophila)	3.919
44	203913_s_at	<i>HPGD</i>	Hydroxyprostaglandin dehydrogenase 15-(NAD)	3.934
45	216016_at	<i>NLRP3</i>	NLR family, pyrin domain containing 3	3.953
46	205819_at	<i>MARCO</i>	Macrophage receptor with collagenous structure	3.958
47	204794_at	<i>DUSP2</i>	Dual specificity phosphatase 2	4.008
48	220091_at	<i>SLC2A6</i>	Solute carrier family 2 (facilitated glucose transporter), member 6	4.01
49	202340_x_at	<i>NR4A1</i>	Nuclear receptor subfamily 4, group A, member 1	4.026
50	212667_at	<i>SPARC</i>	Secreted protein, acidic, cysteine-rich (osteonectin)	4.035
51	207741_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	4.139
52	210997_at	<i>HGF</i>	Hepatocyte growth factor	4.16
53	209576_at	<i>GNAI1</i>	Guanine nucleotide binding protein (G protein), alpha inhibiting activity	4.232
54	207134_x_at	<i>TPSB2</i>	Tryptase beta 2	4.239
55	217849_s_at	<i>BPB</i>	CDC42 binding protein kinase beta (DMPK-like)	4.358
56	216442_x_at	<i>FN1</i>	Fibronectin 1	4.399
57	209487_at	<i>RBPMS</i>	RNA binding protein with multiple splicing	4.432
58	202007_at	<i>NID1</i>	Nidogen 1	4.45
59	212464_s_at	<i>FN1</i>	Fibronectin 1	4.563
60	205683_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	4.612
61	210202_s_at	<i>BIN1</i>	Bridging integrator 1	4.63
62	214020_x_at	<i>ITGB5</i>	Integrin, beta 5	4.64
63	216474_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	4.712
64	206682_at	<i>CLEC10A</i>	C-type lectin domain family 10, member A	4.764
65	202747_s_at	<i>ITM2A</i>	Integral membrane protein 2A	4.771
66	202283_at	<i>SERPINF1</i>	Serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigmen	4.84
67	217023_x_at	<i>TPSB2</i>	Tryptase alpha/beta 1	4.935
68	211719_x_at	<i>FN1</i>	Fibronectin	4.945
69	201496_x_at	<i>MYH11</i>	Myosin, heavy chain 11, smooth muscle	4.967
71	208450_at	<i>LGALS2</i>	Lectin, galactoside-binding, soluble, 2	5.059
70	34210_at	<i>CD52</i>	CD52 molecule	4.977
71	208450_at	<i>LGALS2</i>	Lectin, galactoside-binding, soluble, 2	5.059
72	204163_at	<i>EMILIN1</i>	Elastin microfibril interfacer 1	5.24

73	210139_s_at	<i>PMP22</i>	Peripheral myelin protein 22	5.296
74	215382_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	5.322
75	214390_s_at	<i>BCAT1</i>	Branched chain aminotransferase 1, cytosolic	5.524
76	215666_at	<i>HLA-DRB4</i>	Major histocompatibility complex, class II, DR beta 4	5.615
77	206622_at	<i>TRH</i>	Thyrotropin-releasing hormone	5.722
78	201005_at	<i>CD9</i>	CD9 molecule	5.828
79	205239_at	<i>AREG</i>	Amphiregulin	5.858
80	210495_x_at	<i>FN1</i>	Fibronectin 1	5.859
81	200665_s_at	<i>SPARC</i>	Secreted protein, acidic, cysteine-rich (osteonectin)	5.889
82	210084_x_at	<i>TPSAB1</i>	Tryptase alpha/beta 1	5.904
83	204661_at	<i>CD52</i>	CD52 molecule	5.964
84	221914_at	<i>SYN1</i>	Synapsin I	6.055
85	208978_at	<i>CRIP2</i>	Cysteine-rich protein 2	6.491
86	213479_at	<i>NPTX2</i>	Neuronal pentraxin II	6.53
87	203036_s_at	<i>MTSS1</i>	Metastasis suppressor 1	6.634
88	209387_s_at	<i>TM4SF1</i>	Transmembrane 4 L six family member 1	6.767
89	201324_at	<i>EMP1</i>	Epithelial membrane protein 1	6.792
90	205987_at	<i>CD1C</i>	CD1c molecule	7.03
91	205330_at	<i>MN1</i>	Meningioma (disrupted in balanced translocation) 1	7.947
92	203939_at	<i>NT5E</i>	5'-nucleotidase, ecto (CD73)	8.023
93	208789_at	<i>PTRF</i>	polymerase I and transcript release factor	8.23
94	201325_s_at	<i>EMP1</i>	Epithelial membrane protein 1	8.544
95	209488_s_at	<i>RBPMS</i>	RNA binding protein with multiple splicing	8.791
96	212358_at	<i>CLIP3</i>	CAP-GLY domain containing linker protein 3	9.902
97	201497_x_at	<i>MYH11</i>	Myosin, heavy chain 11, smooth muscle	10.32
98	218876_at	<i>TPPP3</i>	Tubulin polymerization-promoting protein family member 3	10.81
99	204885_s_at	<i>MSLN</i>	Mesothelin	15.47
100	206135_at	<i>ST18</i>	Suppression of tumorigenicity 18 (breast carcinoma) (zinc finger protein)	26.59

Normal Karyotype AML

Genes identified as over-expressed in normal karyotype AML through filtering on fold-change and Bayesian *t*-testing (Chapter 5).

	Probe set	Gene Symbol	Gene Title	Fold Change
1	213147_at	<i>HOXA10</i>	Homeobox A10	4.030
2	218000_s_at	<i>PHLDA1</i>	Pleckstrin homology-like domain, family A, member 1	4.084
3	212803_at	<i>NAB2</i>	NGFI-A binding protein 2 (EGR1 binding protein 2)	4.141
4	214488_at	<i>RAP2B</i>	RAP2B, member of RAS oncogene family	4.171
5	209983_s_at	<i>NRXN2</i>	Neurexin 2	4.195
6	201243_s_at	<i>ATP1B1</i>	ATPase, Na ⁺ /K ⁺	4.252
7	206532_at	-	<i>Human Ini1 mRNA, complete cds</i>	4.296
8	204069_at	<i>MEIS1</i>	Meis homeobox 1	4.446
9	206674_at	<i>FLT3</i>	fms-related tyrosine kinase 3	4.584
10	205453_at	<i>HOXB2</i>	Homeobox B2	4.621
11	214880_x_at	<i>CALD1</i>	Caldesmon 1	4.988
12	211039_at	<i>CHRNA1</i>	Cholinergic receptor, nicotinic, alpha 1 (muscle)	5.538
13	219908_at	<i>DKK2</i>	Dickkopf homolog 2 (<i>Xenopus laevis</i>)	5.705

14	213150_at	<i>HOXA10</i>	Homeobox A10	5.737
15	211125_x_at	<i>GRIN1</i>	Glutamate receptor, ionotropic, N-methyl D-aspartate 1	6.637
16	215441_at	-	<i>Human clone 23842 mRNA sequence</i>	7.490
17	216011_at	<i>SLC39A9</i>	Solute carrier family 39 (zinc transporter), member 9	7.594
18	206310_at	<i>SPINK2</i>	Serine peptidase inhibitor, Kazal type 2	8.029
19	213110_s_at	<i>COL4A5</i>	Collagen, type IV, alpha 5	10.59
20	214651_s_at	<i>HOXA9</i>	Homeobox A9	11.04
21	213844_at	<i>HOXA5</i>	Homeobox A5	11.27
22	206067_s_at	<i>WT1</i>	Wilms tumor 1	14.05
23	217698_at	-	/DB_XREF=AV651668 /CLONE=GLCCSC04	16.94
24	209905_at	<i>HOXA9</i>	Homeobox A9	17.45
25	205366_s_at	<i>HOXB6</i>	Homeobox B6	36.73

Appendix III: Cel Files

See attached DVD

The image files (Cel files) used for the favourable AML gene expression profiling analysis (Master Experiment) (Chapter 3, 4).

The image files (Cel files) used for the normal karyotype AML gene expression profiling analysis (Master Experiment, MILE and GEO) (Chapter 5).

References

- Abkowitz, J.L., Linenberger, M.L., Newton, M.A., Shelton, G.H., Ott, R.L., & Gutter, P. (1990) Evidence for the maintenance of hematopoiesis in a large animal by the sequential activation of stem-cell clones. *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 9062-9066.
- Abramovich, C. & Humphries, R.K. (2005) Hox regulation of normal and leukemic hematopoietic stem cells. *Curr. Opin. Hematol.*, **12**, 210-216.
- Abu-Duhier, F.M., Goodeve, A.C., Wilson, G.A., Gari, M.A., Peake, I.R., Rees, D.C., Vandenberghe, E.A., Winship, P.R., & Reilly, J.T. (2000) FLT3 internal tandem duplication mutations in adult acute myeloid leukaemia define a high-risk group. *Br. J. Haematol.*, **111**, 190-195.
- Ahn, J.H., Brignole, E.J., III, & Hayward, G.S. (1998) Disruption of PML subnuclear domains by the acidic IE1 protein of human cytomegalovirus is mediated through interaction with PML and may modulate a RING finger-dependent cryptic transactivator function of PML. *Mol. Cell Biol.*, **18**, 4899-4913.
- Akashi, K., Traver, D., Miyamoto, T., & Weissman, I.L. (2000) A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*, **404**, 193-197.
- Alcalay, M., Tiacci, E., Bergomas, R., Bigerna, B., Venturini, E., Minardi, S.P., Meani, N., Diverio, D., Bernard, L., Tizzoni, L., Volorio, S., Luzi, L., Colombo, E., Lo, C.F., Mecucci, C., Falini, B., & Pelicci, P.G. (2005) Acute myeloid leukemia bearing cytoplasmic nucleophosmin (NPMc+ AML) shows a distinct gene expression profile characterized by up-regulation of genes involved in stem-cell maintenance. *Blood*, **106**, 899-902.
- Andersson, A., Eden, P., Lindgren, D., Nilsson, J., Lassen, C., Heldrup, J., Fontes, M., Borg, A., Mitelman, F., Johansson, B., Hoglund, M., & Fioretos, T. (2005) Gene expression profiling of leukemic cell lines reveals conserved molecular signatures among subtypes with specific genetic aberrations. *Leukemia*, **19**, 1042-1050.
- Andreeff, M. (1986) Cell kinetics of leukemia. *Semin. Hematol.*, **23**, 300-314.
- Andreeff, M., Ruvolo, V., Gadgil, S., Zeng, C., Coombes, K., Chen, W., Kornblau, S., Baron, A.E., & Drabkin, H.A. (2008) HOX expression patterns identify a common signature for favorable AML. *Leukemia*, **22**, 2041-2047.

- Andrews,R.G., Torok-Storb,B., & Bernstein,I.D. (1983) Myeloid-associated differentiation antigens on stem cells and their progeny identified by monoclonal antibodies. *Blood*, **62**, 124-132.
- Appelbaum,F.R., Kopecky,K.J., Tallman,M.S., Slovak,M.L., Gundacker,H.M., Kim,H.T., Dewald,G.W., Kantarjian,H.M., Pierce,S.R., & Estey,E.H. (2006) The clinical spectrum of adult acute myeloid leukaemia associated with core binding factor translocations. *Br.J.Haematol.*, **135**, 165-173.
- Arai,F., Hirao,A., Ohmura,M., Sato,H., Matsuoka,S., Takubo,K., Ito,K., Koh,G.Y., & Suda,T. (2004) Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche. *Cell*, **118**, 149-161.
- Arber,D.A. (2001) Realistic pathologic classification of acute myeloid leukemias. *Am.J.Clin.Pathol.*, **115**, 552-560.
- Argiropoulos,B. & Humphries,R.K. (2007) Hox genes in hematopoiesis and leukemogenesis. *Oncogene*, **26**, 6766-6776.
- Ariyaratana,S. & Loeb,D.M. (2007) The role of the Wilms tumour gene (WT1) in normal and malignant haematopoiesis. *Expert.Rev.Mol.Med.*, **9**, 1-17.
- Arthur,D.C. & Bloomfield,C.D. (1983) Association of partial deletion of the long arm of chromosome 16 and bone marrow eosinophilia in acute non-lymphocytic leukemia. *Blood*, **62**, 931.
- Attar,E.C. & Scadden,D.T. (2004) Regulation of hematopoietic stem cell growth. *Leukemia*, **18**, 1760-1768.
- Avecilla,S.T., Hattori,K., Heissig,B., Tejada,R., Liao,F., Shido,K., Jin,D.K., Dias,S., Zhang,F., Hartman,T.E., Hackett,N.R., Crystal,R.G., Witte,L., Hicklin,D.J., Bohlen,P., Eaton,D., Lyden,D., de,S.F., & Rafii,S. (2004) Chemokine-mediated interaction of hematopoietic progenitors with the bone marrow vascular niche is required for thrombopoiesis. *Nat.Med.*, **10**, 64-71.
- Ayton,P.M. & Cleary,M.L. (2001) Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene*, **20**, 5695-5707.
- Baldi,P. & Long,A.D. (2001) A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. *Bioinformatics.*, **17**, 509-519.
- Baldus,C.D., Mrozek,K., Marcucci,G., & Bloomfield,C.D. (2007) Clinical outcome of de novo acute myeloid leukaemia patients with normal cytogenetics is affected by molecular genetic alterations: a concise review. *Br.J.Haematol.*, **137**, 387-400.

- Baldus,C.D., Tanner,S.M., Kusewitt,D.F., Liyanarachchi,S., Choi,C., Caligiuri,M.A., Bloomfield,C.D., & de la,C.A. (2003a) BAALC, a novel marker of human hematopoietic progenitor cells. *Exp.Hematol.*, **31**, 1051-1056.
- Baldus,C.D., Tanner,S.M., Ruppert,A.S., Whitman,S.P., Archer,K.J., Marcucci,G., Caligiuri,M.A., Carroll,A.J., Vardiman,J.W., Powell,B.L., Allen,S.L., Moore,J.O., Larson,R.A., Kolitz,J.E., de la,C.A., & Bloomfield,C.D. (2003b) BAALC expression predicts clinical outcome of de novo acute myeloid leukemia patients with normal cytogenetics: a Cancer and Leukemia Group B Study. *Blood*, **102**, 1613-1618.
- Barber,N., Belov,L., & Christopherson,R.I. (2008) All-trans retinoic acid induces different immunophenotypic changes on human HL60 and NB4 myeloid leukaemias. *Leuk.Res.*, **32**, 315-322.
- Barjesteh van Waalwijk van Doorn-Khosrovani, Erpelinck,C., van Putten,W.L., Valk,P.J., van der Poel-van de Luytgaarde, Hack,R., Slater,R., Smit,E.M., Beverloo,H.B., Verhoef,G., Verdonck,L.F., Ossenkoppele,G.J., Sonneveld,P., de Greef,G.E., Lowenberg,B., & Delwel,R. (2003) High EVI1 expression predicts poor survival in acute myeloid leukemia: a study of 319 de novo AML patients. *Blood*, **101**, 837-845.
- Barragan,E., Cervera,J., Bolufer,P., Ballester,S., Martin,G., Fernandez,P., Collado,R., Sayas,M.J., & Sanz,M.A. (2004) Prognostic implications of Wilms' tumor gene (WT1) expression in patients with de novo acute myeloid leukemia. *Haematologica*, **89**, 926-933.
- Barreda,D.R., Hanington,P.C., & Belosevic,M. (2004) Regulation of myeloid development and function by colony stimulating factors. *Dev.Comp Immunol.*, **28**, 509-554.
- Basecke,J., Whelan,J.T., Griesinger,F., & Bertrand,F.E. (2006) The MLL partial tandem duplication in acute myeloid leukaemia. *Br.J.Haematol.*, **135**, 438-449.
- Baxter,E.J., Scott,L.M., Campbell,P.J., East,C., Fourouclas,N., Swanton,S., Vassiliou,G.S., Bench,A.J., Boyd,E.M., Curtin,N., Scott,M.A., Erber,W.N., & Green,A.R. (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet*, **365**, 1054-1061.
- Beillard,E., Pallisgaard,N., van,d.V., V, Bi,W., Dee,R., van der,S.E., Delabesse,E., Macintyre,E., Gottardi,E., Saglio,G., Watzinger,F., Lion,T., van Dongen,J.J., Hokland,P., & Gabert,J. (2003) Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction

- (RQ-PCR) - a Europe against cancer program. *Leukemia*, **17**, 2474-2486.
- Bellantuono, I. (2004) Haemopoietic stem cells. *Int.J Biochem.Cell Biol.*, **36**, 607-620.
- Bennett, J.M., Catovsky, D., Daniel, M.T., Flandrin, G., Galton, D.A., Gralnick, H.R., & Sultan, C. (1976) Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br.J.Haematol.*, **33**, 451-458.
- Bennett, J.M., Catovsky, D., Daniel, M.T., Flandrin, G., Galton, D.A., Gralnick, H.R., & Sultan, C. (1982) Proposals for the classification of the myelodysplastic syndromes. *Br.J.Haematol.*, **51**, 189-199.
- Beretta, C., Gaipa, G., Rossi, V., Bernasconi, S., Spinelli, O., Dell'oro, M.G., Rizzari, C., Rambaldi, A., Biondi, A., & Cazzaniga, G. (2004) Development of a quantitative-PCR method for specific FLT3/ITD monitoring in acute myeloid leukemia. *Leukemia*, **18**, 1441-1444.
- Bergmann, L., Miething, C., Maurer, U., Brieger, J., Karakas, T., Weidmann, E., & Hoelzer, D. (1997) High levels of Wilms' tumor gene (wt1) mRNA in acute myeloid leukemias are associated with a worse long-term outcome. *Blood*, **90**, 1217-1225.
- Beutler, E., Yeh, M., & Fairbanks, V.F. (1962) The normal human female as a mosaic of X-chromosome activity: studies using the gene for C-6-PD-deficiency as a marker. *Proc.Natl.Acad.Sci.U.S.A.*, **48**, 9-16.
- Bienz, M., Ludwig, M., Leibundgut, E.O., Mueller, B.U., Ratschiller, D., Solenthaler, M., Fey, M.F., & Pabst, T. (2005) Risk assessment in patients with acute myeloid leukemia and a normal karyotype. *Clin.Cancer Res.*, **11**, 1416-1424.
- Biondi, A., Rambaldi, A., Pandolfi, P.P., Rossi, V., Giudici, G., Alcalay, M., Lo, C.F., Diverio, D., Pogliani, E.M., Lanzi, E.M., & . (1992) Molecular monitoring of the myl/retinoic acid receptor-alpha fusion gene in acute promyelocytic leukemia by polymerase chain reaction. *Blood*, **80**, 492-497.
- Birg, F., Courcoul, M., Rosnet, O., Bardin, F., Pebusque, M.J., Marchetto, S., Tabilio, A., Mannoni, P., & Birnbaum, D. (1992) Expression of the FMS/KIT-like gene FLT3 in human acute leukemias of the myeloid and lymphoid lineages. *Blood*, **80**, 2584-2593.
- Birnie, G.D. (1988) The HL60 cell line: a model system for studying human myeloid cell differentiation. *Br.J.Cancer Suppl*, **9**, 41-45.
- Bjornsson, J.M., Larsson, N., Brun, A.C., Magnusson, M., Andersson, E., Lundstrom, P., Larsson, J., Repetowska, E., Ehinger, M.,

- Humphries,R.K., & Karlsson,S. (2003) Reduced proliferative capacity of hematopoietic stem cells deficient in Hoxb3 and Hoxb4. *Mol.Cell Biol.*, **23**, 3872-3883.
- Boissel,N., Renneville,A., Biggio,V., Philippe,N., Thomas,X., Cayuela,J.M., Terre,C., Tigaud,I., Castaigne,S., Raffoux,E., De,B.S., Fenaux,P., Dombret,H., & Preudhomme,C. (2005) Prevalence, clinical profile, and prognosis of NPM mutations in AML with normal karyotype. *Blood*, **106**, 3618-3620.
- Bonnet,D. (2005) Leukemic stem cells show the way. *Folia Histochem.Cytobiol.*, **43**, 183-186.
- Bonnet,D. & Dick,J.E. (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat.Med.*, **3**, 730-737.
- Borset,M., Hjorth-Hansen,H., Seidel,C., Sundan,A., & Waage,A. (1996a) Hepatocyte growth factor and its receptor c-met in multiple myeloma. *Blood*, **88**, 3998-4004.
- Borset,M., Lien,E., Espevik,T., Helseth,E., Waage,A., & Sundan,A. (1996b) Concomitant expression of hepatocyte growth factor/scatter factor and the receptor c-MET in human myeloma cell lines. *J.Biol.Chem.*, **271**, 24655-24661.
- Breitman,T.R., Collins,S.J., & Keene,B.R. (1981) Terminal differentiation of human promyelocytic leukemic cells in primary culture in response to retinoic acid. *Blood*, **57**, 1000-1004.
- Breitman,T.R., Selonick,S.E., & Collins,S.J. (1980) Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. *Proc.Natl.Acad.Sci.U.S.A*, **77**, 2936-2940.
- Brun,A.C., Bjornsson,J.M., Magnusson,M., Larsson,N., Leveen,P., Ehinger,M., Nilsson,E., & Karlsson,S. (2004) Hoxb4-deficient mice undergo normal hematopoietic development but exhibit a mild proliferation defect in hematopoietic stem cells. *Blood*, **103**, 4126-4133.
- Buick,R.N., Minden,M.D., & McCulloch,E.A. (1979) Self-renewal in culture of proliferative blast progenitor cells in acute myeloblastic leukemia. *Blood*, **54**, 95-104.
- Buijs,A., van,R.L., Molijn,A.C., Davis,J.N., Vertegaal,A.C., Potter,M.D., Adams,C., van,B.S., Zwarthoff,E.C., Roussel,M.F., & Grosveld,G.C. (2000) The MN1-TEL fusion protein, encoded by the translocation (12;22)(p13;q11) in myeloid leukemia, is a transcription factor with transforming activity. *Mol.Cell Biol.*, **20**, 9281-9293.

- Bullinger,L., Dohner,K., Bair,E., Frohling,S., Schlenk,R.F., Tibshirani,R., Dohner,H., & Pollack,J.R. (2004) Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N.Engl.J.Med.*, **350**, 1605-1616.
- Bullinger,L., Rucker,F.G., Kurz,S., Du,J., Scholl,C., Sander,S., Corbacioglu,A., Lottaz,C., Krauter,J., Frohling,S., Ganser,A., Schlenk,R.F., Dohner,K., Pollack,J.R., & Dohner,H. (2007) Gene-expression profiling identifies distinct subclasses of core binding factor acute myeloid leukemia. *Blood*, **110**, 1291-1300.
- Burel,S.A., Harakawa,N., Zhou,L., Pabst,T., Tenen,D.G., & Zhang,D.E. (2001) Dichotomy of AML1-ETO functions: growth arrest versus block of differentiation. *Mol.Cell Biol.*, **21**, 5577-5590.
- Burnett,A.K., Grimwade,D., Solomon,E., Wheatley,K., & Goldstone,A.H. (1999) Presenting white blood cell count and kinetics of molecular remission predict prognosis in acute promyelocytic leukemia treated with all-trans retinoic acid: result of the Randomized MRC Trial. *Blood*, **93**, 4131-4143.
- Bussolino,F., Di Renzo,M.F., Ziche,M., Bocchietto,E., Olivero,M., Naldini,L., Gaudino,G., Tamagnone,L., Coffer,A., & Comoglio,P.M. (1992) Hepatocyte growth factor is a potent angiogenic factor which stimulates endothelial cell motility and growth. *J.Cell Biol.*, **119**, 629-641.
- Byrd,J.C., Mrozek,K., Dodge,R.K., Carroll,A.J., Edwards,C.G., Arthur,D.C., Pettenati,M.J., Patil,S.R., Rao,K.W., Watson,M.S., Koduru,P.R., Moore,J.O., Stone,R.M., Mayer,R.J., Feldman,E.J., Davey,F.R., Schiffer,C.A., Larson,R.A., & Bloomfield,C.D. (2002) Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). *Blood*, **100**, 4325-4336.
- Call,K.M., Glaser,T., Ito,C.Y., Buckler,A.J., Pelletier,J., Haber,D.A., Rose,E.A., Kral,A., Yeger,H., Lewis,W.H., & . (1990) Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell*, **60**, 509-520.
- Calvi,L.M., Adams,G.B., Weibrecht,K.W., Weber,J.M., Olson,D.P., Knight,M.C., Martin,R.P., Schipani,E., Divieti,P., Bringhurst,F.R., Milner,L.A., Kronenberg,H.M., & Scadden,D.T. (2003) Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature*, **425**, 841-846.
- Cameron,S., Taylor,D.S., TePas,E.C., Speck,N.A., & Mathey-Prevot,B. (1994) Identification of a critical regulatory site in the human

- interleukin-3 promoter by in vivo footprinting. *Blood*, **83**, 2851-2859.
- Campana,D. (2003) Determination of minimal residual disease in leukaemia patients. *Br.J Haematol.*, **121**, 823-838.
- Campana,D. & Coustan-Smith,E. (1999) Detection of minimal residual disease in acute leukemia by flow cytometry. *Cytometry*, **38**, 139-152.
- Campana,D. & Coustan-Smith,E. (2002) Advances in the immunological monitoring of childhood acute lymphoblastic leukaemia. *Best.Pract.Res.Clin.Haematol.*, **15**, 1-19.
- Campana,D. & Pui,C.H. (1995) Detection of minimal residual disease in acute leukemia: methodologic advances and clinical significance. *Blood*, **85**, 1416-1434.
- Campo Dell'Orto,M., Zangrando,A., Trentin,L., Li,R., Liu,W.M., te,K.G., Basso,G., & Kohlmann,A. (2007) New data on robustness of gene expression signatures in leukemia: comparison of three distinct total RNA preparation procedures. *BMC.Genomics*, **8**, 188.
- Candoni,A., Tiribelli,M., Toffoletti,E., Cilloni,D., Chiarvesio,A., Michelutti,A., Simeone,E., Pipan,C., Saglio,G., & Fanin,R. (2009) Quantitative assessment of WT1 gene expression after allogeneic stem cell transplantation is a useful tool for monitoring minimal residual disease in acute myeloid leukemia. *Eur.J.Haematol.*, **82**, 61-68.
- Carella,C., Bonten,J., Sirma,S., Kranenburg,T.A., Terranova,S., Klein-Geltink,R., Shurtleff,S., Downing,J.R., Zwarthoff,E.C., Liu,P.P., & Grosveld,G.C. (2007) MN1 overexpression is an important step in the development of inv(16) AML. *Leukemia*, **21**, 1679-1690.
- Carow,C.E., Levenstein,M., Kaufmann,S.H., Chen,J., Amin,S., Rockwell,P., Witte,L., Borowitz,M.J., Civin,C.I., & Small,D. (1996) Expression of the hematopoietic growth factor receptor FLT3 (STK-1/Flk2) in human leukemias. *Blood*, **87**, 1089-1096.
- Cassinat,B., De,B.S., Kelaidi,C., Ades,L., Zassadowski,F., Guillemot,I., Schlageter,M.H., Raffoux,E., Harousseau,J.L., Legrand,O., Escoffre-Barbe,M., Reman,O., Gardembas,M., Himberlin,C., Cahn,J.Y., Guyotat,D., Bouscary,D., Parry,A., Rousselot,P., Baruchel,A., Dombret,H., Chevret,S., Fenaux,P., & Chomienne,C. (2009) When can real-time quantitative RT-PCR effectively define molecular relapse in acute promyelocytic leukemia patients? (Results of the French Belgian Swiss APL Group). *Leuk.Res.*, **33**, 1178-1182.

- Castaigne,S., Balitrand,N., de,T.H., Dejean,A., Degos,L., & Chomienne,C. (1992) A PML/retinoic acid receptor alpha fusion transcript is constantly detected by RNA-based polymerase chain reaction in acute promyelocytic leukemia. *Blood*, **79**, 3110-3115.
- Castaigne,S., Chomienne,C., Daniel,M.T., Ballerini,P., Berger,R., Fenaux,P., & Degos,L. (1990) All-trans retinoic acid as a differentiation therapy for acute promyelocytic leukemia. I. Clinical results. *Blood*, **76**, 1704-1709.
- Castilla,L.H., Garrett,L., Adya,N., Orlic,D., Dutra,A., Anderson,S., Owens,J., Eckhaus,M., Bodine,D., & Liu,P.P. (1999) The fusion gene Cbfb-MYH11 blocks myeloid differentiation and predisposes mice to acute myelomonocytic leukaemia. *Nat.Genet.*, **23**, 144-146.
- Chang,K.S., Lu,J.F., Wang,G., Trujillo,J.M., Estey,E., Cork,A., Chu,D.T., Freireich,E.J., & Stass,S.A. (1992) The t(15;17) breakpoint in acute promyelocytic leukemia cluster within two different sites of the myl gene: targets for the detection of minimal residual disease by the polymerase chain reaction. *Blood*, **79**, 554-558.
- Chariot,A. & Castronovo,V. (1996) Detection of HOXA1 expression in human breast cancer. *Biochem.Biophys.Res.Comm.*, **222**, 292-297.
- Charrad,R.S., Li,Y., Delpech,B., Balitrand,N., Clay,D., Jasmin,C., Chomienne,C., & Smadja-Joffe,F. (1999) Ligation of the CD44 adhesion molecule reverses blockage of differentiation in human acute myeloid leukemia. *Nat.Med.*, **5**, 669-676.
- Chattopadhyay,N., Butters,R.R., & Brown,E.M. (2001) Agonists of the retinoic acid- and retinoid X-receptors inhibit hepatocyte growth factor secretion and expression in U87 human astrocytoma cells. *Brain Res.Mol.Brain Res.*, **87**, 100-108.
- Chen,J.S., Coustan-Smith,E., Suzuki,T., Neale,G.A., Mihara,K., Pui,C.H., & Campana,D. (2001) Identification of novel markers for monitoring minimal residual disease in acute lymphoblastic leukemia. *Blood*, **97**, 2115-2120.
- Chen,W., Jones,D., Medeiros,L.J., Luthra,R., & Lin,P. (2005) Acute myeloid leukaemia with FLT3 gene mutations of both internal tandem duplication and point mutation type. *Br.J.Haematol.*, **130**, 726-728.
- Chen,W., Rassidakis,G.Z., & Medeiros,L.J. (2006) Nucleophosmin gene mutations in acute myeloid leukemia. *Arch.Pathol.Lab Med.*, **130**, 1687-1692.

- Cheng,T., Rodrigues,N., Shen,H., Yang,Y., Dombkowski,D., Sykes,M., & Scadden,D.T. (2000) Hematopoietic stem cell quiescence maintained by p21cip1/waf1. *Science*, **287**, 1804-1808.
- Cheshier,S.H., Morrison,S.J., Liao,X., & Weissman,I.L. (1999) In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proc.Natl.Acad.Sci.U.S.A*, **96**, 3120-3125.
- Chiba,S., Shibuya,K., Piao,Y.F., Tojo,A., Sasaki,N., Matsuki,S., Miyagawa,K., Miyazono,K., & Takaku,F. (1990) Identification and cellular distribution of distinct proteins forming human GM-CSF receptor. *Cell Regul.*, **1**, 327-335.
- Chiorazzi,N., Rai,K.R., & Ferrarini,M. (2005) Chronic lymphocytic leukemia. *N.Engl.J.Med.*, **352**, 804-815.
- Chirgadze,D.Y., Hepple,J., Byrd,R.A., Sowdhamini,R., Blundell,T.L., & Gherardi,E. (1998) Insights into the structure of hepatocyte growth factor/scatter factor (HGF/SF) and implications for receptor activation. *FEBS Lett.*, **430**, 126-129.
- Chou,W.C., Tang,J.L., Wu,S.J., Tsay,W., Yao,M., Huang,S.Y., Huang,K.C., Chen,C.Y., Huang,C.F., & Tien,H.F. (2007) Clinical implications of minimal residual disease monitoring by quantitative polymerase chain reaction in acute myeloid leukemia patients bearing nucleophosmin (NPM1) mutations. *Leukemia*, **21**, 998-1004.
- Chung,N.G., Buxhofer-Ausch,V., & Radich,J.P. (2006) The detection and significance of minimal residual disease in acute and chronic leukemia. *Tissue Antigens*, **68**, 371-385.
- Cilloni,D., Gottardi,E., De,M.D., Serra,A., Volpe,G., Messa,F., Rege-Cambrin,G., Guerrasio,A., Divona,M., Lo,C.F., & Saglio,G. (2002) Quantitative assessment of WT1 expression by real time quantitative PCR may be a useful tool for monitoring minimal residual disease in acute leukemia patients. *Leukemia*, **16**, 2115-2121.
- Cilloni,D., Gottardi,E., Fava,M., Messa,F., Carturan,S., Busca,A., Guerrasio,A., & Saglio,G. (2003) Usefulness of quantitative assessment of the WT1 gene transcript as a marker for minimal residual disease detection. *Blood*, **102**, 773-774.
- Claxton,D.F., Liu,P., Hsu,H.B., Marlton,P., Hester,J., Collins,F., Deisseroth,A.B., Rowley,J.D., & Siciliano,M.J. (1994) Detection of fusion transcripts generated by the inversion 16 chromosome in acute myelogenous leukemia. *Blood*, **83**, 1750-1756.

- Coppes,M.J., Campbell,C.E., & Williams,B.R. (1993) The role of WT1 in Wilms tumorigenesis. *FASEB J*, **7**, 886-895.
- Coustan-Smith,E., Behm,F.G., Sanchez,J., Boyett,J.M., Hancock,M.L., Raimondi,S.C., Rubnitz,J.E., Rivera,G.K., Sandlund,J.T., Pui,C.H., & Campana,D. (1998) Immunological detection of minimal residual disease in children with acute lymphoblastic leukaemia. *Lancet*, **351**, 550-554.
- Cozzio,A., Passegue,E., Ayton,P.M., Karsunky,H., Cleary,M.L., & Weissman,I.L. (2003) Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. *Genes Dev.*, **17**, 3029-3035.
- Cross,N.C. (1995) Quantitative PCR techniques and applications. *Br.J.Haematol.*, **89**, 693-697.
- Cross,N.C., Feng,L., Chase,A., Bungey,J., Hughes,T.P., & Goldman,J.M. (1993) Competitive polymerase chain reaction to estimate the number of BCR-ABL transcripts in chronic myeloid leukemia patients after bone marrow transplantation. *Blood*, **82**, 1929-1936.
- Dameshek,W. (1951) The "humors" and idiopathic thrombocytopenic purpura. *Blood*, **6**, 954-955.
- Date,K., Matsumoto,K., Kuba,K., Shimura,H., Tanaka,M., & Nakamura,T. (1998) Inhibition of tumor growth and invasion by a four-kringle antagonist (HGF/NK4) for hepatocyte growth factor. *Oncogene*, **17**, 3045-3054.
- Date,K., Matsumoto,K., Shimura,H., Tanaka,M., & Nakamura,T. (1997) HGF/NK4 is a specific antagonist for pleiotrophic actions of hepatocyte growth factor. *FEBS Lett.*, **420**, 1-6.
- Dawczynski,K., Kauf,E., & Zintl,F. (2003) Changes of serum growth factors (IGF-I,-II and IGFBP-2,-3) prior to and after stem cell transplantation in children with acute leukemia. *Bone Marrow Transplant.*, **32**, 411-415.
- De Botton S., Dombret,H., Sanz,M., Miguel,J.S., Caillot,D., Zittoun,R., Gardembas,M., Stamatoulas,A., Conde,E., Guerzi,A., Gardin,C., Geiser,K., Makhoul,D.C., Reman,O., de la,S.J., Lefrere,F., Chomienne,C., Chastang,C., Degos,L., & Fenaux,P. (1998) Incidence, clinical features, and outcome of all trans-retinoic acid syndrome in 413 cases of newly diagnosed acute promyelocytic leukemia. The European APL Group. *Blood*, **92**, 2712-2718.
- Debernardi,S., Lillington,D.M., Chaplin,T., Tomlinson,S., Amess,J., Rohatiner,A., Lister,T.A., & Young,B.D. (2003) Genome-wide analysis of acute myeloid leukemia with normal karyotype

reveals a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. *Genes Chromosomes.Cancer*, **37**, 149-158.

- Delaunay,J., Vey,N., Leblanc,T., Fenaux,P., Rigal-Huguet,F., Witz,F., Lamy,T., Auvrignon,A., Blaise,D., Pigneux,A., Mugneret,F., Bastard,C., Dastugue,N., Van den,A.J., Fiere,D., Reiffers,J., Castaigne,S., Leverger,G., Harousseau,J.L., & Dombret,H. (2003) Prognosis of inv(16)/t(16;16) acute myeloid leukemia (AML): a survey of 110 cases from the French AML Intergroup. *Blood*, **102**, 462-469.
- Deng,C., Zhang,P., Harper,J.W., Elledge,S.J., & Leder,P. (1995) Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell*, **82**, 675-684.
- Deschamps,J., van den,A.E., Forlani,S., De,G.W., Oosterveen,T., Roelen,B., & Roelfsema,J. (1999) Initiation, establishment and maintenance of Hox gene expression patterns in the mouse. *Int.J.Dev.Biol.*, **43**, 635-650.
- Deutsch,V.R. & Tomer,A. (2006) Megakaryocyte development and platelet production. *Br.J.Haematol.*, **134**, 453-466.
- Devaraj,P.E., Foroni,L., Prentice,G.H., Hoffbrand,V.A., & Secker-Walker,L.M. (1996) Relapse of acute promyelocytic leukemia follows serial negative RT-PCR assays: a cautionary tale. *Leuk.Res.*, **20**, 733-737.
- Dighiero,G. (2005) CLL biology and prognosis. *Hematology.Am.Soc.Hematol.Educ.Program.*, 278-284.
- Dighiero,G. & Binet,J.L. (2000) When and how to treat chronic lymphocytic leukemia. *N.Engl.J.Med.*, **343**, 1799-1801.
- Diverio,D., Pandolfi,P.P., Biondi,A., Avvisati,G., Petti,M.C., Mandelli,F., Pelicci,G., & Lo,C.F. (1993) Absence of reverse transcription-polymerase chain reaction detectable residual disease in patients with acute promyelocytic leukemia in long-term remission. *Blood*, **82**, 3556-3559.
- Diverio,D., Rossi,V., Avvisati,G., De,S.S., Pistilli,A., Pane,F., Saglio,G., Martinelli,G., Petti,M.C., Santoro,A., Pelicci,P.G., Mandelli,F., Biondi,A., & Lo,C.F. (1998) Early detection of relapse by prospective reverse transcriptase-polymerase chain reaction analysis of the PML/RARalpha fusion gene in patients with acute promyelocytic leukemia enrolled in the GIMEMA-AIEOP multicenter "AIDA" trial. GIMEMA-AIEOP Multicenter "AIDA" Trial. *Blood*, **92**, 784-789.

- Dohner,K., Schlenk,R.F., Habdank,M., Scholl,C., Rucker,F.G., Corbacioglu,A., Bullinger,L., Frohling,S., & Dohner,H. (2005) Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood*, **106**, 3740-3746.
- Downing,J.R., Head,D.R., Curcio-Brint,A.M., Hulshof,M.G., Motroni,T.A., Raimondi,S.C., Carroll,A.J., Drabkin,H.A., Willman,C., Theil,K.S., & . (1993) An AML1/ETO fusion transcript is consistently detected by RNA-based polymerase chain reaction in acute myelogenous leukemia containing the (8;21)(q22;q22) translocation. *Blood*, **81**, 2860-2865.
- Drabkin,H.A., Parsy,C., Ferguson,K., Guilhot,F., Lacotte,L., Roy,L., Zeng,C., Baron,A., Hunger,S.P., Varella-Garcia,M., Gemmill,R., Brizard,F., Brizard,A., & Roche,J. (2002) Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. *Leukemia*, **16**, 186-195.
- Drach,J., McQueen,T., Engel,H., Andreeff,M., Robertson,K.A., Collins,S.J., Malavasi,F., & Mehta,K. (1994) Retinoic acid-induced expression of CD38 antigen in myeloid cells is mediated through retinoic acid receptor-alpha. *Cancer Res.*, **54**, 1746-1752.
- Drach,J., Zhao,S., Malavasi,F., & Mehta,K. (1993) Rapid induction of CD38 antigen on myeloid leukemia cells by all trans-retinoic acid. *Biochem.Biophys.Res.Comm.*, **195**, 545-550.
- Drayson,M.T., Michell,R.H., Durham,J., & Brown,G. (2001) Cell proliferation and CD11b expression are controlled independently during HL60 cell differentiation initiated by 1,25 alpha-dihydroxyvitamin D(3) or all-trans-retinoic acid. *Exp.Cell Res.*, **266**, 126-134.
- Du,W., Hattori,Y., Yamada,T., Matsumoto,K., Nakamura,T., Sagawa,M., Otsuki,T., Niikura,T., Nukiwa,T., & Ikeda,Y. (2007) NK4, an antagonist of hepatocyte growth factor (HGF), inhibits growth of multiple myeloma cells: molecular targeting of angiogenic growth factor. *Blood*, **109**, 3042-3049.
- Dyck,J.A., Maul,G.G., Miller,W.H., Jr., Chen,J.D., Kakizuka,A., & Evans,R.M. (1994) A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. *Cell*, **76**, 333-343.
- Ebihara,Y., Tsuji,K., Lyman,S.D., Sui,X., Yoshida,M., Muraoka,K., Yamada,K., Tanaka,R., & Nakahata,T. (1997) Synergistic action of Flt3 and gp130 signalings in human hematopoiesis. *Blood*, **90**, 4363-4368.

- Elbashir,S.M., Lendeckel,W., & Tuschl,T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.*, **15**, 188-200.
- Elmaagacli,A.H., Beelen,D.W., TrenscheI,R., & Schaefer,U.W. (2000) The detection of wt-1 transcripts is not associated with an increased leukemic relapse rate in patients with acute leukemia after allogeneic bone marrow or peripheral blood stem cell transplantation. *Bone Marrow Transplant.*, **25**, 91-96.
- Erickson,P., Gao,J., Chang,K.S., Look,T., Whisenant,E., Raimondi,S., Lasher,R., Trujillo,J., Rowley,J., & Drabkin,H. (1992) Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood*, **80**, 1825-1831.
- Ernst,P., Fisher,J.K., Avery,W., Wade,S., Foy,D., & Korsmeyer,S.J. (2004) Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev.Cell*, **6**, 437-443.
- Estey,E. & Dohner,H. (2006) Acute myeloid leukaemia. *Lancet*, **368**, 1894-1907.
- Estey,E.H. (2009) Treatment of acute myeloid leukemia. *Haematologica*, **94**, 10-16.
- Estey,E.H., Keating,M.J., McCredie,K.B., Bodey,G.P., & Freireich,E.J. (1982) Causes of initial remission induction failure in acute myelogenous leukemia. *Blood*, **60**, 309-315.
- Evans,P.A., Short,M.A., Jack,A.S., Norfolk,D.R., Child,J.A., Shiach,C.R., Davies,F., Tobal,K., Liu Yin,J.A., & Morgan,G.J. (1997) Detection and quantitation of the CBFbeta/MYH11 transcripts associated with the inv(16) in presentation and follow-up samples from patients with AML. *Leukemia*, **11**, 364-369.
- Evarts,R.P., Hu,Z., Fujio,K., Marsden,E.R., & Thorgeirsson,S.S. (1993) Activation of hepatic stem cell compartment in the rat: role of transforming growth factor alpha, hepatocyte growth factor, and acidic fibroblast growth factor in early proliferation. *Cell Growth Differ.*, **4**, 555-561.
- Falini,B., Bolli,N., Liso,A., Martelli,M.P., Mannucci,R., Pileri,S., & Nicoletti,I. (2009) Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: molecular basis and clinical implications. *Leukemia*, **23**, 1731-1743.
- Falini,B. & Mason,D.Y. (2002) Proteins encoded by genes involved in chromosomal alterations in lymphoma and leukemia: clinical

value of their detection by immunocytochemistry. *Blood*, **99**, 409-426.

Falini,B., Mecucci,C., Tiacci,E., Alcalay,M., Rosati,R., Pasqualucci,L., La,S.R., Diverio,D., Colombo,E., Santucci,A., Bigerna,B., Pacini,R., Pucciarini,A., Liso,A., Vignetti,M., Fazi,P., Meani,N., Pettrossi,V., Saglio,G., Mandelli,F., Lo-Coco,F., Pelicci,P.G., & Martelli,M.F. (2005) Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N.Engl.J.Med.*, **352**, 254-266.

Falini,B., Nicoletti,I., Bolli,N., Martelli,M.P., Liso,A., Gorello,P., Mandelli,F., Mecucci,C., & Martelli,M.F. (2007a) Translocations and mutations involving the nucleophosmin (NPM1) gene in lymphomas and leukemias. *Haematologica*, **92**, 519-532.

Falini,B., Nicoletti,I., Martelli,M.F., & Mecucci,C. (2007b) Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. *Blood*, **109**, 874-885.

Fauser,A.A. & Messner,H.A. (1979) Identification of megakaryocytes, macrophages, and eosinophils in colonies of human bone marrow containing neutrophilic granulocytes and erythroblasts. *Blood*, **53**, 1023-1027.

Fauser,A.A. & Messner,H.A. (1978) Granuloerythropoietic colonies in human bone marrow, peripheral blood, and cord blood. *Blood*, **52**, 1243-1248.

Fenaux,P., Chomienne,C., & Degos,L. (2001) Treatment of acute promyelocytic leukaemia. *Best.Pract.Res.Clin.Haematol.*, **14**, 153-174.

Fenaux,P., Le Deley,M.C., Castaigne,S., Archimbaud,E., Chomienne,C., Link,H., Guerzi,A., Duarte,M., Daniel,M.T., Bowen,D., & . (1993) Effect of all transretinoic acid in newly diagnosed acute promyelocytic leukemia. Results of a multicenter randomized trial. European APL 91 Group. *Blood*, **82**, 3241-3249.

Fenaux,P., Wattel,E., Archimbaud,E., Sanz,M., Hecquet,B., Fegueux,N., Guerzi,A., Link,H., Fey,M., Castaigne,S., & . (1994) Prolonged follow-up confirms that all-trans retinoic acid followed by chemotherapy reduces the risk of relapse in newly diagnosed acute promyelocytic leukemia. The French APL Group. *Blood*, **84**, 666-667.

Fialkow,P.J., Gartler,S.M., & Yoshida,A. (1967) Clonal origin of chronic myelocytic leukemia in man. *Proc.Natl.Acad.Sci.U.S.A*, **58**, 1468-1471.

- Fialkow,P.J., Singer,J.W., Adamson,J.W., Vaidya,K., Dow,L.W., Ochs,J., & Moohr,J.W. (1981) Acute nonlymphocytic leukemia: heterogeneity of stem cell origin. *Blood*, **57**, 1068-1073.
- Ford,C.E., Hamerton,J.L., Barnes,D.W., & Loutit,J.F. (1956) Cytological identification of radiation-chimaeras. *Nature*, **177**, 452-454.
- Foulds,L. (1954) The experimental study of tumor progression: a review. *Cancer Res.*, **14**, 327-339.
- Fraizer,G.C., Patmasiriwat,P., Zhang,X., & Saunders,G.F. (1995) Expression of the tumor suppressor gene WT1 in both human and mouse bone marrow. *Blood*, **86**, 4704-4706.
- Frohling,S., Schlenk,R.F., Breitruck,J., Benner,A., Kreitmeier,S., Tobis,K., Dohner,H., & Dohner,K. (2002) Prognostic significance of activating FLT3 mutations in younger adults (16 to 60 years) with acute myeloid leukemia and normal cytogenetics: a study of the AML Study Group Ulm. *Blood*, **100**, 4372-4380.
- Frohling,S., Schlenk,R.F., Stolze,I., Bihlmayr,J., Benner,A., Kreitmeier,S., Tobis,K., Dohner,H., & Dohner,K. (2004) CEBPA mutations in younger adults with acute myeloid leukemia and normal cytogenetics: prognostic relevance and analysis of cooperating mutations. *J.Clin.Oncol.*, **22**, 624-633.
- Fuchs,O. (2007) Growth-inhibiting activity of transcription factor C/EBPalpha, its role in haematopoiesis and its tumour suppressor or oncogenic properties in leukaemias. *Folia Biol.(Praha)*, **53**, 97-108.
- Gabert,J., Beillard,E., van,d.V., V, Bi,W., Grimwade,D., Pallisgaard,N., Barbany,G., Cazzaniga,G., Cayuela,J.M., Cave,H., Pane,F., Aerts,J.L., De,M.D., Thirion,X., Pradel,V., Gonzalez,M., Viehmann,S., Malec,M., Saglio,G., & van Dongen,J.J. (2003) Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. *Leukemia*, **17**, 2318-2357.
- Gadhoum,Z., Delaunay,J., Maquarre,E., Durand,L., Lancereaux,V., Qi,J., Robert-Lezenes,J., Chomienne,C., & Smadja-Joffe,F. (2004) The effect of anti-CD44 monoclonal antibodies on differentiation and proliferation of human acute myeloid leukemia cells. *Leuk.Lymphoma*, **45**, 1501-1510.
- Gaiger,A., Schmid,D., Heinze,G., Linnerth,B., Greinix,H., Kalhs,P., Tisljar,K., Priglinger,S., Laczika,K., Mitterbauer,M., Novak,M., Mitterbauer,G., Mannhalter,C., Haas,O.A., Lechner,K., & Jager,U. (1998) Detection of the WT1 transcript by RT-PCR in

- complete remission has no prognostic relevance in de novo acute myeloid leukemia. *Leukemia*, **12**, 1886-1894.
- Gallagher,R.E., Yeap,B.Y., Bi,W., Livak,K.J., Beaubier,N., Rao,S., Bloomfield,C.D., Appelbaum,F.R., Tallman,M.S., Slack,J.L., & Willman,C.L. (2003) Quantitative real-time RT-PCR analysis of PML-RAR alpha mRNA in acute promyelocytic leukemia: assessment of prognostic significance in adult patients from intergroup protocol 0129. *Blood*, **101**, 2521-2528.
- Gameiro,P., Vieira,S., Carrara,P., Silva,A.L., Diamond,J., Botelho de,S.A., Mehta,A.B., Prentice,H.G., Guimaraes,J.E., Hoffbrand,A.V., Foroni,L., & Parreira,A. (2001) The PML-RAR alpha transcript in long-term follow-up of acute promyelocytic leukemia patients. *Haematologica*, **86**, 577-585.
- Gao,J., Erickson,P., Gardiner,K., Le Beau,M.M., Diaz,M.O., Patterson,D., Rowley,J.D., & Drabkin,H.A. (1991) Isolation of a yeast artificial chromosome spanning the 8;21 translocation breakpoint t(8;21)(q22;q22.3) in acute myelogenous leukemia. *Proc.Natl.Acad.Sci.U.S.A*, **88**, 4882-4886.
- Gelmetti,V., Zhang,J., Fanelli,M., Minucci,S., Pelicci,P.G., & Lazar,M.A. (1998) Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Mol.Cell Biol.*, **18**, 7185-7191.
- Ghaffari,S., Dougherty,G.J., Lansdorp,P.M., Eaves,A.C., & Eaves,C.J. (1995) Differentiation-associated changes in CD44 isoform expression during normal hematopoiesis and their alteration in chronic myeloid leukemia. *Blood*, **86**, 2976-2985.
- Giampaolo,A., Felli,N., Diverio,D., Morsilli,O., Samoggia,P., Breccia,M., Lo,C.F., Peschle,C., & Testa,U. (2002) Expression pattern of HOXB6 homeobox gene in myelomonocytic differentiation and acute myeloid leukemia. *Leukemia*, **16**, 1293-1301.
- Gilliland,D.G. & Griffin,J.D. (2002) The roles of FLT3 in hematopoiesis and leukemia. *Blood*, **100**, 1532-1542.
- Goldstone,A.H., Burnett,A.K., Wheatley,K., Smith,A.G., Hutchinson,R.M., & Clark,R.E. (2001) Attempts to improve treatment outcomes in acute myeloid leukemia (AML) in older patients: the results of the United Kingdom Medical Research Council AML11 trial. *Blood*, **98**, 1302-1311.
- Golub,T.R., Slonim,D.K., Tamayo,P., Huard,C., Gaasenbeek,M., Mesirov,J.P., Coller,H., Loh,M.L., Downing,J.R., Caligiuri,M.A., Bloomfield,C.D., & Lander,E.S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531-537.

- Gorello,P., Cazzaniga,G., Alberti,F., Dell'oro,M.G., Gottardi,E., Specchia,G., Roti,G., Rosati,R., Martelli,M.F., Diverio,D., Coco,F.L., Biondi,A., Saglio,G., Mecucci,C., & Falini,B. (2006) Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying nucleophosmin (NPM1) gene mutations. *Leukemia*.
- Goyama,S. & Kurokawa,M. (2009) Pathogenetic significance of ecotropic viral integration site-1 in hematological malignancies. *Cancer Sci.*, **100**, 990-995.
- Griffin,J.D., Davis,R., Nelson,D.A., Davey,F.R., Mayer,R.J., Schiffer,C., McIntyre,O.R., & Bloomfield,C.D. (1986) Use of surface marker analysis to predict outcome of adult acute myeloblastic leukemia. *Blood*, **68**, 1232-1241.
- Griffin,J.D. & Lowenberg,B. (1986) Clonogenic cells in acute myeloblastic leukemia. *Blood*, **68**, 1185-1195.
- Grignani,F., De,M.S., Nervi,C., Tomassoni,L., Gelmetti,V., Cioce,M., Fanelli,M., Ruthardt,M., Ferrara,F.F., Zamir,I., Seiser,C., Grignani,F., Lazar,M.A., Minucci,S., & Pelicci,P.G. (1998) Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. *Nature*, **391**, 815-818.
- Grignani,F., Valtieri,M., Gabbianelli,M., Gelmetti,V., Botta,R., Luchetti,L., Masella,B., Morsilli,O., Pelosi,E., Samoggia,P., Pelicci,P.G., & Peschle,C. (2000) PML/RAR alpha fusion protein expression in normal human hematopoietic progenitors dictates myeloid commitment and the promyelocytic phenotype. *Blood*, **96**, 1531-1537.
- Grimwade,D. (2002) The significance of minimal residual disease in patients with t(15;17). *Best.Pract.Res.Clin.Haematol.*, **15**, 137-158.
- Grimwade,D. & Lo Coco,F. (2002) Acute promyelocytic leukemia: a model for the role of molecular diagnosis and residual disease monitoring in directing treatment approach in acute myeloid leukemia. *Leukemia*, **16**, 1959-1973.
- Grimwade,D., Walker,H., Oliver,F., Wheatley,K., Harrison,C., Harrison,G., Rees,J., Hann,I., Stevens,R., Burnett,A., & Goldstone,A. (1998) The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood*, **92**, 2322-2333.
- Guan,Y., Gerhard,B., & Hogge,D.E. (2003) Detection, isolation, and stimulation of quiescent primitive leukemic progenitor cells from

- patients with acute myeloid leukemia (AML). *Blood*, **101**, 3142-3149.
- Guerrasio,A., Pilatrino,C., De,M.D., Cilloni,D., Serra,A., Gottardi,E., Parziale,A., Marmont,F., Diverio,D., Divona,M., Lo,C.F., & Saglio,G. (2002) Assessment of minimal residual disease (MRD) in CBFbeta/MYH11-positive acute myeloid leukemias by qualitative and quantitative RT-PCR amplification of fusion transcripts. *Leukemia*, **16**, 1176-1181.
- Guerrasio,A., Rosso,C., Martinelli,G., Lo,C.F., Pampinella,M., Santoro,A., Lanza,C., Allione,B., Resegotti,L., & Saglio,G. (1995) Polyclonal haemopoieses associated with long-term persistence of the AML1-ETO transcript in patients with FAB M2 acute myeloid leukaemia in continuous clinical remission. *Br.J.Haematol.*, **90**, 364-368.
- Guinn,B.A., Gilkes,A.F., Woodward,E., Westwood,N.B., Mufti,G.J., Linch,D., Burnett,A.K., & Mills,K.I. (2005) Microarray analysis of tumour antigen expression in presentation acute myeloid leukaemia. *Biochem.Biophys.Res.Comm.*, **333**, 703-713.
- Gutierrez,N.C., Lopez-Perez,R., Hernandez,J.M., Isidro,I., Gonzalez,B., Delgado,M., Ferminan,E., Garcia,J.L., Vazquez,L., Gonzalez,M., & San Miguel,J.F. (2005) Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. *Leukemia*, **19**, 402-409.
- Guttorp,P., Newton,M.A., & Abkowitz,J.L. (1990) A stochastic model for haematopoiesis in cats. *IMA J.Math.Appl.Med.Biol.*, **7**, 125-143.
- Gwang,K.J., Kyun,S.S., Hwan,K.D., Ho,B.J., Young,L.N., Soo,S.J., Chae,S.C., Soo,L.K., & Bo,L.K. (2005) Clinical implications of angiogenic factors in patients with acute or chronic leukemia: Hepatocyte growth factor levels have prognostic impact, especially in patients with acute myeloid leukemia. *Leuk.Lymphoma*, **46**, 885-891.
- Haase,D., Feuring-Buske,M., Konemann,S., Fonatsch,C., Troff,C., Verbeek,W., Pekrun,A., Hiddemann,W., & Wormann,B. (1995) Evidence for malignant transformation in acute myeloid leukemia at the level of early hematopoietic stem cells by cytogenetic analysis of CD34+ subpopulations. *Blood*, **86**, 2906-2912.
- Haber,D.A., Buckler,A.J., Glaser,T., Call,K.M., Pelletier,J., Sohn,R.L., Douglass,E.C., & Housman,D.E. (1990) An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor. *Cell*, **61**, 1257-1269.
- Hamalainen,M.M., Kairisto,V., Juvonen,V., Johansson,J., Auren,J., Kohonen,K., Remes,K., Salmi,T.T., Helenius,H., & Pelliniemi,T.T.

- (2008) Wilms tumour gene 1 overexpression in bone marrow as a marker for minimal residual disease in acute myeloid leukaemia. *Eur.J.Haematol.*, **80**, 201-207.
- Harper,D.P. & Aplan,P.D. (2008) Chromosomal rearrangements leading to MLL gene fusions: clinical and biological aspects. *Cancer Res.*, **68**, 10024-10027.
- Harper,J.W., Elledge,S.J., Keyomarsi,K., Dynlacht,B., Tsai,L.H., Zhang,P., Dobrowolski,S., Bai,C., Connell-Crowley,L., Swindell,E., & . (1995) Inhibition of cyclin-dependent kinases by p21. *Mol.Biol.Cell*, **6**, 387-400.
- Hayakawa,F., Towatari,M., Kiyoi,H., Tanimoto,M., Kitamura,T., Saito,H., & Naoe,T. (2000) Tandem-duplicated Flt3 constitutively activates STAT5 and MAP kinase and introduces autonomous cell growth in IL-3-dependent cell lines. *Oncogene*, **19**, 624-631.
- Heard,J.M., Fichelson,S., Sola,B., Martial,M.A., Varet,B., & Levy,J.P. (1984) Multistep virus-induced leukemogenesis in vitro: description of a model specifying three steps within the myeloblastic malignant process. *Mol.Cell Biol.*, **4**, 216-220.
- Hess,J.L. (2004) MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol.Med.*, **10**, 500-507.
- Hettmer,S., Dannecker,L., Foell,J., Elmlinger,M.W., & Dannecker,G.E. (2005) Effects of insulin-like growth factors and insulin-like growth factor binding protein-2 on the in vitro proliferation of peripheral blood mononuclear cells. *Hum.Immunol.*, **66**, 95-103.
- Heuser,M., Argiropoulos,B., Kuchenbauer,F., Yung,E., Piper,J., Fung,S., Schlenk,R.F., Dohner,K., Hinrichsen,T., Rudolph,C., Schambach,A., Baum,C., Schlegelberger,B., Dohner,H., Ganser,A., & Humphries,R.K. (2007) MN1 overexpression induces acute myeloid leukemia in mice and predicts ATRA resistance in patients with AML. *Blood*, **110**, 1639-1647.
- Heuser,M., Beutel,G., Krauter,J., Dohner,K., von,N.N., Schlegelberger,B., & Ganser,A. (2006) High meningeoma 1 (MN1) expression as a predictor for poor outcome in acute myeloid leukemia with normal cytogenetics. *Blood*, **108**, 3898-3905.
- Hiddemann,W., Kern,W., Schoch,C., Fonatsch,C., Heinecke,A., Wormann,B., & Buchner,T. (1999) Management of acute myeloid leukemia in elderly patients. *J.Clin.Oncol.*, **17**, 3569-3576.
- Hingorani,K., Szebeni,A., & Olson,M.O. (2000) Mapping the functional domains of nucleolar protein B23. *J.Biol.Chem.*, **275**, 24451-24457.

- Hino,M., Inaba,M., Goto,H., Nishizawa,Y., Tatsumi,N., Nishino,T., & Morii,H. (1996) Hepatocyte growth factor levels in bone marrow plasma of patients with leukaemia and its gene expression in leukaemic blast cells. *Br.J.Cancer*, **73**, 119-123.
- Huang,M.E., Ye,Y.C., Chen,S.R., Chai,J.R., Lu,J.X., Zhao,L., Gu,L.J., & Wang,Z.Y. (1988) Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood*, **72**, 567-572.
- Huang,W., Sun,G.L., Li,X.S., Cao,Q., Lu,Y., Jang,G.S., Zhang,F.Q., Chai,J.R., Wang,Z.Y., Waxman,S., & . (1993) Acute promyelocytic leukemia: clinical relevance of two major PML-RAR alpha isoforms and detection of minimal residual disease by retrotranscriptase/polymerase chain reaction to predict relapse. *Blood*, **82**, 1264-1269.
- Huntly,B.J., Shigematsu,H., Deguchi,K., Lee,B.H., Mizuno,S., Duclos,N., Rowan,R., Amaral,S., Curley,D., Williams,I.R., Akashi,K., & Gilliland,D.G. (2004) MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell*, **6**, 587-596.
- Ichikawa,H., Tanabe,K., Mizushima,H., Hayashi,Y., Mizutani,S., Ishii,E., Hongo,T., Kikuchi,A., & Satake,M. (2006) Common gene expression signatures in t(8;21)- and inv(16)-acute myeloid leukaemia. *Br.J.Haematol.*, **135**, 336-347.
- Ihle,J.N. & Kerr,I.M. (1995) Jaks and Stats in signaling by the cytokine receptor superfamily. *Trends Genet.*, **11**, 69-74.
- Ikeda,H., Lethe,B., Lehmann,F., van,B.N., Baurain,J.F., de,S.C., Chambost,H., Vitale,M., Moretta,A., Boon,T., & Coulie,P.G. (1997) Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity.*, **6**, 199-208.
- Inoue,K., Ogawa,H., Sonoda,Y., Kimura,T., Sakabe,H., Oka,Y., Miyake,S., Tamaki,H., Oji,Y., Yamagami,T., Tatekawa,T., Soma,T., Kishimoto,T., & Sugiyama,H. (1997) Aberrant overexpression of the Wilms tumor gene (WT1) in human leukemia. *Blood*, **89**, 1405-1412.
- Inoue,K., Ogawa,H., Yamagami,T., Soma,T., Tani,Y., Tatekawa,T., Oji,Y., Tamaki,H., Kyo,T., Dohy,H., Hiraoka,A., Masaoka,T., Kishimoto,T., & Sugiyama,H. (1996) Long-term follow-up of minimal residual disease in leukemia patients by monitoring WT1 (Wilms tumor gene) expression levels. *Blood*, **88**, 2267-2278.
- Inoue,K., Sugiyama,H., Ogawa,H., Nakagawa,M., Yamagami,T., Miwa,H., Kita,K., Hiraoka,A., Masaoka,T., Nasu,K., & . (1994) WT1 as a new prognostic factor and a new marker for the

- detection of minimal residual disease in acute leukemia. *Blood*, **84**, 3071-3079.
- Jabbour,E.J., Faderl,S., & Kantarjian,H.M. (2005) Adult acute lymphoblastic leukemia. *Mayo Clin.Proc.*, **80**, 1517-1527.
- James,C., Ugo,V., Casadevall,N., Constantinescu,S.N., & Vainchenker,W. (2005a) A JAK2 mutation in myeloproliferative disorders: pathogenesis and therapeutic and scientific prospects. *Trends Mol.Med.*, **11**, 546-554.
- James,C., Ugo,V., Le Couedic,J.P., Staerk,J., Delhommeau,F., Lacout,C., Garcon,L., Raslova,H., Berger,R., naceur-Griscelli,A., Villeval,J.L., Constantinescu,S.N., Casadevall,N., & Vainchenker,W. (2005b) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature*, **434**, 1144-1148.
- Janssen,J.W., Buschle,M., Layton,M., Drexler,H.G., Lyons,J., van den,B.H., Heimpel,H., Kubanek,B., Kleihauer,E., Mufti,G.J., & . (1989) Clonal analysis of myelodysplastic syndromes: evidence of multipotent stem cell origin. *Blood*, **73**, 248-254.
- Jensen,K., Shiels,C., & Freemont,P.S. (2001) PML protein isoforms and the RBCC/TRIM motif. *Oncogene*, **20**, 7223-7233.
- Johnson,P.F. (2005) Molecular stop signs: regulation of cell-cycle arrest by C/EBP transcription factors. *J.Cell Sci.*, **118**, 2545-2555.
- Johnson,W.E., Li,C., & Rabinovic,A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.*, **8**, 118-127.
- Jurcic,J.G., Nimer,S.D., Scheinberg,D.A., DeBlasio,T., Warrell,R.P., Jr., & Miller,W.H., Jr. (2001) Prognostic significance of minimal residual disease detection and PML/RAR-alpha isoform type: long-term follow-up in acute promyelocytic leukemia. *Blood*, **98**, 2651-2656.
- Juric,D., Sale,S., Hromas,R.A., Yu,R., Wang,Y., Duran,G.E., Tibshirani,R., Einhorn,L.H., & Sikic,B.I. (2005) Gene expression profiling differentiates germ cell tumors from other cancers and defines subtype-specific signatures. *Proc.Natl.Acad.Sci.U.S.A.*, **102**, 17763-17768.
- Jurlander,J., Caligiuri,M.A., Ruutu,T., Baer,M.R., Strout,M.P., Oberkircher,A.R., Hoffmann,L., Ball,E.D., Frei-Lahr,D.A., Christiansen,N.P., Block,A.M., Knuutila,S., Herzig,G.P., & Bloomfield,C.D. (1996) Persistence of the AML1/ETO fusion transcript in patients treated with allogeneic bone marrow transplantation for t(8;21) leukemia. *Blood*, **88**, 2183-2191.

- Kainz,B., Fonatsch,C., Schwarzingen,I., Sperr,W.R., Jager,U., & Gaiger,A. (2005) Limited value of FLT3 mRNA expression in the bone marrow for prognosis and monitoring of patients with acute myeloid leukemia. *Haematologica*, **90**, 695-696.
- Kaushansky,K., Lok,S., Holly,R.D., Broudy,V.C., Lin,N., Bailey,M.C., Forstrom,J.W., Buddle,M.M., Oort,P.J., Hagen,F.S., & . (1994) Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. *Nature*, **369**, 568-571.
- Kawagoe,H. & Grosveld,G.C. (2005a) Conditional MN1-TEL knock-in mice develop acute myeloid leukemia in conjunction with overexpression of HOXA9. *Blood*, **106**, 4269-4277.
- Kawagoe,H. & Grosveld,G.C. (2005b) MN1-TEL myeloid oncoprotein expressed in multipotent progenitors perturbs both myeloid and lymphoid growth and causes T-lymphoid tumors in mice. *Blood*, **106**, 4278-4286.
- Kawagoe,H., Humphries,R.K., Blair,A., Sutherland,H.J., & Hogge,D.E. (1999) Expression of HOX genes, HOX cofactors, and MLL in phenotypically and functionally defined subpopulations of leukemic and normal human hematopoietic cells. *Leukemia*, **13**, 687-698.
- Kay,H.E. (1965) How many cell-generations? *Lancet*, **2**, 418-419.
- Kern,W., Haferlach,C., Haferlach,T., & Schnittger,S. (2008) Monitoring of minimal residual disease in acute myeloid leukemia. *Cancer*, **112**, 4-16.
- Kern,W., Schoch,C., Haferlach,T., & Schnittger,S. (2005) Monitoring of minimal residual disease in acute myeloid leukemia. *Crit Rev.Oncol.Hematol.*, **56**, 283-309.
- Kiel,M.J., Yilmaz,O.H., Iwashita,T., Yilmaz,O.H., Terhorst,C., & Morrison,S.J. (2005) SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell*, **121**, 1109-1121.
- Kirstetter,P., Schuster,M.B., Bereshchenko,O., Moore,S., Dvinge,H., Kurz,E., Theilgaard-Monch,K., Mansson,R., Pedersen,T.A., Pabst,T., Schrock,E., Porse,B.T., Jacobsen,S.E., Bertone,P., Tenen,D.G., & Nerlov,C. (2008) Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. *Cancer Cell*, **13**, 299-310.

- Kitabayashi,I., Yokoyama,A., Shimizu,K., & Ohki,M. (1998) Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. *EMBO J.*, **17**, 2994-3004.
- Kiyoi,H., Naoe,T., Nakano,Y., Yokota,S., Minami,S., Miyawaki,S., Asou,N., Kuriyama,K., Jinnai,I., Shimazaki,C., Akiyama,H., Saito,K., Oh,H., Motoji,T., Omoto,E., Saito,H., Ohno,R., & Ueda,R. (1999) Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. *Blood*, **93**, 3074-3080.
- Kiyoi,H., Towatari,M., Yokota,S., Hamaguchi,M., Ohno,R., Saito,H., & Naoe,T. (1998) Internal tandem duplication of the FLT3 gene is a novel modality of elongation mutation which causes constitutive activation of the product. *Leukemia*, **12**, 1333-1337.
- Kiyoi,H., Yanada,M., & Ozekia,K. (2005) Clinical significance of FLT3 in leukemia. *Int.J.Hematol.*, **82**, 85-92.
- Knudson,A.G. (2001) Two genetic hits (more or less) to cancer. *Nat.Rev.Cancer*, **1**, 157-162.
- Kondo,M., Weissman,I.L., & Akashi,K. (1997) Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*, **91**, 661-672.
- Kottaridis,P.D., Gale,R.E., Frew,M.E., Harrison,G., Langabeer,S.E., Belton,A.A., Walker,H., Wheatley,K., Bowen,D.T., Burnett,A.K., Goldstone,A.H., & Linch,D.C. (2001) The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. *Blood*, **98**, 1752-1759.
- Kottaridis,P.D., Gale,R.E., Langabeer,S.E., Frew,M.E., Bowen,D.T., & Linch,D.C. (2002) Studies of FLT3 mutations in paired presentation and relapse samples from patients with acute myeloid leukemia: implications for the role of FLT3 mutations in leukemogenesis, minimal residual disease detection, and possible therapy with FLT3 inhibitors. *Blood*, **100**, 2393-2398.
- Krauter,J., Gorlich,K., Ottmann,O., Lubbert,M., Dohner,H., Heit,W., Kanz,L., Ganser,A., & Heil,G. (2003) Prognostic value of minimal residual disease quantification by real-time reverse transcriptase polymerase chain reaction in patients with core binding factor leukemias. *J.Clin.Oncol.*, **21**, 4413-4422.
- Krauter,J., Hoellge,W., Wattjes,M.P., Nagel,S., Heidenreich,O., Bunjes,D., Ganser,A., & Heil,G. (2001) Detection and

- quantification of CBFB/MYH11 fusion transcripts in patients with inv(16)-positive acute myeloblastic leukemia by real-time RT-PCR. *Genes Chromosomes.Cancer*, **30**, 342-348.
- Kroon,E., Thorsteinsdottir,U., Mayotte,N., Nakamura,T., & Sauvageau,G. (2001) NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. *EMBO J.*, **20**, 350-361.
- Krumlauf,R. (1994) Hox genes in vertebrate development. *Cell*, **78**, 191-201.
- Kuba,K., Matsumoto,K., Date,K., Shimura,H., Tanaka,M., & Nakamura,T. (2000) HGF/NK4, a four-kringle antagonist of hepatocyte growth factor, is an angiogenesis inhibitor that suppresses tumor growth and metastasis in mice. *Cancer Res.*, **60**, 6737-6743.
- Kundu,M., Chen,A., Anderson,S., Kirby,M., Xu,L., Castilla,L.H., Bodine,D., & Liu,P.P. (2002) Role of Cbfb in hematopoiesis and perturbations resulting from expression of the leukemogenic fusion gene Cbfb-MYH11. *Blood*, **100**, 2449-2456.
- Kwong,Y.L., Wong,K.F., Chan,V., & Chan,C.H. (1996) Persistence of AML1 rearrangement in peripheral blood cells in t(8;21). *Cancer Genet.Cytogenet.*, **88**, 151-154.
- La Motte-Mohs,R.N., Herer,E., & Zuniga-Pflucker,J.C. (2005) Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. *Blood*, **105**, 1431-1439.
- Lacayo,N.J., Meshinchi,S., Kinnunen,P., Yu,R., Wang,Y., Stuber,C.M., Douglas,L., Wahab,R., Becton,D.L., Weinstein,H., Chang,M.N., Willman,C.L., Radich,J.P., Tibshirani,R., Ravindranath,Y., Sikic,B.I., & Dahl,G.V. (2004) Gene expression profiles at diagnosis in de novo childhood AML patients identify FLT3 mutations with good clinical outcomes. *Blood*, **104**, 2646-2654.
- Lane,S., Saal,R., Mollee,P., Jones,M., Grigg,A., Taylor,K., Seymour,J., Kennedy,G., Williams,B., Grimmett,K., Griffiths,V., Gill,D., Hourigan,M., & Marlton,P. (2008) A ≥ 1 log rise in RQ-PCR transcript levels defines molecular relapse in core binding factor acute myeloid leukemia and predicts subsequent morphologic relapse. *Leuk.Lymphoma*, **49**, 517-523.
- Langer,C., Radmacher,M.D., Ruppert,A.S., Whitman,S.P., Paschka,P., Mrozek,K., Baldus,C.D., Vukosavljevic,T., Liu,C.G., Ross,M.E., Powell,B.L., de la,C.A., Kolitz,J.E., Larson,R.A., Marcucci,G., & Bloomfield,C.D. (2008) High BAALC expression associates with other molecular prognostic markers, poor outcome, and a distinct gene-expression signature in cytogenetically normal patients

- younger than 60 years with acute myeloid leukemia: a Cancer and Leukemia Group B (CALGB) study. *Blood*, **111**, 5371-5379.
- Lapidot,T., Sirard,C., Vormoor,J., Murdoch,B., Hoang,T., Caceres-Cortes,J., Minden,M., Paterson,B., Caligiuri,M.A., & Dick,J.E. (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, **367**, 645-648.
- Larosa,D.F. & Orange,J.S. (2008) 1. Lymphocytes. *J.Allergy Clin.Immunol.*, **121**, S364-S369.
- Larson,R.A. (2007) Is secondary leukemia an independent poor prognostic factor in acute myeloid leukemia? *Best.Pract.Res.Clin.Haematol.*, **20**, 29-37.
- Larson,R.A., Williams,S.F., Le Beau,M.M., Bitter,M.A., Vardiman,J.W., & Rowley,J.D. (1986) Acute myelomonocytic leukemia with abnormal eosinophils and inv(16) or t(16;16) has a favorable prognosis. *Blood*, **68**, 1242-1249.
- Latchman,D.S. (1998) The Brn-3a transcription factor. *Int.J.Biochem.Cell Biol.*, **30**, 1153-1157.
- Lawrence,H.J., Rozenfeld,S., Cruz,C., Matsukuma,K., Kwong,A., Komuves,L., Buchberg,A.M., & Largman,C. (1999) Frequent co-expression of the HOXA9 and MEIS1 homeobox genes in human myeloid leukemias. *Leukemia*, **13**, 1993-1999.
- Lawrence,H.J., Sauvageau,G., Humphries,R.K., & Largman,C. (1996) The role of HOX homeobox genes in normal and leukemic hematopoiesis. *Stem Cells*, **14**, 281-291.
- Lee,M.H. & Yang,H.Y. (2001) Negative regulators of cyclin-dependent kinases and their roles in cancers. *Cell Mol.Life Sci.*, **58**, 1907-1922.
- Lee,S., Kim,Y.J., Eom,K.S., Min,C.K., Kim,H.J., Cho,S.G., Lee,J.W., Min,W.S., & Kim,C.C. (2006) The significance of minimal residual disease kinetics in adults with newly diagnosed PML-RARalpha-positive acute promyelocytic leukemia: results of a prospective trial. *Haematologica*, **91**, 671-674.
- Legras,S., Gunthert,U., Stauder,R., Curt,F., Oliferenko,S., Kluin-Nelemans,H.C., Marie,J.P., Proctor,S., Jasmin,C., & Smadja-Joffe,F. (1998) A strong expression of CD44-6v correlates with shorter survival of patients with acute myeloid leukemia. *Blood*, **91**, 3401-3413.
- Lekanne Deprez,R.H., Groen,N.A., van Biezen,N.A., Hagemmeijer,A., van,D.E., Koper,J.W., Avezaat,C.J., Bootsma,D., & Zwarthoff,E.C. (1991) A t(4;22) in a meningioma points to the

- localization of a putative tumor-suppressor gene. *Am.J.Hum.Genet.*, **48**, 783-790.
- Lemmink,H.H., Schroder,C.H., Monnens,L.A., & Smeets,H.J. (1997) The clinical spectrum of type IV collagen mutations. *Hum.Mutat.*, **9**, 477-499.
- Leone,G., Mele,L., Pulsoni,A., Equitani,F., & Pagano,L. (1999) The incidence of secondary leukemias. *Haematologica*, **84**, 937-945.
- Leroy,H., De,B.S., Grardel-Duflos,N., Darre,S., Leleu,X., Roumier,C., Morschhauser,F., Lai,J.L., Bauters,F., Fenaux,P., & Preudhomme,C. (2005) Prognostic value of real-time quantitative PCR (RQ-PCR) in AML with t(8;21). *Leukemia*, **19**, 367-372.
- Levine,R.L., Wadleigh,M., Cools,J., Ebert,B.L., Wernig,G., Huntly,B.J., Boggon,T.J., Wlodarska,I., Clark,J.J., Moore,S., Adelsperger,J., Koo,S., Lee,J.C., Gabriel,S., Mercher,T., D'Andrea,A., Frohling,S., Dohner,K., Marynen,P., Vandenberghe,P., Mesa,R.A., Tefferi,A., Griffin,J.D., Eck,M.J., Sellers,W.R., Meyerson,M., Golub,T.R., Lee,S.J., & Gilliland,D.G. (2005) Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell*, **7**, 387-397.
- Levis,M., Allebach,J., Tse,K.F., Zheng,R., Baldwin,B.R., Smith,B.D., Jones-Bolin,S., Ruggeri,B., Dionne,C., & Small,D. (2002) A FLT3-targeted tyrosine kinase inhibitor is cytotoxic to leukemia cells in vitro and in vivo. *Blood*, **99**, 3885-3891.
- Liu Yin,J.A. (2002) Minimal residual disease in acute myeloid leukaemia. *Best Pract.Res.Clin.Haematol.*, **15**, 119-135.
- Liu,P., Tarle,S.A., Hajra,A., Claxton,D.F., Marlton,P., Freedman,M., Siciliano,M.J., & Collins,F.S. (1993) Fusion between transcription factor CBF beta/PEBP2 beta and a myosin heavy chain in acute myeloid leukemia. *Science*, **261**, 1041-1044.
- Lo Coco,F., Diverio,D., Pandolfi,P.P., Biondi,A., Rossi,V., Avvisati,G., Rambaldi,A., Arcese,W., Petti,M.C., Meloni,G., & . (1992) Molecular evaluation of residual disease as a predictor of relapse in acute promyelocytic leukaemia. *Lancet*, **340**, 1437-1438.
- Lo Cocco,F., Diverio,D., Falini,B., Biondi,A., Nervi,C., & Pelicci,P.G. (1999) Genetic diagnosis and molecular monitoring in the management of acute promyelocytic leukemia. *Blood*, **94**, 12-22.
- Lo Coco,F. & Ammatuna,E. (2006) The biology of acute promyelocytic leukemia and its impact on diagnosis and treatment. *Hematology.Am.Soc.Hematol.Educ.Program.*, 156-61, 514.

- Lok,S., Kaushansky,K., Holly,R.D., Kuijper,J.L., Lofton-Day,C.E., Oort,P.J., Grant,F.J., Heipel,M.D., Burkhead,S.K., Kramer,J.M., & . (1994) Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet production in vivo. *Nature*, **369**, 565-568.
- Lutterbach,B., Hou,Y., Durst,K.L., & Hiebert,S.W. (1999) The inv(16) encodes an acute myeloid leukemia 1 transcriptional corepressor. *Proc.Natl.Acad.Sci.U.S.A*, **96**, 12822-12827.
- Lutterbach,B., Westendorf,J.J., Linggi,B., Patten,A., Moniwa,M., Davie,J.R., Huynh,K.D., Bardwell,V.J., Lavinsky,R.M., Rosenfeld,M.G., Glass,C., Seto,E., & Hiebert,S.W. (1998) ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Mol.Cell Biol.*, **18**, 7176-7184.
- Lyman,S.D. & Jacobsen,S.E. (1998) c-kit ligand and Flt3 ligand: stem/progenitor cell factors with overlapping yet distinct activities. *Blood*, **91**, 1101-1134.
- Mackay,C.R., Terpe,H.J., Stauder,R., Marston,W.L., Stark,H., & Gunthert,U. (1994) Expression and modulation of CD44 variant isoforms in humans. *J.Cell Biol.*, **124**, 71-82.
- Magli,M.C., Barba,P., Celetti,A., De,V.G., Cillo,C., & Boncinelli,E. (1991) Coordinate regulation of HOX genes in human hematopoietic cells. *Proc.Natl.Acad.Sci.U.S.A*, **88**, 6348-6352.
- Magli,M.C., Largman,C., & Lawrence,H.J. (1997) Effects of HOX homeobox genes in blood cell differentiation. *J.Cell Physiol*, **173**, 168-177.
- Malcovati,L. & Nimer,S.D. (2008) Myelodysplastic syndromes: diagnosis and staging. *Cancer Control*, **15 Suppl**, 4-13.
- Mandelli,F., Diverio,D., Avvisati,G., Luciano,A., Barbui,T., Bernasconi,C., Broccia,G., Cerri,R., Falda,M., Fioritoni,G., Leoni,F., Liso,V., Petti,M.C., Rodeghiero,F., Saglio,G., Vegna,M.L., Visani,G., Jehn,U., Willemze,R., Muus,P., Pelicci,P.G., Biondi,A., & Lo,C.F. (1997) Molecular remission in PML/RAR alpha-positive acute promyelocytic leukemia by combined all-trans retinoic acid and idarubicin (AIDA) therapy. Gruppo Italiano-Malattie Ematologiche Maligne dell'Adulto and Associazione Italiana di Ematologia ed Oncologia Pediatrica Cooperative Groups. *Blood*, **90**, 1014-1021.
- Marcucci,G., Caligiuri,M.A., Dohner,H., Archer,K.J., Schlenk,R.F., Dohner,K., Maghraby,E.A., & Bloomfield,C.D. (2001) Quantification of CBFbeta/MYH11 fusion transcript by real time RT-PCR in patients with INV(16) acute myeloid leukemia. *Leukemia*, **15**, 1072-1080.

- Marcucci,G., Mrozek,K., & Bloomfield,C.D. (2005) Molecular heterogeneity and prognostic biomarkers in adults with acute myeloid leukemia and normal cytogenetics. *Curr.Opin.Hematol.*, **12**, 68-75.
- Martin,F.H., Suggs,S.V., Langley,K.E., Lu,H.S., Ting,J., Okino,K.H., Morris,C.F., McNiece,I.K., Jacobsen,F.W., Mendiaz,E.A., & . (1990) Primary structure and functional expression of rat and human stem cell factor DNAs. *Cell*, **63**, 203-211.
- Martin-Villar,E., Megias,D., Castel,S., Yurrita,M.M., Vilaro,S., & Quintanilla,M. (2006) Podoplanin binds ERM proteins to activate RhoA and promote epithelial-mesenchymal transition. *J.Cell Sci.*, **119**, 4541-4553.
- Martin-Villar,E., Scholl,F.G., Gamallo,C., Yurrita,M.M., Munoz-Guerra,M., Cruces,J., & Quintanilla,M. (2005) Characterization of human PA2.26 antigen (T1alpha-2, podoplanin), a small membrane mucin induced in oral squamous cell carcinomas. *Int.J.Cancer*, **113**, 899-910.
- Martineau,M., Berger,R., Lillington,D.M., Moorman,A.V., & Secker-Walker,L.M. (1998) The t(6;11)(q27;q23) translocation in acute leukemia: a laboratory and clinical study of 30 cases. EU Concerted Action 11q23 Workshop participants. *Leukemia*, **12**, 788-791.
- Mason,K.D., Juneja,S.K., & Szer,J. (2006) The immunophenotype of acute myeloid leukemia: is there a relationship with prognosis? *Blood Rev.*, **20**, 71-82.
- Matsushita,M., Ikeda,H., Kizaki,M., Okamoto,S., Ogasawara,M., Ikeda,Y., & Kawakami,Y. (2001) Quantitative monitoring of the PRAME gene for the detection of minimal residual disease in leukaemia. *Br.J.Haematol.*, **112**, 916-926.
- McCulloch,E.A. (1983) Stem cells in normal and leukemic hemopoiesis (Henry Stratton Lecture, 1982). *Blood*, **62**, 1-13.
- Mead,A.J., Gale,R.E., Hills,R.K., Gupta,M., Young,B.D., Burnett,A.K., & Linch,D.C. (2008) Conflicting data on the prognostic significance of FLT3/TKD mutations in acute myeloid leukemia might be related to the incidence of biallelic disease. *Blood*, **112**, 444-445.
- Mehrotra,B., George,T.I., Kavanau,K., vet-Loiseau,H., Moore,D., Willman,C.L., Slovak,M.L., Atwater,S., Head,D.R., & Pallavicini,M.G. (1995) Cytogenetically aberrant cells in the stem cell compartment (CD34+lin-) in acute myeloid leukemia. *Blood*, **86**, 1139-1147.

- Melnick,A. & Licht,J.D. (1999) Deconstructing a disease: RARalpha, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia. *Blood*, **93**, 3167-3215.
- Meshinchi,S. & Appelbaum,F.R. (2009) Structural and functional alterations of FLT3 in acute myeloid leukemia. *Clin.Cancer Res.*, **15**, 4263-4269.
- Meyers,S., Lenny,N., & Hiebert,S.W. (1995) The t(8;21) fusion protein interferes with AML-1B-dependent transcriptional activation. *Mol.Cell Biol.*, **15**, 1974-1982.
- Mijovic,A. & Mufti,G.J. (1998) The myelodysplastic syndromes: towards a functional classification. *Blood Rev.*, **12**, 73-83.
- Miller,W.H., Jr., Kakizuka,A., Frankel,S.R., Warrell,R.P., Jr., DeBlasio,A., Levine,K., Evans,R.M., & Dmitrovsky,E. (1992) Reverse transcription polymerase chain reaction for the rearranged retinoic acid receptor alpha clarifies diagnosis and detects minimal residual disease in acute promyelocytic leukemia. *Proc.Natl.Acad.Sci.U.S.A*, **89**, 2694-2698.
- Minden,M.D., Till,J.E., & McCulloch,E.A. (1978) Proliferative state of blast cell progenitors in acute myeloblastic leukemia (AML). *Blood*, **52**, 592-600.
- Misaghian,N., Ligresti,G., Steelman,L.S., Bertrand,F.E., Basecke,J., Libra,M., Nicoletti,F., Stivala,F., Milella,M., Tafuri,A., Cervello,M., Martelli,A.M., & McCubrey,J.A. (2009) Targeting the leukemic stem cell: the Holy Grail of leukemia therapy. *Leukemia*, **23**, 25-42.
- Mistry,A.R., Pedersen,E.W., Solomon,E., & Grimwade,D. (2003) The molecular pathogenesis of acute promyelocytic leukaemia: implications for the clinical management of the disease. *Blood Rev.*, **17**, 71-97.
- Miyamoto,T., Nagafuji,K., Akashi,K., Harada,M., Kyo,T., Akashi,T., Takenaka,K., Mizuno,S., Gondo,H., Okamura,T., Dohy,H., & Niho,Y. (1996) Persistence of multipotent progenitors expressing AML1/ETO transcripts in long-term remission patients with t(8;21) acute myelogenous leukemia. *Blood*, **87**, 4789-4796.
- Miyoshi,H., Shimizu,K., Kozu,T., Maseki,N., Kaneko,Y., & Ohki,M. (1991) t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proc.Natl.Acad.Sci.U.S.A*, **88**, 10431-10434.
- Mizuki,M., Fenski,R., Halfter,H., Matsumura,I., Schmidt,R., Muller,C., Gruning,W., Kratz-Albers,K., Serve,S., Steur,C., Buchner,T., Kienast,J., Kanakura,Y., Berdel,W.E., & Serve,H. (2000) Flt3

- mutations from patients with acute myeloid leukemia induce transformation of 32D cells mediated by the Ras and STAT5 pathways. *Blood*, **96**, 3907-3914.
- Montesano,R., Matsumoto,K., Nakamura,T., & Orci,L. (1991) Identification of a fibroblast-derived epithelial morphogen as hepatocyte growth factor. *Cell*, **67**, 901-908.
- Montserrat,E. & Moreno,C. (2008) Chronic lymphocytic leukaemia: a short overview. *Ann.Oncol.*, **19 Suppl 7**, vii320-vii325.
- Moore,M.A., Chung,K.Y., Plasilova,M., Schuringa,J.J., Shieh,J.H., Zhou,P., & Morrone,G. (2007) NUP98 dysregulation in myeloid leukemogenesis. *Ann.N.Y.Acad.Sci.*, **1106**, 114-142.
- Moreno,I., Martin,G., Bolufer,P., Barragan,E., Rueda,E., Roman,J., Fernandez,P., Leon,P., Mena,A., Cervera,J., Torres,A., & Sanz,M.A. (2003) Incidence and prognostic value of FLT3 internal tandem duplication and D835 mutations in acute myeloid leukemia. *Haematologica*, **88**, 19-24.
- Moretti,P., Simmons,P., Thomas,P., Haylock,D., Rathjen,P., Vadas,M., & D'Andrea,R. (1994) Identification of homeobox genes expressed in human haemopoietic progenitor cells. *Gene*, **144**, 213-219.
- Morris,S.W., Kirstein,M.N., Valentine,M.B., Dittmer,K.G., Shapiro,D.N., Saltman,D.L., & Look,A.T. (1994) Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science*, **263**, 1281-1284.
- Mozziconacci,M.J., Liberatore,C., Brunel,V., Grignani,F., Arnoulet,C., Ferrucci,P.F., Fernandez,F., Sainty,D., Pelicci,P.G., Birg,F., & Lafage-Pochitaloff,M. (1998) In vitro response to all-trans retinoic acid of acute promyelocytic leukemias with nonreciprocal PML/RARA or RARA/PML fusion genes. *Genes Chromosomes.Cancer*, **22**, 241-250.
- Mrozek,K. & Bloomfield,C.D. (2006) Chromosome aberrations, gene mutations and expression changes, and prognosis in adult acute myeloid leukemia. *Hematology.Am.Soc.Hematol.Educ.Program.*, 169-177.
- Mrozek,K., Dohner,H., & Bloomfield,C.D. (2007a) Influence of new molecular prognostic markers in patients with karyotypically normal acute myeloid leukemia: recent advances. *Curr.Opin.Hematol.*, **14**, 106-114.
- Mrozek,K., Heerema,N.A., & Bloomfield,C.D. (2004) Cytogenetics in acute leukemia. *Blood Rev.*, **18**, 115-136.

- Mrozek,K., Marcucci,G., Paschka,P., Whitman,S.P., & Bloomfield,C.D. (2007b) Clinical relevance of mutations and gene-expression changes in adult acute myeloid leukemia with normal cytogenetics: are we ready for a prognostically prioritized molecular classification? *Blood*, **109**, 431-448.
- Mu,Z.M., Chin,K.V., Liu,J.H., Lozano,G., & Chang,K.S. (1994) PML, a growth suppressor disrupted in acute promyelocytic leukemia. *Mol.Cell Biol.*, **14**, 6858-6867.
- Mueller,B.U. & Pabst,T. (2006) C/EBPalpha and the pathophysiology of acute myeloid leukemia. *Curr.Opin.Hematol.*, **13**, 7-14.
- Mufti,G.J. (1992) Chromosomal deletions in the myelodysplastic syndrome. *Leuk.Res.*, **16**, 35-41.
- Mulford,D. (2008) Antibody therapy for acute myeloid leukemia. *Semin.Hematol.*, **45**, 104-109.
- Mutch,D.M., Berger,A., Mansourian,R., Rytz,A., & Roberts,M.A. (2002) The limit fold change model: a practical approach for selecting differentially expressed genes from microarray data. *BMC.Bioinformatics.*, **3**, 17.
- Nakamura,S., Gohda,E., Matsuo,Y., Yamamoto,I., & Minowada,J. (1994) Significant amount of hepatocyte growth factor detected in blood and bone marrow plasma of leukaemia patients. *Br.J.Haematol.*, **87**, 640-642.
- Nakano,Y., Kiyoi,H., Miyawaki,S., Asou,N., Ohno,R., Saito,H., & Naoe,T. (1999) Molecular evolution of acute myeloid leukaemia in relapse: unstable N-ras and FLT3 genes compared with p53 gene. *Br.J.Haematol.*, **104**, 659-664.
- Neben,K., Schnittger,S., Brors,B., Tews,B., Kokocinski,F., Haferlach,T., Muller,J., Hahn,M., Hiddemann,W., Eils,R., Lichter,P., & Schoch,C. (2005) Distinct gene expression patterns associated with. *Oncogene*, **24**, 1580-1588.
- Nishii,K., Kita,K., Nadim,M., Miwa,H., Ohoishi,K., Yamaguchi,M., & Shirakawa,S. (1995) Expression of interleukin-5 receptors on acute myeloid leukaemia cells: association with immunophenotype and karyotype. *Br.J.Haematol.*, **91**, 169-172.
- Nishino,T., Hisha,H., Nishino,N., Adachi,M., & Ikehara,S. (1995) Hepatocyte growth factor as a hematopoietic regulator. *Blood*, **85**, 3093-3100.
- Nucifora,G. (1997) The EVI1 gene in myeloid leukemia. *Leukemia*, **11**, 2022-2031.

- Nucifora,G., Larson,R.A., & Rowley,J.D. (1993) Persistence of the 8;21 translocation in patients with acute myeloid leukemia type M2 in long-term remission. *Blood*, **82**, 712-715.
- Ogawa,H., Tamaki,H., Ikegame,K., Soma,T., Kawakami,M., Tsuboi,A., Kim,E.H., Hosen,N., Murakami,M., Fujioka,T., Masuda,T., Taniguchi,Y., Nishida,S., Oji,Y., Oka,Y., & Sugiyama,H. (2003) The usefulness of monitoring WT1 gene transcripts for the prediction and management of relapse following allogeneic stem cell transplantation in acute type leukemia. *Blood*, **101**, 1698-1704.
- Okuda,T., van,D.J., Hiebert,S.W., Grosveld,G., & Downing,J.R. (1996) AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, **84**, 321-330.
- Ommen,H.B., Nyvold,C.G., Braendstrup,K., Andersen,B.L., Ommen,I.B., Hasle,H., Hokland,P., & Ostergaard,M. (2008) Relapse prediction in acute myeloid leukaemia patients in complete remission using WT1 as a molecular marker: development of a mathematical model to predict time from molecular to clinical relapse and define optimal sampling intervals. *Br.J.Haematol.*, **141**, 782-791.
- Oohashi,T., Sugimoto,M., Mattei,M.G., & Ninomiya,Y. (1994) Identification of a new collagen IV chain, alpha 6(IV), by cDNA isolation and assignment of the gene to chromosome Xq22, which is the same locus for COL4A5. *J.Biol.Chem.*, **269**, 7520-7526.
- Orange,J.S. & Ballas,Z.K. (2006) Natural killer cells in human health and disease. *Clin.Immunol.*, **118**, 1-10.
- Ostergaard,M., Olesen,L.H., Hasle,H., Kjeldsen,E., & Hokland,P. (2004) WT1 gene expression: an excellent tool for monitoring minimal residual disease in 70% of acute myeloid leukaemia patients - results from a single-centre study. *Br.J.Haematol.*, **125**, 590-600.
- Ozeki,M. & Shively,J.E. (2008) Differential cell fates induced by all-trans retinoic acid-treated HL-60 human leukemia cells. *J.Leukoc.Biol.*
- Pabst,T., Mueller,B.U., Zhang,P., Radomska,H.S., Narravula,S., Schnittger,S., Behre,G., Hiddemann,W., & Tenen,D.G. (2001) Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat.Genet.*, **27**, 263-270.
- Pandolfi,P.P. (2001) In vivo analysis of the molecular genetics of acute promyelocytic leukemia. *Oncogene*, **20**, 5726-5735.

- Pandolfi, P.P., Alcalay, M., Fagioli, M., Zangrilli, D., Mencarelli, A., Diverio, D., Biondi, A., Lo, C.F., Rambaldi, A., Grignani, F., & . (1992) Genomic variability and alternative splicing generate multiple PML/RAR alpha transcripts that encode aberrant PML proteins and PML/RAR alpha isoforms in acute promyelocytic leukaemia. *EMBO J.*, **11**, 1397-1407.
- Papadaki, C., Dufour, A., Seibl, M., Schneider, S., Bohlander, S.K., Zellmeier, E., Mellert, G., Hiddemann, W., & Spiekermann, K. (2009) Monitoring minimal residual disease in acute myeloid leukaemia with NPM1 mutations by quantitative PCR: clonal evolution is a limiting factor. *Br.J.Haematol.*, **144**, 517-523.
- Park, M., Dean, M., Cooper, C.S., Schmidt, M., O'Brien, S.J., Blair, D.G., & Vande Woude, G.F. (1986) Mechanism of met oncogene activation. *Cell*, **45**, 895-904.
- Paydas, S., Tanriverdi, K., Yavuz, S., Disel, U., Baslamisli, F., & Burgut, R. (2005) PRAME mRNA levels in cases with acute leukemia: clinical importance and future prospects. *Am.J.Hematol.*, **79**, 257-261.
- Perea, G., Lasa, A., Aventin, A., Domingo, A., Villamor, N., Queipo de Llano, M.P., Llorente, A., Junca, J., Palacios, C., Fernandez, C., Gallart, M., Font, L., Tormo, M., Florensa, L., Bargay, J., Marti, J.M., Vivancos, P., Torres, P., Berlanga, J.J., Badell, I., Brunet, S., Sierra, J., & Nomdedeu, J.F. (2006) Prognostic value of minimal residual disease (MRD) in acute myeloid leukemia (AML) with favorable cytogenetics [t(8;21) and inv(16)]. *Leukemia*, **20**, 87-94.
- Perez, A., Kastner, P., Sethi, S., Lutz, Y., Reibel, C., & Chambon, P. (1993) PMLRAR homodimers: distinct DNA binding properties and heteromeric interactions with RXR. *EMBO J.*, **12**, 3171-3182.
- Perry, C., Eldor, A., & Soreq, H. (2002) Runx1/AML1 in leukemia: disrupted association with diverse protein partners. *Leuk.Res.*, **26**, 221-228.
- Pessano, S., Palumbo, A., Ferrero, D., Pagliardi, G.L., Bottero, L., Lai, S.K., Meo, P., Carter, C., Hubbell, H., Lange, B., & . (1984) Subpopulation heterogeneity in human acute myeloid leukemia determined by monoclonal antibodies. *Blood*, **64**, 275-281.
- Peterson, L.F. & Zhang, D.E. (2004) The 8;21 translocation in leukemogenesis. *Oncogene*, **23**, 4255-4262.
- Pevny, L., Simon, M.C., Robertson, E., Klein, W.H., Tsai, S.F., D'Agati, V., Orkin, S.H., & Costantini, F. (1991) Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, **349**, 257-260.

- Pierce,A., Whetton,A.D., Owen-Lynch,P.J., Tavernier,J., Spooncer,E., Dexter,T.M., & Heyworth,C.M. (1998) Ectopic interleukin-5 receptor expression promotes proliferation without development in a multipotent hematopoietic cell line. *J.Cell Sci.*, **111** (Pt 6), 815-823.
- Pineault,N., Helgason,C.D., Lawrence,H.J., & Humphries,R.K. (2002) Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Exp.Hematol.*, **30**, 49-57.
- Polliack,A. (2002) Hairy cell leukemia: biology, clinical diagnosis, unusual manifestations and associated disorders. *Rev.Clin.Exp.Hematol.*, **6**, 366-388.
- Preudhomme,C., Sagot,C., Boissel,N., Cayuela,J.M., Tigaud,I., De,B.S., Thomas,X., Raffoux,E., Lamandin,C., Castaigne,S., Fenaux,P., & Dombret,H. (2002) Favorable prognostic significance of CEBPA mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). *Blood*, **100**, 2717-2723.
- Puccetti,E. & Ruthardt,M. (2004) Acute promyelocytic leukemia: PML/RARalpha and the leukemic stem cell. *Leukemia*, **18**, 1169-1175.
- Qin,Y., Zhu,H., Jiang,B., Li,J., Lu,X., Li,L., Ruan,G., Liu,Y., Chen,S., & Huang,X. (2009) Expression patterns of WT1 and PRAME in acute myeloid leukemia patients and their usefulness for monitoring minimal residual disease. *Leuk.Res.*, **33**, 384-390.
- Quentmeier,H., Martelli,M.P., Dirks,W.G., Bolli,N., Liso,A., Macleod,R.A., Nicoletti,I., Mannucci,R., Pucciarini,A., Bigerna,B., Martelli,M.F., Mecucci,C., Drexler,H.G., & Falini,B. (2005) Cell line OCI/AML3 bears exon-12 NPM gene mutation-A and cytoplasmic expression of nucleophosmin. *Leukemia*, **19**, 1760-1767.
- Radmacher,M.D., Marcucci,G., Ruppert,A.S., Mrozek,K., Whitman,S.P., Vardiman,J.W., Paschka,P., Vukosavljevic,T., Baldus,C.D., Kolitz,J.E., Caligiuri,M.A., Larson,R.A., & Bloomfield,C.D. (2006) Independent confirmation of a prognostic gene-expression signature in adult acute myeloid leukemia with a normal karyotype: a Cancer and Leukemia Group B study. *Blood*, **108**, 1677-1683.
- Radomska,H.S., Huettner,C.S., Zhang,P., Cheng,T., Scadden,D.T., & Tenen,D.G. (1998) CCAAT/enhancer binding protein alpha is a regulatory switch sufficient for induction of granulocytic development from bipotential myeloid progenitors. *Mol.Cell Biol.*, **18**, 4301-4314.

- Raponi,M., Harousseau,J.L., Lancet,J.E., Lowenberg,B., Stone,R., Zhang,Y., Rackoff,W., Wang,Y., & Atkins,D. (2007) Identification of molecular predictors of response in a study of tipifarnib treatment in relapsed and refractory acute myelogenous leukemia. *Clin.Cancer Res.*, **13**, 2254-2260.
- Raskind,W.H., Steinmann,L., & Najfeld,V. (1998) Clonal development of myeloproliferative disorders: clues to hematopoietic differentiation and multistep pathogenesis of cancer. *Leukemia*, **12**, 108-116.
- Ravandi,F. & Estrov,Z. (2006) Eradication of leukemia stem cells as a new goal of therapy in leukemia. *Clin.Cancer Res.*, **12**, 340-344.
- Redner,R.L., Rush,E.A., Faas,S., Rudert,W.A., & Corey,S.J. (1996) The t(5;17) variant of acute promyelocytic leukemia expresses a nucleophosmin-retinoic acid receptor fusion. *Blood*, **87**, 882-886.
- Reilly,J.T. (2005) Pathogenesis of acute myeloid leukaemia and inv(16)(p13;q22): a paradigm for understanding leukaemogenesis? *Br.J.Haematol.*, **128**, 18-34.
- Rhoades,K.L., Hetherington,C.J., Rowley,J.D., Hiebert,S.W., Nucifora,G., Tenen,D.G., & Zhang,D.E. (1996) Synergistic up-regulation of the myeloid-specific promoter for the macrophage colony-stimulating factor receptor by AML1 and the t(8;21) fusion protein may contribute to leukemogenesis. *Proc.Natl.Acad.Sci.U.S.A*, **93**, 11895-11900.
- Rhoads,D.D., Dixit,A., & Roufa,D.J. (1986) Primary structure of human ribosomal protein S14 and the gene that encodes it. *Mol.Cell Biol.*, **6**, 2774-2783.
- Riedt,T., Ebinger,M., Salih,H.R., Tomiuk,J., Handgretinger,R., Kanz,L., Grunebach,F., & Lengerke,C. (2009) Aberrant expression of the homeobox gene CDX2 in pediatric acute lymphoblastic leukemia. *Blood*, **113**, 4049-4051.
- Roche,J., Zeng,C., Baron,A., Gadgil,S., Gemmill,R.M., Tigaud,I., Thomas,X., & Drabkin,H.A. (2004) Hox expression in AML identifies a distinct subset of patients with intermediate cytogenetics. *Leukemia*, **18**, 1059-1063.
- Rodriguez,S., Gaunt,T.R., & Day,I.N. (2007) Molecular genetics of human growth hormone, insulin-like growth factors and their pathways in common disease. *Hum.Genet.*, **122**, 1-21.
- Romana,S.P., Radford-Weiss,I., Ben,A.R., Schluth,C., Petit,A., Dastugue,N., Talmant,P., Bilhou-Nabera,C., Mugneret,F., Lafage-Pochitaloff,M., Mozziconacci,M.J., Andrieu,J., Lai,J.L., Terre,C., Rack,K., Cornillet-Lefebvre,P., Luquet,I., Nadal,N.,

- Nguyen-Khac,F., Perot,C., Van den,A.J., Fert-Ferrer,S., Cabrol,C., Charrin,C., Tigaud,I., Poirel,H., Vekemans,M., Bernard,O.A., & Berger,R. (2006) NUP98 rearrangements in hematopoietic malignancies: a study of the Groupe Francophone de Cytogenetique Hematologique. *Leukemia*, **20**, 696-706.
- Romeo,P.H., Prandini,M.H., Joulin,V., Mignotte,V., Prenant,M., Vainchenker,W., Marguerie,G., & Uzan,G. (1990) Megakaryocytic and erythrocytic lineages share specific transcription factors. *Nature*, **344**, 447-449.
- Rosen,J.M. & Jordan,C.T. (2009) The increasing complexity of the cancer stem cell paradigm. *Science*, **324**, 1670-1673.
- Ross,D.M., Branford,S., Melo,J.V., & Hughes,T.P. (2009) Reply to 'What do we mean by sensitivity when we talk about detecting minimal residual disease?' by Steinbach and Debatin. *Leukemia*, **23**, 819-820.
- Ross,M.E., Mahfouz,R., Onciu,M., Liu,H.C., Zhou,X., Song,G., Shurtleff,S.A., Pounds,S., Cheng,C., Ma,J., Ribeiro,R.C., Rubnitz,J.E., Girtman,K., Williams,W.K., Raimondi,S.C., Liang,D.C., Shih,L.Y., Pui,C.H., & Downing,J.R. (2004) Gene expression profiling of pediatric acute myelogenous leukemia. *Blood*, **104**, 3679-3687.
- Rowley,J.D. (1973) Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*, **243**, 290-293.
- Rubin,J.S., Chan,A.M., Bottaro,D.P., Burgess,W.H., Taylor,W.G., Cech,A.C., Hirschfield,D.W., Wong,J., Miki,T., Finch,P.W., & . (1991) A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte growth factor. *Proc.Natl.Acad.Sci.U.S.A*, **88**, 415-419.
- Ruggero,D., Wang,Z.G., & Pandolfi,P.P. (2000) The puzzling multiple lives of PML and its role in the genesis of cancer. *Bioessays*, **22**, 827-835.
- Rusten,L.S., Lyman,S.D., Veiby,O.P., & Jacobsen,S.E. (1996) The FLT3 ligand is a direct and potent stimulator of the growth of primitive and committed human CD34+ bone marrow progenitor cells in vitro. *Blood*, **87**, 1317-1325.
- San Miguel,J.F., Martinez,A., Macedo,A., Vidriales,M.B., Lopez-Berges,C., Gonzalez,M., Caballero,D., Garcia-Marcos,M.A., Ramos,F., Fernandez-Calvo,J., Calmuntia,M.J., az-Mediavilla,J., & Orfao,A. (1997) Immunophenotyping investigation of minimal residual disease is a useful approach for predicting relapse in acute myeloid leukemia patients. *Blood*, **90**, 2465-2470.

- Saunders,M.J., Tobal,K., & Yin,J.A. (1994) Detection of t(8;21) by reverse transcriptase polymerase chain reaction in patients in remission of acute myeloid leukaemia type M2 after chemotherapy or bone marrow transplantation. *Leuk.Res.*, **18**, 891-895.
- Sauvageau,G., Lansdorp,P.M., Eaves,C.J., Hogge,D.E., Dragowska,W.H., Reid,D.S., Largman,C., Lawrence,H.J., & Humphries,R.K. (1994) Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. *Proc.Natl.Acad.Sci.U.S.A*, **91**, 12223-12227.
- Sauvageau,G., Thorsteinsdottir,U., Eaves,C.J., Lawrence,H.J., Largman,C., Lansdorp,P.M., & Humphries,R.K. (1995) Overexpression of HOXB4 in hematopoietic cells causes the selective expansion of more primitive populations in vitro and in vivo. *Genes Dev.*, **9**, 1753-1765.
- Sauvageau,G., Thorsteinsdottir,U., Hough,M.R., Hugo,P., Lawrence,H.J., Largman,C., & Humphries,R.K. (1997) Overexpression of HOXB3 in hematopoietic cells causes defective lymphoid development and progressive myeloproliferation. *Immunity.*, **6**, 13-22.
- Sawyers,C.L. (1999) Chronic myeloid leukemia. *N.Engl.J.Med.*, **340**, 1330-1340.
- Sawyers,C.L., Denny,C.T., & Witte,O.N. (1991) Leukemia and the disruption of normal hematopoiesis. *Cell*, **64**, 337-350.
- Schmid,D., Heinze,G., Linnerth,B., Tisljar,K., Kusec,R., Geissler,K., Sillaber,C., Laczika,K., Mitterbauer,M., Zochbauer,S., Mannhalter,C., Haas,O.A., Lechner,K., Jager,U., & Gaiger,A. (1997) Prognostic significance of WT1 gene expression at diagnosis in adult de novo acute myeloid leukemia. *Leukemia*, **11**, 639-643.
- Schnittger,S., Schoch,C., Dugas,M., Kern,W., Staib,P., Wuchter,C., Loffler,H., Sauerland,C.M., Serve,H., Buchner,T., Haferlach,T., & Hiddemann,W. (2002) Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. *Blood*, **100**, 59-66.
- Schnittger,S., Schoch,C., Kern,W., Mecucci,C., Tschulik,C., Martelli,M.F., Haferlach,T., Hiddemann,W., & Falini,B. (2005) Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. *Blood*, **106**, 3733-3739.

- Schnittger,S., Weisser,M., Schoch,C., Hiddemann,W., Haferlach,T., & Kern,W. (2003) New score predicting for prognosis in PML-RARA+, AML1-ETO+, or CBFMBYH11+ acute myeloid leukemia based on quantification of fusion transcripts. *Blood*, **102**, 2746-2755.
- Schoch,C., Kohlmann,A., Schnittger,S., Brors,B., Dugas,M., Mergenthaler,S., Kern,W., Hiddemann,W., Eils,R., & Haferlach,T. (2002) Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proc.Natl.Acad.Sci.U.S.A*, **99**, 10008-10013.
- Scholl,C., Bansal,D., Dohner,K., Eiwen,K., Huntly,B.J., Lee,B.H., Rucker,F.G., Schlenk,R.F., Bullinger,L., Dohner,H., Gilliland,D.G., & Frohling,S. (2007b) The homeobox gene CDX2 is aberrantly expressed in most cases of acute myeloid leukemia and promotes leukemogenesis. *J.Clin.Invest*, **117**, 1037-1048.
- Scholl,C., Bansal,D., Dohner,K., Eiwen,K., Huntly,B.J., Lee,B.H., Rucker,F.G., Schlenk,R.F., Bullinger,L., Dohner,H., Gilliland,D.G., & Frohling,S. (2007a) The homeobox gene CDX2 is aberrantly expressed in most cases of acute myeloid leukemia and promotes leukemogenesis. *J.Clin.Invest*, **117**, 1037-1048.
- Scholl,S., Loncarevic,I.F., Krause,C., Kunert,C., Clement,J.H., & Hoffken,K. (2005) Minimal residual disease based on patient specific Flt3-ITD and -ITT mutations in acute myeloid leukemia. *Leuk.Res.*, **29**, 849-853.
- Schuler,F. & Dolken,G. (2006) Detection and monitoring of minimal residual disease by quantitative real-time PCR. *Clin.Chim.Acta*, **363**, 147-156.
- Scott,E.W., Simon,M.C., Anastasi,J., & Singh,H. (1994) Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science*, **265**, 1573-1577.
- Seidel,C., Borset,M., Turesson,I., Abildgaard,N., Sundan,A., & Waage,A. (1998) Elevated serum concentrations of hepatocyte growth factor in patients with multiple myeloma. The Nordic Myeloma Study Group. *Blood*, **91**, 806-812.
- Sessions,J. (2007) Chronic myeloid leukemia in 2007. *J.Manag.Care Pharm.*, **13**, 4-7.
- Sherr,C.J. & Roberts,J.M. (1999) CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev.*, **13**, 1501-1512.
- Shih,L.Y., Huang,C.F., Wu,J.H., Lin,T.L., Dunn,P., Wang,P.N., Kuo,M.C., Lai,C.L., & Hsu,H.C. (2002) Internal tandem duplication of FLT3 in relapsed acute myeloid leukemia: a

- comparative analysis of bone marrow samples from 108 adult patients at diagnosis and relapse. *Blood*, **100**, 2387-2392.
- Shih,L.Y., Huang,C.F., Wu,J.H., Wang,P.N., Lin,T.L., Dunn,P., Chou,M.C., Kuo,M.C., & Tang,C.C. (2004a) Heterogeneous patterns of FLT3 Asp(835) mutations in relapsed de novo acute myeloid leukemia: a comparative analysis of 120 paired diagnostic and relapse bone marrow samples. *Clin.Cancer Res.*, **10**, 1326-1332.
- Shih,L.Y., Huang,C.F., Wu,J.H., Wang,P.N., Lin,T.L., Dunn,P., Chou,M.C., Kuo,M.C., & Tang,C.C. (2004b) Heterogeneous patterns of FLT3 Asp(835) mutations in relapsed de novo acute myeloid leukemia: a comparative analysis of 120 paired diagnostic and relapse bone marrow samples. *Clin.Cancer Res.*, **10**, 1326-1332.
- Shivdasani,R.A., Fujiwara,Y., McDevitt,M.A., & Orkin,S.H. (1997) A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J.*, **16**, 3965-3973.
- Sievers,E.L., Lange,B.J., Buckley,J.D., Smith,F.O., Wells,D.A., igneault-Creech,C.A., Shults,K.E., Bernstein,I.D., & Loken,M.R. (1996) Prediction of relapse of pediatric acute myeloid leukemia by use of multidimensional flow cytometry. *J.Natl.Cancer Inst.*, **88**, 1483-1488.
- Skinninger,B.F., Connors,J.M., Sutcliffe,S.B., & Gascoyne,R.D. (1999) Anaplastic large cell lymphoma: a clinicopathologic analysis. *Hematol.Oncol.*, **17**, 137-148.
- Slovak,M.L., Kopecky,K.J., Cassileth,P.A., Harrington,D.H., Theil,K.S., Mohamed,A., Paietta,E., Willman,C.L., Head,D.R., Rowe,J.M., Forman,S.J., & Appelbaum,F.R. (2000) Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group Study. *Blood*, **96**, 4075-4083.
- Small,D. (2006) FLT3 mutations: biology and treatment. *Hematology.Am.Soc.Hematol.Educ.Program.*, 178-184.
- Small,D., Levenstein,M., Kim,E., Carow,C., Amin,S., Rockwell,P., Witte,L., Burrow,C., Ratajczak,M.Z., Gewirtz,A.M., & . (1994) STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells. *Proc.Natl.Acad.Sci.U.S.A.*, **91**, 459-463.

- Smith,M.D., Dawson,S.J., Boxer,L.M., & Latchman,D.S. (1998) The N-terminal domain unique to the long form of the Brn-3a transcription factor is essential to protect neuronal cells from apoptosis and for the activation of Bcl-2 gene expression. *Nucleic Acids Res.*, **26**, 4100-4107.
- Sorrentino,B.P. (2004) Clinical strategies for expansion of haematopoietic stem cells. *Nat.Rev.Immunol.*, **4**, 878-888.
- Speck,N.A. & Gilliland,D.G. (2002) Core-binding factors in haematopoiesis and leukaemia. *Nat.Rev.Cancer*, **2**, 502-513.
- Stasi,R., Evangelista,M.L., Buccisano,F., Venditti,A., & Amadori,S. (2008) Gemtuzumab ozogamicin in the treatment of acute myeloid leukemia. *Cancer Treat.Rev.*, **34**, 49-60.
- Steinbach,D. & Debatin,K.M. (2008) What do we mean by sensitivity when we talk about detecting minimal residual disease? *Leukemia*, **22**, 1638-1639.
- Steinbach,D., Schramm,A., Eggert,A., Onda,M., Dawczynski,K., Rump,A., Pastan,I., Wittig,S., Pfaffendorf,N., Voigt,A., Zintl,F., & Gruhn,B. (2006) Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. *Clin.Cancer Res.*, **12**, 2434-2441.
- Stirewalt,D.L., Meshinchi,S., Kopecky,K.J., Fan,W., Pogossova-Agadjanyan,E.L., Engel,J.H., Cronk,M.R., Dorcy,K.S., McQuary,A.R., Hockenbery,D., Wood,B., Heimfeld,S., & Radich,J.P. (2008) Identification of genes with abnormal expression changes in acute myeloid leukemia. *Genes Chromosomes.Cancer*, **47**, 8-20.
- Stirewalt,D.L., Willman,C.L., & Radich,J.P. (2001) Quantitative, real-time polymerase chain reactions for FLT3 internal tandem duplications are highly sensitive and specific. *Leuk.Res.*, **25**, 1085-1088.
- Stoker,M., Gherardi,E., Perryman,M., & Gray,J. (1987) Scatter factor is a fibroblast-derived modulator of epithelial cell mobility. *Nature*, **327**, 239-242.
- Stoker,M. & Perryman,M. (1985) An epithelial scatter factor released by embryo fibroblasts. *J.Cell Sci.*, **77**, 209-223.
- Stone,R.M., O'Donnell,M.R., & Sekeres,M.A. (2004) Acute myeloid leukemia. *Hematology.Am.Soc.Hematol.Educ.Program.*, 98-117.
- Sugimoto,M., Oohashi,T., & Ninomiya,Y. (1994) The genes COL4A5 and COL4A6, coding for basement membrane collagen chains alpha 5(IV) and alpha 6(IV), are located head-to-head in close proximity on human chromosome Xq22 and COL4A6 is

transcribed from two alternative promoters. *Proc.Natl.Acad.Sci.U.S.A*, **91**, 11679-11683.

Suzuki,T., Kiyoi,H., Ozeki,K., Tomita,A., Yamaji,S., Suzuki,R., Kodera,Y., Miyawaki,S., Asou,N., Kuriyama,K., Yagasaki,F., Shimazaki,C., Akiyama,H., Nishimura,M., Motoji,T., Shinagawa,K., Takeshita,A., Ueda,R., Kinoshita,T., Emi,N., & Naoe,T. (2005) Clinical characteristics and prognostic implications of NPM1 mutations in acute myeloid leukemia. *Blood*, **106**, 2854-2861.

Szebeni,A., Herrera,J.E., & Olson,M.O. (1995) Interaction of nucleolar protein B23 with peptides related to nuclear localization signals. *Biochemistry*, **34**, 8037-8042.

Takahashi,A., Satake,M., Yamaguchi-Iwai,Y., Bae,S.C., Lu,J., Maruyama,M., Zhang,Y.W., Oka,H., Arai,N., Arai,K., & . (1995) Positive and negative regulation of granulocyte-macrophage colony-stimulating factor promoter activity by AML1-related transcription factor, PEBP2. *Blood*, **86**, 607-616.

Takeda,A., Goolsby,C., & Yaseen,N.R. (2006) NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. *Cancer Res.*, **66**, 6628-6637.

Takenokuchi,M., Yasuda,C., Takeuchi,K., Nakamachi,Y., Mukai,M., Kondo,S., Kumagai,S., Saigo,K., Murayama,T., Koizumi,T., & Tatsumi,E. (2004) Quantitative nested reverse transcriptase PCR vs. real-time PCR for measuring AML1/ETO (MTG8) transcripts. *Clin.Lab Haematol.*, **26**, 107-114.

Tallman,M.S., Gilliland,D.G., & Rowe,J.M. (2005) Drug therapy for acute myeloid leukemia. *Blood*, **106**, 1154-1163.

Tanner,S.M., Austin,J.L., Leone,G., Rush,L.J., Plass,C., Heinonen,K., Mrozek,K., Sill,H., Knuutila,S., Kolitz,J.E., Archer,K.J., Caligiuri,M.A., Bloomfield,C.D., & de la,C.A. (2001) BAALC, the human member of a novel mammalian neuroectoderm gene lineage, is implicated in hematopoiesis and acute leukemia. *Proc.Natl.Acad.Sci.U.S.A*, **98**, 13901-13906.

Tantravahi,R., Schwenn,M., Henkle,C., Nell,M., Leavitt,P.R., Griffin,J.D., & Weinstein,H.J. (1984) A pericentric inversion of chromosome 16 is associated with dysplastic marrow eosinophils in acute myelomonocytic leukemia. *Blood*, **63**, 800-802.

Teitell,M.A. & Pandolfi,P.P. (2009) Molecular genetics of acute lymphoblastic leukemia. *Annu.Rev.Pathol.*, **4**, 175-198.

- Testa, J.R., Mintz, U., Rowley, J.D., Vardiman, J.W., & Golomb, H.M. (1979) Evolution of karyotypes in acute nonlymphocytic leukemia. *Cancer Res.*, **39**, 3619-3627.
- Thiede, C., Koch, S., Creutzig, E., Steudel, C., Illmer, T., Schaich, M., & Ehninger, G. (2006) Prevalence and prognostic impact of NPM1 mutations in 1485 adult patients with acute myeloid leukemia (AML). *Blood*, **107**, 4011-4020.
- Thiede, C., Steudel, C., Mohr, B., Schaich, M., Schakel, U., Platzbecker, U., Wermke, M., Bornhauser, M., Ritter, M., Neubauer, A., Ehninger, G., & Illmer, T. (2002) Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. *Blood*, **99**, 4326-4335.
- Thompson, A., Quinn, M.F., Grimwade, D., O'Neill, C.M., Ahmed, M.R., Grimes, S., McMullin, M.F., Cotter, F., & Lappin, T.R. (2003) Global down-regulation of HOX gene expression in PML-RARalpha + acute promyelocytic leukemia identified by small-array real-time PCR. *Blood*, **101**, 1558-1565.
- Tiacci, E., Liso, A., Piris, M., & Falini, B. (2006) Evolving concepts in the pathogenesis of hairy-cell leukaemia. *Nat.Rev.Cancer*, **6**, 437-448.
- Tobal, K., Johnson, P.R., Saunders, M.J., Harrison, C.J., & Liu Yin, J.A. (1995) Detection of CBFB/MYH11 transcripts in patients with inversion and other abnormalities of chromosome 16 at presentation and remission. *Br.J.Haematol.*, **91**, 104-108.
- Tobal, K., Moore, H., Macheta, M., & Yin, J.A. (2001) Monitoring minimal residual disease and predicting relapse in APL by quantitating PML-RARalpha transcripts with a sensitive competitive RT-PCR method. *Leukemia*, **15**, 1060-1065.
- Tobal, K., Newton, J., Macheta, M., Chang, J., Morgenstern, G., Evans, P.A., Morgan, G., Lucas, G.S., & Liu Yin, J.A. (2000) Molecular quantitation of minimal residual disease in acute myeloid leukemia with t(8;21) can identify patients in durable remission and predict clinical relapse. *Blood*, **95**, 815-819.
- Tobal, K. & Yin, J.A. (1996) Monitoring of minimal residual disease by quantitative reverse transcriptase-polymerase chain reaction for AML1-MTG8 transcripts in AML-M2 with t(8; 21). *Blood*, **88**, 3704-3709.
- Trka, J., Kalinova, M., Hrusak, O., Zuna, J., Krejci, O., Madzo, J., Sedlacek, P., Vavra, V., Michalova, K., Jarosova, M., & Stary, J. (2002) Real-time quantitative PCR detection of WT1 gene expression in children with AML: prognostic significance,

- correlation with disease status and residual disease detection by flow cytometry. *Leukemia*, **16**, 1381-1389.
- Tsai,S.F., Martin,D.I., Zon,L.I., D'Andrea,A.D., Wong,G.G., & Orkin,S.H. (1989) Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature*, **339**, 446-451.
- Tusher,V.G., Tibshirani,R., & Chu,G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc.Natl.Acad.Sci.U.S.A*, **98**, 5116-5121.
- Ullrich,A. & Schlessinger,J. (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell*, **61**, 203-212.
- Valk,P.J., Verhaak,R.G., Beijen,M.A., Erpelinck,C.A., Barjesteh van Waalwijk van Doorn-Khosrovani, Boer,J.M., Beverloo,H.B., Moorhouse,M.J., van der Spek,P.J., Lowenberg,B., & Delwel,R. (2004) Prognostically useful gene-expression profiles in acute myeloid leukemia. *N.Engl.J.Med.*, **350**, 1617-1628.
- van Baren,N., Chambost,H., Ferrant,A., Michaux,L., Ikeda,H., Millard,I., Olive,D., Boon,T., & Coulie,P.G. (1998) PRAME, a gene encoding an antigen recognized on a human melanoma by cytolytic T cells, is expressed in acute leukaemia cells. *Br.J Haematol.*, **102**, 1376-1379.
- van der Velden,V., Hochhaus,A., Cazzaniga,G., Szczepanski,T., Gabert,J., & van Dongen,J.J. (2003) Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. *Leukemia*, **17**, 1013-1034.
- van Dongen,J.J., Macintyre,E.A., Gabert,J.A., Delabesse,E., Rossi,V., Saglio,G., Gottardi,E., Rambaldi,A., Dotti,G., Griesinger,F., Parreira,A., Gameiro,P., Diaz,M.G., Malec,M., Langerak,A.W., San Miguel,J.F., & Biondi,A. (1999) Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. *Leukemia*, **13**, 1901-1928.
- van Wely,K.H., Meester-Smoor,M.A., Janssen,M.J., Aarnoudse,A.J., Grosveld,G.C., & Zwarthoff,E.C. (2007) The MN1-TEL myeloid leukemia-associated fusion protein has a dominant-negative effect on RAR-RXR-mediated transcription. *Oncogene*, **26**, 5733-5740.
- van Wely,K.H., Molijn,A.C., Buijs,A., Meester-Smoor,M.A., Aarnoudse,A.J., Hellemons,A., den Besten,P., Grosveld,G.C., &

- Zwarthoff,E.C. (2003) The MN1 oncoprotein synergizes with coactivators RAC3 and p300 in RAR-RXR-mediated transcription. *Oncogene*, **22**, 699-709.
- Vannucchi,A.M., Guglielmelli,P., & Tefferi,A. (2009) Advances in understanding and management of myeloproliferative neoplasms. *CA Cancer J.Clin.*, **59**, 171-191.
- Vardiman,J., Catovsky,D., Flandrin,G.R., & Pierre,R. (1984) Fourth International Workshop on Chromosomes in Leukemia 1982: Correlation between morphology and karyotype. *Cancer Genet.Cytogenet.*, **11**, 275-281.
- Vardiman,J.W. (2009) The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: An overview with emphasis on the myeloid neoplasms. *Chem.Biol.Interact.*
- Vardiman,J.W., Harris,N.L., & Brunning,R.D. (2002) The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood*, **100**, 2292-2302.
- Vardiman,J.W., Thiele,J., Arber,D.A., Brunning,R.D., Borowitz,M.J., Porwit,A., Harris,N.L., Le Beau,M.M., Hellstrom-Lindberg,E., Tefferi,A., & Bloomfield,C.D. (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*, **114**, 937-951.
- Vaughan,W.P. & Karp,J.E. (2008) The Long Road to a Cure for Acute Myelocytic Leukemia: From Intensity to Specificity. *Journal of Clinical Oncology*, **26**, 3475-3477.
- Veiby,O.P., Lyman,S.D., & Jacobsen,S.E. (1996) Combined signaling through interleukin-7 receptors and flt3 but not c-kit potently and selectively promotes B-cell commitment and differentiation from uncommitted murine bone marrow progenitor cells. *Blood*, **88**, 1256-1265.
- Verhaak,R.G., Goudswaard,C.S., van,P.W., Bijl,M.A., Sanders,M.A., Hagens,W., Uitterlinden,A.G., Erpelinck,C.A., Delwel,R., Lowenberg,B., & Valk,P.J. (2005) Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. *Blood*, **106**, 3747-3754.
- Verhaak,R.G., Wouters,B.J., Erpelinck,C.A., Abbas,S., Beverloo,H.B., Lugthart,S., Lowenberg,B., Delwel,R., & Valk,P.J. (2009) Prediction of molecular subtypes in acute myeloid leukemia

- based on gene expression profiling. *Haematologica*, **94**, 131-134.
- Verstovsek,S., Kantarjian,H., Estey,E., Aguayo,A., Giles,F.J., Manshouri,T., Koller,C., Estrov,Z., Freireich,E., Keating,M., & Albitar,M. (2001) Plasma hepatocyte growth factor is a prognostic factor in patients with acute myeloid leukemia but not in patients with myelodysplastic syndrome. *Leukemia*, **15**, 1165-1170.
- Vitoux,D., Nasr,R., & de,T.H. (2007) Acute promyelocytic leukemia: new issues on pathogenesis and treatment response. *Int.J.Biochem.Cell Biol.*, **39**, 1063-1070.
- Vorwerk,P., Mohnike,K., Wex,H., Rohl,F.W., Zimmermann,M., Blum,W.F., & Mittler,U. (2005) Insulin-like growth factor binding protein-2 at diagnosis of childhood acute lymphoblastic leukemia and the prediction of relapse risk. *J.Clin.Endocrinol.Metab*, **90**, 3022-3027.
- Vorwerk,P., Wex,H., Hohmann,B., Mohnike,K., Schmidt,U., & Mittler,U. (2002) Expression of components of the IGF signalling system in childhood acute lymphoblastic leukaemia. *Mol.Pathol.*, **55**, 40-45.
- Walsby,E., Walsh,V., Pepper,C., Burnett,A., & Mills,K. (2008) Effects of the aurora kinase inhibitors AZD1152-HQPA and ZM447439 on growth arrest and polyploidy in acute myeloid leukemia cell lines and primary blasts. *Haematologica*, **93**, 662-669.
- Wang,Q., Stacy,T., Binder,M., Marin-Padilla,M., Sharpe,A.H., & Speck,N.A. (1996a) Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc.Natl.Acad.Sci.U.S.A*, **93**, 3444-3449.
- Wang,Q., Stacy,T., Miller,J.D., Lewis,A.F., Gu,T.L., Huang,X., Bushweller,J.H., Bories,J.C., Alt,F.W., Ryan,G., Liu,P.P., Wynshaw-Boris,A., Binder,M., Marin-Padilla,M., Sharpe,A.H., & Speck,N.A. (1996b) The CBFbeta subunit is essential for CBFalpha2 (AML1) function in vivo. *Cell*, **87**, 697-708.
- Wang,S., Wang,Q., Crute,B.E., Melnikova,I.N., Keller,S.R., & Speck,N.A. (1993) Cloning and characterization of subunits of the T-cell receptor and murine leukemia virus enhancer core-binding factor. *Mol.Cell Biol.*, **13**, 3324-3339.
- Wang,W., Budhu,A., Forgues,M., & Wang,X.W. (2005) Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. *Nat.Cell Biol.*, **7**, 823-830.

- Wang,Z.G., Delva,L., Gaboli,M., Rivi,R., Giorgio,M., Cordon-Cardo,C., Grosveld,F., & Pandolfi,P.P. (1998) Role of PML in cell growth and the retinoic acid pathway. *Science*, **279**, 1547-1551.
- Weidner,K.M., Arakaki,N., Hartmann,G., Vandekerckhove,J., Weingart,S., Rieder,H., Fonatsch,C., Tsubouchi,H., Hishida,T., Daikuhara,Y., & . (1991) Evidence for the identity of human scatter factor and human hepatocyte growth factor. *Proc.Natl.Acad.Sci.U.S.A*, **88**, 7001-7005.
- Weimar,I.S., Voermans,C., Bourhis,J.H., Miranda,N., van den Berk,P.C., Nakamura,T., de Gast,G.C., & Gerritsen,W.R. (1998) Hepatocyte growth factor/scatter factor (HGF/SF) affects proliferation and migration of myeloid leukemic cells. *Leukemia*, **12**, 1195-1203.
- Weisberg,E., Boulton,C., Kelly,L.M., Manley,P., Fabbro,D., Meyer,T., Gilliland,D.G., & Griffin,J.D. (2002) Inhibition of mutant FLT3 receptors in leukemia cells by the small molecule tyrosine kinase inhibitor PKC412. *Cancer Cell*, **1**, 433-443.
- Weisel,K.C., Yildirim,S., Schweikle,E., Kanz,L., & Mohle,R. (2009) Regulation of FLT3 and its ligand in normal hematopoietic progenitor cells. *Ann.Hematol.*, **88**, 203-211.
- Weiss,A. & Schlessinger,J. (1998) Switching signals on or off by receptor dimerization. *Cell*, **94**, 277-280.
- Weiss,M.J. & Orkin,S.H. (1995) Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proc.Natl.Acad.Sci.U.S.A*, **92**, 9623-9627.
- Weisser,M., Kern,W., Rauhut,S., Schoch,C., Hiddemann,W., Haferlach,T., & Schnittger,S. (2005) Prognostic impact of RT-PCR-based quantification of WT1 gene expression during MRD monitoring of acute myeloid leukemia. *Leukemia*, **19**, 1416-1423.
- Wen,J., Matsumoto,K., Taniura,N., Tomioka,D., & Nakamura,T. (2007) Inhibition of colon cancer growth and metastasis by NK4 gene repetitive delivery in mice. *Biochem.Biophys.Res.Comm.*, **358**, 117-123.
- West,R.B., Nuyten,D.S., Subramanian,S., Nielsen,T.O., Corless,C.L., Rubin,B.P., Montgomery,K., Zhu,S., Patel,R., Hernandez-Boussard,T., Goldblum,J.R., Brown,P.O., van,d., V, & van de,R.M. (2005) Determination of stromal signatures in breast carcinoma. *PLoS.Biol.*, **3**, e187.
- Whitman,S.P., Archer,K.J., Feng,L., Baldus,C., Becknell,B., Carlson,B.D., Carroll,A.J., Mrozek,K., Vardiman,J.W., George,S.L., Kolitz,J.E., Larson,R.A., Bloomfield,C.D., &

- Caligiuri, M.A. (2001) Absence of the wild-type allele predicts poor prognosis in adult de novo acute myeloid leukemia with normal cytogenetics and the internal tandem duplication of FLT3: a cancer and leukemia group B study. *Cancer Res.*, **61**, 7233-7239.
- Whitman, S.P., Ruppert, A.S., Radmacher, M.D., Mrozek, K., Paschka, P., Langer, C., Baldus, C.D., Wen, J., Racke, F., Powell, B.L., Kolitz, J.E., Larson, R.A., Caligiuri, M.A., Marcucci, G., & Bloomfield, C.D. (2008) FLT3 D835/I836 mutations are associated with poor disease-free survival and a distinct gene-expression signature among younger adults with de novo cytogenetically normal acute myeloid leukemia lacking FLT3 internal tandem duplications. *Blood*, **111**, 1552-1559.
- Wicki, A., Lehembre, F., Wick, N., Hantusch, B., Kerjaschki, D., & Christofori, G. (2006) Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer Cell*, **9**, 261-272.
- Wieser, R. (2007) The oncogene and developmental regulator EVI1: expression, biochemical properties, and biological functions. *Gene*, **396**, 346-357.
- Wilson, C.S., Davidson, G.S., Martin, S.B., Andries, E., Potter, J., Harvey, R., Ar, K., Xu, Y., Kopecky, K.J., Ankerst, D.P., Gundacker, H., Slovak, M.L., Mosquera-Caro, M., Chen, I.M., Stirewalt, D.L., Murphy, M., Schultz, F.A., Kang, H., Wang, X., Radich, J.P., Appelbaum, F.R., Atlas, S.R., Godwin, J., & Willman, C.L. (2006) Gene expression profiling of adult acute myeloid leukemia identifies novel biologic clusters for risk classification and outcome prediction. *Blood*, **108**, 685-696.
- Yagi, R., Chen, L.F., Shigesada, K., Murakami, Y., & Ito, Y. (1999) A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO J.*, **18**, 2551-2562.
- Yamamoto, Y., Kiyoi, H., Nakano, Y., Suzuki, R., Kidera, Y., Miyawaki, S., Asou, N., Kuriyama, K., Yagasaki, F., Shimazaki, C., Akiyama, H., Saito, K., Nishimura, M., Motoji, T., Shinagawa, K., Takeshita, A., Saito, H., Ueda, R., Ohno, R., & Naoe, T. (2001) Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood*, **97**, 2434-2439.
- Yang, L., Han, Y., Suarez, S.F., & Minden, M.D. (2007) A tumor suppressor and oncogene: the WT1 story. *Leukemia*, **21**, 868-876.
- Yao, B., Rakhade, S.N., Li, Q., Ahmed, S., Krauss, R., Draghici, S., & Loeb, J.A. (2004) Accuracy of cDNA microarray methods to

- detect small gene expression changes induced by neuregulin on breast epithelial cells. *BMC.Bioinformatics.*, **5**, 99.
- Yarden,Y., Kuang,W.J., Yang-Feng,T., Coussens,L., Munemitsu,S., Dull,T.J., Chen,E., Schlessinger,J., Francke,U., & Ullrich,A. (1987) Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand. *EMBO J.*, **6**, 3341-3351.
- Yee,K.W., O'Farrell,A.M., Smolich,B.D., Cherrington,J.M., McMahon,G., Wait,C.L., McGreevey,L.S., Griffith,D.J., & Heinrich,M.C. (2002) SU5416 and SU5614 inhibit kinase activity of wild-type and mutant FLT3 receptor tyrosine kinase. *Blood*, **100**, 2941-2949.
- Yin,J.A. & Grimwade,D. (2002) Minimal residual disease evaluation in acute myeloid leukaemia. *Lancet*, **360**, 160-162.
- Yin,J.A. & Tobal,K. (1999) Detection of minimal residual disease in acute myeloid leukaemia: methodologies, clinical and biological significance. *Br.J Haematol.*, **106**, 578-590.
- Yoneda-Kato,N., Fukuhara,S., & Kato,J. (1999) Apoptosis induced by the myelodysplastic syndrome-associated NPM-MLF1 chimeric protein. *Oncogene*, **18**, 3716-3724.
- Yuasa,H., Oike,Y., Iwama,A., Nishikata,I., Sugiyama,D., Perkins,A., Mucenski,M.L., Suda,T., & Morishita,K. (2005) Oncogenic transcription factor Evi1 regulates hematopoietic stem cell proliferation through GATA-2 expression. *EMBO J.*, **24**, 1976-1987.
- Yuen,T., Wurnbach,E., Pfeffer,R.L., Ebersole,B.J., & Sealfon,S.C. (2002) Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res.*, **30**, e48.
- Zamore,P.D., Tuschl,T., Sharp,P.A., & Bartel,D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101**, 25-33.
- Zhang,J., Niu,C., Ye,L., Huang,H., He,X., Tong,W.G., Ross,J., Haug,J., Johnson,T., Feng,J.Q., Harris,S., Wiedemann,L.M., Mishina,Y., & Li,L. (2003a) Identification of the haematopoietic stem cell niche and control of the niche size. *Nature*, **425**, 836-841.
- Zhang,X., Zhu,T., Chen,Y., Mertani,H.C., Lee,K.O., & Lobie,P.E. (2003b) Human growth hormone-regulated HOXA1 is a human mammary epithelial oncogene. *J.Biol.Chem.*, **278**, 7580-7590.
- Zhong,S., Salomoni,P., & Pandolfi,P.P. (2000) The transcriptional role of PML and the nuclear body. *Nat.Cell Biol.*, **2**, E85-E90.

- Zon,L.I., Yamaguchi,Y., Yee,K., Albee,E.A., Kimura,A., Bennett,J.C., Orkin,S.H., & Ackerman,S.J. (1993) Expression of mRNA for the GATA-binding proteins in human eosinophils and basophils: potential role in gene transcription. *Blood*, **81**, 3234-3241.
- Zucker-Franklin,D. & Philipp,C.S. (2000) Platelet production in the pulmonary capillary bed: new ultrastructural evidence for an old concept. *Am.J.Pathol.*, **157**, 69-74.