The Characterisation of Telomere Dynamics in Myelodysplastic Syndromes

Jenna Rachel Williams

PhD

January 2014

Contents

Title page i

Acknowledgements vii

Abstract viii

Abbreviations ix

Chapter 1: Introduction

Part 1: Myelodysplastic Syndromes (MDSs) and Acute Myeloid Leukaemia (AML)

1.1 Haematopoietic System 1-2

1.2 MDS Pathology 2-4

1.3 French-American-British (FAB) Classification 5-6

1.4 The World Health Organisation (WHO) Classification 6-12

1.4.1 WHO Classification for Myelodysplastic Syndrome (MDS)

1.4.2 WHO Classification for Acute Myeloid Leukaemia (AML)

1.5 International Prognostic Scoring System (IPSS) for MDS 12-13

1.6 Revised-International Prognostic Scoring System (IPSS-R) for MDS 13-15

1.7 Therapeutic Options in MDS 15-16

1.8 Prognostic Scoring in AML 16-20

1.9 Therapeutic Studies in AML 20-21

1.10 MDS Cell of Origin 21-23

1.11 AML Cell of Origin 23-25

1.12 Chromosomal Abnormalities 25-28

1.13 Paradox 28-31

Part 2: Telomeres

1.14	**History of Telomeres**	32
1.15	**The End Replication Problem**	32-34
1.16	**Telomerase**	34-35
1.17	**The Mechanism of Telomerase**	35-36
1.18	**'Capping' Linear Chromosomes**	36
1.19	**The Shelterin Complex**	37-39
1.20	**The Protein 'Counting' Model**	39-40
1.21	**Telomere Length Homeostasis**	40-41
1.22	**The G-quadruplex**	41-42
1.23	**Telomerase and Cancer Therapeutics**	42-44
1.24	**Genetic Determination of Telomere Size**	44
1.25	**Sub-telomere Structure**	44-45
1.26	**Methods of Telomere Measurement**	45-47
	1.26.1 Terminal Restriction Fragment (TRF)	
	1.26.2 Quantitative Fluorescence *in situ* Hybridisation (Q-FISH)	
	1.26.3 Flow Fluorescence *in situ* Hybridisation (Flow-FISH)	
	1.26.4 Quantitative telomere-specific PCR (Q-PCR)	
	1.26.5 Single Telomere Length Analysis (STELA)	
1.27	**Telomeres and Homologous Recombination**	47-49
1.28	**Non-Homologous End Joining/Microhomology Mediated End Joining**	49-52
1.29	**Telomeres and Cancer**	53-54
1.30	**Telomeres and Haematological Disorders**	54-57
1.31	**Telomere Length Analysis in MDS/AML**	57-61
1.32	**Aim of Research**	62-63

Chapter 2: Materials and Methods

2.1	**Patient samples**	64-66
2.2	**DNA Extraction**	67-68
2.3	**Single Telomere Length Analysis (STELA)**	68-69
2.4	**Hybridisation**	69

2.5 Fusion Assay 70-71
2.6 TVR Mapping 71
2.7 CD34 Cell Purification using Magnetic Beads 71-72
2.8 Telomerase Assay 72-73
2.9 Statistical Analysis 73
2.10 Array-CGH 73
2.11 Development of 6q STELA 73-74
2.12 Oligonucleotides 74-75

Chapter 3: Telomere Dynamics in MDS and AML

3.1 Abstract 76-77
3.2 Introduction 78-80
3.3 Telomere Length in MDS and AML 81
3.4 Inter-individual Telomere Length Variation 82-83
3.5 Telomere Length Correlation 84
3.6 Intra-clonal Variation 85-90
3.7 Telomere Rapid Deletion (TRD) Events 91-92
3.8 Telomerase Up-regulation 93-94
3.9 Telomere Erosion 95-96
3.10 TVR: Telomere Variant Repeat 97-100
3.11 Cell Fractionation 101-102
3.12 Discussion 103-107

Chapter 4: Telomere Length and Prognosis in MDS

4.1 Abstract 108-109
4.2 Introduction 110-112
4.3 Telomere Length and Age at Onset 113-115
4.4 Blast Cell Percentage and Telomere Length 116-118
4.5 Age-Adjusted Telomere Length in MDS Patients 119
4.6 Telomere Length and Peripheral Blood Cytopenia 120-125
4.7 Telomere Length and Cytogenetic Risk in MDS 126-130
4.8 Telomere Length and the IPSS Scoring System 131-135

4.9 Telomere Length and Survival in MDS 136-139

4.10 Discussion 140-145

Chapter 5: Telomere Length and Prognosis in AML

5.1 Abstract 146

5.2 Introduction 147-149

5.3 Telomere Length and Age at Diagnosis 150-151

5.4 Telomere Length and Gender 152

5.5 Telomere Length and Marrow Blast Count 153-154

5.6 Telomere Length and Presenting White Blood Cell (WBC) Count 154-155

5.7 Telomere Length and AML Type (*De novo*/ Secondary) 156

5.8 Telomere Length and WHO Performance Status 157

5.9 Telomere Length and Cytogenetic Risk Group in AML 158

5.10 Telomere Length and FLT3/ITD Mutation Status 159

5.11 Telomere Length and FLT3/TKD Mutation Status 160

5.12 Telomere Length and NPM1 and FLT3/ITD Mutation Status 161

5.13 Telomere Length and Status After 1st Cycle of Intensive Chemotherapy 162

5.14 Telomere Length and Disease-Free Survival 163-164

5.15 Telomere Length and Overall Survival 165-166

5.16 Discussion 167-171

Chapter 6: Telomere Dysfunction and its Potential Role in Promoting Genetic Instability in MDS and AML

6.1 Abstract 172-173

6.2 Introduction 174-176

6.3 MDS/AML and Telomere Fusion 177-181

6.4 Putative Telomere Fusion Events and Sequencing 182-193

6.5 The Development of 6q STELA 194-199

6.6 Array-CGH 200-202

6.7 Discussion 203-207

Chapter 7: Discussion

7.1 Telomere Length and Intra-Clonal Variation in MDS/AML 210

7.2 Telomere Length and Age in MDS/AML 211

7.3 Clinical Parameters in MDS/AML 212-213

7.4 Telomere Rapid Deletion and Telomere Fusion 213-214

7.5 Pure TTAGGG and Telomere Dysfunction 214-215

7.6 Mechanism of Telomere Fusion in MDS/AML 215-216

7.7 Complex Telomeric Fusion Events 216-217

7.8 Telomere Length and Clonal Expansion 217-218

7.9 The IPSS Scoring System for MDS Prognosis and Telomere Length 218

7.10 Telomere Length and Molecular Mutations in AML 219-220

7.11 Telomere Length and Survival Parameters 221

7.12 In Conclusion 222

7.13 Future Work and Implications 223-224

Appendix-6q family cluster 225-259

References 260-283

Acknowledgements

I am very thankful to Professor Duncan Baird and Professor Chris Pepper for their time spent reading the pages of my thesis. They have been particularly informative over the course of my PhD and have guided me all the way to the end.

I would also like to offer my regards to all of those who supported me in any respect during and up to the completion of my project.

I am immensely grateful to them all.

Abstract

The Myelodysplastic syndromes (MDSs) are comprised of a heterogeneous group of clonal disorders characterised by ineffective haematopoiesis. Although 30 to 35% of MDS cases progress to Acute Myeloid Leukaemia (AML), the majority of patients die from blood related ailments caused by progressive bone marrow failure. Large-scale genomic rearrangements are a key feature of MDS, with different aberrations conferring specific risks of progression. Telomere erosion, dysfunction and fusion, creating cycles of anaphase bridging breakage and fusion is a mechanism that has the potential to drive genomic instability in many tumour types including MDS. The key aim of this project was to examine the role that telomere dysfunction may play in the generation of genomic rearrangements observed in MDS/AML.

High resolution Single Telomere Length Analysis (STELA) revealed telomere shortening when compared to age-matched individuals in two cohorts of MDS and AML patients; this included large-scale telomeric deletion events observed within the MDS cohort. A PCR based telomere fusion assay detected telomere-telomere fusion events at a frequency that was consistent with sporadic fusion arising as a consequence of large-scale deletion. Telomerase activity was up-regulated in AML which may contribute to the reduction of deletion and fusion events in these cells.

Sequence analysis revealed that telomere fusion was associated with microhomology and sub-telomeric deletion; this profile was consistent with error-prone Ku-independent alternative end joining processes.

Telomere length at diagnosis irrespective of conventional markers appeared to influence the overall survival of MDS patients, but this was not apparent in AML. More importantly, telomere length was able to refine favourable prognostic markers, specifically good risk cytogenetics, uni-lineage cytopenia and low-risk IPSS (International Prognostic Scoring System) scores of which MDS patients bearing shorter telomeres for their respective age displayed reduced overall survival. This may be a particularly important finding given the heterogeneous clinical outcomes observed within low-risk MDS patients.

Abbreviations

- AA: Aplastic Anaemia
- ABL1: Abelson murine leukaemia viral oncogene homology 1
- a-CGH/Array-CGH: Array-Comparative Genomic Hybridisation
- ADP: Adenosine Diphosphate
- AKT2: Acute transforming retrovirus Thymoma
- ALL: Acute Lymphoblastic Leukaemia
- ALT: Alternative Lengthening of Telomeres
- AML: Acute Myeloid Leukaemia
- AML1: Acute Myeloid Leukaemia 1 protein
- AML-M0: Minimally Myeloid Differentiation
- AML-M1: Myeloblastic Leukaemia without Maturation
- AML-M2: Myeloblastic Leukaemia with Maturation
- AML-M3: Hypergranular Promyelocytic Leukaemia
- AML-M4: Myelomonocytic Leukaemia
- AML-M5: Monocytic Leukaemia
- AML-M6: Erythroleukaemia
- AML-M7: Megakaryoblastic Leukaemia
- A-NHEJ: Alternative-Non-Homologous End Joining
- ATLD: Ataxia-Telangiectasia Like Disorder
- ATM: Ataxia-Telangiectasia Mutated
- ATP: Adenosine Triphosphate
- ATR: Ataxia-Telangiectasia and Rad3-related
- BCL-2: B-cell Lymphoma 2
- BCR: Breakpoint Cluster Region
- BER: Base Excision Repair
- BFB: Breakage-Fusion-Bridge
- BIR: Break Induced Replication
- BLM: Bloom Syndrome
- BM: Bone Marrow
- BMI: Body Mass Index
- BMMNC: Bone Marrow Mononuclear Cells
- bp: Base pair
- BSA: Bovine Serum Albumin
- CBC: Complete Blood Count
- CBFβ: Core Binding Factor subunit β
- CD: Cluster of Differentiation

- CDC25A: Cell Division Cycle 25 homolog A
- CDK: Cyclin Dependent Kinase
- CDR: Critical minimally Deleted Region
- CHAPS: 3-[(3-cholamidopropyl) dimethylammonio]-1-propane sulphonate
- CHK-1: Checkpoint Kinase 1
- CI: Confidence Interval
- CLL: Chronic Lymphocytic Leukaemia
- CLP: Common Lymphoid Progenitor
- CML: Chronic Myeloid Leukaemia
- CMML: Chronic Myelomonocytic Leukaemia
- CMP: Common Myeloid Progenitor
- C-NHEJ: Classical/Canonical-Non-Homologous End Joining
- CTNNA1: Catenin (cadherin-associated protein) alpha 1
- DC: Dyskeratosis Congenita
- DDR: DNA Damage Response
- DFS: Disease-Free Survival
- DKC1: Dyskeratosis Congenita 1, dyskerin
- DNA-PKcs: DNA dependent Protein Kinase catalytic subunit
- dNTPs: Deoxynucleotide Triphosphates
- DOCK4: Dedicator Of Cytokinesis 4
- DSB: Double Strand Break
- EBV: Esptein-Barr Virus
- EDTA: Ethylenediaminetetraacetic acid
- EGR1: Early Growth Response 1
- ERCC1: Excision Repair Cross-Complementing Group 1
- ETO: Eight Twenty-One
- FAB: French-American-British
- FDA: Food and Drug Administration
- FISH: Fluorescence *in situ* Hybridisation
- Flow-FISH: Flow cytometry combined with FISH
- FLT3/ITD: FLT3 with Internal Tandem Duplication
- FLT3/TKD: FLT3 with Tyrosine Kinase Domain mutation
- FLT3: FMS-Like Tyrosine Kinase
- g: gram
- G6PD: Glucose 6 Phosphate Dehydrogenase
- GDP: Guanosine Diphosphate
- GM-CSF: Granulocyte-Monocyte Colony Stimulating Factor
- GTP: Guanosine Triphosphate
- H2AX: H2A histone family, member X
- HCl: Hydrochloric Acid

- HDM2: Human Double Minute 2
- HML: Hidden MAT Left
- HMR: Hidden MAT Right
- HOXD13: Homeobox D13
- HR: Hazard Ratio
- HSC: Haematopoietic Stem Cell
- hTERT: human Telomerase Reverse Transcriptase
- hTR: human Telomerase RNA component
- HUMARA: Human Androgen Receptor
- Int-1: Intermediate 1
- Int-2: Intermediate 2
- IPSS: International Prognostic Scoring System
- ISEL: *In situ* End Labelling
- KLP4: Kinesin-Like Protein 4
- Let-7: Lethal-7
- LOH: Loss Of Heterozygosity
- LT-HSC: Long Term-Haematopoietic Stem Cell
- M: Moles
- MAFβ: v-Maf avian Musculoaponeurotic Fibrosarcoma oncogene homolog β
- MAT: Mating Type
- MDSs: Myelodysplastic Syndromes
- MEFs: Mouse Embryonic Fibroblasts
- mg: milligram
- $MgCl_2$: Magnesium Chloride
- mins: Minutes
- ml: Millilitre
- MLL: Myeloid/Lymphoid or Mixed Lineage Leukaemia
- MLL2: Myeloid/Lymphoid or Mixed Lineage Leukaemia 2
- MLLT3: Myeloid/Lymphoid or Mixed Lineage Leukaemia (trithorax homolog, Drosophila); translocated to, 3
- mM: Millimole
- MMEJ: Microhomology Mediated End Joining
- MNC: Mononuclear Cells
- MPD: Myeloproliferative Disorder
- MPO: Myeloperoxidase
- MPP: Multipotent Progenitor
- MRE11: Meiotic Recombination 11
- MRN: Mre11/Rad50/Nbs1
- MRX: Mre11/Rad50/Xrs2
- MSi2: Musashi-2

- MYC/c-Myc: v-Myc avian Myelocytomatosis viral oncogene homolog
- MYHII: Myosin Heavy chain II
- NaCl: Sodium Chloride
- NADPH: Nicotinamide Adenine Dinucleotide Phosphate
- $NaHPO_4$ [pH 7.2]: Sodium Phosphate buffer
- NaOAc: Sodium Acetate
- NaOH: Sodium Hydroxide
- Nbs1: Nibrin
- NES: Nucleus Export Signal
- NF-κB: Nuclear Factor Kappa-B
- ng: nanogram
- NOD-SCID: Non-Obese Diabetic-Severe Compromised Immuno-Deficiency
- NPM1: Nucleophosmin
- $NPMc^+$: Cytoplasmic NPM1 mutant
- NRTs: Non-Reciprocal Translocations
- nt/nts: Nucleotides
- NuLS: Nucleolar Localisation Signal
- NUP98: Nucleoporin 98kDa
- OS: Overall Survival
- PAR: Psuedoautosomal Region
- PARP1: Poly (ADP-Ribose) Polymerase I
- PB: Peripheral Blood
- PBS: Phosphate Buffered Saline
- PCR: Polymerase Chain Reaction
- PD: Population Doubling
- PD-1: Programmed Cell Death 1
- pg: Picogram
- PGK: Phosphoglycerate Kinase
- PHA: Phytohaemagglutinin
- PML: Promyelocytic Leukaemia
- PNA: Peptide Nucleic Acid
- POT1: Protection Of Telomeres 1
- Q-FISH: Quantitative-Fluorescence *in situ* Hybridisation
- Q-PCR: Quantitative-Polymerase Chain Reaction
- RA: Refractory Anaemia
- RAC1: Ras-related C3 botulinum toxin substrate 1
- RAEB: Refractory Anaemia with Excess Blasts
- RAEB-T: Refractory Anaemia with Excess Blasts in Transformation
- RAP1: Repressor/Activator Protein 1
- RARS: Refractory Anaemia with Ringed Sideroblasts

- RARα: Retinoic Acid Receptor α
- RCMD: Refractory Cytopenia with Multi-lineage Dysplasia
- RCUD: Refractory Cytopenia with Uni-lineage Dysplasia
- RFLP: Restriction Fragment Length Polymorphism
- RMD: Regions of Minimal Deletion
- ROS: Reactive Oxygen Species
- RPS14: Ribosomal Protein S14
- RS: Ringed Sideroblasts
- SCID: Severe Compromised Immuno-Deficiency
- SD: Standard Deviation
- SDS: Sodium Dodecyl Sulphate
- Secs: Seconds
- SEN6: Senescence (cellular)-related 6
- SNP-A: Single Nucleotide Polymorphism Analysis
- SSA: Single Stand Annealing
- SSBR: Single Strand Break Repair
- SSC: Saline-Sodium Citrate
- STAT5: Signal Transducer and Activator of Transcription-5
- STELA: Single Telomere Length Analysis
- ST-HSC: Short Term-Haematopoietic Stem Cell
- TAE: Tris-acetate-ethylenediaminetetraacetic acid
- TDMs: Telomere DNA containing Double Minutes
- TE: Tris-EDTA buffer
- TEBPα/β: Telomere Binding Protein α/β
- TET2: Ten Eleven Translocation 2
- TIFs: Telomere dysfunction Induced Foci
- TIN2: TRF1 and TRF2 Interacting Protein 2
- TNFRI/II: Tumour Necrosis Factor Receptor I/II
- TNFα: Tumour Necrosis Factor α
- TRAP: Telomeric Repeat Amplification Protocol
- TRD: Telomere Rapid Deletion
- TRF: Terminal Restriction Fragment
- TRF1: TTAGGG Repeat Factor 1
- TRF2: TTAGGG Repeat Factor 2
- T-SCE: Telomere unequal Sister Chromatid Exchange
- TSG: Tumour Suppressor Gene
- TVR: Telomere Variant Repeats
- U: Units
- UPD: Uniparental Disomy
- WHO: World Health Organisation

- WRN: Werner
- WT: Wild-Type
- XPF: Xeroderma Pigmentosum group F
- XRCC1: X-ray Repair Complementing defective repair in Chinese hamster cells
- μg: Microgram
- μl: Microlitre
- μM: Micromole

Chapter 1: Introduction

Part 1: Myelodysplastic Syndromes (MDSs) and Acute Myeloid Leukaemia (AML)

1.1 Haematopoietic System

The haematopoietic system is currently displayed as a multi-step hierarchy governed by a primitive haematopoietic stem cell (HSC) (Figure 1.1). HSCs have the unique ability to both self-renew and differentiate in order to maintain haematopoietic homeostasis.[1,2] They are a heterogeneous pool of cells consisting of at least two functionally distinct HSC populations, long-term self-renewing (LT-HSCs) and short-term self-renewing (ST-HSCs). Whereas, LT-HSCs have life-long self-renewing potential, ST-HSCs are more restricted in self-renewal capacity.[3,4] A low frequency of HSCs exist with 2 to 5 HSCs per 10^5 total adult bone marrow cells[5] and such cells are thought to be enriched within the $Lin^-CD34^+CD38^-CD90^+CD45RA^-$ population of human marrow.[6] Following loss of self-renewal potential, HSCs give rise to Multipotent Progenitors (MPPs) which commit to either myeloid or lymphoid lineages.[7]

MPPs differentiate into common myeloid progenitors (CMPs) or common lymphoid progenitors (CLPs) that undergo further differentiation along their lineage to generate mature, functional peripheral blood cells. CLPs and CMPs show very limited or no self-renewal activity[8,9] and it has been proposed that the CLP and CMP populations represent the earliest branch points between lymphoid and myeloid lineages.[9] T and B lymphocytes and natural killer cells are released from the CLP pathway[8] whereas, platelets, erythrocytes, granulocytes and monocytes are derived from the CMP lineage.[9]

It has been previously demonstrated in young mice that the percentage of LT-HSCs that enter $S/G_2/M$ of the cell cycle is approximately 5% per unit time.[10] However, this fraction of cells increases significantly with age *in vivo*.[11] Quantitative and qualitative abnormalities of haematopoietic cells can arise under conditions of haematological stress. It has been proposed that such abnormalities might arise as part of the normal aging process where increased replication of haematopoietic cells may act in an attempt to compensate for the haematopoietic deficits that develop with age. Alternatively, an accumulation of genetic lesions with age might induce increased proliferation of HSCs and disrupt the regulation of

differentiation. Such processes have the potential to induce neoplastic transformation of HSCs. The increase in proliferation observed in older mice may share some relation to the higher incidence of leukaemia found in aging individuals.[11]

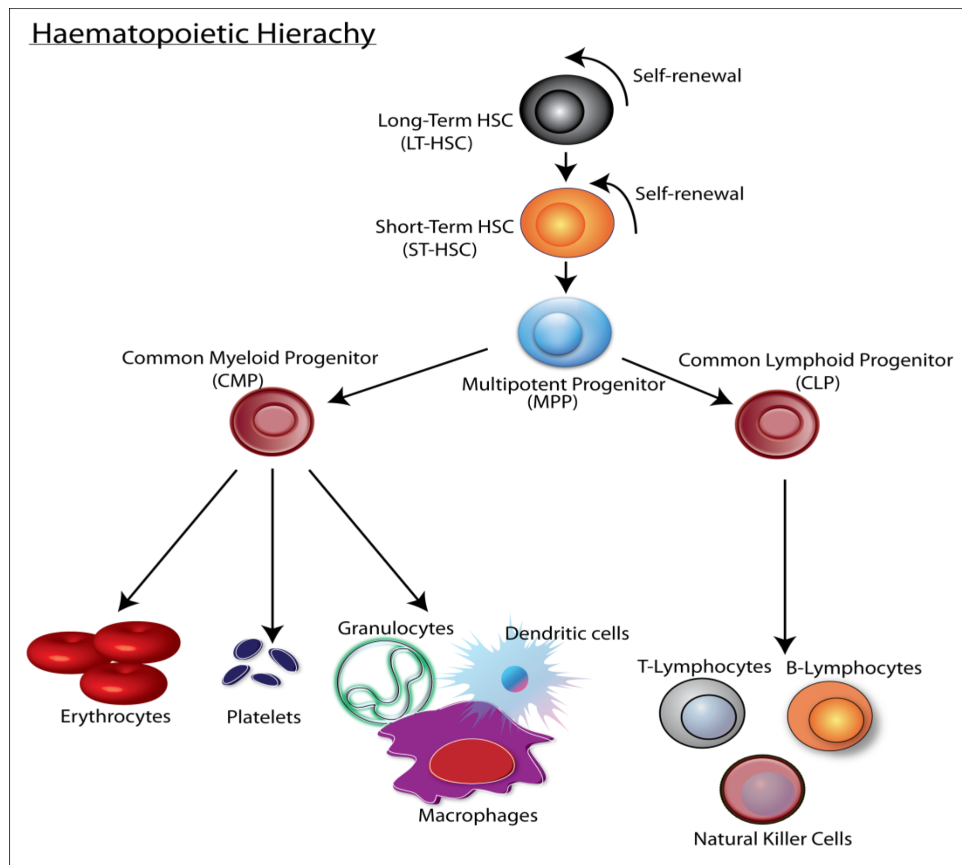


Figure 1.1: The haematopoietic system is currently displayed as a multi-step hierarchy governed by a primitive haematopoietic stem cell (HSC). Cells differentiate along their respective lineages in order to generate mature, functional blood cells.

Notably, haematological disorders including Chronic Lymphocytic Leukaemia (CLL), Chronic Myeloid Leukaemia, Acute Lymphoblastic Leukaemia (ALL), Myelodysplastic Syndromes (MDSs) and Acute Myeloid Leukaemia (AML) are common in the elderly population[12-16] and have been proposed to arise under such circumstances.[12,17]

1.2 MDS Pathology

The Myelodysplastic Syndromes (MDSs) are comprised of a heterogeneous group of clonal disorders with ineffective haematopoiesis.[18] It is considered to exist as a premalignant condition that has a 30 to 35% chance of transformation to Acute Myeloid Leukaemia (AML).[19]

According to the World Health Organisation (WHO) classification system for haematological cancers MDS is one of the five major categories of myeloid neoplasms[20] with an estimated incidence of 2 to 12 new cases per 100,000 people each year which has been noted to increase among persons aged 70 or older.[15]

The diagnosis of MDS is suspected from an abnormal complete blood count (CBC) but is confirmed following a bone marrow (BM) aspiration and biopsy. The BM aspirate allows for a detailed evaluation on cellular morphology and can evaluate the percentage of blast cells present in the marrow.[21]

The marrow cellularity is normal or hypercellular in 90% of MDS patients[22] but is hypocellular in 5 to 10% of cases.[23] Haematopoietic failure disrupts homeostasis resulting in cytopenias, i.e. peripheral blood cell counts lower than the expected range for the healthy population. Accordingly, differentiated cells and their precursors are either dysfunctional or are eradicated by apoptosis.[15] Consequently, dyserythropoiesis, dysgranulopoiesis and dysmegakaryocytopoiesis result in an insufficient production in erythrocytes, granulocytes and platelets along their respective lineage.[24] Patients often suffer with anaemia that is refractory to therapy, i.e. transfusion dependent and can become immune compromised increasing their chance of a recurrent infection. Patients may also haemorrhage more readily as a consequence of reduced platelet counts. MDS has an unpredictable course but a tendency to worsen overtime and can range from an indolent disease spanning years to a type that rapidly evolves to overt leukaemia.[25] Despite being a disease of the elderly, the majority of patients die as a consequence of blood cytopenias and not from age-related co-morbidity or AML progression.[26] Accordingly, it has been previously implicated that the majority of patients with low-risk disease (85%) die of MDS-related causes.[26] Infection, either caused by pneumonia or sepsis accounted for the majority of deaths (38%), whereas AML transformation arose in only 15% of patients.[26] Patients that endure co-lineage cytopenias show an increase in morbidity and reduced overall survival. Such individuals also have a reduced latency period prior to AML transformation.[27]

A subset of patients with MDS present with a hypocellular bone marrow.[23] Clinically it shares similar manifestations with normo/hypercellular MDS including cytopenias and bone marrow dyspoiesis.[28] However, it has been described as an independent parameter of

survival among low-risk MDS patients. Whereas, hypocellularity did not influence the overall survival within high-risk MDS patients, low-risk MDS patients presenting with hypocellularity showed longer overall survival when compared to patients presenting with normo/hypercellular MDS.[28]

Classification and Prognostic Scoring Systems for MDS/AML

Classification systems including those devised by the French-American-British (FAB) and World Health Organisation (WHO) are generally used as diagnostic tools which can be used to define specific disease entities of clinical significance. In contrast, patient prognosis is determined by disease-specific characteristics such as chromosomal abnormalities and haematopoietic insufficiency that are combined into a risk scoring system so as to predict patient outcome and facilitate in therapeutic decisions. Such risk scores include the International Prognostic Scoring System (IPSS) and the Revised-IPSS (IPSS-R) for MDS patients and the Hill's Scoring system for patients with AML.

1.3 French-American-British (FAB) Classification

The FAB criterion for the classification of the Myelodysplastic Syndromes (MDS) and Acute Myeloid Leukaemia (AML) was proposed around 30 years ago. MDS was divided into five subgroups (Table 1.1) based largely on the percentage of blasts in the bone marrow (BM) and peripheral blood (PB) and the presence or absence of ringed sideroblasts or increased circulatory monocytes.[24] In the FAB classification, AML was defined as a BM composed of >30% blasts.

Table 1.1: FAB classification of MDS

FAB Classification	PB Findings	BM Findings
Refractory Anaemia (RA)	<1% blasts	<5% blasts
RA with Ringed Sideroblasts (RARS)	<1% blasts	<5% blasts ≥15% ringed sideroblasts
RA with Excess Blasts (RAEB)	0 to 4% blasts	5% to ≤20% blasts
RAEB in Transformation (RAEB-T)	≥5% blasts	21% to ≤30% blasts
Chronic Myelomonocytic Leukaemia (CMML)	<5% blasts $>10^9$/l monocytes	1% to ≤20%

AML was divided into the subgroups M0 to M7[29-31] (Table 1.2) characterised by the maturation of the major cell lineage(s) involved:

Table 1.2: FAB classification of AML

Subtype	Note
M0: AML with Minimal Myeloid Differentiation	
M1: Myeloblastic Leukaemia without Maturation	Show various degrees of granulocytic differentiation
M2: Myeloblastic Leukaemia with Maturation	
M3: Hypergranular Promyelocytic Leukaemia	
M4: Myelomonocytic Leukaemia -[Granulocytic and monocytic differentiation]	
M5: Monocytic Leukaemia -[Monocyte differentiation]	
M6: Erythroleukaemia -[Erythrocyte differentiation]	
M7: Megakaryoblastic Leukaemia -[Megakaryocyte differentiation]	

1.4 The World Health Organisation (WHO) Classification

The WHO Classification relies on a combination of clinical, morphologic, immunophenotypic, genetic and other biologic features to define specific disease entities. The WHO criteria apply to initial diagnostic peripheral blood (PB) and bone marrow (BM) obtained prior to any definitive therapy for a suspected haematological neoplasm.

1.4.1 WHO Classification for the Myelodysplastic Syndromes (MDS)

At least 10% of cells derived from one myeloid bone marrow lineage i.e. erythroid, granulocytic or megakaryocytic must show dysplasia for the lineage to be considered as dysplastic.[32] However, causes of secondary dysplasia such as nutritional deficiencies, medications or infection should be excluded before a diagnosis of MDS can be confirmed.

In the WHO classification (Table 1.3), the blast threshold for the diagnosis of AML was reduced from 30% to 20% in the PB or BM which therefore eliminated the FAB category RAEB-T.[33] The FAB category of RAEB was also refined into RAEB-1 and RAEB-2 depending on the blast percentage in the blood and marrow and the presence or absence of Auer rods.[33] RAEB-1 has also been redefined to include patients who present with a 2 to 4% blast percentage in the blood even if there is less than 5% blasts in the marrow.[32] In contrast, patients with 5 to 19% blasts in the blood or 10 to 19% blasts in the bone marrow are categorised as RAEB-2, the highest grade of MDS.

Additionally, Chronic Myelomonocytic Leukaemia (CMML) was incorporated into a separate category termed 'the Myelodysplastic/Myeloproliferative Neoplasms' since it demonstrates clinical, laboratory and morphologic features associated with both a Myelodysplastic Syndrome (MDS) and a Myeloproliferative Neoplasm (MPN).[33] CMML has also been separated into two entities: CMML-1 and CMML-2 that can be distinguished by the percentage of blast cells in the marrow and peripheral blood. Promonocytes and <10% and <5% blasts in the BM and PB, respectively define CMML-1 whereas promonocytes and ≥10% blasts in the BM and ≥5% blasts in the PB describe CMML-2.

Table 1.3: WHO classification of MDS

MDS Subtype	PB Findings	BM Findings
Refractory Cytopenia with Uni-lineage Dysplasia (RCUD) - Refractory Anaemia (RA) - Refractory Neutropenia (RN) - Refractory Thrombocytopenia (RT)	- <1% blasts - Uni-/bi-cytopenia	- Unilineage dysplasia: ≥10 of cells in one myeloid lineage - <5% blasts - Ring sideroblasts account for less than 15% of erythroid precursors
Refractory Anaemia with Ring Sideroblasts (RARS)	- Anaemia - No blasts	- ≥15% of erythroid precursors are ring sideroblasts - Dyserythropoiesis only - <5% blasts
Refractory Cytopenia with Multi-lineage Dysplasia (RCMD)	- Cytopenia(s) - No or rare blasts - No Auer rods - <1x 10^9/l monocytes	- Dysplasia in ≥10% of cells in 2 or more lineages - <5% blasts - No Auer rods

Refractory Anaemia with Excess Blasts-1 (RAEB-1)	- Cytopenia(s) - <5% blasts - No Auer rods - <1x 10^9/l monocytes	- Uni-lineage or multi-lineage dysplasia - 5 to 9% blasts - No Auer rods
Refractory Anaemia with Excess Blasts-2 (RAEB-2)	- Cytopenia(s) - 5 to 19% blasts - Auer rods - <1x 10^9/l monocytes	- Uni-lineage or multi-lineage dysplasia - 10 to 19% blasts - Auer rods
Myelodysplastic Syndrome associated with isolated del(5q): "the 5q-Syndrome" - Favourable outcome - Low incidence of AML	- Anaemia - Normal/ elevated platelets - No or rare blasts	- Normal to increased megakaryocytes with hypolobated nuclei - <5% blasts - del(5q) is the sole cytogenetic abnormality - No Auer rods
Myelodysplastic Syndrome, Unclassifiable (MDS-U)	- Cytopenia - <1% blasts	- Unequivocal dysplasia in <10% of cells in one or more myeloid lineages when accompanied by a cytogenetic abnormality considered as presumptive evidence for a diagnosis of MDS (Table 1.2). - <5% blasts

Clonal cytogenetic abnormalities occur in about 50% of MDS cases.[34,35] If a patient presents with persistent cytopenia in the absence of conclusive morphologic features, a presumptive diagnosis of MDS can be made if a specific clonal chromosomal abnormality is present[32] (Table 1.4).

Table 1.4: A presumptive diagnosis of MDS can be made if a specific clonal chromosomal abnormality is present.

Unbalanced Abnormalities	Balanced Abnormalities
-7 or del(7q)	t(11;16)(q23;p13.3)
-5 or del(5q)	t(1;3)(p36.3;q21.2)
-13 or del(13q)	t(3;21)(q26.2;q22.1)
i(17q) or t(17p)	t(2;11)(p21;q23)
del(11q)	t(6;9)(p23;q34)
idic(X)(q13)	inv(3)(q21;q26.2)
del(9q)	
del(12p) or t(12p)	

Diagnosis can be problematic for patients who present MDS with hypocellularity (h-MDS) which arises in ~10% of adult MDS.[36] When the marrow is normal or hypercellular and dysplasia is detected, myelodysplasia can be distinguished from Aplastic Anaemia (AA).[37] However, in cases where the bone marrow cellularity is low (<20%),[36,38] an accurate morphological analysis may be difficult to perform because of inadequate material from hypoplastic specimens.[36,39] The majority of h-MDS cases are classed as Refractory Anaemia (RA) and such characteristics including an absence of increased blast count and mild dyserythropoiesis can overlap with what is seen in AA.[36] Clonal cytogenetic abnormalities are usually considered diagnostic of MDS but 50% of cases present with a normal karyotype.[37,39] Additionally, cytogenetic testing may be less reliable because of a low number of cells[37,39] making it more difficult to differentiate between these disorders.

1.4.2 WHO Classification for Acute Myeloid Leukaemia (AML)

In the WHO scheme, a myeloid neoplasm with 20% or more blasts in the PB or BM is considered to be AML whether it arises *de novo*, in the setting of a previously diagnosed Myelodysplastic Syndrome (MDS), Myelodysplastic/Myeloproliferative neoplasm (MDS/MPN), or blast transformation in a previously diagnosed Myeloproliferative neoplasm (MPN), such as Primary Myelofibrosis, Polycythemia Vera or Essential Thrombocythemia. It may also occur following therapy to a non-haematological malignancy.[32,33] However, a diagnosis of AML can be made regardless of the blast percentage in such cases associated with specific genetic abnormalities, i.e. t(8;21)(q22;q22); *RUNX1-RUNX1T1*, inv(16)(p13.1;q22) or t(16;16)(p13.1;q22); *CBFβ-MYH11* and t(15;17)(q22;q21); *PML-RARα*.[33] Table 1.5 represents the WHO classification of Acute Myeloid Leukaemia (AML).

Table 1.5: WHO classification of AML

Acute Myeloid Leukaemia (AML) with Recurrent Genetic Abnormalities (Variant MLL translocations should be specified at diagnosis since over 80 partner genes can participate in the translocation with MLL therefore resulting in variable biological characteristics)	- AML with t(8;21)(q22;q22); *RUNX1-RUNX1T1* - AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFβ-MYH11* - Acute Promyelocytic Leukaemia (APL) with t(15;17)(q22;q21); *PML-RARα* - AML with t(9;11)(p22;q23); *MLLT3-MLL* - AML with t(6;9)(p23;q34); *DEK-NUP214* - AML with inv(3)(q21;q26.2) or t(3;3)(q21;q26.2); *RPN1-EVI1* - AML (megakaryoblastic) with t(1;22)(p13;q13); *RBM15-MKL1* - AML with mutated *NPM1* (Provisional entity) - AML with mutated *CEBPA* (Provisional entity)
Acute Myeloid Leukaemia (AML) with Myelodysplasia-related changes (AML-MRC)	- ≥ 20% blasts with, - Morphologic dysplasia in ≥ 50% of at least 2 myeloid lineages or, - A history of MDS or MDS/MPN or - with MDS related cytogenetic abn (Table 1.6)

Therapy-related Myeloid Neoplasms (t-AML/t-MDS) and (t-AML/t-MDS/MPN)	Occurring as a late complication of cytotoxic chemotherapy and/or radiotherapy

Acute Myeloid Leukaemia, Not Otherwise Specified (NOS)
 	
Cases that do not fulfil the WHO criteria of the other AML categories. Account for 25% to 30% of all cases. However, this group will continue to reduce with the recognition of more genetic subgroups.
	- 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 Leukaemia
	- Acute Panmyelosis with Myelofibrosis
	- Myeloid Sarcoma

Acute Leukaemia of Ambiguous Lineage
 	
Show no clear evidence of differentiation along a single lineage.
	- Undifferentiated Acute Leukaemia (AUL)
	  - Cases with no lineage specific markers
	- Mixed Phenotype Acute Leukaemia (MPAL)
	  - Blasts co-express certain antigens of more than one lineage on the same cell or that have separate populations of blasts that are of different lineages

It has been argued that 90% of patients with therapy-related disease share cytogenetic abnormalities with those observed in 'AML with Myelodysplasia-related changes' or in 'AML with recurrent cytogenetic abnormalities' and therefore could be more appropriately classified into those categories,[32] however patients with therapy-related myeloid neoplasms have significantly worse outcomes than their *de novo* counterparts with the same genetic abnormality suggesting that that there are biological differences.[40-42]

Table 1.6: The cytogenetic abnormalities sufficient for a diagnosis of AML with myelodysplasia-related changes when 20% or more BM and PB blasts are present.

Unbalanced Abnormalities	Balanced Abnormalities
-7 or del(7q)	t(11;16)(q23;p13.3)
-5 or del(5q)	t(3;5)(q25;q34)
del(11q)	t(5;7)(q33;q11.2)
-13 or del(13q)	t(2;11)(p21;q23)
i(17q) or t(17p)	t(5;12)(q33;p12)
del(12p) or t(12p)	t(1;3)(p36.3;q21.1)
del(9q)	t(5;17)(q33;p13)
idic(X)(q13)	t(5;10)(q33;q21)
	t(3;21)(q26.2;22.1)

1.5 International Prognostic Scoring System (IPSS) for MDS

Patient outlook can be predicted using the IPSS[43] or the more recent IPSS-R[44] scoring system and facilitate in making therapeutic decisions.

The prognostic outlook of MDS patients can be determined using the IPSS scoring system devised in 1997.[43] It relies on three major variables including bone marrow blast percentage (BM blast %), the number of cytopenias and karyotypic complexity with which can be divided into Good, Intermediate or Poor (Table 1.7). Cytopenias are defined as a haemoglobin level of under 10g/dl, an absolute neutrophil count (ANC) of less than 1800/μl and a platelet count of less than 100,000/μl.

Table 1.7: Chromosomal abnormalities associated with a good, intermediate or poor risk in the IPSS.

Good	Normal, del(5q) only, del(20q) only, -Y only
Intermediate	Other abnormalities
Poor	Complex Abnormalities (≥ 3 anomalies), chromosome 7 abnormalities

By combining the risk scores listed in Table 1.8, patient outlook can be stratified into four prognostic subgroups: Low, Intermediate-1, Intermediate-2 or High (Table 1.9). Overall survival decreases in advanced subgroups with low-risk patients showing prolonged survival. Moreover, Low and Int-1 subgroups could be further refined based on patient age. Notably,

individuals 60 years or less showed improved overall survival in contrast to older patients. High-risk patients show increased mortality as a consequence of Acute Myeloid Leukaemia (AML) development whereas low-risk patients more likely die of complications associated with bone marrow failure. In contrast to high-risk patients, leukaemic evolution was prolonged in patients that did develop AML in lower risk groups.

Table 1.8: The severity of each prognostic variable gives a score value which is later summed to predict patient outlook.

Prognostic variable	Score				
	0	**0.5**	**1.0**	**1.5**	**2.0**
BM blast %	<5	5-10	-	11-20	21-30
Karyotype	Good	Intermediate	Poor	-	-
Cytopenia	0/1	2/3	-	-	-

Table 1.9: The total value of the risk scores predicts patient outlook which varies from low, int-1, int-2 or high risk

Prognostic Risk Group	Combined Risk Score
Low	0
Intermediate-1 (Int-1)	0.5-1.0
Intermediate-2 (Int-2)	1.5-2.0
High	≥2.5

1.6 Revised-International Prognostic Scoring System (IPSS-R) for MDS

As in IPSS, the cytogenetic subgroup, bone marrow blast percentage and cytopenia remain prognostically relevant within the revised IPSS score (IPSS-R),[44] however patient outlook has been further refined by incorporating novel chromosomal abnormalities (Table 1.10) and by analysing the depth of blood cell cytopenia. Accordingly, instead of the number of cytopenias, i.e. 1 to 3; haemoglobin, platelet or absolute neutrophil counts were stratified using relative cut-off points as listed in Table 1.11.

Table 1.10: Chromosomal abnormalities associated either very good to very poor risk scores in the revised-IPSS.

Very Good	-Y, del(11q)
Good	Normal, del(5q), del(12p), del(20q), double anomalies including del(5q)
Intermediate	del(7q), +8, +19, i(17q), any other single or double independent clones
Poor	-7, inv(3)/t(3q)/del(3q), double anomalies including -7/del(7q), complex: 3 abnormalities
Very Poor	Complex: >3 abnormalities

The IPSS-R further refined bone marrow blast percentage (Table 1.11). Patients presenting with 0% to ≤2% bone marrow blasts showed prolonged overall survival and time to AML evolution when compared to patients presenting with >2% to <5% blasts, thereby refining the IPSS blast category of <5% blasts. It was also observed that clinical outcomes in terms of overall survival and AML evolution were similar in individuals who presented with >10% to ≤20% bone marrow blasts versus 21% to ≤30% blasts therefore the newly revised IPSS-R combined these two parameters into >10% to ≤30%.

Table 1.11: The severity of each prognostic variable gives a score value which is later summed to predict patient outlook. Haemoglobin, platelet or absolute neutrophil counts have been stratified using relative cut-off points and blast counts have been further refined in the IPSS-R.

Prognostic Variable	**Score**						
	0	**0.5**	**1.0**	**1.5**	**2.0**	**3.0**	**4.0**
Cytogenetics	Very Good	-	Good	-	Intermediate	Poor	Very Poor
BM Blast %	≤2	-	>2 to <5	-	5 to 10	>10	-
Haemoglobin (g/dL)	≥10	-	8 to <10	< 8	-	-	-
Platelets (10^9/L)	≥100	50 to <100	<50	-	-	-	-
ANC (10^9/L)	≥0.8	<0.8	-	-	-	-	-

Thus, using additional features the revised-IPSS scoring system could be differentiated into 5 prognostic subgroups (Table 1.12) for overall survival and AML evolution: Very Low, Low, Intermediate, High and Very High. Moreover, the IPSS-R could be further adjusted for

patient age categorising patients based on the IPSS-RA scoring system, however this was only applicable for overall survival and not AML evolution. Accordingly, the median age of the patient cohort from which the prognostic risk score categories were calculated from was 70 years, therefore risk scores could be age-adjusted using the following formula: (Years - 70) x [0.05 - (IPSS-R Risk Score x 0.0005)], added onto the IPSS-R score. Notably, younger individuals showed prolonged overall survival in contrast to older patients in which overall survival reduced with aging.

Table 1.12: The total value of the risk scores predicts patient outlook. Five subgroups were devised in the IPSS-R from very low to very high risk.

Prognostic Risk Group	Combined Risk Score
Very Low	≤1.5
Low	>1.5 to 3
Intermediate	>3 to 4.5
High	>4.5 to 6
Very High	>6

1.7 Therapeutic Options in MDS

The pathogenesis and prognostic outlook of MDS among the population is very diverse, thus hindering the choice of therapeutic options. A large fraction of MDS patients receive supportive care including transfusions or growth factors for cytopenias, such as erythropoietin or granulocyte stimulating factor rather than a disease-specific therapy. However, cytogenetic analysis has facilitated in predicting the patient's risk of AML transformation and provides a basis for drug selection.[21] Thus, patients initially diagnosed with a lower risk MDS may be identified as having a poorer outcome and might benefit from early therapeutic intervention. Furthermore, patients who are ineligible for transplantation, such that they may be of an unfavourable age, may benefit from an MDS-specific therapy. Accordingly, the FDA has approved three such treatments for use in the USA. These consist of two hypomethylating drugs decitabine[45] and 5-azacitidine[46] and the thalidomide derivative lenalidomide.[47] Inactivation of tumour suppressor genes (TSG) by promoter hypermethylation can be reversed during DNA synthesis. Decitabine is a cytosine nucleoside analogue that can inhibit DNA methylation when incorporated into DNA, thus reactivating

the TSG.[45] Transfusion independence,[45] a significant increase in progression-free survival and reduced AML transformation rate have been observed in a phase III randomised study comparing decitabine with supportive care only.[45] Improvements in the quality of life have also been documented.[48] When undergoing a phase III trial, the DNA methyl-transferase inhibitor, 5-azacitidine has demonstrated prolonged patient survival and a reduced risk of AML transformation in higher risk patients.[46] Lenalidomide is the third drug that has been approved in the USA. This drug works favourably with patients that endure the chromosomal abnormality 5q31 deletion since it has shown a selective cytotoxicity towards the del(5q) clone.[47] It has been observed that the drug is able to suppress the del(5q) clone restoring transfusion independence in this group of patients. Although accepted in the USA, Lenalidomide has not been approved in Europe due to the frequency of treatment-related AML transformation.[21] Accordingly, a third of patients who are refractory to treatment have a high risk of AML progression. These patients often develop complex karyotypes as a result of genetic instability.[49]

1.8 Prognostic Scoring in Acute Myeloid Leukaemia

Patients who are eligible to receive standard induction chemotherapy for Acute Myeloid Leukaemia (AML) are treated under the “7 + 3” regimen that includes 7 days of cytosine arabinoside (Ara-C) and 3 days of anthracycline.[50] However, this excludes patients diagnosed with Acute Promyelocytic Leukaemia (APL) who specifically receive a combination of anthracycline and all-*trans* Retinoic Acid (ATRA). The presence of *PML/RAR*α in APL cells denotes sensitivity to the differentiation inducing agent ATRA.[51] This subset of patients has a favourable prognosis with sustained long term remission and excellent overall survival.[52]

Following the first course of induction chemotherapy, therapeutic management depends on variable clinical parameters which assess the patient’s response to treatment and risk of relapse in CR (complete remission). Such parameters include age, leukaemia cytogenetics (Table 1.14), and response status after course 1, presenting white blood cell (WBC) count and AML type (*de novo*/ secondary AML). Secondary AML can either follow prior cytotoxic chemotherapy or radiotherapy for other cancers or arise subsequently to an antecedent haematological disorder.[53] The response status after the first course of induction chemotherapy is categorised as:

- CR [complete remission; BM is regenerating normal haematopoietic cells and contains <5% blast cells. An absolute neutrophil count of >1x10^9/l and platelet count of ≥100x10^9/l]
- PR [partial remission; BM is regenerating normal haematopoietic cells and blast count has reduced by at least half to a value between 5 and 25% leukaemic cells]
- RD [resistant disease; BM shows persistent AML and patient survives at least 7 days beyond the end of course 1]

Cox Regression analysis has been undertaken on patients derived from the Medical Research Council (MRC) AML trials 10 and 12 to provide a number of weighted factors which could be used to define patients as good, standard or high-risk. Table 1.13 shows how this index can be calculated:

Table 1.13: Patient outlook can be calculated using a number of weighted factors that calculate a risk score

Index = **0.01325***Age (in years) + **0.16994***Sex (1=Male, 0=Female) + **0.22131***Diagnosis(1=*De novo*, 2=Secondary) + **0.65082***Cytogenetics (1=Favourable, 2=Intermediate, 3=Adverse) + **0.19529***Status Post C1 (1=Complete Response, 2=Partial Response, 3=No Response) + **0.00169*** WBC Count (x10^9/l)

Defined cut-off points for dividing patients into good, standard or high-risk are arbitrary since outcome probabilities are produced as a continuum.[53,54] Patients with low-risk AML may continue induction chemotherapy for a further 3 or 4 cycles as a curative treatment (consolidation chemotherapy).[55] Conversely, high-risk patients may be eligible candidates for an allogeneic or autologous Stem Cell Transplant (SCT) provided that a suitable donor is available.[50] If AML cases fail to respond to conventional chemotherapy (resistant AML), patients may be offered alternative or investigational treatment.[53]

Elderly patients (>70 years) who do receive intensive chemotherapy show poor 5 year survival rates of less than 10% in contrast to over 50% of cases in children.[54] Notably, older individuals are more likely to show a poorer tolerance to chemotherapeutic drugs. However, secondary AML arising from an antecedent haematological disorder is more prevalent in the elderly[16] and patients commonly present with unbalanced and complex karyotypes[16,56] including abnormalities of chromosomes 5, 7 and 17.[57]

Cytogenetic and Molecular Genetic Characteristics in AML

The key determinant in influencing patient outcome is the diagnostic karyotype, of which ~60% of patients present with an abnormal karyotype.[58] Specific biological entities of AML that include APL with t(15;17)(q22;q21), AML with t(8;21)(q22;q22) and AML with inv(16)(p13.1;q22)/t(16;16)(p13.1;q22) can be treated using tailored therapy with a relatively favourable prognosis.[52,55] These individuals show low relapse rates in complete remission (CR) are therefore unlikely candidates for bone marrow transplantation. Conversely, AML patients that present with an adverse karyotype have a very poor prognosis with conventional chemotherapy[58] and are therefore considered for a bone marrow transplant (Table.1.14).

Table 1.14: Chromosomal abnormalities which define favourable, intermediate or adverse cytogenetics

Favourable	Irrespective of the presence of additional cytogenetics: • t(8;21)(q22;q22) • inv(16)(p13.1;q22)/ t(16;16)(p13.1;q22) • t(15;17)(q22;q21)
Intermediate	• Normal karyotype • Structural or numerical changes not encompassed by favourable/adverse risk groups
Adverse	• [abn(3q)] • del(5q)/-5 • -7 • Complex Karyotype: ≥ 5 unrelated abnormalities

The mutation status of specific genes can also influence patient outcome such that their identification may further refine patient prognosis, particularly within patients that present with a normal karyotype detected amongst 40% of AML cases.[58] Such molecular markers associated with AML include *FLT3*, *NPM1*, *CEBPA* and *KIT*.

The FLT3 (FMS-like tyrosine kinase 3) receptor plays a role in the survival, proliferation and differentiation of haematopoietic cells.[59,60] Mutations in the receptor, in the form of an

internal tandem duplication (ITD) of the juxtamembrane domain and point mutations within the tyrosine kinase domain (TKD) both result in its constitutive activation.[61] Tandem duplications are thought to disrupt the interaction between the juxtamembrane domain and the activation loop destabilising the inactive configuration of the kinase. The conformational change causes cytokine independent proliferation of haematopoietic cells.[62] Mutations in the TKD alter the configuration of the activation loop to enable increased access to ATP and substrates to the kinase.[62] The ITD mutation has been detected in 25% of AML cases, whereas the TKD mutation has been detected in 5 to 10% of cases.[62,63] FLT3/ITD has been frequently documented in AML cases with an inferior outcome[64,65] whereas the prognostic impact of the FLT3/TKD mutation is less clear. [60,66-68]

NPM1 functions as a molecular chaperone that shuttles between the nucleus and cytoplasm.[69] It is predominantly nucleolar, however 30% of AML patients bear a mutated NPM1 resulting in its aberrant localisation in the cytoplasm (NPMc^{+}).[70,71] The NPMc^{+} mutation is a marker of a favourable prognosis[72] however, these patients regularly harbour the FLT3/ITD and thus the favourable outcome is diminished in these patients.[73,74]

CEBPα (CCAAT/enhancer binding protein alpha) is a transcription factor that is essential for granulocytic development.[75] Loss of CEBPα function in myeloid cells causes a block in granulocytic differentiation.[75,76] Mutations in the CEBPα occur in approximately 15% of cytogenetically normal AML cases and can present as either biallelic or monoallelic mutations. Biallelic mutations frequently involve a combination of an N-terminal frame-shift mutation on one allele and a C-terminal in-frame mutation on the other which result in protein truncation and an impairment of DNA binding activity, respectively. In contrast, a monoalleic mutation presents with either an N or C-terminal mutation. A more favourable prognosis has been indicated in cytogenetically normal AML patients that present with the biallelic mutation. Patients with monoallelic CEBPα show a similar outlook to patients with wild-type CEBPα. Although monoallelic CEBPα is commonly associated with additional mutations, i.e. NPM1 and FLT3 ITD/TKD, biallelic CEBPα continues to be associated with improved prognosis with no difference between monoallelic and wild-type CEBPα following the exclusion of these concurrent abnormalities.[77] However, the prognostic significance of

monoallelic CEBPα is controversial since it has been recently reported to possibly confer a favourable prognosis in patients with both wild-type FLT3/ITD and wild-type NPM1.[75]

Favourable risk groups may be further stratified into prognostic subgroups with a less favourable outcome. Accordingly, AML patients presenting with Core-Binding Factor (CBF) AMLs i.e. AML with inv(16) or AML with t(8;21) have an adverse prognosis in the presence of a *KIT* mutation. This subset of patients have a 6-fold increase in relapse in the first CR when compared to CBF AML without the *KIT* mutation.[78]

1.9 Therapeutic Studies in AML

It has been reported that the addition of a purine analogue, Cladribine to standard induction therapy can increase complete remission rates and improve overall survival when compared with induction chemotherapy alone. Haematological and non-haematological toxicity were comparable among treatment groups. Moreover, in contrast to induction therapy alone, the addition of Cladribine achieved complete remission and improved overall survival in patients presenting with adverse karyotypes, higher initial white blood cell counts and aged over 50.[79]

Phase II trials have demonstrated that Clofarabine, a second generation purine analogue is well tolerated in older adults with AML who are considered unfit for conventional chemotherapy. In contrast to patients treated with low dose Ara-C (LDAC), patients treated with Clofarabine as a single agent show improved complete remission. However, despite improved remission rates survival was inferior in patients with refractory AML treated with Clofarabine (107 days; LDAC vs. 60 days; Clofarabine) and in patients with relapsed disease (40 weeks; LDAC vs. 20 weeks; Clofarabine).[80] Yet Clofarabine has been shown to achieve improved complete remission in individuals presenting with adverse cytogenetics (44%; Clofarabine vs. 0%; LDAC) and secondary AML (31%; Clofarabine vs. 4%; LDAC).[81]

Pre-treatment with DNA-hypomethylating agents prior to the standard “7+3” induction has been demonstrated to increase the efficacy of induction chemotherapy in AML. It was proposed that the inactivation of TSGs by aberrant DNA methylation during carcinogenesis may contribute to the resistance of leukaemic cells to cytotoxic treatment. In a phase I trial, AML patients (median age 55 years) with a less than favourable risk were pre-treated with

the hypomethylating agent Decitabine.[82] 83% of patients achieved complete remission after two cycles of induction chemotherapy and 53% were still alive after a median 32 month follow-up. It has been proposed that Decitabine may act as a chemosensitiser and complement the cytotoxic effects of standard induction chemotherapy by reactivation of TSG expression.

Allogeneic-Haematopoietic Stem Cell Transplantation (Allo-HSCT) is a widely used approach for a treatment of advanced AML and high-risk MDS however, only 20 to 30% of patients with high-risk AML become long-term survivors after a BMT with the most common causes of treatment failure including relapse, non-relapse mortality (NRT) and Graft vs. Host Disease (GVHD). Conventional preparative regimens for Allo-HSCT are often high-dose and thus older patients or those with attendant co-morbidities are ineligible candidates due to treatment associated complications. Patients presenting with >5% blasts in the marrow prior to Allo-HSCT conditioning show relapse rates greater than 50% when treated with a reduced intensity regimen composed of Total body irradiation (TBI) and Fludarabine.[83] Efforts to decrease relapse rates have been focused on therapy intensification, such as increasing TBI dosage. Although this has been successful in reducing relapse rates, non-relapse mortality is escalated as a consequence of surpassing normal organ tolerability.[84] A phase I trial of targeted haematopoietic irradiation with ^{131}I-labelled anti-CD45 antibodies (^{131}I-BC8 Ab) has been demonstrated to reduce relapse rates to 40% in elderly patients presenting with advanced AML or high-risk MDS in the marrow prior to Allo-HSCT conditioning.[85] Of note, 86% of patients in this study presented with >5% blasts at the beginning of the conditioning regimen. The one year survival estimate of the entire cohort in this trial was 41%, among those 46% presented with AML in remission, 46% with relapsed AML, 38% with refractory AML and 33% with high-risk MDS prior to Allo-HSCT conditioning.

1.10 MDS Cell of Origin

Cytogenetic abnormalities associated with the neoplastic clone are often observed in multiple myeloid lineages including peripheral blood granulocytes, monocytes and erythrocytes. Accordingly, it has been assumed that the primary neoplastic event originates in a committed myeloid progenitor, particularly since MDS rarely transforms into Acute Lymphoblastic Leukaemia (ALL).[19,86] However, the clonal involvement of non-myeloid cells

has been detected in a subset of patients raising the possibility that the initial transformation event can occur in a more primitive stem cell with multi-lineage potential.[34] The apparent myeloid lineage restriction could be a consequence of genetic (and epigenetic) abnormalities that have developed in a HSC causing suppression of lymphoid differentiation and providing a false observation.[19,87] Alternatively, a sustainable lymphocytic population may arise from a long-lived lymphoid progenitor generated before the cytogenetic abnormality occurred.[88] It has also been postulated that an efficient compensatory mechanism from a low fraction of normal stem cells might be supporting T- and B-cell production.[19]

X-chromosome inactivation may provide information in determining cellular clonality in female patients on condition that constitutional skewing of X-inactivation is excluded.[89] Cells derived from the same progenitor would retain the X chromosome inactivation pattern, thus this population of cells would have a monoclonal distribution. Conversely, a polyclonal pattern of clonality would be established if cells were derived from alternate progenitors.[90,91] The digestion of the un-methylated X-chromosome at heterozygous loci, i.e. HUMARA (Human Androgen Receptor) or PGK (Phosphoglycerate Kinase) can provide information on the clonal nature of haematopoiesis by means of visualising the clonality of the inactive X chromosome by Restriction Fragment Length Polymorphism (RFLP) analysis. Monoclonal X-inactivation patterns of PGK have been detected in the bone marrow, granulocytic and T-lymphocytic fractions of the peripheral blood in MDS patients that show a polyclonal X-inactivation pattern in corresponding skin tissue.[90] RFLP analysis of the HUMARA locus from sorted haematopoietic cells demonstrated a monoclonal distribution within $CD34^{+}CD38^{-}$, $CD34^{+}CD38^{+}$ and in mature myeloid cells. Although a polyclonal pattern of T and B-lymphocytes was detected in the majority of patients, a CMML and RAEB patient showed monoclonality of the B-lymphocyte and T-lymphocyte populations, respectively suggesting the clonal involvement of an MDS precursor common to both myeloid and lymphoid lineages. The identification of a polyclonality in T-lymphocytes derived from the CMML patient is suggestive of an MDS precursor that is common to both myeloid and B-lymphocytic lineages.[92] Accordingly, single lymphohaematopoietic progenitors have demonstrated their ability to yield progeny committed to either the myeloid or B-lymphoid lineages *in vitro*.[93,94]

FISH (Fluorescence *in situ* Hybridisation) analysis has demonstrated the clonal nature of common chromosome abnormalities within purified bone marrow cells. The 5q deletion has been previously detected in the vast majority of cells derived from the $CD34^+CD38^-$ (94 to 98%) and $CD34^+CD38^+$ (>88%) fractions. Although commonly detected within myeloid progenitors, the deletion was also detected in 25 to 90% of pro-B and 98% of pro-T-cell progenitors within three of five investigated patients and in one case respectively. This is consistent with the initial transformation event arising in a cell with multi-potent potential.[19] Furthermore, patient-derived $CD34^+CD38^-$ cells failed to reconstitute haematopoiesis *in vitro* and *in vivo* in contrast to normal controls.[19]

A high percentage of pluripotent stem cells ($CD34^+$Thy-1^+), pro-B cell progenitors ($CD34^+CD19^+$) and T/natural killer progenitor cells ($CD34^+CD7^+$) have been observed to bear an isolated monosomy 7 whereas its detection was below the cut-off value in T-and B-lymphocytes. However, 60% of natural killer cells retained the monosomy 7 suggesting that T-and B-lymphocyte progenitors positive for the aneuploidy may only undergo limited differentiation.[95] This is consistent with previous observations that have identified the 5q deletion in pro-B and pro-T cell progenitors but not within mature T- and B-lymphocytes.[19] However, *in vitro* expansion may facilitate in the detection of a minor monoclonal lymphocyte pool. Accordingly, a monoclonal pattern of the X-linked Glucose-6-phosphate dehydrogenase (G6PD) has been detected in Esptein-Barr virus (EBV) transformed B-lymphocytes. The B-lymphoblastoid cell lines carried the identical isoenzyme that was detected in myeloid cells.[96] Furthermore, clonal chromosomal markers on B-lymphoblastoid cell lines have shown the presence of an identical 20q deletion to that observed in myeloid cells.[97] Although this indicates the involvement of a cell with both myeloid and lymphoid potentiality, Phytohaemagglutinin (PHA)-stimulated cells did not show the 20q deletion. Although a deletion at 20q may prevent T-lymphocyte differentiation, it has also been postulated that a close relationship may exist between B-lymphocytes and myeloid cells.[97]

1.11 AML Cell of Origin

It has been proposed that two models may provide insight into the development of the AML clone. The first model implicates that the initial event originates within a committed progenitor in which the phenotype of the leukaemic blasts is dependent on the degree of

differentiation. Thus, the degree of commitment influences the AML FAB characteristic.[98] However, for a committed progenitor that has lost its ability to self-replicate; leukaemic transformation would need to acquire further genetic changes that are already intrinsic to the HSC.[1,4] The alternative model proposes that leukaemic transformation arises within a primitive stem cell and the characteristic of that genetic event determines the pathway of differentiation.[98] This concept may explain the absence or appearance of lymphoid differentiation in a subset of AML patients.

The transplantation of AML cells derived from donor patients into NOD-SCID (Non-obese Diabetic/Severe Combined Immunodeficiency) mice has demonstrated that the AML cell population exists as a hierarchy which is comparable to that found within normal bone marrow. A small population of primitive $CD34^{++}CD38^{-}$ (0.2% of the total leukaemic cell population) can successfully engraft AML cells in NOD-SCID mice and resemble the differentiation pathway specific to the patient donor FAB subtype.[98] Furthermore, these cells had the capability to engraft human cells to the equivalent level following serial transplantation in secondary recipients.[98] Persistent and transient leukaemic clones were established throughout serial transplantation. While it was established that persistent clones have a long-term repopulation capacity, which rarely commit along a lineage, transient clones were concluded to have a short-term repopulation capacity and commit more regularly eventually resulting in 'clonal extinction'. Accordingly, it was assumed that AML cells form a highly organised hierarchy that is comparable to that of the stem cell pool retaining function and regulation in which the leukaemia-initiating event occurs in primitive cells and not in committed progenitors.[98] However, it has been previously observed that AML-M1 can be engrafted into SCID mice with both $CD34^{+}$ and $CD34^{-}$ fractions.[99] Although it disfavours this model, it was proposed that multiple genetic events arose in the CD34 fraction uncoupling function from lineage expression.[99] Despite the apparent propensity of HSC transformation, AML-M3 may be an exception; the *PML/RAR*α fusion gene was observed only within the $CD34^{+}CD38^{+}$ population, thus transformation was probably acquired at the level of a committed progenitor.[100]

Normal primitive HSCs defined as $CD34^{+}CD38^{-/lo}$ Thy-1^{+} were found depleted in AML patients in remission carrying the *AML1/ETO* [t(8;21)(q22;q22)] translocation. Instead, the leukaemic bone marrow contained a small population of primitive leukaemic cells that had

the ability to self-renew with a $CD34^{+}CD38^{-/lo}$ Thy-1^{-} phenotype. In this study, this population of cells were not leukaemic and still retained the ability to differentiate into mature myeloid cells *in vitro*. Differentiated myeloid cells were positive for *AML1/ETO* and the translocation was also detected in B-lymphocytes[101] strongly suggesting that the initial event occurred at the stem cell level however, the downregulation of Thy-1 may implicate that additional event(s) occurred in a normal Thy-1^{+} HSC causing the loss of the Thy-1 phenotype.[4]

1.12 Chromosomal Abnormalities

The karyotype of abnormal cells is an independent predictor of therapy response, duration of remission and survival. Acquired cytogenetic aberrations are detected in 50 to 60% of newly diagnosed patients with MDS with a predominance of non-random chromosome copy number alterations.[34,35] However, low sensitivity methods such as conventional cytogenetic G-banding analysis fails to detect karyotypic abnormalities in ~50% of patients.[102] Although the karyotype is termed 'normal',[103] patient outcome is heterogeneous with some individuals' rapidly deteriorating following diagnosis.[104]

15 to 30% of patients with *de novo* MDS/AML have complex chromosome aberrations with no specific rearrangement involving three or more chromosomes.[105] These patients have a significantly inferior prognosis and show a poor response to treatment including a considerable reduction in the success of a bone marrow transplant.[106]

The 5q deletion is the most frequently reported chromosome deletion in MDS occurring in 10 to 15% of patients.[107] An interstitial deletion of 5q, the del(5)(q13q33) is regularly detected although other variants including del(5)(q31q35) and del(5)(q13q35) have been identified.[108] Although, it is cytogenetically indistinguishable from deleted chromosome 5 of other myeloid disorders, the critical minimally deleted region (CDR) associated with the indolent 5q- syndrome is distinct from the CDR associated with more aggressive types of MDS or AML, thus specifying two separate genomic intervals on chromosome 5q.

The 1.5Mb CDR mapped between 5q32 and 5q33[109] has been documented to contain various genes including *RPS14* and miRNAs including *miR-145*.[110,111] The haploinsufficiency of these genes has been implicated in the pathogenesis of the 5q- syndrome with regard to

defective erythropoiesis and megakaryocytic dysplasia,[110,111] respectively. Haploinsufficiency of *RPS14* results in defects in ribosome biogenesis and translation. In response to ribosome dysfunction, HDM2 is prevented from inducing *p53* ubiquitination.[112] Consequently, p53 accumulates and cell cycle arrest or apoptosis ensues. It has been previously demonstrated that $CD34^+$ cells with a *RPS14* knockdown show a reduction in the capacity to differentiate *in vitro* along the erythroid lineage. Notably, p53 was prevalent in early and late $CD71^+$ erythroid progenitors consistent with an elevated percentage of the cells restricted to the G_0/G_1 phase of the cell cycle. Other myeloid inclined lineages including early myeloid progenitors ($CD13^+$, $CD33^+$, $CD45^+$), leukocytes ($CD11b^+$) and megakaryocytes ($CD41^+$) did not show increased levels of p53. Defective erythropoiesis was also detected in the absence of ribosome dysfunction i.e. under conditions where HDM2 was chemically inhibited *in vitro* and *in vivo*.[112] This implicates that erythroid cells show increased sensitivity to p53 than other myeloid derived lineages.[112] Notably, DNA damage or telomere dysfunction could be readily detected in erythroid cells and therefore readily removed from the cell cycle. Consistent with this, Acute Erythroid Leukaemia occurs infrequently in patients manifesting the 5q- syndrome[113] however, *p53* inactivation may play a role in the leukaemic progression in these cases.

Conversely, the 1-1.5Mb CDR commonly identified in AML and more advanced forms of MDS is mapped to 5q31[114] which includes *EGR1* and *CTNNA1*. 5q deletions observed in AML are usually associated with a complex karyotype and an inferior prognosis.[115,116] Consistent with this, EGR1 (Early Growth Response 1) has shown to be a direct transcriptional regulator of many tumour suppressor genes (TSG) including *p53* and *p21*.[117] Moreover, it has been proposed that loss of function of EGR1 plays an initiating role in the development of MDS/AML as observed in $EGR1^{+/-}$ mice. Notably, the haploinsufficiency of *EGR1* led to the development of lymphoid and myeloid malignancies in the murine model.[118]

Monosomy 7 and a 7q deletion have been implicated in refractoriness to therapy and short survival. It was previously demonstrated that the loss at 7q31 is associated with the development of larger 7q- clones and short survival in patients with haematological disease including MDS and AML. Furthermore, a lower frequency of complex karyotypes was detected in patients who had retained 7q31[119] suggesting that a candidate TSG is mapped to this locus. It has been recently documented that the loss of *DOCK4* located at 7q31 may play

a role in myeloid disease. Accordingly, haematopoiesis colony assays demonstrated that reduced expression of *DOCK4* in primary $CD34^+$ cells resulted in a significant decrease in erythroid and myeloid colony formation as well as a significant increase in apoptosis of $CD34^+$ cells.[120] Furthermore, *DOCK4* has been implicated to play a role in the formation of cellular adheren junctions in which its loss has been associated with enhanced tumour invasiveness *in vivo*.[121]

Advanced MDS is characterised by high levels of genomic instability in comparison to early MDS. This is consistent with the downregulation *Chk1* and *Rad51* which has been noted to occur by at least 2-fold in advanced MDS.[122] This implicates a dysfunction in cell cycle control and homologous recombination repair of DNA double strand breaks respectively. Cells show an increase in genomic instability and acquire a growth advantage leading to malignant transformation.[122]

Aberrant DNA methylation in the promoter region of TSGs is an alternative to chromosome deletion for silencing TSGs. Concordantly, MDS patients show a significantly greater number of aberrantly methylated loci that include TSGs and genes involved in cell differentiation.[123]

The $p15^{INK4B}$ cyclin-dependent kinase inhibitor is commonly hypermethylated in MDS with increased methylation status corresponding to disease progression.[124,125] Accordingly, $p15^{INK4B}$ hypermethylation may provide a growth advantage by enabling cells to progress through the G_1 phase of the cell cycle. Consistent with this, advanced MDS is associated with genomic alterations and loss of cell cycle control that can enable the clone to acquire additional genomic aberrations. Clonal variation and positive selection can then provide neoplastic advantage.

Mutations of the Ten-Eleven Translocation-2 (*Tet2*) arise in 26% of MDS cases.[126] It has been implicated to regulate the DNA methylation of genes important for myelopoiesis and leukaemogenesis.[127] $Tet2^{-/-}$ mice show characteristics typical to CMML patients including neutrophilia and monocytosis with 33% developing pronounced splenomegaly and hepatomegaly caused by either erythroblast or myeloid cell infiltration, i.e. myeloblasts, monocytes, macrophages and neutrophils. This is suggestive that Tet2 functions as a tumour suppressor in myelopoiesis. *Tet2* haploinsufficiency induced myeloid malignancy in 8% of mice; however, erythroblast infiltration was not detected. Furthermore, disease latency was

longer when compared to *Tet2*$^{-/-}$ mice.[127] Thus, it has been implicated that *Tet2* alters the disease phenotype in a dose-dependent manner. In keeping with this, loss of its function has been associated with an inferior prognosis e.g. a patient with a low IPSS risk score of 0 at diagnosis only survived for three months.[128]

LOH (Loss of Heterozygosity) on 6q and 10p have been detected in a third of AML patients following MDS progression. Furthermore, allelic loss on 7p, 11q, 14q and 20q are also frequent events that have been observed in 23 to 27% of cases. It was proposed that recombination may be the mechanism responsible for LOH since no deletions were present on these arms.[129] Consistent with this, Uniparental Disomy (UPD) has been recognised to occur in MDS.[130]

UPD extending to the telomere has been identified in MDS on multiple chromosomes including 7p, 4q and 3q.[130] The duplication of a pre-existing mutated TSG has the potential to completely inactivate its function, thus allowing clonal progression. Alternatively, activating mutations that are duplicated by UPD, i.e. *FLT3* mutations can provide a growth advantage. Consistent with this, terminal UPD within the 13q12.11-qter comprised of the *FLT3* locus has been described in AML cases with a normal karyotype.[131] Furthermore, segmental UPD at *FLT3/ITD* has been observed in AML patients that have relapsed.[132] It has been predicted that terminal LOH occurs in 10 to 20% of normal karyotype AML.[131] Furthermore, UPD has also been detected within the terminal region of 17p consistent with the loss of functional *p53*. Consistently, this patient presented with a complex karyotype.[130] The prevalence of *p53* mutations has been studied extensively in MDS.[133-136] Patients harbouring 17p monosomy have been found to have a higher propensity to develop a *p53* missense mutation on the remaining allele.[133-135] It has been implicated that the complete abrogation of *p53* results in a significantly shorter survival, leukaemic transformation and enhanced resistance to chemotherapy.[135] *p53* mutations have been found to accompany abnormal cytogenetics, particularly abnormalities involving chromosome 5 or 7.[136]

1.13 Paradox

MDS is a highly proliferative disorder with almost a third of marrow cells engaged in DNA synthesis.[137] However, the bone marrow is simultaneously undergoing a high rate of cell death, which is particularly apparent in the early stages of the disease.[138,139] *In situ* end

labelling (ISEL) of fragmented DNA in MDS bone marrow aspirates has shown that all three myeloid lineages, i.e. erythroid, megakaryocytic and granulocytic undergo apoptosis as well as stromal cells including fat cells, endothelial cells and fibroblasts.[137] Accordingly, it was observed that over 75% of haematopoietic cells in the bone marrow were undergoing apoptosis in 50% of MDS cases.[137] Notably, elevated apoptosis of differentiating cells may contribute to the existing paradox of hypercellularity in conjunction with peripheral blood cytopenias.[22] Furthermore, it has been observed that cells which entered the S-phase of the cell cycle were also apoptotic implicating that an intact p53 pathway may be responsible for initiating apoptosis in replicating cells.[137] Consistent with this, increased H2AX phosphorylation has been observed in the $Lin^-c\text{-}kit^+Sca\text{-}1^+$ stem and progenitor cell population in an MDS mouse model in contrast to a wild-type.[140] It has been proposed that the elevated level of apoptosis in MDS may function as a protective mechanism by reducing the number of premalignant cells that can acquire additional genetic mutations limiting the progress to AML.[140]

Dual-labelled flow cytometry has been utilised to analyse the extent of apoptosis and proliferation of $CD34^+$ cells following their transformation from MDS to Acute Myeloid Leukaemia.[141] The level of $CD34^+$ apoptosis exceeds the percentage of cells in the S-phase in RA/RARS patients. However, the evolution to RAEB-T/AML has been demonstrated to accompany a decrease in apoptosis.[141] Accordingly, altered oncoprotein expression has been shown to accompany advanced disease supporting the accumulation of neoplastic $CD34^+$ cells. C-myc has been observed to be a potent inducer of apoptosis under certain microenvironmental conditions, i.e. in the absence of stimulatory growth factors or in the presence of inhibitory cytokines,[142] such as TNFα. The expression of C-myc to Bcl-2 (blocks apoptosis) has been quantified in $CD34^+$ cells derived from MDS patients of different stages as well as those from AML. It was observed that the degree of apoptosis occurring in $CD34^+$ cells from MDS patients correlated with the relative C-myc:Bcl-2 oncoprotein ratio. Accordingly, apoptosis and the C-myc:Bcl-2 ratio was highest in RA/RARS but reduced sequentially with increased Bcl-2 expression, i.e. RA/RARS > Normal > RAEB/RAEB-T > AML.[143] It has been documented that increased C-myc expression enhances cell cytotoxicity to TNFα.[144] Higher levels of TNFα have been shown in MDS[145] which may contribute to the apoptotic nature of this disease. However, the predisposition to apoptosis is reduced when

Bcl-2 expression is elevated.[144] Whilst a near 4-fold reduction of *bcl-2* expression has been observed in mouse models of early MDS when compared to wild-type mice[140] an increase in *bcl-2* expression has been associated with AML progression as well as a poor response to chemotherapy.[146,147]

TNFα has been found at elevated levels in marrow plasma from MDS patients.[145] It has the ability to regulate anti-apoptotic or pro-apoptotic effects on the cell by interacting with cell surface receptors TNFRI or TNFRII. Interaction with TNFRI can initiate anti- or pro-apoptotic responses by means of Nuclear Factor kappa B (NF-κB) or induction of the caspase cascade respectively.[148] TNFRII lacks a death domain and therefore it is only able to provide cellular protection. It has been previously observed that RA patients show a significant increase in TNFRI expression as compared to controls. In contrast, the expression level in late stage RAEB/ RAEB-T was similar or even lower to that found in controls.[148] This is consistent with the higher apoptotic capacity in early stage disease. Furthermore, late stage MDS was associated with a significant increase in TNFRII expression consistent with blast cell accumulation and reduced apoptosis in advanced disease.[148] Thus, a switch in favour of TNFRII from TNFRI plays a role in promoting MDS progression by reducing apoptosis of transformed cells.

Activated cytotoxic CD8$^+$ T-cells may contribute to the degree of myelosuppression by inducing apoptosis in by-standing normal haematopoietic cells[149] and also prevent the propagation of the MDS clone. However, it has been proposed that in later stages MDS blasts have the advantage over T-cells and escape immune detection. Notably, B7-H1 (CD274) molecules have been detected more often on MDS blasts in high-risk patients when compared to low-risk. B7-H1$^+$ cells deliver an inhibitory signal to activated T-cells that express the Programmed cell Death-1 (PD-1) transmembrane protein. It has been observed that B7-H1$^+$ blasts have greater proliferative capacity than B7-H1$^-$ blasts and have the ability to suppress T-cell proliferation and induce T-cell apoptosis.[150] Thus, B7-H1 expression is associated with immune evasion and possibly with MDS progression.

Normal CD34$^+$ cells do not spontaneously express Fas but can be induced to do so with TNFα in culture.[151] Fas is a cell surface receptor that induces apoptosis when ligated by the Fas ligand (FasL), a cell surface molecule. Up-regulated Fas expression has been detected on

total BMMNC (bone marrow mononuclear cells) and on different BM subpopulations including $CD34^+$, $CD33^+$, glycophorin$^+$ and $CD14^+$ cells in MDS compared to normal controls. A strong negative correlation has been previously detected between Fas expression on $CD34^+$ cells and the percentage of BM blast cells.[152] Accordingly, patients with advanced disease had lower Fas expression. It has been suggested that low Fas expression in the BM of AML patients is associated with a low remission rate of induction chemotherapy. Thus, the ratio of Fas^+ to Fas^- cells was predictive of treatment outcome i.e. disease resistance to apoptosis.[153]

Up-regulated TNFα in MDS can induce Fas expression on normal cells increasing their susceptibility to apoptosis by FasL-expressing cells. Accordingly, the growth of clonogenic progenitors was inhibited by FasL-expression in MDS.[154] FasL-expression is more pronounced on blasts cells in advanced MDS cases and has been observed to increase by at least three fold upon transformation to AML.[154] Additionally, increased FasL-expression on malignant cells may facilitate in the escape from T-cell mediated immunological surveillance by inducing apoptosis in Fas^+ T-cells.[155] Nevertheless, MDS clonal cells expressing Fas on their surface are also susceptible to Fas-mediated cell death therefore, elevated anti-apoptotic signals may provide a protective mechanism for these cells.

Accordingly, an increase in Nuclear Factor kappa B (NF-κB) activation has been found to correlate with disease stage[156] with it being constitutively active in AML.[157] NF-κB is a transcription factor that regulates the expression of a variety of proteins that inhibit apoptosis and promote cell proliferation and survival.[158] Notably, it has been implicated in the pharmacological resistance to many chemotherapeutic agents.[159]

Part 2: Telomeres

1.14 History of Telomeres

Early cytogenetic work demonstrated that X-ray induced chromosome breakage resulted in the production of chromosomal fusion between broken ends.[160,161] The centromeres of the dicentric chromosome were pulled to opposite poles during anaphase generating a chromatin bridge which was subsequently broken following pole-ward migration of the centromeres.[161] Thus, a loss or gain of genetic information was passed on to each daughter cell. It was proposed that these breakage-fusion-bridge (BFB) cycles continued indefinitely unless the broken end is 'healed'.[161] This was consistent with a previous observation that the natural chromosome ends did not take part in end-to-end fusions.[160] Thus, it was proposed that the terminal end has a special function in sealing the end of the chromosome. The word 'telomere' was coined by Herman Muller in the late 1930s and derived from the Greek translation of 'telos' and 'meros' meaning, 'end' and 'part' respectively.

The End Replication problem was later discovered following the elucidation of conventional DNA synthesis.[162,163] Loss of chromosome terminal sequence with each round of replication was proposed to result from the inability of DNA polymerase to completely replicate the linear DNA molecule. Thus, it was proposed that telomere shortening acts as a cell-intrinsic clock that would eventually lead to replicative senescence.[163] Consistent with this, a strong correlation between telomere length and cellular proliferative capacity has been documented.[164] Sequencing the ends of chromosomes derived from *Tetrahymena* revealed that the terminal DNA sequence was composed of simple tracks of T and G residues.[165] These tandem copies of 6nt sequence TTGGGG were presumed to defend chromosomes against the end replication problem and other assaults on their integrity. It was later identified that the hexanucleotide TTAGGG is the telomere repeat sequence found in humans[166] demonstrating the conservation of these repeats through evolution.

1.15 The End Replication Problem

The replication fork paves the way for DNA replication facilitating the synthesis of the 5' to 3' leading strand and the 3' to 5' lagging strand. Assuming DNA synthesis initiates within the

molecule; the leading strand can be synthesised continuously in the direction of the replication fork completing synthesis to the 3'termini. An alternative process is required to synthesise the lagging strand since DNA polymerase is unable to initiate DNA replication in the 3' to 5' direction (Figure 1.2). Instead RNA primers are utilised and extended in the 5' to 3' direction to generate a succession of Okazaki fragments. Prior to completion, the RNA primers are converted into DNA and the fragments are subsequently ligated. However, the most distal RNA primer is not converted to DNA due to the incapability of DNA polymerase to initiate replication *de novo*. Its subsequent degradation results in a 5' gap in the newly synthesised strand. In principle if the most distal RNA primer is located at the terminus the minimum loss of sequence at the lagging strand would be the size of the RNA primer, i.e. 7 to 10 nt[167] however, it has been demonstrated *in vitro* that this loss may increase substantially (~250nt) as a result of the priming initiation site.[167,168]

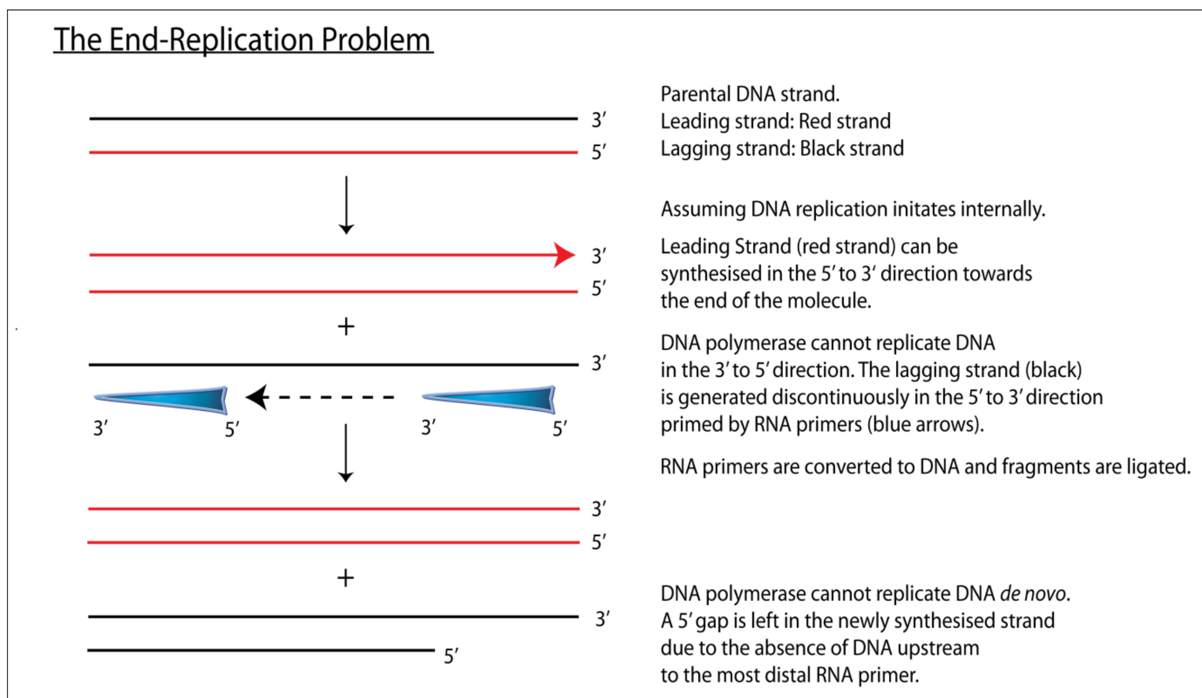


Figure 1.2: Since DNA polymerase is unable to replicate DNA *de novo*, semi-conservative replication leaves a 5' gap in the newly synthesised strand.

In the absence of 3' overhang resection, it has been proposed that telomeric double stranded DNA is lost at 0.25 the length of the single strand per population doubling.[169] Accordingly, a single stranded loss of 200nt per generation would account for a double strand loss of 50bp per cell doubling as described in human fibroblasts.[169,170] The reason being it was assumed that only half of the cells amongst the distribution were losing single stranded DNA per generation, thus it was proposed that a single strand deletion would be

attained upon the second generation and a double strand deletion every fourth generation. Moreover, this model also assumed that the variance in telomere length distributions would increase with each generation. Notably, an accumulation of variable single stranded and double stranded deletions in each daughter cell would result in an increase of the telomere length variance.[169] This increase in the telomere length distribution has been previously demonstrated in fibroblast cells in culture.[171]

A variation of telomere shortening can result in the division heterogeneity of cells derived from the same precursor however; prolonged telomere erosion ultimately leads to replicative senescence at which cells stop dividing in order to prevent further telomere loss. Accordingly, the telomere length distribution inevitably homogenises as the number of cells reaching cellular senescence accumulates with progressive replication.[172] It has been observed that the heterogeneity of telomere length distributions is more pronounced within multiple clonal populations in contrast to single cell clones that show less variation in telomere length.[173-175] Thus, telomere length distributions may be indicative of the relative clonality of the cell population.

1.16 Telomerase

The enzyme Telomerase was initially identified in *Tetrahymena* by its ability to add tandem TTGGGG repeats onto the 3'end of synthetic telomere primers.[176] It has reverse transcriptase activity and synthesises telomeric DNA onto chromosome ends using an internal RNA template.[177,178] Thus, telomerase is capable of compensating for the loss of terminal sequence that occurs as a result of the end replication problem. Telomerase activity has been detected in the germ line but is undetectable in normal somatic cells excluding proliferative cells of renewable tissues, i.e. haematopoietic stem cells, activated lymphocytes and intestinal crypt cells.[179] Telomerase activity has been documented in 85% human malignancies, including colon cancer, neuroblastoma and in lung carcinoma. A large majority of cultured immortal cell lines also present detectable telomerase activity.[180] Thus, the up-regulation of telomerase in malignant cells may contribute to their immortality by maintaining telomere length.

Telomerase is composed of a catalytic core (hTERT) and an RNA template (hTR). hTR is present in cell lines and tissues that lack telomerase activity and has been detected in both

tumour and normal samples.[181] Conversely, hTERT is detected in telomerase positive tissue, cancer cell lines and tumours but it is not detected in tissues that lack enzymatic activity.[182,183] The level of hTR expression does not always correlate with telomerase activity, i.e. cellular immortalisation has been associated with a 100 to 2000 fold increase in telomerase activity but only a 2 fold increase in hTR expression.[181] However, hTERT expression has been observed to reflect the presence of telomerase activity and has been implicated as the rate limiting determinant of the enzyme.[182-186] Accordingly, ectopic expression of hTERT mRNA can induce telomerase activity in telomerase negative cells[184] and extend their cellular lifespan.[185] hTERT expression has been detected in the earliest stages of cancer development including the breast,[187] colon[187] and lung squamous cell[188] carcinoma. Furthermore, it has been noted to gradually increase with progression into an invasive carcinoma showing prominent hTERT expression in later stages of cancer development. This suggests that telomerase activation is regulated during the progression of tumourigenesis but moreover, hTERT is a barrier that has to be overcome for telomerase activation required by cancer cells to gain telomerase dependent immortalisation.

1.17 The Mechanism of Telomerase

The template region is longer that the telomeric repeat they encode. The human template of telomerase RNA reads 3'-CAAUCCCAAUC- 5' providing a coding region for the human 5'-TTAGGG- 3'.[177] A portion of the template aligns through base pairing with the 3'overhang and DNA elongation ensues through copying of the template (Figure 1.3). However, the RNA templating region is restricted, such that the appropriate nucleotides are added to complete only one telomeric repeat unit.[189] Upon completion, the newly extended DNA terminus dissociates from the RNA template and the enzyme complex translocates to enable another round of telomere repeat unit replication.[189] The complementary DNA strand is then synthesised by means of lagging strand synthesis, thus completing telomere extension by telomerase.

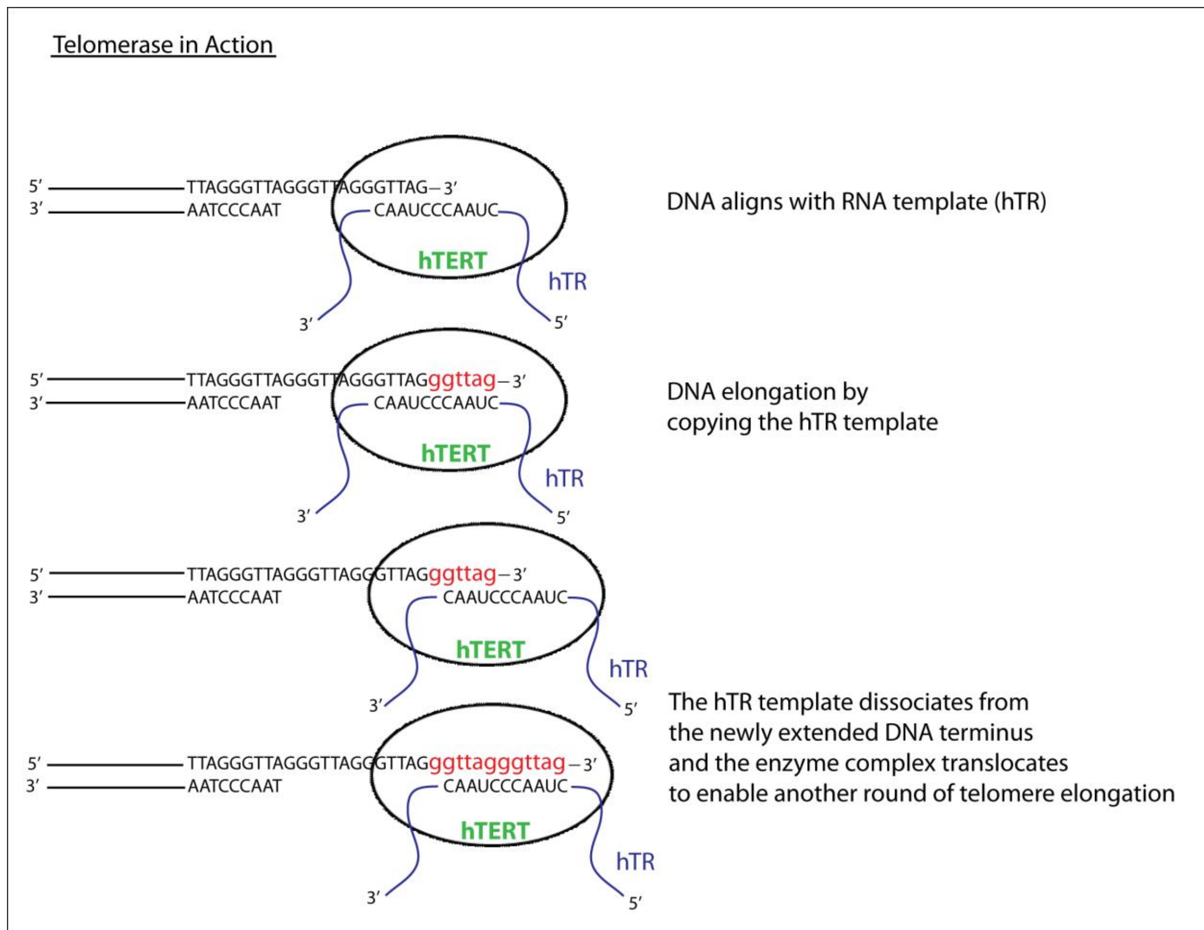


Figure 1.3: An illustration showing the process behind telomerase extension of telomeric ends (bases in red are copied from hTR (RNA template)).

1.18 'Capping' Linear Chromosomes

It has been suggested that chromosome protection can be achieved by a telomere through a looping back mechanism. The 3'overhang loops back and invades the preceding telomeric tract which is thought to sequester DNA ends from DNA repair pathways. Electron microscope analysis has visualised this occurrence as large lariat structures or T-loops that exist at the ends of chromosomes.[190] It was also observed that there is a close correlation between the length of the telomeric repeat array and the size of the T-loop implicating that the T-loop may encompass the whole telomere.[190] It has been proposed that T-loops 'seal' the end of telomeres thus enabling cells to distinguish random DNA breaks from natural chromosome ends.[190] The remodelling of telomere ends is enabled through a number of specialised telomere proteins that work in synchrony to generate a 'cap' at the ends of linear chromosomes.

1.19 The Shelterin Complex

A multi-subunit protein complex that is referred to as Shelterin binds to telomere ends[191,192] to facilitate the formation of the telomere 'cap'.[193] The complex is composed of six main proteins that mediate associations between the double and single stranded portions of the telomere (Figure 1.4).

The TTAGGG Repeat Factors 1 and 2 (TRF1 and TRF2) proteins bind to the 5'YTAGGGTTR 3' sequence in double stranded DNA and exist as either homodimers or oligomers.[192] Although they do not interact directly;[192] they are interconnected by the TRF1 and TRF2 Interacting Protein 2 (TIN2).[194] TIN2 has a central position in the Shelterin complex as it facilitates in bridging the double stranded interacting proteins with those on the single stranded overhang. Its depletion results in destabilisation of the complex and consequential reductions of TRF1 and TRF2 at the telomere.[195] TIN2 indirectly interacts with the Protection of Telomeres 1 (POT1) protein through the TIN2 and POT1 Interacting Protein 1 (TPP1).[191] This interaction is crucial since TPP1 tethers POT1 to TRF1 and TRF2 by the TPP1-TIN2 bridge.[196] POT1 binds directly to the 5' TAGGGTTAG 3' sequence on the single stranded overhang.[192] It forms a heterodimer with TPP1 that functions to enhance its DNA binding affinity.[192,194]

TRF1 and TRF2 are the core proteins in maintaining the structure of the T-loop. TRF1 has been shown to induce bending, looping and pairing of duplex DNA *in vitro.*[197,198] However, the fundamental role of chromosome end 'capping' is dependent on TRF2.[199] TRF2 has the ability to generate T-loops *in vitro* and it was proposed that it may stabilise[190] and induce strand invasion of the 3'overhang into duplex telomeric DNA. Intra-telomeric synapsing of TTAGGG repeat arrays by TRF1 may also facilitate strand invasion by shaping of the T-loop.[190] As a result, natural chromosome ends would be inhibited from inducing cell cycle arrest or entering into deleterious inadvertent double strand break repair pathways.[191]

Human cell lines expressing a dysfunctional TRF2 show telomere end-to-end fusion events including multiple fused, ring and dicentric chromosomes. It was demonstrated that the frequency of end-to-end fusions increased by 10 fold in these cells comparable to controls that rarely presented fusion events. Accordingly, these cells were TRF2 proficient and the protective function at telomeres was maintained.[199] POT1 has also been implicated in

chromosome end protection. Its deficiency has been described to result in senescence and telomere fusions in mouse embryonic fibroblasts (MEFs) as well as enhancing tumourigenesis *in vitro* and *in vivo* in a p53 deficient background.[200] However, the frequency of telomere fusions in POT1 deficient MEFs is considerably less than that in cells deleted for TRF2, i.e. 2% of chromosomes per cell[201] vs. genome wide.[202] This implicates that TRF2 has a dominant role in protecting telomeres against end joining by maintaining the T-loop structure preventing aberrant chromosome fusions. Notably, it has been proposed that although infrequent, the fusion events that arise with POT1 depletion may be associated with a transient open configuration of the T-loop during DNA replication. POT1 bound to the 3'overhang may prevent end joining proteins from accessing telomere ends.[201]

The Repressor/Activator Protein 1 (RAP1) is the final protein that associates with the Shelterin complex through interactions with TRF2. It has been proposed to interact with non-telomeric proteins that may provide maintenance of telomeric integrity. Furthermore, it has been suggested that RAP1 may act at telomeres in order to repress homologous recombination and unequal telomeric sister chromatid exchange (T-SCE). Accordingly, mouse embryonic fibroblast (MEFs) cells that express a mutant TRF2 incapable of interacting with RAP1 demonstrate an increased propensity of T-SCE at a similar frequency to TRF2-null cells.[203] Unequal telomeric sister chromatid exchange (T-SCE) can threaten the integrity of individual telomeres by inducing abrupt telomere shortening. This can significantly reduce the proliferative capacity and viability of the daughter cell that acquires the shortened telomere.[203]

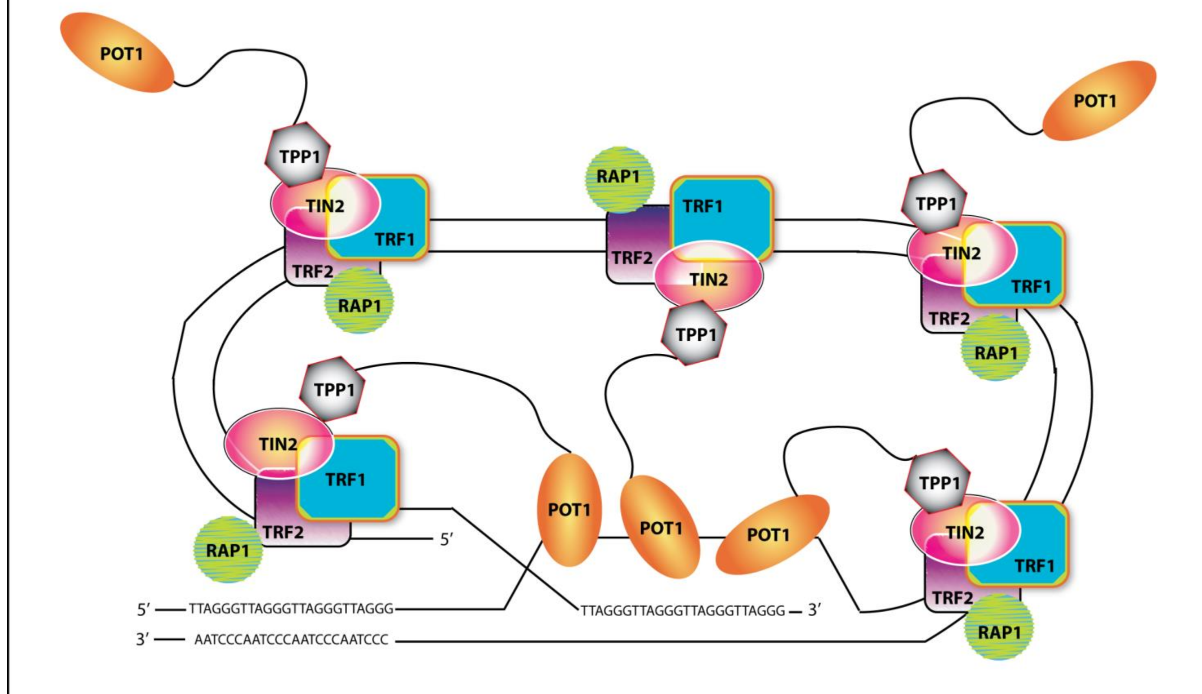


Figure 1.4: Shelterin makes up the telomere 'cap'. It is composed of 6 proteins: TRF1, TRF2, POT1, TIN2, TPP1 and RAP1. Associations between the double and single stranded portions of the telomere are mediated through protein interactions to facilitate in the shaping of the T-loop.

1.20 The Protein 'Counting' Model

It has been proposed that the presence of POT1 on the 3' single strand can negatively influence the telomerase activity by hindering base pairing between the RNA template and the 3' terminus of the overhang.[204] This has been suggested to be achieved through a protein 'counting' model[198] that enables duplex telomeric DNA to relay information about the telomere length to the 3'overhang where telomerase activity is regulated.[205]

The inhibition of TRF1 in telomerase expressing cell lines leads to telomere elongation and it has been assumed to act in *cis* by inhibiting telomerase activity at the telomere termini.[206] Conversely, overexpression of TRF1 leads to gradual telomere shortening in telomerase positive cell lines.[207] Telomeres can exist in either an 'open' or 'closed' state. Telomerase accessibility is granted when telomeres are in the 'open' state whereas it is prohibited when telomeres are 'closed'.[198] The alteration in telomere access has been proposed to be governed by the amount of TRF1 bound to the telomere.

It has been suggested that an accumulation of POT1 to the 3'overhang is dependent upon its interaction with TRF1. Longer telomeres will bind more TRF1 and therefore may facilitate

in the accumulation of POT1 to single stranded ends prohibiting telomerase accessibility. Accordingly, very short telomeres will be unable to bind sufficient amounts of TRF1 (and also POT1) enabling single stranded access to telomerase.[191,198,205] This results in telomere length homeostasis that is regulated in *cis* by telomere associated proteins.

1.21 Telomere Length Homeostasis

Telomere length homeostasis is the result of the balance between telomere shortening and telomere lengthening. If telomeres were selected at random for telomerase extension then this may have catastrophic effects on the cell. Accordingly, if telomerase only acted on long but not short telomeres then cells would enter premature senescence or alternatively short telomeres could initiate oncogenesis. In contrast, if telomerase acted on short telomeres then cell viability and proliferation capacity would be maintained.

It has been described that telomerase does not act on every telomere in the cell cycle. Accordingly, the frequency of telomere extension in *Saccharomyces Cerevisiae* increases steadily as a function of telomere length. Only 6 to 8% of telomeres were extended within long telomere distributions; whereas telomerase extended 42 to 46% of telomeres within short distributions, thus favouring the elongation of short telomeres.[208] Telomerase has shown a preference for short telomeres in mice[209] and human cells[210] and it has been proposed that the mechanisms acting in *cis* favour telomerase association with short telomeres. However, this regulation is reduced if telomerase is in abundance. Telomere elongation at 0.25-0.8kb/PD far above physiological length has been observed in human primary and cancer cell lines overexpressing both *hTERT* and *hTR*.[211] It was suggested that telomeres elongate in a length independent manner which further implicates that long telomeres have the potential to undergo telomerase extension if telomerase levels are high. Thus, to maintain telomere length homeostasis and preferable elongation of short telomeres then telomerase must be limited to enable regulation by telomeric proteins. Consistent with this, the elongation of short telomeres is favoured in mice with limiting telomerase activity.[212]

It has been suggested that long telomeres switch back more rapidly into a non-extendible state (i.e. T-loop or 3’ overhang coverage by POT1) following DNA replication, thus increasing the probability of telomerase to extend short telomeres which are more likely to

be in an extendible state, i.e. open configuration.[213] However, this regulation is reduced if telomerase is overexpressed enabling the elongation of long telomeres in a length independent manner.

Thus, Shelterin not only provides protection against DSB (Double Strand Break) repair but also functions as a length sensing mechanism to provide control over telomere length homeostasis in telomerase positive cells.

1.22 The G-quadruplex

G-rich oligonucleotides that contain at least four short runs of G residues can fold into compact forms that are stabilised through the association of a K^+ or Na^+ cation. This structure is referred to as the G-quadruplex. The building blocks of the G-quadruplex are G-quartets that arise from four guanines held in plane by Hoogsteen-hydrogen bonding in which each guanine serves as a hydrogen bond acceptor and donor.[214,215] The planar G-quartets stack on top of one another giving rise to four stranded helical structures. G-rich DNA oligonucleotides can form aggregates through G-G base pairing *in vitro*. Under non-denaturing conditions, electrophoresis has demonstrated that multiple intramolecular folded forms can originate from a DNA G-rich single strand.[216] The structure of the G-quadruplex is highly polymorphic since it can assemble in an intramolecular or intermolecular configuration with strands in a parallel or antiparallel orientation. Intermolecular structures can arise from the association of two dimers or four independent G-rich strands.[215]

The guanine rich sequences within the 3'overhang have the propensity to form G-quadruplexes. The dynamics of the folding and unfolding of the telomeric G-quadruplexes have been observed *in vivo*. The 16nt long G-overhang in the ciliate protozoan *Stylonychia lemnae* is bound by telomere end binding proteins TEBPα and TEBPβ, which are both required for generating G-quadruplexes *in vivo*. TEBPα can bind directly to the 3'overhang however, TEBPβ requires the interaction with TEBPα to associate with the single strand to engage in G-quadruplex formation. The phosphorylation of the C-terminus of TEBPβ by Cyclin Dependent Kinase during the S-phase of the cell cycle causes TEBPβ to dissociate from TEBPα and induces G-quadruplex unfolding for the end replication machinery.[217] Vertebrate homologues of TEBPα and TEBPβ are POT1 and TPP1 respectively.[218] However,

the C-terminus of TEBPβ is not conserved in TPP1 which may suggest an alternate role. Alternatively, longer vertebrate overhangs may have the capability of generating G-quadruplexes without the aid of telomere proteins.[219] However, it has recently been proposed that the POT1-TPP1 complex can fold and unfold the G-quadruplex by operating as a 'sliding clamp' that initiates near the more accessible 3'end of the overhang.[220] The association with Shelterin via TIN2 enables POT1-TPP1 to slide along the single strand while tethered to the duplex DNA. This sliding mechanism may regulate the accessibility of the 3'end and consequently telomerase elongation.

G-quadruplexes can also resemble telomeric fragile sites that have the propensity in inducing replication fork stalling. TRF1 has been implicated in having a specific role in facilitating telomere replication. In its absence, multiple telomeric signals separated from chromatid ends were observed by means of Fluorescence *in situ* Hybridisation (FISH) analysis.[221] It had been suggested that TRF1 may repress replication problems by recruiting helicases, e.g. BLM RecQ that can remove G4-DNA. Notably, BLM-deficient mouse cells show an increased frequency of spontaneous fragile telomeres.[221]

Therefore it can be assumed that POT1 and G-quadruplex DNA play important roles in regulating telomere length homeostasis. Intermolecular formation of the G-quadruplex may also favour telomere-telomere interactions, i.e. recombination. However, G4-DNA has the propensity to generate extensively short telomeres as a result of telomere fragility and replication fork stalling.

1.23 Telomerase and Cancer Therapeutics

Telomerase activity is absent in most somatic cells therefore telomerase inhibition is unlikely to induce an adverse effect. Furthermore, telomerase competent cells, i.e. germ and stem cells have long telomeres as opposed to short telomeres in telomerase positive cancer cells. Moreover, elevated *hTERT* expression in cancer cells as opposed to low expression in normal cells has been demonstrated to provide a valuable target for immunological therapy. Thus, it has been proposed that cancer cells may be more sensitive to telomerase based therapies with the reduced probability of toxicity to normal tissue.

Telomerase immunotherapy is an alternative approach to destroying tumour cells. Antigen presenting cells i.e. autologous dendritic or B-cells are exposed to high levels of synthetic hTERT peptides or genetic components *ex vivo* which are subsequently transplanted into the patient. The antigen presenting cell then interacts with $CD8^+$ cytotoxic T cells or $CD4^+$ helper cells to elicit an immune response against TERT expressing cancer cells.[222] One such immunotherapy is GRNVAC1. This is a preparation of autologous dendritic cells that are transfected with TERT mRNA. mRNA processing by dendritic cells encodes multiple variations of the TERT peptide inducing a polyclonal immune response.[223] Therefore, GRNVAC1 treatment promotes an anti-tumour effect since it may target any variation of TERT expressed on tumour cells. GRNVAC1 has completed Phase II clinical trials in patients with AML and metastatic prostate cancer.[224]

Compounds inhibiting telomerase directly can induce telomere shortening and subsequent genetic instability returning malignant cells to proliferative crisis. However, this therapeutic approach may be associated with a lag period that is dependent on telomere length. Furthermore, cancer cell may initiate ALT (Alternative Lengthening of Telomeres) and maintain telomere length through recombination. It has also been proposed that a functional DDR may be required to induce the apoptotic response in cells where telomerase inhibition has resulted in critically short telomeres.[225] Thus, DDR (DNA Damage Response) components such as p53 may serve as an important genetic marker to help determine effective treatment options.

GRN163L (Imetelstat) is one such inhibitor that acts as a 'telomerase template antagonist.' It is a small molecule oligonucleotide that hybridises to the hTR template region preventing it from forming an active complex with hTERT. GRN163L has been shown to limit the growth of cancer in multiple tumour types including breast,[226] bladder,[227] lung[228] and haematological cancers such as lymphoma and Multiple Myeloma.[229] GRN163L has already completed several Phase I trials in patients with CLL (Chronic Lymphocytic Leukaemia) and solid tumors including breast cancer and non-small cell lung cancer.[230]

Translocation and re-initiation is required to take place after each cycle of template copying. BIBR1532 is another telomerase inhibitor that reduces the number of added TTAGGG repeats but maintains the periodicity of six nucleotides. It has been implicated that

BIBR1532 affects the translocation of the enzyme DNA substrate complex or promotes the dissociation of the enzyme from the DNA following template copying.[231] It has been shown to inhibit telomerase in several human cancer cell lines including breast, prostate and lung carcinoma.[232] Additionally, in combination with chemotherapeutics it has been shown to induce enhanced sensitivity in human promyelocytic leukaemia resistant cell lines.[233]

G-quadruplexes have also been discussed as a potential therapeutic approach. The RHPS4 ligand can bind and stabilise G-quadruplexes hindering both telomerase and POT1 association to the 3'overhang. Consistent with this, RHPS4 can induce telomere dysfunction by means of telomere shortening but can also induce short term effects by interfering with the Shelterin complex when provided at high concentrations. Accordingly, RHPS4 can lead to POT1 dissociation from the telomere and induce a DDR in melanoma cell lines.[234]

1.24 Genetic Determination of Telomere Size

There is considerable telomere length variation in the human population.[235,236,237] This is underpinned by genetic variation. Previous studies have reported heritability estimates ranging from 36 to 90%.[238] It has been proposed that common environmental factors may be responsible for the apparent discrepancy in reported heritability estimates, such as Body Mass Index (BMI) or smoking,[239] however when compared to dizygotic and unrelated individuals of the same age the variation of telomere length between monozygotic twins was found to be the smallest.[240]

1.25 Sub-telomere Structure

The human genome is composed of dense concentrations of inter-chromosomal segmental sub-telomeric duplications that display considerable variation.[241]

Human sub-telomeres are bounded proximally by chromosome specific sequences and distally by an array of least three telomere variant repeats (TVR) composed of TTGGGG, TGAGGG and TCAGGG that extend up to 3kb into the telomere repeat array. Distal to the TVR region lies a continuous block of TTAGGG repeats that extend to the chromosomal terminus.[242-244]

Sub-telomeric regions are highly polymorphic with respect to copy number and chromosomal location.[241] The presence or absence of these tracts can result in size

variations of homologous chromosomes by several 100kbs.[245] Notably, allelic variations of up to 260kb have been observed to exist at the 16p telomere.[246] Multiple chromosome ends show very high sequence similarity that may range between 90 and >99.5%.[247] Moreover, these repeat regions may extend for up to 200kb in humans.[247]

Human sub-telomeres contain members of 25 small gene families including those of odorant and cytokine receptors, tubulins, transcription factors and others of unknown function. Rearrangements within sub-telomeric regions may induce phenotypic variation and have the potential to contribute to human disorders. Mental retardation has previously been reported to originate from sub-telomere alterations that create sub-terminal microdeletions which are healed by the action of telomerase creating a new telomere.[241]

1.26 Methods of Telomere Measurement

1.26.1 Terminal Restriction Fragment (TRF)

TRF analysis is the traditional method of measuring telomere length in samples of total human genomic DNA.[248]

Genomic DNA is digested with frequently cutting restriction enzymes that digest the majority of the genome without cutting into the telomere repeat array. This leaves intact terminal restriction fragments (TRFs) which can be resolved by agarose gel electrophoresis and detected by Southern hybridisation with a telomere repeat containing probe.[249] TRF requires microgram quantities of DNA per individual (~10^5 cells) and therefore this may hinder the measurement of small tissue or cell samples.[249]

TRF analysis is also biased towards longer telomere length.[248] Accordingly, it is dependent on hybridisation and therefore may fail to detect short telomeres which are below the telomere length threshold. Moreover, it determines the genome wide telomere length of all ends simultaneously from a large number of cells. Although this reduces its resolution capacity, it also measures an overestimate of the average telomere length since TRF analysis includes heterogeneous quantities of sub-telomeric and TVR DNA.[249]

1.26.2 Quantitative-Fluorescence *in situ* Hybridisation (Q-FISH)

Fluorescently labelled $(CCCTAA)_3$ peptide nucleic acid (PNA) probes are hybridised to metaphase chromosome preparations.[250] Under conditions of low ionic strength, the neutral PNA probe can anneal to complementary single stranded DNA sequences whereas single stranded DNA cannot therefore this allows the quantification of the fluorescent signal derived from telomeric sequences.[251]

Although it is able to measure chromosome specific telomere length at single cell resolution, values are relative telomere measurements and not actual telomere length. Moreover, since this procedure relies on hybridisation a telomere length threshold exists such that 'signal free ends' may be misinterpreted as completely denuded telomeres which are actually those that contain telomere repeats below the detection threshold.[249]

Q-FISH is unable to measure the telomere length in terminally senescent cells or cells with low mitotic indices.[248]

1.26.3 Flow-Fluorescence *in situ* Hybridisation (Flow-FISH)

Similarly to Q-FISH, flow FISH uses directly fluorescently labelled $(CCCTAA)_3$ peptide nucleic acid (PNA) probes,[251] however unlike Q-FISH the 'genome wide' telomere length is measured in individual cells in suspension including metaphase, interphase and senescent cells. Immuno-phenotyping can be utilised to measure the telomere length in distinct cell populations within a single sample.[248] However, telomere length values are relative telomere measurements, not actual telomere length.

1.26.4 Quantitative telomere-specific PCR (Q-PCR)

Telomeres (T) are PCR-amplified using primers that anneal to both the C- and G-rich strands of the telomere; however mismatches are present along their length in order to reduce the formation of primer dimer derived products. The amplification is measured quantitatively and compared to that of a single copy gene (S) to generate a T/S ratio that is proportional to the average telomere length.[252] Although this approach only requires nanogram quantities of DNA, it only provides the mean relative telomere length and not the actual length.

1.26.5 Single Telomere Length Analysis (STELA)

STELA is a long-range single-molecule PCR approach that amplifies the double-stranded region of telomere repeats from specific chromosome ends.[171]

The initial step in STELA consists of annealing a 'telorette' linker that is comprised of seven bases of TTAGGG repeat homology followed by a 20 nucleotide non-complementary tail to the G-rich 3'overhang of the telomere. Ligation to the 5'end of the duplex telomeric C-rich strand provides a telorette 'tag' to the end of the chromosome. PCR can then be performed on chromosome-specific telomeres utilising a 'teltail' primer that is complementary to the telorette tail and a chromosome-specific upstream primer located within the subtelomeric region. STELA products are identified as single bands on a gel following Southern hybridisation with a TTAGGG specific probe with each band representing a single telomeric molecule.[171]

This approach has the ability to detect the full spectrum of telomere lengths at specific telomeres,[171] particularly those that are very short which have the potential to initiate telomere fusion or replicative senescence.[253-255]

Small quantities of input DNA (~250pg/μl) or as few as 42 cells can be analysed using STELA therefore enabling the measurement of telomere length in rare cell types. Whereas TRF analysis measures an overestimate of the average telomere length, STELA is able to measure the telomere length accurately since the exact distance of telomere adjacent DNA (sub-telomeric and TVR) can be determined.[171,249]
However, only a subset of chromosome ends can be measured using STELA due to the presence of extensive sub-telomeric homology.[241,247] Notably, ends that encompass a unique telomere-adjacent sequence can only be analysed by STELA.

1.27 Telomeres and Homologous Recombination

Telomere length homeostasis involves *cis* acting regulators that negatively influence telomerase mediated extension[256] to ensure limited heterogeneity of individual telomere lengths within a clone.[235,257] Maintenance of homeostasis is apparent in immortalised cells and vital to ensure consistency of telomere length within the germ line.[258] Telomerase has

the capability to add telomere repeats *de novo*, however, telomerase deficient *Saccharomyces Cerevisiae*[257] and *Kluyveromyces Lactis*[259] mutants are able to generate post-senescent survivors that maintain telomeric length through a RAD52 dependent homologous recombination pathway.

The RAD52 epistasis group, composed of Rad51, Rad54 and Rad57 is required for the resistance against DSB instability in which a deletion renders *Saccharomyces Cerevisiae* highly sensitive to γ-radiation.[260,261] Early studies have demonstrated RAD52 involvement in the homothallic mating type gene switching in *Saccharomyces Cerevisiae*[262] which involves the exchange of genetic information initiated by a HO-endonuclease site specific DSB at the MAT locus. MAT is replaced by either the HML (Hidden MAT Left) or HMR (Hidden MAT Right) gene to confer sexual differentiation within a population of yeast cells.[263] Incomplete HO-induced mating type switching can induce lethality in *Rad52* yeast mutants.

Post-senescent survivors with heterogeneous telomere and sub-telomere profiles[259] can be generated by telomere/telomere or sub-telomere/sub-telomere recombination via the Rad50 and Rad51 subgroups respectively.[264] It has been proposed that the Rad51 subgroup enables strand exchange in regions where the chromatin structure is repressed, such as sub-telomeric regions.[265,266] This heterogeneity of telomere lengths has also been observed in telomerase independent immortalised human tumour cell lines.[267] It is thought that RAD52 dependent telomere elongation is conferred through a Break Induced Replication (BIR) mechanism that enables homologous recombination between sequences on opposite arms.[268] The resulting telomere capture provides the broken arm with a replica of sequence adjacent to the homology site up to the end of the telomere.[269] This is referred to as ALT; an Alternative Lengthening of Telomeres. Although telomere length can be preserved though this mechanism; the high frequency of interstitial sub-telomeric homologs within the genome could be inadvertently used to generate non-reciprocal translocations[269,270] and initiate a sequence of genomic rearrangements that may be detrimental to the cell. This is comparable to the high incidence of Alu repeats that have been proposed to account for 0.3% of human diseases.[271]

Additionally, recombination can occur intra-chromosomally[269] and it has been postulated that the resolution of a recombination 'T-loop like' intermediate can generate extensive

telomere shortening (Telomere Rapid Deletion; TRD) within a single round of cell division. These telomeric events have been observed in yeast mutants that lack the ability to control telomerase accessibility through a deficiency in Rap1.[272] TRDs maintained telomere homeostasis by restoring elongated telomeres back to their wild-type length.[273] It has been proposed that yeast acquires a 'yardstick' mechanism that measures telomere lengths relative to each other. Furthermore, this approach of telomere truncation generates extra-chromosomal telomeric circles that can be exploited by short telomeres to elongate by telomere/telomere recombination via a 'rolling loop' mechanism.[274] The mammalian ERCC1/XPF complex has been located at the telomere which is thought to prevent T-loop formation within interstitial telomere repeats that may result in TRDs and extrachromosomal circles (Telomere DNA containing Double Minutes; TDMs). Consistent with this, ERCC1 null MEFs (Mouse Embryonic Fibroblasts) undergo senescence prematurely.[275]

The Mre11 complex is composed of the Mre11, Rad50 and Xrs2 (MRX) proteins in yeast or Mre11, Rad50 and Nbs1 (MRN) in humans.[273] It has the ability to stabilise DNA strand interactions[276] and thus playing a structural role in telomere/ telomere circle directed telomere elongation.[274] Moreover, *Saccharomyces Cerevisiae* mutants in Rad50 show a reduction in TRDs suggesting it may also play a structural role in generating terminal deletions.[273] Additionally, mutations within the BRCT (BRCA1 carboxyl terminal) domain on Rap1 result in a reduction of TRD events,[256] possibly through a loss of interaction with the MRN complex. It has also been suggested that the nuclease activity of Mre11 generates the 3'overhang on the newly replicated leading strand to provide protection against telomere end-to-end fusions by Non-Homologous End Joining (NHEJ). Mre11 mutant mouse fibroblasts deficient in TRF2 exhibit leading strand initiated sister chromatid fusions as a consequence of their transient blunt ends during synthesis.[277] The MRN complex has also been proposed to facilitate in the generation of the telomerase substrate for telomere elongation.[278,279]

1.28 Non-Homologous End Joining/Microhomology Mediated End Joining

Non-Homologous End Joining (NHEJ) appears to be the most common pathway of DSB repair in mammalian cells.[280] In the context of a TRF2 deficiency[281] cells attempt to repair

damaged telomeres using NHEJ generating telomere-telomere fusions, dicentric or multicentric chromosomes. Classical NHEJ is dependent on the heterodimeric nuclear protein Ku[282] that is comprised of subunits Ku70 and Ku80.[282,283] In the event of a DSB, the Ku heterodimer binds to DNA ends and recruits several accessory factors that serve to process the broken ends.[284] Ku mediates the recruitment of the catalytic subunit of an ATM-related DNA protein kinase (DNA-PKcs)[285] to DSBs that becomes activated on DNA binding.[286] DNA-PKcs phosphorylates Artemis which functions as an endonuclease that cleaves at DNA structures containing 5' or 3' overhangs[287] enabling subsequent ligation catalysed by the XRCC4/DNA Ligase IV complex.[288,289] Intermediate gaps existing within the DNA break can be filled by the DNA Polymerases polμ and polλ.

Mice deficient in Ku or DNA-PKcs exhibit pronounced growth retardation and endure a *scid* phenotype that is characteristic of impaired lymphoid cell development due to a deficiency in the DNA NHEJ step of V(D)J recombination.[261,285] Additionally, fibroblasts derived from Ku deficient mice demonstrated a reduced proliferation capacity and entered senescence prematurely.[285] High levels of aneuploidy and chromosomal abnormalities have been detected in mice embryonic fibroblast cells deficient in Ku.[290] In accordance with this the loss of Ku has been linked to telomere length reduction[291,290] and a large increase in the frequency of TRD events.[291] Furthermore, Ku deficiency in human somatic cells is lethal whereas a knockdown confers significant telomere shortening and increased apoptosis.[292] Fusion between short dysfunctional telomeres have been detected in human and mouse cells deficient in Ku80[292]and DNA-PKcs[293,294] respectively. Telomere fusions were found to occur in DNA-PKcs deficient cells at a rate approximately 1 in 60 cells per round of cell division; whereas in DNA-PKcs proficient cells spontaneous fusions are nearly undetectable. This implies that Ku and DNA-PKcs make up a part of the telomere cap to protect against recombination or end degradation[291,293,129] possibly facilitating in production of the 3'overhang.

The observation that telomere end-to-end fusions occur despite deficiencies in the NHEJ components suggests that an alternative 'repair' mechanism may operate at dysfunctional telomeres that is independent to NHEJ. Components of these alternative-NHEJ mechanisms have recently been characterised. Mutants of the DT40 chicken B cell defective in homologous repair and Ku-dependent NHEJ are able to repair DSBs.[295] However,

deficiencies confer significant radiosensitivity and genomic instability. A ligase IV haploinsufficiency in mice mutants' deficient in p16^{INK4a} and p19^{ARF} are prone to clonal chromosomal translocations involving deletions and amplifications which promote the development of soft tissue sarcomas.[296]

The inactivation of *HDF1* (Ku70 homolog) in *rad52* mutant yeast strains increases the sensitivity to γ radiation.[261] *Rad52* yeast mutants unable to complete HO-induced mating type switching were found to escape lethality by deletion of their HO cut site and subsequent ligation.[297] However, chromosome loss of up to 700bps has been observed flanking the break sites in *rad52* yeast mutants with a significant number of events with 2 or more nucleotides of homology at a given junction.[280] Hamster cell lines deficient in Ku80 or Xrcc4 exhibit sites of microhomology at DSB junctions implicating that their use for efficient alignment is compensated with the stability of microhomology base pairing in their absence.[298] About 40%[299] of large deletions in human disorders including cancers such as retinoblastoma[300] and bladder cancer[301] have been characterised by the presence of very short sequence homologies (2 to 6bp) at the breakpoints.[302,303] This alternative pathway is independent of canonical NHEJ factors and is referred to as Microhomology Mediated End Joining (MMEJ). It favours the repair of DSBs through extensive deletion but may also result in insertions. Interestingly, regions of microhomology have been detected within telomere fusion junctions in telomerase null mice,[209] in human cells undergoing crisis in culture[255] and tumour cells.[304]

It has been postulated that Ku-independent MMEJ requires a protein that is involved in the detection and alignment of DNA microhomology prior to the resection of the 3'overhangs. It has been suggested that the MRN complex may be associated with this process. Notably, Mre11 in conjunction with Rad50 may function to maintain a synapsis between the two ends. Telomere fusions in Mre11 deficient ATLD (Ataxia-Telangiectasia Like Disorder) cells carry short tracts of nucleotide insertions.[305] It has been implicated that deficiencies in Mre11 may promote further genetic instability in a MMEJ dependent manner. Instead a translesion polymerase may extend an annealed sequence using templated error prone synthesis which may subsequently realign at another site of microhomology following dissociation and potentially result in genetic amplification.[306] Templated nucleotides have also been detected in *Drosophila* mutants deficient in Rad51 and DNA ligase IV. These

repeats shared similarity with adjacent DNA and it was speculated that they are generated in an attempt to create microhomology that can be utilised for end joining.[307]

Poly (ADP-ribose) Polymerase I [PARP-1] becomes activated by single strand DNA breaks and its activity has been proposed to promote the synapsis and ligation of double strand breaks in MMEJ. The DNA ligase III/XRCC1 complex previously implicated in Base Excision Repair/ Single-Strand Break Repair (BER/SSBR) has also been proposed to play a role in MMEJ.[308] Elevated levels of DNA ligase III has been detected in BCR/ABL$^+$ CML (Chronic Myeloid Leukaemia) cells accompanying the down-regulation of the NHEJ factors, DNA ligase IV and Artemis. The majority of DSBs (80%) in these cells were repaired using microhomology at the break junctions; moreover a reduction of DNA ligase III resulted in a decrease in the frequency and size of microhomology.[309] DNA ligase III has been postulated to form a complex with WRN, a 3' to 5' exonuclease of which its depletion has been found to cause the premature aging syndrome, Werner.[310] Notably, a knockdown of WRN was observed to increase the size of deletions at repair junctions and thus it was proposed that WRN may play a role in limiting the extent of resection that takes place during DNA repair.[309]

The efficiency of PARP-1 DSB synapse formation can be modulated by the stability of the microhomology overhang.[308] An overhang composed of G:C bases is repaired much more efficiently.[308,311] Furthermore, the dependence of canonical NHEJ in DSB repair is inversely correlated to the G:C content of the overhang suggesting that NHEJ is mostly required when the annealing of the overhangs is not energetically favourable.[311] Accordingly, it has been proposed that NHEJ catalyses the synapsis of short overhangs through protein-protein interactions in which the energy provided by base pairing is limited.[311]

Saccharomyces Cerevisiae utilises the Rad10-Rad1 endonuclease to remove the 3'overhang prior to ligation for single strand annealing (SSA) and MMEJ DNA repair.[311] The mammalian ortholog ERCC1/XPF complex has been demonstrated *in vivo* to remove the 3'flap in mammalian cells in a pathway that is independent to the canonical NHEJ.[312] Furthermore, the ERCC1/XPF complex facilitates DSB repair by removing the 3'overhang at the end of uncapped telomeres[275] raising the possibility that telomere fusions may arise in an alternative pathway to NHEJ. Accordingly, the frequency of telomere fusion events has been shown to reduce significantly in cells deficient for both TRF2 and either ERCC1 or XPF.[275]

1.29 Telomeres and Cancer

Telomeres within human somatic cells do not express telomerase and will continue to shorten at 50 to 100bp per cell division.[294] Progressive telomere shortening will initiate replicative senescence in which cells stop dividing disenabling further telomere loss. It has been proposed that the onset of senescence is triggered by the shortest telomere within a distribution rather than the global average.[209] Consistent with this, an accumulation of short telomeres has been observed in cells that are approaching senescence.[313] A decline in the 'capping' function of the telomere by loss of TRF2 and POT1 or when telomeres become critically short can compromise chromosome end protection and initiate an ATM (Ataxia-Telangiectasia Mutated) or ATR (ATM and Rad3-related) kinase dependent DNA Damage Response (DDR) respectively.[202] Activated ATM or ATR kinase phosphorylate regions of histones H2AX flanking into sub-telomeric regions to generate DNA damage foci at dysfunctional telomeres (Telomere dysfunction Induced Foci [TIFs]).[314,315] Cells respond to TIFs by either entering into a p21 mediated permanent senescent state or become proapoptotic in response to p53.[314] Consistent with this, *hTERT* immortalisation reduces the generation of TIFs in human fibroblasts in culture.[316]

Efficient maintenance of cell cycle arrest in senescent cells depends on the continued activity of the DDR components since negatively interfering with their actions can restore cell cycle progression into the S phase of the cell cycle.[317,314] Accordingly, carcinogen or virus-induced transformation can enable cells to acquire an extended lifespan and overcome cellular senescence.[318] Post-senescent cellular replication will drive further telomere loss that may denude telomeres of all TTAGGG repeats. It has been implicated that telomere dysfunction may provide a stepwise accumulation of cytogenetic changes during cancer development through the formation of dicentric chromosomes by telomere-telomere fusion. Dicentric chromosomes that are pulled apart during anaphase cause chromosome breakage and new recombinogenic free ends that can initiate and continue Breakage-Fusion-Bridge (BFB) cycles.[319] Continuous BFB cycles can generate a wide spectrum of novel genetic alterations including non-reciprocal translocations (NRTs). This paradigm has been previously implicated in many human solid tumours including renal cell carcinoma,[320] pancreatic carcinoma and osteosarcomas.[321] Consequently, an accumulation

of non-reciprocal translocations (NRTs) may facilitate in carcinogenesis through positive selection.[322]

Although, the vast majority of cells enduring genetic instability will cease to divide and die; a small proportion of cells may acquire a mechanism that enables telomere length stability.[318]

Telomerase is expressed in >90% of human cancers.[180] However, it is not expressed to excess in the majority of cases. Consequently, telomeres are either the same or shorter than those in adjacent normal tissue. Thus, it has been suggested that telomerase is required to maintain telomere function.[323] Consistent with this, telomerase activity has shown to reduce the frequency of end-to-end chromosome fusions in the presence of critically short telomeres.[212] It has been described that telomere dysfunction could promote carcinogenesis prior to telomerase activation. Accordingly, numerous end-to-end fusions and signal free ends have been observed in metaphase spreads derived from late generation $mTR^{-/-}$ $p53^{-/-}$ mutant mice.[322]

Telomerase activation or elongation via homologous recombination can provide telomere length and genome stability to facilitate the outgrowth of a sub-clonal population with an altered phenotype.

This model suggests that telomere dysfunction may initiate the neoplastic process to oncogenesis[322] and the shortest telomere within the distribution may acquire this role through chromatid-chromatid or inter-chromosomal fusion following replication.[209]

Therefore, it can be assumed that replicative senescence functions as a protection mechanism against the initiation of BFB cycles following the generation of dicentric chromosomes.[294]

Accordingly, deficiencies in ATM can cause disorders that fail to suppress DNA replication following DSBs consequently predisposing patients with a higher incidence of malignancies due to an elevated frequency of genetic instability and chromosomal translocations.[138,324]

1.30 Telomeres and Haematological Disorders

Despite being telomerase competent; $CD34^+$ haematopoietic cells show a reduction in telomere length with age[325] losing up to 33bp per year.[326] Consistent with this, a transient

phase of rapid telomere shortening has been observed following an allogeneic bone marrow transplant prior to stabilisation and haematopoietic reconstitution. It was also observed that the extent of telomere shortening is inversely correlated with the number of mononuclear cells received by the recipient.[327] This suggests that the proliferative pressure is less intense with a larger population of progenitor cells. Thus, increased telomere attrition may cause the diminished capacity of haematopoietic cell reserves through senescence or apoptosis. Furthermore, telomere dysfunction could lead to the acquisition of clonal translocations that could promote leukaemogenesis. Accordingly, late clonal disorders, such as MDS or acute leukaemia have noted to arise from a bone marrow transplant or high dose chemotherapy following the successful elimination of a former leukaemia, lymphoma or solid tumour.[328]

Disease anticipation in families with Dyskeratosis Congenita (DC) is associated with telomere shortening.[329] Mutated telomerase in DC fails to maintain telomere function causing accelerated exhaustion of stem cell reserves and consequential bone marrow failure. Marked similarities of clinical phenotypes observed in DC patients have been documented in $mTR^{-/-}$ mice where an age dependent compromise of haematopoietic reconstitution is observed following HSC ablation.[330] Furthermore, a high incidence of teratocarinomas and lymphomas have been described in $mTR^{-/-}$ mice as well as reduced tumour latency following successive $mTR^{-/-}$ generations.[330] Accordingly, it has been documented that a 196 fold increase in AML predisposition is present in patients with DC.[331] It has been proposed that the limited level of telomerase activity within the haematopoietic system targets short telomeres in an act to maintain the stability of the genome of the highly proliferative bone marrow.[212]

Terminal restriction fragment (TRF) length analysis has demonstrated that telomere shortening is a frequent observation in MDS and AML relative to aged matched controls[332-334] with the observation of significant telomere attrition in AML.[335,336] Although this may indicate the mitotic history of the disease; telomere shortening has been associated with complex chromosomal rearrangements and poorer prognoses in a number of haematological disorders that include acquired Aplastic Anaemia (AA),[337,338] Chronic Myeloid Leukaemia (CML),[339,340] and Chronic Lymphocytic Leukaemia (CLL)[341] and MDS/AML.[333,334,342] However, telomere dysfunction is not coupled to the global average

length within the cell. Instead, it is the shortest telomere within the distribution that is vital for cell viability and chromosome stability. Therefore, the shortest telomere within the distribution can initiate chromosomal fusions[209] that may play a role in the pathogenesis of MDS and its subsequent transition to AML.

The activation of ATM/ATR regulated checkpoint pathways may collectively contribute to ineffective haematopoiesis in this disease and the ‘latency’ period prior to AML, in which a compromised DDR enables clonal expansion. Consistent with this, a DDR response that is commonly present at pre-invasive stages of major human cancers becomes abrogated with development. This concept has been applied to breast, colon, lung and bladder carcinomas.[343] Furthermore, disease progression in MDS has been associated with inactivation of components required for a DDR. LOH at the *ATM* locus has been previously observed in AML[344] and the loss of *p53* has been implicated to result in significantly shorter survival, AML transformation and enhanced resistance to chemotherapy.[135] Additionally, *p53* mutations have been found to accompany abnormal cytogenetics.[136] An ‘uncoupling’ effect has also been observed between functional ATM and its DDR response in AML. A positive correlation was observed between ATM activation and the number of blast cells present in the bone marrow; however downstream targets were uncoupled from inactive checkpoint kinases (chk-1 and chk-2) in AML enabling cells to evade cell cycle checkpoints and apoptosis.[345] This is consistent with the presence of a DDR response during the pre-malignant stage which becomes down-regulated with transformation.

Short telomeres have also been associated with low haemoglobin concentrations and multiple cytopenias[332] including a significant negative correlation between telomere length and the number of blasts in the bone marrow.[336] Consistent with this, progressive telomere shortening has been associated with a consensual decrease in apoptosis of MDS $CD34^+$ cells.[346] It was proposed that in the absence of functional cell cycle checkpoints prolonged telomere reduction can initiate genetic instability which may promote blast cell proliferation and leukaemogenesis through positive selection.

Most MDS patients have normal to low levels of telomerase[332] that can increase by 18 fold in AML.[335] Consistent with this; RAEB-T associated with high telomerase activity has shown progression into AML shortly after initial diagnosis.[332] Furthermore, it has been

demonstrated that *hTERT* expression[347] and telomerase activity[342] are more pronounced in AML with complex rearrangements associated with the loss or gain of chromosome parts. Low levels of telomerase have been demonstrated to correlate with telomere shortening during the *ex vivo* expansion of MDS CD34$^+$ cells. Despite telomerase activity, telomere maintenance is insufficient to completely prevent telomere loss in cells that are capable of increased cycling, i.e. 20 fold. However, increased telomerase activity in AML may be vital in preventing excessive telomere attrition in malignant cells that can proliferate up to 45 to 50 fold of normal haematopoietic cells.[335]

Despite an accumulation of telomerase competent blast cells in late MDS and AML; increased telomerase activity may also be a feature of the malignant phenotype that is vital to stabilise the genome acquired in AML cells. Thus, the reactivation of telomerase could occur as a late genetic event to promote colony formation, as reported in the blastic phase in CML.[340]

Telomere shortening may not only represent a consequence but also be a predisposing factor for the development of MDS. The increased cell turnover in this disease may promote replicative senescence and genetic instability during the malignant transformation of haematopoietic cells.[348]

1.31 Telomere Length Analysis in MDS/AML

Telomere length has been analysed using various approaches in MDS and AML. Such techniques have included Terminal Restriction Fragment (TRF), Quantitative-Fluorescence *in situ* Hybridisation (Q-FISH), Flow-Fluorescence *in situ* Hybridisation (Flow-FISH) and Quantitative-Polymerase Chain Reaction (Q-PCR). Some of the studies and the results obtained regarding telomere length in MDS and AML have been discussed:

Telomere length was measured using Flow-Fluorescence *in situ* Hybridisation (Flow-FISH) on fractionated peripheral blood (PB) granulocytes and CD34$^+$ cellular populations derived from 55 MDS patients (40 at diagnosis; 15 derived at 6 to 12 months following diagnosis).[346] In contrast to what was observed in normal controls, telomere fluorescence did not decline with age in MDS derived PB granulocytes and CD34$^+$ cells. Moreover, telomere fluorescence was shorter amongst MDS patients, particularly in patients presenting with

intermediate/unfavourable karyotypes and IPSS Intermediate-2/High-risk scores. The degree of apoptosis in the $CD34^+$ population was directly correlated with telomere fluorescence and thus it was proposed that cell cycle checkpoints had been bypassed in these cases. Consistently, the extent of apoptosis was inversely correlated with the percentage of bone marrow blasts. This study observed that patients presenting with lower telomere fluorescence had significantly worse prognostic characteristics and suggested that telomere length might represent an independent prognostic factor that could be linked to an unfavourable outcome.

Telomere dynamics has been analysed using Q-FISH analysis in a subset of 13 MDS patients at diagnosis.[349] Telomere length was measured at the single cellular level and compared to telomere measurements derived from TRF analysis. The frequency of the telomere fluorescence intensity in each metaphase was summed to determine patterns of telomere length. In healthy aged-matched individuals, telomere length distributions were wide and skewed more towards longer telomeres. In contrast telomere length distributions derived from MDS metaphases were narrow and accumulated at shorter telomere lengths. Telomere length measurements analysed by TRF and Q-FISH analysis were comparable in normal controls. In contrast, telomere length measured by Q-FISH was shorter in MDS patients when compared to TRF analysis. Accordingly, MDS patients presented with telomere length in the range of 3.2kb to 17.5kb by TRF analysis; however peak telomere fluorescence values only ranged between 3.6kb to 4.8kb. MDS marrow cells contain at least two populations; one derived from the MDS clone and the other from normal background cells therefore, it was speculated that the superimposition of normal cells (with potentially longer telomeres) may be accountable for the longer telomeric length in TRF measurements. This study suggested that short telomeres in MDS may play a role in the pancytopenia that is feature of MDS patients and further proposed that some MDS cells may survive and divide with short telomere lengths, particularly since telomere signals accumulated at short length in MDS metaphases. Accordingly, it was speculated that telomerase activation may enable these cells to escape checkpoint mechanisms and enable clonal expansion.

Telomere length, telomerase activity and mRNA expression of critical telomeric proteins was analysed in age-matched groups of AML (>90% blasts) with distinct chromosome

aberrations namely; AML with three or more gains and/or losses of chromosome parts, terminal deletions or non-reciprocal translocations (13 patients) and two control groups presenting with chromosomal abnormalities unlikely related to telomere dysfunction; AML with reciprocal translocations or inversions (17 patients) and AML without chromosome aberrations (8 patients).[342] The aim of the study was to determine if telomere shortening (analysed using Q-FISH) played a role in generating extensive chromosomal instability. Telomere length was significantly shorter in AML with gains and/or losses in contrast to the control groups. Moreover, Q-FISH revealed extensive telomere shortening on individual chromosome arms in this subset of patients. It was speculated that critically short telomeres in these cells may have a role in generating chromosomal instability. Control groups presented with comparable telomerase activity to that within normal marrow however, AML patients with gains and/or losses showed elevated telomerase activity by 3 to 8 fold. It was speculated that telomere elongation by telomerase may be repressed by TRF1. Notably, high expression of *TRF1* mRNA was observed in AML with gains and/or losses when compared with AML patients in control subgroups. This study suggested that critical telomere shortening and subsequent telomere dysfunction in AML patients may contribute to the development of chromosomal abnormalities that are detected in patients with gains and/or losses. It was further proposed that telomerase up-regulation in patients with gains and/or losses may indicate that these cells have already bypassed cellular crisis.

Telomere length was analysed using Quantitative-Polymerase Chain Reaction (Q-PCR) in 167 *de novo* paediatric AML patients presenting with at least 80% leukaemic cells at diagnosis in order to investigate if short telomeres were associated with AML characteristics and whether they contributed to an adverse outcome.[350] Telomere length was not associated with age in AML patients. Moreover, extensive telomere shortening was observed among leukaemic cell populations in contrast to PB leukocytes derived from healthy controls. Telomere length was not significantly different among FAB subgroups and was not associated with the number of cytogenetic aberrations (numerical or structural), including complex karyotypes ($\geq$3). Q-PCR fails to measure specific telomere length and therefore the authors considered that the chromosomal instability identified in these cases may be influenced by the shortest telomere within the cell. AML patients harbouring the FLT3/ITD mutation had shorter telomeres than patients without FLT3/ITD implicating that this might

reflect an extended mitotic history of leukaemic cells within this subgroup of patients. Telomere length did not influence the overall survival, event-free survival, the incidence of relapse or response to treatment. Accordingly, short telomeres did not appear to correlate with an unfavourable outcome.

Telomere parameters were analysed in 137 patients diagnosed with *de novo* or secondary AML presenting with at least 80% blasts at diagnosis.[347] Telomere length was measured using flow-FISH in peripheral blood (PB) granulocytes and compared with telomere measurements derived from PB granulocytes from a control group of healthy individuals. Telomere length was significantly shorter in AML patients when compared to aged-matched healthy controls and no correlation was identified between telomere length and disease status, i.e. *de novo* AML or secondary disease. Patients presenting with a high FLT3 ratio (mutant FLT3/ wild type ratio) had significantly shorter telomeres than those with a low FLT3 ratio and this was particularly pronounced in individuals exhibiting a normal karyotype. Patients with multiple cytogenetic abnormalities had the shortest telomeres when compared to those presenting with a normal karyotype. Moreover, the degree of chromosomal abnormalities correlated with *hTERT* expression. It was suggested that extensive telomere loss may contribute to the development of genetic instability in this subset of patients and that telomerase up-regulation permits the expansion of the AML clone by preventing cellular senescence. Telomere length did not have a significant influence on treatment response, overall survival or disease-free survival in AML patients.

Telomere length was evaluated by TRF analysis in bone marrow and peripheral blood cells obtained from 50 patients presenting with *de novo* MDS and AML (arising from MDS) and 21 untreated *de novo* AML.[336] The aim of the study was to determine whether telomere shortening was associated with the progression of early to advanced MDS and subsequently towards acute leukaemia. Patients presenting with shorter TRFs than aged-matched controls increased with MDS progression such that 30% of early RA/RARS, 65% of advanced RAEB/RAEB-T and 80% of AML patients (secondary from MDS) had reductions in TRF length. Patients with *de novo* AML presented with the shortest TRF measurements. A significant negative correlation was detected between TRF measurement and the percentage of blasts. Moreover, a significant negative correlation was observed between TRF and IPSS score and thus it was suggested that telomere length may serve as a useful prognostic variable in the

assessment of risk and facilitate in making therapeutic decisions. Shorter TRFs were also detected amongst patients presenting with an abnormal karyotype when compared to patients with a normal karyotype. However, this difference was only significant in early MDS and diminished in advanced MDS and AML.

Telomere length was measured using TRF analysis in 93 patients with MDS at the time of diagnosis[332] and compared with TRF measurements derived from aged-matched controls. It was observed that 57% of patients presented with TRFs of aged-matched normal range whereas 38% and 5% had shortened TRFs and elongated TRFs, respectively. TRF measurements were not associated with FAB subgroups such that 34% of RA/RARS and 47% of RAEB-T patients showed shortened TRFs relative to age. 66% of patients with shortened TRFs presented with an abnormal karyotype in contrast to only 34% of patients with normal range TRFs. MDS patients with shortened TRFs had a significantly low haemoglobin concentration, a high marrow blast percentage, multiple cytopenia and poor cytogenetic changes. Additionally, the incidence of leukaemic transformation was significantly higher in MDS patients with shortened TRFs compared to those with normal range TRFs. Patients presenting with intermediate-2/ high IPSS scores had significantly shorter TRFs than those with low/ Intermediate-1 IPSS scores. It was also identified that patient outcome was significantly influenced by TRF length such that patients presenting with normal range TRFs had a favourable prognosis in contrast to those with abnormal TRFs (shortened/elongated TRFs). The majority of MDS patients analysed had normal-to-low levels of telomerase activity. Therefore, it was suggested that accelerated telomere erosion due to rapid cell division is not restored in patients presenting with shortened TRF measurements because telomerase activity is insufficient to maintain telomere length. 2 RAEB-T patients showed high telomerase activity and subsequently developed acute leukaemia shortly after diagnosis. Thus, it was proposed that accelerated telomere erosion may be an early event in MDS pathogenesis and that telomerase reactivation may be a late genetic event that enables AML transformation. This study suggested that TRF measurements at the time of diagnosis may refine individuals according to risk.

1.32 Aim of Research

High resolution methodology will be utilised to analyse telomere length in a cohort of 80 MDS and 144 AML patients. Single Telomere Length Analysis (STELA) is a long-range single molecule PCR approach that has the ability to amplify the double-stranded region of telomere repeats from specific chromosome ends.[171] Telomere length analysis has previously been employed on MDS and AML however, STELA has the ability to detect the full spectrum of telomere lengths and is able to identify critically short telomeres which have the potential to induce replicative senescence or initiate cycles of telomere fusion and breakage.[255,304,351]

STELA will be used to analyse whether telomere length at the XpYp and 17p telomere is associated with specific clinical parameters in MDS and AML. Such parameters in MDS will include age, marrow blast percentage, peripheral blood cytopenia, cytogenetics, International Prognostic Scoring System (IPSS) and overall survival. Whereas age, gender, marrow blast percentage, presenting white blood cell (WBC) count, AML type (*de novo*/ secondary), WHO performance status, FLT3 and NPM1 mutation status, response after first cycle of intensive chemotherapy and finally clinical outcome data including overall survival and disease-free survival will in analysed in AML.

Additionally, specific features including atypical Telomere Rapid Deletions (TRDs), bimodal populations and Telomeric-Loss of Heterozygosity (Telomeric-LOH) will be checked for within STELA profiles. In the event of detecting bimodal populations at the XpYp telomere, allele-specific STELA will be employed by using previously characterised heterozygosities in the sub-telomeric region of XpYp.[244] This will facilitate in determining whether this observation is a reflection of differing maternal and paternal contributions in the zygote. The non-functional Telomere Variant Repeat (TVR) region of the XpYp telomere will also be measured using TVR-PCR in order to calibrate the STELA data such that the length of the uninterrupted tandem repeats of TTAGGG can be established.

Further studies will include TRAP (Telomere Repeat Amplification Protocol) to measure telomerase activity between a cohort of MDS and AML patients and the rate of telomere loss will be analysed in a small subset of MDS patients in which serial samples are available. Magnetic Dynabeads specific for the $CD34^+$ antigen will be utilised to separate the $CD34^+$

and $CD34^-$ fractions derived from a set of 20 MDS patients and STELA will be subsequently performed to analyse the telomere length within each fraction so as to determine whether telomere erosion is more pronounced in the more primitive $CD34^+$ population.

In order to analyse the extent of telomere dysfunction in MDS and AML, a PCR-based telomere fusion assay will be used to detect and quantify single telomere-telomere fusion molecules between specific chromosome ends. Putative single fusion molecules will be re-amplified with nested PCR primers to identify the participating telomeres and coexisting fusion junction. Finally, the nature of telomere fusions will be characterised using direct sequencing analysis.

Chapter 2:

Materials and Methods

2.1 Patient samples

Bone marrow mononuclear cells (BMMNC) derived from 80 MDS (median age 70 years; range 21 to 86 years) and 144 AML patients were collected at the time of diagnosis. All samples were acquired after written informed consent was obtained from each patient. Comprehensive clinical information available for the MDS and AML cohorts are listed in Table 2.1 and 2.2, respectively of which demographic information was only available for 110 of these AML cases who had a median age of 60 years (range 17 to 82 years). MDS samples were kindly provided by Dr Paul Baines from a local tissue bank held in the Department of Haematology, Cardiff University, School of Medicine or were paid for via the Dundee tissue bank. AML samples were from the Cardiff AML Biobank, Cardiff University, School of Medicine and were derived from patients enrolled in the Medical Research Council (MRC) trials AML 14, 15, 16 and 17.

Of the 144 AML samples analysed in this study, mean XpYp telomere length ±SD was provided for 87/144 patients by Bethan Britt-Compton. In addition, a comparator dataset derived from peripheral blood leukocytes from 68 healthy individuals (median age 31.5 years; range 0 to 72.5 years) was provided by Duncan Baird.

Serial samples were available from 4 MDS patients; these samples were analysed to quantify temporal telomere erosion.

Table 2.1: Demographic information of MDS patients who were analysed within the current study.

Demographic information of MDS patients

MDS patients: n = 80; Median age 70 years (range: 21 years to 86 years)

Gender: n = 80

Male: n = 50   Female: n = 30

French-American-British (FAB) Score: n = 29

RA/RARS: n = 17

RAEB: n = 10

RAEB-T: n = 2

Cytopenia Number: n = 71

1 Cytopenia: n = 26

2 Cytopenia: n = 17

3 Cytopenia: n = 28

Cytogenetic Profile: n = 47

Good: n = 23

Intermediate: n = 10

Poor: n = 14

International Prognostic Scoring System (IPSS Score): n = 72

Low: n = 20

Intermediate-1: n = 14

Intermediate-2: n = 6

High: n = 32

Table 2.2: Demographic information of AML patients who were analysed within the current study.

<u>Demographic information on AML patients</u>

Total AML patients: n = 144; Demographic information available for 110 of 144 patients. Median age 60 years (range: 17 years to 82 years)

Gender: n = 110

Male: n = 57   Female: n = 53

AML type: n = 110

De novo: n = 95   Secondary: n = 15

Bone Marrow Blast %: n = 80

1st Quartile ≤50.25%: n = 20

2nd Quartile >50.25% to ≤78%: n = 20

3rd Quartile >78% to ≤90.75%: n = 20

4th Quartile >90.75%: n = 20

Presenting White Blood Cell (WBC) count ($x10^9$/l): n = 110

1st Quartile ≤11.45: n = 27

2nd Quartile >11.45 to ≤30.50: n = 28

3rd Quartile >30.50 to ≤76.53: n = 28

4th Quartile >76.53: n = 27

Cytogenetic Profile: n = 94

Favourable: n = 7   Intermediate: n = 83   Adverse: n = 4

World Health Organisation (WHO) Performance Status (PS): n = 110

PS 0: n = 61   PS 1: n = 34   PS 2: n = 7   PS 3: n = 8

Mutation Status

$FLT3/ITD^+$: n = 46 $FLT3/ITD^-$: n = 53 $FLT3/TKD^+$: n = 19 $FLT3/TKD^-$: n = 65

$NPM1^+$: n = 52 $NPM1^-$: n = 44

2.2 DNA Extraction

As described previously,[171] DNA was extracted from bone marrow mononuclear cells using standard Proteinase K, RNase A and phenol/chloroform protocols. Accordingly, samples stored at -80°C were thawed quickly in a 37°C water bath and centrifuged at 6000g for 5 minutes. The supernatant was removed and details belonging to each sample were noted. Cells were re-suspended in 1ml of 1xPBS (Phosphate Buffered Saline) and centrifuged for a second time at 6000g for 5 minutes. PBS was removed. This process was repeated until the medium was removed.

20ml of Lysis buffer was made containing 100mM NaCl, 10mM Tris-HCl (pH 8.0), 5mM EDTA (pH 8.0), 0.50% SDS and filtered H_2O. Lysis buffer was added to the cell pellet in accordance to the number of cells. Notably, 300µl, 500µl or 1ml of buffer was added to the sample if the cell population exceeded $1x10^5$, $3x10^6$ and $28x10^6$, respectively. 3µl, 5µl or 10µl of RNAse A (10mg/ml) and Proteinase K (20mg/ml) were subsequently added to each sample if the cellular population exceeded $1x10^5$, $3x10^6$ and $28x10^6$, respectively. Samples were incubated at 45°C in a hot block overnight.

Cells were centrifuged briefly and phenol/chloroform extraction commenced. 500µl of ultra-high purity phenol/chloroform was added and then each sample was placed on a tube rotator for 20 to 30 minutes. Samples were subsequently centrifuged at 13K for 4 to 6 minutes.

The top aqueous and interphase layers were transferred into a fresh tube containing 500µl of phenol/chloroform. Samples were rotated for a second time for 20 to 30 minutes and centrifuged at 13K for 4 to 6 minutes. This process was repeated if the phenol/chloroform layer remained cloudy. Finally, the aqueous phase of each sample was transferred into a fresh tube.

3M NaOAc (pH 5.2) was added to each sample. The 1:10 volume was required; salt to lysis buffer, i.e. used 50µl of NaOAc if 500µl lysis buffer was used. Ice cold 100% Ethanol was subsequently added and samples were left to stand on ice to enable DNA precipitation.

The DNA was pelleted by centrifugation at 13K for 1 minute and the ethanol was removed. Residual salt was washed away by leaving the DNA pellet to stand in 70% ice cold ethanol

over ice for 10 minutes. Samples were centrifuged at 13K for 1 minute and supernatant was removed. Finally, the DNA pellet was allowed to air dry and re-suspended in 50µl, 25µl or 15µl 10mM Tris-HCl (pH 8.0) depending on whether the pellet was large, medium or invisible respectively and placed in the fridge overnight.

DNA was quantified in triplicate by Hoechst 33258 Fluorometry (BioRad). The instrument was warmed up for approximately 30 minutes prior to calibration. A 20ml working solution composed of 1xTNE and 0.001mg/ml Hoechst was prepared in filtered H_2O and mixed.

Standards (i.e. blank and calf thymus) and patient samples were prepared in 1ml working solution. Notably, 5µl of calf thymus DNA (100ng/µl) was utilised to achieve a standard of 500ng. The fluorometer was calibrated and 1µl of patient samples were measured in triplicate. Finally, samples were diluted in 10mM Tris-HCl (pH 8.0) to 10ng/µl.

2.3 Single Telomere Length Analysis (STELA)

10ng/µl genomic DNA diluted to 250pg/µl in a 40µl reaction containing 1µl Telorette-2 linker (10µM) and 38µl 10mM Tris-HCl (pH 8.0). Multiple PCRs (typically 6 reactions per sample) were carried out for each test DNA in 10µl containing Taq Buffer (x1), 2mM of $MgCl_2$, 1.2mM dNTPs, 0.5µM Teltail, 0.5µM XpYpE2 or 0.5µM 17pseqrev1 (XpYp or 17p telomere adjacent primers) or 0.5µM XpYp-427G/415C or 0.5µM XpYp-427A/415T allele-specific primers, 0.5 Units of a 10:1 mixture of Taq (Thermoscientific) and Pwo Polymerase (Roche) and 250pg diluted DNA. Reactions were cycled with an MJ PTC-225 thermocycler (MJ research) under the following conditions: 22 cycles of 94°C for 20secs, 65°C (XpYpE2) or 66.5°C (XpYp-427G/415C and XpYp-427A/415T allele-specific primers) or 59°C (17pseqrev1) for 30secs, and 68°C 8 minutes. Allele specific haplotypes (AT or GC) were identified by genotyping of the XpYp telomere-adjacent DNA prior to STELA.

Products were resolved by 0.5% Tris-acetate-ethylenediaminetetraacetic acid agarose gel electrophoresis. Prior to blotting, the DNA was depurinated in 0.25M HCl (Hydrochloric Acid) and denatured in 1.5M NaCl (Sodium Chloride) and 0.5M NaOH (Sodium Hydroxide).
Products were identified by two separate Southern blot hybridisations using a random-primed [α-^{33}P]dCTP-labelled (GE Healthcare) telomere adjacent probe and TTAGGG repeat

probe. Probes that specifically target the 1kb (Stratagene) and 2.5kb (BioRad) molecular weight markers were also used in each hybridisation reaction.

2.4 Hybridisation

15ml of Church Buffer (1% BSA, 7% SDS, 1mM EDTA, 0.5M $NaHPO_4$ [pH 7.2]) was transferred into a hybridisation bottle before being placed into a hybridisation oven set at 65°C. The DNA blot was washed in H_2O and inserted into the bottle. To ensure sufficient buffer coverage of the membrane and block of non-specific binding sites the membrane was pre-hybridised prior to adding the DNA probe.

DNA was labelled using random oligomer labelling upon using the GE Healthcare DNA labelling beads (composed of buffer, dATP, dGTP, dTTP, random nonamers [$d(N)_9$] and the Klenow fragment (*E. coli* DNA polymerase I). 25ng of probe was made up in a volume of 45μl 1xTE buffer (pH 8.0). The probe was denatured at 96°C on a heating block for 5 minutes and subsequently snap-cooled for 5 minutes over ice. In a radiation cabinet, 4μl of [α-^{33}P]dCTP (GE Healthcare) was added to the probe and the mixture was re-suspended. The probe mixture was placed into a 37°C water bath for at least 30 minutes before transferring 1μl of a pre-made 1kb/2.5kb probe marker. 50μl of H_2O was added to the probe mixture and re-suspended to ensure efficient mixing. Finally, the probe was denatured at 96°C on a heating block for 5 minutes and 25μl was inserted into the centre of the pre-hybridised bottle. The bottle was placed back into the oven overnight.

After hybridisation, excess probe was washed at high stringency from the nylon membrane in a buffer solution containing 0.1% SSC (saline-sodium citrate) and 0.1% SDS (sodium dodecyl sulphate).
The hybridised fragments were detected by phorphorimaging with a Molecular Dynamics Storm 860 phosphorimager (GE Healthcare) and the molecular weights of DNA fragments were calculated using the Phoretix 1D quantifier (Nonlinear Dynamics).

2.5 Fusion Assay

The telomere fusion assay was carried out as described previously.[255] Multiple PCRs (typically 18 reactions per sample) were carried out for each test DNA in 10μl containing Taq Buffer (x1), 2mM $MgCl_2$, 1.2mM dNTPs, 0.5μM XpYpM, 0.5μM 17p6, 0.5μM 21q1, 0.5 Units Taq/Pwo (10:1) and 100ng gDNA. PCR reactions utilised the XpYpM and 17p6 primers in conjunction with the 21q1 primer. The 21q1 primer targets multiple chromosome ends which share homology with the 21q sub-telomere allowing the simultaneous detection of 13 chromosome ends including 21q, 1q, 2q, 5q, 6q, 6p, 8p, 10q, 13q, 17p, 19p, 19q, 22q as well as the 2q13 interstitial telomeric locus.[351]

Products were resolved by 0.5% Tris-acetate-ethylenediaminetetraacetic acid agarose gel electrophoresis and identified by Southern Hybridisation with random-primed [α-^{33}P]dCTP labelled (GE Healthcare) XpYp and 17p telomere-adjacent probes; XpYpOG and 17p7, respectively. Probes were generated by PCR as described previously.[255,351]

Probes that specifically target the 1kb (Stratagene) and 2.5kb (BioRad) molecular weight markers were also used in each hybridisation reaction. The hybridised fragments were detected by phorphorimaging with a Molecular Dynamics Storm 860 phosphorimager (GE Healthcare).

Taking into account the DNA content of a single diploid cell at 6pg; the frequency of fusion could be calculated from the number of positive molecules compared with the number of input molecules. Further hybridisation was performed with the 21q1-adjacent probe;[351] however since this probe yields additional non-specific products the fusion frequency was not quantified.

Fusion reactions were re-amplified using nested PCR primers (XpYpMb, 17p6b and 21q1C). Each re-amplification reaction (30μl) contained Taq Buffer (x1), 2mM $MgCl_2$, 1.2mM dNTPs, 0.5 Units Taq/Pwo (10:1) and a combination of 0.5μM XpYpMb, 0.5μM 17p6b or 0.5μM 21q1C. Fusion reactions were diluted by 1:20 in H_2O and 3μl of the product was added to the re-amplification reaction.

Products were resolved by 0.7% Tris-acetate-ethylenediaminetetraacetic acid agarose gel electrophoresis and were subsequently gel-purified (QIA quick Gel Extraction Kit; Qiagen) for direct sequencing analysis using BigDye 3.1 (Applied Biosystems) in conjunction with designed primers specific to the target sequence.

2.6 TVR Mapping

TVR-PCR was used to determine the distal extent of the XpYp TVR region. Three separate 10μl reactions per sample contained Taq Buffer (x1), 3mM $MgCl_2$, 1.2mM dNTPs, 1μM of the XpYpE2 primer, 1μM of the TVR primers: TTAGGG (TAG-TelW), TGAGGG (TAG-TelX) or TCAGGG (TAG-TelY), 100ng DNA and 0.5 Units Taq polymerase (no Pwo). Each PCR reaction was cycled 20 times at 96°C for 20secs, 67°C for 30secs and 70°C for 3 minutes.

Products were resolved by 0.9% Tris-acetate-ethylenediaminetetraacetic acid agarose gel electrophoresis and detected by Southern blot hybridisation using a random-primed [α-^{33}P]dCTP-labelled (GE Healthcare) probe using the XpYpEB probe generated by PCR between primers XpYpE2 and XpYpB2. Probes that specifically target the 1kb (Stratagene) and 2.5kb (BioRad) molecular weight markers were also used in each hybridisation reaction.

The hybridised fragments were detected by phorphorimaging with a Molecular Dynamics Storm 860 phosphorimager (GE Healthcare).

2.7 CD34 Cell Purification using Magnetic Beads

CD34 cells were purified from a subset of 20 MDS patients using the Dynal CD34 Progenitor Cell Selection System (Invitrogen). 100μl of re-suspended CD34 Dynabeads were transferred into a 5ml FACS tube and washed in 900μl of buffer (PBS with 0.1% Bovine Serum Albumin (BSA)). The tube was placed in a magnet (Magna bot 96 Promega Separation Device) for 1 minute and the supernatant was discarded. After removing the tube from the magnet, the Dynabeads were re-suspended in 100μl of buffer.

Cryopreserved samples were rapidly thawed in a 37°C water bath and re-suspended. An aliquot equivalent to $1x10^6$ cells was placed in a fresh tube marked 'unsorted' and taken aside. The remaining cells from the sample were transferred to the washed Dynabeads and

mixed briefly. Cells were then incubated at 2 to 8°C for 30 minutes on a shaker enabling gentle tilting and rotation.

Following incubation, the tube was placed on the magnet for 2 minutes and the supernatant was transferred into a fresh tube marked 'CD34 negative'. The bead-bound cells were re-suspended in 1ml of cold buffer and returned to the magnet for 2 minutes. The supernatant was discarded. This process was repeated for a second time in order to wash the bead-bound cells. Finally, the bead-bound cells were re-suspended in 400μl of buffer and placed into a fresh tube marked 'CD34 positive'.

TRAP assay, DNA extraction and STELA at the XpYp and 17p telomere were subsequently performed.

2.8 Telomerase Assay

Telomerase activity was assessed according to the Telomeric Repeat Amplification Protocol (TRAP) TRAPeze RT Telomerase Detection Kit (Chemicon International). Cells (samples and telomerase positive control, i.e. 293 cell line) stored at -80°C were washed in 1xPBS prior to the TRAP assay. 10^5 to 10^6 MNCs were lysed in 200μl of CHAPS Lysis Buffer and incubated on ice for 30 minutes. Samples were centrifuged at 12K for 20 minutes at 4°C and 160μl of the supernatant was collected. The protein concentrations were measured and samples (as well as 293 cells) were diluted to 250ng/μl in CHAPS lysis buffer. 10μl of each sample (and 293 cells) were heat treated at 85°C for 10 minutes in order to inactivate telomerase. Notably, these were used as telomerase negative controls for the TRAP assay.

To generate a standard curve, stock TSR8 control template (20amoles/μl) was serially diluted by 1:10 in CHAPS lysis buffer in order to prepare 2amoles/μl, 0.2amoles/μl and 0.02amoles/μl.

Each 25μl reaction consisted of 1x$TRAP_{EZE}$ RT Reaction Mix, 2 units of Taq Polymerase and 2μl of each sample (listed below) in nuclease free H_2O.

- TSR8 dilutions (20amoles/μl, 2amoles/μl, 0.2amoles/μl and 0.02amoles/μl)
- 293 positive control
- Minus Telomerase Control: CHAPS lysis buffer only

- No template control: Nuclease free H_2O
- Experimental samples: 2μl of 250ng/μl (i.e. 500ng/reaction)
- Heat treated samples (and 293) telomerase negative controls

Positive controls were also utilised for each sample to rule out false-negative results. The TSK control template was diluted in CHAPS lysis buffer down to 0.002amole/μl for use as a positive control. Each 25μl reaction consisted of 1x$TRAP_{EZE}$ RT Control Reaction Mix, 2 units of Taq Polymerase, 0.004 amoles of TSK control template and 2μl of each sample (listed below) in nuclease free H_2O.

- Experimental samples: 2μl of 250ng/μl (i.e. 500ng/reaction)
- 293 positive control

Each PCR reaction was carried out in triplicate in a 96 clear-well plate format using the ABI PRISM 7900. Reactions were subjected to 30°C for 30 minutes for 1 cycle, 95°C for 5 minutes for 1 cycle and (94°C 15sec, 59°C 60sec, 45°C 20sec) X 45 cycles.

2.9 Statistical Analysis

The Kaplan-Meier method was used to calculate survival curves. Telomere data was expressed as the mean ±SD. Comparisons between groups were analysed using the Unpaired *t*-test. Values of $p < 0.05$ were considered significant. Statistical tests were performed utilising GraphPad Prism.

2.10 Array-Comparative Genomic Hybridisation (Array-CGH)

Array CGH was carried out as a service by Nimblegen (Roche NimbleGen, Madison, WI). 2μg of tumour genomic DNA and reference DNA samples were independently labelled with either Cy3 or Cy5 dyes. Labelled DNA was co-hybridised to a NimbleGen comparative genomic hybridisation (CGH) array format 12 x 130000 (130000 probes). Reference samples were labelled with Cy 5 and test samples with Cy 3.

2.11 Development of 6q STELA

Sequence alignment of 5457 bases proximal to the published 6q terminus against a family of multiple chromosome ends was achieved using Clustal W. A 6q specific primer (6qspecific)

was designed to target a region 4070 bases proximal to the published end and a temperature gradient STELA reaction was subsequently performed under the following conditions: 22 cycles of 94°C for 20secs, (temperature gradient 55°C to 70°C) for 30secs, and 68°C 8 minutes. STELA at 17p was utilised as a positive control. Southern blot hybridisations were performed using only the TTAGGG repeat probe.

A sub-telomeric region sequenced to the adjacent telomeric TTAGGG tract was utilised to design primers that may be shared by the unidentified 6q terminus. STELA at 17p was used as a positive control. The favoured annealing temperature was identified by a gradient temperature STELA. Southern blot hybridisations were performed using only the TTAGGG repeat probe.

A telomere-adjacent reverse primer used in conjunction with a 6q specific primer amplified a putative 6q specific product using conditions: 33 cycles of 94°C for 20secs, (temperature gradient 55°C to 70°C) for 30secs, and 68°C 8 minutes.

The probe specific to this product was generated by re-amplification PCR (used gradient PCR since annealing temperature was unknown) between the forward primer '6qspecific' and a downstream (~1kb) reverse primer shared by the 6q family.

The product was resolved by agarose gel electrophoresis, gel-purified (QIA quick Gel Extraction Kit; Qiagen) and quantified (ng/μl) by comparing the intensity of the probe band against a range of 1kb ladder concentrations. 25ng/μl was subsequently used for hybridisation.

Direct sequencing of the region downstream to the 6q published terminus was performed with designed sub-telomeric primers present within the telomere adjacent region shared by members of the 6q family.

2.12 Oligonucleotides

STELA/ XpYp Genotyping

XpYpE2: 5' -TTGTCTCAGGGTCCTAGTG- 3'

XpYpB2: 5' -TCTGAAAGTGGACC(A/T)ATCAG- 3'

XpYp-427G/415C: 5’ -GGTTATCGACCAGGTGCTCC- 3’

XpYp-427A/415T: 5’ -GGTTATCAACCAGGTGCTCT- 3’

17pseqrev1: 5’ -GAATCCACGGATTGCTTTGTGTAC- 3’

Telorette2: 5’ -TGCTCCGTGCATCTGGCATCTAACCCT- 3’

Teltail: 5’ -TGCTCCGTGCATCTGGCATC- 3’

TVR-Mapping

TAG-TELW: 5' -TCATGCGTCCATGGTCCGGACCCTTACCCTTACCCTNACCCTA- 3’

TAG-TELX: 5' -TCATGCGTCCATGGTCCGGACCCTTACCCTTACCCTNACCCTC- 3’

TAG-TELY: 5' -TCATGCGTCCATGGTCCGGACCCTTACCCTTACCCTNACCCTG- 3’

Fusion

XpYpM: 5’ -ACCAGGTTTTCCAGTGTGTT- 3’

17p6: 5’ -GGCTGAACTATAGCCTCTGC- 3’

21q1: 5’ -CTTGGTGTCGAGAGAGGTAG- 3’

Re-amplification

XpYpMb: 5’ -AGGTTTTCCAGTGTGTTATC- 3’

17p6b: 5’ -CTGAACTATAGCCTCTGCAATG- 3’

21q1C: 5’ -GTGTCGAGAGAGGTAGCTTTTAAATG- 3’

6q

6qSpecific: 5’ -CTAGTTTCCTCTCTTATTAACATA- 3’

Chapter 3:

Telomere Dynamics in MDS and AML

3.1 Abstract

Specific telomere length profiles were analysed using STELA at XpYp and 17p to determine the nature of telomere dynamics in a cohort of MDS and AML patients. STELA is a long-range single-molecule PCR approach that amplifies the double-stranded region of telomere repeats from specific chromosome ends. Uniquely, STELA is capable of detecting short telomeres within the length ranges that have the potential to initiate telomere fusion.

Patients with AML showed significantly shorter telomeres than those with MDS ($p < 0.0001$). Telomere length distributions appeared to homogenise in AML suggestive of blast accumulation and clonal expansion.

Telomere loss is dictated by the end-replication problem but also by sporadic, atypical large-scale deletion events. A significantly higher percentage of truncated telomeres were observed at 17p in MDS which may have been associated with the abrogation of *p53* and damage response mechanisms. An RAEB2 patient presented 9.1% of short telomeres at the 17p telomere. A reduction of these events was apparent within the AML cohort; this may reflect the up-regulation of telomerase activity.

TVR-PCR revealed extensive variation in measurements of TVR regions ranging from 0 to 3kb from individuals analysed. Extreme telomere shortening was apparent following correction for the TVR region with some patients demonstrating a lower 25th percentile of only 8 TTAGGG repeats; these telomeres are within the length ranges at which telomere fusion is detected.

Bimodal distributions were detected using STELA. These could be confirmed as a consequence of biallelic telomere length variation in a subset of patients.

CD34^{+} cells failed to show increased telomere shortening, however provided the majority of these patients were low-risk, this may have been associated with the subsequent differentiation into CD34^{-} cells. High-risk patients may have demonstrated a larger degree of

telomere shortening within the $CD34^+$ fraction due to clonal expansion but associated with blocked differentiation.

3.2 Introduction

Telomere shortening has been associated with advanced MDS and its progression into AML[332-334] with the additional observation of a co-existing complex karyotype.[333,334,342] However, these studies have relied upon the global average of telomere length and have failed to show the prognostic implications that single, critically short telomeres may have with disease progression. Notably, it has been proposed that the shortest telomere within a cell has the propensity to initiate genomic instability.[209]

STELA is a high resolution technique for telomere length analysis which has provided insights into the highly dynamic nature of single telomeric molecules. It is a long-range single-molecule PCR approach that amplifies the double-stranded region of telomere repeats from specific chromosome ends (Figure 3.1).[171]

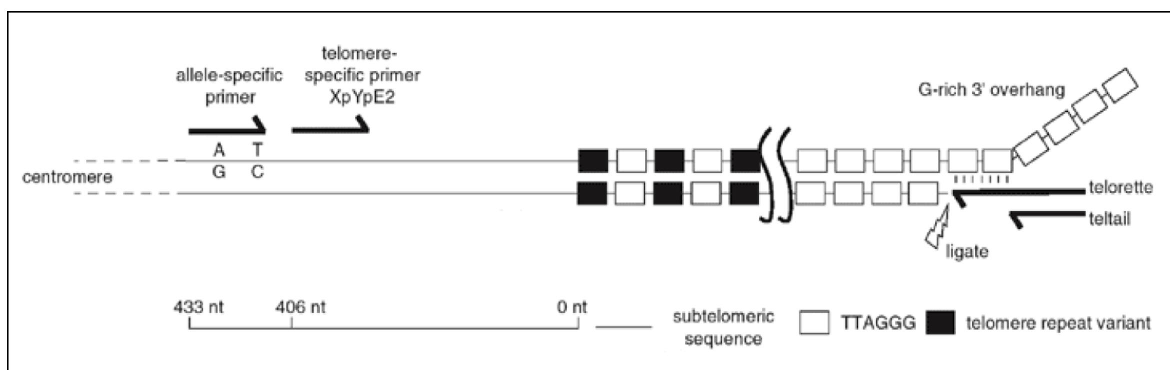


Figure 3.1: A representation of STELA extracted from: Baird DM, Rowson J, Wynford-Thomas D, Kipling D. Extensive allelic variation and ultrashort telomeres in senescent human cells. Nat Genet. 2003;33:203-207. Single telomere amplification by means of utilising an upsteam chromosome or allele specific primer in conjunction with the teltail.

The initial step in STELA consists of annealing a 'telorette' linker that is comprised of seven bases of TTAGGG repeat homology followed by a 20 nucleotide non-complementary tail to the G-rich 3'overhang of the telomere. Ligation to the 5'end of the duplex telomeric C-rich strand provides a telorette 'tag' to the end of the chromosome. PCR can then be performed on chromosome-specific telomeres utilising a 'teltail' primer that is complementary to the telorette tail and a chromosome-specific upstream primer located within the sub-telomeric region. Such sub-telomeric primers that have been used in STELA reactions are XpYp and 17p specific which are complementary to a region 433bp and 311bp from the start of the telomere, respectively. STELA products are identified as single bands on a gel following Southern hybridisation with a TTAGGG specific probe with each band representing a single

telomeric molecule (also containing a region of sub-telomeric DNA, i.e. 433bp at XpYp or 311bp at 17p). This approach has the ability to detect the full spectrum of telomere lengths at specific telomeres,[171] particularly those that are very short which have the potential to initiate telomere fusion.[253-255]

Previous demonstrations using STELA have shown that telomeres have the propensity to undergo stochastic large-scale deletion events. Abrupt telomere deletion may be sufficient to generate a dysfunctional telomere which can reduce cell viability. More importantly, severely truncated telomeres have the potential to initiate a sequence of breakage-fusion-bridge (BFB) cycles by means of telomere fusion.[254,255] It is possible that recombination or replication slippage may generate extensive telomere shortening.[352]

By employing previously characterised heterozygosities (-427G/A and -415C/T) in DNA adjacent to the XpYp telomere,[244] allele-specific STELA can determine whether bimodal distributions are composed of differing maternal and paternal contributions in the zygote. Differential telomere size distributions of up to 6.5kb have been observed at the XpYp telomere.[171] Bimodal distributions may also be characteristic of the presence of non-tumour cells, particularly in bulk populations or alternatively by means of recombination creating a new sub-telomeric distribution of longer or shorter propensity.

Blast accumulation in MDS and AML may present the tendency for reduced heterogeneity within telomere distributions. Increased homogenisation has been previously detected in CLL (Chronic Lymphocytic Leukaemia), particularly with disease progression and severity.[253] Telomeric loss of heterozygosity (LOH) can also be detected with STELA since it enables the quantification of single telomeric molecules. Loss at the 17p telomere was observed within CLL cells and further FISH analysis confirmed LOH at the *p53* locus.[253]

It has also been observed that the proximal end of the telomere contains an interspersion pattern of TTAGGG and telomere variant repeats (TVR) including TCAGGG and TGAGGG.[244] These TVR regions can vary in length from 0 to 3kb in the human population and their distribution can be determined by TVR repeat mapping by PCR (TVR-PCR).[244] It has been proposed that the TVRs are a non-functional region of the telomere as TRF1, TRF2 and POT1 show strong specificity for the TTAGGG tract and not to any repeat variants.[192] Additionally, telomere fusions have been noted to include the TVR region which indicate that these

regions may not be capable of forming the protective telomeric cap.[255] Therefore STELA profiles provide an overestimate of the telomere length distributions due to variable measurements of TVRs. Thus, by means of exploiting both STELA and TVR-PCR the functional telomeric tract composed of pure TTAGGG repeats can be analysed. Previous observations have implicated that a tract composed of only 13 TTAGGG repeats is sufficient to enter into breakage-fusion-bridge (BFB) cycles which can generate non-reciprocal translocations.[255]

In this study, STELA was employed to determine the nature of telomere dynamics in MDS and AML. Telomere length was shorter within AML cells, consistent with previous observations.[335,336] There was a tendency for telomere shortening at XpYp and telomeric LOH was confirmed at this chromosome in an AML patient. STELA profiles homogenised in AML demonstrating the clonal expansion of blast cells. Bimodal distributions had been characterised as allele-specific in a number of individuals but also events that were consistent with telomere elongation by additional mechanisms such as recombination were noted to occur. Severely truncated telomeres were present within some individuals, for example one RAEB2 patient presented with 9.1% of these events at 17p. Dysfunction at the 17p telomere has the potential to initiate loss of genetic material on the 17p chromosome arm including the *p53* gene. Telomerase activity was up-regulated in AML cells, consistent with cancer propagation.[180] $CD34^+$ cells failed to show elevated telomere shortening to its $CD34^-$ counterpart. TVR-PCR confirmed the presence of extensive TVRs (up to 3kb) such that the 25th percentile of specific telomere profiles displayed telomeres that were only 48bp or only 8 TTAGGG repeats.

3.3 Telomere Length in MDS and AML

STELA analysis was undertaken on a panel of 80 MDS and 144 AML patients and the mean of each of the telomere length distributions was calculated. The mean of the whole MDS cohort was determined as 4.45kb (±1.49kb) [Mean (±SD)] and 4.47kb (±1.23kb) for the XpYp and 17p telomeres, respectively. Conversely, the mean telomere length distributions recorded within the AML cohort were 3.54kb (±1.35kb) at XpYp and 3.67kb (±0.89kb) at 17p (Figure 3.2). The telomere length within the AML cohort was significantly shorter than that recorded within the MDS cohort ($p < 0.0001$) illustrating that telomere shortening is a common feature in AML cells; this was apparent at both the XpYp and 17p telomeres.

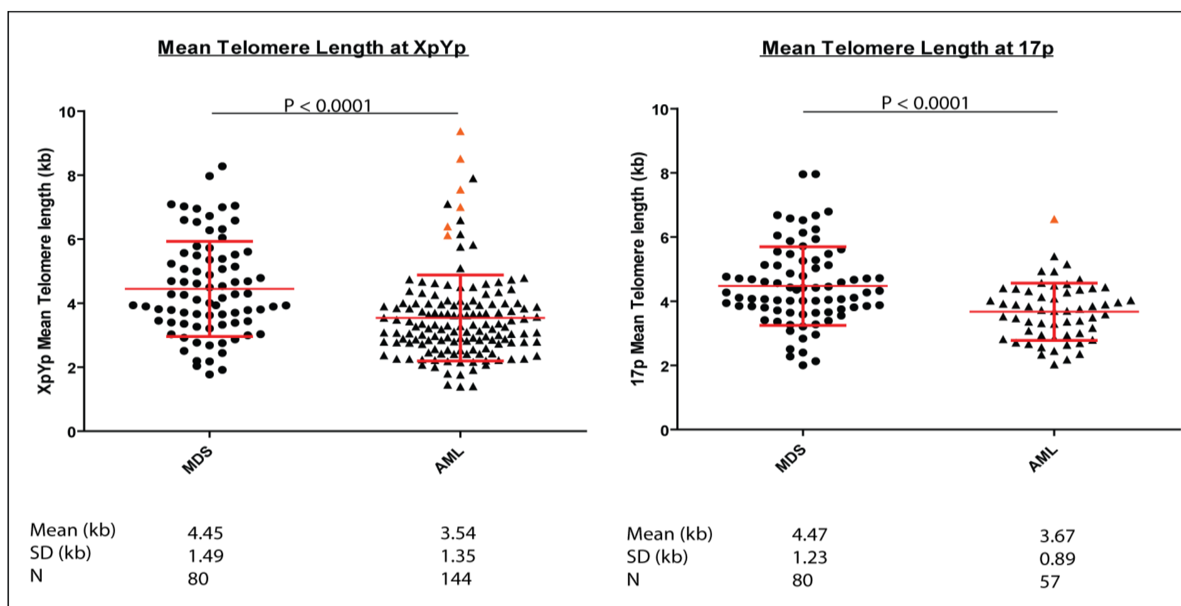
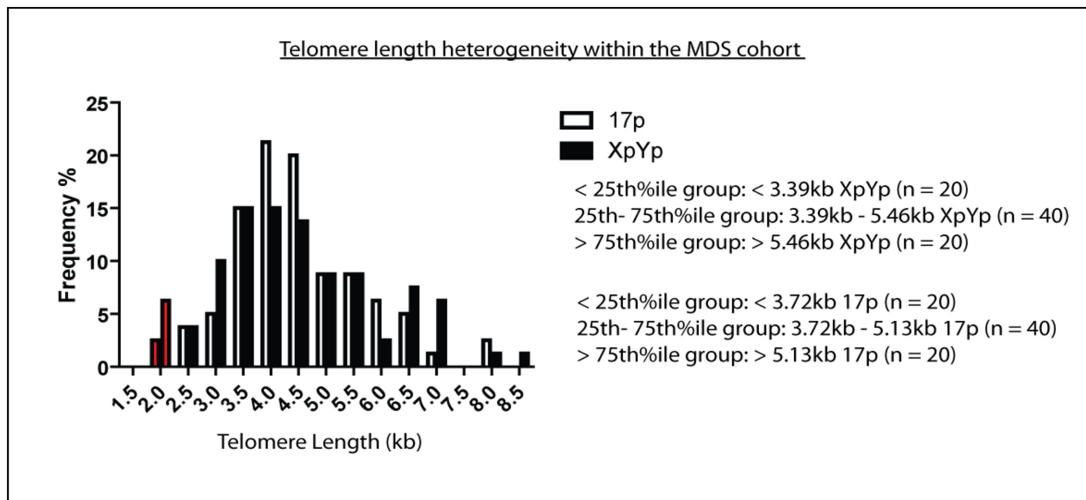


Figure 3.2: Telomere length distibutions within MDS and AML derived cells. Telomere length was statistically shorter in AML ($p < 0.0001$). There appeared to be a tendency for telomere shortening at XpYp within the AML cohort. Orange points signify patients that may have induced telomere skewing within the distribution. It was noted that these patients had large standard deviation values (>2.50kb) implicating the presence of large intra-clonal heterogeneity.

There was no significant difference between the XpYp and 17p telomere length within the MDS cohort ($p = 0.8960$), however, the telomere length at XpYp appeared shorter than the 17p telomere in the AML cohort. Whilst this was not significant ($p = 0.4848$) it may have been skewed by the presence of a subset of samples ($n = 6$) with longer XpYp telomere lengths that have been specified as orange points in Figure 3.2. It was noted that these individual STELA profiles were more heterogeneous with a standard deviation of over 2.50kb.

3.4 Inter-individual Telomere Length Variation

Within the patient cohorts, the XpYp and 17p telomeres displayed heterogeneous length distributions amongst the individuals analysed, particularly within the MDS cohort. The 25th and 75th percentiles enabled cohort subdivision into three groups based on telomere length distributions at XpYp and 17p (Figure 3.3). 20 MDS patients had presented telomere length profiles below the 25th percentile of the cohort.



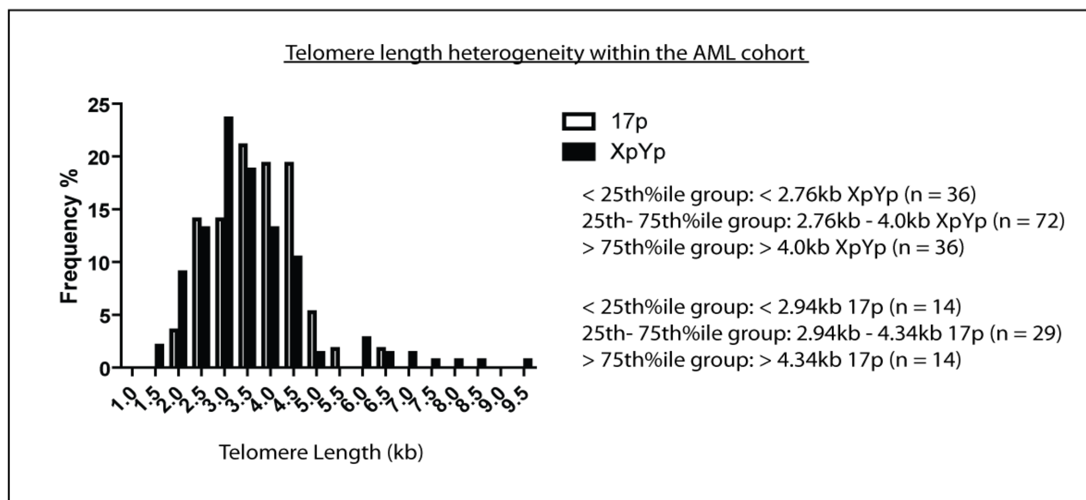


Figure 3.3: Inter-individual telomere length heterogeneity at XpYp and 17p within MDS and AML. Telomere length appears to be more variable within the MDS cohort-this may have prognostic implications. Of particular interest, an MDS patient showed telomere distributions at XpYp and 17p that were shorter than the lower 25th percentile of the AML cohort (patient within highlighted red bars). Consistently, this patient progressed to AML.

A subset of MDS patients displayed telomere length distributions below the lower 25th percentile of the AML cohort, in particular one individual that had presented a telomere length distribution of 1.77kb at the XpYp telomere and 2.28kb at 17p (Figure 3.3; red bars). This had represented the shortest telomere distributions recorded within the MDS cohort (Figure 3.4; Patient#2); this patient progressed to AML and survived for under a year post diagnosis.

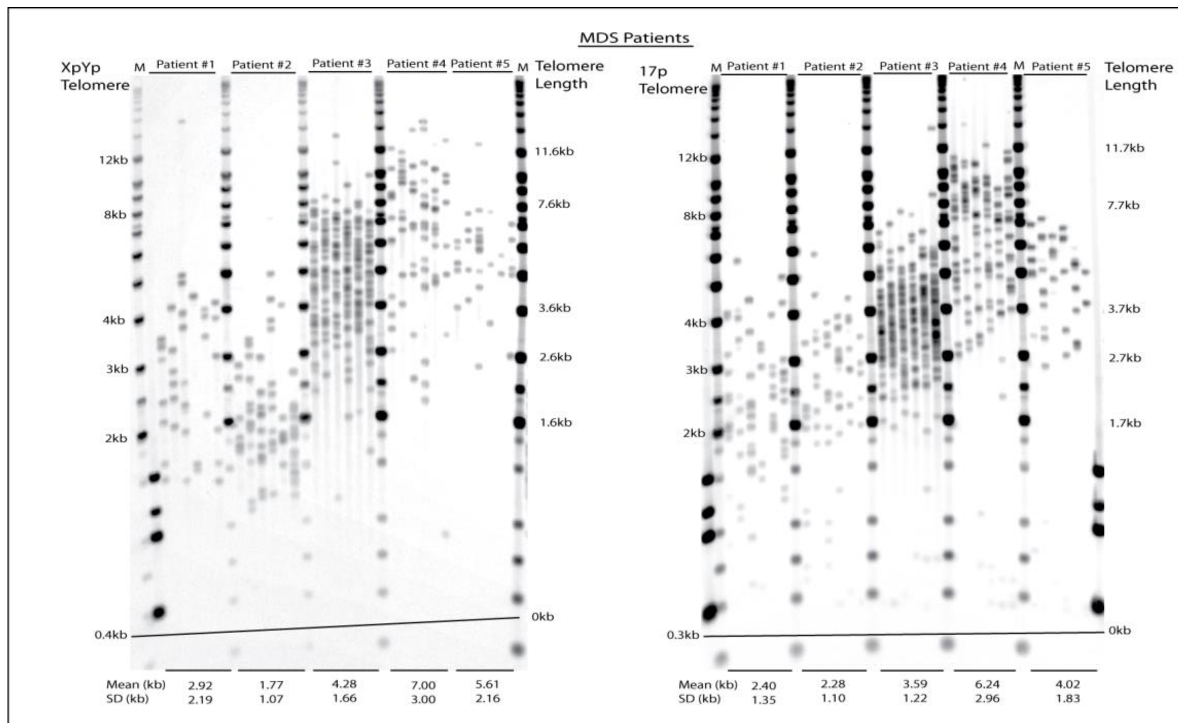


Figure 3.4: STELA profiles of 5 MDS patients. Patient #2 demonstrated the shortest telomere length distributions within the MDS cohort which was shorter than the 25th percentile of the AML cohort.

With this information at hand, it may be fair to speculate that telomere length has the potential to provide prognostic information, including the propensity of AML transformation.

It appeared that telomere length homogenised between individuals within the AML cohort (Figures 3.2 and 3.3). This indicates the possibility that telomere length may display less prognostic potential in AML compared to MDS. Conversely, the heterogeneity of telomere length in MDS may have prognostic implications in itself. However, further insight into the clinical data associated with these samples may provide a more detailed analysis (Chapter 4 and 5).

3.5 Telomere Length Correlation

The mean XpYp and 17p telomere lengths from each sample were strongly correlated within the MDS ($r^2 = 0.5992$; $p < 0.0001$) and AML ($r^2 = 0.6036$; $p < 0.0001$) cohorts (Figure 3.5).

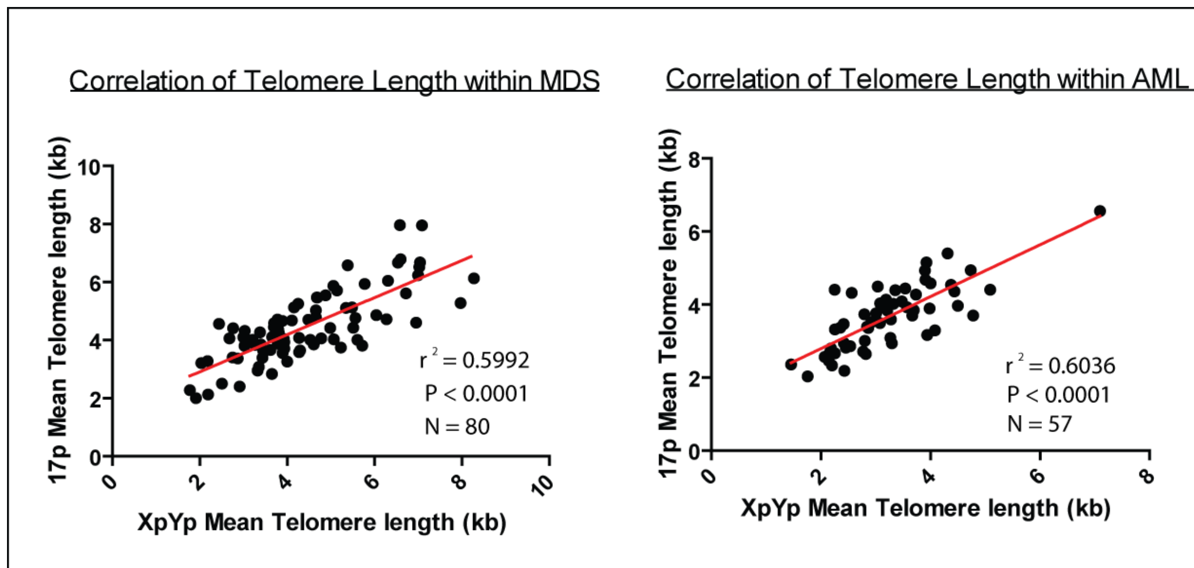


Figure 3.5: Regression analysis between the XpYp and 17p STELA distributions. There was a strong correlation between telomere length in both the MDS and AML cohort ($r^2 = 0.60$, $p < 0.0001$).

The intersection of the axis revealed the 17p telomere to be 1.63kb and 1.37kb longer than that at XpYp in the MDS and AML cohorts, respectively. Whilst is it possible that the XpYp telomere may have a propensity for elevated rate of telomere shortening, this phenomenon hasn't been observed in previous studies that indicated that telomeres at different ends erode at a constant rate.[353] Instead, contrary to previous reports,[354] it is possible that the 17p telomere may have been set longer in the zygote. Alternatively, if the non-functional telomere variant repeat (TVRs) region is longer at the 17p telomere compared to XpYp then this will have the effect of making the length of the STELA distributions longer at 17p compared to XpYp.

3.6 Intra-clonal variation

The heterogeneity of a telomere distribution can reflect the replicative history of the cell sample analysed, with cell populations derived from single cells exhibiting homogeneous distributions. In addition, the presence of non-tumour material with differing replicative histories, telomerase activity and telomere dynamics will render the telomere-length profiles more heterogeneous.

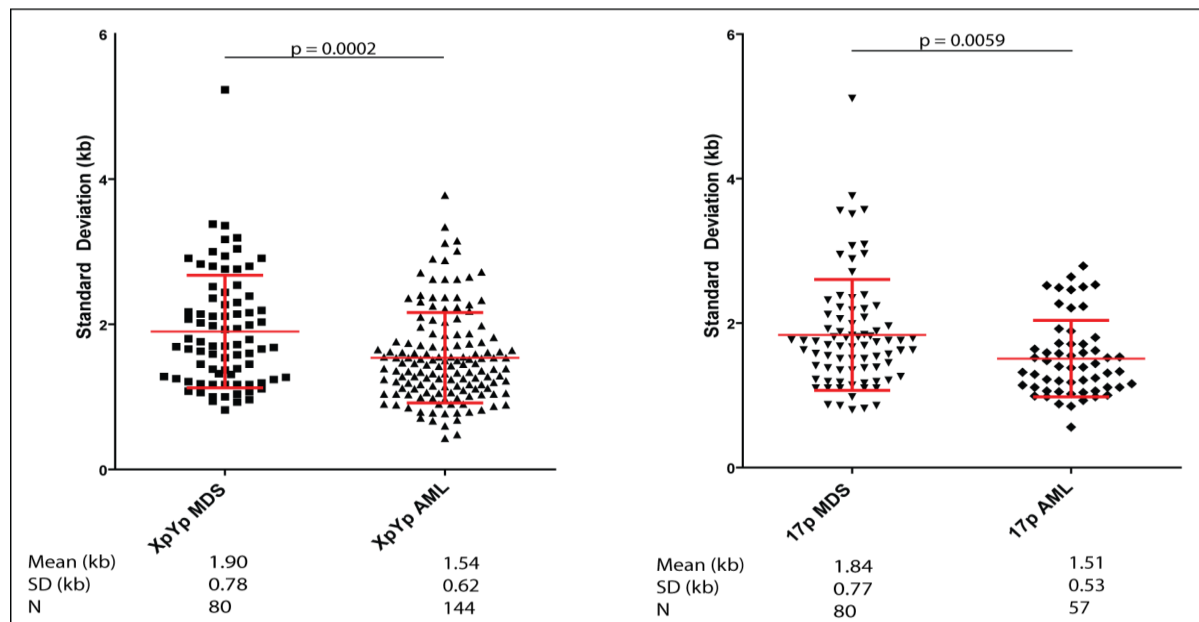


Figure 3.6: Telomere length heterogeneity within each STELA profile. Telomere distributions homogenised within AML profiles, consistent with blast accumulation. Large SD values may be explained by allelic specific differences or recombination events.

The standard deviations (SD) of the XpYp telomere were significantly different between the MDS and AML cohorts (p = 0.0002) measuring 1.90kb (±0.78kb) and 1.54kb (±0.62kb), respectively. The difference in the standard deviation at the 17p telomere also reached significance with 1.84kb (±0.77kb) and 1.51kb (±0.53kb) (p = 0.0059) within the MDS and AML cohorts, respectively (Figure 3.6). Thus, it appears that telomere length distributions homogenise with progression to AML, consistent with malignant clonal evolution.

Some individuals displayed extremely heterogeneous telomere length profiles with some displaying clear bimodal distributions; this may be explained by differential clonal distributions or by the presence of healthy and malignant cells. Alternatively, bimodal populations may also be the result of telomere length differences between maternal and paternal alleles. One further possibility is that a mutational event resulting in a large-scale change in telomere length could create a bimodal distribution. The mechanisms underlying

this are not clear but they could include a recombination event that either elongates or shortens a telomeric molecule. Recombination can abruptly generate a sub-clone of cells with an altered proliferation potential.

An example of a bimodal telomere-length distribution is shown in patient #1 in Figure 3.7.

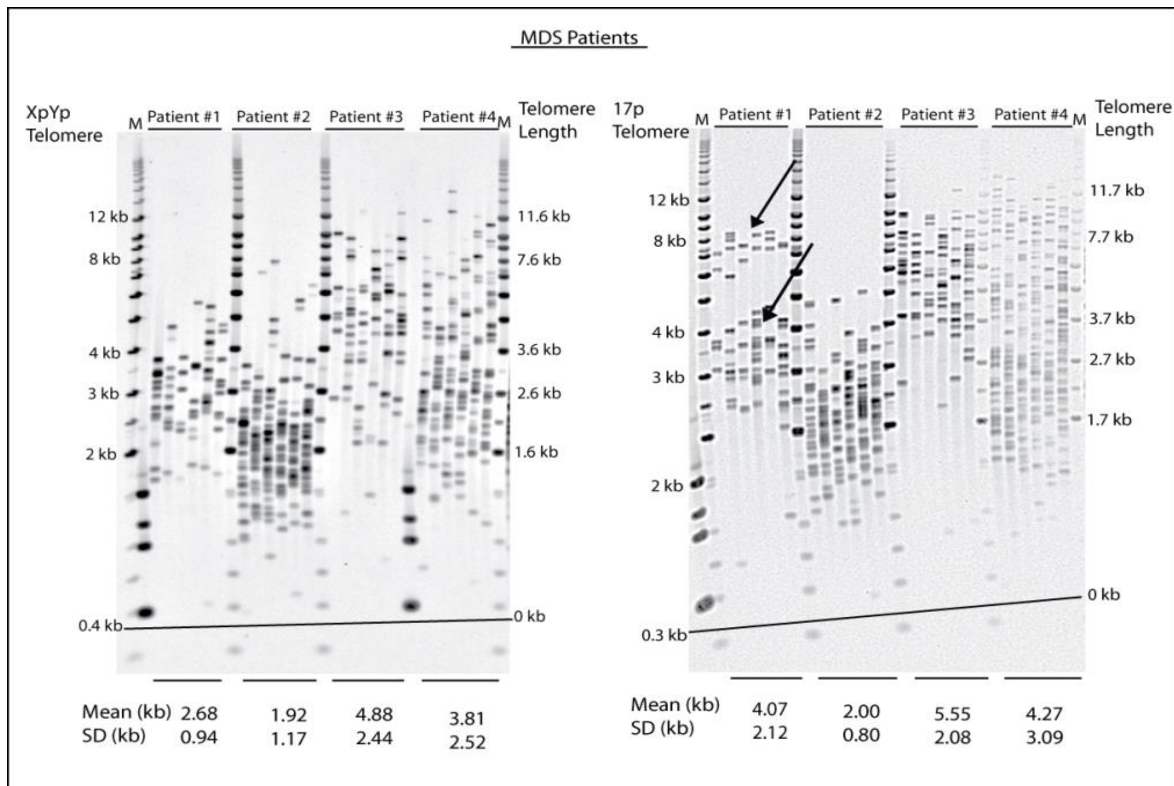


Figure 3.7: STELA profiles derived from 4 MDS patients. The bimodal distribution at 17p (represented with arrows) within the telomere profile of patient #1 illustrates a potential recombination event that has extended the average length of the telomere profile. In contrast, the telomere distribution at XpYp appears to be fairly homogeneous which is described by its much lower SD value.

Here bimodality is observed at 17p but not at the XpYp telomere. The lack of an allele specific 17p polymorphism prevented the ability to rule out allele-specific length differences at 17p however, the unequal number of telomere molecules within each population (17 within the longer and 47 molecules in the shorter) is inconsistent with allele-specific distributions, in which an equal number of molecules per distribution would be expected.

This is also unlikely to represent a population of healthy cells exhibiting longer telomeres at 17p and a shorter telomere at XpYp. Thus, the longer distribution may have arisen as a consequence of a mutation event that resulted in telomeric elongation specifically at 17p. This mechanism is not clear from this analysis; however one could speculate that this arose following a recombination event at the 17p telomere in a single cell of either normal or

malignant propensity, thus creating a sub-clone of cells that carry a 17p telomere longer than the overall distribution.

A subset of bimodal distributions (n = 28) were a consequence of biallelic telomere length variation. Allele-specific STELA was performed on patients that were heterozygous for the AT and GC haplotypes adjacent to the XpYp telomere;[244] an example of which is presented in Figure 3.8.

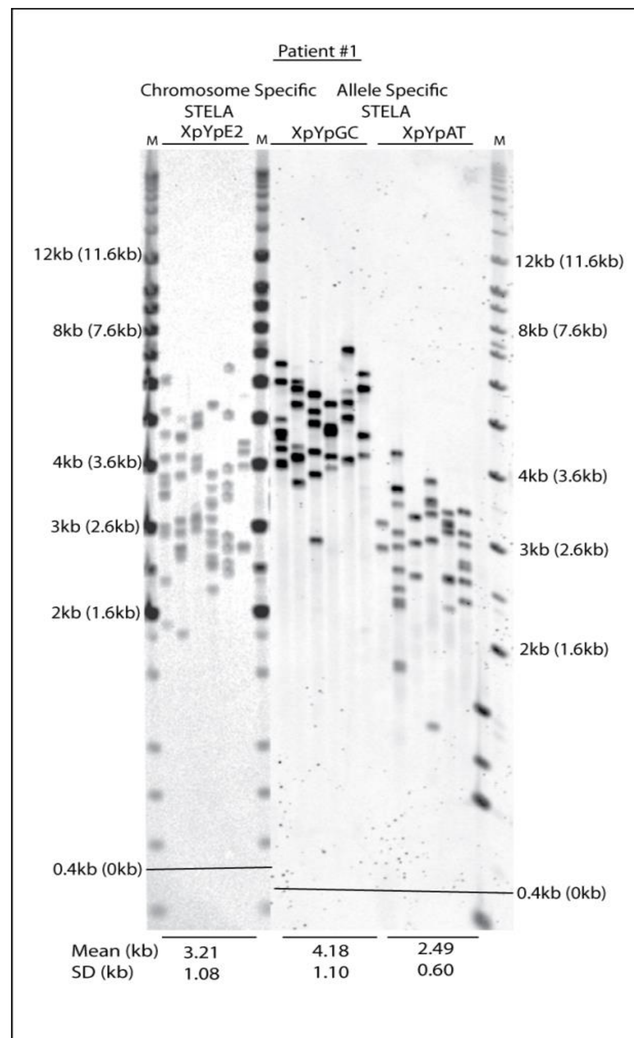


Figure 3.8: A STELA profile illustrating a clear bimodal separation of allele specific telomeres. Analysing the telomere length of a specific allele revealed a short distribution of only 2.49kb at XpYp AT.

The difference between the two distributions reached statistical significance ($p < 0.0001$). This patient was exhibiting a mean telomeric distribution of 3.21kb at the XpYp telomere; however, the shorter AT allele was only 2.49kb. Differentials between the telomere length

distributions are illustrated in Figure 3.9; those that are highlighted exhibited significant differences ($p < 0.05$) between the sizes of the telomere distributions. Patient #18 in Figure 3.9 presented with the clear bimodal distribution which is shown in Figure 3.8.

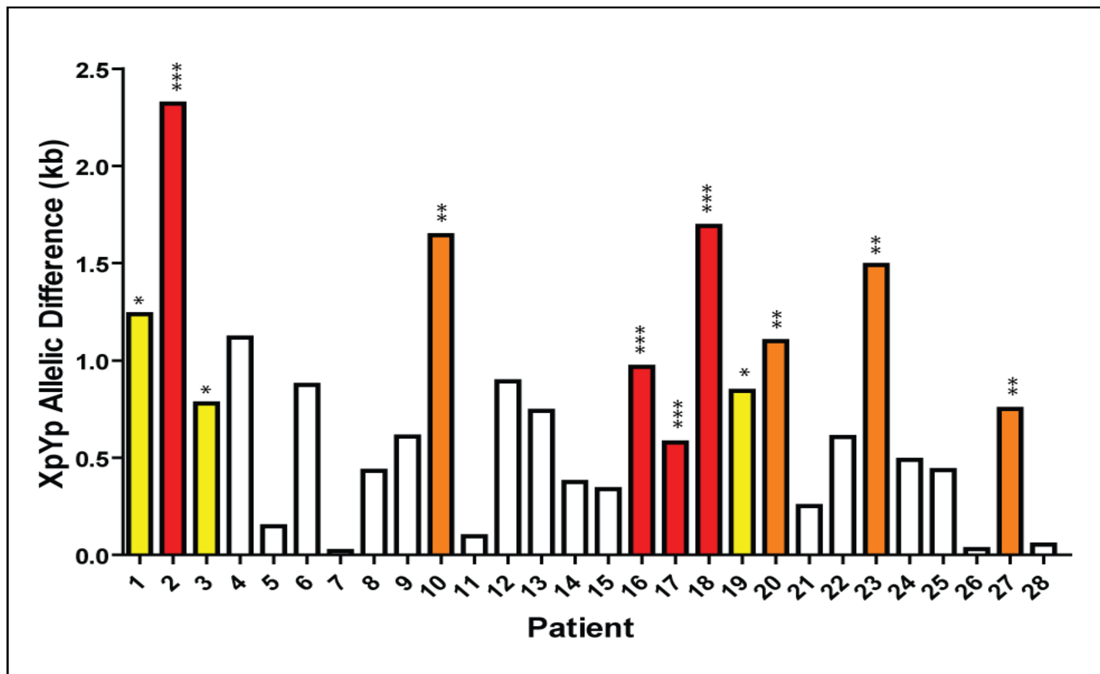


Figure 3.9: Differentials in length of allele specific telomere distributions. Significant differences ($p < 0.05$) have been highlighted. More specifically yellow, orange and red bars signify those differences that are either mildly, intermediately or highly significant ($p < 0.0001$), respectively.

The genotyping (Figure 3.10) of the XpYp telomere-adjacent DNA in the population was undertaken to identify individuals that were heterozygous and could be used for allele-specific STELA. This revealed that 36% of patients were AT/GC heterozygous. However, data on these haplotypes within the Caucasian population[244] indicated that 45% (based on the Hardy-Weinberg principle) of individuals are heterozygous at this telomere. Using a Chi-squared test it was apparent that the difference between that observed and what was expected reached statistical significance ($p = 0.0363$).

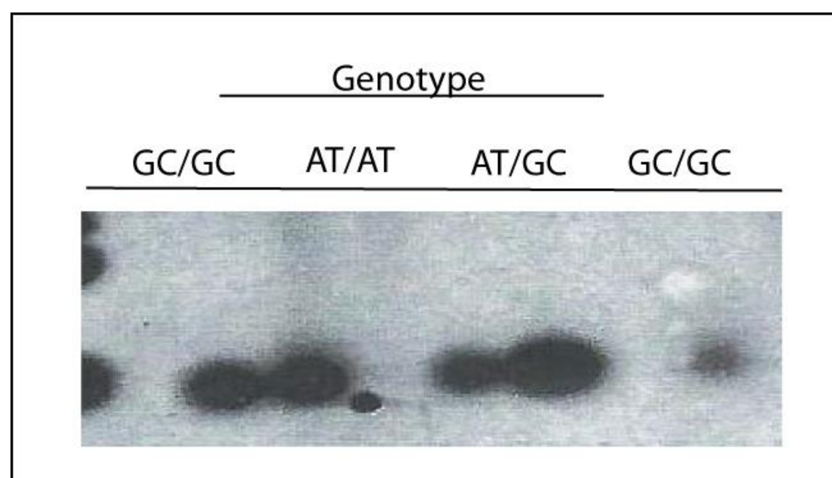


Figure 3.10: Genotyping of the XpYp telomere-adjacent DNA. It was identified that only 36% of individuals were heterozygous at this telomere. ($p = 0.0363$).

LOH at a specific telomere can also be demonstrated by a reduction in the number of amplifiable molecules detected by STELA. Given an equal input amount of DNA, STELA will typically yield a similar number of amplifiable molecules. However, some individuals such as patient #3 illustrated in Figure 3.11 displayed fewer molecules at one telomere compared to the other. This patient's telomere profile presented a population of 29 telomere molecules at 17p but only 2 at XpYp.

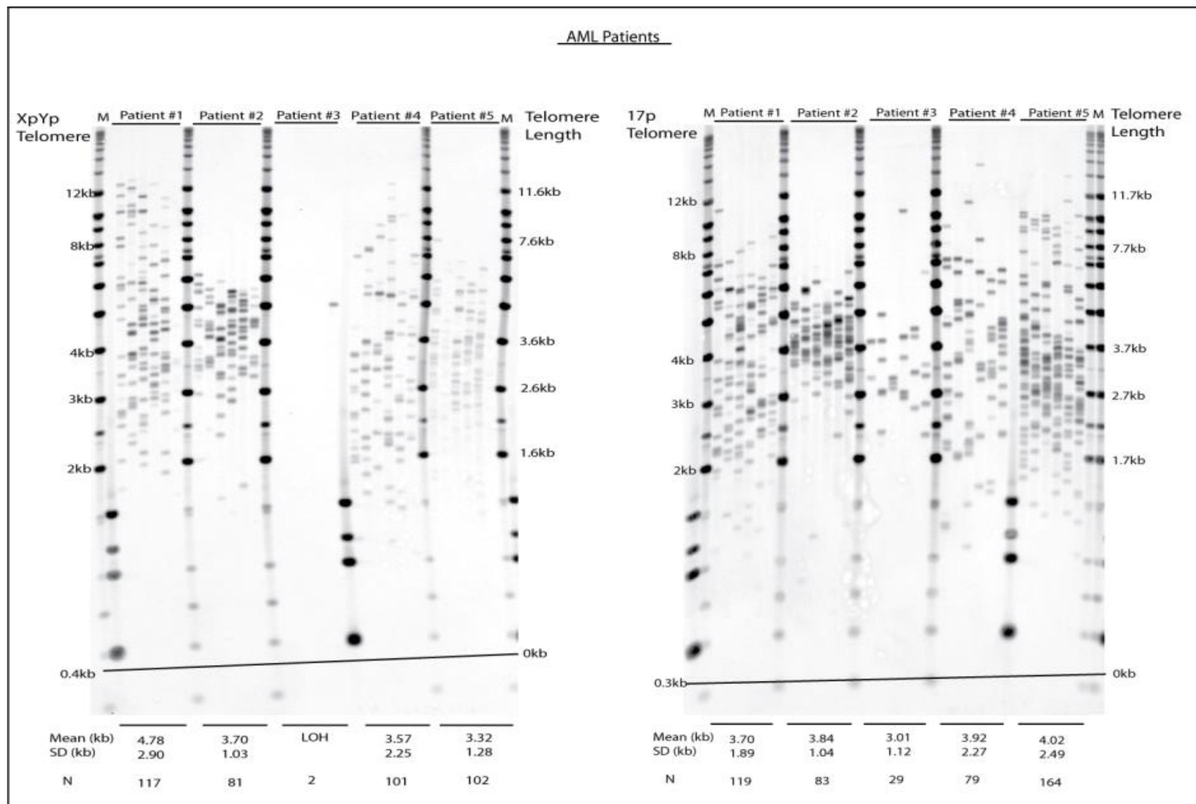


Figure 3.11: STELA telomere profiles derived from 5 AML patients. Patient #3 illustrates that telomeric LOH has occurred at the XpYp telomere.

To confirm this represented an LOH event and was not the result of a natural polymorphism at the XpYp telomere, an alternative XpYp primer (XpYpM) that binds a further 1.2kb (1666bp from the telomere) into the sub-telomeric region was utilised for the STELA reaction. Again loss of the XpYp telomere was also detected using this primer (Figure 3.12) providing further evidence in support of a telomeric LOH event.

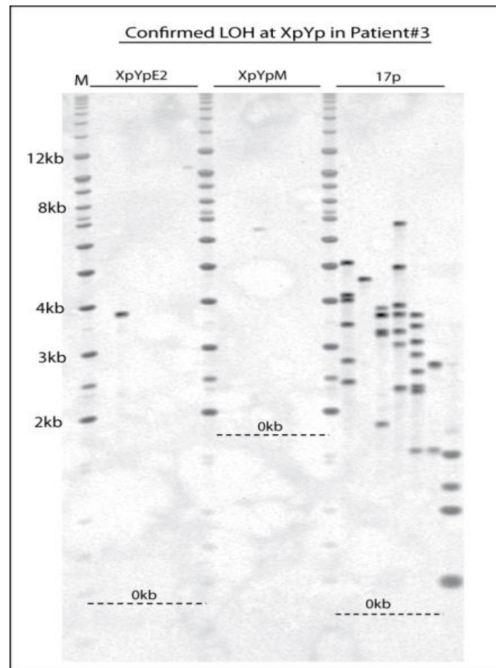


Figure 3.12: The presence of LOH at the XpYp telomere was confirmed using an alternative XpYp primer (XpYpM).

Several individuals presented bimodal allelic telomere distributions (Figures 3.8 and Figure 3.9); however it was revealed that the telomere length of the separated XpYp alleles was strongly correlated within the population (r^2 = 0.5386; p < 0.0001; Figure 3.13). Accordingly, no apparent difference existed between the length of the two alleles (p = 0.1050; Figure 3.13); this was in contrast to the differences observed between XpYp and the 17p telomeres (Figure 3.5).

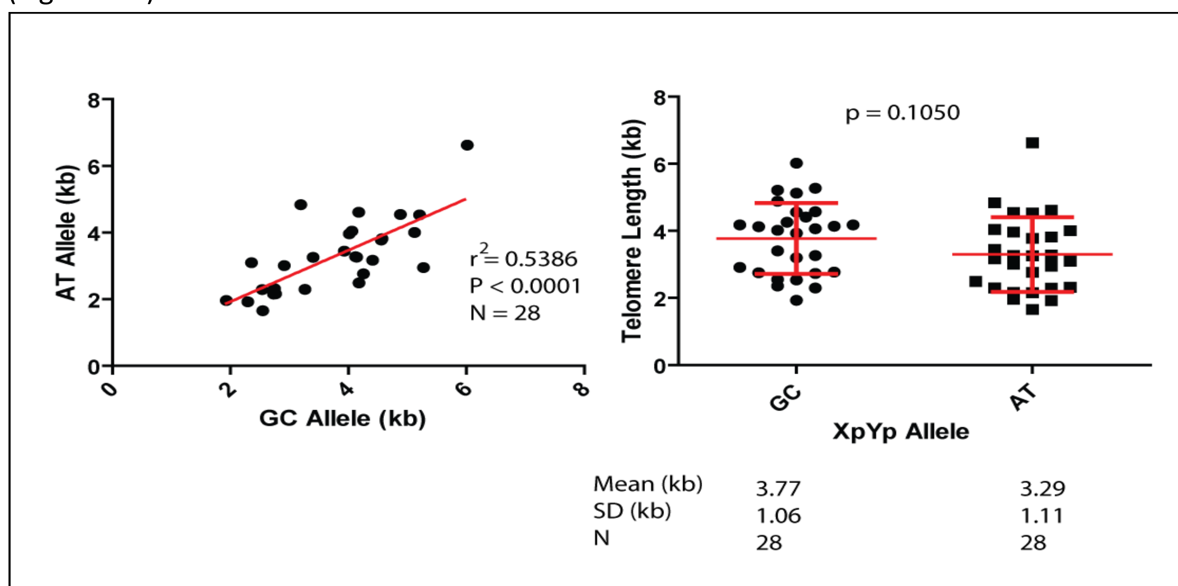


Figure 3.13: Similar telomere length was recorded at the different XpYp alleles with the AT allele showing a strong correlation with the GC allele (r^2 = 0.5386; p < 0.0001). Telomere length was not significantly different between the alleles (p = 0.1050).

3.7 Telomere Rapid Deletion (TRD) Events

The telomere length distributions were superimposed by telomere rapid deletion (TRD) events (Figure 3.14; indicated by arrows). These telomeres had been identified as those that fall outside the normal range of telomere length by means of abrupt telomere deletion. The underlying mechanism that generates these rapid changes in telomere length have not been established but may include unequal-recombination/sister chromatid exchange, replication fork stalling or replication slippage.[352] These events were quantified by quantifying the number of telomeric molecules 2.33SDs below the mean length of the distribution and recording them as a percentage of the total number of telomere molecules within each profile.

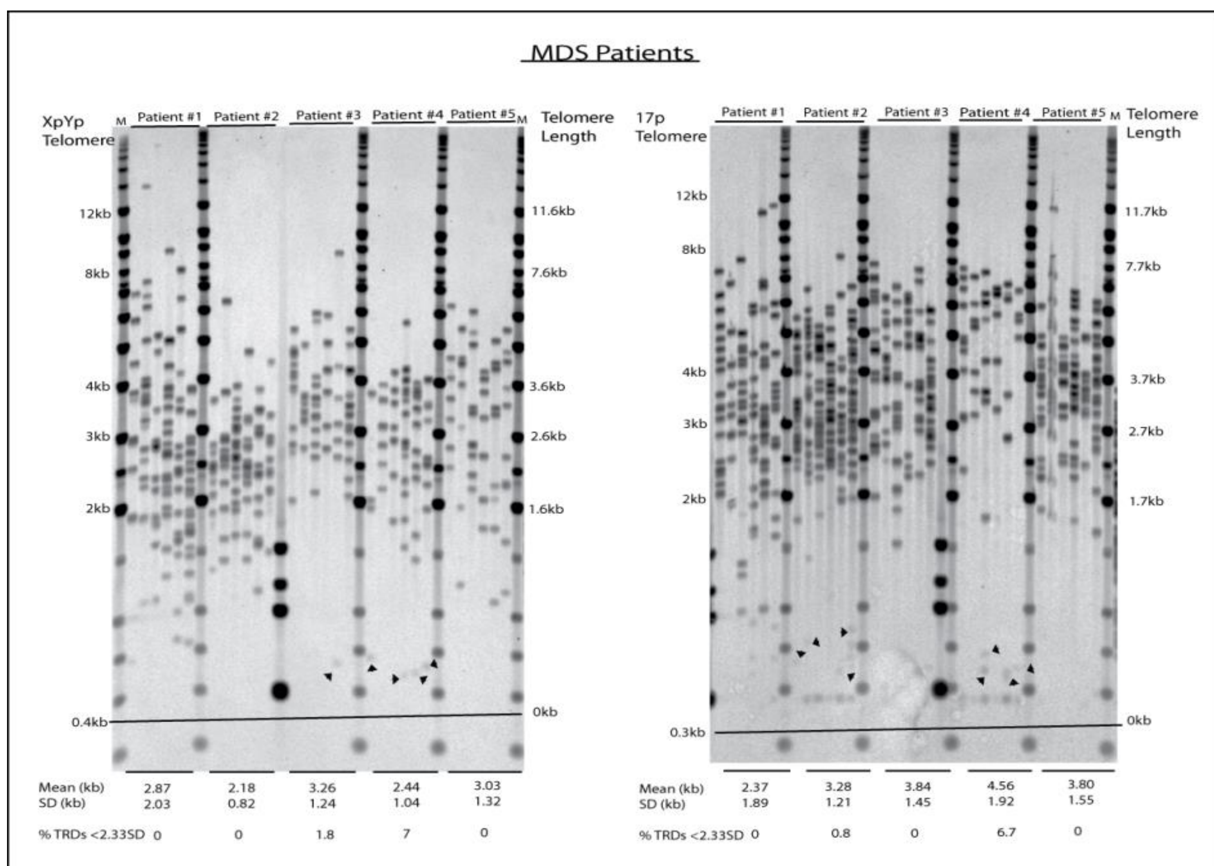


Figure 3.14: Telomere rapid deletions were detected within STELA profiles. These occurred 2.33SDs below the mean of the telomere distributions (represented by small arrows).

The hypercellular nature of MDS may provoke the generation of TRDs by means of faulty DNA replication. Furthermore, the presence of telomerase in haematopoietic cells may not be sufficient to prevent the appearance of severely truncated telomeres which may be exacerbated by the hypercellularity in MDS. These telomeres can play a crucial role in

determining cell viability and proliferative potential. It is the shortest telomere within a cell that has the potential to initiate senescence or an apoptotic response reducing the proliferative capacity of the cell.[209] Furthermore, stochastically deleted telomeres have been shown to be directly involved in telomere fusion events.[255] The mean frequency of TRD events observed at the XpYp telomere was recorded as 1.11% and 0.83% within the MDS and AML cohort, respectively, this difference was not significant (p = 0.2748). However at the 17p telomere a significant difference was observed, with 1.74% of TRD events within the MDS cohort and 0.94% within the AML cohort (p = 0.0404).

3.8 Telomerase Up-regulation

Telomerase activity was determined using the Telomeric Repeat Amplification Protocol (TRAP) assay. Despite an up-regulation of telomerase in over 90% of human malignancies,[180] cancer cells are commonly reported to possess shorter telomere length distributions when compared to adjacent normal tissue.[323]

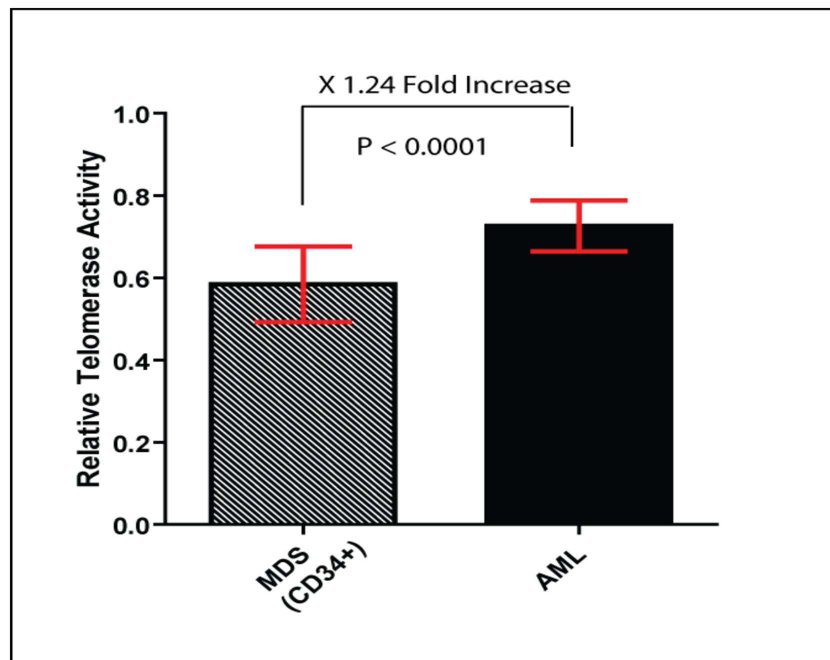


Figure 3.15: Telomerase activity was analysed using TRAP and the relative activity measured within MDS derived $CD34^+$ cells and AML derived samples. It is expressed as relative activity to the positive control (293 cell line). A significant increase by 1.24 was detected in AML ($p < 0.0001$).

Telomerase activity was analysed in a subset ($n = 12$) of AML patients and compared to the telomerase activity detected in cell sorted $CD34^+$ cells from a group of MDS patients ($n = 20$). Consistent with previous observations,[328,332,342,347] telomerase was significantly up-regulated in AML ($p < 0.0001$; Figure 3.15).

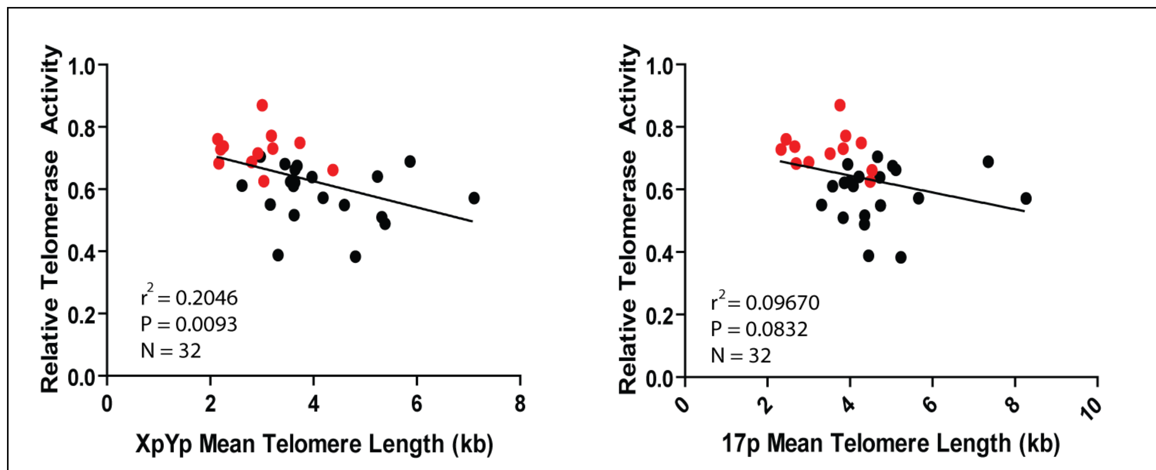


Figure 3.16: There was a general propensity of telomere shortening with elevated telomerase activity. Red markers signify AML patients that show reduced telomere length and up-regulated telomerase.

Telomerase activity was inversely correlated with telomere length at both chromosome ends (Figure 3.16). Furthermore, it was apparent that the AML patients exhibited both short telomeres and telomerase up-regulation; these patients are indicated as red points in Figure 3.16. This is consistent with the current notion that although cancer cells are commonly associated with shorter telomeres than their pre-cancerous counterparts, telomerase activity is elevated to maintain telomere stability[212,323] and prevent replicative senescence or apoptosis to enable unlimited proliferation during carcinogenesis.

3.9 Telomere Erosion

Serial samples were provided from a small subset of MDS patients (n = 4). Although the sample size was small and the time between sampling was relatively short, it was of interest to see whether accelerated telomere erosion could be detected. Telomere length at XpYp and 17p were analysed at the time of diagnosis and also at a post diagnosis time period. Overall, no significant difference was detected between the telomere length recorded at XpYp (p = 0.3092) or 17p (p = 0.6373) at diagnosis and after a subsequent follow-up (Figure 3.17). However, one patient exhibited significant telomere loss ($p < 0.05$) 3 years following diagnosis.

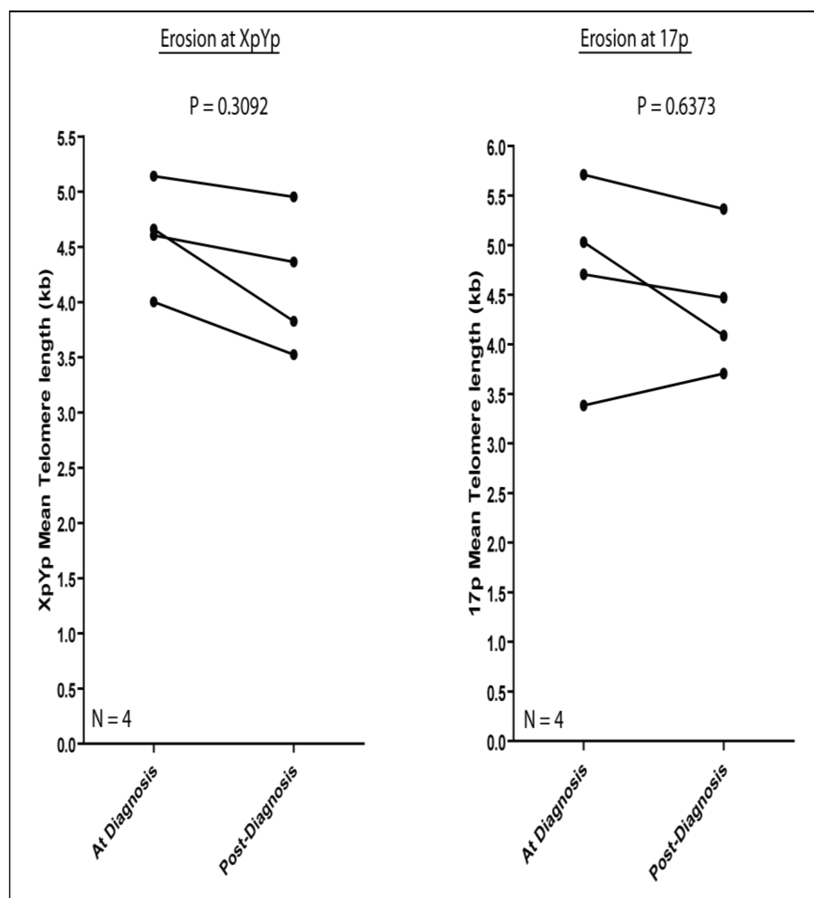


Figure 3.17: No significant difference was detected between the telomere length of XpYp (p = 0.3092) and 17p (p = 0.6373) measured at diagnosis and after subsequent sampling.

Telomeres erode by 33bp/year in proliferating haematopoietic cells.[326] The STELA presenting telomere distributions in patient #1 (Figure 3.18) illustrates a clear case of elevated telomere shortening. XpYp and 17p shortened at a roughly similar rate at

300bp/year. The bone marrow within this patient was losing telomere length 10 times faster than compared to healthy marrow.

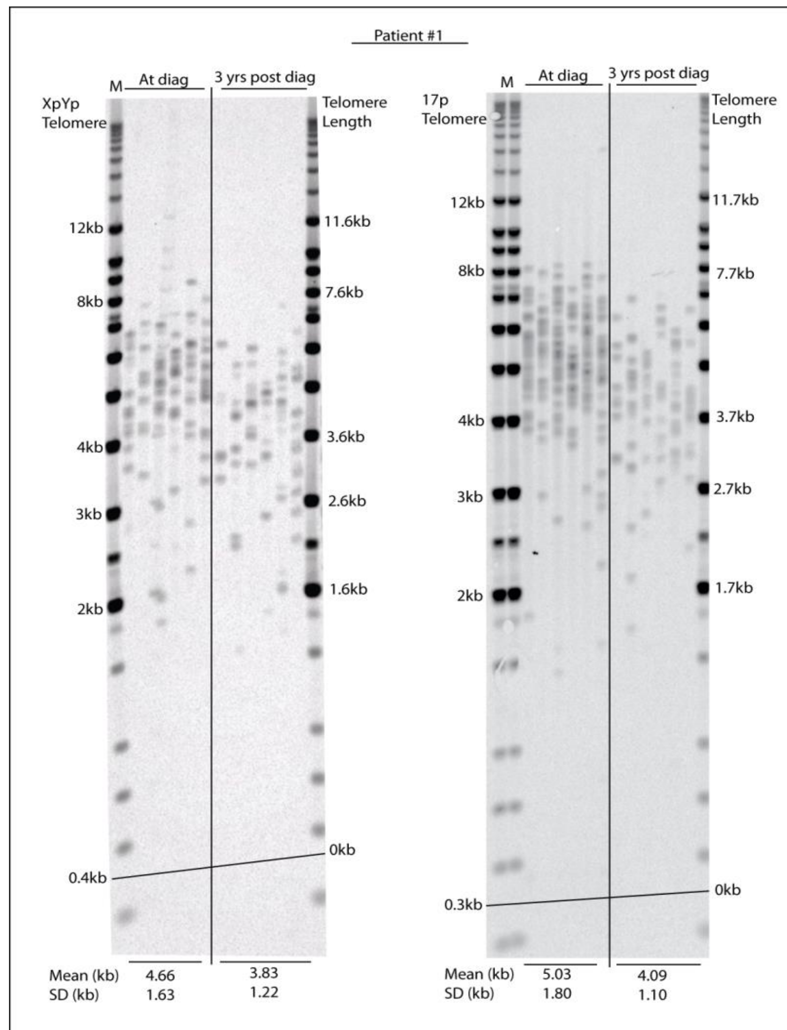


Figure 3.18: STELA profile illustrating extensive telomere erosion in an MDS patient. The rate of attrition within this patient was 10 fold higher than within normal marrow.

3.10 TVR: Telomere Variant Repeat

STELA provides an overestimate of the telomere length since it also measures a total of a sub-telomere region (i.e. 433bp into XpYp) and a Telomere Variant Repeat (TVR) region located within the first 1 to 3kb of the telomere repeat array. The length of this region varies considerably between individuals.[244] Thus, the start of the pure TTAGGG repeat region of the telomere repeat array is not identified using STELA analysis alone.

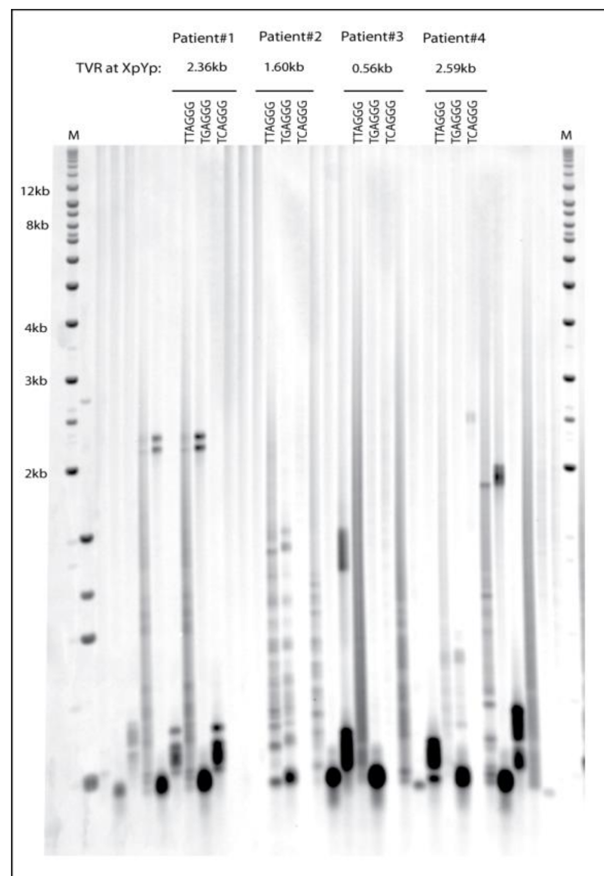


Figure 3.19: **The representation of TVR-PCR. Every three lanes represent the composition of the TVR region within each individual. The first of the three show the length of the TTAGGG repeat tract. The second and third measure the extent of TGAGGG and TCAGGG repeats, respectively and provide a measurement of their internal composition within the telomere.**

The TVR regions within each patient were analysed by using TVR-PCR (Figure 3.19) which characterises the full extent of the TVR distribution within the XpYp telomere. This

information can be used to calibrate the STELA data so that the length of uninterrupted tandem repeats of TTAGGG can be established.

The TVR repeats recorded ranged from 0 to 3kb into the telomere repeat array within the total population of individuals analysed. The mean size of the TVR region was 1.09kb (±0.72kb).

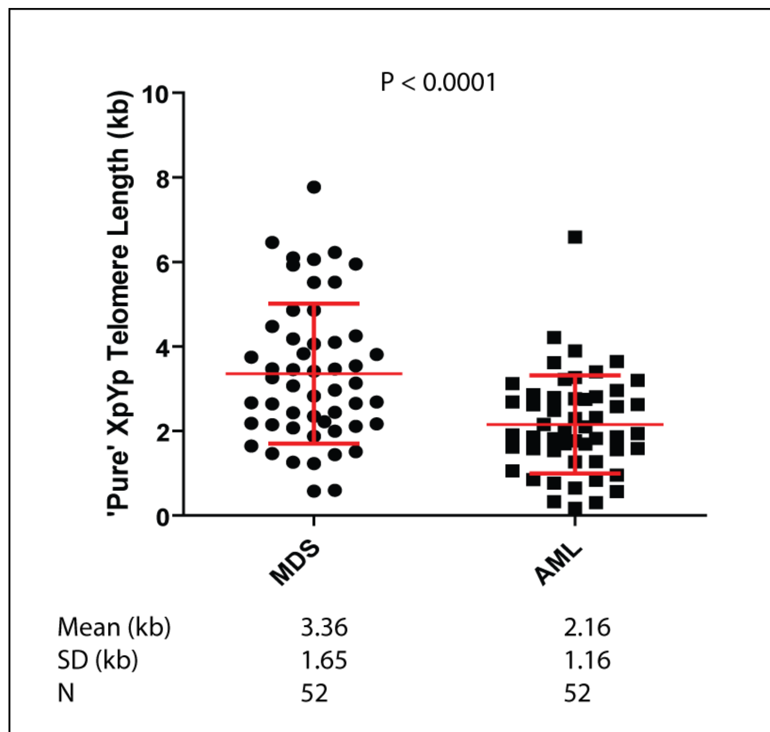


Figure 3.20: The length of the pure TTAGGG tract is significantly shorter within the AML cohort (p < 0.0001).

The telomere length distribution at XpYp was shorter within the AML cohort following the exclusion of the TVR region (Figure 3.20). The mean telomere lengths were 3.36kb (±1.65kb) and 2.16kb (±1.16kb) within the MDS and AML cohort, respectively, this difference was significant ($p < 0.0001$; Figure 3.20).

The shortest STELA profile that had been recorded within the MDS cohort was 1.77kb at the XpYp telomere (Figure 3.21; Patient #2); however, exclusion of the TVR region demonstrated that the pure TTAGGG tract was 1.25kb. This measurement exceeded that from an alternative MDS patient who showed the shortest pure XpYp TTAGGG tract at 0.58kb. Prior to TVR mapping, STELA recorded this telomere profile as 2.92kb (Figure 3.21; Patient #1) and thus was excluded as one of the shortest. However, it was revealed that this

patient only had 97 pure TTAGGG tracts on average at the XpYp telomere; this patient progressed to AML and was diagnosed with a poor cytogenetic risk score.

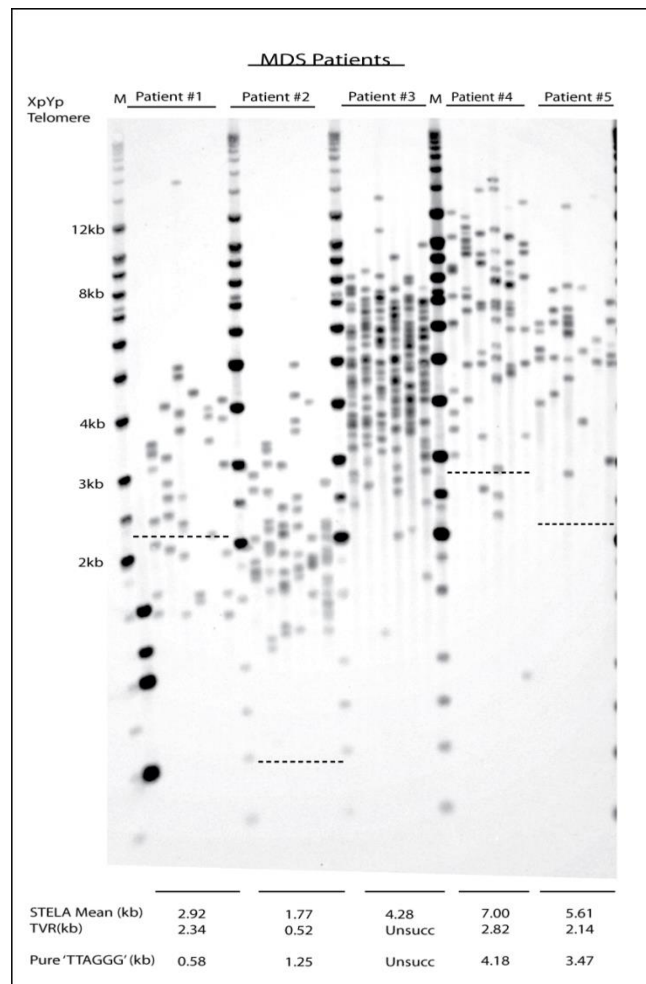


Figure 3.21: Patient #2 was originally described as having the shortest telomere distribution within the MDS cohort; however, exclusion of the TVR region demonstrated that patient #1 had a shorter pure TTAGGG tract. Dotted lines illustrate the length of the TVR.

The shortest telomere length distribution that had been recorded following adjustment for TVR content was only 0.16kb (Figure 3.22; Patient #1). This was derived from an AML patient who had presented a STELA profile of 2.20kb prior to TVR exclusion. Thus, removal of the TVR region has the potential to reveal patients with critically short telomeres despite STELA profiles presenting a telomere length of a functional range. This patient only had 27 TTAGGG repeats on average at the XpYp telomere.

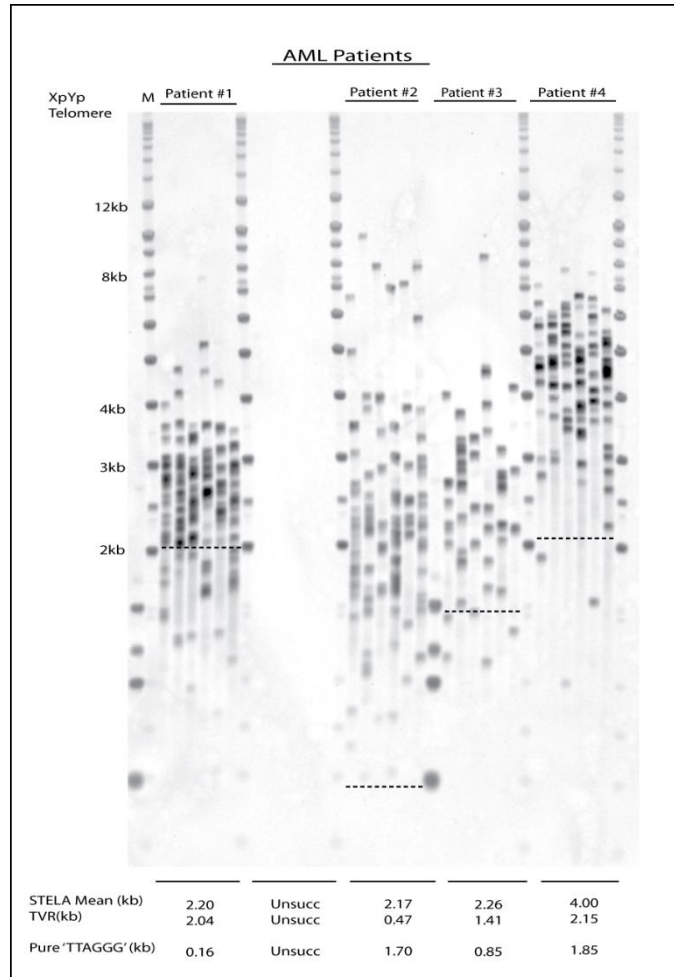


Figure 3.22: The shortest functional TTAGGG tract that was recorded had been derived from an AML patient (patient #1). The pure telomere length was only 0.16kb. Dotted lines illustrate the length of the TVR.

3.11 Cell Fractionation

MDS is considered to arise within cells capable of self-renewal,[19,34] thus it has been proposed that the initial transformation event occurs within a primitive haematopoietic cell. Cells lose their CD34 expression with increased differentiation along their designated pathways, i.e. myeloid or lymphoid.[4,8,9] Cells also progressively lose their self-renewal capacity as they differentiate into their more mature lineage specific counterparts.[4,8,9]

Whole bone marrow aspirates were cell fractionated based on their CD34 cell surface expression from 20 MDS patients. CD34$^+$ cells could be captured by using magnetic beads coated with an antibody specific for the CD34 antigen (Dynabeads) and subsequently isolated using a magnet. Telomere length was analysed in primitive CD34$^+$ and more differentiated CD34$^-$ cells to determine if telomere shortening was more extensive in purified CD34$^+$ cells. Analysis of the telomere distributions within each fractionation showed no difference in telomere length between the CD34$^+$ and CD34$^-$ fractions (Figure 3.23).

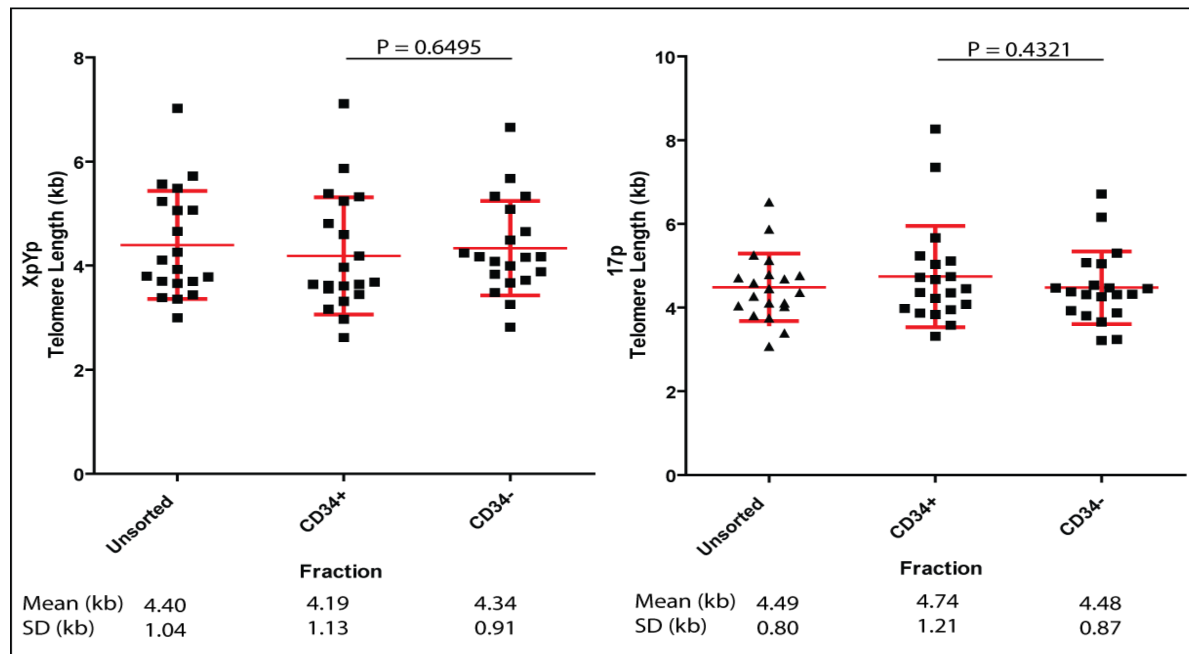


Figure 3.23: Bulk cell populations from a panel of 20 MDS patients were purified into CD34$^+$ and CD34$^-$ sub-populations. Telomere length distributions failed to show a significant difference at the XpYp (p = 0.6495) and 17p (p = 0.4321) telomere within the CD34$^+$ and CD34$^-$ fractions.

The paired analysis of CD34$^+$ and CD34$^-$ sub-populations were not significantly different within the XpYp and 17p telomere distributions presenting p values of 0.6495 and 0.4321, respectively (Figures 3.23 and 3.24). However, the majority of this small group of patients

had been diagnosed with low-risk MDS and the $CD34^+$ population of cells could be differentiating into their $CD34^-$ counterparts, thus an extensive loss of telomere repeats may not have been detected. Conversely, high-risk MDS has been associated with the inability of $CD34^+$ cells to differentiate into their lineage specific cells. Thus, late stage MDS may show an elevated decline in telomere shortening within the $CD34^+$ fraction. Alternatively, excessive apoptosis in early MDS[138,139,141] may lead to the elimination of cells with shorter telomeres also preventing their detection.

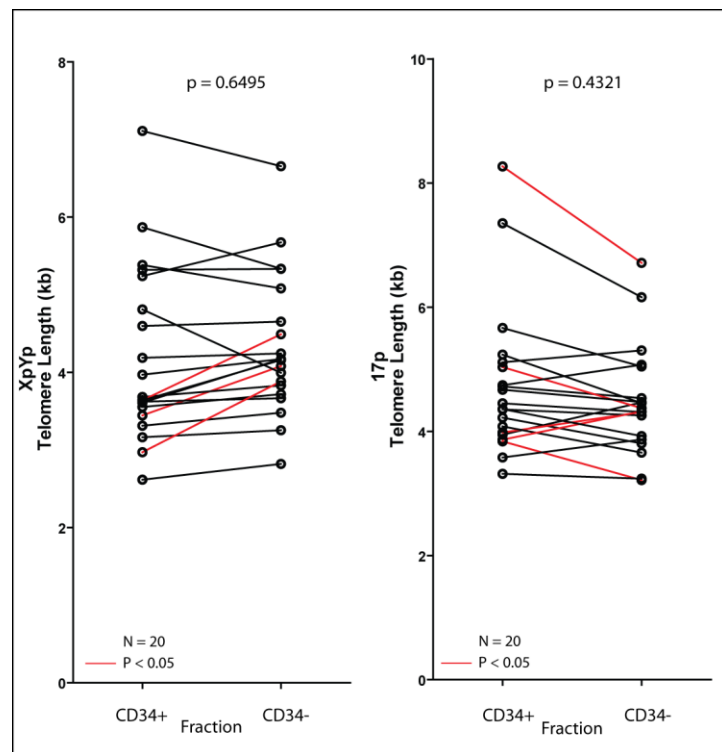


Figure 3.24: No significant difference was detected between paired values of telomere length within each cellular sub-population (p = 0.6495 and p = 0.4321 at XpYp and 17p, respectively. Of note, significant differences (p < 0.05) in telomere length between sub-populations have been highlighted in red.

3.12 Discussion

STELA was employed to determine the nature of telomere dynamics in MDS and AML. STELA is a long-range single molecule PCR approach that amplifies the double-stranded region of telomere repeats from specific chromosome ends. Of particular importance, it is capable of observing the shortest telomere in the distribution that may play a role in disease pathogenesis by means of genetic instability.

It was demonstrated that telomere length is significantly shorter within AML cells ($p < 0.0001$); this was consistent with previous observations.[335,336] The hyperproliferative nature of MDS may enhance the rate of telomere attrition which can eventually initiate genomic instability by means of telomere dysfunction. Notably, short telomeres in MDS have been associated with complex rearrangements and a poorer prognosis.[333,334,342] MDS has been classed as a heterogeneous disease and patient prognosis has been dependent on particular clinical features such as blast count, cytopenia number and cytogenetics.[21,24,43,355] The heterogeneity of telomere length profiles at diagnosis observed within the MDS cohort raises speculation that it may provide prognostic implications including patient susceptibility to AML. This was particularly apparent in one individual that had presented with a telomere length distribution at diagnosis which was shorter than the 25th percentile of the AML cohort; this patient progressed to AML. Further insight into the clinical data associated with patient samples may provide a more detailed analysis on telomere instability in MDS and AML prognosis. This will be referred to in Chapter 4 and 5, respectively.

Homogenisation of telomere length distributions as a function of clinical staging has been observed in CLL cells.[304] The heterogeneity of the telomere length profiles was significantly reduced in patients with AML compared to those in MDS. A reduction in telomere length heterogeneity in AML may be consistent with an accumulation of blast cells and clonal expansion. Additionally, it has been previously implicated that telomerase up-regulation is a feature of AML,[328,332,342,347] thus the maintenance of telomere length homeostasis may result in the reduced telomere heterogeneity within this population of cells.

Telomere length appeared to show a significant correlation between that at XpYp and 17p. Although it appeared that there was a general propensity of telomere shortening at XpYp, it should be noted that variable regions of telomere variant repeats (TVR) may extend the length of the pure telomeric tract. Alternatively, the length of the TTAGGG repeats may have been set longer at 17p. Furthermore, in addition to previous suggestions, *cis*-acting[353] mechanisms may skew the average length of the distribution. Such alterations may arise by means of telomere recombination which was apparent within a bimodal STELA profile at the 17p telomere.

Genotyping of the XpYp telomere-adjacent DNA revealed that 36% of all patients in this study were heterozygous for the GC and AT allele. This was a significant reduction to the expected 45% identified within the Caucasian population.[244] Although this might reflect MDS/AML pathogenesis; loss of the Y chromosome is commonly observed in the bone marrow of elderly males who do not present with any haematological disease.[356] Therefore, it is possible that this finding was a normal age-related phenomenon and not related to disease pathogenesis. However, the loss of a sex chromosome occurs in 50 to 60% of AML presenting with t(8;21).[357] It has been proposed that haploinsufficiency within the pseudoautosomal region (PAR) may support the development of this subtype. The GM-CSF (granulocyte-monocyte colony stimulating factor) receptor, mapped 1180 to 1300kb from the XpYp telomere[358] has been demonstrated to inhibit the immortalisation of t(8;21)$^{+}$ murine and human leukaemic cell lines.[359] Notably, it was proposed that it may play a role as a candidate tumour suppressor by enhancing the differentiation of preleukaemic myeloid cells.[359]

In addition to end-replication losses, telomeres are also subjected to large scale deletion events that occur sporadically amongst the telomere distribution. Telomere rapid deletion (TRD) events have the propensity to induce the abrupt transition of a clone with a high proliferative capacity to a very low replication potential. The highly proliferative nature of this disease may induce replication slippage which has the potential to significantly reduce the proliferative capacity of a haematopoietic precursor. Thus, TRD events may contribute to the cytopenia severity in individual patients, particularly in the presence of a functional DNA damage response (DDR).

Furthermore, these telomeres can be quickly reduced to a critical length which may provoke genetic instability and AML transformation. Interestingly, a RAEB2 patient showed a high number of these deletion events (9.1%) at the 17p telomere. This may have been a mechanism in order to inactivate particular components of the DDR, in particular *p53* which has been mapped to this chromosome. This raises speculation that the loss of the telomere at 17p in a pre-leukaemic state, i.e. MDS may allow for a more proliferative disease that has the potential to transform into leukaemia following the abrogation of *p53*. Consistently, the loss of *p53* has been implicated to result in significantly shorter survival and AML transformation.[135]

A reduction of TRD events was apparent in AML cells and may be associated with the up-regulation of telomerase. Telomerase has shown a preference for short telomeres in mice[209] and human cells,[210] thus its up-regulation may be of importance in the AML clone to reduce the level of genomic instability in the presence of critically short telomeres.

Telomerase up-regulation was detected in AML cells ($p < 0.0001$), consistent with previous observations.[328,332,342,347] Telomerase activity has been detected at low levels in MDS cells[332] and it has been proposed that the level of activity is unable to prevent extensive telomere attrition.[335] However, increased telomerase activity in AML may be vital in preventing excessive telomere erosion in order to circumvent cellular senescence or apoptosis. Despite an accumulation of telomerase competent blast cells in late MDS and AML; increased telomerase may also be a feature of the malignant phenotype that is vital to stabilise the genome acquired in AML cells. Thus, promoting continued proliferation and clonal expansion of leukaemic cells.

Telomeres erode by 33bp/year in proliferating haematopoietic cells;[326] however, this would be expected to increase in cells that are actively proliferating. Accordingly, MDS has been associated with a hypercellular bone marrow; thus it may be expected that the rate of telomere shortening is elevated in haematopoietic cells. The proliferation rate of healthy cells may increase substantially in an attempt to maintain a peripheral blood homeostasis as a consequence of increased apoptosis. Notably, elevated telomere erosion was apparent within the STELA profiles derived from MDS patients. One such extreme case showed telomere erosion at 300bp/year. This is 10 fold higher than that described in healthy cells. Increased cellular proliferation can reduce the capacity of haematopoietic cell reserves by

means of telomere induced apoptosis or senescence increasing the extent of cytopenia severity. Furthermore, accelerated telomere erosion may reduce the latency period prior to AML transformation. Thus, serial samples may be able to determine the extent of cellular proliferation and rate of telomere attrition and provide prognostic information.

The start of the functional TTAGGG tract varies considerably due to the presence of variable quantities of TVRs. The TVR region has been regarded as a non-functional region of the telomere since the Shelterin proteins TRF1, TRF2 and POT1 have a high specificity for the TTAGGG repeat tract.[192] Furthermore, TVR regions have been observed within telomere fusions implicating that they provide no protective capping.[255] The TVR region within each patient could be established by TVR-PCR. Variable measurements from 0 to 3kb had been observed. The shortest pure TTAGGG tract recorded was only 0.16kb and had been derived from an AML patient. Thus, on average the telomere distribution at XpYp was only composed of 27 repeats. More importantly, the STELA distribution derived from this individual revealed that the lower 25th percentile of telomeres was comprised of only 8 TTAGGG repeats. Previous observation has implicated that a telomere of only 13 TTAGGG repeats is sufficient to induce telomere fusion.[255]

It has been previously proposed that MDS initiates from cells capable of self-renewal,[19,34] thus it was speculated that CD34^{+} cells would show accelerated telomere erosion due to extensive cellular proliferation. However, paired analysis of telomere length between CD34^{+} and CD34^{-} purified cells derived from the same patient failed to show a difference in telomere length. The majority of the patients analysed had been diagnosed with low-risk MDS and it was proposed that the apoptotic nature of early MDS[138,139,141] may eliminate the clone that has the potential to show prolonged telomere attrition. MDS CD34^{+} cells may also still be capable of cellular differentiation into their CD34^{-} counterparts and significant differences between the subcellular compartments could go undetected. It may be suggested that extensive telomere attrition of the CD34^{+} fraction would be more apparent in high-risk individuals where reduced apoptosis and a block in differentiation would ensue. Alternatively, an accumulation of genetic events may arise within the CD34 fraction uncoupling function from lineage expression. Thus, a CD34^{+} cell could lose its positive expression and clonal expansion and telomere attrition may be observed within the CD34^{-}

fraction. It is also possible that the novel microenvironment generated by the disease has the propensity to promote cellular replication of stromal cells accelerating telomere attrition within the $CD34^-$ fraction. However, it also raises speculation that telomerase activity might have become up-regulated within a sub-population of cells providing telomere stability and rendering the inability to detect accelerated telomere erosion.

Chapter 4:

Telomere Length and Prognosis in MDS

4.1 Abstract

The Myelodysplastic syndromes (MDSs) are comprised of a heterogeneous group of clonal disorders characterised by ineffective haematopoiesis. Although 30 to 35% of MDS cases progress to Acute Myeloid Leukaemia (AML); the majority of patients die from blood related ailments caused by progressive bone marrow failure.

Telomere length conformed only weakly to the aging dogma in the MDS cohort suggesting that chronological age plays a minimal role in influencing the telomere length of haematological cells within MDS patients.

There was a trend for decreasing telomere length with increasing blast count; however telomere length did not appear to influence the severity of cytopenia, cytogenetics or IPSS risk scores.

When MDS patients were divided into subgroups based on diagnostic telomere length, patients bearing shorter telomeres for their respective age showed shorter overall survival rates. However, telomere length was only able to refine favourable prognostic markers that included good risk cytogenetics ($p < 0.0001$; HR = 27.26; 95%CI 5.538-134.2), uni-lineage cytopenia ($p = 0.0144$; HR = 7.457; 95%CI 1.492-37.26) and low-risk IPSS scores ($p = 0.0489$; HR = 3.026; 95%CI 1.006-9.109). Telomere length did not appear to influence the mortality rate within more unfavourable prognostic subgroups. Thus, telomere length may provide an independent prognostic indicator that could determine disease outlook in patients presenting with low-risk markers. This may be particularly informative provided that patient outcome is heterogeneous within low-risk cohorts. Moreover, the telomere length at diagnosis influenced the overall survival of MDS patients irrespective of conventional markers.

In conclusion, this data suggests that the telomere length at diagnosis may provide a novel independent prognostic marker for delineating MDS patients into those with a favourable or

unfavourable outcome and quite possibly characterise low-risk individuals with the possibility of developing a more aggressive disease.

4.2 Introduction

According to the World Health Organisation (WHO) classification system for haematological cancers the Myelodysplastic Syndromes (MDS) are one of the five major categories of myeloid neoplasms[20] with an estimated incidence of 2 to 12 new cases per 100,000 people each year which has been noted to increase among persons aged 70 or older.[15]

MDS is comprised of a heterogeneous group of clonal disorders associated with ineffective haematopoiesis.[18] One third of patients with MDS will undergo leukaemic transformation into Acute Myeloid Leukaemia (AML), however a significant number die from complications of bone marrow failure without progressing into acute leukaemia.[19]

MDS is suspected following an abnormal complete blood cell (CBC) count but is confirmed by means of a bone marrow (BM) biopsy that evaluates the cellular morphology and percentage of blast cells.[21] Clinically, MDS results in an insufficient production of peripheral blood cells increasing the susceptibility of blood related disorders. Patients develop complications characterised by anaemia, haemorrhage and immune-compromise.[15,24,26]

MDS marrow is normal or hypercellular in 90% of cases but differentiating cells are either dysfunctional or prematurely removed by apoptosis.[22] This results in low peripheral blood counts (blood cytopenia(s)) commonly involving those of the myeloid lineage. Notably, dyserythropoiesis, dysgranulopoiesis and dysmegakaryocytopoiesis causes a reduction of erythrocytes, granulocytes and platelets respectively.[24]

The FAB criterion for the classification of the Myelodysplastic Syndromes (MDS) was devised over 30 years ago. MDS was divided into five subgroups based on morphological features in the bone marrow (BM) and peripheral blood (PB).[24] Marrow composed of fewer than 5% blast cells was regarded as either Refractory Anaemia (RA) or RA with Ringed Sideroblasts (RARS) with the latter composed of over 15% morphologically abnormal erythroid cells.[43] Refractory Anaemia with Excess Blasts (RAEB) was characterised by the appearance of 5 to 20% blast cells with RAEB in transformation (RAEB-T) showing 21 to 30%.[43] A blast cell count of over 30% was referred to as a leukaemic AML marrow by means of the FAB criterion.[43] The World Health Organisation (WHO) later proposed that RAEB be subdivided into RAEBI and RAEBII with regard to marrow blast cell percentage, i.e. 5 to 9% would be classed as

RAEBI whereas 10 to 19% as RAEBII.[355] The WHO lowered the FAB standard for AML classification to 20% blasts provided that patients with 20 to 30% blasts (previously classified as RAEB-T) endured a similar prognosis to those with over 30% blasts.[355] Although the presence of monocytosis could delineate CMML (Chronic Myelomonocytic Leukaemia) from other FAB subgroups,[360] the WHO divided CMML into CMML-1 and CMML-2 due to its heterogeneous outcome.[361] Less that 10% blasts represents that of CMML1 whereas CMML2 has been considered as bone marrow composed of 11 to 20% blasts.[361]

The International Prognostic Scoring System (IPSS) was devised to predict patient outcome and facilitate in making therapeutic decisions, particularly since the prognosis of MDS is heterogeneous.[103,104] This combines diagnostic clinical parameters including the percentage of blasts, number of cytopenia(s) and cytogenetics so as to calculate a risk score which can separate patients into distinctive subgroups comprised of low, intermediate-1 (Int-1), intermediate-2 (Int-2) and high-risk based on their mortality rate and risk of AML transformation.[43] High-risk patients show increased mortality as a consequence of Acute Myeloid Leukaemia (AML) development whereas low-risk patients more likely die of complications associated with bone marrow failure.

Acquired cytogenetic alterations are detected in 50 to 60% of newly diagnosed patients with MDS;[34,35] however low sensitivity methods such as conventional G-banding analysis fail to detect karyotypic abnormalities in a substantial proportion of patients (~50%).[102] Although the karyotype is termed 'normal', patient outcome is heterogeneous[103] with some individuals rapidly deteriorating following diagnosis.[104] In contrast, the overall prognosis of patients who present with a complex karyotype is poor. These have been noted to occur within 15 to 30% of patients at diagnosis[105] with the majority of these patients undergoing leukaemic transformation into AML.[43]

Previous reports have shown increased telomere attrition in MDS patients with multiple cytopenias and low haemoglobin concentrations.[332] It has also been documented that telomere shortening is commonly associated with the presence of complex chromosomal rearrangements.[333,334,342] These data suggest that short telomeres may correlate with an inferior prognostic outlook. Accordingly, the importance of telomere length with respect to

clinical parameters of MDS was investigated in this chapter. Furthermore, telomere length was also analysed to determine if it predicted overall survival of MDS patients.

Results

4.3 Telomere Length and Age at Onset

Healthy individuals show telomere length variation that is influenced by a variety of determinants including age, lifestyle and genetic factors in the zygote, e.g. *hTERT* expression. A genetic contribution has been implicated in Dyskeratosis Congenita (DC) such that telomere shortening and disease anticipation have been linked to a reduction in telomerase activity.[329] However, despite the level of heritability between individuals, environmental influences such as lifestyle and disease status can contribute to the rate of telomere attrition.[238]

Figure 4.1 demonstrates a significant correlation between age and telomere length within cells derived from the peripheral blood of healthy individuals ($r^2 = 0.6001$; $p < 0.0001$). Telomere attrition appears to occur at roughly 53bp/year; a similar erosion rate has been previously documented in haematopoietic $CD34^+$ cells, i.e. 33bp/year.[326] However, considerable telomere length variation is apparent within individuals of a similar age.

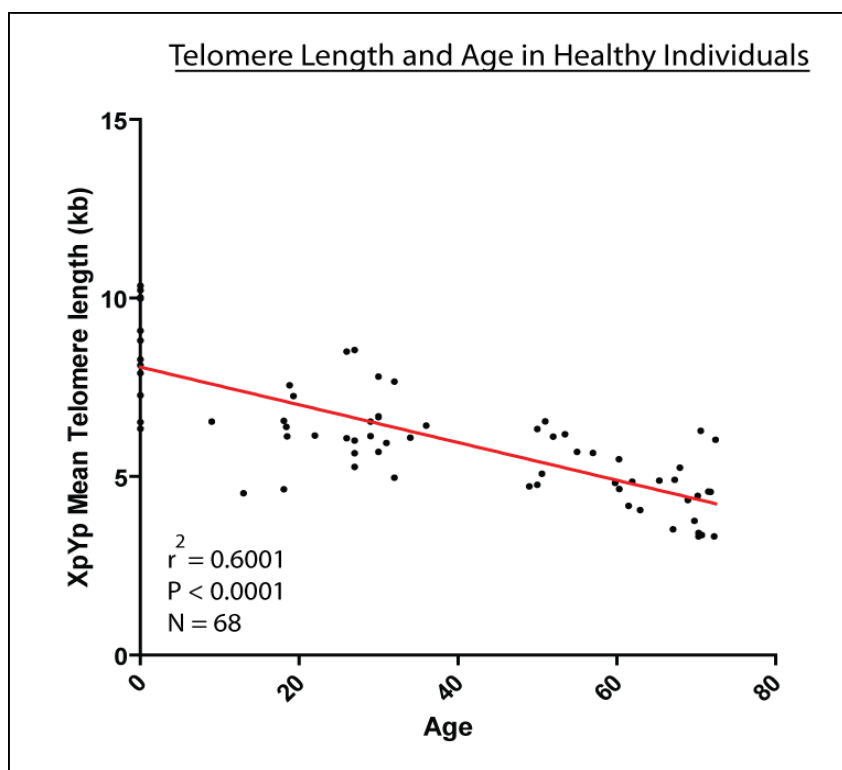


Figure 4.1: A strong negative correlation exists between telomere length and age ($r^2 = 0.6001$; $p < 0.0001$). Progressive telomere shortening occurs with aging. Telomere length measurements of healthy individuals were provided by Duncan Baird.

In the context of MDS and AML, a weak correlation between age at diagnosis and telomere length at the XpYp telomere in MDS (r^2 = 0.06250; p = 0.0253) and AML cells (r^2 = 0.03662; p = 0.0442) was detected (Figure 4.2).

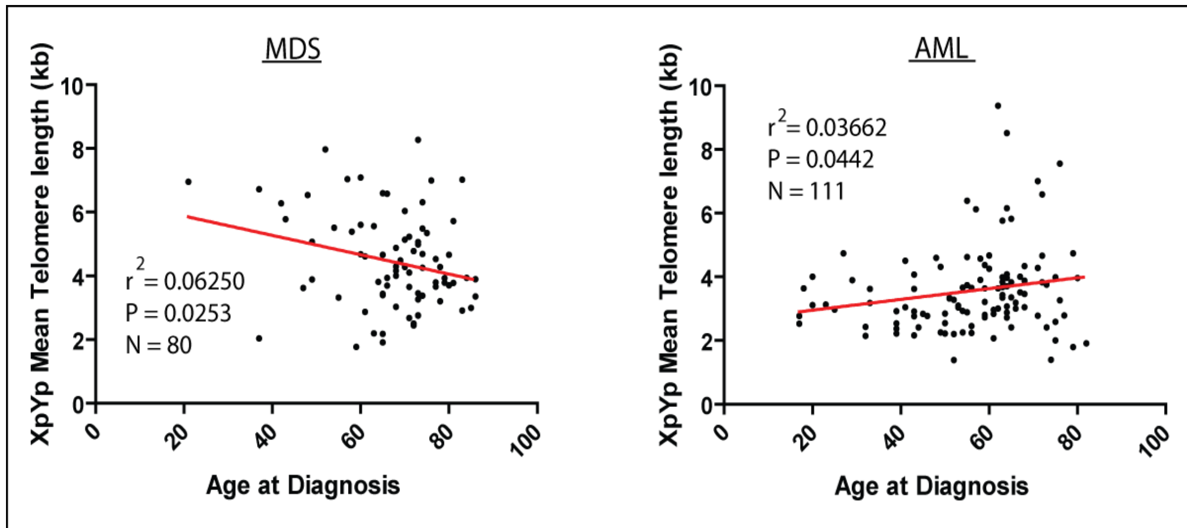


Figure 4.2: Telomere length at XpYp shows a weak negative correlation with age in MDS (r^2 = 0.06250; p = 0.0253) and a weak positive correlation with age in AML (r^2 = 0.03662; p = 0.0442).

Similarly, a weak correlation was detected between age at diagnosis and an age-related decline in MDS (r^2 = 0.01770; p = 0.2395) and AML at the 17p telomere (r^2 = 0.04068; p = 0.3337) (Figure 4.3). Telomere length and age within the MDS cohort continues to conform to the aging dogma. Although only weakly correlated, shorter telomeric length was associated with increasing age in the MDS patient cohort. In contrast, a positive correlation was apparent within the AML cohort, i.e. telomere length appeared to increase with age.

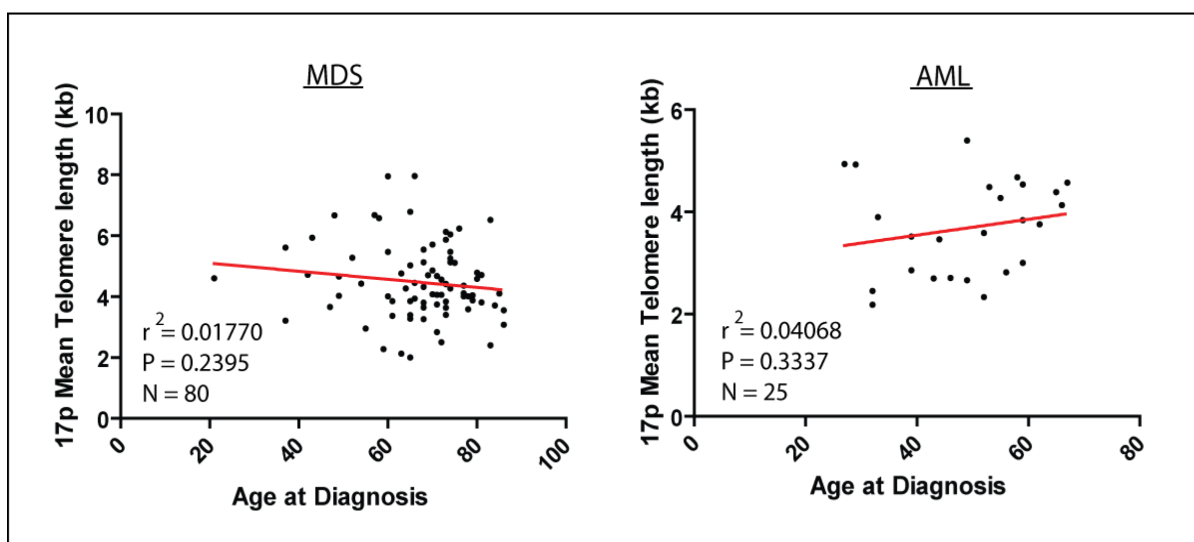


Figure 4.3: Telomere length at 17p shares a weak negative correlation with age in MDS (r^2 = 0.01770; p = 0.2395 and a weak positive correlation with age in AML (r^2 = 0.04068; p = 0.3337).

Extensive heterogeneity in telomere length was observed within the MDS and AML cohorts (Figures 4.2 and 4.3). Many patients showed similar age-related telomere length whereas others presented with much longer and shorter telomere profiles when compared to age-matched healthy controls. Although this may be associated with the duration and/or severity of the disease it also may be consistent with the process behind disease development.

A comparison of age-related decline in MDS and control individuals clearly illustrated that normal age-adjusted telomere erosion cannot explain the radical telomere shortening observed in some patients. For example, one MDS patient presented at the age of 37 years but had a telomere length equivalent to a normal individual >100 years of age (Figure 4.4; highlighted in blue).

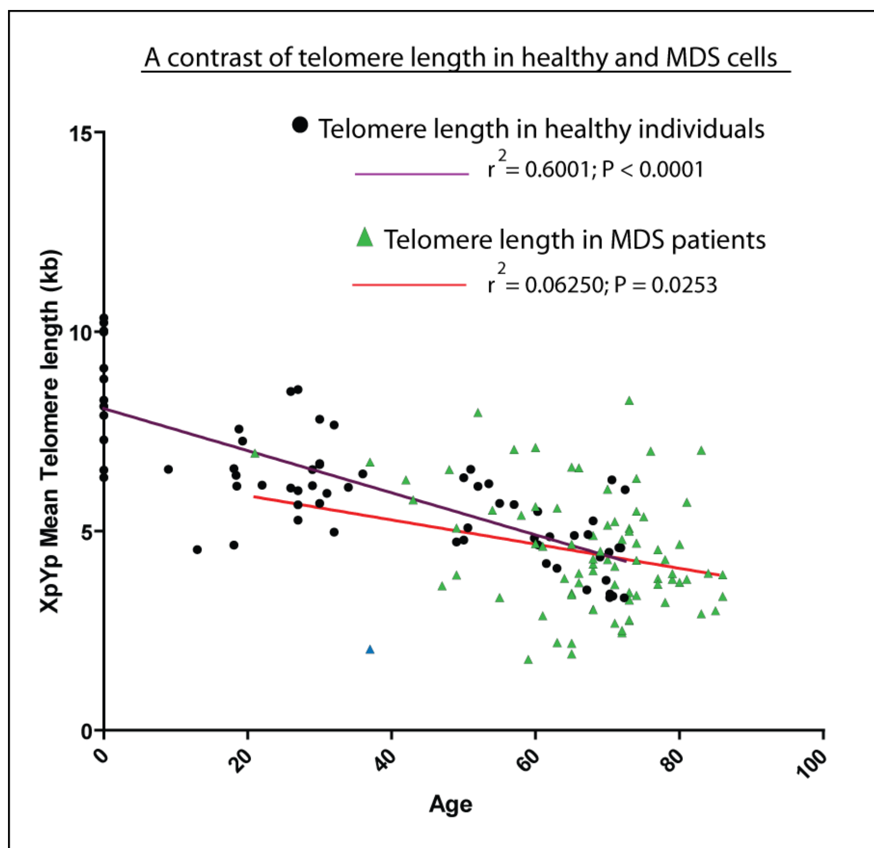


Figure 4.4: Comparing the correlation between telomere length with age in MDS patients and healthy individuals. Elevated telomere attrition is apparent in several MDS cases relative to the telomere length of healthy individuals of a comparative age. A 37 year old patient (highlighted in blue) presented with a telomere length profile equivalent to that expected in a healthy individual >100 years of age.

4.4 Blast Cell Percentage and Telomere Length

Telomere length was significantly reduced in patients with MDS compared to healthy individuals ($p < 0.0001$). Furthermore, when patient cohorts (MDS and AML) were broken down into subgroups classified on blast cell percentage there was a trend towards reduced telomere length with increasing blast count (Figure 4.5). Accordingly, the telomere length at XpYp was recorded as 4.68kb (±1.09kb) within a bone marrow composed of fewer than 5% blast cells, 4.49kb (±1.59kb) in marrow composed of 5 to 20% blast cells and 3.53kb (±1.34kb) within marrow that had a high percentage of blast cells, i.e. over 20%. A significant difference in telomere length was apparent between divisions of <5% and >20% ($p = 0.0009$) as well as 5 to 20% and >20% ($p = 0.0011$). The telomeric length also reduced at the 17p telomere with blast cell expansion. Consistently, telomere length reduced from 4.90kb (±1.01kb) within populations composed of <5% blasts to 3.67kb (±0.88kb) after an accumulation of over 20% blast cells ($p < 0.0001$). Previous reports have demonstrated the existence of a significant negative correlation between blast cell accumulation and telomere length.[336]

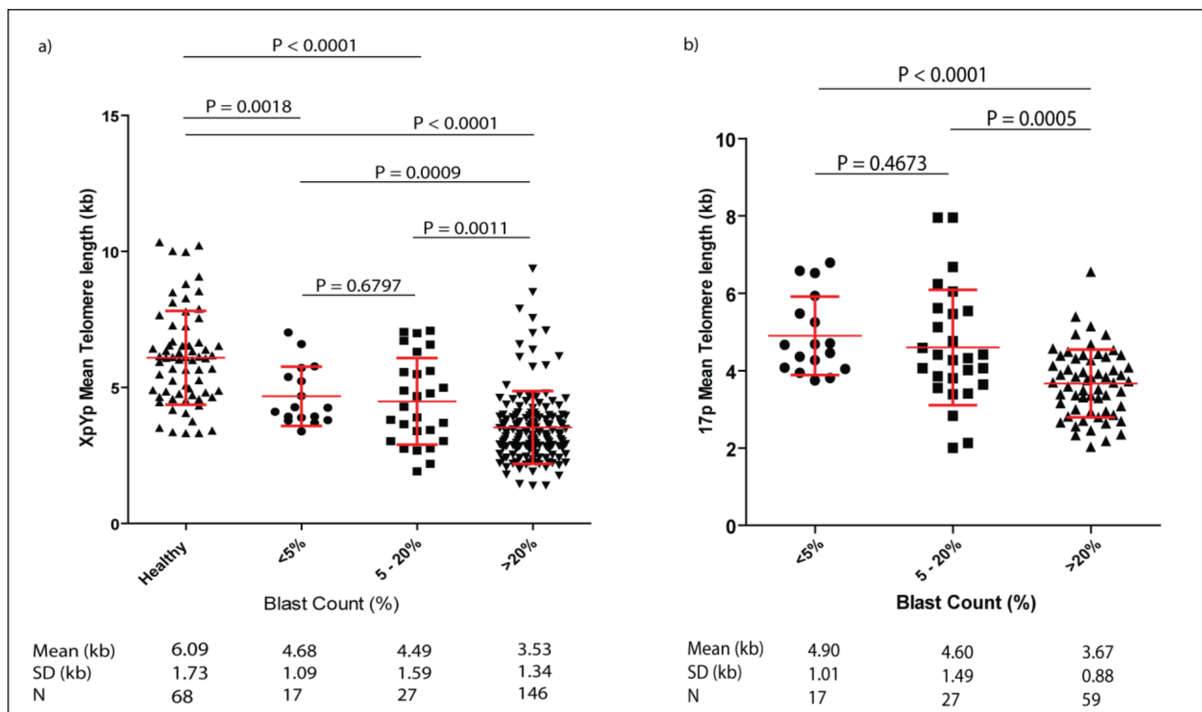


Figure 4.5: Telomere length decreases with blast cell accumulation at XpYp (a) and 17p (b). A significant ($p < 0.05$) reduction in telomere length is apparent between that of 5% and 20% as well as between 5-20% and 20% at the XpYp and 17p telomere.

The intra-clonal variation in telomere length also reduced with increasing blast count (Figure 4.6) with a decline in the standard deviation (SD) from 2.06kb to 1.54kb at the XpYp telomere (p = 0.0017). A significant difference was detected between the SD of the XpYp telomeres of healthy age-matched marrow and that of prognostic subgroups presenting with >20% blasts (p < 0.0001). This reduction in SD probably reflects the increasing clonality of MDS with increasing blast count. The same phenomenon was also apparent at the 17p telomere, which demonstrated a significant reduction in the SD from 2.04kb to 1.51kb with increasing blast count (p = 0.0019).

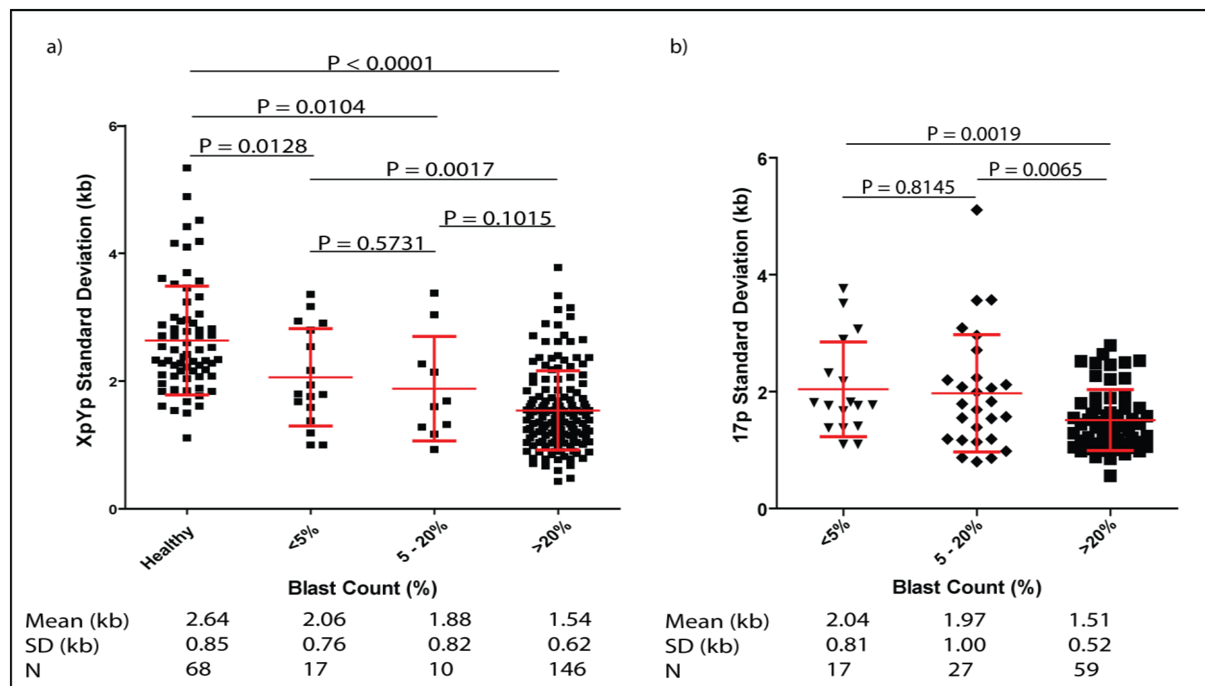


Figure 4.6: Intra-clonal telomere length variation decreased with elevated blast count percentage at (a) the XpYp and (b) the 17p telomere. The difference in telomere length variation was highly significant (p < 0.0001) between that of healthy marrow and a marrow composed of over 20% blast cells. This is consistent with the clonal expansion of a single cell.

It is interesting to note that some patients presenting with long telomere profiles within the 5 to 20% and >20% blast count subgroups, showed similar telomere length to individuals with <5% blast cells (Figure 4.5). To ensure that intra-clonal heterogeneity within each telomere profile was not responsible for this apparent finding, the lower 25th percentile was analysed to reduce any skewing caused by the inclusion of healthy cells (normal marrow) in the profile. Despite plotting the lower 25th percentile, a number of patients within the 5 to 20% and >20% prognostic groups showed long mean telomere length (Figure 4.7). Given the known heterogeneity in the size of the TVR (Telomere Variant Repeats) from one individual to another,[244] it seemed possible that these 'outlier' patients may have longer TVRs thereby

raising the overall telomere length. In patients where the TVR data was available, it was apparent that patients exhibiting long telomere profiles also had long TVR regions that would be influencing the overall telomere length. Those samples with large TVR regions (>1.5kb) are highlighted in red and blue with individuals that presented a pure TTAGGG tract of under 1kb within the lower 25th percentile of telomere distributions highlighted in red (Figure 4.7). Notably, several patients had a pure TTAGGG tract of under 1kb within the >20% blast cohort.

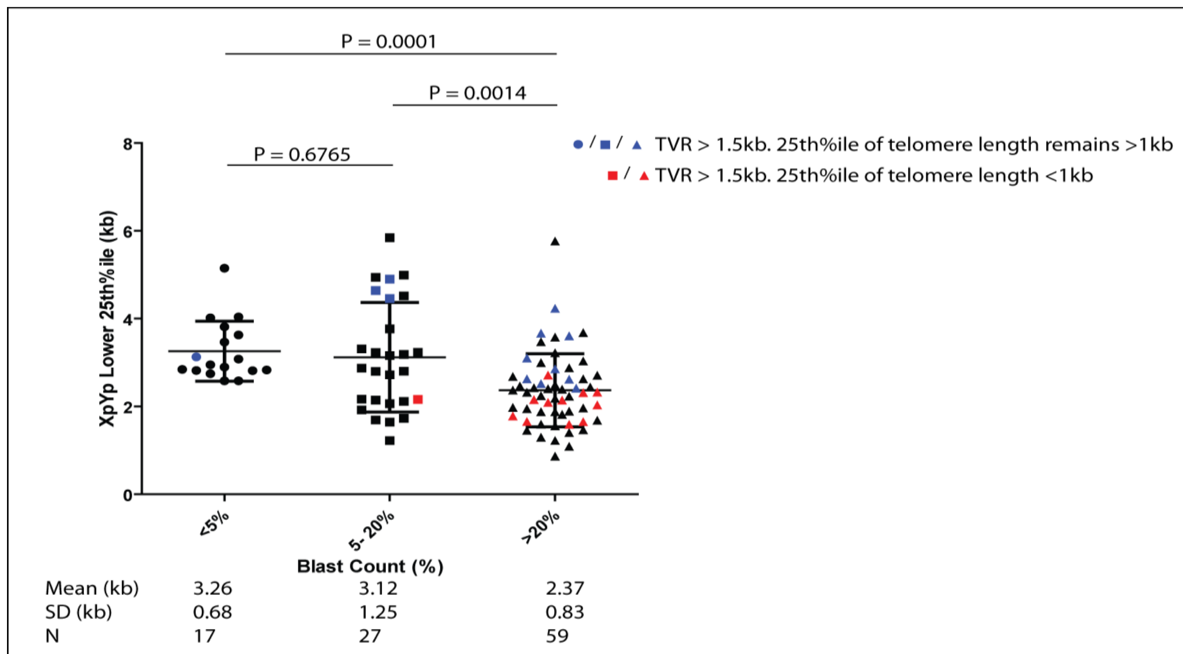


Figure 4.7: Specific telomeric length profiles within blast percentage prognostic subgroups were influenced by expansive TVR regions of >1.5kb. Several patients with over 20% blast cells presented extensive shortening of pure TTAGGG repeats (under 1kb) within the lower 25th percentile of the telomere population (highlighted in red).

4.5 Age-Adjusted Telomere Length in MDS Patients

MDS patients presenting with a telomere length within the normal age-adjusted range did not have a better outcome (p = 0.1278; Hazard Ratio (HR) = 1.606; HR 95% Confidence Interval (CI) 0.8729-2.953) when compared to individuals bearing shorter telomeres than expected in a healthy age-matched individual (Figure 4.8). However, a biphasic distribution was detected within the superior curve which may indicate a possibility that a subset of MDS patients presenting with a telomere length within the normal age-adjusted range may be associated with a better outcome. The relationship between telomere length and overall survival was therefore examined within the currently available prognostic subgroups within the MDS cohort.

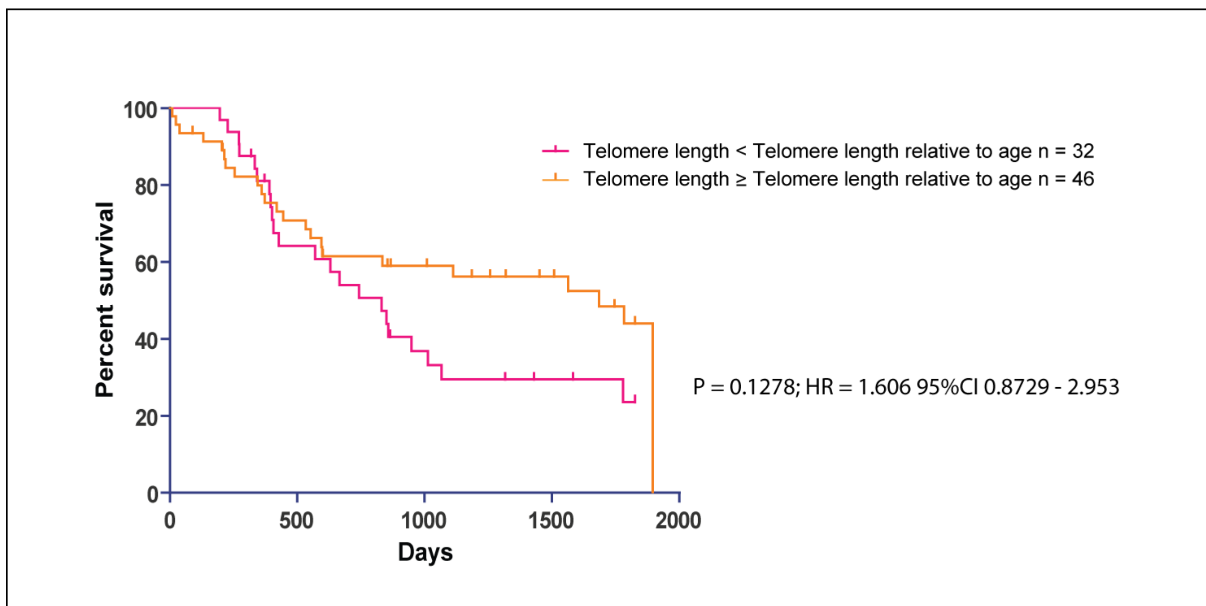


Figure 4.8: Overall survival censored at 5 years. MDS patients presenting with a telomere length within the normal age-adjusted range did not show a more favourable prognosis (p = 0.1278; HR = 1.606; 95%CI 0.8729-2.953).

4.6 Telomere Length and Peripheral Blood Cytopenia

Patients were next stratified based on the number of cytopenias initially present at diagnosis (Figure 4.9). The Kaplan Meier curves demonstrate that the presence of multiple cytopenias, i.e. 2 or 3 is associated with a poorer prognosis and significantly reduced overall survival. Accordingly, the median survival declined significantly from 2110 days in patients with 1 cytopenia to only 595 days and 445 days with 2 and 3 cytopenias respectively ($p < 0.0001$). Given this finding, it was of interest to determine whether a decrease in telomere length may associate with an increased number of cytopenia.

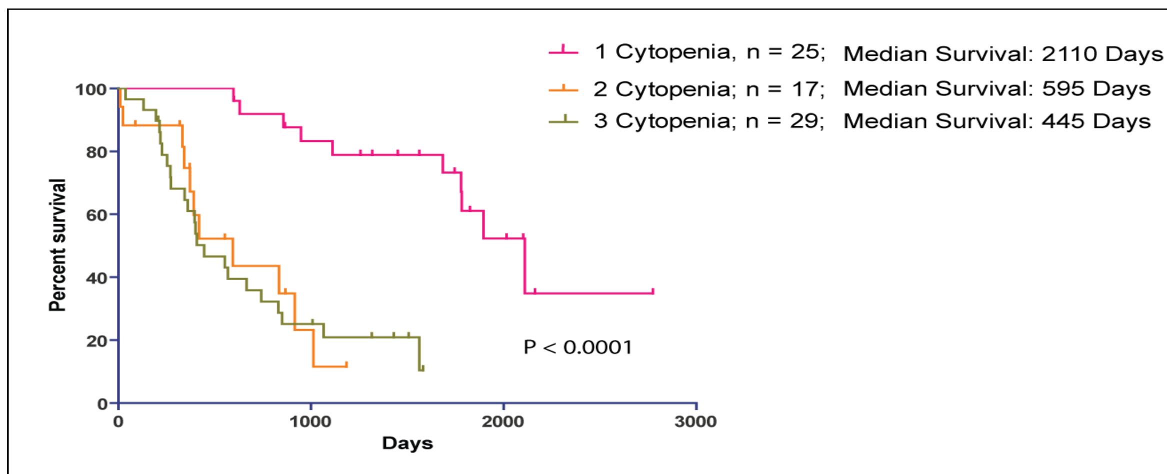


Figure 4.9: Kaplan Meier curves illustrating a significant reduction ($p < 0.0001$) of overall survival in patients presenting multiple cytopenias.

It might be assumed that elevated haematopoietic pressure accelerates the rate of telomere attrition as a consequence of increased $CD34^+$ cell proliferation. Telomere shortening would inadvertently be the result of attempted haematopoietic reconstitution. However, a dichotomy exists as accelerated telomere shortening has the propensity to exacerbate cytopenia severity by reducing the capacity of haematopoietic cell reserves through telomere induced cell cycle arrest or apoptosis.[314]

In this study, the number of cytopenias did not appear to influence the mean telomere length at XpYp or 17p (Figure 4.10). Accordingly, the average telomere length at XpYp was 4.35kb (±1.01kb) within patients showing uni-lineage cytopenia and 4.22kb (±1.76kb) within those that shared tri-lineage cytopenia ($p = 0.7405$). A similar picture emerged when examining the telomere length at 17p. Patients with uni-lineage cytopenia showed a mean telomere length of 4.36kb (±0.90kb) versus 4.32kb (±1.49kb) within those that shared tri-lineage cytopenias ($p = 0.8908$).

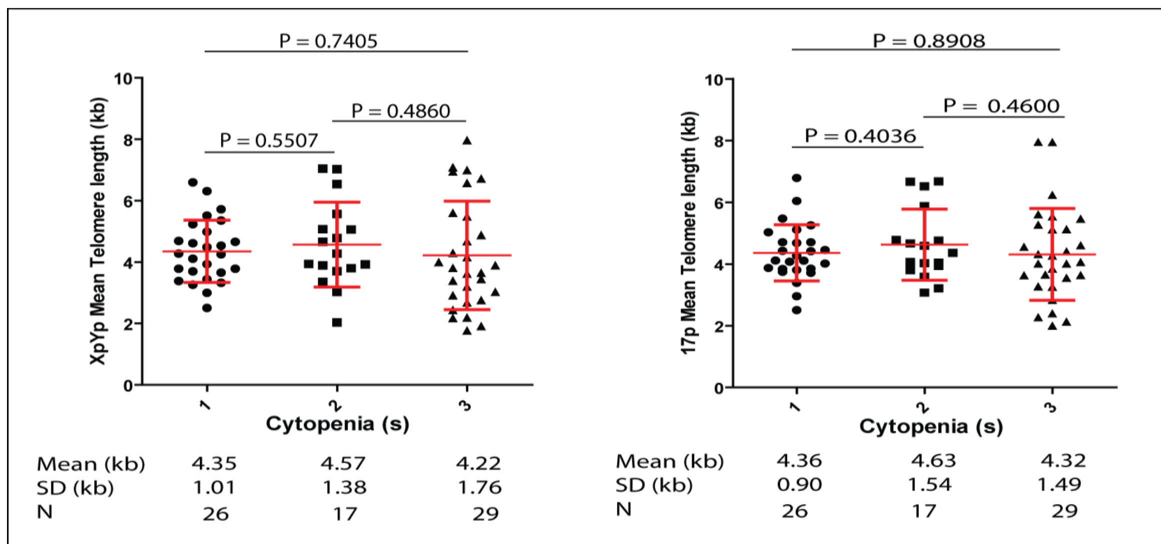


Figure 4.10: Telomere length did not show increased attrition with increasing cytopenia. No significant difference was detected between either of the cytopenia subgroups.

It was apparent that a subset of patients with longer telomeres, particularly within the bi- and tri-lineage prognostic subgroups, might have been responsible for skewing the mean values. Since the number of haematopoietic precursors that undergo increased proliferation is unknown, evaluation of the lower 25th percentile of the telomere profiles may enable a more accurate assessment into cytopenia severity and telomere length. Accordingly, the superimposition of quiescent cells (and potentially long telomeres) has the propensity to skew the average value of the telomere length profiles. It may be assumed that the lower 25th percentile of the telomere distributions could be comprised of a cellular population with shortened telomeres induced by prolonged replication or stochastic telomere rapid deletion (TRD). Notably, short telomeres superimposed within an apparently larger average telomeric distribution could contribute to the cytopenia severity in MDS patients.

However, telomere shortening was not significant between the cytopenia subgroups even when comparing the lower 25th percentiles at the XpYp and 17p telomere (Figure 4.11). The uni-lineage cytopenia group had a mean lower 25th percentile of 3.13kb (±0.79kb) at XpYp and 3.18kb (±0.66kb) at 17p, whereas the mean lower 25th percentile of the tri-lineage group was 2.88kb (±1.24kb) at the XpYp telomere and 3.00kb (±0.95kb) at the 17p telomere. These differences did not reach statistical significance with p = 0.3797 at XpYp and p = 0.4167 at the 17p telomere.

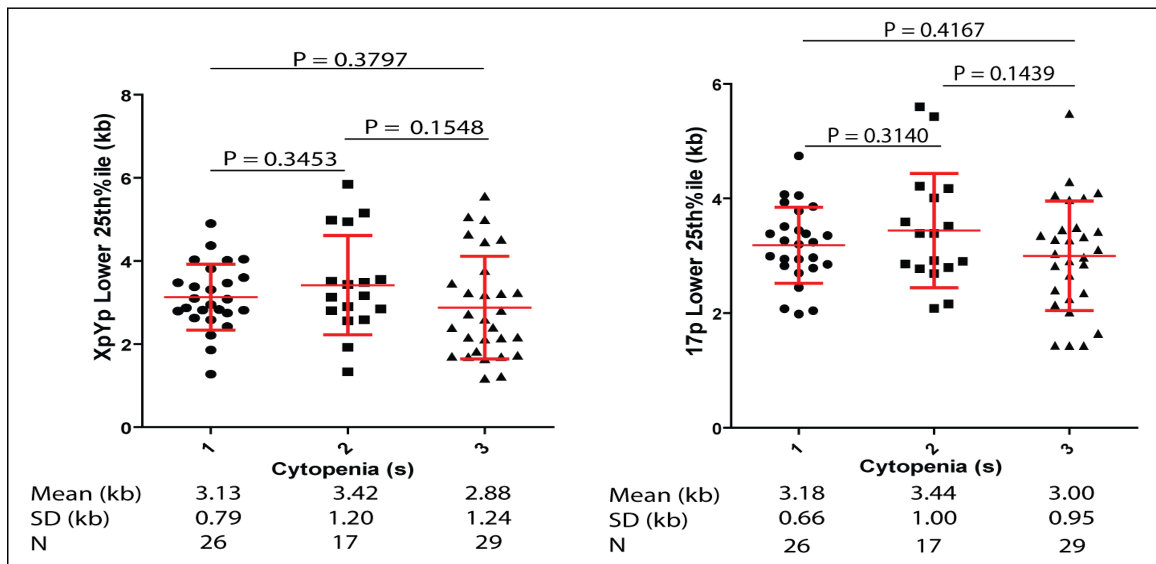


Figure 4.11: Telomere length within the lower 25th percentile was not significantly different within prognostic cohorts with multi-lineage cytopenia not showing elevated telomere shortening.

Longer telomeres were present within each of the cohorts, which may have been influenced by the length of the TVR region. Patients showing TVR regions above 1.5kb were noted within each cohort (Figure 4.12).

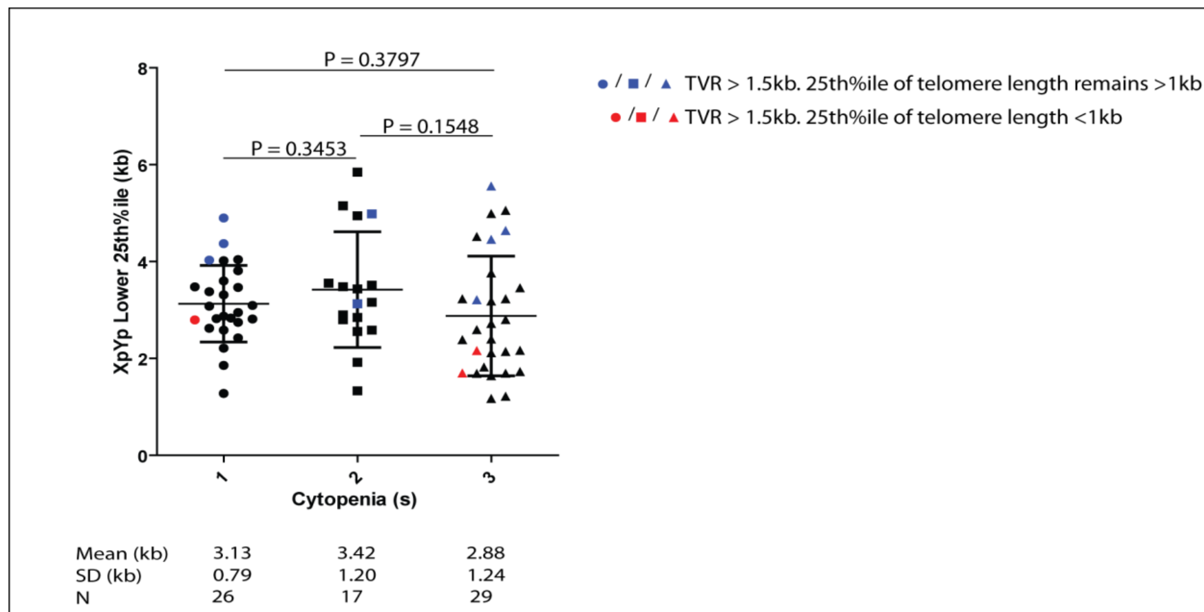


Figure 4.12: Specific telomeric length profiles within cytopenia subgroups were influenced by expansive TVR regions of >1.5kb (red/blue). It was apparent that several individuals continued to show long pure TTAGGG tracts despite presenting with multi-lineage cytopenias.

To determine whether the pure TTAGGG tract declined with cytopenia severity, TVR regions were substracted from the average telomere length and analysed within each of the cytopenia cohorts (Figure 4.13). Although not significant (p = 0.1014), there was a general propensity of telomere shortening with progressing cytopenia. Accordingly, the pure

telomeric length within the lower 25th percentile was 2.86kb (±0.85kb) and 2.21kb (±1.32kb) within the uni- and tri-lineage cytopenias respectively.

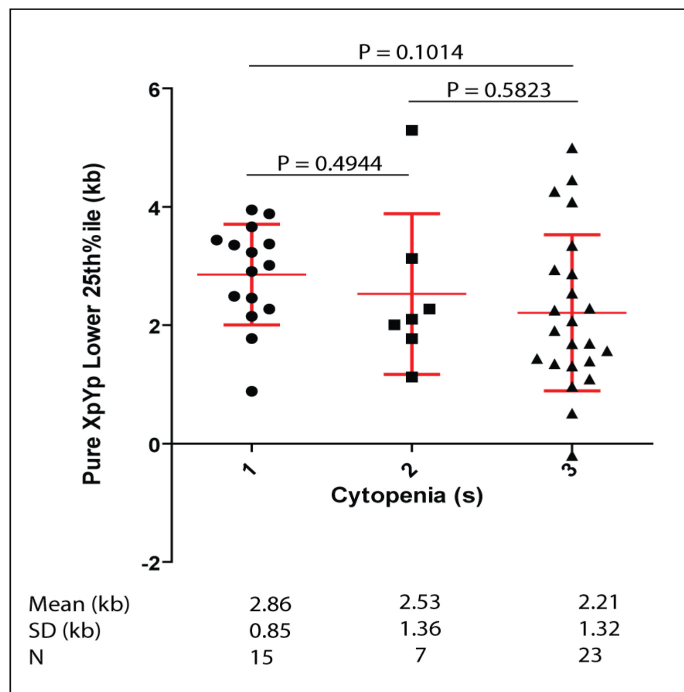


Figure 4.13: Although not statistically significant there was a general tendency of pure TTAGGG shortening with increased cytopenia.

The telomere length profiles were looked into further by taking into account the patient age and respective telomere length (Figure 4.14).

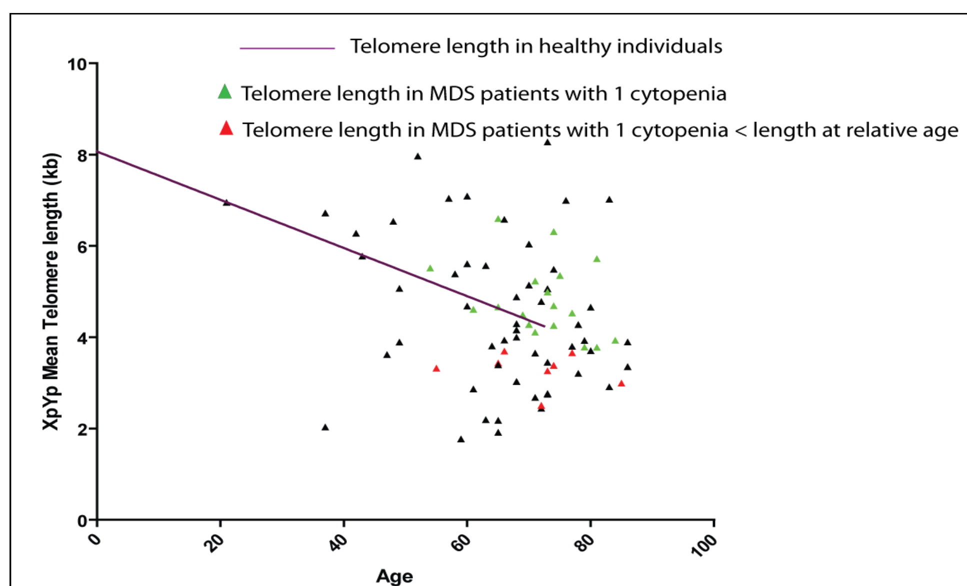


Figure 4.14: Telomere length was analysed further by taking into account relative age. Individuals diagnosed with a single cytopenia are highlighted in green. Those showing shorter telomeres relative to chronological age are presented in red.

Patients presenting with a single cytopenia who showed a telomere length below that expected in a healthy age-matched individual had a much shorter survival time (Figure 4.15). Consistently, some of these patients shared similar mean telomere lengths to those individuals presenting with two or three cytopenias. Differences in survival times were particularly apparent within the first subgroup (cytopenia 1; p = 0.0144) with a hazard ratio (HR) of 7.457 (HR 95% confidence interval (CI) 1.492-37.26).

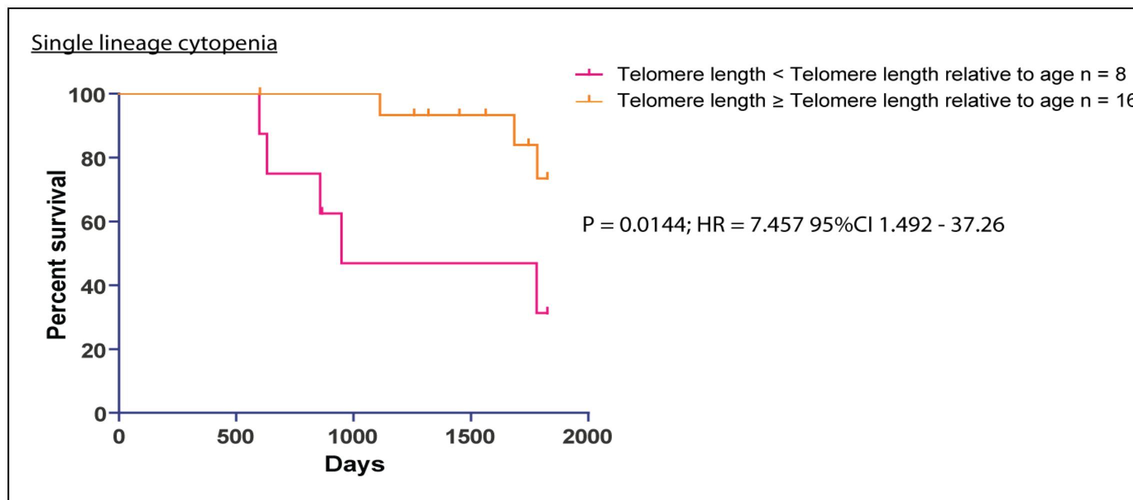
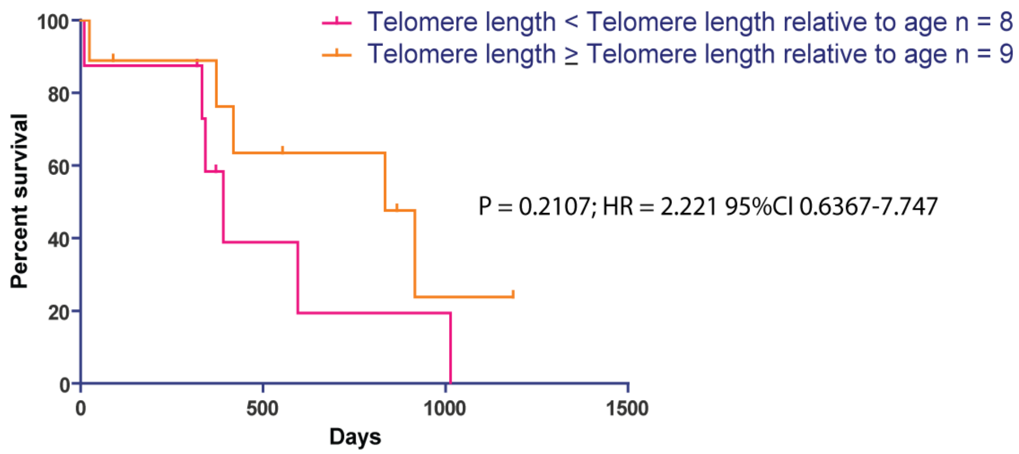


Figure 4.15: Overall survival censored at 5 years. Telomere length relative to age was able to stratify patients with a single cytopenia into favourable and unfavourable prognostic subgroups (p = 0.0144; HR = 7.457; 95%CI 1.492-37.26).

In contrast, the ability of telomere length to refine subgroups with altered overall survival was lost in prognostic groups showing multi-lineage cytopenias (Figure 4.16). Accordingly, there was no significant difference in survival curves when comparing the overall survival within the cytopenia subgroups two (p = 0.2107) and three (p = 0.7499).

2 Cytopenias



3 Cytopenias

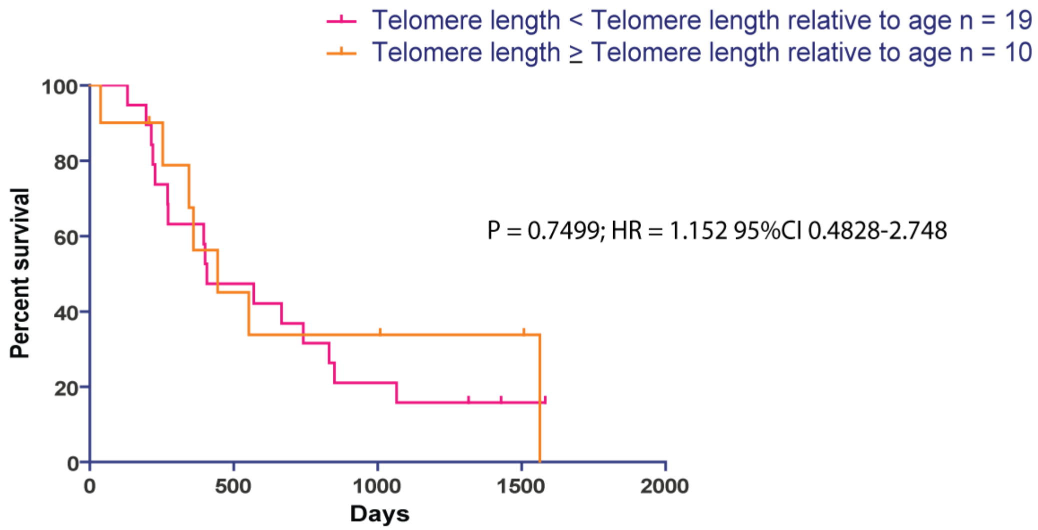


Figure 4.16: Telomere length was unable to delineate patients presenting multiple cytopenias into favourble and unfavourable prognostic subgroups.

4.7 Telomere Length and Cytogenetic Risk in MDS

Overall survival was determined within cytogenetic risk groups using the log rank test and was depicted in the form of Kaplan Meier curves. It was observed that the median survival declined within patients presenting with a poor cytogenetic profile. Notably, the median survival time for patients within the good cytogenetic subgroup was 1779 days in contrast to only 421 days within the poor cytogenetic subgroup (Figure 4.17).

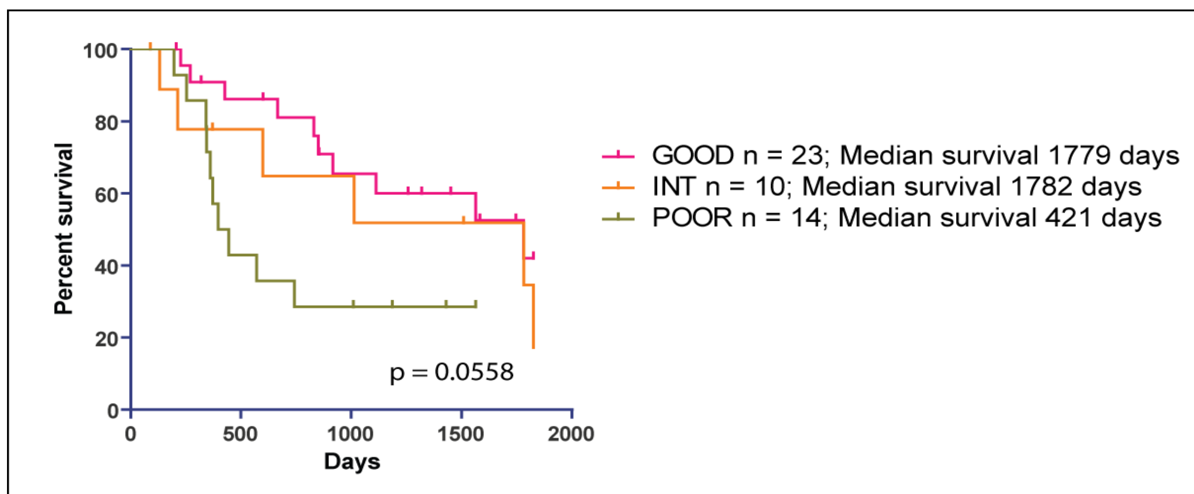


Figure 4.17: Kaplan Meier curves illustrating reduced survival within patients that present with poor-risk karyotype at diagnosis (p = 0.0558).

It has been implicated that telomere dysfunction may provide a stepwise accumulation of cytogenetic changes during cancer development through the formation of dicentric chromosomes generated by successive breakage-fusion-bridge (BFB) cycles.[319] Thus, it was speculated that increased telomere attrition may be more apparent within the poor cytogenetic subgroup provided that several of these patients present an abnormal karyotype composed of three or more chromosomal alterations.[105]

However, this was not the case when analysing telomere length at XpYp and 17p (Figure 4.18). There was no significant difference in telomere length between either of the cytogenetic cohorts and the telomere length within the poor prognostic group did not show increased telomere attrition. Accordingly, the telomere length at XpYp was 4.38kb (±1.18kb) and 4.87kb (±2.03kb) within groups presenting with good and poor risk cytogenetic profiles, respectively. Conversely, the telomere length at 17p was 4.43kb (±0.99kb) and 4.95kb (±1.93kb) in the good and poor risk cohorts, respectively.

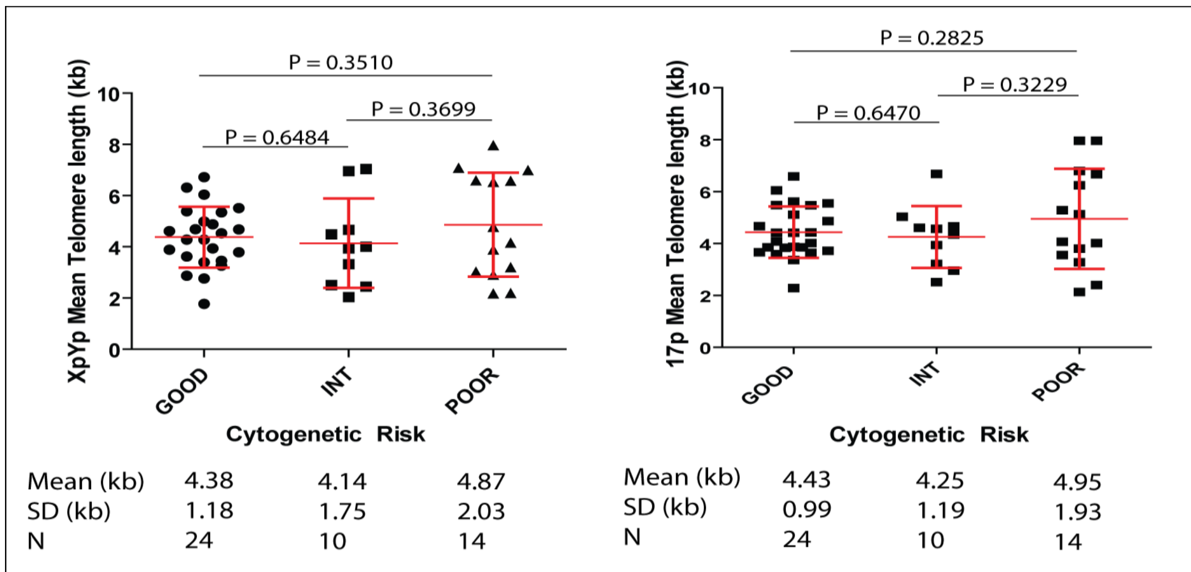


Figure 4.18: Telomere length did not show increased attrition with cytogenetic severity. No significant difference was detected between either of the cytogenetic profiles.

A substantial number of MDS patients (~50%) show a normal karyotype when analysed using conventional cytogenetic analysis.[102] However it has been noted that this group can be fairly heterogeneous with respect to progression.[103] A comparison of telomere length relative to age (Figure 4.19) demonstrated a significant difference ($p < 0.0001$; HR = 27.26; 95%CI 5.538-134.2) in mortality rate. Patients exhibiting telomere profiles shorter than expected for their chronological age had a reduction in overall survival when compared with patients that presented with longer telomeres.

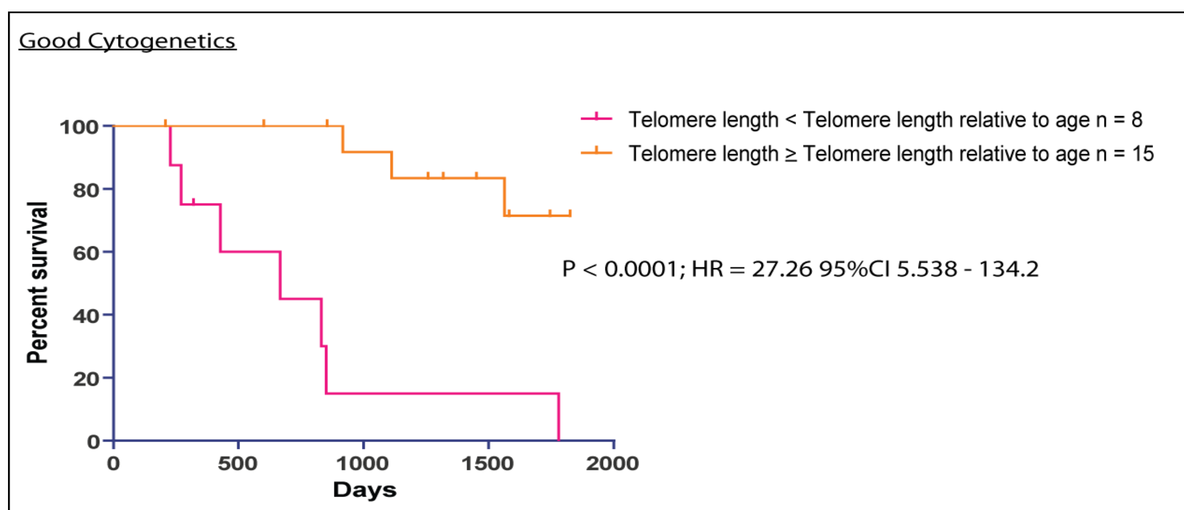


Figure 4.19: Telomere length relative to age was able to delineate patients with a favourable and unfavourable outcome within the good prognostic cohort. Overall survival was significantly reduced with elevated telomere shortening (p < 0.0001; HR = 27.26; 95%CI 5.528-134.2).

The impact of telomere length was lost in the remaining cytogenetic prognostic groups; p = 0.8618 and p = 0.1911 within the intermediate and poor cohorts, respectively (Figure 4.20).

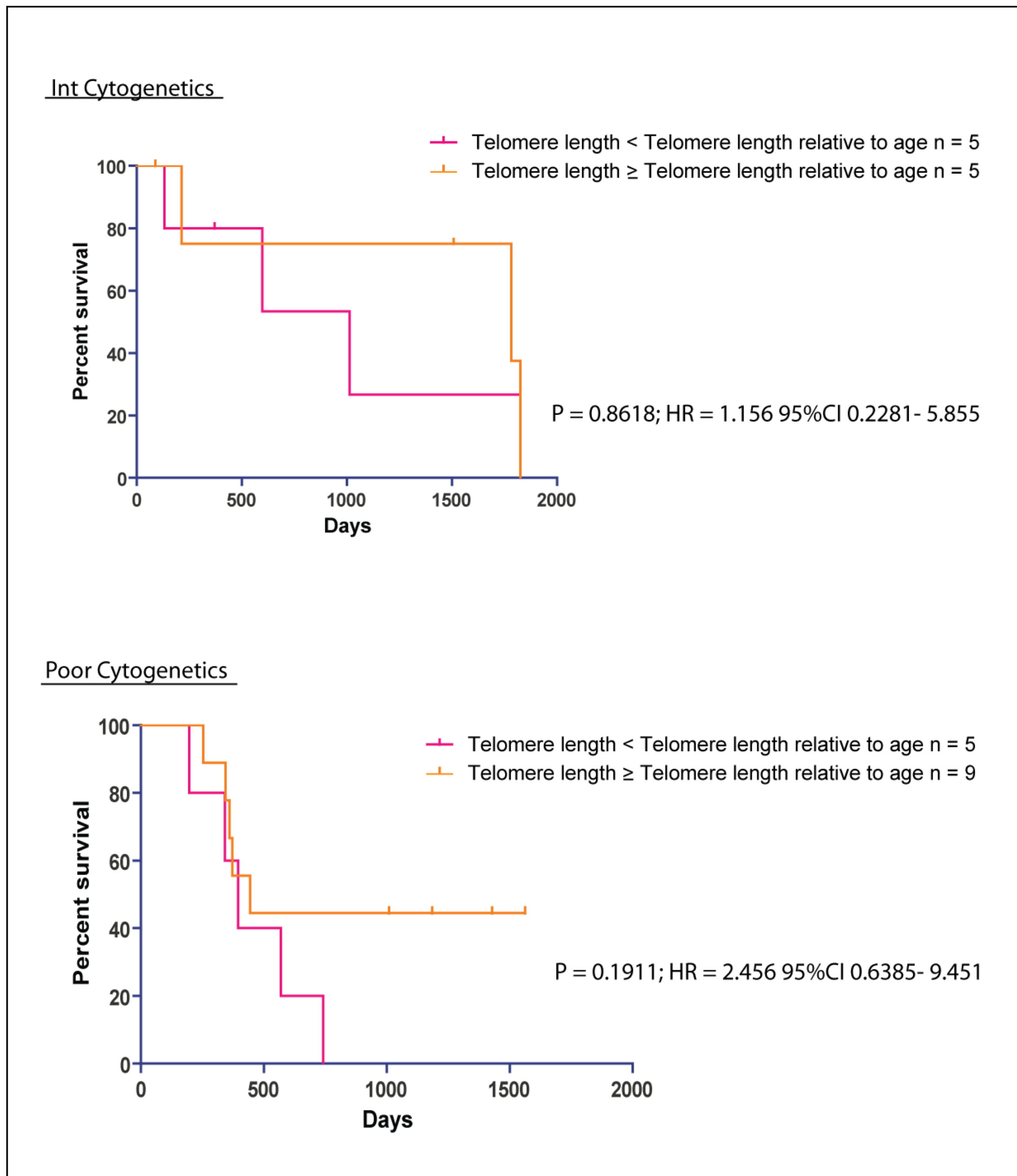


Figure 4.20: Telomere length was unable to stratify patients that showed an intermediate or poor karyotype into favourable or unfavourable subgroups.

The risk of AML transformation has been observed to be very high in patients that show a poor cytogenetic profile.[43] This is often consistent with the accumulation of genetic alterations and positive selection of leukaemia-favouring mutations. The lower 25th percentile of the telomere distribution is most likely enriched with cells with an increased propensity to undergo telomere fusion.[255] Accordingly, this population of telomeres were

analysed at the XpYp and 17p telomere within each prognostic subgroup (Figure 4.21). Although no significant difference was detected between each group, patients that developed AML were identified to determine if telomere length played a role in AML transformation. Several individuals exhibiting short telomeres developed AML (Figure 4.21; highlighted in red), this was particularly apparent at the XpYp telomere within the good prognostic subgroup. Telomere shortening may not have been as pronounced at the 17p telomere within these individuals possibly due to a larger TVR. This is consistent with the regression analysis in chapter 3 (Figure 3.5) which showed that the telomere at 17p was longer by 1.63kb within the MDS cohort. Furthermore, the majority of these individuals (apart from one) exhibited multiple cytopenias. This may suggest that increased proliferative pressure accelerating telomere attrition predisposes individuals to AML transformation by means of telomere induced genetic instability. Telomere shortening has been linked to non-reciprocal translocations,[322] thus patients presenting with long telomere profiles despite AML progression may have acquired genetic alterations independent of telomere dysfunction, i.e. reciprocal translocations, inversions and whole chromosome losses and gains.[342] One such individual gained an additional chromosome 8 (Trisomy 8-highlighted in blue in Figure 4.21) which is unlikely to have developed from telomere attrition.

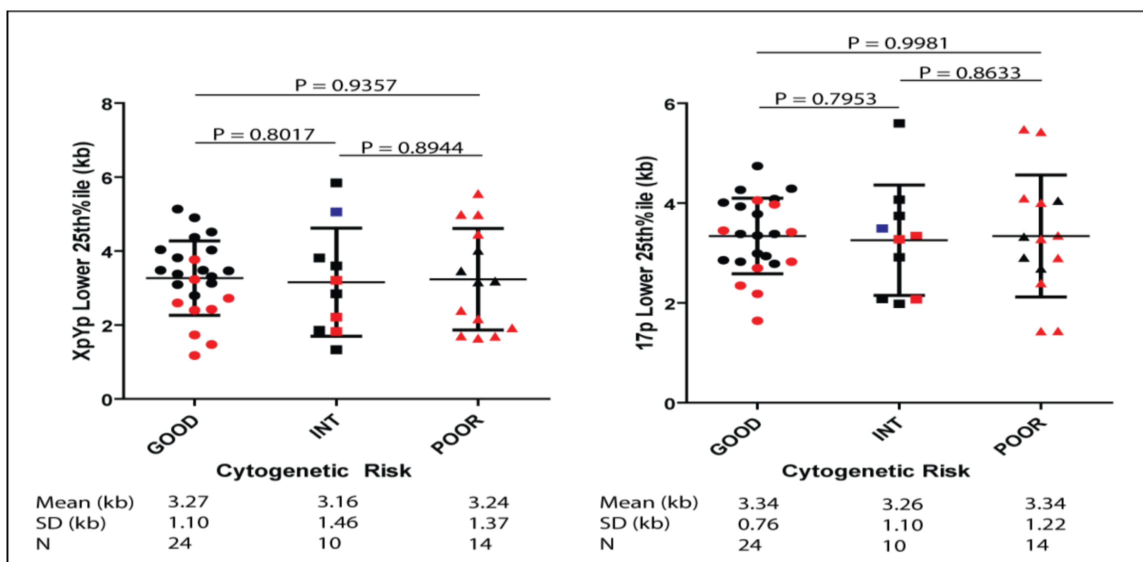


Figure 4.21: Telomere length was not significantly different within either of the cytogenetic risk groups. Several individuals progressed to AML (red). Furthermore, patients that presented long telomeres at diagnosis might have developed a karyotypic abnormality not associated with telomere dysfunction, i.e. reciprocal translocations, inversions and whole chromosome losses and gains. One such example is highlighted in blue. This patient presented with Trisomy 8 at diagnosis.

Telomere shortening may have an important prognostic role independent of the cytogenetic profile provided that short telomeres have the propensity to enter into BFB (breakage-fusion-bridge) cycles that could initiate the development of an abnormal clone and enable AML propagation. Patients exhibiting telomere shortening may be identified as those with a more inferior prognosis. This was apparent within the good prognostic subgroup with respect to survival and AML transformation.

4.8 Telomere Length and the IPSS Scoring System

Patients were separated into prognostic cohorts based on their IPSS (International Prognostic Scoring System) risk score. The risk score stratifies individuals into low, intermediate-1 (Int-1), intermediate-2 (Int-2) or high by combining independent variables including blast percentage, cytopenia(s) and cytogenetics in order to predict patient outcome and facilitate in making therapeutic decisions.[43]

Telomere length was analysed within different risk groups to determine if it correlated with the severity of prognostic features. The average telomere length at XpYp and 17p failed to show any correlation with IPSS prognostic risk groups (Figure 4.22). Accordingly, the telomere length at XpYp was 4.28kb (±0.82kb) within the low-risk group and was 4.46kb (±1.71kb) within patients that had a high IPSS score. Similarly, the telomere length at 17p was 4.38kb (±0.76kb) and 4.51kb (±1.45kb) within the low and high-risk prognostic groups respectively. However, there appeared to be an increase in the heterogeneity with increasing risk score. Accordingly, the standard deviation recorded at the XpYp telomere was ±0.82kb and ±1.71kb within the low and high-risk, respectively. Similarly, that at 17p was observed as ±0.76kb in patients presenting low-risk and ±1.45kb within those presenting high-risk IPSS scores. This may share some relation with how patients are stratified into prognostic subgroups. Notably, risk scores are defined by cytogenetics, cytopenia and blast count.[43]

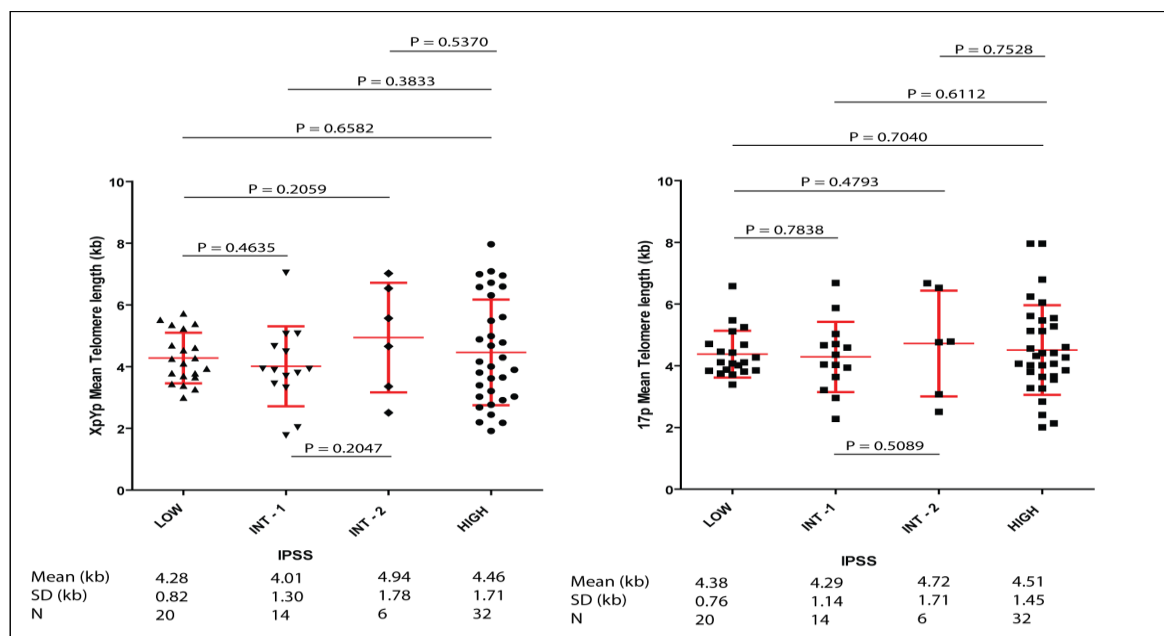


Figure 4.22: An increase in the IPSS score was not associated with elevated telomere attrition. Statistical significance was not reached between any of the prognostic cohorts.

The lower 25th percentile of the telomere profiles within each subgroup was determined to see if the shorter distributions within a population of telomere molecules had the potential to influence the prognostic scoring of each patient (Figure 4.23). It appeared there was a trend for telomere shortening in high-risk cohorts. Notably, the telomere length at XpYp was 3.07kb (±0.69kb) and 3.03kb (±1.21kb) within the low and high-risk, respectively and that at 17p was 3.22kb (±0.61kb) within the low-risk group and 3.09kb (±0.90kb) within the high-risk cohort. However, statistical significance was not achieved between any of the prognostic groups. Patient groups were examined further by determining which individuals had undergone AML transformation (Figure 4.23; AML transformation highlighted in red). The majority of patients developed AML within the high-risk cohort but it was also apparent that inter-patient telomere length heterogeneity existed. Notably, several high-risk patients presented a population of telomeres within the 25th percentile that were longer than those detected in the low-risk cohort. However, it was also noted that a minority of individuals with short telomeres developed AML within the low and Int-1 prognostic cohorts (Figure 4.23).

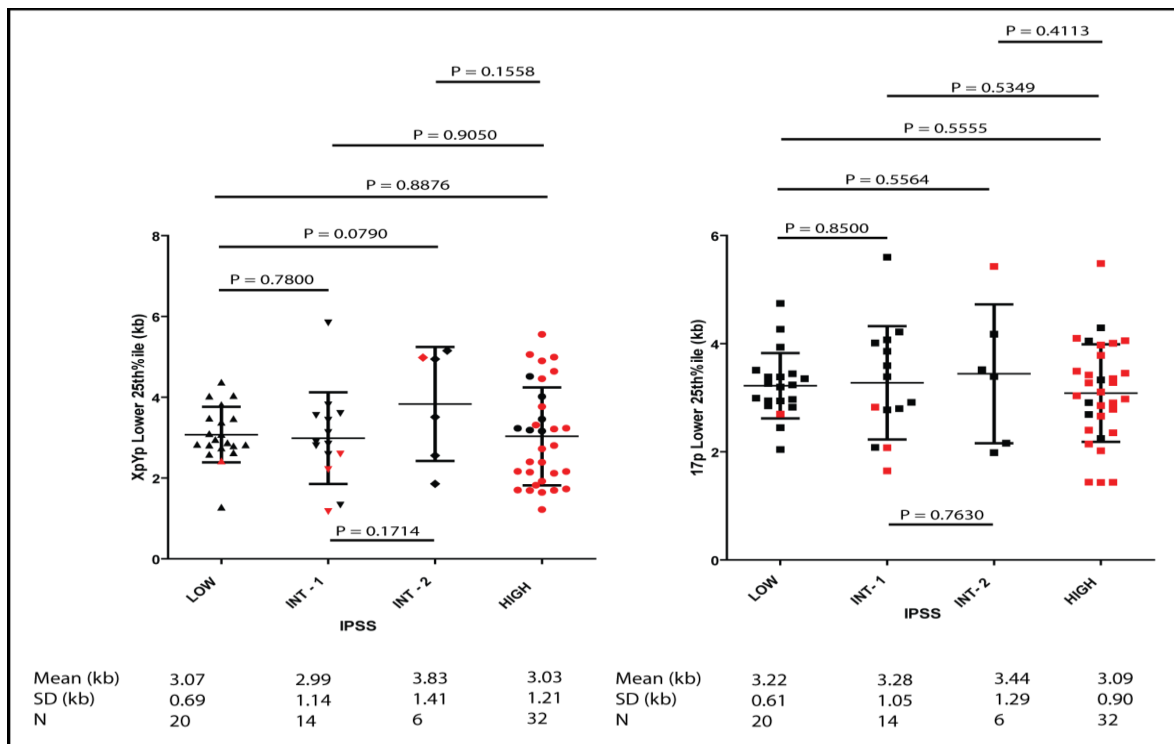


Figure 4.23: An increase in the IPSS score was not associated with elevated telomere attrition within the lower 25th percentile of STELA distributions. Statistical significance was not reached between any of the prognostic cohorts. Patients that progressed to AML have been highlighted in red with the majority of individuals developing overt leukaemia within the high prognostic subgroup.

It was speculated that the TVR region might have influenced the telomere length within some of these patients. In order to examine this, low and Int-1 subgroups were combined together as 'Low-risk MDS' and the Int-2 and high-risk were considered 'High-risk MDS'. The telomere length did not show any significant difference again between the low and high-risk groups (p = 0.4460). However, the high-risk MDS group had a lower median telomere length than the low-risk MDS group. The median telomere length within the 25th percentile of the low-risk group following TVR removal was 2.52kb and that within the high-risk group was 2.17kb (Figure.4.24).

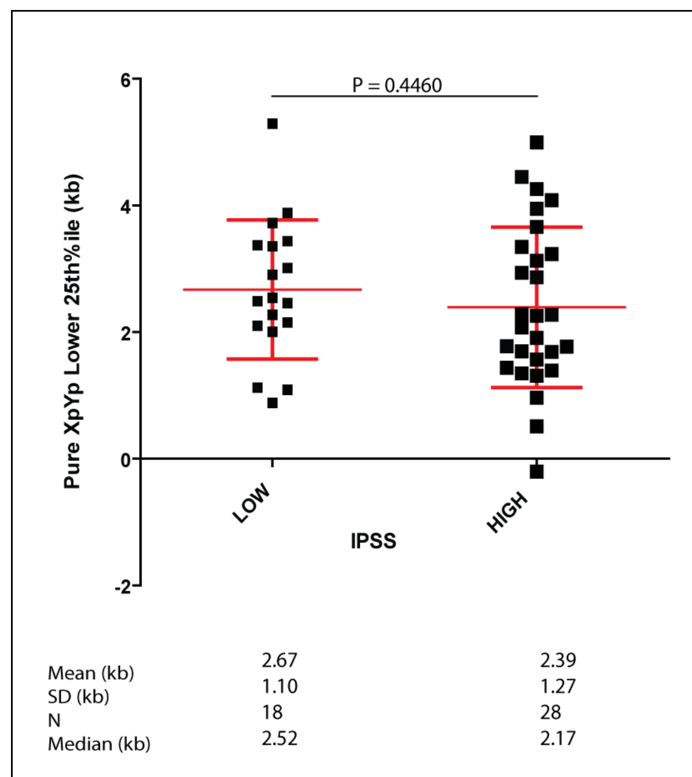


Figure 4.24: Statistical significance was not achieved between the pure TTAGGG tract within the low and high IPSS cohorts (p = 0.4460). However, the median pure TTAGGG length within the 25th percentile was shorter in high-risk MDS.

Patients were further analysed to determine telomere length relative to age (Figure 4.25). Low-risk (Low/Int-1) patients with long telomeres did not progress to AML whereas a minority of individuals with short telomeres had undergone leukaemic transformation. This may be of particular importance when stratifying low-risk patients with regard to AML risk. In contrast, the telomere length was heterogeneous with respect to AML progression in patients presenting high-risk profiles (Int-2/High).

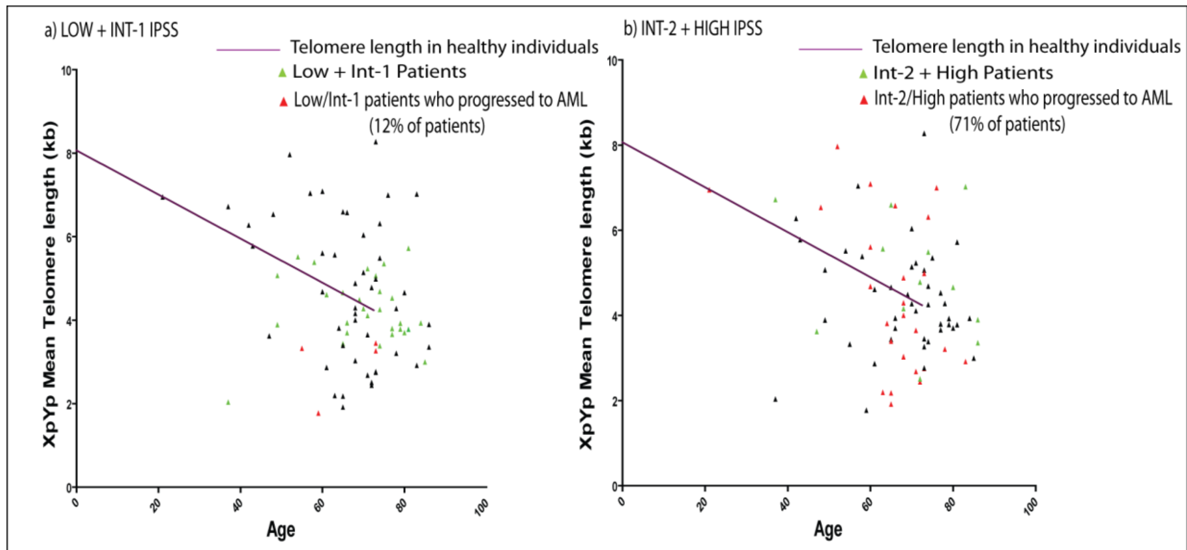


Figure 4.25: 12% of patients within the low-risk prognostic group developed AML. All of which had shorter telomere distributions when compared to age-matched controls. Conversely, the majority of individuals (71%) progressed to AML within the high-risk group but telomere length distributions were heterogeneous.

Furthermore, there appeared to be a tendency for reduced survival (p = 0.0489; HR = 3.026; 95%CI 1.006-9.109) within the low-risk prognostic group in patients with short telomeres for their respective age, however this was not apparent within the high-risk group (p = 0.0938) (Figure 4.26). Since the majority of patients (Figure 4.25) within the high-risk prognostic group developed AML (71% in contrast to only 12% low-risk individuals) it is possible that telomerase up-regulation is induced in some high-risk individuals prior to AML transformation.

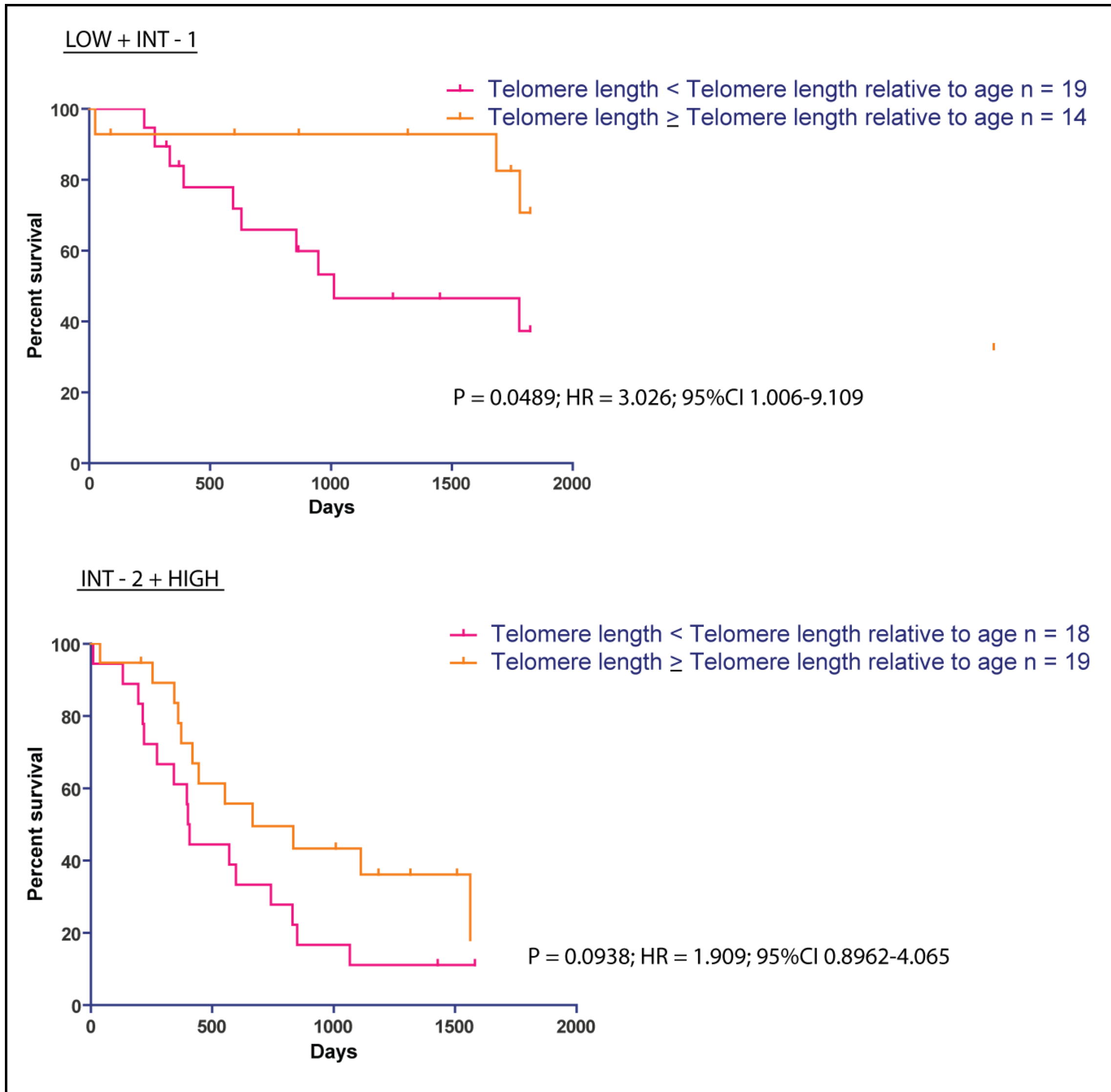


Figure 4.26: There appeared to be a tendency for reduced survival (p = 0.0489; HR = 3.026; 95%CI 1.006-9.109) within the low-risk prognostic group in patients with short telomeres for their respective age, however this was not apparent within the high-risk group (p = 0.0938; HR = 1.909; 95%CI 0.8962-4.065).

4.9 Telomere Length and Survival in MDS

Telomere length may provide an independent prognostic factor that could refine the ability to provide accurate prognostic information for individual patients. Accordingly, survival was analysed with respect to telomere length in a group of 77 MDS patients; the data was censored at 5 years follow-up as data beyond this point was not available for the majority of individuals. Recursive partitioning was used to divide patients into groups above and below different telomere length cut-off points. In doing so, the influence of telomere length on overall survival was established.

It was apparent that overall survival was influenced by telomere length. In comparison to patients with longer telomeres, those with shorter telomere profiles had a significantly worse overall survival (Figure 4.27 and Figure 4.28).

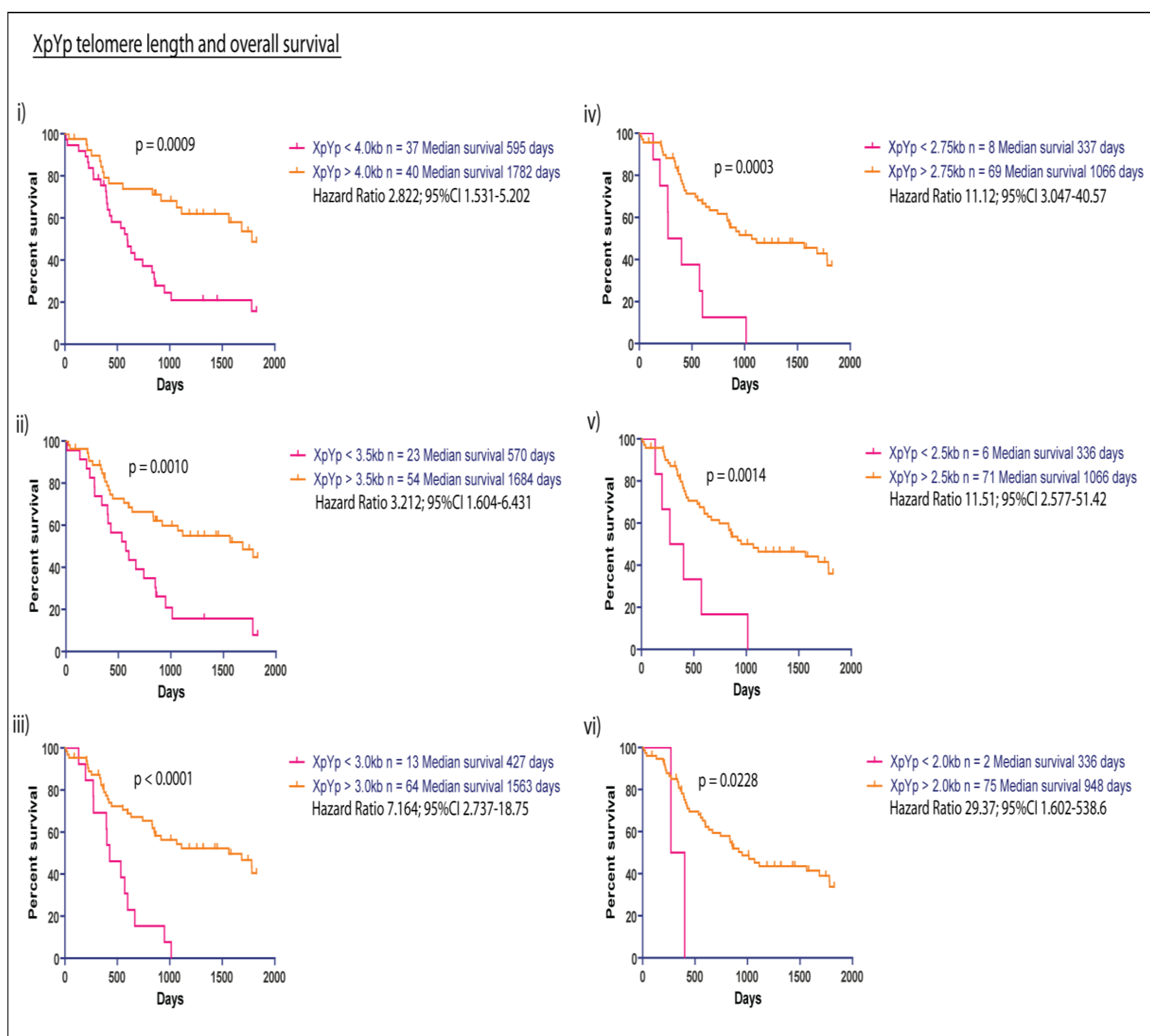


Figure 4.27: Overall survival censored at 5 years. A significant reduction ($p < 0.05$) in overall survival was apparent with elevated telomere attrition at XpYp.

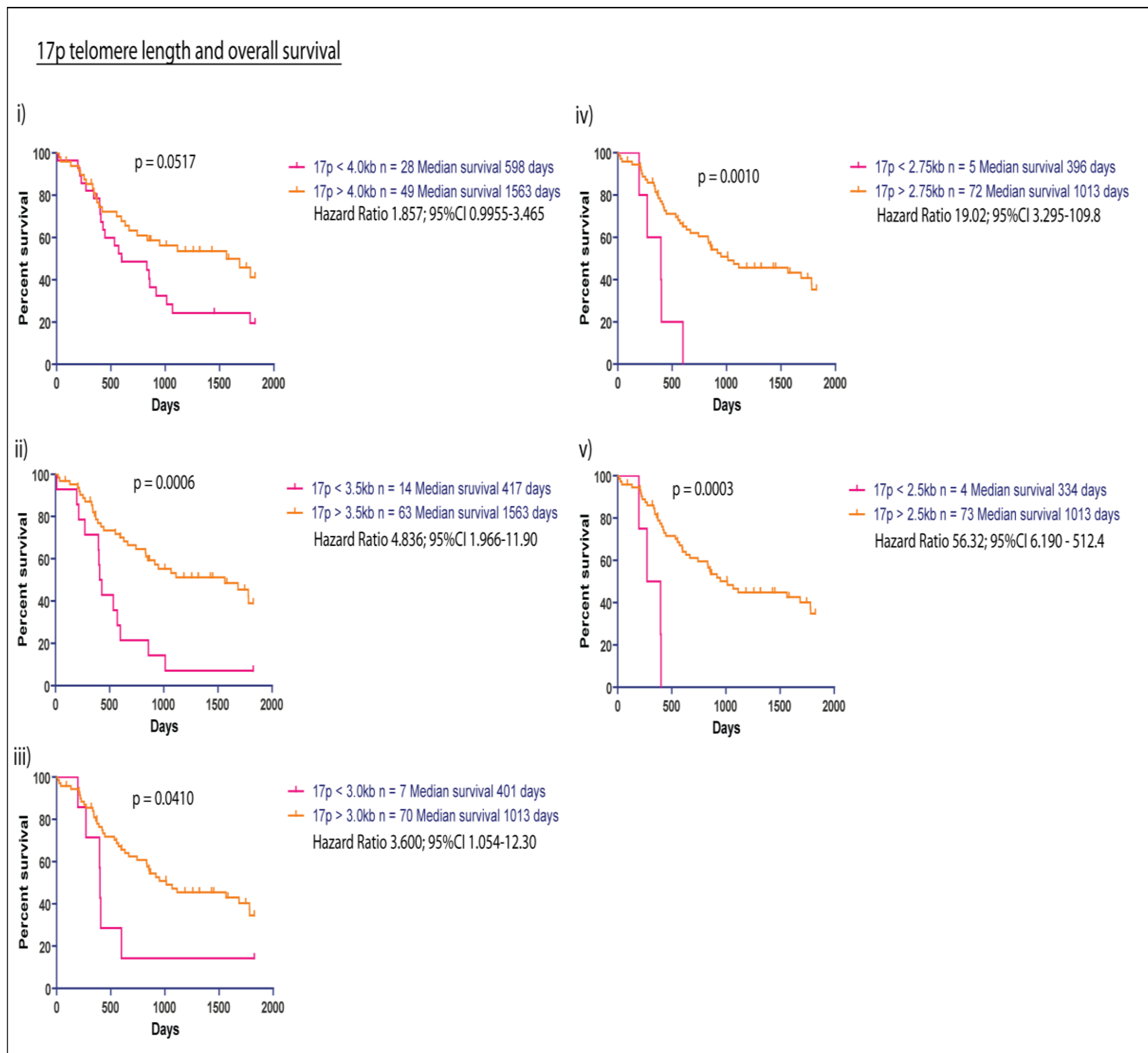


Figure 4.28: Overall survival censored at 5 years. A significant reduction ($p < 0.05$) in overall survival was apparent with elevated telomere attrition at 17p.

It was noted that the hazard ratio (HR) increased significantly with progressive telomere shortening. The HR increased to 11.51 (95%CI 2.577-51.42; p = 0.0014) and 56.32 (95%CI 6.190-512.4; p = 0.0003) at the XpYp and 17p telomere, respectively in patients that presented with a telomere length below 2.5kb at diagnosis (Figure 4.29). The variation of the hazard ratio may be related to longer TVR (Telomere Variant Repeat) regions at the 17p telomere. Extensive TVR regions would make the apparent telomere length longer raising the HR value as a consequence of a shorter functional TTAGGG tract. Furthermore, the greater mortality risk at 17p may also share some association with LOH at the *p53* locus, which has been shown to be correlated with a poor prognosis in MDS.[135,136] However, despite the apparent variation in HR between the XpYp and 17p telomere, it should be taken into consideration that only a small number of patients, i.e. n = 6 at XpYp and n = 4 at

17p presented with a telomere distribution of below 2.5kb at diagnosis. Accordingly, the 95% confidence intervals were very wide, i.e. HR 95%CI 2.577-51.42 at XpYp and HR 95%CI 6.190-512.4 at 17p (Figures 4.27 and 4.28). A larger number of patients would be necessary to determine whether telomere shortening at 17p elevates the mortality risk.

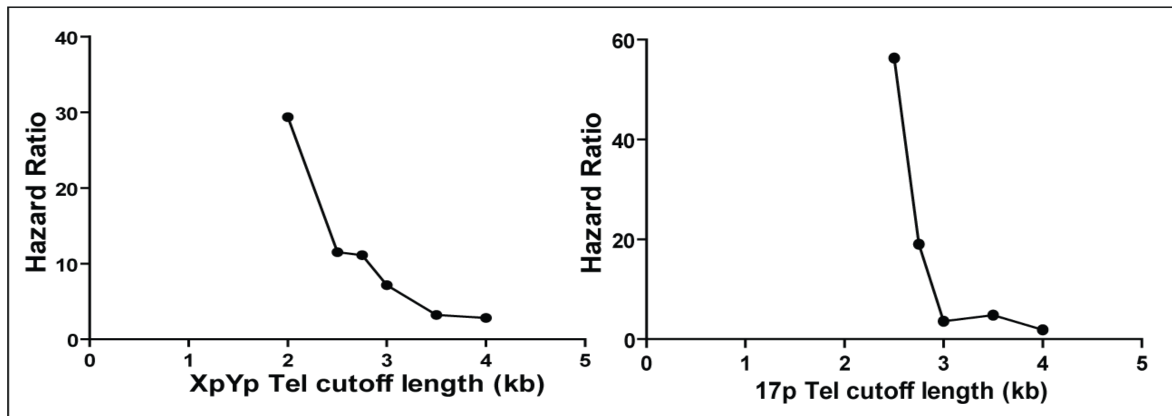


Figure 4.29: A reduction in telomere length increased the value of the hazard ratio (HR) so that the mortality rate increased by 11.51 (p = 0.0014) and 56.32 (p = 0.0003) times at the XpYp and 17p telomere, respectively within patients that had a telomere length below 2.5kb at diagnosis.

Survival analysis was taken further by analysing how the overall outlook is influenced by the actual length of the pure TTAGGG tract, i.e. excluding the TVR region from the telomere length (Table 4.1 and Figure 4.30). It was apparent that a pure TTAGGG tract of 2.0kb at the XpYp telomere was the cut-off point that defined the worst clinical outcome in the MDS cohort (HR = 7.123; 95%CI 2.365-21.45; p = 0.0005).

Table 4.1: It appeared that a pure TTAGGG tract of 2kb at the XpYp telomere was the cut-off point that defined the worst clinical outcome of MDS patients (p = 0.0005; HR = 7.123; 95%CI 2.365-21.45).

TTAGGG Cutoff (kb) [Cen: 5 years]	Patients (n) [Above/ Below]	p value	Hazard Ratio (HR)	HR 95% Confidence Interval (CI)
4.00	14/35	0.0721	2.142	0.934-4.914
3.50	17/32	0.0061	3.014	1.369-6.635
3.00	24/25	< 0.0001	5.134	2.306-11.43
2.75	26/23	< 0.0001	5.725	2.513-13.04
2.50	30/19	< 0.0001	5.945	2.463-14.35
2.00	38/11	0.0005	7.123	2.365-21.45
1.75	40/9	0.0044	5.519	1.705-17.87

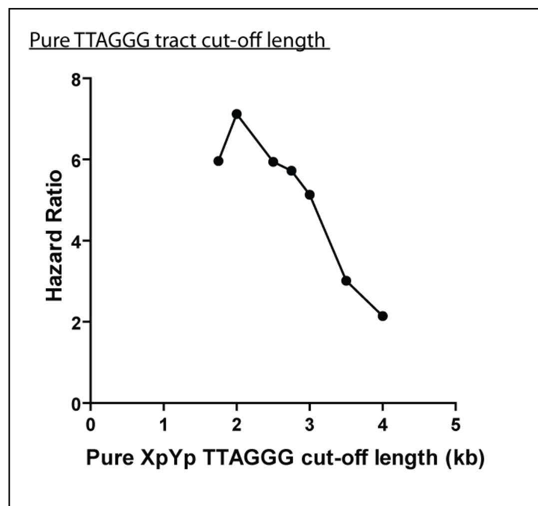


Figure 4.30: Hazard ratio values increased with pure TTAGGG shortening. The cut-off point recorded was a pure TTAGGG tract repeat of 2.0kb (p = 0.0005; HR = 7.123).

Unfortunately, TVR data could not be achieved at the 17p telomere and thus the value of the hazard ratio in patients with a short pure functional TTAGGG tract could not be determined and compared to the TTAGGG tract at the XpYp telomere.

4.10 Discussion

Quantitative and qualitative abnormalities of haematopoietic cells arise under conditions of deregulated haematological homeostasis and it has been suggested that increased cellular proliferation ensues in an attempt to compensate for haematological deficits.[11] However, leukaemia becomes more prevalent under such settings likely due to an accumulation of genetic insults.[11] Telomere dysfunction has been implicated as a mechanism that can induce genetic alterations and cancer progression by means of entering into breakage-fusion-bridge (BFB) cycles.[319] This paradigm has been observed in many human solid tumours including renal cell carcinoma, pancreatic carcinoma and osteosarcomas.[320,321] In order to maintain telomere stability and function selected cells up-regulate telomerase to facilitate the outgrowth of the malignant clone.[180,318]

The Myelodysplastic syndromes (MDSs) are comprised of a heterogeneous group of clonal disorders associated with ineffective haematopoiesis.[18] Although one third of patients develop Acute Myeloid Leukaemia (AML), mortality commonly arises through complications of bone marrow failure, i.e. anaemia, bleeding complications or the patient becoming immune-compromised.[15,24,26]

Telomere length is negatively correlated with age within healthy individuals however, telomere length heterogeneity exists within the human population.[236,240] The reported impact of heritability in humans has been noted to vary from 36 to 90% with biological and environmental factors influencing telomere maintenance and shortening.[238] Such biological factors may include components involved in telomere homeostasis (e.g. telomerase activity) and Shelterin function (e.g.TRF1/TRF2/POT1).[238] Variations in telomere length maintenance may result in the reduction of telomere length which has the potential to limit tissue renewal and increase the predisposition to organ failure, particularly those of high cellular turnover, i.e. the bone marrow.[362] Consistently, late generation $mTR^{-/-}$ or $mTR^{+/-}$ mice develop clinical manifestations similar to those observed within Dyskeratosis Congenita (DC) patients. Short telomeres were associated with a reduction in stem cell reserves and correlated with a decrease in erythrocyte, platelet and white blood cell count.[362,363]

Telomere length continued to conform to the aging dogma, albeit only weakly, in the MDS cohort examined in this study. This suggests that chronological age plays a minimal role in

influencing the telomere length of haematological cells within MDS patients. Accordingly, a 37-year-old MDS patient demonstrated extensive telomere attrition at diagnosis that was comparable to the expected telomere length of a healthy 114-year-old. This also suggests that despite being telomerase competent,[325] haematopoietic cells continue to lose telomeric repeats, particularly under conditions that exert haematological stress. Although the variability observed within the MDS cohort may reflect inheritable traits of telomere maintenance and environmental influences,[238] the rate of telomere attrition can be affected by the replicative history and damage to the stem cell pool as well as by acquired mutations that alter telomere length regulation. Moreover, telomere shortening in healthy individuals could increase the predisposition for MDS development.

Telomere length was significantly shorter in MDS when compared to a healthy marrow which is possibly associated with an increase in the proliferative pressure on haematopoietic cells in a pathological environment. There also appeared to be a decline in telomere length with elevated bone marrow blast count. This is consistent with previous observations[336] which have also identified an association between progressive telomere shortening and a decrease in apoptosis of MDS $CD34^+$ cells.[346] It is possible that functional DNA damage response (DDR) mechanisms have become abrogated with development enabling cells with short telomeres to bypass cell cycle arrest and prolong telomere attrition. There was also a tendency towards telomere length homogenisation with elevated blast count which reached significance in marrow composed of over 20% blasts. Differentiating cells in early stage MDS is probably accountable for the extensive telomere length heterogeneity, i.e. division heterogeneity of differentiated precursors in the bone marrow. The acquisition of molecular changes prevents the differentiation of haematological cells resulting in the accumulation of immature cells in the marrow. The progressive decline in telomere length within the remaining healthy cells reduces the observation of long telomeres and the telomere length distribution tends towards homogenisation with disease progression. It is also plausible to suggest that the up-regulation of telomerase in a subset of patients[335,342,347] may provide maintenance of short telomeres and reduce the extent of telomere length variation. It should be taken into account, however that this analysis was based on percentage subgroups, i.e. <5% and 5 to 20% and thus in a subsequent study, microscopic analysis of

bone marrow samples and blast quantification should be ensued so as to make a direct correlation between such parameters.

A minority of samples exhibited long telomeres within the lower 25th percentile of distributions even in the presence of high blast cell counts. Although this may be compatible with early telomerase up-regulation, extensive TVR (telomere variant repeat) regions may also be responsible for the appearance of longer telomeres. Notably, several patients showed large TVR regions of over 1.5kb and it was also apparent that numerous patients with over 20% blast cells in the bone marrow presented extensive shortening of pure TTAGGG repeats (under 1kb) within the lower 25th percentile of telomere distributions. More importantly, a telomere length distribution of less than 1kb has been previously described in fibroblasts undergoing crisis *in vitro*.[255]

MDS patients who endured multi-lineage cytopenia, i.e. 2 or 3 demonstrated a significantly elevated mortality rate when compared with those who presented with a single lineage cytopenia. Previous reports have implicated that telomeres lose repeats following an allogeneic bone marrow transplant prior to haematopoietic reconstitution by which, the extent of telomere shortening is inversely correlated with the number of cells received by the recipient.[327] With this information at hand it was originally speculated that increased $CD34^+$ cellular proliferation (and increased telomere attrition) may coincide with elevated haematological pressure in an attempt to regain reconstitution. Additionally, telomere shortening has the propensity to exacerbate cytopenia severity, particularly when a DDR response prevails. Accordingly, telomere induced apoptosis or senescence[314] has the potential to abolish differentiating cells and reduce peripheral blood cell numbers.[314] However, in contrast with previous reports,[332] telomere length was not significantly different within patients that endured a single or multi-lineage cytopenia. This was also reiterated when measuring the lower 25th percentile of the telomere distributions. Despite there being a tendency of pure TTAGGG repeat shortening within the lower end of the distributions (25th percentile), statistical significance was not reached with increasing cytopenia severity. This analysis was based on the number and not the actual depth of cytopenia and therefore a direct correlation between telomere length and cytopenia severity was not made. In a subsequent study it will be more informative to establish

whether telomere length was associated with the intensity of haematopoietic insufficiency along a specific lineage.

The association between telomere length and cytopenia severity was analysed further by stratifying MDS patients into groups that were dependable on the telomere length for age in healthy individuals. Kaplan Meier curves illustrated that overall survival was significantly increased in patients who presented with normal or longer telomeres for age at diagnosis than those with shorter telomeric lengths. However, telomeric separation only had a prognostic influence within patients that endured a single lineage cytopenia ($p = 0.0144$; HR = 7.457; 95%CI 1.492-37.26). Patients who presented with short telomeres might have had a more severe degree of haematopoietic insufficiency along a single lineage and it may be speculated that this subset of patients may be at risk of developing multi-lineage cytopenia. Accordingly, elevated telomere shortening can exacerbate the problem by promoting premature cell cycle arrest and induce a further decline in haematopoietic cell reserves.

Consistent with previous studies, overall survival deteriorated within patients who presented with a poor cytogenetic profile at diagnosis.[105] Telomere dysfunction is a mechanism that can generate genetic instability by entering into breakage-fusion-bridge (BFB) cycles.[319] Complex karyotypes involving genetic alterations of three or more chromosomes are characteristic of a poor cytogenetic profile,[105] thus it was proposed that telomere length may appear shorter within this cohort. Although telomere shortening has been previously associated with complex chromosomal rearrangements in many haematological diseases, including MDS and AML[333,334,342] the current data failed to show increased telomere attrition in patients presenting with a poor cytogenetic profile.

Conventional cytogenetic G-banding fails to detect karyotypic alterations in a substantial number of MDS patients (~50%) as a result of its low resolution.[102] Consequently, this group of individuals are considered to present with a normal karyotype. Separating telomere length into groups with respect to age revealed that patients could be delineated into those with an unfavourable and favourable prognosis. Accordingly, telomere length greatly influenced the overall survival of patients who presented with a good karyotype ($p < 0.0001$; HR = 27.26; 95%CI 5.538-134.2). Cryptic chromosomal lesions have been detected using high-resolution single-nucleotide polymorphism analysis (SNP-A) in cytogenetically normal

MDS[102,130,364] and AML.[365,366] Such alterations include terminal uniparental disomy (UPD), deletions and amplifications of which patients with copy number changes, particularly deletions, have a significantly poorer prognosis.[102,364] Telomere dysfunction may initiate early events for neoplastic development and predispose to the evolution of pathogenetic chromosome aberrations that are detected by conventional cytogenetics. Moreover, due to the heterogeneity of patient outcome within the good cytogenetic cohort,[103] telomere length may provide a prognostic marker that has the potential to delineate subgroups with an unfavourable prognosis. This may be particularly important in classifying patients early that have the potential to rapidly succumb to their disease. In contrast, overall survival was not influenced by telomere length within patients that presented with intermediate or poor cytogenetic profiles. It is possible that telomerase may have been up-regulated in order to stabilise the novel genome, particularly in cells which present with poor cytogenetics.

Patients were separated into prognostic cohorts based on their IPSS (International Prognostic Scoring System) risk score. The risk score stratifies individuals into low, intermediate-1 (Int-1), intermediate-2 (Int-2) or high by combining independent variables including blast percentage, cytopenia(s) and cytogenetics in order to predict patient outcome and facilitate in making treatment decisions.[43] Telomere length was not significantly different within any of the prognostic cohorts, however because the risk of disease development is based on categorical features and not on the actual depth of cytopenia severity, presenting blast count or specific karyotypic abnormality it may be speculated that variations of these features result in fluctuations in telomere length and contribute to the absence of an association between telomere length and IPSS score. It is also possible that telomerase is up-regulated in numerous patients removing the prognostic signature of telomere length, particularly within higher risk groups. This may also explain, in part the heterogeneity of telomere length that was detected amongst the higher risk scores.

In spite of this, the current data suggest that the telomere length at diagnosis may be able to stratify low-risk (Low/Int-1) patients into those with a favourable or unfavourable outcome. Patients presenting with short telomeres relative to age at diagnosis had a reduced overall survival rate ($p = 0.0489$; HR = 3.026; 95%CI 1.006-9.109). It was also apparent that a minority of these individuals progressed to AML. In contrast, the telomere length was fairly heterogeneous with respect to AML progression in patients presenting

with high-risk IPSS (Int-2/High). Additionally, telomere length did not appear to influence the mortality rate within this group. The majority of individuals within the high-risk prognostic group developed AML (71% high-risk to 12% low-risk) therefore it is possible that telomerase was up-regulated in some high-risk patients prior to AML transformation in order to provide telomere stability.

This study revealed that telomere length at diagnosis significantly influenced the overall survival of MDS patients irrespective of conventional markers. Accordingly, patients who presented with shorter telomere length showed a reduction in overall survival. The telomere length at diagnosis may be identified as a prognostic variable that is able to classify patients who have the potential to rapidly deteriorate. Notably, low-risk MDS cases are heterogeneous[103] with respect to outcome and a fraction of these patients show poor prognoses similar to high-risk individuals.[104]

Although these observations suggest that telomere length at diagnosis may have potential in refining patient outlook and facilitate in making therapeutic decisions it is highlighted that individual therapy was unknown for the MDS patients in this study and therefore these data are inconclusive. Notably, overall survival can be greatly influenced depending on individual therapy. Further analysis on a more robust cohort of patients who are undergoing uniform treatment would substantially improve these data and potentially show an association between diagnostic telomere length and prognosis in MDS patients.

Chapter 5:

Telomere Length and Prognosis in AML

5.1 Abstract

Acute Myeloid Leukaemia (AML) is an aggressive myeloid neoplasm characterised by the clonal proliferation of undifferentiated myeloid precursor cells and represents a group of heterogeneous conditions with a diversity of clinical and biological features.

Regression analysis revealed a significant correlation between the telomere length at XpYp and 17p ($r^2 = 0.6036$; $p < 0.0001$). Telomere length was significantly reduced in patients with AML compared to healthy aged-matched individuals ($p < 0.0001$). However, a weak positive correlation was detected between telomere length and age at diagnosis suggesting that telomere length in AML fails to conform to the aging dogma. Categorically, patients ≤60 years of age showed significantly shorter telomere length when compared to patients older than 60 years of age ($p = 0.037$).

With respect to other clinical parameters telomere length was not correlated with gender, presenting white blood cell (WBC) count, blast percentage at diagnosis and WHO performance status. In contrast, significantly shorter telomeres were identified in cases with *de novo* AML ($p = 0.0356$) when compared to secondary AML cases. There was no significant difference between telomere length and cytogenetic subgroups; however this data set was limited. There was a tendency for telomere shortening in patients that had the FLT3/ITD mutation ($p = 0.0724$) possibly associated with an increase in mitotic history, whilst significantly longer telomeres were observed in patients positive for the FLT3/TKD ($p = 0.0344$). Patients with the TKD mutation were significantly older when compared to patients without the mutation ($p = 0.0018$). Finally, the log-rank test failed to show any prognostic impact of telomere length on the number of disease-free days and overall survival at both the XpYp and 17p telomere.

5.2 Introduction

Acute Myeloid Leukaemia (AML) is an aggressive myeloid neoplasm characterised by the clonal proliferation of undifferentiated myeloid precursor cells and represents a group of heterogeneous conditions with a diversity of clinical and biological features.[367] According to the World Health Organisation (WHO) AML is predicted when a count of 20% or more myeloblasts are present in the bone marrow or peripheral blood.[20] AML can evolve *de novo* or secondary which is defined as having an anecdotal haematopoietic disorder or following prior chemotherapy or radiotherapy for a non-haematological disorder.[53]

A patient's risk score has been devised that takes into account specific prognostic parameters including age, presenting white blood cell (WBC) count, cytogenetics, AML type (*de novo*/ secondary) and response status after the first cycle of induction chemotherapy (Complete Remission [CR]/Partial Remission [PR]/Resistant Disease [RD]). The calculated score can be used to categorise patients into good, standard or high-risk subgroups and thus facilitate in making therapeutic decisions, e.g. consolidation chemotherapy for good-risk or bone marrow transplant (BMT) for high-risk.

Clinical management and decision making in AML relies strongly on risk stratification based on conventional karyotyping.[368] Cytogenetic risk has been divided into three subcategories which include favourable, intermediate and adverse. Favourable abnormalities are composed of balanced translocations such as *AML1-ETO* t(8; 21), *PML-RAR*α t(15;17) and *CBF*β*-MYH11* inv(16)/t(16;16) whereas unbalanced translocations and complex karyotypes (3 or more clonal abnormalities) are adverse cytogenetic characteristics.[369] Patients presenting with favourable cytogenetics generally have good outcomes with conventional chemotherapy; whereas patients with an unfavourable karyotype have a very poor prognosis with conventional chemotherapy and are therefore considered for an allogeneic transplant.[370] 10% to 20% of AML cases present with a complex aberrant karyotype which has been identified as the worst group prognostically.[371]

Balanced translocations are more commonly detected in younger AML patients whereas unbalanced and complex abnormalities are more frequent in older patients.[57] A 24.6 fold increase in the incidence of complex karyotypes was observed in patients of age 61 to 70 years in contrast to patients of age 21 to 30 years while only a 1.7 fold increase in balanced

translocations was detected with age.[371] The loss of 5q, 7q or a deletion of 17p is regularly observed in patients with unbalanced karyotypes. Moreover, while *p53* alterations are rare in other AML subtypes; they have been identified in more than 90% of cases with a complex karyotype.[371]

Secondary AML (sAML) can arise from a preceding chronic phase such as a Myelodysplastic Syndrome (MDS) or Myeloproliferative disease (MPD).[53] Unlike *de novo* AML patients who often achieve CR after treatment, patients who transform to sAML have a very poor prognosis and die within a few months following AML diagnosis.[372] 44.2% of sAML cases present with complex aberrations in contrast to only 11.4% of *de novo* AML patients. However genetic features typical of sAML with complex karyotypes have also been identified in *de novo* AML with complex aberrations. Such abnormalities include del7q, del5q or loss of *p53*.[372] Thus, it has been proposed that *de novo* AML presenting with a complex karyotype may in fact be sAML evolved from a previous undiagnosed MDS/MPD.[372,373]

AML with intermediate risk cytogenetics encompass a heterogeneous population of patients of which most are cytogenetically normal.[60] Molecular markers have been used to further refine patient prognosis within this cytogenetic subgroup. Such molecular abnormalities include mutations of the FMS-like receptor tyrosine kinase-3 (FLT3) receptor[374] and Nucleophosmin (NPM1).[70,71]

FLT3 is expressed by cells found in the haematopoietic stem cell compartment and early committed progenitors.[59,60] The stimulation of FLT3 by its ligand has been proposed to play a role in cell proliferation.[59,60] Approximately 30% of AML patients present with mutations in the FLT3 receptor leading to its autophosphorylation and constitutive activation. 25% of adult AML patients present with an internal tandem duplication (FLT3/ITD) of the juxtamembrane domain whereas 7% carry point mutations within the tyrosine kinase domain (FLT3/TKD).[62,63] The prognosis of patients harbouring the FLT3/ITD is poor with individuals exhibiting high relapse rates and an inferior overall survival (OS).[64,65] Disease-free survival (DFS) is significantly shorter in patients bearing the FLT3/TKD than FLT3 wild-type cases and tend to have a worse OS and high relapse rates.[66,67] However, other studies have

observed no influence of the TKD mutation on OS or DFS[60,68] and thus its prognostic implication continues to be controversial.

The NPM1 protein functions as a molecular chaperone that shuttles between the nucleus and cytoplasm.[69] It is predominately nucleolar but 30% of AML cases bear the cytoplasmic NPMc^{+}.[70,71]NPM1 is composed of an NES (nucleus export signal) and NuLS (nucleolar localisation signal) sequence motif that reside at its C-terminal. In AML, the NuLS is substituted into an extra NES generating increased nuclear export and cytoplasmic accumulation of the protein.[71] Mutations in NPM1 are associated with a more favourable outcome when compared to patients presenting with wild-type NPM1, however patients with the NPM1 mutation also present with FLT3/ITD more frequently than patients with wild-type NPM1.[72] Therefore the favourable outcome is no longer applicable to patients presenting with the dual NPM1 and FLT3/ITD mutation.[73,74]

Extensive telomere shortening has been observed among AML patients when compared to aged matched healthy controls.[347,350] Patients with multiple cytogenetic abnormalities have been associated with shorter telomeres when compared to those presenting with reciprocal translocations or a normal karyotype.[342,347] Moreover, Q-FISH analysis revealed extensive telomere shortening on individual chromosome arms in AML patients who presented with gains and or/losses.[342] This raises speculation that critically short telomeres in these cells may have a role in generating chromosomal instability. Additionally, the extent of chromosomal abnormalities was correlated with *hTERT* expression[347] and patients who presented with gains/losses of chromosomes showed elevated telomerase activity when compared to patients with normal or balanced karyotypes.[342]

FLT3/ITD has been associated with shorter telomeric length which is likely associated with an extended proliferative history of AML cells[347,350] but telomere length has failed to show an influence on overall survival, disease-free survival, the incidence of relapse or response to treatment.[347,350]

The importance of telomere length with respect to clinical parameters of AML has been investigated in this chapter. Furthermore, telomere length was analysed to determine whether it influenced a chemotherapeutic response, overall survival and disease-free survival of AML patients following a cycle of intensive chemotherapy.

Results

5.3 Telomere Length and Age at Diagnosis

STELA analysis at both the XpYp and 17p telomere was available for 57 AML patients. Regression analysis revealed that the telomere length at XpYp and 17p were strongly correlated (Figure 5.1) with $r^2 = 0.6036$; $p < 0.0001$. However, it was clear from the intersection of the axis that the 17p telomere was 1.37kb longer than XpYp.

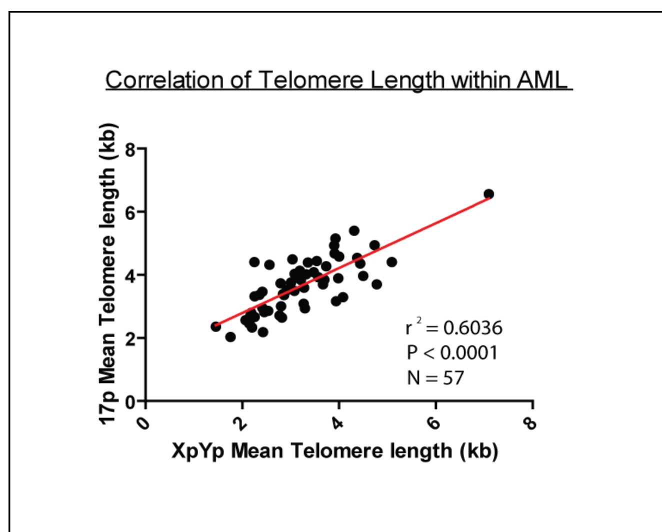


Figure 5.1: Telomere length at XpYp and 17p are strongly correlated $r^2 = 0.6036$; $p < 0.0001$.

Telomere length was significantly reduced in patients with AML compared to healthy aged-matched individuals ($p < 0.0001$). The telomere length at the XpYp telomere was analysed in a cohort of 110 AML patients (95 primary and 15 secondary) with a median age of 60 years (range 17 to 82 years) however, because the clinical information was only available for a subset of these individuals at 17p, telomere length was examined in only 23 of these individuals with a median age of 50 years (range 29 to 67 years). It was apparent that the telomere length at the XpYp and 17p telomere failed to conform to the aging dogma detected in healthy individuals, i.e. increased telomere attrition with age.[240]

Figure 5.2 illustrates the tendency towards a weak positive correlation with telomere length at XpYp and age at diagnosis ($r^2 = 0.03017$; $p = 0.0696$). It appeared that this was also the case at the 17p telomere ($r^2 = 0.09874$; $p = 0.1442$) however, the limited number of patients is likely associated with a reduction in significance.

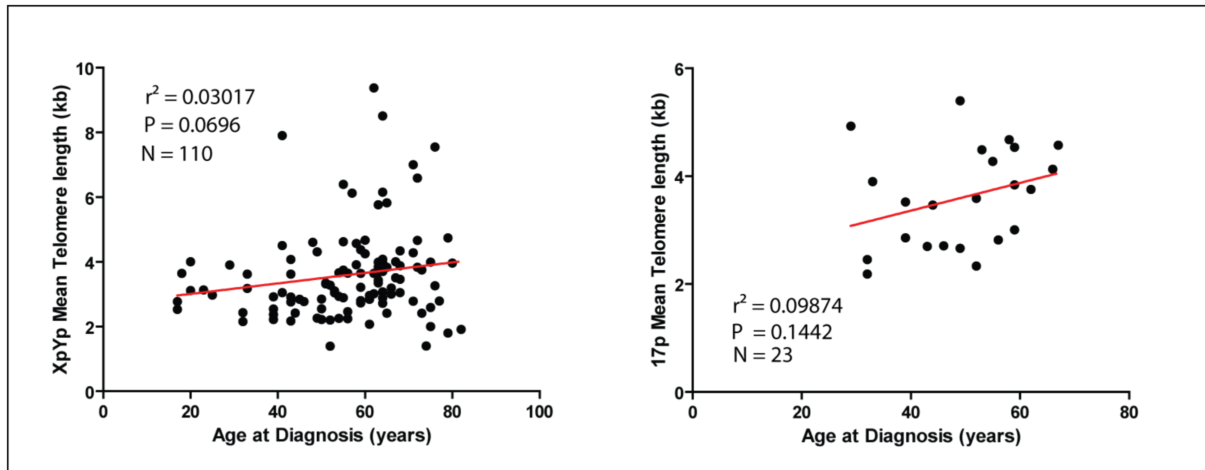


Figure 5.2: Telomere length at XpYp (r^2 = 0.03017; p = 0.0696) and 17p (r^2 = 0.09874; p = 0.1442) show a tendency towards a weak positive correlation with patient age at diagnosis.

Patients were further categorised into subgroups divided by the median age within the XpYp (60 years) and 17p (50 years) cohorts (Figure 5.3). Younger patients showed significantly shorter telomere length at XpYp when compared to patients older than 60 years of age (p = 0.037). The mean telomere length recorded in patients ≤60 years of age and those older than 60 years was 3.33kb (±1.12kb) and 3.90kb (±1.66kb), respectively. In contrast, the difference in telomere length did not reach statistical significance (p = 0.2075) at the 17p telomere; however there was again a trend for short telomeres within the younger cohort. Accordingly, the mean telomere length in patients ≤50 years of age was 3.34kb (±1.04kb) and 3.84kb (±0.77kb) in patients older than 50. It is conceivable that the lack of statistical significance between age and telomere length at the 17p chromosome was likely due to the limited number of patients analysed.

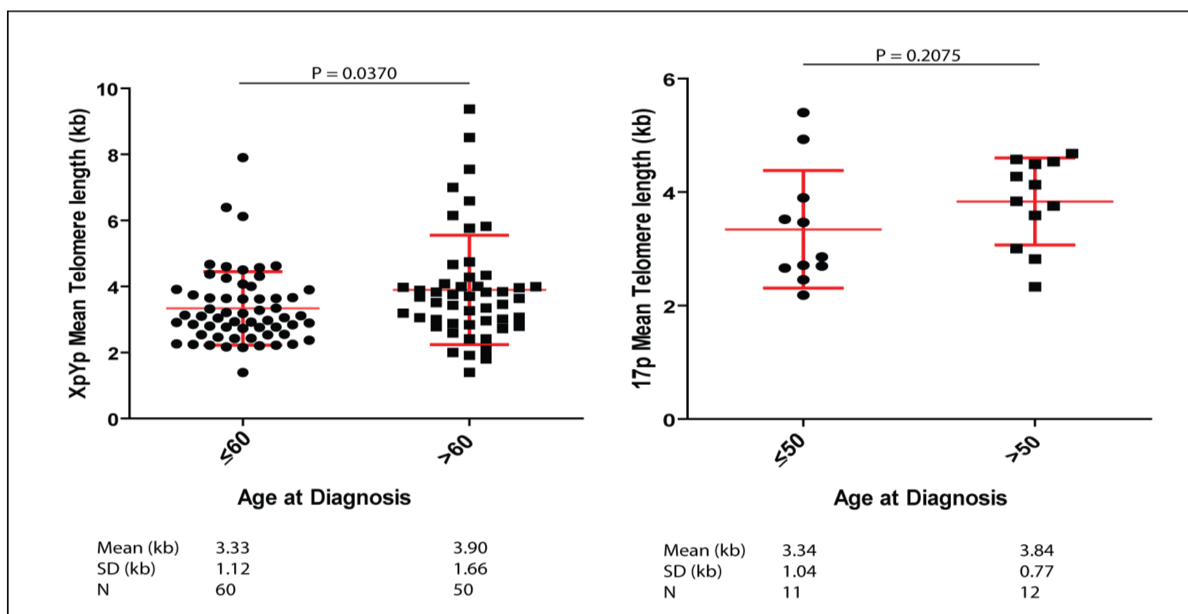


Figure 5.3: Telomere length appeared to be shorter in younger patients at the XpYp and 17p telomere. This reached statistical significance at the XpYp telomere (p = 0.0370).

5.4 Telomere Length and Gender

Of the 110 AML patients analysed 52% were male (median age 62 years; range 17 to 79 years) and 48% were female (median age 57 years; range 18 to 82 years). The difference in age did not reach statistical significance (p = 0.3796). There was no significant difference between the two gender subgroups at the XpYp (p = 0.8513) or 17p telomere (0.4771). Accordingly, the telomere length at XpYp was 3.61kb (±1.46kb) and 3.56kb (±1.37kb) within the male and female cohort, respectively. Conversely, the telomere length at 17p was 3.42kb (±0.93kb) within the male cohort and 3.71kb (±0.93kb) within the female cohort (Figure 5.4).

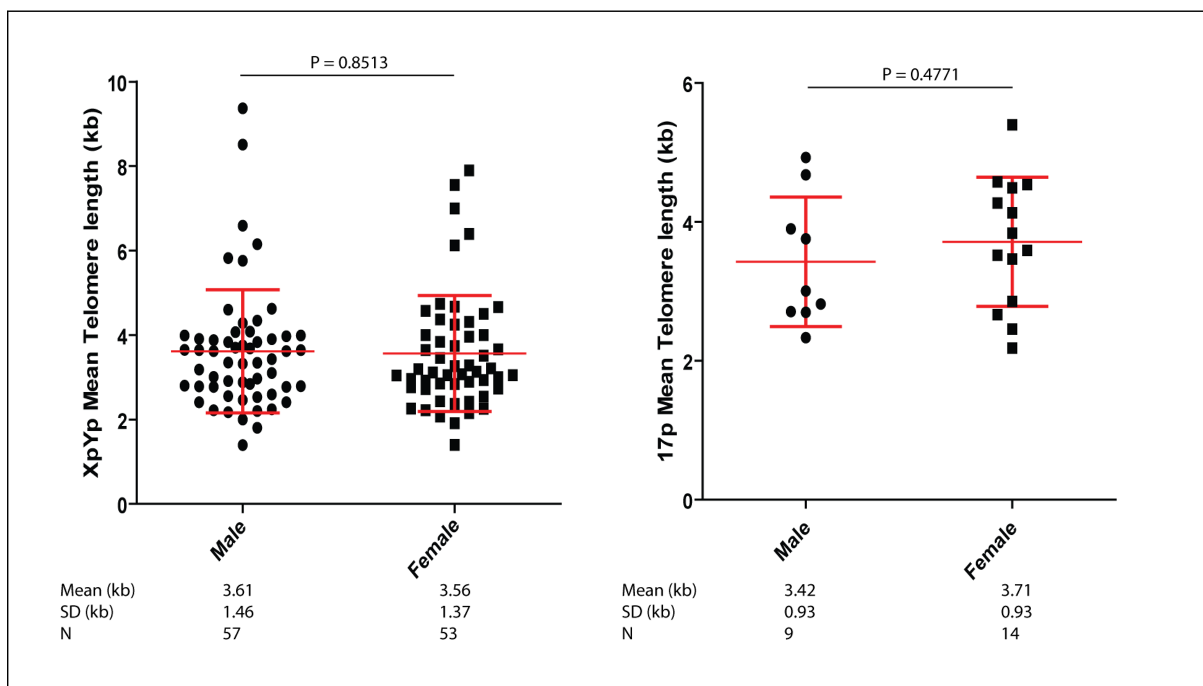


Figure 5.4: There was no significant difference in telomere length at XpYp and 17p between male and female patients.

5.5 Telomere Length and Marrow Blast Count

Of the 110 AML patients, the bone marrow blast count was available for 80 of these individuals at XpYp whereas the blast count was available for 21 of the 23 patients at 17p. Figure 5.5 shows no correlation between telomere length and bone marrow blasts at XpYp (r^2 = 0.01385; p = 0.2985) and 17p (r^2 = 0.04071; p = 0.3804).

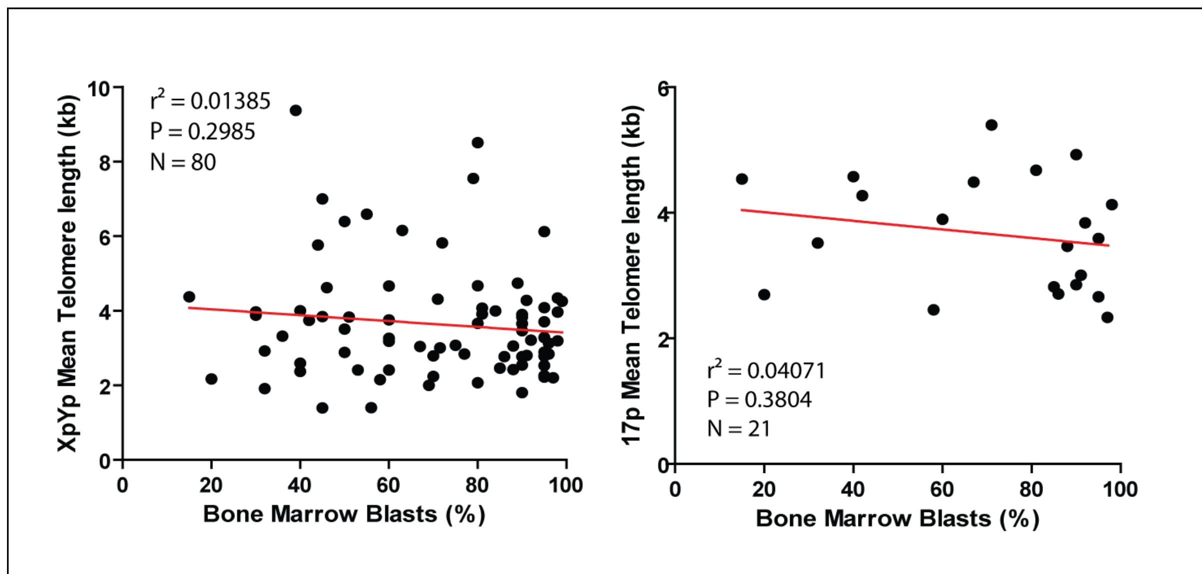


Figure 5.5: Telomere length at XpYp (r^2 = 0.01385; p = 0.2985) and 17p (r^2 = 0.04071; p = 0.3804) show no correlation with bone marrow blasts (%).

Patients were further categorised into quartiles based on the bone marrow blast percentage (Figure 5.6). The mean telomere length at XpYp was 4.00kb (±1.91kb) within the 1st quartile (bone marrow blasts ≤50.25%) and 3.31kb (±0.98kb) within the 4th quartile (bone marrow blasts >90.75%). However, this did not reach statistical significance (p = 0.1618). Due to the limited number of patients measured at the 17p telomere (n = 21) the median (85%) blast percentage was used as a cut-off point. Again, telomere length was not significantly (p = 0.1427) shorter in patients that presented with a higher blast count. Notably, the mean telomere length at 17p was 3.94kb (±0.95kb) and 3.35kb (±0.80kb) in patients presenting with ≤85% blasts in the bone marrow and those with >85% blasts, respectively.

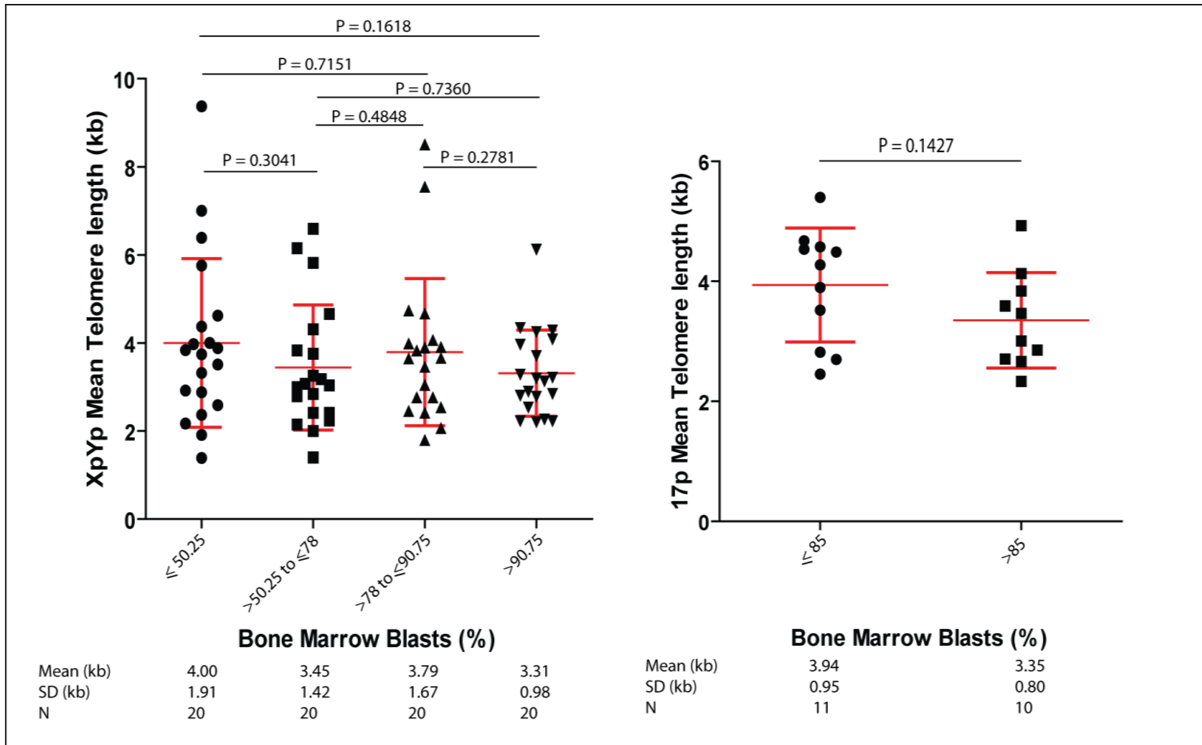


Figure 5.6: A decline in telomere length was not detected with increasing blast count. Telomere length was not significantly different between subgroups.

5.6 Telomere Length and Presenting White Blood Cell (WBC) Count

In the context of white blood cell (WBC) count, telomere length was not correlated with WBC count at presentation. Figure 5.7 shows no correlation between telomere length and presenting white blood cell count at XpYp (r^2 = 0.004645; p = 0.4793) and 17p (r^2 = 0.02809; p = 0.4447).

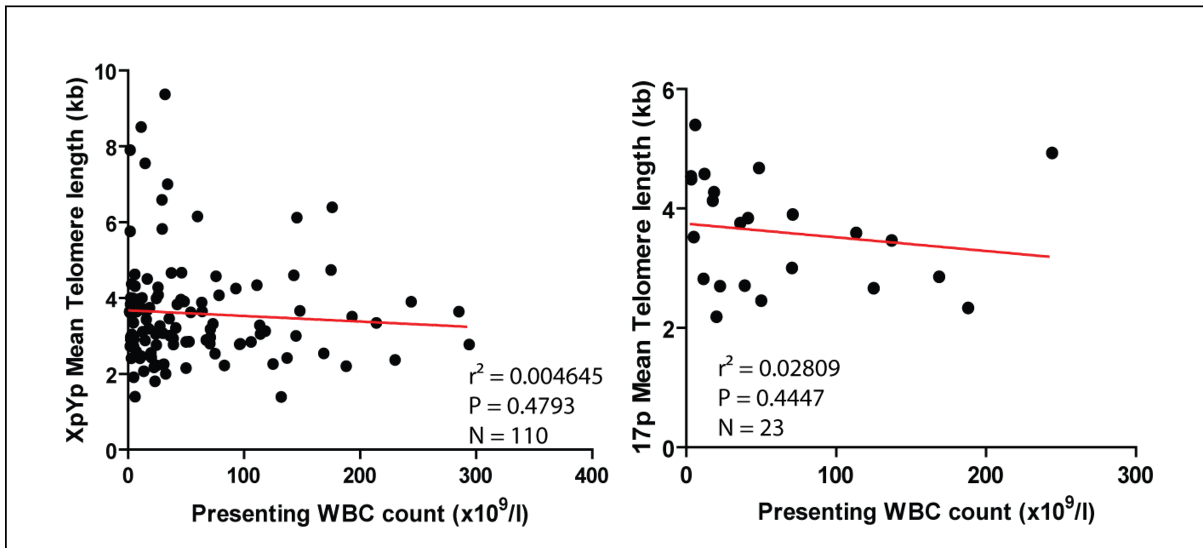


Figure 5.7: Telomere length at XpYp (r^2 = 0.004645; p = 0.4793) and 17p (r^2 = 0.02809; p = 0.4447) show no correlation with presenting WBC.

Patients were further categorised into quartiles based on the white blood cell (WBC) count at presentation (Figure 5.8). Individuals categorised within the 1st quartile presenting with a WBC count of ≤11.45 x10^9/l had a telomere length of 3.75kb (±1.57kb) whereas patients categorised within the 4th quartile had a telomere length of 3.39kb (±1.16kb) presenting with a WBC count of >76.53 x10^9/l. However this did not reach significance (p = 0.3526). Due to the limited number of patients measured at the 17p telomere (n = 23) the median (39.1 x10^9/l) WBC at presentation was used as a cut-off point. Similarly, the telomere length was not significantly different (p = 0.4040) between patients presenting with a WBC count of 39.1 x10^9/l or less and over 39.1 x10^9/l at 17p. Accordingly, the telomere length was 3.76kb (±0.98kb) and 3.43kb (±0.86kb), respectively.

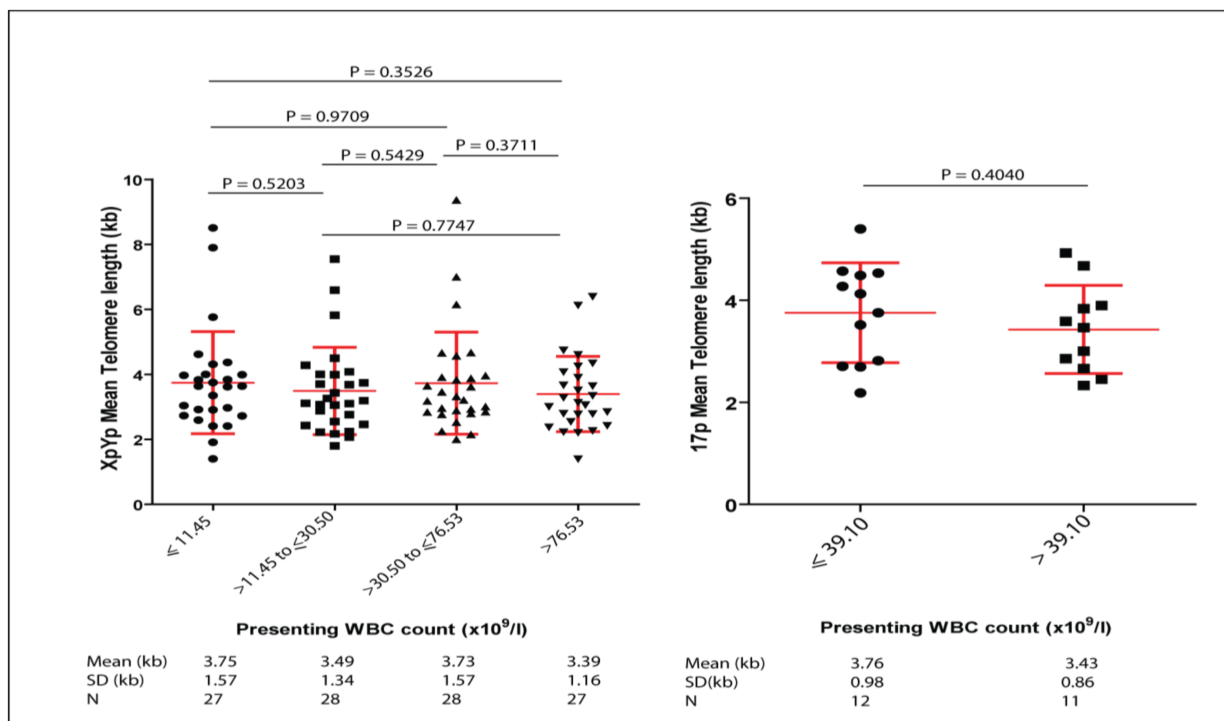


Figure 5.8: A decline in telomere length was not detected with increasing WBC count at presentation. Telomere length was not significantly different between the subgroups.

5.7 Telomere Length and AML Type (*De novo*/Secondary)

Patients with *de novo* AML were significantly (p = 0.0021) younger (median 56 years; range 17 to 80 years) when compared to secondary AML cases (median 68 years; range 41 to 82 years). The telomere length at XpYp was analysed within patients presenting with *de novo* AML or secondary AML (Figure 5.9). *De novo* AML cases had significantly shorter telomeres (p = 0.0356) when compared to secondary AML cases. The mean telomere length was 3.48kb (±1.17kb) and 4.30kb (±2.37kb) within patients presenting with *de novo* or secondary AML, respectively. The standard deviation (SD) was analysed at the XpYp telomere to determine whether this was related to intra-clonal variation of the disease subsets. The mean SD of the *de novo* cohort was 1.59kb (±0.60kb) and 1.70kb (±0.89kb) within secondary cases, however this was not significantly different (p = 0.5583). Unfortunately, all patients analysed at 17p (n = 23) had been diagnosed with *de novo* AML and therefore a comparison between AML type could not be performed at this telomere.

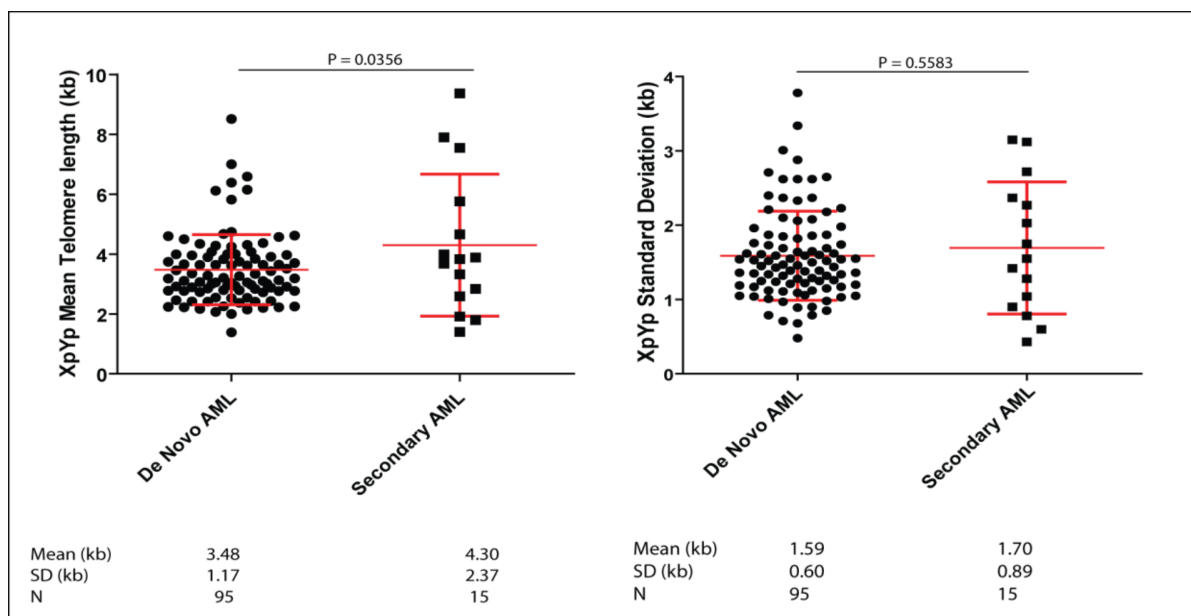


Figure 5.9: Telomere length was significantly shorter in patients presenting with *de novo* AML (p = 0.0356). Further analysis revealed that the standard deviation was not significantly different (p = 0.5583) suggesting that intra-clonal variation has no role in this finding.

5.8 Telomere Length and WHO Performance Status

Most of the AML patients analysed in this cohort did not present with an inferior WHO performance status at diagnosis; 55.5% of individuals had a performance status (PS) of 0 (fully active) and 30.1% with a PS of 1 (ambulatory). In this study, telomere length was not associated (p = 0.7159) with WHO performance status (Figure 5.10) such that the mean telomere length recorded at XpYp was 3.47kb (±1.10kb) in patients with a PS of 0 and 3.63kb (±1.36kb) in patients with a PS of 3 (in bed >50% of the time), respectively. There appeared to be a non-significant decline in telomere length with increasing PS at the 17p telomere such that a PS of 0 had a telomere length of 3.83kb (±1.02kb) and a PS of 1 and 2 had a telomere length of 3.36kb (±0.70kb) and 2.98kb (±0.43kb), respectively. However, this is likely associated with the limited data set at this telomere.

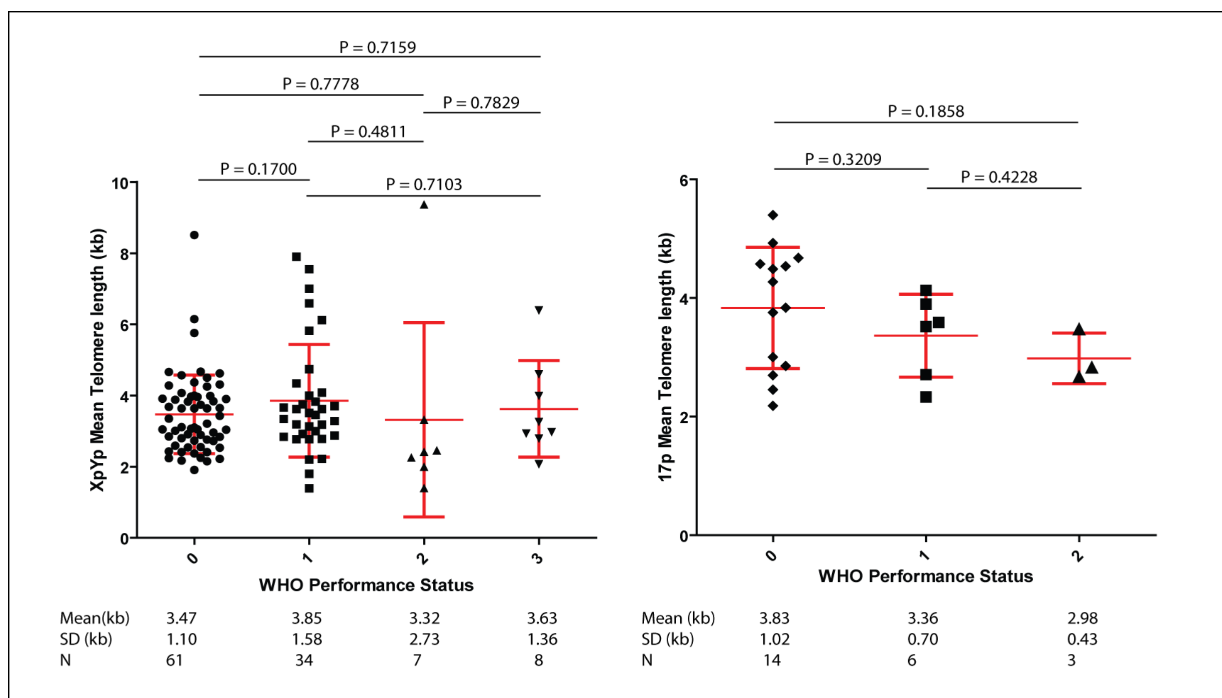


Figure 5.10: Telomere length was not associated with increasing WHO performance status (PS) at XpYp (p = 0.7159). Although there appeared to be a trend of increasing telomere attrition at the 17p telomere, the limited number of patients is likely accountable for this finding.

5.9 Telomere Length and Cytogenetic Risk Group in AML

The cytogenetic profile derived from 94 AML patients (84 *de novo* and 10 secondary AML) was available for comparing telomere length at XpYp within each prognostic subgroup (Figure 5.11). Although not significant, telomere length at XpYp was longer within the adverse subgroup at 4.65kb (±2.25kb) when compared to intermediate with a mean of 3.52kb (±1.41kb); p = 0.1324 and favourable with a mean of 3.86kb (±1.03kb); p = 0.4360. However, a conclusion cannot be made due to the limited number of patients presenting with an adverse (n = 4) and favourable (n = 7) karyotype.

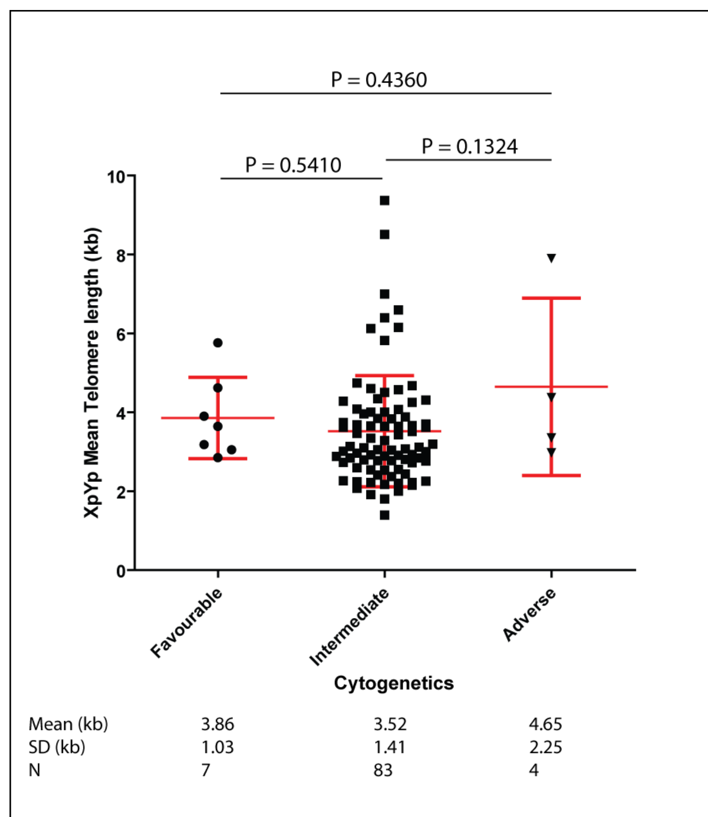


Figure 5.11: Telomere length at XpYp was not associated with cytogenetic risk group. No significant difference was detected between either subgroup.

5.10 Telomere Length and FLT3/ITD Mutation Status

Telomere length was analysed at XpYp and 17p to determine whether the FLT3/ITD mutation was associated with telomere length (Figure 5.12). Patients presenting with the FLT3/ITD mutation were younger than individuals who did not present with the mutation, however this did not reach significance (p = 0.7989). The median age of patients was 55.5 years (range 20 to 82 years) and 62 years (range 17 to 80 years) in patients presenting with the ITD mutation and those who did not, respectively. Telomere length appeared to be shorter in cases with the FLT3/ITD however, this did not reach statistical significance at both XpYp (p = 0.0724) and 17p (p = 0.2410). The mean value at XpYp was 3.27kb (±1.31kb) in patients who presented with the ITD mutation whereas it was 3.77kb (±1.42kb) within patients that did not. Similarly, the mean value at the 17p telomere was 3.14kb (±0.83kb) and 3.68kb (±0.94kb) within patients that were positive and negative for the FLT3/ITD, respectively.

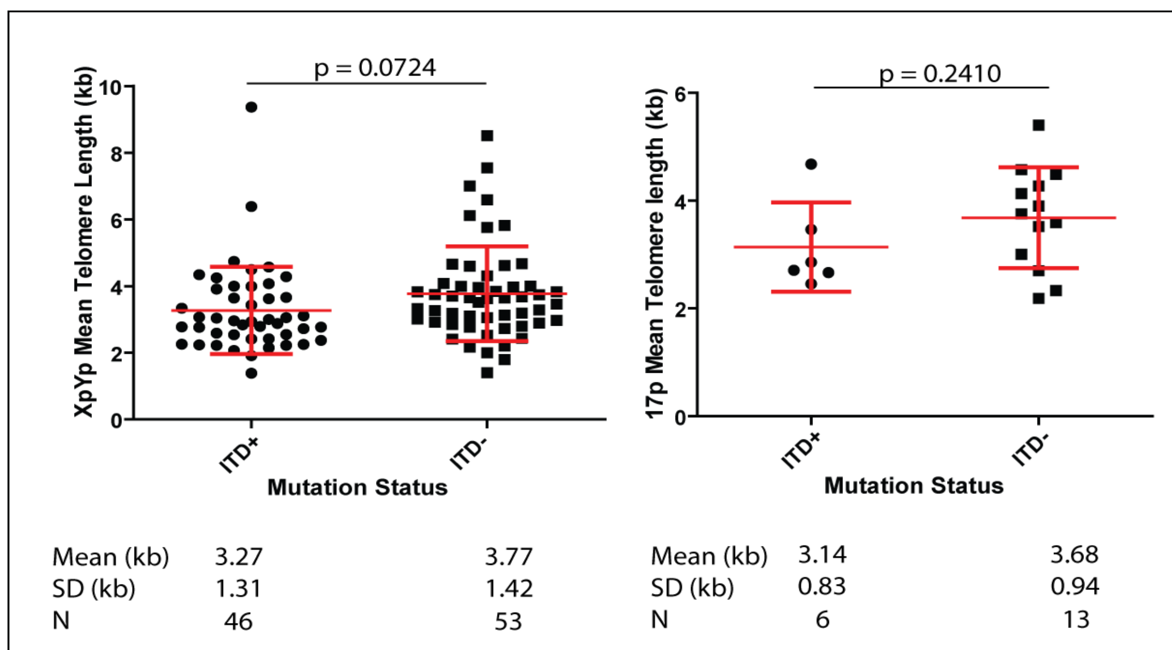


Figure 5.12: Although not significantly different, there appeared to be a tendency for telomere shortening within patients that presented with the FLT3/ITD mutation.

5.11 Telomere Length and FLT3/TKD Mutation Status

Telomere length was also analysed within patients that presented with the FLT3/TKD mutation (Figure 5.13). The telomere length at XpYp was 4.20kb (±1.80kb) within TKD positive patients and 3.39kb (±1.34kb) within those who did not present with the mutation. The difference in telomere length reached statistical significance (p = 0.0344). Regression analysis revealed a weak positive correlation between age and telomere length within the AML cohort (Figure 5.2). Notably, individuals presenting with the TKD mutation were significantly older (p = 0.0018) by 11 years. The median age of the patients within each prognostic cohort was 67 years (range 39 to 80 years) and 56 years (range 17 to 82 years) within the $FLT3/TKD^{+}$ and $FLT3/TKD^{-}$ subgroup, respectively. Seeing, that there was a tendency of telomere shortening within patients who presented with the $FLT3/ITD^{+}$ mutation (Figure 5.12) further analysis was performed to reassure this observation was based solely on the presence of the TKD mutation (Figure 5.13b). In patients in which both FLT3/ITD and FLT3/TKD status were established, telomere length was longer (although with reduced significance) within patients presenting with the TKD mutation following the exclusion of those also bearing an ITD mutation (p = 0.0463).

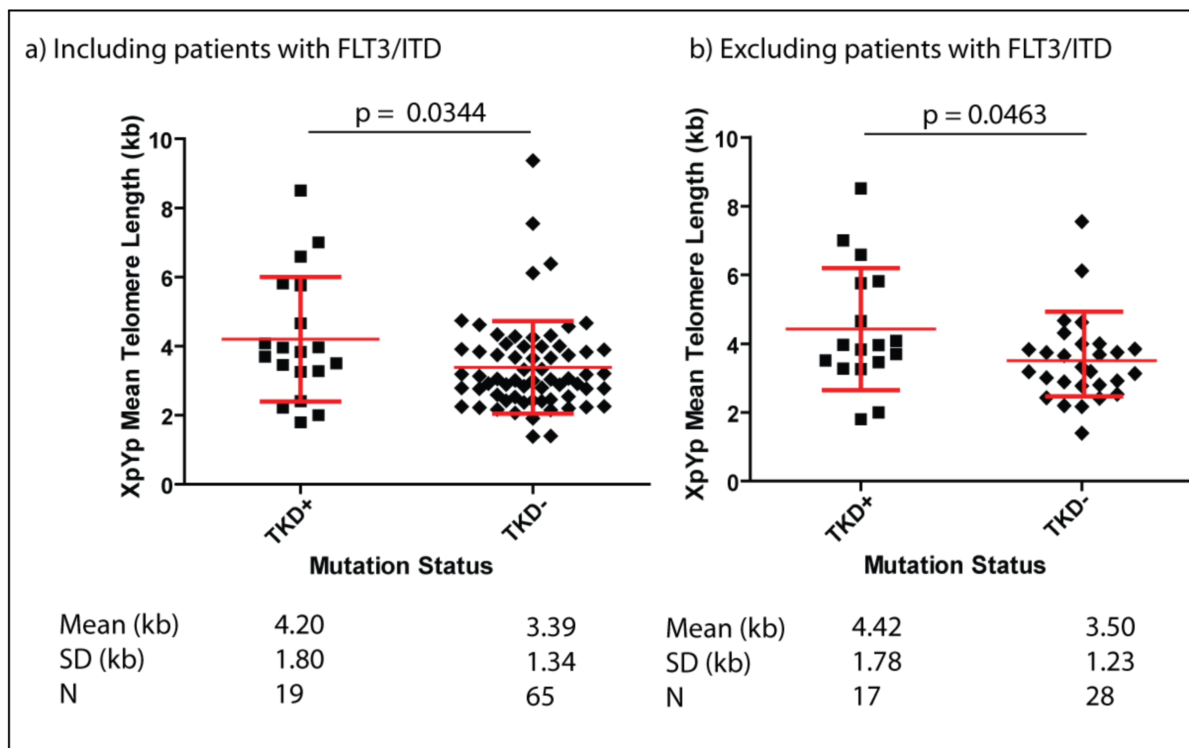


Figure 5.13: Patients who presented with the FLT3/ITD mutation had significantly longer telomeres ($p < 0.05$) compared to those that did not. Exclusion of patients who also shared the FLT3/ITD mutation also revealed that telomere length was longer in those exhibiting the FLT3/TKD mutation (b).

5.12 Telomere Length and NPM1 and FLT3/ITD Mutation Status

Patients presenting with the $NPM1^+$ mutation have a favourable prognosis however, this outcome is abrogated if AML cells also share the $FLT3/ITD^+$.[72] Data was categorised into subgroups associated with FLT3/ITD and NPM1 status so that telomere length within each cohort could be analysed (Figure 5.14). Telomere length at XpYp was 3.27kb (±1.62kb) in the ITD^+NPM1^+ subgroup whereas the telomere length in cases presenting with either a sole ITD^+ or $NPM1^+$ mutation was 3.30kb (±0.80kb) and 3.88kb (±1.45kb), respectively. However, these differences in telomere length did not reach significance with p = 0.9418 and p = 0.1576, respectively. The mean telomere length was similar among patients with the ITD^+NPM1^+ and ITD^+NPM1^- however, further analysis revealed that the ITD^+NPM1^+ subgroup presented with the shortest median telomere length at 2.91kb. The $NPM1^+$ is associated with favourable prognosis[72] and although not significant, patients presenting with this mutation had longer telomeres.

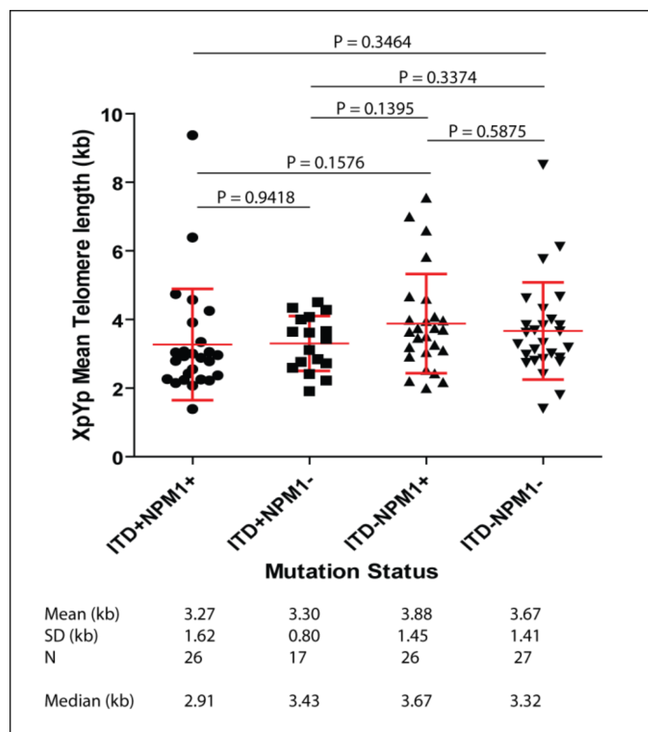


Figure 5.14: There was no association between telomere length with ITD and NPM1 status, however the median was shorter in patients presenting with ITD^+NPM1^+.

5.13 Telomere Length and Status after First Cycle of Intensive Chemotherapy

97 of 110 patients received intensive chemotherapy at diagnosis. 10 of these cases died of treatment related causes, i.e. induction death whereas 84 patients achieved complete remission and 3 showed resistant disease. Diagnostic telomere length was analysed to determine whether it influenced the patient's response after the first cycle of intensive chemotherapy. Short telomere length appeared to be associated with chemo-resistance (Figure 5.15) however this result is inconclusive because only 3 patients were unsuccessful at achieving remission. The telomere length at XpYp was 3.54kb ± 1.34kb in patients who entered complete remission (CR) and 2.67kb ± 0.88kb in patients who were showing resistant disease. This did not reach significance (p = 0.2674).

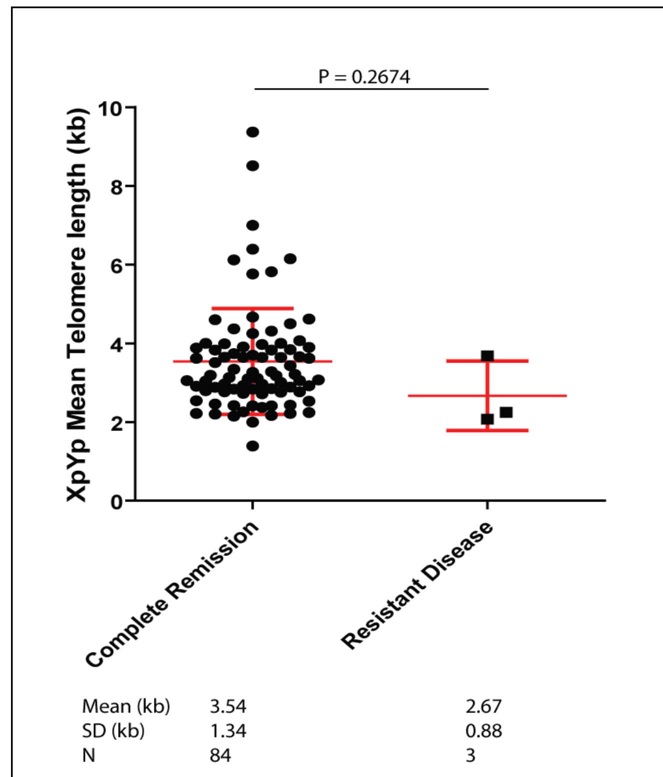


Figure 5.15: Telomere length was shorter in patients presenting with resistant AML however this was not significant and the data set was limited.

5.14 Telomere Length and Disease-Free Survival

Disease-free survival is defined as the time from entering complete remission up to the point of relapse or death in 1st remission. The diagnostic telomere length at XpYp (Figure 5.16) and 17p (Figure 5.17) divided patients into subgroups that were defined by telomere length cut-off points. Kaplan Meier curves were subsequently generated to determine whether telomere length could influence disease-free survival. In this study 84 of 97 patients who received intensive chemotherapy entered complete remission however, the telomere length at 17p was only available for 23 of these cases. The log-rank test failed to show any prognostic impact of telomere length on the number of disease-free days at both the XpYp (Figure 5.16) and 17p (Figure 5.17) telomere.

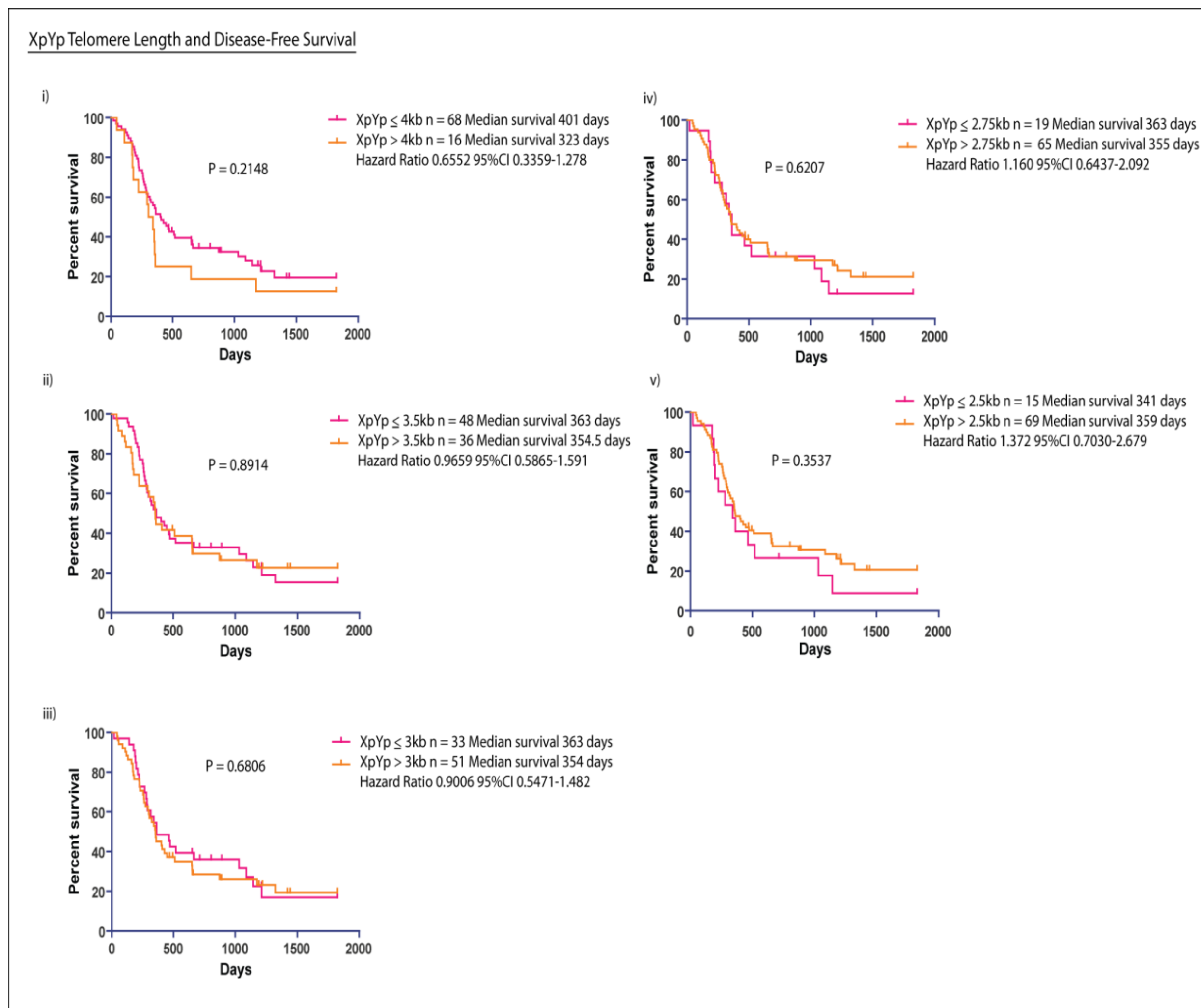


Figure 5.16: Kaplan Meier curves failed to show any prognostic impact of telomere length on the number of disease-free days at the XpYp telomere.

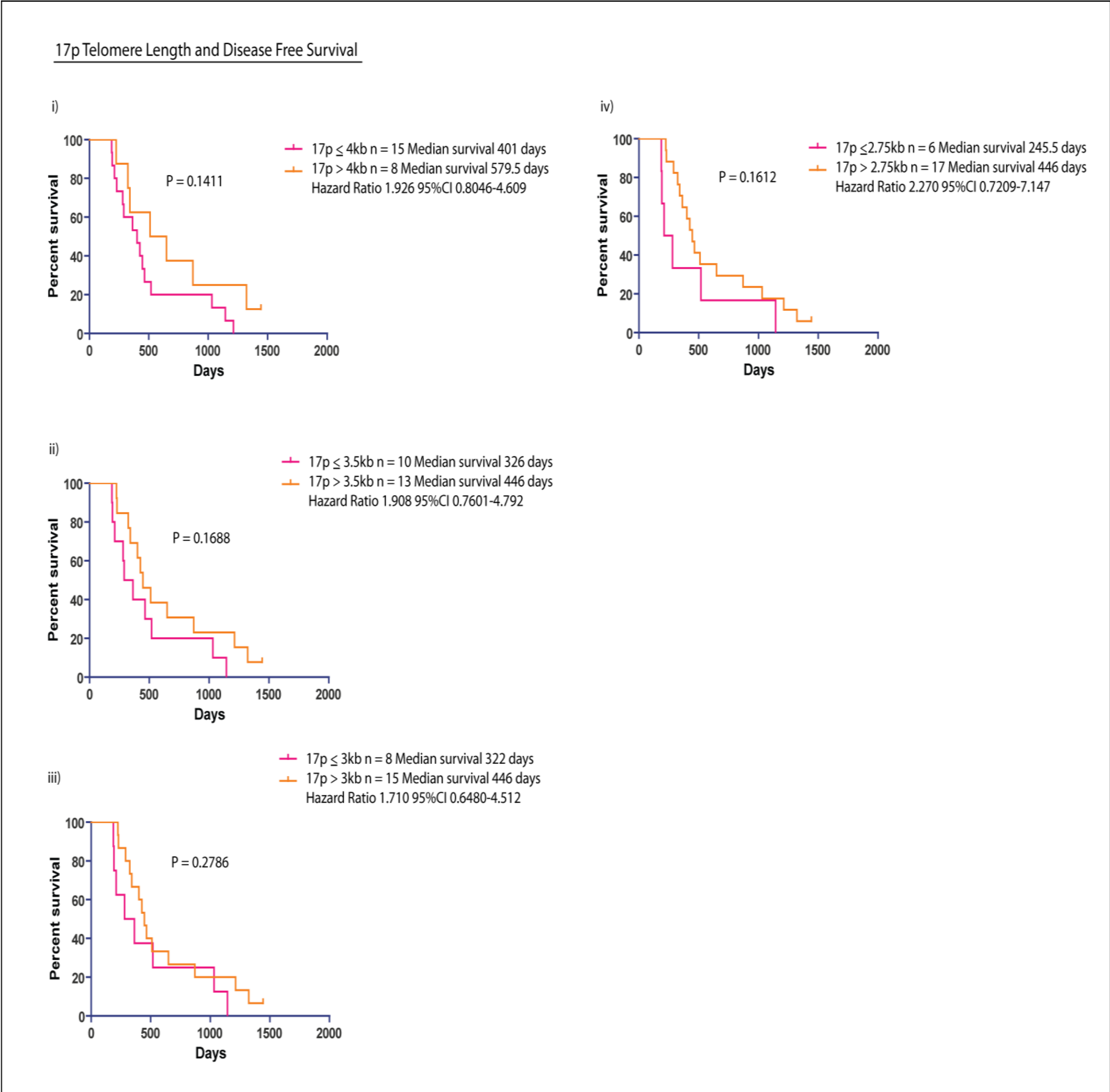


Figure 5.17: Kaplan Meier curves failed to show any prognostic impact of telomere length on the number of disease-free days at the 17p telomere.

5.15 Telomere Length and Overall Survival

Overall survival was defined as time from diagnosis up to the point of death. Recursive partitioning was used to divide patients into groups above and below defined telomere length cut-off points at XpYp (Figure 5.18) and 17p (Figure 5.19). Kaplan Meier curves were subsequently generated to determine whether the telomere length at diagnosis influenced patient's overall survival. 97 cases were analysed at the XpYp telomere whereas only 23 of these individuals were analysed at 17p. The log-rank test failed to show any prognostic impact of telomere length on overall survival at both the XpYp (Figure 5.18) and 17p (Figure 5.19) telomere.

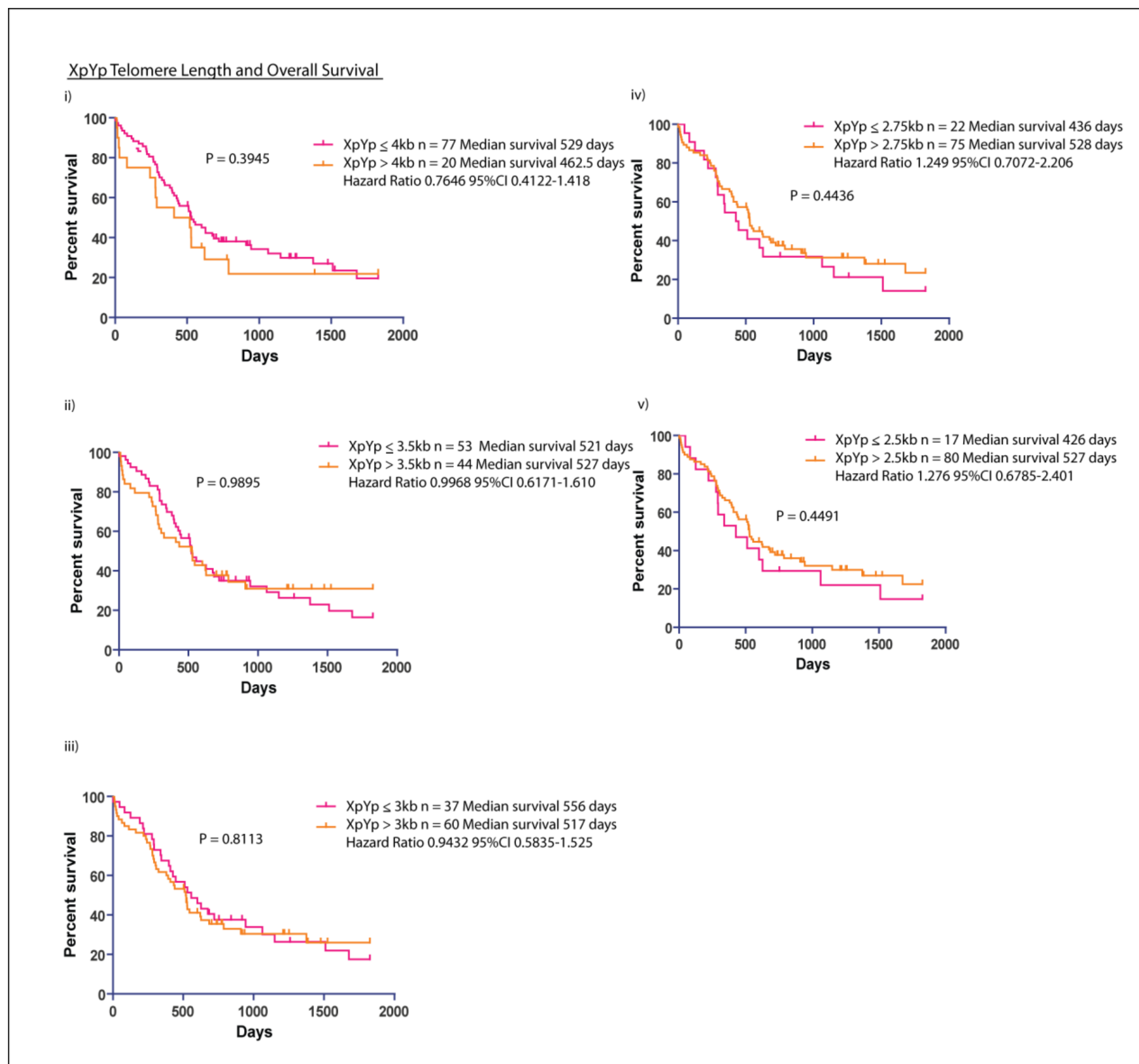


Figure 5.18: Kaplan Meier curves failed to show any prognostic impact of telomere length on overall survival at the XpYp telomere.

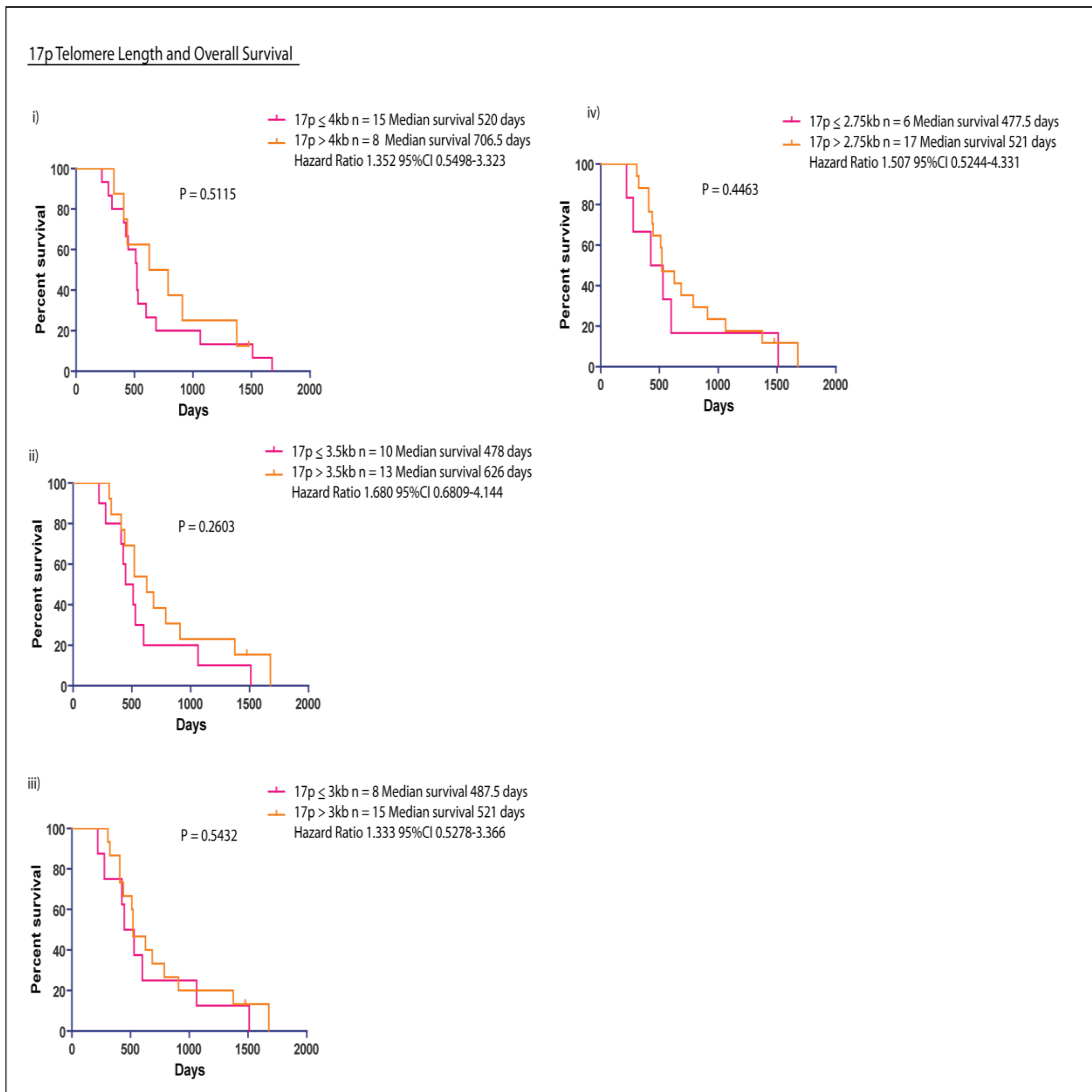


Figure 5.19: Kaplan Meier curves failed to show any prognostic impact of telomere length on overall survival at the 17p telomere.

Although the rate of entry into 1st remission was unavailable for this study it may be tempting to speculate that telomere length does not influence this parameter since the results from both disease-free survival and overall survival were consistent. Although Figure 5.15 might indicate that telomere length at diagnosis could have a role in determining the patient's response after the 1st cycle of intensive chemotherapy with short telomeres associated with resistant AML, the data set was very limited and thus no conclusion can be made.

5.16 Discussion

Acute Myeloid Leukaemia (AML) is an aggressive myeloid neoplasm characterised by the clonal proliferation of undifferentiated myeloid precursor cells and represents a group of heterogeneous conditions with a diversity of clinical and biological features.[367] AML can evolve *de novo* or secondary to an anecdotal haematopoietic disorder or to prior chemotherapy or radiotherapy for a non-haematological disorder.[53]

Telomere length was analysed within a cohort of AML patients (*de novo* and secondary) to determine whether it was associated with clinical parameters at presentation. Additionally, in an attempt to establish if diagnostic telomere length influences patient outlook; response status after the 1st cycle of intensive chemotherapy, disease-free survival and overall survival were investigated with respect to telomere length.

Consistent with these data, extensive telomere shortening has been observed among AML patients when compared to aged-matched healthy controls.[347,350] Moreover, it was evident that telomere length did not comply with the aging dogma. Notably, younger cases were showing elevated telomere attrition when compared to older AML patients. Telomere length was significantly longer in patients >60 years in contrast to their younger counterparts (≤60 years). It has been previously shown that a large proportion of older AML cases present with poorly differentiated subtypes of AML (FAB: M0/M1) with the probability of these AML subtypes increasing with age.[375] Interestingly, telomerase activity is correlated with FAB subgroups with activity following M1>M2>M5>M4[376] with a significant increase in M1.[377] Accordingly, longer telomeres within the aging population might be associated with a bias towards M1 cases however, FAB subgroups were not available for this study and therefore this cannot be concluded.

Telomere length was not correlated with white blood cell (WBC) count and blast count at presentation. This was expressed as both categorical and continuous variables. The process behind clonal expansion is likely accountable for this finding, i.e. telomere dynamics and the process behind disease development. Notably, this study (Chapter 3; Figure 3.15) as well as others has shown a significant up-regulation in telomerase activity[342,347] in AML and thus it is possible that in the presence of up-regulated telomerase and *hTERT* expression;[342,347,376,377]

the access of telomerase to the telomere is deregulated resulting in fluctuations in telomere length control within individual cases.

The prognosis of patients harbouring the FLT3/ITD is poor with individuals exhibiting high relapse rates and an inferior overall survival (OS).[64,65] However, the prognostic impact of FLT3/TKD remains controversial.[60,66-68] This study, as well as others has identified telomere shortening within patients who express the FLT3/ITD mutation.[347,350] FLT3 is expressed by cells found in the haematopoietic stem cell compartment and early committed progenitors.[59,60] The stimulation of FLT3 by its ligand has been proposed to play a role in cell proliferation.[59,60] The ITD and TKD mutation both result in an increase in cellular proliferation however, several studies[378,379] have shown a weaker proliferative effect induced by the FLT3/TKD when compared to patients with the FLT3/ITD. The clonogenic ability of TKD expressing cells *in vitro* has been observed to be significantly less when compared to ITD cells.[380] Notably, the ITD but not TKD mutation (or wild-type FLT3) has been shown to induce robust activation of the STAT5 signalling pathway *in vitro*.[380,381] STAT5 (signal transducer and activator of transcription-5) is one of the principal pathways that regulates gene expression in response to FLT3.[382] Downstream targets of STAT5 include Pim-1 and CCND3 which both play a role in cell cycle progression and are increased in FLT3/ITD cells.[383] Thus, up-regulation of the STAT5 pathway may be in part accountable for the tendency towards shorter telomeric length in patients presenting with the ITD mutation. However, because this did not reach statistical significance it is tempting to speculate that telomerase activity is preferentially up-regulated in ITD$^+$ AML cells which may mask the increased proliferation associated with this genetic mutation. Patients presenting with the TKD mutation had significantly longer telomeres when compared to those who did not. It is possible that this difference is in part, associated with differential STAT5 signalling between ITD$^+$ and TKD$^+$ AML cells.

Different age profiles for molecular markers have been described such that the ITD occurs at a constant frequency irrespective of age[384] and the incidence of the TKD increases by 29.4 fold from 21 to 70 years.[56] Patients who expressed the FLT3/TKD mutation were significantly older when compared to those who were negative for FLT3/TKD. Moreover, this cohort of patients presenting with the FLT3/TKD mutation had longer telomeric length. Therefore, it is

possible that the prevalence of FLT3/TKD in the elderly population may be in part accountable for the finding of longer telomeres within older patients.

Patients were sub-divided based on the NPM1 and FLT3/ITD mutation status. Although not significantly different, patients presenting with a sole NPM1 mutation showed the longest telomere length when compared to the other three subgroups. Mutations in NPM1 are associated with a more favourable outcome when compared to patients presenting with wild-type NPM1.[72] MicroRNA expression profiling of $NPM1^+$ and $NPM1_{wild\text{-}type}$ AML cells has identified different patterns of microRNA (miRNA) expression. Consistent with a favourable prognosis, the tumour suppressor family of microRNAs *let-7* is up-regulated in $NPM1^+$.[72] The overexpression of *let-7* has been observed to inhibit cell proliferation in human lung cancer cell lines.[385] Notably, it has been documented to interact with *MYC* and *CDC25A*[386] regulating cell proliferation and cell cycle progression, respectively. It is possible that the tendency for longer telomeres within the ITD^-NPM1^+ subgroup may attribute to the tumour suppressor effect, i.e. *let-7* on inhibiting cellular proliferation. The reason behind the apparent elevated telomere attrition of the median length is inconclusive in patients presenting with both mutations. Inconsistent with the literature, $NPM1^+$ has been proposed to oppose the FLT3/ITD dependent activation of STAT5[387] however; cells expressing ITD^+ without the NPM1 mutation may have already undergone an extensive period of proliferation and have preferentially up-regulated telomerase in order to maintain telomere stability. Thus, telomerase activity may in part be accountable for longer median telomeric length within the ITD^+ $NPM1^-$ subgroup.

Secondary AML can arise from a preceding chronic phase such as a Myelodysplastic Syndrome (MDS) or Myeloproliferative disease (MPD).[53] This study identified significantly longer telomeres in patients with secondary AML when compared to *de novo* cases. Moreover, the standard deviation of the telomere distribution was not significantly different between the two subsets suggesting that this difference was not related to intra-clonal variation. It is possible that secondary AML cells accumulated multiple genetic alterations during its antecedent pathological phase and telomerase up-regulation was an early event before AML evolution which may be greater in patients with secondary AML.

Clinical management and decision making in AML relies strongly on risk stratification based on conventional karyotyping.[368] Cytogenetic risk has been divided into three subcategories which include favourable, intermediate and adverse.[369] Favourable abnormalities are composed of balanced translocations such as *AML1-ETO* t(8; 21), *PML-RAR*α t(15;17) and *CBF*β*-MYH11* inv(16)/t(16;16) whereas unbalanced translocations and complex karyotypes (3 or more clonal abnormalities) are adverse cytogenetic characteristics.[369] In contrast to previous studies,[342,347] telomere length was not significantly different between any of the cytogenetic subgroups. However, it is possible that telomere dysfunction arose at an early stage in disease development, initiating BFB (Breakage-Fusion-Bridge) cycles and generated the cytogenetic complexity found in these cells. Subsequent telomerase activity would stabilise the telomeres possibly resulting in the longer telomere length identified in patients with an adverse karyotype. Notably, chromosomal complexity has been shown to be correlated with *hTERT* expression[347] and patients who present with gains/losses of chromosomes show elevated telomerase activity when compared to patients with a normal or balanced karyotype.[342] However, the limited number of cases presenting with an adverse or favourable karyotype in this study renders this finding inconclusive.

Although not significant it appeared that shorter diagnostic telomere length was associated with chemo-resistance after the first cycle of intensive chemotherapy. However, due to the limited number of individuals' refractory to therapy, a conclusion of whether telomere length can be used as prognostic marker in this instance is uncertain.

Consistent with the literature,[347,350] the log-rank test in this study failed to show any prognostic impact of telomere length on the number of disease-free days and overall survival. Although the rate of entry into 1st remission was unavailable for this study it may be tempting to speculate that telomere length does not influence this parameter since the results from both disease-free survival and overall survival were consistent. It may be speculated that telomerase up-regulation in AML cells provides telomere length stability thereby removing the prognostic signature of telomere length. This is in contrast with what was detected within the MDS cohort in which shorter telomeric length appeared to be associated with a poorer outlook.

In conclusion, telomere length at diagnosis appears to have a minimal role in influencing the outlook of AML patients. However, due to the limited number of patients presenting with specific clinical parameters such as adverse cytogenetics and resistant disease, a conclusion based on these factors cannot be made.

Chapter 6:

Telomere Dysfunction and its Potential Role in Promoting Genetic Instability in MDS and AML

6.1 Abstract

Telomere induced genetic instability may contribute to the development of AML by means of promoting an accumulation of non-reciprocal translocations (NRTs) through multiple breakage-fusion-bridge (BFB) cycles resulting in gross chromosomal rearrangements.

To examine the extent of telomere dysfunction in MDS and AML, telomere fusion events were quantified at the XpYp and 17p telomere using a PCR based fusion assay. The mean frequency of fusion events detected within the MDS cohort was 2.33×10^{-6} ($\pm 5.80 \times 10^{-6}$) and 6.67×10^{-6} ($\pm 1.27 \times 10^{-5}$) at the XpYp and 17p telomere, respectively. There appeared to be a reduction in the frequency of fusion within the AML cohort possibly reflecting the up-regulation of telomerase activity in AML cells. The mean frequency of telomere adjacent fusions at the XpYp and 17p telomere in AML was 1.76×10^{-6} ($\pm 4.64 \times 10^{-6}$) and 1.86×10^{-6} ($\pm 6.61 \times 10^{-6}$), respectively. In one individual the fusion assay revealed the presence of a clonal telomere fusion that was shared by a 40% minimum of AML cells.

Direct sequencing revealed that telomere fusion partners aligned at short patches of 100% homology that ranged in length from 2 to 33nts following the sub-telomeric deletion at one or both telomeres. This profile was consistent with error-prone Ku-independent alternative end joining processes. Sequencing also revealed the existence of complex fusion events involving insertions of non-telomeric genomic loci, including the common fragile sites FRA17B and FRA19A mapped to 17q23 and 19q13, respectively.

Array-CGH demonstrated that telomere dysfunction may contribute to chromosomal instability and disease progression. This was emphasised in an array profile that presented gross chromosomal loss that extended to the telomeres.

These data are consistent with the view that telomere dysfunction may contribute to the progression of MDS and AML via telomere fusion. The resultant Breakage-Fusion-Bridge

(BFB) cycles have the propensity to generate gross chromosomal rearrangements which may be detected by means of CGH analysis.

6.2 Introduction

Chromosome end protection is maintained by the telomere cap, i.e. Shelterin in which TRF2 has been proposed in playing a vital role in suppressing aberrant telomere-telomere fusion.[199] It has been proposed that critically short telomeres are unable to provide sufficient stability of telomere binding proteins[255] which in turn reduces the proficiency of Shelterin to maintain the integrity of chromosome ends. Consequently, chromosome ends may no longer be protected by the telomere 'cap' and are recognised as double strand breaks (DSBs).[202] In the presence of a functional DNA damage response (DDR), the ATM or ATR pathway[202] either eliminates the cell by apoptosis or supresses cell cycle progression by means of p53 or p21 activation.[314] In the absence of a functional DNA damage response cells may re-enter the S phase and continue into post–senescent replication.[314,317] Continued telomere shortening in these cells may lead to telomeres that are denuded of all TTAGGG repeats and induce further erosion into the sub-telomeric region of the chromosome. Short dysfunctional telomeres generated either as a consequence of cellular replication or stochastic telomere rapid deletion (TRDs) can be subjected to fusion with other telomeres or non-telomeric double stranded DNA breaks.[255,351] This is consistent with observations in mouse models where numerous end-to-end fusions and signal free ends have been observed in late generation $mTR^{-/-}$ $p53^{-/-}$ mutant mice.[322] Therefore telomere dysfunction has the potential to drive genetic instability during oncogenesis by initiating breakage-fusion-bridge (BFB) cycles upon the formation of dicentric chromosomes.[319]

Dicentric chromosomes are pulled apart to opposite poles during anaphase causing chromosome breakage.[319] This generates new recombinogenic free ends that have the propensity to enter into another cycle of BFB. This paradigm has been documented to induce genetic instability in many human solid tumours including pancreatic carcinoma[321] and osteosarcomas[321] as well as in human leukaemias including Chronic Lymphocytic Leukaemia (CLL).[304] Ultimately, cells acquire a novel genetic profile that stabilises following telomere maintenance which can be achieved by the expression or up-regulation of *hTERT* which has been detected in over 90% of human cancers.[180] Thus, this process has the ability to create gross chromosomal rearrangements driving the path towards cancer development prior to telomerase activation or up-regulation.

Cell cycle control is commonly abrogated in advanced MDS; such examples include the downregulation of Cdk1[122] and hypermethylation of p15^{ink4B}.[124,388] Furthermore, the loss of ATM[344] and p53[130] has been observed in AML and MDS, respectively with which has been noted to accompany abnormal cytogenetics.[130,136] It has also been demonstrated that *hTERT* expression[347] and telomerase activity[342] are more pronounced in AML with complex rearrangements. This may result in the stabilisation of telomere length and structure and a reduction in the frequency of end-to-end fusion events.[212]

Classical Non-Homologous End Joining (C-NHEJ) appears to be the most common pathway of DSB repair in mammalian cells.[280] It is dependent on the heterodimeric protein Ku,[282] DNA-PKcs[285] and Ligase IV[288,289] and appears to play a dominant role in telomere fusion following the experimental abrogation of TRF2.[202] Conversely, fusion events between short dysfunctional telomeres have been detected in human and mouse cells deficient in NHEJ components including Ku[292] and DNA-PKcs,[293,294] respectively suggesting that an alternative repair mechanism may operate at dysfunctional telomeres. Alternative Non-Homologous End Joining (A-NHEJ) is error prone, resulting in large deletion events and is dependent on sites of microhomology at the fusion junction. It has been proposed that base pairing at sites of microhomology, particularly within those comprised of G:C bases,[308,311] may compensate for the loss of stability in the absence of NHEJ components.[298] Small regions of microhomology have been previously detected within telomere fusion junction points in human cells undergoing crisis in culture[255] and tumour cells.[304] Subsequent sequencing revealed that these events involved the deletion of one or both of the participating telomeres with deletions extending into sub-telomeric DNA.

Genome-wide telomere shortening has been associated with complex chromosomal rearrangements and a poorer prognosis in MDS and AML.[333,334,342] However; cell viability and chromosomal stability are coupled to the shortest telomere within the distribution and it is the shortest telomere within a distribution that has the propensity to initiate telomere fusion.[209] Thus, it is of interest to determine whether telomere fusion events arise in MDS and AML provided that telomere dysfunction may play a role in the development and pathogenesis of these diseases.

Given the potential role of telomere length in the pathogenesis of MDS/AML, the extent of telomere dysfunction was examined in this study using a PCR-based telomere fusion assay. This assay enables the detection and quantification of single telomere-telomere fusion molecules between specific chromosome ends.[255] Oligonucleotide primers are utilised that target the telomere-adjacent sequence of specific telomeres (Figure 6.1). Long-range single molecule amplification prior to Southern Hybridisation with telomere adjacent probes enables the detection of specific products. In order to identify the participating telomeres and coexisting fusion junction, putative single fusion molecules are re-amplified with nested PCR primers. Direct sequence analysis can determine the nature of the telomere fusions which is achieved over multiple sequence reactions.

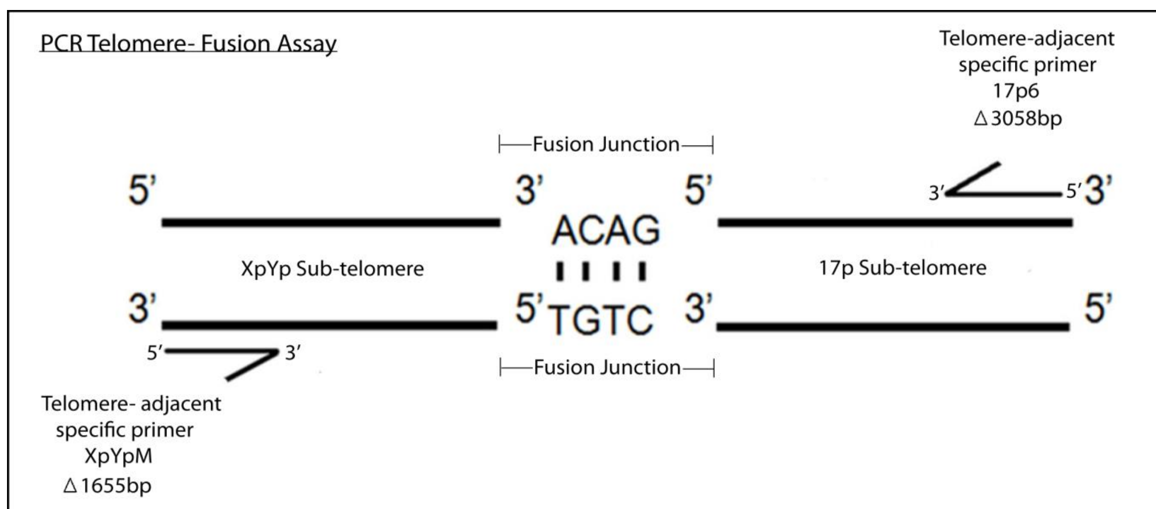


Figure 6.1: PCR based telomere fusion assay. XpYpM specific for the XpYp telomere extends 1655bp into telomere-adjacent DNA. 17p6 specific to 17p targets sub-telomeric DNA 3058bp from the start of the telomere. Primers extend in the 5' to 3' direction towards the chromosome terminus in order to cross the fusion junction.

Results

6.3 MDS/AML and Telomere Fusion

A subset of MDS patients (n = 55) were examined for fusion events at the XpYp and 17p telomeres. Of those that showed telomere fusions, 10 patients (18%) presented fusions involving the XpYp telomere and 22 (40%) demonstrated fusions involving the 17p telomere (Figure 6.2; dictated by red and yellow points). Each fusion reaction consisted of 100ng/μl of DNA and upon taking into account the weight of nuclear DNA within a diploid cell, i.e. 6pg the frequency of fusion events could be quantified within each patient. The mean frequency of fusion events was 2.33×10^{-6} ($\pm 5.80 \times 10^{-6}$) and 6.67×10^{-6} ($\pm 1.27 \times 10^{-5}$) at the XpYp and 17p telomere, respectively. Furthermore, it was apparent that the frequency of fusion events at the 17p telomere occurred at a significantly higher frequency than at the XpYp telomere (p = 0.0232).

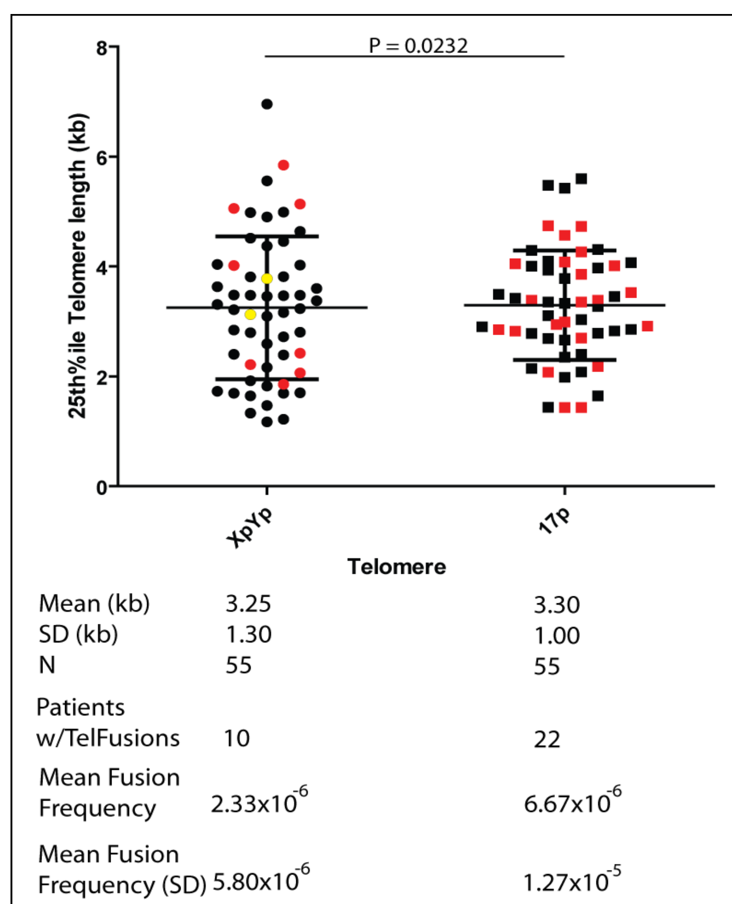


Figure 6.2: MDS patients highlighted in red and yellow show fusions at either the XpYp or 17p telomere. There appeared to be a significant increase in the number of fusion events involving the 17p telomere (p = 0.0232). The patients marked in yellow present with extensive TVR regions that ranged between 1.5kb to 3kb.

Telomere fusion was also analysed within a subset of AML patients (n = 58; secondary and *de novo* AML). Of those that showed telomere fusions, 10 patients (18%) presented fusions involving the XpYp telomere and 7 (12%) demonstrated fusions involving the 17p telomere (Figure 6.3; dictated by coloured points). The mean frequency of fusion events at the XpYp and 17p telomere had been recorded at 1.76×10^{-6} ($\pm 4.64 \times 10^{-6}$) and 1.86×10^{-6} ($\pm 6.61 \times 10^{-6}$), respectively. The difference in the frequency of telomere fusion events at the XpYp and 17p telomeres did not reach statistical significance (p = 0.9225).

The frequency of telomere fusion events was reduced within the AML cohort compared to the MDS cohort, although this was not significant at XpYp (p = 0.5652), but was at 17p (p = 0.0125).

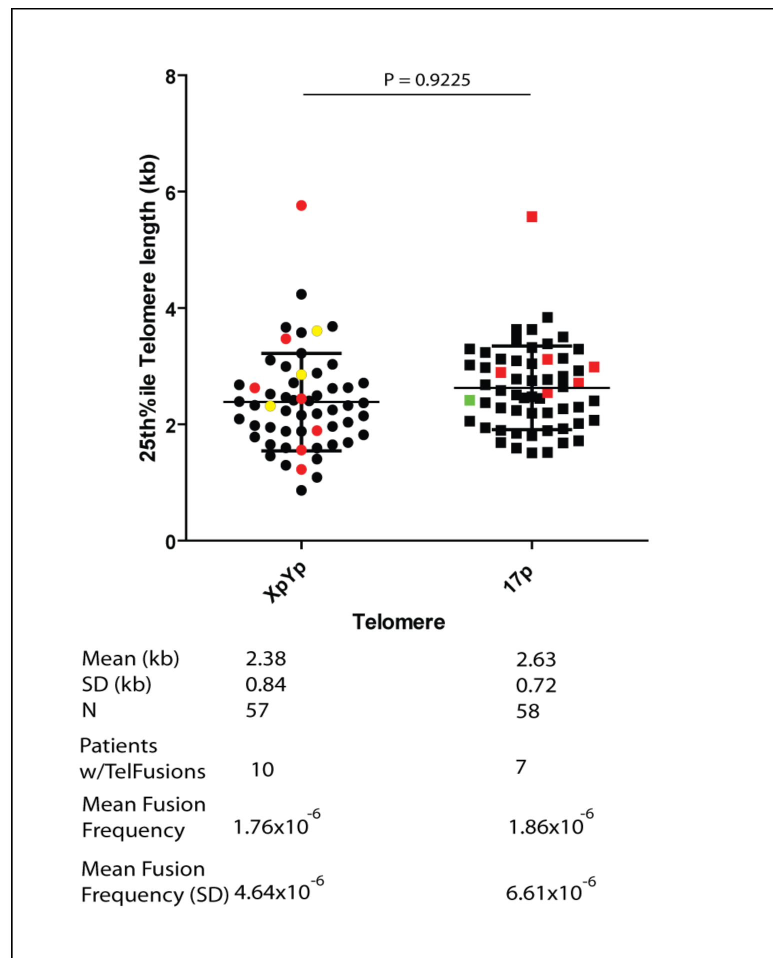


Figure 6.3: AML patients highlighted in colour show fusions at either the XpYp or 17p telomere. There was not a significant difference in the number of fusion events between XpYp and 17p (p = 0.9225). The patients marked in yellow presented an extensive TVR region between 1.5 and 2.5kb. A clonal fusion event was detected at the 17p telomere within the patient marked in green.

Patients exhibiting apparently long, normal telomere lengths demonstrated the presence of telomere fusion events. It is possible that extensive TVR regions within the telomere played a role in this observation. Further analysis revealed that the length of the pure TTAGGG tract did not appear to be associated with the presence of telomere fusion events within the MDS and AML cohort (Figure 6.4). Accordingly, telomere fusion events were apparent within patients exhibiting long pure TTAGGG tracts whilst telomere fusion was not detected in several patients showing short pure TTAGGG tracts.

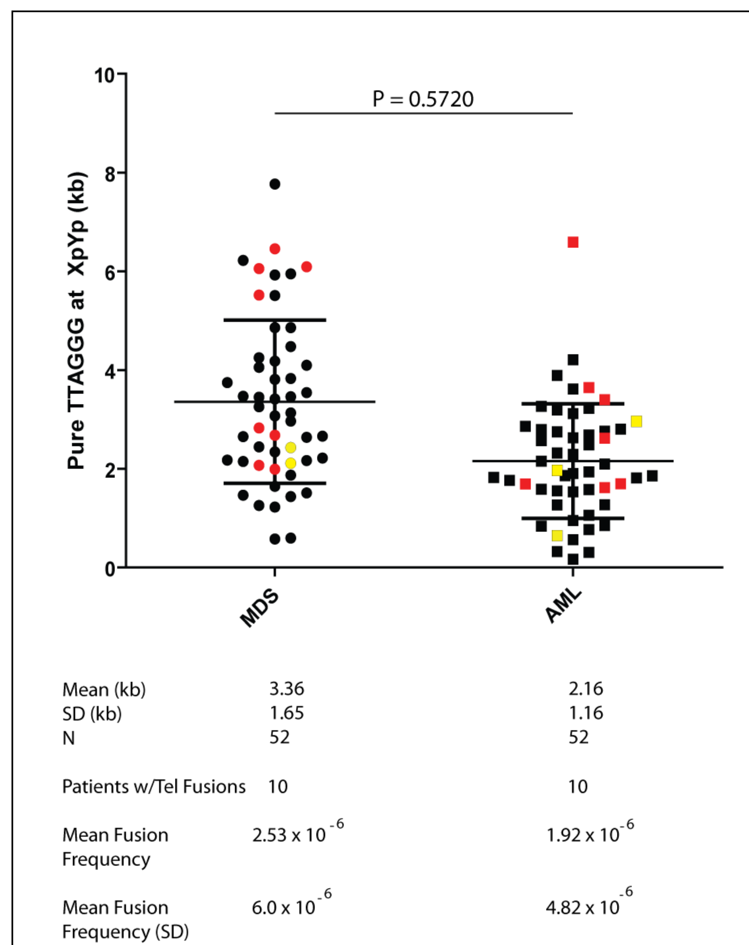


Figure 6.4: Telomere fusion did not appear to be related to the length of the pure TTAGGG tract within the MDS and AML cohort. Fusion events were detected in patients highlighted in red and yellow. Those marked in yellow presented with large TVR regions that ranged between 1.5kb to 3kb and simultaneously presented with telomere fusions. There was not a significant difference (p = 0.5720) between the frequency of fusion events at XpYp detected between the MDS and AML cohort in which TVR data was available.

However, it may be possible that a subset of individuals might be predisposed to telomere fusion if TVR regions are extensive. Notably, one MDS patient exhibited a TVR of 3.03kb

(Figure 6.5). This MDS patient presented numerous telomere fusion events at the XpYp and 17p telomere occurring at a frequency of 2.7×10^{-5} and 3.3×10^{-5}, respectively.

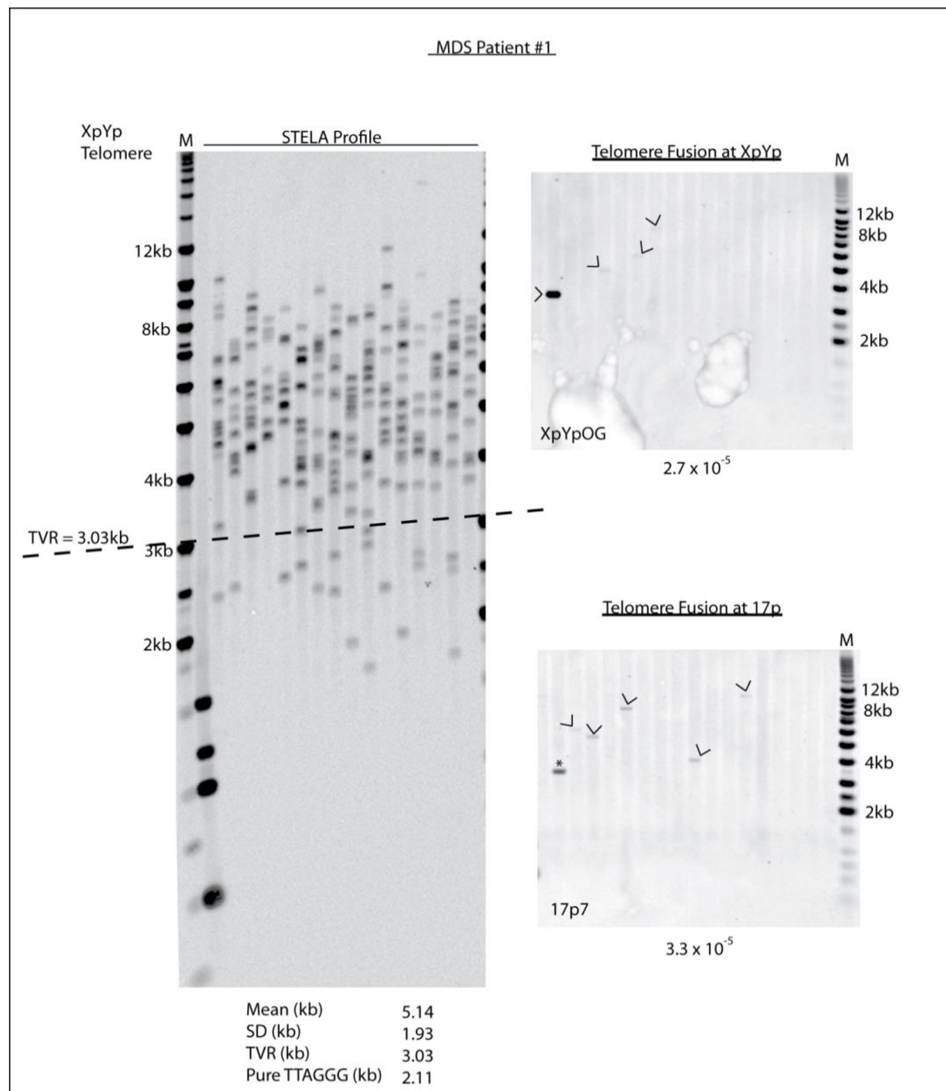


Figure 6.5: The STELA profile at XpYp presented long telomeres with a mean of 5.14kb however; this patient exhibited numerous telomere fusion events. Notably, the TVR was 3.03kb which indicated that the pure TTAGGG tract was only 2.11kb.

A putative clonal telomere adjacent fusion event (~5.5kb) was detected within an AML patient (Figure 6.3 [marked in green] and Figure 6.6). The fusion assay amplifies single DNA molecules thus if the same fusion is identified in separate reactions, it would have been derived from the same precursor. Further analysis revealed that this clonal telomere fusion specifically involved the 17p telomere combined with the 21q family of telomeres as indicated in Figure 6.6(ii); the fusion product could only be detected in the fusion reaction containing the 17p6 + 21q1 reactions. In order to quantify the number of cells comprised of

this telomeric fusion event, serial dilution was required to reduce the fusion event down to single molecule level (Figure 6.6(iii)). A 1:5120 dilution achieved this objective. Taking into account the DNA content of a single diploid cell at 6pg it was calculated that the percentage of cells comprised of this fusion event was 40%. Considering that these were not purified cells it is possible that a large fraction of the AML clone consisted of this telomeric fusion.

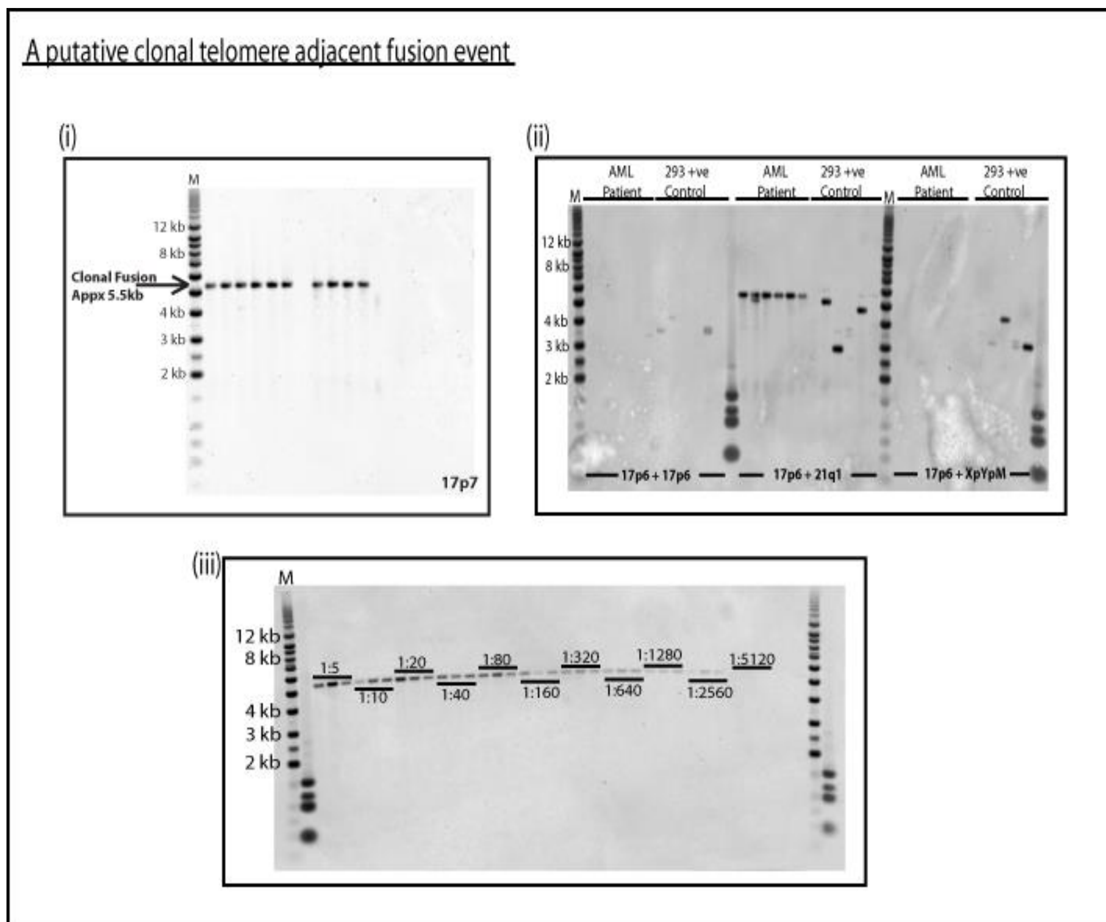


Figure 6.6: (i) A clonal telomere fusion product was detected within an AML patient (~5.5kb). This was detected following the use of the 17p7 probe. (ii) The telomeres participating in the clonal fusion event included specifically 17p and a member of the 21q family. The 293 cell line was used as a positive control. (iii) The clonal fusion was diluted by 1:5120 to the single molecule level and it was calculated that a 40% minimum of cells presented this fusion product.

6.4 Putative Telomere Fusion Events and Sequencing

Putative single fusion molecules were re-amplified by means of nested PCR and were subsequently subjected to direct sequencing in order to determine the internal structure of individual fusion products. Following extensive investigation (Figure 6.7) into retrieving telomeric fusions, a total of 11 fusion events were characterised by DNA sequencing. Re-amplification failed to detect fusion products coexisting within individual fusion profiles, however occasional re-amplification products could be generated some of which yielded interpretable DNA sequence.

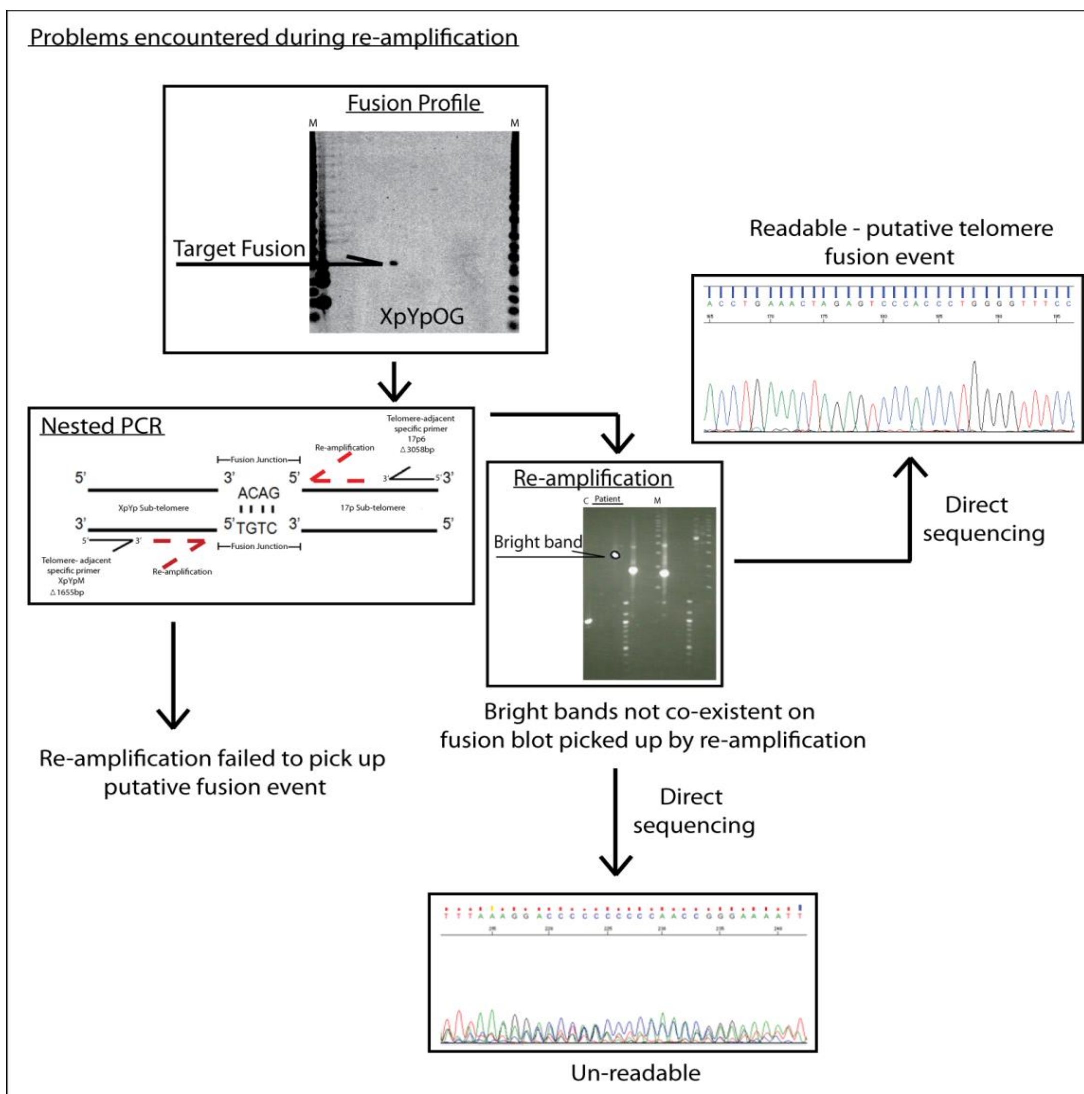


Figure 6.7: Extensive investigation into retrieving telomere fusion products. Nested PCR was performed in an attempt to re-amplify target fusions however, re-amplification failed to detect target fusion products. Conversely, bright bands not coexistent on the fusion blot had been exposed following re-amplification. Direct sequencing revealed that these products were either un-readable or presented successful putative telomere fusion events. Accordingly, 11 fusion products were sequenced.

Unsuccessful Sequence Products

PCR amplification from genomic DNA using primers that can potentially hybridise with several regions throughout the genome runs the risk of producing recombination artefacts. Accordingly, several re-amplified products which appeared promising throughout the course of the study were either non-telomeric, uninterpretable or PCR artefacts (Figures 6.8 to 6.10). Putative telomere fusion events which shared the same primer sequence on each participating telomere were identified such that products could be read successfully up to the point of divergence from which the DNA downstream was composed of merged sequences (Figure 6.8); these products are consistent with telomere fusion, however the fusion point cannot be identified. Interpreting such products was particularly difficult in the case of analysing fusion events involving the 21q telomere in which 13 chromosome ends could be amplified. PCR artefacts consisted of uninterpretable sequence composed of multiple products (Figure 6.9). Non-telomeric interstitial DNA located megabases from the telomere were also observed (Figure 6.10); one example of which was mapped to chromosome 3 (~29Mb from the telomere). Whilst these events may represent genuine telomere fusion events, the absence of a fusion point with the sub-telomeric sequences targeted in the fusion assay, meant that these events could not be verified as arising as a consequence of telomere fusion. Putative telomeric fusions similar to that shown in Figure 6.10i/ii were regularly detected in the current study however, it was uncertain as to whether they were possible fusion events since only the primer sequence was identified in one of the participating telomeres, i.e. XpYpMb or 17p6b. Additional sequence beyond the primer DNA would have supported the characterisation of such a sequence. Reducing the number of cycles or increasing the annealing temperature in a subsequent study may facilitate the reduction of non-specific re-amplification products.

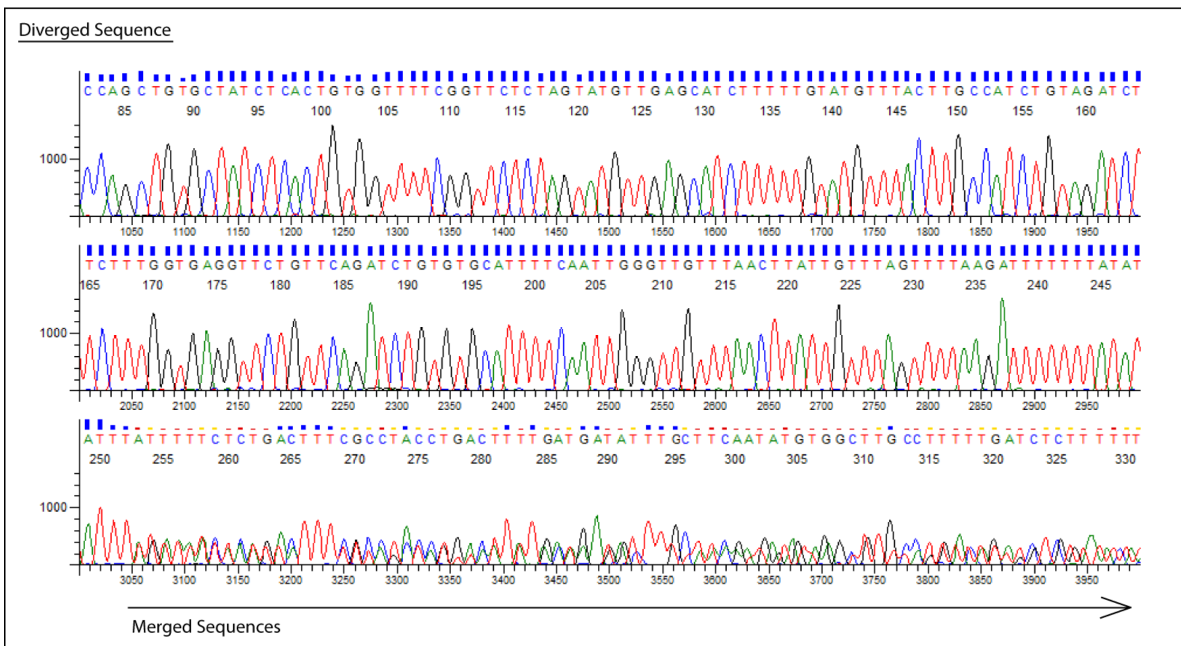


Figure 6.8: The same primer sequence on each participating telomere can successfully amplify a putative fusion product up to the point of divergence (at base 253). Downstream DNA from this site is no longer readable due to the presence of merged sequences

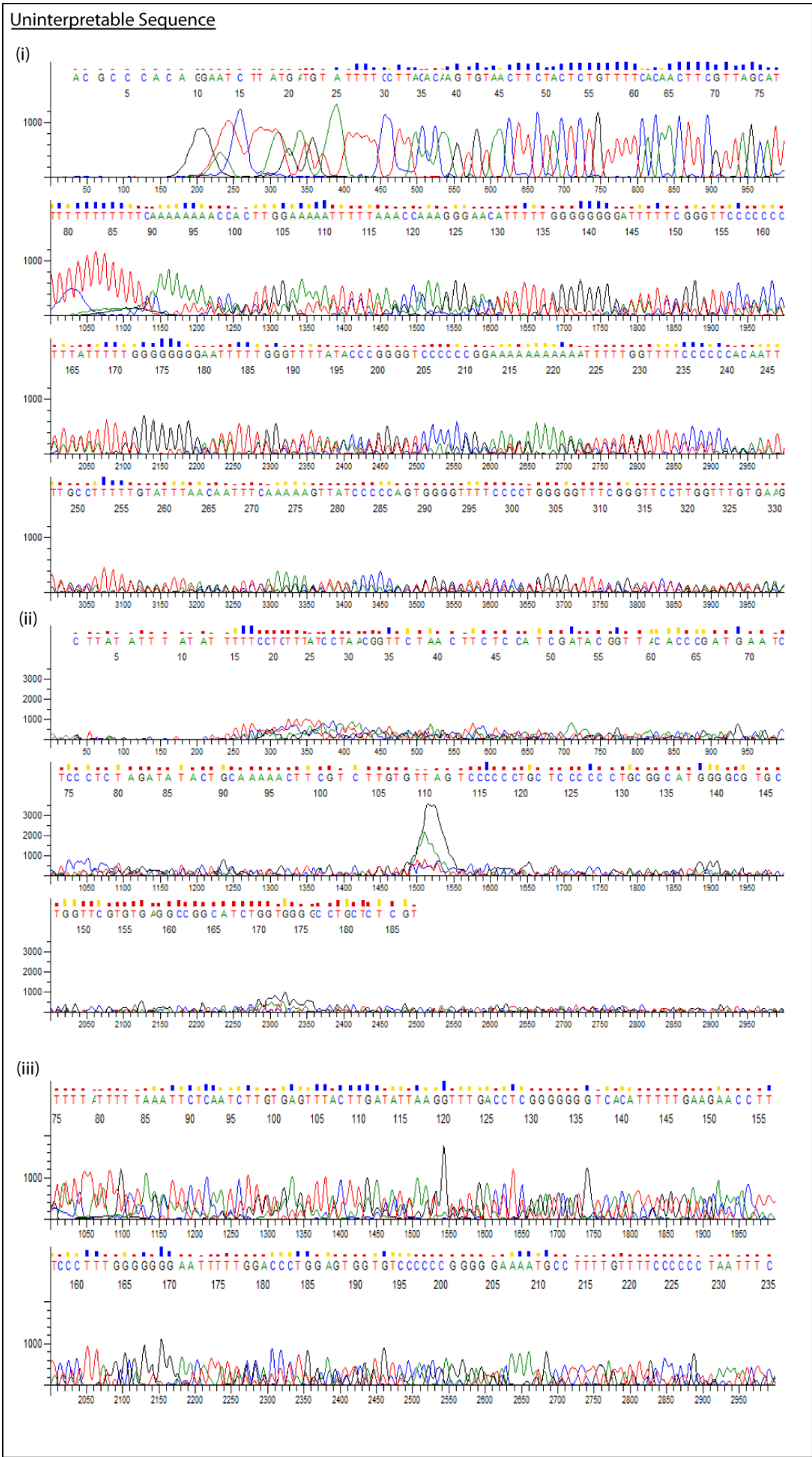


Figure 6.9: Uninterpretable sequences detected throughout the study

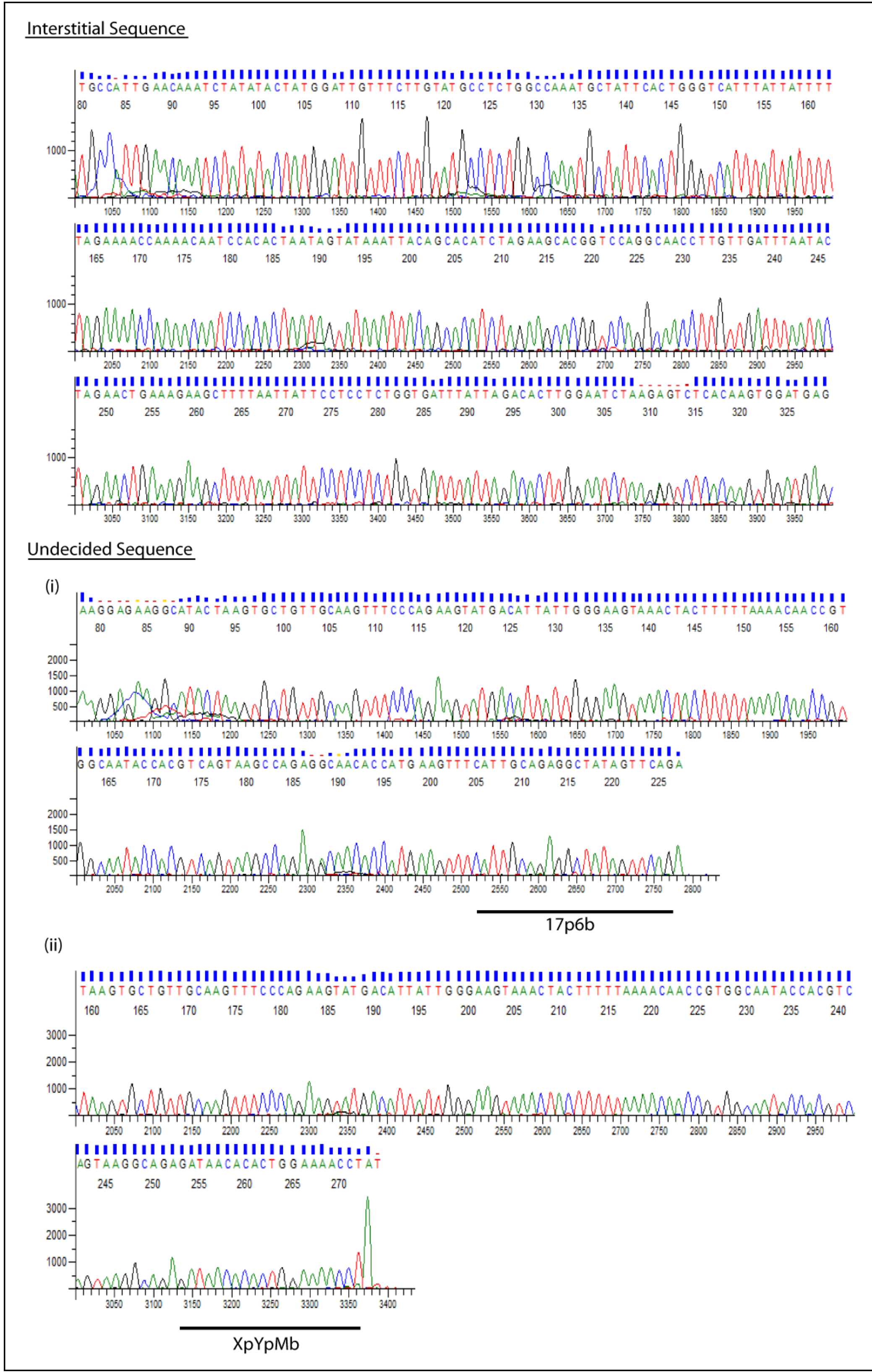


Figure 6.10: Although clear, several sequences were interstitial and not located at the telomere. This particular reading was derived from chromosome 3, mapped ~29Mb from telomere. Several sequences were 'undecided' due to the identification of only the primer sequence in the participating telomere.

Successful Sequencing Data

Direct sequencing of re-amplified fusion products revealed the existence of putative telomeric fusion events.

Upon examining the internal structure of several DNA sequences (Figures 6.12-17 and 6.19-20) it was apparent that the fusion products involved the deletion of one or both telomeres. On average, the mean sub-telomeric deletion at the XpYp and 17p telomere was recorded as 1241bp (±810bp) and 826bp (±1388bp), respectively. The mean length of sub-telomeric deletion at the 21q family of telomeres was 2493bp (±1096bp) (Figure 6.11(i)). Moreover, several deletion events extended into the telomere adjacent DNA close to the limits of the assay, i.e. 1655bp at XpYp and 3058bp at 17p. Interestingly, there appeared to be a frequent breakpoint observed 3640 to 3645bp from the 21q telomere.

100% homology was apparent at the fusion junction which ranged from 2 to 33nts in length. Analysis revealed that the mean length of 100% homology was 8.3nt (±10.5nt) (Figure 6.11(ii), however it was also observed that 2 fusion products displayed no homology at the junction point (Figure 6.11).

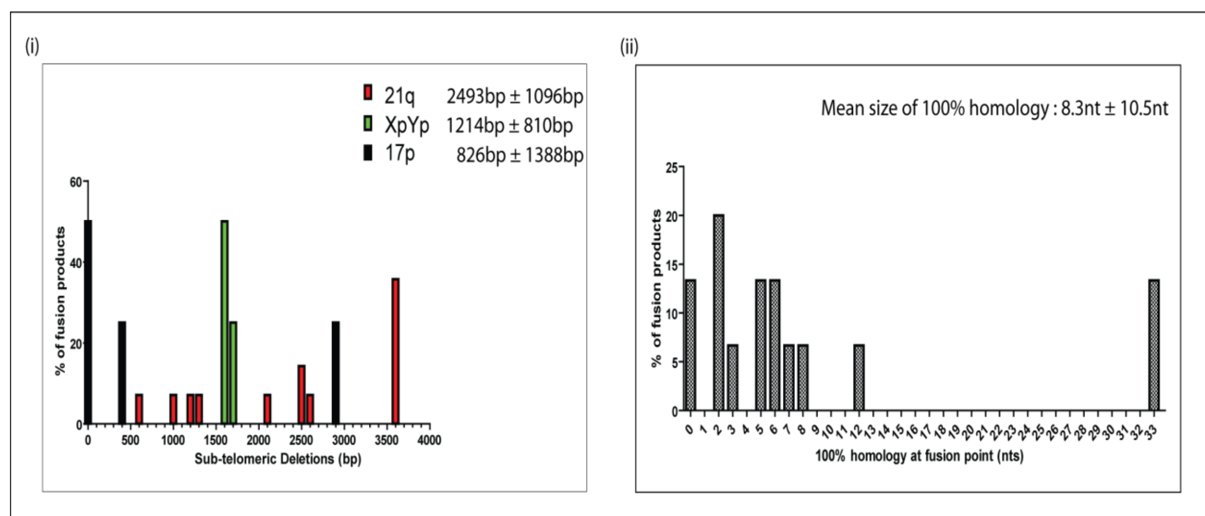


Figure 6.11: (i) A summary of the size of the sub-telomeric deletions at the XpYp and 17p telomere. Deletions at the 21q family of telomeres are also illustrated. (ii) A summary of the size of 100% homology detected at the fusion junction.

Three fusion products (Figure 6.12 and Figure 6.13) revealed an absence of sub-telomeric erosion which was apparent by the presence of telomeric TTAGGG or TVR (TTGGGG, TGAGGG or TCAGGG) repeats adjacent to the fusion junction. This included the clonal product that was detected within an AML patient (Figure 6.12).

Clonal telomere-telomere fusion detected in AML

21q SubTel family Δ 1248bp?
17p SubTel Δ 0bp
21q Sub-telomere family
17p Telomere
Fusion: GAAGCCGGCTTTTCAGATTACTCAGGTTCAATAACCCTAACCCTAACCCTCACCCTCACCCTCACCCT
GAAGCCGGCTTTTCAGATTACTCAGGTTCCACTCGTCTCTGCGCCGCCGGGGACGTGTGTCTCTGCGC
CTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTCACCCTCACCCTCACCCT
Putative Fusion Junction
Telomere Variant Repeats (TVR)
CCCTCA
Telomeric Repeats
CCCTAA

Figure 6.12: A 40% minimum of AML cells in this one individual contained this fusion event involving the telomere of 17p and that of the 21q family. There was no erosion into the sub-telomeric DNA of 17p but a deletion was evident at 21q. TVR and TTAGGG repeats of the 17p telomere were evident within the fusion product. Underlined bases illustrate sequence homology flanking the fusion junction.

Fusion events resulting from the joining of homologous chromosomes or possible sister chromatid fusion, i.e. XpYp:XpYp (Figure 6.13 (i)) and 17p:17p (Figure 6.13 (ii)) had also been observed however, insufficient material at the XpYp fusion meant it was not possible to formally establish which XpYp alleles were involved. Interestingly, the fusion junction highlighted in Figure 6.13 (ii) appears to have occurred by means of nucleotide insertion. It was apparent that the sequence across the junction (highlighted in yellow) was duplicated in the adjacent DNA of one of the participating telomeres (17p telomere highlighted in green). This may have resulted in an increase in the local homology at the fusion site strengthening the synpasis between the two DNA ends.

(i)
XpYpTel-XpYpTel
XpYp SubTel Δ 0bp
XpYp SubTel Δ 1643bp
XpYp Telomere (allele?)
XpYp Sub-telomere (allele?)
Fusion: GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTGTAGATAACACACTGGAAAACCT
GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGG
AACAAAGATTGCTGAAACATTATGTACCTAACAGCGTTACAGGGTGTAGATAACACACTGGAAAACCT
7x Pure TTAGGG repeats to fusion junction
100% Microhomology

(ii)
17pTel-17pTel
17p SubTel Δ 0bp
17p SubTel Δ 2886bp
17p Telomere
17p Sub-telomere
Fusion: TTAGGGTCAGGGTTAGGGTTAGGGTTGGGGTGGCGCGTATTTTGGCAAATCTGGTACAGCAATAACTG
TCCATGCTTTCAGTACAACATACCGGAGTATTTTGTGGGACACTGCAAATCTGGTACAGCAATAACTG
TTAGGGTCAGGGTTAGGGTTAGGGTTGGGGTGGCGCGTATTTTGTTAGGGTTAGGGTTAGGGTTAGGG
(Possibly continuing into pure TTAGGG tract?)
Fusion Junction
TTAGGG + TVR: TCAGGG/TTGGGG

Figure 6.13: Telomere fusion involving either a pair of homologous chromosomes or sister chromatids at (i) XpYp and (ii) 17p. Microhomology is depicted in purple lettering with underlined bases illustrating sequence homology flanking the fusion junction.

Fusion was also observed involving the 21q family of telomeres (Figure 6.14(i) and (ii) however, due to the extensive homology shared by the 21q family of telomeres,[351] only one of the fusion partners in Figure 6.14(ii) could be identified as the 5q telomere.

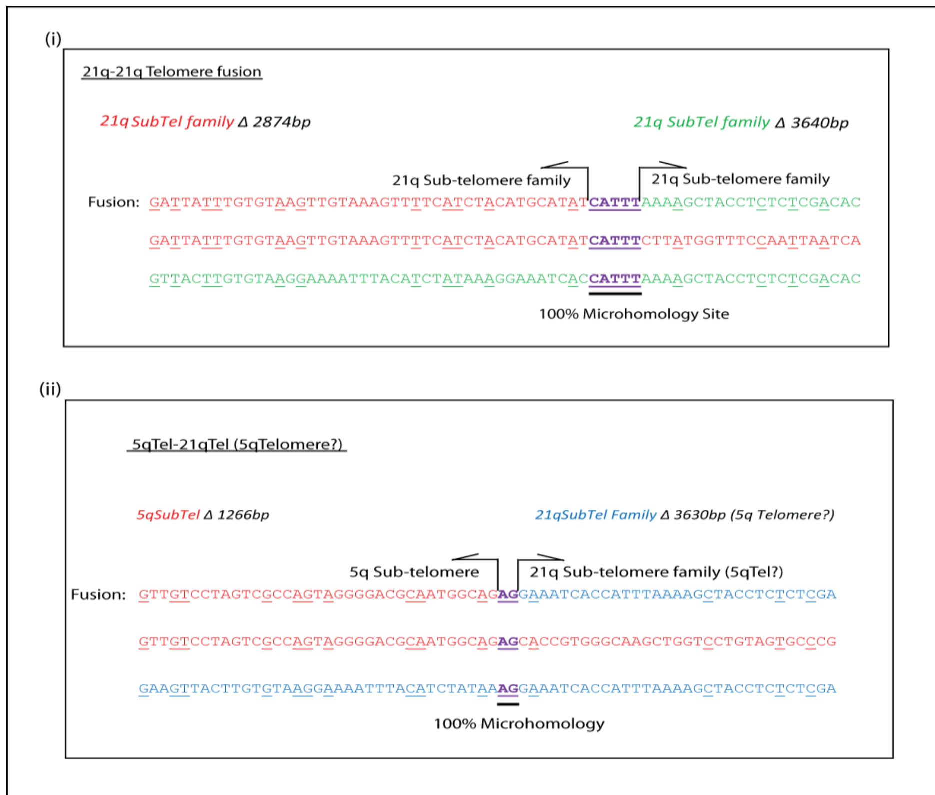


Figure 6.14: Telomere fusion possibly involving a pair of homologous chromosomes or sister chromatids of the 21q family. The 5q telomere could be identified in Figure (ii). Deletions of the 21q family are predicted from specifically utilising the 21q telomere as a reference. Microhomology is depicted in purple lettering with underlined bases illustrating sequence homology flanking the fusion junction.

Sequencing also revealed the existence of complex fusion events involving insertions of non-telomeric genomic loci. It was noted that the mean length of the inserted loci was recorded as 707bp (±248bp). Insertions could be identified as 17q23, 21q11, 5q33 and 19q13 within the fusion products illustrated in Figures 6.15-17 and 19, respectively. Notably, 17q23 and 19q13 correspond to common fragile sites FRA17B[389] and FRA19A,[389] respectively. Common fragile sites have been noted to coincide with chromosomal breakpoints in cancer cells resulting in the possible deletion of tumour suppressor genes and in the generation of Breakage-Fusion-Bridge cycles (BFB).

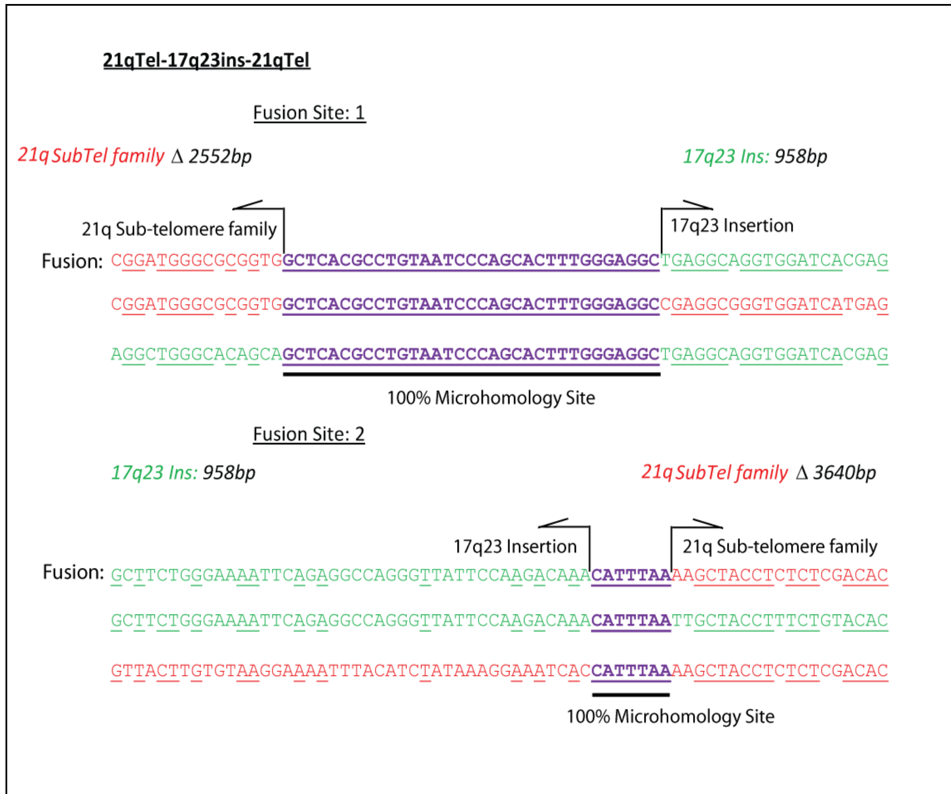


Figure 6.15: A complex fusion event that includes 958bp interstitial non-telomeric DNA mapped to 17q23. Notably, this has been documented as the common fragile site listed as FRA17B. 100% microhomology is depicted in purple lettering with underlined bases illustrating sequence homology flanking the fusion junctions. Deletions of the 21q family are predicted from specifically utilising the 21q telomere as a reference.

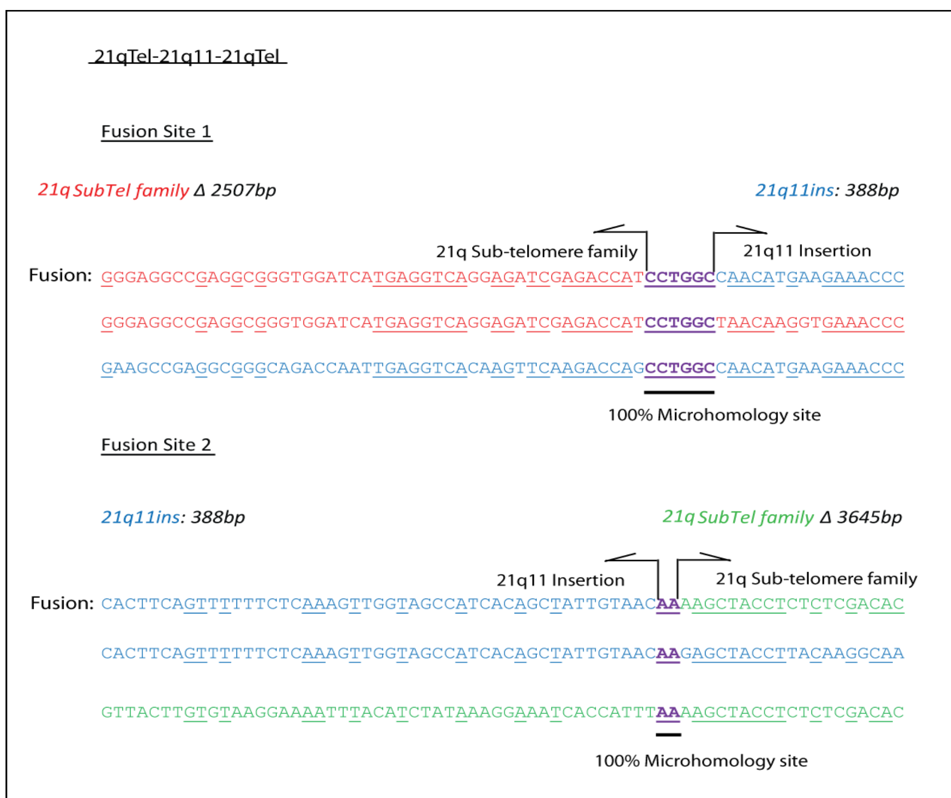


Figure 6.16: A complex fusion event that includes 388bp interstitial non-telomeric DNA mapped to 21q11. 100% microhomology is depicted in purple lettering with underlined bases illustrating sequence homology flanking the fusion junctions. Deletions of the 21q family are predicted from specifically utilising the 21q telomere as a reference.

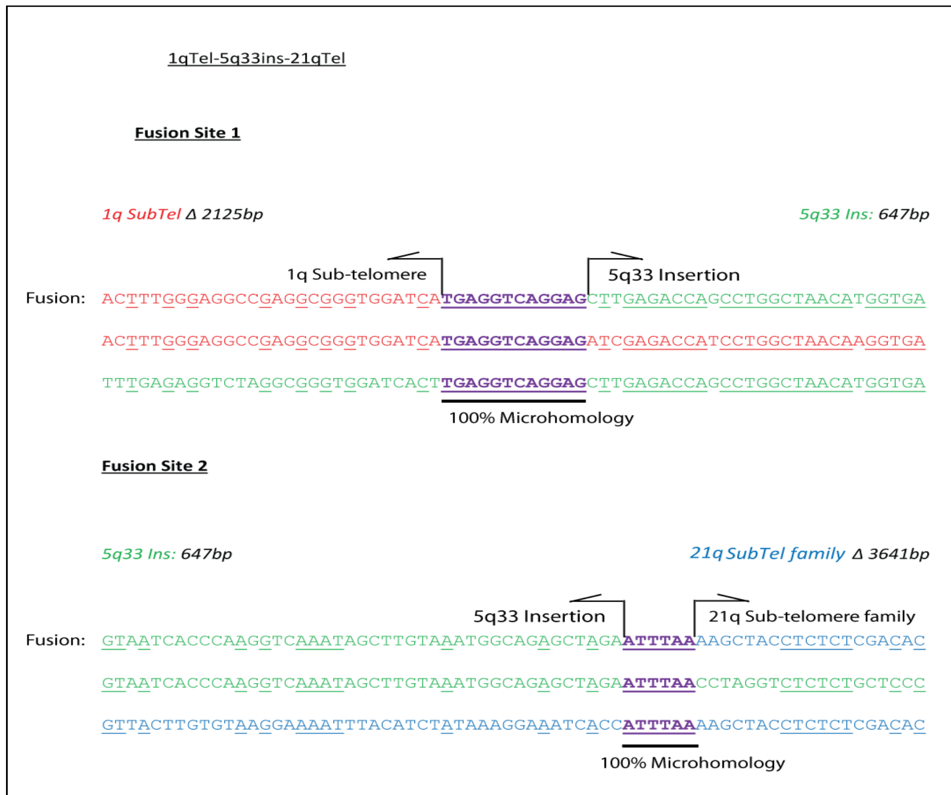


Figure 6.17: A complex fusion event that includes 647bp interstitial non-telomeric DNA mapped to 5q33. 100% microhomology is depicted in purple lettering with underlined bases illustrating sequence homology flanking the fusion junctions. Deletions of the 21q family are predicted from specifically utilising the 21q telomere as a reference.

Direct sequencing revealed a complex fusion product with interstitial loci mapped to 5q33 (Figure 6.17). 5q33 has been implicated as one of the most frequent breakpoints in MDS and AML and is commonly associated with complex chromosomal rearrangements[390,391] thus, it was speculated that it may be a clonal event. However, further insight revealed that despite the documented persistence of breakpoints located at the 5q33 locus[390,391] the fusion event was not detected with a custom assay utilising specific primers across the fusion junctions. This suggests that it was a single sporadic event and not clonal within this patient. This was confirmed by means of gradient PCR (Figure 6.18) in which increasing temperature failed to show the existence of the fusion event within the patient sample.

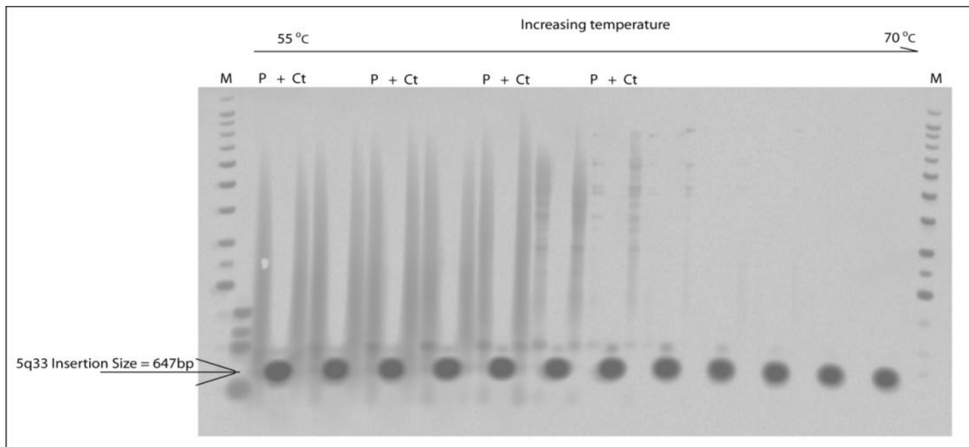


Figure 6.18: Primers that target along the fusion junctions were utilised in order to amplify the inserted sequence (647bp). A gradient PCR was performed to determine whether the fusion involving the 5q33 locus was present in multiple cells. Patient (P) and the 293 cell line (Ct) did not show the fusion product along the temperature gradient. Conversely, the amplified 5q33 fusion product (+) continued to amplify even at higher temperatures. This suggests that the 5q33 fusion event was only sporadic and occurred in a single cell.

One event showed evidence of additional processing at the fusion point (Figure 6.19). It was apparent that further homology (bold orange lines) was apparent downstream of the highlighted A, thus it is possible that the misaligned A was removed by endonucleolytic activity resulting in an increase in the DNA sequence homology at the fusion point and stabilisation of the synapse.

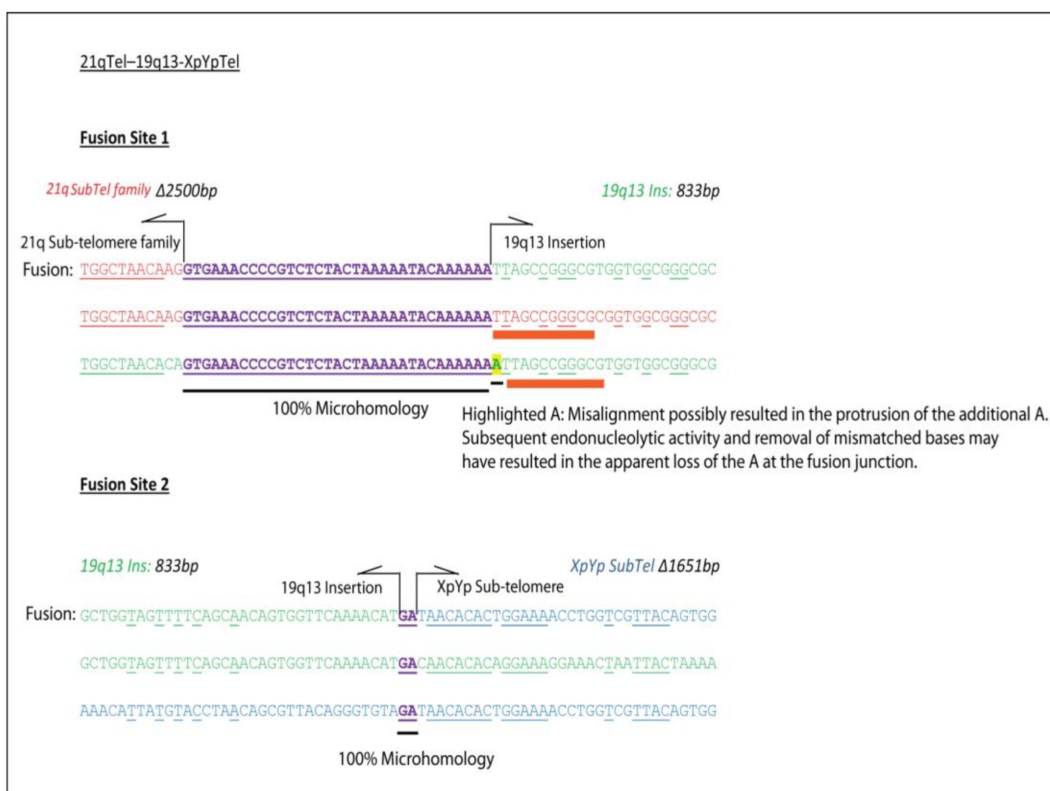


Figure 6.19: A complex fusion event that includes 833bp interstitial non-telomeric DNA mapped to 19q13 documented as the common fragile site listed as FRA19A. 100% microhomology is depicted in purple lettering with underlined bases illustrating sequence homology flanking the fusion junctions. Deletions of the 21q family are predicted from specifically utilising the 21q telomere as a reference. Bold orange lines indicate the homology shared downstream from the highlighted A.

Despite the observation of 100% homology at the fusion junction within the majority of products obtained (n = 9), two events displayed no homology at the junction point (Figures 6.20 (i) and (ii)).

(i)

12q-17p Telomere fusion

12qSubTel Δ 621bp

17pSubTel Δ 416bp

No homology

12q Sub-telomere | 17p Sub-telomere

Fusion: CTTTTTGCACGTTTACTTGCCATAGATCTTCTCTGCGAAAACAGTGGGATCCTTGCACAGGAATAAAC

CTTTTTGCACGTTTACTTGCCATAGATCTTCTCTGCGCCTGCGCCGGCGCGGGGGGCCTTTGCGAGGG

CTCCCAATTTCAAATTTTATTACAAAGGAACAGTAATAAAACAGTGGGATCCTTGCACAGGAATAAAC

(ii)

21qTelFam(1q?)-XpYpTel

XpYp SubTel Δ1563bp

21qSubTel Family (1qTelomere?) Δ1008bp

XpYp Sub-telomere | 21q Sub-telomere family (1qTel?)

Fusion: TTCACCAGCACTCTGAGTACATGAAAAAGGCCAAGATGCAGTGTTGCAGTCGCACAGTCGCGCGCCGC

TTCACCAGCACTCTGAGTACATGAAAAAGGCCAAGATGCTTCTTCAGGGATGAATTTTGCTACTTTTT

GGCGCAGAGACGCACGTCGCTGGGCTGAGGGTGGCGGGGAGTGTTGCAGTCGCACAGTCGCGCGCCGC

Possibly no microhomology at fusion junction?

Figure 6.20: No homology was detected at the fusion junction in (i) involving the 12q telomere and 17p telomere and (ii) involving the XpYp telomere and possibly the 1q telomere.

6.5 The Development of 6q STELA

Interestingly, multiple sequences identified over the course of this study appeared to involve the 6q telomere but unfortunately the sequencing of these putative fusion events was unsuccessful. However, as they were regularly detected it was of interest to examine the telomere dynamics at the 6q telomere.

Single base differences within sub-telomeric regions are utilised by STELA to enable the amplification of specific chromosome ends however, due to the extensive sub-telomeric homology that exists among different chromosomes the development of a STELA specific to the 6q telomere is impeded.

The available telomere-adjacent sequence of 6q did not extend to the start of the telomere repeat array and thus the distance from the published terminus to the start of the pure TTAGGG tract was unknown. However, STELA has long-range capacity of up to 25kb[171] and thus it was speculated that it may be able to detect a 6q specific profile provided that base polymorphisms could be detected within the sequence adjacent to the published terminus.

The 5457 bases proximal to the end of 6q were clustered in order to determine whether any 6q specific bases were present. Sequence alignment (Figure 6.21) indicated that only a single base was present 4070bp proximal to the 6q terminus. This illustrates the extensive homology that exists within sub-telomeric regions, i.e. the published end of 6q shared nearly 100% homology with multiple chromosome ends. A STELA assay was subsequently designed making use of this single polymorphism to determine whether it was able to detect 6q telomeres. Figure 6.21 illustrates a sub-section of the alignment that shows the single base difference (highlighted in yellow).

```
6q          ACAGAAATCGTCAAAAAAAAAAAAAAAAGCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
1p          ACAGAAATCGTCAAAAAAAAAAAA----GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
2q          ACGGAAATTGTCAA--------AAAAAGCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
5q          ACAGAAATCGTCAAAAAAAAAAAA-------GAGATTTCCCATGTAGCCGCAACCTAGTTT
17q         ACAGAAATCGTCAAAAAAAAAAAAAA--GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
2q13        ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
19q         ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
22q         ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
10q         ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
21q         ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
4q          ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
1q          ACAGAAATTGTCAAAAAAAAA-------GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
5p          ACAGAAATTGTC-------AAAAAAAAGCAGATATTTCCCATGTAGCCGCAACCTAGTTT
19p         ACAGAAATTGTC-------AAAAAAAAGCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
16q         ACAGAAATCGTCAAAAAAAAAAAA----GCAGAGATTTCCCATGTAGCCGCAACCTAGTTT
8p          ACAGAAATTGTCAAAAAAAAAAAAAAA-----AGATTTCCCATGTAGCCGCAACCTAGTTT
14q         ACAGAAATTGTC--------AAAAAAAGCAGAAATTTCCCATGTAGCCGCAACCTAGTTT
            **.***** ***                   * ***************************

6q          CCTCTCTTATTAACATATTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
1p          CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
2q          CCTCTCTTATTAACATCTTCTATCAGTGTGTT---TCACATGGCTTATTAATATCTTACA
5q          CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
17q         CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
2q13        CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
19q         CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
22q         CCTCTCTTATTAACATCTCCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
10q         CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
21q         CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
4q          CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
1q          CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
5p          CCTCTCTTATTAACATCTTTTATCAGTGTGTTCCTTCACATGGCTTATTAATATCTTACA
19p         CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
16q         CCTCTCTTATTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
8p          CCTCTCTTATTAACATCTTCTATTAGTGTGTC---TCACATGGCTTATTAATATCTTACA
14q         CCTCTC---TTAACATCTTCTATCAGTGTGTC---TCACATGGCTTATTAATATCTTACA
            ******   *******.*  *** *******    *************************

6q          TAATTTGCCACAGTTAATGAACCAATACTGATAGACTGTTATTAACTAAAGTTCATATTT
1p          TAATTTGCCGCAGTTAATGAACCAATACTGATAGACTGTTATTAACTGAAGTTCATATTT
2q          TAATTTGCCGCAGTTAATGAACCAATACTGATAGACTGTTATTAACTAAAGTTCATATTT
5q          TAATTTGCCGCAGTTAATGAACCAATACTGATAGACTGTTATTAACTAAAGTTCATATTT
17q         TAATTTGCCGCAGTTAATGAACCAATACTGATAGACTGTTATTAACTAAAGTTCATATTT
2q13        TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAACTGAAGTTCATATTT
19q         TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAATTGAAGTTCATATTT
22q         TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAACTGAAGTTCATATTT
10q         TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAACCGAAGTTCATATTT
21q         TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAATTGAAGTTCATATTT
4q          TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAACCGAAGTTCATATTT
1q          TAATTTGTCACAGTTAATGAACCAATACTGATAGACTATTATTAACTGAAGTTCATATTT
5p          CAATTTGTCACAGTTAATGAACCAATACTGATAG----------ACTAAAGTTCATATTT
19p         TAATTTGTCACAGTTAATGAACCAATACTG--AGACTATTATTAACTAAAGTTCATATTT
16q         TAATTTGCCGCAGTTAATGAACCAATACTGATAGACTGTTATTAACTGAAGTTCATATTT
8p          CAATTTGTCACAGTTAATGAACCAATACTGATAG----------ACTAAAGTTCATATTT
14q         TAATTTGTCACAATTAATGAACCAATACTGATAGACTATTATTAACTAAAGTTCATATTT
             ****** *.**.*****************  **          *  .************
```

Figure 6.21: 5457 bases proximal to the end of 6q were clustered in order to determine whether base polymorphisms were present. Sequence alignment indicated that only a single base polymorphism (highlighted in yellow) was present 4070bp proximal to the 6q terminus.

STELA failed to show a product following the use of the 6q specific oligonucleotide with alternate annealing temperatures (Figure 6.22). It was postulated that the pure TTAGGG tract is at a distance from the published end and therefore unable to be amplified by STELA. Thus, in an attempt to develop a 6q specific STELA direct sequencing of the region downstream to its published terminus was performed in order to identify single base polymorphisms that may be located nearer to its terminus.

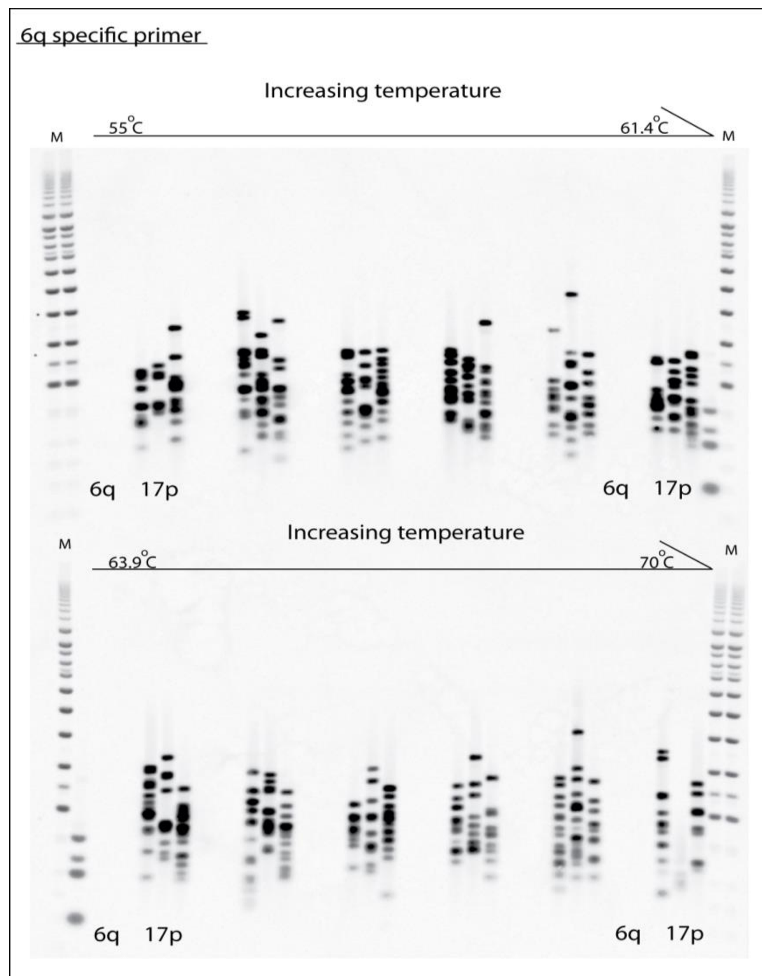


Figure 6.22: A primer specific to 6q failed to show telomeric molecules following a STELA reaction at increasing temperature. 17p presents the positive control for the PCR.

A related sub-telomeric region that had been sequenced to the adjacent telomeric TTAGGG tract was utilised to design oligonucleotides that may be shared by the unidentified 6q terminus. Notably, the sub-telomeric region within the 6q family was required in order to achieve this. To confirm the telomeric location (and not interstitial) of the proposed sub-telomere region, STELA was performed with a TTAGGG adjacent primer that is shared by

multiple chromosome ends within the 6q family. A homogeneous telomere profile was detected amongst 24 individual profiles that was consistent with the amplification of a specific telomere within the 6q family (Figure 6.23). However, the true identity of the telomere detected with this assay is unknown provided that STELA has the capacity to amplify single telomeric molecules of up to 25kb.[171]

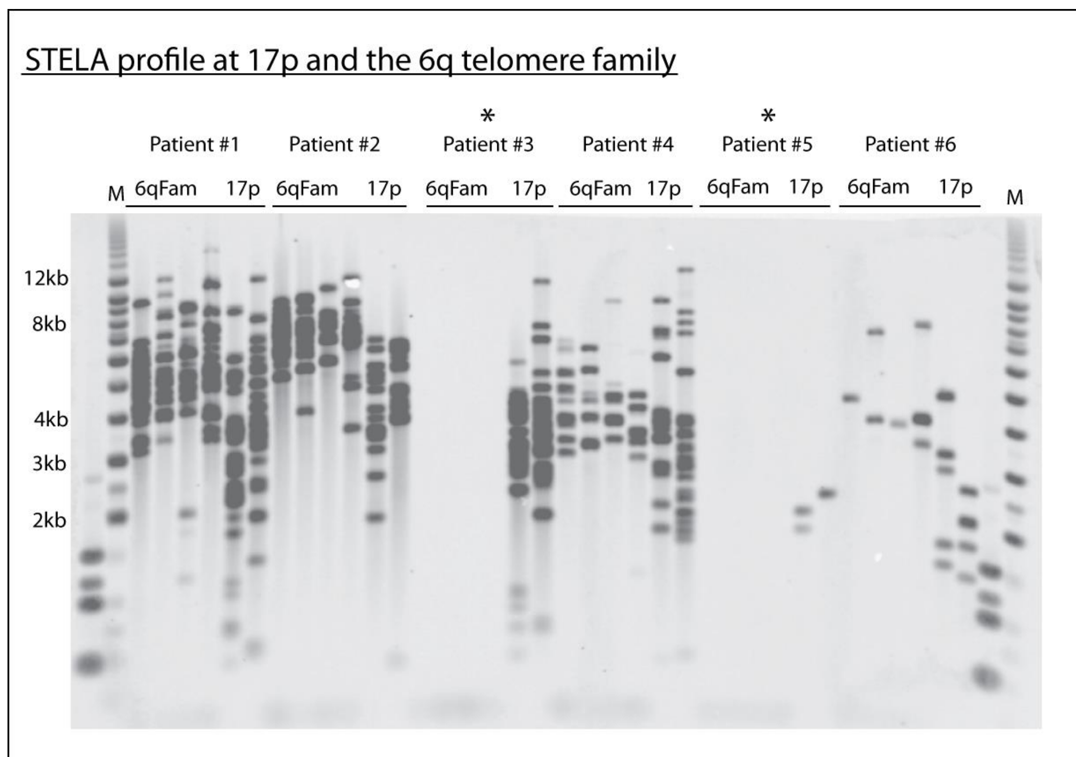


Figure 6.23: Homogeneous telomere profiles were detected upon utilising a 6q family target sequence. Polymorphisms were also identified amongst the population which are presented in patient #3 and #5. The STELA at 17p is a positive control for the PCR reaction.

A 6q specific product was amplified upon using a telomere-adjacent reverse oligonucleotide shared by the family in conjunction with the 6q specific primer previously identified following alignment (Figure 6.21). The presence of a product (Figure 6.24; size ~10kb) following the use of a 6q specific probe suggested that the sequence homology shared by multiple chromosome ends may extend downstream to the published 6q terminus.

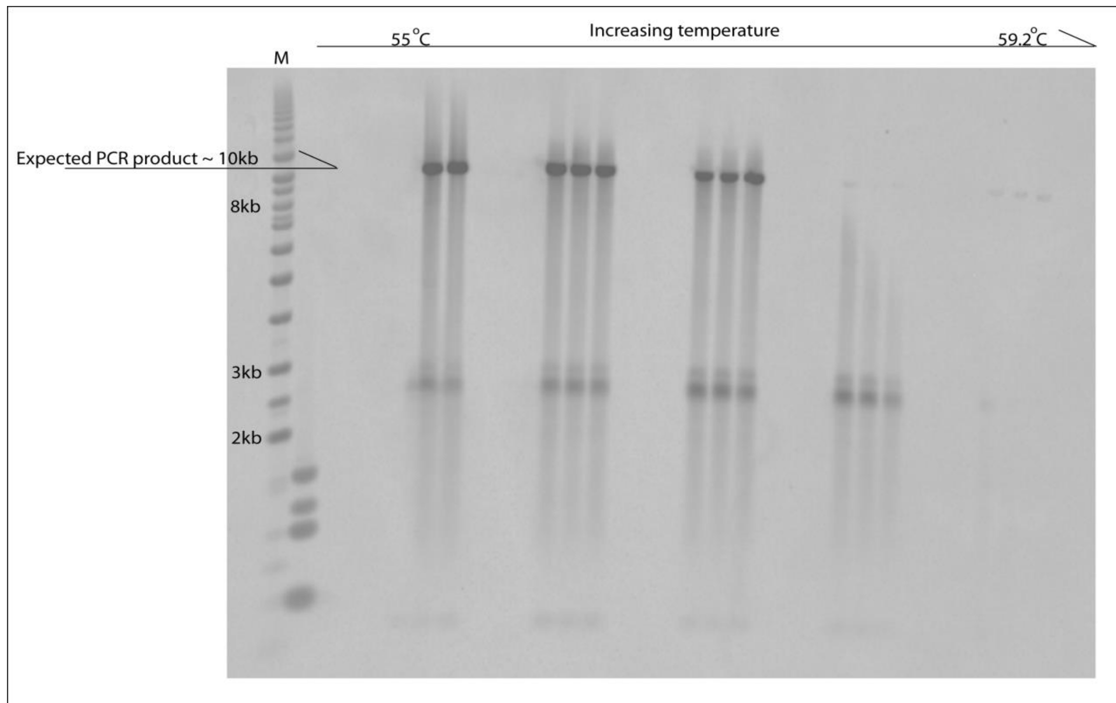


Figure 6.24: The amplification of a 6q specific product (~10kb) by utilising the 6q specific primer previously identified in conjunction with a TTAGGG adjacent reverse primer shared by the family.

In an attempt to identify 6q specific nucleotides, direct sequencing of the 6q specific product was performed by designing oligonucleotides already present within the telomere adjacent region shared by members of the 6q family. Unfortunately no 6q specific sequences could be identified and thus a 6q specific STELA could not be achieved (Sub-section Figure 6.25; Appendix 1).

```
6qPredictedSubTel      GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
10qSubTel              GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
22qSubTel              GGGGTGGCATAGTTTGGCCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
1pSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
5qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
17qSubTel              GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
4qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
19qSubTel              GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
21qSubTel              GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
1qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
2qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
19p'End'               GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACCGGCAATGAAAAGAGTTCTTGTT
8p'End'                GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACCGGCAATGAAAAGAGTTCTTGTT
13qSubTel              GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACCGGCAATGAAAAGAGTTCTTGTT
20qSubTel              ----TGGCATATTCTGGTCTTATACACTGTCTTCCACCAGCAATGAAAAGAGTTATTGTC
12qSubTel              GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACAGGCAACGAAAAGAGTTCTTGTT
7qSubTel               GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACAGGCAACGAAAAGAGTTCTTGTT
                           ******* * *** ************ *** ** .**** **********.****

6qPredictedSubTel      TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
10qSubTel              TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
22qSubTel              TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
1pSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
5qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
17qSubTel              TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
4qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGGAGTTCTAAGAGATATAGACCAG
19qSubTel              TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
21qSubTel              TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
1qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
2qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
19p'End'               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGTTTAGCAGTTCTAAGAGATATAGACCAG
8p'End'                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGTTTAGCAGTTCTAAGAGATATAGACCAG
13qSubTel              TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGTTTAGCAGTTCTAAGAGATACAGACCAG
20qSubTel              TTTCTTGCAGCAATTTGTGATTT-TTTTAGAGTTTAGCAGTTCTAATAGATATAGACCAG
12qSubTel              TTTCCTGCAGCAATTTGTCATTTTTAAAAGAGCTTAGCAGTTCTAAGAGCTATAGAGTAG
7qSubTel               TTTCCTGCAGCAATTTGTCATTTTTAAAAGAGCTTAGCAGTTCTAAGAGCTATAGAGTAG
                       **** * *********** **** *:::**** **** ******** **.** ***  **

6qPredictedSubTel      CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
10qSubTel              CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTAAGCATCTTTTTGTAGGTGT
22qSubTel              CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
1pSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
5qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
17qSubTel              CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
4qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
19qSubTel              CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
21qSubTel              CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
1qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
2qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
19p'End'               CTGTGCTATCTTTTTGTGGTTTTTAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
8p'End'                CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
13qSubTel              CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
20qSubTel              CTGTGCTATCTCC---TGGTTTTCAGTTCTGTAGTATGTTGAGCATCTTTTTGTATGTTT
12qSubTel              CTGTGCTATCTCATTGTGGTTTTCAATTCTCTAGTATGTTGAGCATCTTTTTGCACGTTT
7qSubTel               CTGTGCTATCTCATTGTGGTTTTCAATTCTCTAGTATGTTGAGCGTCTTTTTGCACGTTT
                       ***********     ******* *.**** *********.***.******** * ** *
```

Figure 6.25: Direct sequencing of the 6q specific product revealed that a 6q specific STELA could not be achieved provided that single base polymorphisms were not present within the 6kb region succeeding its published terminus (Full sequence cluster in Appendix 1).

6.6 Array-Comparative Genomic Hybridisation

Whole genome tiling using NimbleGen 130k-array CGH (aCGH) was carried out on 12 MDS samples (Figure 6.27) to determine whether telomere dysfunction correlated with genomic instability. Telomere dysfunction was assumed by the presence of fusion events at the XpYp and 17p telomere.

Genomic alterations in the form of chromosome aneuploidy (Trisomy 8; patient #1, #12 and Monosomy 7; patient #9) and gains and losses (patients #5 and #8) were detected using aCGH analysis. Patients #5 and #8 displayed a complex karyotype (Figure 6.27). Compared to those in which no large genomic lesions could be detected; the two individuals with a complex karyotype showed shorter telomeres (Figure 6.26). Accordingly, the mean telomere length recorded within the cohorts presenting with a normal and complex karyotype was 4.71kb (±1.63kb) and 2.54kb (±0.70kb), respectively. However, this did not reach statistical significance ($p = 0.1200$). In addition, individuals presenting with chromosomal aneuploidy, i.e. Trisomy 8 or Monosomy 7 displayed longer telomeres than those with complex rearrangements. Notably, the telomeric mean recorded was 5.31kb (±2.49kb) but again the difference was not significant ($p = 0.2393$).

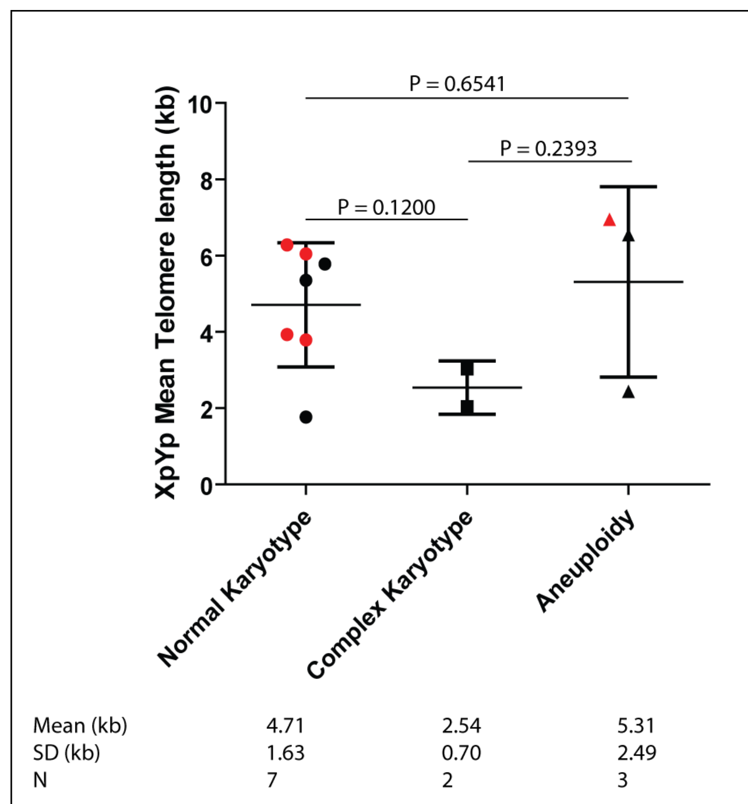
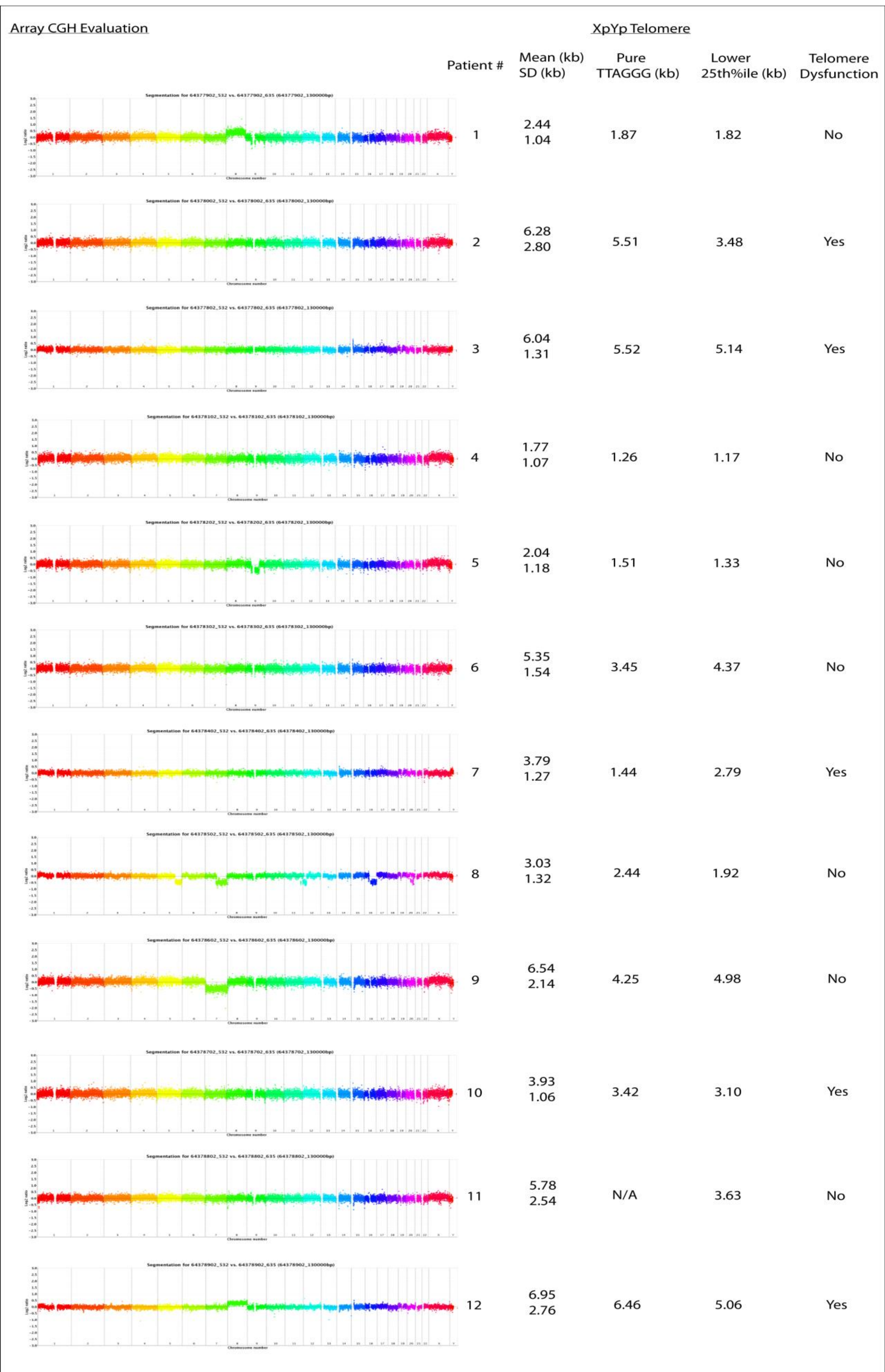


Figure 6.26: Although not significant, telomere length appeared shorter within patients presenting with a complex karyotype. Fusion events were evident in patients highlighted in red.

Telomere fusions were particularly apparent within individuals that presented with a normal karyotype. Telomere dysfunction may initiate early events for neoplastic development and predispose to the evolution of gross chromosomal rearrangements that were apparent within those showing complex karyotypes. Higher resolution may have revealed cryptic alterations that had failed detection upon utilising 130000 probes or events that were not fully clonal in the population and therefore not detectable with aCGH. It is plausible to suggest that telomere fusion events were not observed within those individuals presenting a complex karyotype provided that telomere dysfunction is an early event and thus the fusion assay may be unable to detect subsequent rearrangements.

These data indicate the possibility that telomere length is associated with karyotypic complexity. Further analysis using a larger cohort may provide a more detailed account of whether telomere dysfunction accompanies gross chromosomal rearrangements in MDS/AML.


Array CGH Evaluation

XpYp Telomere

Patient #	Mean (kb) SD (kb)	Pure TTAGGG (kb)	Lower 25th%ile (kb)	Telomere Dysfunction
1	2.44 1.04	1.87	1.82	No
2	6.28 2.80	5.51	3.48	Yes
3	6.04 1.31	5.52	5.14	Yes
4	1.77 1.07	1.26	1.17	No
5	2.04 1.18	1.51	1.33	No
6	5.35 1.54	3.45	4.37	No
7	3.79 1.27	1.44	2.79	Yes
8	3.03 1.32	2.44	1.92	No
9	6.54 2.14	4.25	4.98	No
10	3.93 1.06	3.42	3.10	Yes
11	5.78 2.54	N/A	3.63	No
12	6.95 2.76	6.46	5.06	Yes

Figure 6.27: Array-CGH was carried out on 12 MDS patients. Gross chromosomal rearrangements were evident in patient #8 that also showed extensive loss that extended to the telomeres.

6.7 Discussion

Short telomeres have been observed to induce genetic instability by means of entering into breakage-fusion-bridge (BFB) cycles. Dysfunctional telomeres are recognised as DNA double strand breaks (DSBs) which can be inadvertently repaired by means of chromosomal fusion with an alternate recombinogenic free end. Aberrant chromosomal fusion results in the production of dicentric chromosomes which are subsequently pulled apart to opposite poles during mitosis.[319] This has the propensity to cause a break along the chromatin bridge and generate new recombinogenic ends. Further chromosomal fusion and breakage can result in a wide spectrum of non-reciprocal translocations (NRTs) that may have the ability to drive neoplastic transformation. This paradigm has been considered as a mechanism to induce genetic instability in many human solid tumours including pancreatic carcinoma and osteosarcomas[321] as well as in human leukaemias such as Chronic Lymphocytic Leukaemia (CLL).[304]

There was no obvious relationship between telomere length and the presence of telomere fusion events. However, long telomere STELA profiles may be superimposed by large scale telomere deletion events (TRDs) that can result in a truncated telomere which has the propensity to enter into fusion[171,352] with other chromosome ends or non-telomeric loci. The low frequency of fusion events was consistent with sporadic fusion arising as a consequence of TRD events. TRDs were detected within the MDS and AML cohorts (Chapter 3; Section 3.7) part and there appeared to be a significant reduction in the mean frequency at 17p within the AML cohort when compared to those in MDS. Consistently, there appeared to be a significant reduction in the frequency of fusion events involving the 17p telomere within the AML cohort. Moreover, the tendency for telomere fusion at 17p has the propensity to induce loss of heterozygosity (LOH) at the *p53* locus which may enable the progression to overt leukaemia. Notably, the extensive telomere deletion at a single telomere has the propensity to initiate genetic instability and play a role in the neoplastic transformation of normal cells.

The reduction in the frequency of fusion events may be related to the up-regulation of telomerase activity in AML cells. MDS may represent the early episodic phase of telomere induced genetic instability that contributes to the development of AML by means of

promoting the generation of non-reciprocal translocations (NRTs). Telomerase activity[342] and *hTERT* expression[347] are more pronounced in AML with complex rearrangements; it may be speculated that the reduction in telomeric fusion in AML cells could be associated with the up-regulation of telomerase and stabilisation of telomere length to enable the transition to a more stable genome.

A putative clonal telomere adjacent fusion product was detected within an AML patient involving a telomere belonging to the 21q family and that of 17p. The persistence of this fusion event suggests that it was maintained throughout the cell cycle and has been propagated to daughter cells. This could arise through the inactivation of a single centromere on the dicentric chromosome. Alternatively this fusion event may have resulted in large-scale rearrangements that led to the loss of key tumour suppressors or amplification of oncogenes and clonal evolution. Provided that these were not purified cells it may be speculated that a large fraction of the AML clone consisted of this telomeric fusion.

The amplification and characterisation of chromosomal fusions arising from telomere dysfunction had revealed that fusion junctions are processed in a manner reminiscent of Ku-independent Alternative-Non-Homologous End Joining (A-NHEJ).

A-NHEJ has been implicated to utilise Poly (ADP-ribose) Polymerase I [PARP-1] as well as the DNA ligase III/ XRCC1 complex[392] and has been previously implicated in DNA double strand break repair in both yeast[280] and hamster cell lines.[298] Furthermore, A-NHEJ has also been implicated in the production of telomere fusion products in human cells undergoing crisis *in vitro*[255] and within tumour cells.[304]

Chromosomal loss has been observed to flank the fusion junction site which is composed of a small region of perfect overlapping homology. Consistent with this, the current data show that telomere fusion partners align at short patches of 100% microhomology ranging in length from 2 to 33nts following the sub-telomeric deletion at one or both telomeres. Extensive deletion was apparent in several cases suggesting the possibility that exonucleolytic resection could extend beyond the proximal limit of the assay for which is 1655bp at XpYp and 3058bp at 17p. It has been proposed that the extent of 5’ to 3’ exonuclease activity may be dependent on the compatibility of DNA ends. If microhomology

is exposed during the process of strand resection further degradation is inhibited and base pairing ensues.[393] Mre11 has been proposed to play a role in the detection of microhomology and it has been suggested that it may function in conjunction with a 5' to 3' exonuclease[255] in order to reveal sites of microhomology which can be utilised for strand alignment.

One fusion event revealed a small inserted DNA tract at the fusion junction which appeared to arise from a duplication of the adjacent DNA from one of the participating telomeres. This has been previously observed at the telomere fusion junction within ATLD (Ataxia Telangiectasia-Like Disorder) cells in which hypomorphic mutations of Mre11 have been proposed in preventing the efficient synapsis of the two DNA ends.[305] Templated nucleotides have also been detected in *Drosophila* mutants deficient in Rad51 and DNA ligase IV.[307] It has been proposed that these duplicated regions may play a role in increasing the local homology and stability for DNA end joining.[307] Moreover, an identified fusion product appeared to have been misaligned prior to successful joining. These events may be indicative to additional processing at the fusion point, for example via the removal of specific nucleotides resulting in an increase in the localised DNA sequence homology and an increase in the stability of DNA synapse.

Fusion products were identified that presented TTAGGG and Telomere Variant Repeats (TVR) adjacent to the fusion junction. This is consistent with previous implications that the TVR region is incapable of providing termini protection.[255] Notably, it has been previously documented that only 13 TTAGGG repeats is sufficient to initiate telomere fusion[255] possibly due to the insufficient binding of Shelterin to maintain the terminal cap.

Sequencing also revealed the existence of complex fusion events involving insertions of non-telomeric genomic loci. Insertions could be identified as 17q23, 21q11, 5q33 and 19q13. 17q23 and 19q13 correspond to common fragile sites FRA17B[389] and FRA19A,[389] respectively. Not only have common fragile sites been noted to coincide with cancer breakpoints, interstitial TTAGGG and TVR repeats within fusion products may also promote chromosomal fragility by means of generating G-quadruplex structures. These secondary structures have the propensity to induce stalling of the replication fork[221] and potentially result in the production of a double strand break.

Although 100% homology was detected at the fusion junction in the majority of fusions identified, two products showed no homology at the fusion junction. This may be suggestive of DSB repair by means of the classical Ku dependent Non-Homologous End Joining (C-NHEJ).

Deletions involving the long arm of chromosome 6 (6q) have been detected in various solid tumours and haematological disorders including Acute Lymphocytic Leukaemia (ALL),[394] Non-Hodgkin's Lymphoma[395] and secondary Acute Myeloid Leukaemia (AML).[129] Molecular analysis has detected two regions of minimal deletion (RMD) that occur between 6q21-q23 and 6q25-q27[396,397] suggesting the presence of candidate tumour suppressor genes within these regions. Notably, *SEN6*; a cellular senescence gene has been previously mapped to 6q27.[398] Furthermore, 6q has been noted to harbour multiple common fragile sites which include FRA6F and FRA6E localised to 6q21 and 6q26, respectively.[389] Due to the prevalence of 6q deletions it was of interest to determine whether telomere dysfunction contributed to the loss of genetic material on 6q. However, direct sequencing revealed that a 6q specific STELA could not be achieved since single base polymorphisms were not available within the region analysed. Homologous repeat tracts (90 to >99.5%) of sub-telomeric DNA have been identified at multiple chromosome ends which have been observed to extend for up to 200kb in humans.[241,247] The extensive sub-telomeric homology hinders the development of a chromosome specific STELA as a specific target sequence encompassing a single base polymorphism is required.

The data provided from array-CGH analysis demonstrated that telomere dysfunction may contribute to chromosomal instability that can enable disease progression and AML transformation. This was particularly emphasised in a single individual that presented gross chromosomal loss which extended to the telomeres. Telomere dysfunction can result in an accumulation of non-reciprocal translocations[322] producing gross chromosomal rearrangements as detected upon this array profile. Interestingly, whilst not statistically significant, telomere length appeared to be related to karyotypic complexity. This is consistent with previous reports which show telomere shortening coupled with complex chromosomal rearrangements.[333,334,342] MDS samples with stable genomes displayed longer telomeres than those with complex karyotypes, characterised by gains and losses that

included the telomeres. Moreover, telomere length was not related to the presence of trisomy or monosomy, indicating that mechanisms distinct from telomere dysfunction may be implicated in this process, i.e. sister chromatid non-disjunction.[399,400] However, it was also observed that patients presenting with a normal karyotype in conjunction with long STELA distributions exhibited telomere fusion events. This may be consistent with stochastic truncation events contributing to the development of a complex karyotype. It is possible that the presence of telomere fusion events is an early event in the neoplastic process which may confer an inferior prognosis following the generation of a complex pathological genomic profile.

The data shown in this chapter indicate that telomere dysfunction may contribute to the progression of MDS and AML via telomere fusion. The resultant breakage-fusion-bridge (BFB) cycles have the propensity to generate gross chromosomal rearrangements that may be detected by means of CGH analysis.

Chapter 7:

Discussion

The Myelodysplastic Syndromes (MDS) and Acute Myeloid Leukaemia (AML) each represent a group of heterogeneous haematological diseases characterised by the clonal proliferation of undifferentiated myeloid cells.[18,367] According to the WHO criteria, a dysplastic marrow presents with dysplasia in at least 10% of cells derived from a single myeloid lineage.[32] Myelodysplasia carries a 30% risk of developing Acute Myeloid Leukaemia,[19] defined by the clonal expansion (≥20%) of myeloblasts in the bone marrow or peripheral blood.[20] AML can also arise *de novo* or follow prior chemotherapy or radiotherapy for a non-haematological disorder.[53] Bone marrow failure in MDS and AML induces peripheral blood cytopenia(s)[15] defined by an insufficient production of erythrocytes, granulocytes and/or platelets increasing an individual's susceptibility to anaemia, infection or to internal haemorrhaging, respectively.[26]

Telomeres enable cells to distinguish their natural chromosome ends from double strand breaks (DSBs) in order to maintain genomic integrity and prevent premature senescence.[190] Telomeres retain this function by means of providing a specialised nucleoprotein 'cap' that prevents the chromosome terminus from initiating DNA repair pathways.[258,401] A strong correlation between telomere length and cellular proliferative capacity has been documented[164] with progressive telomere shortening associated with a concomitant reduction in a cell's proliferative potential. However, previous studies have implicated that the shortest telomere within a distribution is vital for cell viability and chromosome stability[209] and thus single short telomeres generated by sporadic telomere rapid deletion (TRD) may induce premature cell cycle arrest or initiate cycles of Breakage-Fusion-Bridge (BFB) creating gross chromosomal rearrangements.[255,319,351]

Haematopoietic CD34$^+$ cells are telomerase competent; however they lose up to 33bp of telomere repeats per year.[325,326] Several studies have shown that telomere length is significantly shorter in various myeloid disorders including the Myelodysplastic Syndromes (MDS) and Acute Myeloid Leukaemia (AML),[336,342,346,347,349,350] Aplastic (AA)[337,338] and

Fanconi (FA) Anaemia[402] and Chronic Myeloid Leukaemia (CML).[339,340] Moreover, the physiological consequences of telomere shortening have been analysed in $mTR^{-/-}$ murine models in which late generation animals exhibit defects in highly proliferative tissues including the haematopoietic system.[330] Telomere shortening is also associated with disease anticipation in Dyskeratosis Congenita (DC) in which the proliferative potential of haematopoietic stem cells is compromised with each generation due to insufficient telomerase activity.[329] Such patients with DC show a 196-fold increase in AML progression.[331] It has been speculated that telomere dysfunction may be in part accountable for disease progression and neoplastic transformation. Supporting this theory is the observation of numerous end-to-end fusions and signal free ends in epithelial cancers derived from late generation $mTR^{-/-}p53^{-/-}$ mutant mice.[322] In the presence of up-regulated telomerase activity, the frequency of end-to-end fusions is reduced[212] so as to facilitate the outgrowth of a sub-clonal population.

Owing to the heterogeneous clinical features and outcome of MDS and AML, prognostic scoring systems including the IPSS/IPSS-R[43,44] and Hill's Risk Score generated from data derived from the Medical Research Council (MRC) AML 10 and 12 trials are used to help predict patient outlook and facilitate in making therapeutic decisions. These prognostic systems are particularly informative when referring to high-risk individuals who are more likely to benefit from a bone marrow transplant in contrast to low-risk patients for whom supportive care or conventional treatment would be more appropriate.[50,58,370] The IPSS calculates a risk score from cytogenetics, blast percentage and number of cytopenia(s) to predict overall survival and risk of AML transformation.[43] In contrast, the Hill's Risk score takes into account variable clinical parameters including age, presenting white blood cell (WBC) count, cytogenetics, AML type (*de novo*/ secondary) as well as the patient's response status after the first cycle of induction chemotherapy (Complete Remission [CR]/Partial Remission [PR]/Resistant Disease [RD]). AML patients are subsequently categorised into subgroups based on their response to treatment and/or risk of relapse after the first cycle of induction chemotherapy. However, further refinement of the current prognostic scores would improve decisions in therapeutic intervention, particularly due to the heterogeneous outcomes of individual patients.[18,25,103,104,367]

High resolution methodology was utilised to analyse telomere length in MDS and AML. Single Telomere Length Analysis (STELA) is a long-range single molecule PCR approach that has the ability to amplify the double-stranded region of telomere repeats from specific chromosome ends.[171] STELA has the ability to detect the full spectrum of telomere lengths and is able to identify critically short telomeres which have the potential to induce replicative senescence or initiate cycles of telomere fusion and breakage.[255,304,351] In this study, telomere length was analysed using STELA at XpYp and 17p to determine the nature of telomere dynamics in a cohort of 80 MDS and 144 AML patients. Telomeric features specific to a single chromosome were identified with STELA including bimodal distributions, Telomere Rapid Deletion (TRD) events and Telomeric-Loss of Heterozygosity (Telomeric-LOH); features that would not be readily detected or occult to other available assays such as Terminal Restriction Fragment (TRF) analysis, Quantitative-Fluorescence *in situ* Hybridisation (Q-FISH), Flow-FISH or Quantitative telomere-specific Polymerase Chain Reaction (Q-PCR).

7.1 Telomere Length and Intra-Clonal Variation in MDS/ AML

This study showed a strong correlation between the telomere length at XpYp and 17p within the MDS and AML cohorts. As previously reported,[332-334] the current data showed a significant reduction in telomere length among MDS and AML patients when compared to aged-matched healthy individuals. However, this study as well as others,[335,336] showed significantly shorter telomeres in AML when compared to MDS. It was also apparent that telomere length distributions were significantly more homogeneous in AML profiles suggestive of clonal expansion. This has been described in Chronic Lymphocytic Leukaemia (CLL) such that telomere shortening and homogenisation are concurrent with disease progression and severity.[304] Heterogeneity of individual telomere length distributions could be identified on several STELA profiles amongst the patients analysed. Such variation could be characterised by differential maternal and paternal contributions in the zygote of patients heterozygous at the XpYp sub-telomeric region. However, it was conceivable that recombination at a specific telomere may be accountable for the dynamics at the 17p telomere in an MDS patient. This STELA profile was characterised with a small population of long telomeric molecules and a greater population of short telomeres, consequently generating a sub-clonal population.

7.2 Telomere Length and Age in MDS/ AML

In the MDS cohort, telomere length conformed only weakly to the aging dogma with several patients showing extensive telomere attrition relative to the telomere length of healthy individuals of a comparative age. One extreme case was that of a patient who at the age of 37 years presented with a telomere length profile equivalent to a healthy individual of over 100 years. Furthermore, assuming haematopoietic cells lose 33bp each year,[325,326] the current study revealed a 10-fold increase in the rate of telomere erosion in an MDS patient. This may reflect the stressful conditions that are endured by the haematopoietic system in an attempt to regain haematopoietic reconstitution. Accordingly, the fraction of proliferating LT-HSCs (Long-Term mitotic capacity-HSCs) has been demonstrated to increase significantly with age *in vivo*[11] under conditions where haematological stress may be prevalent. Moreover, a transient phase of rapid telomere shortening has been observed following an allogeneic bone marrow transplant prior to stabilisation and haematopoietic reconstitution.[327]

In contrast, telomere length increased with age in AML with significantly longer telomeres in patients >60 years in contrast to their younger counterparts (≤60 years). It is possible that this may reflect a bias towards the incidence of FAB (French-American-British) M1 cases in the elderly population. A large proportion of older AML cases present with poorly differentiated subtypes (FAB: M0/M1) with the probability of these AML subtypes increasing with age.[375] Telomerase activity is correlated with the FAB AML subgroups with activity following M1>M2>M5>M4[376] with a significant increase in M1[377] and significantly lower activity in M0 and M3 when compared to the other subgroups.[403,404] Therefore, it might be speculated that the elevated incidence of longer telomeres with age might be associated with a bias towards M1 cases. Unfortunately the degree of differentiation was unavailable for this study and therefore this cannot be concluded. However, it might have been informative to compare telomere length and telomerase activity in AML cells showing differential levels of haematopoietic differentiation.

7.3 Clinical Parameters in MDS/ AML

A diagnosis of a blood cytopenia can be made when the peripheral blood presents with a reduction in cell count derived from a specific lineage.[24] The development of cytopenia in the early stages of MDS is the result of elevated apoptosis of differentiating cells in the marrow.[22,139] This is in contrast to late-stage MDS and AML which show a reduction in apoptosis but a block in haematopoietic differentiation and myeloblast expansion.[141] The International Prognostic Scoring System (IPSS) defines anaemia, neutropenia and thrombocytopenia by a haemoglobin level of under 10g/dl, an absolute neutrophil count (ANC) of less than 1800/μl and a platelet count of less than 100,000/μl.[43] In this study, the prognosis of MDS patients presenting with multiple cytopenias was poor and showed significantly reduced overall survival when compared with individuals enduring uni-lineage cytopenia.

It was speculated that telomere shortening in haematopoietic cells would arise under conditions of haematopoietic stress to account for the deficits in peripheral blood count, i.e. compensatory cell divisions of stem cell reserves. Moreover, telomere shortening in itself might contribute to the development of cytopenia by telomere-induced cell cycle arrest. In the current study, uni-variate analysis failed to show a relationship between telomere length and multiple cytopenia(s). This is in contrast to previous reports which have identified a significant reduction of telomere length in MDS patients with multiple cytopenias.[332] It might have been informative to establish whether telomere length was associated with the intensity of haematopoietic insufficiency along a specific lineage, particularly since the IPSS score only weighs the number of cytopenia(s) present.[43] Interestingly, it has been observed *in vitro* that the telomere length of cord blood derived $CD34^+$ cells is a determinant of erythroid proliferative potential. This is in contrast to other myeloid lineages of which no relationship between telomere length and the production of granulocytes, megakaryocytes or monocytes could be found.[405] This raises speculation that telomere dynamics may differ in cells derived from patients presenting with different lineage cytopenia and therefore telomere loss may not be as extensive in such cases.

Cytogenetics is an independent predictor of patient outcome for MDS and AML patients such that poor cytogenetic profiles are associated with an inferior outlook.[106,368,370,371]

Telomere shortening has been observed to accompany complex chromosomal rearrangements in a variety of haematological disorders, including acquired Aplastic Anaemia (AA),[337,338] Chronic Myeloid Leukaemia (CML)[340] and MDS/AML,[333,334,342] however, in this study telomere length was not associated with cytogenetics in MDS. It is possible that telomere dysfunction and subsequent BFB cycles occurred early in a subset of patients presenting with an adverse karyotype. The preferential up-regulation of telomerase would stabilise telomere length and enable clonal expansion allowing for its subsequent detection using conventional G-banding. Supporting this theory was that loss up to the telomere was identified in an array-CGH profile derived from an MDS patient presenting with losses and gains. Telomere length was not associated with cytogenetic complexity in AML; however, the data set was limited for patients who presented with a favourable or adverse karyotype. Nonetheless, it may be speculated that telomere dysfunction also occurred early in AML cells which later present with an adverse karyotype at diagnosis.

7.4 Telomere Rapid Deletion and Telomere Fusion

Telomere length distributions in MDS and AML were superimposed by sporadic, atypical large-scale telomere rapid deletion events (TRDs).[255,352] Such events may play a significant role in cytopenia development and karyotypic complexity found in a subset of MDS and AML patients since the shortest telomere in a cell has been implicated in inducing premature arrest and genetic instability.[209] It is possible that TRDs are generated as a consequence of increased Reactive Oxygen Species (ROS) or replicative stress endured by the haematopoietic system under MDS/AML pathological conditions. In this study a greater frequency of TRD events were detected at the XpYp and 17p telomere in MDS when compared to AML with a significantly greater frequency at 17p in MDS. It should be taken into account that only the XpYp and 17p telomere were analysed over the course of this study and therefore the actual incidence of TRDs is unknown. The probability of this event happening within a single cell could be considerable since 92 telomeres are susceptible to this process and possibly contribute greatly to the development of these diseases.

To examine the extent of telomere dysfunction in MDS and AML, telomere fusion events were quantified at the XpYp and 17p telomere using a PCR based fusion assay.[255] The modest amount of telomere fusion detected within the MDS and AML cohort appeared

consistent with sporadic fusion arising as a consequence of telomere rapid deletion. MDS patients appeared to show a higher frequency of telomere fusion when compared to the AML cohort. Moreover, a significantly greater percentage of TRD events at 17p appeared to accompany a significantly greater number of fusion events involving the 17p telomere within MDS patients. It is possible this was associated with the targeted abrogation of *p53* and development of the pathological clone in MDS. The prevalence of *p53* mutations has been extensively studied in MDS[133-136] and patients presenting with the complete abrogation of *p53* show a significantly shorter overall survival, increased propensity to leukaemic progression and inferior response to chemotherapy.[135]

The incidence of TRD events might be a predisposing factor to AML transformation in MDS patients prior to telomerase up-regulation, particularly since telomerase activity is low in MDS cells[332] and may be insufficient in preventing an accumulation of TRD events. This study as well as others[328,332,342,347] has reported telomerase up-regulation in AML cells which could be accountable for the reduction in TRDs and fusion events found among the AML cohort. Notably, up-regulated telomerase activity would enable unlimited proliferation by maintaining telomere stability[212,224] and reduce the frequency of DSBs by adding TTAGGG repeats *de novo* onto broken ends.[186]

7.5 Pure TTAGGG and Telomere Dysfunction

STELA profiles provide an overestimate of the pure telomeric length due to variable measurements of Telomere Variant Repeats (TVR). The TVR is a non-functional region[192,255] within the proximal end of the telomere composed of an interspersion pattern of TTAGGG and variant repeats including TCAGGG and TGAGGG,[242-244] which in this study ranged from 0 to 3kb. Individuals had been identified as having extreme telomere shortening following the correction for the TVR region. An AML patient presented with a mean STELA profile of 2.20kb but an uninterrupted TTAGGG tract with a mean of only 0.16kb or 26.7 TTAGGG repeats. This is shorter than what has been previously identified in a late-stage CLL patient (0.36kb) and possibly the shortest known in the literature.[304] Also, an MDS patient with a STELA profile of 2.92kb presented with the shortest pure TTAGGG tract of only 0.58kb amongst the MDS cohort. This patient presented with a poor cytogenetic risk score at diagnosis and progressed to AML. The pure telomeric length measured in these individuals

is shorter than what has been previously detected in fibroblast cells undergoing crisis in culture (<1kb).[255] Therefore, correction for the TVR region might reflect the true extent of telomere loss in these cells and potentially identify patients at risk of disease progression.

7.6 Mechanism of Telomere Fusion in MDS/ AML

It appeared that the chromosomal fusion events arising from short dysfunctional telomeres were processed in a manner reminiscent of Ku-independent Alternative Non-homologous end joining (A-NHEJ). Chromosomal loss adjacent to the fusion junction characterised by perfect overlapping homology has been observed to accompany A-NHEJ. Yeast and hamster cell lines deficient in *rad52*[280]and *Ku80/Xrcc4*,[298] respectively have been observed to utilise this approach in double strand break repair. Moreover, its involvement has been detected in the production of telomere fusion products in human cells undergoing crisis *in vitro*[255] and also within human leukaemic cells.[304] A-NHEJ has been proposed to utilise Poly (ADP-ribose) Polymerase I [PARP-1] along with the DNA ligase III/ XRCC1 complex.[392] Whereas the DNA ligase III/XRCC1 complex has been previously implicated in Base Excision Repair/Single Strand Break Repair (BER/SSBR); PARP-1 has been suggested in promoting the synapsis and ligation of double stand breaks following its activation by single stranded DNA.[308]

Increased A-NHEJ activity, accompanied by elevated levels of DNA ligase III has been observed in BCR-ABL^{+} Chronic Myeloid Leukaemia (CML) cells.[406] Notably, DSBs were repaired using regions of microhomology and characterised with a high frequency of deletions adjacent to the junction sites. Aberrant NHEJ components have also been detected in MDS and AML cells. The expression levels of major Classical-NHEJ (C-NHEJ) factors including DNA-PKcs, Ligase IV and Xrcc4 have been noted to decrease in NUP98/HOXD13 transgenic mice (a mouse model displaying phenotypical features of MDS)[407] and in *de novo* MDS patients in which Ligase IV expression was negatively correlated with karyotypic complexity.[408] The FLT3/ITD mutation, detected in AML patients has been observed to accompany a reduction of Ku in conjunction with an increase in DNA ligase III both *in vitro* and *in vivo*.[409] Furthermore, DSB repair utilised sites of microhomology which also accompanied deletions adjacent to the junction point, of which were both reduced in the presence of a FLT3 inhibitor.[409]

It may be speculated that aberrant expression of C-NHEJ factors may also favour the induction of A-NHEJ in the fusion between dysfunctional telomeres. Notably, the current data illustrated that telomere fusion partners aligned at short patches of 100% homology which varied in length between 2 to 33nts (mean 8.3nts). It was also apparent that adjacent telomeric or sub-telomeric deletions occurred at one or both of the participating telomeres. Sub-telomeric deletion near the limit of the fusion assay was apparent in several cases. Thus, it may be postulated that telomeric loss may extend beyond the proximal limit of the assay for which was 1655bp and 3058bp at the XpYp and 17p telomere, respectively. Sub-telomeric regions have demonstrated marked sensitivity to DNA double-strand breakage under drug-induced replicative stress[221] such that it has been proposed that they may represent fragile sites in which fork stalling would result in DSB formation.[221] Additionally, the ratio of C-NHEJ to other processes of DNA repair has been observed to decrease progressively towards telomeric loci whereby the joining of I-*Sce*I endonucleolytic induced DSBs involves extensive resection up to 9kb near telomeric regions in haploid yeast strains.[410] Similarly, large deletions of sub-telomeric DNA up to 30kb have been identified in the joining of DSBs in mouse embryonic stem cells.[411] In keeping with the current study, 50% of cytogenetically normal AML cases have been identified with cryptic sub-telomeric aberrations that include deletions and gains which can encompass up to 600kb.[412]

7.7 Complex Telomeric Fusion Events

Sequencing revealed the existence of complex fusion events involving insertions of non-telomeric genomic loci. Insertions could be identified as 17q23, 21q11, 5q33 and 19q13 whereby 17q23 and 19q13 correspond to the common fragile sites FRA17B and FRA19A, respectively. Chromosomal translocations involving 21q11 have been identified in *de novo* MDS and AML[413,414] suggesting the localisation of candidate genes involved in disease pathogenesis. Recurrent breakpoints mapped to 19q13 and 21q11 have been observed in *de novo* erythroid leukaemia (AML-M6) that present with complex karyotypes.[415] Common breakpoints at the chromosomal region 19q13 have been previously observed in various solid tumours including pancreatic and glioblastoma.[416] Notably, this region has been described to harbour candidate tumour suppressor genes including *MLL2*[417] and *AKT2.*[418] Translocations involving 17q23 has been defined as a prognostic marker whereby more aggressive forms of AML have been described in patients displaying this abnormality. The

MSi2 (*Musashi-2*) gene has been mapped to 17q23[419] in which its expression level in human myeloid leukaemia is directly correlated with overall survival.[420] Overexpression of *MSi2* increases haematopoietic cell cycle progression in mouse models, whilst its depletion leads to decreased proliferation and increased apoptosis.[420] Finally, 5q deletions are commonly detected in MDS and AML in which critical minimally deleted regions (CDR) have been identified between 5q32 to 5q33[421,422] and at 5q31,[423,424] associated with the indolent 5q- syndrome and aggressive forms of MDS/AML, respectively. Despite the failure in detecting a clonality of the telomeric fusion involving the 5q33 insertion; it is possible that this event may contribute in the development of del(5q).

7.8 Telomere Length and Clonal Expansion

Consistent with previous observations,[336] it appeared that telomere length was decreasing with elevated marrow blast count in MDS. The MDSs may be considered as a chronic phase in pathological development in which telomerase is insufficient[332] in preventing prolonged telomere loss with ongoing disease.[335] Progressive telomere shortening has been associated with a decrease in apoptosis of MDS $CD34^+$ cells.[346] Accordingly, the DNA Damage Response (DDR) that is commonly present at pre-invasive stages of major human cancers becomes abrogated with development[343] enabling ongoing entry into the cell cycle and prolonging telomere attrition. Disease progression in MDS is associated with a reduction or inactivation of such components involved in the DDR, including the $p15^{INK4B}$cyclin dependent kinase inhibitor,[124,125] *Chk1*,[122] *p53*[135,136,117,114] and *ATM*.[344] However, it should be pointed out that such results were based on percentage subgroups, i.e. <5% and 5 to 20% and thus in a subsequent study, microscopic analysis of bone marrow samples and blast quantification should be ensued so as to make a direct correlation between such parameters.

Marrow blast and presenting white blood cell (WBC) count failed to show an association with telomere length in AML cells. This study, as well as others[332,335,342,347] has shown an increase in telomerase activity in AML and therefore it is conceivable that a heterogeneous level of telomerase activity is in part accountable for this observation. However, in the presence of up-regulated telomerase and *hTERT* expression;[342,347,376,377] the regulation of telomerase access to the telomere may be deregulated in AML cells. This has been previously suggested in a study that identified elevated *TRF1* with *hTERT* expression in acute

leukaemias in which the authors suspected that *TRF1* may be deregulated in a subset of patients.[425] If this is true, telomere length would not be controlled and extensive telomere elongation or shortening may ensue.[207] Moreover, shelterin genes, i.e. *PTOP*, *RAP1* and *TRF2* as well as non-shelterin genes, i.e. *Ku70* and *Pinx1* have also been shown to be deregulated in acute leukaemias (AML and ALL) when compared to normal BMMNCs.[404] Therefore, the regulation of telomerase access to the telomere may be deregulated resulting in fluctuations in telomere length control within individual cases.

7.9 The IPSS Scoring System for MDS prognosis and Telomere Length

The International Prognostic Scoring System (IPSS) accommodates specific variables including blast cell percentage, presenting cytopenia(s) and cytogenetic abnormalities to stratify patients into low, Intermediate-1 (Int-1), Intermediate-2 (Int-2) or high in order to predict overall survival, risk of AML evolution and facilitate in therapeutic decisions.[43] In contrast to previous cases,[332,346] telomere length failed to show increased attrition with greater IPSS scores, however because the risk of disease development is based on categorical features and not on the actual depth of cytopenia severity, presenting blast count or specific karyotypic abnormality it may be speculated that variations of these features result in fluctuations in telomere length and contribute to the absence of an association between telomere length and IPSS score. It is also possible that telomerase is up-regulated in numerous patients removing the prognostic signature of telomere length, particularly within higher risk groups. This may also explain, in part the heterogeneity of telomere length that was detected amongst the higher risk scores.

It is possible that telomere length may be associated with increased risk in the more recent revised-IPSS (IPSS-R)[44] which incorporates novel chromosomal abnormalities, refines bone marrow blast percentage and analyses the depth of blood cell cytopenia by using relative cut-off points. In the IPSS-R, prognostic outlook is differentiated into 5 prognostic subgroups: Very Low, Low, Intermediate, High and Very High. However, as in the IPSS, higher risk groups may also present with up-regulated telomerase activity removing the prognostic signature of telomere length.

7.10 Telomere Length and Molecular Mutations in AML

The identification of specific molecular mutations can influence patient outcome so as to further refine patient prognosis, particularly in patients who present with a normal cytogenetic profile. Molecular markers that are commonly presented in AML patients are mutations involving the FLT3 (FMS-like tyrosine kinase 3)[374] receptor and NPM1 (nucleophosmin).[70,71]

FLT3 is expressed by cells found in the haematopoietic stem cell compartment and early committed progenitors[59,60] and has been proposed to play a role in cell proliferation survival, and differentiation.[59,60,62] Mutations either in the form of an internal tandem duplication (ITD) of the juxtamembrane domain or point mutations within the tyrosine kinase domain (TKD) result in the constitutive activation of the receptor.[61] The prognosis of patients harbouring the FLT3/ITD is poor with individuals exhibiting high rates of relapse and inferior overall survival (OS)[64,65] however, the prognostic impact of FLT3/TKD remains controversial.[60,66-68] Telomere shortening has been identified in patients who present with the FLT3/ITD mutation[347,350] and this study also showed a tendency for telomere shortening in FLT3/ITD^{+} patients. In contrast, patients presenting with the TKD mutation had significantly longer telomeres when compared to those who did not. Interestingly, several studies[378,379] have shown a weaker proliferative effect induced by the FLT3/TKD when compared to patients with the FLT3/ITD. The clonogenic ability of TKD expressing cells *in vitro* has been observed to be significantly less when compared to ITD^{+} cells.[380] Accordingly, the ITD but not TKD mutation (or wild-type FLT3) has been shown to induce robust activation of the STAT5 signalling pathway *in vitro*.[380,381] STAT5 (signal transducer and activator of transcription-5) is one of the principal pathways that regulates gene expression in response to FLT3.[382] Downstream targets of STAT5 include Pim-1 and CCND3 which both play a role in cell cycle progression and are increased in FLT3/ITD^{+} cells.[383] Thus, up-regulation of the STAT5 pathway may be in part accountable for the tendency towards shorter telomeric length in patients presenting with the ITD mutation and its activation may be associated with the observed differences of telomere length in patients with the FLT3/ITD and FLT3/TKD mutation. Surprisingly, the difference in telomere length between the ITD^{+} and ITD^{-} cohorts did not reach statistical significance; therefore it is tempting to

speculate that telomerase activity is preferentially up-regulated in ITD^+ AML cells masking the increased proliferation associated with this genetic mutation.

The current data illustrated significantly longer telomeres in AML patients carrying the FLT3/TKD mutation irrespective of the FLT3/ITD status. Regression analysis revealed a weak positive correlation between age and telomere length within the AML cohort and patients presenting with the TKD mutation were significantly older when compared to those who were negative for FLT3/TKD. Interestingly, different age profiles for molecular markers have been described such that the FLT3/ITD occurs at a constant frequency irrespective of age[384] and the incidence of FLT3/TKD increases by 29.4 fold from 21 to 70 years.[56] Therefore, it is possible that in the current study, the prevalence of FLT3/TKD in the older AML population is also accountable for the finding of longer telomeres in patients over 60 years of age.

The NPM1 protein functions as a molecular chaperone that shuttles between the nucleus and cytoplasm.[69] It is predominately nucleolar but 30% of AML cases bear the cytoplasmic NPM1 (termed: $NPMc^+$).[70,71] Mutations in NPM1 are associated with a more favourable outcome when compared to patients presenting with wild-type NPM1.[72] MicroRNA expression profiling of $NPM1^+$ and $NPM1_{wild\text{-}type}$ AML cells has identified differential expression patterns of microRNAs including the up-regulation of *let-7* in $NPM1^+$.[72] The overexpression of *let-7* has been observed to inhibit cell proliferation in human lung cancer cell lines.[385] Notably, it has been documented to interact with *MYC* and *CDC25A*[386] regulating cell proliferation and cell cycle progression, respectively. Although not significant, AML cells presenting with the ITD^-NPM1^+ had longer telomeres when analysed among ITD and NPM1 subgroups. It is possible that the tendency for longer telomeres within this subgroup may attribute to the tumour suppressor effect, i.e. *let-7* on inhibiting cellular proliferation. The reason behind the apparent elevated telomere attrition of the median length is inconclusive in patients presenting with both the ITD^+ and $NPM1^+$ mutations. This is inconsistent with the literature which has suggested a role of $NPM1^+$ to oppose the FLT3/ITD dependent activation of STAT5[387] however; it may be speculated that cells negative for the NPM1 mutation and positive for the FLT3/ITD have already undergone an extensive period of proliferation and have preferentially up-regulated telomerase in order to maintain telomere stability. Thus, telomerase activity may, in part be accountable for longer telomeres within this subgroup.

7.11 Telomere length and Survival Parameters

When telomere length was adjusted for age, MDS patients within the normal range failed to present with a better outcome, however a biphasic distribution was identified within the superior curve suggesting that a subset of individuals presenting with a telomere length within the normal age-adjusted range may have a more favourable prognosis. Further analysis revealed that the mortality rate within favourable prognostic subgroups including good-risk cytogenetics, uni-lineage cytopenia and low-risk IPSS scores was influenced by telomere length. Notably, patients who presented with shorter age-adjusted telomere length also showed a reduction in overall survival. Its prognostic potential was lost within more unfavourable prognostic categories. Moreover, when analysed by means of recursive partitioning, telomere length at diagnosis significantly influenced the overall survival of MDS patients irrespective of conventional markers.

These observations may have potential in refining patient outlook and facilitate in making therapeutic decisions, however these data are inconclusive because individual therapy was unknown for patients in the MDS cohort. Notably, overall survival can be greatly influenced depending on individual therapy. Further analysis on a cohort of patients who are undergoing uniform treatment would substantially improve these data and potentially show an association between diagnostic telomere length and prognosis in MDS patients.

Consistent with the literature,[347,350] telomere length did not influence the number of disease-free days or overall survival of AML patients who had received intensive chemotherapy. Although the rate of entry into 1st remission was unavailable for this study it is also possible that telomere length does not influence this parameter since the results from both disease-free survival and overall survival were consistent. It may be speculated that telomerase up-regulation in AML cells provides telomere length stability and removes the prognostic signature of telomere length.

Telomere length at diagnosis appeared to be shorter in AML patients who failed to achieve complete remission after the 1st cycle of intensive chemotherapy. However, this difference was not significant and the data set was limited. Accordingly, only a very small number of cases were resistant to therapy rendering this analysis inconclusive.

7.12 In Conclusion

The current data showed a significant reduction in telomere length among MDS and AML patients when compared to aged-matched healthy individuals with significantly shorter and homogeneous telomere distributions detected amongst the AML cohort. In this study, telomere length conformed only weakly to the aging dogma in the MDS cohort with several individuals showing extensive telomere loss. One such patient presented with a 10-fold increase in the rate of telomere shortening at follow-up. In contrast, telomere length appeared to increase with age in patients with AML.

Telomere length appeared to decline with increasing bone marrow blasts in the MDS cohort however, the marrow blast percentage and presenting white blood cell (WBC) count failed to show an association with telomere length in AML cells. With respect to other clinical parameters, telomere length was not associated with the number of cytopenia in MDS or cytogenetic complexity in MDS/AML. Throughout the study, large-scale telomere rapid deletion events (TRDs) were detected among STELA profiles. Such telomeres may play a role in the development of cytopenia severity and in the karyotypic complexity that is later detected in MDS and AML cells using low resolution conventional G-banding. Of note, telomere-telomere fusion events which were processed in a manner reminiscent of Ku-independent Alternative Non-Homologous end joining (A-NHEJ) were detected within the MDS cohort and to a lesser degree in the AML cohort. A reduction in telomere fusion events in AML possibly reflects the up-regulation of telomerase activity that was detected in AML cells.

Elevated telomere shortening in patients with MDS may be a prognostic indicator with shorter telomeres associated with an inferior outcome, however because individual therapy was unknown this finding is inconclusive. In contrast, telomere length did not influence the number of disease-free days or overall survival of AML patients who received intensive chemotherapy.

7.13 Future Work and Implications

This study was limited in particular areas and therefore in order to improve these data specific questions should be answered.

One such weakness included blast quantification in the MDS cohort. In a subsequent study the correlation between telomere length and presenting blast count should be determined, particularly since this analysis could only be carried out as categorical data composed of wide intervals, i.e. 5 to 20%. The morphological features of a myeloblast have been described by the International Working Group on Morphology of MDS[426] and as such are used to detect and quantify the percentage of myeloblasts in the marrow. Blast count as a percentage of all marrow nucleated cells can be determined from a 500-cell differential performed on marrow aspirate smears analysed by light microscope. STELA will be used to determine marrow telomere length of individual patients and a correlation between the presenting marrow blast percentages will be performed so as to establish an association between such parameters.

In the current study, telomere length appeared to increase with age in the AML cohort. Although the reason for this is unknown it was speculated that it might share some association with the subtype of AML. AML-M1 is frequently identified in older patients which has also been associated with significantly higher telomerase activity when compared to the remaining maturation subgroups.[375-377] In a subsequent study, AML blasts will be subtyped by immunohistochemical investigation using a panel of antibodies directed against specific cellular antigens, e.g. MPO, CD61 or Glycophorin A. Of the maturation subgroups available for analysis, TRAP will be utilised to quantify telomerase activity and telomere length will be measured by means of STELA so as to determine the relationship between marrow telomere length and telomerase activity within such subgroups. This analysis might add to the apparent finding of longer telomeres in older patients with AML if telomere length also increases with age in a subsequent study.

Telomere length was not associated with the number of cytopenia(s) in the current study; however the depth of cytopenia was unavailable for patients within the MDS cohort and therefore this analysis is inconclusive. In a follow-up study, marrow telomere length should be correlated with the intensity of haematopoietic insufficiency along a specific lineage in

order to determine whether an association exists between such parameters. Telomere length can be measured using STELA and peripheral blood cell counts can be analysed by means of automated cell counting.

Telomere length was significantly longer in patients with secondary AML when compared to their *de novo* counterparts. In order to determine whether this finding was associated with greater telomerase activity in secondary AML cells, TRAP will be carried out in a subsequent study to elucidate this theory. Additionally, telomerase activity should be measured in FLT3/ITD^{+} cases and compared with patients who do not present with the FLT3/ITD mutation, particularly since the difference in telomere length between such subgroups was not statistically significant in this study. The FLT3/ITD mutation can be characterised by gene mutational screening as previously described[379] and then TRAP and STELA can be utilised to quantify telomerase activity and telomere length, respectively between patients presenting with the ITD and patients who do not.

In this retrospective study, short telomeres were associated with significantly shorter overall survival of MDS patients. In particular, age-adjusted telomere length could refine patients with favourable prognostic markers and distinguish low-risk individuals with a poorer outlook. In a prospective study telomere length should be analysed on patients receiving uniform treatment in a more robust cohort in order to validate the prognostic significance of telomere length. Accordingly, telomere length (including patient work-up, i.e. marrow and peripheral blast counts, complete blood count (CBC) and cytogenetics) should be analysed initially at diagnosis and at regular intervals throughout the course of the disease in a homogeneous cohort of low-risk patients in order to establish whether telomere dynamics correlate with disease progression.

Appendix: 1

Cluster analysis of the 6kb region succeeding the published terminus of 6q:

```
6qPredictedSubTel       ATCTAAACTGAGTCCAGCTGGCTAACTCTAAATATATGTGTATTTTTTCAGCATAAAAAA
10qSubTel               ATCTAAACTGAGTCCAGCTGGCTAACTCTAAATATATGTGTATTTTTTCAGCATAAAAAA
22qSubTel               --CTAAACTGAGTCCAGCTGGCTAACTCTAAATATATGTGTATCTTTTCAGCATAAAAAA
1pSubTel                ------------------------------------------------------------
5qSubTel                ------------------------------------------------------------
17qSubTel               ------------------------------------------------------------
4qSubTel                ------------------------------------------------------------
19qSubTel               ------------------------------------------------------------
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------


6qPredictedSubTel       ATAATGTTTTTCATAAGAATGACAACTTAATTAGAATCAAATCTATAAGCTTTAAGATTT
10qSubTel               ATAATGTTTTTCATAAGAATGACAACTTAATTAGAATCAAATCTATAAGCTTTAAGATTT
22qSubTel               ATAATGTTTTTCATAAGAATGACAACTTAATTAGAATCAAATCTATAAGCTTTAAGATTT
1pSubTel                ------------------------------------------------------------
5qSubTel                ------------------------------------------------------------
17qSubTel               ------------------------------------------------------------
4qSubTel                ------------------------------------------------------------
19qSubTel               ------------------------------------------------------------
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------


6qPredictedSubTel       TACATTTCTAGTAAGTATAATATTAGCTTATTTGACTAGAACTCAAGCAGAATAGGAATT
10qSubTel               TACATTTCTAGTAAGTATAATATTAGCTTATTTGACTAGAACTCAAGCAGAATAGGAATT
22qSubTel               TACGTTTCTAGTAAGTATAATATTAGCTTATTTGACTAGAACTCAAGCAGAATAGGAATT
1pSubTel                ------------------------------------------------------------
5qSubTel                ------------------------------------------------------------
17qSubTel               ------------------------------------------------------------
4qSubTel                ------------------------------------------------------------
19qSubTel               ------------------------------------------------------------
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel      TATGCTTGTTTTATATTCAATAATGATAATTTTGAAGATATAGTTGTTTTATTACACCAA
10qSubTel              TATGCTTGTTTTATATTCAATAATGATAATTTTGAAGATATAGTTGTTTTATTACACCAA
22qSubTel              TATGCTTGTTTTATATTCAATAATGATAATTTTGAAGATATAGTTGTTTTATTACACCAA
1pSubTel               ------------------------------------------------------------
5qSubTel               ------------------------------------------------------------
17qSubTel              ------------------------------------------------------------
4qSubTel               ------------------------------------------------------------
19qSubTel              ------------------------------------------------------------
21qSubTel              ------------------------------------------------------------
1qSubTel               ------------------------------------------------------------
2qSubTel               ------------------------------------------------------------
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      AAATACTATATTAATCTTATTTAACTAAGTTTTATCCAAATCATGTTAACTTAAGAAACA
10qSubTel              AAATACTATATTAATCTTATTTAACTAAGTTTTATCCAAATCATGTTAACTTAAGAAACA
22qSubTel              AAATACTATATTAATCTTATTTAACTAAGTTTTATCCAAATCATGTTAACTTAAGAAACA
1pSubTel               ------------------------------------------------------------
5qSubTel               ------------------------------------------------------------
17qSubTel              ------------------------------------------------------------
4qSubTel               ------------------------------------------------------------
19qSubTel              ------------------------------------------------------------
21qSubTel              ------------------------------------------------------------
1qSubTel               ------------------------------------------------------------
2qSubTel               ------------------------------------------------------------
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TTTGATCAGTTCCTATATTTCTAGGAGTTTGGTGAATATTTATTTATAAATGCTTATTTT
10qSubTel              TTTGATCAGTTCCTATATTTCTAGGAGTTTGGTGAATATTTATTTATAAATGCTTATTTT
22qSubTel              TTTGATCAGTTCCTATATTTCTAGGAGTTTGGTGAATATTTATTTATAAATGCTTATTTT
1pSubTel               ------------------------------------------------------------
5qSubTel               ------------------------------------------------------------
17qSubTel              ------------------------------------------------------------
4qSubTel               ------------------------------------------------------------
19qSubTel              ------------------------------------------------------------
21qSubTel              ------------------------------------------------------------
1qSubTel               ------------------------------------------------------------
2qSubTel               ------------------------------------------------------------
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel      TTTCCAAGCCAAGTTAGAATAGAGCACTTTTAGAGGATTTCATAAATGAATTTTGCAATG
10qSubTel              TTTCCAAGCCAAGTTAGAATAGAGCACTTTTAGAGGATTTCATAAATGAATTTTGCAATG
22qSubTel              TTTCCAAGCCAAGTTAGAATAGAGCACTTTTAGAGGATTTCATAAATGAATTTTGCAATG
1pSubTel               ------------------------------------------------------------
5qSubTel               ------------------------------------------------------------
17qSubTel              ------------------------------------------------------------
4qSubTel               ------------------------------------------------------------
19qSubTel              ------------------------------------------------------------
21qSubTel              ------------------------------------------------------------
1qSubTel               ------------------------------------------------------------
2qSubTel               ------------------------------------------------------------
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      CTCTCTGGAGTTAAGAAAATATCACATATACATAACATACATTAATAGATATACAAACAC
10qSubTel              CTCTCTGGAGTTAAGAAAATATCACATATACATAACATACATTAATAGATATACAAACAC
22qSubTel              CTCTCTGGAGTTAAGAAAATATCACATATACATAACATACATTAATAGATACACAAACAC
1pSubTel               ------------------------------------------------------------
5qSubTel               ------------------------------------------------------------
17qSubTel              ------------------------------------------------------------
4qSubTel               ------------------------------------------------------------
19qSubTel              ------------------------------------------------------------
21qSubTel              ------------------------------------------------------------
1qSubTel               ------------------------------------------------------------
2qSubTel               ------------------------------------------------------------
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      AAATAGAGATTTCATAGCTTTCATCCTGAAATTTCAGCCTTGAATCAGGCATAAATATTC
10qSubTel              AAATAGAGATTTCATAGCTTTCATCCTGAAATTTCAGCCATGAATCAGGCATAAATATTC
22qSubTel              AAATAGAGATTTCATAGCTTTCATCCTGAAATTTCAGCCATGAATCAGGCATAAATATTC
1pSubTel               ------------------------------ATTTCAGCCATGAATCAGGCATAAATATTC
5qSubTel               ------------------------------ATTTCAGCCTTGAATCAGGCATAAATATTC
17qSubTel              ------------------------------ATTTCAGCCATGAATCAGGCATAAATATTC
4qSubTel               ------------------------------------------------------------
19qSubTel              ------------------------------------------------------------
21qSubTel              ------------------------------------------------------------
1qSubTel               ------------------------------------------------------------
2qSubTel               ------------------------------------------------------------
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel       TGATGGTTAATTTCAGACATCTACTTGATCCGACTGAGAGACACACATAGCTGGTCAAAC
10qSubTel               TGATGGTTAATTTCAGACATCTACTTGATCCGATTGAGAGACACACATAGCTGGTCAAAC
22qSubTel               TGATGGTTAATTTCAGACATCTACTTGATCGGATTGAGAGACACGCATAGCTGGTCAAAC
1pSubTel                TGACGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAAC
5qSubTel                TGATGGTTAATTTCAGACATCTACTTGATCCGACTGAGAGACACACATAGCTGGTCAAAC
17qSubTel               TGT-GGTTAATTTTAGACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAAC
4qSubTel                ------------------------------------------------------------
19qSubTel               ------------------------------------------------------------
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       ACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTGTAGACATCTACTTG
10qSubTel               ACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTG
22qSubTel               ACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTG
1pSubTel                AAGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTG
5qSubTel                ACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTGTAGACATCTACTTG
17qSubTel               --AATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTACTTG
4qSubTel                ---ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTCAGATATCTACTTG
19qSubTel               ------------------------------------------------------------
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       ACTGGATTAAGAGACACACATAGCTGGTCAAACAAGATTTCAGCCATGAATCAGGCATAA
10qSubTel               ACTGGATTAAGAGACACACATAGCTGGTCAAACAAGATTTCAGCCATGAATCAGGCATAA
22qSubTel               ACTGGATTAAGAGACACACATAGCTGGTCAAACATGATTTCTGACATGAATCAGGCATAA
1pSubTel                ACTGGATTAAGAGACACACATAGCTGGTCTAACACGATTTCAGCCATGAATCAGGCATAA
5qSubTel                ACTGGATTAAGAGACACACATAGCTGGTCAAACAAGATTTCAGCCATGAATCAGGCATAA
17qSubTel               AGTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAA
4qSubTel                ATCCGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAA
19qSubTel               ------------------------------------------------------------
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel     ATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGG
10qSubTel             ATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGG
22qSubTel             ATATTCTGACGGTTAATTTTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGG
1pSubTel              ATATTCTGATGGTTAACTTTAGGCATCTACTTGATTGGATTGAGAGACACACATAGCTGG
5qSubTel              ATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGG
17qSubTel             ATATTCTGACGGTTAATTTTAGACATCAACTTGACTGGATTAAGGGACACACATAGCTGG
4qSubTel              ATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGG
19qSubTel             ------------------------------------------------------------
21qSubTel             ------------------------------------------------------------
1qSubTel              ------------------------------------------------------------
2qSubTel              ------------------------------------------------------------
19p'End'              ------------------------------------------------------------
8p'End'               ------------------------------------------------------------
13qSubTel             ------------------------------------------------------------
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------

6qPredictedSubTel     TCTAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAACTTTAGACATC
10qSubTel             TCTAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAACTTTAGGCATC
22qSubTel             TCAAACATGATTTCAGCCATGAATCAGCCATAAATATTCTGACGGTTAATTTTAGACATC
1pSubTel              TCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTGTAGACATC
5qSubTel              TCTAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAACTTTAGACATC
17qSubTel             TCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTGTAGACATC
4qSubTel              TCAAACAAGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTGTAGACATC
19qSubTel             ---------ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTCAGACATC
21qSubTel             ---------ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTCAGACATC
1qSubTel              ------------------------------------------------------------
2qSubTel              ------------------------------------------------------------
19p'End'              ------------------------------------------------------------
8p'End'               ------------------------------------------------------------
13qSubTel             ------------------------------------------------------------
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------

6qPredictedSubTel     TACTTGATTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
10qSubTel             TACTTGATTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
22qSubTel             TACTTGATCGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
1pSubTel              TACTTGACTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
5qSubTel              TACTTGATTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
17qSubTel             TACTTGACTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
4qSubTel              TACTTGACTGGATTAAGAGACACACATAGCTGGTCTAACACGATTTCAGCCATGAATCAG
19qSubTel             TACTTGATCGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
21qSubTel             TACTTGATCGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAG
1qSubTel              ------------------------------------------------------------
2qSubTel              ------------------------------------------------------------
19p'End'              ------------------------------------------------------------
8p'End'               ------------------------------------------------------------
13qSubTel             ------------------------------------------------------------
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------
```

```
6qPredictedSubTel       GCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGACACACAT
10qSubTel               -CATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGACACACAT
22qSubTel               GCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTGAGAGACACACAT
1pSubTel                GCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGACTGGATTAAGGGACACACAC
5qSubTel                GCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGACACACAT
17qSubTel               GCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGACACACAT
4qSubTel                GCATAAATATTCTGATGGTTAACTTTAGGCATCTACTTGATTGGATTGAGAGACACACAT
19qSubTel               GCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACAT
21qSubTel               GCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGACTGGATTAAGAGACACACAT
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       AGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTA
10qSubTel               AGCTGGCCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTA
22qSubTel               AGCTGGTCAAACACAATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTA
1pSubTel                AGCTGGTCAAACACAATTTCAGCCATGAATCAGGCATAAATATTCTGACAGTTAATTTTA
5qSubTel                AGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTA
17qSubTel               AGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTA
4qSubTel                AGCTGGTCAAACACGATTTCAGCCATGAAGCAGGCATAAATATTCTGATGGTTAATTGTA
19qSubTel               AGCTGGTCAAACATGATTTCAGCCATGAATCAGCCATAAATATTCTGATGGTTAATTTTA
21qSubTel               AGCTGGTCAAACATGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTA
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       GACATCTACTTGACTGGATTAAGGGACACACACAGCTGGTCAAACACAATTTCAGCCATG
10qSubTel               GACATCTACTTGACTGGATTAAGGGACACACACAGCTGGTCAAACACAATTTCAGCCATG
22qSubTel               GACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAACA--ATTTCAGCCATG
1pSubTel                GACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCGTG
5qSubTel                GACATCTACTTGACTGGATTAAGGGACACACACAGCTGGTCAAACACAATTTCAGCCATG
17qSubTel               GACATCAACTTGACTGGATTAAGGGACACACATAGCTGGTCAAACACGATTTCAGCCATG
4qSubTel                GACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATG
19qSubTel               GACATCTACTTGAGTGGATTAAGAGACACACATAGCTGGTCAAACACAATTTCAGCCATG
21qSubTel               GATATCTACTTGAGTGGATTAAGAGACACACATAGCTGGTCAAACACAATTTCAGCCATG
1qSubTel                ------------------TTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATG
2qSubTel                ------------------------------------------------------------
19p'End'                ------------------------------------------------------------
8p'End'                 ------------------------------------------------------------
13qSubTel               ------------------------------------------------------------
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel    AATCAGGCATAAATATTCTGACAGTTAATTTTAGACATCTACTTGACTGGATTAAGAGAC
10qSubTel            AATCAGGCATAAATATTCTGACAGTTAATTTTAGACATCTACTTGACTGGATTAAGAGAC
22qSubTel            AATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTACTTGAGTGGATTGAGAGAC
1pSubTel             AAGCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGAC
5qSubTel             AATCAGGCATAAATATTCTGACAGTTAATTTTAGACATCTACTTGACTGGATTAAGAGAC
17qSubTel            AATCAGGCATAA-TATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTGAGAGAC
4qSubTel             AAGCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGAC
19qSubTel            AATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTACTTGATCGGATTGAGAGAC
21qSubTel            AATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTACTTGATCGGATTGAGAGAC
1qSubTel             AATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGACTGGATTAAGAGAC
2qSubTel             ------------------------------------------------------------
19p'End'             ------------------------------------------------------------
8p'End'              ------------------------------------------------------------
13qSubTel            ------------------------------------------------------------
20qSubTel            ------------------------------------------------------------
12qSubTel            ------------------------------------------------------------
7qSubTel             ------------------------------------------------------------

6qPredictedSubTel    ACACATAGCTGGTCAAACACGATTTCAGCCGTGAAGCAGGCATAAATATTCTGATGGTTA
10qSubTel            ACACATAGCTGGTCAAACATGATTTCAGCCGTGAAGCAGGCATAAATATTCTGATGGTTA
22qSubTel            ACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTA
1pSubTel             ACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCGTAAATATTCTGATGGTTA
5qSubTel             ACACATAGCTGGTCAAACACGATTTCAGCCGTGAAGCAGGCATAAATATTCTGATGGTTA
17qSubTel            ACACATAGCTGGTCAAACATGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTA
4qSubTel             ACACATAGCTGGTCAAACACAATTTCAGCCATGAATCAGGCGTAAATATTCTGATGGTTA
19qSubTel            ACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTA
21qSubTel            ACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTA
1qSubTel             ACACATAGCTGGTCAAACATGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTA
2qSubTel             --------------------AATTTCAGCCTTGAATCAGGCATAAATATTCTGATGGTTA
19p'End'             ------------------------------------------------------------
8p'End'              ------------------------------------------------------------
13qSubTel            ------------------------------------------------------------
20qSubTel            ------------------------------------------------------------
12qSubTel            ------------------------------------------------------------
7qSubTel             ------------------------------------------------------------

6qPredictedSubTel    ATTGTAGACATCTACTTGACTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCA
10qSubTel            ATTGTAGACATCTACTTGACTGGATTGAGAGACACACATAGCTGGTCAAACACAATTTCA
22qSubTel            ATTTTAGACATCAACTTGACTGGATTAAGGGACACACATAGCTGGTCAAACACGATTTCA
1pSubTel             ATTGTAGACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAACACAATTTCA
5qSubTel             ATTGTAGACATCTACTTGACTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCA
17qSubTel            ATTTTAGGCATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAAC--AATTTCA
4qSubTel             ATTGTAGACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCA
19qSubTel            ATTTTAGACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAACACAATTTCA
21qSubTel            ATTTTAGACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAACACAATTTCA
1qSubTel             ATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGCTGGTCAAACACGATTTCA
2qSubTel             ATTTCAGACATCTACTTGATCCGATTGAGAGACACACATAGCTGGTCAAACACGATTTCA
19p'End'             ------------------------------------------------------------
8p'End'              ------------------------------------------------------------
13qSubTel            ------------------------------------------------------------
20qSubTel            ------------------------------------------------------------
12qSubTel            ------------------------------------------------------------
7qSubTel             ------------------------------------------------------------
```

```
6qPredictedSubTel      GCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTG
10qSubTel              GCCATGAATCAGGCGTAAATATTCTGATGGTTAATTGTAGACATCTACTTGAGTGGATTG
22qSubTel              GCCATGAATCAGGCATAAATATTCTGACGGTTAATCGTAGACGTCTACTTGACTGGATTG
1pSubTel               GCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATGTACTTGACTGGATTA
5qSubTel               GCCATGAATCAGGCGTAAATATTCTGATGGTTAATTGTAGACATCTACTTGAGTGGATTG
17qSubTel              GCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTG
4qSubTel               GCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATGTACTTGACTGGATTA
19qSubTel              GCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTACTTGAGTGGATTG
21qSubTel              GCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTG
1qSubTel               GCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGTGGATTA
2qSubTel               GCCATGAATCAGGCATAAATATTCTGACGGTTAATTGTAGACATCTACTTGACTGGATTA
19p'End'               ------------------------------------------------------------
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      AGAGACACACATAGCTGGTCAAACACAATTTCAGCCATGAATCAGGCATAAATATTCTGA
10qSubTel              AGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGA
22qSubTel              AGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGA
1pSubTel               AGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAAGCAGGCATAAATATTCTGA
5qSubTel               AGAGACACACATAGCTGGTCAAACACAATTTCAGCCATGAATCAGGCATAAATATTCTGA
17qSubTel              AGAGACACACATAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCATAAATATTCTGA
4qSubTel               AGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAAGCAGGCATAAATATTCTGA
19qSubTel              AGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGA
21qSubTel              AGAGACACACATAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCATAAATATTCTGA
1qSubTel               AGAGACATACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATATTCTGA
2qSubTel               AGAGACACACATAGCTGGTCAAACAAGATTTCAGCCATGAATCAGGCATAAATATTCTGA
19p'End'               --------------------------AATTTCAGCCATGAATCAGGCATAAATATTCTGA
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TGGTTAATTTTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACG
10qSubTel              TGGTTAATTTTAGACATGTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACG
22qSubTel              TGGTTAATCGTAGACATCTACTTAACTGGATTGAGAGACACACATAGCTGGTCAAACACG
1pSubTel               TGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACG
5qSubTel               TGGTTAATTTTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACG
17qSubTel              CGGTTAATTTTAGACATCTACTTGAGTGGATTGAGAGACACACATAGCTGGTCAAACACG
4qSubTel               TGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACG
19qSubTel              CGGTTAATTTTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTGAAACACG
21qSubTel              CGGTTAATTTTAGACATCAACTTGACTGGATTAAGGGACACACATAGCTGGTCAAACACG
1qSubTel               TGGTTAATTGTAGACATCTACTTGAGTGGATTAAGAGACACGCATAGCTGGTCAAACACG
2qSubTel               TGGTTAATTGTAGACATCTACTTGACTGGATTAAGAGACACACATAGCTGGTCAAACACG
19p'End'               TGGTTAATTTTAGACATCCACTTGATCGGATTAAGAGACACACATAGCTGGTCAAACACG
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel      ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGACT
10qSubTel              ATTTCAGCCATGAAGCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACT
22qSubTel              ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATCGTAGACATCTACTTGACT
1pSubTel               ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAACTTTAGGCATCTACTTGATT
5qSubTel               ATTTCAGCCATGAAGCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACT
17qSubTel              ATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCAACTTGACT
4qSubTel               ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAACTTTAGGCATCTACTTGATT
19qSubTel              ATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAACTTTAGACATCTACTTGACT
21qSubTel              ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTGTAGACATCTACTTGACT
1qSubTel               ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGAGT
2qSubTel               ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAACTTTAGGCATCTACTTGATT
19p'End'               ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTACTTGACT
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel      GGATTGAGAGACACACATAGCTGGTCAAACACAATTTCAGCCATGAATCAGGCATAAATA
10qSubTel              GGATTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
22qSubTel              GGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
1pSubTel               GGATTGAGAGACACACATAGCTGATCAAACACAATTTCAGCCATGAATCAGGCATAAATA
5qSubTel               GGATTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
17qSubTel              GGATTAAGGGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATACATA
4qSubTel               GGATTGAGAGACACACATAGCTGATCAAACACAATTTCAGCCATGAATCAGGCATAAATA
19qSubTel              GGATTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
21qSubTel              GGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
1qSubTel               GGATTAAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
2qSubTel               GGATTGAGAGACACACATAGCTGATCAAACACAATTTCAGCCATGAATCAGGCATAAATA
19p'End'               GGATTGAGAGACACACACAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCATAAATA
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel      TTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGACACACATAGCTGATCA
10qSubTel              TTCTGATGGTTAACTTTAGGCATCTACTTGATTGGATTGAGAGACACACATAGCTGATCA
22qSubTel              TTCTGATGGTTAATCGTAGACATCTACTTGACTGGATTGAGAGACACACATAGCTGGTCA
1pSubTel               TTCTGACAGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACACAGCTGGTCA
5qSubTel               TTCTGATGGTTAACTTTAGGCATCTACTTGATTGGATTGAGAGACACACATAGCTGATCA
17qSubTel              TTCTGATGGTTAATTGTAGACATCTACTTGACTGGATTGAGAGACACACATAGCTGGTCA
4qSubTel               TTCTGACGGTTAATTTTAGA----------------------------------------
19qSubTel              TTCTGACGGTTAACTTTAGACATCTACTTGATTGGATTGAGAGACACACATAGCTGGTCA
21qSubTel              TTCTGATGGTTAATCGTAGACGTCTACTTGACTGGATTGAGAGACACACATAGCTGGTCA
1qSubTel               TTCTGATGGTTAATTGTAGACATCTACTTGAGTGGATTAAGAGACACGCATAGCTGGTCA
2qSubTel               TTCTGACAGT--------------------------------------------------
19p'End'               TTCTGATGGTTAATTTTAGACATCTACTTGATCGGATTA-AGGACACACACAGCTGGTCA
8p'End'                ------------------------------------------------------------
13qSubTel              ------------------------------------------------------------
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel       AACACGATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTAC
10qSubTel               AACACAATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTAC
22qSubTel               AACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATCGTAGACATCTAC
1pSubTel                AACA--ATTTCAGCCATGAATCAGGCATAAATATTCTGACAGTTAATTTTAGACATCTAC
5qSubTel                AACACAATTTCAGCCATGAATCAGGCATAAATATTCT-----------------------
17qSubTel               AACAGGATTTCAGCCATGAATCAGACATAAATATTCTGATGGTTAATTGTAGACGTCTAC
4qSubTel                -----------------------------------------------------CATCTAC
19qSubTel               AACACAATTTCAGCCATGAATCAGGCATAAATATTCTGACGGTTAATTTTAGACATCTAC
21qSubTel               AACACGATTTCAGCCATGAATCAGGCATAAATATTCT-----------------------
1qSubTel                AACACGATTTCAGCCATGAATCAGGCATAAATATTCTG----------------------
2qSubTel                ------------------------------------------------------------
19p'End'                AACACGATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTAC
8p'End'                 ------ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTAC
13qSubTel               ------ATTTCAGCCATGAATCAGGCATAAATATTCTGATGGTTAATTTTAGACATCTAC
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------


6qPredictedSubTel       TTGACTGGATTAAGGGACACACACAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCA
10qSubTel               TTGACTGGATTAAGGGACACACACAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCA
22qSubTel               TTGACTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCA
1pSubTel                TTGAGTGGATTAAGAGACACACACAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCA
5qSubTel                ------------------------------------------------------------
17qSubTel               TTGACTGGATTGAGAGACACACATAGCTGGTCAAACACGATTTCAGCCATGAATCAGGCA
4qSubTel                TTGACTGGATTAAGGGACACACACAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCA
19qSubTel               TTGACTGGATTAAGGGACACACACAGCTGGTCAAAC--AATTTCAGCCATGAATCAGGCA
21qSubTel               ------------------------------------------------------------
1qSubTel                ------------------------------------------------------------
2qSubTel                ------------------------------------------------------------
19p'End'                TTGACTGGATTAAGAGACACACATAGCTGGTCAAACACAATTTCAGCCATGAATCTAGCA
8p'End'                 TTGATTGGATTAAGAAACATACATAGCTGGTCAAAC--AATTTCAGCCATGAAACAGGCA
13qSubTel               TTGACTGGACTAAGAGACACACATAGCTGGTCAAACACAATTTCAGCCATGAATCAGGCA
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------


6qPredictedSubTel       TAAATATTCTGACAGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGC
10qSubTel               TAAATATTCTGACAGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGC
22qSubTel               TAAATATTCTGATGGTTAATCGTAGACGTCTACTTGACTGGATTGAGAGACACACACAGC
1pSubTel                TAAATATTCTGACAGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGC
5qSubTel                ----------GACAGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGC
17qSubTel               TAAATATTCTGATGGTTAATCGTAGACGTCTACTTGACTGGATTGAGAGACACACACAGC
4qSubTel                TAAATATTCTGACAGTTAATTTTAGACATCTATTTGAGTGGATTAAGAGACACACATAGC
19qSubTel               TAAATATTCTGACAGTTAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGC
21qSubTel               ----------GATGGTTAATCGTAGACGTCTACTTGACTGGATTGAGAGACACACACAGC
1qSubTel                -----------ATGGTTAATCGTAGACGTCTACTTGACTGGATTGAGAGACACACACAGC
2qSubTel                ----------------TAATTTTAGACATCTACTTGAGTGGATTAAGAGACACACATAGC
19p'End'                TAAATATTCTGATGGTTAATTGTAGACATCTACTTGGCTGGATTGAGAGACACACACAGC
8p'End'                 TAAATATTCTGATGGTTAATTTTAGACATCTACTTGACTGGATTGAGAGACACACATAGC
13qSubTel               TAAATATTCTGATGGTTAAGTTTAGACATCTACTTGATTGGATTAAGAGACACACATAGC
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel        TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
10qSubTel                TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
22qSubTel                TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
1pSubTel                 TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
5qSubTel                 TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
17qSubTel                TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
4qSubTel                 TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
19qSubTel                TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
21qSubTel                TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
1qSubTel                 TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
2qSubTel                 TGGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAAGACACTGAGATAAC
19p'End'                 TGGTCAAACACAATTGCTGGGCATATCTGTGAGGGTGTTTCTGGAAGACACTGAGATAAG
8p'End'                  TGGTGAAACACAATTTCTGGGCATATCTGTGAAGGTGTTTCTGGAAGACACTGAGATAAC
13qSubTel                TGGTCAAACACAATTTCTGGGCATATCTGTGAGGGTGTTTCTGGAAGACACTGAGATAAC
20qSubTel                ------------------------------------------------------------
12qSubTel                ------------------------------------------------------------
7qSubTel                 ------------------------------------------------------------


6qPredictedSubTel        CATGACC-CAATGTGGATGGGCACTGAT---ATGGTTTGGCTGTGTCCCCACCCAGATCT
10qSubTel                CAGTCAAACACGATTTCTGGGCATATCTATGAGGGTGTTTCTGGAA-GACA--CTGA-GA
22qSubTel                CATGACC-CAATGTGGATGGGCACTGA---TATGGTTTGGCTGTGTCCCCACCCAGATCT
1pSubTel                 CATGACC-CAATGTGGATGGGCACTGA---TATGGTTTGGCTGTGTCCCCACCCAGATCT
5qSubTel                 CATGACC-CAATGTGGATGGGCACTGAT---ATGGTTTGGCTGTGTCCCCACCCAGATCT
17qSubTel                CATGACC-CAATGTGGATGGGCACTGAT---ATGGTTTGGCTGTGTCCCCACCCAGATCT
4qSubTel                 CATGACC-CAATGTGGATGGGCACTGA---TATGGTTTGGCTGTGTCCCCACCCAGATCT
19qSubTel                CATGACC-CAATGTGGATGGGCACTGA---TATGGTTTGGCTGTGTCCCCACCCAGATCT
21qSubTel                CATGACC-CAATGTGGATGGGCACTGAT---ATGGTTTGGCTGTGTCCCCACCCAGATCT
1qSubTel                 CATGACC-CAATGTGGATGGGCACTGA---TATGGTTTGGCTGTGTCCCCACCCAGATCT
2qSubTel                 CATGACC-CAATGTGGATGGGCACTGA---TATGGTTTGGCTGTGTCCCCACCCAGATCT
19p'End'                 CATGATC-CAGTGTGGATGGGCACTGA---TAGGGTTTGGCTGTGTCCCCACCCAGATCT
8p'End'                  CCTGACC-CAGTGTGGATGGGCACTGA---TATGGTTTGCCTGTGTCCCCACCCAGATCT
13qSubTel                CATGACC-CAGTGTAGATGGGAACTGAT---ATGGTTTGCCTGTTTCCCCACCCAGATCT
20qSubTel                ------------------------------------------------------------
12qSubTel                ------------------------------------------------------------
7qSubTel                 ------------------------------------------------------------


6qPredictedSubTel        CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
10qSubTel                TAACCAGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
22qSubTel                CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAATGGGAGGTG
1pSubTel                 CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
5qSubTel                 CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
17qSubTel                CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAATGGGAGGTG
4qSubTel                 CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
19qSubTel                CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
21qSubTel                CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAATGGGAGGTG
1qSubTel                 CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAATGGGAGGTG
2qSubTel                 CATCTTGAATTGTAGTTCCTGTAATACCTACATGTCGTGGGAGGGACCCAATGGGAGGTG
19p'End'                 CATCTTGAATTGTAGTTCCTATAATCCCTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
8p'End'                  CATCTTGAATTGTAGTTCCTATAATCCGTACATGTCGTGGGAGGGACCCAGTGGGAGGTG
13qSubTel                CATCTTGAATTGTAGTTCCTATAATCCCTAAATGTCGTGGGAGGGACCCGGTGGGAGGTG
20qSubTel                ------------------------------------------------------------
12qSubTel                ------------------------------------------------------------
7qSubTel                 ------------------------------------------------------------
```

```
6qPredictedSubTel      ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
10qSubTel              ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
22qSubTel              ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
1pSubTel               ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
5qSubTel               ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
17qSubTel              ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
4qSubTel               ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
19qSubTel              ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
21qSubTel              ACTGAATCATGGTGGTGGTTACTGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
1qSubTel               ACTGAATCATGGTGGTGGTTACTGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
2qSubTel               ACTGAATCATGGTGGTGGTTACCGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCATG
19p'End'               ACTGAATCATGGTGGTGGTTACTGCCATGCTGTTCCCATGACAGTGAGTGAGTTCTCACG
8p'End'                ATTGAATCATGGTGGTTGTTACTGCCATTCTGTTTCATGGCAGTGAGTGAGTTCTCATG
13qSubTel              ATTGAATCATGGTGGTGGTTACTGCCATGCTGTTCTCATGACAGTGAGTGAGTTCTCACG
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
10qSubTel              ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTACC
22qSubTel              ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGGTCAGCACTT---CTTGTTGCTGCC
1pSubTel               ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
5qSubTel               ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
17qSubTel              ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
4qSubTel               ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
19qSubTel              ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGGTCAGCACTT---CTTGTTGCTGCC
21qSubTel              ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
1qSubTel               ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
2qSubTel               ATCTGATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTT---CTTGTTGCTGCC
19p'End'               ATCTCATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTTCTCCTTGTTGCTGCC
8p'End'                ATCCAATGGTTTCATAAGGGGCTGTTCCCCTTTGGCTCAGCACTTCTTCTTGTTGCTGCC
13qSubTel              ATCTCATGGTTTTATAAGGGGCTTTTCCCCTTTGGCTCAGCACTTCTCCTTGTTGCTGCC
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
10qSubTel              ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
22qSubTel              ATGTGAAGAGGGACAGCTTTGCTTCCCCTTCTGCCATGATTTTGAGGCCCCTGCAGCCAT
1pSubTel               ATGTGAAGAGGGACAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAC
5qSubTel               ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
17qSubTel              ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
4qSubTel               ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
19qSubTel              ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
21qSubTel              ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
1qSubTel               ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCCGCAGCCAT
2qSubTel               ATGTGAAGAGGGATAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
19p'End'               TTGTGAAGAAGGACAGCTTTGCTTCCCCTTCTGCCATGATTGTGAGGGCCCTGCAGCCAT
8p'End'                ATGTGAAGAAGGACGTCTTTGTTTCCCCTTCTGCCATGATTGTGAGGCCTCTGCAGCCAC
13qSubTel              TTGTGAAGAGGGATAGATTTGCTTCCCCTTCTGCCATGATTGTGAGGCCCCTGCAGCCAT
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel       GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
10qSubTel               GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
22qSubTel               GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
1pSubTel                GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGAATT
5qSubTel                GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
17qSubTel               GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGATATT
4qSubTel                GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
19qSubTel               GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
21qSubTel               GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
1qSubTel                GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
2qSubTel                GTGGAACTGTCAGCCCATTAAACCCCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
19p'End'                GTGGAACTGTCAGCCCATTCAACCTCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
8p'End'                 GTGGAACTGTCAGCCCATTTAACCTCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
13qSubTel               GTGGAACTGTCAGCCCATTAAACCTCTTTGTTCTTTATAAATTGCTCAGACTCAGGTATT
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
10qSubTel               TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
22qSubTel               TCTTCATAGCTGTATAAAAATGGATGAATACAGGCAGCATCCAATTGGTTGAGAGCCCAG
1pSubTel                TCTTCATAGCTGTATAAAAATGGATGAATACAGGCAGCATCCAATTGGTTGAGAGCCCAG
5qSubTel                TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
17qSubTel               TCTTCATAGCTGTATAAAAATGGATGAATACAGGCAGCATCCAATTGGTTGAGAGCCCAG
4qSubTel                TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
19qSubTel               TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
21qSubTel               TCTTCATAGCTGTATAAAAATGGATGAATACAGGCAGCATCCAATTGGTTGAGAGCCCAG
1qSubTel                TCTTCATAGCTGTATAAAAATGGATGAATACAGGCAGCATCCAATTGGTTGAGAGCCCAG
2qSubTel                TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
19p'End'                TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
8p'End'                 TCTTCACAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
13qSubTel               TCTTCATAGCTGTATAAAAATGGATGAATACAGGCACCATCCAATTGGTTGAGAGCCCAG
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
10qSubTel               ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
22qSubTel               ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
1pSubTel                ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAGTGGAAACATCCTTCT
5qSubTel                ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
17qSubTel               ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
4qSubTel                ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
19qSubTel               ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
21qSubTel               ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
1qSubTel                ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
2qSubTel                ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAGTGGAAACATCCTTCT
19p'End'                ATAAAACAACAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAATGGAAACATCCTTCT
8p'End'                 ATAGAATAACAAGGAAGAGGAAAGGTGAATTATCTCC---TGAAATTGAAACATCCTTCT
13qSubTel               ATAGAACAAAAAGGAAGAGGAAAGGTGAATTATCTCCTTCTGAAACGGAAACATCCTTCT
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel      TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
10qSubTel              TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
22qSubTel              TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
1pSubTel               TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
5qSubTel               TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
17qSubTel              TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
4qSubTel               TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
19qSubTel              TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
21qSubTel              TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
1qSubTel               TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
2qSubTel               TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
19p'End'               TCTCCTGCCCTTGACATCAGAACTTCAGGGTCTCAGACCTTTGGCCTCAGAATCAGAGTT
8p'End'                TCTCCTGCCCTTGACATCAGAA--TCAGTGTCTCAGAGCTTTGGCCTCAGAATCAGAGTT
13qSubTel              TCTCCTGCCCTTGACATCAGAACTTCGGGGTCTCAGACCTTTGGCCTCACAATCAGAGTT
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      ACACCATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
10qSubTel              ACACCATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
22qSubTel              ACACCATTGGCTTCCCCGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
1pSubTel               ACACCATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
5qSubTel               ACACCATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
17qSubTel              ACACCATTGGCTTCCCCGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
4qSubTel               ACACCATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
19qSubTel              ACACCATTGGCTTCCCCGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
21qSubTel              ACACCATTGGCTTCCCCGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
1qSubTel               ACACCATTGGCTTCCCCGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
2qSubTel               ACACCATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCATGCTACCAGCT
19p'End'               ACACTATTGGCTTCCCTGATTCTGAGTCCTTTGTATCTGGAGTGAGCCACGCTACCAGCT
8p'End'                ACACCATTGGCTTCCCTGATTCTGAGTCCTTTATATCTGGAGTGAGTCATGCTGCCAGCT
13qSubTel              ACACCATTGGCTTCCCTGATTCTGAGTCCTTTATATCTGGAGTGAGCCATGCTGCCAGCT
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
10qSubTel              TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAAATTCTCAGCCTCCATAATTATG
22qSubTel              TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
1pSubTel               TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
5qSubTel               TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
17qSubTel              TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
4qSubTel               TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
19qSubTel              TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
21qSubTel              TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
1qSubTel               TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAGTTATG
2qSubTel               TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
19p'End'               TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGGAACTTCTCAGCCTCCATAATTATG
8p'End'                TTCCTGGTTCTCCAACTTGGAGACAGGCTATTGTGTAACTTCTCAGCCTCCATAATTATG
13qSubTel              TTCCTGGTTCTCGAACTTGGAGACAGGCTATTGTGTAACTTCTCAGCCTCCATAATTATG
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel       TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
10qSubTel               TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
22qSubTel               TGAACCAGTTCCCCTAATGAATCTTCTGTCATCTG----TCTACATATATCCTATTGATT
1pSubTel                TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
5qSubTel                TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
17qSubTel               TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTG----TCTACATATATCCTATTGATT
4qSubTel                TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
19qSubTel               TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
21qSubTel               TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTG----TCTACATATATCCTATTGATT
1qSubTel                TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTG----TCTACATATATCCTATTGATT
2qSubTel                TGAACCAGTTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
19p'End'                TGAACCAATTCCCCTAATGAATCTTCTCTCATCTA----TCTACATATATCCTATTGATT
8p'End'                 TGAACCAATTCCCCTAATGAGTCTTCTCTCATCTA----TCTACATATATCCTATTGATT
13qSubTel               TGAACCAATTCCCCTAATGAGTCTTCTCTCATCTATCTATCTACATATATCCTATTGATT
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       CTGCCTTTCTGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACTACTTTA
10qSubTel               CTGCCTTTATGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACTACTTTA
22qSubTel               CTGCCTTTCTGGAGACCCCTGACTAATGTGATTACAATAACTACACAATTCACTAGTTTA
1pSubTel                CTGCCTTTCTGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACTACTTTA
5qSubTel                CTGCCTTTCTGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACTACTTTA
17qSubTel               CTGCCTTTCTGGAGACCCCTGACTAATGTGATTACAATAACTACACAATTCACTAGTTTA
4qSubTel                CTGCCTTTATGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACCACTTTA
19qSubTel               CTGCCTTTATGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACTAGTTTA
21qSubTel               CTGCCTTTCTGGAGACCCCTGACTAATGTGATTACAATAACTACACAATTCACTAGTTTA
1qSubTel                CTGCCTTTCTGGAGACCCCTGACTAATGTGATTACAATAACTACACAATTCACTAGTTTA
2qSubTel                CTGCCTTTCTGGAGACCCCTGCCTAATGTGATTACAATAACTACAAAATTCACTACTTTA
19p'End'                CTGCCT-TCTGGAGAACCCTGACTAATGTGATTGCAAT-ACTACAAAATTCACTAGTTTA
8p'End'                 CTGCCTTTCTGGAGAACCCTGACTAATGTTATTACAATA-ATACAAAATTCACTAGTTTA
13qSubTel               CTGCCTTTCTGGAGAACCCTGACTAATGTTATTACAATA-ATACTAAATTCACTAGTTTA
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
10qSubTel               TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
22qSubTel               TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
1pSubTel                TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
5qSubTel                TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
17qSubTel               TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
4qSubTel                TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTGTCT
19qSubTel               TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
21qSubTel               TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
1qSubTel                TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
2qSubTel                TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTATATTTGTATTATAACTATGTATCT
19p'End'                TATAGAAGACTTGGTTTTTGTCTTTGCCTCATTTTATATTTGTATTATAACTGTGTAACT
8p'End'                 TATAGAAGACTTGGTTTTTGTCTTTGCCCCATTTTGTATTTGTATTATAACTGTGTATCT
13qSubTel               TGTAGAAGACTTGGTTTTTGTCTTTGCCCAATTTTATATTTGTATTATAACTGTGTATCT
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel      GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
10qSubTel              GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
22qSubTel              GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
1pSubTel               GGAAAGTGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
5qSubTel               GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
17qSubTel              GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
4qSubTel               GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
19qSubTel              GGAAAATGGAACAAGTTT-TTTCTTCTTTATATGAGGGCTAAGGCTTTTTTCTCACCAAT
21qSubTel              GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
1qSubTel               GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
2qSubTel               GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
19p'End'               GGAAAATGGAACAAGTTT-TTTCTTCTTCATATGAGGGCTAAGGCTTTTTTCTCACCAAT
8p'End'                GGAAAATGGAACAAGTTTTTATCTTCTTCATATGAGGGCCAAAGCTTTTTTCTCACCAAT
13qSubTel              GGAAAATGGAACAAGTTTTTATCTTCTTCATATGAGCGCCAAAGCTTTGTCTCACCAAT
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      ATTTTTGGAGATTTTAAAGATTTTCTTTTT-TTTTTGACATAGAATCTTATGGAGGCTGA
10qSubTel              ATTTTTGGAGATTTTAAAGATTTTCTTTTT-TTTTTGACATAGAATCTTATGGAGGCTGA
22qSubTel              ATTTTTGGAGATTTTAAAGATTTTCTTTT-TTTTT-GACATAGAATCTTATGGAGGCTGA
1pSubTel               ATTTTTGGAGATTTTAAAGATTTTC-TTT-TTTTTTGACATAGAATCTTATGGAGGCTGA
5qSubTel               ATTTTTGGAGATTTTAAAGATTTTCTTTTT-TTTT-GACATAGAATCTTATGGAGGCTGA
17qSubTel              ATTTTTGGAGATTTTAAAGATTTTCTTTT--TTTTTGACATAGAATCTTATGGAGGCTGA
4qSubTel               ATTTTTGGAGATTTTAAAGATTTTCTTTT-TTTTTTGACATAGAATCTTATGGAGGCTGA
19qSubTel              ATTTTTGGAGATTTTAAAGATTTTCTTTTTTTTTTTGACATAGAATCTTATGGAGGCTGA
21qSubTel              ATTTTTGGAGATTTTAAAGATTTTCTTTT--TTTTTGACATAGAATCTTATGGAGGCTGA
1qSubTel               ATTTTTGGAGATTTTAAAGATTTTCTTT--TTTTTTGACATAGAATCTTATGGAGGCTGA
2qSubTel               ATTTTTGGAGATTTTAAAGATTTTCTTTT--TTTTTGACATAGAATCTTATGGAGACTGA
19p'End'               ATTTTTGGAGATTTTTAAGGTTTTCTTTT--GTTTTCACATACAATCTTATGGAGGATGA
8p'End'                ATTTTTGGAGATTTTTAAGATTTTCTTTT--GTTTGGACATACAATCTTATGGAGGCTGA
13qSubTel              ATTTTTGGAGATTTTTAAGATTTTCTTTT--GTTTGGACATACAATCTTATGGAGGATGA
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
10qSubTel              GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
22qSubTel              GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
1pSubTel               GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
5qSubTel               GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
17qSubTel              GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
4qSubTel               GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
19qSubTel              GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
21qSubTel              GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
1qSubTel               GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
2qSubTel               GAAATAATT-TTTTTTCTATTTTATTCTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
19p'End'               GAAATATTTTTTTTTTCTATTTTATTTTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
8p'End'                GAAATAAAA-TTTTTTCTATTTTATTTTTCAGCCCCAGATGTTTGCTTTTGCAGATTCTT
13qSubTel              GAAATATTT-TTTTTTCTATTTTATTTTTCAGCCCCAGGTGTTTGCTTTTGCAGATTCTT
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel     GAGCACATTGAGAGC----------TTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
10qSubTel             GAGCACATTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
22qSubTel             GAGCACATTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
1pSubTel              GAGCACATTGAGAGC----------TTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
5qSubTel              GAGCACATTGAGAGC----------TTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
17qSubTel             GAGCACATTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
4qSubTel              GAGCACATTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
19qSubTel             GAGCACATTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
21qSubTel             GAGCACACTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
1qSubTel              GAGCACACTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
2qSubTel              GAGCACATTGAGAGC----------TTCCAAGGCATGGAGTGGGGTGCCTGAAGTTTCAG
19p'End'              GAGCACATTGAGAGC----------CTCCAAGGCATGGAGTGGGGTGCCTAAAGTTTCAG
8p'End'               GAGCACATTGAGAGCATGGAGAGCACTCCAAGGCATGGAGTGGGGTGCCTAAAGTTTCAG
13qSubTel             GAGCACATTGAGAGCATGGAGAGCACTCCAAGGCATGGAGTGGGGTGCCTAAAGTTTCAG
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------

6qPredictedSubTel     TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
10qSubTel             TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
22qSubTel             TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
1pSubTel              TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
5qSubTel              TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
17qSubTel             TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
4qSubTel              TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
19qSubTel             TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
21qSubTel             TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
1qSubTel              TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
2qSubTel              TGATTATAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAGGTCTAAAAGGAGGCAATTTG
19p'End'              TGATTACAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAAGTCTAAAAGGAGGCAATTTG
8p'End'               TGATTACAGGGAGTTGAGAGACTCAACTGGGAAAGGAAAAGTCTAAAAGGAGGCAATTTG
13qSubTel             TGATTACAGGGAGTTGAGATACTCAACTGGGAAAGGAAAAGTCTAAAAGGAGGCAATTTG
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------

6qPredictedSubTel     GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
10qSubTel             GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
22qSubTel             GAAGATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
1pSubTel              GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
5qSubTel              GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
17qSubTel             GAAGATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
4qSubTel              GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
19qSubTel             GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
21qSubTel             GAAGATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
1qSubTel              GAAGATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
2qSubTel              GAAAATAAAAATTTTCTCAAAGGAGCCATTAAAGTTGTAAATAATTCTTAGTAAAGTCAT
19p'End'              GAAGATAAAAATTTTCTCAAAGGAGCAATTAAATTTCTAAATAATTCTTAGTAAAATCAT
8p'End'               GAAGATAAAAATTTTCTCAAAGGAGCGATTAAATTTCTAAATAATTCTTAGTAAAATCAT
13qSubTel             GAAGATAAAAATTTTCTCAAAAGAGCGATTAAATTTCTAAATAATTCTTAGTAAAATCAT
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------
```

```
6qPredictedSubTel      GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
10qSubTel              GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
22qSubTel              GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
1pSubTel               GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
5qSubTel               GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
17qSubTel              GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
4qSubTel               GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
19qSubTel              GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
21qSubTel              GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
1qSubTel               GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
2qSubTel               GCAAACAGGAAAAGAAGTAGAATTAGTTCCATATTGGTGGAACACATAGTCAGCAGAGGT
19p'End'               GTAAACAGGAAAGGAAGTAGAATTAGTTCCATATTGGTGGAACACATA----GCAGAGGT
8p'End'                GTAAACAGGAAAGGAAATAGAATTAGTTCCATATTGGTGGAACACATA----GCAGAGGT
13qSubTel              GCAAACAGGAAAGGAAGTAGAATTAGTTCCATATTGTTGGAACACATAGTCAGCAGAGGT
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TGGAGAAGGGAGAA-TTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
10qSubTel              TGGAGAAGGGAG-AATTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
22qSubTel              TTGAGAAGGGAGAA-TTTAGTGAAGTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
1pSubTel               TGGAGAAGGGAG-AATTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
5qSubTel               TGGAGAAGGGAG-AATTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
17qSubTel              TTGAGAAGGGAGAA-TTTAGTGAAGTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
4qSubTel               TGGAGAAGGGAGAA-TTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
19qSubTel              TGGAGAAGGGAGAA-TTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
21qSubTel              TTGAGAAGGGAGAA-TTTAGTGAAGTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
1qSubTel               TTGAGAAGGGAGAA-TTTAGTGAAGTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
2qSubTel               TGGAGAAGGGAGAA-TTTAGTGAACTGAGAAGTTCCCATGAAAGCAGCAAGATCAAGATC
19p'End'               TTGAGAAGGGAGAA-TTTAGTCAACTGAGAAGTTCTCATGAAAGGAGCAAGTTCAAGATC
8p'End'                TTGAGAAGGGAGAA-TTTAGTCAACTGAGAAGTTCTCATGAAAGGAGCAAGTTCAAGATC
13qSubTel              TTGAGAAGGGAGAA-TTTAGTCAACTGAGAAGTTCCCATGAAAGGAGCAAGTTCAAGATC
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
10qSubTel              ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
22qSubTel              ACAGA--CACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
1pSubTel               ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
5qSubTel               ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
17qSubTel              ACA--GACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
4qSubTel               ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
19qSubTel              ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
21qSubTel              ACA--GACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
1qSubTel               ACA--GACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
2qSubTel               ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
19p'End'               ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
8p'End'                ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCAACCCAAGAGGAGAACAGAGA
13qSubTel              ACAGAGACACCTTGAAACAAAAAGCCAGGAATAACTTCCGACCCAAGAAGAGAACAGAGA
20qSubTel              ------------------------------------------------------------
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel       GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
10qSubTel               GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
22qSubTel               GGCCTCAAAACCAAAGGTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
1pSubTel                GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
5qSubTel                GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
17qSubTel               GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
4qSubTel                GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
19qSubTel               GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
21qSubTel               GGCCTCAAAACCAAAGGTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
1qSubTel                GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
2qSubTel                GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
19p'End'                GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
8p'End'                 GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
13qSubTel               GGCCTCAAAACCAAAGCTAGGATAAGAAACTTGTAGCCCAAGAGTTATCTTCCAGACAAA
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
10qSubTel               GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
22qSubTel               GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTGCTCACTCAAAATGTTACTGAAACTGT
1pSubTel                GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
5qSubTel                GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
17qSubTel               GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTGCTCACTCAAAATGTTACTAAAACTGT
4qSubTel                GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
19qSubTel               GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTGCTCACTCAAAATGTTACTGAAACTGT
21qSubTel               GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTGCTCACTCAAAATGTTACTGAAACTGT
1qSubTel                GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTGCTCACTCAAAATGTTACTGAAACTGT
2qSubTel                GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
19p'End'                GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGT
8p'End'                 GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTACTCACTCAAAATGTTACTGAAACTGA
13qSubTel               GAAGCCTGAGATTCCAACGCAGCTTCAGAGAGTGCTCACTCAAAATGTTACTGAAACTGT
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
10qSubTel               AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
22qSubTel               AGGCTTTTTAATGAGTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
1pSubTel                AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
5qSubTel                AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
17qSubTel               AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
4qSubTel                AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
19qSubTel               AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAT
21qSubTel               AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
1qSubTel                AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
2qSubTel                AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
19p'End'                AGGCTTTTTAATGACTTAGCCATGCATGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
8p'End'                 AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
13qSubTel               AGGCTTTTTAATGACTTAGCCATGCCTGCAAAAGGCATTCCCTAAGGTGGCACAGAAGAC
20qSubTel               ------------------------------------------------------------
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------
```

```
6qPredictedSubTel     GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
10qSubTel             GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
22qSubTel             GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
1pSubTel              GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
5qSubTel              GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
17qSubTel             GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
4qSubTel              GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
19qSubTel             GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
21qSubTel             GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
1qSubTel              GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
2qSubTel              GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
19p'End'              GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
8p'End'               GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
13qSubTel             GGAGCCCCCATATCCAAAGATAGCCAAGGAGAAAGAAAGACCCCTGTTGCCAGAGCCAGT
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------

6qPredictedSubTel     GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGGTGAGATCCTTTCGGATTTA
10qSubTel             GGGAAAAGGTAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
22qSubTel             GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
1pSubTel              GGGAAAAGGTAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
5qSubTel              GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGGTGAGATCCTTTCGGATTTA
17qSubTel             GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
4qSubTel              GGGAAAAGGTAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
19qSubTel             GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
21qSubTel             GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
1qSubTel              GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
2qSubTel              GGGAAAAGGTAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
19p'End'              GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
8p'End'               GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
13qSubTel             GGGCAAAGGCAACAGAAAAGGAGACAAGGGTCCTAATGGGATGAGATCCTTTCGGATTTA
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------

6qPredictedSubTel     GGCTTT--TACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
10qSubTel             GGCTTA--TACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAATTGGGGTACCAAACT
22qSubTel             GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
1pSubTel              GGCTTT--TACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
5qSubTel              GGCTTT--TACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
17qSubTel             GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
4qSubTel              GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAATTGGGGTACCAAACT
19qSubTel             GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
21qSubTel             GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
1qSubTel              GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
2qSubTel              GGCTTT--TACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
19p'End'              GGCTTT--TACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
8p'End'               GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
13qSubTel             GGCTTTTATACAAACTCCTGAGAACTGGCAGGTTGACAGCCATAAATGGGGTACCAAACT
20qSubTel             ------------------------------------------------------------
12qSubTel             ------------------------------------------------------------
7qSubTel              ------------------------------------------------------------
```

```
6qPredictedSubTel      TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGACAATTTC
10qSubTel              TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGATAATTTC
22qSubTel              TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATGAACAAAATGACAATTTC
1pSubTel               TTCTACTCATTGGATTACAAGTTCCCAGGCATCCAGAATGATGAACAAAATGACAATTTC
5qSubTel               TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGACAATTTC
17qSubTel              TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATGAACAAAATGACAATTTC
4qSubTel               TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGATAATTTC
19qSubTel              TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGACAATTTC
21qSubTel              TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGACAATTTC
1qSubTel               TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGACAATTTC
2qSubTel               TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATTAACAAAATGACAATTTC
19p'End'               TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATGAACAAAATGACAATTTC
8p'End'                TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATGAACAAAATGACAATTTC
13qSubTel              TTCTACTCATTGGATTACAAGTTCTCAGGCATCCAGAATGATGAACAAAATGACAATTTC
20qSubTel              -------------------AGTTCTCAAGCATCCAAAATGATTAACAAAATGACAATTTC
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
10qSubTel              TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
22qSubTel              TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTGAACGGA
1pSubTel               TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
5qSubTel               TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
17qSubTel              TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTGAACGGA
4qSubTel               TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
19qSubTel              TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAATGGA
21qSubTel              TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTGAACGGA
1qSubTel               TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTGAACGGA
2qSubTel               TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAATGGA
19p'End'               TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
8p'End'                TAGGGCTTCTGTGGGAGAGTATGGAAAGGTCTTTTTGAACCTTTTAATGCTGTCAACGGA
13qSubTel              TAGGGCTTCTGTGGGAGAGTATGGAAAGATCTTTTTGAACCTTTTAATGCTATCAGTGGA
20qSubTel              TAGGGCTTCTGTGGGAGAGTATGGA-AGGTCTTTTTGAACTTTTTAATGCTGTCAACTAA
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
10qSubTel              AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
22qSubTel              AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
1pSubTel               AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
5qSubTel               AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
17qSubTel              AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
4qSubTel               AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
19qSubTel              AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
21qSubTel              AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCACTTTTATGCTTATTTTT
1qSubTel               AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
2qSubTel               AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGCTTATTTTT
19p'End'               AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGTTTATTTTT
8p'End'                AGAATGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGTTTATTTTT
13qSubTel              AGAACGATGAGGTTCATAAATTTGGAAAGGAGACATTTCTTCATTTTTATGTTTATTTTT
20qSubTel              AGAATGATGAGGTTCATAAATATGGAAAGGAGAGATTTCTATATTTTTGT--TTA-----
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel      AT---TTTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTATG
10qSubTel              AT---TCTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTATG
22qSubTel              AT---TTTTTTT-GAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGCGTG
1pSubTel               ATT--TTTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTATG
5qSubTel               AT---TTTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTATG
17qSubTel              AT----TTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGCATG
4qSubTel               ATT--TTTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTATG
19qSubTel              ATT---TTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTAGG
21qSubTel              AT---TTTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGCAGTGCAATGGCATG
1qSubTel               ATT---TTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGCATG
2qSubTel               ATT---TTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTATG
19p'End'               A----TTTTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGCGTG
8p'End'                A----TTTTTTGTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGCGTG
13qSubTel              ATTTTTATTTTTTGAGACAGAGTTTCACTCTTGTTGCCCAGGCTGGAGTGCAATGGTGTG
20qSubTel              TTTTTATTTTTTTGAGACGGAGTTTCACTCTTCTTGCCCAGGCTGGAGTGCAATGGCGTA
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
10qSubTel              ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
22qSubTel              ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
1pSubTel               ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
5qSubTel               ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
17qSubTel              ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
4qSubTel               ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
19qSubTel              ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
21qSubTel              ATCTTGGTTCACTGCAACCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
1qSubTel               ATCTTGGTTCACTGCAACCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
2qSubTel               ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTGAAGCGATTCTCCTGCCTCAGCCTCC
19p'End'               ATCTTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
8p'End'                ATCTTGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
13qSubTel              ATCTTGGTTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC
20qSubTel              ATCTTGGCTCACTGCAACCTCCACCTCCCAGGTTCAAGTGATTCTCCTGCATCAGCCTCT
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
10qSubTel              TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
22qSubTel              TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
1pSubTel               TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
5qSubTel               TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
17qSubTel              TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
4qSubTel               TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
19qSubTel              TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
21qSubTel              TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
1qSubTel               TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
2qSubTel               TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
19p'End'               TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
8p'End'                TGATTAGCTGGGATTACAGATGCCCACCACCACACCTGGCTAATTTTT---TGCAGTTTT
13qSubTel              TGATTAGCTGGGATTACAGATACCCACCACCACACCTGGATAATTTTTTTTTGTATTTTT
20qSubTel              CGATTAGCTGGGATTACAGACCCTCACCACTATGCCTGGCTAATTTT----TGTTTTTTT
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel      GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
10qSubTel              GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTCACCTCAGGTGA
22qSubTel              GGTAGAGACAGGATTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
1pSubTel               GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
5qSubTel               GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
17qSubTel              GGTAGAGACAGGATTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
4qSubTel               GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
19qSubTel              GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
21qSubTel              GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
1qSubTel               GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
2qSubTel               GGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
19p'End'               AGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCGTGACCTCAGGTGA
8p'End'                AGTAGAGACAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCGTGAGCTCAGGTGA
13qSubTel              AGTAGAGACAGGGTGTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
20qSubTel              AGTAGAGATGAGGTTTCACCATGTTGGCCAGTCTGGTCTTGAACGCCTGACCTCAGGTGA
12qSubTel              ---AGAGGAAGGGTTTCATCATGTTGGCCAGGCTGGTCTGAAACTCCTGACCTCAGGTGA
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCA-CCCACCCAGTG
10qSubTel              TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
22qSubTel              TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
1pSubTel               TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
5qSubTel               TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCA-CCCACCCAGTG
17qSubTel              TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
4qSubTel               TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
19qSubTel              TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
21qSubTel              TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
1qSubTel               TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
2qSubTel               TCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCA-CCCACCCAGTG
19p'End'               TCCACTCGCCTTAGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCCA-CCCAACCAGTG
8p'End'                TCCACTTGCCTTAGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCCA-CCCAACCAGTG
13qSubTel              TCCACCTGCCTTGGCCTCCCAAAGTGCTGGGTTTACAGGCATGAGCCACCACACCCAATG
20qSubTel              TCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTATAGGC-------ACTGCACCCAGTG
12qSubTel              TCCACCTGCCTCAGCATCCCAAAGTGCTGGGATTACAGGCATGAGCCATCACAACCAGTG
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
10qSubTel              AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
22qSubTel              AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
1pSubTel               AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
5qSubTel               AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
17qSubTel              AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
4qSubTel               AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
19qSubTel              AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
21qSubTel              AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
1qSubTel               AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
2qSubTel               AGAGATTTATTTTCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
19p'End'               AGAGATTTATTTCCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
8p'End'                AGAGATTTATTTCCTATAAAGGGTTGTAGCCTGCAGGGTTGTCCTTCTGACAGGCTGGGA
13qSubTel              AGAGATTTATTTCTTATAAAGGGTTGCAGCCTGAAGGGTTGTCCTTCTGACAGGCTGGGA
20qSubTel              AGGTATTTATTTCCTGTAAAGGGTTGCAGCCTGCAGGGTAGTCCTTCTGATAGGCTGGGA
12qSubTel              AGAGATTTGTTTCCTATAAAGGGTTGCAGCCTTCAGGGTTGTCCTTCTGACAGACTGGGA
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel     AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGTTTCAAGGAGGAGGTAAAGGAAATA
10qSubTel             AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
22qSubTel             AGCATAACCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
1pSubTel              AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
5qSubTel              AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGTTTCAAGGAGGAGGTAAAGGAAATA
17qSubTel             AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
4qSubTel              AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
19qSubTel             AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
21qSubTel             AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
1qSubTel              AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
2qSubTel              AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
19p'End'              AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GACTCTTCAAGGAGGAGGTAAAGGAAATA
8p'End'               AGCATAGCCTCCAGCCAGAAGCCAGAAACA-GACTCTTCAAGGAGGAGGTAAAGGAAATA
13qSubTel             AGCACAGCCTCCAGCCAGAAGCCAGAAACA-GATGCTTCAAGGAGGAGGTAAAGGAAATA
20qSubTel             AGCATAGCCTCCAGCCAGAAACTAGAAACAAGACACTTCAAGGAAGAGGTGAAGGAAACA
12qSubTel             AGCATAGCTTCCAGGCAGAAGCCAGAAACG-GATGCTTCAAGGAGGATGCAAAGGAAATA
7qSubTel              --------------------------------------------------------AATA
                                                                              ** *

6qPredictedSubTel     GCAATTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
10qSubTel             GCAATTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
22qSubTel             GCAACTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
1pSubTel              GCAATTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
5qSubTel              GCAATTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
17qSubTel             GCAACTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
4qSubTel              GCAATTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
19qSubTel             GCAACTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
21qSubTel             GCAACTTATGCTGAGTGGAATGGCCAATTAGATTTATTTAATAAGCTCTAGGAGGAGTCA
1qSubTel              GCAACTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
2qSubTel              GCAATTTATGCTGAGTGGAATGGCCAAATAGATTTATTTAATAAGCTCTAGGAGGAGTCA
19p'End'              GCAATTTATGCTGACTGGAATGGCCAAATACATTTATTTAATAAGCTCCAGGAGGAGTCA
8p'End'               GCAATTTATGCTGACTGGAATGGCCAAATACATTTATTTAATAAGCTCCAGGAGGAGTCA
13qSubTel             GTAATTTATGCTGATTGGAATGGCCAAATGCATTTATTTAATAAGCTCTAGGAGGAGTCA
20qSubTel             GTAATTTATGCTGAGTGGCATGGCCAAATACACATCTTTAATAAGCTCTAGGCGGAGTCA
12qSubTel             GTAATTTATGCTGAGTGGAATGGCCAAATACATTTATTTAATATGCTCTAGGAGGAGTCA
7qSubTel              GTAATTTATGCTGAGTGGAATGGCCAAATACATTTATTTAATATGCTCTAGGAGGAGTCA
                      * ** ********* ***.********:*. * :*.*******:**** ***.*******

6qPredictedSubTel     TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
10qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
22qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
1pSubTel              TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
5qSubTel              TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
17qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
4qSubTel              TGAATATTTATGGAAGGAGAAATGCATGCACACACA-ATTGAGTTTCTTGCTTCTTCATG
19qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
21qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
1qSubTel              TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
2qSubTel              TGAATATTTATGGAAGGAGAAATGCATGCACGCACA-ATTGAGTTTCTTGCTTCTTCATG
19p'End'              TGAATATTTATGGAAGGAGAAATGCATGCATGCACA-ATTGAGTTTCTGGCTTCTTCATG
8p'End'               TGAATATTTATGGAAGGAGAAATGCATGCATGCACA-ATTGAGTTTCTGGCTTCTTCATG
13qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCATGTGCA-ATTGAGTTTCTTGCTTCTTCATG
20qSubTel             TGAATATTTATGAAAGGGGAAATGCGTGCATGCACA-ACTGAGTTCCTTGGTTCTTCATG
12qSubTel             TGAATATTTATGGAAGGAGAAATGCATGCATGCTTTAATTGAGTTTCTTGCTTCTTCATG
7qSubTel              TGAATATTTATGGAAGGAGAAATGCATGCATGCGCA-ATTGAGTTTCTTGCTTCTTCATG
                      ************.****.*******.**** .   : * ****** ** * *********
```

```
6qPredictedSubTel       GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
10qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
22qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
1pSubTel                GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
5qSubTel                GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
17qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
4qSubTel                GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
19qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
21qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
1qSubTel                GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
2qSubTel                GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
19p'End'                GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTGGAGTTTTCAGCCCTC
8p'End'                 GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCAGGGTTGAGTTTTCAGCCCTC
13qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTATCATGATCCCAGGGTGGAGTTTTCAGCTATC
20qSubTel               GGTCCCATGTACAAAAAATAGCAGTGTTAGCATCATCCCAGGGTGGAGTTTTCAGCCCTC
12qSubTel               GGTCCCATGTACAAAAAATGGCAGTGTTAGCATGATCCCATGGTGGAGTTTTCAGCCCTC
7qSubTel                GGTCCCATGTACAAAAAATGGCGGTGTTAGCATGATCCCATGGTGGAGTTTTCAGCCCTC
                        *******************.**.****** *** ****** *** *********** .**

6qPredictedSubTel       TGACATTAAAAGGTGAAGCAGAGGAAATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
10qSubTel               TGACATTAAAAGGTGAAGCAGAGGAAATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
22qSubTel               TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
1pSubTel                TGACGTTAAAAGGTGAAGCAGAGGAAATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
5qSubTel                TGACATTAAAAGGTGAAGCAGAGGAAATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
17qSubTel               TGACATTAAAAGGTGAAGTAGAGGACATGAAAACTTGCTCTGTGCATCCTCTGTACGCTG
4qSubTel                TGACATTAAAAGGTGAAGCAGAGGAAATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
19qSubTel               TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTTGCTCTGTGCATCCTCTGTACGCTG
21qSubTel               TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTTGCTCTGTGCATCCTCTGTACACTG
1qSubTel                TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTTGCTCTGTGCATCCTCTGTACGCTG
2qSubTel                TGACATTAAAAGGTGAAGCAGAGGAAATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
19p'End'                TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
8p'End'                 TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTCGCTCTGTGCATCCTCTGTACGCTG
13qSubTel               TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTCGCTCTGTGCATCCTCTGTACACTG
20qSubTel               TGACATTAAAAGGTGAAGCAGAGGACATGAAAACTCGCTCTGTGCATCCTCTGTGTGCTG
12qSubTel               TGACATTAAAAGGTGAAGCAGAGTACATGAAAACTCGCTCTGTGCATCCTCTGTATGCTG
7qSubTel                TGACATTAAAAGGTGAAGCAGAGTACATGAAAACTCGCTCGGTGCATCCTCTGTATGCTG
                        ****.************* **** *.********* **** *************. .***

6qPredictedSubTel       GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
10qSubTel               GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
22qSubTel               GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
1pSubTel                GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
5qSubTel                GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
17qSubTel               GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGGAAGAAAGGAGAGGTTGATATCA
4qSubTel                GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
19qSubTel               GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
21qSubTel               GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
1qSubTel                GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
2qSubTel                GCCAGAACCTCTCCATCGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
19p'End'                GCCAGAACCTCTCTGTTGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
8p'End'                 GCCAGAACCTCTCTGTTGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCG
13qSubTel               GCCAGAACCTCTCCATTGTGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
20qSubTel               TCCAGAACCTCTCCGTCATGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
12qSubTel               GCCAGAACCTCTCCATCATGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
7qSubTel                GCCAGAACCTCTCCATCATGGGTGGTCTCTTATCAGGCAAGAAAGGAGAGGTTGATATCA
                         ************ .* .****************** **********************.
```

```
6qPredictedSubTel          GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
10qSubTel                  GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
22qSubTel                  GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
1pSubTel                   GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
5qSubTel                   GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
17qSubTel                  GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
4qSubTel                   GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
19qSubTel                  GTAGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
21qSubTel                  GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
1qSubTel                   GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
2qSubTel                   GTGGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
19p'End'                   GTAGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
8p'End'                    GTAGTGGAGGCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
13qSubTel                  GTGGTAGAGCCTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
20qSubTel                  GTGGT---GATTTTGAAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
12qSubTel                  GTGGTGTAGCCTTTGGAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
7qSubTel                   GTGGTGTAGCCTTTGGAAGGGCTGGTTTCTGTTAAATCCTTAGGGAAGAAAGCCTCATCA
                           **.**   *  ****.********************************************

6qPredictedSubTel          TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
10qSubTel                  TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCGCATCCTAGCA
22qSubTel                  TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
1pSubTel                   TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCGCATCCTAGCA
5qSubTel                   TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
17qSubTel                  TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
4qSubTel                   TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCGCATCCTAGCA
19qSubTel                  TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
21qSubTel                  TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
1qSubTel                   TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCCCATCCTAGCA
2qSubTel                   TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTTAACCCCATCATCGCATCCTAGCA
19p'End'                   TGGTTAGCAAAGGAGTGGGTTTAACGATGTGTATGTTAACCCCATCATCCCATCCTAGCA
8p'End'                    TGGTTAGCAAAGGAGTGGGTTTAACGATGTGTATGTTAACCCCATCATCCCATCCTAGCA
13qSubTel                  AGGTTAGCAAAGGAAGGGGTATAACGATGTGTATCTGACCTCCATCATCCCATCCTAGCA
20qSubTel                  TGGTTTGCAAAGCAGGGGGTATAATGGGGTGTACCTGACTCCCATCATCCCATTCTGGCC
12qSubTel                  TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTGACCCCCATCATCCCATCCTAGCA
7qSubTel                   TGGTTAGCAAAGGAGGGGGTATAACGATGTGTATCTGACCCCCATCATCCCATCCTAGCA
                           :****:****** *. ****:*** *. *****  * *.  ******** *** **.**.

6qPredictedSubTel          AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
10qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
22qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
1pSubTel                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
5qSubTel                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
17qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
4qSubTel                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
19qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
21qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
1qSubTel                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
2qSubTel                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
19p'End'                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGTGGTCCCCTCAGCCAAGAGTGGTTCTGT
8p'End'                    AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGTGGTCCCCTCAGCCAAGAGTGGTTCTGT
13qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
20qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
12qSubTel                  AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
7qSubTel                   AAGCTGAGAACTCAGTTTTGAAAGTTACTCTGGGGTCCCCTCAGCCAAGAGTGGGTCTGT
                           ******************************** ********************* *****
```

```
6qPredictedSubTel      TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGCGAAAGAGTAC
10qSubTel              TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGCGAAAGAGTAC
22qSubTel              TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
1pSubTel               TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGCGAAAGAGTAC
5qSubTel               TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGCGAAAGAGTAC
17qSubTel              TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
4qSubTel               TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGCGAAAGAGTAC
19qSubTel              TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
21qSubTel              TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
1qSubTel               TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
2qSubTel               TCAGTCAGTTGGGAGCTTAGAATTTAATTTTCATTTATCAATGCTAATGCGAAAGAGTAC
19p'End'               TCAGTCAGTTGGGAGCTTATAATTTAATTTTTGTTTATCAATGCTAATGCGAAAGAGTAC
8p'End'                TCAGTCAGTTGGGAGCTTATAATTTAATTTTTGTTTATCAATGCTAATGCGAAAGAGTAC
13qSubTel              TCAGTCAGTTGGGAGCTTAGGATTTTATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
20qSubTel              TCAGTCAGCTGGGAGCTTAGGATTTCATTTTCATTTATCAATGCTAATGGGAATGAGCAG
12qSubTel              TCAGTCAGTTGGGAGCTTAGGATTTCATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
7qSubTel               TCAGTCAGTTGGGAGCTTAGGATTTCATTTTCATTTATCAATGCTAATGGGAAAGAGTAC
                       ******** ********** .**** ***** .**************** ***:*** * 

6qPredictedSubTel      GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
10qSubTel              GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
22qSubTel              GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
1pSubTel               GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
5qSubTel               GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
17qSubTel              GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
4qSubTel               GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
19qSubTel              GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
21qSubTel              GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
1qSubTel               GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
2qSubTel               GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
19p'End'               GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
8p'End'                GCTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
13qSubTel              GCTGTGTCTATGGCAGCTGAATTTGCAAGAAAC---------------------------
20qSubTel              TCTATCTTCATGGCAGCTGAATTTGCAAGAAACTTCGTGGATGGGGTTAACGGCAGCTGT
12qSubTel              ACTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
7qSubTel               ACTGTCTTCATGGCAGCTGAATTTGCAAGAAAC---------------------------
                        **.* *  ************************

6qPredictedSubTel      -----------------------------------------------------TCCTTGG
10qSubTel              -----------------------------------------------------TCCTTGG
22qSubTel              -----------------------------------------------------TCCTTGG
1pSubTel               -----------------------------------------------------TCCTTGG
5qSubTel               -----------------------------------------------------TCCTTGG
17qSubTel              -----------------------------------------------------TCCTTGG
4qSubTel               -----------------------------------------------------TCCTTGG
19qSubTel              -----------------------------------------------------TCCTTGG
21qSubTel              -----------------------------------------------------TCCTTGG
1qSubTel               -----------------------------------------------------TCCTTGG
2qSubTel               -----------------------------------------------------TCCTTGG
19p'End'               -----------------------------------------------------TCCTTGG
8p'End'                -----------------------------------------------------TCCTTGG
13qSubTel              -----------------------------------------------------TCCTTGG
20qSubTel              ATTTTCTGGGCGCTCCATGGATGGGGTTAACGGCAGCTGAATTTGCAAGGAACTCCGTGG
12qSubTel              -----------------------------------------------------TCCTTGG
7qSubTel               -----------------------------------------------------TCCTTGG
                                                                            *** ***
```

```
6qPredictedSubTel      ATGGGGTTAATGGCAGCTGTATTTTACTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
10qSubTel              ATGGGGTTAATGGCAGCTGTATTTTACTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
22qSubTel              ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTGTGCTTTAATTGGATAAAGTAAGTT
1pSubTel               ATGGGGTTAATGGCAGCTGTATTTTACTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
5qSubTel               ATGGGGTTAATGGCAGCTGTATTTTACTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
17qSubTel              ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTGTGCTTTAATTGGATAAAGTAAGTT
4qSubTel               ATGGGGTTAATGGCAGCTGTATTTTACTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
19qSubTel              ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTGTGCTTTAATTGGATAAAGTAAGTT
21qSubTel              ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTGTGCTTTAATTGGATAAAGTAAGTT
1qSubTel               ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTGTGCTTTAATTGGATAAAGTAAGTT
2qSubTel               ATGGGGTTAATGCCAGCTGTATTTTACTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
19p'End'               ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTCTGCTTTAATTGGACAAAGTAAGTT
8p'End'                ATGGGGTTAATGGCAGCTGTATTTTTCTGGGAGCTCTGCTTTAATTGGACAAAGTAAGTT
13qSubTel              ATGGGGTTAATGGCAGTTGTATTTTTCTGGGAGTTCTGCTTTAATTGGATAAAGTAAGTT
20qSubTel              ATGGGGTTAATGGCAGCTGTAT-TTTCTGGGAGCTGTGCTTTAATTGGATAAAGTAAGTT
12qSubTel              ATGGGGTTAA-TGCAGCTGTATTTTTCTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
7qSubTel               ATGGGGTTAA-TGCAGCTGTATTTTTCTGGGAGCTCTGCTTTAATTGGATAAAGTAAGTT
                       **********   *** ***** **:******* * ************* **********

6qPredictedSubTel      CTGGTAAGATTTCTTC----TTCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
10qSubTel              CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
22qSubTel              CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
1pSubTel               CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
5qSubTel               CTGGTAAGATTTCTTC----TTCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
17qSubTel              CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
4qSubTel               CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
19qSubTel              CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
21qSubTel              CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
1qSubTel               CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
2qSubTel               CTGGTAAGATTTCTTC----ATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
19p'End'               GTGGTAAGATTTCTTCTTTTATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
8p'End'                GTGGTAAGATTTCTTCTTTTATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTTTA
13qSubTel              CTGGTAAGATTTCTTCCTTTATCTTCAGTATCTCAAGTGTTTTCACTTAAATAATCTTTA
20qSubTel              CTGGTAAGATTTCTTCCTTCATCTTCGGTATCTCAAATGTTTTCATTTAAATAATCTTTA
12qSubTel              CTGGTAAGATTTCTTCCTTTATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTGTG
7qSubTel               CTGGTATGATTTCTTCCTTTATCTTCAGTATCTCAAATGTTTTCATTTAAATAATCTGTG
                        *****:*********    :*****.*********.******** *********** *.

6qPredictedSubTel      TAACAACTTTTGATGTCTGAGTGGAGTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
10qSubTel              TAACAACTTTTGATGTCTGAGTGGAGTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
22qSubTel              TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
1pSubTel               TAACAACTTTTGATGTCTGAGTGGAGTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
5qSubTel               TAACAACTTTTGATGTCTGAGTGGAGTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
17qSubTel              TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
4qSubTel               TAACAACTTTTGATGTCTGAGTGGAGTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
19qSubTel              TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
21qSubTel              TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
1qSubTel               TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
2qSubTel               TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGGCTTTCTGA
19p'End'               TAACAACTTTTGATGTCTGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
8p'End'                TAACAACTTTTGATGTCTGAGTGGATTCCCATACAGTCATCTATTGT-AAGACTTTCTGA
13qSubTel              TAACAACTTTTGATGTCTGAGTGGAGTCCCACACAGTCGTCTATTGT-AAGACTTTCTGA
20qSubTel              TAATAACTCTTGATGTCTGAGGGGGTTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
12qSubTel              TAAAAATTTTTGATGTCAGAGTGGATTCCCACACAGTCATCTATTGT-AAGACTTTCTGA
7qSubTel               TAACAATTTTTGATGTCAGAGTAGATTCCCACACAGTCATCTATTGTTAAGACTTTCTGA
                       *** ** * ********:*** .*. ***** ******.******** ***.********
```

```
6qPredictedSubTel    TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
10qSubTel            TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
22qSubTel            TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
1pSubTel             TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
5qSubTel             TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
17qSubTel            TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
4qSubTel             TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
19qSubTel            TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
21qSubTel            TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
1qSubTel             TTCATTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
2qSubTel             TTCCTTTTTTTTCCTTTGGTCATTATGAATAGGGCTTCTGTAAATAACTGCATGGTAGCT
19p'End'             TTCCGTTTTTTTCCTTTGGTCATTATGAATAGGGCTTATGTAAATAATTGCATGGTAGCT
8p'End'              TTCCGTTTTTTTCCTTTGGTCATTATGAATAGGGCTTATGTAAATAATTGCATGGTAGCT
13qSubTel            TTCCTTTTTTTTCCTTTAGTCATTCTGAATAGGGCTTCTGTAAATAATTGCATGGTAGCT
20qSubTel            TTCCTTT-TTTTCCTTTGGTCATTATGAATAGGGCTTCTGTACATAATTGCATGGTAGCT
12qSubTel            TTCCTTTTTTTTCCTCTGGTCATTATGAATAGGGCTTCTGTAAATAATGGCATGGTAGCT
7qSubTel             TTCCCTTTTTTTCCTCTGGTCATTATGAATAGGGCTTATGTAAATAATGGCATGGTAGCT
                     ***. ** ******* *.******.************.****.****  ***********

6qPredictedSubTel    TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
10qSubTel            TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
22qSubTel            TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACCTAGGAATGTTATTTTTGGATT
1pSubTel             TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
5qSubTel             TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
17qSubTel            TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACCTAGGAATGTTATTTTTGGATT
4qSubTel             TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATCCTTAGGAATGTTATTTTTGGATT
19qSubTel            TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACCTAGGAATGTTATTTTTGGATT
21qSubTel            TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACCTAGGAATGTTATTTTTGGATT
1qSubTel             TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACCTAGGAATGTTATTTTTGGATT
2qSubTel             TTTGATGGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
19p'End'             TTTGATTGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
8p'End'              TTTGATTGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTTATTTTTGGATT
13qSubTel            TTTGTTTGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTGATTTTTGGATT
20qSubTel            TTTGTTTGGAAATCACATCAAAGTGGTTGTCAAAATACTTAGGAATGTCGTTTTTGGATT
12qSubTel            TTTGATCGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTGATTTTTGGATT
7qSubTel             TTTGATTGGAAATAACATCAAAGTAGTTGTCAAAATACTTAGGAATGTGATTTTTGGATT
                     ****:* ******.**********.***********.* ********* .**********

6qPredictedSubTel    GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
10qSubTel            GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
22qSubTel            GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGACCAACTTGTAATAGGGGAGGAAA
1pSubTel             GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
5qSubTel             GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
17qSubTel            GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
4qSubTel             GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
19qSubTel            GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
21qSubTel            GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
1qSubTel             GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
2qSubTel             GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
19p'End'             GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
8p'End'              GTAAGGTGAGACTTGTTTAGCTTTGGAA-AAAAATGCCCAACTTGTAATAGGGGAGGAAA
13qSubTel            GTAAGGTGAGACTTGTTTAGCTTTAGAAAAA-ACTGCCGAAATTT-TAATGGGGAGGAA-
20qSubTel            GTAAGGTGAGACTTATTTAGCTTTGGAA-AAAACTGCCCAACTTGTAATAGGTGAGGAAA
12qSubTel            GTAAGGTGAGACTTGTTTAGCTTTGGGGAAAAAATGCCCAAGTTGTAATAGGGGAGGAAA
7qSubTel             GTAAGGTGAGACTTGTTTAGCTTTGGGGAAAAAATGCCCAACTTGTAATAGGGGAGGAAA
                     **************.*********.*.. ** *.**.* ** **   :*::** ******
```

```
6qPredictedSubTel        AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAAGTGACAG
10qSubTel                AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
22qSubTel                AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
1pSubTel                 AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAAGTGACAG
5qSubTel                 AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAAGTGACAG
17qSubTel                AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
4qSubTel                 AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
19qSubTel                AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
21qSubTel                AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
1qSubTel                 AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAACTGACAG
2qSubTel                 AATAATTTTC------TGTTTTTGGAATTCTTAGATGGAACGCTCTGTAAAAAGTGACAG
19p'End'                 AATAATTTTC------TGTTTTCAGAATTCTTAGATGGGATGCTCTGTAAAAACTGACAG
8p'End'                  AATAATTTTC------TGTTTTCGGAATTCTTAGATGGGATGCTCTGTAAAAACTGACAG
13qSubTel                AAAAATTTTC----TATGTTTTTGGAATTCTTAGATGGGACCCGCTGTAAAAACTGACAG
20qSubTel                AATAATTTTCTCTCTATCTTTTCAGAATTCTTAGATTGGACCCTCTGTAACAAATGACAG
12qSubTel                AATAATTTTC----TATGTTTTCAGAATTCTTAGATGGGACCCTCTGTAAAAACTGACGG
7qSubTel                 AATAATTTTC----TATGTTTTCAGAATTCTTAGATGGGACCTTCTGTAAAAACTGACGG
                         **:*******      * **** .************ *.*    ******.** ****.*

6qPredictedSubTel        ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
10qSubTel                ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
22qSubTel                ATTAAAATGAGAAAA--AGAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
1pSubTel                 ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
5qSubTel                 ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
17qSubTel                ATTAAAATGAGAAAA--AGAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
4qSubTel                 ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
19qSubTel                ATTAAAATGAGAAAA--AGAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
21qSubTel                ATTAAAATGAGAAAA--AGAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
1qSubTel                 ATTAAAATGAGAAAA--AGAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
2qSubTel                 ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATATCTTATGGTTACATGGGA
19p'End'                 ATTAAAATGAGAAAA--ACAGAAAAGTAGAAAAACACGTATACCTTATGGTTACATGGGA
8p'End'                  ATTAAAATGAGAAAA--ACAGAAAAGTTGAAAAACACGTATACCTTATGGTTACATGGGA
13qSubTel                ATTAAAATGAGAAAA--ACAGAAAAGTTTAAAAACATGTATACCTTATGGATACATGGGA
20qSubTel                ATTAAAATGAGAAAAATGTAGAAAAGTTTAAAAACATGTATACCTTATGGATACATGGGC
12qSubTel                ATTAAAATTAGAAAA--ATAGAAAAGTTTAAAAACATGTATACCTTATGGATACATGGGA
7qSubTel                 ATTAAAATTAGAAAA--ATAGAAAAGTTTAGAAACATGTATACCTTATGGATACATGGGA
                         ******** ******  . ********: *.***** ***** *******:********.

6qPredictedSubTel        TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
10qSubTel                TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
22qSubTel                GATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
1pSubTel                 TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
5qSubTel                 TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
17qSubTel                GATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
4qSubTel                 TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
19qSubTel                TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
21qSubTel                GATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
1qSubTel                 GATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
2qSubTel                 TATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
19p'End'                 GATACTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAGTAC
8p'End'                  GATATTCAGGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
13qSubTel                GATACTCAAGGAAAAATGAGTAAATCTCCAACAGGTGGCTTTCAATTAAAGCATAAATAC
20qSubTel                GATACTCAGGGAAAAATGAGTAAATCTCCAGCAGGTGGTTTTCAATTCAAGTGTAAATAC
12qSubTel                GATACTCAGGAAAAAACGAGTAAATCTCCAACAGGTGGCTTTCAATTCAAGCATAAATAC
7qSubTel                 GATACTCAGGAAAAAACGAGTAAATCTGCAACAGGTGGCTTTCAATTCAAGCATAAATAC
                          *** ***.*.***** ********** **.******* ********.***  .***.***
```

```
6qPredictedSubTel        TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
10qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
22qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGGAGTTAACCAGCA
1pSubTel                 TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
5qSubTel                 TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
17qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGGAGTTAACCAGCA
4qSubTel                 TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
19qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
21qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGGAGTTAACCAGCA
1qSubTel                 TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGGAGTTAACCAGCA
2qSubTel                 TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGAAGTTAACCAGCA
19p'End'                 TGTCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAT----TGGGGAGTTAACCAGCA
8p'End'                  TGTCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAT----TGGGGAGTTAACCAGCA
13qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTGGTGAGTGGGGAGTTAGCCAGCA
20qSubTel                TATCTTCAACTTGAA-------GATTTGAGGCACAGTAG----TGGGGAGTTAACCAGCA
12qSubTel                TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGGATCTAACCAGCA
7qSubTel                 TATCTTCAACTTAAAGAAAGAAGATTTGAGGTGCAGTAG----TGGGGATCTAACTAGCA
                         *.**********.**       ********* .****.     ****.*  **.* ****

6qPredictedSubTel        AAAGCACATTAGACAGGGGTAAGGTTCATTATACAGAGTTAAGTCCATGCATTCTCCATT
10qSubTel                AAAGCACATTAGACAGGGGTAAGGTTCGTTATACAGAGTTAAGTCCATGCATTCTCCATT
22qSubTel                AAAGCACATTAGACAAGGGTAAGGTTCGTTATACAGACTTAAGTCCATGCATTCTCCATT
1pSubTel                 AAAGCACATTAGACAGGGGTAAGGTTCATTATACAGAGTTAAGTCCATGCATTCTTCATT
5qSubTel                 AAAGCACATTAGACAGAGGTAAGGTTCATTATACAGAGTTAAGTCCATGCATTCTCCATT
17qSubTel                AAAGCACATTAGACAAGGGTAAGGTTCGTTATACAGACTTAAGTCCATGCATTCTCCATT
4qSubTel                 AAAGCACATTAGACAGGGGTAAGGTTCGTTATACAGAGTTAAGTCCATGCATTCTCCATT
19qSubTel                AAAGCACATTAGACAGGGGTAAGGTTTGTTATACAGAGTTAAGTCCATGCAGTCTCCATT
21qSubTel                AAAGCACATTAGACAAGAGTAAGGTACGTTATACAGACTTAAGTCCATGCATTCTCCATT
1qSubTel                 AAAGCACATTAGACAAGGGTAAGGTTCGTTATACAGACTTAAGTCCATGCATTCTCCATT
2qSubTel                 AAAGCACATTAGACAGGGGTAAGGTTCATTATACAGAGTTAAGTCCATGCATTCTTCATT
19p'End'                 AAAGCACATTAGACAAGGGTAAGGTTCGTTATACAGACTTAAGTCCATGCATTCTCCATT
8p'End'                  AAAGCACATTAGACAAGGGTAAGGTTCATTATACAGACTTAAGTCCATGCATTCTCCATT
13qSubTel                AAAGCACATTAGACAAGGGTAAGGTTCGTTATACAGACTTAAGTCCATGCATTCTCCATT
20qSubTel                AAACCACATTAGACAAGAGTAACGTTTGTTATGCAGACTGAAGTCCATGCATCCTCCATT
12qSubTel                AAAGCACAGTAGACAAGGGTAAGGTTTGTTATACAGACTGAAGTCTATGCATTCTCCATT
7qSubTel                 AAAGCACAGTAGACAAGGGTAAGGTTTGTTATACAGACTGAAGTCTATGCATTCTCCATT
                         *** **** ******...**** **: .****.**** * ***** *****  ** ****

6qPredictedSubTel        GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
10qSubTel                GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
22qSubTel                GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
1pSubTel                 GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
5qSubTel                 GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
17qSubTel                GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTTGAGA--GAGGTAGCT
4qSubTel                 GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
19qSubTel                GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
21qSubTel                GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
1qSubTel                 GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
2qSubTel                 GATAAGACTCTTCAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
19p'End'                 GATAAGACTCTTTAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
8p'End'                  GATAAGACTCTTTAGTGATTTAGTTATCCTTCTCTTCTTGGTGTCGAGA--GAGGTAGCT
13qSubTel                GATAAGAGTCTTTAGTGATTTAGTTATCCT-----TCTTGGTGTCGAGA--GAGGTAGCT
20qSubTel                GATAAGACTCT--AGTGATTTAGTTATCCTTCTCTTCTTGGTGCCGAGAGAGAGGTAGCT
12qSubTel                GATAAGACTCT--AGTGATTTACTTATCCTTCTCTTCTTGATGTCGACA--GAGGTAGCT
7qSubTel                 GATAAGACTCT--AGTGATTTACTTATCCTTCTCTTCTTGATGTTGACA--GAGGTAGCT
                         ******* ***  ********* *******     *****.**  ** *  *********
```

```
6qPredictedSubTel      TTTAAATGGTGATTTCCTTTATAGGTGTAAATTTTCCTTACACAAGT--AACTTTTACTC
10qSubTel              TTTAAATGGTGATTTCCTTTATAGGTGTAAATTTTCCTTACACAAGT--AACTTTTACTC
22qSubTel              TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGT--AACTTCTACTC
1pSubTel               TTTAAATGGTGATTTCCTTTATAGGTGTAAATTTTCCTTACACAAGT--AACTTTTA--C
5qSubTel               TTTAAATGGTGATTTCCTTTATAGGTGTAAATTTTCCTTACACAAGT--AACTTTTACTC
17qSubTel              TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGT--AACTTCTACTT
4qSubTel               TTTAAATGGTGATTTCCTTTATAGGTGTAAATTTTCCTTACACAAGT--AACTTTTACTC
19qSubTel              TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGT--AACTTCTACTC
21qSubTel              TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGT--AACTTCTACTC
1qSubTel               TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGT--AACTTCTACTC
2qSubTel               TTTAAATGGTGATTTCCTTTATAGGTGTAAATTTTCCTTACACAAGT--AACTTTTACTC
19p'End'               TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGTGTAACTTCTACTC
8p'End'                TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGT--GT-AACTTCTC
13qSubTel              TTTAAATGGTGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGTGTAACTTCTACTC
20qSubTel              TTTAAATGGGGATTTCCTTTATAGATGTAAATTTTCCTTACAGAAGGGTAACTTCTACTC
12qSubTel              TTTAAATGGGGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGGGTAACTTCTACTC
7qSubTel               TTTAAATGGGGATTTCCTTTATAGATGTAAATTTTCCTTACACAAGGGTAACTTCTACTC
                       ********* **************.***************** ***    .: :: *:

6qPredictedSubTel      TGTTTTCACAACTTCCTTTGTTAGCATTTTTTTTTT-CAAAATAATTAGCTTGGAATAAT
10qSubTel              TGTTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
22qSubTel              TATTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
1pSubTel               TGTTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
5qSubTel               TGTTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
17qSubTel              TATTTTCACAACTTCCTTTGTTAGCATTTTTTTTTTCCAAAATAATTAGCTTGGAATAAT
4qSubTel               TGTTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
19qSubTel              TGTTTTCACAACTTCCTTTGTTAGCATTTTT--TTTTCAAAATAATTAGCTTGGAATAAT
21qSubTel              TATTTTCACAACTTCCTTTGTTAGCATTTTTTTTTTTCAAAATAATTAGCTTGGAATAAT
1qSubTel               TATTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
2qSubTel               TGTTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
19p'End'               TGTTTTCACAACTTCCTTTGTTAGCGTTTTTTTTTTTCAAAATAATTAGCTTGGAATAAT
8p'End'                TGTTTTCACAACTTCCTTTGTTAGCATTTTTTTTTTTCAAAATAATTAGCTTGGAATAAT
13qSubTel              TGTTTTCACAACTTCCTTTGTTAGCATTTTTT-TTTTCAAAATAATTAGCTTGGAATAAT
20qSubTel              TGGTTTTGGAACTTCCTTTGTTAGCATTTTT---TTTCAAAATAATCATCTTGGAATAAT
12qSubTel              TGTTTTCAGAACT----TTGTTAGCATTTTTTT--------TAAAATAGCTTGGAATAAT
7qSubTel               TGTTTTCAGAACG----TTGTTAGCATTTTTTT--------TAAAATAGCTTGGAATAAT
                       *. *** . ***     ********.*****          ::**: * ***********

6qPredictedSubTel      TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
10qSubTel              TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
22qSubTel              TCTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
1pSubTel               TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCCGGTTTCCTACCATTATATTTT
5qSubTel               TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
17qSubTel              TCTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
4qSubTel               TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
19qSubTel              TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
21qSubTel              TCTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
1qSubTel               TCTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
2qSubTel               TTTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
19p'End'               TCTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
8p'End'                TCTTAAGCCAAAGGGACATATTTTGGGGTTGCATATTCTGGTTTCCTACCATTATATTTT
13qSubTel              TGTTAAGCCAAAGGGACATATTTTGGGGTTTCATATTCTGGTTTCCTACCATTATATTTT
20qSubTel              TCTTATGCCAAAGGGACATATTTTGGGG--------------------------------
12qSubTel              TCTTAAGCCAAAGGGACATATTTTGGGACAGCATATTCTGGTTACCTACCATTATATTTT
7qSubTel               TCTTAAGCCAAAGGGACATATTTTGGGACAGCATATTCTGGTTTCCTACCATTATATTTT
                       * ***:*********************.
```

```
6qPredictedSubTel       GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
10qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
22qSubTel               GGGGTGGCATAGTTTGGCCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
1pSubTel                GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
5qSubTel                GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
17qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
4qSubTel                GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
19qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
21qSubTel               GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
1qSubTel                GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCCACTGGCAATGAAAAGAGTTCTTGTT
2qSubTel                GGGGTGGCATAGTTTGGTCTTATACACTGTGTTCTACTGGCAATGAAAAGAGTTCTTGTT
19p'End'                GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACCGGCAATGAAAAGAGTTCTTGTT
8p'End'                 GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACCGGCAATGAAAAGAGTTCTTGTT
13qSubTel               GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACCGGCAATGAAAAGAGTTCTTGTT
20qSubTel               ----TGGCATATTCTGGTCTTATACACTGTCTTCCACCAGCAATGAAAAGAGTTATTGTC
12qSubTel               GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACAGGCAACGAAAAGAGTTCTTGTT
7qSubTel                GGGGTGGCATATTTTGGTCTTATACACTGTGTTCCACAGGCAACGAAAAGAGTTCTTGTT
                            ******* * *** ************ *** ** .**** **********.****

6qPredictedSubTel       TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
10qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
22qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
1pSubTel                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
5qSubTel                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
17qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
4qSubTel                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGGAGTTCTAAGAGATATAGACCAG
19qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
21qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
1qSubTel                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
2qSubTel                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGCTTAGCAGTTCTAAGAGATATAGACCAG
19p'End'                TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGTTTAGCAGTTCTAAGAGATATAGACCAG
8p'End'                 TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGTTTAGCAGTTCTAAGAGATATAGACCAG
13qSubTel               TTTCCTCCAGCAATTTGTCATTTGTTAAAGAGTTTAGCAGTTCTAAGAGATACAGACCAG
20qSubTel               TTTCTTGCAGCAATTTGTGATTT-TTTTAGAGTTTAGCAGTTCTAATAGATATAGACCAG
12qSubTel               TTTCCTGCAGCAATTTGTCATTTTTAAAAGAGCTTAGCAGTTCTAAGAGCTATAGAGTAG
7qSubTel                TTTCCTGCAGCAATTTGTCATTTTTTAAAGAGCTTAGCAGTTCTAAGAGCTATAGAGTAG
                        **** * *********** **** *:::**** **** ******** **.** ***  **

6qPredictedSubTel       CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
10qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTAAGCATCTTTTTGTAGGTGT
22qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
1pSubTel                CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
5qSubTel                CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
17qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
4qSubTel                CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
19qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
21qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
1qSubTel                CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
2qSubTel                CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAGGTGT
19p'End'                CTGTGCTATCTTTTTGTGGTTTTTAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
8p'End'                 CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
13qSubTel               CTGTGCTATCTTTTTGTGGTTTTCAGTTCTCTAGTATGTTGAGCATCTTTTTGTAAGTGT
20qSubTel               CTGTGCTATCTCC---TGGTTTTCAGTTCTGTAGTATGTTGAGCATCTTTTTGTATGTTT
12qSubTel               CTGTGCTATCTCATTGTGGTTTTCAATTCTCTAGTATGTTGAGCATCTTTTTGCACGTTT
7qSubTel                CTGTGCTATCTCATTGTGGTTTTCAATTCTCTAGTATGTTGAGCGTCTTTTTGCACGTTT
                        ***********     ******* *.**** *********.***.******** * ** *
```

```
6qPredictedSubTel      ACTTGCCATCTGTAGATCTTCTTTGATGAGGCGTCTGTTCAGATCTGTGTGCATTTTTAA
10qSubTel              ACTTGCCATCTGTAGATCTTCTTTGATGAGGCATCTGTTCAGATCTGTGTGCATTTTTAA
22qSubTel              ACTTGCCATCTGTAGATCTTCTTTGGTGAAGTGTCTGTTCAGATCTGTGTGCATTTTTAA
1pSubTel               ACTTGCCATCTGTAGATCTTCTTTGATGAGGCATCTGTTCAGATCTGTGTGCATTTTTAA
5qSubTel               ACTTGCCATCTGTAGATCTTCTTTGATGAGGCGTCTGTTCAGATCTGTGTGCATTTTTAA
17qSubTel              ACTTGCCATCTGTAGATCTTCTTTGGTGAAGTGTCTGTTCAGATCTGTGTGCATTTTTAA
4qSubTel               ACTTGCCATCTGTAGATCTTCTTTGATGAGGCGTCTGTTCAGATCTGTGTGCATTTTTAA
19qSubTel              ACTTGCCATCTGTAGATCTTCTTTGGTGAGGTGTCTGTTCAGATCTGTGTGCATTTTTAA
21qSubTel              ACTTGCCATCTGTAGATCTTCTTTGGTGAAGTGTCTGTTCAGATCTGTGTGCATTTTTAA
1qSubTel               ACTTGCCATCTGTAGATCTTCTTTGGTGAAGTGTCTGTTCAGATCTGTGTGCATTTTTAA
2qSubTel               ACTTGCCATCTGTAGATCTTCTTTGATGAGGCGTCTGTTCAGATCTGTGTGCATTTTTAA
19p'End'               ACTTGCCATCTGTAGATCTTCTTTGGTGAGGCGTCTGTTCAGATCTGTGTGCATTTTTAA
8p'End'                ACTTGCCATCTGTAGATCTTCTTTGGTGAGGCGTCTGTTCAGATCTGTGTGCATTTTTAA
13qSubTel              ACTTGCCATCTGTAGATCTTCTTTGGTGAGGTGTCTGTTTAGATCTGTGTGT--TTTTAA
20qSubTel              ACTTGCCATCTGTAGATCTTCTTTGGTGAGGTGTCTGTTCAGATCTGTGTGCATTTTTAA
12qSubTel              ACTTGCCATA----GATCTTCTCTGC----------------------------------
7qSubTel               ACTTGCCATA----GATCTTTTT-------------------------------------
                       *********.    ****** *

6qPredictedSubTel      TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
10qSubTel              TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
22qSubTel              TTGGGTTGTTTAACTTA---TTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
1pSubTel               TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
5qSubTel               TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
17qSubTel              TTGGGTTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
4qSubTel               TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
19qSubTel              TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
21qSubTel              TTGGGTTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
1qSubTel               TTGGGTTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
2qSubTel               TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATATATTTTGAATACAAA--
19p'End'               TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATGTATTTTGAATACAAA--
8p'End'                TTGGGCTGTTTAACTTATTGTTTAGTTTTAACAATTTTTTATGTATTTTGAATACAAA--
13qSubTel              TTGGGCTGTTTAACTTATTGTTTAGTTTTAACACTTTTTTATATATTTTGAATACAAA--
20qSubTel              TTGTGCTGTTTAACTTGT---TTAGTTTTAAGAATTTTTTATATATTTTGAATACAAATT
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------

6qPredictedSubTel      --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
10qSubTel              --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
22qSubTel              --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
1pSubTel               --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
5qSubTel               --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
17qSubTel              --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTGTTTTTGTTCT
4qSubTel               --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
19qSubTel              --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
21qSubTel              --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
1qSubTel               --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
2qSubTel               --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
19p'End'               --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTGGTTCT
8p'End'                --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTGGTTCT
13qSubTel              --TTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
20qSubTel              CTTTCTCAGATCTGTATTTTGCAAATATTTTCTTCAATATGTGGCTTGTCTTTTTGTTCT
12qSubTel              ------------------------------------------------------------
7qSubTel               ------------------------------------------------------------
```

```
6qPredictedSubTel       CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
10qSubTel               CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
22qSubTel               CTTGACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
1pSubTel                CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
5qSubTel                CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
17qSubTel               CTTGACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
4qSubTel                CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
19qSubTel               CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
21qSubTel               CTTGACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
1qSubTel                CTTGACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
2qSubTel                CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
19p'End'                CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
8p'End'                 CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
13qSubTel               CTTAACAAGGTCTCTTCCAGAGTATAAACTGTAAATATTAAGAAATCCACATTGTCATTT
20qSubTel               CTTAAGAAGGTCTCTTCCAGAGTATAAACTTTAAAGATTAAGAAATCCACATTGTCATTT
12qSubTel               ------------------------------------------------------------
7qSubTel                ------------------------------------------------------------

6qPredictedSubTel       CTTCTGTGTA------------------------
10qSubTel               CTTCTGTGTA------------------------
22qSubTel               CTTCTGTGTA------------------------
1pSubTel                CTTCTGTGTA------------------------
5qSubTel                CTTCTGTGTA------------------------
17qSubTel               CTTCTGTGTA------------------------
4qSubTel                CTTCTGTGTA------------------------
19qSubTel               CTTCTGTGTA------------------------
21qSubTel               CTTCTGTGTA------------------------
1qSubTel                CTTCTGTGTA------------------------
2qSubTel                CTTCTGTGTA------------------------
19p'End'                CTTCTGTGTA------------------------
8p'End'                 CTTCTGTGTATATCAACCTTCTGTGTCATTTGTT
13qSubTel               CTTCTGTGTATATC--------------------
20qSubTel               CTTCTGTGTA------------------------
12qSubTel               ----------------------------------
7qSubTel                ----------------------------------
```

References

1. Wang JCY, Dick JE. Cancer stem cells: lessons from leukemia. Trends in Cell Biology. 2005;15:494-501.
2. Gazit R, Weissman IL, Rossi DJ. Hematopoietic Stem Cells and the Aging Hematopoietic System. Seminars in Hematology. 2008;45:218 - 224.
3. Yang L, Bryder D, Adolfsson Jr, et al. Identification of Lin-Sca1+kit+CD34+Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. Blood. 2005;105:2717-2723.
4. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature. 2001;414:105-111.
5. Geiger H, de Haan G, Florian MC. The ageing haematopoietic stem cell compartment. Nat Rev Immunol. 2013;13:376-389.
6. Pang WW, Price EA, Sahoo D, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. Proc Natl Acad Sci U S A. 2011;108:20012-20017.
7. Ema H, Morita Y, Suda T. Heterogeneity and hierarchy of hematopoietic stem cells. Exp Hematol. 2013;2013:00863-00861.
8. Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell. 1997;91:661-672.
9. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature. 2000;404:193-197.
10. Cheshier SH, Morrison SJ, Liao X, Weissman IL. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. Proc Natl Acad Sci U S A. 1999;96:3120-3125.
11. Morrison SJ, Wandycz AM, Akashi K, Globerson A, Weissman IL. The aging of hematopoietic stem cells. Nat Med. 1996;2:1011-1016.
12. Henry CJ, Marusyk A, DeGregori J. Aging-associated changes in haematopoiesis and leukaemogenesis: What's the connection? Aging. 2011;3:643-656.
13. Ren R. Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. Nature Reviews Cancer. 2005;5:172-183.
14. Vickers M. Estimation of the number of mutations necessary to cause chronic myeloid leukaemia from epidemiological data. British Journal of Haematology. 1996;94:1-4.
15. Corey SJ, Minden MD, Barber DL, Kantarjian H, Wang JCY, Schimmer AD. Myelodysplastic syndromes: the complexity of stem-cell diseases. Nat Rev Cancer. 2007;7:118-129.
16. Pollyea DA, Kohrt HE, Medeiros BC. Acute myeloid leukaemia in the elderly: a review. British Journal of Haematology. 2011;152:524-542.
17. Rossi DJ, Bryder D, Weissman IL. Haematopoietic stem cell aging: Mechanism and consequence. Experimental Gerontology. 2007;42:385-390.
18. Greenberg PL, Young NS, Gattermann N. Myelodysplastic Syndromes. ASH Education Program Book. 2002;2002:136-161.
19. Nilsson L, Astrand-Grundtrom I, Arvidsson I, et al. Isolation and characterization of hematopoietic progenitor/stem cells in 5q-deleted myelodysplastic syndromes: evidence for involvement at the hematopoietic stem cell level. Blood. 2000;96:2012-2021.

20. Vardiman JW, Thiele Je, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114:937-951.
21. Garcia-Manero G. Myelodysplastic syndromes: 2012 update on diagnosis, risk-stratification, and management. American Journal of Hematology. 2012;87:692-701.
22. Shimazaki K, Ohshima K, Suzumiya J, Kawasaki C, Kikuchi M. Evaluation of apoptosis as a prognostic factor in myelodysplastic syndromes. British Journal of Haematology. 2000;110:584-590.
23. Tong W-G, Quintás-Cardama A, Kadia T, et al. Predicting survival of patients with hypocellular myelodysplastic syndrome. Cancer. 2012: 118: 4462-4470.
24. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the classification of the myelodysplastic syndromes. British Journal of Haematology. 1982;51:189-199.
25. Giannouli S, Kanellopoulou T, Voulgarelis M. Myelodysplasia and autoimmunity. Current Opinion in Rheumatology. 2012;24:97-102.
26. Dayyani F, Conley AP, Strom SS, et al. Cause of death in patients with lower-risk myelodysplastic syndrome. Cancer. 2010;116:2174-2179.
27. Kao JM, McMillan A, Greenberg PL. International MDS risk analysis workshop (IMRAW)/IPSS reanalyzed: Impact of cytopenias on clinical outcomes in myelodysplastic syndromes. American Journal of Hematology. 2008;83:765-770.
28. Huang TC, Ko BS, Tang JL, et al. Comparison of hypoplastic myelodysplastic syndrome (MDS) with normo-//hypercellular MDS by International Prognostic Scoring System, cytogenetic and genetic studies. Leukemia. 2007;22:544-550.
29. Bennett JM, Catovsky D, Daniel M-T, et al. Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. British Journal of Haematology. 1976;33:451-458.
30. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the recognition of minimally differentiated acute myeloid leukaemia (AML-M0). British Journal of Haematology. 1991;78.
31. Matsuo T, Bennett JM. Editorial: Acute leukemia of megakaryocyte lineage (M7). Cancer Genetics and Cytogenetics. 1988;34:1-3.
32. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114:937-951.
33. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood. 2002;100:2292-2302.
34. Liesveld JL, Jordan CT, Phillips GL. The Hematopoietic Stem Cell in Myelodysplasia. Stem Cells. 2004;22:590-599.
35. Wegrzyn J, Lam JC, Karsan A. Mouse models of myelodysplastic syndromes. Leukemia Research. 2011;35:853-862.
36. Bennett JM, Orazi A. Diagnostic criteria to distinguish hypocellular acute myeloid leukemia from hypocellular myelodysplastic syndromes and aplastic anemia: recommendations for a standardized approach. Haematologica. 2009;94:264-268.
37. Young NS. Acquired Aplastic Anemia. Annals of Internal Medicine. 2002;136:534-546.
38. Afable MG, Wlodarski M, Makishima H, et al. SNP array-based karyotyping: differences and similarities between aplastic anemia and hypocellular myelodysplastic syndromes. Blood. 2011;117:6876-6884.

39. Matsui WH, Brodsky RA, Smith BD, Borowitz MJ, Jones RJ. Quantitative analysis of bone marrow CD34 cells in aplastic anemia and hypoplastic myelodysplastic syndromes. Leukemia. 2006;20:458-462.
40. Schoch C, Kern W, Schnittger S, Hiddemann W, Haferlach T. Karyotype is an independent prognostic parameter in therapy-related acute myeloid leukemia (t-AML): an analysis of 93 patients with t-AML in comparison to 1091 patients with de novo AML. Leukemia. 2003;18:120-125.
41. Godley LA, Larson RA. Therapy-Related Myeloid Leukemia. Seminars in oncology. 2008;35:418-429.
42. Kayser S, Dohner K, Krauter J, et al. The impact of therapy-related acute myeloid leukemia (AML) on outcome in 2853 adult patients with newly diagnosed AML. Blood.2011;117:2137-2145.
43. Greenberg P, Cox C, LeBeau MM, et al. International Scoring System for Evaluating Prognosis in Myelodysplastic Syndromes. Blood. 1997;89:2079-2088.
44. Greenberg PL, Tuechler H, Schanz J, et al. Revised International Prognostic Scoring System for Myelodysplastic Syndromes. Blood. 2012;120:2454-2465.
45. Kantarjian H, Issa J-PJ, Rosenfeld CS, et al. Decitabine improves patient outcomes in myelodysplastic syndromes. Cancer. 2006;106:1794-1803.
46. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. The Lancet Oncology. 2009;10:223-232.
47. List A, Dewald G, Bennett J, et al. Lenalidomide in the Myelodysplastic Syndrome with Chromosome 5q Deletion. New England Journal of Medicine. 2006;355:1456-1465.
48. Lubbert M, Suciu S, Baila L, et al. Low-Dose Decitabine Versus Best Supportive Care in Elderly Patients With Intermediate- or High-Risk Myelodysplastic Syndrome (MDS) Ineligible for Intensive Chemotherapy: Final Results of the Randomized Phase III Study of the European Organisation for Research and Treatment of Cancer Leukemia Group and the German MDS Study Group. Journal of Clinical Oncology. 2011;29:1987-1996.
49. Gohring G, Giagounidis A, Busche G, et al. Patients with del(5q) MDS who fail to achieve sustained erythroid or cytogenetic remission after treatment with lenalidomide have an increased risk for clonal evolution and AML progression. Ann Hematol. 2009;89:365-374.
50. Dohner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood;2010;115:453-474.
51. Burnett AK, Milligan D, Prentice AG, et al. A comparison of low-dose cytarabine and hydroxyurea with or without all-trans retinoic acid for acute myeloid leukemia and high-risk myelodysplastic syndrome in patients not considered fit for intensive treatment. Cancer. 2007;109:1114-1124.
52. Hasserjian RP. Acute myeloid leukemia: advances in diagnosis and classification. International Journal of Laboratory Hematology;2013;35:358-366.
53. Wheatley K, Burnett AK, Goldstone AH, et al. A simple, robust, validated and highly predictive index for the determination of risk-directed therapy in acute myeloid leukaemia derived from the MRC AML 10 trial. United Kingdom Medical Research Council's Adult and Childhood Leukaemia Working Parties. Br J Haematol. 1999;107:69-79.

54. Wheatley K, Brookes CL, Howman AJ, et al. Prognostic factor analysis of the survival of elderly patients with AML in the MRC AML11 and LRF AML14 trials. British Journal of Haematology. 2009;145:598-605.
55. Dohner H, Gaidzik VI. Impact of Genetic Features on Treatment Decisions in AML. Hematology;2011:36-42.
56. Bacher U, Kern W, Schnittger S, Hiddemann W, Haferlach T, Schoch C. Population-based age-specific incidences of cytogenetic subgroups of acute myeloid leukemia. Haematologica. 2005;90:1502-1510.
57. Appelbaum FR, Gundacker H, Head DR, et al. Age and acute myeloid leukemia. Blood. 2006;107:3481-3485.
58. Grimwade D, Hills RK. Independent prognostic factors for AML outcome. Hematology. 2009;2009:385-395.
59. Kelly LM, Liu Q, Kutok JL, Williams IR, Boulton CL, Gilliland DG. FLT3 internal tandem duplication mutations associated with human acute myeloid leukemias induce myeloproliferative disease in a murine bone marrow transplant model. Blood. 2002;99:310-318.
60. Bacher U, Haferlach C, Kern W, Haferlach T, Schnittger S. Prognostic relevance of FLT3-TKD mutations in AML: the combination matters-- an analysis of 3082 patients. Blood. 2008;111:2527-2537.
61. Sallmyr A, Fan J, Datta K, et al. Internal tandem duplication of FLT3 (FLT3/ITD) induces increased ROS production, DNA damage, and misrepair: implications for poor prognosis in AML. Blood. 2008;111:3173-3182.
62. Markovic A, MacKenzie KL, Lock RB. FLT-3: a new focus in the understanding of acute leukemia. The International Journal of Biochemistry & Cell Biology. 2005;37:1168-1172.
63. Bagrintseva K, Geisenhof S, Kern R, et al. FLT3-ITD-TKD dual mutants associated with AML confer resistance to FLT3 PTK inhibitors and cytotoxic agents by overexpression of Bcl-x(L). Blood. 2005;105:3679-3685.
64. Meshinchi S, Woods WG, Stirewalt DL, et al. Prevalence and prognostic significance of Flt3 internal tandem duplication in pediatric acute myeloid leukemia. Blood. 2001;97:89-94.
65. Kiyoi H, Naoe T, Nakano Y, et al. Prognostic Implication of FLT3 and N-RAS Gene Mutations in Acute Myeloid Leukemia. Blood. 1999;93:3074-3080.
66. Whitman SP, Ruppert AS, Radmacher MD, et al. 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. 2008;111:1552-1559.
67. Moreno I, Martin G, Bolufer P, et al. Incidence and prognostic value of FLT3 internal tandem duplication and D835 mutations in acute myeloid leukemia. Haematologica. 2003;88:19-24.
68. Mead AJ, Linch DC, Hills RK, Wheatley K, Burnett AK, Gale RE. FLT3 tyrosine kinase domain mutations are biologically distinct from and have a significantly more favorable prognosis than FLT3 internal tandem duplications in patients with acute myeloid leukemia. Blood. 2007;110:1262-1270.
69. Borer RA, Lehner CF, Eppenberger HM, Nigg EA. Major nucleolar proteins shuttle between nucleus and cytoplasm. Cell. 1989;56:379-390.
70. Boissel N, Renneville A, Biggio V, et al. Prevalence, clinical profile, and prognosis of NPM mutations in AML with normal karyotype. Blood. 2005;106:3618-3620.

71. Falini B, Mecucci C, Tiacci E, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med. 2005;352:254-266.
72. Becker H, Marcucci G, Maharry K, et al. Favorable Prognostic Impact of NPM1 Mutations in Older Patients With Cytogenetically Normal De Novo Acute Myeloid Leukemia and Associated Gene- and MicroRNA-Expression Signatures: A Cancer and Leukemia Group B Study. Journal of Clinical Oncology. 2010;28:596-604.
73. Verhaak RGW, Goudswaard CS, van Putten W, et al. 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. 2005;106:3747-3754.
74. Dohner K, Schlenk RF, Habdank M, et al. Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. Blood. 2005;106:3740-3746.
75. Park SH, Chi H-S, Cho Y-U, Jang S, Park C-J. CEBPA single mutation can be a possible favorable prognostic indicator in NPM1 and FLT3-ITD wild-type acute myeloid leukemia patients with intermediate cytogenetic risk. Leukemia Research; 2013; 37:1488-1494.
76. Mueller BU, Pabst T. C/EBPA and the pathophysiology of acute myeloid leukemia. Current Opinion in Hematology. 2006;13:7-14.
77. Dufour A, Schneider F, Metzeler KH, et al. Acute Myeloid Leukemia With Biallelic CEBPA Gene Mutations and Normal Karyotype Represents a Distinct Genetic Entity Associated With a Favorable Clinical Outcome. Journal of Clinical Oncology;28:570-577.
78. Paschka P, Marcucci G, Ruppert AS, et al. Adverse Prognostic Significance of KIT Mutations in Adult Acute Myeloid Leukemia With inv(16) and t(8;21): A Cancer and Leukemia Group B Study. Journal of Clinical Oncology. 2006;24:3904-3911.
79. Holowiecki J, Grosicki S, Giebel S, et al. Cladribine, But Not Fludarabine, Added to Daunorubicin and Cytarabine During Induction Prolongs Survival of Patients With Acute Myeloid Leukemia: A Multicenter, Randomized Phase III Study. Journal of Clinical Oncology. 2012;30:2441-2448.
80. Burnett AK, Russell NH, Hunter AE, et al. Clofarabine doubles the response rate in older patients with acute myeloid leukemia but does not improve survival. Blood. 2013;122:1384-1394.
81. Burnett AK, Russell NH, Kell J, et al. European Development of Clofarabine as Treatment for Older Patients With Acute Myeloid Leukemia Considered Unsuitable for Intensive Chemotherapy. Journal of Clinical Oncology. 2010;28:2389-2395.
82. Scandura JM, Roboz GJ, Moh M, et al. Phase 1 study of epigenetic priming with decitabine prior to standard induction chemotherapy for patients with AML. Blood. 2011;118:1472-1480.
83. Laport GG, Sandmaier BM, Storer BE, et al. Reduced-intensity conditioning followed by allogeneic hematopoietic cell transplantation for adult patients with myelodysplastic syndrome and myeloproliferative disorders. Biol Blood Marrow Transplant. 2008;14:246-255.
84. Clift RA, Buckner CD, Appelbaum FR, et al. Allogeneic marrow transplantation in patients with acute myeloid leukemia in first remission: a randomized trial of two irradiation regimens. Blood. 1990;76:1867-1871.
85. Pagel JM, Gooley TA, Rajendran J, et al. Allogeneic hematopoietic cell transplantation after conditioning with ^{131}I-anti-CD45 antibody plus fludarabine and low-dose total body

irradiation for elderly patients with advanced acute myeloid leukemia or high-risk myelodysplastic syndrome. Blood. 2009;114:5444-5453.
86. Serefhanoglu S, Goker H, Buyukasik Y, Sayinalp N, Ozcebe OI. Transformation of adult myelodysplastic syndrome-refractory anemia to acute T-cell lymphoblastic leukemia. J Natl Med Assoc. 2009;101:370-372.
87. Nimer SD. MDS: A stem cell disorder--but what exactly is wrong with the primitive hematopoietic cells in this disease? ASH Education Program Book. 2008;2008:43-51.
88. Buckton KE, Brown WMC, Smith PG. Lymphocyte Survival in Men treated with X-rays for Ankylosing Spondylitis. Nature. 1967;214:470-473.
89. Vogelstein B, Fearon ER, Hamilton SR. Clonal analysis using recombinant DNA probes from the X-chromosome Cancer Research. 1987;47:4806-4813.
90. Tefferi A, Thibodeau S, Solberg LJ. Clonal studies in the myelodysplastic syndrome using X-linked restriction fragment length polymorphisms. Blood. 1990;75:1770-1773.
91. Vogelstein B, Fearon ER, Hamilton SR, Feinberg AP. Use of restriction fragment length polymorphisms to determine the clonal origin of human tumours Science. 1985;227:642-645.
92. Delforge M, Demuynck H, Verhoef G, et al. Patients with high-risk myelodysplastic syndrome can have polyclonal or clonal haemopoiesis in complete haematological remission. British Journal of Haematology. 1998;102:486-494.
93. Cumano A, Paige CJ, Iscover NN, Brady G. Bipotential precursors of B cells and macrophages in murine fetal liver. Nature. 1992;356:612-615.
94. Hirayama F, Shih J-P, Awgulewitsch A, Warr GW, Clark SC, Ogawa M. Clonal proliferation of murine lymphohemopoietic progenitors in culture. Proceedings of the National Academy of Sciences. 1992;89:5907-5911.
95. Miura I, Kobayashi Y, Takahashi N, Saitoh K, Miura AB. Involvement of natural killer cells in patients with myelodysplastic syndrome carrying monosomy 7 revealed by the application of fluorescence in situ hybridization to cells collected by means of fluorescence-activated cell sorting. British Journal of Haematology. 2000;110:876-879.
96. Raskind W, Tirumali N, Jacobson R, Singer J, Fialkow P. Evidence for a multistep pathogenesis of a myelodysplastic syndrome. Blood. 1984;63:1318-1323.
97. White N, Nacheva E, Asimakopoulos F, Bloxham D, Paul B, Green A. Deletion of chromosome 20q in myelodysplasia can occur in a multipotent precursor of both myeloid cells and B cells. Blood. 1994;83:2809-2816.
98. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med. 1997;3:730-737.
99. Terpstra W, Prins A, Ploemacher RE, et al. Long-term leukemia-initiating capacity of a CD34-subpopulation of acute myeloid leukemia. Blood. 1996;87:2187-2194.
100. Turhan AG, Lemoine FM, Debert C, et al. Highly purified primitive hematopoietic stem cells are PML-RARA negative and generate nonclonal progenitors in acute promyelocytic leukemia. Blood. 1995;85:2154-2161.
101. Miyamoto T, Weissman IL, Akashi K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. Proceedings of the National Academy of Sciences. 2000;97:7521-7526.
102. Gondek LP, Haddad AS, O'Keefe CL, et al. Detection of cryptic chromosomal lesions including acquired segmental uniparental disomy in advanced and low-risk myelodysplastic syndromes. Experimental Hematology. 2007;35:1728-1738.

103. Breccia M, Mancini M, Nanni M, et al. Clinical features of prognostic significance in myelodysplastic patients with normal karyotype at high risk of transformation. Leukemia Research. 2005;29:33-39.
104. Iwase O, Iwama H, Okabe S, et al. Refractory anemia with ringed sideroblasts with a low IPSS score progressed rapidly with de novo appearance of multiple karyotypic abnormalities and into acute erythroleukemia (AML-M6A). Leukemia Research. 2000;24:597-600.
105. Martínez-Ramírez A, Urioste M, Melchor L, et al. Analysis of myelodysplastic syndromes with complex karyotypes by high-resolution comparative genomic hybridization and subtelomeric CGH array. Genes, Chromosomes and Cancer. 2005;42:287-298.
106. Nevill TJ, Fung HC, Shepherd JD, et al. Cytogenetic Abnormalities in Primary Myelodysplastic Syndrome Are Highly Predictive of Outcome After Allogeneic Bone Marrow Transplantation. Blood. 1998;92:1910-1917.
107. Giagounidis AAN, Germing U, Haase S, et al. Clinical, morphological, cytogenetic, and prognostic features of patients with myelodysplastic syndromes and del(5q) including band q31. Leukemia. 2003;18:113-119.
108. Dewald G, Davis M, Pierre R, O'Fallon J, Hoagland H. Clinical characteristics and prognosis of 50 patients with a myeloproliferative syndrome and deletion of part of the long arm of chromosome 5. Blood. 1985;66:189-197.
109. Boultwood J, Fidler C, Strickson AJ, et al. Narrowing and genomic annotation of the commonly deleted region of the 5q- syndrome. Blood. 2002;99:4638-4641.
110. Starczynowski DT, Kuchenbauer F, Argiropoulos B, et al. Identification of miR-145 and miR-146a as mediators of the 5q- syndrome phenotype. Nat Med. 2010;16:49-58.
111. Kumar MS, Narla A, Nonami A, et al. Coordinate loss of a microRNA and protein-coding gene cooperate in the pathogenesis of 5q- syndrome. Blood. 2011;118:4666-4673.
112. Dutt S, Narla A, Lin K, et al. Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. Blood. 2011;117:2567-2576.
113. Hasserjian RP, Zuo Z, Garcia C, et al. Acute erythroid leukemia: a reassessment using criteria refined in the 2008 WHO classification. Blood. 2010;115:1985-1992.
114. Lai F, Godley LA, Joslin J, et al. Transcript Map and Comparative Analysis of the 1.5-Mb Commonly Deleted Segment of Human 5q31 in Malignant Myeloid Diseases with a del(5q). Genomics. 2001;71:235-245.
115. Giagounidis AAN, Germing U, Aul C. Biological and Prognostic Significance of Chromosome 5q Deletions in Myeloid Malignancies. Clinical Cancer Research. 2006;12:5-10.
116. Boultwood J, Pellagatti A, Cattan H, et al. Gene expression profiling of CD34+ cells in patients with the 5q− syndrome. British Journal of Haematology. 2007;139:578-589.
117. Baron V, Adamson ED, Calogero A, Ragona G, Mercola D. The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGFbeta1, PTEN, p53, and fibronectin. Cancer Gene Ther. 2006;13:115-124.
118. Joslin J, Fernald AA, Tennant TR, et al. Haploinsufficiency of *EGR1*, a candidate gene in the del(5q), leads to the development of myeloid disorders. Blood. 2007;110:719-726.
119. Pedersen B, Ellegaard J. A factor encoded by 7q31 suppresses expansion of the 7q- clone and delays cytogenetic progression. Cancer Genetics and Cytogenetics. 1994;78:181-188.
120. Zhou L, Opalinska J, Sohal D, et al. Aberrant Epigenetic and Genetic Marks Are Seen in Myelodysplastic Leukocytes and Reveal Dock4 as a Candidate Pathogenic Gene on Chromosome 7q. Journal of Biological Chemistry. 2011;286:25211-25223.

121. Yajnik V, Paulding C, Sordella R, et al. DOCK4, a GTPase Activator, Is Disrupted during Tumorigenesis. Cell. 2003;112:673-684.
122. Pellagatti A, Cazzola M, Giagounidis A, et al. Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells. Leukemia. 2010;24:756-764.
123. Jiang Y, Dunbar A, Gondek LP, et al. Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. Blood. 2009;113:1315-1325.
124. Hofmann WK, Takeuchi S, Takeuchi N, Thiel E, Hoelzer D, Koeffler HP. Comparative analysis of hypermethylation of cell cycle control and DNA-mismatch repair genes in low-density and CD34+ bone marrow cells from patients with myelodysplastic syndrome. Leuk Res. 2006;30:1347-1353.
125. Uchida T, Kinoshita T, Nagai H, et al. Hypermethylation of the p15INK4B Gene in Myelodysplastic Syndromes. Blood. 1997;90:1403-1409.
126. Langemeijer SMC, Kuiper RP, Berends M, et al. Acquired mutations in *TET2* are common in myelodysplastic syndromes Nature Genetics. 2009;41:838-842.
127. Li Z, Cai X, Cai C-L, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood. 2011;118:4509-4518.
128. Thiel A, Beier M, Ingenhag D, et al. Comprehensive array CGH of normal karyotype myelodysplastic syndromes reveals hidden recurrent and individual genomic copy number alterations with prognostic relevance. Leukemia. 2011;25:387-399.
129. Mori N, Morosetti R, Hoflehner E, Lubbert M, Mizoguchi H, Koeffler HP. Allelic Loss in the Progression of Myelodysplastic Syndrome. Cancer Research. 2000;60:3039-3042.
130. Heinrichs S, Kulkarni RV, Bueso-Ramos CE, et al. Accurate detection of uniparental disomy and microdeletions by SNP array analysis in myelodysplastic syndromes with normal cytogenetics. Leukemia. 2009;23:1605-1613.
131. Gorletta TA, Gasparini P, D'Elios MM, Trubia M, Pelicci PG, Di Fiore PP. Frequent loss of heterozygosity without loss of genetic material in acute myeloid leukemia with a normal karyotype. Genes, Chromosomes and Cancer. 2005;44:334-337.
132. Raghavan M, Smith L-L, Lillington DM, et al. Segmental uniparental disomy is a commonly acquired genetic event in relapsed acute myeloid leukemia. Blood. 2008;112:814-821.
133. Sugimoto K, Hirano N, Toyoshima H, et al. Mutations of the p53 gene in myelodysplastic syndrome (MDS) and MDS- derived leukemia. Blood. 1993;81:3022-3026.
134. Fenaux P, Jonveaux P, Quiquandon I, et al. P53 gene mutations in acute myeloid leukemia with 17p monosomy. Blood. 1991;78:1652-1657.
135. Kikukawa M, Aoki N, Sakamoto Y, Mori M. Study of p53 in Elderly Patients with Myelodysplastic Syndromes by Immunohistochemistry and DNA Analysis. The American Journal of Pathology. 1999;155:717-721.
136. Stirewalt DL, Kopecky KJ, Meshinchi S, et al. FLT3, RAS, and TP53 mutations in elderly patients with acute myeloid leukemia. Blood. 2001;97:3589-3595.
137. Raza A, Gezer S, Mundle S, et al. Apoptosis in bone marrow biopsy samples involving stromal and hematopoietic cells in 50 patients with myelodysplastic syndromes [see comments]. Blood. 1995;86:268-276.
138. Peter L G. Apoptosis and its role in the myelodysplastic syndromes: implications for disease natural history and treatment. Leukemia Research. 1998;22:1123-1136.
139. Greenberg PL. Apoptosis and its role in the myelodysplastic syndromes: implications for disease natural history and treatment. Leukemia Research. 1998;22:1123-1136.

140. Slape CI, Saw J, Jowett JBM, et al. Inhibition of apoptosis by BCL2 prevents leukemic transformation of a murine myelodysplastic syndrome. Blood. 2012;120:2475-2483.
141. Parker JE, Mufti GJ, Rasool F, Mijovic A, Devereux S, Pagliuca A. The role of apoptosis, proliferation, and the Bcl-2-related proteins in the myelodysplastic syndromes and acute myeloid leukemia secondary to MDS. Blood. 2000;96:3932-3938.
142. Evan GI, Wyllie AH, Gilbert CS, et al. Induction of apoptosis in fibroblasts by c-myc protein. Cell. 1992;69:119-128.
143. Rajapaksa R, Ginzton N, Rott L, Greenberg P. Altered oncoprotein expression and apoptosis in myelodysplastic syndrome marrow cells. Blood. 1996;88:4275-4287.
144. Klefstrom J, Vastrik I, Saksela E, Valle J, Eilers M, Alitalo K. c-Myc induces cellular susceptibility to the cytotoxic action of TNF-alpha. EMBO J. 1994;13:5442-5450.
145. Gersuk G, Lee J, Beckham C, Anderson J, Deeg H. Fas (CD95) receptor and Fas-ligand expression in bone marrow cells from patients with myelodysplastic syndrome [letter]. Blood. 1996;88:1122-1123.
146. Campos L, Rouault JP, Sabido O, et al. High expression of bcl-2 protein in acute myeloid leukemia cells is associated with poor response to chemotherapy. Blood. 1993;81:3091-3096.
147. Boudard D, Vasselon C, Bertheas MF, et al. Expression and prognostic significance of Bcl-2 family proteins in Myelodysplastic Syndromes. American Journal of Hematology. 2002;70:115-125.
148. Sawanobori M, Yamaguchi S, Hasegawa M, et al. Expression of TNF receptors and related signaling molecules in the bone marrow from patients with myelodysplastic syndromes. Leukemia Research. 2003;27:583-591.
149. Sloand EM, Kim S, Fuhrer M, et al. Fas-mediated apoptosis is important in regulating cell replication and death in trisomy 8 haematopoietic cells but not in cells with other cytogenetic abnormalites Blood. 2002;100:4427-4432.
150. Kondo A, Yamashita T, Tamura H, et al. Interferon-γ and tumor necrosis factor-α induce an immunoinhibitory molecule, B7-H1, via nuclear factor- κB activation in blasts in myelodysplastic syndromes. Blood. 2010;116:1124-1131.
151. Maciejewski J, Selleri C, Anderson S, Young N. Fas antigen expression on CD34+ human marrow cells is induced by interferon gamma and tumor necrosis factor alpha and potentiates cytokine-mediated hematopoietic suppression in vitro. Blood. 1995;85:3183-3190.
152. Bouscary D, De Vos J, Guesnu M, et al. Fas/Apo-1 (CD95) expression and apoptosis in patients with myelodysplastic syndromes. Leukemia. 1997;11:839-845.
153. Min YH, Lee S, Lee JW, Chong SY, Hahn JS, Ko YW. Expression of Fas antigen in acute myeloid leukaemia is associated with therapeutic response to chemotherapy. Br J Haematol. 1996;93:928-930.
154. Gupta P, Niehans GA, LeRoy SC, et al. Fas ligand expression in the bone marrow in myelodysplastic syndromes correlates with FAB subtype and anemia, and predicts survival. Leukemia. 1999;13:44-53.
155. Strand S, Hofmann WJ, Hug H, et al. Lymphocyte apoptosis induced by CD95 (APO - 1/ Fas) ligand- expressing tumor cells - A mechanism of immune invasion? Nature Medicine. 1996;2:1361 - 1366.
156. Kerbauy DMB, Lesnikov V, Abbasi N, Seal S, Scott B, Deeg HJ. NF-κB and FLIP in arsenic trioxide (ATO)-induced apoptosis in myelodysplastic syndromes (MDSs). Blood. 2005;106:3917-3925.

157. Guzman ML, Neering SJ, Upchurch D, et al. Nuclear factor-kB is constitutively activated in primitive human acute myelogenous leukaemia cells Blood. 2001;98:2301-2307.
158. Karin M. Nuclear factor-kB in cancer development and progression. Nature. 2006;441:431-436.
159. Cilloni D, Messa F, Rosso V, et al. Increase sensitivity to chemotherapeutical agents and cytoplasmatic interaction between NPM leukemic mutant and NF-kappaB in AML carrying NPM1 mutations. Leukemia. 2008;22:1234-1240.
160. Muller HJ. An Analysis of the Process of Structural Change in Chromosomes of *Drosophila.* Journal of Genetics. 1940;40:1-66.
161. McClintock B. The stability of broken ends of chromosomes in *Zea Mays*. Genetics. 1941;26:234-282.
162. Watson JD. Origin of concatemeric T7 DNA. Nature: New Biology 1972;239:197-201.
163. Olovnikov AM. A theory of marginotomy. Journal of Theoretical Biology. 1973;41:181-190.
164. Allsopp RC, Vaziri H, Patterson C, et al. Telomere length predicts replicative capacity of human fibroblasts. Proc Natl Acad Sci U S A. 1992;89:10114-10118.
165. Blackburn EH, Gall JG. A tandemly repeated sequence at the termini of the extachromosomal ribosomal RNA genes in Tetrahymena. Journal of Molecular Biology. 1978;120:33-53.
166. Moyzis RK, Buckingham JM, Cram LS, et al. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proceedings of the National Academy of Sciences. 1988;85:6622-6626.
167. Bullock PA, Tevosian S, Jones C, Denis D. Mapping initiation sites for simian virus 40 DNA synthesis events in vitro. Molecular and Cellular Biology. 1994;14:5043-5055.
168. Ohki R, Tsurimoto T, Ishikawa F. In vitro reconstitution of the end replication problem. Molecular and Cellular Biology. 2001;21:5753-5766.
169. Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB. Telomere end-replication problem and cell aging. Journal of Molecular Biology. 1992;225:951-960.
170. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. Nature. 1990;345:458-460.
171. Baird DM, Rowson J, Wynford-Thomas D, Kipling D. Extensive allelic variation and ultrashort telomeres in senescent human cells. Nat Genet. 2003;33:203-207.
172. Blackburn EH. Telomere states and cell fates. Nature. 2000;408:53-56.
173. Rubelj I, Vondracek Z. Stochastic Mechanism of Cellular Aging-Abrupt Telomere Shortening as a Model for Stochastic Nature of Cellular Aging. Journal of Theoretical Biology. 1999;197:425-438.
174. Smith, Whitney R. Intraclonal variation in proliferative potential of human diploid fibroblasts: stochastic mechanism for cellular aging. Science. 1980;207:82-84.
175. Smith JR, Hayflick L. Variation in the life-span of clones derived from human diploid cell strains. The Journal of Cell Biology. 1974;62:48-53.
176. Greider CW, Blackburn EH. Identification of a specific telomere terminal transferase activity in tetrahymena extracts. Cell. 1985;43:405-413.
177. O'Reilly M, Teichmann SA, Rhodes D. Telomerases. Current Opinion in Structural Biology. 1999;9:56-65.
178. Theimer CA, Feigon J. Structure and function of telomerase RNA. Current Opinion in Structural Biology. 2006;16:307-318.

179. Shay JW, Gazdar AF. Telomerase in the early detection of cancer. J Clin Pathol. 1997;50:106-109.
180. Kim NW, Piatyszek MA, Prowse KR, et al. Specific association of human telomerase activity with immortal cells and cancer. Science. 1994;266:2011-2015.
181. Avilion AA, Piatyszek MA, Gupta J, Shay JW, Bacchetti S, Greider CW. Human Telomerase RNA and Telomerase Activity in Immortal Cell Lines and Tumor Tissues. Cancer Research. 1996;56:645-650.
182. Kilian A, Bowtell DDL, Abud HE, et al. Isolation of a Candidate Human Telomerase Catalytic Subunit Gene, Which Reveals Complex Splicing Patterns in Different Cell Types. Human Molecular Genetics. 1997;6:2011-2019.
183. Meyerson M, Counter CM, Eaton EN, et al. hEST2, the Putative Human Telomerase Catalytic Subunit Gene, Is Up-Regulated in Tumor Cells and during Immortalization. Cell. 1997;90:785-795.
184. Nakayama J-i, Tahara H, Tahara E, et al. Telomerase activation by hTRT in human normal fibroblasts and hepatocellular carcinomas. Nat Genet. 1998;18:65-68.
185. Bodnar AG, Ouellette M, Frolkis M, et al. Extension of life-span by introduction of telomerase into normal human cells. Science. 1998;279:349-352.
186. Counter CM, Meyerson M, Eaton EN, et al. Telomerase activity is restored in human cells by ectopic expression of hTERT (hEST2), the catalytic subunit of telomerase. Oncogene. 1998;16:1217-1222.
187. Kolquist KA, Ellisen LW, Counter CM, et al. Expression of TERT in early premalignant lesions and a subset of cells in normal tissues. Nat Genet. 1998;19:182-186.
188. Shibuya K, Fujisawa T, Hoshino H, et al. Increased telomerase activity and elevated hTERT mRNA expression during multistage carcinogenesis of squamous cell carcinoma of the lung. Cancer. 2001;92:849-855.
189. Blackburn EH. Telomerases. Annual Review of Biochemistry. 1992;61:113-129.
190. Griffith JD, Comeau L, Rosenfield S, et al. Mammalian Telomeres End in a Large Duplex Loop. Cell. 1999;97:503-514.
191. De Boeck G, Forsyth RG, Praet M, Hogendoorn PCW. Telomere-associated proteins: cross-talk between telomere maintenance and telomere-lengthening mechanisms. The Journal of Pathology. 2009;217:327-344.
192. Palm W, de Lange T. How Shelterin Protects Mammalian Telomeres. Annual Review of Genetics. 2008;42:301-334.
193. Churikov D, Wei C, Price CM. Vertebrate POT1 Restricts G-Overhang Length and Prevents Activation of a Telomeric DNA Damage Checkpoint but Is Dispensable for Overhang Protection. Molecular and Cellular Biology. 2006;26:6971-6982.
194. Barrientos KS, Kendellen MF, Freibaum BD, Armbruster BN, Etheridge KT, Counter CM. Distinct Functions of POT1 at Telomeres. Molecular and Cellular Biology. 2008;28:5251-5264.
195. Kim S-h, Davalos AR, Heo S-J, et al. Telomere dysfunction and cell survival: roles for distinct TIN2-containing complexes. The Journal of Cell Biology. 2008;181:447-460.
196. Palm W, Hockemeyer D, Kibe T, de Lange T. Functional Dissection of Human and Mouse POT1 Proteins. Molecular and Cellular Biology. 2009;29:471-482.
197. Griffith J, Bianchi A, de Lange T. TRF1 promotes parallel pairing of telomeric tracts in vitro. Journal of Molecular Biology. 1998;278:79-88.
198. Smogorzewska A, van Steensel B, Bianchi A, et al. Control of Human Telomere Length by TRF1 and TRF2. Molecular and Cellular Biology. 2000;20:1659-1668.

199. van Steensel B, Smogorzewska A, de Lange T. TRF2 Protects Human Telomeres from End-to-End Fusions. Cell. 1998;92:401-413.
200. Wu L, Multani AS, He H, et al. Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. Cell. 2006;126:49-62.
201. Hockemeyer D, Daniels JP, Takai H, de Lange T. Recent expansion of the telomeric complex in rodents: Two distinct POT1 proteins protect mouse telomeres. Cell. 2006;126:63-77.
202. Celli GB, de Lange T. DNA processing is not required for ATM - mediated damage response after TRF2 deletion. Nature Cell Biology. 2005;7:712- 718.
203. Sfeir A, Kabir S, van Overbeek M, Celli GB, de Lange T. Loss of Rap1 Induces Telomere Recombination in the Absence of NHEJ or a DNA Damage Signal. Science. 2010;327:1657-1661.
204. Kelleher C, Kurth I, Lingner J. Human Protection of Telomeres 1 (POT1) is a Negative Regulator of Telomerase Activity In Vitro. Molecular and Cellular Biology. 2005;25:808-818.
205. Loayza D, de Lange T. POT1 as a terminal transducer of TRF1 telomere length control. Nature. 2003;423:1013-1018.
206. Seimiya H, Muramatsu Y, Ohishi T, Tsuruo T. Tankyrase 1 as a target for telomere-directed molecular cancer therapeutics. Cancer Cell. 2005;7:25-37.
207. van Steensel B, de Lange T. Control of telomere length by the human telomeric protein TRF1. Nature. 1997;385:740-743.
208. Teixeira MT, Arneric M, Sperisen P, Lingner J. Telomere length homeostasis is achieved via a switch between telomerase- extendible and -nonextendible states. Cell. 2004;117:323-335.
209. Hemann MT, Strong MA, Hao L-Y, Greider CW. The Shortest Telomere, Not Average Telomere Length, Is Critical for Cell Viability and Chromosome Stability. Cell. 2001;107:67-77.
210. Britt-Compton B, Capper R, Rowson J, Baird DM. Short telomeres are preferentially elongated by telomerase in human cells. FEBS Letters. 2009;583:3076-3080.
211. Cristofari G, Lingner J. Telomere length homeostasis requires that telomerase levels are limiting. EMBO J. 2006;25:565-574.
212. Liu Y, Kha H, Ungrin M, Robinson MO, Harrington L. Preferential maintenance of critically short telomeres in mammalian cells heterozygous for mTert. Proc Natl Acad Sci U S A. 2002;99:3597-3602.
213. Hug N, Lingner J. Telomere length homeostasis. Chromosoma. 2006;115:413-425.
214. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. Genes Dev. 2004;18:1618-1629.
215. Lipps HJ, Rhodes D. G-quadruplex structures: in vivo evidence and function. Trends Cell Biol. 2009;19:414-422.
216. Sen D, Gilbert W. A sodium-potassium switch in the formation of four-stranded G4-DNA. Nature. 1990;344:410-414.
217. Paeschke K, Simonsson T, Postberg J, Rhodes D, Lipps HJ. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. Nat Struct Mol Biol. 2005;12:847-854.
218. Xin H, Liu D, Wan M, et al. TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. Nature. 2007;445:559-562.

219. Paeschke K, Juranek S, Simonsson T, Hempel A, Rhodes D, Lipps HJ. Telomerase recruitment by the telomere end binding protein-[beta] facilitates G-quadruplex DNA unfolding in ciliates. Nat Struct Mol Biol. 2008;15:598-604.
220. Hwang H, Buncher N, Opresko PL, Myong S. POT1-TPP1 Regulates Telomeric Overhang Structural Dynamics. Structure. 2012;20:1-9.
221. Sfeir A, Kosiyatrakul ST, Hockemeyer D, et al. Mammalian Telomeres Resemble Fragile Sites and Require TRF1 for Efficient Replication. Cell. 2009;138:90-103.
222. Harley CB. Telomerase and cancer therapeutics. Nat Rev Cancer. 2008;8:167-179.
223. Su Z, Dannull J, Yang BK, et al. Telomerase mRNA-Transfected Dendritic Cells Stimulate Antigen-Specific CD8+ and CD4+ T Cell Responses in Patients with Metastatic Prostate Cancer. The Journal of Immunology. 2005;174:3798-3807.
224. Buseman CM, Wright WE, Shay JW. Is telomerase a viable target in cancer? Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2012;730:90-97.
225. Brassat U, Balabanov S, Bali D, et al. Functional p53 is required for effective execution of telomerase inhibition in BCR-ABL-positive CML cells. Experimental Hematology. 2011;39:66-76.e62.
226. Hochreiter AE, Xiao H, Goldblatt EM, et al. Telomerase Template Antagonist GRN163L Disrupts Telomere Maintenance, Tumor Growth, and Metastasis of Breast Cancer. Clinical Cancer Research. 2006;12:3184-3192.
227. Dikmen ZG, Wright WE, Shay JW, Gryaznov SM. Telomerase targeted oligonucleotide thio-phosphoramidates in T24-luc bladder cancer cells. J Cell Biochem. 2008;104:444-452.
228. Dikmen ZG, Gellert GC, Jackson S, et al. In vivo Inhibition of Lung Cancer by GRN163L: A Novel Human Telomerase Inhibitor. Cancer Research. 2005;65:7866-7873.
229. Wang ES, Wu K, Chin AC, et al. Telomerase inhibition with an oligonucleotide telomerase template antagonist: in vitro and in vivo studies in multiple myeloma and lymphoma. Blood. 2004;103:258-266.
230. Ruden M, Puri N. Novel anticancer therapeutics targeting telomerase. Cancer Treatment Reviews. 2012.
231. Pascolo E, Wenz C, Lingner J, et al. Mechanism of human telomerase inhibition by BIBR1532, a synthetic, non-nucleosidic drug candidate. J Biol Chem. 2002;277:15566-15572.
232. Damm K, Hemmann U, Garin-Chesa P, et al. A highly selective telomerase inhibitor limiting human cancer cell proliferation. EMBO J. 2001;20:6958-6968.
233. Ward RJ, Autexier C. Pharmacological Telomerase Inhibition Can Sensitize Drug-Resistant and Drug-Sensitive Cells to Chemotherapeutic Treatment. Molecular Pharmacology. 2005;68:779-786.
234. Salvati E, Leonetti C, Rizzo A, et al. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. The Journal of Clinical Investigation. 2007;117:3236-3247.
235. Zijlmans JMJM, Martens UM, Poon SSS, et al. Telomeres in the mouse have large inter-chromosomal variations in the number of T2AG3 repeats. Proceedings of the National Academy of Sciences. 1997;94:7423-7428.
236. Kipling D, Cooke HJ. Hypervariable ultra-long telomeres in mice. Nature. 1990;347:400-402.
237. Starling JA, Maule J, Hastie ND, Allshire RC. Extensive telomere repeat arrays in mouse are hypervariable. Nucleic Acids Research. 1990;18:6881-6888.

238. Njajou OT, Cawthon RM, Damcott CM, et al. Telomere length is paternally inherited and is associated with parental lifespan. Proceedings of the National Academy of Sciences. 2007;104:12135-12139.
239. Andrew T, Aviv A, Falchi M, et al. Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. Am J Hum Genet. 2006;78:480-486.
240. Slagboom PE, Droog S, Boomsma DI. Genetic determination of telomere size in humans: a twin study of three age groups. The American Journal of Human Genetics. 1994;55:876- 882.
241. Linardopoulou EV, Williams EM, Fan Y, Friedman C, Young JM, Trask BJ. Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. Nature. 2005;437:94-100.
242. Baird DM, Coleman J, Rosser ZH, Royle NJ. High Levels of Sequence Polymorphism and Linkage Disequilibrium at the Telomere of 12q: Implications for Telomere Biology and Human Evolution. The American Journal of Human Genetics. 2000;66:235-250.
243. Allshire RC, Dempster M, Hastie ND. Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucleic Acids Research. 1989;17:4611-4627.
244. Baird DM, Jeffreys AJ, Royle NJ. Mechanisms underlying telomere repeat turnover, revealed by hypervariable variant repeat distribution patterns in the human Xp/Yp telomere. Embo J. 1995;14:5433-5443.
245. Mefford HC, Linardopoulou E, Coil D, van den Engh G, Trask BJ. Comparative sequencing of a multicopy subtelomeric region containing olfactory receptor genes reveals multiple interactions between non-homologous chromosomes. Human Molecular Genetics. 2001;10:2363-2372.
246. Wilkie AOM, Higgs DR, Rack KA, et al. Stable length polymorphism of up to 260 kb at the tip of the short arm of human chromosome 16. Cell. 1991;64:595-606.
247. Riethman H, Ambrosini A, Castaneda C, et al. Mapping and Initial Analysis of Human Subtelomeric Sequence Assemblies. Genome Research. 2004;14:18-28.
248. Aubert G, Hills M, Lansdorp PM. Telomere length measurement - caveats and a critical assessment of the available technologies and tools. Mutation Research. 2012;730:59-67.
249. Baird DM. New developments in telomere length analysis. Experimental Gerontology. 2005;40:363-368.
250. Lansdorp PM, Verwoerd NP, Van de Rijke FM, et al. Heterogeneity in telomere length of human chromosomes. Human Molecular Genetics. 1996;5:685-691.
251. Baerlocher GM, Vulto I, de Jong G, Lansdorp PM. Flow cytometry and FISH to measure the average length of telomeres (flow FISH). Nature Protocols. 2006;1:2365-2376.
252. Cawthon R. Telomere measurement by quantitative PCR. Nucleic Acids Research. 2002;30:e47.
253. Lin TT, Letsolo BT, Jones RE, et al. Telomere dysfunction and fusion during the progression of chronic lymphocytic leukemia: evidence for a telomere crisis. Blood. 2010;116:1899- 1907.
254. Letsolo BT, Rowson J, Baird DM. Fusion of short telomeres in human cells is characterized by extensive deletion and microhomology, and can result in complex rearrangements. Nucleic Acids Research;38:1841-1852.

255. Capper R, Britt-Compton B, Tankimanova M, et al. The nature of telomere fusion and a definition of the critical telomere length in human cells. Genes & Development. 2007;21:2495-2508.
256. Li B, de Lange T. Rap1 Affects the Length and Heterogeneity of Human Telomeres. Molecular Biology of the Cell. 2003;14:5060-5068.
257. Shampay J, Blackburn EH. Generation of telomere-length heterogeneity in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences. 1988;85:534-538.
258. Stansel RM, de Lange T, Griffith JD. T-loop assembly in vitro involves binding of TRF2 near the 3[prime] telomeric overhang. EMBO J. 2001;20:5532-5540.
259. McEachern MJ, Blackburn EH. Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in *Kluyveromyces lactis* lacking telomerase. Genes & Development. 1996;10:1822-1834.
260. Basile G, Aker M, Mortimer RK. Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. Molecular and Cellular Biology. 1992;12:3235-3246.
261. Siede W, Friedl AA, Dianova I, Eckardt-Schupp F, Friedberg EC. The *Saccharomyces cerevisiae* Ku Antoantigen Homologue Affects Radiosensitivity Only in the Absence of Homologous Recombination. Genetics. 1996;142:91- 102.
262. Malone RE, Esposito RE. The RAD52 gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast. Proceedings of the National Academy of Sciences. 1980;77:503 - 507.
263. Herskowitz I. Life cycle of the budding yeast Saccharomyces cerevisiae. Microbiological Reviews. 1988;52:536- 553.
264. Le S, Moore JK, Haber JE, Greider CW. RAD50 and RAD51 Define Two Pathways That Collaborate to Maintain Telomeres in the Absence of Telomerase. Genetics. 1999;152:143-152.
265. Wyman C, Kanaar R. Homologous Recombination: Down to the Wire. Current Biology. 2004;14:629 - 631.
266. Teng S-C, Chang J, McCowan B, Zakian VA. Telomerase - Independent Lengthening of Yeast Telomeres Occurs by an Abrupt Rad50p - Dependent Rif - Inhibited Recombinational Process. Molecular Cell. 2000;6:947 - 952.
267. Bryan TM, Englezou A, Dalla-Pozza L. Evidence for an alternative mechanism for maintaining telomere length in human tumours and tumour-derived cell lines. Nature Medicine. 1997;3:1271-1274.
268. Kraus E, Leung W-Y, Haber JE. Break-induced replication: A review and an example in budding yeast. Proceedings of the National Academy of Sciences. 2001;98:8255-8262.
269. Bosco G, Haber JE. Chromosome Break-Induced DNA Replication Leads to Nonreciprocal Translocations and Telomere Capture. Genetics. 1998;150:1037-1047.
270. Smith CE, Llorente B, Symington LS. Template switching during break-induced replication. Nature. 2007;447:102-105.
271. Deininger PL, Batzer MA. Alu Repeats and Human Disease. Molecular Genetics and Metabolism. 1999;67:183-193.
272. Li B, Lustig AJ. A novel mechanism for telomere size control in *Saccharomyces cerevisiae*. Genes & Development. 1996;10:1310-1326.
273. Lustig AJ. Clues to catastrophic telomere loss in mammals from yeast telomere rapid deletion. Nat Rev Genet. 2003;4:916-923.

274. Lin C-Y, Chang H-H, Wu K-J, et al. Extrachromosomal Telomeric Circles Contribute to Rad52-, Rad50-, and Polymerase δ-Mediated Telomere-Telomere Recombination in *Saccharomyces cerevisiae*. Eukaryotic Cell. 2005;4:327-336.
275. Zhu X-D, Niedernhofer L, Kuster B, Mann M, Hoeijmakers JHJ, de Lange T. ERCC1/XPF Removes the 3' Overhang from Uncapped Telomeres and Represses Formation of Telomeric DNA - Containing Double Minute Chromosomes. Molecular Cell. 2003;12:1489- 1498.
276. Bressan DA, Baxter BK, Petrini JHJ. The Mre11-Rad50-Xrs2 Protein Complex Facilitates Homologous Recombination-Based Double-Strand Break Repair in *Saccharomyces cerevisiae*. Molecular and Cellular Biology. 1999;19:7681-7687.
277. Deng Y, Guo X, Ferguson DO, Chang S. Multiple Roles for MRE11 at uncapped telomeres. Nature. 2009;460:914- 918.
278. Zhu X-D, Küster B, Mann M, Petrini JHJ, de Lange T. Cell- cycle- regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. Nature Genetics. 2000;25:347-352.
279. Nugent CI, Bosco G, Ross LO, et al. Telomere maintenance is dependent on activities required for end repair of double-strand breaks. Current biology : CB. 1998;8:657-662.
280. Kramer KM, Brock JA, Bloom K, Moore JK, Haber JE. Two different types of double-strand breaks in *Saccharomyces cerevisiae* are repaired by similar RAD52-independent, nonhomologous recombination events. Molecular and Cellular Biology. 1994;14:1293-1301.
281. Smogorzewska A, Karlseder J, Holtgreve-Grez H, Jauch A, de Lange T. DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. Current Biology. 2002;12:1635-1644.
282. Weaver DT. What to do at an end: DNA double-strand-break repair. Trends in Genetics. 1995;11:388- 392.
283. Lustig AJ. The Kudos of non-homologous end-joining. Nat Genet. 1999;23:130-131.
284. McVey M, Lee SE. MMEJ repair of double- strand breaks (director's cut): deleted sequences and alternative endings. Cell 2008.
285. Gu Y, Seidl KJ, Rathbun GA, et al. Growth Retardation and Leaky SCID Phenotype of Ku70-Deficient Mice. Immunity. 1997;7:653-665.
286. Meek K, Gupta S, Ramsden DA, Lees-Miller SP. The DNA- dependent protein kinase: the director at the end. Immunological reviews 2004;200:132-141.
287. Ma Y, Schwarz K, Lieber MR. The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. DNA Repair. 2005;4:845-851.
288. Critchlow SE, Bowater RP, Jackson SP. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. Current biology : CB. 1997;7:588-598.
289. Kuhfittig- Kulle S, Fledmann E, Odersky A, et al. The mutagenic potential of non-homologous end joining in the absence of the NHEJ core factors Ku70/80, DNA -PKcs and XRCC4- LigIV. Mutagenesis. 2007;22:217- 233.
290. d'Adda di Fagagna F, Hande MP, Tong W-M, et al. Effects of DNA nonhomologous end-joining factors on telomere length and chromosomal stability in mammalian cells. Current biology : CB. 2001;11:1192-1196.
291. Bucholc M, Park Y, Lustig AJ. Intrachromatid Excision of Telomeric DNA as a Mechanism for Telomere Size Control in Saccharomyces cerevisiae. Molecular and Cellular Biology. 2001;21:6559-6573.
292. Jaco I, Munoz Pn, Blasco MaA. Role of Human Ku86 in Telomere Length Maintenance and Telomere Capping. Cancer Research. 2004;64:7271-7278.

293. Bailey SM, Brenneman MA, Halbrook J, Nickoloff JA, Ullrich RL, Goodwin EH. The kinase activity of DNA-PK is required to protect mammalian telomeres. DNA Repair. 2004;3:225-233.
294. Maser RS, Wong K-K, Sahin E, et al. DNA-Dependent Protein Kinase Catalytic Subunit Is Not Required for Dysfunctional Telomere Fusion and Checkpoint Response in the Telomerase-Deficient Mouse. Molecular and Cellular Biology. 2007;27:2253-2265.
295. Wang H, Zeng Z-C, Bui T, et al. Efficient rejoining of radiation-induced DNA double-strand breaks in vertebrate cells deficient in genes of the *RAD52* epistasis group. Oncogene. 2001;20:2212-2224.
296. Sharpless NE, Ferguson DO, O'Hagan RC, et al. Impaired Nonhomologous End-Joining Provokes Soft Tissue Sarcomas Harboring Chromosomal Translocations, Amplifications, and Deletions. Molecular Cell. 2001;8:1187-1196.
297. Mortimer RK, Contopoulou R, Schild D. Mitotic chromosome loss in a radiation-sensitive strain of the yeast *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences. 1981;78:5778-5782.
298. Kabotyanski EB, Gomelsky L, Han J, Stamato TD, Roth DB. Double - strand break repair in Ku86 - and XRCC4- deficient cells Nucleic Acids Research. 1998;26:5333-5342.
299. Morris T, Thacker J. Formation of large deletions by illegitimate recombination in the HPRT gene of primary human fibroblasts. Proceedings of the National Academy of Sciences. 1993;90:1392-1396.
300. Canning S, Dryja TP. Short, direct repeats at the breakpoints of deletions of the retinoblastoma gene. Proceedings of the National Academy of Sciences. 1989;86:5044-5048.
301. Bentley J, Diggle CP, Harnden P, Knowles MA, Kiltie AE. DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. Nucleic Acids Research. 2004;32:5249-5259.
302. Kornreich R, Bishop DF, Desnick RJ. Alpha-galactosidase A gene rearrangements causing Fabry disease. Identification of short direct repeats at breakpoints in an Alu-rich gene. Journal of Biological Chemistry. 1990;265:9319-9326.
303. Woods-Samuels P, Kazazian Jr HH, Antonarakis SE. Nonhomologous recombination in the human genome: Deletions in the human factor VIII gene. Genomics. 1991;10:94-101.
304. Lin TT, Letsolo BT, Jones RE, et al. Telomere dysfunction and fusion during the progression of chronic lymphocytic leukemia: evidence for a telomere crisis. Blood. 2010;116:1899-1907.
305. Tankimanova M, Capper R, Letsolo BT, et al. Mre11 modulates the fidelity of fusion between short telomeres in human cells. Nucleic Acids Research. 2012;40:2518-2526.
306. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends in Genetics. 2008;24:529-538.
307. McVey M, Radut D, Sekelsky JJ. End-Joining Repair of Double-Strand Breaks in Drosophila melanogaster Is Largely DNA Ligase IV Independent. Genetics. 2004;168:2067-2076.
308. Audebert M, Salles B, Calsou P. Effect of double-strand break DNA sequence on the PARP-1 NHEJ pathway. Biochemical and Biophysical Research Communications. 2008;369:982-988.
309. Sallmyr A, Tomkinson AE, Rassool FV. Up-regulation of WRN and DNA ligase IIIα in chronic myeloid leukemia: consequences for the repair of DNA double-strand breaks. Blood. 2008;112:1413-1423.

310. Bachrati CZ, Hickson ID. RecQ helicases: suppressors of tumourigenesis and premature aging. The Biochemical Journal. 2003;374:577-606.
311. Daley JM, Wilson TE. Rejoining of DNA Double-Strand Breaks as a Function of Overhang Length. Molecular and Cellular Biology. 2005;25:896-906.
312. Ahmad A, Robinson AR, Duensing A, et al. ERCC1-XPF Endonuclease Facilitates DNA Double-Strand Break Repair. Molecular and Cellular Biology. 2008;28:5082-5092.
313. Martens UM, Chavez EA, Poon SSS, Schmoor C, Lansdorp PM. Accumulation of Short Telomeres in Human Fibroblasts Prior to Replicative Senescence. Experimental Cell Research. 2000;256:291-299.
314. d' Adda di Fagagna F, Teo SH, Jackson SP. Functional links between telomeres and proteins of the DNA - damage response. Genes & Development. 2004;18:1781- 1799.
315. Takai H, Smogorzewska A, De Lange T. DNA Damage Foci at Dysfunctional Telomeres. Current Biology. 2003;13:1549- 1556.
316. Herbig U, Jobling WA, Chen BP, Chen DJ, Sedivy JM. Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). Mol Cell. 2004;14:501-513.
317. d' Adda di Fagagna F, Reaper PM, Clay - Farrace L, et al. A DNA damage checkpoint response in telomere - initiated senescence. Nature. 2003;426:194- 197.
318. Counter CM, Avilion AA, LeFeuvre CE, et al. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J. 1992;11:1921- 1929.
319. McClintock B. The Stability of Broken Ends of Chromosomes in Zea Mays. Genetics. 1941;26:234-282.
320. Gisselsson D, Gorunova L, Hoglund M, Mandahl N, Elfving P. Telomere shortening and mitotic dysfunction generate cytogenetic heterogeneity in a subgroup of renal cell carcinomas. Br J Cancer. 2004;91:327-332.
321. Gisselsson D, Jonson T, Petersen A, et al. Telomere dysfunction triggers extensive DNA fragmentation and evolution of complex chromosome abnormalities in human malignant tumors. Proceedings of the National Academy of Sciences. 2001;98:12683-12688.
322. Artandi SE, Chang S, Lee S-L, et al. Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. Nature. 2000;406:641-645.
323. Shay JW, Wright WE. Role of telomeres and telomerase in cancer. Seminars in Cancer Biology. 2011;21:349-353.
324. Petrini JHJ, Bressan DA, Yao MS. The RAD52 epistasis group in mammalian double strand break repair. Seminars in Immunology. 1997;9:181-188.
325. Vaziri H, Dragowska W, Allsopp RC, Thomas TE, Harley CB, Lansdorp PM. Evidence for a mitotic clock in human hematopoietic stem cells: loss of telomeric DNA with age. Proceedings of the National Academy of Sciences. 1994;91:9857-9860.
326. Hastie ND, Dempster M, Dunlop MG, Thompson AM, Green DK, Allshire RC. Telomere reduction in human colorectal carcinoma and with ageing. Nature. 1990;346:866-868.
327. Notaro R, Cimmino A, Tabarini D, Rotoli B, Luzzatto L. In vivo telomere dynamics of human hematopoietic stem cells. Proc Natl Acad Sci U S A. 1997;94:13782-13785.
328. Engelhardt M, Wasch R, Guo Y. Telomeres and telomerase in normal and leukemic hematopoietic cells. Leuk Res. 2004;28:1001-1004.

329. Vulliamy T, Marrone A, Szydlo R, Walne A, Mason PJ, Dokal I. Disease anticipation is associated with progressive telomere shortening in families with dyskeratosis congenita due to mutations in TERC. Nat Genet. 2004;36:447-449.
330. Rudolph KL, Chang S, Lee HW, et al. Longevity, stress response, and cancer in aging telomerase-deficient mice. Cell. 1999;96:701-712.
331. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in dyskeratosis congenita. Blood. 2009;113:6549-6557.
332. Ohyashiki JH, Iwama H, Yahata N, et al. Telomere Stability Is Frequently Impaired in High-Risk Groups of Patients with Myelodysplastic Syndromes. Clinical Cancer Research. 1999;5:1155-1160.
333. Boultwood J, Fidler C, Kusec R, et al. Telomere length in myelodysplastic syndromes. American Journal of Hematology. 1997;56:266-271.
334. Ohyashiki JH, Ohyashiki K, Fujimura T, et al. Telomere Shortening Associated with Disease Evolution Patterns in Myelodysplastic Syndromes. Cancer Research. 1994;54:3557-3560.
335. Engelhardt M, Mackenzie K, Drullinsky P, Silver RT, Moore MAS. Telomerase Activity and Telomere Length in Acute and Chronic Leukemia, Pre- and Post-ex Vivo Culture. Cancer Research. 2000;60:610-617.
336. Sieglova Z, Zilovcova S, Cermak J, et al. Dynamics of telomere erosion and its association with genome instability in myelodysplastic syndromes (MDS) and acute myelogenous leukemia arising from MDS: a marker of disease prognosis? Leuk Res. 2004;28:1013-1021.
337. Ball SE, Gibson FM, Rizzo Sn, Tooze JA, Marsh JCW, Gordon-Smith EC. Progressive Telomere Shortening in Aplastic Anemia. Blood. 1998;91:3582-3592.
338. Calado RT, Cooper JN, Padilla-Nash HM, et al. Short telomeres result in chromosomal instability in hematopoietic cells and precede malignant evolution in human aplastic anemia. Leukemia. 2012;26:700-707.
339. Boultwood J, Peniket A, Watkins F, et al. Telomere length shortening in chronic myelogenous leukemia is associated with reduced time to accelerated phase. Blood. 2000;96:358-361.
340. Brummendorf TH, Holyoake TL, Rufer N, et al. Prognostic implications of differences in telomere length between normal and malignant cells from patients with chronic myeloid leukemia measured by flow cytometry. Blood. 2000;95:1883-1890.
341. Roos Gr, Krober A, Grabowski P, et al. Short telomeres are associated with genetic complexity, high-risk genomic aberrations, and short survival in chronic lymphocytic leukemia. Blood. 2008;111:2246-2252.
342. Swiggers SJJ, Kuijpers MA, de Cort MJM, Beverloo HB, Zijlmans JMJM. Critically short telomeres in acute myeloid leukemia with loss or gain of parts of chromosomes. Genes, Chromosomes and Cancer. 2006;45:247-256.
343. Bartkova J, Horejsi Z, Koed K, et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature. 2005;434:864-870.
344. Horibe S, Takagi M, Unno J, et al. DNA damage check points prevent leukemic transformation in myelodysplastic syndrome: Leukemia. 2007 Oct;21(10):2195-8. Epub 2007 May 10.; 2007.
345. Boehrer S, Ades L, Tajeddine N, et al. Suppression of the DNA damage response in acute myeloid leukemia versus myelodysplastic syndrome. Oncogene. 2009;28:2205-2218.

346. Rigolin GM, Porta MD, Bugli AM, et al. Flow cytometric detection of accelerated telomere shortening in myelodysplastic syndromes: correlations with aetiological and clinical–biological findings*. European Journal of Haematology. 2004;73:351-358.
347. Hartmann U, Brummendorf T, Balabanov S, Thiede C, Illme T, Schaich M. Telomere length and hTERT expression in patients with acute myeloid leukemia correlates with chromosomal abnormalities. Haematologica. 2005;90:307-316.
348. Lange K, Holm L, Vang Nielsen K, et al. Telomere shortening and chromosomal instability in myelodysplastic syndromes. Genes, Chromosomes and Cancer. 2010;49:260-269.
349. Sashida G, Ohyashiki JH, Nakajima A, et al. Telomere Dynamics in Myelodysplastic Syndrome Determined by Telomere Measurement of Marrow Metaphases. Clinical Cancer Research. 2003;9:1489-1496.
350. Aalbers AM, Calado RT, Young NS, et al. Telomere length and telomerase complex mutations in pediatric acute myeloid leukemia. Leukemia;27:1786-1789.
351. Letsolo BT, Rowson J, Baird DM. Fusion of short telomeres in human cells is characterized by extensive deletion and microhomology, and can result in complex rearrangements. Nucleic Acids Research. 2010;38:1841-1852.
352. Baird DM, Britt-Compton B, Rowson J, Amso NN, Gregory L, Kipling D. Telomere instability in the male germline. Human Molecular Genetics. 2006;15:45-51.
353. Britt-Compton B, Rowson J, Locke M, Mackenzie I, Kipling D, Baird DM. Structural stability and chromosome-specific telomere length is governed by cis-acting determinants in humans. Human Molecular Genetics. 2006;15:725-733.
354. Martens UM, Zijlmans JM, Poon SS, et al. Short telomeres on human chromosome 17p. Nat Genet. 1998;18:76-80.
355. Harris NL, Jaffe ES, Diebold J, et al. The World Health Organization Classification of Neoplastic Diseases of the Hematopoietic and Lymphoid Tissues. Annals of Oncology. 1999;10:1419-1432.
356. Herens C, Brasseur E, Jamar M, Vierset L, Schoenen I, Koulischer L. Loss of the Y chromosome in bone marrow cells: results on 1907 consecutive cases of leukaemia and preleukaemia. Clinical & Laboratory Haematology. 1999;21:17-20.
357. Kuchenbauer F, Schnittger S, Look T, et al. Identification of additional cytogenetic and molecular genetic abnormalities in acute myeloid leukaemia with t(8;21)/AML1-ETO. British Journal of Haematology. 2006;134:616-619.
358. Rappold G, Willson TA, Henke A, Gough NM. Arrangement and localization of the human GM-CSF receptor alpha chain gene CSF2RA within the X-Y pseudoautosomal region. Genomics. 1992;14:455-461.
359. Matsuura S, Yan M, Lo M-C, et al. Negative effects of GM-CSF signaling in a murine model of t(8;21)-induced leukemia. Blood. 2012;119:3155-3163.
360. Mijovic A, Mufti GJ. The myelodysplastic syndromes: towards a functional classification. Blood Reviews. 1998;12:73-83.
361. Bennett JM. Myelodysplastic syndromes: update on classification. Leukemia Research. 2009;33, Supplement 1:S6-S8.
362. Hao L-Y, Armanios M, Strong MA, et al. Short Telomeres, even in the Presence of Telomerase, Limit Tissue Renewal Capacity. Cell. 2005;123:1121-1131.
363. Armanios M, Alder JK, Parry EM, Karim B, Strong MA, Greider CW. Short Telomeres are Sufficient to Cause the Degenerative Defects Associated with Aging. American journal of human genetics. 2009;85:823-832.

364. Mohamedali A, Gaken J, Twine NA, et al. Prevalence and prognostic significance of allelic imbalance by single-nucleotide polymorphism analysis in low-risk myelodysplastic syndromes. Blood. 2007;110:3365-3373.
365. Gondek LP, Tiu R, O'Keefe CL, Sekeres MA, Theil KS, Maciejewski JP. Chromosomal lesions and uniparental disomy detected by SNP arrays in MDS, MDS/MPD, and MDS-derived AML. Blood. 2008;111:1534-1542.
366. Walter MJ, Payton JE, Ries RE, et al. Acquired copy number alterations in adult acute myeloid leukemia genomes. Proceedings of the National Academy of Sciences. 2009;106:12950-12955.
367. Colovic N, Tomin D, Vidovic A, et al. Pretreatment prognostic factors for overall survival in primary resistant acute myeloid leukemia. Biomed Pharmacother. 2012;66:578-582.
368. Parkin B, Erba H, Ouillette P, et al. Acquired genomic copy number aberrations and survival in adult acute myelogenous leukemia. Blood;116:4958-4967.
369. Grimwade D, Walker H, Oliver F, et al. The Importance of Diagnostic Cytogenetics on Outcome in AML: Analysis of 1,612 Patients Entered Into the MRC AML 10 Trial. Blood. 1998;92:2322-2333.
370. Slovak ML, Kopecky KJ, Cassileth PA, et al. Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group study. Blood. 2000;96:4075-4083.
371. Schoch C, Kern W, Kohlmann A, Hiddemann W, Schnittger S, Haferlach T. Acute myeloid leukemia with a complex aberrant karyotype is a distinct biological entity characterized by genomic imbalances and a specific gene expression profile. Genes, Chromosomes and Cancer. 2005;43:227-238.
372. Milosevic JD, Puda A, Malcovati L, et al. Clinical significance of genetic aberrations in secondary acute myeloid leukemia. Am J Hematol. 2012;87:1010-1016.
373. Leith CP, Kopecky KJ, Godwin J, et al. Acute Myeloid Leukemia in the Elderly: Assessment of Multidrug Resistance (MDR1) and Cytogenetics Distinguishes Biologic Subgroups With Remarkably Distinct Responses to Standard Chemotherapy. A Southwest Oncology Group Study. Blood. 1997;89:3323-3329.
374. Gilliland DG, Griffin JD. Role of FLT3 in leukemia. Curr Opin Hematol. 2002;9:274-281.
375. Paietta E, Neuberg D, Bennett JM, et al. Low expression of the myeloid differentiation antigen CD65s, a feature of poorly differentiated AML in older adults: study of 711 patients enrolled in ECOG trials. Leukemia. 2003;17:1544-1550.
376. Xu D, Gruber A, Peterson C, Pisa P. Telomerase activity and the expression of telomerase components in acute myelogenous leukaemia. British Journal of Haematology. 1998;102:1367-1375.
377. Counter C, Gupta J, Harley C, Leber B, Bacchetti S. Telomerase activity in normal leukocytes and in hematologic malignancies. Blood. 1995;85:2315-2320.
378. Yamamoto Y, Kiyoi H, Nakano Y, et al. Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. Blood. 2001;97:2434-2439.
379. Frohling S, Schlenk RF, Breitruck J, et al. 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. 2002;100:4372-4380.

380. Choudhary C, Schwable J, Brandts C, et al. AML-associated Flt3 kinase domain mutations show signal transduction differences compared with Flt3 ITD mutations. Blood. 2005;106:265-273.
381. Leischner H, Albers C, Grundler R, et al. SRC is a signaling mediator in FLT3-ITD- but not in FLT3-TKD- positive AML. Blood;119:4026-4033.
382. Piloto O, Wright M, Brown P, Kim KT, Levis M, Small D. Prolonged exposure to FLT3 inhibitors leads to resistance via activation of parallel signaling pathways. Blood. 2007;109:1643-1652.
383. Li L, Piloto O, Kim K-T, et al. FLT3/ITD expression increases expansion, survival and entry into cell cycle of human haematopoietic stem/progenitor cells. British Journal of Haematology. 2007;137:64-75.
384. Thiede C, Steudel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis: Presented in part at the 42nd Annual Meeting of the American Society of Hematology, December 1-5, 2000, San Francisco, CA (abstract 2334). Blood. 2002;99:4326-4335.
385. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Research. 2004;64:3753-3756.
386. Bussing I, Slack FJ, Grosshans H. let-7 microRNAs in development, stem cells and cancer. Trends in Molecular Medicine. 2008;14:400-409.
387. Seedhouse CH, Pallis M, Grundy M, Shang S, Russell NH. FLT3-ITD expression levels and their effect on STAT5 in AML with and without NPM mutations. Br J Haematol. 2009;147:653-661.
388. Quesnel B, Guillerm G, Vereecque R, et al. Methylation of the p15(INK4b) gene in myelodysplastic syndromes is frequent and acquired during disease progression. Blood. 1998;91:2985-2990.
389. Debacker K, Kooy RF. Fragile Sites and Human Disease. Human Molecular Genetics. 2007;16:150-158.
390. Barouk-Simonet E, Soenen-Cornu Vr, Roumier C, et al. Role of multiplex FISH in identifying chromosome involvement in myelodysplastic syndromes and acute myeloid leukemias with complex karyotypes: a report on 28 cases. Cancer Genetics and Cytogenetics. 2005;157:118-126.
391. Babicka L, Ransdorfova S, Brezinova J, et al. Analysis of complex chromosomal rearrangements in adult patients with MDS and AML by multicolor FISH. Leukemia Research. 2007;31:39 - 47.
392. Audebert M, Salles B, Weinfeld M, Calsou P. Involvement of Polynucleotide Kinase in a Poly(ADP-ribose) Polymerase-1-dependent DNA Double-strand Breaks Rejoining Pathway. Journal of Molecular Biology. 2006;356:257-265.
393. Paull TT, Gellert M. A mechanistic basis for Mre11-directed DNA joining at microhomologies. Proceedings of the National Academy of Sciences. 2000;97:6409-6414.
394. Taborelli M, Tibiletti MG, Martin V, Pozzi B, Bertoni F, Capella C. Chromosome band 6q deletion pattern in malignant lymphomas. Cancer Genetics and Cytogenetics. 2006;165:106-113.
395. Merup M, Moreno TC, Heyman M, et al. 6q Deletions in Acute Lymphoblastic Leukemia and Non-Hodgkin's Lymphomas. Blood. 1998;91:3397-3400.

396. Jackson A, Carrara P, Duke V, et al. Deletion of 6q16-q21 in Human Lymphoid Malignancies: A Mapping and Deletion Analysis. Cancer Research. 2000;60:2775-2779.
397. Mancini M, Vegna ML, Castoldi GL, et al. Partial deletions of long arm of chromosome 6: biologic and clinical implications in adult acute lymphoblastic leukemia. Leukemia. 2002;16:2055-2061.
398. Banga SS, Kim S, Hubbard K, et al. SEN6, a locus for SV40-mediated immortalization of human cells, maps to 6q26-27. Oncogene. 1997;14:313-321.
399. Gisselsson D. Mechanisms of whole chromosome gains in tumors--many answers to a simple question. Cytogenetic and Genome Research. 2011;133:190-201.
400. Klein A, Zang KD, Steudel WI, Urbschat S. Different mechanisms of mitotic instability in cancer cell lines. Int J Oncol. 2006;29:1389-1396.
401. Lindsey J, McGill NI, Lindsey LA, Green DK, Cooke HJ. In vivo loss of telomeric repeats with age in humans. Mutation Research/DNAging. 1991;256:45-48.
402. Li X, Leteurtre F, Rocha V, et al. Abnormal telomere metabolism in Fanconi's anaemia correlates with genomic instability and the probability of developing severe aplastic anaemia. Br J Haematol. 2003;120:836-845.
403. Wang Y, Fang M, Sun X, Sun J. Telomerase activity and telomere length in acute leukemia: correlations with disease progression, subtypes and overall survival. International Journal of Laboratory Hematology. 2009;32:230-238.
404. Capraro V, Zane L, Poncet D, et al. Telomere deregulations possess cytogenetic, phenotype, and prognostic specificities in acute leukemias. Experimental Hematology. 2011;39:195-202.e192.
405. Schuller CE, Jankowski K, Mackenzie KL. Telomere length of cord blood-derived CD34(+) progenitors predicts erythroid proliferative potential. Leukemia. 2007;21:983-991.
406. Sallmyr A, Tomkinson AE, Rassool FV. Up-regulation of WRN and DNA ligase III in chronic myeloid leukaemia: consequences for the repair of DNA double-strand breaks. Blood. 2008;112:1413-1423.
407. Puthiyaveetil AG, Reilly CM, Pardee TS, Caudell DL. Non-homologous end joining mediated DNA repair is impaired in the NUP98-HOXD13 mouse model for myelodysplastic syndrome. Leukemia Research. 2013;37:112-116.
408. Economopoulou P, Pappa V, Kontsioti F, et al. Expression analysis of proteins involved in the non homologous end joining DNA repair mechanism, in the bone marrow of adult de novo myelodysplastic syndromes. Ann Hematol. 2010;89:233-239.
409. Fan J, Li L, Small D, Rassool F. Cells expressing FLT3/ITD mutations exhibit elevated repair errors generated through alternative NHEJ pathways: implications for genomic instability and therapy. Blood. 2010;116:5298-5305.
410. Ricchetti M, Dujon B, Fairhead C. Distance from the chromosome end determines the efficiency of double strand break repair in subtelomeres of haploid yeast. J Mol Biol. 2003;328:847-862.
411. Gao Q, Reynolds GE, Wilcox A, et al. Telomerase-dependent and -independent chromosome healing in mouse embryonic stem cells. DNA Repair. 2008;7:1233-1249.
412. Gross M, Mkrtchyan H, Glaser M, et al. Delineation of yet unknown cryptic subtelomere aberrations in 50% of acute myeloid leukemia with normal GTG-banding karyotype. Int J Oncol. 2009;34:417-423.
413. Matteucci C, La Starza R, Crescenzi B, et al. Different mechanisms lead to a karyotypically identical t(20;21) in myelodysplastic syndrome and in acute myelocytic leukemia. Cancer Genetics and Cytogenetics. 2003;140:13-17.

414. Wlodarska I, Selleri L, La Starza R, et al. Molecular cytogenetics localizes two new breakpoints on 11q23.3 and 21q11.2 in myelodysplastic syndrome with t(11;21) translocation. Genes, Chromosomes and Cancer. 1999;24:199-206.
415. Cigudosa JC, Odero MD, Calasanz MJ, et al. De novo erythroleukemia chromosome features include multiple rearrangements, with special involvement of chromosomes 11 and 19. Genes, Chromosomes and Cancer. 2003;36:406-412.
416. Mitelman F, Mertens F, Johansson B. A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. Nature Genetics. 1997;15:417-474.
417. Huntsman DG, Chin SF, Muleris M, et al. *MLL2*, the second human homolog of the *Drosophila trithorax* gene, maps to 19q13.1 and is amplified in solid tumor cell lines. Oncogene. 1999;18:7975-7984.
418. Nimer SD, MacGrogan D, Jhanwar S, Alvarez S. Chromosome 19 abnormalities are commonly seen in AML, M7. Blood. 2002;100:3838-3839.
419. Barbouti A, Hoglund M, Johansson B, et al. A Novel Gene, *MSI2*, Encoding a Putative RNA-binding Protein Is Recurrently Rearranged at Disease Progression of Chronic Myeloid Leukemia and Forms a Fusion Gene with *HOXA9* as a Result of the Cryptic t(7;17)(p15;q23) Cancer Research. 2003;63:1202- 1206.
420. Kharas MG, Lengner CJ, Al-Shahrour F, et al. Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. Nature Medicine. 2010;16:903-908.
421. Jaju RJ, Boultwood J, Oliver FJ, et al. Molecular Cytogenetic Delineation of the Critical Deleted Region in the 5q- Syndrome. Genes, Chromosomes and Cancer. 1998;22:251-256.
422. Boultwood J, Fidler C, Strickson AJ, et al. Narrowing and genomic annotation of the commonly deleted region of the 5q− syndrome. Blood. 2002;99:4638-4641.
423. Zhao N, Stoffel A, Wang PW, et al. Molecular delineation of the smallest commonly deleted region of chromosome 5 in malignant myeloid diseases to 1-1.5 Mb and preparation of a PAC-based physical map. Proc Natl Acad Sci U S A. 1997;94:6948-6953.
424. Lai F, Godley LA, Joslin J, et al. Transcript map and comparative analysis of the 1.5-Mb commonly deleted segment of human 5q31 in malignant myeloid diseases with a del(5q). Genomics. 2001;71:235-245.
425. Ohyashiki JH, Hayashi S, Yahata N, et al. Impaired telomere regulation mechanism by TRF1 (telomere-binding protein), but not TRF2 expression, in acute leukemia cells. Int J Oncol. 2001;18:593-598.
426. Mufti GJ, Bennett JM, Goasguen J, et al. Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of myelodysplastic syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts. Haematologica. 2008;93:1712-1717.