Queen Mary
University of London

Investigation of distal repetitive sequences in the genus allium

Chester, Michael

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author

For additional information about this publication click this link.
https://qmro.qmul.ac.uk/jspui/handle/123456789/444

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

Investigation of Distal Repetitive Sequences in the Genus *Allium*

Michael Chester

Thesis submitted for the degree of

Doctor of Philosophy

Queen Mary, University of London

2009

Abstract

The telomere is a DNA/protein structure required to maintain the ends of linear chromosomes. Usually the DNA component comprises a highly conserved tandemly repeated minisatellite sequence. In most plants the minisatellite sequence is typically present in several hundred copies at each chromosome end, and is extended primarily by telomerase, which adds telomere repeats to the 3' end. In the plant genus *Allium*, which contains around 700 species, there is an absence of typical telomeric DNA repeats. It is of great interest to determine what sequence or sequences have replaced the ancestral repeats and how they are lengthened.

A range of molecular cloning methods were used to isolate candidate telomere sequences from the genomes of two diverged species, *Allium cernuum* and *Allium cepa*. I analyse several putative telomere sequences, isolated in this work and by others, but no proven candidate sequence has emerged. Nevertheless, one of those sequences, 35S ribosomal DNA (rDNA) encoding 35S rRNA, proved to have a structure that is previously not described for plants. I show that some units have a Ty1/*copia* retrotransposon fragment in the intergenic spacer region. Sequence analysis indicates that there was a single insertion followed by amplification, probably involving homogenisation mechanisms. Furthermore, I show high levels of rDNA length heterogeneity and rDNA unit divergence both within species and across the genus, respectively.

Acknowledgements

At QMUL, I thank Andrew Leitch for his excellent supervision, endless enthusiasm and support throughout the project. I thank Yoong Lim who taught me chromosome squashing, provided many ideas and inspiration. I thank Elizabeth McCarthy, Monika Bozek, Paris Veltsos, Simon Renny-Byfield, Felipe Sánchez Teyer and Danaraj Joseph for their help and companionship.

At The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) in Gatersleben, Germany, I thank Katrin Kumke for carrying out the chromosome microdissection and Andreas Houben who kindly provided training as well making me welcome in his laboratory during a short visit in the summer of 2005.

At The Department of Functional Genomics and Proteomics, Masaryk University in Bruno, Czech Republic I thank Eva Sýkorová who shared her expertise and contributed several important pieces of experimental data for this thesis. I thank Zuzana Kunická for teaching me several molecular techniques. I thank Martina Dvořáčková, Kamila Neplechová, Ctirad Hofr, Petra Schrumpfová and Marcela Šefčíková for training and laboratory assistance. I am very grateful to Jiri Fajkus allowing me to spend three months in his laboratory in the spring of 2006 and for his continued encouragement.

At The Royal Botanic Gardens, Kew, I wish to thank Mike Fay for valuable discussions concerning Asparagales taxonomy and Richard Wilford for providing seeds of *Tulbaghia fragrans*.

At The French National Institute for Agricultural Research (INRA) I wish to thank Marie-Angèle Grandbastien for sharing her expertise in plant retrotransposons.

In memory of Yoong Lim.

Declaration

I declare that all the work in this thesis is my own with the exception of the work presented in the following figures: 4.3.4 on page 81, 4.3.4 on page 82, 5.3.3 on page 91, 5.3.3 on page 93, which were carried out be Eva Sýkorová.

Contents

1 General Introduction **15**

1.1 The Telomere . . . 15

1.2 Telomerase . . . 18

1.3 Evolution of Telomerase and the Telomere
in the Asparagales . . . 19

1.4 Non-Telomerase Maintained Telomeres . . . 21

1.4.1 Alternative Lengthening of Telomeres (ALT) . . . 22

1.4.1.1 ALT mechanisms . . . 22

1.4.1.2 Long-Term ALT in Insects . . . 23

1.4.2 Retrotransposons as Telomeres . . . 24

1.5 Isolation of Telomeric Sequences . . . 25

1.6 Aims and Overview . . . 26

2 Materials and Methods **27**

2.1 Plant materials . . . 27

2.1.1 Living Plant Accessions (QMUL) . . . 27

2.1.2 DNA from Living Collections . . . 27

2.2 DNA Extraction and Purification . . . 28

2.2.1 Genomic DNA Extraction . . . 28

2.2.2 RNA and Protein Digestion . . . 28

2.2.3 Phenol:Chloroform Extraction . . . 28

2.2.4 DNA Precipitation . . . 29

2.3 Chromosome Microdissection . . . 29

2.4 Polymerase Chain Reaction . . . 29

2.4.1 Standard PCR . . . 29

2.4.2 PCR Labelling for FISH and EDF-FISH . . . 30

2.4.3 Long Range PCR 30
2.4.4 DNA Gel Electrophoresis 30
2.4.5 Degenerate Oligonucleotide Primed-PCR (DOP-PCR) 31
2.5 Southern Hybridisation 31
2.5.1 DNA Restriction Digests 31
2.5.2 Southern Transfer 31
2.5.3 Radioactive DNA Southern Hybridisation 32
2.6 Cloning and Colony Screening 33
2.6.1 Ligation and Transformation 33
2.6.2 Minipreparation Plasmid DNA Extraction 33
2.6.3 Glycerol Stocks 33
2.6.4 Plasmid Library Preparation 33
2.6.5 Plasmid Library Screening 34
2.7 DNA Sequencing and Analysis 34
2.7.1 DNA Sequencing 34
2.7.2 DNA Sequence Alignment 34
2.7.3 Phylogenetic Analysis 35
2.7.4 Median Joining Networks 35
2.7.5 Basic Local Alignment Search Tool (BLAST) 35
2.7.6 Sequence Randomisation 35
2.7.7 Secondary Structure Prediction 35
2.8 Fluorescent *In Situ* Hybridisation 36
2.8.1 Chromosome Preparation 36
2.8.2 Probe Preparation 37
2.8.3 Hybridisation Solution 37
2.8.4 Hybridisation 37
2.8.5 Post-Hybridisation Washing 37
2.8.6 Detection 37
2.9 Extended DNA Fibre (EDF) FISH 38
2.9.1 Nuclei Extraction 38
2.9.2 DNA Fibre Extension 39
2.9.3 Hybridisation Solution 39
2.9.4 Hybridisation 39
2.9.5 Post-Hybridisation Washing 39

2.9.6 Detection 40

3 *Allium* Tandem Repeats and Sequences Isolated by DOP-PCR 41

3.1 Introduction 41

3.1.1 Repetitive DNA 41

3.1.1.1 Tandem Satellite Repeats 42

3.1.2 Characterisation of Repetitive DNA 43

3.1.3 *Allium* Genome Organisation 43

3.1.4 *Allium* Genome Evolution 44

3.2 Evolution of Repetitive DNA Families 45

3.2.1 Concerted Evolution of 35S rDNA 46

3.2.2 35S rDNA Position Variability in *Allium* 48

3.2.3 *Candida albicans* Telomere Repeat Sequence 48

3.3 Materials and Methods 49

3.3.1 Microdissection, DOP-PCR and Colony Blotting 49

3.3.2 FISH 50

3.3.3 EDF-FISH 51

3.3.4 Tandem Repeat Isolation 51

3.3.5 *C. albicans* Repeat Isolation - PCR 51

3.3.6 *C. albicans* Repeat Isolation - DNA Capture 51

3.4 Results 52

3.4.1 rDNA Localisation 52

3.4.2 Tandem Repeat Isolation 53

3.4.3 Microdissection / DOP-PCR Isolated Sequences 53

3.4.3.1 Sequencing and FISH 53

3.4.3.2 Secondary Structure 55

3.4.4 Telomere Sequence from *Candida albicans* 58

3.5 Discussion 60

3.5.1 DOP-PCR Isolated Repetitive DNA 60

3.5.2 Possible Roles of DNA Secondary Structures 62

3.5.3 Minor 35S rDNA Sites 63

3.5.4 35S rDNA Mobility 63

3.5.5 Links between Telomeres and rDNA 65

3.5.5.1 rDNA as a Recombination Hot Spot 66

3.5.5.2 Interstitial NORs 66
3.5.5.3 rDNA Involvement in Chromosome Fission and Fusion . . . 67
3.5.5.4 rDNA behaviour in Telomerase Knock-Out Plants 68
3.5.5.5 rDNA and Telomere G-Quadruplexes 69
3.5.5.6 rDNA and Telomere DNA Unwinding 69
3.5.5.7 Telomeres and the Nucleolus 71
3.5.5.8 Summary of Links between rDNA and Telomeres: 71

4 Organisation of the Ribosomal DNA Intergenic Spacer in *Allium cernuum* and *Allium cepa* 72

4.1 Introduction 72
4.1.1 IGS Organisation 72
4.1.2 IGS Evolution 74
4.2 Materials and Methods 76
4.2.1 IGS Amplification and Sequencing 76
4.2.2 DNA Restriction Digests and Southern Hybridisation 76
4.2.3 Nucleotide Sequence Analysis 77
4.3 Results 78
4.3.1 IGS Structure of *Allium cepa* 78
4.3.2 IGS Structure of *Allium cernuum* 78
4.3.3 Overview of the *Allium* IGS 79
4.3.4 Southern Hybridisation of *Allium* species 80
4.4 Discussion 81
4.4.1 *Allium* 35S rDNA IGS 83

5 Amplification of a Retrotransposon Fragment within the rDNA of *Allium cernuum* 85

5.1 Introduction 85
5.1.1 rDNA Homogeneity 85
5.1.2 Targeted Insertion of Non-LTR Retroelements into the 26S/28S rDNA Unit 86
5.2 Materials and Methods 86
5.2.1 26S-IGS Sequencing 86
5.2.2 DNA Sequence Analysis 87
5.2.3 Southern Hybridisation Analysis 88

5.2.4 FISH . . . 88
5.2.5 EDF-FISH . . . 89
5.3 Results . . . 89
5.3.1 *Allium cernuum* IGS Sequence Variability . . . 89
5.3.2 Ty1/*copia*-like Sequence within the IGS of *Allium cernuum* . . . 89
5.3.3 Southern Analysis of Ty1/*copia* rDNA Insertions . . . 91
5.3.4 FISH of Retroelement, rDNA and A-Subrepeats . . . 92
5.3.5 Relationships and Diversity of the 26S-IGS Sequences . . . 96
5.4 Discussion . . . 98
5.4.1 Integration of a Ty1/*copia*-like Element into 35S rDNA . . . 98
5.4.2 Evolution of the class II IGS Region . . . 100

6 General Discussion **102**
6.0.3 35S rDNA Evolution . . . 103
6.1 Future Work . . . 103
6.1.1 A New Approach to Isolating the Telomere of *A. cepa* . . . 103
6.1.2 Obtaining an Overview of Total rDNA Variation and its Evolution . . . 104

A Southern Hybridisation Summary Table **105**

B GenBank Accession Numbers **107**

C DNA Sequence Alignments **109**

List of Figures

1.1 Telomere and G-Rich DNA structures 17

1.2 Changes in plant telomere sequences 20

3.1 Karyotype of *Allium cernuum* 52

3.2 35S rDNA minor sites in *Allium cernuum* 53

3.3 35S rDNA and ACSAT localisation in *Allium cepa* 54

3.4 FISH localisation of cer540 sequence to *Allium cernuum* 56

3.5 FISH localisation of cer703 sequence to *Allium cernuum* 57

3.6 Calculated Gibbs free energy values for predicted secondary structures of *A. cernuum* DOP-PCR repeats and 100 randomisations 58

3.7 FISH of *Allium cepa* var. Zweibel with rDNA (pTA71) and the *Candida albicans* telomere repeat 59

3.8 EDF-FISH of *Allium cepa* var. Bedfordshire Champion with 18S rDNA and the *Candida albicans* telomere repeat 60

3.9 EDF-FISH of *Allium cepa* var. Ailsa Craig with 18S rDNA and the *Candida albicans* telomere repeat 61

4.1 *Allium cepa* IGS structure 78

4.2 *Allium cernuum* IGS structure 79

4.3 Southern hybridisations of S1B1 BAC, AcH1932 probe and 26S rDNA to *A. cepa* and *A. cernuum* 81

4.4 Southern hybridisation of AcH1932 probe and 26S rDNA to *Allium* species . 82

5.1 Alignment of Ty1/*copia*-like polypeptide sequences 90

5.2 Southern hybridisation showing colocalisation of the Ty1/*copia* sequence in some rDNA units of *A. cernuum* 91

5.3 Southern hybridisation of the Ty1/*copia*-like sequence to different *Allium* species 93
5.4 FISH of rDNA derived Ty1/*copia* retroelement sequence 94
5.5 EDF-FISH of rDNA and Ty1/*copia* IGS sequence 95
5.6 FISH of IGS A-repeats 96
5.7 Pairwise nucleotide diversity (π) plot for the 26S-IGS region of class II IGSs . 97
5.8 Median joining network of intraspecific 26S-IGS sequences 98

C.1 *A. cepa* IGS A-subrepeat alignment 110
C.2 *A. cepa* IGS B-subrepeat alignment 110
C.3 *A. cernuum* IGS A-subrepeat alignment 110

List of Tables

2.1 Plant materials 27

3.1 Haploid genome sizes of *Allium* species and other plants 44

3.2 Telomere repeats of fungi (phylum Ascomycota) 49

3.3 *C. albicans* PCR amplification primers 51

3.4 DOP-PCR microdissection sequences 55

4.1 *Allium cepa* specific IGS sequencing primers 77

4.2 *Allium cernuum* specific IGS resequencing primers 77

A.1 Summary of 35S rDNA Southern hybridisation results 106

B.1 GenBank accession numbers 108

Nomenclature

ACS ARS consensus sequence

ALT alternative lengthening of telomeres

ARS autonomously replicating sequence

BIR break-induced replication

DOP-PCR degenerate oligonucleotide primed polymerase chain reaction

ECCs extrachomosomal circles

EST expressed sequence tag

ETS external transcribed spacer

NTS non-transcribed spacer

RCR rolling circle replication

SCE sister chromatid exchange

SDSA synthesis dependent strand annealing

ssDNA single strand DNA

TIS transcription initiation site

TRIM terminal repeat retrotransposon in miniature

TTS transcription termination site

Chapter 1

General Introduction

1.1 The Telomere

All eukaryotes analysed to date have linear chromosomes. Since the ancestral state for prokaryotes was a circular genome, chromosome linearisation must have taken place at some point in the lineage that gave rise to the eukaryotes. A transition from circular to linear DNA presents several difficulties for the cell including: (i) incomplete replication of the lagging strand leading to progressive end shortening; (ii) inadvertent DNA repair of chromosome ends, leading to chromosome fusion or circularisation; (iii) recombination induced instability of DNA ends; (iv) exposure to exonuclease degradation. These problems needed to be overcome in order to maintain genome integrity in eukaryotes and also independently in some prokaryotes, organelles and viruses (Lo et al. 2002, Nosek et al. 2006). Telomeres first cytologically described by Muller (1938) have evolved to address each of these problems. Telomeres are usually made up of a short repetitive DNA sequence at the chromosome end complexed with proteins. Proteins found at the telomere of plants and animals include those required for DNA condensation (histones) (Fajkus et al. 1995, Grunstein 1997), telomere regulation and structure, which bind directly to telomere DNA (e.g. telomere repeat factor 1 - TRF1) (Bianchi et al. 1997) and those which bind indirectly via protein-protein interactions (e.g. DNA repair proteins: Ku70 and Ku80) (Fisher and Zakian 2005).

The DNA component of the telomere was discovered in 1978 (Blackburn and Gall 1978), in the ciliated protozoan *Tetrahymena thermophila*, by isolating extrachromosomal ribosomal DNA (rDNA) fragments called minichromosomes, which are present in the macronucleus. Restriction endonuclease digestion confirmed the presence of a $(TTGGGG)_{20-70}$ hexamer at both ends of the minichromosomes, with an orientation of 5'-GGGGTT-3' / 3'-CCCCAA-

5' (Blackburn and Gall 1978). In many eukaryotes telomere DNA consists of a tandemly repeated minisatellite sequence varying in length between species e.g. 2-8 kb in *Arabidopsis thaliana* (Richards and Ausubel 1988, Maillet et al. 2006), 10 kb in *Solanum lycopersicum* (tomato) (Ganal et al. 1991), 20-80 kb in *S. tuberosum* (potato) (Fajkus et al. 2002) and 90-130 kb in *Nicotiana tabacum* (tobacco). In *A. thaliana* a minimal functional telomere length of 300-400 bp has been proposed based on the amount of telomere sequence present at chromosome fusion sites in mutants that have short telomeres (Heacock et al. 2004).

A common feature present across eukaryotes is a 3' single-strand overhang of the guanine-rich (G-rich) DNA strand (Henderson and Blackburn 1989, Henderson et al. 1990, Wright et al. 1997). This overhang is able to invade the the duplex telomere DNA further back along the same chromosome creating a displacement-loop (d-loop), which in effect hides the chromosome end (fig. 1.1 A)(Greider 1999, Wei and Price 2003). This results in the formation of a terminal-loop (t-loop) 'lasso' structure, which again has been reported across a wide range of eukaryotes (fig. 1.1 A) (Griffith et al. 1999, Munoz-Jordan et al. 2001, Tomaska et al. 2002, Cesare et al. 2003).

Another type of folded structure that telomeres may adopt is the G-quadruplex (G4), formed between parallel or anti-parallel runs of G-rich DNA via guanine-guanine double hydrogen bonds (fig. 1.1 B) (Zimmerman et al. 1975, Sen and Gilbert 1988, Shida et al. 1989, Kim et al. 1991). G4 DNA can be formed by four runs of $G_{n>3}$ DNA, and may be formed on a single strand or between up to four separate strands, perhaps between G-rich sequences at telomeres, in rDNA and immunoglobulin genes (Dempsey et al. 1999, Hanakahi et al. 1999, Maizels 2006). In the case of immunoglobulin genes the G4 structures may induce recombination within a specific region allowing switching of the constant region from IgM to IgA, IgG or IgE types (Yu et al. 2003). The telomere repeat of ciliates of the genus *Oxytricha* is TTTTGGGG, oligonucleotides containing multiples of this repeat can form quadruplexes *in vitro* between two self-paired oligonucleotides (Smith and Feigon 1992, Kang et al. 1992, Schultze et al. 1994). However, G4 DNA has only been localised *in vivo* in the ciliate *Stylonychia lemnae* using antibodies that specifically recognise this structure (Schaffitzel et al. 2001).

The fission yeast *Schizosaccharomyces cerevisiae* has been genetically engineered to maintain mitotically stable circular chromosomes by knocking out two telomere regulating proteins, suggesting that it is possible for linear chromosomes to re-circularise, although meiosis is severely affected (Naito et al. 1998). The low number of viable haploid spores produced is likely due to the inherent problem of meiotic crossing-over between circular chromatids.

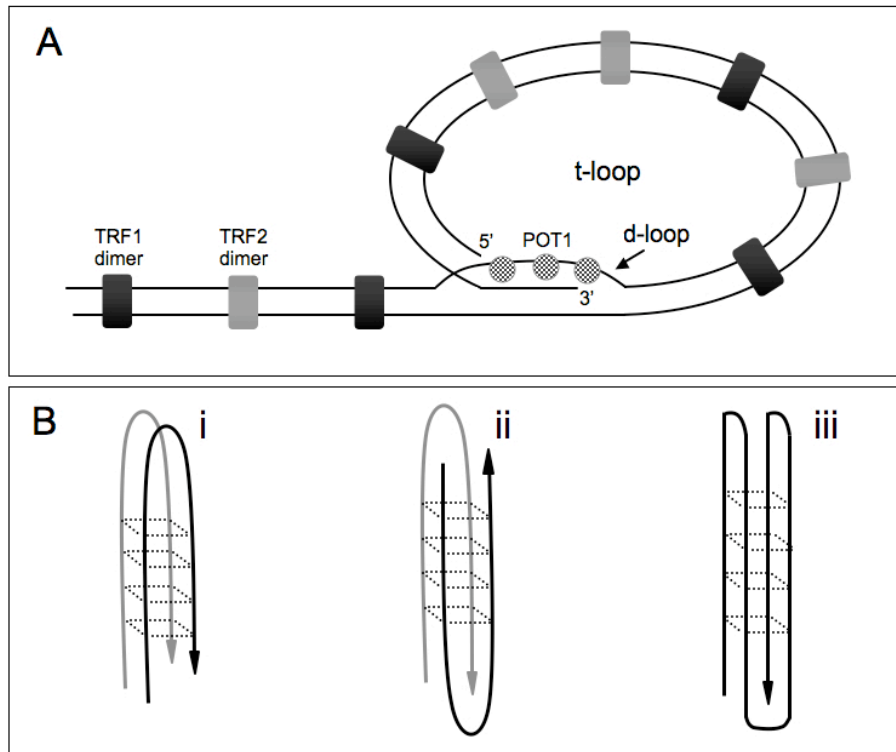

Figure 1.1: Telomere and G-Rich DNA structures
(A) A simplified telomere terminal loop structure, two dsDNA telomere binding proteins and a ssDNA telomere binding protein are indicated, note that unlike mammals, in plants multiple homologues are present for each of these proteins (Kuchar and Fajkus 2004, Schrumpfova et al. 2004, 2008). (B) G-quadruplex/G4 structures that can be formed by G-rich DNA, arrowheads indicate 5' ends of DNA strands: (i) parallel G4 structure between two self-paired DNA strands (Schultze et al. 1994); (ii) antiparallel G4 between two self-paired strands (modified from Dapic et al. 2003); (iii) parallel/antiparallel G4 structure between G-runs on the same DNA strand (modified from Dapic et al. 2003).

An unequal number of crossovers would generate concatenated chromosomes that if pulled towards opposite poles, would break during segregation. There are reports of chromosome circularisation occurring in humans, although these chromosomes are very rarely transmitted to the next generation (Jenderny et al. 1993). Having circular chromosomes would be highly disadvantageous during meiosis, explaining perhaps why circular chromosomes are so rarely seen in eukaryotes, where gamete production by meiosis is so prevalent (Ishikawa and Naito 1999). Telomeres are also implicated in the formation of the meiotic bouquet, this is assembled in many eukaryotes at the leptotene-zygotene transition, with telomeres clustering together at the nuclear envelope prior to synapsis and bivalent formation (Bass et al. 2000, Scherthan 2001, Harper et al. 2004).

1.2 Telomerase

Telomerase activity was first identified in *Tetrahymena* cell extracts (Greider and Blackburn 1985). Telomerase was subsequently found to have both protein and RNA components (Greider and Blackburn 1987), it is the RNA component that acts as a template dictating which nucleotides are added (Greider and Blackburn 1989). Mutations in the template region in the gene encoding the RNA component lead to corresponding mutations in the telomere repeat synthesised (Yu et al. 1990). The catalytic protein component is related in structure and activity to the reverse transcriptases from non-long terminal repeat (non-LTR) retrotransposons (Eickbush 1997, Nakamura and Cech 1998).

In plants active telomerase is normally present in undeveloped seeds, early embryogenesis and meristematic tissues with little to no activity in differentiated tissues such as leaves, roots and quiescent seeds (Kilian et al. 1995, Heller et al. 1996, Killan et al. 1998, Riha et al. 1998). Telomere extension is controlled at the transcriptional level (Oguchi et al. 1999) and through the accessibility of chromosome ends (Fulneckova and Fajkus 2000). Telomerase is able to extend non-specifically from the 3' ends of DNA *in vitro* using the telomeric repeat amplification protocol (TRAP) assay (Wright et al. 1997, Bednenko et al. 1997) and *in vivo* (Melek and Shippen 1996).

1.3 Evolution of Telomerase and the Telomere in the Asparagales

Allium is classified within the order Asparagales (Fay et al. 2000). The order is estimated to be 60-69 million years old (Good-Avila et al. 2006) includes several well known plant groups such as the orchids (family Orchidaceae), hyacinths (in family Hyacinthaceae) and Irises (in family Iridaceae) (Angiosperm Phylogeny Group, II 2003). Following the discovery that some plants lacked the 'plant type' TTTAGGG telomere repeat (Fuchs et al. 1995), 16 species from across the Asparagales were checked for the presence or absence of the plant type telomere (Adams et al. 2001). Adams et al. (2001) found that a number of families forming a monophyletic clade lacked the typical plant telomere. It was subsequently found that species of the genus *Aloe* had the TTAGGG minisatellite telomere (Weiss and Scherthan 2002). Later most of the Asparagales species lacking the typical plant telomere repeat were found to have the TTAGGG repeat, but the exception was *Allium* where the telomere remained unknown (fig. 1.3) (Sýkorová et al. 2003, 2006).

No RNA component of any plant telomerase has been isolated to date, but it is possible that a mutation occurred in the template region giving rise to the TTAGGG repeat (Sýkorová et al. 2003). Another possibility is that mechanism by which telomerase adds new telomere repeats to chromosome ends may have altered, leading to the loss of the first thymine (Sýkorová, Leitch and Fajkus 2006). In the green algae *Chlamydomonas reinhardtii* the telomere is made up of TTTTAGGG repeats (Petracek et al. 1990), showing another change in plant telomeres due to the number of thymidines added by telomerase.

In vitro telomere synthesis using telomerase extracts indicates that the telomerase in *Bulbine glauca* (rock lily) and *Ornithogalum virens* (star-of-Bethlehem), which synthesise the TTAGGG repeat have a low fidelity, synthesising up to a quarter of repeats erroneously (Sýkorová et al. 2003). Telomere variants are also seen in the diversity of telomere variants detected by slot blot hybridisation and fluorescence *in situ* hybridisation (FISH) to DNA of species across the Asparagales, i.e. *Ornithogalum umbellatum* has TTGGGG repeats detectable at many of its telomeres (Sýkorová et al. 2003). Work addressing the telomere proteins of Asparagales plants with the derived TTAGGG telomeres indicates that G-rich single strand DNA (ssDNA) binding proteins are able to bind both the TTTAGGG repeat and the TTAGGG repeat (Rotkova et al. 2004, 2007). There is a 25 kDa G-rich ssDNA binding protein in *Muscari armeniacum* (grape hyacinth) and *Scilla peruviana* (Portuguese squill) (both in family Hyacinthaceae (Angiosperm Phylogeny Group II 2003)) which only

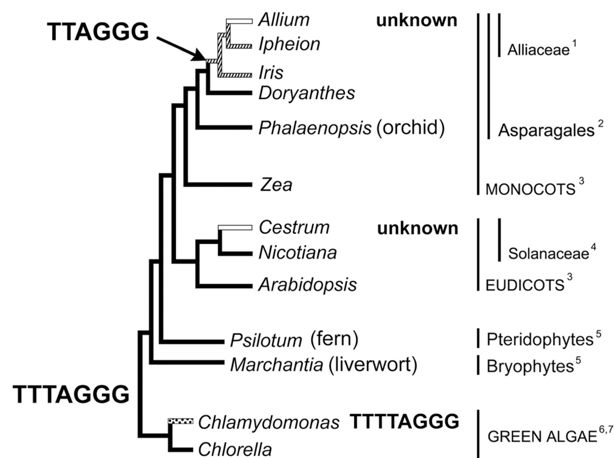

Figure 1.2: Changes in plant telomere sequences
Cladogram based on published phylogenetic relationships (Angiosperm Phylogeny Group II 2003) summarising the known transition points from the TTTAGGG telomere repeat (solid branches), to the TTAGGG telomere repeat (hashed branches) or to unknown telomere repeats (unfilled branches). Reproduced from Sýkorová et al. (2006). Superscript numbers refer to the following papers: 1, Sýkorová et al. (2006); 2, Sýkorová et al. (2003); 3, Richards and Ausubel (1988); 4, Sýkorová et al. (2003); 5, Suzuki (2004); 6, Petracek et al. (1990); 7, Higashiyama et al. (1995).

binds to the ancestral TTTAGGG telomere repeat and not the derived TTAGGG repeat (Rotkova et al. 2004, 2007). From these observations it appears that the low fidelity of the telomerase synthesising the derived TTAGGG repeat may have permitted a change in telomere repeat, through availability of at least a few ancestral TTTAGGG binding sites (Sýkorová et al. 2003) and binding tolerance of telomere proteins (Rotkova et al. 2004, 2007). Other plant groups have been found to have variant repeats within the telomeric arrays. *Pisum sativum* (pea) has both TTAGGG and TTTAGG repeats present, but rather than being dispersed as is probably the case in the Asparagales, sequencing has shown that several of these variants occur in tandem (Macas et al. 2007). In rice, a thorough analysis of telomeres has also showed that telomeric repeat variants are often found either adjacent to each other or close together, particularly towards the centromere end of the telomeres (Mizuno et al. 2006, 2008a,b). As the distal ends of telomeres have less variant repeats, the authors of this work favoured the idea that replication slippage or unequal recombination based processes have been primarily responsible for localised expansion of atypical repeats, rather than repeated telomerase errors (Mizuno et al. 2008a).

It has been established that a switch from the TTAGGG telomere to an unknown sequence must have occurred in an ancestor to the genus *Allium* (Sýkorová et al. 2006). Telomerase activity *in vitro* (TRAP assay) has not been detected using *Allium cepa* protein

extracts, suggesting that the active gene is no longer present (Sýkorová et al. 2003). PCR amplification using Asparagales consensus primers for different parts of telomerase have also failed to produce any products (E. Sýkorová pers. comm.). It is therefore likely that the telomerase-based telomere elongation mechanisms may also be missing in *Allium.* Several repetitive DNA sequences have been mapped to the terminal regions of *A. cepa* chromosomes. These are: a tandem repeat called *Allium cepa* satellite (ACSAT) (Barnes et al. 1985, Pich, Fritsch and Schubert 1996), Ty1/*copia* retrotransposons (Pearce et al. 1996, Cunado et al. 2001), an *En*/*Spm*-like transposon (Pich and Schubert 1998) and/or 35S rDNA (Pich and Schubert 1998, Pich, Fuchs and Schubert 1996). But as yet none of these, nor any other DNA sequence has been confirmed as the sequence(s) comprising the telomere in *Allium.*

An independent loss of the TTTAGGG telomere has been found in a clade of three genera, *Cestrum, Vestia* and *Sessea* (family Solanaceae) (Sýkorová et al. 2003). It is not known what sequence has replaced the TTTAGGG telomere repeat in this group, although in *Cestrum* interstitial telomeric repeats are still present in the genome Sýkorová et al. (2003).

1.4 Non-Telomerase Maintained Telomeres

In the absence of telomerase, it is possible to elongate the terminal DNA of a chromosome by using DNA recombination based mechanisms (Wang and Zakian 1990a, Teng and Zakian 1999, Dunham et al. 2000, McEachern and Iyer 2001). This has been adopted as a mechanism for telomere extension in some *Anopheles*, *Chironomus* and *Drosophila* species probably following the loss of telomerase in an ancestor of the *Diptera.* Recombination-based maintenance of telomeres has been termed alternative lengthening of telomeres (ALT) in humans, where it can arise spontaneously in carcinomas allowing them to become immortalised (Bryan et al. 1995, 1997, Hakin-Smith et al. 2003, Londono-Vallejo et al. 2004). Recently there has been evidence for ALT occurring in *A. thaliana* telomerase null mutants, leading to the view that ALT may be widespread and could be occuring occur alongside telomerase activity (Ruckova et al. 2008).

Another mechanism for elongating telomeres is based on repeated transposition to the terminal regions by retrotransposons (e.g. *Drosophila*). ALT and retrotransposon-based telomere maintenance may provide hints towards ancestral mechanisms of telomere maintenance and shed light on how telomerase arose in early eukaryotes (Nosek et al. 2006).

Both ALT and retrotransposon-based mechanisms have been suggested as candidates for the maintenance of telomeres in *Allium* (Pich, Fuchs and Schubert 1996, Fajkus et al. 2005).

1.4.1 Alternative Lengthening of Telomeres (ALT)

1.4.1.1 ALT mechanisms

ALT is used to describe lengthening of telomeres via DNA recombination (Bryan et al. 1995). These mechanisms fall into two broad groups (i) reciprocal and (ii) non-reciprocal DNA recombination (Szostak and Wu 1980, Paques and Haber 1999).

Reciprocal recombination (responsible for meiotic crossovers) leads to no net increase of telomere repeats but can prolong the replicative lifespan of cells by swapping DNA unequally between chromosome ends. This has been shown to occur in human and murine cancer cells (Bailey et al. 2004, Londono-Vallejo et al. 2004, Wang et al. 2005), indeed in human cancer cells, sister chromatid exchanges (SCEs) have been shown to be particularly concentrated at the telomeres.

Non-reciprocal DNA recombination or gene conversion may occur in several ways, and in some cases can lead to net increases in telomere length (Wang and Zakian 1990b). At the telomere there are two models which could explain telomere extension. Synthesis dependent strand annealing (SDSA) involves a 3' ssDNA end invading similar duplex DNA on the same or another chromatid/chromosome and extension of the 3' end by DNA polymerase. The extended strand is eventually displaced from the duplex and lagging strand synthesis can then occur on the other strand (Nasmyth 1982, Nassif et al. 1994). Telomerase negative *Saccharomyces cerevisiae* cells that undergo this type of telomere recombination have much longer telomeres and greater length heterogeneity (0.3-12 kb) compared to wild type cells (350 bp $\pm$ 75 bp) (Teng and Zakian 1999). Another mechanism is break-induced replication (BIR) which also involves ssDNA invading another duplex and using it as a template for DNA synthesis, but this cascade is initiated by a double strand break, with the distal telomere DNA being lost (McEachern and Haber 2006). It is likely that at least two types of BIR can occur in *Saccharomyces* yeasts enabling telomere lengthening independently of telomerase (Bosco and Haber 1998, Nakamura et al. 1998). Extrachromosomal circles (ECCs) of telomeric DNA may provide another template for ALT mechanisms to elongate from (Natarajan and McEachern 2002, McEachern and Haber 2006), additionally ECCs may replicate by rolling-circle replication/amplification (RCR) (Nosek et al. 2005). ECCs may arise by resolution (endonuclease digestion) of homologous recombination inter-

mediates between chromosome ends (Compton et al. 2007), ECCs can also be released from chromosome ends by endonuclease digestion of the T-loop (Lin et al. 2005, Groff-Vindman et al. 2005, Nabetani and Ishikawa 2009).

The widespread occurrence of ALT mechanisms in telomerase knock-out cell lines and organisms is not surprising considering that proteins responsible are involved in normal DNA repair, replication and recombination. In insects there is substantial data showing that ALT has been adopted as a principle mode of telomere extension (Biessmann and Mason 2003).

1.4.1.2 Long-Term ALT in Insects

The likely ancestral telomere sequence for phylum Arthopoda is TTAGG and appears to be present in the majority of species in this phylum indicating that in many lineages it has remained unchanged for approximately 545 million years (Vitkova et al. 2005). However in the class Insecta there hava been several independent losses of the TTAGG repeat including in order Coleoptera (beetles), but the replacement telomere sequence is unknown (Frydrychova and Marec 2002). To date no species in a clade containing three insect orders, Diptera (true flies), Mecoptera (includes scorpion flies) and Siphonaptera (fleas) (Whiting et al. 1997) have shown positive hybridisation to TTAGG probes (Frydrychova et al. 2004). Within Diptera the telomeres have been partly or fully characterised for some species of *Chironomus* and a species of *Anopheles*. In these genera telomeres appear to be made up of telomeric satellite repeats. *Drosophila* species (also dipterans) are thought to have telomeres consisting of retrotransposons (Traverse and Pardue 1988, Levis et al. 1993, Abad et al. 2004).

The *Anopheles gambiae* telomere on the long arm of chromosome II (2L) comprises a unique 820 bp satellite repeat, which is maintained by recombination (Biessmann et al. 1996, Roth et al. 1997). Several different fragments of middle repetitive sequences, including a 2.3 kb 28S rDNA fragment, were found adjacent to the 820 bp telomere repeat in different individuals (Biessmann et al. 1998). It is thought that the telomere 'picked up' these other sequences via gene conversion events initiated in regions of sequence similarity (Biessmann et al. 1998). Non-reciprocal transfer of a marker integrated into the end of chromosome 2L supports the idea that this mechanism occurs normally to extend shortened telomeres (Roth et al. 1997). The amount of replicative loss at the telomere on chromosome arm 2L was measured over 44 generations, a steady loss of DNA was shown (55 bp per generation) (Walter et al. 2001). During this time no increases in telomere length were seen in any individual (Walter et al. 2001), perhaps because the telomeres did not drop below a critical

threshold to trigger telomere lengthening.

The terminal DNA sequences of several species of *Chironomus* (chironomid flies) have been analysed, namely in *C. pallidivitatus*, *C. tentans* and *C. thummi* (Saiga and Edstrom 1985, Nielsen et al. 1990, Zhang et al. 1994, Lopez et al. 1996). Most of the telomeres in these species consist of ~350 bp repeats, which can be subdivided into subfamilies based on minor sequence variation (Nielsen and Edstrom 1993). In *C. pallidivitatus*, repeat subfamilies can occur in blocks and interdispersed with one another, the latter being explained as a result of multiple short gene conversion events (Zhang et al. 1994). The 350 bp repeats show G-rich bias on one strand and are probably derived from smaller subrepeated sequences (including telomere repeats) that have degenerated (Nielsen and Edstrom 1993). The G-rich strand as with minisatellite telomeres is orientated in a 5' to 3' direction (Kamnert et al. 1997) with telomeric arrays measuring approximately 50-200 kb in length (Zhang et al. 1994). Comparative analysis between *C. pallidivitatus* and *C. thummi* stocks shows that efficient gene conversion of telomeric repeats has lead to rapid fixation of different mutations between species (Kamnert et al. 1998).

1.4.2 Retrotransposons as Telomeres

In the genus *Drosophila* the telomeres consist of arrays of non-LTR retrotransposons. In *D. melanogaster* these are named *HeT-A*, *TART* and *TAHRE* (Traverse and Pardue 1988, Levis et al. 1993, Abad et al. 2004). These elements are LINEs (long interspersed nuclear elements), which typically have two ORFs, however in the case of *HeT-A* the ORF2 is missing and so the reverse transcriptase, endonuclease and RNase activities which are normally encoded by this region must be provided by other autonomous elements, if they are to transpose (Biessmann et al. 1994). The presence of LINEs at the telomere in species across the genus suggests telomeric LINEs were present in the common ancestor of the genus (Casacubierta and Pardue 2003a,b). Furthermore, 5' truncated LINEs, analogous to *Het-A* in *D. melanogaster* with only ORF1 encoding GAG (group specific antigen), have arisen independently multiple times in different *Drosophila* lineages (Villasante et al. 2007). Where chromosomes have been experimentally broken in some stocks of *D. melanogaster* multiple *Het-A* elements have transposed to the chromosome end in a process analogous to *de novo* telomerase extension (Biessmann et al. 1990, 1994). Despite the great DNA structural differences between telomeres composed of minisatellites and those of LINEs, several proteins show a conservation of function including ATM, Ku70 and Ku80 homologues, which are important for telomere regulation and function (Oikemus et al. 2004, Melnikova et al. 2005).

Another distinct group of telomeric LINEs has been characterised in the alga *Chlorella vulgaris*, this family is named Zepp and as in the case of *Drosophila* elements are arranged in arrays at the chromosome termini (Higashiyama et al. 1997, Noutoshi et al. 1998). Analysis of different minichromosomes released following cell irradiation shows that Zepp elements actively transpose to chromosome ends by integrating into other Zepp elements (Yamamoto et al. 2000, 2003). Unlike in *Drosophila*, TTTAGGG repeats are also often present at chromosome ends suggesting that telomerase is still active (Higashiyama et al. 1997, Yamamoto et al. 2000). *C. vulgaris* telomeres appear to represent a transition point between telomerase maintained termini and telomeres maintained solely via targeted LINE transposition (Higashiyama et al. 1997, Yamamoto et al. 2003).

Bombyx mori (silkworm; order Lepidoptera) has two families of LINEs which are associated with telomeres. The families are named SART1 and TRAS1 both of which specifically target and insert within telomere repeats (Okazaki et al. 1995, Takahashi et al. 1997). SART1 inserts on the 5'-TT*AGG-3' strand (asterisk indicates insertion site) and TRAS1 inserts on the opposite 3'-AA*CCT-5' strand (Okazaki et al. 1995, Takahashi et al. 1997). Within the ORF2 between the endonuclease and reverse transcriptase domains is a region with similarity to the myb domain which is present in some telomere DNA binding proteins, suggesting that the protein encoded may allow elements to target telomeres (Kubo et al. 2001).

1.5 Isolation of Telomeric Sequences

The telomere sequence of a wide spectrum of eukaryotic organisms is now known or have been predicted due to the high conservation of telomere repeats. However in angiosperms and insects, surprises have emerged which question the predictive power of sparse sampling repeats across such large radiations of species. In cases where telomeres have been isolated without *a priori* knowledge of the repeat, the organisms examined have a relatively small genome. In *Anopheles gambiae* a chance integration of a transgenic construct allowed the telomere on one chromosome arm to be discovered (Biessmann et al. 1996). In *A. thaliana* an elegant end-sequence enriched cloning method was used to isolate the telomeric repeats, even with this enrichment method only 5 of 765 clones contained telomeric DNA (Richards and Ausubel 1988).

End-cloning based strategies, where adapters are ligated onto chromosome termini, are highly dependent on genomic DNA being intact, as interstitial breaks compete with chro-

mosome ends for adaptors. This may have contributed to Biessmann et al. (2000) wrongly identifing the *Drosohphila virilis* satellite repeat as telomeric by using this type of an end-cloning method. Subsequent sequencing showed that this repeat was in fact subtelomeric and LINEs form the telomeres as in other *Drosophila* species (Villasante et al. 2007).

1.6 Aims and Overview

This thesis set out to determine the telomere structure in *Allium*. Whilst a definite conclusion was not found, a potential telomere repeat has been partially characterised for *A. cepa*. A considerable effort was focussed on rDNA, which was found to be in the terminal domain more frequently than expected. During the course of this work a novel relic retrotransposon was uncovered in the rDNA of *A. cernuum*. These analyses lead to a greater understanding of *Allium* genomic structure, but did not lead to greater insight into telomere biology in the genus.

Chapter 2

Materials and Methods

2.1 Plant materials

2.1.1 Living Plant Accessions (QMUL)

Living material was grown from seed in a temperature controlled greenhouse in the School of Biological and Chemical Sciences, Queen Mary.

Species	Supplier	Voucher
Allium cernuum	Chiltern Seeds, Cumbria, UK	Chester 0701 QMUL
Allium cernuum	Pottertons Nursery, Nettleton, UK	Chester 0702 QMUL
Allium cernuum	Rose Cottage Plants, Essex, UK	Lim 04-08 QMUL
Allium cepa cv. Ailsa Craig	Suttons Seeds, Devon, UK	QMUL
Allium cepa cv. Bedfordshire Champion	Suttons Seeds, Devon, UK	QMUL

Table 2.1: Plant materials

2.1.2 DNA from Living Collections

All other plant material was from either Botanical Garden of Osnabruck University, Osnabruck, Germany (provided by Dr. Nikolai Friesen) or the IPK, Gatersleben, Germany, in each case vouchers are deposited in their respective herbaria. For identification of IPK accessions, TAX accession numbers are used as in Sýkorová et al. (2006).

2.2 DNA Extraction and Purification

2.2.1 Genomic DNA Extraction

DNA was extracted following Dellaporta et al. (1983) with minor modifications. Fresh leaves were collected and washed with distilled water and dried, and 1 g was weighed for extraction. For each gram of leaf material 5 ml of extraction buffer (100 mM Tris pH 8.0, 50mM EDTA, 500 mM NaCl, 0.1% (v/v) β-mercaptoethanol) was pre-chilled on ice. Using a pre-chilled pestle and mortar, dry leaf material was ground to a fine powder together with liquid nitrogen, ensuring the material was maintained in a frozen state. The powder was mixed thoroughly by vortexing with the extraction buffer, to this 0.5 ml of 20% (w/v) sodium dodecyl sulphate was added and mixed well to lyse the cells. Falcon tubes containing the lysed plant cells were left on ice while other extractions were prepared. Samples were then incubated at 65°C for 30-40 min, gently inverting intermittently. After this 1.7 ml of high salt solution (3 M potassium acetate, pH 5.5) was added, and the solution was left on ice for 30 min. The samples were then centrifuged at 12,000 rpm for 14 min, the supernatant was then transferred to an Erlenmeyer flask, to which an equal volume of isopropanol was added. The solution was mixed gently and allowed to stand for approximately 10 min, a glass rod was then used to spool out DNA from the interface and collected; this was repeated several times to maximise the yield. The harvested DNA was then washed in 70% ethanol several times before air-drying. Dried DNA was resuspended in 1 ml TE buffer (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0) and then further purified by removing RNA and protein.

2.2.2 RNA and Protein Digestion

Extracted genomic DNA underwent RNA digestion by adding 110 μl of 10x RNAse buffer (100 mM Tris pH 8.0, 500 mM NaCl, 10 mM dithiothreitol), 200 μg RNAse (Sigma) and incubating at 37°C for 6 hrs. The reaction was then terminated by adding 6 μl 1 M EDTA (pH 8.0). Proteins were then digested by adding 30 μl 20% (w/v) sodium dodecyl sulphate and 3.5 μl 20 mg/ml proteinase K, this was left at 50°C overnight, the following day the sample was purified by phenol:chloroform extraction.

2.2.3 Phenol:Chloroform Extraction

DNA in aqueous solution was mixed with an equal volume of phenol:chloroform (Sigma) in a 2 ml eppendorf tube. The solution was then centrifuged at 13,500 rpm in a microcentrifuge

for 10 min. The upper layer was then transferred to a 2 ml eppendorf and mixed with an equal volume of -20°C chloroform. The solution was then centrifuged at 13,500 rpm in a microcentrifuge for 10 min. The upper layer was then transferred to a 2 ml eppendorf tube and precipitated.

2.2.4 DNA Precipitation

To precipitate DNA from solution it was mixed with $1/10^{th}$ the original volume of 1 M sodium acetate (pH 5.2) and 2.5 volumes of -20°C 100% ethanol. Samples were left at -20°C overnight. The following day these were then centrifuged for 30 minutes at 13,000 rpm in a microcentrifuge. Following removal of the 100% ethanol the pellet was carefully washed with -20°C 70% ethanol, the pellet was then air-dried. Alternatively, to avoid the use of sodium acetate for salt sensitive applications, DNA was precipitated from aqueous solution with 1 volume of -20°C 100% isopropanol.

2.3 Chromosome Microdissection

Terminal regions of chromosomes, were collected from metaphase squashes prepared as in section 2.8.1, with the root tip fixation step reduced to 5 minutes. Microscope slides were defrosted (from -20° C storage) and then rehydrated by adding one drop of sterile distilled water. The microdissection needle was prepared from a 1.0 mm diameter glass rod that was initially cut into two by simultaneously pulling each end and heating the middle of the rod (Bachofer). The newly cut ends were then melted to the correct diameter using an MF-9 Narashige-2 instrument (Greenvale, NY), which consisted of a heating element mounted below an inverted microscope to allow precise melting of the glass tip. Chromosomes were dissected/collected using an electronic micromanipulator (Eppendorf 5170) under a 32x objective. Chromosome tips were transferred and stored together on ice in a sterile 0.5 ml tube.

2.4 Polymerase Chain Reaction

2.4.1 Standard PCR

For amplifying DNA fragments less than 2 kb in length PCR was carried out with BIOTAQ Polymerase (Bioline, UK) with supplied 10x NH_4 buffer and 50 mM $MgCl_2$ solutions. A typical PCR for a 1 kb fragment in a 50 μl reaction contained the following: 1.5 mM $MgCl_2$,

1x NH_4 buffer (1x NH_4 buffer: 16 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 67 mM Tris-HCl, 0.01% Tween-20), 100 nM primer, 100 nM forward primer, 200 μM of each dNTP, 3% DMSO (v/v), 0.2-1 ng of low complexity DNA template (e.g. plasmid DNA) or 10-200 ng of high complexity DNA template (e.g. genomic DNA), 2 U BIOTAQ DNA polymerase (Bioline).

2.4.2 PCR Labelling for FISH and EDF-FISH

Biotin or digoxigenin conjugated to dUTP were incorporated into DNA amplicons for generating FISH and EDF-FISH probes. This method is best used for fragments less than 600 bp in length to avoid non-specific hybridisation to preparations i.e. causing 'background' clumping of probe DNA. In all cases plasmid DNA was used as a template for PCR labelling in order to avoid amplifiying and labelling unwanted DNA sequences. The following method is based on that described by Leitch et al. (1994), it was found to work well for a variety of cloned sequences.

A 50 μl reaction contained the following: 1.5 mM $MgCl_2$, 1x NH_4 buffer (1x NH_4 buffer: 16 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 67 mM Tris-HCl, 0.01% Tween-20), 400 nM primer, 400 nM forward primer, 100 μM of each unmodified dNTP, 50 μM of either digoxigenin-11-dUTP or biotin-16-dUTP, 3% DMSO (v/v), 0.2-1 ng of plasmid DNA, 2.5 U BIOTAQ DNA polymerase (Bioline).

2.4.3 Long Range PCR

In order to amplify DNA sequences greater than 2 kb in length a high fidelity DNA polymerase was used with proof-reading capability. Either Expand High Fidelity Polymerase (Roche) or DNA Phusion Polymerase (New England Biolabs) was used. Amplifications were carried out following the manufacturers instructions using a DYAD DNA Engine (MJ Research).

2.4.4 DNA Gel Electrophoresis

PCR products were checked on 0.8-2% agarose gels containing 1x TAE (40 mM Tris-acetate, 1 mM EDTA) depending on expected size range and 0.5 μg/ml ethidium bromide. Typically 5 μl of PCR product was combined with 1 μl of 6x loading dye (Fermantas) Agarose gels were viewed and photographed with UV illumination using a Chemigenius Bioimaging System (Syngene, UK). PCR products were purified either directly from solution or were cut out, following separation on an agarose gel, using a scalpel and extracted using a QIAquick©

PCR Purification Kit (Qiagen).

2.4.5 Degenerate Oligonucleotide Primed-PCR (DOP-PCR)

DOP-PCR was carried out in two stages (primary & secondary) using DOP-PCR kit reagents (Roche Diagnostics, GmbH., Germany), with minor modifications to the protocols used in terms of primers used (Telenius et al. 1992). Primary PCR primers were designed and tested by E. Sýkorová to not amplify 35S ribosomal DNA.

Primary PCRs contained the following: 25 μl of 1x mastermix: 1.25U Taq. DNA polymerase (0.005% V/W); 0.2mM dNTPs; 50mM KCl; 1.5mM MgCl2; 10mM Tris-HCl), 2.5µl of degenerate oligonucleotide primer (0.2 μM) (DOPG1: 5'- CTAATACGACTCACTATAGGG NNNNNNTTAGG-3'), the final volume was made up to 50 μl with sterile water. Reagents were added directly to the PCR tube containing the microdissected DNA. Genomic DNA was diluted to 5 ng for use as a positive control, a negative control containing no template was included.

Reactions were carried out in sterile 0.5 ml tubes and covered with 2 drops of mineral oil. The following steps were used in the thermocycler: 94° C for 0.5 min; 57° C for 1.2 min; 72° C for 1.2 min. This was followed by 25 (higher stringency) cycles at 94° C for 20 sec; 60° C for 1 min; 72°C for 25 sec followed by a final extension of 72° C for 5 min. Secondary PCRs were carried out with 25μl of 1x mastermix, 2.5 μl of 0.2 μM degenerate oligonucleotide primer T7: 5'-TAATACGACTCACTATAGGG-3', 5 μl of primary PCR product was added as the template.

2.5 Southern Hybridisation

2.5.1 DNA Restriction Digests

Genomic DNA was digested overnight using restriction endonucleases and supplied buffers (New England Biolabs) following the manufacturers protocol. The digestion products were separated on a 0.8% agarose gel.

2.5.2 Southern Transfer

Agarose gels were alkali blotted onto Hybond N+ nylon membrane (Amersham Pharmacia Biotech) following the protocol of Sambrook and Russell (2001). Briefly, the gel was depurinated in 0.25 M HCl for 30 min, drained and rinsed in distilled water. The gel was

then neutralised in transfer buffer (1M NaCl, 0.4M NaOH) for 30 min. The blotting apparatus was assembled as follows: three layers of Whatman© 3MM paper were cut to act as a wick between a reservoir of transfer buffer and the blotting platform. The gel was then placed face down on the wick and covered with a nylon membrane ~3 mm larger than the gel, ensuring that no air bubbles were trapped between layers. To provide a good contact between the gel and cellulose wadding three sheets of pre-wetted and then three sheets of dry Whatman© 3MM paper cut to the same size of the membrane were placed over the nylon membrane. This arrangement allowed the transfer buffer to flow from the reservoir through the gel, nylon membrane and Whatman© paper into Cellosene© cellulose wadding (Thomas and Green Ltd, Hereford, UK), close contact between layers was maintained by placing a glass plate with weights over the blotting apparatus. Following DNA transfer for at least 24 hrs, the nylon membrane was placed in neutralisation buffer (1 M NaCl, 0.5 M Tris pH 7.2) for 15 minutes, the membrane was then dried in a oven for 1 hour.

2.5.3 Radioactive DNA Southern Hybridisation

Specific probes for Southern hybridisation were amplified by PCR from clones using universal M13 primers, otherwise genomic DNA was used for genomic hybridisations. The probes were radioactively labelled using DecaLabel™ DNA Labeling Kit (MBI Fermentas) or Ready-To-Go™ DNA Labelling Beads (Amersham Biosciences).

The procedure for the Ready-To-Go™ DNA labelled beads (Amersham Biosciences, Buckinghamshire, UK) is outlined briefly. DNA (50ng) was heated to 100 $^\circ$C for 15 min and then snap chilled on ice to produce single stranded DNA fragments. The labelling mixture was made following the manufacturers protocol; 5 µl α-P^{32}-dCTP (50 μCi activity), 50 ng genomic DNA, double distilled water to a final volume of 50µl and a reaction bead (containing: Klenow Fragment, dATP, dTTP, dGTP and priming oligomers). The reaction was allowed to proceed for 60 min, the labelling products were then denatured (94 $^\circ$C) for 10 min and then snap chilled on ice.

Hybridisation was carried out overnight in hybridisation buffer (0.25 M sodium phosphate buffer (pH 7.0), 7% (w/v) SDS) at an appropriate temperature (typically between 55-65°C). The membranes were washed at high stringency (0.2x SSC, 0.1% SDS) or low stringency (2x SSC, 0.1% SDS) using the same temperature as used for hybridisation and exposed to a phosphoimager screen (Amersham Biosciences). Screens were analysed on a Typhoon™ Variable Mode Imager (Amersham Biosciences).

2.6 Cloning and Colony Screening

2.6.1 Ligation and Transformation

DNA amplicons of interest were generally ligated into A-overhang accepting vectors; either pCR2.1-TOPO or pCR4-TOPO (Invitrogen) were used according to manufacturer's protocol. Blunt-ended DNA fragments were ligated into pZErO-2 (Invitrogen) plasmid that had been digested with *Eco*RV (New England Biolabs), for inserts with any other overhang either pZErO-2 or pBluescript KS+ (Stratagene) was digested with a compatible enzyme. For vectors with the multiple cloning site embedded within the β-galactosidase gene (e.g. pCR2.1-TOPO, pBluescript KS+), blue/white selection of colonies was carried out using 40 mg/ml X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) substrate which was added to solidified LB plates prior to spreading *E. coli* cells.

2.6.2 Minipreparation Plasmid DNA Extraction

Colonies for plasmid extraction were grown overnight in 2-5 ml of liquid Luria-Bertani (LB) medium (tryptone 10 g, yeast extract 5 g, NaCl 10 g, agar 15 g, made up to 1 litre with distilled water, pH 7.0) supplemented with either 100 μg/μl ampicillin or 50 μg/μl kanamycin at 37 $^\circ$C on a orbital shaker at 220 rpm. Following overnight incubation, cells were harvested by centrifugation in a microcentrifuge for 1 min at 13,200 rpm, discarding the supernatant. The Mini-Prep Extraction Kit (Qiagen GmbH, Germany) was then used to extract plasmid DNA from the resulting pellet according to the manufacturer's protocol.

2.6.3 Glycerol Stocks

In some cases where long term storage of *E. coli* cells was required, following growth overnight in LB medium (as above) an equal volume of sterile glycerol was added, vortexed briefly and frozen in liquid nitrogen before being stored at -70 $^\circ$C.

2.6.4 Plasmid Library Preparation

Plasmid libraries were prepared following the protocols in Sambrook and Russell (2001). Individual colonies were transferred with sterile cocktail sticks firstly to positively charged 88 mm diameter nylon filters (Electran, BDH Laboratory Supplies, UK), mounted over solid LB medium plates supplemented with 50 μg/μl ampicillin and secondly to a master plate of solid LB medium without a nylon filter. The colonies were placed with identical

arrangements so as to refer between the blotting test plate and the master plate. The test and master plates were then incubated at 37 °C for 10 hours, after this master plates were stored at 4 °C. Nylon filters were removed from test plates and screened as described below.

2.6.5 Plasmid Library Screening

Plasmid libraries were screened following the protocols in Sambrook and Russell (2001). DNA was released from the *E. coli* and bound to the filters by placing them sequentially onto Whatman 3MM paper saturated in the following solutions. (1) 10% (w/v) SDS (sodium dodecyl sulfate) for 3 min. (2) Denaturing solution (1.5 M NaCl; 0.5 M NaOH) for 5 min. (3) Neutralising solution (0.5 M Tris-Cl; 1.5 M NaCl pH 7.4) for 5 min. (4) 2x SSPE (20x saline sodium phosphate EDTA: 3 M sodium chloride, 0.2 M sodium hydrogen phosphate, 0.02 M EDTA, pH 7.4) for 5 min. Filters were then dried and on a sheet of Whatman 3MM (VWR, UK) paper at room temperature for 30 min.

Cross-linking was done by exposing the filters to UV light for 5 min. DNA Filters were immersed in 2x SSC until saturated, they were then moved to a dish on a rotating platform in hybridisation buffer (0.25 M sodium phosphate buffer (pH 7.0), 7% (w/v) SDS) for 30 min at 50 °C. Cellular debris still remaining at this stage was completely removed with Kimwipes soaked in 6x SSPE. Filters were transferred to a hybridisation flask containing 150 ml prehybridisation solution for 2 hours at 68 °C. The filters were then used directly for radioactive DNA Southern hybridisation described in section 2.5.3.

2.7 DNA Sequencing and Analysis

2.7.1 DNA Sequencing

Sequencing was carried out commercially by sending 1 μg of dried plasmid DNA to Cogenics (UK) or MWG (Germany). Electropherograms were trimmed to remove vector sequences, assembled and checked for errors in Aligner version 2.0.4 (CodonCode, USA).

2.7.2 DNA Sequence Alignment

Sequences were exported from Aligner in FASTA format and imported into Geneious Pro version 3.6 (Drummond et al. 2007) in which sequences were aligned by implementing CUSTALW version 1.83 (Thompson et al. 1994). For putative protein coding regions DNA sequences were inspected in MacClade version 4.08 (Maddison and Maddison 2000) using

the translated view in order to establish the boundaries of putative coding regions.

2.7.3 Phylogenetic Analysis

The neighbour-joining (NJ) (Saitou and Nei 1987) method was used for rapidly approximating phylogenetic relationships between DNA sequences using the HKY substitution model (Hasegawa et al. 1985) implemented in the program Geneious. Another approach was to use maximum-likelihood (ML) with the program PHYML (Guindon and Gascuel 2003).

2.7.4 Median Joining Networks

To examine the relationships between DNA sequences showing a low level of divergence the median joining (MJ) method was used. MJ networks were constructed using the program Network version 4.5.0 (Fluxus-Engineering.com). Aligned DNA sequences were analysed using the median-joining (MJ) algorithm which allows for multistate characters (Bandelt et al. 1999). MJ networks were initially made with default settings for epsilon (value: 0) and all characters were equally weighted (value: 10).

2.7.5 Basic Local Alignment Search Tool (BLAST)

DNA sequences were checked against similar sequences on the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) nucleotide and protein databases using BLAST (Altschul et al. 1990) via the online interface.

2.7.6 Sequence Randomisation

DNA sequences were randomised using the web-based EMBOSS (Rice et al. 2000) program "shuffleseq" hosted on the Pasteur Institute website (http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py). Sequence randomisation was used to assess the stability of secondary DNA structures for different arrangements. The program takes a DNA sequence and randomises the order of the nucleotides, but retains the total length and frequencies of nucleotides. This program was used to generate 100 randomised sequences for cer129, cer352, cer330, cer540 and cer703.

2.7.7 Secondary Structure Prediction

In order to predict possible DNA secondary structures, sequences were submitted via the web-based program mfold on the Rensselaer bioinformatics web server (http://dinamelt.bioin

fo.rpi.edu/quikfold.php) (Zuker 2003).

2.8 Fluorescent *In Situ* Hybridisation

2.8.1 Chromosome Preparation

Root tips were collected from plants growing in a temperate controlled greenhouse. Pre-treatment to accumulate metaphase nuclei was carried out by imbibing in saturated aqueous solution of hexachlorocyclohexane (Lindane; Sigma-Aldrich) for 3-4 h at room temperature. Root tips were then fixed in modified Carnoy's solution (1:3 glacial acetic acid:absolute alcohol) for at least 24 hrs, and then stored in 70% ethanol at -20 $^\circ$C. For squashing, root tips were equilibrated in 1x citric buffer (10x citric buffer: 10 mM citric acid, 60 mM sodium citrate pH 4.8) three times for 5 min. Root tips were trimmed to around 10 mm in length and incubated in 500 μl of enzyme solution (0.3% (w/v) driselase (Sigma), 0.3% (w/v) cellulase R-10 (Yakult Hinsha Co. Ltd, Japan), 0.3% (w/v) pectolyase Y-23 (MP Biomedicals, Solon, Ohio, USA) suspended in 1x enzyme solution) for approximately 25 min. Root tips were then dissected under a binocular microscope to free the meristematic tissue, this was pipetted onto a chromic acid washed slide. With the addition of a drop of 45% acetic acid the tissue was macerated with a fine needle until fully dispersed and then squashed under a circular glass cover slip. Nuclei were inspected under a phase contrast microscope for cytoplasm, which if remaining was removed by the addition of another drop of 45% acetic acid and re-squashing. Slides were cooled on dry ice or in liquid nitrogen allowing the coverslip to be flicked off with a scalpel. Slides were then stored at -20 $^\circ$C.

Slides of root tip squashes were placed in a 37 $^\circ$C drying oven overnight. Following overnight drying 100 μl RNase A [200 μg/ml] was added, covered by a 22 x 22 mm piece of Parafilm© and left for 1 hr at 37 $^\circ$C in a humid chamber. Slides were then washed in 2x SSC (20x SSC: 3 M NaCl, 0.3 M sodium citrate) for 5 min, three times ensuring to remove coverslip. Slides were then incubated in 0.01 M HCl, and then were left on their side to drain. 100 μl of pepsin [1 μg/ml] was added, slides were covered with Parafilm© and left to incubate for 6 min at 37 $^\circ$C in a humid chamber. Slides were washed in distilled water for 1 min and then transferred to 2x SSC twice for 5 min. Paraformaldehyde was prepared in a fume hood by mixing 2 g in 40 ml H_2O, heated to 65-70 $^\circ$C and cleared with 10 ml 0.1 M NaOH. Slides were then left to fix in the paraformaldehyde solution at room temperature for 10 min. Washing was carried out twice in 2x SSC for 5 min. Dehydration was done by incubating slides sequentially in 70%, 90% and 100% ethanol, slides were then drained and

left to air dry.

2.8.2 Probe Preparation

Probes were generated by incorporating modified nucleotides using either a Nick Translation Kit (Roche) or by PCR (see 2.4.2). Probes were labelled with either digoxigenin-11-2'-deoxy-uridine-5'-triphosphate (digoxigenin-11-dUTP) or biotin-16-2'-deoxy-uridine-5'-triphosphate (biotin-16-dUTP).

2.8.3 Hybridisation Solution

The hybridisation solution was made up with the following components: 20 μl 100% formamide, 8 μl 50% dextran sulphate, 4 μl 20x SSC, 0.5 μl 10% sodium dodecyl sulfate, 10-50 ng of each DNA probe and sterile water to make a total volume of 40 μl. The hybridisation solution was denatured by heating at 75 $^\circ$C for 12 min followed by rapid cooling on ice, to ensure that the DNA was distributed equally the probe was vortexed and centrifuged briefly.

2.8.4 Hybridisation

The 40 μl hybridisation solution was added to the squashed material and covered using a Parafilm© coverslip. The slide was placed in a Alpha Unit block assembly mounted in a DYAD DNA Engine (MJ Research), the following temperatures were used for heating the slides: 71 $^\circ$C for 3.5 min, 55 $^\circ$C for 2 min, 50 $^\circ$C for 30 sec, 45 $^\circ$C for 1 min, 42 $^\circ$C for 2 min, 40 $^\circ$C for 5 min, 38 $^\circ$C for 5 min; the slides were then incubated in a humid chamber at 37 $^\circ$C overnight.

2.8.5 Post-Hybridisation Washing

Slides were placed in 2x SSC three times for 5 min at room temperature, ensuring to remove coverslips. Stringent washes were carried out by incubating the slides at 42 $^\circ$C for 5 min in 0.1x SSC, 20% formamide (v/v); this was repeated once more. Slides were then washed in 2x SSC at 42 $^\circ$C for 3 min, this was repeated twice more. Slides were then placed in 4x SSC/0.2% (v/v) Tween-20 for 5 min and then briefly left to drain.

2.8.6 Detection

Slides were prepared for detection by adding 80 μl bovine serum albumin (BSA) blocking solution (5% BSA (w/v) in 4x SSC Tween-20) and covered with a Parafilm© coverslip for 5

min. Biotin labelled DNA probe was detected using Cy3-avidin; digoxigenin labelled DNA probe was detected using anti-digoxigenin-FITC. After draining the BSA blocking solution 50 μl of detection reagent (for two colour FISH: 3 μl anti-digoxigenin-FITC [200 μg/ml] and 3 μl Cy3-avidin [1 μg/ml] made up to 600μl with BSA block solution) was applied. The slide was then incubated for 1 hr at 37 $^\circ$C in a humid chamber. Slides were then washed briefly in 4x SSC/Tween. Following draining, a drop of Vectashield© (Vector laboratories, California, USA) containing 4',6-diamidino-2-phenylindole (DAPI) was applied and a 24 x 50 mm glass coverslip was placed over the material.

Images were acquired using a Leica DMRA2 epifluorescence microscope and ORCA-ER digital camera (Hamamatsu, Japan) with Openlab (Improvision, Coventry, UK). All images were treated uniformly for colour balance, contrast and brightness. Images were exported as 72 ppi TIFF files were resized to 150 ppi, cropped and collated on a 300 ppi canvas with Photoshop CS2 (Adobe Systems Incorporated, USA). A scale bar was imported for scaling, this was made by photographing a 10 μm graticule with the same objective lens.

For measuring chromosome arm lengths and DNA fibres images, ImageJ version 1.36b (http://rsb.info.nih.gov/ij/) was used. Images were converted to 8-bit and the scale was set using the scale bar, the selection tool was then used to trace along the feature, this was repeated three times and the average measurement was recorded.

2.9 Extended DNA Fibre (EDF) FISH

EDF-FISH was carried out following the protocol by Fransz et al. (1996) where nuclei are extracted and spread via lysis with some minor modifications carrying out the following procedure.

2.9.1 Nuclei Extraction

1 g of fresh leaf material was chopped in a Petri dish with 500 μl of freshly prepared nuclear isolation buffer (NIB: 10 mM Tris-HCl, pH 9.5; 10 mM EDTA; 100 mM KCl; 0.5 M sucrose; 4 mM spermidine; 1 mM spermine; 0.1% (v/v) 2-mercaptoethanol) with a razor blade over ice. The suspension was filtered through a 120 μm nylon mesh into a 1.5 ml eppendorf tube. The filtrate was then passed through a 50 μm nylon mesh into a 1.5 ml eppendorf tube, $1/20^{th}$ the volume of 10% Triton-X100 in NIB was added to remove plastids. The solution was then centrifuged in a microcentrifuge at 2000 rpm for 3 min at 4 $^\circ$C. The supernatent was discarded and the white pellet was resuspended with 20 μl NIB. The integrity of nuclei

could be checked at this stage by adding 2 μl of the filtered nuclei to one drop of Vectashield© (Vector laboratories, California, USA) containing DAPI and covering with a 24 x 50 mm glass coverslip. For best results filtered nuclei were extended immediately.

2.9.2 DNA Fibre Extension

1 μl of filtered nuclei suspension was pipetted in two places at one end of a polylysine coated slide and allowed to dry at room temperature for about 3 minutes. 30 μl of STE buffer (0.5% (w/v) SDS; 5 mM EDTA; 100 mM Tris, pH 7.0) was then added over the dried spots to lyse the nuclei for 4 min. The slide was then placed on a slope to allow the chromatin to extend along the length of the slide. After drying, material was fixed in modified Carnoy's solution (3:1 ethanol:glacial) for 2 min and dried. After aging the slides at 60 °C for 30 min, slides were stored at 4 °C.

2.9.3 Hybridisation Solution

The hybridisation solution was made up as above for standard FISH with the following components: 20 μl 100% formamide, 8 μl 50% dextran sulphate, 4 μl 20x SSC, 0.5 μl 10% SDS, 10-50 ng of each DNA probe and sterile water to make a total volume of 40 μl.

2.9.4 Hybridisation

The 40 μl hybridisation solution was added to the extended material and covered using a Parafilm© coverslip. The slide was placed in a Alpha Unit block assembly mounted in a DYAD DNA Engine (MJ Research), the following temperatures were used for heating the slides: 80 °C for 2 min, 55 °C for 2 min, 50 °C for 30 sec, 45 °C for 1 min, 42 °C for 2 min, 40 °C for 5 min, 38 °C for 5 min; the slides were then incubated in a humid chamber at 37 °C overnight.

2.9.5 Post-Hybridisation Washing

Slides were placed in 2x SSC three times for 5 min at room temperature, ensuring to remove coverslips. Stringent washes were carried out by incubating the slides at 42 °C for 5 min in 0.1x SSC, 50% formamide (v/v); this was repeated once more. Slides were then washed in 2x SSC at 42 °C for 3 min, this was repeated twice more. Slides were then placed in 0.1x SSC pH 7.0 at 55 °C repeating twice more. Slides were then placed in 4x SSC/0.2% (v/v) Tween-20 for 5 min and then briefly left to drain.

2.9.6 Detection

Slides were prepared for detection by adding 80 μl BSA blocking solution (5% BSA (w/v) in 4x SSC/0.2% (v/v) Tween-20) and covered with a Parafilm© coverslip for 5 min. Biotin labelled DNA probe was detected using Cy3-avidin; digoxigenin labelled DNA probe was detected using anti-digoxigenin-FITC. After draining the BSA blocking solution 50 μl of detection reagent (3 μl anti-digoxigenin-FITC [200 μg/ml] and 3 μl Cy3-avidin [1 μg/ml] made up to 600μl with BSA block solution) was applied. The slide was then incubated for 1 hr at 37 $^\circ$C in a humid chamber. Slides were then washed briefly in 4x SSC/0.2% (v/v) Tween-20. Following draining, a drop of Vectashield© (Vector laboratories, California, USA) containing DAPI was applied and a 24 x 50 mm glass coverslip was placed over the material.

Images were acquired using a Leica DMRA2 epifluorescence microscope and ORCA-ER digital camera (Hamamatsu, Japan) with Openlab (Improvision, Coventry, UK). All images were treated uniformly for colour balance, contrast and brightness. Images were exported from Openlab as 72 ppi TIFF files, these were converted to 150 ppi, cropped and collated in Photoshop CS2 (Adobe Systems Incorporated, USA) on a 300 ppi canvas.

Chapter 3

Allium Tandem Repeats and Sequences Isolated by DOP-PCR

3.1 Introduction

This chapter collates data from DOP-PCR cloning techniques designed to isolate terminal DNA sequences, and some DNA sequences previously proposed to be telomeric in *Allium*. The chapter discusses structure and genomic organisation of tandem repeats found in *A. cepa* and discusses their potential role in telomere biology.

3.1.1 Repetitive DNA

Repetitive DNA is a major constituent of eukaryote genomes with the proportion increasing as genomes size increases (Bennett and Leitch 2005, Bennetzen et al. 2005). A large proportion of repetitive DNA is dispersed and comprises active and inactive mobile genetic elements; in plants they can account for more than 50% of the total genome size (Kumar and Bennetzen 1999, Kidwell and Lisch 2000, Vitte and Panaud 2005, Hawkins et al. 2006). Mobile genetic elements are divided into 2 main classes, retrotransposons (class I) and transposons (class II) (Finnegan 1989, Wicker et al. 2007). Retrotransposons move around the genome via an RNA copy produced by reverse transcriptase (Kumar and Bennetzen 1999). Transposons move either directly or indirectly from a DNA intermediate (Feschotte and Pritham 2007).

Repetitive DNA can also occur in arrays as simple repeats with monomers of various sizes, as microsatellites (*c.* 1 - 5 bp), minisatellites (*c.* 5 - 25 bp) and satellite repeats (*c.*

>25 bp) (Franck et al. 1991, Ramel 1997). Satellite repeats are typically heterochromatic and can be located at centromeres, pericentromeres, interstitial sites and at subtelomeres, they can show rapid evolutionary rates and copy number variation (Guerra 2000, Heslop-Harrison 2000, Lamb et al. 2007, Davison et al. 2007). Heterochromatic regions containing repetitive sequences such as tandem repeats, retrotransposons and transposable elements are thought to be a major cause of changes in chromosome structure (Shapiro and von Sternberg 2005, Raskina et al. 2008).

3.1.1.1 Tandem Satellite Repeats

A combined analysis of 152 plant satellite repeat families revealed that although repeats have been identified with lengths up to 4 kb, most repeats fall into size ranges of 135-195 bp and 315-375 bp (Macas et al. 2002). These sizes correspond to the DNA folding periodicity of chromatin where approximately 146 bp of DNA wraps around the core histone (the octamer has two subunits of H2A, H2B, H3, H4) and another 15-100 bp of DNA is bound by a linker histone (histone H1 variants) (Kornberg 1977, Noll and Kornberg 1977, Vignali and Workman 1998, McGhee and Felsenfeld 1980), this DNA/protein structure is called a nucleosome (Van Holde et al. 1974). In a review of satellite repeats Sharma and Raina (2005) take the view that DNA repair/recombination processes are important factors leading to selection for mono- and dinucleosome sized repeat lengths. In addition to monomer length, within a repeat there are often features that influence nucleosome positioning, such as certain regularly spaced dinucleotides or larger sequence motifs influencing DNA curvature (Trifonov and Sussman 1980, Trifonov 1985) and bendability (Gabrielian and Pongor 1996, Vinogradov 2003).

Certain satellite repeats have been found to show sequence similarity to other tandem or dispersed repetitive DNA families. In *Anemone hortensis* a 743 bp subtelomeric satellite repeat family (AhTR2) contains the telomeric minisatellite motif (TTTAGGG) and degenerate variants (e.g. TTTAGTG, TTTTGGG) (Mlinarec et al. 2009). A family of mammalian centromere repeats contain a motif similar to part of the terminal inverted repeat of a transposable element (Kipling and Warburton 1997). Observations such as these support the idea that satellite repeats undergo a cyclical mode of evolution where new repeats arise through recombination of existing sequences and over time are lost through mutation/deletion/recombination (Flavell 1980).

3.1.2 Characterisation of Repetitive DNA

A number of methods exist to amplify repetitive DNA including genomic DNA digestion and ligation (Hemleben et al. 1982, Kato et al. 1984, Gazdova et al. 1995) and self-primed PCR (Buntjer and Lenstra 1998, Macas et al. 2000). Another technique called degenerate oligonucleotide primed PCR (DOP-PCR) has been developed to isolate repeated sequences, and has proved effective for a variety of repeat types ranging from low to high copy number. The technique can be applied to total genomic DNA (Telenius et al. 1992) as well as from flow sorted chromosomes, microdissected chromosome fragments or entire chromosomes (Langford et al. 1992, Viersbach et al. 1994, Jamilena et al. 1995, Houben et al. 1996, Macas et al. 1996, Buzek et al. 1997). To date the DOP-PCR technique has not been used to isolate telomere sequences *de novo*. A technique which has isolated telomere repeats successfully is 454-sequencing (Margulies et al. 2005) of pea total genomic DNA, albeit in low numbers (14 reads out of a total 319,402) (Macas et al. 2007). This technique also identified blocks of variant telomere TTAGG and TTTAGG minisatellites (Macas et al. 2007). FISH using concatamers of these variant telomere motifs confirmed that these do indeed occur at the chromosome ends(Macas et al. 2007).

3.1.3 *Allium* Genome Organisation

Full-scale genome sequencing has not until recently been practical in *Allium* because of the large genome sizes which are approximately two orders larger than that of *A. thaliana* (fig. 3.1.3). So far only a partial BAC genomic DNA library of *Allium cepa* has been made and one BAC sequence characterised in detail (Suzuki et al. 2001, Do et al. 2004). Prior to the development of DNA sequencing, genome content was inferred from denaturation/reassociation studies. From these analyses it was estimated from *A. cepa* that 83.3% of DNA can be placed in one of three fractions: (I) 41% highly repetitive DNA (average of 21,600 copies), (II) 36.4% medium repetitive fraction (average of 225 copies), (III) 5.9% single copy (non-repeated) DNA (Stack and Comings 1979). Annealing of different repetitive DNA fractions to each other indicated that much of the single copy and repetitive DNA sequences are interdispersed between each other (Stack and Comings 1979).

Although initial attempts were unsuccessful at isolating satellite DNA by CsCl density centrifugation (Ingle et al. 1973, Stack and Comings 1979), a modification by using the AT-binding Hoechst 33258 to bind DNA instead of actinomycin D, which binds methyl-CpG, enabled the separation of several satellite bands (Barnes et al. 1985). The authors then

Species	1C value (pg)	1C value (Mb)	Reference
Arabidopsis thaliana	0.16	156	1
Oryza sativa	0.50	489	2
Pisum sativum	4.88	4773	2
Triticum aestivum	17.33	16949	3
Allium cernuum	17.10	16724	4
Allium cepa	16.75	16381	5
Allium genus (average)	19.61	19179	6

Table 3.1: Haploid genome sizes of *Allium* species and other plants
Published genome sizes for some plant species and an average based on available measurements for 157 different *Allium* species. The genome size in Mb was calculated using 1 pg = 978,000 bases. References: 1, Bennett et al. (2003); 2, Bennett and Smith (1991); 3, Bennett and Smith (1976); 4, Jones and Rees (1968); 5, Van't Hof (1965); 6, Bennett and Leitch (2004).

used a restriction enzyme to release a *c.* 375 bp satellite, later designated *A. cepa* satellite (ACSAT) (Pich, Fritsch and Schubert 1996). Homologous subtelomeric sequences are present in all members of subgenus *Cepa* (Pich, Fritsch and Schubert 1996). This repeat accounts for 4.4% of the *A. cepa* genome with an average pairwise heterogeneity of 8.5%, and was found on all subtelomeres apart from the NOR bearing arm of chromosome VI (Barnes et al. 1985). The subtelomeric position of ACSAT and *A. fistulosum* satellite (AFISAT) agrees with the position of heterochromatic Giemsa C-bands in the respective species (El-Gadi and Elkington 1975, Vosa 1976, Barnes et al. 1985, Irifune et al. 1995).

3.1.4 *Allium* Genome Evolution

Recently, physical mapping and EST libraries have been assembled for some *Allium* species. Both of these methods reveal that considerable genome turnover has occurred within the genus *Allium*, and in the Asparagales as a whole (by comparisons to asparagus *(Asparagus officinalis*). Genome turnover is used to describe the divergence in genomic sequences following speciation, due to changes such as sequence shuffling, transposition, amplification, mutation and deletion (Murray et al. 1981, Rose and Doolittle 1983, Lim et al. 2007, Leitch 2007). RFLP and AFLP mapping studies have shown that the genomes of the cultivated *A. sativum* (garlic) (subgenus *Allium*), *A. cepa* and *A. fistulosum* (bunching onion) (both subgenus *Cepa*) contain substantial amounts of duplicated loci that are unlinked (King et al. 1998, van Heusden et al. 2000, Ipek et al. 2005). In *A. cepa,* from duplicated AFLP markers there was evidence for this occurring between chromosomes II and IV; and III and VI (van Heusden et al. 2000). When a sample of coding DNA sequences derived from *A. cepa, A. sativum* and asparagus ESTs were compared to rice (order Poales) there was an aver-

age similarity of 78% (Kuhl et al. 2004). However, when comparisons of *A. cepa* and rice genomic sequences were made, there was at best only minimal evidence for microsynteny within genic regions, albeit based on limited gene sampling (Martin et al. 2005). Further work selecting different sets of highly similar coding sequences, which are tightly linked in asparagus, showed a complete absence of synteny, when compared to rice and even to *A. cepa* (Jakse et al. 2006). This comparative data suggests that whilst coding DNA may be generally conserved, extensive genome turnover has occurred within the Asparagales lineage since it diverged from its sister lineage the Poales (Martin et al. 2005, Jakse et al. 2006).

3.2 Evolution of Repetitive DNA Families

In some repeated DNA families such as 5S and 35S rRNA, genes diverge between species, yet show high levels of homogeneity within species (Brown et al. 1972, Coen, Strachan and Dover 1982, Nei 1987, Hillis and Dixon 1991, Kellogg and Appels 1995, Elder and Turner 1995). If each unit were evolving independently, the divergence from the common ancestor within a species would be expected to be comparable to that between the two species. This is not the pattern observed, hence the homogeneity indicates that they are not diverging independently; this pattern is described as 'concerted evolution'. However, this term is also used in some of the literature to describe the molecular processes responsible for the homogeneity which has led to considerable confusion. The view that DNA homogenisation mechanisms are often responsible for the concerted evolution has been supported by empirical sequencing studies on *Drosophila* and fungi 35S rRNA, plant 5S rRNA and also for non-coding tandem repetitive DNA such as centromere repeats (Ganley and Kobayashi 2007, Stage and Eickbush 2007, Kellogg and Appels 1995, Hall et al. 2003, 2005). Nei and Hughes (1991) put forward an alternative model to explain low sequence divergence of multigene families within species named birth-and-death evolution. This process is thought occur in genes including the mammalian major histocompatability complex (Hughes and Nei 1989), eukaryote histone H1 (Eirin-Lopez et al. 2004), vertebrate immunoglobulin (Nei et al. 1997) and plant MADs-box (Nam et al. 2004) gene families. When evolving through birth-and-death, genes are thought to be 'born' through duplication events with selection on individual genes being the predominant force resulting in homogeneity, gene 'death' may occur through gene pseudogenisation and/or deletion (Nei and Rooney 2005). However, gene duplication is a DNA homogenisation mechanism that results in concerted evolution. Therefore, it is not clear if there is any qualitative distinction between the pattern of homogenisation caused by

gene duplication under the birth-and-death model from the other homogenisation processes that result in concerted evolution.

3.2.1 Concerted Evolution of 35S rDNA

35S rDNA is a multicopy nuclear gene encoding 18S-5.8S-26S ribosomal RNA (rRNA) which is a structural, non-coding RNA (ncRNA), which together with 5S rRNA form the RNA components of ribosomes. The DNA sequences which encode it are organised as one or several arrays of tandem repeats within the nucleolar organising region(s) (NOR) (Reviewed by: Pikaard 2002). During interphase, one or several NORs form a nucleolus (or nucleoli), where rDNA is transcribed by RNA polymerase I (Pol I) in the dense fibrillar component (DFC) producing precursor RNA (pre-RNA) (Gonzalez-Melendi et al. 2001), in plants ribosome assembly probably takes place outside the DFC in the surrounding granular component (GC) (Olmedilla et al. 1993, Shaw and Doonan 2005). The pre-RNA then undergoes cleavage by endonucleases, chemical modifications (e.g. methylation, pseudouridylation) some of which are directed by small nucleolar RNAs (snoRNAs), before mature rRNAs assemble with ribosomal proteins to form ribosomes (Brown and Shaw 1998).

rDNA is highly transcribed and if rDNA units are deleted in *Drosophila melanogaster* and drop below a lower limit (*c.* 150 units per genome), they are rapidly amplified (Tartof 1971, 1974). A large number of inter-/intramolecular interactions are required for the processing and assembly of the different components of ribosomes, it may be for this reason that transcribed rDNA undergoes concerted evolution, so as to produce a homogenous 'population' of rDNA units (Long and Dawid 1980). The molecular drive hypothesis is used to explain the spread of variants such as those in a multi-gene family, as long as recombination events are more frequent than the mutation rate (Dover 1982, Dover et al. 1982, Dover 1984, Ohta and Dover 1984).

The effects of concerted evolution in rDNA were first studied by Brown et al. (1972) in *Xenopus laevis* and *X.* [*borealis*] *mulleri*; where they identified an apparently paradoxical occurence of intraspecific homogeneity between units as well as interspecific divergence in the spacer regions. Work on primate and mouse rDNA showed that homogenisation can occur between rDNA arrays on non-homologous chromosomes (Arnheim et al. 1980, 1982). Although it may be that homogenisation occurs more efficiently between homologous than non-homologous arrays (Saghai-Maroof et al. 1984, Ellis et al. 1984, Seperack et al. 1988, Schlotterer and Tautz 1994). There are a number of likely homogenisation mechanisms that result in concerted evolution:

- unequal crossing-over (between sister, non-sister chromatids and non-homologous chromosomes) (Smith 1976, Szostak and Wu 1980, Arnheim et al. 1980, Rogers and Bendich 1987)

- gene conversion (Klein and Petes 1981, Coen, Thoday and Dover 1982, Linares et al. 1994, Gangloff et al. 1996, Ganley and Scott 1998)

- intra-array recombination leading to copy number expansion/contraction (Tartof 1974, Petes 1980, Kobayashi and Ganley 2005, Davison et al. 2007)

Homogenisation mechanisms can act to both maintain uniformity (e.g. 5.8S gene) within a family, as well as to rapidly spread new mutations/variants (e.g. intergenic spacer (IGS) region). This may occur because recombination does not act on a cistron in a uniform manner, but rather, there are "modules" or "segments" within a cistron that can evolve semi-independently due to recombination hotspots; as shown for *D. melanogaster* IGS (Polanco et al. 1998, 2000) and the microalga *Prototheca wickerhamii* 18S gene (Ueno et al. 2007). This type of data is still sparse due to the necessity for long tracts (10-15kb) of sequence and sufficient polymorphisms for detecting conflicting phylogenetic signals.

An important factor affecting the efficiency of concerted evolution is rDNA unit copy number. rDNA sequences from several eukaryotic genome sequences have been compiled separately to estimate intragenomic variation. From the five yeast species (Ganley and Kobayashi 2007) and the 12 *Drosophila* species (Stage and Eickbush 2007) analysed there appears to be a marked difference in the degree of intragenomic homogeneity between the groups. Yeast species show very low levels of sequence heterogeneity throughout the locus (Ganley and Kobayashi 2007). In contrast, *Drosophila* species tend to show more polymorphisms, particularly in the intergenic regions, where polymorphisms can occur in more than 5% of units (Stage and Eickbush 2007). Polymorphisms present in less than 5% of units are distributed evenly throughout the rDNA unit, it is thought that these infrequent mutations may be below the level of selection (Stage and Eickbush 2007). Whereas the fungi species studied have on average from 45 units (*Aspergillus nidulans*) up to 150 units (*Saccharomyces cerevisiae*) per haploid genome, *Drosophila* have several hundred units per genome, but a similar proportion are transcribed in both groups. The increased copy number redundancy in *Drosophila* species may be one important factor for increased rDNA variability (Long and Dawid 1980, Stage and Eickbush 2007). Considering that there is a positive correlation between rDNA copy number and genome size in both plants and animals (Prokopowich et al. 2003), one might expect organisms that have relatively large genomes (e.g. *Allium cepa*) to

harbour higher numbers of rDNA variants.

3.2.2 35S rDNA Position Variability in *Allium*

The chromosome number for *A. cepa* is $2n = 2x = 16$ (Stack and Comings 1979), with one to four NORs reported on the short arms of chromosomes VI and VIII (Stack and Comings 1979, Sato 1981). Later, Panzera et al. (1996) reconfirmed by FISH with a 35S rDNA probe (pTA71; Gerlach and Bedbrook 1979) the position of NORs on chromosomes VI and VIII; additionally a fifth minor 35S rDNA site was also found on the long arm of chromosome VIII. There are two homologous pairs of 5S rDNA loci on chromosome VII, with each locus bearing a distinct 5S rDNA family (Shibata and Hizume 2002).

Within genus *Allium* a number of species have been reported to exhibit variation in the position of rDNA loci. In *A. cepa*, rDNA has been detected at either two, three or four sites on the short arms of chromosomes VI and VIII, varying both between and within individual plants (Schubert and Wobus 1985). This variation in the numbers of rDNA sites gave the appearance that the NORs were "jumping" between chromosomes. In F1 "top onion" pseudodiploid hybrids between *A. cepa* and *A. fistulosum* rDNA position varied amongst the NOR bearing arms and probably other distal sites (Schubert 1984, Schubert and Wobus 1985, Pich, Fuchs and Schubert 1996). In *Allium sphaerocephalon* rDNA loci were also found to vary between and within individuals (Garrido-Ramos et al. 1992). While in *Allium schoenoprasum* Garrido et al. (1994) observed rDNA locus variability in plants from different populations. These data collectively reveal that *Allium* species exhibit unusually high levels of rDNA position variability, even within individual plants. In other plant groups the transposition of rDNA loci (which may involve array expansion) has been inferred from comparing the linkage maps and/or karyotypes of closely related species (Dubcovsky and Dvorak 1995, Shishido et al. 2000, Adams et al. 2000, Datson and Murray 2006, Cai et al. 2006). Studies of synthetic tobacco polyploids (*Nicotiana sylvestris x Nicotiana tomentosiformis*) have shown that new arrays can appear within a few generations (Skalická et al. 2003).

3.2.3 *Candida albicans* Telomere Repeat Sequence

The telomere sequence from *Candida albicans* was found to weakly hybridise, by Southern hybridisation, to genomic DNA of *Allium* species from subgenus *Cepa* (Eva Sýkorová, unpublished). *C. albicans* belongs to the group of yeast-like fungi (phylum Ascomycota) (Liu

Species	Monomer sequence	Reference
Saccharomyces cerevisiae	**TGGTGT**G**TGGGT**G	1
Saccharomyces exiguus	**TGGTGT**G**TGGGT**G	2
Saccharomyces kluyveri	GACATGCGTACTG**TGA**G**GT**C**TGGGT**G	2
Kluyveromyces lactis	TGATTAGGTATG**TGGTGTACGGAT**T	4
Candida glabrata	CTG**TGG**G**GT**C**TGGGT**G	4
Candida tropicalis	TCACGATCAT**TGGTGTA**M**GGAT**G	4
Candida albicans	TCTAACTTCT**TGGTGTACGGAT**G	3

Table 3.2: Telomere repeats of fungi (phylum Ascomycota)
Selection of published telomere monomer sequences of yeasts from across the phylum Ascomycota. Nucleotides matching the conserved core region determined by Cohn et al. (1998) are indicated in bold. Data from the sources as indicated: 1, Szostak and Blackburn (1982); 2, Cohn et al. (1998); 3, McEachern and Hicks (1993); 4, McEachern and Blackburn (1994).

et al. 1999), which exhibit an extraordinary amount of telomere sequence variation between species and have some of the longest telomere repeat monomers known. The telomere repeat of *C. albicans* is 23 bp with the sequence (TCTAACTTCT**TGGTGTACGGAT**G)$_n$ (McEachern and Hicks 1993). The sequence contains a conserved core region, shown in bold, which is well conserved in other fungi of the class Hemiascomycetes (phylum Ascomycota) (see 3.2.3). This conserved region corresponds to a sequence which is bound by *Saccharomyces cerevisiae* RAP1 telomere regulating protein (Larson et al. 1994, Cohn et al. 1998).

3.3 Materials and Methods

3.3.1 Microdissection, DOP-PCR and Colony Blotting

Allium cernuum root tip squashes were microdissected following section 2.3, 30 metaphase chromosome ends (estimated to be less than 2-4 μm in length) were collected with the exception of the two pairs that have NORs. These were pooled and used for degenerate oligonucleotide-PCR following 2.4.5.

5 μl of primary and secondary DOP-PCR products were ran on 1.2% agarose gel with a size marker. The DNA was then alkaline transferred onto a nylon membrane and probed with *Allium fistulosum* genomic DNA by Southern hybridisation as in section 2.5.3. As *A. fistulosum* (subgenus *Cepa*) is not closely related to *A. cernuum*, this was used to crudely identify whether highly conserved DNA sequences were present among secondary DOP-PCR products.

In order to isolate repetitive sequences that are present in medium to high copy number in the *A. cernuum* genome the secondary DOP-PCR products were cloned into pCR4-TOPO

plasmid (Invitrogen) (following section 2.6.1). The ninety individual colonies produced were screened by colony blotting, using *A. cernuum* genomic DNA as a probe (following 2.6.4 and 2.6.5). Seventeen colonies with signals ranging from relatively weak to strong (i.e. low to high copy) were selected for replication in liquid LB and plasmid extraction (following 2.6.2). The plasmids were then sequenced using universal M13 primers (see 2.7.1), sequences were compared against each other and checked for similar sequences on GenBank (see 2.7.5).

3.3.2 FISH

DNA templates for making FISH probes were isolated and labelled as follows: (1) The ACSAT probe was isolated by PCR from *A. cepa* genomic DNA using the forward primer X02572F: 5'-CCACGTGACGAAAAAACGAAGGGT-3' and reverse primer X02572R: 5'-CGGGATCCCCGTGGCCGGTCTATG-3' described by Pich, Fritsch and Schubert (1996). Products of monomer and dimer repeat size were then cloned as described in section 2.6.1, plasmid DNA was used as a template for PCR labelling with biotin-16-dUTP. (2) For the 35S rDNA probe a plasmid, pTA71, containing an entire 35S rDNA cistron from *Triticum aestivum* (wheat) (Gerlach and Bedbrook 1979) was labelled by nick translation with digoxigenin-11-dUTP. (3) For the 5S rDNA probe a plasmid, pTZ19-R, containing a 5S rDNA unit from *Nicotiana rustica* (Venkateswarlu et al. 1991) was labelled with biotin-16-dUTP. (4) The 18S rDNA probe was isolated by PCR from *A. cepa* genomic DNA using the forward primer A.18S2F: 5'-CGGAGAATTAGGGTTCGATTC-3' and the reverse primer 17_rev_SE: 5'-ACGAATTCATGGTCCGGTGAAGTGTTCG-3', which is derived from the reverse complement of primer 17SE from Sun et al. (1994). A plasmid containing a 1321 bp 18S rDNA fragment was then used as a template for nick translation labelling with digoxigenin-11-dUTP. (5) The *C. albicans* telomere sequence was in a plasmid pSK-BSA, provided by J. Nosek (Commenius Univ., Bratislava, Slovakia), it contains a *c.* 635 bp *Cla*I/*Sma*I insert containing around 25 TCTAACTTCTTGGTGTACGGATG telomere repeats and 60 bp of subtelomere DNA cloned in pBluescript SK-, originally cloned by M. McEachern (McEachern and Hicks 1993, McEachern and Blackburn 1994). The *C. albicans* telomere sequence was used as a template for nick translation labelling with biotin-16-dUTP. FISH was carried out according to section 2.8.

Primer name	Primer sequence (5'-3')
ca1_for	TCTAACTTCTTGGTGTACGGATG
ca1_rev	CATCCGTACACCAAGAAGTTAGA
ca2_for	ACTTCTTGGTGTACGGATGTCTA
ca2_rev	TAGACATCCGTACACCAAGAAGT
ca3_for	GTACGGATGTCTAACTTCTTGGT
ca3_rev	ACCAAGAAGTTAGACATCCGTAC
ca4_for	GGATGTCTAACTTCTTGGTGTAC
ca4_rev	GTACACCAAGAAGTTAGACATCC

Table 3.3: *C. albicans* PCR amplification primers

3.3.3 EDF-FISH

DNA fibres were prepared and used for *in situ* following section 2.9, the following DNA templates were used. (1) A plasmid containing 1321 bp of 18S rDNA from *A. cepa* (see previous section: 3.3.2) was as a template for PCR labelling with digoxigenin-11-dUTP. (2) The plasmid pSK-BSA was used as a template for nick-translation labelling with biotin-16-dUTP.

3.3.4 Tandem Repeat Isolation

In order to isolate tandemly repeated DNA sequences genomic DNA from *A. cernuum* was digested with 22 different restriction enzymes (*Bsr*DI, *Eco*RI, *Bam*HI, *Eco*RV, *Xba*I, *Tsp*509I, *Hin*dIII, *Xho*I, *Taq*I, *Bst*NI, *Mbo*I, *Dra*I, *Pst*I, *Msp*I, *Rsa*I, *Nde*I, *Mse*I, *Alu*I, *Ase*I, *Bcl*I, *Bam*HI, *Pvu*II) following section 2.5.1.

3.3.5 *C. albicans* Repeat Isolation - PCR

Oligonucleotides were designed to PCR amplify a putative *C. albicans* telomere-like repeat sequence ($\text{TCTAACTTCTTGGTGTACGGATG}_n$) from the genome of *A. cepa* (see table 3.3.5). All primers were made up of a single repeat unit but with a different start and end point in the repeat unit. This was done to vary the point at which the 3' end of the primer terminated as mismatches are least tolerated in this region for polymerase extension. A range of PCRs were carried out with either a single primer or a primer pair following a standard PCR protocol (2.4.1).

3.3.6 *C. albicans* Repeat Isolation - DNA Capture

To isolate *A. cepa* genomic sequences that were similar to the *C. albicans* telomere sequence a magnetic bead based capture protocol was carried out according to St John and Quinn

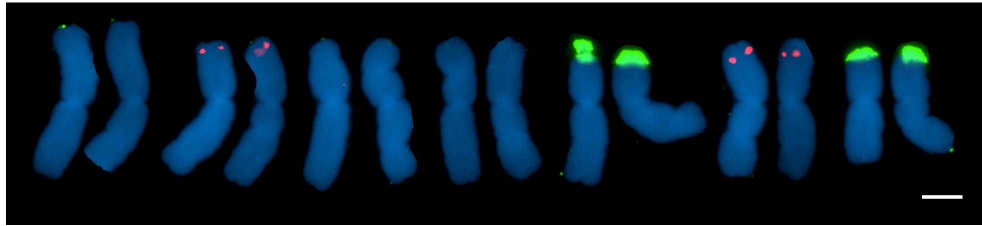

Figure 3.1: Karyotype of *Allium cernuum*
FISH karyotype with 18S rDNA (green) and 5S rDNA (red). Chromosomes are arranged by size as in Friebe (1989). Scale bar indicates 5 μm.

(2008). Briefly, this protocol requires genomic DNA to be digested and linkers ligated to both ends. This ssDNA pool is then mixed with a biotinylated ssDNA probe (a single *C. albicans* repeat unit), DNA molecules which form duplexes with the probe can be retained using streptavidin coated magnetic beads. This fraction is then used for cloning by digesting the adapters and ligating into a digested pZErO-2 plasmid vector ((2.6.1)).

3.4 Results

3.4.1 rDNA Localisation

The chromosome number for *A. cernuum* is $2n = 2x = 14$ (Chinnappa and Basappa 1986), with two pairs of major 35S rDNA loci and two pairs of 5S rDNA loci (fig. 3.4.1). The major 35S rDNA loci correspond to NORs previously detected as Giemsa positive/silver staining secondary constrictions (Friebe 1989). Any additional rDNA sites were referred to as minor rDNA sites, there were detected predominantly at, or near the ends of metaphase chromosome arms (fig. 3.4.1). The number of detectable minor rDNA sites varied as they were often so small that they were at the threshold of detection. It is estimated that at least one third of *A. cernuum* chromosome arms carry distal minor rDNA sites.

For *A. cepa* cv. Ailsa Craig there are two pairs of major 35S rDNA loci, within the range that has been detected previously for this species (Stack and Comings 1979, Sato 1981, Ricroch et al. 1992). But, as in *A. cernuum*, there were also several minor 35S rDNA sites detected in the distal regions of metaphase chromosome arms (fig. 3.4.1). On interphase spreads 35S minor rDNA sites are frequently detectable close to blocks of ACSAT repeats (fig. 3.4.1C). On metaphase chromosomes when 35S rDNA is localised simultaneously with ACSAT, 35S minor rDNA sitesappear to be in a more distal position relative to ACSAT (fig. 3.4.1).

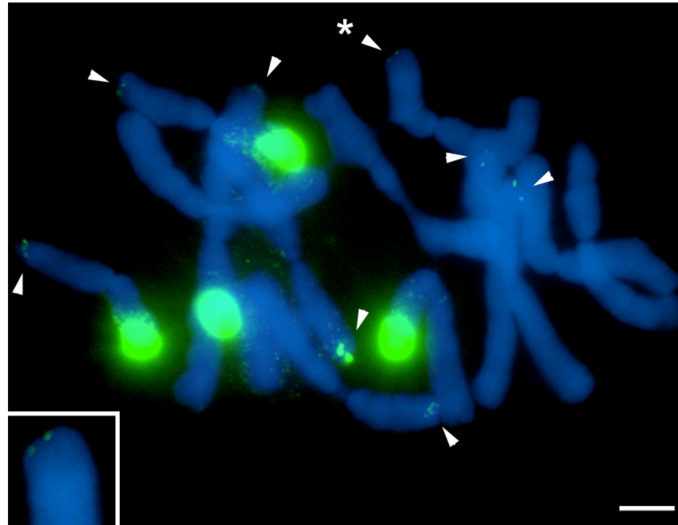

Figure 3.2: 35S rDNA minor sites in *Allium cernuum*
Metaphase chromosomes probed with pTA71 (green) 35S rDNA with prolonged exposure and counterstained with DAPI (blue). Arrowheads indicate 8 minor rDNA sites in addition to the 4 major rDNA sites, inset - 2x magnification and enhancement of a chromosome end (asterisk) with a minor site. Scale bar indicates 5 μm.

3.4.2 Tandem Repeat Isolation

None of 22 individual genomic DNA digestions produced any prominent bands below 3 kb, where satellite repeats would be expected. In the majority of cases digestions generated a broad range of restriction fragments visible as a smear. In the case of *Pst*I, *Msp*I, *Pvu*II and *Bcl*I the genomic DNA showed little to no digestion. This was probably due to the presence of at least one cytosine residue in each of the recognition sequences, methylation of these cytosines may have inhibited digestion.

3.4.3 Microdissection / DOP-PCR Isolated Sequences

3.4.3.1 Sequencing and FISH

The results from the sequencing of cloned microdissection sequences are summarised in table 3.4.3.1. Several sequences were found to be represented more than once, namely: cer540 (four times), cer703 (four times), cer330 (two times), cer352 (two times), cer129 was found repeated head to tail five times in a single clone. These clones were found to correspond to strongly amplified products of the same size generated in the DOP-PCR. None of the sequences retrieved showed any significant matches against the NCBI GenBank nucleotide and protein databases.

Sequences cer540 and cer703 were used as FISH probes against *A. cernuum* chromosomes,

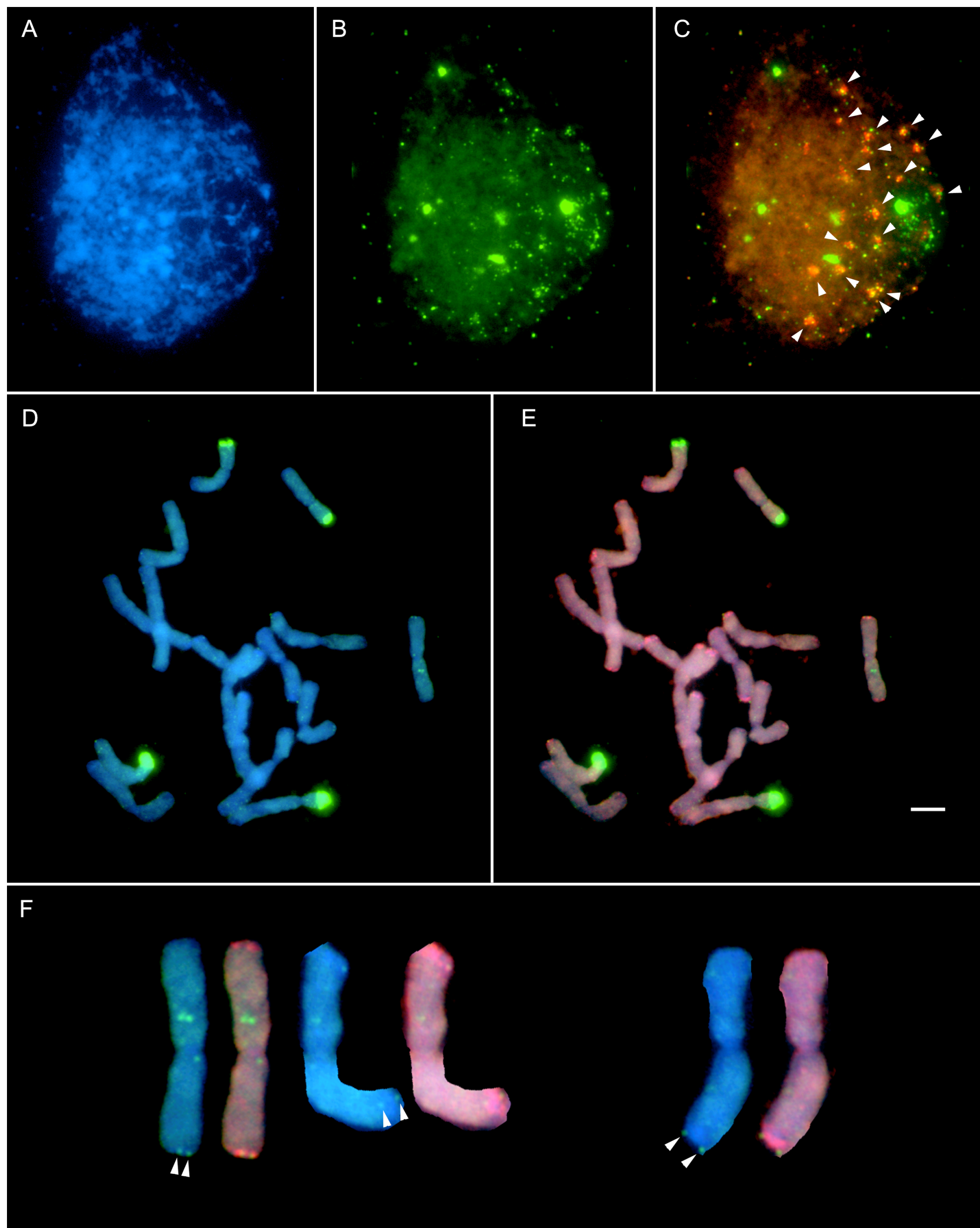

Figure 3.3: 35S rDNA and ACSAT localisation in *Allium cepa*
A. cepa root tip cell spreads. (A-C) Interphase spread. (A) DAPI (blue); (B) 18S rDNA, used to detect 35S sites (green); (C) ACSAT (red), 18S rDNA (green) and DAPI (blue), arrowheads indicate colocalised signals. (D-E) Metaphase spread. (D) 18S rDNA (green) and DAPI (blue); (E) ACSAT (red), 18S rDNA (green) and DAPI (blue). (F) 2.5x magnification of chromosomes from (D) and (E) highlighting the differences in minor 35S rDNA site position (indicated by arrowheads), a pair of homologous chromosomes are shown on the left and a single chromosome is shown on the right, note that each chromosome is shown twice with different probes, as described for D and E. Scale bar indicates 5 μm except for in (F) where it indicates 2 μm.

Sequence Number	Sequence Type
1	cer540
2	cer703
3	cer540
4	unique
5	cer352
6	unique
7	cer540
8	cer330
9	cer540
10	cer352
11	cer330
12	cer129
13	cer703
14	unique
15	cer703
16	cer703

Table 3.4: DOP-PCR microdissection sequences
Sequence number indicates the DOP-PCR clone. Sequence types are categorised by their insert fragment length with the exception of cer129, which was a 648 bp fragment containing 5 tandem repeat units of 129 bp.

shown in fig. 3.4.3.1 and fig. 3.4.3.1, respectively. Both sequences are widely dispersed throughout the genome of *A. cernuum*. However, whereas cer703 has a uniform distribution, cer540 has an increased concentration in the pericentromeric regions of some chromosomes. cer330 and cer352 showed dispersed ditributions to interphase nuclei (data not shown). Primers were designed to amplify the cer129 repeat from genomic DNA, however attempts to PCR amplify the repeat were unsuccessful.

3.4.3.2 Secondary Structure

Although no high scoring matches were obtained against public databases with the DOP-PCR isolated sequences, there is still a possibilty that the sequences obtained were derived from mobile elements. In the absence of any evidence for these sequences being part of mobile elements, other possibilities were explored to explain the abundance and dispersal throughout the genome of *A. cernuum*. Foreach of the *A. cernuum* repeats isolated by microdissection and for 100 randomisations of each sequence, secondary structure was predicted computationally. This was done to determine whether the computed thermodynamic stability was higher than would be expected by chance, given the same number of nucleotides but with a different order. Fig. 3.4.3.2 shows the lowest calculated Gibbs free energy (ΔG) values for the 100 randomised sequences and the position of the original isolated repeat.

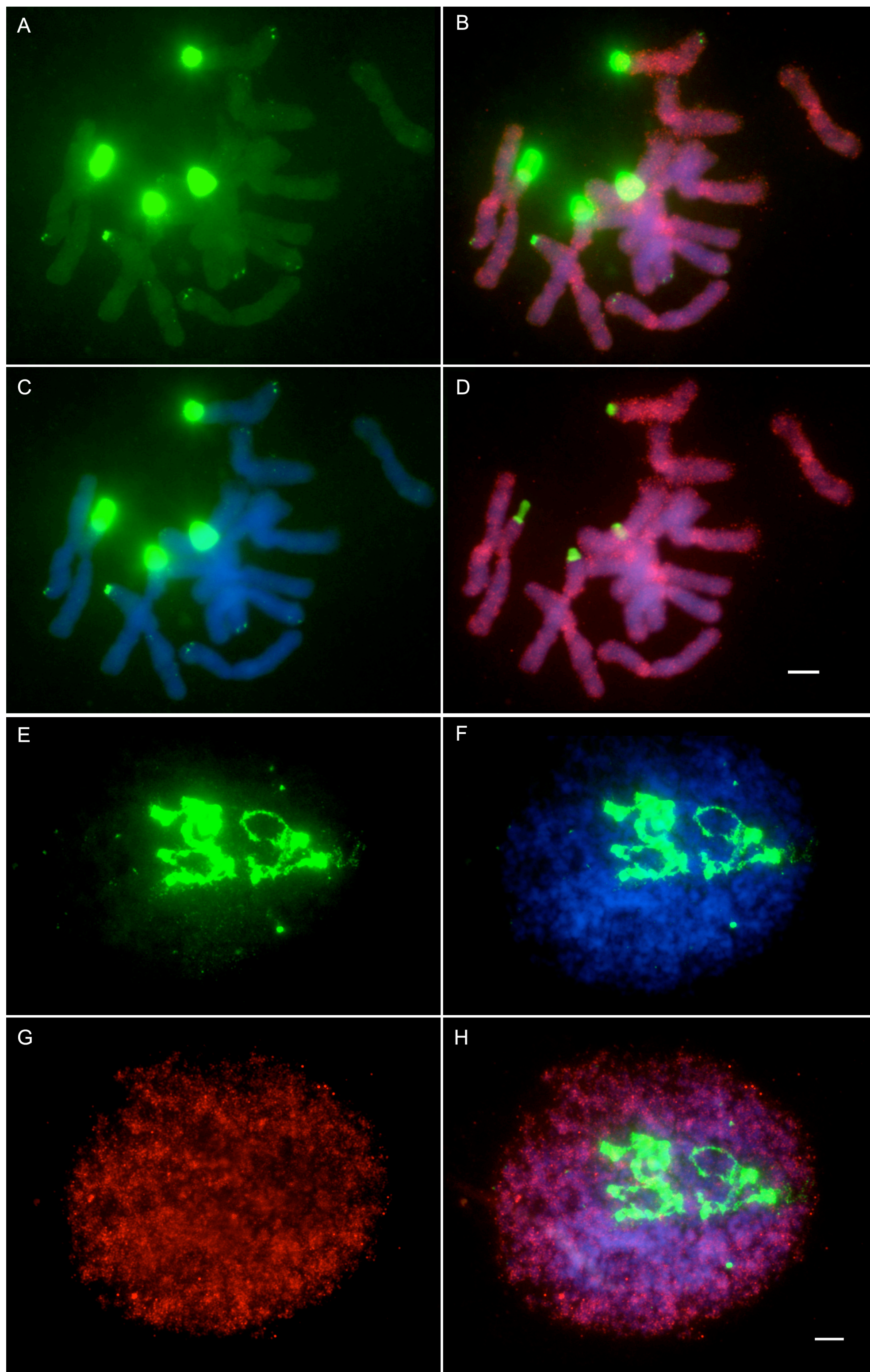

Figure 3.4: FISH localisation of cer540 sequence to *Allium cernuum*
A. cernuum root tip cell spreads. (A-D) Metaphase spread. (A) 35S rDNA (green) - long exposure; (B) cer540 (red), 35S rDNA (green) - long exposure and DAPI (blue); (C) 35S rDNA (green) - long exposure and DAPI (blue); (D) cer540 (red), 35S rDNA (green) and DAPI (blue). (E-H) Interphase spread. (E) 35S rDNA (green) - long exposure; (F) 35S rDNA (green) - long exposure and DAPI (blue); (G) cer540 (red); (H) cer540 (red), 35S rDNA (green) - long exposure and DAPI (blue). Scale bars indicate 5 μm.

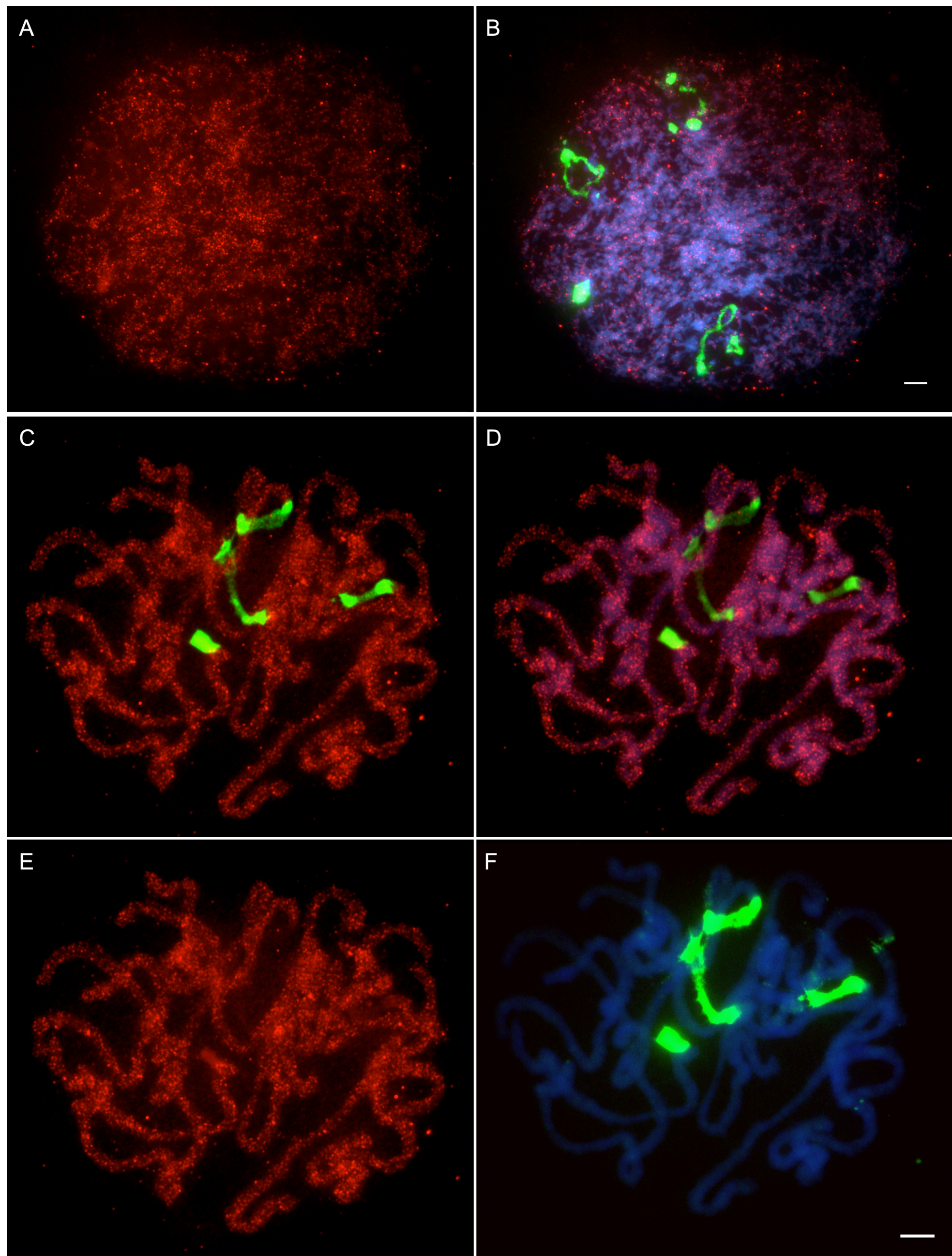

Figure 3.5: FISH localisation of cer703 sequence to *Allium cernuum*
A. cernuum root tip cell spreads. (A-B) Interphase spread. (A) cer703 (red); (B) cer703 (red), 35S rDNA (green) and DAPI (blue); (C-F) prophase spread. (C) cer703 (red) and 35S rDNA (green); (D) cer703 (red), 35S rDNA (green) and DAPI (blue); (E) cer703 (red); (F) cer703 (red), 35S rDNA (green) - long exposure and DAPI (blue). Scale bars indicate 5 μm.

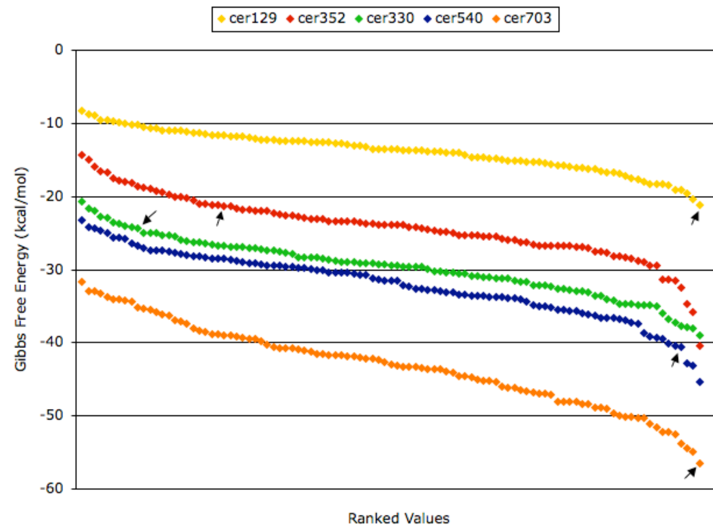

Figure 3.6: Calculated Gibbs free energy values for predicted secondary structures of *A. cernuum* DOP-PCR repeats and 100 randomisations
Graph shows Gibbs free energy (ΔG) values ordered from highest (left) to lowest (right). In each series the ΔG value corresponding to the original repeat (prior to randomisations) is indicated by an arrow.

3.4.4 Telomere Sequence from *Candida albicans*

FISH revealed colocalisation between the *C. albicans* telomere and 35S major and minor sites in *A. cepa*. The signal intensity varied between nuclei but tended to be strongest in interphase spreads, one of these is shown in fig. 3.4.4. Earlier FISHs have already shown that rDNA is found close to chromosome termini in *A. cepa* thus confirming, even at interphase, the localisation of the *C. albicans* telomere-like sequence in the distal chromosome region. Furthermore, the rDNA and *C. albicans* telomere repeat in most cases were colocalised, usually with similar signal strengths. To further examine the organisation at a higher resolution, EDF-FISH was used.

DNA fibres showing *C. albicans* telomere signal were infrequently found on slides, the majority of 18S rDNA probe did not colocalise with *C. albicans* telomere probe, but localised to long tracts of rDNA. Where colocalisation was found, the transition into colocalised regions was in some cases immediate (fig. 3.4.4), and the probes were found closely interspersed in an alternating pattern (fig. 3.4.4 and fig. 3.4.4). In most cases the *C. albicans* telomere probe was found at the ends of fibres, it is not clear if these ends correspond to chromosome termini or to broken fibres. As a consequence of this, fibres were stretched to a lesser degree than is typically seen (fig. 3.4.4 and fig. 3.4.4), making length estimation unreliable.

Two molecular methods were used to try to isolate the *C. albicans*-like sequence residing

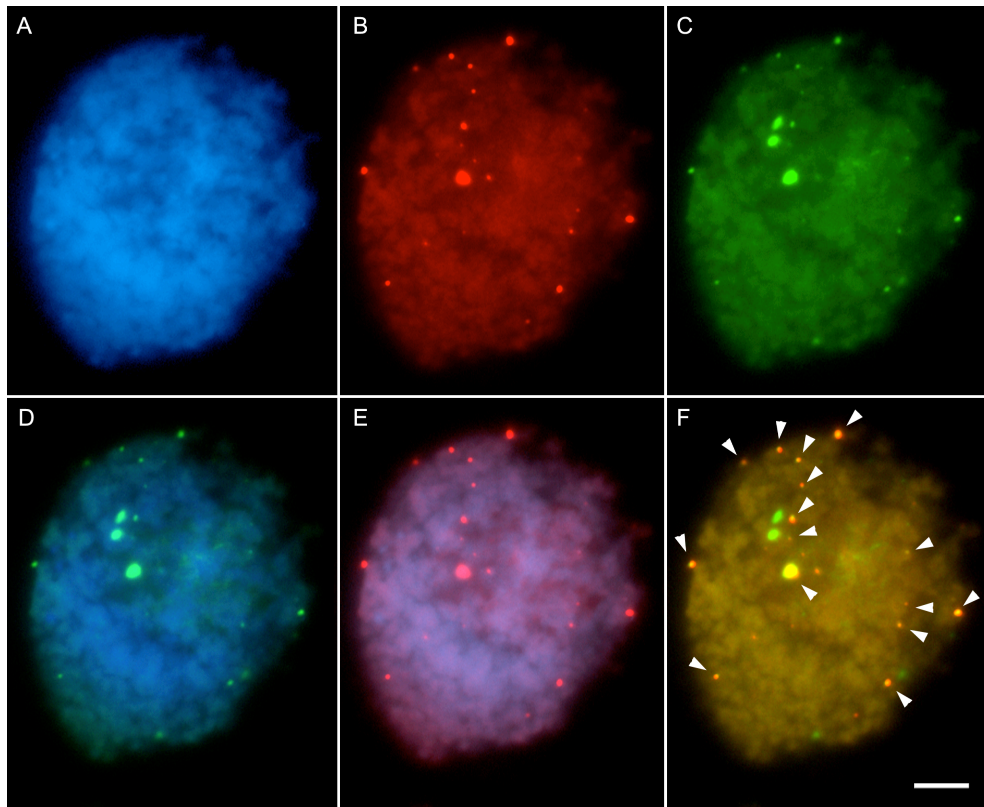

Figure 3.7: FISH of *Allium cepa* var. Zweibel with rDNA (pTA71) and the *Candida albicans* telomere repeat
Allium cepa var. Zweibel root tip cell spreads. (A-F) Interphase spread. (A) DAPI (blue); (B) *C. albicans* telomere repeat (red); (C) 35S rDNA (green); (D) 35S rDNA (green) and DAPI (blue) merged; (E) *C. albicans* telomere repeat (red) and DAPI merged; (F) *C. albicans* telomere repeat (red), 35S rDNA (green) and DAPI merged, arrowheads indicate colocalised signals. Scale bar indicates 5 μm.

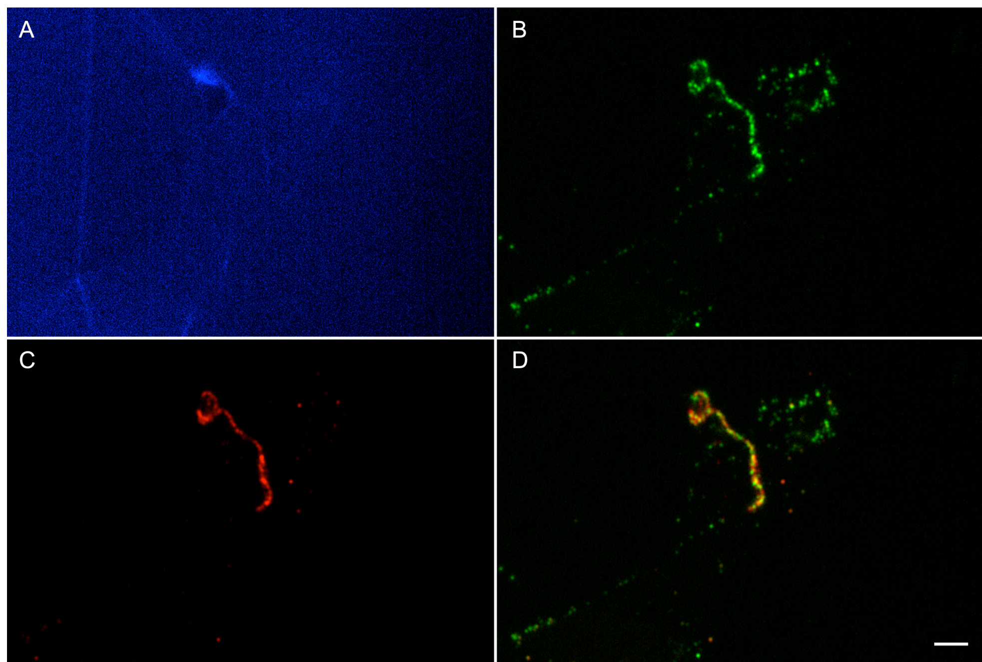

Figure 3.8: EDF-FISH of *Allium cepa* var. Bedfordshire Champion with 18S rDNA and the *Candida albicans* telomere repeat
Allium cepa var. Bedfordshire Champion leaf DNA fibres. (A) DAPI (blue); (B) 18S rDNA (green); (C) *C. albicans* telomere repeat (red); (D) 18S rDNA (green) and *C. albicans* telomere repeat (red) merged. Scale bar indicates 5 μm.

in the genome of *A. cepa*, namely PCR and magnetic bead DNA capture by DNA-DNA hybridisation. Due to the *Allium* sequence probably not having 100% complimentarity with the *C. albicans* telomere, PCRs were carried out using primers against the repeat that were in different phases in order to find a primer pair that could prime successfully at the 3' end. However no products were obtained from PCRs other than weak primer concatenation products. Magnetic bead DNA capture was also unsuccessful at producing sufficient PCR product for cloning.

3.5 Discussion

3.5.1 DOP-PCR Isolated Repetitive DNA

DOP-PCR successfully amplified repetitive DNA from the genome of *A. cernuum*, however the microdissection process was not suffciently effective to enrich terminal sequences, due to the limits on the minimum size that can be dissected. The repeated sequences that were localised using FISH were dispersed throughout the genome, it is likely that these sequences were in high enough copy number within the microdissected chromosome ends to be amplified

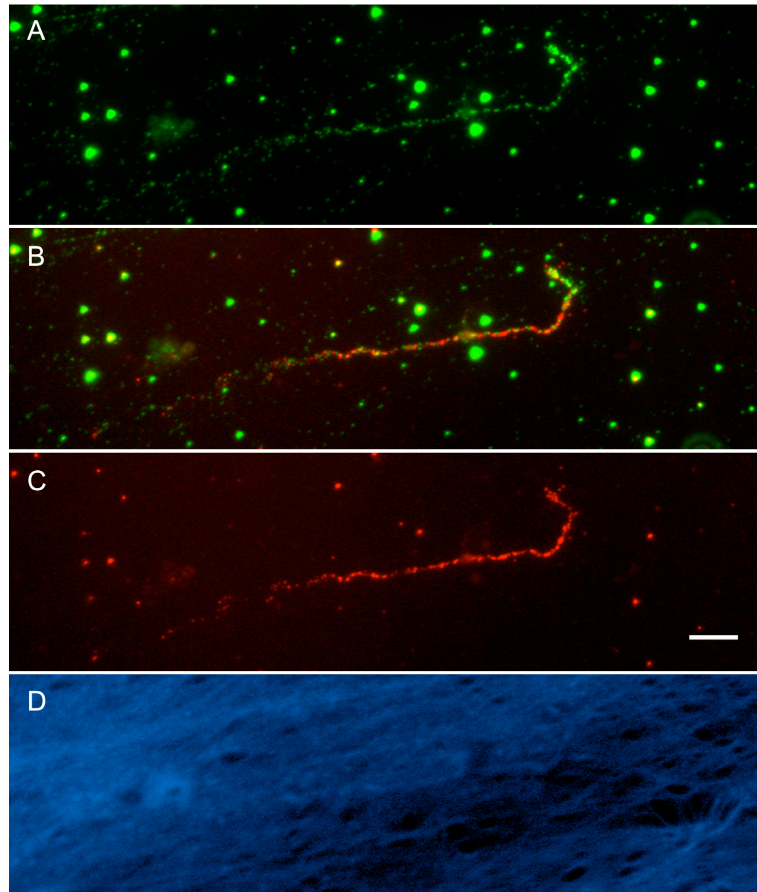

Figure 3.9: EDF-FISH of *Allium cepa* var. Ailsa Craig with 18S rDNA and the *Candida albicans* telomere repeat
Allium cepa var. Ailsa Craig leaf DNA fibres. (A) 18S rDNA (green); (B) 35S rDNA (green) and *C. albicans* telomere repeat (red) merged; (C) *C. albicans* telomere repeat (red); (D) DAPI (blue). Scale bar indicates 5 μm.

by the degenerate oligonucleotide primers. When considering that microdissected fragments were approximately two to four microns in length, based on genome size I estimate that up to 500 Mb of DNA were present in each microdissected chromosome fragment. Furthermore medium to highly repetitive sequences in *A. cepa*, which has a similar genome size to *A. cernuum*, account for almost 80% of the genome (Stack and Comings 1979).

3.5.2 Possible Roles of DNA Secondary Structures

Some of the repeats (cer129, cer540, cer703) isolated from *A. cernuum* by microdissection have the potential to form stable secondary structures given their sequence composition. Secondary structure in a DNA molecule occurs when hydrogen bonds form between adenine-thymine and guanine-cytosine bases, leading to the formation of right-handed helices, paired stems and unpaired loops (Gillespie 2004). DNA secondary structure formation may be the result of selection on its transcript to form secondary structures by forming hydrogen bonds between adenine-thymine, guanine-cytosine and guanine-uracil (RNA) bases. These structures guide the splicing machinery and/or alter mRNA splicing efficiency (Meyer and Miklos 2005). In non-transcribed DNA, secondary structures may be present as they could form aberrant structures when duplex DNA is separated, e.g. during DNA replication (Muller et al. 1999). In bacteria, trinucleotide repeats have been shown to stall replication, requiring processing by the DNA repair machinery, which can lead to repeat expansion and greater instability (Samadashwily et al. 1997).

Secondary structure may have a potential role in recombination based on observations of repeats DNA of rye (*Secale cereale*). In rye there are two major subtelomere repeat families, pSc200 (379 bp) and pSc250 (571 bp), these are found at the subtelomeres of most chromosomes (Vershinin et al. 1995). Molecular analyses revealed that the repeats can be found in uninterrupted blocks of several kb in length and as mixed arrays containing both types of repeats; where blocks of different repeats meet at transition zones they can also be interspersed (Vershinin et al. 1995, Alkhimova et al. 2004). Sequencing of cloned pSc200 and pSc250 spacer regions identified various mobile elements and in one case a spacer sequence which contained several imperfect direct repeats up to 25 bp in length (Alkhimova et al. 2004). Computations of secondary structure of this repetitive spacer and flanking pSc repeats indicate that pSc200 and pSc250 DNA have some complementarity, forming a stem, with most of the spacer itself forming an unpaired loop (Alkhimova et al. 2004). It is thought that a combination of the stem/loop structure and the short subrepeated region within the loop may induce recombination (Alkhimova et al. 2004). The putative secondary structures

in the cer540 and cer703 repeats may perhaps have contributed to their abundance and dispersed distribution around the genome of *A. cernuum.*

3.5.3 Minor 35S rDNA Sites

Using FISH to *A. cepa* chromosomes reconfirmed the presence of four major rDNA sites reported previously (Stack and Comings 1979, Sato 1981, Ricroch et al. 1992). However many additional minor sites were detected at distal positions (fig. 3.4.1). Minor 35S rDNA sites are not unusual and have been reported widely in plants (Pedersen and Linde-Laursen 1994, Dubcovsky and Dvorak 1995, Ali et al. 2005, Vaio et al. 2005, Chung et al. 2008). The *A. cepa* cultivar analysed here has the ACSAT repeat at a subtelomeric location on all chromosome ends, with the exceptions of the NOR bearing short arms of chromosomes VI and VIII, ACSAT was subtelomeric with respect to the minor rDNA loci (fig. 3.4.1). Pich, Fuchs and Schubert (1996) found that in some individuals of *A. cepa* x *A. fistulosum* pseudodiploid hybrids (strain 413) the satellite repeat can occur on the short arm of chromosome VIII both distal and subtelomeric with respect to the NOR position. It is possible that this distal satellite repeat arose through homeologous or non-homologous recombination with a *A. fistulosum* chromsome in the F1 hybrid line. In light of the observations reported here and previously (Pich and Schubert 1998), it seems likely that ACSAT is a subtelomeric repeat especially as it shows properties common to other satellite repeats including typical nucleosome positioning signals and more compaction relative to bulk chromatin (Fajkus et al. 2005). Accordingly, the minor rDNA sites were more distal than ACSAT, making 35S rDNA a useful marker for chromosome ends and potentially useful for isolating telomeric sequences.

3.5.4 35S rDNA Mobility

It was shown that within and between individual *A. cepa* plants that the NORs/major 35S rDNA loci can show apparent position variability amongst these four chromosomes and probably several other chromosomes that do not normally bear NORs (Schubert et al. 1983, Schubert 1984, Schubert and Wobus 1985). This condition was also found within individual pseudodiploid *A. cepa* x *A. fistulosum* hybrids (Schubert and Wobus 1985, Schubert et al. 1983, Schubert 1984). At the time, the authors speculated that this was due to homologous/non-homologous exchange induced by either transposable elements or recombination hotspots. Another possibility considered was that new rDNA loci can arise

from single copy rDNA loci through rapid amplification (Schubert and Wobus 1985). The multiple minor rDNA sites present in *A. cepa*, which were not previously known to exist, give support for either rapid amplification or to homologous/non-homologous chromosomal recombination between rDNA sequences. In both cases new loci could appear or disappear rapidly, giving rise to the rDNA variation observed previously.

In *A. cepa*, total rDNA copy number variation occurs normally during root development (Avanzi et al. 1973, Durante et al. 1977, 1990), therefore both the observed position variability and copy number variation may not require meiotic recombination. Similarily data in studies on asexually reproducing *Daphnia* and triploid *Heteronotia* lizards show that concerted evolution takes place, even between rDNA loci on non-homologous chromosomes (Crease and Lynch 1991, Hillis et al. 1991, Shufran et al. 1997). Within plant cell lineages, intra-array recombination is one process that could result in net rDNA copy number variation between cells. Another possibility is unequal sister chromatid exchange, resulting in differing amounts of rDNA segregating into daughter nuclei. rDNA copy number variation has been shown to occur in many different tissues in *Vicia faba*, varying up to 12-fold within some individuals (Rogers and Bendich 1987). In the case of *V. faba*, there is only one pair of homologous 35S rDNA loci (Knälmann and Burger 1977), so in this species only recombination within or between homologous rDNA loci is possible (Rogers and Bendich 1987).

In species of the tribe Triticeae (order Poales), NORs show position variability between species whilst neighbouring genes show a conservation of synteny. From this data Dubcovsky and Dvorak (1995) suggested that rDNA loci could be translocated or amplified from minor rDNA loci, in processes that do not affect the overall chromosome structure. Considering that other studies have also shown mobility in rDNA loci (Shishido et al. 2000, Adams et al. 2000, Datson and Murray 2006, Cai et al. 2006) whilst the overall chromosome structure remains unchanged, it is feasible that the intraindividual (Schubert and Wobus 1985, Dubcovsky and Dvorak 1995) and interindividual variation (Schubert and Wobus 1985, Garrido-Ramos et al. 1992, Garrido et al. 1994) are the result of similar processes.

Allium somatic cells are known to exhibit the Rabl configuration (Rabl 1885), where telomeres and centromeres are polarised at opposite hemispheres of the nucleus (Fussell 1975, Roy and Ghosh 1977). Furthermore, the somatic chromosome organisation is maintained by attachment of telomeric heterochromatin to the nuclear envelope (Roy and Ghosh 1977, Fussell 1992). It is easy to envisage how this could facilitate mitotic homologous/non-homologous rDNA recombination, if termini are brought into proximity with each other

at the nuclear periphery. In the replicated interphase nucleus shown in fig. 3.4.1, some chromosome ends can be seen in close association. Another structure that brings together active rDNA loci is the nucleolus. Nucleoli have also been suggested as sites of somatic rDNA recombination, possibly as DNA is decondensed and undermethylated (Lim et al. 2000). The apparent end-to-end fusion of different chromosomes at the NORs, at metaphase, has been observed previously in *Allium* species (Stack and Clarke 1973, Schubert 1984, Fussell 1977) and may also be due to events in the nucleolus (Maggini et al. 1978), since resolution of nucleolar recombination events may result in ligated chromosome termini.

In summary: (1) the observed rDNA position variability reported previously could be due either genuine NOR mobility through unequal exchange or it may simply appear as mobility, due to rapid rDNA array expansion and contraction, both models require the presence of minor rDNA sites reported here; (2) the NOR "jumping" is occurring in somatic root tip cells (Schubert and Wobus 1985), so meiotic recombination is not strictly required.

Frequent recombination occurring at distal chromosome positions leading to rDNA variability in *Allium* may affect processes that maintain telomeres (Pich, Fuchs and Schubert 1996, Pich and Schubert 1998, Fajkus et al. 2007), this is discussed in the following section.

3.5.5 Links between Telomeres and rDNA

It has not been possible to determine how close the minor rDNA sites are to the chromosome ends. However, in the case of *A. cepa*, minor rDNA sites are distal to the subtelomere repeat, ACSAT (fig. 3.4.1), and appear to be interdispersed with a sequence, which is similar to the telomere of *C. albicans* (fig. 3.4.4 and fig. 3.4.4). There are various examples where 35S rDNA appears to exert effects on adjacent sequences and chromosome stability in general, these are reviewed below.

Sequences which flank the NORs distally on the five human acrocentric chromosomes are very similar, indicating that homogenisation between non-homologous chromosomes also takes place in the regions flanking the rDNA arrays (Arnheim et al. 1980, Worton et al. 1988, Gonzalez and Sylvester 2001). Similar sequences are also found in other primates (Gonzalez and Sylvester 1997a). An unrelated sequence containing dispersed and low-copy DNA sequences is located proximal to all NORs, as well as some other non-rDNA sites (Bodega et al. 2006). These findings suggest that gene conversion may potentially run-on into homologous sequences flanking the NORs. It is thought that these sequences were initially dispersed by "duplicative transposition events" along with the rDNA during primate divergence (Samonte and Eichler 2002, Bodega et al. 2006). The remains of the NOR

proximal sequence on other chromosomes is thought to mark sites where NORs existed previously, based on ancestral karyotype construction, interestingly these occur mostly at centromeric locations implicating NORs in chromosome fusion events (Bodega et al. 2006).

3.5.5.1 rDNA as a Recombination Hot Spot

rDNA could be harnessed for telomere maintenance as gene conversion mechanisms may affect neighbouring sequences such as the telomere. In some organisms rDNA is known to be closely associated with some or all of the telomeres, as e.g. in *Guillardia theta* (Zauner et al. 2000), *Dictyostelium discoideum* (Emery and Weiner 1981), *Encephalitozoon cuniculi* (Brugere et al. 2000) and *Giardia lamblia* (Adam et al. 1991). In *Guillardia theta* the terminal 13 kb region that includes an rDNA cistron, is identical between chromosome ends, suggesting that these regions are maintained by homogenisation mechanisms (Zauner et al. 2000). Similarly in *Dictyostelium discoideum*, the single rDNA units that are adjacent to a terminal AG_{1-8} repeat at each chromosome end are highly similar or identical, indicating frequent exchange between homologous and non-homologous loci (Eichinger et al. 2005). *In Giardia lamblia* the chromosome ends that have subtelomeric rDNA undergo frequent meiotic recombination, leading to variability in cistron copy number (Le Blancq et al. 1991, Adam 1992). The presence of rDNA adjacent to the telomere may be more important in cases where telomerase is not available to lengthen chromosome termini. This may well be the case in *Dictyostelium discoideum* where the variable terminal $AG_{(1-8)}$ repeat is unlikely to be synthesised from an RNA template and furthermore the gene resembling TERT appears non-functional based on its putative amino acid sequence (Fajkus et al. 2007).

3.5.5.2 Interstitial NORs

In organisms with more than a single pair of homologous NORs it may be beneficial to have 35S rDNA at subterminal sites. This is because translocations between non-homologous chromosomes may arise if an odd number of crossovers occur within interstitial NORs between non-homologous chromosomes. However, if NORs are positioned adjacent to the telomere (subterminally), an odd number of crossovers between NORs on non-homologous chromosomes would limit the exchange of DNA to only 35S rDNA and telomeric sequences. Where NOR position has been finely mapped in plants, minimal amounts of DNA separates the rDNA array from the telomere. In the genome of *A. thaliana,* the telomere on the short arm of chromosome IV joins the rDNA in the IGS downstream of the transcription initiation site of the first rDNA unit (Copenhaver and Pikaard 1996). The junction is made up of two

13 bp head to tail repeats (Copenhaver and Pikaard 1996). In rice, the NOR on the short arm of chromosome IX is separated by only 53 bp of DNA (consisting mostly of inverted telomere repeats) before adjoining the telomere (Fujisawa et al. 2006).

Molecular data of this sort on the size of rDNA-telomere junctions in other plants is not available, but it has been noted that 35S rDNA is often found at terminal or subterminal chromosome positions (Zhdanova, Rubtsov and Minina 2007, Raskina et al. 2008). A recent review of the Brassicaceae family by Ali et al. (2005) revealed that out of 45 species examined only eight species had interstitial 35S rDNA. In one of these eight species, *Brassica oleracea*, there are three rDNA loci, with two loci being distal and the third locus being interstitial. In different accessions the third locus size is either small, minor or undetectable (Maluszynska and Heslop-Harrison 1993, Snowdon et al. 1997, Armstrong et al. 1998, Hasterok et al. 2005). Another three species of the eight with interstitial sites are recently derived species (Yang et al. 2002), with *Brassica juncea* (AABB, 2n = 36) and *Brassica napus* (AACC, 2n = 38) being allopolyploids derived in part from *Brassica rapa* [syn. *campestris*] (AA, 2n = 20) (Maluszynska and Heslop-Harrison 1993, Snowdon et al. 1997, Armstrong et al. 1998, Hasterok et al. 2005). In the case of *Thlaspi arvense* there is only a single pair of homologous rDNA loci (Ali et al. 2005), so rDNA recombination between non-homologous chromsomes cannot occur. With the remaining three species it is not known how old they are, but I would predict that interstitial 35S rDNA loci in these species (*Conringia orientalis*, *Raphanus sativus* and *Arabis alpina* (Ali et al. 2005)) have been recently acquired.

3.5.5.3 rDNA Involvement in Chromosome Fission and Fusion

In a few cases there are tentative links between 35S rDNA amplification and chromosome fission, and it is possible that rDNA expansion aided chromosome end stabilisation. In *Hypochoeris radicata*, the chromosome number is normally $2n = 8$, however in a single plant, a spontaneous centromeric fission of one homologue of chromosome I resulted in a $2n = 9$ complement (Hall and Parker 1995). Interestingly, NORs were found at the fission ends on both the resulting acrocentric chromosomes, but no rDNA was detected on the intact chromosome I. The authors suggested three possible scenarios for its occurrence: (1) rDNA transfer into the centromere caused the fission; (2) rDNA transfer occurred after the fission; (3) rDNA already present at the centromere in low copy number underwent expansion following fission. A similar situation has been reported in the sawfly, *Neodiprion abietis* (Hymenoptera: Diprioninae), and it is likely that this species ($n = 8$) is derived from a ($n = 7$) ancestor (Rousselet et al. 2000). A centric fission of an acrocentric ancestral chromosome

bearing an rDNA site close to the centromere gave rise to two smaller chromosomes, one of which has an rDNA locus at the fission site (Rousselet et al. 2000).

In a group of shrews (genus *Sorex*), repeated metacentric fission and acrocentric fusion events are thought to have a major role in the karyotype evolution of this group (Wojcik and Searle 1988, Ratkiewicz et al. 2002, Wójcik et al. 2002). The karyotype of *Sorex granarius* is as follows: (2n = 36/37, females/males), consisting of 16 acrocentric autosome pairs (designated *a*-*c*, *f*-*r*), one metacentric autosome pair (*tu*), in females this is accompanied by a pair of X chromsomes (*de*) and in males this is accompanied by an X (*de*) and two Y chromosomes, Y_1 (*s*) and Y_2 (*d*) (Zhdanova et al. 2005). Interestingly, distal NORs are found on all short arms of acrocentric chromsomes, sometimes also at the presumed fusion sites on *de* and *tu* chromsomes (Zhadnova et al. 2007). Zhadnova (2007) investigated why all the short arms of acrocentric chromosomes have unusually long telomeres. EDF-FISH showed that these regions contained long (*c.* 200 kb) interdispersed tracts of rDNA and telomere repeats. Also, it is likely that chromsomes *a, b*, *c* and *f* are derived from ancestral metacentrics (*bc* and *af*), with rDNA amplifying at the site of the fission (Fumagalli et al. 1999, Zhadnova et al. 2007). This situation in *Sorex* provides another case of rDNA being involved in telomere amplification and chromosome stability (Zhadnova et al. 2007). The interdispersed distribution of rDNA and telomere repeat is reminiscent of the results shown here in *A. cepa,* of *C. albicans*-like telomeric repeat interdispersed with rDNA (figs. 3.4.4, 3.4.4 and 3.4.4).

3.5.5.4 rDNA behaviour in Telomerase Knock-Out Plants

In *A. thaliana* eighth generation telomerase null mutants, most plants exhibit developmental problems due to insufficient telomere length resulting in chromosome instability (Riha et al. 2001). In mutant plants that did produce infloresences, pistil cells were squashed and analysed by FISH (Siroky et al. 2003). Screening of mitotic anaphases showed that rDNA to rDNA and rDNA to non-rDNA bridges were significantly over represented in anaphase bridges compared to the other DNA sequences (Siroky et al. 2003). The high frequency of fusions involving rDNA could have been due to its abundance and subtelomeric location on chromsomes II and IV and/or some other unknown property that makes it prone to involvement in breakage-fusion-bridge events (Siroky et al. 2003).

3.5.5.5 rDNA and Telomere G-Quadruplexes

A feature shared by rDNA and telomeres is the presence of G-rich DNA that in both sequences may form quartet/quadruplex structures (Zimmerman et al. 1975, Williamson et al. 1989, Sundquist and Klug 1989, Hanakahi et al. 1999). This may be significant as these structures can be formed from one to four DNA strands, and their formation may bring chromosomes into close proximity. Indeed, it has been postulated that this association initiates immunoglobulin gene recombination, through recognition by G quartet binding proteins (Dempsey et al. 1999). One of the proteins implicated in binding G4 DNA is nucleolin, this conserved protein is predominantly localised in the plant nucleolus, as in other eukaryotes (Tong et al. 1997, Petricka and Nelson 2007). In *A. cepa*, nucleolin has been localised by TEM within the DFC (where rDNA transcription is carried out) and at the periphery of the fribrillar center (FC), where it meets the surrounding DFC (Minguez and Moreno Diaz de la Espina 1996). *In vitro*, human nucleolin is able to bind strongly to G-quartet DNA, several runs of three or more G residues are present in the IGS, and are more frequent on the non-transcribed strand than the transcribed DNA strand (Hanakahi et al. 1999). The authors suggested that nucleolin could bind to G-quartet rDNA and have specific functions during strand separation at transcription, replication or recombination (Hanakahi et al. 1999). Many functions have been attributed to nucleolin (reviewed by: Ginisty et al. 1999, Mongelard and Bouvet 2007), including roles in rDNA transcription (Roger et al. 2003, Rickards et al. 2007) and processing (Ginisty et al. 1998); studies in humans have shown nucleolin binding to telomere DNA (Ishikawa et al. 1993, Pollice et al. 2000), telomerase (Khurts et al. 2004) and to topoisomerase I (Bharti et al. 1996). Thus G-quartet formation and associated proteins may link telomere and rDNA biology.

3.5.5.6 rDNA and Telomere DNA Unwinding

Another protein family linking telomere and rDNA biology in yeast (Sun et al. 1999) and humans (Schawalder et al. 2003) are the recQ DNA helicases. The helicase domain is conserved across a wide range of organisms including; *Escherichia coli* (recQ) (Umezu et al. 1990), *Saccharomyces cerevisiae* (SRS2, SGS1) (Rong and Klein 1993, Gangloff et al. 1994) and human (RECQ1, BLM, WRN, RECQ4, RECQ5β) (Ellis et al. 1995, Yu et al. 1996, Kitao et al. 1998, Garcia et al. 2004, Hu et al. 2007). Human premature aging disorders have been attributed to mutations in *BLM*, *WRN* and *RECQ4* causing Bloom syndrome, Werner syndrome and Rothmund-Thompson syndrome respectively (Epstein and Motul-

sky 1996, Yamagata et al. 1998, Sharma et al. 2006). The recQ helicase domain encodes DNA unwinding activity, typically in a 3' to 5' direction, liberating the 3' end of the opposite strands (Umezu et al. 1990). The loss of RecQ proteins leads to genome instability, due to impaired DNA metabolism and hyper-recombination (Gordienko and Rupp 1997, Hanada et al. 1997, Yamagata et al. 1998, Hanada et al. 2000). Reduced lifespan phenotypes have been reported in budding yeast and human RecQ knock-out cell lines (Martin et al. 1970, Sinclair et al. 1997). Activities attributed to RecQ helicases that may have a role in replication/repair/recombination include unwinding duplex, triplex or quadruplex DNA (Shen et al. 1998, Brosh et al. 2001, Wu and Maizels 2001, Huber et al. 2002, 2006). Recombination specific activities include suppressing illegitimate recombination, perhaps by unwinding newly formed Holliday junctions (Bennett et al. 1999) or by separating short paired non-homologous DNA strands (Hanada et al. 2000, Janscak et al. 2003) and assisting branch migration, where DNA strands are swapped between duplexes during Holliday junction movement (Cheok et al. 2005). RecQ helicase SRS2 in budding yeast has been shown to strongly bias the outcome of mitotic homologous recombination events towards gene conversion, rather than crossing over (Ira et al. 2003).

RecQ proteins have a number of activities specific for telomere maintenance including: the removal of G-quartet structures during DNA replication (Azam et al. 2006); resolving recombination intermediates that arise during the replication of telomeric DNA (Lee et al. 2006); and removing illegitimate d-loops, formed by strand invasion by free 3' telomere ends (Orren et al. 2002, Opresko et al. 2004, Bachrati et al. 2006). RecQ helicases are also important in the maintenance of rDNA sequences and have been shown to have specific roles including: aiding the progression of Pol I during transcription (Shiratori et al. 2002), and ensuring proper DNA replication (Kaliraman and Brill 2002). Interestingly, if the RecQ homologue (*RQH1*) in fission yeast is knocked out, this leads to impaired DNA replication and mitotic anaphase bridges, with rDNA being frequently involved (Win et al. 2005). Human BLM protein localises predominantly within the IGS (Schawalder et al. 2003), and budding yeast SGS1 protein preferentially unwinds G-quartet rDNA Sun et al. (1999).

Human WRN protein and its likely orthologue in *Xenopus leavis* (FFA-1) (Yan and Newport 1995, Yan et al. 1998, Chen et al. 2001) are unusual as in addition to the helicase domain, they also have a 3' to 5' DNA exonuclease domain, further increasing functionality (Huang et al. 2000). In *A. thaliana* six RecQ helicase proteins and one homologous WRN-like exonuclease protein are encoded (Hartung et al. 2000).

3.5.5.7 Telomeres and the Nucleolus

The nucleolus itself as well as being the site of rDNA transcription and ribosome production may also harbour molecules associated with telomere biology. Within the nucleolus the following have been found: telomerase RNA component (vertebrates) (Lukowiak et al. 2001, Tomlinson et al. 2006), telomerase protein (TERT) (*Plasmodium falciparum*, human and *A. thaliana*) (Figueiredo et al. 2005, Tomlinson et al. 2006, Kannan et al. 2008), telomeric repeat binding factor 2 (TRF2) (humans) (Zhang et al. 2004). In *Chironomus thummi* and *C. pallidivitatus,* which do not have telomerase maintained minisatellite telomere repeats (Zhang et al. 1994, Lopez et al. 1996, Martinez et al. 2001), a reverse transcriptase protein associated with telomere synthesis in this group has been localised in the nucleolus (Díez et al. 2006). The importance of nucleolar localisation is still not understood, in the case of human TERT, it was shown that deletion of the nucleolar targeting signal does not abolish telomerase function in transformed fibroblast cells (Lin et al. 2008).

3.5.5.8 Summary of Links between rDNA and Telomeres:

It is clear that there are many links between rDNA and telomere biology. Possible reasons for their close physical association are briefly summarised below. The finding that *A. cepa* distal minor rDNA sites are closely associated with a sequence similar to the telomere of *C. albicans* may be a highly significant finding.

- 35S rDNA may be located close to telomeres as a result of the frequent recombination events that are necessary to maintain rDNA homogeneity. It therefore may be advantageous to have NORs located adjacent to telomeres as crossing-over initiated in the rDNA may lead to the translocation of flanking DNA between non-homologous chromosomes.

- 35S rDNA recombination may also have an affect on flanking DNA such as telomeres, allowing sequence homogenisation (e.g. via gene conversion) and/or extension of telomere DNA (e.g. via strand invasion and extension).

- 35S rDNA may also be found close to telomeres as a result of similarities in chromatin structure and proteins required for replication, transcription or repair. If chromosome breakage occurs in rDNA, it may perhaps be more readily stabilised by telomerase due to the localisation of telomere proteins (e.g. telomerase, TRF2, nucleolin) in the nucleolus, where active 35S rDNA is localised.

Chapter 4

Organisation of the Ribosomal DNA Intergenic Spacer in *Allium cernuum* and *Allium cepa*

4.1 Introduction

A consequence of finding the 35S rDNA in many distal chromosome regions of two diverged species of *Allium* and this likely physical association and potentially functional association led to further studies of this gene. With this work I aimed to address two main questions: (1) does rDNA show any unusual features that could support its involvement in telomere function (e.g. the presence of telomere-like repeats within the IGS), (2) does rDNA show variability within the genome of an individual. The latter question could provide a way of isolating distal rDNA units, and a potential means to walk out to telomeric sequences. Within the 35S rDNA cistron, the IGS is the most likely region for harbouring telomere-like repeats and to show intraspecific variability. This region was PCR amplified and cloned from *A. cepa* and *A. cernuum*, although no telomere-like repeats were found in the IGS, there are marked differences in the sequence organisation between species.

4.1.1 IGS Organisation

Within the IGS are a number of features which are required for transcription by RNA-polymerase I (Pol I) into precursor rRNA and its subsequent processing. The transcription

initiation site (TIS), which may be present several times per IGS (Gruendler et al. 1991, Doelling et al. 1993), can be found several hundred to several thousand bases upstream of the 18S coding region (reviewed in: Bena et al. 1998). The TIS consensus TATA(G)TA (the final A usually being the first transcribed base), (Doelling and Pikaard 1995, 1996) appears to be the only part of the promoter that is conserved across plants (Cordesse et al. 1993, Bena et al. 1998, Piller et al. 1990). The minimal promoter sequence has been mapped in *A. thaliana* between positions -55 and -33 to +6 (Doelling and Pikaard 1995). Outside this motif, sequence conservation is low, even when comparing species of the same genus such as *Vicia* (Nickrent and Patrick 1998). Upstream of the TIS there is no evidence in plants for an upstream promoter/control element as has been shown in yeasts, mammals and insects (Doelling et al. 1993, Pikaard 2002).

In fungi and animals, repetitive sequences upstream of the TIS have been demonstrated to act as transcriptional enhancers (reviewed in: Moss and Stefanovsky 1995). In plants, arrays of subrepeats within the IGS are commonly found but in *A. thaliana* subrepeats were shown to only weakly enhance transcription transiently in protoplasts (Doelling et al. 1993) and in *vivo* the absence of subrepeats made no difference to transcriptional activity (Wanzenbock et al. 1997). The IGS subrepeats often vary in number between cistrons of a single genome (Rogers and Bendich 1987, Gruendler et al. 1991, Cordesse et al. 1993, Lakshmikumaran and Negi 1994) and between individuals of a species (Ellis et al. 1984, Flavell et al. 1986, Rogers and Bendich 1987, Lakshmikumaran and Negi 1994, Fukunaga et al. 2005). The transcription termination site (TTS), which can be present several times in an IGS may also act as an enhancer (Zentgraf and Hemleben 1992, Echeverria et al. 1992).

One or several TTSs can be found several tens to hundreds of bases downstream of the 26S coding region, in plants this is probably bound by a protein, to halt RNA Pol I (Zentgraf and Hemleben 1992). There is often an AT-rich region upstream of the TIS (Delcasso-Tremousaygue et al. 1988, Gruendler et al. 1991, Echeverria et al. 1992, Borisjuk and Hemleben 1993, Borisjuk et al. 1997) that may contain both terminators and enhancers (Zentgraf and Hemleben 1992, Echeverria et al. 1992). In pea, upstream of the TIS there is an AT-rich region containing sequences that are similar to the autonomously replicating sequence (ARS) (Hernandez et al. 1993), which in *Saccharomyces cerevisiae* has been demonstrated to be an essential part of the origin of replication (Van Houten and Newlon 1990, Campbell and Newton 1991). In pea there are four sequences, in each of which 9-10 bp match the 11 bp ARS core consensus sequence (ACS) (A/T)TTTAT(A/G)TTT(A/T) of *S. cerevisiae* (Hernandez et al. 1993). A similar 440 bp AT-rich region in the IGS of *Nicotiana*

tabacum contains nine sequences, each of which match 9-11 bp of the ACS (Borisjuk et al. 2000). When this 440 bp region was included in a transformation cassette it was found to induce transgene copy number amplification and transcription in tobacco, thought to be due partly to the presence of several DNA bending motifs and scaffold-attached regions (Borisjuk et al. 2000).

Another feature of the IGS found across eukaryotes is the presence of a replication fork barrier (RFB), leading to rDNA mostly being replicated unidirectionally, in the same orientation as transcription (Brewer and Fangman 1988, Linskens and Huberman 1988, Little et al. 1993, Rothstein et al. 2000). In pea a RFB is found 156 bp downstream of the 26S gene and contains three or nine imperfect 27 bp repeats, depending on the IGS length class (Lopez-Estrano et al. 1999, Hernandez et al. 1993). This sequence frequently stops replication forks moving into the 26S gene, by acting as a binding site for blocking protein(s) (Lopez-Estrano et al. 1999). An interesting model for copy number amplification in genes that are highly transcribed, has been proposed for the yeast rDNA cistron (Kobayashi et al. 1998). This was demonstrated in a *fob1* mutant, where the fork blocking protein which normally binds the RFB was rendered non-functional; the resulting collisions between transcription and replication machinery initiated recombination resulting in copy number variation (Kobayashi et al. 1998, Takeuchi et al. 2003, Kobayashi 2003).

4.1.2 IGS Evolution

When IGS sequences from related species from the same genus have been compared, such as from *Nicotiana* (Borisjuk et al. 1997) and *Brassicaceae* (Da Rocha and Bertrand 1995), subrepeats in the NTS part of the IGS show the most rapid change, whereas the TIS and the region adjoining the 18S region (i.e. 5' ETS) are more conserved (Appels and Dvořák 1982). These interspecific differences in IGS subrepeats highlight the recombinogenic nature of the IGS subrepeats. This can result in both heterogeneity i.e. in species such as *A. thaliana* and rice where some recombinant NTS subrepeat variants have been generated (Gruendler et al. 1991, Cordesse et al. 1993). In other plants, recombination has resulted in subrepeats upstream of the TIS that are very similar, suggesting that repeat homogenisation acts within cistrons (intracistronic) as well as between cistrons (intercistronic) (Dvorak et al. 1987). An analysis of rDNA sequences cloned from species of *Mitella* (Saxifragaceae), showed that polymorphisms arising through introgression were homogenised more rapidly in the ETS than the ITS (Okuyama et al. 2005).

Several phylogenetic analyses of ETS and ITS1/ITS2 sequences from plants show phylo-

genetic concordance, suggesting that at least these regions are evolving as a single unit, albeit with the ETS typically showing greater divergence between species (Baldwin and Markos 1998, Bena et al. 1998, Clevinger and Panero 2000, Linder et al. 2000, Vander Stappen et al. 2003, Okuyama et al. 2005). This may be because transcribed regions of the rDNA cistron could undergo localised recombination and homogenisation separately from non-transcribed regions, due to the displacement of DNA strands during transcription (Appels and Dvořák 1982). Where deeper analyses of an individual's rDNA population have been made, including pseudogenes, there is evidence for recombination in ITS1 and ITS2 between variant rDNA units, generating chimeric/mosaic units (Suh et al. 1993, Wendel et al. 1995, Buckler and Holtsford 1996, Buckler et al. 1997).

Different parts of the IGS region can occur outside the 35S rDNA loci as independent tandem repeats, including IGS subrepeats (Unfried et al. 1991, Nouzova et al. 2001, Macas et al. 2003, Lim et al. 2004), partial subrepeats together with the AT-rich region containing ACS, SAR and DNA bending motifs (Stupar et al. 2002), or the AT-rich region within the TIS (Raina et al. 2005). There are many possible scenarios in which IGS sequences could become liberated from rDNA loci including: by reverse transcription of rRNA transcripts and integration back into the genome, direct translocation by non-homologous recombination from an rDNA array, indirect translocation by integration of an rDNA recombination intermediate or partial deletion(s) of a disjunct 'orphan' rDNA unit (Lohe and Roberts 1990, Maggini et al. 1991, Gonzalez and Sylvester 1997b, Lim et al. 2004, Kovarik et al. 2004). These scenarios would all result in a low number or perhaps a single copy of a sequence inserted into a single homologue of a chromosome (i.e. it is in a hemizygous state), and as a result sister chromatid recombination would not be possible for copy number amplification (Stupar et al. 2002). Therefore recombination with a homologous sequence at a different location in the genome or an extrachromosomal process, such as rolling circle amplification/replication (RCR), would be required for copy number amplification (Stark et al. 1989, Stupar et al. 2002).

To date, in plants, there is only evidence for 5S rDNA, telomere DNA and satellite DNA being found as extrachromosomal circles (ECCs) (Cohen et al. 2008, Navratilova et al. 2008). In *Drosophila melanogaster* DNA circles corresponding to the IGS "240 subrepeat" exist, probably resulting from intracistron recombination (looping-out) (Pont et al. 1987, 1988, Cohen et al. 2003). Also, in *D. melanogaster*, 35S rDNA circles greater than 10 kb have been detected, probably resulting from intercistron recombination (Cohen et al. 2003, 2005). In *Xenopus leavis* whole cistrons are amplified by RCR in early development (Brown

and Dawid 1968, Hourcade et al. 1973, Rochaix et al. 1974), but there is no evidence for rDNA circles being able to replicate autonomously in other organisms. All these processes of rDNA translocation, RCR from ECCs and homologous recombination could potentially be involved in telomere biology.

4.2 Materials and Methods

4.2.1 IGS Amplification and Sequencing

The IGSs were amplified from *A. cepa* cv. Ailsa Craig and *A. cernuum* (0701 QMUL) genomic DNA using the Expand High Fidelity PCR System (Roche) or BIO-X-ACT long DNA polymerase (Bioline, London, UK) respectively with primers, 26S: 5'-GGGAACGTGAGCTG GGTTTAGACCGTC-3' and 18S: 5'-GCCTGCTGCCTTCCTTGGATGTGG-3' as published by Bena et al. (1998). Reactions were performed in 50 µl volumes and were supplemented with 3% (v/v) dimethyl sulphoxide (DMSO) (Sigma-Aldrich). The thermocycler parameters were based on the parameters suggested by Roche: initial denaturation of 94°C for 2 minutes followed by 10 cycles of 94°C for 15 seconds, 56.5°C for 45 seconds, 68°C for 4 minutes this was followed by 18 cycles of 94°C for 15 sec, 56.5°C for 30 sec, 68°C for 4 min and a final extension of 72°C for 7 min.

PCR products were cleaned using PCR purification columns (Qiagen) and cloned into the pCR2.1-TOPO vector (Invitrogen). Colonies were screened by blue-white selection on plates supplemented with ampicillin and X-Gal. Colonies selected for sequencing were cultured overnight in liquid LB medium following section 2.6.2 on page 33. *A. cernuum* IGSs were sequenced commercially by primer walking service (Macrogen, South Korea), some repeat regions were checked by resequencing. *A. cepa* was sequenced by several rounds of primer synthesis and sequencing (MWG Biotech, Germany) using the program Primer3 to design suitable primers (Rozen and Skaletsky 2000). A list of effective primers used to sequence the IGSs of *A. cepa* and *A. cernuum* are listed in tables 4.2.1 and 4.2.1, respectively.

4.2.2 DNA Restriction Digests and Southern Hybridisation

DNA Restriction digests and Southern transfer was carried out following section 2.5.1 on page 31, 2.5.2.

Probes for detection were made as follows: (1) The 26S probe was made by PCR using the IGS clone cerD8 from *A. cernuum* as a template with the following primers: "26S" from

Primer name	Primer sequence (5'-3')
ACIGS1F	TTCATGAGTCGTCCAATCCA
ACIGS2F	AGGTAGGAGCGCAACTTCAA
ACIGS2bF	GGTCATGGAGTAAGCCGAAA
ACIGS3F	TAGAAAACTCGGAACTGGCTTC
ACIGS4F	CCGAGTGTCGGGTCTAGAAAT
ACIGS5F	TGGCAATAAATAGGTCCTCCA
ACIGS6F	GAGAATGATATTTGCCCGGAT
ACIGS1R	AGAATCGAACAATTCCCTCTCA
ACIGS3R	ATCCGGGCAAATATCATTCTC

Table 4.1: *Allium cepa* specific IGS sequencing primers

Primer name	Primer sequence (5'-3')
Acer1F	CTGAGATCCAGCCCTTTGTC
5'Rep_AAA	CGTCTGACCAACCAAGATACAAAA
IGS_3'Rep1	ATACCGTTGTGCCCTTGAAC

Table 4.2: *Allium cernuum* specific IGS resequencing primers

Bena et al. (1998) and 440_26S_rev: 5'-GCGTATTTAAGTCGTCTGCAAAG-3'. (2) The *A. cepa Hin*dIII probe was isolated by *Hin*dIII digestion of a BAC (S1B1) from *A. cepa* cv. Cheonjudaego, provided by G. Suzuki (c.f. Suzuki et al. 2001), which has a 105kb insert containing several copies of the complete 35S rDNA unit. The insert was cloned into pBluescript KS+. Sequencing confirmed that the 1932 bp fragment referred to hereafter as "AcH1932" was identical in sequence to the other IGS sequence amplified from *A. cepa* cv. Ailsa Craig.

The temperatures used for the Southern hybridisation were as follows: AcH1932 probe 62°C, 26S probe 65°C. The membranes were washed at high stringency (0.2x SSC, 0.1% SDS) using the same temperature as used for hybridisation (section 2.5.3).

4.2.3 Nucleotide Sequence Analysis

Gene/spacer boundaries were predicted on the basis of existing rDNA sequences for *Zea mays* (GenBank: X03990.1) and *Citrus limon* (GenBank: X05910.1). IGS sequences were checked for coding and non-coding repeats using RepeatMasker (http://www.repeatmasker .org/). Sequences were checked on GenBank (http://www.ncbi.nlm.nih.gov/blast/) non-redundant nucleotide, non-redundant protein or EST databases using BLASTN, TBLASTX and BLASTP (Altschul et al. 1997, 1990). Repetitive sequences within the IGS were found using the dotplot and BLAST output from PipMaker (Schwartz et al. 2000) and confirmed by aligning subrepeats.

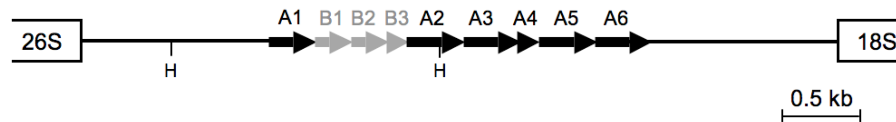

Figure 4.1: *Allium cepa* IGS structure
rDNA IGS structure in *A. cepa*:the two major subrepeat families are indicated with black arrows (A-subrepeats) or grey arrows (B-subrepeats). *Hin*dIII restriction sites are indicated (H), this fragment corresponds to the AcH1932 probe.

4.3 Results

4.3.1 IGS Structure of *Allium cepa*

DNA sequence analysis indicated that subrepeats are located in one contiguous cluster towards the middle of the IGS. Two types of imperfect repeats were found, A and B, with the B subrepeats embedded within the A-subrepeats (fig. 4.3.1).

An alignment of the six A-subrepeat monomers indicated the presence of internal and 3' deletions relative to the longest monomer A2 (fig. 6.1.2). A large 3' deletion was found in monomer A1, where there is an abrupt transition to the B-subrepeats, another 3' deletion was found in monomer A6, the final A-subrepeat. There are three B-subrepeats downstream of monomer A1. An alignment of those three B-subrepeats revealed a large deletion in the final B-monomer where there is an abrupt transition back to the A-subrepeats (fig. 6.1.2).

4.3.2 IGS Structure of *Allium cernuum*

High fidelity PCR amplification of the IGS generated PCR products ranging from 0.7 kb to 6 kb in length. The complete mixture of amplification products was cloned and 5 variants were sequenced to see whether particular regions were prone to length variability. A diagrammatic summary of four IGS sequences (sizes: 0.7, 1.9, 4.1, 4.1 kb) is shown (4.3.2), the two clones of 0.7 kb had identical sequences. IGSs are divided into class I and II subfamilies based on the absence or presence, respectively, of a 363 bp region similar to part of a Ty1/*copia* element.

Two domains of repeated sequence are present in the IGS. The first domain contains 257 bp A-subrepeats arranged head to tail in tandem (alignment in appendix 6.1.2), the second complex repeated domain contains three subrepeats in various organisations. The average pairwise similarity is above 99% between clones within the sequenced 26S and 18S

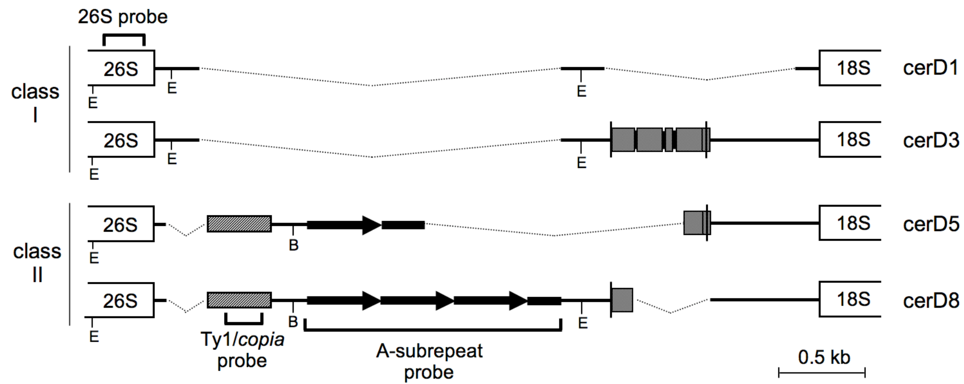

Figure 4.2: *Allium cernuum* IGS structure
Diagram of four IGS sequences from *A. cernuum* with regions of high similarity aligned, gaps are shown as dotted lines (clones: cerD1, cerD3, cerD5, cerD8). Two sequences contain a partial putative Ty1/*copia* ORF; this is represented as a hashed box downstream of the 26S gene. Two repetitive domains are present; the 457 bp A-subrepeat domain (black arrows, where partial arrows represent different sized 5' truncations) and the complex repeated domain (vertical lines (12 bp), grey boxes (9-160 bp) and black boxes (23 bp)). *Eco*RI and *Bsr*DI restriction sites are represented by sites marked E and B.

coding regions. Within the IGS, clones have low levels of sequence divergence due to base substitutions, although there are differences in length due to insertions/deletions (indels) involving regions with both repeated and unique sequences.

4.3.3 Overview of the *Allium* IGS

The 26S and 18S coding regions of the *A. cepa* IGS and longest *A. cernuum* IGS (cerD8) sequenced showed 89% and 95% similarity respectively. The only alignable part of the IGS between species was in the region directly upstream of the 18S gene between the *A. cepa* clone and the three longest *A. cernuum* clones (cerD3, cerD5, cerD8) showing 72% similarity. The length was 554 bp in *A. cepa* and 563 bp in the three *A. cernuum* IGSs, it is likely that much of this conserved 3' IGS region corresponds to the ETS, previous reports from plants have found the length to be between *c.* 0.5-5 kb (reviewed by Bena et al. 1998). Within this conserved 3' IGS region we searched for the TATA(G)TA motif of the TIS that is thought to be the conserved in plants (Cordesse et al. 1993, Bena et al. 1998, Piller et al. 1990), however in both species it is absent. When the search is widened to include the whole IGS region, in *A. cernuum* there are two TATATA motifs in each complete A-subrepeat and one motif in the truncated A-subrepeats. An additional TATATA motif was found within the Ty1/*copia* sequence. Typically, a run of four to six guanines are present immediately downstream of the TIS consensus (Doelling and Pikaard 1995, Bena et al. 1998), these were

not found next to any of the TATATA motifs. In *A. cepa* the closest match to the TIS consensus when including the subrepeats and the non-repeated sequence downstream, was a TATATC motif, with the cytosine located 623 bp from the start of the 18S coding region.

There is a sequence conserved in the IGS between the two *Allium* species, with the consensus TGAGTGGT(C/T), of which the first 8 residues match a conserved 9 bp ETS element reported previously for several plants (TGAGT(G/T)GTA) (Bena et al. 1998).

4.3.4 Southern Hybridisation of *Allium* species

A 1932 bp partial IGS probe (called AcH1932) was made by digesting *A. cepa* BAC S1B1 with *Hin*dIII, the restriction sites are shown in fig. 4.3.1. A 1.9 kb band corresponding to this part of the *A. cepa* IGS can be seen in fig. 4.3.4, this fragment could be detected with both the S1B1 probe containing several entire 35S rDNA units and with the *Hin*dIII fragment itself. When the *Hin*dIII fragment was used as a probe (AcH1932), a higher molecular weight band probably corresponding to entire units was detected. As the upper band is about 2 kb longer than the main 10 kb fragment, this is probably due to units only being cut once in the IGS, perhaps due to point mutations in the *Hin*dIII sites.

A. cernuum genomic DNA digested with *Hin*dIII produced a slightly diffuse band when probed with the 26S probe, there was also a weaker smear between approximately five and 12 kb (fig. 4.3.4). This result indicates that *A. cernuum* rDNA is cut once per rDNA unit based on the S1B1 probe showing only one band. It is unlikely that small fragments of the spacer regions were liberated, as sequencing of the IGS and ITS (not shown) revealed no *Hin*dIII sites in *A. cernuum*. If fragments of genic regions had been cut out this would have been detectable with the S1B1 probe. The 26S probe labels a prominent fragment size of approximately 12 kb (arrowed fig.4.3.4) indicating that the most rDNA units exhibit a small amount of length variation. A weaker background smear in the size range of 5-13 kb (indicated by a 'bar' in fig. 4.3.4) indicates a small proportion of rDNA units show a large amount of length variation.

The 1932 bp *Hin*dIII IGS fragment from *A. cepa* was used as a probe (AcH1932) to examine the conservation of the IGS across *Allium* (fig. 4.3.4; for more distantly related species, data not shown). The *Hin*dIII fragment contained 1280 bp of non-repetitive DNA upstream of the subrepeats and 652 bp of the subrepeat region itself (containing the subrepeats A1, B1, B2, B3 and part of A2). Genomic DNAs from 33 species from across the genus *Allium* were probed with AcH1932 from *A. cepa* and gave the following results based on a recent classification of the genus by Friesen et al. (2006): (i) strong signal was revealed

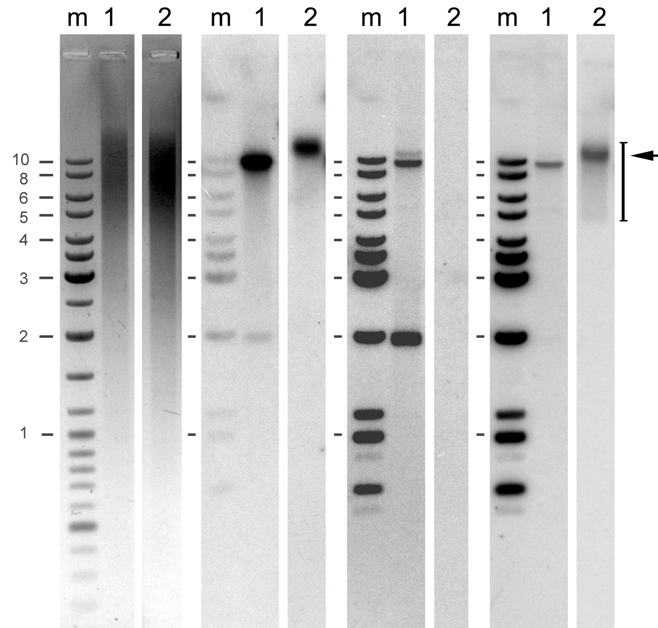

Figure 4.3: Southern hybridisations of S1B1 BAC, AcH1932 probe and 26S rDNA to *A. cepa* and *A. cernuum*
Ethidium bromide stained agarose gel (left most set) and Southern hybridisations (three remaining sets) of *Hin*dIII digested genomic DNA of *A. cepa* (1) and *A. cernuum* (2). The nylon membrane was successively re-probed with the following from left to right: the S1B1 BAC (contains several entire 35S rDNA units from *A. cepa*) (high stringency); AcH1932 probe from *A. cepa* (low stringency wash) and the 26S genic probe (high stringency wash). m-marker, with sizes shown in kbp. Arrow indicates the main rDNA unit size in *A. cernuum*, bar indicates the extent of rDNA unit length variation above and below the main rDNA unit size in *A. cernuum*. (Southern hybridisation carried out by E. Sýkorová.)

at high stringency for all investigated species from subgenus *Cepa* and including *A. roylei* (ii) *A. schoenoprasum*, *A. chinense* from subgenus *Cepa* and *A. splendens* from subgenus *Reticulatobulbosa* gave a weak signal in low stringency conditions; (iii) 24 other species gave no signal (see table 6.1.2 on page 106 for a summary).

4.4 Discussion

Within the sequenced IGSs of *A. cepa* and *A. cernuum* there are no motifs similar to known minisatellite telomere sequences. Such sequences might have provided evidence of direct binding site for telomere proteins. There are several findings in terms of intraspecific variability that have not been previously reported in plants.

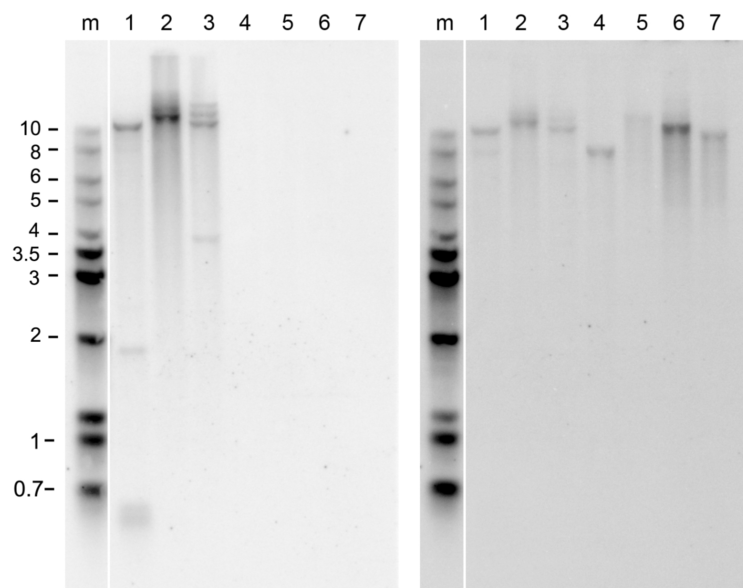

Figure 4.4: Southern hybridisation of AcH1932 probe and 26S rDNA to *Allium* species Southern hybridisation of *Hin*dIII digested *Allium* genomic DNA probed with AcH1932 with low stringency washing (left) and re-probed with the 26S genic probe with high stringency washing (right). 1, *A. galanthum* (subgenus *Cepa*); 2, *A. altaicum* (subgenus *Cepa*); 3, *A. roylei* (subgenus *Cepa*); 4, *A. mairei* (subgenus *Cyathophora*); 5, *A. tuberosum* (subgenus *Butomissa*); 6, *A. cernuum* (subgenus *Amerallium*); 7, *A. neapolitanum* (subgenus *Amerallium*); m, marker with sizes shown in kbp. (Southern hybridisation carried out by E. Sýkorová.)

4.4.1 *Allium* 35S rDNA IGS

The sequencing data shows that although *A. cernuum* and *A. cepa* diverged early on in the radiation of *Allium* (Friesen et al. 2006), as expected they still retain extensive homology in the 18S and 26S coding regions. This conservation is lost within a few bases at the beginning of the IGS and no similarity is detectable until *c.* 0.5 kb upstream of the 18S region. The extensive molecular divergence generally agrees with the high divergence reported for the ITS regions, where pairwise genetic distances (Kimura 1980) between *Allium* species are as high as 53% compared to a typical value of 10% for species of other plant genera (Klaas and Friesen 2002). One possible reason for this unusually high divergence is the age of the genus (Dubouzet and Shinoda 1999, Klaas and Friesen 2002).

In terms of intragenomic variability the two species analysed show striking differences. In *A. cepa* a single IGS type is present in the majority of rDNA cistrons, this was confirmed by the low variation in fragments generated by restriction digestion and probing with the *Hin*dIII fragment (fig. 4.3.4). The larger smear generated by probing with the S1B1 probe may be due to some cross-hybridisation to similar sequences in the genome, although these were not detected by FISH. In contrast, within the genome of *A. cernuum* rDNA units show length variation based on southern hybridisation data (fig.4.3.4). Specifically, heterogeneity was observed at three levels in the IGSs sequenced: (1) IGS subrepeat copy number variability, (2) the absence or presence of a truncated retroelement and (3) the presence or absence of indels. Extensive IGS subrepeat variability has been found in a variety of plants most notably in *Oryza sativa* (Cordesse et al. 1993), *Avena sativa* (Polanco and Perez de la Vega 1997), *A. thaliana* (Copenhaver and Pikaard 1996) and *Setaria italica* (Fukunaga et al. 2005). Previous studies on three *Allium* species have shown that pairs of rDNA IGS length variants can exist within individuals (Friesen et al. 1997). This is the first case of a Ty1/*copia* related sequence being amplified within eukaryote rDNA, this sequence is analysed in detail in chapter 5.

Most of the short rDNA units released by *Hin*dIII digestion, visible as a smear in fig. 4.3.4, can be attributed to units with short IGSs as was found in the sequencing data (fig. 4.3.2). However, the region spanning from the 18S through to the 26S coding region in rice is 5.8 kb long (Takaiwa et al. 1984, 1985, 1990), this region is likely to be of a similar size in *A. cernuum*. Therefore any fragments that are shorter than *c.* 5.8 kb in length, are likely to be from rDNA units that are some missing parts of the coding regions as well as the IGS.

Importantly, *A. cernuum* rDNA units with longer (6-7 kb) IGSs account for the majority

of rDNA units visible as the prominent *c.* 13 kb fragment hybridised by the 26S probe (arrowed in fig. 4.3.4). Shorter IGS sequences were probably preferentially amplified by polymerase enzymes *in vitro* as they are shorter, and may perhaps lack certain GC-rich tracts and/or regions forming highly stable secondary structures. It is likely that these longer IGSs have the full complement of rDNA subrepeats of which different parts were identified in different IGS clones. Based on the shortfall between the maximum lengths of IGS sequenced (3.4 kb IGS in clone cerD8 or 4.1 kb of non-overlapping sequence in the IGS alignment fig. 4.3.2) and the length of the typical IGS based on restriction digestion (6-7 kb), some parts of the IGS remain unsequenced.

In *A. cepa* although the A-subrepeats show internal and 3' deletions, and are interrupted by B-subrepeats, these are perhaps the result of unequal recombination events between units. However this structure arose, it is clear that efficient homogenisation processes have led to this IGS type being spread close to fixation with the genome. It is possible that the onion domestication process has contributed to the observed homogeneity in the rDNA, as less defined bands were revealed by hybridisation in other species analysed (fig. 4.3.4). This result is interesting in light of the unusual organisation at many chromosome ends in *A. cepa*, where there appears to be inter dispersion of 35S rDNA units and telomere-like sequences resembling the *C. albicans* telomere (see 3.4.4). This would suggest that either 35S rDNA can be kept homogenous even when not in a contiguous tandem arrangement, or that the interspersion has occurred recently and so mutations have not accumulated. Another posssibility is that rDNA units present on minor sites are in a low copy number and so may not been detected by Southern hybridisation.

Chapter 5

Amplification of a Retrotransposon Fragment within the rDNA of *Allium cernuum*

5.1 Introduction

The rDNA of *A. cernuum* shows unusually high levels of length variability in the IGS region, which is in part due to the presence of a partial retrotransposon coding sequence. As this phenomenon has never been shown before in plants, the rDNA of *A. cernuum* was studied in detail with the aim of understanding how this IGS sequence became integrated and subsequently evolved. In another species, *A. cepa*, it has been shown previously using FISH that some Ty1/*copia* retrotransposons are concentrated at the chromosome termini, suggesting that they may be involved in *Allium* telomere biology.

5.1.1 rDNA Homogeneity

Typically rDNA units show low levels of genetic variation within and between individuals of the same species (Hillis and Dixon 1991, Ganley and Kobayashi 2007, Stage and Eickbush 2007) but there are usually fixed differences between species (Brown et al. 1972). This is due to rDNA units evolve by concerted evolution, where homogenisation mechanisms act to reduce levels of genetic variation, giving the appearance of all units evolving in a unified manner (Arnheim 1983). These mechanisms can include: unequal crossing-over (Smith

1976, Szostak and Wu 1980), gene conversion (Gangloff et al. 1996, Ganley and Scott 1998), and copy number expansion/contraction (Tartof 1974, Petes 1980, Dover and Coen 1981, Kobayashi and Ganley 2005). In organisms that have more than one rDNA locus, these processes must occur between both homologous and non-homologous NOR loci, in order to homogenise all members of the genome's rDNA population (Arnheim et al. 1980). In some recently formed allopolyploid plants, it has been observed that only a few generations generate a homogenous pool of several thousand rDNA units (Skalická et al. 2003, Kovarik et al. 2005). In other cases rDNA homogenisation is not so apparent and divergent rDNA paralogues persist (Brownell et al. 1983, Buckler et al. 1997, Gonzalez and Sylvester 2001, Keller et al. 2006, Zheng et al. 2008). The efficiency of concerted evolution in eukaryotes may be affected by factors including: NOR position on the chromosome (Zhang and Sang 1999), generation time (Sang et al. 1995, Richardson et al. 2001), genome size (Keller et al. 2006), frequency of hybridisation (Keller et al. 2008), epigenetic state (Lim et al. 2000) and corresponding transcriptional activity (Jupe and Zimmer 1993, Koukalova et al. 2005, Dadejova et al. 2007).

5.1.2 Targeted Insertion of Non-LTR Retroelements into the 26S/28S rDNA Unit

Large variation between rDNA units can be brought about by the insertion of mobile genetic elements. The best documented are non-long terminal repeat (non-LTR) retrotransposons called *R* elements that are an ancient group found in many animal lineages (Jakubczak et al. 1991, Kojima and Fujiwara 2005, Kojima et al. 2006). Mobile elements that insert into coding regions will render the rDNA unit non-functional, so there is a necessity to continually remove these variants from the rDNA population as they arise (Zhang et al. 2008).

5.2 Materials and Methods

5.2.1 26S-IGS Sequencing

To sample intragenomic diversity within *A. cernuum* (0701 QMUL) in a 1117-1118 bp region containing the final *c.* 454 bp of the 26S gene and the adjacent 3' IGS region containing the Ty1/*copia*-like sequence. A PCR was carried out using the forward primer 26S: 5'-GGGAACGTGAGCTGGGTTTAGACCGTC-3' from Bena et al. (1998) and a re-

verse primer, designed from sequenced *A. cernuum* IGS clones, at the 5' boundary of the A-subrepeats 26Scop3R: 5'-TTTTTGGATTAGACTTTTGTATCTTGG-3'. The amplification was carried out in a 25 µl reaction using *A. cernuum* (0701 QMUL) genomic DNA as a template, using BIOTAQ polymerase (Bioline), including 3% (v/v) DMSO (Sigma-Aldrich) (2.4.1). The thermocycling parameters were as follows: initial denaturation step 94°C for 3 minutes followed by 30 cycles of 94°C for 40 seconds, 55°C for 30 seconds, 72°C for 1 minute; this was followed by a final extension step of 72°C for 7 min.

PCR products were cleaned using a QIAquick PCR Purification Kit (Qiagen GmbH, Germany) and cloned into the pCR2.1-TOPO vector (Invitrogen) (following 2.6.1). Colonies were screened by blue-white selection on ampicillin and X-Gal supplemented LB plates. Following selection, cells were cultured overnight in liquid LB medium and plasmids were isolated using a QIAprep Spin Miniprep Kit (Qiagen) (following 2.6.2).

5.2.2 DNA Sequence Analysis

The 363 bp sequence identified as similar to a Ty1/*copia* element was checked for similar sequences in the NCBI non-redundant protein database (http://www.ncbi.nlm.nih.gov/blast/) using BLASTX (Altschul et al. 1997). DNA sequences of retrotransposons were aligned manually using the nucleotide with translated amino acid view in MacClade 4.08 (Maddison and Maddison 2000).

An alignment of 37 intragenomic rDNA sequences was made after removing vector sequences and priming sites, leaving sequences of 1064-1065 bp. Due to the high similarity, an alignment was made by eye. This was used to find the shortest maximum parsimony tree using PAUP version 4.10b (Swofford 2002). A heuristic search was carried out with 1000 replicates of random sequence addition and the tree bisection-reconnection branch-swapping algorithm. The same alignment was used to make a median joining network using the program Network version 4.50.1 (Bandelt et al. 1999) (www.fluxus-engineering.com), using an equal weighting on all characters and an epsilon value of 30. The average pairwise nucleotide diversity, π, (Nei 1987) along the intragenomic rDNA alignment was calculated with the program DnaSP version 4.50.2 (Rozas et al. 2003), with a sliding window of size of 25 bp and step size of 10 bp.

5.2.3 Southern Hybridisation Analysis

DNA Restriction digests, Southern transfer and DNA hybridisation was carried out following methods in section 2.5.1 and 2.5.2.

Probes for detection were made as follows: (1) The 26S probe was made by PCR using the IGS clone cerD8 from *A. cernuum* as a template with the following primers, 26S: 5'-GGGAACGTGAGCTGGGTTTAGACCGTC-3' from Bena et al. (1998) and 440_26S_rev: 5'-GCGTATTTAAGTCGTCTGCAAAG-3'. (2) A probe containing 338 bp of the Ty1/*copia*-like sequence was made by PCR from a diluted plasmid (clone cerD8) of an *A. cernuum* IGS sequence using the forward primer 26Scop1F: 5'-TAGATTTCCTGGTGGCCTTG-3' and the reverse primer 26Scop2R: 5'-GGACTTCAAACGATGGGAGA-3'.

The temperatures used for the Southern hybridisation were as follows: 26S probe at 65°C, Ty1/*copia* probe at 55°C. The membranes were washed at high stringency (0.2x SSC, 0.1% SDS) for the 26S probe or low stringency (2x SSC, 0.1% SDS) for the Ty1/*copia* probe, using the same temperature as used for hybridisation.

5.2.4 FISH

DNA templates for making FISH probes were isolated and labelled as follows:

(1) The 18S rDNA probe was isolated by PCR from *A. cernuum* genomic DNA using the forward primer A.18S2F: 5'-CGGAGAATTAGGGTTCGATTC-3' and the reverse primer 17_rev_SE: 5'-ACGAATTCATGGTCCGGTGAAGTGTTCG-3', which is derived from the reverse complement of primer 17SE from Sun et al. (1994). A plasmid containing a 1321 bp 18S rDNA fragment was then used as a template for nick translation labelling with digoxigenin-11-dUTP.

(2) As above for the Southern hybridisation (section: 5.2.3), a probe containing 338 bp of the Ty1/*copia*-like sequence was made by PCR from a diluted plasmid (clone cerD8) of an *A. cernuum* IGS sequence using the forward primer 26Scop1F: 5'-TAGATTTCCTGGTGGCCT TG-3' and the reverse primer 26Scop2R: 5'-GGACTTCAAACGATGGGAGA-3'. The resulting Ty1/*copia* sequence was purified and used for PCR labelling with biotin-11-dUTP.

(3) The *A. cernuum* IGS A-subrepeat probe was made from a 1.6 kb clone isolated by PCR from a diluted plasmid (clone cerD8) using the forward primer 5'Rep_AAA: 5'-CGTCTAACCAACCAAGATACAAAA-3' and reverse primer IGS_3'Rep1: 5'-ATACCGTT GTGCCCTTGAAC-3', which was labelled by nick translation with biotin-16-dUTP. FISH was carried out according to section 2.8.

5.2.5 EDF-FISH

DNA fibres were prepared following the method in section 2.9 on page 38. The probes were made as follows: (1) The 18S rDNA probe was made from the clone described above (section 5.2.4 on the preceding page), the labelling method used was PCR incorporation of digoxigenin-11-dUTP. (2) The Ty1/*copia* probe was made using the clone described above (section: 5.2.4 on the previous page) by PCR labelling with biotin-16-dUTP.

5.3 Results

5.3.1 *Allium cernuum* IGS Sequence Variability

Details of rDNA unit variation are described previously in chapter 4, and a summary of several *A. cernuum* IGS sequences is shown in fig. 4.3.2 on page 79.

5.3.2 Ty1/*copia*-like Sequence within the IGS of *Allium cernuum*

IGSs are divided into class I and II subfamilies based on the absence or presence, respectively, of a 363 bp region similar to part of a Ty1/*copia*-like elements. Using the program RepeatMasker, this to region was found to be similar to a retrotransposon sequence from *Solanum lycopersicum* (Tont1; GenBank: AF220603). The reading frame occurs in reverse orientation relative to the rDNA coding regions.

A search amongst protein sequences on GenBank (NCBI) using BLASTX identified 22 similar plant sequences (E value: $< 1\text{x}10^{-10}$) including five that were annotated as being possible retrotransposon polyproteins. The sequence with most significant similarity (E value: 4×10^{-34}) was a hypothetical 1316 peptide sequence from *Vitis vinifera* (Protein/CDS GenBank: CAN66637/AM424683) (Velasco et al. 2007). This protein shows regions corresponding to the protease, integrase and reverse transcriptase core domains. The region of similarity between the sequences from *A. cernuum* and *V. vinifera* is upstream of these domains and so may be responsible for the GAG protein, but as this has a much shorter conserved core sequence it is difficult to confirm if this is the case.

A putative ORF containing the Ty1/*copia*-like sequence in the class II IGSs was 414 bp in length. This putative peptide sequence is shown aligned to other similar putative Ty1/*copia* sequences from plants in fig. 5.3.2. The stop codon that terminates this ORF is precisely at the transition point where the class I and II IGSs become unalignable (see fig. 4.3.2).

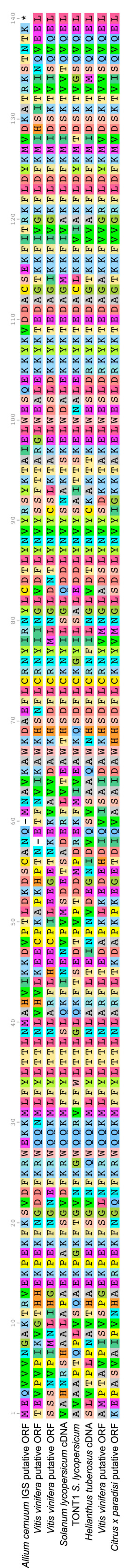

Figure 5.1: Alignment of Ty1/*copia*-like polypeptide sequences
Alignment of partial polypeptides derived from genomic or EST sequences to the translated Ty1/*copia*-like ORF in *A. cernuum* type II IGSs. GenBank accession numbers in order from top to bottom: EU256498, AM431335, AM489208, BM411626, AF220603, EL450546, AM424683 and AF220603.

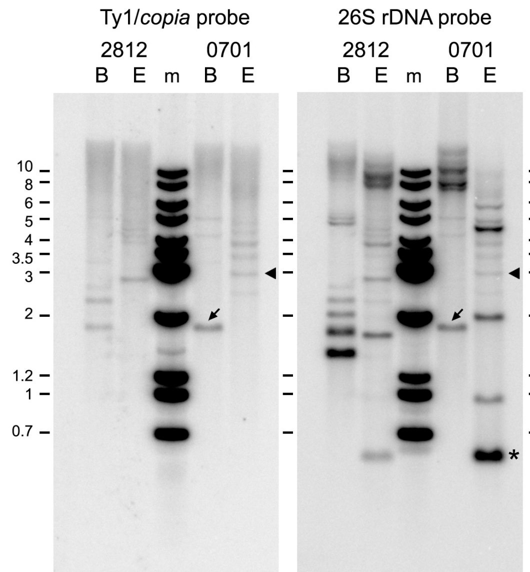

Figure 5.2: Southern hybridisation showing colocalisation of the Ty1/*copia* sequence in some rDNA units of *A. cernuum*
Southern hybridisation of two *A. cernuum* accessions (2812 IPK and 0701 QMUL) digested with *Bsr*DI (B lanes) and *Eco*RI (E lanes). Shown on the left is a low stringency hybridisation with the Ty1/*copia* probe. The same membrane is shown on the right after being stripped and reprobed with the 26S probe using high stringency conditions. An arrow indicates a 1.8 kb fragment corresponding to the rDNA fragment sequenced in class II IGS clones (cerD5/cerD8). An asterisk indicates a 0.5 kb fragment expected from sequenced type I IGS clones; an arrowhead indicates a 2.9 kb fragment expected from the longest class II IGS sequenced (cerD8). See fig. 4.3.2 for restriction site positions and probe annealing sites. (Southern hybridisation carried out by E. Sýkorová.)

5.3.3 Southern Analysis of Ty1/*copia* rDNA Insertions

To confirm the findings from high fidelity PCR that some *A. cernuum* 35S rDNA units harbour partial Ty1/*copia*-like elements, we carried out Southern hybridisation on digested genomic DNA of *A. cernuum*, two accessions were analysed to assess intraspecific variation. *Eco*RI was chosen to release fragments containing the 3' end of the 26S gene, the Ty1/*copia*-like sequence and the A-subrepeats from class II IGSs. *Bsr*DI was chosen as there is a conserved cut site 1.2 kb from the 3' end of the 26S rDNA gene, as a result class II IGSs would be expected to release a 1.8 kb fragment without the A-subrepeats (see fig. 4.3.2).

Digestion of *A. cernuum* (0701 QMUL) with *Eco*RI produced discrete fragments between 2.5-5 kb that were hybridised by the Ty1/*copia* probe, this included one of 2.9 kb that was expected based on the longest class II IGS sequenced (cerD8) (indicated by an arrowhead in fig. 5.3.3). Reprobing with the 26S probe hybridised a 0.5 kb fragment most strongly, this corresponds to the fragment expected from type I IGSs (indicated by an asterisk in fig.

5.3.3). Additionally several longer fragments were hybridised including a 2.9 kb fragment corresponding to that expected from the longest class II IGS clone (cerD8) (indicated in fig. 4.3.2). The hybridisation of the 2.9 kb fragment with both the Ty1/*copia* probe and the 26S probe confirms that both sequences are adjacent to each other in some rDNA units.

Digestion of *A. cernuum* (0701 QMUL) with *Bsr*DI released a 1.8 kb fragment (indicated by an arrow in fig. 5.3.3), as predicted from the class II IGS clone, cerD8. Also, weaker fragments of approximately 4.2 kb and 5 kb that were hybridised by the Ty1/*copia* probe, the sequence giving rise to these fragments is unknown. Reprobing with the 26S probe confirmed that these three fragments also contained the 3' end of the 26S gene as expected. However, the majority of 26S probe signal was from larger fragments of 7-13 kb that are not associated with the Ty1/*copia* sequence. From the *Bsr*DI digestion of accession 0701 QMUL the 26S probe signal from Ty1/*copia*-associated fragments (1.8 kb size) and non-Ty1/*copia*-associated fragments (7-13 kb size range) was quantified. From these two values I estimate that approximately 16% of rDNA units contain the Ty1/*copia*-like sequence in the individual analysed.

Southern hybridisation of the other *A. cernuum* accession (2812 IPK) revealed few fragments to be preserved between species, making interpretation difficult. In the *Bsr*DI digestion the Ty1/*copia* probe hybridised mainly to two *Bsr*DI fragments, approximately 1.8 kb and 2.3 kb long (fig. 5.3.3), confirming that a similar sequence is present. The 1.8 kb fragment probably corresponds to the fragment of the same size identified in the class II IGSs of sequenced *A. cernuum* accession (0701 QMUL).

To see if the Ty1/*copia* fragment was present in species other than *A. cernuum*, 17 other *Allium* species were tested by low stringency Southern hybridisation, some of which are shown in fig. 5.3.3 and summarised in table 6.1.2 on page 106. In all cases the Ty1/*copia* probe failed to hybridise to the genomic DNA of other *Allium* species, including species of the same subgenus, *Amerallium*.

5.3.4 FISH of Retroelement, rDNA and A-Subrepeats

The Ty1/*copia* probe showed weak colocalisation through both major rDNA loci, with the signal showing an unevenly dispersed distribution (fig. 5.3.4 A-C). Outside of the NORs, a weaker signal could also be detected across most of the metaphase chromosomes (fig. 5.3.4 B-C) and throughout interphase euchromatin (fig. 5.3.4 E-F). This signal is similar to what has been seen with retrotransposon sequences that have dispersed genomic distributions e.g. Ty1/*copia* elements in *Hypochaeris* species (family Asteraceae) (Ruas et al. 2008). However,

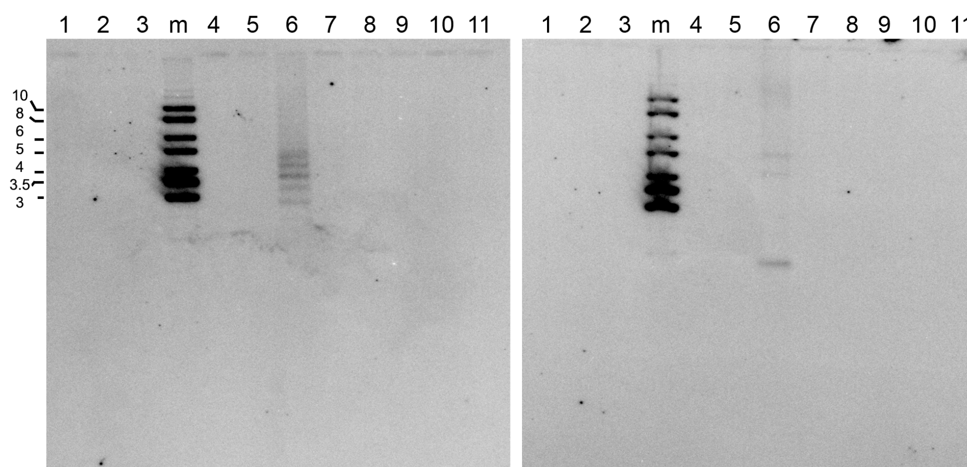

Figure 5.3: Southern hybridisation of the Ty1/*copia*-like sequence to different *Allium* species High stringency Southern hybridisation of the Ty1/*copia*-like sequence from *A. cernuum* to different *Allium* species. *Eco*RI digested genomic DNA (left) and *Bsr*DI digested DNA (right). 1-*Leucojum aestivum*, 2-*Tulbaghia fragrans*, 3-*Ipheion uniflorum*, 4-*A. neapolitanum**, 5-*A. triquetrum**, 6-*A. cernuum**, 7-*A. ursinum**, 8-*A. bulgaricum*, 9-*A. hookeri*, 10-*A. schuberti*, 11-*A. christophii*. Asterisk indicates species in subgenus *Amerallium* following Friesen et al. (2006). (Southern hybridisation carried out by E. Sýkorová.)

in both metaphase (fig. 5.3.4 B) and interphase nuclei (fig. 5.3.4 E), at several subtelomeric positions the Ty1/*copia* sequence was not detected (arrowed in fig. 5.3.4 B). As this may be due to subtelomeric tandem arrays in these regions, we attempted to isolate repeats by restriction digestion of genomic DNA, however no frequently occurring restriction fragments were produced with 22 different enzymes (details in section 3.3.4).

EDF-FISH was used to examine the fine scale association of the retrotransposon sequence within rDNA units. A variation in Ty1/*copia* sequence frequency was observed across different rDNA fibres within an individual, samples of the most densely arranged rDNA arrays are shown in fig. 5.3.4. There appears to be a variable spacing between successive 18S rDNA units, perhaps the result of length variability within IGSs, although varibility in fibre extension may also account for this heterogeneity. In some cases the Ty1/*copia* sequence was not detected between 18S genes indicative of its partial or complete absence in the respective rDNA unit, in agreement with sequenced clones.

The location of the IGS A-subrepeats was confirmed by FISH, this was done to test whether there was any evidence for IGS sequences located outside rDNA loci as found with the Ty1/*copia* probe. Only hybridisation to rDNA loci was detected (fig. 5.3.4 A), indicating that the Ty1/*copia* probe was not detecting extra-rDNA copies of the IGS (fig. 5.3.4). However, along the length of the rDNA array the signal intensity of the A-subrepeat

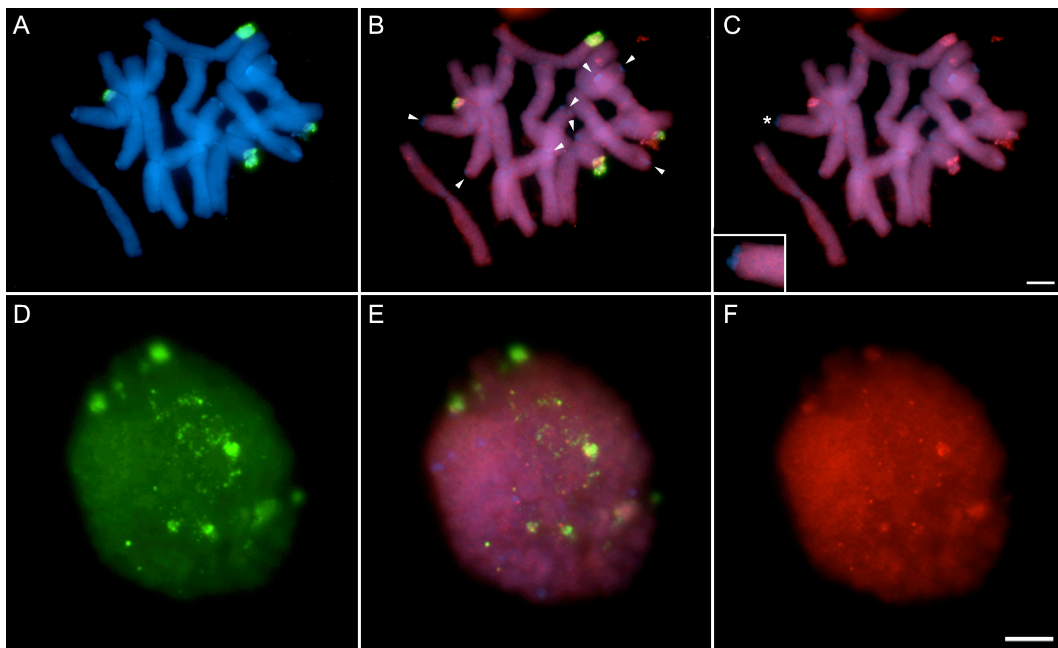

Figure 5.4: FISH of rDNA derived Ty1/*copia* retroelement sequence
FISH to *Allium cernuum* root tip cells. (A-C) Metaphase spread. (A) 18S rDNA probe (green); (B) 18S rDNA probe (green) and the Ty1/*copia* probe (red), arrowheads indicate subterminal regions with low/absent Ty1/*copia* probe signal; and (C) the Ty1/*copia* probe (red). Inset - 2x magnification and enhancement of a chromosome end (asterisk) with an absence of hybridisation in the subtelomere region. (D-F) Interphase nucleus. (D) 18S rDNA probe (green); (E) DAPI (blue), 18S rDNA probe (green) and the Ty1/*copia* probe (red); and (F) the Ty1/*copia* probe (red). Scale bar indicates 5 μm.

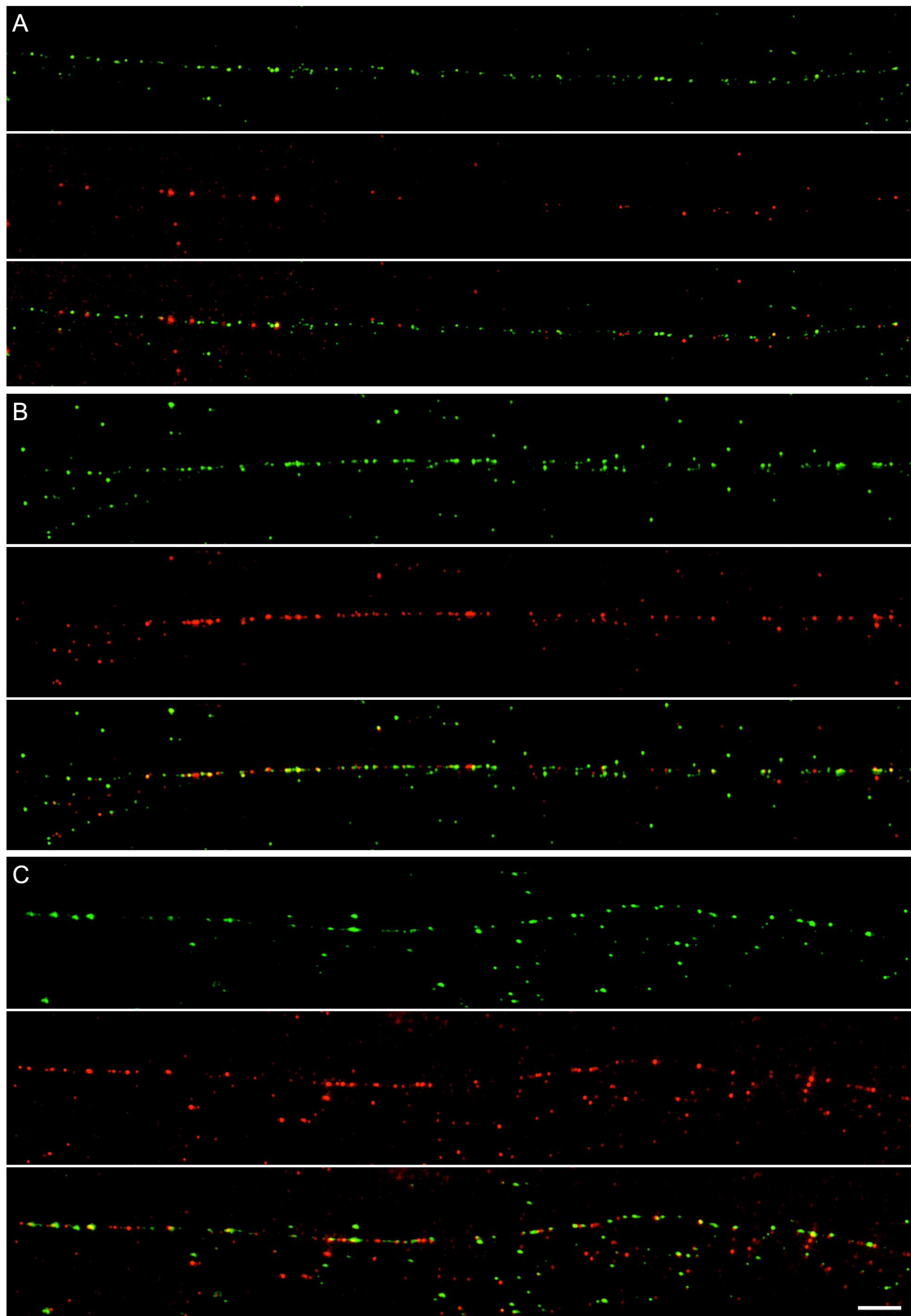

Figure 5.5: EDF-FISH of rDNA and Ty1/*copia* IGS sequence
EDF-FISH to leaf nuclei of *Allium cernuum*, showing the fine-scale arrangement of rDNA repeats in regions where the Ty1/*copia* sequence is in highest abundance. (A-C) Three images showing fibres probed with 18S rDNA (green), the Ty1/*copia* sequence (red) and a merged image of both probes. Scale bar indicates 5 μm.

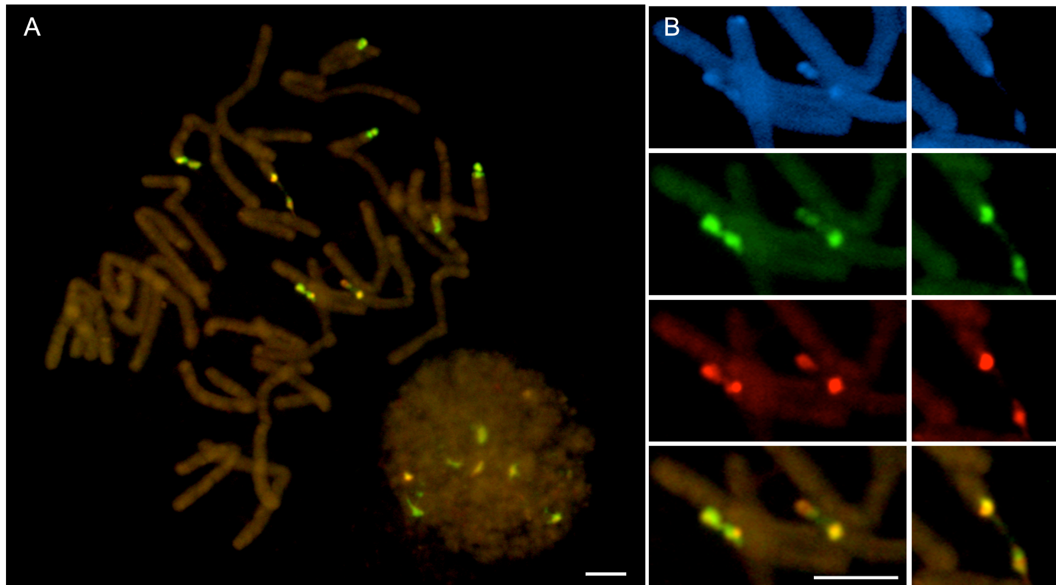

Figure 5.6: FISH of IGS A-repeats
FISH to *Allium cernuum* root tip cells. (A-B) Anaphase and interphase spreads, probed with (A) the IGS A-subrepeat (red) and 18S rDNA (green); (B) 2x magnification of three NORs from the anaphase in (A), shown from top to bottom: DAPI, 18S rDNA probe, IGS A-subrepeat probe and a merged image of both 18S rDNA probe and IGS A-subrepeat probe. Scale bar indicates 5 μm.

probe was not equal to that of the 18S rDNA probe (fig. 5.3.4 B). This is probably due to localised variations in the copy number of A-subrepeats, which was observed with sequencing and Southern hybridisation of class II IGSs.

5.3.5 Relationships and Diversity of the 26S-IGS Sequences

On the basis that rRNA genes typically show low levels of divergence, we were interested in looking at variation between rDNA units with a class II IGS. Primers were used to amplify a 1.1 kb sequence comprising the 3' end of the 26S gene (estimated to be 427 bp) and the adjacent 3' IGS region containing the Ty1/*copia*-like sequence. In total 36 clones were sequenced, one sequence (CR7) was found to be highly divergent from all others, as this was suspected to be a potential artifact of the PCR amplification it was excluded from further analysis. The intragenomic rDNA data set consisted of the remaining 35 clones and the equivalent region from the two class II IGSs sequenced (cerD5, cerD8), providing 37 sequences in total. In 4 cases the same sequence was sampled more than once, producing 32 distinct IGS-types.

When comparing the 37 sequences, the average number of polymorphic sites present between paired sequences varied three-fold along the sequence as calculated using π, to

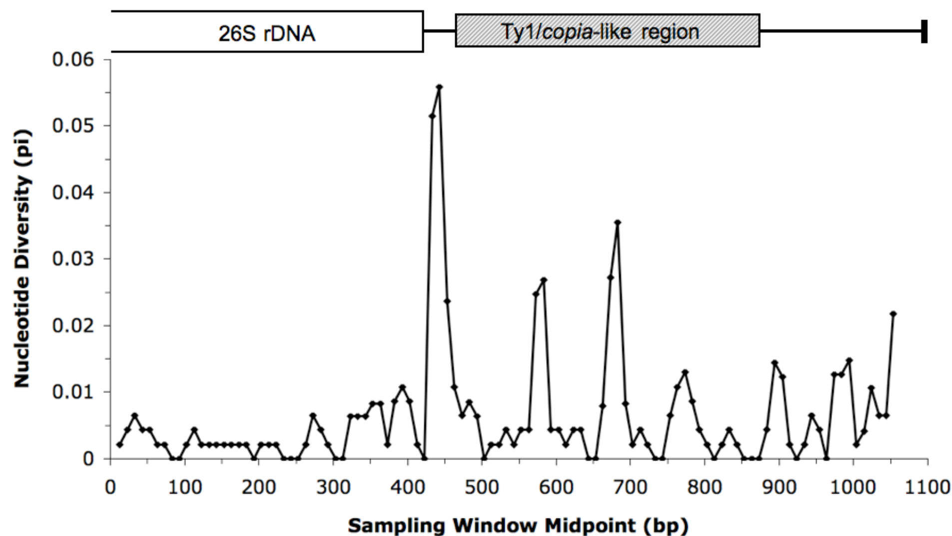

Figure 5.7: Pairwise nucleotide diversity (π) plot for the 26S-IGS region of class II IGSs Graph showing nucleotide diversity (π) on the y-axis plotted against sliding window midpoint in base pairs on the x-axis. The calculation was based on an alignment of 37 sequenced rDNA sequences with a class II IGS. Above the graph is a diagrammatic representation of key features: the 26S gene and the IGS region (black line) that includes the putative ORF with similarity to Ty1/*copia* elements (shaded rectangle).

measure nucleotide diversity (Nei 1987) (fig. 5.3.5). The average nucleotide diversity across the whole 26S-IGS region was 0.0068 (standard deviation (SD) = 0.0005). When calculated separately for the 26S gene and IGS region, the average diversity was 0.0031 (SD = 0.0007) and 0.0093 (SD = 0.0007) respectively. The highest nucleotide diversity with sliding window sampling was at the beginning of the IGS (nucleotide window: 431-455) (fig. 5.3.5). This peak is associated with the presence of three linked mutations found in a short stretch in 24 of the 35 sequences, shown here in bold: T**T**T**A**TTTTAA**G**. In its reverse complement the underlined part of the latter mentioned sequence contains a AATAAA polyadenylation signal (Krishnan 1995, Shen et al. 2008). Despite being in the correct orientation and downstream of the Ty1/*copia*-like ORF, sequenced class I IGSs also have this motif despite lacking the element.

A maximum parsimony (MP) tree was constructed to determine relationships between sequences, a single shortest tree was generated on which several IGS-types had a terminal branch length of zero. Due to this and the low levels of nucleotide variability found (average pairwise similarity overall was 99%), a median joining network was constructed. The topology was essentially the same as the MP tree, with all IGS-types that were on a zero branch length on the MP tree placed on interior nodes (fig. 5.3.5). Only in the cases where

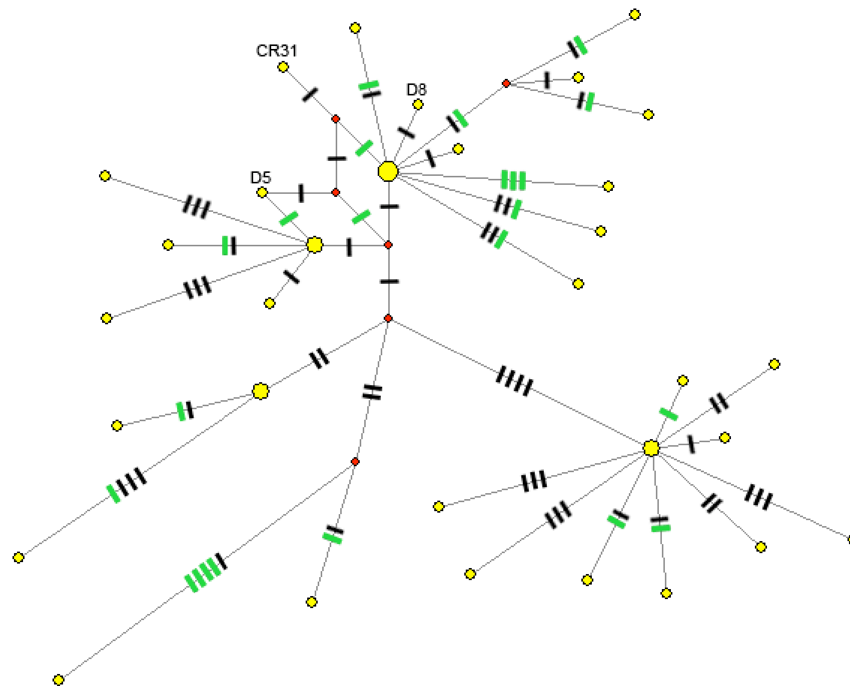

Figure 5.8: Median joining network of intraspecific 26S-IGS sequences
Median joining network showing the relationships amongst 37 intragenomic rDNA sequences (with a class II IGS) based on a 1.1 kb region containing part of the 26S gene and adjacent IGS region. Nodes represent sampled IGS-types (yellow), where the same IGS-types were sampled 2 or 3 times these nodes are shown proportionally larger. Labelled nodes correspond to full-length IGS clones (cerD5, cerD8) and clone CR31 (see results). Unsampled nodes at multifurcating branch points are indicated (red circles). Bars on branches indicate single nucleotide changes in the 26S rDNA gene (green) or the IGS (black).

IGS-types were sampled from interior nodes were they recovered more than once, indicating that these ancestral IGS-types are likely to be present at a higher frequency in the rDNA population. The data show only one homoplasious mutation in two IGS-types (cerD5 and CR31). There are no nucleotide differences in the 26S region between 14 of the 32 IGS-types, including all those placed on internal nodes (fig. 5.3.5).

5.4 Discussion

5.4.1 Integration of a Ty1/*copia*-like Element into 35S rDNA

It seems likely that all the class II IGSs sequenced here are derived from the same progenitor rDNA unit containing a Ty1/*copia* related sequence, this is based on the clones having the same length ORF and the common origin indicated by sequence similarity and the median joining network. The majority of this putative ORF is probably derived from a

retrotransposon based on the significant similarity with several plant Ty1/*copia* polyprotein sequences. The dispersed FISH signal across the chromosomes indicates that sequences similar to the Ty1/*copia* element in the IGS, are spread throughout the genome of *A. cernuum*, the only exception being some of the subtelomeres where the signal is very weak or absent.

It is not clear how the original retrotransposon copy became integrated into the rDNA, but retrotransposition or recombination are both plausible. In rice, retrotransposons are prevalent in the region that is proximal to the rDNA array on chromosome IX (Fujisawa et al. 2006). In the grape genome three different sized retroelements (2870, 2950, and 5800 bp) were found inserted into 35S rDNA units, it is not known how frequently the inserted units occur or where they are located, namely in the NOR locus or as disjunct orphan copies (Velasco et al. 2007). It may be that a longer Ty1/*copia* element became integrated into an rDNA unit and deletions were responsible for the loss of typical retrotransposon features such as the LTRs, the GAG ORF and the conserved domains of the polyprotein gene (e.g. reverse transcriptase) (Kumar et al. 1997), which were not found in any IGS sequenced from *A. cernuum*. It may be that the longer, lower copy 4.2 kb and 5 kb restriction fragments produced by the *Bsr*DI digestion in both accessions originate from IGSs with more of the retrotransposon remaining.

Other reports of retrotransposons found in rDNA include, *Hideaway*, an element with similarity to LTR retrotransposons found within the IGS of a fungus, *Ascobolus immersus* (Kempken 2001). The only case reported for plants is a Ty3/*gypsy* LTR retrotransposon named *monkey*, of which most copies reside at unknown locations within the rDNA loci of some *Musa* (banana) species (Balint-Kurti et al. 2000). During the evolution of primates, multiple non-LTR *Alu* retroelements have inserted into the non-transcribed part of the IGS and become fixed in the rDNA population (Brownell et al. 1983, Gonzalez et al. 1992, 1993). In addition a DNA transposable element, *Pokey*, which targets the large rDNA subunit gene has been found in several *Daphnia* species (Sullender and Crease 2001, Penton et al. 2002, Penton and Crease 2004).

Within the other rDNA gene which encodes 5S rRNA, there are also reports of associations with transposable elements. A family of non-autonomous terminal repeat retrotransposon in miniature (TRIM), named *Cassandra* elements have a 5S-like coding region within each LTR, they have been detected across a wide range of plants (Kalendar et al. 2008). *Cassandra* elements as with other TRIMs are reduced retroelements that lack a coding region (Yang et al. 2007, Kwon et al. 2007), it is thought that the retention of 5S rDNA

provides a way of ensuring transcription of the element, as there is sequence and folding conservation of the internal RNA polymerase III promoter region (Kalendar et al. 2008). In *Allium cepa* and some other related *Allium* species, there are two 5S rDNA families with either short or long spacers, which are arranged in separate arrays (Shibata and Hizume 2002). The 5S units with the longer spacer are thought to be sequences derived from mobile elements (Shibata and Hizume 2002), although in this case it is not known whether *Cassandra* elements are responsible. In the genus *Artemisia* (in family Asteraceae) plants have 35S rDNA loci with a functional 5S rDNA unit inserted downstream of the 26S coding region in an inverted orientation with respect to the rDNA genes (Garcia et al. 2009). It is possible that *Cassandra* elements were responsible for the transposition of 5S rDNA gene into the 26S-18S IGS as there are *c.* 30 bp flanking sequences that resemble LTR reverse transcriptase primer binding sites (Garcia et al. 2009).

5.4.2 Evolution of the class II IGS Region

Following intergration of the Ty1/*copia* element into a single IGS there must have been an increase in copy number of the class II IGS perhaps coinciding with its dispersal through the rDNA arrays. Perhaps a similar mechanism was responsible for the spread and amplification of a novel IGS type in a *Nicotiana tabacum* x *Atropa belladonna* somatic hybrid line (Borisjuk et al. 1988). A mode of dispersal based on homogenisation mechanism(s) seems much more likely than repeated integration events, as the Ty1/*copia* ORF was the same length in all IGS-types. There is evidence that homogenisation acts more efficiently between homologous rDNA arrays than between non-homologous arrays e.g. between sister chromatids at meiosis (Saghai-Maroof et al. 1984, Ellis et al. 1984, Seperack et al. 1988, Schlotterer and Tautz 1994). However, the Ty1/*copia*-like sequence was dispersed throughout both pairs of NORs, despite only accounting for approximately 16% of rDNA units.

More class II IGS-types would inevitably be recovered with a deeper sampling, but if those sampled here are representative, there is no IGS-type that is in a vast excess over the others. The sequence diversity and lack of reticulation (fig. 5.3.5) suggests that concerted evolution is not maintaining homogeneity of class II IGS sequences, although homogenisation mechanisms have probably been responsible for dispersing the Ty1/*copia* sequence though the rDNA arrays. The situation seen here contrasts that seen in *Drosophila* where *R* retroelements repeatedly insert into 26S rDNA genes (Zhang et al. 2008). In contrast, the evolution of primate rDNA is much more reminiscent of the situation in *A. cernuum*, in this case several non-LTR *Alu* retroelements have become incorporated in the non-transcribed

part of the IGS and have become fixed across all units (Brownell et al. 1983, Gonzalez et al. 1992, 1993). The presence of fragmented and even some full-length *Alu* retroelements (Dickson et al. 1989, Gonzalez et al. 1993) demonstrates that concerted evolution is able to act efficiently upon such inserted elements (Gonzalez et al. 1989, 1993).

One finding that remains unresolved is whether the Ty1/*copia* sequence has spread through the existing rDNA units via modification of class I IGSs (e.g. by gene conversion) or by repeated exchanges and/or amplifications of whole rDNA units with class II IGSs. Due to the low numbers of polymorphic sites present in the 26S gene, it was not possible to determine if the genic region is evolving independently of the IGS. Putative topoisomerase I sites have been proposed to drive recombination in the 3' IGS of *Drosophila melanogaster*, enabling parts of the same rDNA unit to take "different evolutionary trajectories" due to repeated localised recombination events (Polanco et al. 1998, 2000). If it is the case that the 26S gene is being homogenised more frequently, the lower nucleotide diversity found in the 26S gene (fig. 5.3.5) could be explained by localised gene conversion in the gene (Hibner et al. 1991).

To better understand how rDNA has evolved in *A. cernuum*, it would be insightful to pinpoint when the putative retrotransposon became integrated into the rDNA. The closest related species of *A. cernuum, A. stellatum* (Nguyen et al. 2008), was not analysed here. However, none of the species sampled from the subgenus *Amerallium* have a related sequence, suggesting that the Ty1/*copia* element integration was relatively recent. FISH indicates that sequences similar to the retrotransposon appear to be distributed throughout most of the genome, but there is no evidence for it being telomeric, in fact FISH hybridisation signals appear to be absent from some distal regions. This contrasts the report in *A. cepa* that some Ty1/*copia* elements are enriched at the chromosome termini (Pearce et al. 1996, Cunado et al. 2001).

Chapter 6

General Discussion

Since the discovery that the telomeres of *Allium* species are unusual (Fuchs et al. 1995) they have been the subject of research to determine how the DNA component is organised. Whilst a range of different distal repetitive sequences have been described for *A. cepa*, none has been unequivocally demonstrated to be telomeric. One sequence which is a satellite repeat called ACSAT (Pich, Fritsch and Schubert 1996) in *A. cepa* was previously proposed to be telomeric (Pich and Schubert 1998), it is likely that this is a subtelomeric sequence as minor 35S rDNA sites were found here to be more distal in *A. cepa.* Distal minor rDNA sites are also detectable in *A. cernuum* which is distantly related to *A. cepa*, suggesting that these may be a common feature of *Allium* genomes. As the minor rDNA sites are located at, or close to telomeric regions they are potentially useful markers for isolating terminal sequences.

During this work a sequence similar to the *Candida albicans* telomere repeat was identified by E. Sýkorová as being present in the *A. cepa* genome. Using FISH this sequence was found to colocalise with rDNA minor sites and parts of the major rDNA sites at interphase and metaphase. EDF-FISH indicates that where these two sequences were detected on the same DNA fibre they generally occur interspersed with each another in an alternating fashion. However, work to isolate the *C. albicans*-like repeat from *A. cepa* was unsuccessful. Nonetheless, the fact that a G-rich sequence was found associated with rDNA which is known to occur close to chromosome termini is a significant finding.

6.0.3 35S rDNA Evolution

Due to the possible inclusion of 35S rDNA in *Allium* telomeres this multicopy family was characterised for *A. cepa* and *A. cernuum*. *A. cepa* 35S rDNA is highly homogenous and has a small proportion of rDNA IGS sequence variants present. *A. cernuum* in contrast has a large proportion of IGS length and sequence variants, these were divided into two classes (I/II) based on the absence/presence of a relic Ty1/*copia*-like element, respectively. In one individual analysed the class II IGS types account for around 16% of the *A. cernuum* genome and are physically dispersed throughout the major rDNA arrays. The median joining network of class II IGS sequences indicates that this fraction is derived from a single sequence, which has spread and amplified, but may not been maintained by homogenisation mechanisms allowing mutations to accumulate in the process.

6.1 Future Work

During this study two important findings have been made that would benefit from further work. These are the identification of a *C. albicans*-like telomere sequence from *A. cepa* and understanding how the relic Ty1/*copia* sequence was incorported into the 35S rDNA IGS of *A. cernuum*.

6.1.1 A New Approach to Isolating the Telomere of *A. cepa*

A method that has recently become available for sequencing repetitive DNA is 454-sequencing. This approach could be used to recover the *C. albicans*-like sequence of *A. cepa*. When considering that a single run generates 500 Mb of sequence, with *A. cepa* having a haploid genome size of 16381 Mb, assuming random sampling, a run would produce the equivalent of 1/33 of the genome. Therefore if a sequence is repeated 330 times it should be sampled approximately 10 times, assuming the sampling size (sequence length) is the same as the repeat length. This method has been shown to be effective for isolating repeats such as the 35S rDNA unit and highly repeated retrotransposon families even when sampling only 1/129 of the genome (Macas et al. 2007). The drawback to using this type of whole genome sequencing is that pyrosequencing is limited to only several hundred bases read length and genomic DNA must be sonicated into corresponding lengths (Margulies et al. 2005). As a result repetitive regions can be difficult to assemble, but consensus reads can be constructed which summarise these repetitive DNA regions.

Discovering the telomere of *Allium* is important as it would open up a new research field in which new questions could then be asked. These include addressing how the telomere DNA is replicated, how recombination is involved in length regulation (e.g. ALT) and how telomere proteins have responded to the change in telomere sequence. Completion of this work will also reveal the extent to which rDNA is involved in telomere biology and allow the phenomenon of rDNA mobility in *Allium* to be resolved.

6.1.2 Obtaining an Overview of Total rDNA Variation and its Evolution

The 35S rDNA of *A. cernuum* is unusually complex with a range of sequence and length variants being present. A complete characterisation of full length IGSs in this species and its closest relatives would allow the evolution of this gene to be better understood. Cloning of rDNA into high capacity vectors such as phage vectors would enable the longest of IGSs to be sequenced by walking. Obtaining longer IGSs with and without the Ty1/*copia* insertion should also be informative on pinpointing the retroelement insertion site. This would make an interesting comparison to an analogous situation in primates where *Alu* elements have become incorporated into the rDNA IGS at different points in their radiation.

Another line of study that the variability in *A.cernuum* opens up is based on there being a relatively high number of single nucleotide polymorphisms (SNPs) between rDNA units. These SNPs could allow recombination breakpoints to be mapped and linkage disequilibrium to be assessed, collectively these would help to shed light if concerted evolution acts within different parts of the rDNA unit. To date this approach has not been done across the whole rDNA unit, partly because sequencing over multiple full length rDNA units still requires extensive sequencing effort, but also because the low levels of mutations normally present do not provide sufficient resolution or sensitivity.

Appendix A

Southern Hybridisation Summary Table

Clade	Accession	Species	Subgenus	*Hind*III probe	Ty1/*copia* probe
3	QMUL	*A. cepa* cv. Ailsa Craig	*Cepa*	Strong	Absent
	0500	*A. proliferum*	*Cepa*	Strong*	n.a.
	5010	*A. galanthum*	*Cepa*	Strong*	n.a.
	1667	*A. altaicum*	*Cepa*	Strong*	n.a.
	1854	*A. schoenoprasum*	*Cepa*	Absent	Absent
	QMUL	*A. schoenoprasum*	*Cepa*	Weak*	Absent
	0988	*A. chinense*	*Cepa*	Weak*	n.a.
	5661	*A. splendens*	*Reticulatobulbosa*	Weak*	n.a.
	5152	*A. roylei*	*Polyprason*	Strong*	Absent
	3179	*A. obliquum*	*Polyprason*	Absent	n.a.
	2909	*A. carolinianum*	*Polyprason*	Absent	n.a.
	QMUL	*A. flavum*	*Allium*	Absent*	Absent
	1797	*A. sativum*	*Allium*	Absent	Absent
	2373	*A. mongolicum*	*Rhizirideum*	Absent	n.a.
	1853	*A. angulosum*	*Rhizirideum*	Absent	n.a.
	5895	*A. mairei*	*Cyathophora*	Absent	n.a.
	4247	*A. cyathophorum*	*Cyathophora*	Absent	n.a.
	4246	*A. tuberosum*	*Butomissa*	Absent*	Absent
	OSBG 01-17-0128-10	*A. oreoprasum*	*Butomissa*	Absent	n.a.
2	2800	*A. hollandicum*	*Melanocrommyum*	Absent	Absent
	2552	*A. rosenorum*	*Melanocrommyum*	Absent	n.a.
	0515	*A. nigrum*	*Melanocrommyum*	Absent	n.a.
	QMUL	*A. schubertii*	*Melanocrommyum*	Absent	Absent
	1388	*A. christophii*	*Melanocrommyum*	n.a.	Absent
	0465	*A. macleanii*	*Melanocrommyum*	Absent	n.a.
	2264	*A. stipitatum*	*Melanocrommyum*	Absent	n.a.
	5942	*A. giganteum*	*Melanocrommyum*	Absent	n.a.
	2732	*A. oreophilum*	*Porphyroprason*	Absent	n.a.
	0773	*A. victorialis*	*Anguinum*	Absent	Absent
	2797	*A. neriniflorum*	*Caloscordum*	Absent	Absent
1	2812	*A. cernuum*	*Amerallium*	Absent	strong
	Chester 0701 QMUL	*A. cernuum*	*Amerallium*	Absent	strong
	0914	*A. ursinum*	*Amerallium*	Absent	Absent
	QMUL	*A. neapolitanum*	*Amerallium*	Absent	n.a.
	QMUL	*A. triquetrum*	*Amerallium*	n.a.	Absent
	3220	*A. bulgaricum*	*Nectaroscordum*	Absent	Absent
	QMUL	*Tulbaghia fragrans*	outgroup	Absent	Absent
	QMUL	*Ipheion sp.*	outgroup	Absent	Absent

Table A.1: Summary of 35S rDNA Southern hybridisation results
Hybridisation intensity is listed for two probes to the genomic DNA of *Allium* species and species from closely related genera. The IPK TAX accession code is shown for species unless they are from OSBG or QMUL, subgenus name is included following Friesen et al. (2006). An asterisk indicates that low stringency washing was used, if absent high stringency washing was used, n.a. = not analysed).

Appendix B

GenBank Accession Numbers

Species	Voucher	Code	Region	GenBank
A. cepa Ailsa Craig	QMUL	ACIGS4	26S-IGS-18S	EU256494
A. cernuum	Chester 0701 QMUL	CerDI	26S-IGS-18S	EU256495
A. cernuum	Chester 0701 QMUL	CerD3	26S-IGS-18S	EU256496
A. cernuum	Chester 0701 QMUL	CerD5	26S-IGS-18S	EU256497
A. cernuum	Chester 0701 QMUL	CerD8	26S-IGS-18S	EU256498
A. cernuum	Chester 0701 QMUL	2Cer26S	partial 26S	EU256499
A. cernuum	Chester 0701 QMUL	CR1	26S-IGS	EU980328
A. cernuum	Chester 0701 QMUL	CR2	26S-IGS	EU980329
A. cernuum	Chester 0701 QMUL	CR3	26S-IGS	EU980330
A. cernuum	Chester 0701 QMUL	CR4	26S-IGS	EU980331
A. cernuum	Chester 0701 QMUL	CR5	26S-IGS	EU980332
A. cernuum	Chester 0701 QMUL	CR6	26S-IGS	EU980333
A. cernuum	Chester 0701 QMUL	CR8	26S-IGS	EU980334
A. cernuum	Chester 0701 QMUL	CR9	26S-IGS	EU980335
A. cernuum	Chester 0701 QMUL	CR10	26S-IGS	EU980336
A. cernuum	Chester 0701 QMUL	CR11	26S-IGS	EU980337
A. cernuum	Chester 0701 QMUL	CR12	26S-IGS	EU980338
A. cernuum	Chester 0701 QMUL	CR13	26S-IGS	EU980339
A. cernuum	Chester 0701 QMUL	CR14	26S-IGS	EU980340
A. cernuum	Chester 0701 QMUL	CR15	26S-IGS	EU980341
A. cernuum	Chester 0701 QMUL	CR16	26S-IGS	EU980342
A. cernuum	Chester 0701 QMUL	CR17	26S-IGS	EU980343
A. cernuum	Chester 0701 QMUL	CR18	26S-IGS	EU980344
A. cernuum	Chester 0701 QMUL	CR19	26S-IGS	EU980345
A. cernuum	Chester 0701 QMUL	CR20	26S-IGS	EU980346
A. cernuum	Chester 0701 QMUL	CR21	26S-IGS	EU980347
A. cernuum	Chester 0701 QMUL	CR22	26S-IGS	EU980348
A. cernuum	Chester 0701 QMUL	CR23	26S-IGS	EU980349
A. cernuum	Chester 0701 QMUL	CR24	26S-IGS	EU980350
A. cernuum	Chester 0701 QMUL	CR25	26S-IGS	EU980351
A. cernuum	Chester 0701 QMUL	CR26	26S-IGS	EU980352
A. cernuum	Chester 0701 QMUL	CR27	26S-IGS	EU980353
A. cernuum	Chester 0701 QMUL	CR28	26S-IGS	EU980354
A. cernuum	Chester 0701 QMUL	CR29	26S-IGS	EU980355
A. cernuum	Chester 0701 QMUL	CR30	26S-IGS	EU980356
A. cernuum	Chester 0701 QMUL	CR31	26S-IGS	EU980357
A. cernuum	Chester 0701 QMUL	CR32	26S-IGS	EU980358
A. cernuum	Chester 0701 QMUL	CR33	26S-IGS	EU980359
A. cernuum	Chester 0701 QMUL	CR34	26S-IGS	EU980360
A. cernuum	Chester 0701 QMUL	CR35	26S-IGS	EU980361
A. cernuum	Chester 0701 QMUL	CR36	26S-IGS	EU980362

Table B.1: GenBank accession numbers

Appendix C

DNA Sequence Alignments

```
   1        10        20        30        40        50        60        70        80        90        100       110       120       130       140
a1 TGCAGTCGGTAGTGCGCTCGGAAAATGGACACTCATTTAGGCCCCGCTCGAGTTGGCAATCACCCGACGCTCTCCCCGGTCCGGCCCGATGTTTAGGCTAAGGGCGCAACGAGG-TAGGAGCGCAACTTCAAGCCCGACG
a2 TGCACTCGGTCGTGCGTCCGGCAAATCGACACGCTTTTGGGAACCGCACGAGTTGGCAATCACCCGGCGCTCGCCCCGGTCCGGTCCGATGCTCCCGCCCAAA---CATCGGGGGTCGGAGCGCAAGTTCACGCCCGACG
a3 TGCACTCGGTCGTGCGCCCGGCAAATCGACACGCTTTTGGGAACCGCACGAGTTGGCAATCACCCGGCGCTCGTCCCGGTCCGGCCCGATGCTCCCGCCCAAA---CATCGGGG-TCGGAGCGCAAGTT-AACCCCGGCC
a4 TGCGCTCGGTCGTGCGTCCGGCAAATCGAAACGCTTTTGAAAACCGCTCGAGTTGGCAAT--------------------------------------------------------------------------------
a5 TGCACTCGGTCGTGCGTCCGGCAAATCGACACGCTTTTGGGTACCGCAAGAGTTGGCAATCACCGGGCGCTCGCCCTGGTCCGGCCCGATGCTCCCGCAAAAA---CATCGGGG-TCGGAGCGCAAGTT-AACCCCGGCC
a6 TGTGCTCGGTCGTGCGTCCGGCAAATCGACACGCTTTTTGGTACCGCAAGAGTTGGCAATCGCCCGGCGCTCGCCTCGGTCCGGCCCGATGCTCCTGCCCAAA---CATCGGGG-TCGGAGCGCAAGTT-AACCCCGTCC
             150       160       170       180       190       200       210       220       230       240       250       260       270       280
a1 ACGGCCTTGCTACGAGTCGGGCACGTGGAAACTCGGGGAGCGCTTCATGTCGGAGTGGTAAATACCCGAAGCATGCACTTGGAT-GCTGCTGCTTCGAGTCGCGCGATGGAAAACACGAGGCGCGCTTCGGGTCATTCGG
a2 ACGGCCTTGCTTCGAGTCTGGGACGTAAAAACTCGGAGAGCGCTTCGGGTCAGAGCGGCAAATACCCTAAGCGTGCACATGGAT-GCCGCTGCTTCGAGTCGCGCGAAGGAAAACTCGGAGCGCGCTTCGGGTCATTAGG
a3 ACGGCCTTGCTTCGAGTCGCGCACGTGGAAACTCGGGTAGCGCTTCGGGTCAGAGTGGAAAATACCCGAAGCGCGCACTTGGAT-GCCGCTGCTTCGAGTCGCGCAATGGAAAACTCGAAGCGAACTTCGGGTCATTCGG
a4 --------------------------------------------------------------------------------------------------------------------------------------------
a5 ACGGCATTGCTTCGAGTCGTGCGCGTGGAAAATCGGCCTGCGCTTCGGGTCAGAGTGGGAAAAACCCGAAGCACGCACTTGGAT-GCCGCTGCTTCGAGACGCGCAGTGGAAAACTCGAAGCGCACTTCGGGTCATTCGT
a6 ACGGCCTTGCTTCGAGTCGTGCGCGTGGAAAATCGGCCTGCGCTTCGGGTCAGAGTGGGAAATACCCGAAGCGCGCACTTGGTTTGCCTCTGCTTCGAGTCGTTCTGTGGAAAACTCGGAGCGCACCTCGGGTCATTCGG
             290       300       310       320       330       340       350       360       370       380       390       400       410       420
a1 CAACTTTACCGAAGGAAGCCCGAGGAAGTCCCTA
a2 CAACATTCCCGAACCAAGCCCGAGGAGGTCACTGCTCCGAG--TCGAGCACGAAAAAACCGCAAGCGTGCTTCGGGTCATGCGGCTACGTACCCGAAGC-TCGCCCCGGACGGTATTGCTCCGAGAGAACTCGGAAGTTG
a3 CAAATTTCCCGAAGCAAGCCGGAGGAAGTCCCCTCTCCGATAGCCGAGCGCGAAAAAACTGCCAGCGTGCTTCGGGTCGAGCGGCTTCGTACCCGGAGC-TCGTCCTGGACGATATTGCTCAGAGAGAACTCGGAACTTG
a4 --------------------------------------------------------------------------------------------------------------------------------------------
a5 CAAATTTCCCGAAGCAAGCCCGAGGTAGTCCCCGCTCCGAGAGCCGAGCACGTGAAAACTGCCAGCGTTCTTCGGGTCAAGAGGCTACGTACCTGAAGC-TCGTCCTGGACGATATTGCTCCGAGAGAACTTGGAACTTG
a6 CAAATTTCCCGAAGCAGGCCAGAA-ACGTCCCTGCTCCGAG--TCGAGCGCGAGAAAAGTGCCTGCGTGCTTCGAGTCAGTGCATCATTTACCCGAAGTATTCAAGTGGCCGAGTTGGCTTCGGGTCGAGCTTGAAAAAA
             430       440       450       460       470       480       490       500       510       520       530       540       550  555
a1
a2 CTTTGGCACGGGCATTCGATAACTAGAAGCTTGGTGTGGAACGAGGATATGAATCGAGCGTACGAAGGACAGAAGCGCGCTCCGAGTCGGGCCGAGTGTCGGGTCTAGAAATATGTCCGCCATACCCCGTAACCG
a3 CTTTGGCTCGGGCATTCGAAAATTTGAAGCTGGGTATGGAAAGTTGATACGAATCGAGCGTACAAACGACAGACGCGTGCTCCG------------TGTCGGGTCGGGAAATATGTCCGCCCTCCTCCGTAAGCA
a4 --------------------------------------------------------------------------------------------------------AAATAGGTCCTCCATACCCCGTAACCG
a5 CTTTGGCTCGGGCATTCGATAATTTGAAGCTGGGTATGGAAAGTGGATACGAATCGAGCGTATAGACGACAGAAGCGTGCTCCG------------TGTCGGGTCGAGAAATATGTCCGCCATCCCTCGTAACCG
a6 CTCGG
```

Figure C.1: *A. cepa* IGS A-subrepeat alignment

```
   1        10        20        30        40        50        60        70        80        90        100       110       120       130       140
b1     CTTCGAGTCGAGCAAGAAAAAACTGCTAGCGTGCTTCGGGTGTTGCGGCGACATACCCGAAGCTCGCCC-TGGACGATATT-GCTCCGAGAAAACTCGGAACTTGCTACGGCTCGGGTATTCGGAAATTTGAAGCT
b2 CCCGCTTCGAGTCGTGCACGTGAAAACTGCTAGCGTGCTAAGGGTCATGCGGCCATGTGCCAGAAGCTCGCGCGCGGACAATATTTGCTCCTAGAAAACTCGGAACTGGCTTCGGCTCGGGCATTCGAAAATTTGAAGCT
b3 CCCGCTTCGAGTCGTGCACGTGGAAAACGCCAGCGTGCTTTGGGTCATGCGGCTACATGCCCAAAGCTCGCGCCTGGT-----------------------------------------------------------------
             150       160       170       180       190       200       210       220  224
b1 GGGTCGGGAACGAGCATACAAACGTTTGGATCGTGCTCCGGGTCGAGCAGCGGCAACCTGCCCGAAGTAAGCCCGAGGTAG-CC
b2 GGGCCAAAAACGAGCATACCAACGTTTGGATCGTGCTCCAGGTCGACCCGCGGCAACCT-CCCGAAGCAAGCCGGAGGAAGTCC
b3 -----------------------------------------------------AACCG
```

Figure C.2: *A. cepa* IGS B-subrepeat alignment

```
   1        10        20        30        40        50        60        70        80        90        100       110       120       130       140
a1 AACAAGCACTTGTTTTTGGATTCTTCTCGAAGTAGCAAGTGTTTCTCTCTCTTTTTTGTGGGGAGAACCGTGGTGAAATGTTTTTAGAAGAATAATGCAGAAGAAGGAGAAGAAGTTGTTATTCTGTAAATAGAAAATGG
a2 AACAAACACTTGTTTTTGGATTCTTCTCGAAGTAGCAAGTGTTTCTCTCTCTTTTTTGTGGGGAGAACCGTGGTGAAATGTTTTTAGAAGAATAATGCAGAAGAAGGAGAAGAAGTTGTTATTCTGTAAATAGAAAATGG
a3 AACAAACACTTGTTTTTGGATTCTTCTCGAAGTAGCAAGTGTTTCTCTCTCTTTTTTGTGGGGAGAACCGTGGTGAAATGTTTTTAGAAGAATAATGCAGAAGAAGGAGAAGAAGTTGTTATTCTGTAAATAGAAAATGG
a4 AACAAGCACTTGTGTTTGGATTCTTCTCGAAGTAGTAAGTGTTTCTCTCTCTTTTTTGTGGGGATAACCGTGGTGAAAAGTTTTTAGAAGAATAATGCAGAAGAAGGAGAAGAAGTTGTTATTCTGTAAATAGAAAATGG
             150       160       170       180       190       200       210       220       230       240       250       260       270       280
a1 ACTGTTTATATAAGAAGTGAAAGAGTATATTTGTAACAATCAAAACTATTCAAAAAAGATGTGTTTGTTCGAATTTAAAAAGGTCAGGTTTTTCCCTCTTTGATTCCAATTTAACCCCAGGTACCCCAAGGTCCCCAAGT
a2 ACTGTTTATATAAGAAGTGAAAGGGTATATTTGTAACAATCAAAACTGTTCAAAAAAGACGTGTTTGTTCGAATTTAAAAAGGTCAAGTTTTTCCCTCTTTGATTCCAATTTAACCCCAGGTACCCCAAGGTCCCCAAGT
a3 ACTGTTTATATAAGAAGTGAAAGGGTATATTTGTAACAACCAAAACTGTTCAAAAAAGACGTGTTTGTTCGAATTTAAAAAGGTCAAGTTTTTCCCTCTTTGATTCCAATTTAACCCCAGGTACCCCAAGGTCCCCAAGT
a4 ACTGTTTATATAAGAAGTGAAAGGGTATATTTGTAACAATCAAAACTGTTCAA
             290       300       310       320       330       340       350       360       370       380       390       400       410       420
a1 CCCAAGTCCCATCTTTTGACTATTAATATATATGTGTGTGAACCCATGTGAGCCCCAAGTCAAGCCCCTTTTCATTGTTTGTTTCTATTCAAAACATGAATAATGATATGACATATAAACAAACAGAAAACATTATTCTT
a2 CCCAAGTCCCATCTTTTGACTATTAATATATATGTGTGTGAACCCATGTGAGCCCCAAGTCAAGCCCCTTTTCATTGTTTGTTTCTATTCAAAACATGAATAATGATATGACATATAAACAAACAGAAAACATTATTCTT
a3 CCCAAGTCCCACCTTTTGACTATTAATATATATGTGTGTGAACCCATGTGAGCCCCAAGTCAAGCCCCTTTTCATTGTTTGTTTCTATTCAAGATATGAATAATGATATGACATATAAACAAACAGAAAACATTATTCTT
a4
             430       440       450  457
a1 TGTGATGTGGGACAAATGCTTTTCATTTAAAAACACC
a2 TGTGATGTGGGACAAATGCTTTTCATTTAAAAACACC
a3 TGTGATGTGGGACAAATGCTTTTCATTTAAAAACACC
a4
```

Figure C.3: *A. cernuum* IGS A-subrepeat alignment

Bibliography

Abad, J. P., de Pablos, B., Osoegawa, K., de Jong, P. J., Martin-Gallardo, A. and Villasante, A.: 2004, Genomic analysis of drosophila melanogaster telomeres: Full-length copies of het-a and tart elements at telomeres, *Molecular Biology and Evolution* **21**(9), 1613–1619.

Adam, R. D.: 1992, Chromosome-size variation in giardia lamblia - the role of rdna repeats, *Nucleic Acids Research* **20**(12), 3057–3061.

Adam, R., Nash, T. and Wellems, T.: 1991, Telomeric location of giardia rdna genes, *Molecular Cell Biology* **11**(6).

Adams, S., Hartman, T., Lim, K. Y., Chase, M. W., Bennett, M. D., Leitch, I. J. and Leitch, A. R.: 2001, Loss and recovery of arabidopsis-type telomere repeat sequences 5'-(tttaggg)n-3' in the evolution of a major radiation of flowering plants, *Proceedings of the Royal Society of London B* **268**, 1541–1546.

Adams, S. P., Leitch, I. J., Bennet, M. D., Chase, M. W. and Leitch, A. R.: 2000, Ribosomal dna evolution and phylogeny in aloe (asphodelaceae), *American Journal of Botany* **87**(11), 1578–1583.

Ali, H. B. M., Lysak, M. A. and Schubert, I.: 2005, Chromosomal localization of rdna in the brassicaceae, *Genome* **48**(2), 341–346.

Alkhimova, O. G., Mazurok, N. A., Potapova, T. A., Zakian, S. M., Heslop-Harrison, J. S. and Vershinin, A. V.: 2004, Diverse patterns of the tandem repeats organization in rye chromosomes, *Chromosoma* **113**(1), 42–52.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J.: 1990, Basic local alignment search tool, *Journal of molecular biology* **215**(3), 403–410.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J.: 1997, Gapped blast and psi-blast: a new generation of protein database search programs, *Nucleic Acids Research* **25**(17), 3389–3402.

Appels, R. and Dvořák, J.: 1982, Relative rates of divergence of spacer and gene sequences within the rdna region of species in the triticeae: Implications for the maintenance of homogeneity of a repeated gene family, *Theoretical and Applied Genetics* **63**(4), 361–365.

Armstrong, S. J., Fransz, P., Marshall, D. F. and Jones, G. H.: 1998, Physical mapping of dna repetitive sequences to mitotic and meiotic chromosomes of brassica oleracea var. alboglabra by fluorescence in situ hybridization, *Heredity* **81**(6), 666–673.

Arnheim, N.: 1983, *Concerted evolution in multigene families*, Sinauer, Sunderland, Mass., pp. 38–61.

Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O. and Zimmer, E.: 1980, Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes, *Proceedings of the National Academy of Science USA* **77**(12), 7323–7327.

Arnheim, N., Treco, D., Taylor, B. and Eicher, E. M.: 1982, Distribution of ribosomal gene length variants among mouse chromosomes, *Proceedings of the National Academy of Science USA* **79**(15), 4677–4680.

Avanzi, S., Maggini, F. and Innocenti, A. M.: 1973, Amplification of ribosomal cistrons during the maturation of metaxylem in the root of allium cepa, *Protoplasma* **76**(2), 197–210.

Azam, M., Lee, J. Y., Abraham, V., Chanoux, R., Schoenly, K. A. and Johnson, F. B.: 2006, Evidence that the s.cerevisiae sgs1 protein facilitates recombinational repair of telomeres during senescence, *Nucleic Acids Research* **34**(2), 506–516.

Bachrati, C. Z., Borts, R. H. and Hickson, I. D.: 2006, Mobile d-loops are a preferred substrate for the bloom's syndrome helicase, *Nucleic Acids Research* **34**(8), 2269–2279.

Bailey, S. M., Brenneman, M. A. and Goodwin, E. H.: 2004, Frequent recombination in telomeric dna may extend the proliferative life of telomerase-negative cells, *Nucleic Acids Research* **32**(12), 3743–3751.

Baldwin, B. G. and Markos, S.: 1998, Phylogenetic utility of the external transcribed spacer (ets) of 18s-26s rdna: congruence of ets and its trees of calycadenia (compositae), *Molecular Phylogenetics and Evolution* **10**(3), 449–463.

Balint-Kurti, P. J., Clendennen, S. K., Doleželová, M., Valárik, M., Doležel, J., Beetham, P. R. and May, G. D.: 2000, Identification and chromosomal localization of the monkey retrotransposon in musa sp, *Molecular and General Genetics* **263**(6), 908–915.

Bandelt, H., Forster, P. and Rohl, A.: 1999, Median-joining networks for inferring intraspecific phylogenies, *Molecular Biology and Evolution* **16**, 37–48.

Barnes, S., James, A. and Jamieson, G.: 1985, The organisation, nucleotide sequence, and chromosomal distribution of a satellite dna from allium cepa, *Chromosoma* **92**, 185–192.

Bass, H. W., Riera-Lizarazu, O., Ananiev, E. V., Bordoli, S. J., Rines, H. W., Phillips, R. L., Sedat, J. W., Agard, D. A. and Cande, W. Z.: 2000, Evidence for the coincident initiation of homolog pairing and synapsis during the telomere-clustering (bouquet) stage of meiotic prophase, *Journal of Cell Science* **113**(Pt 6), 1033–1042.

Bednenko, J., Melek, M., Greene, E. C. and Shippen, D. E.: 1997, Developmentally regulated initiation of dna synthesis by telomerase: evidence for factor-assisted de novo telomere formation, *EMBO Journal* **16**(9), 2507–2518.

Bena, G., Jubier, M.-F., Olivieri, I. and Lejeune, B.: 1998, Ribosomal external and internal transcribed spacers: combined use in the phylogenetic analysis of medicago (leguminosae), *Journal of Molecular Evolution* **46**(3), 299–306.

Bennett, M. D. and Leitch, I. J.: 2004, Angiosperm dna c-values database, http://data.kew.org/cvalues/.

Bennett, M. D. and Leitch, I. J.: 2005, Plant genome size research: a field in focus, *Annals of Botany* **95**(1), 1–6.

Bennett, M. D., Leitch, I. J., Price, H. J. and Johnston, J. S.: 2003, Comparisons with caenorhabditis (100 mb) and drosophila (175 mb) using flow cytometry show genome size in arabidopsis to be 157 mb and thus 25 than the arabidopsis genome initiative estimate of 125 mb, *Annals of Botany* **91**(5), 547–557.

Bennett, M. D. and Smith, J.: 1976, Nuclear dna amounts in angiosperms, *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **274**, 227–274.

Bennett, M. and Smith, J.: 1991, Nuclear-dna amounts in angiosperms, *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **334**(1271), 309–345.

Bennett, R. J., Keck, J. L. and Wang, J. C.: 1999, Binding specificity determines polarity of dna unwinding by the sgs1 protein of s. cerevisiae, *Journal of Molecular Biology* **289**(2), 235–248.

Bennetzen, J. L., Ma, J. and Devos, K. M.: 2005, Mechanisms of recent genome size variation in flowering plants, *Annals of Botany* **95**(1), 127–132.

Bharti, A. K., Olson, M. O., Kufe, D. W. and Rubin, E. H.: 1996, Identification of a nucleolin binding site in human topoisomerase i, *Journal of Biological Chemistry* **271**(4), 1993–1997.

Bianchi, A., Smith, S., Chong, L., Elias, P. and de Lange, T.: 1997, Trf1 is a dimer and bends telomeric dna, *EMBO Journal* **16**(7), 1785–1794.

Biessmann, H., Donath, J. and Walter, M. F.: 1996, Molecular characterization of the anopheles gambiae 2l telomeric region via an integrated transgene, *Insect Molecular Biology* **5**(1), 11–20.

Biessmann, H., Kasravi, B., Bui, T., Fujiwara, G., Champion, L. E. and Mason, J. M.: 1994, Comparison of two active het-a retroposons of drosophila melanogaster, *Chromosoma* **103**(2), 90–98.

Biessmann, H., Kobeski, F., Walter, M. F., Kasravi, A. and Roth, C. W.: 1998, Dna organization and length polymorphism at the 2l telomeric region of anopheles gambiae, *Insect Molecular Biology* **7**(1), 83–93.

Biessmann, H. and Mason, J.: 2003, Telomerase-independent mechanisms of telomere elongation, *Cellular and Molecular Life Sciences* **60**(11), 2325–2333.

Biessmann, H., Mason, J. M., Ferry, K., d'Hulst, M., Valgeirsdottir, K., Traverse, K. L. and Pardue, M. L.: 1990, Addition of telomere-associated het dna sequences "heals" broken chromosome ends in drosophila, *Cell* **61**(4), 663–673.

Biessmann, H., Zurovcova, M., Yao, J. G., Lozovskaya, E. and Walter, M. F.: 2000, A telomeric satellite in drosophila virilis and its sibling species., *Chromosoma* **109**(6), 372–380.

Blackburn, E. H. and Gall, J. G.: 1978, A tandemly repeated sequence at the termini of the extrachromosomal ribosomal rna genes in tetrahymena, *Journal of Molecular Biology* **120**(1), 33–53.

Bodega, B., Cardone, M. F., Rocchi, M., Meneveri, R., Marozzi, A. and Ginelli, E.: 2006, The boundary of macaque rdna is constituted by low-copy sequences conserved during evolution, *Genomics* **88**(5), 564–571.

Borisjuk, N., Borisjuk, L., Komarnytsky, S., Timeva, S., Hemleben, V., Gleba, Y. and Raskin, I.: 2000, Tobacco ribosomal dna spacer element stimulates amplification and expression of heterologous genes, *Nature Biotechnology* **18**(12), 1303–1306.

Borisjuk, N. and Hemleben, V.: 1993, Nucleotide sequence of the potato rdna intergenic spacer, *Plant Molecular Biology* **21**(2), 381–384.

Borisjuk, N. V., Davidjuk, Y. M., Kostishin, S. S., Miroshnichenco, G. P., Velasco, R., Hemleben, V. and Volkov, R. A.: 1997, Structural analysis of rdna in the genus nicotiana, *Plant Molecular Biology* **35**(5), 655–660.

Borisjuk, N. V., Momot, V. P. and Gleba, Y.: 1988, Novel class of rdna repeat units in somatic hybrids between nicotiana and atropa, *Theoretical and Applied Genetics* **76**(1), 108–112.

Bosco, G. and Haber, J. E.: 1998, Chromosome break-induced dna replication leads to nonreciprocal translocations and telomere capture, *Genetics* **150**(3), 1037–1047.

Brewer, B. J. and Fangman, W. L.: 1988, A replication fork barrier at the 3' end of yeast ribosomal rna genes, *Cell* **55**(4), 637–643.

Brosh, R. M. J., Majumdar, A., Desai, S., Hickson, I. D., Bohr, V. A. and Seidman, M. M.: 2001, Unwinding of a dna triple helix by the werner and bloom syndrome helicases, *The Journal of Biological Chemistry* **276**(5), 3024–3030.

Brown, D. D. and Dawid, I. B.: 1968, Specific gene amplification in oocytes. oocyte nuclei contain extrachromosomal replicas of the genes for ribosomal rna, *Science* **160**(825), 272–280.

Brown, D. D., Wensink, P. C. and Jordan, E.: 1972, A comparison of the ribosomal dna's of xenopus laevis and xenopus mulleri: the evolution of tandem genes, *Journal of Molecular Biology* **63**(1), 57–73.

Brown, J. W. and Shaw, P. J.: 1998, Small nucleolar rnas and pre-rrna processing in plants, *Plant Cell* **10**(5), 649–657.

Brownell, E., Krystal, M. and Arnheim, N.: 1983, Structure and evolution of human and african ape rdna pseudogenes., *Molecular Biology and Evolution* **1**(1), 29–37.

Brugere, J.-F., Cornillot, E., Metenier, G., Bensimon, A. and Vivares, C. P.: 2000, Encephalitozoon cuniculi (microspora) genome: physical map and evidence for telomere-associated rdna units on all chromosomes, *Nucleic Acids Research* **28**(10), 2026–2033.

Bryan, T., Englezou, A., Dalla-Pozza, L., Dunham, M. and Reddel, R.: 1997, Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines, *Nature Medicine* **3**, 1271–1274.

Bryan, T., Englezou, A., Gupta, J., Bacchetti, S. and Reddel, R.: 1995, Telomere elongation in immortal human cells without detectable telomerase activity, *EMBO Journal* **14**(17), 4240–4248.

Buckler, E. S. t. and Holtsford, T. P.: 1996, Zea systematics: ribosomal its evidence, *Molecular Biology and Evolution* **13**(4), 612–622.

Buckler, E. S. t., Ippolito, A. and Holtsford, T. P.: 1997, The evolution of ribosomal dna: divergent paralogues and phylogenetic implications, *Genetics* **145**(3), 821–832.

Buntjer, J. B. and Lenstra, J. A.: 1998, Self-amplification of satellite dna in vitro, *Genome* **41**(3), 429–434.

Buzek, J., Koutnikova, H., Houben, A., Riha, K., Janousek, B., Siroky, J., Grant, S. and Vyskot, B.: 1997, Isolation and characterization of x chromosome-derived dna sequences from a dioecious plant melandrium album, *Chromosome Research* **5**(1), 57–65.

Cai, Q., Zhang, D., Liu, Z.-L. and Wang, X.-R.: 2006, Chromosomal localization of 5s and 18s rdna in five species of subgenus strobus and their implications for genome evolution of pinus, *Annals of Botany* **97**(5), 715–722.

Campbell, J. and Newton, C.: 1991, *The Molecular and Cellular Biology of Yeast Saccharomyces*, Vol. 1: Genome Dynamics, Protein Synthesis, and Energetics, Cold Spring Harbor Laboratory Press, New York.

Casacubierta, E. and Pardue, M.: 2003a, Het-a elements in drosophila virilis: retrotransposon telomeres conserved across the drosophila genus, *Proceedings of the National Academy of Sciences USA* **100**, 14091–14096.

Casacubierta, E. and Pardue, M.: 2003b, Transposon telomeres are widely distributed in the drosophila genus: Tart elements in the virilis group, *Proceedings of the National Academy of Sciences USA* **100**, 3363–3368.

Cesare, A. J., Quinney, N., Willcox, S., Subramanian, D. and Griffith, J. D.: 2003, Telomere looping in p. sativum (common garden pea), *The Plant Journal* **36**(2), 271–279.

Chen, C. Y., Graham, J. and Yan, H.: 2001, Evidence for a replication function of ffa-1, the xenopus orthologue of werner syndrome protein, *The Journal of Cell Biology* **152**(5), 985–996.

Cheok, C. F., Wu, L., Garcia, P. L., Janscak, P. and Hickson, I. D.: 2005, The bloom's syndrome helicase promotes the annealing of complementary single-stranded dna, *Nucleic Acids Research* **33**(12), 3932–3941.

Chinnappa, C. C., C. and Basappa, G.: 1986, Cytology of some western canadian allium species, *American Journal of Botany* **73**(4), 529–534.

Chung, M.-C., Lee, Y.-I., Cheng, Y.-Y., Chou, Y.-J. and Lu, C.-F.: 2008, Chromosomal polymorphism of ribosomal genes in the genus oryza, *Theoretical and Applied Genetics* **116**(6), 745–753.

Clevinger, J. and Panero, J.: 2000, Phylogenetic analysis of silphium and subtribe engelmanniinae (asteraceae: Heliantheae) based on its and ets sequence data, *American Journal of Botany* **87**(4), 565–572.

Coen, E. S., Thoday, J. M. and Dover, G.: 1982, Rate of turnover of structural variants in the rdna gene family of drosophila melanogaster, *Nature* **295**(5850), 564–568.

Coen, E., Strachan, T. and Dover, G.: 1982, Dynamics of concerted evolution of ribosomal dna and histone gene families in the melanogaster species subgroup of drosophila, *Journal of Molecular Biology* **158**(1), 17–35.

Cohen, S., Agmon, N., Yacobi, K., Mislovati, M. and Segal, D.: 2005, Evidence for rolling circle replication of tandem genes in drosophila, *Nucleic Acids Research* **33**(14), 4519–4526.

Cohen, S., Houben, A. and Segal, D.: 2008, Extrachromosomal circular dna derived from tandemly repeated genomic sequences in plants, *Plant Journal* **53**(6), 1027–1034.

Cohen, S., Yacobi, K. and Segal, D.: 2003, Extrachromosomal circular dna of tandemly repeated genomic sequences in drosophila, *Genome Research* **13**(6A), 1133–1145.

Cohn, M., McEachern, M. J. and Blackburn, E. H.: 1998, Telomeric sequence diversity within the genus saccharomyces, *Current genetics* **33**(2), 83–91.

Compton, S. A., Choi, J.-H., Cesare, A. J., Ozgur, S. and Griffith, J. D.: 2007, Xrcc3 and nbs1 are required for the production of extrachromosomal telomeric circles in human alternative lengthening of telomere cells, *Cancer Research* **67**(4), 1513–1519.

Copenhaver, G. P. and Pikaard, C. S.: 1996, Rflp and physical mapping with an rdna-specific endonuclease reveals that nucleolus organizer regions of arabidopsis thaliana adjoin the telomeres on chromosomes 2 and 4, *The Plant Journal* **9**(2), 259–272.

Cordesse, F., Cooke, R., Tremousaygue, D., Grellet, F. and Delseny, M.: 1993, Fine structure and evolution of the rdna intergenic spacer in rice and other cereals, *Journal of Molecular Evolution* **36**(4), 369–379.

Crease, T. and Lynch, M.: 1991, Ribosomal dna variation in daphnia pulex, *Molecular Biology and Evolution* **8**, 620–640.

Cunado, N., Sanchez-Moran, E., Barrios, J. and Santos, J. L.: 2001, Searching for telomeric sequences in two allium species., *Genome* **44**(4), 640–643.

Da Rocha, P. S. and Bertrand, H.: 1995, Structure and comparative analysis of the rdna intergenic spacer of brassica rapa. implications for the function and evolution of the cruciferae spacer, *European Journal of Biochemistry* **229**(2), 550–557.

Dadejova, M., Lim, K. Y., Souckova-Skalicka, K., Matyasek, R., Grandbastien, M.-A., Leitch, A. and Kovarik, A.: 2007, Transcription activity of rrna genes correlates with a tendency towards intergenomic homogenization in nicotiana allotetraploids, *New Phytologist* **174**(3), 658–668.

Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O. and Bates, P. J.: 2003, Biophysical and biological properties of quadruplex oligodeoxyribonucleotides, *Nucleic Acids Research* **31**(8), 2097–2107.

Datson, P. M. and Murray, B. G.: 2006, Ribosomal dna locus evolution in nemesia: transposition rather than structural rearrangement as the key mechanism?, *Chromosome Research* **14**(8), 845–857.

Davison, J., Tyagi, A. and Comai, L.: 2007, Large-scale polymorphism of heterochromatic repeats in the dna of arabidopsis thaliana, *BMC Plant Biology* **7**, 44.

Delcasso-Tremousaygue, D., Grellet, F., Panabieres, F., Ananiev, E. D. and Delseny, M.: 1988, Structural and transcriptional characterization of the external spacer of a ribosomal rna nuclear gene from a higher plant, *European Journal of Biochemistry* **172**(3), 767–776.

Dellaporta, S., Wood, J. and Hicks, J.: 1983, Plant dna minipreparation: version ii, *Plant Molecular Biology Reporter* **1**, 19–21.

Dempsey, L. A., Sun, H., Hanakahi, L. A. and Maizels, N.: 1999, G4 dna binding by lr1 and its subunits, nucleolin and hnrnp d, a role for g-g pairing in immunoglobulin switch recombination, *Journal Biological Chemistry* **274**(2), 1066–1071.

Dickson, K. R., Braaten, D. C. and Schlessinger, D.: 1989, Human ribosomal dna: conserved sequence elements in a 4.3-kb region downstream from the transcription unit, *Gene* **84**(1), 197–200.

Díez, J. L., Vilariño, V. R., Medina, F. and Morcillo, G.: 2006, Nucleolar localization of a reverse transcriptase related to telomere maintenance in chironomus (diptera), *Histochemistry and Cell Biology* pp. 1–8.

Do, G. S., Suzuki, G. and Mukai, Y.: 2004, Genomic organization of a novel root alliinase gene, all1, in onion, *Gene* **325**, 17–24.

Doelling, J. H., Gaudino, R. J. and Pikaard, C. S.: 1993, Functional analysis of arabidopsis thaliana rrna gene and spacer promoters in vivo and by transient expression, *PNAS* **90**(16), 7528–7532.

Doelling, J. H. and Pikaard, C. S.: 1995, The minimal ribosomal rna gene promoter of arabidopsis thaliana includes a critical element at the transcription initiation site, *The Plant Journal* **8**(5), 683–692.

Doelling, J. H. and Pikaard, C. S.: 1996, Species-specificity of rrna gene transcription in plants manifested as a switch in rna polymerase specificity, *Nucleic Acids Research* **24**(23), 4725–4732.

Dover, G. A.: 1982, Molecular drive a cohesive mode of species evolution, *Nature* **299**, 111–117.

Dover, G. A.: 1984, Evolutionary unease?, *Nature* **309**(5969), 588.

Dover, G. A., Brown, S., Coen, E., Dallas, J., Strachan, T. and Trick, M.: 1982, *Genome Evolution*, Vol. 20, Academic Press, chapter The dynamics of genome evolution and species differentiation, pp. 343–371.

Dover, G. and Coen, E.: 1981, Springcleaning ribosomal dna: a model for multigene evolution?, *Nature* **290**(5809), 731–732.

Drummond, A., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T. and Wilson, A.: 2007, Geneious v3.0, http://www.geneious.com/.

Dubcovsky, J. and Dvorak, J.: 1995, Ribosomal rna multigene loci: nomads of the triticeae genomes, *Genetics* **140**(4), 1367–1377.

Dubouzet, J. G. and Shinoda, K.: 1999, Relationships among old and new world alliums according to its dna sequence analysis, *Theoretical and Applied Genetics* **98**(3-4), 422–433.

Dunham, M. A., Neumann, A. A., Fasching, C. L. and Reddel, R. R.: 2000, Telomere maintenance by recombination in human cells, *Nature Genetics* **26**(4), 447–450.

Durante, M., Cremonini, R., Brunori, A., Avanzi, S. and Innocenti, A. M.: 1977, Differentiation of metaxylem cell line in the root of allium cepa, *Protoplasma* **93**(2-3), 289–303.

Durante, M., Frediani, M., Mariani, L., Citti, L., Geri, C. and Cremonini, R.: 1990, Differentiation of metaxylem cell line in the root of allium cepa l., *Protoplasma* **158**(3), 149–154.

Dvorak, J., Jue, D. and Lassner, M.: 1987, Homogenization of tandemly repeated nucleotide sequences by distance-dependent nucleotide sequence conversion, *Genetics* **116**(3), 487–498.

Echeverria, M., Delcasso-Tremousaygue, D. and Delseny, M.: 1992, A nuclear protein fraction binding to da/dt-rich sequences upstream from the radish rdna promoter, *The Plant Journal* **2**(2), 211–219.

Eichinger, L., Pachebat, J. A., Glockner, G., Rajandream, M.-A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B. A., Rivero, F., Bankier, A. T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M. A., Urushihara, H., Hernandez, J., Rabbinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E. C., Chisholm, R. L., Gibbs, R., Loomis, W. F., Platzer, M., Kay, R. R., Williams, J., Dear, P. H., Noegel, A. A., Barrell, B. and Kuspa, A.: 2005, The genome of the social amoeba dictyostelium discoideum, *Nature* **435**(7038), 43–57.

Eickbush, T. H.: 1997, Telomerase and retrotransposons: which came first?, *Science* **277**(5328), 911–912.

Eirin-Lopez, J. M., Gonzalez-Tizon, A. M., Martinez, A. and Mendez, J.: 2004, Birth-and-death evolution with strong purifying selection in the histone h1 multigene family and the origin of orphon h1 genes, *Molecular Biology and Evolution* **21**(10), 1992–2003.

El-Gadi, A. and Elkington, T. T.: 1975, Comparison of the giemsa c-band karyotypes and the relationships of allium cepa, a. fistulosum and a. galanthum, *Chromosoma* **51**(1), 19–23.

Elder, J. F. J. and Turner, B. J.: 1995, Concerted evolution of repetitive dna sequences in eukaryotes, *The Quarterly Review of Biology* **70**(3), 297–320.

Ellis, N. A., Groden, J., Ye, T. Z., Straughen, J., Lennon, D. J., Ciocci, S., Proytcheva, M. and German, J.: 1995, The bloom's syndrome gene product is homologous to recq helicases, *Cell* **83**(4), 655–666.

Ellis, T. H. N., Davies, D. R., Castleton, J. A. and Bedford, I. D.: 1984, The organization and genetics of rdna length variants in peas, *Chromosoma* **91**(1), 74–81.

Emery, H. and Weiner, A.: 1981, An irregular satellite sequence is found at the termini of the linear extrachromosomal rdna in dictyostelium discoideum, *Cell* **26**, 411–419.

Epstein, C. J. and Motulsky, A. G.: 1996, Werner syndrome: entering the helicase era, *Bioessays* **18**(12), 1025–1027.

Fajkus, J., Kovarik, A., Kralovics, R. and Bezdek, M.: 1995, Organization of telomeric and subtelomeric chromatin in the higher plant nicotiana tabacum, *Molecular and General Genetics* **247**(5), 633–638.

Fajkus, J., Leitch, A. R., Chester, M. and Sýkorová, E.: 2007, *Evolution, composition and interrelated functions of telomeres and subtelomeres: Lessons from plants*, Landes Bioscience, pp. 114–127.

Fajkus, J., Novotná, M. and Ptáček, J.: 2002, Analysis of chromosome termini in potato varieties, *Rostlinna Vyroba* **48**(11), 477–479.

Fajkus, J., Sýkorová, E. and Leitch, A. R.: 2005, Telomeres in evolution and evolution of telomeres, *Chromosome Research* **13**(5), 469–479.

Fay, M., Rudall, P., Stobart, K., de Bruijn, A., Reeves, G., Qamarus-Zaman, F., Hong, W.-P., Joseph, J., Hahn, W., Conran, J. and Chase, M.: 2000, *Phylogenetic studies of Asparagales based on four plastid DNA regions*, CSRIO, Melbourne, Australia.

Feschotte, C. and Pritham, E. J.: 2007, Dna transposons and the evolution of eukaryotic genomes, *Annual Review of Genetics* **41**, 331–368.

Figueiredo, L. M., Rocha, E. P. C., Mancio-Silva, L., Prevost, C., Hernandez-Verdun, D. and Scherf, A.: 2005, The unusually large plasmodium telomerase reverse-transcriptase localizes in a discrete compartment associated with the nucleolus, *Nucleic Acids Research* **33**(3), 1111–1122.

Finnegan, D. J.: 1989, Eukaryotic transposable elements and genome evolution, *Trends in Genetics* **5**(4), 103–107.

Fisher, T. S. and Zakian, V. A.: 2005, Ku: a multifunctional protein involved in telomere maintenance, *DNA Repair (Amst)* **4**(11), 1215–1226.

Flavell, R.: 1980, The molecular characterization and organization of plant chromosomal dna sequences, *Annual Review of Plant Physiology* **31**(569-596).

Flavell, R., O'Dell, M., Sharp, P., Nevo, E. and Beiles, A.: 1986, Variation in the intergenic spacer of ribosomal dna of wild wheat, triticum dicoccoides. in israel, *Molecular Biology and Evolution* **3**(6), 547–558.

Franck, J. P., Harris, A. S., Bentzen, P., Denovan-Wright, E. M. and Wright, J. M.: 1991, Organization and evolution of satellite, minisatellite and microsatellite dnas in teleost fishes, *Oxford Surveys on Eukaryotic Genes* **7**, 51–82.

Fransz, P. F., Alonso-Blanco, C., Liharska, T. B., Peeters, A. J., Zabel, P. and de Jong, J. H.: 1996, High-resolution physical mapping in arabidopsis thaliana and tomato by fluorescence in situ hybridization to extended dna fibres, *The Plant Journal* **9**(3), 421–430.

Friebe, B.: 1989, Nucleolar activity of b-chromosomes in allium cernuum (alliaceae), *Plant Systematics and Evolution* **163**(1-2), 87–92.

Friesen, N., Borisjuk, N., Mes, T. H. M., Klaas, M. and Hanelt, P.: 1997, Allotetraploid origin of allium altyncolicum (alliaceae, allium sect schoenoprasum) as investigated by karyological and molecular markers, *PLANT SYSTEMATICS AND EVOLUTION* **206**(1-4), 317–335.

Friesen, N., Fritsch, R. and Blattner, F.: 2006, *Phylogeny and new intrageneric classification of Allium (Alliaceae) based on nuclear rDNA ITS sequences*, Vol. 22, Aliso, pp. 372–395.

Frydrychova, R., Grossmann, P., Trubac, P., Vitkova, M. and Marec, F.: 2004, Phylogenetic distribution of ttagg telomeric repeats in insects, *Genome* **47**(1), 163–178.

Frydrychova, R. and Marec, F.: 2002, Repeated losses of ttagg telomere repeats in evolution of beetles (coleoptera), *Genetica* **115**(2), 179–187.

Fuchs, J., Brandes, A. and Schubert, I.: 1995, Telomere sequence localization and karyotype evolution in higher plants, *Plant Systematics and Evolution* **196**, 227–241.

Fujisawa, M., Yamagata, H., Kamiya, K., Nakamura, M., Saji, S., Kanamori, H., Wu, J., Matsumoto, T. and Sasaki, T.: 2006, Sequence comparison of distal and proximal ribosomal dna arrays in rice (oryza sativa l.) chromosome 9s and analysis of their flanking regions, *Theoretical and Applied Genetics* **113**(3), 419–428.

Fukunaga, K., Ichitani, K., Taura, S., Sato, M. and Kawase, M.: 2005, Ribosomal dna intergenic spacer sequence in foxtail millet, setaria italica (l.) p. beauv. and its characterization and application to typing of foxtail millet landraces, *Hereditas* **142**(2005), 38–44.

Fulneckova, J. and Fajkus, J.: 2000, Inhibition of plant telomerase by telomere-binding proteins from nuclei of telomerase-negative tissues, *FEBS Letters* **467**(2-3), 305–310.

Fumagalli, L., Taberlet, P., Stewart, D. T., Gielly, L., Hausser, J. and Vogel, P.: 1999, Molecular phylogeny and evolution of sorex shrews (soricidae: insectivora) inferred from mitochondrial dna sequence data, *Molecular Phylogenetics and Evolution* **11**(2), 222–235.

Fussell, C. P.: 1975, The position of interphase chromosomes and late replicating dna in centromere and telomere regions of allium cepa l., *Chromosoma* **50**(2), 201–210.

Fussell, C. P.: 1977, Telomere associations in interphase nuclei of allium cepa demonstrated by c-banding, *Experimental Cell Research* **110**(1), 111–117.

Fussell, C. P.: 1992, Rabl distribution of interphase and prophase telomeres in allium cepa not altered by colchicine and/or ultracentrifugation, *American Journal of Botany* **79**(7), 771–777.

Gabrielian, A. and Pongor, S.: 1996, Correlation of intrinsic dna curvature with dna property periodicity, *FEBS Letters* **393**(1), 65–68.

Ganal, M. W., Lapitan, N. and Tanksley, S. D.: 1991, Macrostructure of the tomato telomeres, *The Plant Cell* **3**(1), 87–94.

Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. and Rothstein, R.: 1994, The yeast type i topoisomerase top3 interacts with sgs1, a dna helicase homolog: a potential eukaryotic reverse gyrase, *Molecular and Cellular Biology* **14**(12), 8391–8398.

Gangloff, S., Zou, H. and Rothstein, R.: 1996, Gene conversion plays the major role in controlling the stability of large tandem repeats in yeast, *EMBO Journal* **15**(7), 1715–1725.

Ganley, A. R. D. and Kobayashi, T.: 2007, Highly efficient concerted evolution in the ribosomal dna repeats: Total rdna repeat variation revealed by whole-genome shotgun sequence data, *Genome Research* **17**(2), 184–191.

Ganley, A. R. D. and Scott, B.: 1998, Extraordinary ribosomal spacer length heterogeneity in a neotyphodium endophyte hybrid: implications for concerted evolution, *Genetics* **150**(4), 1625–1637.

Garcia, P. L., Liu, Y., Jiricny, J., West, S. C. and Janscak, P.: 2004, Human recq5beta, a protein with dna helicase and strand-annealing activities in a single polypeptide, *EMBO Journal* **23**(14), 2882–2891.

Garcia, S., Lim, K. Y., Chester, M., Garnatje, T., Pellicer, J., Valles, J., Leitch, A. R. and Kovarik, A.: 2009, Linkage of 35s and 5s rrna genes in artemisia (family asteraceae): first evidence from angiosperms, *Chromosoma* **118**(1), 85–97.

Garrido, M. A., Jamilena, M., Lozano, R., Ruiz Rejon, C., Ruiz Rejon, M. and Parker, J. S.: 1994, rdna site number polymorphism and nor inactivation in natural populations of allium schoenoprasum, *Genetica* **94**(1), 67–71.

Garrido-Ramos, M. A., Jamilena, M., Lozano, R., Ruiz Rejun, C. and Ruiz Rejun, M.: 1992, A cytogenetical and molecular analysis of ribosomal cistrons of allium sphaerocephalon l. (liliaceae), *Heredity* **69**, 43–49.

Gazdova, B., Siroky, J., Fajkus, J., Brzobohaty, B., Kenton, A., Parokonny, A., Heslop-Harrison, J. S., Palme, K. and Bezdek, M.: 1995, Characterization of a new family of tobacco highly repetitive dna, grs, specific for the nicotiana tomentosiformis genomic component, *Chromosome Research* **3**(4), 245–254.

Gerlach, W. L. and Bedbrook, J. R.: 1979, Cloning and characterization of ribosomal rna genes from wheat and barley., *Nucleic Acids Research* **7**(7), 1869–1885.

Gillespie, J. J.: 2004, Characterizing regions of ambiguous alignment caused by the expansion and contraction of hairpin-stem loops in ribosomal rna molecules, *Molecular Phylogenetics and Evolution* **33**(3), 936–943.

Ginisty, H., Amalric, F. and Bouvet, P.: 1998, Nucleolin functions in the first step of ribosomal rna processing, *EMBO Journal* **17**(5), 1476–1486.

Ginisty, H., Sicard, H., Roger, B. and Bouvet, P.: 1999, Structure and functions of nucleolin, *Journal of Cell Science* **112**((Pt 6)), 761–772.

Gonzalez, I. L., Petersen, R. and Sylvester, J. E.: 1989, Independent insertion of alu elements in the human ribosomal spacer and their concerted evolution, *Molecular Biology and Evolution* **6**(4), 413–423.

Gonzalez, I. L. and Sylvester, J. E.: 1997a, Beyond ribosomal dna: on towards the telomere, *Chromosoma* **105**(7-8), 431–437.

Gonzalez, I. L. and Sylvester, J. E.: 1997b, Incognito rrna and rdna in databases and libraries, *Genome Research* **7**(1), 65–70.

Gonzalez, I. L. and Sylvester, J. E.: 2001, Human rdna: evolutionary patterns within the genes and tandem arrays derived from multiple chromosomes, *Genomics* **73**(3), 255–263.

Gonzalez, I. L., Tugendreich, S., Hieter, P. and Sylvester, J. E.: 1993, Fixation times of retroposons in the ribosomal dna spacer of human and other primates, *Genomics* **18**(1), 29–36.

Gonzalez, I. L., Wu, S., Li, W. M., Kuo, B. A. and Sylvester, J. E.: 1992, Human ribosomal rna intergenic spacer sequence, *Nucleic Acids Research* **20**(21), 5846.

Gonzalez-Melendi, P., Wells, B., Beven, A. F. and Shaw, P. J.: 2001, Single ribosomal transcription units are linear, compacted christmas trees in plant nucleoli, *The Plant Journal* **27**(3), 223–233.

Good-Avila, S. V., Souza, V., Gaut, B. S. and Eguiarte, L. E.: 2006, Timing and rate of speciation in agave (agavaceae), *Proceedings of the National Academy of Science USA* **103**(24), 9124–9129.

Gordienko, I. and Rupp, W. D.: 1997, The limited strand-separating activity of the uvrab protein complex and its role in the recognition of dna damage, *EMBO Journal* **16**(4), 889–895.

Greider, C. W.: 1999, Telomeres do d-loop-t-loop, *Cell* **97**(4), 419–422.

Greider, C. W. and Blackburn, E. H.: 1985, Identification of a specific telomere terminal transferase activity in tetrahymena extracts, *Cell* **43**(2 Pt 1), 405–413.

Greider, C. W. and Blackburn, E. H.: 1987, The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity, *Cell* **51**(6), 887–898.

Greider, C. W. and Blackburn, E. H.: 1989, A telomeric sequence in the rna of tetrahymena telomerase required for telomere repeat synthesis, *Nature* **337**(6205), 331–337.

Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H. and de Lange, T.: 1999, Mammalian telomeres end in a large duplex loop, *Cell* **97**(4), 503–514.

Groff-Vindman, C., Cesare, A. J., Natarajan, S., Griffith, J. D. and McEachern, M. J.: 2005, Recombination at long mutant telomeres produces tiny single- and double-stranded telomeric circles, *Molecular and Cellular Biology* **25**(11), 4406–4412.

Gruendler, P., Unfried, I., Pascher, K. and Schweizer, D.: 1991, rdna intergenic region from arabidopsis thaliana. structural analysis, intraspecific variation and functional implications, *Journal of Molecular Biology* **221**(4), 1209–1222.

Grunstein, M.: 1997, Molecular model for telomeric heterochromatin in yeast, *Current opinion in cell biology* **9**(3), 383–387.

Guerra, M.: 2000, Patterns of heterochromatin distribution in plant chromosomes, *Genetics and Molecular Biology* **23**(4), 1029–1049.

Guindon, S. and Gascuel, O.: 2003, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, *Systematic biology* **52**(5), 696–704.

Hakin-Smith, V., Jellinek, D. A., Levy, D., Carroll, T., Teo, M., Timperley, W. R., McKay, M. J., Reddel, R. R. and Royds, J. A.: 2003, Alternative lengthening of telomeres and survival in patients with glioblastoma multiforme, *Lancet* **361**(9360), 836–838.

Hall, K. J. and Parker, J. S.: 1995, Stable chromosome fission associated with rdna mobility, *Chromosome Research* **3**(7), 417–422.

Hall, S. E., Kettler, G. and Preuss, D.: 2003, Centromere satellites from arabidopsis populations: Maintenance of conserved and variable domains, *Genome Research* **13**(2), 195–205.

Hall, S. E., Luo, S., Hall, A. E. and Preuss, D.: 2005, Differential rates of local and global homogenization in centromere satellites from arabidopsis relatives, *Genetics* **170**(4), 1913–1927.

Hanada, K., Iwasaki, M., Ihashi, S. and Ikeda, H.: 2000, Uvra and uvrb suppress illegitimate recombination: synergistic action with recq helicase, *Proceedings of the National Academy of Science USA* **97**(11), 5989–5994.

Hanada, K., Ukita, T., Kohno, Y., Saito, K., Kato, J. and Ikeda, H.: 1997, Recq dna helicase is a suppressor of illegitimate recombination in escherichia coli, *Proceedings of the National Academy of Science USA* **94**(8), 3860–3865.

Hanakahi, L. A., Sun, H. and Maizels, N.: 1999, High affinity interactions of nucleolin with g-g-paired rdna, *Journal of Biological Chemistry* **274**(22), 15908–15912.

Harper, L., Golubovskaya, I. and Cande, W. Z.: 2004, A bouquet of chromosomes, *Journal of Cell Science* **117**(18), 4025–4032.

Hartung, F., Plchova, H. and Puchta, H.: 2000, Molecular characterisation of recq homologues in arabidopsis thaliana, *Nucleic Acids Research* **28**(21), 4275–4282.

Hasegawa, M., Kishino, H. and Yano, T.: 1985, Dating of the human-ape splitting by a molecular clock of mitochondrial dna, *Journal of molecular evolution* **22**(2), 160–174.

Hasterok, R., Wolny, E., Kulak, S., Zdziechiewicz, A., Maluszynska, J. and Heneen, W. K.: 2005, Molecular cytogenetic analysis of brassica rapa-brassica oleracea var. alboglabra monosomic addition lines, *Theoretical and Applied Genetics* **111**(2), 196–205.

Hawkins, J. S., Kim, H., Nason, J. D., Wing, R. A. and Wendel, J. F.: 2006, Differential lineage-specific amplification of transposable elements is responsible for genome size variation in gossypium, *Genome Research* **16**(10), 1252–1261.

Heacock, M., Spangler, E., Riha, K., Puizina, J. and Shippen, D. E.: 2004, Molecular analysis of telomere fusions in arabidopsis: multiple pathways for chromosome end-joining, *EMBO Journal* **23**(11), 2304–2313.

Heller, K., Kilian, A., Piatyszek, M. A. and Kleinhofs, A.: 1996, Telomerase activity in plant extracts, *Molecular Genetics and Genomics* **252**(3), 342–345.

Hemleben, V., Leweke, B., Roth, A. and Stadler, J.: 1982, Organization of highly repetitive satellite dna of two cucurbitaceae species (cucumis melo and cucumis sativus), *Nucleic Acids Research* **10**(2), 631–644.

Henderson, E. R. and Blackburn, E. H.: 1989, An overhanging 3' terminus is a conserved feature of telomeres, *Molecular and Cellular Biology* **9**(1), 345–348.

Henderson, E. R., Moore, M. and Malcolm, B. A.: 1990, Telomere g-strand structure and function analyzed by chemical protection, base analogue substitution, and utilization by telomerase in vitro, *Biochemistry* **29**(3), 732–737.

Hernandez, P., Martin-Parras, L., Martinez-Robles, M. L. and Schvartzman, J. B.: 1993, Conserved features in the mode of replication of eukaryotic ribosomal rna genes, *EMBO Journal* **12**(4), 1475–1485.

Heslop-Harrison, J. S.: 2000, Comparative genome organization in plants: From sequence and markers to chromatin and chromosomes, *The Plant Cell* **12**(5), 617–636.

Hibner, B., Burke, W. and Eickbush, T.: 1991, Sequence identity in an early chorion multigene family is the result of localized gene conversion, *Genetics* **128**(3), 595–606.

Higashiyama, T., Maki, S. and Yamada, T.: 1995, Molecular organization of chlorella vulgaris chromosome i: presence of telomeric repeats that are conserved in higher plants, *Molecular and General Genetics* **246**(1), 29–36.

Higashiyama, T., Noutoshi, Y., Fujie, M. and Yamada, T.: 1997, Zepp, a line-like retrotransposon accumulated in the chlorella telomeric region, *EMBO Journal* **16**(12), 3715–3723.

Hillis, D. M. and Dixon, M. T.: 1991, Ribosomal dna: molecular evolution and phylogenetic inference, *The Quarterly Review of Biology* **66**(4), 411–453.

Hillis, D. M., Moritz, C., Porter, C. A. and Baker, R. J.: 1991, Evidence for biased gene conversion in concerted evolution of ribosomal dna, *Science* **251**(4991), 308–310.

Houben, A., Kynast, R. G., Heim, U., Hermann, H., Jones, R. N. and Forster, J. W.: 1996, Molecular cytogenetic characterisation of the terminal heterochromatic segment of the b-chromosome of rye (secale cereale), *Chromosoma* **105**(2), 97–103.

Hourcade, D., Dressler, D. and Wolfson, J.: 1973, The amplification of ribosomal rna genes involves a rolling circle intermediate, *Proceedings of the National Academy of Science USA* **70**(10), 2926–2930.

Hu, Y., Raynard, S., Sehorn, M. G., Lu, X., Bussen, W., Zheng, L., Stark, J. M., Barnes, E. L., Chi, P., Janscak, P., Jasin, M., Vogel, H., Sung, P. and Luo, G.: 2007, Recql5/recql5 helicase regulates homologous recombination and suppresses tumor formation via disruption of rad51 presynaptic filaments, *Genes and Development* **21**(23), 3073–3084.

Huang, S., Beresten, S., Li, B., Oshima, J., Ellis, N. A. and Campisi, J.: 2000, Characterization of the human and mouse wrn 3'->5' exonuclease, *Nucleic Acids Research* **28**(12), 2396–2405.

Huber, M. D., Duquette, M. L., Shiels, J. C. and Maizels, N.: 2006, A conserved g4 dna binding domain in recq family helicases, *Journal of Molecular Biology* **358**(4), 1071–1080.

Huber, M. D., Lee, D. C. and Maizels, N.: 2002, G4 dna unwinding by blm and sgs1p: substrate specificity and substrate-specific inhibition, *Nucleic Acids Research* **30**(18), 3954–3961.

Hughes, A. L. and Nei, M.: 1989, Evolution of the major histocompatibility complex: independent origin of nonclassical class i genes in different groups of mammals, *Molecular Biology and Evolution* **6**(6), 559–579.

II, A. P. G.: 2003, An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: Apg ii, *Botanical Journal of the Linnean Society* **141**(4), 399–436.

Ingle, J., Pearson, G. and Sinclair, J.: 1973, Species distribution and properties of nuclear satellite dna in higher plants, *Nature New Biology* **242**, 193–197.

Ipek, M., Ipek, A., Almquist, S. G. and Simon, P. W.: 2005, Demonstration of linkage and development of the first low-density genetic map of garlic, based on aflp markers, *Theoretical and Applied Genetics* **110**(2), 228–236.

Ira, G., Malkova, A., Liberi, G., Foiani, M. and Haber, J. E.: 2003, Srs2 and sgs1-top3 suppress crossovers during double-strand break repair in yeast, *Cell* **115**(4), 401–411.

Irifune, K., Hirai, K., Zheng, J., Tanaka, R. and Morikawa, H.: 1995, Nucleotide sequence of a highly repeated dna sequence and its chromosomal localization in allium fistulosum, *Theoretical and Applied Genetics* **90**(3), 312–316.

Ishikawa, F., Matunis, M. J., Dreyfuss, G. and Cech, T. R.: 1993, Nuclear proteins that bind the pre-mrna 3' splice site sequence r(uuag/g) and the human telomeric dna sequence d(ttaggg)n, *Mol Cell Biol* **13**(7), 4301–4310.

Ishikawa, F. and Naito, T.: 1999, Why do we have linear chromosomes? a matter of adam and eve, *Mutation Research* **434**(2), 99–107.

Jakse, J., Telgmann, A., Jung, C., Khar, A., Melgar, S., Cheung, F., Town, C. D. and Havey, M. J.: 2006, Comparative sequence and genetic analyses of asparagus bacs reveal no microsynteny with onion or rice, *Theoretical and Applied Genetics* **114**(1), 31–39.

Jakubczak, J. L., Burke, W. D. and Eickbush, T. H.: 1991, Retrotransposable elements r1 and r2 interrupt the rrna genes of most insects, *Proceedings of the National Academy of Sciences USA* **88**(8), 3295–3299.

Jamilena, M., Garrido-Ramos, M., Ruiz Rejon, M., Ruiz Rejon, C. and Parker, J. S.: 1995, Characterisation of repeated sequences from microdissected b chromosomes of crepis capillaris, *Chromosoma* **104**(2), 113–120.

Janscak, P., Garcia, P. L., Hamburger, F., Makuta, Y., Shiraishi, K., Imai, Y., Ikeda, H. and Bickle, T. A.: 2003, Characterization and mutational analysis of the recq core of the bloom syndrome protein, *Journal of Molecular Biology* **330**(1), 29–42.

Jenderny, J., Caliebe, A., Beyer, C. and Grote, W.: 1993, Transmission of a ring chromosome 18 from a mother with 46,xx/47,xx, + r(18) mosaicism to her daughter, resulting in a 46,xx,r(18) karyotype, *Journal of Medical Genetics* **30**(11), 964–965.

Jones, R. and Rees, H.: 1968, Nuclear dna variation in allium, *Heredity* **23**, 591–605.

Jupe, E. R. and Zimmer, E. A.: 1993, Dnasei-sensitive and undermethylated rdna is preferentially expressed in a maize hybrid, *Plant Molecular Biology* **21**(5), 805–821.

Kalendar, R., Tanskanen, J., Chang, W., Antonius, K., Sela, H., Peleg, O. and Schulman, A. H.: 2008, Cassandra retrotransposons carry independently transcribed 5s rna, *Proceedings of the National Academy of Science USA* **105**(15), 5833–5838.

Kaliraman, V. and Brill, S. J.: 2002, Role of sgs1 and slx4 in maintaining rdna structure in saccharomyces cerevisiae, *Current Genetics* **41**(6), 389–400.

Kamnert, I., Lopez, C. C., Rosen, M. and Edstrom, J. E.: 1997, Telomeres terminating with long complex tandem repeats, *Hereditas* **127**(3), 175–180.

Kamnert, I., Nielsen, L. and Edström, J.-E.: 1998, A concertedly evolving region in chironomus, unique within the telomere, *Journal of Molecular Evolution* **46**(5), 562–570.

Kang, C., Zhang, X., Ratliff, R., Moyzis, R. and Rich, A.: 1992, Crystal structure of four-stranded oxytricha telomeric dna, *Nature* **356**(6365), 126–131.

Kannan, K., Nelson, A. D. L. and Shippen, D. E.: 2008, Dyskerin is a component of the arabidopsis telomerase rnp required for telomere maintenance, *Molecular and cellular biology* **28**(7), 2332–2341.

Kato, A., Yakura, K. and Tanifuji, S.: 1984, Sequence analysis of vicia faba repeated dna, the foki repeat element, *Nucleic Acids Research* **12**(16), 6415–6426.

Keller, I., Chintauan-Marquier, I. C., Veltsos, P. and Nichols, R. A.: 2006, Ribosomal dna in the grasshopper podisma pedestris: escape from concerted evolution, *Genetics* **174**(2), 863–874.

Keller, I., Veltsos, P. and Nichols, R. A.: 2008, The frequency of rdna variants within individuals provides evidence of population history and gene flow across a grasshopper hybrid zone, *Evolution* **62**(4), 833–844.

Kellogg, E. A. and Appels, R.: 1995, Intraspecific and interspecific variation in 5s rna genes are decoupled in diploid wheat relatives, *Genetics* **140**(1), 325–343.

Kempken, F.: 2001, Hideaway, a repeated element from ascobolus immersus, is rdna-associated and may resemble a retrotransposon, *Current Genetics* **40**(3), 179–185.

Khurts, S., Masutomi, K., Delgermaa, L., Arai, K., Oishi, N., Mizuno, H., Hayashi, N., Hahn, W. C. and Murakami, S.: 2004, Nucleolin interacts with telomerase, *The Journal of Biological Chemistry* **279**(49), 51508–51515.

Kidwell, M. and Lisch, D.: 2000, Transposable elements and host genome evolution, *Trends in Ecology and Evolution* **15**(3), 95–99.

Kilian, A., Stiff, C. and Kleinhofs, A.: 1995, Barley telomeres shorten during differentiation but grow in callus culture, *Proceedings of the National Academy of Sciences USA* **92**, 9555–9559.

Killan, A., Heller, K. and Kleinhofs, A.: 1998, Development patterns of telomerase activity in barley and maize, *Plant Molecular Biology* **37**(4), 621–628.

Kim, J., Cheong, C. and Moore, P. B.: 1991, Tetramerization of an rna oligonucleotide containing a gggg sequence, *Nature* **351**(6324), 331–332.

Kimura, M.: 1980, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, *Journal of Molecular Evolution* **16**(2), 111–120.

King, J. J., Bradeen, J. M., Bark, O., McCallum, J. A. and Havey, M. J.: 1998, A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome, *Theoretical and Applied Genetics* **96**(1), 52–62.

Kipling, D. and Warburton, P. E.: 1997, Centromeres, cenp-b and tigger too, *Trends in Genetics* **13**(4), 141–145.

Kitao, S., Ohsugi, I., Ichikawa, K., Goto, M., Furuichi, Y. and Shimamoto, A.: 1998, Cloning of two new human helicase genes of the recq family: biological significance of multiple species in higher eukaryotes, *Genomics* **54**(3), 443–452.

Klaas, M. and Friesen, N.: 2002, *Allium Crop Science: Recent Advances*, number 8, CAB International, chapter Molecular Markers in Allium, pp. 159–185.

Klein, H. L. and Petes, T. D.: 1981, Intrachromosomal gene conversion in yeast, *Nature* **289**(5794), 144–148.

Knälmann, M. and Burger, E.-C.: 1977, Cytologische lokalisation von 5s und 18/25s rna genorten in mitose-chromosomen von vicia faba, *Chromosoma* **61**(2), 177–192.

Kobayashi, T.: 2003, The replication fork barrier site forms a unique structure with fob1p and inhibits the replication fork, *Molecular and Cellular Biology* **23**(24), 9178–9188.

Kobayashi, T. and Ganley, A. R. D.: 2005, Recombination regulation by transcription-induced cohesin dissociation in rdna repeats, *Science* **309**(5740), 1581–1584.

Kobayashi, T., Heck, D. J., Nomura, M. and Horiuchi, T.: 1998, Expansion and contraction of ribosomal dna repeats in saccharomyces cerevisiae: requirement of replication fork blocking (fob1) protein and the role of rna polymerase i, *Genes and Development* **12**(24), 3821–3830.

Kojima, K. K. and Fujiwara, H.: 2005, Long-term inheritance of the 28s rdna-specific retrotransposon r2, *Molecular Biology and Evolution* **22**(11), 2157–2165.

Kojima, K. K., Kuma, K.-i., Toh, H. and Fujiwara, H.: 2006, Identification of rdna-specific non-ltr retrotransposons in cnidaria, *Molecular Biology and Evolution* **23**(10), 1984–1993.

Kornberg, R. D.: 1977, Structure of chromatin, *Annual Review of Biochemistry* **46**, 931–954.

Koukalova, B., Fojtova, M., Lim, K. Y., Fulnecek, J., Leitch, A. R. and Kovarik, A.: 2005, Dedifferentiation of tobacco cells is associated with ribosomal rna gene hypomethylation, increased transcription, and chromatin alterations., *Plant Physiol* **139**(1), 275–286.

Kovarik, A., Matyasec, R., Lim, K. Y., Skalicka, K., Koukalova, B., Knapp, S., Chase, M. and Leitch, A. R.: 2004, Concerted evolution of 18-5.8-26s rdna repeats in nicotiana allotetraploids, *Biological Journal of the Linnean Society* **82**(4), 615–625.

Kovarik, A., Pires, J. C., Leitch, A. R., Lim, K. Y., Sherwood, A. M., Matyasek, R., Rocca, J., Soltis, D. E. and Soltis, P. S.: 2005, Rapid concerted evolution of nuclear ribosomal dna in two tragopogon allopolyploids of recent and recurrent origin., *Genetics* **169**(2), 931–944.

Krishnan, H. B.: 1995, Nucleotide and primary sequences of a rice endosperm cdna are extensively homologous to elongation factor 1 beta', *Biochemical and Biophysical Research Communications* **209**(3), 1026–1031.

Kubo, Y., Okazaki, S., Anzai, T. and Fujiwara, H.: 2001, Structural and phylogenetic analysis of tras, telomeric repeat-specific non-ltr retrotransposon families in lepidopteran insects, *Molecular Biology and Evolution* **18**(5), 848–857.

Kuchar, M. and Fajkus, J.: 2004, Interactions of putative telomere-binding proteins in arabidopsis thaliana: identification of functional trf2 homolog in plants, *FEBS Letters* **578**(3), 311–315.

Kuhl, J. C., Cheung, F., Yuan, Q., Martin, W., Zewdie, Y., McCallum, J., Catanach, A., Rutherford, P., Sink, K. C., Jenderek, M., Prince, J. P., Town, C. D. and Havey, M. J.: 2004, A unique set of 11,008 onion expressed sequence tags reveals expressed sequence and genomic differences between the monocot orders asparagales and poales, *Plant Cell* **16**(1), 114–125.

Kumar, A. and Bennetzen, J. L.: 1999, Plant retrotransposons, *Annual Review of Genetics* **33**, 479–532.

Kumar, A., Pearce, S. R., McLean, K., Harrison, G., Heslop-Harrison, J. S., Waugh, R. and Flavell, A. J.: 1997, The ty1-copia group of retrotransposons in plants: genomic organisation, evolution, and use as molecular markers, *Genetica* **100**(1-3), 205–217.

Kwon, S.-J., Kim, D.-H., Lim, M.-H., Long, Y., Meng, J.-L., Lim, K.-B., Kim, J.-A., Kim, J. S., Jin, M., Kim, H.-I., Ahn, S.-N., Wessler, S. R., Yang, T.-J. and Park, B.-S.: 2007, Terminal repeat retrotransposon in miniature (trim) as dna markers in brassica relatives, *Molecular genetics and genomics* **278**(4), 361–370.

Lakshmikumaran, M. and Negi, M. S.: 1994, Structural analysis of two length variants of the rdna intergenic spacer from eruca sativa, *Plant Molecular Biology* **24**(6), 915–927.

Lamb, J. C., Yu, W., Han, F. and Birchler, J. A.: 2007, Plant chromosomes from end to end: telomeres, heterochromatin and centromeres, *Current Opinion in Plant Biology* **10**(2), 116–122.

Langford, C. F., Telenius, H., Carter, N. P., Miller, N. G. and Tucker, E. M.: 1992, Chromosome painting using chromosome-specific probes from flow-sorted pig chromosomes, *Cytogenetics and Cell Genetics* **61**(3), 221–223.

Larson, G. P., Castanotto, D., Rossi, J. J. and Malafa, M. P.: 1994, Isolation and functional analysis of a kluyveromyces lactis rap1 homologue, *Gene* **150**(1), 35–41.

Le Blancq, S. M., Korman, S. H. and Van der Ploeg, L. H. T.: 1991, Frequent rearrangements of rrna-encoding chromosomes in giardia lamblia, *Nucleic Acids Research* **19**(16), 4405–4412.

Lee, H.-R., Neumann, P., Macas, J. and Jiang, J.: 2006, Transcription and evolutionary dynamics of the centromeric satellite repeat cento in rice, *Molecular Biology and Evolution* **23**(12), 2505–2520.

Leitch, A. R.: 2007, Conserved gene order belies rapid genome turnover: the dynamic interplay between genomic dna and the outside world, *Heredity* **98**(2), 61–62.

Leitch, A. R., Schwarzacher, T., Jackson, D. and Leitch, I. J.: 1994, *In Situ Hybridisation*, BIOS Scientific Publishers, Oxford.

Levis, R. W., Ganesan, R., Houtchens, K., Tolar, L. A. and Sheen, F. M.: 1993, Transposons in place of telomeric repeats at a drosophila telomere, *Cell* **75**(6), 1083–1093.

Lim, K. Y., Kovarik, A., Matyasek, R., Bezdek, M., Lichtenstein, C. P. and Leitch, A. R.: 2000, Gene conversion of ribosomal dna in nicotiana tabacum is associated with undermethylated, decondensed and probably active gene units, *Chromosoma* **109**(3), 161–172.

Lim, K. Y., Kovarik, A., Matyasek, R., Chase, M. W., Clarkson, J. J., Grandbastien, M. A. and Leitch, A. R.: 2007, Sequence of events leading to near-complete genome turnover in allopolyploid nicotiana within five million years., *New Phytologist* **175**(4), 756–763.

Lim, K. Y., Skalicka, K., Koukalova, B., Volkov, R. A., Matyasek, R., Hemleben, V., Leitch, A. R. and Kovarik, A.: 2004, Dynamic changes in the distribution of a satellite homologous to intergenic 26-18s rdna spacer in the evolution of nicotiana, *Genetics* **166**(4), 1935–1946.

Lin, C.-Y., Chang, H.-H., Wu, K.-J., Tseng, S.-F., Lin, C.-C., Lin, C.-P. and Teng, S.-C.: 2005, Extrachromosomal telomeric circles contribute to rad52-, rad50-, and polymerase delta-mediated telomere-telomere recombination in saccharomyces cerevisiae, *Eukaryotic Cell* **4**(2), 327–336.

Lin, J., Jin, R., Zhang, B., Chen, H., Bai, Y. X., Yang, P. X., Han, S. W., Xie, Y. H., Huang, P. T., Huang, C. and Huang, J. J.: 2008, Nucleolar localization of tert is unrelated to telomerase function in human cells, *Journal of Cell Science* **121**(Pt 13), 2169–2176.

Linares, A. R., Bowen, T. and Dover, G. A.: 1994, Aspects of nonrandom turnover involved in the concerted evolution of intergenic spacers within the ribosomal dna of drosophila melanogaster, *Journal of molecular Evolution* **39**(2), 151–159.

Linder, C. R., Goertzen, L. R., Heuvel, B. V., Francisco-Ortega, J. and Jansen, R. K.: 2000, The complete external transcribed spacer of 18s-26s rdna: amplification and phylogenetic utility at low taxonomic levels in asteraceae and closely allied families, *Molecular Phylogenetics and Evolution* **14**(2), 285–303.

Linskens, M. H. and Huberman, J. A.: 1988, Organization of replication of ribosomal dna in saccharomyces cerevisiae, *Molecular Cell Biology* **8**(11), 4927–4935.

Little, R. D., Platt, T. H. and Schildkraut, C. L.: 1993, Initiation and termination of dna replication in human rrna genes, *Molecular Cell Biology* **13**(10), 6600–6613.

Liu, Y. J., Whelen, S. and Hall, B. D.: 1999, Phylogenetic relationships among ascomycetes: evidence from an rna polymerse ii subunit, *Molecular Biology and Evolution* **16**(12), 1799–1808.

Lo, A. W. I., Sabatier, L., Fouladi, B., Pottier, G., Ricoul, M. and Murnane, J. P.: 2002, Dna amplification by breakage/fusion/bridge cycles initiated by spontaneous telomere loss in a human cancer cell line, *Neoplasia* **4**(6), 531–538.

Lohe, A. R. and Roberts, P. A.: 1990, An unusual y chromosome of drosophila simulans carrying amplified rdna spacer without rrna genes, *Genetics* **125**(2), 399–406.

Londono-Vallejo, J. A., Der-Sarkissian, H., Cazes, L., Bacchetti, S. and Reddel, R. R.: 2004, Alternative lengthening of telomeres is characterized by high rates of telomeric exchange, *Cancer research* **64**(7), 2324–2327.

Long, E. O. and Dawid, I. B.: 1980, Repeated genes in eukaryotes, *Annual Review of Biochemistry* **49**, 727–764.

Lopez, C. C., Nielsen, L. and Edstrom, J. E.: 1996, Terminal long tandem repeats in chromosomes from chironomus pallidivittatus, *Molecular and Cellular Biology* **16**(7), 3285–3290.

Lopez-Estrano, C., Schvartzman, J. B., Krimer, D. B. and Hernandez, P.: 1999, Characterization of the pea rdna replication fork barrier: putative cis-acting and trans-acting factors, *Plant Molecular Biology* **40**(1), 99–110.

Lukowiak, A. A., Narayanan, A., Li, Z. H., Terns, R. M. and Terns, M. P.: 2001, The snorna domain of vertebrate telomerase rna functions to localize the rna within the nucleus, *RNA* **7**(12), 1833–1844.

Macas, J., Gualberti, G., Nouzova, M., Samec, P., Lucretti, S. and Dolezel, J.: 1996, Construction of chromosome-specific dna libraries covering the whole genome of field bean (vicia faba l.), *Chromosome Res* **4**(7), 531–539.

Macas, J., Meszaros, T. and Nouzova, M.: 2002, Plantsat: a specialized database for plant satellite repeats, *Bioinformatics* **18**(1), 28–35.

Macas, J., Navratilova, A. and Meszaros, T.: 2003, Sequence subfamilies of satellite repeats related to rdna intergenic spacer are differentially amplified on vicia sativa chromosomes, *Chromosoma* **112**(3), 152–158.

Macas, J., Neumann, P. and Navratilova, A.: 2007, Repetitive dna in the pea (pisum sativum l.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and medicago truncatula, *BMC Genomics* **8**, 427.

Macas, J., Pozarkova, D., Navratilova, A., Nouzova, M. and Neumann, P.: 2000, Two new families of tandem repeats isolated from genus vicia using genomic self-priming pcr, *Molecular and General Genetics* **263**(5), 741–751.

Maddison, D. R. and Maddison, W. P.: 2000, *MacClade 4.0: Analysis of Phylogeny and Character Evolution*, Sinauer Associates, Sunderland, MA.

Maggini, F., Barsanti, P. and Marazia, T.: 1978, Individual variation of the nucleolus organizer regions in allium cepa and a. sativum, *Chromosoma* **66**(2), 173–183.

Maggini, F., Cremonini, R., Zolfino, C., Tucci, G., D'Ovidio, R., Delre, V., DePace, C., Scarascia Mugnozza, G. and Cionini, P.: 1991, Structure and chromosomal localization of dna sequences related to ribosomal subrepeats in vicia faba, *Chromosoma* **100**(4), 229–234.

Maillet, G., White, C. I. and Gallego, M. E.: 2006, Telomere-length regulation in inter-ecotype crosses of arabidopsis, *Plant Molecular Biology* **62**(6), 859–866.

Maizels, N.: 2006, Dynamic roles for g4 dna in the biology of eukaryotic cells, *Nature Structural and Molecular Biology* **13**(12), 1055–1059.

Maluszynska, J. and Heslop-Harrison, J. S.: 1993, Physical mapping of rdna loci in brassica species, *Genome* **36**(4), 774–781.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. and Rothberg, J. M.: 2005, Genome sequencing in microfabricated high-density picolitre reactors, *Nature* **437**(7057), 376–380.

Martin, G. M., Sprague, C. A. and Epstein, C. J.: 1970, Replicative life-span of cultivated human cells. effects of donor's age, tissue, and genotype, *Laboratory Investigation Invest* **23**(1), 86–92.

Martin, W. J., McCallum, J., Shigyo, M., Jakse, J., Kuhl, J. C., Yamane, N., Pither-Joyce, M., Gokce, A. F., Sink, K. C., Town, C. D. and Havey, M. J.: 2005, Genetic mapping of expressed sequences in onion and in silico comparisons with rice show scant colinearity, *Molecular Genetics and Genomics* **274**(3), 197–204.

Martinez, J., Edström, J.-E., Morcillo, G. and Díez, J.: 2001, Telomeres in chironomus thummi are characterised by different subfamilies of complex dna repeats, *Chromosoma* **110**, 221–227.

McEachern, M. J. and Blackburn, E. H.: 1994, A conserved sequence motif within the exceptionall diverse telomeric sequences of budding yeasts, *Proceedings of the National Academy of Sciences USA* **91**, 3453–3457.

McEachern, M. J. and Haber, J. E.: 2006, Break-induced replication and recombinational telomere elongation in yeast, *Annual Review of Biochemistry* **75**, 111–135.

McEachern, M. J. and Hicks, J. B.: 1993, Unusually large telomeric repeats in the yeast candida albicans, *Molecular Cell Biology* **13**(1), 551–560.

McEachern, M. J. and Iyer, S.: 2001, Short telomeres in yeast are highly recombinogenic, *Molecular Cell* **7**(4), 695–704.

McGhee, J. D. and Felsenfeld, G.: 1980, Nucleosome structure, *Annual Review of Biochemistry* **49**, 1115–1156.

Melek, M. and Shippen, D. E.: 1996, Chromosome healing: spontaneous and programmed de novo telomere formation by telomerase, *Bioessays* **18**(4), 301–308.

Melnikova, L., Biessmann, H. and Georgiev, P.: 2005, The ku protein complex is involved in length regulation of drosophila telomeres, *Genetics* **170**(1), 221–235.

Meyer, I. M. and Miklos, I.: 2005, Statistical evidence for conserved, local secondary structure in the coding regions of eukaryotic mrnas and pre-mrnas, *Nucleic Acids Research* **33**(19), 6338–6348.

Minguez, A. and Moreno Diaz de la Espina, S.: 1996, In situ localization of nucleolin in the plant nucleolar matrix, *Experimental Cell Research* **222**(1), 171–178.

Mizuno, H., Wu, J., Kanamori, H., Fujisawa, M., Namiki, N., Saji, S., Katagiri, S., Katayose, Y., Sasaki, T. and Matsumoto, T.: 2006, Sequencing and characterization of telomere and subtelomere regions on rice chromosomes 1s, 2s, 2l, 6l, 7s, 7l and 8s, *The Plant Journal* **46**(2), 206–217.

Mizuno, H., Wu, J., Katayose, Y., Kanamori, H., Sasaki, T. and Matsumoto, T.: 2008a, Characterization of chromosome ends on the basis of the structure of trsa subtelomeric repeats in rice (oryza sativa l.), *Molecular Genetics and Genomics* **280**(1), 19–24.

Mizuno, H., Wu, J., Katayose, Y., Kanamori, H., Sasaki, T. and Matsumoto, T.: 2008b, Chromosome-specific distribution of nucleotide substitutions in telomeric repeats of rice (oryza sativa l.), *Molecular Biology and Evolution* **25**(1), 62–68.

Mlinarec, J., Chester, M., Siljak-Yakovlev, S., Papes, D., Leitch, A. and Besendorfer, V.: 2009, Molecular structure and chromosome distribution of three repetitive dna families in anemone hortensis l. (ranunculaceae), *Chromosome Research* .

Mongelard, F. and Bouvet, P.: 2007, Nucleolin: a multifaceted protein, *Trends in Cell Biology* **17**(2), 80–86.

Moss, T. and Stefanovsky, V. Y.: 1995, Promotion and regulation of ribosomal transcription in eukaryotes by rna polymerase i, *Progress in Nucleic Acid Research and Molecular Biology* **50**, 25–66.

Muller, A. E., Kamisugi, Y., Gruneberg, R., Niedenhof, I., Horold, R. J. and Meyer, P.: 1999, Palindromic sequences and a+t-rich dna elements promote illegitimate recombination in nicotiana tabacum, *Journal of Molecular Biology* **291**(1), 29–46.

Muller, H.: 1938, The remaking of chromosomes, *The Collecting Net* **13**, 181, 183–195, 198.

Munoz-Jordan, J. L., Cross, G. A., Lange, T. d. and Griffith, J. D.: 2001, t-loops at trypanosome telomeres, *EMBO Journal* **20**(3), 579–588.

Murray, M. G., Peters, D. L. and Thompson, W. F.: 1981, Ancient repeated sequences in the pea and mung bean genomes and implications for genome evolution, *Journal of Molecular Evolution* **17**(1), 31–42.

Nabetani, A. and Ishikawa, F.: 2009, Unusual telomeric dnas in human telomerase-negative immortalized cells, *Molecular and Cellular Biology* **29**(3), 703–713.

Naito, T., Matsuura, A. and Ishikawa, F.: 1998, Circular chromosome formation in a fission yeast mutant defective in two atm homologues, *Nature Genetics* **20**(2), 203–206.

Nakamura, T. M. and Cech, T. R.: 1998, Reversing time: origin of telomerase, *Cell* **92**(5), 587–590.

Nakamura, T. M., Cooper, J. P. and Cech, T. R.: 1998, Two modes of survival of fission yeast without telomerase, *Science* **282**(5388), 493–496.

Nam, J., Kim, J., Lee, S., An, G., Ma, H. and Nei, M.: 2004, Type i mads-box genes have experienced faster birth-and-death evolution than type ii mads-box genes in angiosperms, *Proceedings of the National Academy of Science USA* **101**(7), 1910–1915.

Nasmyth, K. A.: 1982, Molecular genetics of yeast mating type, *Annual Review of Genetics* **16**, 439–500.

Nassif, N., Penney, J., Pal, S., Engels, W. R. and Gloor, G. B.: 1994, Efficient copying of nonhomologous sequences from ectopic sites via p-element-induced gap repair, *Molecular and Cellular Biology* **14**(3), 1613–1625.

Natarajan, S. and McEachern, M. J.: 2002, Recombinational telomere elongation promoted by dna circles, *Molecular and Cellular Biology* **22**(13), 4512–4521.

Navratilova, A., Koblizkova, A. and Macas, J.: 2008, Survey of extrachromosomal circular dna derived from plant satellite repeats, *BMC Plant Biology* **8**, 90.

Nei, M.: 1987, *Molecular Evolutionary Genetics*, Columbia University Press, New York.

Nei, M., Gu, X. and Sitnikova, T.: 1997, Evolution by the birth-and-death process in multigene families of the vertebrate immune system, *Proceedings of the National Academy of Science USA* **94**(15), 7799–7806.

Nei, M. and Hughes, A.: 1991, Balanced polymorphism and evolution by the birth-and-death process in the mhc loci, *in* K. Tsuji, M. Aizawa and T. Sasazuli (eds), *Proc 11th Histocompatibility Workshop and Conference*, Vol. 2, Oxford University Press, pp. 27–38.

Nei, M. and Rooney, A. P.: 2005, Concerted and birth-and-death evolution of multigene families, *Annual Review of Genetics* **39**, 121–152.

Nguyen, N. H., Driscoll, H. E. and Specht, C. D.: 2008, A molecular phylogeny of the wild onions (allium; alliaceae) with a focus on the western north american center of diversity, *Molecular Phylogenetics and Evolution* **47**(3), 1157–1172.

Nickrent, D. L. and Patrick, J. A.: 1998, The nuclear ribosomal dna intergenic spacers of wild and cultivated soybean have low variation and cryptic subrepeats., *Genome* **41**(2), 183–192.

Nielsen, L. and Edstrom, J. E.: 1993, Complex telomere-associated repeat units in members of the genus chironomus evolve from sequences similar to simple telomeric repeats, *Molecular and Cellular Biology* **13**(3), 1583–1589.

Nielsen, L., Schmidt, E. R. and Edstrom, J. E.: 1990, Subrepeats result from regional dna sequence conservation in tandem repeats in chironomus telomeres, *Journal of Molecular Biology* **216**(3), 577–584.

Noll, M. and Kornberg, R. D.: 1977, Action of micrococcal nuclease on chromatin and the location of histone h1, *Journal of Molecular Biology* **109**(3), 393–404.

Nosek, J., Kosa, P. and Tomaska, L.: 2006, On the origin of telomeres: a glimpse at the pre-telomerase world, *Bioessays* **28**(2), 182–190.

Nosek, J., Rycovska, A., Makhov, A. M., Griffith, J. D. and Tomaska, L.: 2005, Amplification of telomeric arrays via rolling-circle mechanism, *The Journal of Biological Chemistry* **280**(11), 10840–10845.

Noutoshi, Y., Arai, R., Fujie, M. and Yamada, T.: 1998, Structure of the chlorella zepp retrotransposon: nested zepp clusters in the genome, *Molecular Genetics and Genomics* **259**(3), 256–263.

Nouzova, M., Neumann, P., Navratilova, A., Galbraith, D. W. and Macas, J.: 2001, Microarray-based survey of repetitive genomic sequences in vicia spp., *Plant Molecular Biology* **45**(2), 229–244.

Oguchi, K., Liu, H., Tamura, K. and Takahashi, H.: 1999, Molecular cloning and characterization of attert, a telomerase reverse transcriptase homolog in arabidopsis thaliana, *FEBS Letters* **457**(3), 465–469.

Ohta, T. and Dover, G. A.: 1984, The cohesive population genetics of molecular drive, *Genetics* **108**(2), 501–521.

Oikemus, S. R., McGinnis, N., Queiroz-Machado, J., Tukachinsky, H., Takada, S., Sunkel, C. E. and Brodsky, M. H.: 2004, Drosophila atm/telomere fusion is required for telomeric localization of hp1 and telomere position effect, *Genes and Development* **18**(15), 1850–1861.

Okazaki, S., Ishikawa, H. and Fujiwara, H.: 1995, Structural analysis of tras1, a novel family of telomeric repeat-associated retrotransposons in the silkworm, bombyx mori, *Molecular and Cellular Biology* **15**(8), 4545–4552.

Okuyama, Y., Fujii, N., Wakabayashi, M., Kawakita, A., Ito, M., Watanabe, M., Murakami, N. and Kato, M.: 2005, Nonuniform concerted evolution and chloroplast capture: heterogeneity of observed introgression patterns in three molecular data partition phylogenies of asian mitella (saxifragaceae), *Molecular Biology and Evolution* **22**(2), 285–296.

Olmedilla, A., Testillano, P. S., Vicente, O., Delseny, M. and Risueno, M. C.: 1993, Ultrastructural rrna localization in plant cell nucleoli. rna/rna in situ hybridization, autoradiography and cytochemistry, *Journal of Cell Science* **106**, 1333–1346.

Opresko, P. L., Otterlei, M., Graakjaer, J., Bruheim, P., Dawut, L., Kolvraa, S., May, A., Seidman, M. M. and Bohr, V. A.: 2004, The werner syndrome helicase and exonuclease cooperate to resolve telomeric d loops in a manner regulated by trf1 and trf2, *Molecular Cell* **14**(6), 763–774.

Orren, D. K., Theodore, S. and Machwe, A.: 2002, The werner syndrome helicase/exonuclease (wrn) disrupts and degrades d-loops in vitro, *Biochemistry* **41**(46), 13483–13488.

Panzera, F., Gimenez-Abian, M. I., Lopez-Saez, J. F., Gimenez-Martin, G., Cuadrado, A., Shaw, P. J., Beven, A. F., Canovas, J. L. and De la Torre, C.: 1996, Nucleolar organizer expression in allium cepa l. chromosomes, *Chromosoma* **105**(1), 12–19.

Paques, F. and Haber, J. E.: 1999, Multiple pathways of recombination induced by double-strand breaks in saccharomyces cerevisiae, *Microbiology and Molecular Biology Reviews* **63**(2), 349–404.

Pearce, S. R., Pich, U., Harrison, G., Flavell, A. J., Heslop-Harrison, J. S., Schubert, I. and Kumar, A.: 1996, The ty1-copia group retrotransposons of allium cepa are distributed throughout the chromosomes but are enriched in the terminal heterochromatin, *Chromosome Research* **4**(5), 357–364.

Pedersen, C. and Linde-Laursen, I.: 1994, Chromosomal locations of four minor rdna loci and a marker microsatellite sequence in barley, *Chromosome Research* **2**(1), 65–71.

Penton, E. H. and Crease, T. J.: 2004, Evolution of the transposable element pokey in the ribosomal dna of species in the subgenus daphnia (crustacea: Cladocera), *Molecular Biology and Evolution* **21**(9), 1727–1739.

Penton, E. H., Sullender, B. W. and Crease, T. J.: 2002, Pokey, a new dna transposon in daphnia (cladocera: crustacea), *Journal of Molecular Biology* **55**(6), 664–673.

Petes, T. D.: 1980, Unequal meiotic recombination within tandem arrays of yeast ribosomal dna genes, *Cell* **19**(3), 765–774.

Petracek, M. E., Lefebvre, P. A., Silflow, C. D. and Berman, J.: 1990, Chlamydomonas telomere sequences are a+t-rich but contain three consecutive g-c base pairs, *Proceedings of the National Academy of Science USA* **87**(21), 8222–8226.

Petricka, J. J. and Nelson, T. M.: 2007, Arabidopsis nucleolin affects plant development and patterning, *Plant Physiology* **144**(1), 173–186.

Pich, U., Fritsch, R. and Schubert, I.: 1996, Closely related allium species (alliaceae) share a very similar satellite sequence, *Plant Systematics and Evolution* **202**(3-4), 255–264.

Pich, U., Fuchs, J. and Schubert, I.: 1996, How do alliaceae stabilize their chromosome ends in the absence of tttaggg sequences?, *Chromosome Res* **4**(3), 207–213.

Pich, U. and Schubert, I.: 1998, Terminal heterochromatin and alternative telomeric sequences in allium cepa, *Chromosome Research* **6**(4), 315–321.

Pikaard, C. S.: 2002, *The Arabidopsis Book*, American Society of Plant Biologists, Rockville, MD, chapter Transcription and tyranny in the nucleolus: the organization, activation, dominance and repression of ribosomal RNA genes, pp. 1–23.
URL: *http://www.aspb.org/publications/arabidopsis/*

Piller, K. J., Baerson, S. R., Polans, N. O. and Kaufman, L. S.: 1990, Structural analysis of the short length ribosomal dna variant from pisum sativum l. cv. alaska, *Nucleic Acids Research* **18**(11), 3135–3145.

Polanco, C., Gonzalez, A. I., de la Fuente, A. and Dover, G. A.: 1998, Multigene family of ribosomal dna in drosophila melanogaster reveals contrasting patterns of homogenization for igs and its spacer regions: A possible mechanism to resolve this paradox, *Genetics* **149**(1), 243–256.

Polanco, C., Gonzalez, A. I. and Dover, G. A.: 2000, Patterns of variation in the intergenic spacers of ribosomal dna in drosophila melanogaster support a model for genetic exchanges during x-y pairing, *Genetics* **155**(3), 1221–1229.

Polanco, C. and Perez de la Vega, M.: 1997, Intergenic ribosomal spacer variability in hexaploid oat cultivars and landraces, *Heredity* **78**, 115–123.

Pollice, A., Zibella, M. P., Bilaud, T., Laroche, T., Pulitzer, J. F. and Gilson, E.: 2000, In vitro binding of nucleolin to double-stranded telomeric dna, *Biochemical and Biophysical Research Communications* **268**(3), 909–915.

Pont, G., Degroote, F. and Picard, G.: 1987, Some extrachromosomal circular dnas from drosophila embryos are homologous to tandemly repeated genes, *Journal of Molecular Biology* **195**(2), 447–451.

Pont, G., Degroote, F. and Picard, G.: 1988, Illegitimate recombination in the histone multigenic family generates circular dnas in drosophila embryos, *Nucleic Acids Res* **16**(18), 8817–8833.

Prokopowich, C. D., Gregory, T. R. and Crease, T. J.: 2003, The correlation between rdna copy number and genome size in eukaryotes, *Genome* **46**(1), 48–50.

Rabl, C.: 1885, Über zelltheilung, *Morphologisches Jahrbuch* **10**, 214–330.

Raina, S. N., Sharma, S., Sasakuma, T., Kishii, M. and Vaishnavi, S.: 2005, Novel repeated dna sequences in safflower (carthamus tinctorius l.) (asteraceae): cloning, sequencing, and physical mapping by fluorescence in situ hybridization, *Journal of Heredity* **96**(4), 424–429.

Ramel, C.: 1997, Mini- and microsatellites, *Environmental Health Perspectives* **105**(Suppl 4), 781–789.

Raskina, O., Barber, J. C., Nevo, E. and Belyayev, A.: 2008, Repetitive dna and chromosomal rearrangements: speciation-related events in plant genomes, *Cytogenetic and Genome Research* **120**(3-4), 351–357.

Ratkiewicz, M., Fedyk, S., Banaszek, A., Gielly, L., Chetnicki, W., Jadwiszczak, K. and Taberlet, P.: 2002, The evolutionary history of the two karyotypic groups of the common shrew, sorex araneus, in poland, *Heredity* **88**(4), 235–242.

Rice, P., Longden, I. and Bleasby, A.: 2000, Emboss: the european molecular biology open software suite, *Trends in Genetics* **16**(6), 276–277.

Richards, E. J. and Ausubel, F. M.: 1988, Isolation of a higher eukaryotic telomere from arabidopsis thaliana, *Cell* **53**(1), 127–136.

Richardson, J. E., Pennington, R. T., Pennington, T. D. and Hollingsworth, P. M.: 2001, Rapid diversification of a species-rich genus of neotropical rain forest trees, *Science* **293**(5538), 2242–2245.

Rickards, B., Flint, S. J., Cole, M. D. and LeRoy, G.: 2007, Nucleolin is required for rna polymerase i transcription in vivo, *Molecular and Cellular Biology* **27**(3), 937–948.

Ricroch, A., Peffley, E. B. and Baker, R. J.: 1992, Chromosomal location of rdna in allium - insitu hybridization using biotin-labeled and fluorescein-labeled probe, *THEORETICAL AND APPLIED GENETICS* **83**(4), 413–418.

Riha, K., Fajkus, J., Siroky, J. and Vyskot, B.: 1998, Developmental control of telomere lengths and telomerase activity in plants, *The Plant Cell* **10**(10), 1691–1698.

Riha, K., McKnight, T. D., Griffing, L. R. and Shippen, D. E.: 2001, Living with genome instability: plant responses to telomere dysfunction, *Science* **291**(5509), 1797–1800.

Rochaix, J. D., Bird, A. and Barkken, A.: 1974, Ribosomal rna gene amplification by rolling circles, *J Mol Biol* **87**(3), 473–487.

Roger, B., Moisand, A., Amalric, F. and Bouvet, P.: 2003, Nucleolin provides a link between rna polymerase i transcription and pre-ribosome assembly, *Chromosoma* **111**(6), 399–407.

Rogers, S. and Bendich, A.: 1987, Heritability and variability in ribosomal rna genes of vicia faba, *Genetics* **117**(2), 285–295.

Rong, L. and Klein, H. L.: 1993, Purification and characterization of the srs2 dna helicase of the yeast saccharomyces cerevisiae, *The Journal of Biological Chemistry* **268**(2), 1252–1259.

Rose, M. and Doolittle, W.: 1983, Molecular biological mechanisms of speciation, *Science* **220**(4593), 157–162.

Roth, C. W., Kobeski, F., Walter, M. F. and Biessmann, H.: 1997, Chromosome end elongation by recombination in the mosquito anopheles gambiae, *Molecular and Cellular Biology* **17**(9), 5176–5183.

Rothstein, R., Michel, B. and Gangloff, S.: 2000, Replication fork pausing and recombination or "gimme a break", *Genes and Development* **14**(1), 1–10.

Rotkova, G., Sklenickova, M., Dvorackova, M., Sykorova, E., Leitch, A. R. and Fajkus, J.: 2004, An evolutionary change in telomere sequence motif within the plant section asparagales had significance for telomere nucleoprotein complexes, *Cytogenetic and Genome Research* **107**(1-2), 132–138.

Rotkova, G., Sykorova, E. and Fajkus, J.: 2007, Characterization of nucleoprotein complexes in plants with human-type telomere motifs, *Plant Physiology and Biochemistry* **45**(9), 716–721.

Rousselet, J., Monti, L., Auger-Rozenberg, M. A., Parker, J. S. and Lemeunier, F.: 2000, Chromosome fission associated with growth of ribosomal dna in neodiprion abietis (hymenoptera: Diprionidae), *Proceedings of the Royal Society of London B* **267**(1455), 1819–1823.

Roy, S. C. and Ghosh, S.: 1977, Orientation of giemsa c-bands in interphase cells of allium cepa l., *Experientia* **33**(4), 432–433.

Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X. and Rozas, R.: 2003, DnaSP, DNA polymorphism analyses by the coalescent and other methods, *Bioinformatics* **19**(18), 2496–2497.

Rozen, S. and Skaletsky, H.: 2000, *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, Vol. 132, Humana Press Inc., Totowa, NJ, chapter Primer3 on the WWW for general users and for biologist programmers, pp. 365–386.

Ruas, C. F., Weiss-Schneeweiss, H., Stuessy, T. F., Samuel, M. R., Pedrosa-Harand, A., Tremetsberger, K., Ruas, P. M., Schluter, P. M., Ortiz Herrera, M. A., Konig, C. and Matzenbacher, N. I.: 2008, Characterization, genomic organization and chromosomal distribution of ty1-copia retrotransposons in species of hypochaeris (asteraceae), *Gene* **412**(1-2), 39–49.

Ruckova, E., Friml, J., Prochazkova Schrumpfova, P. and Fajkus, J.: 2008, Role of alternative telomere lengthening unmasked in telomerase knock-out mutant plants, *Plant Molecular Biology* **66**(6), 637–646.

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. and Allard, R. W.: 1984, Ribosomal dna spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics, *Proceedings of the National Academy of Sciences USA* **81**(24), 8014–8018.

Saiga, H. and Edstrom, J. E.: 1985, Long tandem arrays of complex repeat units in chironomus telomeres, *The EMBO journal* **4**(3), 799–804.

Saitou, N. and Nei, M.: 1987, The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Molecular Biology and Evolution* **4**(4), 406–425.

Samadashwily, G. M., Raca, G. and Mirkin, S. M.: 1997, Trinucleotide repeats affect dna replication in vivo, *Nature Genetics* **17**(3), 298–304.

Sambrook, J. and Russell, D.: 2001, *Molecular Cloning: A Laboratory Manual*, third edition edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Samonte, R. V. and Eichler, E. E.: 2002, Segmental duplications and the evolution of the primate genome, *Nature Reviews Genetics* **3**(1), 65–72.

Sang, T., Crawford, D. J. and Stuessy, T. F.: 1995, Documentation of reticulate evolution in peonies (paeonia) using internal transcribed spacer sequences of nuclear ribosomal dna: implications for biogeography and concerted evolution, *Proceedings of the National Academy of Science USA* **92**(15), 6813–6817.

Sato, S.: 1981, Cytological studies on the satellited chromosomes of allium cepa, *Caryologia* **34**, 431–440.

Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J. and Pluckthun, A.: 2001, In vitro generated antibodies specific for telomeric guanine-quadruplex dna react with stylonychia lemnae macronuclei, *Proceedings of the National Academy of Science USA* **98**(15), 8572–8577.

Schawalder, J., Paric, E. and Neff, N. F.: 2003, Telomere and ribosomal dna repeats are chromosomal targets of the bloom syndrome dna helicase, *BMC Cell Biology* **4**, 15.

Scherthan, H.: 2001, A bouquet makes ends meet, *Nature Reviews Molecular Cell Biology* **2**, 621–627.

Schlotterer, C. and Tautz, D.: 1994, Chromosomal homogeneity of drosophila ribosomal dna arrays suggests intrachromosomal exchanges drive concerted evolution, *Current Biology* **4**(9), 777–783.

Schrumpfova, P., Kuchar, M., Mikova, G., Skrisovska, L., Kubicarova, T. and Fajkus, J.: 2004, Characterization of two arabidopsis thaliana myb-like proteins showing affinity to telomeric dna sequence, *Genome* **47**(2), 316–324.

Schrumpfova, P. P., Kuchar, M., Palecek, J. and Fajkus, J.: 2008, Mapping of interaction domains of putative telomere-binding proteins attrb1 and atpot1b from arabidopsis thaliana, *FEBS Letters* **582**(10), 1400–1406.

Schubert, I.: 1984, Mobile nucleolus organizing regions (nors) in allium (liliaceae s. lat.)? - inferences from the specifity of silver staining, *Plant Systematics and Evolution* **144**(3 - 4), 291–305.

Schubert, I., Ohle, H. and Hanelt, P.: 1983, Phylogenetic conclusions from giemsa banding and nor staining in top onions (liliaceae), *Plant Systematics and Evolution* **143**, 245–256.

Schubert, I. and Wobus, U.: 1985, In situ hybridization confirms jumping nucleolus organizing regions in allium, *Chromosoma* **92**(2), 143–148.

Schultze, P., Smith, F. W. and Feigon, J.: 1994, Refined solution structure of the dimeric quadruplex formed from the oxytricha telomeric oligonucleotide d(ggggttttgggg), *Structure* **2**(3), 221–.

Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. and Miller, W.: 2000, Pipmaker-a web server for aligning two genomic dna sequences, *Genome Research* **10**(4), 577–586.

Sen, D. and Gilbert, W.: 1988, Formation of parallel four-stranded complexes by guanine-rich motifs in dna and its implications for meiosis, *Nature* **334**(6180), 364–366.

Seperack, P., Slatkin, M. and Arnheim, N.: 1988, Linkage disequilibrium in human ribosomal genes: implications for multigene family evolution, *Genetics* **119**(4), 943–949.

Shapiro, J. A. and von Sternberg, R.: 2005, Why repetitive dna is essential to genome function, *Biological reviews of the Cambridge Philosophical Society* **80**(2), 227–250.

Sharma, S., Doherty, K. M. and Brosh, R. M. J.: 2006, Mechanisms of recq helicases in pathways of dna metabolism and maintenance of genomic stability, *The Biochemical Journal* **398**(3), 319–337.

Sharma, S. and Raina, S.: 2005, Organization and evolution of highly repeated satellite dna sequences in plant chromosomes, *Cytogenetic and Genome Research* **109**(1-3), 15–26.

Shaw, P. and Doonan, J.: 2005, The nucleolus. playing by different rules?, *Cell Cycle* **4**(1), 102–105.

Shen, J., Gray, M., Oshima, J. and Loeb, L.: 1998, Characterization of werner syndrome protein dna helicase activity: directionality, substrate dependence and stimulation by replication protein a, *Nucleic Acids Research* **26**(12), 2879–2885.

Shen, Y., Ji, G., Haas, B. J., Wu, X., Zheng, J., Reese, G. J. and Li, Q. Q.: 2008, Genome level analysis of rice mrna 3'-end processing signals and alternative polyadenylation, *Nucleic Acids Res* **36**(9), 3150–3161.

Shibata, F. and Hizume, M.: 2002, Evolution of 5s rdna units and their chromosomal localization in allium cepa and allium schoenoprasum revealed by microdissection and fish, *Theoretical and Applied Genetics* **105**(2-3), 167–172.

Shida, T., Yokoyama, K. and Sekiguchi, J.: 1989, Self-assembly of synthetic deoxyoligonucleotides containing d-g clusters, *Nucleic Acids Symposium Series* (21), 83–84.

Shiratori, M., Suzuki, T., Itoh, C., Goto, M., Furuichi, Y. and Matsumoto, T.: 2002, Wrn helicase accelerates the transcription of ribosomal rna as a component of an rna polymerase i-associated complex, *Oncogene* **21**(16), 2447–2454.

Shishido, R., Sano, Y. and Fukui, K.: 2000, Ribosomal dnas: an exception to the conservation of gene order in rice genomes, *Molecular and General Genetics* **263**(4), 586–591.

Shufran, K. A., Peters, D. C. and Webster, J. A.: 1997, Generation of clonal diversity by sexual reproduction in the greenbug, schizaphis graminum, *Insect Molecular Biology* **6**(3), 203–209.

Sinclair, D. A., Mills, K. and Guarente, L.: 1997, Accelerated aging and nucleolar fragmentation in yeast sgs1 mutants, *Science* **277**(5330), 1313–1316.

Siroky, J., Zluvova, J., Riha, K., Shippen, D. E. and Vyskot, B.: 2003, Rearrangements of ribosomal dna clusters in late generation telomerase-deficient arabidopsis, *Chromosoma* **112**(3), 116–123.

Skalická, K., Lim, K., Matyásek, R., Koukalová, B., Leitch, A. and Kovarík, A.: 2003, Rapid evolution of parental rdna in a synthetic tobacco allotetraploid line, *American Journal of Botany* **90**, 988–996.

Smith, F. W. and Feigon, J.: 1992, Quadruplex structure of oxytricha telomeric dna oligonucleotides, *Nature* **356**(6365), 164–168.

Smith, G. P.: 1976, Evolution of repeated dna sequences by unequal crossover, *Science* **191**(4227), 528–535.

Snowdon, R. J., Kohler, W. and Kohler, A.: 1997, Chromosomal localization and characterization of rdna loci in the brassica a and c genomes, *Genome* **40**(4), 582–587.

St John, J. and Quinn, T. W.: 2008, Rapid capture of dna targets, *Biotechniques* **44**(2), 259–264.

Stack, S. and Clarke, C.: 1973, Differential giemsa staining of the telomeres of allium cepa chromosomes: Observations related to chromosome pairing, *Canadian Journal of Genetics and Cytology* **15**, 619–624.

Stack, S. M. and Comings, D. E.: 1979, The chromosomes and dna of allium cepa, *Chromosoma* **70**(2), 161–181.

Stage, D. E. and Eickbush, T. H.: 2007, Sequence variation within the rrna gene loci of 12 drosophila species, *Genome Research* **17**(12), 1888–1897.

Stark, G. R., Debatisse, M., Giulotto, E. and Wahl, G. M.: 1989, Recent progress in understanding mechanisms of mammalian dna amplification, *Cell* **57**(6), 901–908.

Stupar, R. M., Song, J., Tek, A. L., Cheng, Z., Dong, F. and Jiang, J.: 2002, Highly condensed potato pericentromeric heterochromatin contains rdna-related tandem repeats, *Genetics* **162**(3), 1435–1444.

Suh, Y., Thien, L., Reeve, H. and Zimmer, E.: 1993, Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of ribosomal dna in winteraceae, *American Journal of Botany* **80**(9), 1042–1055.

Sullender, B. W. and Crease, T. J.: 2001, The behavior of a daphnia pulex transposable element in cyclically and obligately parthenogenetic populations, *Journal of Molecular Evolution* **53**(1), 63–69.

Sun, H., Bennett, R. J. and Maizels, N.: 1999, The saccharomyces cerevisiae sgs1 helicase efficiently unwinds g-g paired dnas, *Nucleic Acids Research* **27**(9), 1978–1984.

Sun, Y., Skinner, D. Z., Liang, G. H. and Hulbert, S. H.: 1994, Phylogenetic analysis of sorghum and related taxa using internal transcribed spacers of nuclear ribosomal dna, *Theoretical and Applied Genetics* **89**(1), 26–32.

Sundquist, W. I. and Klug, A.: 1989, Telomeric dna dimerizes by formation of guanine tetrads between hairpin loops, *Nature* **342**(6251), 825–829.

Suzuki, G., Ura, A., Saito, N., Do, G. S., Seo, B. B., Yamamoto, M. and Mukai, Y.: 2001, Bac fish analysis in allium cepa, *Genes and Genetic Systems* **76**, 251–255.

Suzuki, K.: 2004, Characterization of telomere dna among five species of pteridophytes and bryophytes, *Journal of Bryology* **26**, 175–180.

Swofford, D. L.: 2002, *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4.*, Sinauer Associates, Sunderland, Massachusetts.

Sýkorová, E., Fajkus, J., Mezníková, M., Lim, K. Y., Neplechová, K., Blattner, F. R., Chase, M. and Leitch, A. R.: 2006, Minisatellite telomeres occur in the family alliaceae but are lost in allium, *American Journal of Botany* **93**, 814–823.

Sýkorová, E., Leitch, A. R. and Fajkus, J.: 2006, Asparagales telomerases which synthesize the human type of telomeres, *Plant Molecular Biology* **60**(5), 633–646.

Sýkorová, E., Lim, K. Y., Chase, M. W., Knapp, S., Leitch, I. J., Leitch, A. R. and Fajkus, J.: 2003, The absence of arabidopsis-type telomeres in cestrum and closely related genera vestia and sessea (solanaceae): first evidence from eudicots, *The Plant Journal* **34**(3), 283–291.

Sýkorová, E., Lim, K. Y., Fajkus, J. and Leitch, A. R.: 2003, The signature of the cestrum genome suggests an evolutionary response to the loss of (tttaggg)n telomeres, *Chromosoma* **112**(4), 164–172.

Sýkorová, E., Lim, K. Y., Kunickå, Z., Chase, M., Bennett, M. D., Fajkus, J. and Leitch, A. R.: 2003, Telomere variability in the monocotyledonous plant order asparagales, *Proceedings of the Royal Society of London Series B* **270**(1527), 1893–1904.

Szostak, J. W. and Blackburn, E. H.: 1982, Cloning yeast telomeres on linear plasmid vectors, *Cell* **29**(1), 245–255.

Szostak, J. W. and Wu, R.: 1980, Unequal crossing over in the ribosomal dna of saccharomyces cerevisiae, *Nature* **284**(5755), 426–430.

Takahashi, H., Okazaki, S. and Fujiwara, H.: 1997, A new family of site-specific retrotransposons, sart1, is inserted into telomeric repeats of the silkworm, bombyx mori, *Nucleic Acids Research* **25**(8), 1578–1584.

Takaiwa, F., Kikuchi, S. and Oono, K.: 1990, The complete nucleotide sequence of the intergenic spacer between 25s and 17s rdnas in rice, *Plant Molecular Biology* **15**(6), 933–935.

Takaiwa, F., Oono, K., Iida, Y. and Sugiura, M.: 1985, The complete nucleotide sequence of a rice 25s rrna gene, *Gene* **37**(1-3), 255–259.

Takaiwa, F., Oono, K. and Sugiura, M.: 1984, The complete nucleotide sequence of a rice 17s rrna gene, *Nucleic Acids Research* **12**(13), 5441–5448.

Takeuchi, Y., Horiuchi, T. and Kobayashi, T.: 2003, Transcription-dependent recombination and the role of fork collision in yeast rdna, *Genes and Development* **17**(12), 1497–1506.

Tartof, K. D.: 1971, Increasing the multiplicity of ribosomal rna genes in drosophila melanogaster, *Science* **171**(968), 294–297.

Tartof, K. D.: 1974, Unequal mitotic sister chromatin exchange as the mechanism of ribosomal rna gene magnification, *Proceedings of the National Academy of Science USA* **71**(4), 1272–1276.

Telenius, H., Carter, N. P., Bebb, C. E., Nordenskjold, M., Ponder, B. A. J. and Tunnacliffe, A.: 1992, Degenerate oligonucleotide-primed pcr: General amplification of target dna by a single degenerate primer, *Genomics* **13**(3), 718–725.

Teng, S.-C. and Zakian, V. A.: 1999, Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in saccharomyces cerevisiae, *Molecular and Cellular Biology* **19**(12), 8083–8093.

Thompson, J., Higgins, D. and Gibson, T.: 1994, Clustal w: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Research* **22**, 4673–4680.

Tomaska, L., Makhov, A. M., Griffith, J. D. and Nosek, J.: 2002, t-loops in yeast mitochondria, *Mitochondrion* **1**(5), 455–459.

Tomlinson, R. L., Ziegler, T. D., Supakorndej, T., Terns, R. M. and Terns, M. P.: 2006, Cell cycle-regulated trafficking of human telomerase to telomeres, *Molecular biology of the cell* **17**(2), 955–965.

Tong, C. G., Reichler, S., Blumenthal, S., Balk, J., Hsieh, H. L. and Roux, S. J.: 1997, Light regulation of the abundance of mrna encoding a nucleolin-like protein localized in the nucleoli of pea nuclei, *Plant Physiology* **114**(2), 643–652.

Traverse, K. L. and Pardue, M. L.: 1988, A spontaneously opened ring chromosome of drosophila melanogaster has acquired he-t dna sequences at both new telomeres, *Proceedings of the National Academy of Sciences USA* **85**(21), 8116–8120.

Trifonov, E.: 1985, Curved dna, *Critical Reviews in Biochemistry* **19**(2), 89–106.

Trifonov, E. and Sussman, J.: 1980, The pitch of chromatin dna is reflected in its nucleotide sequence, *Proceedings of the National Academy of Sciences USA* **77**(7), 3816–3820.

Ueno, R., Huss, V. A. R., Urano, N. and Watabe, S.: 2007, Direct evidence for redundant segmental replacement between multiple 18s rrna genes in a single prototheca strain, *Microbiology* **153**(Pt 11), 3879–3893.

Umezu, K., Nakayama, K. and Nakayama, H.: 1990, Escherichia coli recq protein is a dna helicase, *Proceedings of the National Academy of Sciences USA* **87**(14), 5363–5367.

Unfried, K., Schiebel, K. and Hemleben, V.: 1991, Subrepeats of rdna intergenic spacer present as prominent independent satellite dna in vigna radiata but not in vigna angularis, *Gene* **99**(1), 63–68.

Vaio, M., Speranza, P., Valls, J. F., Guerra, M. and Mazzella, C.: 2005, Localization of the 5s and 45s rdna sites and cpdna sequence analysis in species of the quadrifaria group of paspalum (poaceae, paniceae), *Annals of Botany* **96**(2), 191–200.

van Heusden, A. W., Shigyo, M., Tashiro, Y., Vrielink-van Ginkel, R. and Kik, C.: 2000, Aflp linkage group assignment to the chromosomes of allium cepa l. via monosomic addition lines, *Theoretical and Applied Genetics* **100**(3), 480–486.

Van Holde, K. E., Sahasrabuddhe, C. G. and Shaw, B. R.: 1974, A model for particulate structure in chromatin, *Nucleic Acids Research* **1**(11), 1579–1586.

Van Houten, J. V. and Newlon, C. S.: 1990, Mutational analysis of the consensus sequence of a replication origin from yeast chromosome iii, *Molecular and Cellular Biology* **10**(8), 3917–3925.

Vander Stappen, J., Marant, S. and Volckaert, G.: 2003, Molecular characterization and phylogenetic utility of the rdna external transcribed spacer region in stylosanthes (fabaceae), *Theoretical and Applied Genetics* **107**(2), 291–298.

Van't Hof, J.: 1965, Relationships between mitotic cycle duration s period duration and average rate of dna synthesis in root meristem cells of several plants, *Experimental Cell Research* **39**, 48–58.

Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D. A., Cestaro, A., Pruss, D., Pindo, M., Fitzgerald, L. M., Vezzulli, S., Reid, J., Malacarne, G., Iliev, D., Coppola, G., Wardell, B., Micheletti, D., Macalma, T., Facci, M., Mitchell, J. T., Perazzolli, M., Eldredge, G., Gatto, P., Oyzerski, R., Moretto, M., Gutin, N., Stefanini, M., Chen, Y., Segala, C., Davenport, C., Dematte, L., Mraz, A., Battilana, J., Stormo, K., Costa, F., Tao, Q., Si-Ammour, A., Harkins, T., Lackey, A., Perbost, C., Taillon, B., Stella, A., Solovyev, V., Fawcett, J. A., Sterck, L., Vandepoele, K., Grando, S. M., Toppo, S., Moser, C., Lanchbury, J., Bogden, R., Skolnick, M., Sgaramella, V., Bhatnagar, S. K., Fontana, P., Gutin, A., Van de Peer, Y., Salamini, F. and Viola, R.: 2007, A high quality draft consensus sequence of the genome of a heterozygous grapevine variety, *PLoS ONE* **2**(12), e1326.

Venkateswarlu, K., Lee, S. W. and Nazar, R. N.: 1991, Conserved upstream sequence elements in plant 5s ribosomal rna-encoding genes, *Gene* **105**(2), 249–254.

Vershinin, A. V., Schwarzacher, T. and Heslop-Harrison, J. S.: 1995, The large-scale genomic organisation of repetitive dna families at the telomeres of rye chromosomes, *The Plant Cell* **7**, 1823–1833.

Viersbach, R., Schwanitz, G. and Nothen, M. M.: 1994, Delineation of marker chromosomes by reverse chromosome painting using only a small number of dop-pcr amplified microdissected chromosomes, *Human Genetics* **93**(6), 663–667.

Vignali, M. and Workman, J. L.: 1998, Location and function of linker histones, *Nature Structural Biology* **5**(12), 1025–1028.

Villasante, A., Abad, J. P., Planello, R., Mendez-Lago, M., Celniker, S. E. and de Pablos, B.: 2007, Drosophila telomeric retrotransposons derived from an ancestral element that was recruited to replace telomerase, *Genome Research* **17**(12), 1909–1918.

Vinogradov, A. E.: 2003, Dna helix: the importance of being gc-rich, *Nucleic Acids Research* **31**(7), 1838–1844.

Vitkova, M., Kral, J., Traut, W., Zrzavy, J. and Marec, F.: 2005, The evolutionary origin of insect telomeric repeats, (ttagg)n, *Chromosome Research* **13**(2), 145–156.

Vitte, C. and Panaud, O.: 2005, Ltr retrotransposons and flowering plant genome size: emergence of the increase/decrease model, *Cytogenet Genome Research* **110**(1-4), 91–107.

Vosa, C. G.: 1976, Heterochromatic patterns in allium 1. the relationship between the species of the cepa group and its allies, *Heredity* **36**, 383–392.

Walter, M. F., Bozorgnia, L., Maheshwari, A. and Biessmann, H.: 2001, The rate of terminal nucleotide loss from a telomere of the mosquito anopheles gambiae, *Insect Molecular Biology* **10**, 105–110.

Wang, S.-S. and Zakian, V.: 1990a, Sequencing of the saccharomyces telomeres cloned using t4 dna polymerase reveals two domains, *Molecular Cell Biology* **10**, 4415–4419.

Wang, S. S. and Zakian, V. A.: 1990b, Telomere-telomere recombination provides an express pathway for telomere acquisition, *Nature* **345**(6274), 456–458.

Wang, Y., Erdmann, N., Giannone, R. J., Wu, J., Gomez, M. and Liu, Y.: 2005, An increase in telomere sister chromatid exchange in murine embryonic stem cells possessing critically shortened telomeres, *Proceedings of the National Academy of Sciences USA* **102**(29), 10256–10260.

Wanzenbock, E. M., Schofer, C., Schweizer, D. and Bachmair, A.: 1997, Ribosomal transcription units integrated via t-dna transformation associate with the nucleolus and do not require upstream repeat sequences for activity in arabidopsis thaliana, *The Plant Journal* **11**(5), 1007–1016.

Wei, C. and Price, M.: 2003, Protecting the terminus: t-loops and telomere end-binding proteins, *Cellular and Molecular Life Sciences* **60**(11), 2283–2294.

Weiss, H. and Scherthan, H.: 2002, Aloe spp. - plants with vertebrate-like telomeric sequences, *Chromosome Research* **10**(2), 155–164.

Wendel, J. F., Schnabel, A. and Seelanan, T.: 1995, An unusual ribosomal dna sequence from gossypium gossypioides reveals ancient, cryptic, intergenomic introgression, *Molecular Phylogenetics and Evolution* **4**(3), 298–313.

Whiting, M. F., Carpenter, J. C., Wheeler, Q. D. and Wheeler, W. C.: 1997, The strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18s and 28s ribosomal dna sequences and morphology, *Systematic Biology* **46**(1), 1–68.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P. and Schulman, A. H.: 2007, A unified classification system for eukaryotic transposable elements, *Nature Reviews Genetics* **8**(12), 973–982.

Williamson, J. R., Raghuraman, M. K. and Cech, T. R.: 1989, Monovalent cation-induced structure of telomeric dna: the g-quartet model, *Cell* **59**(5), 871–880.

Win, T. Z., Mankouri, H. W., Hickson, I. D. and Wang, S.-W.: 2005, A role for the fission yeast rqh1 helicase in chromosome segregation, *Journal of Cell Science* **118**(Pt 24), 5777–5784.

Wójcik, J. M., Ratkiewicz, M. and Searle, J. B.: 2002, Evolution of the common shrew sorex araneus: chromosomal and molecular aspects, *Acta Theriologica* **47**(supp 1), 139–167.

Wojcik, J. M. and Searle, J. B.: 1988, The chromosome complement of sorex granarius - the ancestral karyotype of the common shrew (sorex araneus)?, *Heredity* **61 (Pt 2)**, 225–229.

Worton, R. G., Sutherland, J., Sylvester, J. E., Willard, H. F., Bodrug, S., Dube, I., Duff, C., Kean, V., Ray, P. N. and Schmickel, R. D.: 1988, Human ribosomal rna genes: orientation of the tandem array and conservation of the 5' end, *Science* **239**(4835), 64–68.

Wright, W., Tesmer, V., Huffman, K., Levene, S. and Shay, J.: 1997, Normal human chromosomes have long g-rich telomeric overhangs at one end, *Genes and Development* **11**, 2801–2809.

Wu, X. and Maizels, N.: 2001, Substrate-specific inhibition of recq helicase, *Nucleic Acids Research* **29**(8), 1765–1771.

Yamagata, K., Kato, J., Shimamoto, A., Goto, M., Furuichi, Y. and Ikeda, H.: 1998, Bloom's and werner's syndrome genes suppress hyperrecombination in yeast sgs1 mutant: implication for genomic instability in human diseases, *Proceedings of the National Academy of Sciences USA* **95**(15), 8733–8738.

Yamamoto, Y., Fujimoto, Y., Arai, R., Fujie, M., Usami, S. and Yamada, T.: 2003, Retrotransposon-mediated restoration of chlorella telomeres: accumulation of zepp retrotransposons at termini of newly formed minichromosomes, *Nucleic Acids Research* **31**(15), 4646–4653.

Yamamoto, Y., Noutoshi, Y., Fujie, M., Usami, S. and Yamada, T.: 2000, Analysis of double-strand-break repair by chlorella retrotransposon zepp, *Nucleic Acids Symposium Series* (44), 101–102.

Yan, H., Chen, C. Y., Kobayashi, R. and Newport, J.: 1998, Replication focus-forming activity 1 and the werner syndrome gene product, *Nature Genetics* **19**(4), 375–378.

Yan, H. and Newport, J.: 1995, Ffa-1, a protein that promotes the formation of replication centers within nuclei, *Science* **269**(5232), 1883–1885.

Yang, T.-J., Kwon, S.-J., Choi, B.-S., Kim, J. S., Jin, M., Lim, K.-B., Park, J. Y., Kim, J.-A., Lim, M.-H., Kim, H.-I., Lee, H.-J., Lim, Y. P., Paterson, A. H. and Park, B.-S.: 2007, Characterization of terminal-repeat retrotransposon in miniature (trim) in brassica relatives, *Theoretical and Applied Genetics* **114**(4), 627–636.

Yang, Y.-W., Tai, P.-Y., Chen, Y. and Li, W.-H.: 2002, A study of the phylogeny of brassica rapa, b. nigra, raphanus sativus, and their related genera using noncoding regions of chloroplast dna, *Molecular Phylogenetics and Evolution* **23**(2), 268–275.

Yu, C. E., Oshima, J., Fu, Y. H., Wijsman, E. M., Hisama, F., Alisch, R., Matthews, S., Nakura, J., Miki, T., Ouais, S., Martin, G. M., Mulligan, J. and Schellenberg, G. D.: 1996, Positional cloning of the werner's syndrome gene, *Science* **272**(5259), 258–262.

Yu, G. L., Bradley, J. D., Attardi, L. D. and Blackburn, E. H.: 1990, In vivo alteration of telomere sequences and senescence caused by mutated tetrahymena telomerase rnas, *Nature* **344**(6262), 126–132.

Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T. E. and Lieber, M. R.: 2003, R-loops at immunoglobulin class switch regions in the chromosomes of stimulated b cells, *Nature Immunology* **4**(5), 442–451.

Zauner, S., Fraunholz, M., Wastl, J., Penny, S., Beaton, M., Cavalier-Smith, T., Maier, U.-G. and Douglas, S.: 2000, Chloroplast protein and centrosomal genes, a trna intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph, *Proceedings of the National Academy of Sciences USA* **97**(1), 200–205.

Zentgraf, U. and Hemleben, V.: 1992, Complex formation of nuclear proteins with the rna polymerase i promoter and repeated elements in the external transcribed spacer of cucumis sativus ribosomal dna, *Nucleic Acids Research* **20**(14), 3685–3691.

Zhang, D. and Sang, T.: 1999, Physical mapping of ribosomal rna genes in peonies (paeonia, paeoniaceae) by fluorescent in situ hybridization: implications for phylogeny and concerted evolution, *American Journal of Botany* **86**(5), 735–740.

Zhang, S., Hemmerich, P. and Grosse, F.: 2004, Nucleolar localization of the human telomeric repeat binding factor 2 (trf2), *Journal of Cell Science* **117**(17), 3935–3945.

Zhang, X., Zhou, J. and Eickbush, T. H.: 2008, Rapid r2 retrotransposition leads to the loss of previously inserted copies via large deletions of the rdna locus., *Molecular Biology and Evolution* **25**(1), 229–237.

Zhang, Y., Kamnert, I., López, C., Cohn, M. and Edström, J.: 1994, A family of complex tandem dna repeats in the telomeres of chironomus pallidivittatus, *Molecular Cell Biology* **14**(12), 8028–8036.

Zhdanova, N. S., Karamisheva, T. V., Minina, J., Astakhova, N. M., Lansdorp, P., Kammori, M., Rubtsov, N. B. and Searle, J. B.: 2005, Unusual distribution pattern of telomeric repeats in the shrews sorex araneus and sorex granarius, *Chromosome Research* **13**(6), 617–625.

Zhdanova, N. S., Minina, J., Karamisheva, T., Draskovic, I., Rubtsov, N. and Londono-Vallejo, J.-A.: 2007, The very long telomeres in sorex granarius (soricidae, eulipothyphla) contain ribosomal dna, *Chromosome Research* **15**(7), 881–890.

Zhdanova, N. S., Rubtsov, N. B. and Minina, I. M.: 2007, Terminal regions of mammal chromosomes: plasticity and role in evolution, *Genetika* **43**(7), 873–886.

Zheng, X., Cai, D., Yao, L. and Teng, Y.: 2008, Non-concerted its evolution, early origin and phylogenetic utility of its pseudogenes in pyrus., *Molecular Phylogenetics and Evolution* **48**(3), 892–903.

Zimmerman, S. B., Cohen, G. H. and Davies, D. R.: 1975, X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid, *Journal of Molecular Biology* **92**(2), 181–184.

Zuker, M.: 2003, Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Research* **31**(13), 3406–3415.