

https://dspace.org/

Developing DNA profiling strategies for lithops.

Kruger, Warren Michael.

1994
https://hdl.handle.net/10539/26226

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

DEVELOPING DNA PROFILING STRATEGIES FOR LITHOPS

Warren Michael Kruger

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science

Johannesburg 1994

ABSTRACT

There are a number of recorded difficulties in classing lithop species according to their phenotype alone. DNA profiling, which can provide a quantitative measure of the degree of sequence similarity between any two genomic DNAs, is the method of choice for establishing phylogenetic relationships between species.

Cetyltrimethylammonium bromide extraction was routinely used to isolate over 200 µg of lithop DNA per gram of lyophilised tissue. Only 15 µg of a high molecular weight DNA could be spooled and used in hybridisation studies. High molecular weight DNA was also purified from low molecular weight contaminants by gel filtration chromatography through a small 1 ml column packed with BioGel A15m. The resultant elution profile is a visual representation of the degree of degradation of the crude DNA preparation. The four oligonucleotide sequences $(TCC)_5$, $(GATA)_4$, $(GACA)_4$ and $(GTG)_5$ were synthesised, cold labelled and hybridised to EcoR1 digests of lithop DNA, immobilised in dried agarose gels. However, none of these sequences produced a scoreable banding pattern.

DNA amplification fingerprinting was investigated because profiles can be generated with arbitrary sequence 10mers and with only 25 ng of genomic DNA. These profiles were found to be sensitive to changes in the concentration of $MgCl_2$ and to differences in the "actual" time spent at each temperature during cycling. The amount of primer must be in vast molar excess over the amount of template DNA if short (< 400 bp) regions are to be efficiently amplified. The amplification reactions of three different human DNA preparations with ACGGTACACT produced a homogeneous set of similarly sized markers. The percent band sharing between three lithop species was calculated from the profiles directed by ACGGTACACT and CCCTCTGCGG. There was 32% band sharing between *L. Lesliei & L. Hookeri* (both found in western Transvaal), 18% between *L. Lesliei & L. Terricolor* (found in south western Cape) and 12% between *L. Hookeri & L. Terricolor*

DAF markers generated from any lithop template represent specific lithop sequences. These sequences can be isolated, labelled and used as a source of probes for restriction DNA digest analysis of other lithop species. DAF could be used to "extract" sequence data from lithops which can then be used for more conventional forms of profiling.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University, nor has it been prepared under the aegis or with the assistance of any other body or organisation or person outside the University of the Witwatersrand, Johannesburg.

WARREN MICHAEL KRUGER Kruger
(Name of candidate)

2 day of August, 1994

List of Abbreviations

AFLP - Amplified Fragment Length Polymorphism

bp - base pairs

CTAB - Cetyltrimethylammonium Bromide

DAF - DNA Amplification Fingerprinting

DIG - Digoxigenin

EDTA - Ethylenediaminetetraacetate

IU - Intensity Units (arbitrary)

kb - kilobase pairs

MLP - Multi Locus Probe

NBT - Nitro Blue Tetrazolium

PCR - Polymerase Chain Reaction

RAPD - Random Amplified Polymorphic DNA

SDS - Sodium Dodecylsulfate

SLP - Single Locus Probe

SQR - Simple Quadruplet Repeat

TEMED - N,N,N',N'-Tetramethylethylenediamine

Tris - Tris(hydroxymethyl)aminoethane

UV - Ultraviolet

VNTR - Variable Number of Tandem Repeats

X-phosphate - 5-Bromo-4-chloro-3-indolyl phosphate

Contents

Abstract	i
Declaration	iii
List of Abbreviations	iv
List of Figures	x
CHAPTER 1	1
1 Introduction	1
1.1 The Species as the Unit for Taxonomy	1
1.2 The Role of Variation in Classification	3
1.3 Lithops: Problematic to Classify	4
1.4 DNA Polymorphism	5
1.5 Examining DNA Polymorphism	6
CHAPTER 2	11
2 Preparation of Total Genomic DNA	11
2.1 Introduction	11
2.1.1 Purification by Gel Filtration	13
2.1.2 Estimating the Total DNA Content	13
2.2 Materials and Methods	15
2.2.1 Plant DNA Isolation	15
2.2.2 Total DNA Content of Lithops	15

2.2.2.1 Two Wavelength Method 15
2.2.2.2 Diphenylamine Colour Reaction 16
2.2.3 Gel Filtration Chromatography 16
2.3 Results 18
2.3.1 Lithop DNA Isolation 18
2.3.2 Gel Filtration Chromatography 18
2.4 Discussion 25
2.4.1 Lithop DNA Isolation 25
2.4.2 Total DNA Estimation 26
2.4.3 Gel Filtration 26
2.5 Conclusion 28

CHAPTER 3 29

3 Simple Quadruplet Repeat (SQR) Analysis 29
3.1 Introduction 29
3.1.1 Specificity of Hybridisation 30
3.1.2 Theory of Hybridisation 32
3.1.3 Nonradioactive Hybridisation 33
3.2 Materials and Methods 35
3.2.1 Oligonucleotide Synthesis and Purification 35
3.2.2 3'-end Digoxigenin Labelling 35
3.2.3 Restriction Digestion and Electrophoresis 36
3.2.4 Hybridisation 36
3.2.5 Detecting the Nonradioactive Signal 37
3.3 Results 38
3.3.1 Purifying Synthetic Oligonucleotides by Gel Filtration 38
3.3.2 Labelling Efficiency 38
3.3.3 Hybridisation 39
3.4 Discussion 48

3.4.1 G ying and Purifying Oligonucleotides 48
3.4.2 Labelling Efficiency 48
3.4.3 Hybridisation 49
3.5 Conclusion 51

CHAPTER 4 52

4 DNA Amplification Fingerprinting (DAF) 52
4.1 Introduction 52
4.1.1 Molecular Mechanism of DAF 54
4.1.2 Primer Binding during DAF 55
4.1.3 Estimating the Number of Amplification Products 56
4.1.4 Silver Staining DNA in Acrylamide Gels 57
4.1.5 Nature of DAF Polymorphisms 58
4.2 Materials and Methods 61
4.2.1 Electrophoresis 61
4.2.2 Silver Staining 62
4.2.3 DAF 62
4.3 Results 64
4.3.1 Silver Staining 64
4.3.2 Electrophoresis 64
4.3.3 DAF 65
4.4 Discussion 85
4.4.1 Silver Staining 85
4.4.2 Electrophoresis 85
4.4.3 DAF of Lithops 86
4.5 Conclusion 91

CHAPTER 5 92

5 Summary 92

CHAPTER 6 98

6 References 98

APPENDICES 106

Appendix I 106

Appendix II 107

Appendix III 108

List of Figures

Figure 1: Determination of the void and total column volumes of BioGel A15m column 19

Figure 2: Elution profile of human genomic DNA through BioGel A15m 20

Figure 3: Elution profile of chicken genomic DNA through BioGel A15m 21

Figure 4: Elution profile of potato genomic DNA through BioGel A15m 22

Figure 5: Elution profile of lithop genomic DNA through BioGel A15m 23

Figure 6: Agarose electrophoresis of high molecular weight DNA fractions 24

Figure 7: Purification of 1 A_{260} unit of oligonucleotide $(TCC)_5$ through Sephadex G-15 40

Figure 8: Dot blot of 3′-end labelled 30mer and standard random primed (RP) labelled probe 41

Figure 9: Dot blot of undiluted 30mer and 800x diluted random primed standard 42

Figure 10: Pretreatment of nylon membranes to improve the binding efficiency of oligonucleotides 43

Figure 11: Dot blot of 3′-end labelled synthetic oligonucleotide and 30mer end labelled standard 44

Figure 12: In-gel hybridisation of the standard labelled 30mer to its complementary sequence in the pUC18 plasmid 45

Figure 13: In-gel hybridisation of DIG 3'-end labelled oligonucleotides to restriction digested lithop genomic DNA 46

Figure 14: Hybridisation of DIG-$(TCC)_5$ to EcoR1 digested lithop DNA 47

Figure 15: Sensitivity of silver staining of DNA in acrylamide gels 68

Figure 16: Differential mobility of single and double stranded DNA in denaturing polyacrylamide gels 69

Figure 17: Denaturing electrophoresis of heat denatured DAF products 70

Figure 18: Non-denaturing electrophoresis of heat denatured DAF products 71

Figure 19: Denaturing electrophoresis of undenatured DAF samples 72

Figure 20: DAF of lithop DNA with three different oligonucleotides 73

Figure 21: Effect of increasing the amount of primer during DAF 74

Figure 22: Optimising $MgCl_2$ for the reaction of *L. Lesliei* and CCCTCTGCGG 75

Figure 23: DAF of *L. Lesliei* and human DNA with ACGGTACACT and CCCTCTGCGG respectively 76

Figure 24: Histogram showing fragment length and band intensity polymorphisms between the human and *L. Lesliei* DAF profiles generated with ACGGTACACT 77

Figure 25: Histogram of DAF profiles of human and *L. Lesliei* DNA generated with CCCTCTGCGG 78

Figure 26: Comparing the reproducibility of DAF profiles from different laboratories 79

Figure 27: DAF of three different lithop species with ACGGTACACT and CCCTCTGCGG 80

Figure 28: Histogram showing band mobility and intensity differences for the three lithop species with ACGGTACACT 81

Figure 29: Histogram showing band mobility and intensity differences for the three lithop species with CCCTCTGCGG 82

Figure 30: Comparing DAF profiles of *L. Lesliei* and ACGGTACACT between different gels 83

Figure 31: Measuring pattern variability between different *L. Lesliei*/CCCTCTGCGG profiles 84

Figure 32: Overview of the results of DNA isolation from different species of fungi and plants obtained by Weising *et al.* (1991) 96

Figure 33: Hybridisation patterns of SQR sequences with different plant genomes as obtained by Weising *et al.* (1991) 97

CHAPTER 1

1 INTRODUCTION

1.1 The Species as the Unit for Taxonomy

The study of plant systematics is founded on the idea that, in the tremendous variation in the plant world, conceptually discontinuous units exist - called species - which can be recognised, classified and named. Relationships based on evolution are presumed to exist among these units. Classification is the arrangement of plants into groups having common characteristics. These groups are then arranged into a hierarchy of ranks. Similar species of plants are placed into a genus, similar genera are grouped into families, families are arranged into orders, orders into subclasses, subclasses into classes and classes into divisions. Identification is the recognition of certain characteristics and the application of a name to a plant with those particular characteristics. Recognition occurs when the specimen under consideration is similar to a previously known plant. If comparison of the specimen with similar species reveals that it differs from them, it may be named as a new species.

Taxonomy is based upon the similarities and dissimilarities between organisms. Taxonomic schemes which try to reflect evolution are said to be phylogenetic. The advantage of phylogenetic systems is that they are rich in informational content, since the identity of a plant implies knowledge of its affinities and evolutionary relationships. In practice a group of plants which are fundamentally alike are generally treated as a species. Ideally, a species should be separated by distinct morphological differences from other closely related species. However, it is sometimes difficult to delimit a species precisely. If each minor variation in plant populations were to made the basis of species distinction, there would be no end to the number of species. Differences in features should not be allowed to obscure resemblances. A species is a concept that cannot be defined in exact terms, it is not absolute and inelastic (Jones *et al.*, 1979).

In developing concepts of species, specimens should be regarded as samples of living, reproducing populations of genetically related individuals. Different species have developed by diverse evolutionary and genetic mechanisms. Since species represent lineages, systems of living populations may be found at various stages or levels of morphological divergence from one another in reproductive isolation. As a result the classification of the plant kingdom continues to be modified as new information becomes available. New data used in conjunction with information derived from traditional sources, such as comparative anatomy and gross morphology, further refine the classification. The biological species concept envisages the species as a distinct population system of central importance in nature and in evolutionary biology. Biological species are kept separate from closely related species, inhabiting the same general region, by reproductive isolation mechanisms that prevent or greatly reduce genetic exchange between them. Two such reproductively isolated species can coexist in the same general area and not lose their genetic identity. However, when many plant species come into contact, they are able to exchange genes and may produce fertile progeny. For this reason the ability or lack of ability to produce hybrids and exchange genes cannot be used as a general criterion for erecting species boundaries in most plants.

The categories of subspecies and variety are applied to populations of species in various stages of differentiation. A common process of evolution and speciation in plants is the gradual divergence of a former homogeneous species into two or more diverse population systems. Their divergence is usually related to adaptation to differing geographical areas or climates. During the process of becoming adapted to different habitats, the populations become genetically distinct. This genetic variation is reflected in morphology and physiology. Often differentiated populations, or ecotypes, occupy adjacent ranges where they interbreed at points of contact. There can be gradual or or sharp discontinuities in the variation patterns between divergent populations, depending on the environmental gradient. These ecotypes most often form the basis of a subspecies or variety. Subspecies and varieties are recognisable morphological variations within species. Their populations have their own

patterns of variation correlated with geographical distributions or ecological requirements. Whether "subspecies" or "variety" is used often depends on the systematist.

If two populations are alike in their features the resemblance is usually attributed to either parallelism or convergence. Convergence is the resemblance of two or more distinct phylogenetic lines brought about through adaptation to similar environments. With parallelism the resemblance is due to a common ancestry and genetic background. Two phylogenetic lines may diverge up to some stage, cease to diverge and then run parallel. Parallel evolution can take place in related lineages at any stage in their separation.

1.2 The Role of Variation in Classification

Taxonomic evidence for establishing classifications is gathered from a variety of sources. Because all parts of a plant at all stages of its development can provide taxonomic characters, data must be assembled from many diverse disciplines. The complexity and diversity of plant life is apparent among species and between individuals of the same species. This variation is universal, it is the basis of both evolution and classification, without it natural selection would have nothing upon which to act (Bell *et al.*, 1969). Variation is neither uniform nor constant and physical differences may be obvious or obscure. Differences must be stably inherited and not have arisen by chance. This can be verified by examining differences between plants in a population, between plants of different populations and between plants of related species (Bell *et al.*, 1969). A better understanding of the evolution of plants may be discovered by careful analysis of their variation. Variation patterns give us information about the role that selection and mutation play in evolution. Darwin proposed that species are ever changing entities which are the products of natural selection, this however did not explain the sudden emergence of a variety of new species - mutational evolution sought to account for these anomalies. Discontinuous variation patterns are seen as the result of mutation and continuous patterns as a consequence of environmental pressure (Briggs and Walters, 1984).

Variation within a population possesses three broad components:

1. environmental - the phenotype and behaviour of the plant are determined by "interactions" between the environment and the plant's genotype. Each genotype reacts differently to a given environment, some plants can alter their patterns of growth in response to environmental changes (phenoplasticity),

2. developmental - several genes may determine a plant's phenotype and they are often regulated depending on the age and stage of development of the plant,

3. genetic variation - it is the ultimate source of variation in species and it replenishes the supply of genetic variability (Jones *et al.*, 1979).

The role of genetic mutations on evolutionary change depends on several factors: (1) their effect upon the organism, (2) the adaptive advantage or disadvantage of the mutation, and (3) the role of the mutation in population-environmental interactions. The environment in which a mutation occurs will have a great effect on its incorporation into a population. Unless the environment is changing, mutations will most likely lower the level of adaptation of the species. If a species is well adjusted to its environment, small mutations might allow it to inhabit new or changing environments, but drastic changes are almost certain to make it function poorly.

1.3 Lithops: Problematic to Classify

Lithops are fleshy plants which are generally found in areas of low rainfall. They have an individual plant body where the 'corpusculum' or head is made up of two opposite leaves which have thickened and fused together leaving a fissure across the top surface - flowers bud through this fissure. The corpusculum tapers down to a junction with the root, above which is the growing meristem.

Distinct populations of lithops are found in defined and isolated geographical locations, scattered throughout southern Africa (Cole, 1988). There is considerable correlation between individual species and the nature of their habitat, especially in terms of their topography,

type and colour of stone and soil. Some taxa are remarkably uniform in their appearance and they are thus easily distinguishable and identified, whereas others are highly variable. There are a wide variety of habitats in which lithops occur, they are found on flat plains, gentle inclines, rough and precipitous slopes, in pans, on low ridges, hill tops and mountain summits. They can grow in clay, sandy soil, humus, stony soil and even in fissures between rocks. Lithops occur at different altitudes ranging from a few metres above sea-level to 2 500 metres. Furthermore, different varieties are known to inhabit arid regions where rainfall is low, while others occur in fairly well vegetated bush and grasslands (Cole, 1988).

Cole (1988) presents an interesting insight into the problems associated with classifying lithops. Basically no single criterion is adequate on its own to enable one consistently to differentiate between two species, or even subspecies. An entire series of criteria must be employed for this objective. Of course such a system is beset with inadequacies - which criteria should be used? How should they be defined? And how are they to be continually applied to the entire mass of available material to be classified? Several formal descriptions of lithop species are only based on a small number of specimens and this gives rise to a "half-true" image of that species. Any variability beyond these descriptions might either be ignored or even lead to separate varietal types being falsely recorded. Many instances have been described where two quite distinct taxa have been reported, only for an intermediate form of both to be identified and thereby reveal some degree of relatedness between them (Cole, 1988). It is hoped that DNA profiling, which measures genotypic differences, can be used to complement the existing classification system.

1.4 DNA Polymorphism

DNA sequence polymorphisms are a convenient way of distinguishing between two organisms' genotypes (Saghai-Maroof *et al.*, 1984). Before studying genetic variation, we must understand some things about the nature of DNA. DNA is divided into different classes depending on its abundance throughout the genome. The three main classes of DNA are the unique sequences, the highly repetitive sequences and the moderately repetitive sequences. The unique sequences encompass the protein coding sequences (not exclusively)

which are present in only one or a few copies in the genome. The highly repetitive class, which are transcriptionally inactive, are present in hundreds or thousands of copies which are believed to be clustered. Moderately repetitive sequences are present in both clustered and dispersed forms, this class contains a number of gene families which occur in several copies in the genome. These sequences are present in several to hundreds of copies (Thompson and Thompson, 1986). Swanson *et al.* (1981) proposed that each eukaryotic species may be characterised by its repetitive DNA which varies in its frequency of repetitions, length and nucleotide sequence of each repeated unit and its location in the genome.

There are thousands of variations in DNA sequence which characterise each organism. Genes are continuously changing, albeit slowly. Consider a gene (cistron) of only 1000 nucleotide pairs then there are 4^{1000} possible allelic states for this gene. In reality however, a large number of these variant allelic forms are never seen because functional coding sequences must be maintained - even so the number of possible allelic states is large. In non-coding regions (between genes) this requirement does not exist and these DNA sequences are more tolerant of changes. Bearing this in mind we see that any new mutation is almost always different from the alleles preexisting in the present population (Nei, 1975). The structure of eukaryotic genes is such that there are extensive non-coding, flanking regions which are highly polymorphic (these regions are important for gene regulation). Furthermore, the DNA sequences of non-coding introns are highly variable except at their splice junctions. Gene families contain several closely related gene sequences separated by non-coding, polymorphic spacer DNA. All these sites are "hot spots" for genetic changes, i.e. they differ greatly between unrelated individuals and are said to be hypervariable. They are ideal for studying genetic divergence because their high levels of sequence alteration generate unique differences that characterise individuals.

1.5 Examining DNA Polymorphism

If we are to employ differences in DNA sequences to aid in the classification of species, there must be procedures by which we can quantify such differences in order to show genetic distances between species. The sequencing of the entire genome of each organism is highly

impractical and expensive and so other methods for quantifying differences in DNA are required. If DNA is isolated, freed of all bound proteins and incubated at a temperature of 80-100° C its double helix structure "melts" into single strands. The higher the GC content of the DNA, the higher will be the melting temperature required. When the temperature is then lowered to about 60° C reannealing of the single strands to the double stranded helix occurs (Swanson *et al.*, 1981) - this process is dependent on the random collision of molecules in solution and it is therefore a function of the concentration of the DNA molecules which are capable of base pairing as well as the time allowed for reannealing. Any highly repetitive, simple sequence DNA would be expected to undergo rapid reannealing, while unique sequences occuring in only one or a few copies would reanneal far slower. The reannealing characteristics of an organism's DNA can be roughly quantified using C_0t values. The melting temperatures and C_0t values are at best an approximation of the sequence composition of an organism's DNA. These parameters are of limited use and can only be used for simple comparative purposes of genetic relatedness. The absolute amount of genetic variation is amenable to study at the DNA level by examining changes in restriction digest sites (Hartl and Clark, 1989).

Mutations in the DNA base sequence can either create or destroy the recognition sites of an endonuclease, altering the way in which such an enzyme will digest genomic DNA. In addition to actual sequence changes which create or destroy specific restriction sites, there are other forms of mutation that can be used to distinguish individual genomes. The existence of genetic redundancy, in the form of repetitive DNA sequences in eukaryotic genomes, is well known although the very reason for their existence is unknown. In 1984 Jeffreys showed that the human genome contains a large number of hypervariable regions which are related by a common "core" sequence. These sequences (minisatellites) are made up of short, tandemly repeated DNA sequences (which are thought to arise through unequal crossing-over during recombination). They are believed to have evolved from a common set of ancestral core sequences, therefore each hypervariable sequence will have some sequence homology with many other loci. A sequence probe can recognise several sequences with sufficient sequence similarity to permit hybridisation, especially under low stringency conditions (Debenham, 1992).

The utility of some (but not all) probes, which are complementary to minisatellite regions, derives from their ability to cross hybridise with several genomic minisatellite regions and so generate highly polymorphic fragment lengths when digested with an endonuclease (Rogstad *et al.*, 1988). These probes are called multi-locus probes (MLPs) because of their ability to detect many loci simultaneously. Generally, the closer the match in sequence between probe and core sequence, the more likely the detection of other hypervariable loci in the genome. If all the tandem repeats in one individual were identical with those in other individuals, then the band pattern would be common to all. It was reasoned that the simultaneous analysis of several hypervariable loci was a sufficient measure of genetic variability so as to provide unique individual identification.

Variations in the number of repeats arising from the loss or gain of the "core" sequence are accumulated during evolution and result in length variations between species. This variation can be detected by cleavage with restriction enzymes which cut either side of the repeated sequence (minisatellite), but not within it. The length of the restriction fragment would depend on the number of repeats between sites thus generating DNA fragment length polymorphisms. These restriction fragment length polymorphisms (RFLPs) as they are called, are inherited as simple Mendelian codominant markers (Thompson and Thompson, 1986). In fact the size of each tandem repeat can vary considerably among individuals - many of these loci have 20 or more variants. Since the number and distribution of hypervariable sequences are specific for each person the restriction patterns are also specific. Changes in the DNA are reflected by corresponding changes in the restriction digest pattern of the DNA. Loci with base pair mutations, that create or destroy restriction sites, have only two alleles (a cleavage site is either present or absent). In contrast minisatellite loci, where length variation depends on the number of copies of the tandem repeat, have many possible alleles (Nakamura *et al.*, 1987).

A probe, complementary to a unique sequence found flanking a repeat region, can be used to get information for the allelic composition of that particular locus (single locus

fingerprint). In single locus probe (SLP) analysis DNA sequence variability is examined at a defined genetic locus, thus providing only two bands per individual. A genetic sequence that contains tandem repeats but represents only a single locus is referred to as a VNTR (variable number of tandem repeat) locus. VNTR loci show high levels of length variation and can therefore generate several allele variants. In addition to exhibiting more than two allele sizes VNTR loci, cut by different restriction enzymes, produce fragments of different length. Even though the DNA sequences of VNTR regions are members of a large set of similar sequences located at many different loci, there is sufficient variation in sequence amongst such loci that hybridisation at high stringency permits the unique identification of a single locus. At lower stringencies additional hybridising DNA fragments are detected with single locus probes (Wong *et al.*, 1986). Sequence alterations can be . lied at different sites depending on the sequence of the probe used and the hybridisation stringency employed.

The so-called simple quadruplet repeat (SQR) elements are another class of repetitive DNA. Their definition rests upon their actual sequence content, viz. short motifs, usually four bases, which are reiterated in tandem many times in a head-to-tail fashion (Epplen, 1992). This is in contrast to the minisatellite sequences described earlier which are constructed from larger repeat sequences. A common feature amongst all the repetitive DNA sequences is their high degree of polymorphism between unrelated genomes. Most of the variability found at SQR sites, as with minisatellites, results from different copy numbers of the basic repeat motif. In the plant kingdom a whole variety of these simple repetitive motifs are present, though to a dissimilar extent and organisation in different species. Probes usually reveal interspecific genetic variability, resulting in polymorphic or hypervariable banding patterns (Weising *et al.*, 1991). Depending on the combination of species and oligonucleotide probe species, variety or individual specific DNA fingerprints can be obtained. Unfortunately, the most informative combination (which probe with which species?) has to be determined empirically.

These short, simple sequence stretches (which occur as highly repetitive elements in all eukaryotic genomes) are thought to arise by slippage-like events working randomly on internally repetitive sequences and they generally show length variability. This is different

from the variation seen at minisatellite loci, which is thought to be due to recombination mechanisms. Slippage mutations are sufficiently frequent to maintain a high degree of polymorphism within populations, but not frequent enough to occur in successive generations (Tautz, 1989).

The high degree of polymorphism, found throughout the genome, suggests that there exists a large amount of variability in the DNA sequences of all eukaryotes. A great deal of variation between plant species is found in the repetitive components of their DNA - differences are found in the arrangement of the repeats and polymorphisms in their core sequences occur (Rivin, 1986). Restriction site analysis is used to detect variation amongst alleles as fragment size differences which can result from the presence or absence of restriction sites or insertion or deletion variations. These fragments are separated according to size by gel electrophoresis and visualised by incubation with a labelled probe (Hartl and Clark, 1989) - thus giving a DNA fingerprint. Variation in restriction sites is detected by the change in migration patterns of the gel bands.

Each probe will only detect its complementary (or near complementary) sequence if present. With the use of a series of different probes additional polymorphic loci might be detected because each probe will detect a different sequence. The pattern complexity is not the same for each sequence motif used for hybridisation (Jeffreys *et al.*, 1985).

CHAPTER 2

2. PREPARATION OF TOTAL GENOMIC DNA

2.1 Introduction

The ability to extract and purify high molecular weight genomic DNA is of the utmost importance for DNA profiling studies. The DNA fragments must be at least 20 kb long. They must be free of contaminants and low molecular weight nucleic acids fragments which interfere with restriction digestion and show up as unwanted background noise during hybridisation. To obtain undegraded DNA one must overcome the activity of residual and adventitious DNase enzymes. The isolation of plant DNA is technically demanding because the same forces which are required to break plant cell walls can also shear DNA. Considerable care must be taken and a compromise reached between DNA length and yield (Murray and Thompson, 1980). The general procedure for extracting DNA involves the disruption of the cells to release their DNA, removal of cell debris and polysaccharides, denaturing and precipitating proteins and the final precipitation of the DNA (Robyt and White, 1987).

Lithop genomic DNA has never, to the best of our knowledge, previously been prepared. Given the difficult nature of isolating undegraded plant DNA we decided to canvass the literature for a protocol most likely to be successful. Phenol extraction is most probably the most common method. The fundamental aim of a phenol extraction is the denaturation of proteins and their removal from an aqueous solution containing the desired DNA. Phenol is mixed with the sample under conditions which favour the denaturing and dissociation of the proteins from the DNA usually in the presence of a chaotropic agent like guanidinium or a detergent. Chloroform is usually added to remove the phenol and denatured proteins and any other organic contaminants which will partition into the organic phase. Centrifugation of the mixture yields two phases: a lower organic phase carrying the protein (much of which forms a white flocculent interphase) and the less dense aqueous phase con-

taining the DNA. Lipids which are soluble in chloroform can also be removed. Chloroform's high density enhances the separation of the two phases, while isoamyl alcohol is added to prevent foaming.

Phenol extraction is not the only method for isolating DNA and several other protocols are available. Richards (1990) describes a procedure for isolating plant DNA which employs a detergent for cell lysis, protease treatment and CsCl gradient purification. Whole cells are lysed in this procedure and so total cellular DNA (viz. nuclear, mitochondrial and chloroplast) is obtained. Dellaporta *et al.* (1983) use a similar procedure for isolating plant DNA, only omitting the CsCl gradient purification. DNA is usually precipitated with ethanol by leaving a mixture of the DNA sample, salt and ethanol at -20° C, or lower. The nucleic acid salt is sedimented by centrifugation and the ethanol supernatant is removed. The DNA pellet is dried and resuspended in buffer. Isopropanol can be used to minimise the total volume of the precipitating solution. However, isopropanol is not easily removed because of its low volatility and the sample usually requires an ethanol (or ether) wash after precipitation.

The most common problem affecting plant DNA purification is polysaccharide contamination. Carbohydrates are usually very difficult to separate from the DNA itself. The recommended procedure for polysaccharide removal is by chloroform extraction of cell lysates in the presence of cetyltrimethylammonium bromide (CTAB - a cationic detergent) and 0,7 M NaCl, first described by Murray and Thompson (1980). The technique capitalises on the fact that DNA and RNA are soluble in CTAB and 0,7 M NaCl. However, many polysaccharides are insoluble at this salt concentration and they will separate into the interface which is discarded. Weising *et al.* (1991) isolated DNA from a wide number of plant species, with variable success, using a technique based on that of Murray and Thompson (1980) (see figure 32). The DNA obtained by this procedure can be further purified by CsCl gradient centrifugation. For routine work this step is both costly and time consuming. Weising *et al.* (1991) recommend a RNase treatment followed by Sepharose CL6B spun column chromatography as an effective purification alternative.

2.1.1 Purification by Gel Filtration

A simple gel filtration system can be used for separating high molecular weight DNA from low molecular weight contaminants, e.g. degraded nucleic acids. Chromatographic separation of molecules of different dimensions can be brought about by making use of their varying ability to penetrate into a suitable stationary phase. A porous gel packed into a column serves as the stationary phase. Any solute molecule will partition itself between the solvent in the mobile and stationary phases (Morris and Morris, 1976). Selection of a gel having a suitable size exclusion range means that high molecular weight DNA fragments will be sterically excluded from the stationary phase solvent as they are too large to enter the gel pores. These molecules will elute from the column first in the void (or excluded) volume. Conversely the elution of smaller nucleic acids and other contaminants will be differentially retarded depending on their size. The elution pattern of the nucleic acids can be followed by measuring their absorbance at 260 nm with a UV detector. The operational limit is set by the viscosity of the initial solute solution. An increase in viscosity produces a progressive deterioration in the symmetry of the solute zones.

We decided to use BioGel A15m as the stationary phase because of its operating range, viz. $4x10^4$ - $15x10^6$ Da. BioGel is an agarose gel which has high porosity and mechanical rigidity. Its lack of defined structure restricts its use to temperatures below 40° C and to the pH range 4-9. Agarose gels must be stored in the wet state as they cannot be dried without loss of structure (Morris and Morris, 1976).

2.1.2 Estimating the Total DNA Content

Before attempting to isolate high molecular weight DNA it is useful to estimate the amount of DNA in the tissue. By far the easiest and most commonly used procedure for determining the concentration of DNA is by measuring the absorbance of an aqueous extract at 260 nm (1 mg/ml of double stranded DNA is known to have an A_{260} of about 20). This is generally a sufficiently accurate measure of the quantity of DNA. However it can be confirmed by other quantitative methods.

Two other procedures have also been used to estimate the total DNA content of lithops in the hope of confirming the values obtained by A_{260} estimations. The two wavelength method described by Tsanev and Markov (1960) was developed to calculate the amount of nucleic acid in various tissue extracts spectrophotometrically. Several other procedures for calculating the quantity of DNA and RNA were known e.g. phosphorus analysis and pentose colour reactions, but these estimates were prone to inaccuracies because of contaminating molecules. Tsanev and Markov showed that it was possible to determine suitable wavelengths at which the absorption of contaminants would be eliminated by the two wavelength method. The amount of RNA was estimated following alkali hydrolysis in 1 N KOH for 18 hours at 37° C (DNA is not susceptible to alkali hydrolysis at this temperature). Fleck and Munro (1962) showed that prolonged digestion in strong alkali released considerable amounts of protein into the RNA fraction. This could be avoided by using a lower concentration of alkali (0,5 N) and by hydrolysing for only one hour at 37° C.

The other procedure is the specific colour reaction between deoxyribose and diphenylamine described by Stenesh (1984). Since the sugar moiety of DNA consists exclusively of 2-deoxyribose the colour reaction serves equally well with the intact nucleic acid or the free sugar. The blue colour produced with DNA measures only the purine bound sugars. The intermediate responsible for the blue colour is thought to be hydroxylevulinic aldehyde and only substances that can be transformed into this compound will give the specific reaction.

2.2 Materials and Methods

2.2.1 Plant DNA Isolation

One gram of lyophilised lithop leaf tissue was ground up and dispersed in 15 ml prewarmed (60° C) CTAB buffer (2% w/v CTAB; 1,4 M NaCl; 0,2% mercaptoethanol; 20 mM EDTA and 100 mM Tris-HCl, pH 8,0) and incubated for 60 minutes at 60° C. This mixture was then extracted several times with a chloroform-isoamyl alcohol (24:1) mixture and centrifuged at room temperature for 10 minutes at 10 000xg. The aqueous phase was recovered and the nucleic acids were precipitated with 0,6 volume of isopropanol at -20° C for at least two hours. The precipitate was collected by centrifugation at 4° C for 10 minutes at 10 000xg and the resultant pellet was washed with 2 ml of 76% ethanol, 10 mM ammonium acetate for 20 minutes with gentle stirring. The insoluble DNA pellet was collected by centrifugation at 4° C for 10 minutes at 10 000xg, air dried to remove any traces of ethanol and then dissolved in 1 ml TE buffer (10 mM Tris-HCl; 1 mM EDTA, pH 8,0).

After resuspension the sample was digested overnight with 5 units of RNase at room temperature. The RNase was boiled, prior to use, to denature any contaminating DNases. DNA was precipitated with isopropanol and dissolved in 100 µl of TE buffer before being further purified. High molecular weight DNA could be spooled directly from the RNase digested extract on addition of isopropanol and was used without any further purification.

2.2.2 Total DNA Content of Lithops

2.2.2.1 Two Wavelength Method

The total DNA and RNA content of lithops was estimated by the two wavelength method of Fleck and Munro (1962). 80-120 mg of tissue was broken up in 3 ml 1 N $HClO_4$ at 4° C and finely ground. The insoluble material was collected by centrifugation and washed four times by resuspension in 3 ml 1 N $HClO_4$. RNA was extracted from the pellet by resuspension in 3 ml 0,5 N NaOH at 37° C for 1 hour. Dissolved protein and DNA was repre-

cipitated by addition of 2 ml cold 1 N $HClO_4$. The precipitate was collected by centrifugation and further extracted for DNA. The RNA phosphate content of the supernatant was determined from the following equation:

RNA P (µg/ml) = $11,87(A_{260}) - 10,4(A_{275})$

DNA was extracted from the pellet after first washing in cold 1 N $HClO_4$ by heating in 3 ml 1 N $HClO_4$ for 20 minutes at 90° C. After centrifugation the DNA phosphate content of the supernatant was determined from:

DNA P (µg/ml) = $8(A_{268} - A_{284})$

To convert the values of RNA P and DNA P to the mass of RNA and DNA one must multiply by a factor of ten.

2.2.2.2 Diphenylamine Colour Reaction

Duplicate standards of known DNA concentrations were set up, viz. 0, 5, 10, 25, 50, 100 and 200 µg/ml with duplicates of the DNA preparation from the two wavelength method. 4 ml of diphenylamine reagent (0,6 g diphenylamine; 0,6 ml concentrated sulphuric acid and 40 ml glacial acetic acid) was added to each tube which were then placed into a boiling waterbath for 10 minutes. After this incubation the samples were immersed in cold water and their absorbances at 600 nm were measured. The tubes containing 0 µg/ml of DNA were used as blanks.

2.2.3 Gel Filtration Chromatography

A small 1 ml column, though not possessing many theoretical plates, can readily separate large DNA fragments (in the void volume) from lower molecular weight contaminants which are retarded. A 1 ml syringe (dimensions: height = 6,5 cm and diameter = 0,6 cm) was plugged with glass wool and filled with BioGel A15m. Samples were loaded and eluted in 0,1M KCl. A Linear UVIS 200 HPLC detector, with a flow cell of 8 µl, was used to follow the elution profile at 260 nm. The thin HPLC tubing helped to regulate the flow rate (1 ml/hr) and minimise the "dead space" volume. The high molecular weight DNA fraction

was collected and concentrated by spooling after the addition of 0,6 volume isopropanol. The spooled DNA was washed with ethanol, dried and resuspended in TE buffer.

The elution volume of high molecular weight calf thymus DNA (obtained from Boehringer Mannheim) was taken to be the measure of the void volume. The total column volume was taken as the volume required to elute AMP. High molecular weight DNA was also purified from human, chicken and potato genomic DNA preparations using this system.

2.3 Results

2.3.1 Lithop DNA Isolation

30 A_{260} units were recovered in the first isopropanol precipitation, prior to RNase digestion, which is equivalent to 1,5 mg of nucleic acid. 1 mg was found to be RNA (which accounts for approximately 83% of the A_{260} units) and only 215 µg is recovered as DNA. The amount of DNA determined by the two wavelength and diphenylamine methods is 230 and 180 µg of DNA per gram lyophilised tissue respectively. This is similar to the published yield of Murray and Thompson (1980) which ranges from 200 to 700 µg of high molecular weight DNA per gram lyophilised tissue.

Most of the lithop DNA is degraded and only 15 µg of high molecular weight DNA was recovered by spooling after the addition of isopropanol, the remainder is found in the extract. That this is not the fault of the spooling can be seen by the use of gel filtration chromatography. The amount of high molecular weight lithop DNA recovered from the column was identical to that recovered by spooling, i.e. only 7% of the A_{260} units loaded onto the column were eluted in the void volume (see figure 5). The rest were recovered in the total column volume fraction.

2.3.2 Gel Filtration Chromatography

The void and total column volumes of the BioGel were found to be 300 and 1000 µl respectively (see figure 1). Elution profiles give a visual representation of the degree of DNA degradation that has occured during the isolation of DNA from different organisms (see figures 2 to 5). DNA samples that are not highly degraded will have a greater percentage of A_{260} units eluting in the void volume than degraded samples. The removal of low molecular weight contaminants by gel filtration can be seen from the difference in the electrophoretic characteristics of the purified and unpurified samples (see figure 6).

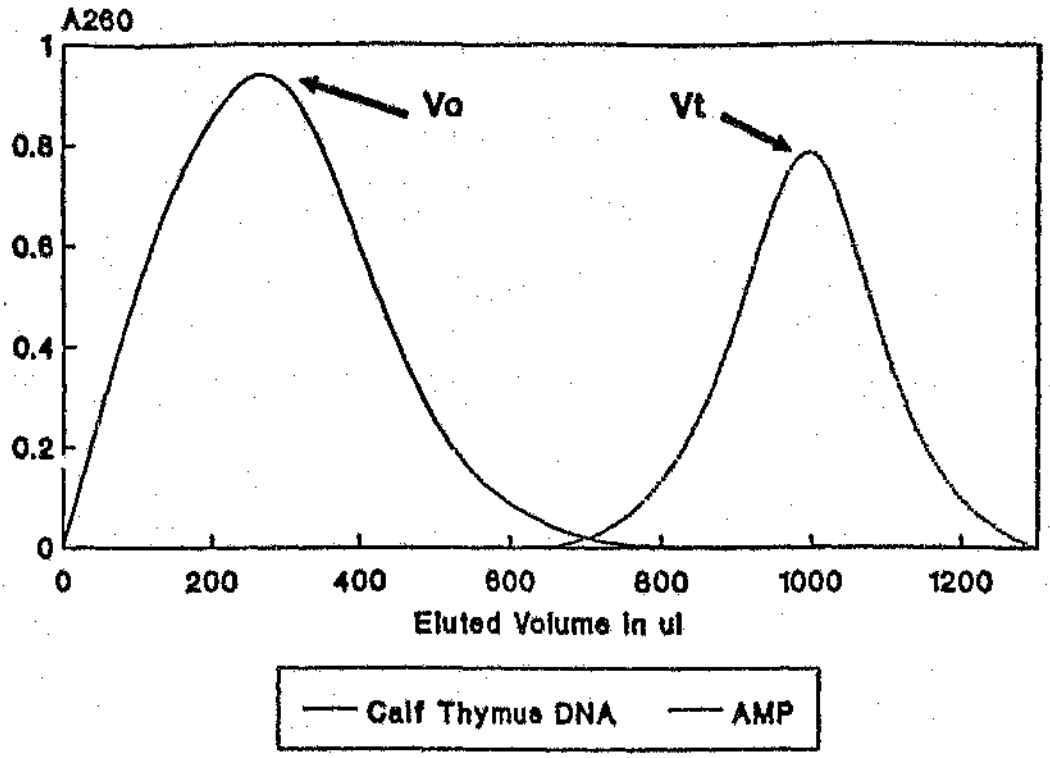


Figure 1: Determination of the void and total column volumes of BioGel A15m column. 50 µg or 1 A_{260} unit of high molecular weight calf thymus DNA was used as the void volume (Vo) indicator. 50 µg of AMP was used to measure the total column volume (Vt).

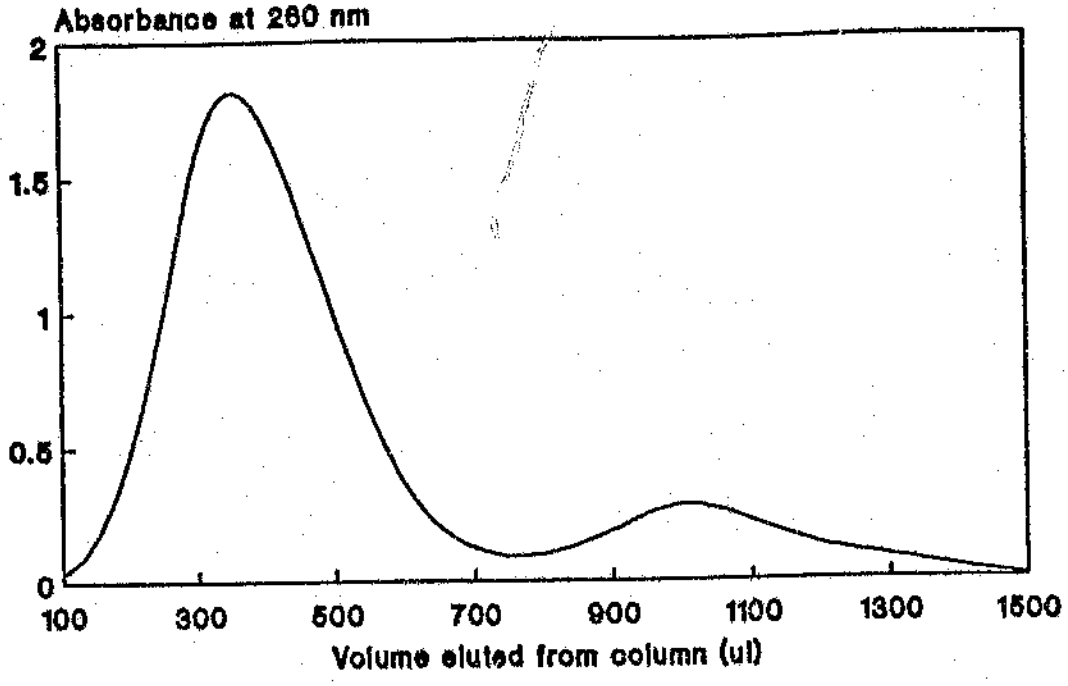


Figure 2: Elution profile of human genomic DNA through BioGel A15m. Human genomic DNA samples were obtained from Human Genetics, SAIMR (Michelle Ramsay's laboratory) and 80 µg (1,6 A_{260} units) fractions were loaded onto the column. 95% of the A_{260} units were recovered in the void volume fraction by isopropanol precipitation.

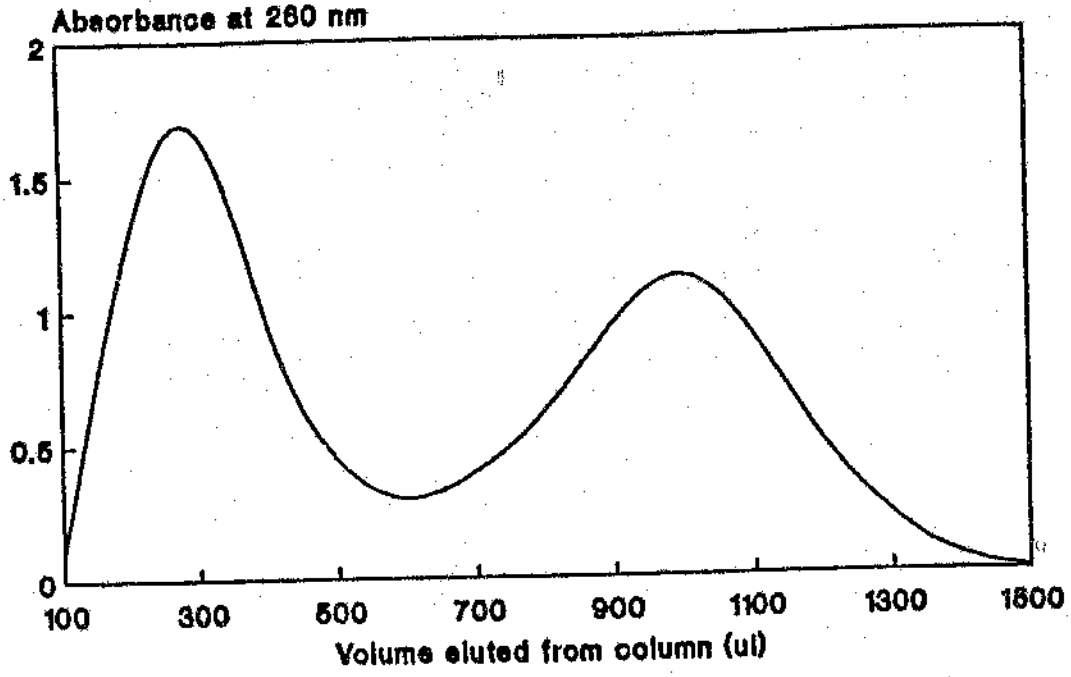


Figure 3: Elution profile of chicken genomic DNA through BioGel A15m. Genomic DNA was isolated from chicken blood (see appendix I). 80 µg (1,6 A_{260} units) was loaded onto the column in 0,1M KCl. This elution profile shows a significant degree of DNA degradation. 60% of the A_{260} units were recovered in the void volume.

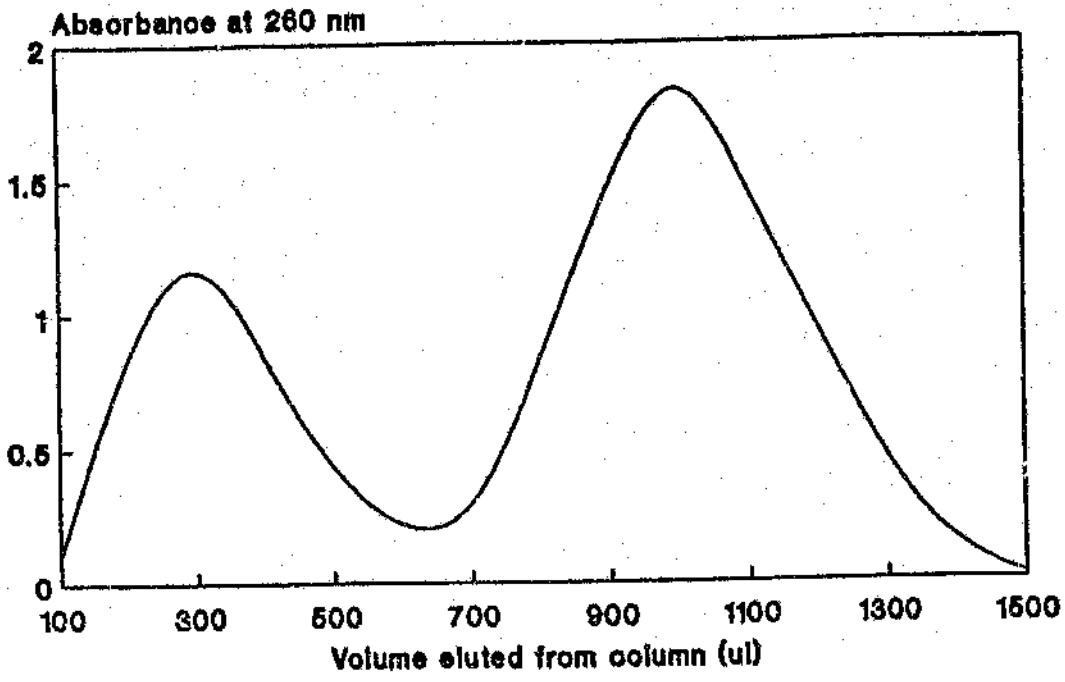


Figure 4: Elution profile of potato genomic DNA through BioGel A15m. Potato genomic DNA was isolated according to the method of Weising *et al.* (1991) and purified by gel filtration. Most of the DNA is degraded and only 0,4 A_{260} units (25% of the A_{260} units) were recovered in the void volume.

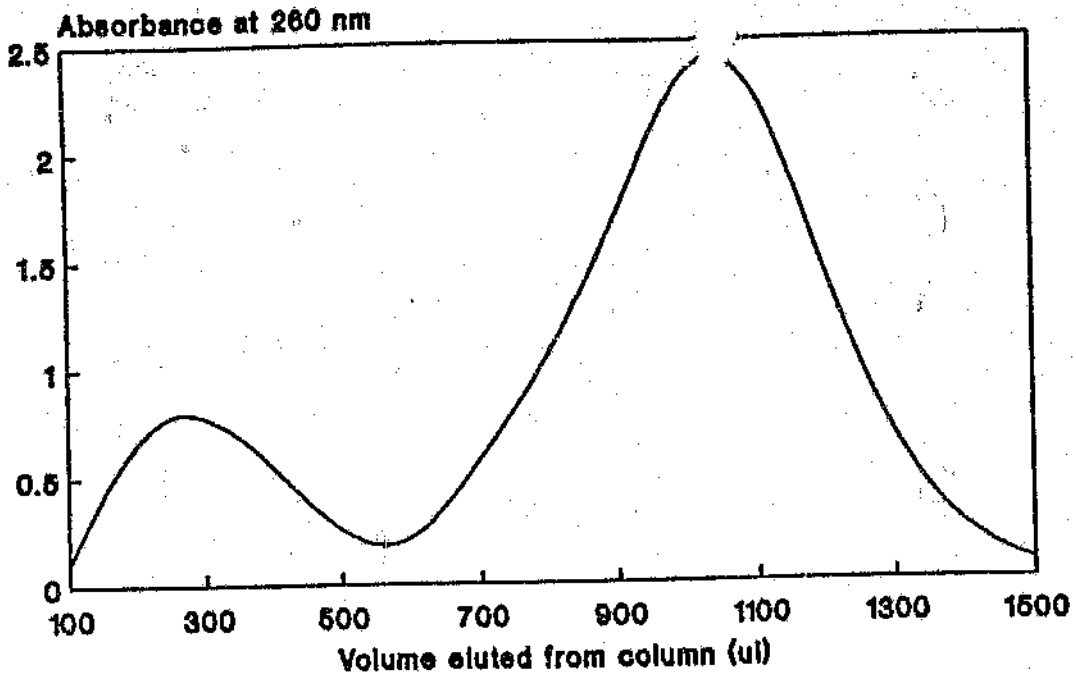


Figure 5: Flution profile of lithop genomic DNA through BioGel A15m. 80 µg of the total DNA preparation was loaded onto the column. There has been extensive DNA degradation and only 7% of the A_{260} units (about 5,5 µg) were recovered in the void volume.

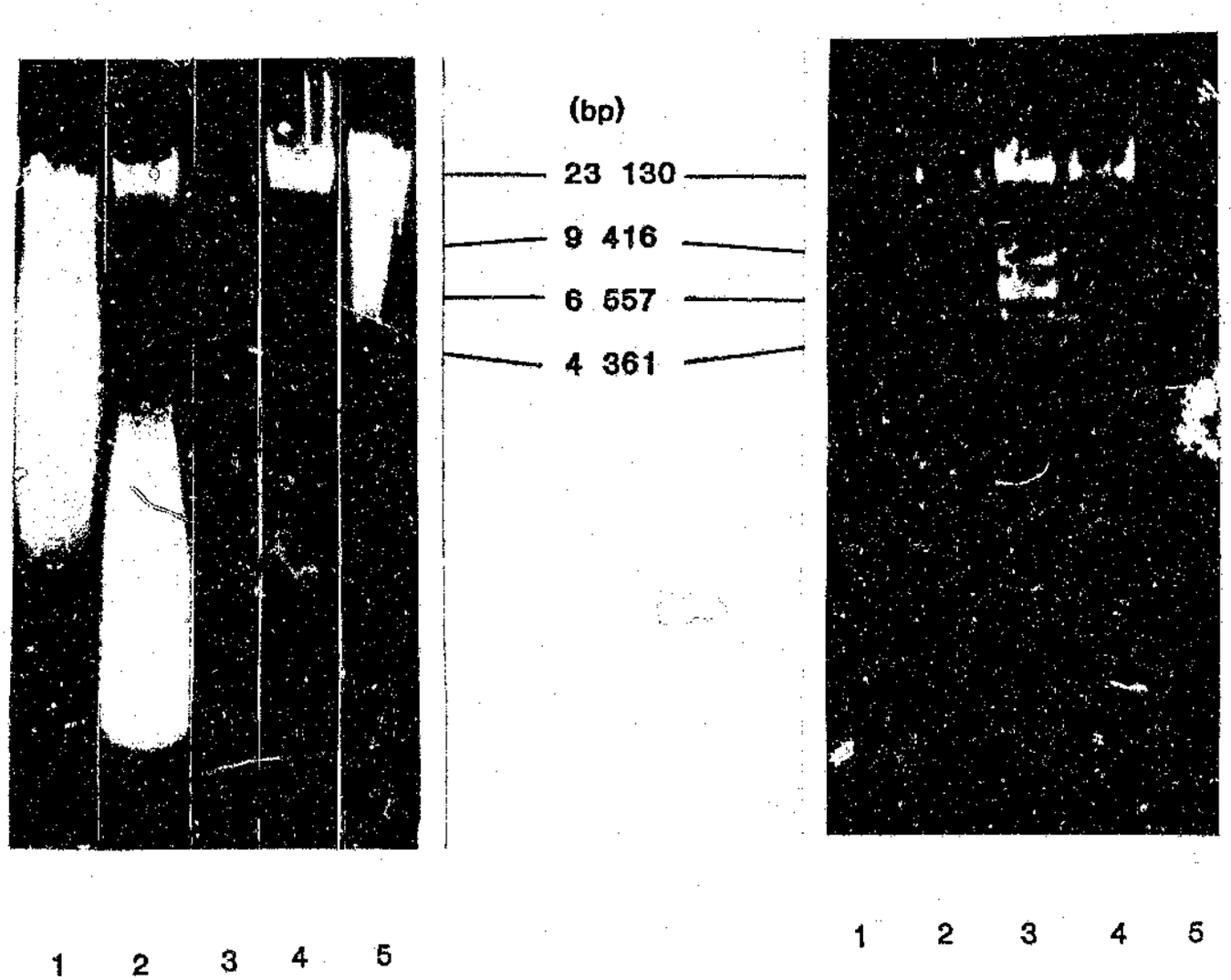


Figure 6: Agarose electrophoresis of high molecular weight DNA fractions. A 0,6% agarose gel was run at 60 V for 1 hour in TAE buffer (40 mM Tris- acetate; 1 mM EDTA pH 8,0) containing ethidium bromide as described by Maniatis *et al.* (1982). Gel A contains the samples before fractionation and gel B shows the same DNA preparations after purification on BioGel A15m. Lane (1) potato DNA, (2) lithop DNA, (3) molecular weight markers, (4) human DNA and (5) chicken DNA.

2.4 Discussion

2.4.1 DNA Isolation

The amount of RNA (determined by the two wavelength method) is 5 fold more than that of DNA, by weight, and makes up 83% of the A_{260} units. The RNA that is coprecipitated with DNA in isopropanol is therefore the major contaminant of the crude DNA preparation. Contaminating RNA can be digested with RNase and the undigested DNA is recovered by isopropanol precipitation.

The low yield of high molecular weight DNA (15 µg/g lyophilised tissue) imposes severe restrictions on the number of profiles that can be generated from a single DNA extraction since 5 to 10 µg of DNA are required for each profile. If the efficacy of the extraction procedure for isolating high molecular weight DNA cannot be improved, then several extractions have to be done in order to obtain sufficient DNA for Southern blotting and hybridisation. It is not unusual to obtain such low yields of DNA. Weising *et al.* (1991) have shown that there is very little uniformity in either the yield or quality of DNA isolated from different plant species. Lithops, like many succulents, have a high water content, the ratio of fresh weight to dry weight is greater than 10 to 1. The low yield of high molecular weight DNA per gram of dry weight tissue means that several plants have to be sacrificed to obtain sufficient DNA. This is not always possible especially for rare specimens.

The published yield of 200 to 700 µg DNA per gram lyophilised tissue by Murray and Thompson (1980) refers to the quantity of high molecular weight DNA. Although the yield of lithop DNA was about 200 µg/g lyophilised tissue, only 15 µg/g lyophilised tissue (or 7%) was high molecular weight DNA. There is no definitive reason to explain the low yield of high molecular weight lithop DNA, but precautions were taken to minimise sheering forces and DNase activity which are thought to be the cause. To reduce the risk of DNA degradation, lyophilised as opposed to fresh tissue is used because DNA is less susceptible to hydrolysis when water is removed. Leaf tissue can be stored for long periods in this de-

hydrated state without any DNA degradation. When the lyophilised tissue is rehydrated it is usually in the presence of EDTA and detergent (CTAB) which inhibit nucleolytic activity. The ability of EDTA to chelate metal ions is vital to inhibit DNase activity.

Polysaccharides are said to be a major contaminant of plant DNA preparations and can interfere in restriction digestion (Richards, 1990). A high salt concentration (1,4 M NaCl) has been introduced by Weising *et al.* during extraction because many proteins and polysaccharides are insoluble at salt concentrations greater than the 0,7 M NaCl used by Murray and Thompson. CTAB, which has a hydrophobic 16 carbon chain and a polar tertiary amine "head", is able to form soluble salts with nucleic acids under these conditions. When chloroform is added the insoluble polysaccharides and proteins form a white interphase while organic contaminants will partition into the chloroform phase. In the original method of Murray and Thompson DNA was recovered from the aqueous phase by adding 1% CTAB and lowering the salt concentration below 0,3 M NaCl. The DNA-CTAB salt obtained is fairly insoluble and difficult to redissolve. In Weising et al.'s variation of this method DNA is precipitated at the higher salt concentration as a sodium salt with isopropanol.

2.4.2 DNA Estimations

The three methods for estimating the total quantity of DNA, viz. A_{260}, diphenylamine and the two wavelength method were all in close agreement. Of particular interest is the diphenylamine estimate which is specific for 2-deoxyribose whether it is found in the intact nucleic acid or as the free sugar. Theoretically, this should give the absolute amount of DNA present because DNA has an unique 2-deoxyribose component.

2.4.3 Gel Filtration

Gel filtration chromatography produces a visual profile of the quality, or degree of degradation, of a total DNA preparation. The small 1 ml column could resolve two elution zones, viz. high molecular weight DNA eluting in the void volume and low molecular weight con-

taminants eluting in the total column volume. The area of the void and total column volume eluents are proportional to the number of A_{260} units in these fractions and can be determined from the elution profile. If the yield of high molecular weight DNA is low, as is the case for lithops, most of the A_{260} units will elute in the total column volume. If the yield is high, as for human DNA, then most of the A_{260} units will be in the void volume.

Gel filtration is a useful purification method because DNA fragments can be separated by size without using agarose gel electrophoresis. It is also a quick, inexpensive procedure for cleaning up DNA samples that have been degraded with time (see figure 6). High molecular weight DNA can be separated from low molecular weight contaminants which absorb at 260 nm but are not readily visualised by ethidium bromide fluorescence. The quantity of DNA, contaminated with low molecular weight contaminants, which is visualised in agarose gels containing ethidium bromide often appears to be less than what is judged by A_{260} measurements. These low molecular weight contaminants can be easily removed by gel filtration.

There are two factors which affect the efficacy of gel filtration as a preparative procedure. Firstly, even though it is possible to scale-up the method the quantity of DNA that can be loaded onto the column is limited by the viscosity of the DNA solution, i.e. concentrated DNA solutions are viscous and they disrupt the symmetry of elution zones, whereas dilute solutions are more readily resolved (Morris and Morris, 1975). Secondly, the volume of DNA that is loaded onto the column can compromise the resolution of adjacent elution zones. Increasing the loading volume (above 10% of the column volume) decreases the separation of the void and total column volume elution zones because the number of theoretical plates is reduced. 40 to 50 µl aliquots (4 to 5% of the column volume) were routinely found to produce distinct elution zones

2.5 Conclusion

The initial objective of this chapter was to isolate purified high molecular weight lithop DNA. This was achieved with partial success using the method of Weising *et al.* (1991), but only 15 µg of high molecular weight lithop DNA could be recovered from the 215 µg of DNA apparently present per gram of lyophilised tissue. Given the inherent difficulties in isolating plant DNA (Weising *et al.*, 1991) and the fact that lithop DNA has never previously been isolated, it was an accomplishment to obtain even some DNA. Although sheering and DNase activity are widely known to cause DNA degradation, the use of EDTA and lyophilised tissue did not increase the yield of high molecular weight lithop DNA. No suitable explanation could account for this anomaly and this meant that only a few DNA profiles could be generated from each DNA extract.

Gel filtration was shown to be a quick and useful method for purifying high molecular weight DNA from low molecular weight contaminants and for generating a visual profile of the "quality" of the extracted DNA.

CHAPTER 3

3 SIMPLE QUADRUPLET REPEAT (SQR) ANALYSIS

3.1 Introduction

DNA profiling is increasingly being applied to areas such as forensic analysis, paternity testing, animal breeding and population genetics. The selection and specific labelling of probes is a crucial factor in hybridisation technology. The target DNA sequences that are recognised by the profiling probes (minisatellites and SQRs) are present not only in animal but also in plant genomes (Bierwerth *et al.*, 1992). Simple repetitive motifs are ubiquitous components of the genome of higher plant species. Weising *et al.* (1991) have shown the unique distribution and organisation of some of these motifs throughout the plant kingdom.

Different combinations of probe and template DNA yield different hybridisation patterns which can vary from a smear pattern (if the motif is abundant) to faintly detectable bands (if the motif is rare or only partially complementary to the probe). Those combinations which generate a scoreable banding pattern can be used for profiling.

Almost any sequence can be cloned into a vector, plasmid or phage and be labelled and used as a probe. In many cases, however, the signal to noise ratio is suboptimal, especially if the sequence is small when compared to the size of the vector (Feinberg and Vogelstein, 1983). Synthetic oligonucleotide probes eliminate the need for cloning "minisatellite" sequences. They can be designed or modified according to a researcher's requirements. Furthermore, oligonucleotide probes hybridise effectively with DNA fragments immobilised in dried agarose gels because they are small enough to diffuse into the dry gel matrix. This means that hybridisation can be carried out directly in the dried agarose gel, eliminating the need for time-consuming Southern blotting and therefore reducing any loss of DNA due to non-optimal transfer conditions (Epplen, 1992 and Z'schler *et al.*, 1991). The technological ad-

vance in the chemical synthesis of oligonucleotides has led to a concurrent increase in their use as complementary probes for SQR "alleles".

It is advantageous if one is able to predict which oligonucleotide sequence is most likely to generate an informative profile pattern with a given genomic DNA template. The presence and overall quantity of simple tandem repeats, in a given genome, can be assessed by slot-blot hybridisation techniques (Epplen, 1992). But this does not prove the suitability of a probe for DNA fingerprinting because it only indicates whether a sequence is present and not its distribution. It is generally better to start out on the basis of previous experiences in related species. However the principle of "trial and error" has to be adopted. Since no sequence information on lithops is available it seemed reasonable to attempt to generate SQR profiles for lithops with the SQR oligonucleotides described by Weising *et al.* (1991).

3.1.1 Specificity of Hybridisation

Probe hybridisation procedures are rarely sufficient to discriminate between perfect probe target hybrids and hybrids formed with close sequence relatives. It can be assumed, for simplicity, that hybridisation doesn't discriminate against random sequences 95% or more homologous to the probe. A probe and its target may share a certain overall homology, yet hybridise in different ways. Incorrectly hydrogen bonded pairs (mismatches) may be dispersed uniformly throughout the length of the hybrid or clustered in one particular region. The relative stability of hybrids formed by dispersed versus clustered homology is an important consideration, since the background of random sequences presents the possibility of dispersed homology whereas the target sequence usually does not. Clustered homology generates a more stable hybrid. The destabilising effect of dispersed homology may be slightly offset by the stability increase that accompanies increased duplex length. A long hybrid with less than 100% homology has a comparable stability to a much shorter hybrid with 100% homology. Short stretches of perfect homology occur at a significant frequency and probe length must be increased to allow discrimination against these short hybrids (Lathe, 1985).

The incorporation of a single mismatched base pair destabilises a DNA duplex to some extent, the amount depending on the nature of the mismatched base pair. Experimentally the mismatched duplex is discriminated from the perfectly matched one in two ways:-

1. the hybridisation efficiency of the mismatched oligonucleotide is lower than that of the perfectly matched duplex and

2. the dissociation of the mismatched duplex is higher than that of the perfectly matched duplex (Ikuta *et al.*, 1987). This stems from the fact that the existence of a single mismatched base pair in an oligonucleotide-DNA duplex significantly reduces its thermal stability. This effect allows for the formation of the perfectly paired DNA duplexes under conditions where DNA duplexes with a single mismatch do not form.

The activation energy for the dissociation of duplexes with various mismatches are similar, suggesting that the difference in dissociation rates is entropic in origin. The activation energy for dissociation seems to be too small for the overall strand separation. This implies that the melting of the first several base pairs is the actual rate limiting step of the overall dissociation process.

Despite these constraints certain "stable" mismatches can be tolerated. Guanine is the usual partner involved in the formation of stable mismatches because it best fits the classic criteria for Watson-Crick base pairing, viz:

- formation of a stable base pair requires at least two hydrogen bonds
- one of these hydrogen bonds utilises the imino proton of guanine or thymine
- none of the bases converts to the enol or imino tautomers
- the classical base pair is constructed from a combination of imino proton acceptor adenine or cytosine with donor thymine or guanine

- combination of the larger purine bases adenine (A) and guanine (G) with the smaller pyrimidine bases cytosine (C) or thymine (T) form sterically well balanced base pairs.

The stable mismatch guanine-thymine has most of the above mentioned features except that both of these bases are imino proton donors. The other stable mismatch guanine-adenine also has most of the above features except both of these bases are purines.

Among the unstable mismatches T-T and C-T destabilise the duplexes more than A-A and C-A do. This is probably due to the fact that the former include only pyrimidines which have a weak stacking effect to neighbouring bases, which results in an excessive distortion of the helical structure (Ikuta *et al.*, 1987).

3.1.2 Theory of Hybridisation

The objective of any hybridisation procedure is to optimise the signal and to reduce all the unwanted background "noise". This is achieved by employing hybridisation conditions that favour the dissociation of unstable mismatches without disrupting perfect hybrids. One of the major problems of various hybridisation techniques that use synthetic oligonucleotides is the lack of well described conditions for each oligonucleotide (Albretsen *et al.*, 1988). The formation of nucleic acid hybrids is a reversible process and an understanding of the parameters which affect their stability enable one to derive the optimal conditions for discriminating between perfect and imperfect hybrids.

Methods for immobilising DNA onto a solid support and for detecting the fixed nucleic acid with a labelled probe are referred to as mixed phase hybridisation techniques (Meinkoth and Wahl, 1984). Mixed phase hybridisation technology was originally limited in its applications by the inability to obtain diverse probes. Molecular cloning and organic synthesis have now eliminated this problem.

The kinetics of hybridisation of nucleic acid probes with DNA tethered to a solid support or free in solution are very similar (Meinkoth and Wahl, 1984). This suggests that parameters

which affect nucleic acid reannealing in solution will have similar effects in mixed phase systems. The rate of hybrid formation with single stranded probes should follow first order kinetics if the concentration of probe is in vast excess over the target sequences. The optimal temperature for hybridisation in solution is determined empirically and usually found to be between 5 and 20 degrees lower than the Tm of the duplex (Tm is the temperature at which half of the hybrids are dissociated). The melting temperature is affected by ionic strength, base composition of the probe and the length of the shortest chain in the duplex. Duplex stability, between strands with mismatched bases, is decreased according to the number and location of the mismatches - this effect is especially pronounced when using short oligonucleotide as probes (Albretson *et al.*, 1988).

Stringency is usually altered by adjusting the salt and/or temperature either during hybridisation or in the posthybridisation washes. It is good technique to perform the hybridisation at low stringency (low temperature and high ionic strength) and wash at progressively higher stringencies, analysing the results after each wash. This enables one to detect related sequences and monitor the efficacy of the washes in removing these sequences. The ratio of the extent of hybridisation of perfect and mismatched duplexes is critically dependent on the stringency conditions at which hybridisation and washing are carried out. A compromise must be reached between the higher stringencies required to achieve dissociation of the mismatched duplexes and the lower stringencies needed to give the greatest degree of sensitivity.

3.1.3 Nonradioactive Hybridisation

Avoiding radioactivity and introducing nonradioactive detection has several inherent advantages. Neither health risks nor disposal problems are encountered. The probes are stable (up to one year) and the hybridisation solutions can be reused several times, allowing for convenient planning of experiments. No special laboratory facilities are required. The improvement in technology for detecting cold-labelled (nonradioactive) probes has increased their degree of sensitivity, e.g. detection with a chemiluminescent substrate or by enzyme linked

colour reaction. All these factors, in addition to the cost of radioactivity, have led to the increase in use of nonradioactive probes.

There are several systems available for nonradioactively labelling probes. The system that is chosen depends on the type of probe that is to be used, viz. whether an oligonucleotide or gene sequence is to be labelled. Whichever the case one or more of the probe's nucleotide bases is labelled by attachment of a "reporter" molecule. Boehringer Mannheim provides several easy to use nonradioactive labelling kits which use a digoxigenin (DIG) modified deoxyuridine triphosphate as the reporter molecule (Boehringer Mannheim - The DIG System User's Guide for Filter Hybridisation). The digoxigenin (steroid hapten) is linked to the modified uridine base via a spacer arm (Feinberg and Vogelstein, 1983). Large probes (100 to 10 000 bp) can be efficiently labelled by random primed labelling. This procedure uses the Klenow enzyme to synthesise a labelled copy of the probe by incorporating one digoxigenin-dUTP every 20-25 nucleotides. Synthetic oligonucleotides (14 to 100 nucleotides long) are 3'-end labelled with digoxigenin-ddUTP using terminal transferase. The hybridised probes are immuno-detected with an alkaline phosphatase conjugated anti-digoxigenin antibody and then visualised with the colorimetric substrates Nitro Blue Tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (X-phosphate).

Blots or gels that have already been hybridised can be stripped of their probe and rehybridised with another probe. This is a big advantage for lithop DNA digests, because each digest can be probed several times with a different sequence probe.
The digoxigenin system had to be shown to be effective for oligonucleotide (SQR) profiling by showing that:

1. oligonucleotides could be efficiently labelled and detected and

2. reproducible hybridisation patterns could be generated either on blots, in gels or both.

3.2 Materials and Methods

3.2.1 Oligonucleotide Synthesis and Purification

Four SQR oligonucleotides were synthesised using a Milligen Biosearch Cyclone Plus DNA synthesiser according to the manufacturer's instructions, viz. the two 15mers $(TCC)_5$ and $(GTG)_5$ and the two 16mers $(GATA)_4$ and $(GACA)_4$. The yield of oligonucleotide was calculated from the A_{260} units and the molar absorption coefficients of the bases using a simple computer program (see appendix II). Newly synthesised oligonucleotides were purified from residual salts by gel filtration chromatography, using a Sephadex G-15 column. The samples were loaded and eluted with water and the eluting species were detected by their absorbance at 260 nm.

3.2.2 3'-end Digoxigenin Labelling

100 pmol of oligonucleotide was added to 4 ul of reaction buffer (1M potassium cacodylate; 125 mM Tris-HCl; BSA 1,25 mg/ml; pH 6,6) and 4 µl of 25 mM $CoCl_2$ solution. 1 µl of 1 mM DIG-ddUTP solution and 1 µl of terminal transferase were added and the volume was made up to 20 µl with water. Thi` mixture was incubated at 37° C for 15 minutes and then placed on ice. The reaction was stopped by the addition of 2 µl of glycogen solution (a 20 mg/ml solution diluted 1 in 200 with 0,2 M EDTA). The labelled oligonucleotides were precipitated with 75 µl prechilled ethanol and 2,5 µl of 4 M LiCl. After being left for at least 2 hours at -20° C the mixture was centrifuged at 12 000 g and the resultant pellet was washed with 50 µl cold 70% ethanol. The pellet was dried and resuspended in 20 µl water.

The efficiency of detecting labelled oligonucleotides bound to nylon membranes was compared to that of larger DNA probes because small DNA fragments were thought to bind to these membranes less effectively. This was done by dot blot comparison of a standard 3'-end labelled 30mer (Boehringer Mannheim DIG Oligonucleotide 3'-End Labelling Kit) with a standard random labelled probe (Boehringer Mannheim DNA Labelling and Detection Kit

Nonradioactive). A 3 fold, stepwise dilution series of the random labelled and end labelled probes, containing equal quantities of label, were prepared and spotted onto a nylon membrane. The diluted probes were fixed to the nylon by cross-linking with UV light and the colour was developed (as described below). The efficiency of labelling the synthetic oligonucleotides was then checked by dot blot comparison with the standard labelled 30mer oligonucleotide.

3.2.3 Restriction Digestion and Electrophoresis

10 µg of lithop genomic DNA was restriction digested with EcoR1 overnight at 37° C before electrophoresis.

Gel electrophoresis was carried out in a minigel apparatus as described by Maniatis *et al.* (1982). Samples were mixed with a loading solution containing 7% sucrose and 0,025% xylene cyanol and bromophenol blue before loading into the gel. A 0,6% agarose gel was run at 70 V for 2 hours. TAE buffer (40 mM Tris-acetate; 1 mM EDTA, pH 8,0) was used as the running buffer during electrophoresis. Nucleic acids were visualised by ethidium bromide fluorescence under a UV light source. Longer agarose gels were also run using the same running conditions.

3.2.4 Hybridisation

After electrophoresis hybridisation was first tried with the restriction digested lithop DNA immobilised in the dried agarose gel. The gel was dried under vacuum and denatured in 0,5 M NaOH; 1,5 M NaCl for 1 hour. This was followed by neutralisation in 0,5 M Tris-HCl (pH 7,5); 2,5 M NaCl for 30 minutes. Hybridisation was carried out overnight with the labelled oligonucleotide in 5xSSPE (0,75 M NaCl; 50 mM NaH_2PO_4 and 5 mM EDTA pH 7,4), 5x Denhardt's solution with 0,5% SDS at the required hybridisation temperature. Probes that gave scoreable patterns were rehybridised to restriction digested lithop DNA which had been run in a longer agarose gel and Southern blotted onto a nylon membrane. The membrane was equilibrated in 6xSSC (0,9 M NaCl; 0,09 M sodium citrate pH 7,0) for

5 minutes. The hybridisation solution (5xSSPE; 5x Denhardt's solution; 0,1% SDS and 10 µg/ml E. Coli DNA) was prewarmed to the correct hybridisation temperature.

The hybridisation temperature that was used depends on the melting temperature (Tm) of the oligonucleotide probe. Tm was calculated from the probe's sequence using the formula:

$$Tm = 2(A+T) + 4(G+C)$$

The final hybridisation temperature was taken as Tm - 15° C. The gel/membrane was placed into a plastic bag with hybridisation solution and 40 pmol of DIG labelled oligonucleotide per millilitre of hybridisation solution and hybridised overnight. The hybridisation solution was decanted and stored at -20° C. The gel/membrane was washed 3 times for 30 minutes in 6xSSC at room temperature and then once in 6xSSC at the hybridisation temperature.

No prehybridisation or carrier DNA was required for in-gel hybridisation (Pena *et al.*, 1991).

3.2.5 Detecting the Nonradioactive Signal

The gels/filters were washed in buffer 1 (0,1 M Tris-HCl, 0,15 M NaCl pH 7,5). Filters were incubated for 30 minutes in buffer 2 (0,5% blocking reagent in buffer 1). The anti-digoxigenin antibody-conjugate was diluted to 150 mU/ml and incubated with the gel/filters after washing in buffer 1 for 30 minutes. The colour signal was developed using NBT and X-phosphate in 0,1 M Tris-HCl; 0,1 M NaCl and 50 mM $MgCl_2$ pH 9,5. This colour reaction was stopped by washing in 10 mM Tris-HCl and 1 mM EDTA pH 8 for 5 minutes.

3.3 Results

3.3.1 Purifying Synthetic Oligonucleotides by Gel Filtration

1 A_{260} unit of oligonucleotide was loaded onto the column and the yield of A_{260} units recovered in each fractioned was expressed as a percentage of this. 95% of the A_{260} units loaded onto the column eluted in the void volume fraction, this is the oligonucleotide fraction. The remainder of the A_{260} units are residual "contaminants" which are distributed amongst three remaining elution zones (see figure 7).

3.3.2 Labelling Efficiency

Short oligonucleotides were found to bind to nylon membranes very inefficiently. When equivalent quantities of digoxigenin label, from the random primed and 8 fold diluted end labelled standards, were blotted onto a nylon membrane the 3'-end labelled 30mer stained very weakly in comparison to the random primed labelled (RP) probe (see panel A of figure 8). If the RP probe is diluted further then only after a 3^8 or 6561 fold dilution, is the sensitivity of the RP comparable with the 8 fold diluted end labelled 30mer (see panel B of figure 8), i.e. a 800 fold dilution of the RP standard should have a similar staining intensity to the undiluted 30mer. I found that even after a 800 fold dilution of the RP standard it was detected with 3 times more efficiency over the undiluted 30mer, i.e. a combined staining difference of about 2400 (see figure 9). This is believed to be an estimate of the differential affinity of the probes for the nylon. To try to improve the low binding affinity of the 30mer for the nylon several pretreated membranes were used. However, there was no significant increase in the retention and subsequent detection of the 30mer when compared to the untreated membrane (see figure 10).

Labelling efficiency is usually determined by dot blot comparisons made between a newly labelled probe and a standard labelled probe. Boehringer suggest using dot blot comparison with their standard end labelled 30mer to determine the labelling efficiency of other end labelled oligonucleotides. Given the fact that the binding affinity of an oligonucleotide for a

nylon membrane is related to its size one cannot estimate the labelling efficiency of 15 and 16mers from the staining intensity of a 30mer because they may not bind to the nylon membrane with the same efficiency (see figure 11).

3.3.3 Hybridisation

The hybridisation duplex between a probe and its complementary sequence can be destabilised in solutions of low ionic strength (high stringency). Colour development takes place after the prescribed stringency washes have been done and a weak signal may be caused by the low ionic strength buffers (see Materials and Methods) used during this procedure. Simply raising the salt concentration of these buffers is not the solution because the colorimetric substrates NBT and X-phosphate become insoluble above 0,2 M NaCl. A staining agar, as suggested by Epplen (1992) and Bierwerth *et al.* (1992), can be used to prevent these substrates from precipitating in buffers containing salt concentrations greater than 0,2 M. When the salt concentration, used during colour development, is increased 9 fold a stronger signal is detected because the oligonucleotide/template hybrid is stabilised (see figure 12).

Of the four SQR sequence probes described by Weising *et al.* (1991) only $(TCC)_5$ gave a resolvable hybrisation pattern with EcoR1 digested lithop DNA immobilised in a dried agarose gel (see figure 13). When the DIG labelled $(TCC)_5$ oligonucleotide was hybridised to a Southern blot of EcoR1 digested lithop DNA only a high molecular weight smear was seen (see figure 14).

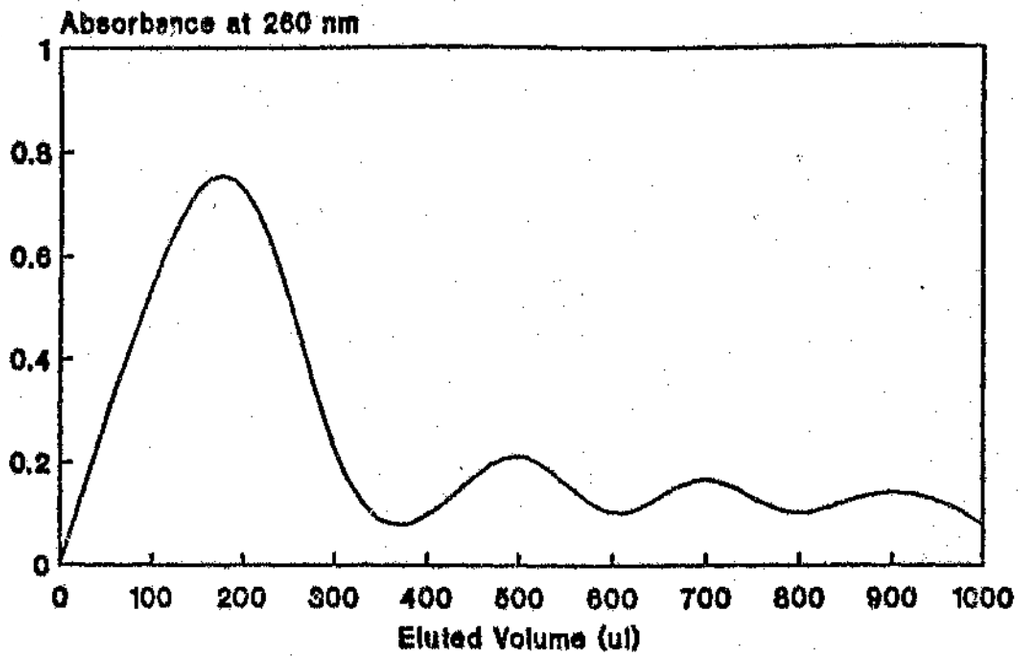


Figure 7: Purification of 1 A_{260} unit of oligonucleotide $(TCC)_5$ through Sephadex G-15. Other oligonucleotides were also purified from residual "blocking" groups by gel filtration chromatography through Sephadex G15.

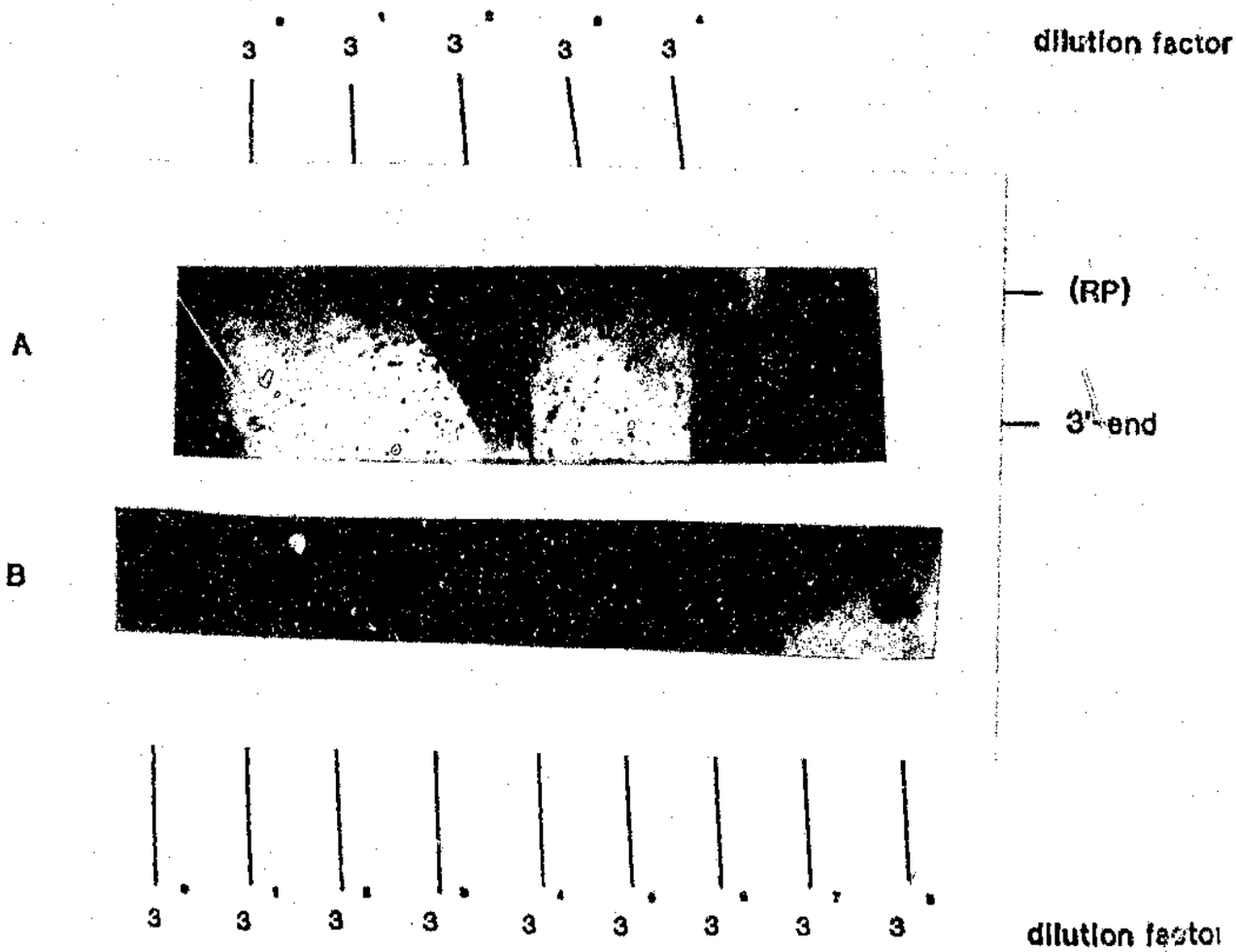


Figure 8: Dot blot of 3'-end labelled 30mer and standard random primed (RP) labelled probe. The concentration of DIG in the end labelled 30mer is 2,5 pmol/µl. This is 8 fold more than that of the random primed labelled probe which is about 0,3 pmol/µl (assuming the rate of DIG-dUTP incorporation to be 1 in every 22 bases), i.e. the 30mer was first diluted 8 fold. Panel A) dot blot of RP and 3'-end labelled probes and panel B) dilution series of RP probe.

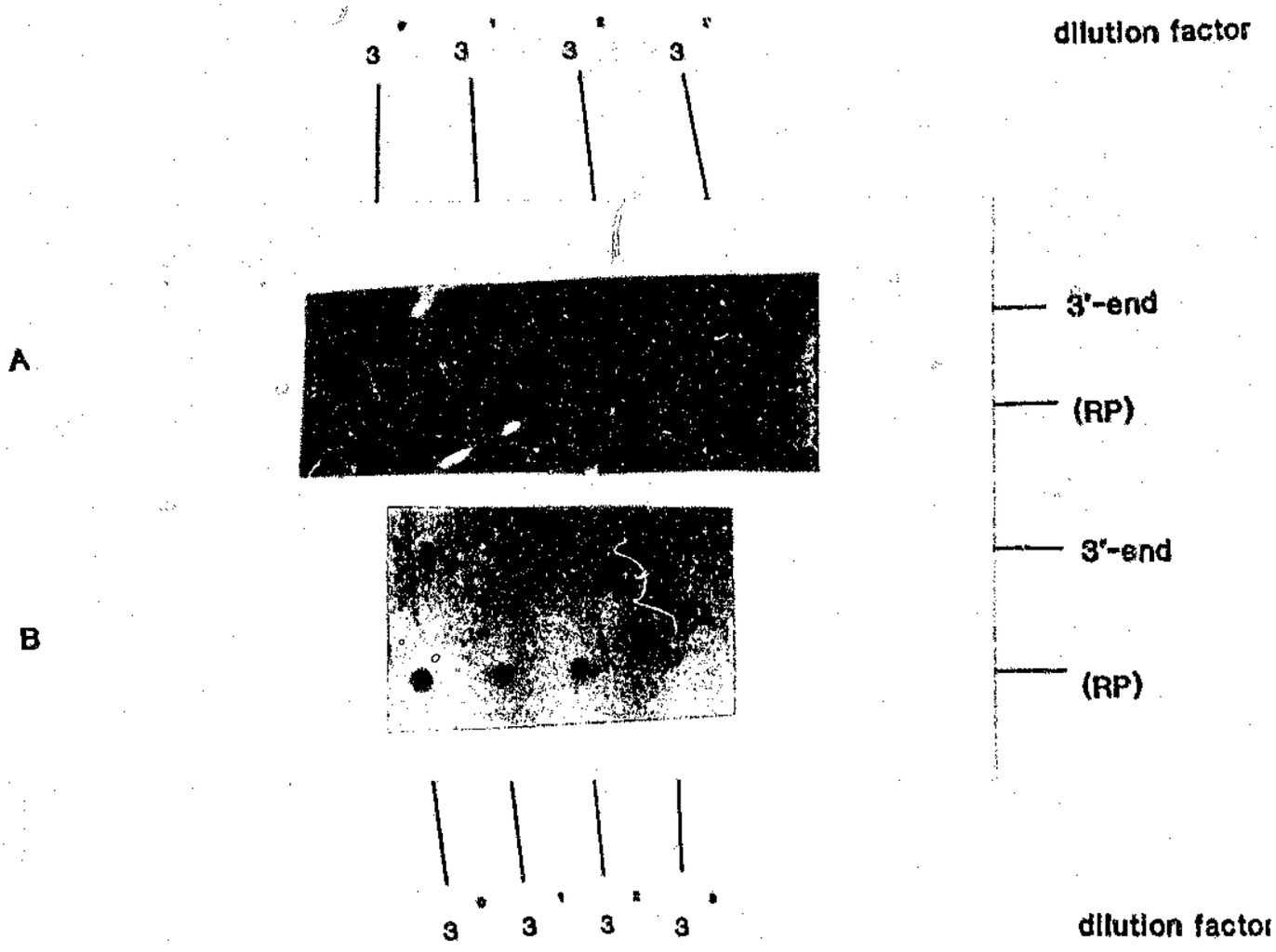


Figure 9: Dot blot of undiluted 30mer and 800x diluted random primed standard. Panels A and B both show that although there is an improved correlation between the two dilution series, the RP standard is still 3 fold more sensitive than the 30mer.

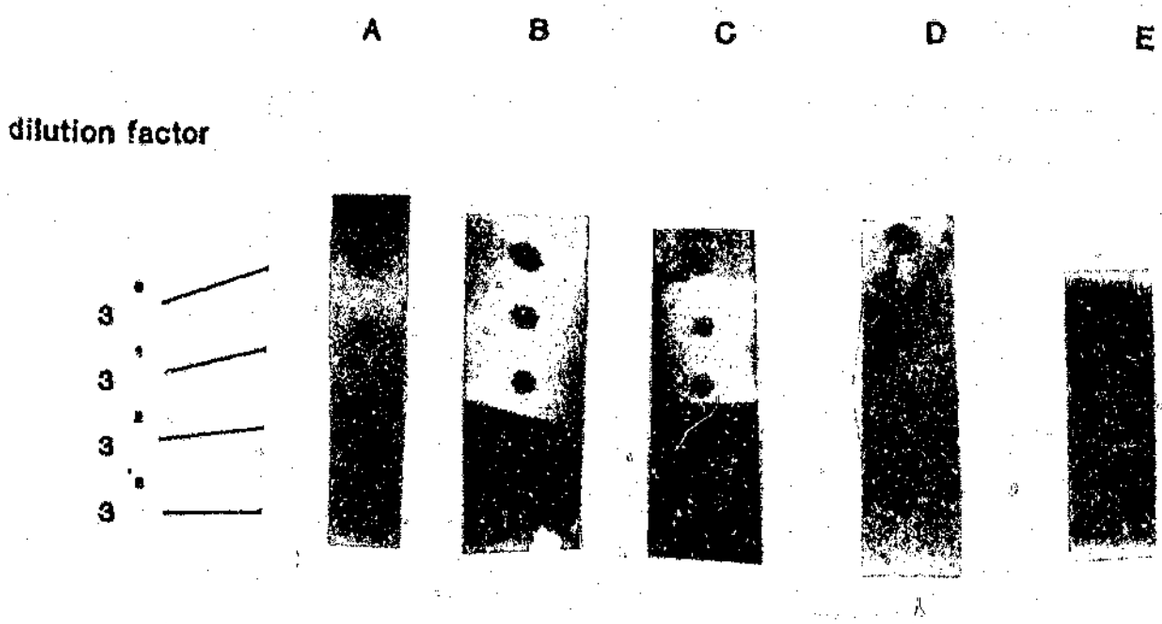


Figure 10: Pretreatment of nylon membranes to improve the binding efficiency of oligonucleotides. The nylon membranes were first soaked in BSA (panel A), gelatin (panel B) and herring sperm DNA (panel C) before cross-linking the 30mer to the nylon membrane. BM positively charged nylon membrane was also used (panel D). Panel E) is the untreated control membrane.

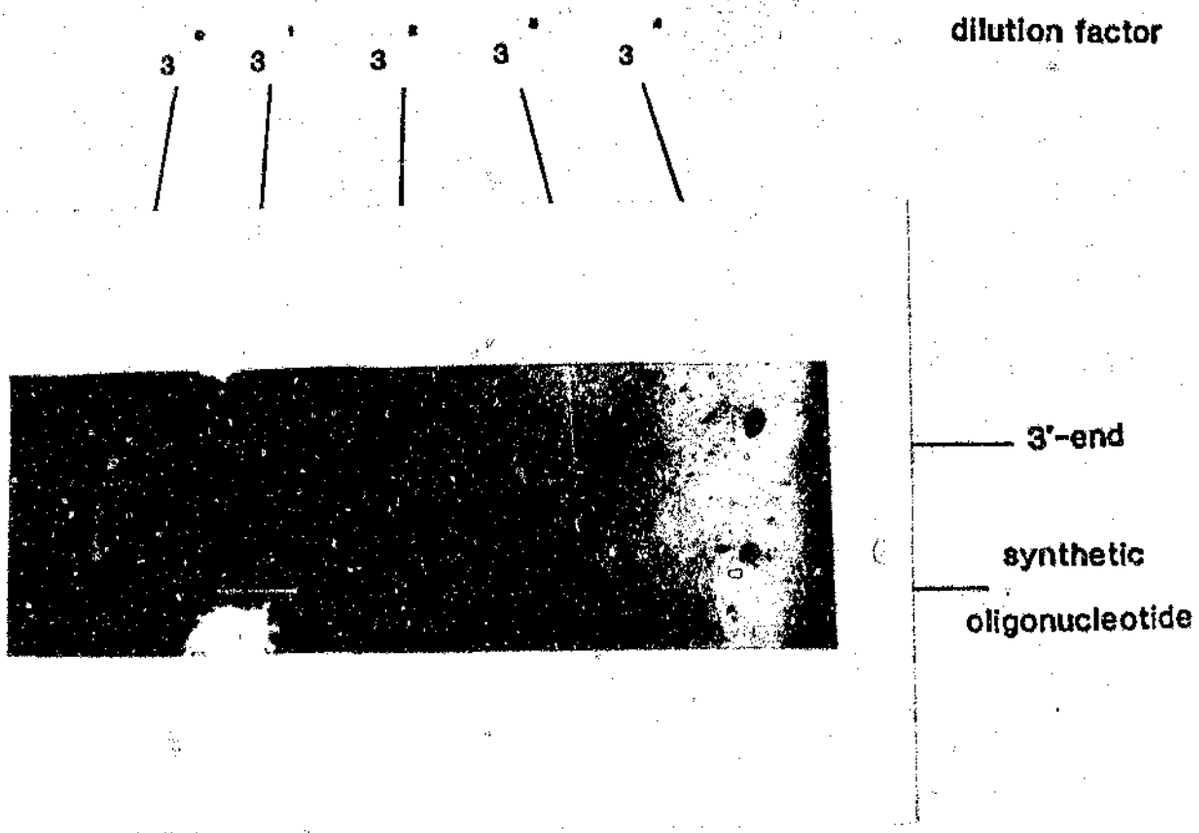


Figure 11: Dot blot of 3'-end labelled synthetic oligonucleotide and 30mer end labelled standard. There appears to be a 3 fold difference in sensitivity between the standard 30mer and the newly labelled oligonucleotide, i.e. 33% labelling efficiency.

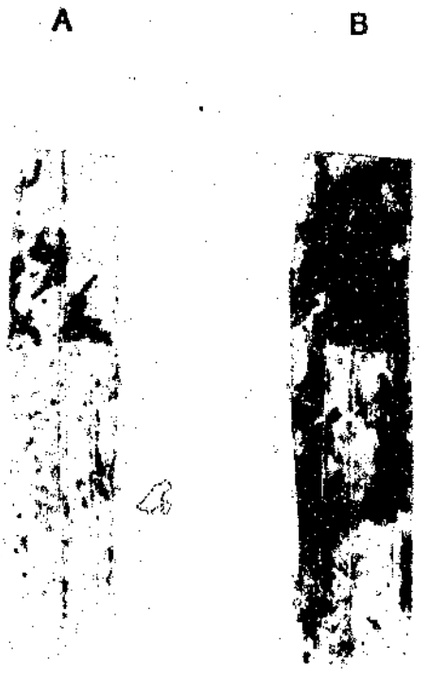


Figure 12: In-gel hybridisation of the standard labelled 30mer to its complementary sequence in the pUC18 plasmid. Two in-gel hybridisation reactions were carried out. The signals were detected separately using 0,1 M NaCl in the detection buffers as originally outlined by Boehringer (gel A) and using 0,9 M NaCl in a staining agar as suggested by Epplen (1992) and Bierwerth *et al.* (1992) (gel B). The colorimetric substrates NBT and X-phosphate are dissolved in the staining agar (see appendix III).

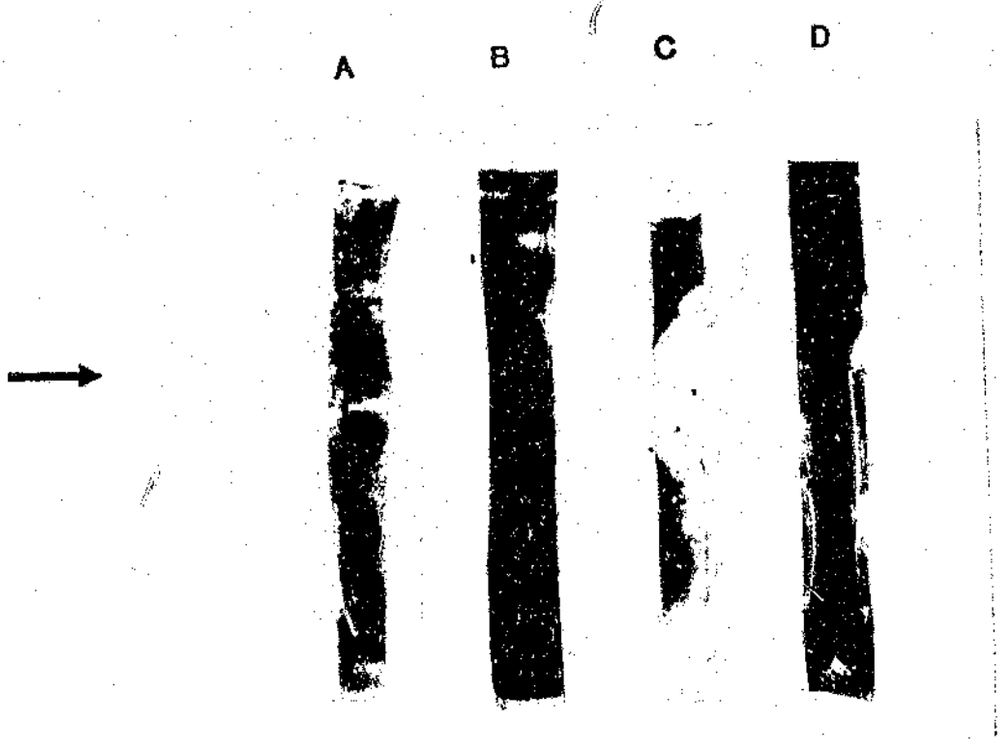


Figure 13: In-gel hybridisation of DIG 3'-end labelled oligonucleotides to restriction digested lithop genomic DNA. Four SQR oligonucleotide sequences were hybridised to EcoR1 digested lithop DNA, immobilised in a mini agarose gel, to see which might generate a scoreable RFLP pattern. $(TCC)_5$, $(GATA)_4$, $(GACA)_4$ and $(GTG)_5$ were tried (gels A, B, C and D respectively). Some high molecular weight bands can be seen with $(TCC)_5$ - see arrow.

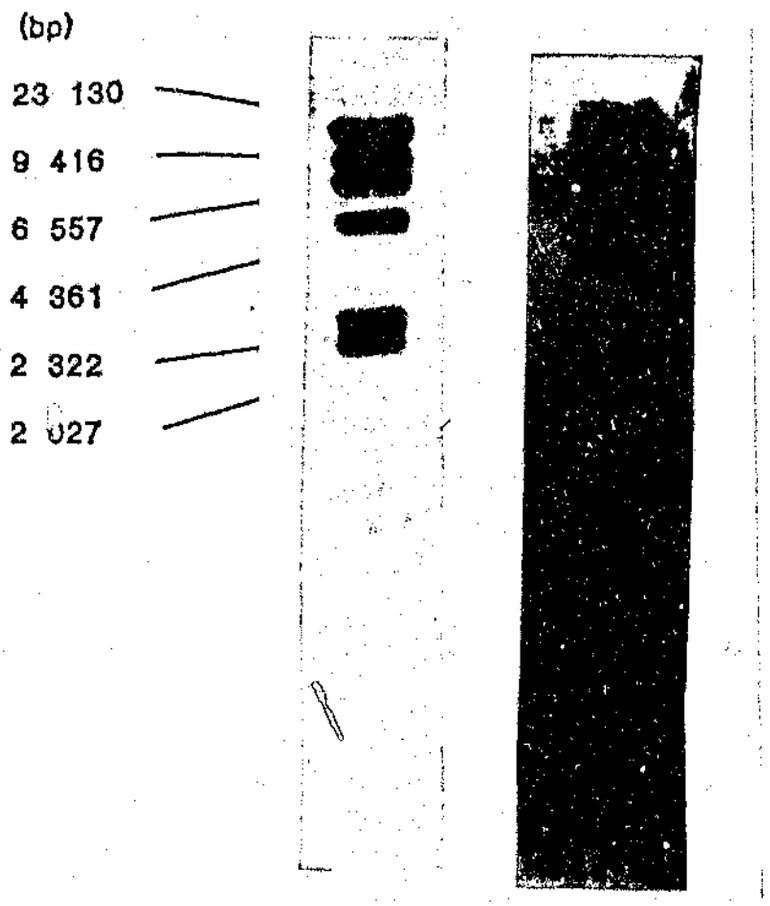


Figure 14: Hybridisation of DIG-$(TCC)_5$ to EcoR1 digested lithop DNA. After electrophoresis on a 14 cm long agarose gel the digested lithop DNA was Southern blotted and cross-linked onto a nylon membrane (Southern, 1975). Hybridisation with 3'-end labelled DIG-$(TCC)_5$ was carried out at 35° C overnight and the colour was developed using a staining agar. Lambda molecular weight markers (marker II Boehringer Mannheim) were random primed labelled and self hybridised after Southern blotting.

3.4 Discussion

3.4.1 Quantifying and Purifying Oligonucleotides

It has been published that 20 µg of a single stranded DNA oligonucleotide is equivalent to 1 A_{260} unit (Maniatis *et al.*, 1982). This is incorrect because the actual number of micrograms of oligonucleotide per A_{260} unit should be calculated using the molar absorption coefficients of the DNA bases, at 260 nm, and from the relative proportion of each base in the oligonucleotide. The amount of oligonucleotide has to be precisely calculated before one can proceed with the labelling and hybridisation reactions because they require specified quantities of oligonucleotide.

Residual blocking groups, which absorb at 260 nm, will interfere in A_{260} estimates of the amount of oligonucleotide. Removal of these "contaminants" will improve the determination of the amount of oligonucleotide. Oligonucleotides can be fractionated into the void volume by gel filtration chromatography using a small 1 ml Sephadex G15 column. This is actually a desalting or more specifically a group separation procedure. The purity of the oligonucleotide can be calculated from the resultant elution profile by measuring the number of A_{260} units eluted in the void volume fraction and expressing it as a fraction of the total A_{260} units loaded onto the column. The oligonucleotides were found to be 95% pure (see figure 7) and could be used in labelling and hybridisation reactions without any purification.

3.4.2 Labelling Efficiency

There is a marked difference between the detection of the DIG RP and the end labelled 30mer probes (see figures 8 and 9). This was believed to be the result of the low efficiency with which small oligonucleotides were bound to the nylon membranes, even after cross-linking with UV light. In an effort to reduce the effective pore size of the membranes and so facilitate the retention of oligonucleotides, the membranes were pretreated by soaking in gelatin, herring sperm DNA or bovine serum albumin (BSA) before the oligonucleotides were applied. Positively charged BM nylon membrane was also tried, but no increase in the

retention of the 30mer could be seen when comparisons were made with the untreated, ordinary nylon membrane (see figure 10). The reason for the decreased retention of the 30mer is now known to be the result of the low binding affinity of short oligonucleotides for nylon membranes (Boehringer Mannheim - personal communications).

Given this anomalous behaviour I decided that it was impossible to calculate the labelling efficiency of the short, synthetic 15 and 16mer SQR oligonucleotides by dot blot comparison with the standard labelled 30mer because they will have an even lower affinity for nylon membranes than the standard end labelled 30mer (see figure 11)

3.4.3 Hybridisation

Short duplexes, like the oligonucleotide-DNA hybrids, are known to be relatively unstable even at 15 degrees below their melting temperature. The stringency washes are carried out with increasing temperature, but at a high salt concentration. The immunological colour detection procedure, prescribed by Boehringer Mannheim, uses a relatively low salt concentration in the detection buffers (0,1 M NaCl). This has the effect of introducing a low salt, high stringency, wash during colour development after the required stringency washes have been done. Given the inherently low stability of the oligonucleotide-DNA hybrids we might expect most (if not all) of these duplexes to denature under these high stringency conditions. Zischler *et al.* (1991) recommend the use of buffers with higher salt concentrations (0,9 M NaCl) during colour development. The colorimetric substrates NBT and X-phosphate, which are insoluble in solutions with salt concentrations above 0,2 M can be prevented from precipitating in 0,9 M NaCl solutions by using a standard agar to develop the coloured signal. The improvement in signal detection, using the higher salt concentration, is evident when the standard 30mer is hybridised to its complementary sequence in the pUC18 plasmid which was immobilised in a mini agarose gel (see figure 12).

Due to the fact that there are no lithop DNA sequence data available, four synthetic oligonucleotides (like those used by Weising *et al.*, 1991) were synthesised to be used as potential SQR probes. Each of these oligonucleotides was 3'-end DIG labelled and

hybridised to EcoR1 digested lithop DNA, immobilised in a mini-agarose gel (see figure 13). Out of the four sequences used, viz. $(TCC)_5$, $(GACA)_4$, $(GATA)_4$ and $(GTG)_5$, only $(TCC)_5$ appeared to generate a distinctive pattern after the stringency washes in 6xSSC at room temperature and at the hybridisation temperature.

The ability to obtain a distinctive banding pattern depends not only on the resolving power of the agarose gel, but also on its length. Many DNA polymorphisms may be of similar length (Lander, 1989) and so the minigel apparatus is not suitable for analytical profiling. Jeffreys *et al.* (1985) used a 14 cm long, 0,6% agarose gel to resolve restriction site polymorphisms. It was hoped that the pattern seen with $(TCC)_5$ could be improved upon if a longer gel was used. The EcoR1 digested lithop DNA was run on a 14 cm agarose gel and the hybridisation with DIG-$(TCC)_5$ was repeated with a Southern blot of this gel. However, only a high molecular weight smear could be seen after the colour was developed (see figure 14).

3.5 Conclusion

Commercially available DNA synthesisers will produce relatively pure oligonucleotides, after deprotection procedures. Gel filtration chromatography, using Sephadex G15, can be used to purify synthesised oligonucleotides from residual "blocking" groups which interfere marginally in A_{260} estimations. The number of micrograms of oligonucleotide represented by 1 A_{260} unit is best calculated using the relative proportion of each deoxyribonucleotide base and its corresponding molar absorption coefficient at 260 nm.

Oligonucleotides are usually end labelled because this does not introduce excessive steric disruptions during the formation of oligonucleotide:DNA hybrids. The efficiency of the 3'-end DIG labelling of the SQR probes was found to be at least 33% by dot blot comparison with a standard labelled 30mer. However, this is by no means conclusive because the 15 and 16mers, that were labelled, have a lower binding affinity for nylon membranes than the 30mer.

Four different SQR oligonucleotides (used with some success by Weising *et al.* (1991) on different plant species) were hybridised to EcoR1 restriction digested lithop DNA, but none generated a distinguishable profile pattern. Nevertheless, minigels can be used as a preliminary in-gel hybridisation strategy for testing which combinations of probe and digested, genomic DNA will generate informative profile patterns. This is a fast and informative strategy, especially when several probes have to be tested, because each in-gel digest can be stripped of its probe and rehybridised with several different probes. Once a suitable combination has been chosen, then the experimenter can use a longer gel for analytical purposes.

CHAPTER 4

4 DNA AMPLIFICATION FINGERPRINTING (DAF)

4.1 Introduction

The polymerase chain reaction (PCR - for which its inventor, Kary Mullis, recently won the Nobel prize for Chemistry) has become a powerful tool in molecular biology because it enables one to amplify thousands of copies of a specified DNA sequence (usually 100 to 2000 bases long) using a thermostable DNA polymerase. The amplification is directed by oligonucleotide primers which are complementary to the regions flanking the sequence that is to be amplified. These oligonucleotides must bind on opposite strands with their 3'-ends facing each other, i.e. there is a defined geometry or orientation that must be satisfied for amplification.

The PCR is a cyclical procedure made up of three distinct stages. Each stage is characterised by the temperature at which it occurs (Gibbs, 1990). The first is a high temperature (96° C) stage which results in the melting of the double stranded template DNA. The second is a low temperature (usually 45 to 60° C) stage which allows the primers to anneal to complementary sites on the single stranded template. Finally during the last stage the enzyme anchors to the primer-template duplexes and extends each chain in one direction only, viz. 5' to 3'. This final stage is carried out at 72° C which is the optimal temperature for the thermostable enzyme. The overall procedure is usually repeated 30 to 50 times using a programmable thermocycler. Once a sequence has been copied from the template, it will serve as an intermediate template during the next round of replication, i.e. once a DNA sequence has been copied, the copy will be copied and so on (Newbury and Ford-Lloyd, 1993). In this way the sequence defined by the two primers will accrue exponentially. This technique has also helped to shape new approaches towards DNA profiling.

Rassmann *et al.* (1991) have described how PCR can be used to detect length polymorphisms by exploiting the hypervariability within the short, simple sequence stretches and SQRs. Oligonucleotide primers, which are complementary to the DNA flanking simple sequence stretches, are used to direct the amplification of these regions from the genomic DNA template. The amplified products are then resolved on a suitable polyacrylamide gel. Alleles that have different copy numbers of the basic repeat sequence will have different mobilities on the gel. If the oligonucleotide primers flank regions of different length then one can test several loci simultaneously on the same gel, thus exploiting the advantages of multilocus probing (outlined in chapter 1). The length polymorphisms detected in this manner are referred to as amplified fragment length polymorphisms (AFLPs). AFLPs are only useful for defined polymorphic loci for which the sequences of the conserved flanking regions are known.

Another DNA profiling technique that makes use of PCR technology is DNA amplification fingerprinting (DAF) which is also known as the RAPD (random amplified polymorphic DNA) assay. This technique was first outlined by Williams *et al.* (1990) and Welsh *et al.* (1990) and later by Caetano-Anolles *et al.* (1991). The DAF strategy involves the enzymic amplification of DNA directed by a single arbitrary sequence oligonucleotide primer (Ba sam *et al.*, 1992). These oligonucleotides prime DNA synthesis from genomic sites to which they are fortuitously matched, or almost matched. The reaction produces a spectrum of amplified DNA products, some of which are polymorphic, that are characteristic of the template DNA. DAF products are resolved by polyacrylamide gel electrophoresis and visualised by silver staining.

Unlike PCR, DAF relies on non-stringent reaction conditions for the amplification of arbitrary target sites. Thus DAF is conceptually and mechanistically distinct from the PCR procedure. It represents a spec'al in vitro reaction in which the oligonucleotide primers are so short that they are thought to approach the functional limits for priming DNA amplification (Caetano-Anolles *et al.*, 1992). DAF is an attractive alternative for generating DNA profiles because it does not require any prior sequence knowledge of the organism. Fur-

thermore, because the method is based on PCR, only small quantities of genomic DNA are required.

4.1.1 Molecular Mechanism of DAF

The molecular nature of the DNA polymorphism detected in the DAF assay is thought to be due to primer binding and orientation. As in the PCR, a DNA sequence can only be amplified if two primers hybridise on opposite strands with their 3'-ends facing each other. Unlike PCR, where the primers are designed to meet this requirement, the single primer used in DAF can have one of four possible orientations. However only one will support amplification of its target sequence, viz. a single primer will support DNA amplification from a genomic template or' if binding sites on opposite strands of the template exist within a distance that can be traversed by the DNA polymerase (Baird *et al.*, 1992). I have called this the critical distance (C_d).

Amplification is dependent on three factors:

1. the formation of stable oligonucleotide-template hybrids to which the polymerase can anchor and extend,

2. binding and orientation of the primers, on opposite strands, within the critical distance (C_d) and

3. the binding stability of the hybridised oligonucleotide.

If two primer binding sites have 100% homology, but they are in the incorrect orientation or they are further apart than C_d, then the region that they border will not be exponentially amplified. If two potential binding sites are in the correct orientation and within the C_d, but they form unstable hybrids then the sequence that they border will not be amplified. The primers are extended from their 3'-ends and, not surprisingly, mismatches in this region are believed to be more disruptive to primer extension than those at the 5'-end.

At sufficiently low temperatures (low stringency), primers will anneal to many sequences with various mismatches some of which will be stable. DAF primers usually have a high GC content (at least 60%) that will enhance the formation of stable mismatched hybrids. Furthermore, guanine forms stable mismatches with adenine and thymine (see Introduction, chapter 3). If two stable annealing sites are within the critical distance to support amplification and on opposite strands then the sequence between these primers can be amplified (Welsh *et al.*, 1990). The extent to which sequences will be amplified will depend on the efficiency of priming at each pair of primer annealing sites and the efficiency of primer extension. DAF profiles are composed of bands of varying intensity that may result from products being amplified to varying extents. Those products that are weakly amplified could result from the weaker mismatch annealing of the primer to one or both of the target sites defining a product. More intense products could arise because they are present in a higher copy number. During the initial rounds of amplification, template conformations which cause steric hindrance of annealing and/or primer extension may also affect product yield.

4.1.2 Primer Binding during DAF

The hybridisation patterns detected on Southern blots or in gels depend on the specific binding of a given probe to its complementary sequence/s in the restriction digested target DNA. In much the same way DAF patterns depend on the stable binding of a primer to a site on the template DNA to direct amplification. Several factors affect the efficiency with which a particular DNA target is amplified, viz. the number and location of primer template mismatches, their stability at various annealing and extension temperatures and the efficiency with which the polymerase can recognise and extend mismatched primer-template duplexes. The length of the primer, the template's conformation and the reaction conditions can affect the kinetics of duplex formation (between primer and template) and thus influence the efficacy with which a particular genomic site is amplified. Not all targeted sites are equally well amplified, some "amplicons" are amplified preferentially. This implies that while primer-template sequence homology defines a set of possible amplification sites there are other factors that act either during primer annealing, enzyme anchoring or elsewhere in the amplification process that can restrict amplification. Once extension from a primer occurs, the

resultant products (and their complements) have termini that are defined by the sequence of the primer, i.e. their ends are perfectly complementary.

DNA amplification is modulated at two levels. In the first, the selection of target sequences for amplification is determined by the sequence of the primer and the reaction conditions. During the first few cycles the primer "screens" for possible DNA targets and anneals to many of them. In this early stage annealing events will include certain stable primer-template mismatches. Primer annealing is followed by enzyme anchoring and primer extension. Therefore, during the first few rounds a subset of many possible target sites are selected to produce a defined population of amplification products. These products are efficiently amplified because their terminal sequences are perfectly complementary to the sequence of the primer. Internal sequences of some target regions and their products, which have a tendency to form secondary structures, may hinder the availability of their annealing sites and lower the efficiency with which they are amplified (Caetano-Anolles *et al.*, 1992).

The differential extent to which an amplification product is further amplified constitutes the second level of modulation. Amplification products initiated by a single primer share the particular characteristic of having terminal symmetry at least as long as the primer itself. These products can form secondary hairpin loop structures. In certain fragments symmetry may extend into internal regions thereby stabilising these hairpin structures. For efficient amplification of such products to occur the primer must displace these hairpin loop complexes long enough to allow the DNA polymerase to anchor and stabilise the duplex for strand extension. The extent of hairpin loop structure interference will be variable for each fragment, allowing some fragments to be preferentially amplified.

4.1.3 Estimating the Number of Amplification Products

Calculating the expected number of amplification products is complex. Theoretically, the number of perfect hybrids (N) of an oligonucleotide of length (L) to a single stranded template of complexity (C) having a random sequence distribution will be given by the equation (Bassam *et al.*, 1992):

$$N = (0{,}25)^{L} \times 2C$$

However, sequence distribution is not random and only a fraction of the expected sites will occur within the range suitable for amplification, i.e. within Cd. Of these sites only about a quarter will anneal with primers on opposite strands with their 3'-ends facing each other. These considerations reduce the number of bonafide products dramatically. Amplification products can arise from mismatched annealing events. Allowing single base mismatches at all positions in a heptamer would produce 21 times more annealing sites, but this number is greatly reduced because mismatches are only tolerated at the 5'-end of the duplex (Bassam *et al.*, 1992).

It is crucial that all potential sites are efficiently amplified during the first few rounds of amplification because any regions that are not will not be stoichiometrically represented in later rounds.

4.1.4 Silver Staining DNA in Acrylamide Gels

Merril *et al.* (1981) were one of the first to establish a protocol for the staining of proteins with silver. Silver was found to bind to proteins under acidic conditions. Image development requires a change in pH which causes the formation of insoluble silver salts (Blum *et al*, 1987). Early protocols using silver nitrate often used oxidising agents in a pretreatment to prevent adventitious silver ion reduction in the gel (Heukeshoven and Dernick, 1985). More recent protocols use reducing agents (like thiosulfate) which are believed to enhance the formation of the silver image (Blum *et al.*, 1987 and Rabilloud, 1992). Beidler *et al.* (1982) have described a procedure for silver staining nucleic acids which is based on the original protein stain developed by Merril. Nucleic acids are detected photochemically when silver ions bind to the bases and are then selectively reduced chemically e.g. with formaldehyde.

Although the mechanism involved in silver staining is still unclear, various authors have tried to optimise this technique for staining DNA. Bassam *et al.* (1991) have described a fast and sensitive procedure for silver staining DNA in polyacrylamide gels. This procedure uses formaldehyde to reduce silver ions to metallic silver under alkaline conditions without the

need for any pretreatment. Thiosulfate is used during image development, as recommended by Blum *et al.* (1987), to prevent the formation of insoluble silver salts in the developer. This protocol was used successfully by Bassam *et al.* (1991, to analyse the DNA banding patterns generated during DAF on polyacrylamide gels containing 7 M urea.

4.1.5 Nature of DAF Polymorphisms

On the face of it DAF is an unique and powerful technique for studying DNA polymorphisms which are detected as mobility differences in polyacrylamide gels. This technique has been used to profile a wide array of organisms to confirm the phylogenetic relationships between species e.g. mosquitoes (Ballinger-Crabtree *et al.*, 1992), wheat aphids (Puterka *et al.*, 1993), celery cultivars (Yang and Quiros, 1993) and schistosome parasites (Neto *et al.*, 1993). DAF has also been used in related fields such as pedigree assessment in cereal crops (Dweikat *et al.*, 1993) and as genetic markers for disease resistance loci (Barua *et al.*, 1993). Thacker (1994) used DAF to distinguish between mouse and hamster cell lines. Differences between species are calculated on the basis of band sharing frequencies, much like in the case of RFLPs.

Unlike RFLPs, DAF products do not represent specific sequences or alleles, they are merely a collection of fortuitously amplified products of different sizes. Furthermore, DAF products are 'ominant "alleles" and so it is impossible to determine whether an individual is homo or heterozygous for that "allele" (Williams *et al.*, 1990). What then is the nature of the polymorphisms detected in this procedure?

The spectrum of DAF products are the result of the extent to which a primer will rec nise and bind to potential sites on the template DNA. A product will only be amplified if a stable primer-template duplex is formed in an orientation that will support amplification. DNA polymorphisms that affect either the stability of the duplexes (e.g. base change mutations) or their orientation (e.g. insertion or deletion mutations) will change the spectrum of products which are amplified. This will be reflected by differences in the electrophoretic mi-

gration pattern, viz. some bar ay disappear, "new" bands may appear or there may be a change in the efficiency with which a product is amplified and detected.

Given the random nature of this procedure it is possible that when comparing two species distinct products may have the same or similar mobilities quite by chance. Although the probability of this seems rather low, no statistical evidence has been published to suggest that it is. Kaemmer *et al.* (1992) employed a novel approach in their work on differentiating banana cultivars. They used D. profiles directed by an arbitrary sequence 10mer to verify the identity of cultivars that had been established from RFLP patterns. Genetic identity was scored on the basis of band sharing frequencies. Furthermore, they concluded that the two techniques used in tandem complemented each other because DAF could be used to score smaller DNA fragments and RFLPs for larger ones.

I have independently been told by Dr.'s S. Grant (University of the Witwatersrand) and E. Harley (University of Cape Town) that DAF data would not be statistically significant, on its own, to ascertain the phylogenetic relatedness of species of lithops. Therefore, I planned to follow a similar strategy to that outlined by Kaemmer *et al.* (1992) by using SQR profiling to confirm the phylogenetic relationships of lithops established by DAF. However, the fact that none of the SQR sequences described by Weising *et al.* (1991) had given interpretable patterns meant that this strategy would have to be altered.

A DAF product, purified from an acrylamide gel, can be labelled and used as a probe for Southern blots (Weaver *et al.*, 1994). In an adaptation of this technique a DAF product can be used to probe a restriction digest of lithop DNA in a RFLP type analysis. Such a probe would consist of two short regions complementary to the DAF primer, which flank a DNA sequence that has been copied from the lithop template. The lithop DNA can be sequenced and used to probe restriction digests of other lithop species. Some of these alleles may prove to be polymorphic and therefore useful for profiling lithops. To pursue this type of analysis a reproducible DAF procedure has to be developed. This means establishing a standard set of DAF reaction conditions that define the amounts of template, primer and $MgCl_2$. DAF profiles must be reproducible with different DNA preparations of the same species. DAF

profiles produced on separate PCR machines in different laboratories should be consistent so that comparisons can be made between laboratories.

Caetano-Anolles *et al.* (1991) incorporated 7 M urea in their polyacrylamide gels to analyse their DAF profiles. This initially suggested to me that their DAF products were analysed as single and not double stranded DNA fragments under denaturing electrophoresis.

4.2 Materials and Methods

4.2.1 Electrophoresis

Continuous polyacrylamide gel electrophoresis was carried out in a vertical gel apparatus, using glass plates 15 cm long and 16,5 cm wide, as described by Maniatis *et al.* (1982). The glass plates were first well cleaned with ethanol. The notched and unnotched plates were then separately handled. The notched plate was treated with Repelsilane (2% w/v dimethyldichlorosilane in 1,1,1 trichloroethane) and the unnotched, backing plate was treated with a Bindsilane solution (1 µl Bindsilane; 5 ml ethanol; 10 µl acetic acid and 150 µl water). After allowing the plates to dry they were wiped of residual silane before being assembled. Thin spacers (0,5 mm thick) were used as recommended by Caetano-Anolles *et al.* (1991). Spacers were placed along the sides of the gel assembly and the system was held in place with masking tape and miniclamps (no spacer was placed along the bottom of the gel).

10 ml of a 5% gel mixture was made up with 1,7 ml 30% acrylamide:bisacrylamide (29:1); 0,2 ml 3% ammonium persulphate; 1 ml TBE buffer (0,09 M Tris bas. 0,09 M boric acid and 2 mM EDTA pH 8,0); 4,2 g of urea and 7,1 ml of water. 10 µl of TEMED was added to the gel mixture and it was poured in between the plates. A comb was inserted at the top of the gel mixture to form the wells and the gel was left to polymerise.

After polymerisation the miniclamps and the masking tape were removed and the gel was fastened into a vertical electrophoresis tank containing TBE running buffer. The comb was removed and the wells were washed out with the running buffer. The samples were first mixed with an equal volume of loading solution (10 M urea and 0,02% xylene cyanol and bromophenol blue) before loading them into the washed wells. Electrophoresis was performed at 200V until the bromophenol blue had moved to within 1 cm of the edge of the gel. Some samples were loaded after being boiled for 5 minutes to ensure that they were denatured, while others were loaded unboiled.

4.2.2 Silver Staining

The procedure used for silver staining was based on that described by Bassam *et al.* (1991). After electrophoresis the gel plates were removed from the tank and the side spacers were removed. The glass plates were carefully separated - the backing plate should retain the gel. The backing plate was placed gel side up into 10% acetic acid for about 5 minutes. Acetic acid was removed by soaking the gel in water for about 30 minutes. The gel was then impregnated with freshly prepared silver stain (0,25 g silver nitrate, 0,4 ml 35% formaldehyde made up to 250 ml with water) for 30 minutes. The gel was washed in water for 1 minute before the stain was developed in developer solution (0,4 ml 35% formaldehyde, 5 µl 0,1 g/ml sodium thiosulfate made up to 250 ml with 30 g/l sodium carbonate solution). The reaction was stopped after 5 to 8 minutes by placing the gel in about 500 ml of water. The gel was left to soak in water overnight before it was allowed to dry.

4.2.3 DAF

The procedure of Williams *et al.* (1990) who had generated DAF profiles of human DNA with primer ACGGTACACT was initially replicated. DAF reaction was carried out in 10 mM Tris-HCl pH 8,3, 50 mM KCl, 2 mM $MgCl_2$, 0,001% gelatin, 0,5% (v/v) Tween 20, 100 µM of each dNTP, 0,2 µM primer, 25 ng genomic DNA and was made up to 24 µl with water. The reaction mixture was overlaid with 30 µl of mineral oil before being placed into the preheated (93° C) thermocycler. After allowing the reaction tubes to heat up at this temperature for 2 minutes, one unit (1 µl) of Taq DNA polymerase enzyme (Promega) was added and the amplification cycle was started. Amplification was performed on a ESU Programmable Temperature Cycler (purchased from UCT) for 45 cycles at 93° C (30 seconds), 30° C (10 seconds) and 72° C (60 seconds).

The heating and cooling rates of the thermocycler were calculated and taken into consideration when determining the "actual" time required to reach each temperature, viz. 60 seconds

at 93° C, 100 seconds at 30° C and 120 seconds at 72° C. The effects of using different DNA preparations of the same species and of varying the amounts of template, primer and $MgCl_2$ were tested with human DNA and AUGGTACACT. The 10mer CCCTCTGCGG that was successfully used by Kaemmer *et al.* (1992) on banana cultivars was also used to analyse lithop species.

4.3 Results

4.3.1 Silver Staining

A molecular weight marker was selected that would cover the expected size range of DAF products, viz. 100 to 2000 base pairs. There are twelve different molecular sized markers which are all uniformly detected with silver when 50 to 200 ng of total marker is analysed (see figure 15).

4.3.2 Electrophoresis

The heat denatured single stranded markers migrate slower than their double stranded counterparts. The undenatured (unboiled) sample does not spontaneously denature in the presence of 7 M urea. It is noticeable that several of the higher molecular weight bands are common to the boiled and unboiled samples and this may be due to reannealing of the heat denatured fragments, i.e. the boiled sample appears to be a mixture of single and double stranded DNA. This would account for the greater number of bands that are detected in the boiled sample and which makes it difficult to assign a molecular size to these bands (see figure 16).

When the amplification reactions with human DNA and ACGGTACACT were denatured before electrophoresis none of the DAF markers conformed to the size range that was reported by Williams *et al.* (1990) to be between 200 to 1000 bp (see figure 17). The size of the boiled DAF samples, which run as single stranded DNA under denaturing electrophoresis, cannot be estimated from the double stranded markers because they have different migration characteristics. The exclusion of urea from the gel facilitates the reannealing of some of the lower molecular weight DAF products which now run as faster moving double stranded fragments. The resultant profile shifts more towards the expected size range (see figure 18).

The human DNA/ACGGTACACT DAF reactions were also analysed without boiling under denaturing conditions (7 M urea) as described by Caetano-Anolles *et al.* (1991). These profiles, although run on acrylamide and not agarose gels, conform to the size range reported by Williams *et al.* (1990). The use of acrylamide gels and silver staining enhances both the resolution of small fragments and their detection. There is an overall pattern homogeneity between each of the human profiles and there are no genuine polymorphic or individual specific products (see figure 19).

4.3.3 DAF of Lithops

Given the arbitrary nature of DAF primers I wanted to see whether any GC-rich SQR oligonucleotides (see figure 13, chapter 3) could direct DAF of lithop DNA. It is usually 10mers that are used to direct DAF and indeed the ACGGTACACT 10mer gave the best profile of the three primers that were tested, even though it has the lowest GC content (see figure 20).

Caetano-Anolles *et al.* (1992) had suggested that an increase in the amount of primer would increase the number of low molecular DAF products because the primer would outcompete shorter, more stable hairpin loop structures for the available primer binding sites. The appearance of low molecular weight bands at higher primer concentrations is in agreement with this model (figure 21). Increasing the amount of $MgCl_2$ reduces the number of higher molecular weight DAF products and leads to a concurrent increase in the number of lower molecular weight products (figure 22). What this shows is that the amount of $MgCl_2$ used during DAF determines the final spectrum of products, although the reason for this is unclear. One should select a concentration at which a good mix of high and low molecular weight products are produced. Once a $MgCl_2$ concentration has been selected it is vital that it is maintained from reaction to reaction.

Human and *Lithop Lesliei* DNA were amplified with ACGGTACACT and CCCTCTGCGG respectively and the resultant profiles were analysed on 5% denaturing acrylamide gels (figure 23). These two distinct organisms have a similar sized product range

and this accentuates the problem with DAF analysis. While the human and lithop profiles can be distinguished, it is difficult to identify DAF products that are uniquely human or lithop because of the number of bands and the extensive overlap of their size domains. Different primers produce different DAF profiles and can thus be used to show different polymorphisms. A DAF product can only be shown to be specific for a species by isolating it and using it to probe a genomic DNA digest of an individual from that species.

Bands were scored on two criteria, viz. their mobility relative to the molecular weight standards and their intensity. Bands detected with high intensity were assigned an arbitrary intensity unit (IU) which is greater than those for less intense bands. Only bands below 1250 bp were scored. Bands from the human and lithop DAF profiles, generated with ACGGTACACT, with similar mobilities can be distinguished by their intensities. There are four prominent bands (see figure 24) which have similar mobilities in both profiles. When the profiles were score with CCCTCTGCGG there were no prominent bands common to both DAF profiles (figure 25). Taking the two primers together and only counting the prominent bands there are four common bands out of a total of 40, i.e. 10% of the major bands are common. This number can be reduced if intensity differences are also considered. As expected there is a very low level of sequence similarity between human and lithop DNA.

There is an overall similarity in the profiles generated in two different laboratories but there are conspicuous variations in the number, distribution and intensity of bands (see figure 26). There are a number of reasons that could account for this, not least of which are variations in the heating and cooling rates of different machines. DAF programs must be worked out according to the actual time spent at each temperature and not the total time taken to reach that temperature (see Materials and Methods). The heating and cooling rates of each PCR machine have to be separately calibrated and cannot be assumed to be the same for any two machines (even if they are the same model).

Analysing the DAF profiles of three lithop species, viz. *L. Lesliei*, *L. Hookeri* and *L. Terricolor*, reveals a nice mix of unique and common bands that enables one to distinguish them as being derived from different DNA templates (see figure 27). This is in stark contrast

to the homogeneity of the patterns derived from individuals from the same species (see figure 19). Different primers recognise different sets of binding sites which produces different sets of DAF markers, i.e. different primers reveal different polymorphisms. Each primer, used separately, will give a different estimate of the degree of sequence similarity between two DNA templates. It is better to use both of the primers to calculate the degree of similarity by measuring the total number of shared bands as a fraction of the total number of bands.

The DAF profiles generated with ACGGTACACT of the *L. Lesliei* and *L. Hookeri* species (which are both located in the western Transvaal) have seven prominent bands in common (see figure 28). *L. Terricolor*, which is found in the south western Cape (near Laingsburg), has at best only two bands in common with both *L. Lesliei* and *L. Hookeri*. This indicates that there can be extensive variation amongst lithop species from distinct locations, although it is odd that there are so few DAF products generated from *L. Terricolor*. The primer CCCTCTGCGG reveals five prominent bands that are common to *L. Lesliei* and *L. Hookeri*. There are four common bands between *L. Lesliei* and *L. Terricolor* and two between *L. Hookeri* and *L. Terricolor* (see figure 29). CCCTCTGCGG which was used to direct the DAF of banana cultivars was more successful than ACGGTACACT in amplifying products for all three lithop DNA templates. Taking both the primers and all three species together the following pattern of band sharing percentages emerges: 32% between *L. Lesliei* and *L. Hookeri*, 18% between *L. Lesliei* and *L. Terricolor* and 12% between *L. Hookeri* and *L. Terricolor*.

The profiles of *L. Lesliei* with ACGGTACACT and CCCTCTGCGG were used as "standard" reactions to measure the reproducibility of DAF between reactions and between gels. The profiles of *L. Lesliei* with both primers were scored from different reactions and from different gels (figures 30 and 31) and were found to be consistent. Twenty five major bands, thirteen with ACGGTACACT and twelve with CCCTCTGCGG, were reproduced with respect to mobility and intensity.

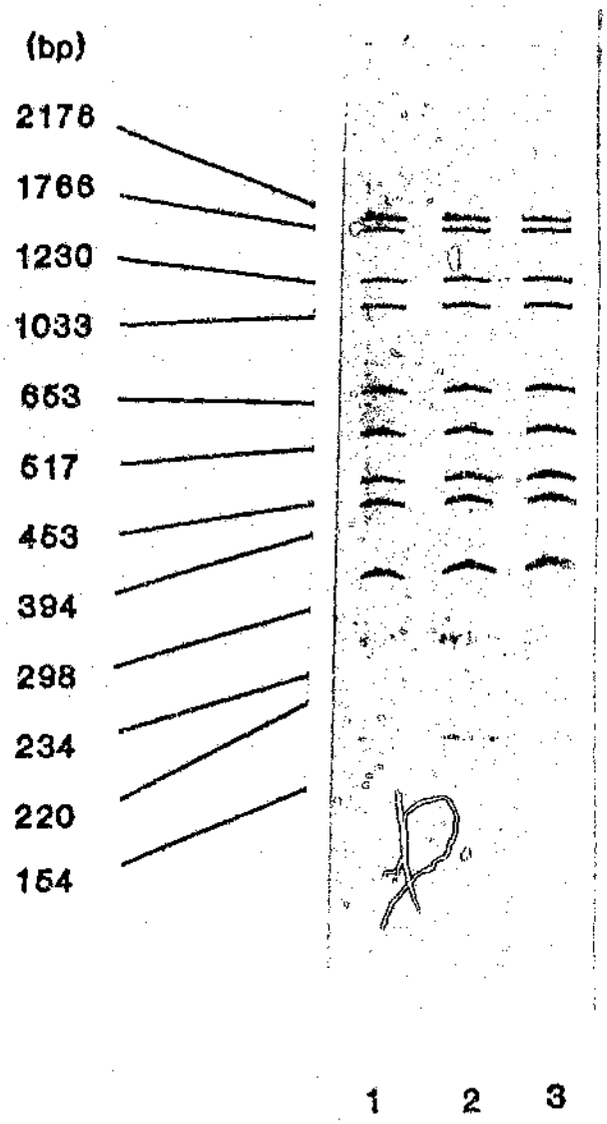


Figure 15: Sensitivity of silver staining of DNA in acrylamide gels. Molecular weight markers (type VI - Boehringer Mannheim) were run at increasing concentration to examine the sensitivity of the silver stain. Lane (1) 50 ng, lane (2) 100 ng and lane (3) 200 ng. The samples were loaded in 5 M urea and run in a 7 M urea gel. Molecular weights are indicated.

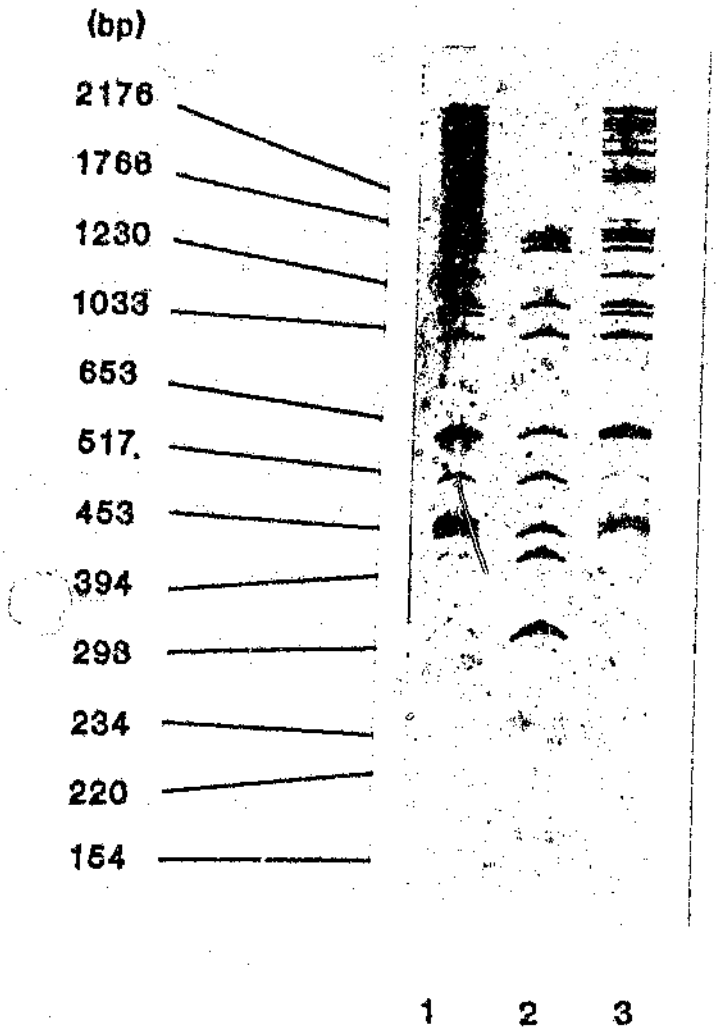


Figure 16: Differential mobility of single and double stranded DNA in denaturing polyacrylamide gels. Two samples were heat denatured by boiling prior to loading and another was not. Lane (1) boiled sample, lane (2) unboiled and lane (3) boiled. Molecular weights of the unboiled marker are indicated.

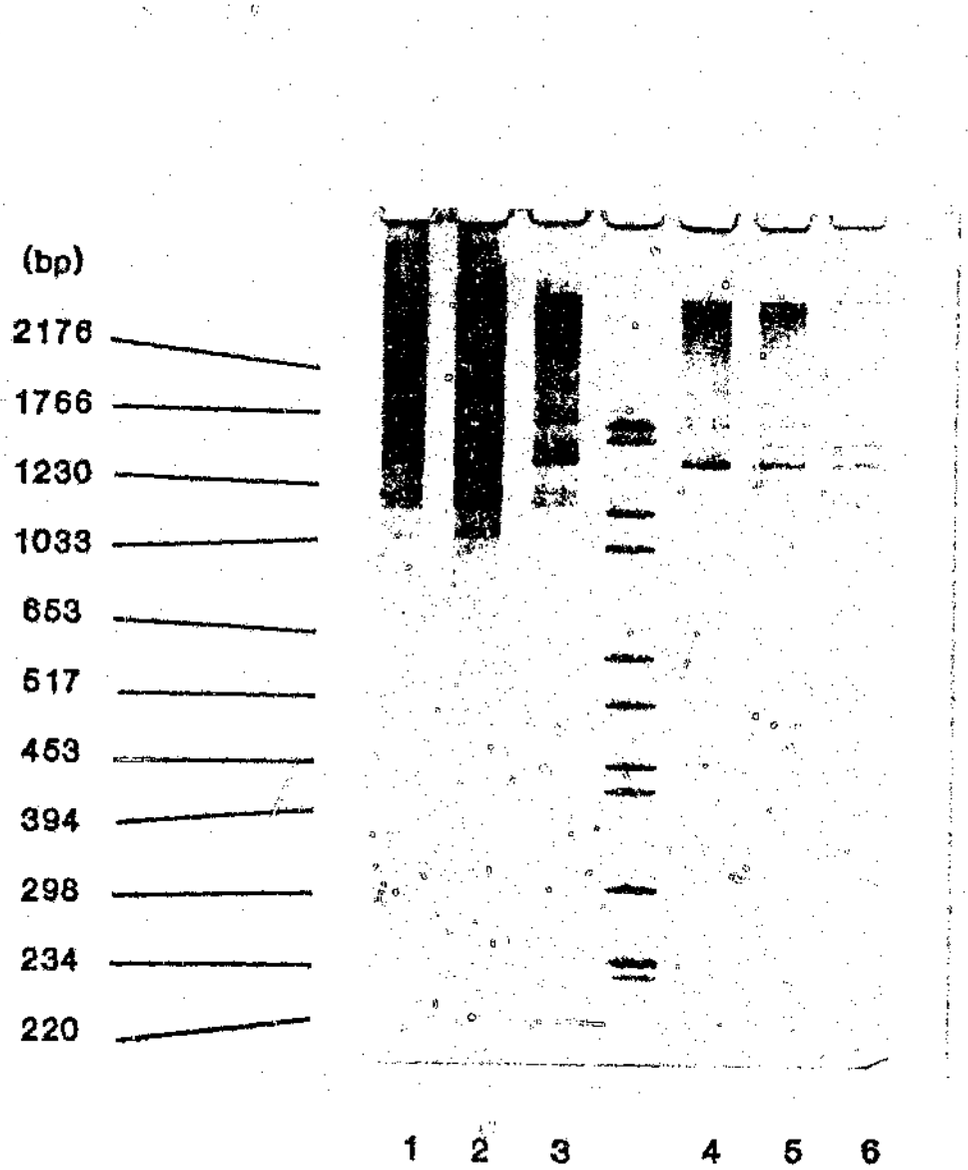


Figure 17: Denaturing electrophoresis of heat denatured DAF products. Three different human DNA preparations were amplified with ACGGTACACT as described by Williams *et al.* (1990). The samples were boiled and loaded in loading solution. Lanes (1-3) 50 ng human DNA and lanes (4-6) 12,5 ng human DNA. Markers were loaded without being boiled. There are no bands detected below the 1033 bp marker.

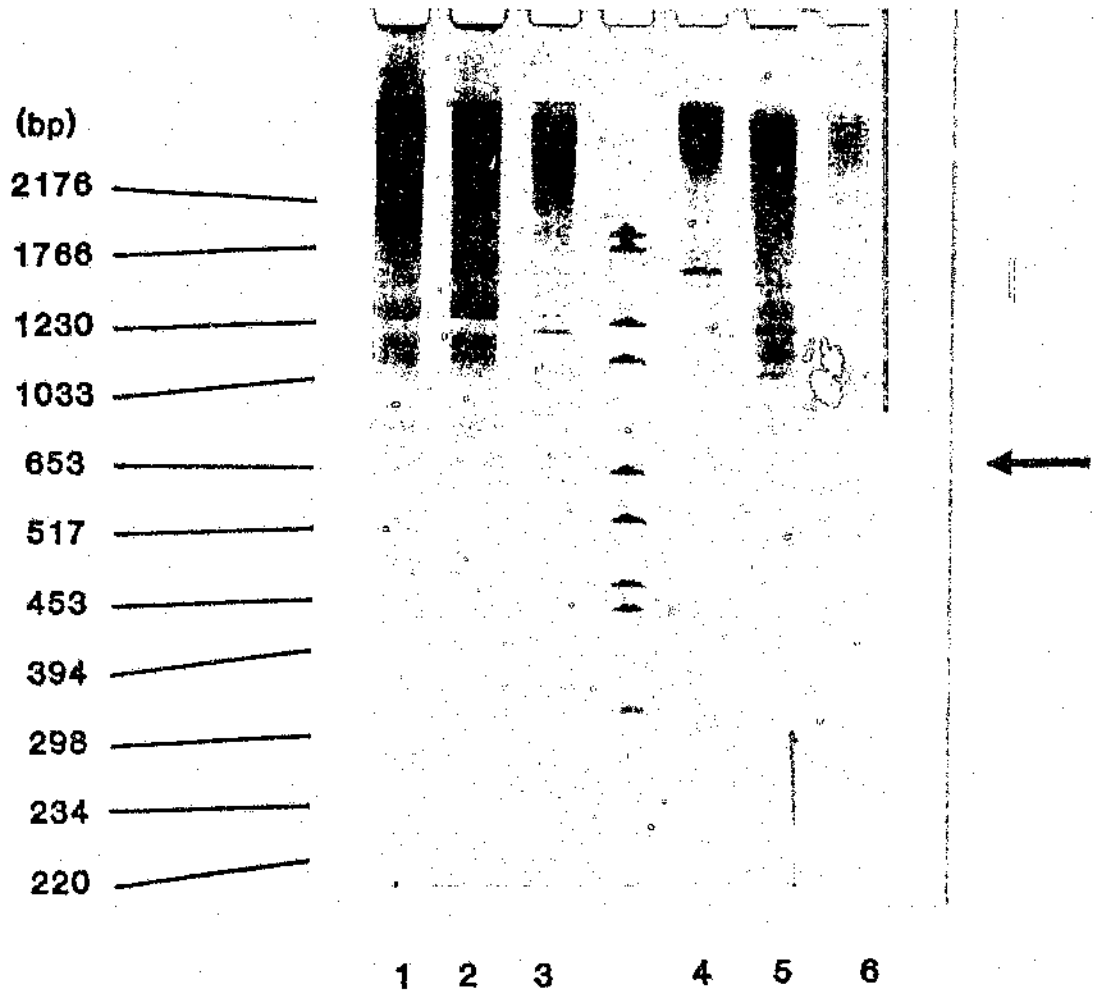


Figure 18: Non-denaturing electrophoresis of heat denatured DAF products. The three human DNA preparations, amplified with ACGGTACACT and mixed with loading solution, were boiled and loaded onto a non-denaturing gel, i.e. no urea. Lanes (1-3) 50 ng human DNA and lanes (4-6) 12,5 ng human DNA. Markers were not boiled. Some bands can be seen at about 653 bp (see arrow).

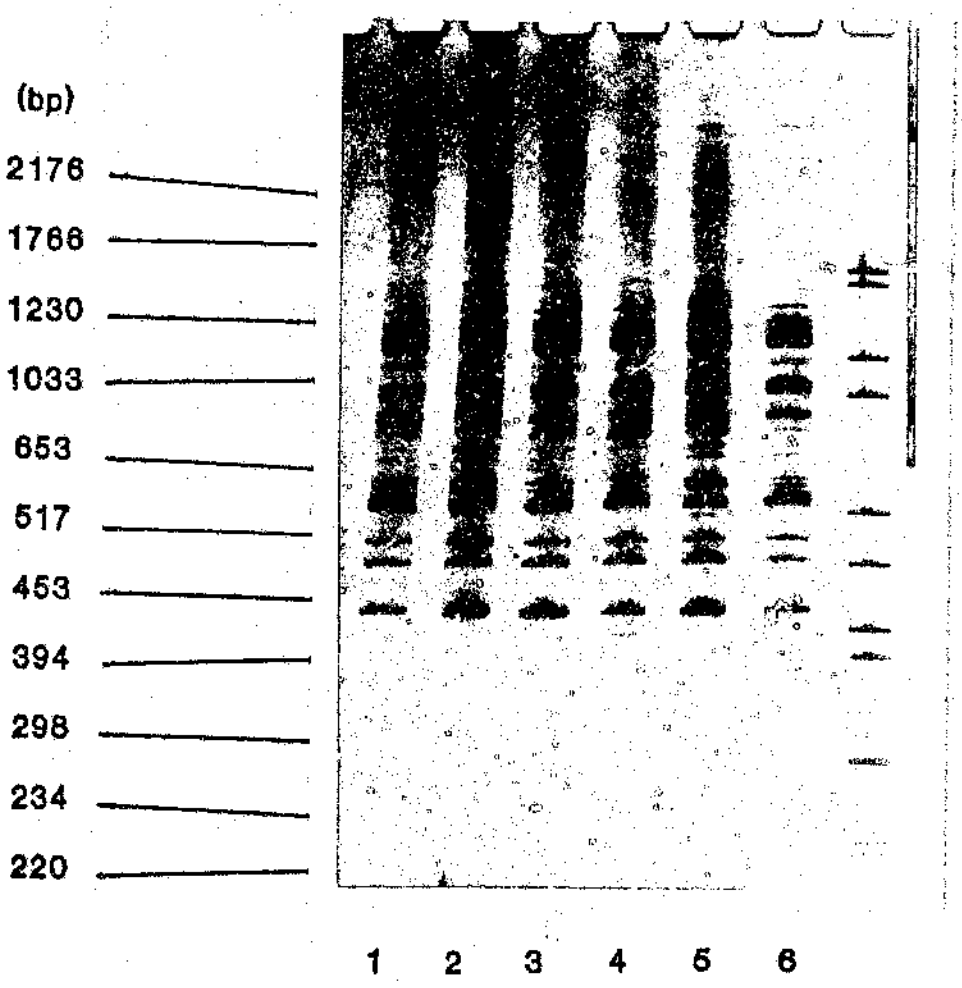


Figure 19: Denaturing electrophoresis of undenatured DAF samples. Three human DNA DAF reactions were analysed on a denaturing gel, but they were not heat denatured prior to loading. The DAF reactions were repeated with a four fold decrease in the amount of template DNA for each of the human samples. Lane (1 to 3) 50 ng template DNA and lane (4 to 6) 12,5 ng of template DNA. The markers were not boiled.

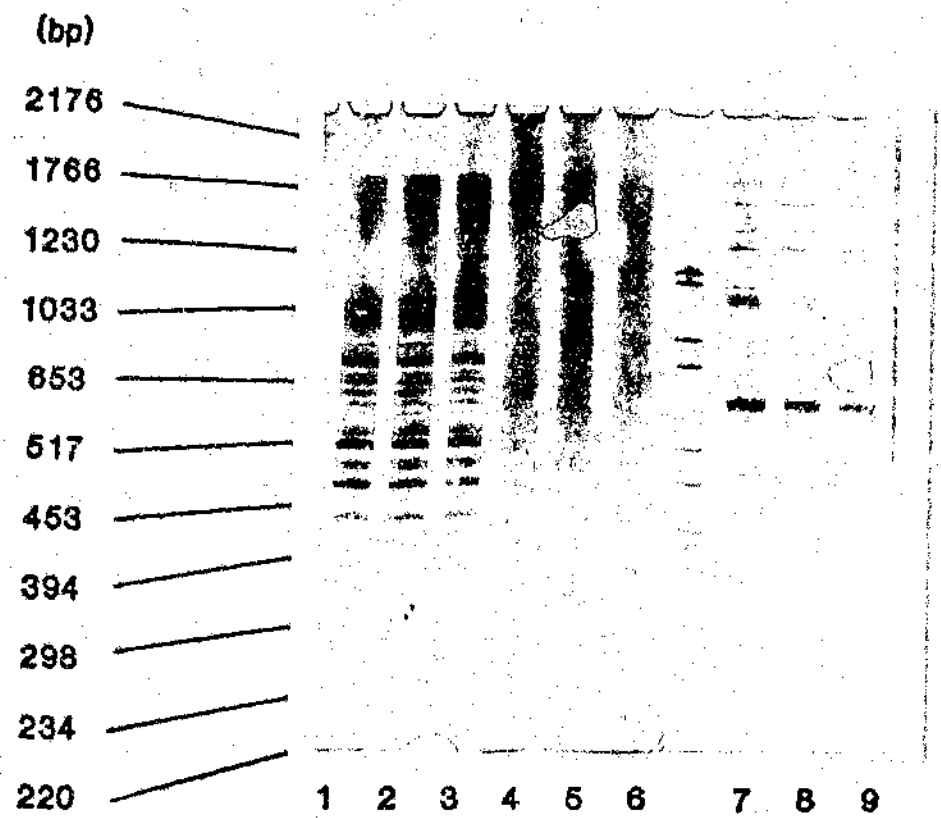


Figure 20: DAF of lithop DNA with different oligonucleotides. The amplification of 25 ng of *L. Lesliei* DNA was directed by 20 ng of ACGGTACACT (lanes 1-3), $(TCC)_5$ (lanes 4-6) and $(GTG)_5$ (lanes 7-9). The profiles were analysed on 7 M urea gels without prior boiling.

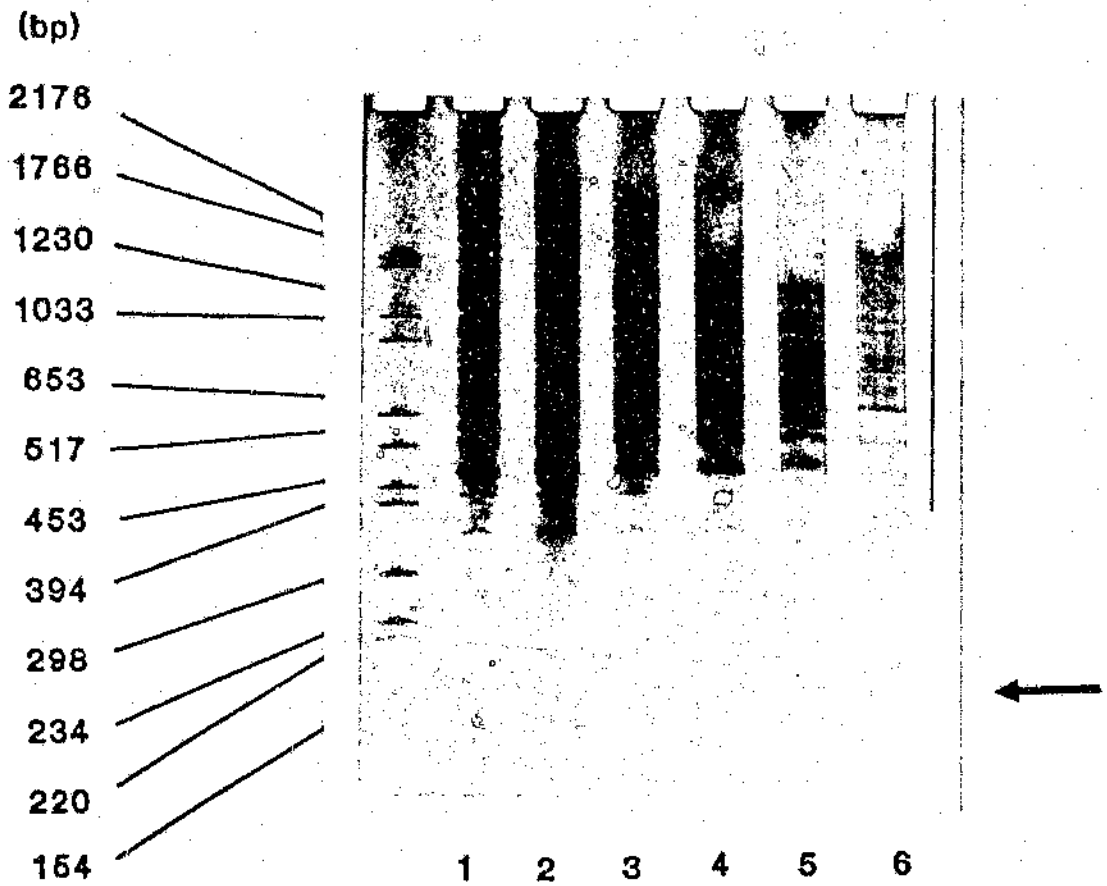


Figure 21: Effect of increasing the amount of primer during DAF. Increasing amounts of primer ACGGTACACT were used to direct the amplification of *L. Lesliei* DNA. Lane (1) 20 ng, lane (2) 50 ng, lane (3) 100 ng, lane (4) 150 ng, lane (5) 250 ng and lane (6) 500 ng. There is an obvious decrease in the number and intensity of the higher molecular weight bands as the amount of primer is increased. A band can be seen at about 154 bp when 500 ng of primer is used (see arrow), unfortunately this band is not very clear in the photograph.

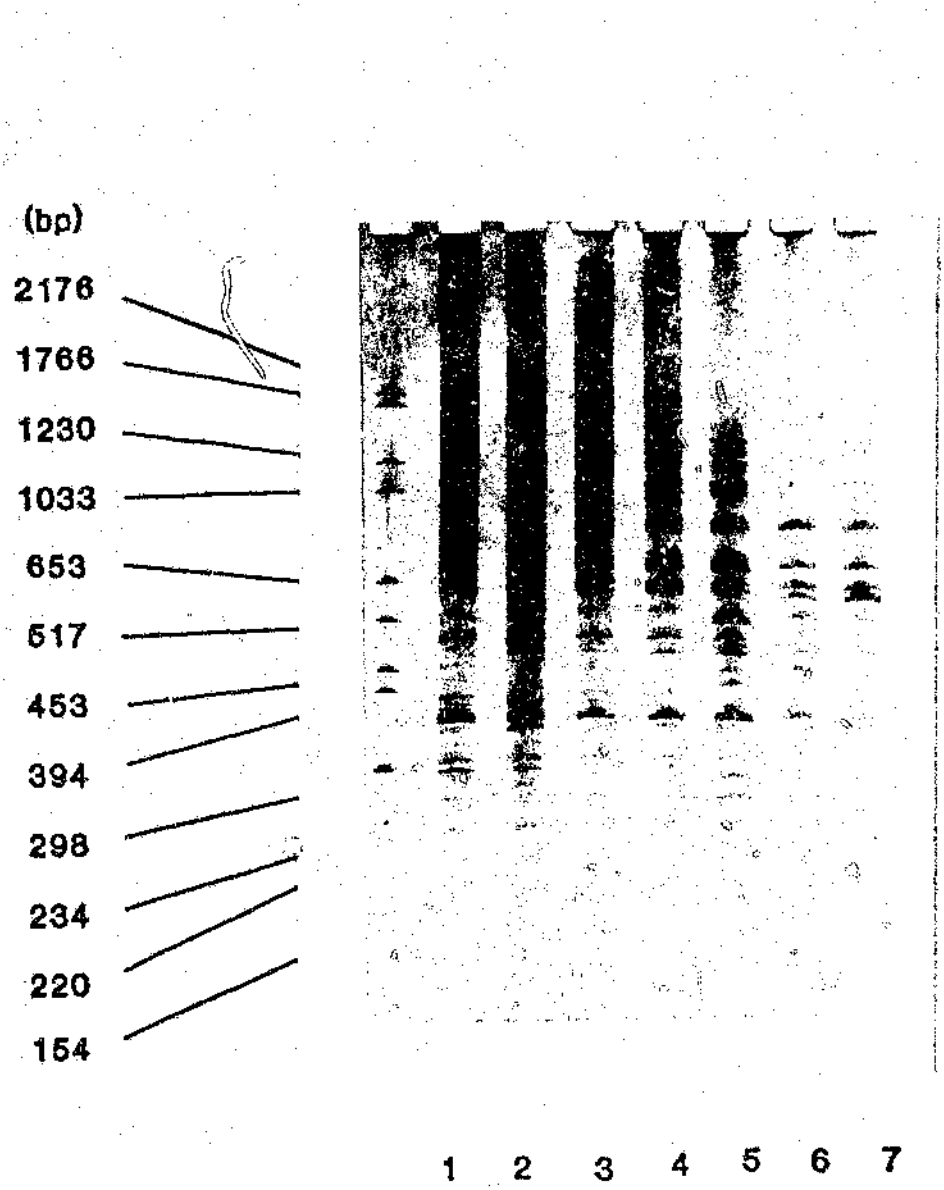


Figure 22: Optimising $MgCl_2$ for the reaction of *L. Lesliei* and CCCTCTGCGG. The amount of $MgCl_2$ was increased from 2mM (lane 1), 3 mM (lane 2), 4 mM (lane 3), 5 mM (lane 4), 6 mM (lane 5), 7 mM (lane 6) to 8 mM (lane 7).

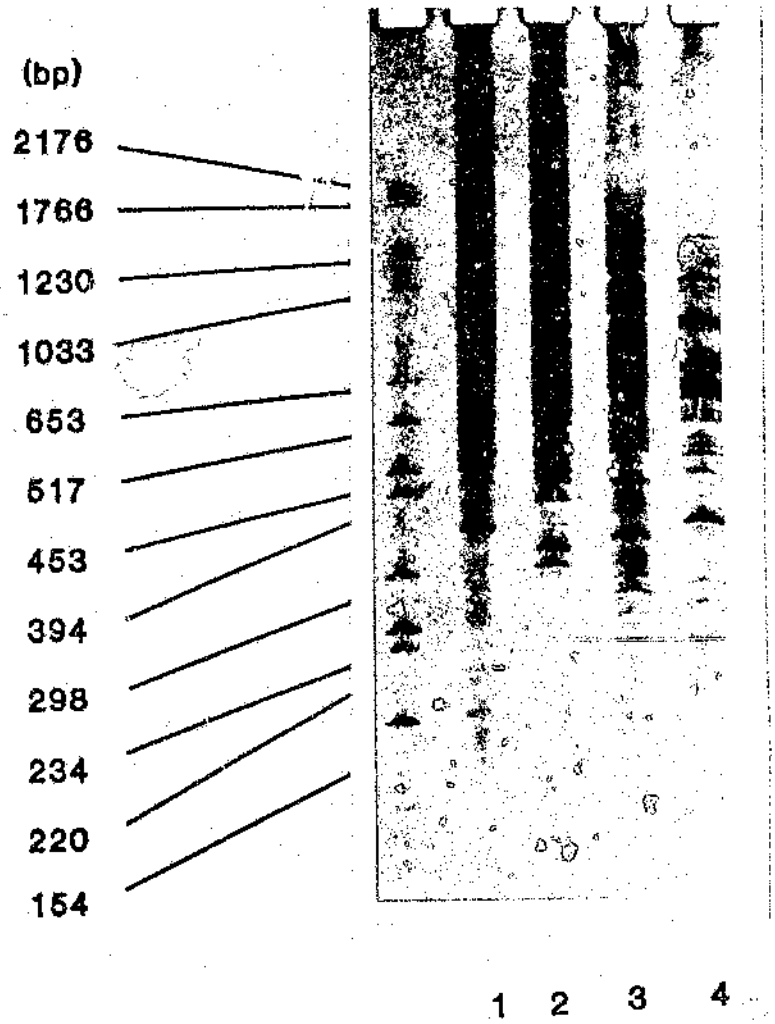


Figure 23: DAF of *L. Lesliei* and human DNA with ACGGTACACT and CCCTCTGCGG respectively. DAF profiles of human and *L. Lesliei* were directed with 500 ng of ACGGTACACT and CCCTCTGCGG respectively, at 5 mM $MgCl_2$. Lane (1) human with ACGGTACACT, lane (2) *L. Lesliei* with ACGGTACACT, lane (3) human with CCCTCTGCGG and lane (4) *L. Lesliei* with CCCTCTGCGG.

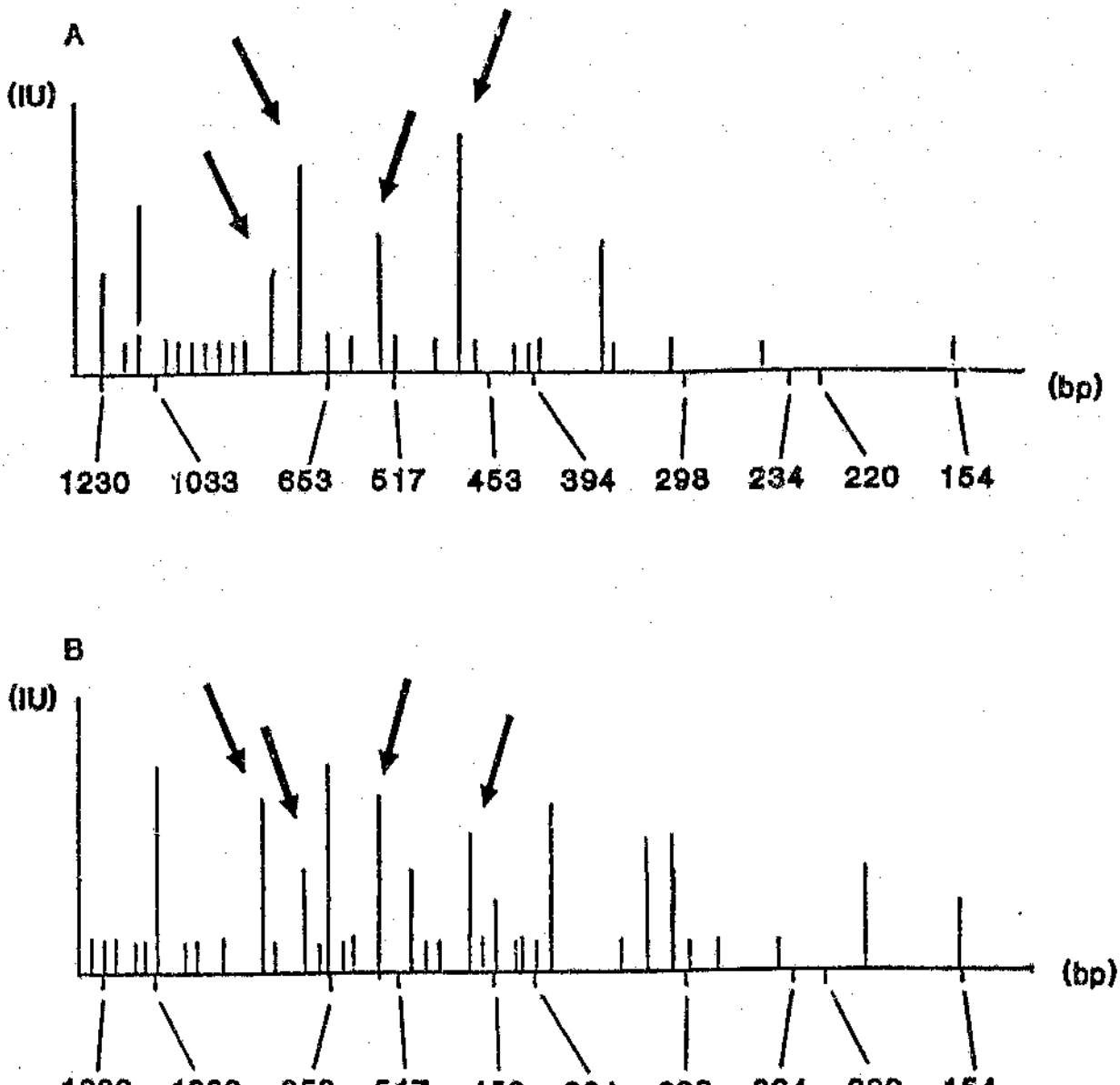


Figure 24: Histogram showing fragment length and band intensity polymorphisms between the human and *L. Lesliei* DAF profiles generated with ACGGTACACT. Migration relative to molecular weight markers is plotted on the horizontal axis and staining intensity (IU - arbitrary intensity units) on the vertical axis. Panel A) human DNA and panel B) *L. Lesliei* DNA. Four prominent bands with similar mobilities can be seen (see arrows).

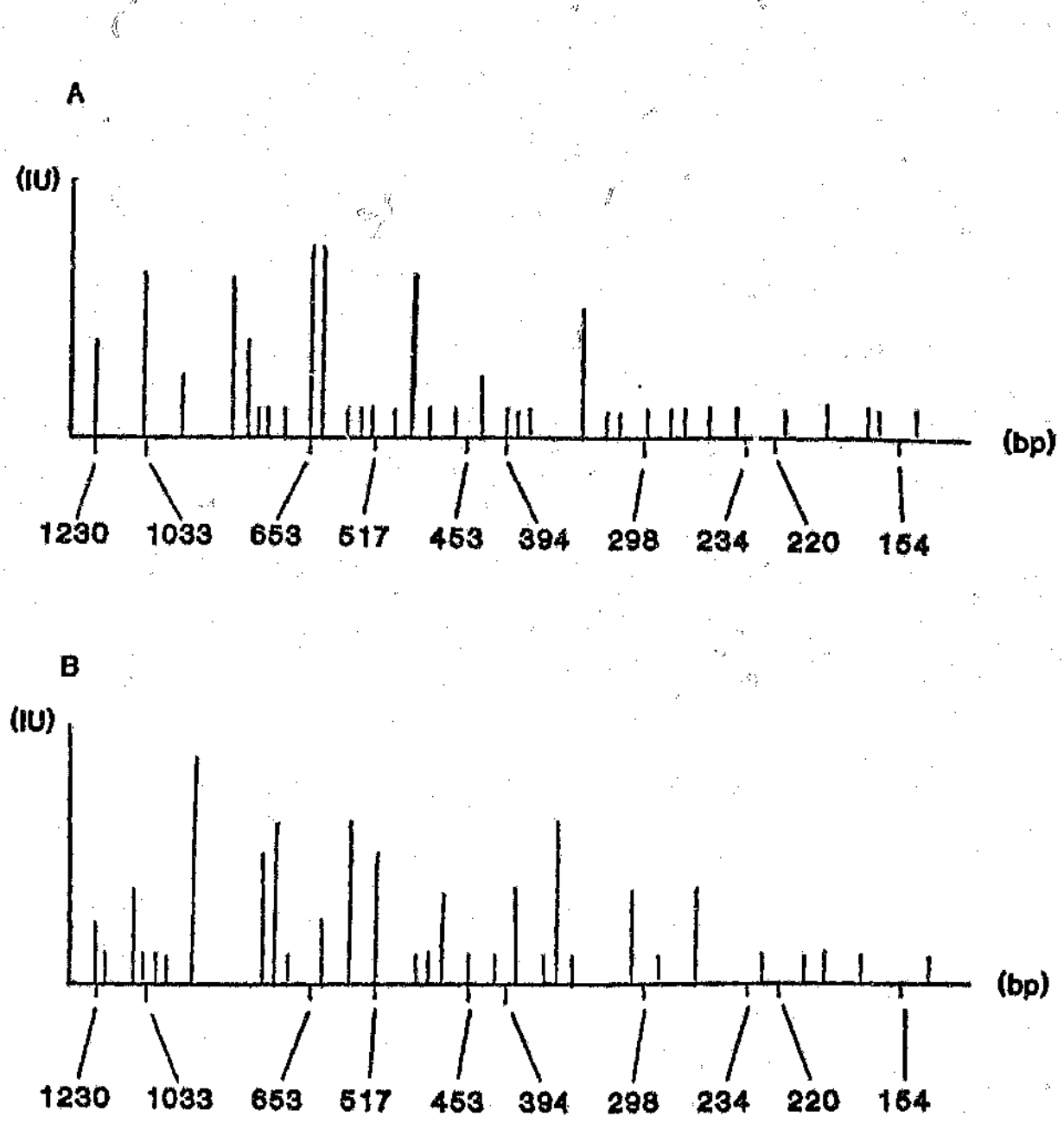


Figure 25: Histogram of DAF profiles of human and *L. Lesliei* DNA generated with CCCTCTGCGG. Panel A) human and panel B) *L. Leliei* DNA. There are no common bands.

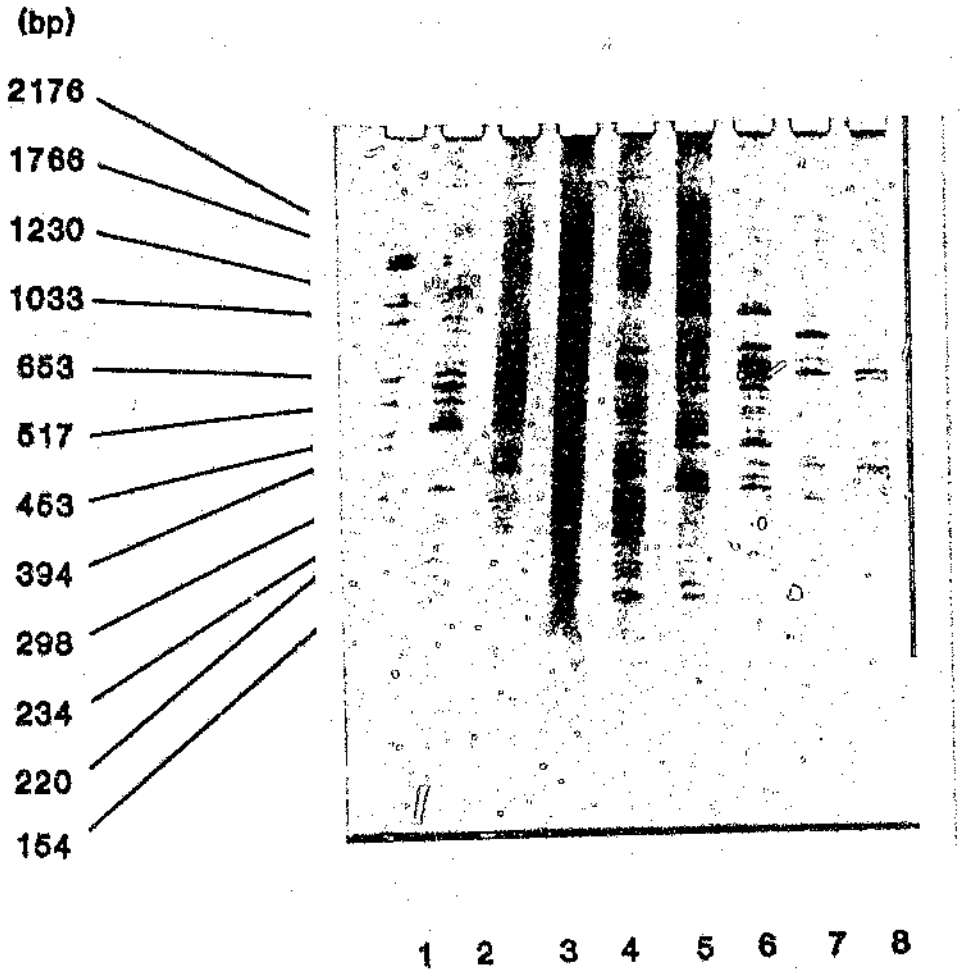


Figure 26: Comparing the reproducibility of DAF profiles from different laboratories. The DAF reactions of humen and lithop DNA, with ACGGTACACT and CCCTCTGCGG respectively, were repeated on the same model thermocycler but in separate laboratories. Lane (1 and 2) human with ACGGTACACT, lane (3 and 4) lithop with ACGGTACACT, lane (5 and 6) human with CCCTCTGCGG and lane (7 and 8) lithop with CCCTCTGCGG. The even and odd numbered lanes were run on separate machines.

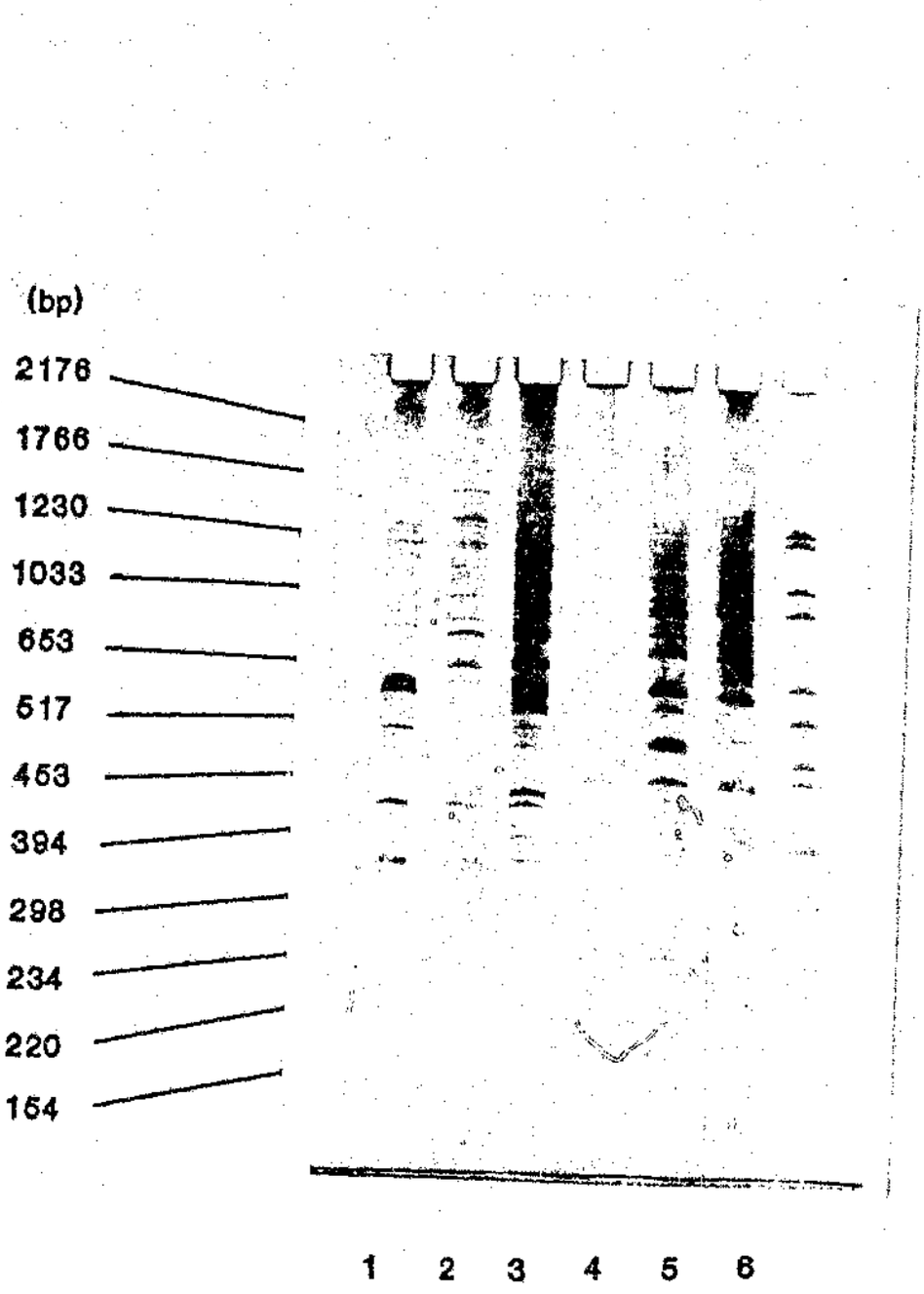


Figure 27: DAF of three different lithop species with ACGGTACACT and CCCTCTGCGG. The amplification of three different lithop species was directed by 500 ng of CCCTCTGCGG and ACGGTACACT respectively, using 5 mM $MgCl_2$. *L. Terricolor* (lane 1 and 4), *L. Hookeri* (lane 2 and 5) and *L. Lesliei* (lane 3 and 6).

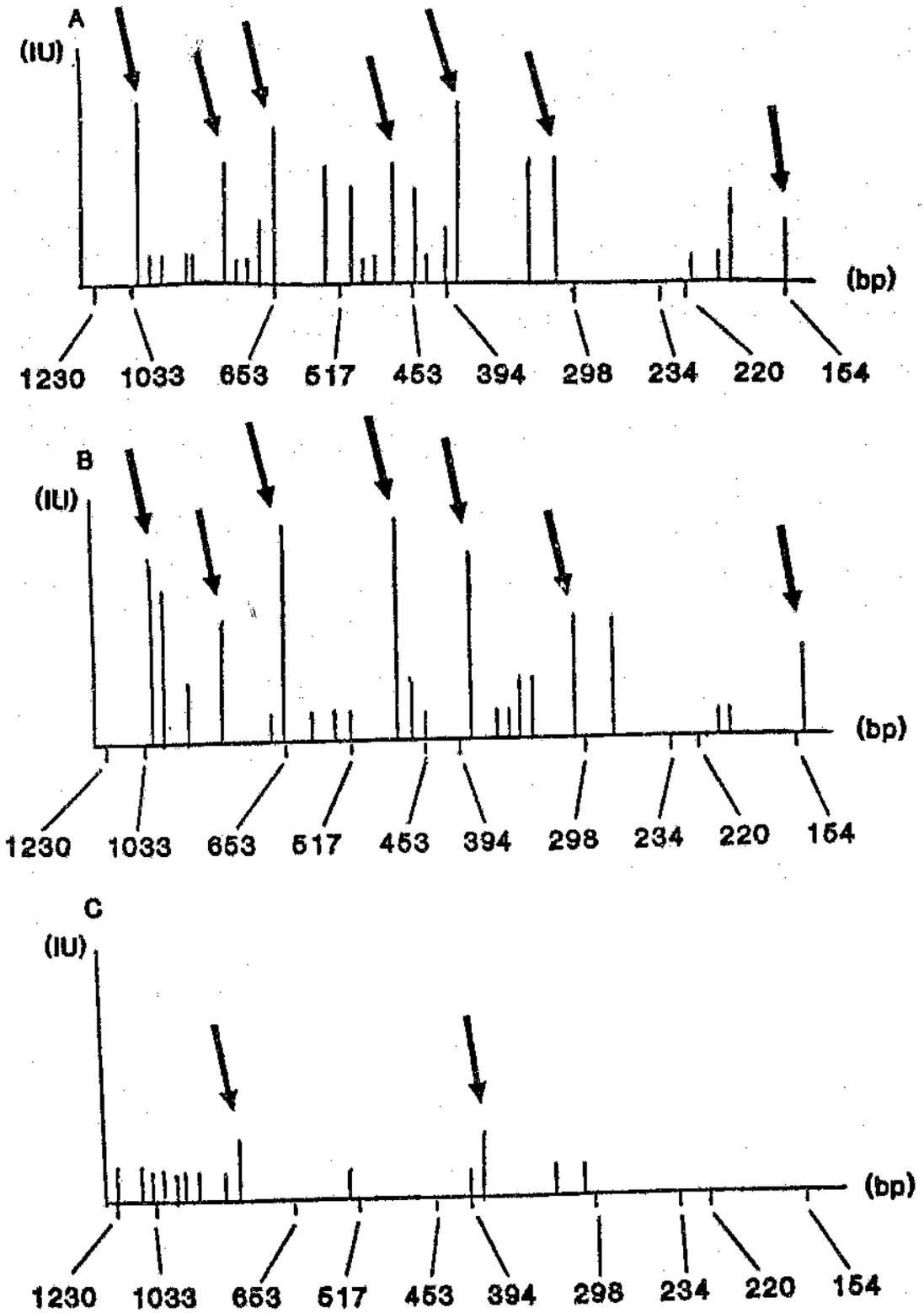


Figure 28: Histogram showing band mobility and intensity differences for the three lithop species with ACGGTACACT. DAF was directed with ACGGTACACT and bands below 1250 bp were scored. Panel A) *L. Lesliei*, panel B) *L. Hookeri* and panel C) *L. Terricolor*. All common bands are indicated with arrows.

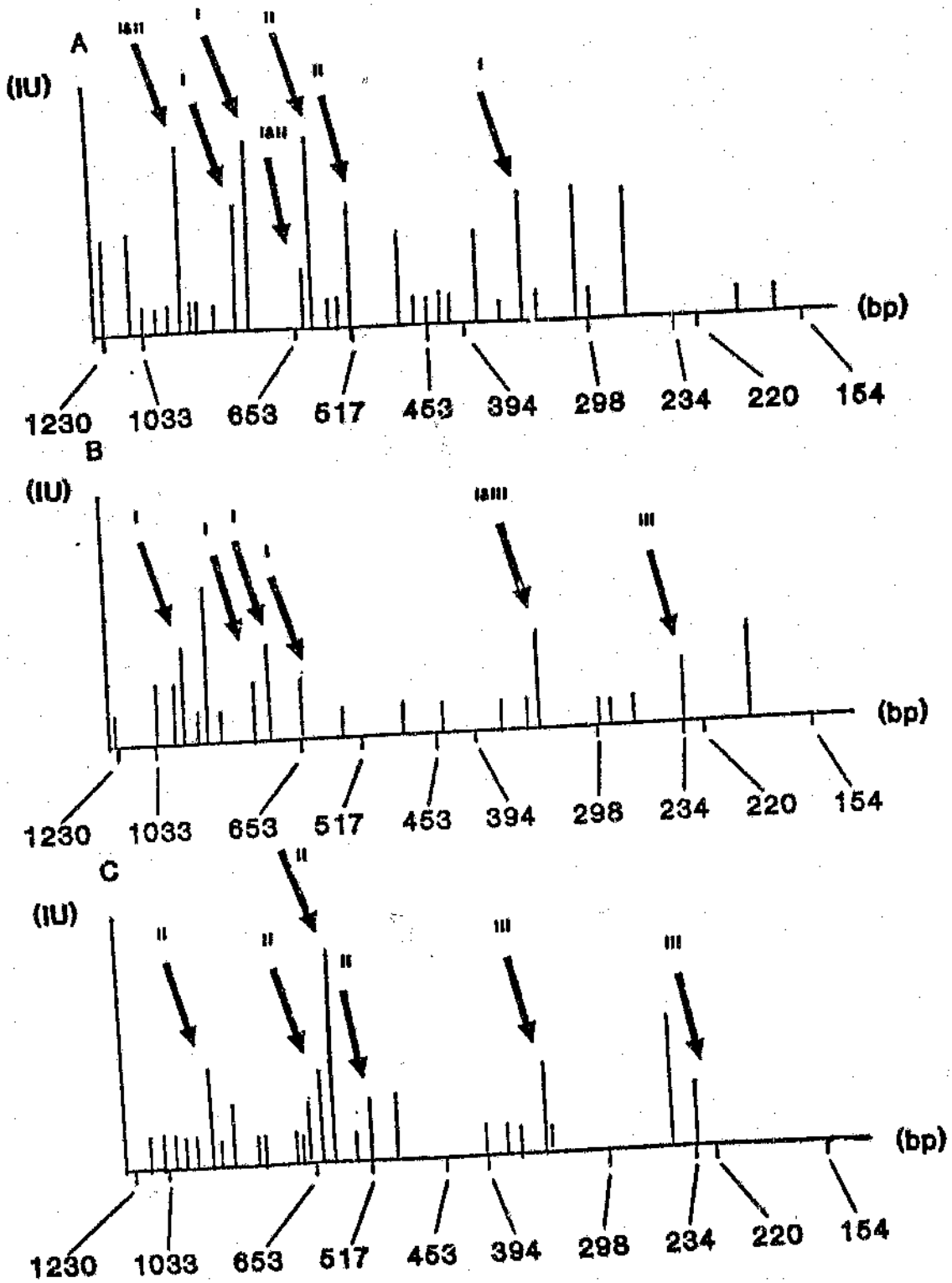


Figure 29: Histogram showing band mobility and intensity differences for the three lithop species with CCCTCTGCGG. Band scoring of the DAF profiles of *L. Lesliei* (panel A), *L. Hookeri* (panel B) and *L. Terricolor* (panel C) generated with CCCTCTGCGG. Arrows indicate common bands between (i) *L. Lesliei*/*L. Hookeri*, (ii) *L. Lesliei*/*L. Terricolor* and (iii) *L. Hookeri*/*L. Terricolor*.

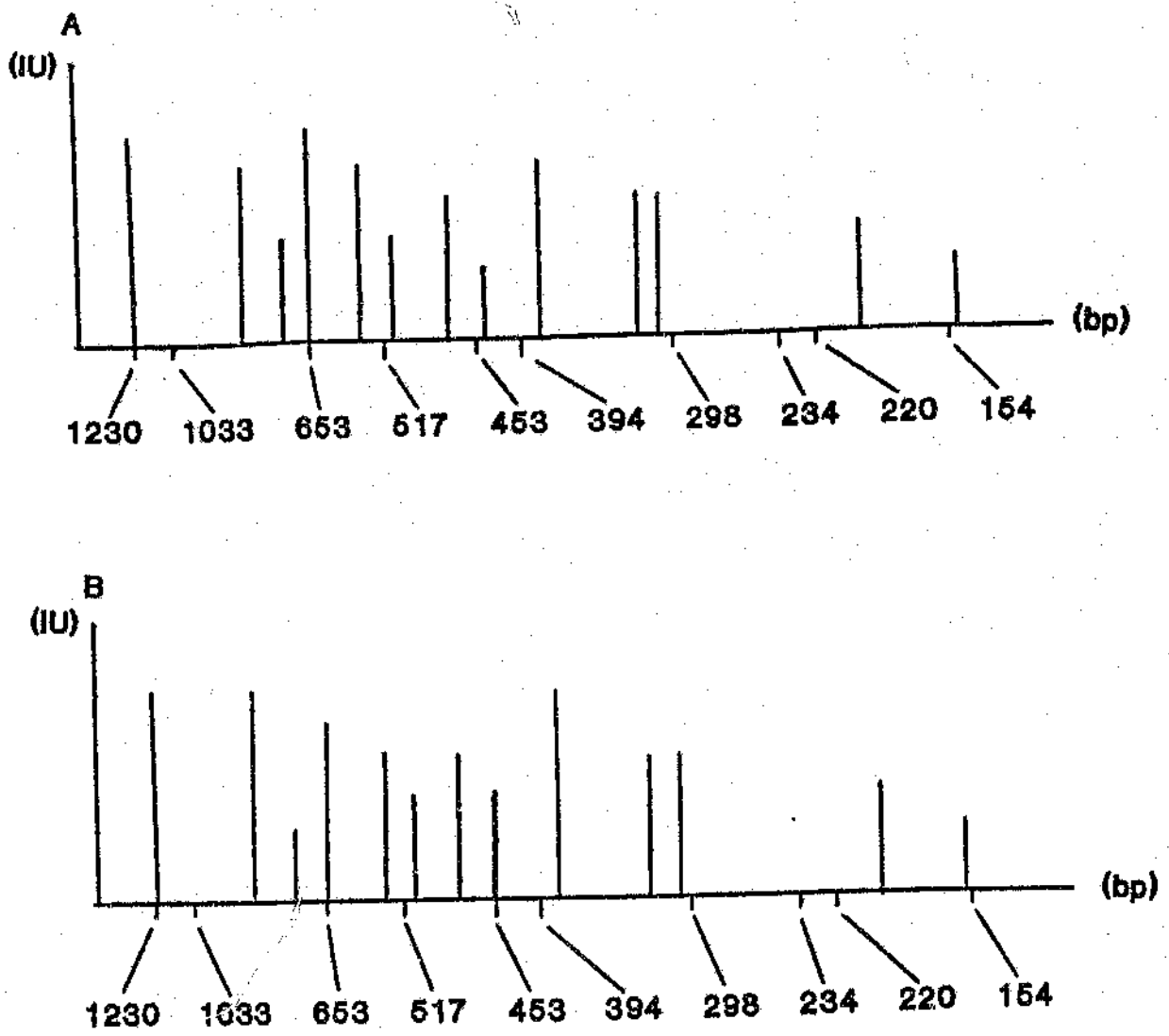


Figure 30: Comparing DAF profiles of *L. Lesliei* and ACGGTACACT between different gels. The resultant pattern was corrected for differences in the overall mobilities of each gel. Major bands from figure 24B (panel A) and from figure 28A (panel B) were counted. All thirteen major bands showed similar relative mobility and intensity.

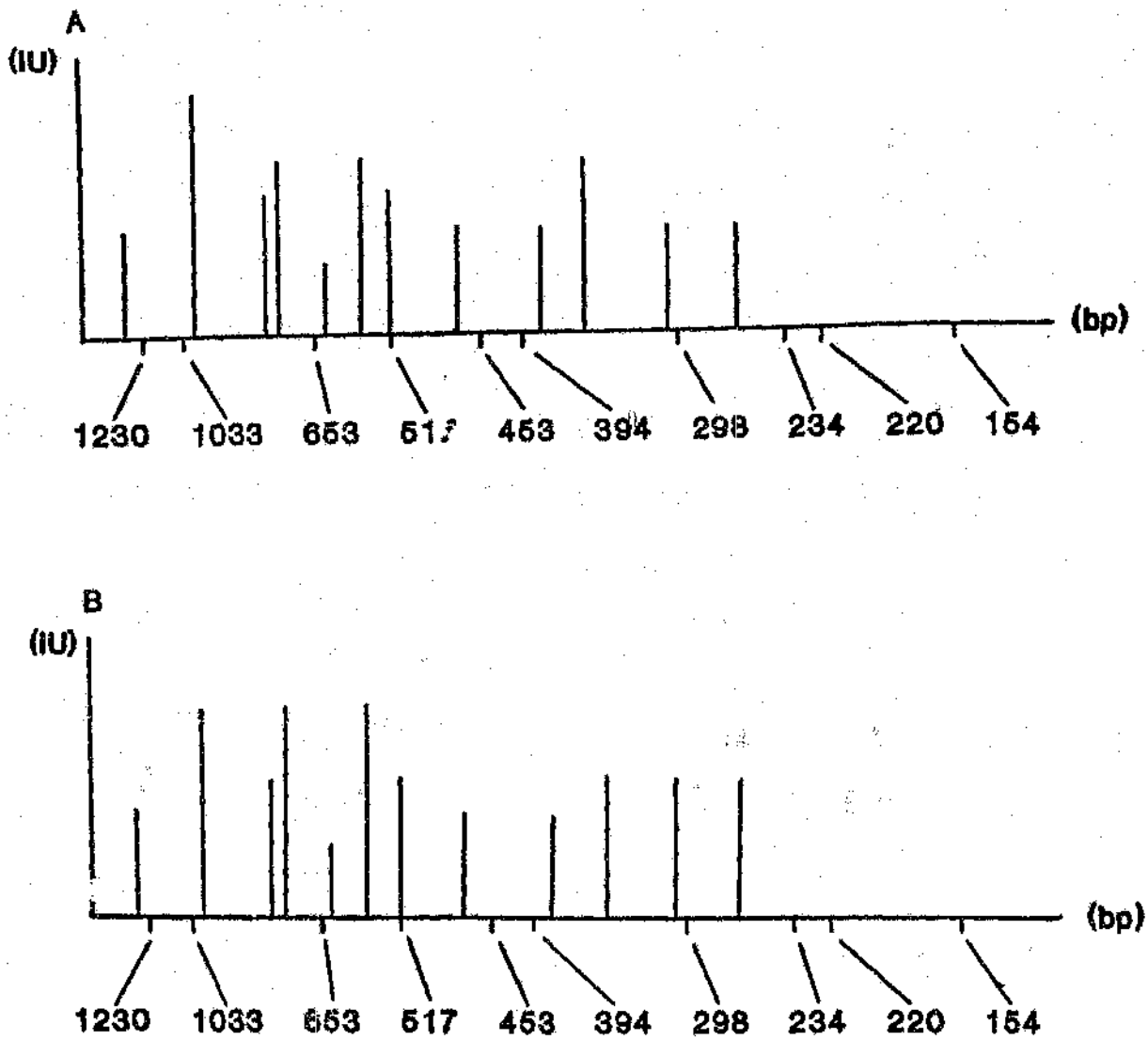


Figure 31: Measuring pattern variability between different *L. Leslie*i/CCCTCTGCGG profiles. Major bands from figure 25A (panel A) and figure 29A (panel B) were scored. All twelve major bands showed consistent mobility and intensity.

4.4 Discussion

4.4.1 Silver Staining

The molecular weight marker that has been used covers the expected size range for amplification products. It consists of 15 fragments of the following size (bp): 2176, 1766, 1230, 1033, 653, 517, 453, 394, 2x298, 2x234, 220 and 2x154. The 220 bp fragment represents only about 2% of the total number of base pairs and hence molecular weight. A total of 50 ng of marker was used during electrophoresis which means that approximately 1 ng of double stranded DNA is visualised in the 220 bp fragment. If ethidium bromide staining, which is far less sensitive, were used a considerable loss of information could result.

4.4.2 Electrophoresis

Although it is common for DAF profiles to be examined on agarose gels it cannot be disputed that polyacrylamide gels provide a higher degree of band resolution for DNA molecules up to 2000 bp. Caetano-Anolles *et al.* (1991) described the incorporation of 7 M urea into their acrylamide gels to ensure good band resolution. Their DAF reactions were also loaded in 5 M urea suggesting that their DAF products were being analysed as single stranded and not double stranded DNA fragments. They had, however, listed their molecular weight markers in terms of base pairs and not nucleotides as would be the case for single stranded DNA. To avoid similar confusion I wanted to ensure that the markers and DAF products were both either run as single or double stranded fragments.

When the boiled and unboiled markers are analysed on a 7 M urea (denaturing) gel the first thing that strikes one is the decreased mobility of single stranded DNA relative to the double stranded DNA. The unboiled markers do not spontaneously "melt" under the denaturing conditions (see figure 16). In fact even after being boiled in 5 M urea, it appeared as if some of the denatured fragments (>298 bp) were able to reanneal and "comigrate" with their undenatured counterparts. Maniatis *et al.* (1975) had noted that DNA molecules (>200 bp) could not be consistently denatured in 7 M urea gels at room temperature. It seemed

that the boiled samples were actually a mixture of single and double stranded DNA molecules, although I must add that this had not been established for certain. Turning my attention back to the DAF separation issue I decided to investigate all possibilities. The amplification of human DNA, directed with ACGGTACACT as described by Williams *et al.* (1990), was used as a reference sample. Some samples were boiled before being loaded onto a denaturing gel (see figure 17), others were boiled and loaded onto a non-denaturing gel (see figure 18) and still others were loaded onto a denaturing gel without being boiled (see figure 19).

When boiled samples were analysed under denaturing electrophoresis there was an obvious absence of very small fragments and the profiles did not resemble anything like those obtained by Caetano-Anolles *et al.* (1991). The denaturing conditions had obviously ensured that the DNA fragments were "melted" and hence running as single stranded molecules. Removal of urea from the gel facilitated the reannealing of some denatured DAF products and some "fast" migrating molecules were seen. When the samples were analysed in 7 M urea gels without being boiled, the DAF profiles of the human DNA came the closest to portraying the type of pattern that had been generated by Caetano-Anolles *et al.* (1991). Furthermore these profiles, although analysed on acrylamide and not agarose gels, were the same size as those originally described by Williams *et al.* (1990).

The human DNA DAF profiles, using template DNA prepared from different individuals, are very similar although a few individual specific bands can be seen. In addition the profile patterns were consistent over a four fold variation in template concentration, i.e. very small differences in the template concentration do not distort the overall profile pattern.

4.4.3 DAF of Lithops

Most DAF research starts with a screening process in which a kit of about twenty or so short (usually 10mers) random sequence oligonucleotides are used to direct amplification of a particular template DNA. Not all of these primers will produce interpretable patterns, but more often than not at least some will. Without large resources I decided to screen some of

my own oligonucleotides which I had tried to use for SQR profiling (see figure 13, chapter 3). The only selection criterion was that an oligonucleotide had to have a high GC content. $(TCC)_5$, $(GTG)_5$ and the 10mer ACGGTACACT (which had worked well with human DNA) were chosen (see figure 20). The 10mer, which had the lowest GC content (50%), gave the best profile of the lithop DNA illustrating that GC content should not be the only criterion for primer selection.

Although I was using a 10mer primer, which would be expected to direct the amplification of short DNA regions, there were no DAF fragments smaller than approximately 400 bp. Indeed there appeared to be a sharp cut off of detectable bands below the 394 bp marker. This was explained in a recent publication by Caetano-Anolles *et al.* (1992). They put forward a model in which a single primer amplifies certain products due to competition for annealing sites between the primer and terminal hairpin loop structures of the intermediate template molecules.

Short primers are expected to recognise a large number of annealing sites during primer screening. Amplification of these sites produces products with terminal sequence symmetry which are able to form hairpin loop structures and thus "block" primer binding. The primers will have great difficulty in displacing hairpin loop structures, especially those that are stabilised by additional base pair matches in internal regions of the loops. This can reduce the number of amplification products. Large loops have unstable hairpin structures, i.e. hairpin structures are less stable the longer the amplification product and the shorter the primer and thus the region of symmetry. These large "amplicons" will amplify efficiently because they do not form stable hairpin structures that can interfere in primer binding. Hairpin structures, formed from short loops, are more stable and will disrupt primer binding and therefore lower the efficacy with which these sites are amplified. Only by increasing the primer concentration will the primer be able to outcompete stable hairpin structures for potential binding sites. Increasing the amount of primer 2.5 times more than that suggested by Williams *et al.* (1990) was found to greatly increase the efficiency with which smaller DNA regions were amplified and hence detected (see figure 21). This whole exercise dem-

onstrated the importance of constantly reviewing all of the available literature regarding developments in a field.

Kaemmer *et al.* (1992) had been successful in using the 10mer CCCTCTGCGG to direct the DAF of banana cultivars. I decided that this primer, which had shown itself to be useful in plants, might be effective for distinguishing lithop species. A DAF profile directed by a single primer is very sensitive to the amount of $MgCl_2$ that is used. The resultant patterns that are detected when different concentrations of $MgCl_2$ are used can be quite varied (see figure 22). It is therefore crucial that once a $MgCl_2$ concentration has been selected it must be strictly adhered to. In this regard I have found it useful to include the $MgCl_2$ into a reaction "master mix" that contains all the buffer components. This reduces the possibility of different samples having different quantities of $MgCl_2$.

Most DAF profiles generally consist of a number of bands, anywhere from 10 to 20, which cover a similar size domain. Many of these bands will be found within a few millimetres of each other when examined on a 15 cm long acrylamide gel. This means that two bands representing distinct DNA sequences may be deemed to be the same because they have similar mobilities. This can be illustrated when the DAF profiles of human and lithop DNA, directed with ACGGTACACT and CCCTCTGCGG respectively, are compared (see figure 23). Each DAF profile can be plotted as a histogram (see figure 24 and 25). The distance along the horizontal axis is a measure of molecular weight, which is estimated from the relative migration of molecular weight standards. The height of each bar line indicates a difference in the band staining intensity. In this way bands with similar mobilities can be distinguished by their staining intensities.

Prominent bands from different DAF profiles having the same or similar mobilities (e.g. seen with human and lithop DNA) cannot be assumed to be identical, i.e. they should not be scored as common bands. The confusion arises from the fact that DAF markers represent size variations between arbitrarily amplified sequences. Markers with similar mobilities may only be assumed to represent the same locus if they are derived from related DNA templates or species. For this reason the degree of band sharing between *L. Lesliei* and human DNA,

although seemingly low, is probably not accurate. In contrast RFLP-type markers measure size variations of a particular sequence/s or allele/s that is/are found at a defined locus/loci, i.e. common bands must have the same size and sequence. DAF markers can be isolated, amplified, labelled and used to probe genomic DNA digests. If two DAF markers produce the same hybridisation pattern with the same genomic digest only then they can be assumed to represent the same genetic locus. The actual sequence of this allele will simply be the complement of the marker. In this way DAF markers can be used to "extract" sequence data from lithops.

The literature abounds with different time and temperature cycles for DAF that make no mention of the heating and cooling rates of their PCR machines. Some researchers claim that DAF profiles are not reproducible in different laboratories (see figure 26). DAF profiles are sensitive to variations in both time and temperature and so it is the actual time and temperature spent at each step that has to be followed. Thermocyclers have different heating and cooling rates and therefore different programs may have to be used on different machines to effect the same times and temperatures during cycling.

Three lithop species were profiled, viz. *L. Lesliei*, *L. Hookeri* and *L. Terricolor* using the two 10mers ACGGTACACT and CCCTCTGCGG (see figure 27). The large number of similarly sized markers makes it technically difficult to score the degree of band sharing. Other researchers who use DAF have also noted this and many have used ethidium bromide stained agarose gels to reduce the number of detectable bands and thereby resolve the profiles. As I have already indicated this leads to a loss of data which may be potentially useful, especially since these markers can be recovered and sequenced. Instead the profiles were reproduced as histograms and the most prominent bands were scored (see figures 28 and 29). The highest percent of band sharing was 32% between the *L. Lesliei* and *L. Hookeri* species which are both found in the western Transvaal. Both of these species have significantly lower band sharing frequencies with *L. Terricolor*, 18% and 12% respectively, which is found in the southern Cape. The number of unique bands easily exceeds the number of shared bands which is in contrast to the homogeneity seen amongst the different human DNA templates

(see figure 19). The significance of the ratio of unique to common bands should be examined further with hybridisation studies.

It is crucial in DAF analysis that some type of standard reaction is always used so that comparisons can be made between different gels run on different days. By using the DAF reactions of *L. Lesliei* and the respective primer as the standard I was able to check the consistency of the reaction with time, i.e. the relative mobility and intensity of the *L. Lesliei* markers were shown to be consistent (see figures 30 and 31). Slight variations in the differential mobilities of gels that are being compared must be accounted for.

4.5 Conclusion

Although it seems to be generally accepted that DAF can be used to indicate differences between genomic templates, these differences cannot be scored like RFLP markers because they do not represent a specific genetic locus, i.e. randomly amplified markers cannot be used to distinguish between genomes just because they exhibit size variations. RFLP-type analyses use differences in the distribution as well as length variations of a defined sequence to differentiate genomes, i.e. RFLP markers representing a given sequence (or close variants thereof) are scored according to size variation. As a result, it is my belief that either single or multiple locus hybridisation studies should be carried out to verify any phylogenetic relationships that have been derived from DAF profiles.

Two options for generating RFLP markers exist. Firstly, a range of different probes could be designed and screened to to see if they generate scoreable hybridisation patterns. This is costly and time consuming process. The second and better alternative is to use DAF markers to extract sequence data from lithops, which can then be amplified, labelled and used to probe genomic digests of other lithop species. In this way one can ensure that the probes that are used will hybridise to lithop sequences. Some of these lithop sequences may prove to be polymorphic while others will be monomorphic thus opening the door for a more definitive form of profiling.

CHAPTER 5

5 SUMMARY

DNA profiling is a general term that encompasses several distinct techniques or rationales for measuring genetic similarity. The basis for all these methodologies is that genetic differences (polymorphisms) are synonymous with genetic relatedness, i.e. closely related organisms will have greater sequence similarity (or fewer differences) than unrelated individuals. These methods differ in the way in which they detect DNA polymorphisms. In the original DNA profiling strategy, described by Jeffreys, a labelled DNA probe is used to hybridise to its complementary sequence/s in an immobilised genomic DNA digest. The resultant hybridisation pattern reflects the distribution and organisation of that sequence in an organism's genome. Certain sequences, which are found at specific loci, will occur in different copy numbers in unrelated individuals. These length variations will be seen as mobility differences in their respective hybridisation patterns (profiles). Different sequence probes are used to study different types of genetic loci, e.g. single and multiple locus, VNTR and SQR probes.

The discovery of the polymerase chain reaction opened up new approaches for DNA profiling. DNA amplification fingerprinting uses short, arbitrary sequence oligonucleotides (primers) to direct the amplification of random regions of DNA from genomic DNA sites to which they have fortuitously hybridised. The result is a spectrum of amplified products which are characteristic of the genomic DNA template. Unrelated genomes will produce different amplification spectra when analysed on silver stained acrylamide gels.

To be able to make use of the analytical advantages of DNA profiling I wanted to establish a procedure that could be used with lithops. The were two initial problems, firstly lithop DNA had never before been isolated and because of this no sequence data were available. To overcome the first problem I used a plant DNA isolation technique, first described by

Murray and Thompson (1980), which had been used with variable success by Weising *et al.* (1991) to isolate DNA from a wide variety of plant types. This method uses lyophilised and not fresh tissue which is actually a great advantage because whole plants can be stored in a desiccated state until they are needed. One can also build up sufficient stocks of plant material without having to wait until the growing season. Nucleic acids are purified at high salt concentration (1,4 M NaCl) by exploiting their solubility in a 2% (w/v) CTAB buffer. Denatured proteins and polysaccharides are insoluble under these conditions.

I was encouraged with the reasonably high yield of lithop DNA which was routinely found to be greater than 200 µg/g lyophilised tissue. However, these preparations were found to be highly degraded when analysed on agarose gels. Only 15 µg (about 7% of the total DNA) of high molecular weight DNA could be spooled. Although this was a very low yield it is in line with the difficulties that one can experience in isolating plant DNA. Weising *et al.* (1991) could not isolate high molecular weight DNA from several plant species. There were some cases where they could not isolate any DNA (see figure 12). A small 1 ml column, packed with a gel having suitable size exclusion limits, was used to purify high molecular weight DNA from low molecular weight contaminants. The resultant elution profile is a visual representation of the extent of DNA degradation and contamination of the crude preparation.

Weising *et al.* (1991) had shown that short, quadruplet repeat sequences were present throughout the plant kingdom. Some of these sequences displayed length polymorphisms and could therefore be used as DNA profile markers. Since there were no known polymorphic lithop sequences I decided to test whether some of these SQR sequences would yield scoreable hybridisation patterns with lithop DNA. The oligonucleotides $(TCC)_5$, $(GATA)_4$, $(GACA)_4$ and $(GTG)_5$ were synthesised, end labelled with digoxigenin and hybridised to EcoR I digested lithop DNA, immobilised in a dried agarose gel matrix. The in-gel hybridisation technique was shown to work with a standard hybridisation reaction (see figure 12), but the SQR oligonucleotides used only generated smeared patterns (see figure 13).

Although this was disappointing it was, in retrospect, the most probable outcome especially in view of the fact that I was trying to find a polymorphic sequence empirically, much like looking for a needle in a haystack. There was some consolation in the fact that Weising *et al.* (1991) were also unable to produce scoreable profiles with several combinations of plant DNA and SQR probe. They confirmed that in some cases they could not generate any hybridisation pattern while in others they could only generate a smeared pattern which was impossible to score (see figure 33).

The discovery of PCR introduced several new strategies for profiling. DNA amplification fingerprinting can be carried out with very little template DNA (25 ng) and requires no sequence data because amplification is directed by short arbitrary sequence primers. This technique offered the most realistic way for overcoming the initial problems for generating DNA profiles of lithops. For this reason a lot of time was spent trying to understand and standardise a procedure that was informative and reproducible. DAF was found to be sensitive to the actual time and temperature spent at each stage during cycling. Changes in the concentrations of primer and $MgCl_2$ were reflected by differences in the spectrum of amplified products. The procedure was not sensitive to small differences (4 fold) in the amount of template DNA. Different individuals from the same species had homogeneous profiles whilst individuals from distinct species produced unrelated profiles.

There are two levels of polymorphisms that can be measured, firstly there are length variations between markers and secondly there are variations in the staining intensity of similar size markers. These differences can be represented on a histogram where the distance along the horizontal axis is related to a markers mobility relative to molecular weight markers and the height of each bar line represents a difference in staining intensity. In this way common bands will be scored only if they have similar mobilities and intensities. Bands from organisms that are not known to be closely related cannot be scored on this basis alone because two markers which have the same or similar mobilities might represent two unrelated sequences or loci. The banding pattern generated with a particular combination of template

and primer was shown to be consistent for different DAF reactions analysed on different gels. This is important because it enables one to use a particular reaction as a standard measure of the uniformity of the reaction conditions over time, i.e. one can be sure that any differences are only the result of differences in the template DNAs and not the procedure.

RFLP type markers measure genetic variation as a function of length polymorphisms of a particular sequence/s. As a result of their specificity for a particular sequence/s (or closely related derivatives) these markers can be quantified on the basis of band sharing frequencies alone, i.e. they represent an absolute and not an arbitrary measure of sequence divergence. DNA amplification can be used to develop RFLP type strategies by "extracting" DNA sequence information from lithops. Amplified markers can be isolated, amplified and labelled and used to probe restriction DNA digests of other lithops. These markers can be sequenced and thereby used to extract sequence information from lithops. If any of these markers are found to be polymorphic they can be used to score length polymorphisms of specific sequences or loci. This is a better strategy for developing RFLP type markers than simply screening a series of SQR oligonucleotides because one will be using actual lithop DNA sequences.

Fungi	
Phycomyces blakesleeanus (Zygomycetes)	+
Saccharomyces cerevisiae (Ascomycetes)	-
Ascochyta rabiei (Ascomycetes)	+
Coprinus comatus (Basidiomycetes)	+
Nematoloma sublateritium (Basidiomycetes)	-
Algae	
Oedogonium spec. (Chlorophyceae)	(+)
Spirogyra spec. (Chlorophyceae)	-
Stenogramme interrupta (Rhodophyceae)	+
Fucus serratus (Phaeophyceae)	(+)
Liverworts and Mosses	
Marchantia polymorpha (Hepaticae)	(+)
Polytrichum formosum (Musci)	+
Ferns	
Equisetum arvense (Equisetatae)	+
Dodia caudata (Filicatae)	+
Polypodium vulgare (Filicatae)	+
Osmunda vulgaris (Filicatae)	+
Nephrolepis exaltata (Filicatae)	(+)
Gymnosperms	
Ephedra distachya (Gnetatae)	-
Gingko biloba (Gingkoatae)	-
Taxus baccata (Pinatae)	(+)
Taxodium distichum (Pinatae)	-
Thuiopsis dolabrata (Pinatae)	-
Juniperus communis (Pinatae)	*
Cryptomeria japonica (Pinatae)	(+)
Monocotyledonous angiosperms	
Echinodorus osiris (Alismataceae; Alismatidae)	+
Musa acuminata (Musaceae; Liliidae)	+
Hordeum vulgare (Poaceae; Liliidae)	+
Hordeum spontaneum (Poaceae; Liliidae)	+
Asparagus densiflorus (Asparagaceae; Liliidae)	+
Dioscorea bulbifera (Dioscoreaceae; Liliidae)	+
Cocos nucifera (Arecaceae; Arecidae)	(+)
Chamaedorea cataractarum (Arecaceae; Arecidae)	+
Monstera deliciosa (Araceae; Arecidae)	+
Dicotyledonous angiosperms	
Persea americana (Lauraceae; Magnoliidae)	-
Laurus nobilis (Lauraceae; Magnoliidae)	(+)
Helleborus niger (Ranunculaceae; Magnoliidae)	+
Urtica dioica (Urticaceae; Hamamelidae)	+
Ficus benjamina (Moraceae; Hamamelidae)	+
Humulus lupulus (Cannabinaceae; Hamamelidae)	+
Aruncus silvester (Rosaceae; Rosidae)	+
Cicer arietinum (Fabaceae; Rosidae)	+
Lens culinaris (Fabaceae; Rosidae)	+
Lens orientalis (Fabaceae; Rosidae)	+
Lens nigricans (Fabaceae; Rosidae)	+
Hippophae rhamnoides (Eleagnaceae; Rosidae)	-
Aucuba japonica (Cornaceae; Rosidae)	-
Simmondsia sinensis (Buxaceae; Rosidae)	+
Macaranga hullett (Euphorbiaceae; Rosidae)	+
Macaranga triloba (Euphorbiaceae; Rosidae)	(+)
Macaranga tanarius (Euphorbiaceae; Rosidae)	+
Macaranga hypoleuca (Euphorbiaceae; Rosidae)	-
Camellia sinensis (Theaceae; Dilleniidae)	+
Brassica oleracea (Brassicaceae; Dilleniidae)	+
Brassica napus (Brassicaceae; Dilleniidae)	+
Brassica campestris (Brassicaceae; Dilleniidae)	+
Bryonia dioica (Cucurbitaceae; Dilleniidae)	+
Silene alba (Caryophyllaceae; Caryophyllidae)	+
Silene dioica (Caryophyllaceae; Caryophyllidae)	+
Beta vulgaris (Chenopodiaceae; Caryophyllidae)	+
Rumex acetosella (Polygonaceae; Caryophyllidae)	+
Nicotiana tabacum (Solanaceae; Asteridae)	+
Nicotiana silvestris (Solanaceae; Asteridae)	+
Nicotiana acuminata (Solanaceae; Asteridae)	+
Nicotiana otophora (Solanaceae; Asteridae)	+
Nicotiana glutinosa (Solanaceae; Asteridae)	+
Nicotiana paniculata (Solanaceae; Asteridae)	+
Lycopersicum esculentum (Solanaceae; Asteridae)	+
Lycopersicum hirsutum (Solanaceae; Asteridae)	+
Solanum tuberosum (Solanaceae; Asteridae)	+
Helianthus annuus (Asteraceae; Asteridae)	+
Lactuca sativa (Cichoriaceae; Asteridae)	+

Figure 32: Overview of the results of DNA isolation from different species of fungi and plants obtained by Weising *et al.* (1991).

+ - high molecular weight DNA obtained

(-) - no DNA obtained or DNA largely degraded

(+) - lo yield of DNA or DNA partially degraded

* - DNA obtained from fruit but not leaf tissue.

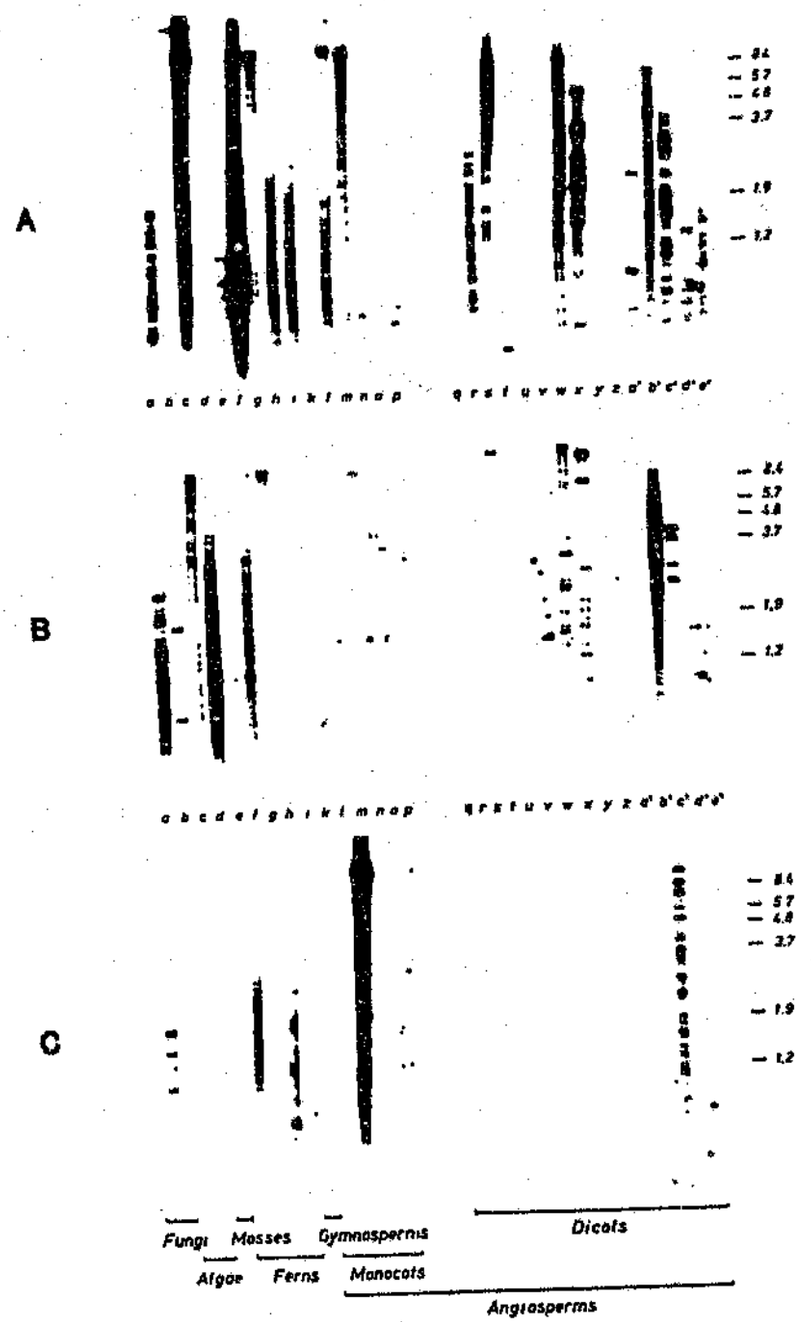


Figure 33: Hybridisation patterns of SQR sequences with different plant genomes as obtained by Weising *et al.* (1991). The oligonucleotides were labelled with ^{32}P and hybridised to Hinf1 digested genomic DNA in 1% dried agarose gels. Panel A $(GATA)_4$, panel B $(GACA)_4$ and panel C $(TCC)_5$.

CHAPTER 6

6 REFERENCES

Albretson C., Haukanes B., Aasland R. and Kleppe K. (1988) *Analytical Biochemistry* **170**, 193-202

Baird E., Coope-Bland S., Waugh R., DeMaine M. and Powell W. (1992) *Molecular and General Genetics* **233**, 469-475

Ballinger-Crabtree M.E., Black W.C. IV and Miller B.R. (1992) *American Journal of Tropical Medicine and Hygiene* **47**, 893-901

Barua A.M., Chalmers K.J., Hackett C.A., Thomas W.T.B., Powell W. and Waugh R. (1993) *Heredity* **71**, 177-184

Bassam B.J., Caetano-Anolles G. and Gresshoff P.M. (1991) *Analytical Biochemistry* **196**, 80-83

Bassam B.J., Caetano-Anolles G. and Gresshoff P.M. (1992) *Applied Microbiology and Biotechnology* **38**, 70-76

Beidler J.L., Hilliard P.R. and Rill R.L. (1982) *Analytical Biochemistry* **126**, 374-380

Bell C.R. (1969) Plant Variation and Classification, Macmillan, London, 40-49

Bierwerth S., Kahl G., Weigand F. and Weising K. (1992) *Electrophoresis* **13**, 115-122

Blum H., Beier H. and Gross H.J. (1987) *Electrophoresis* **8**, 93-99

Boehringer Mannheim (1993) The DIG System User's Guide for Filter Hybridisation, Boehringer Mannheim GmbH - Biochemica. Mannheim, 1-17

Briggs D. and Walters S.M. (1984) Plant Variation and Evolution 2nd edition, Cambridge University Press, Cambridge, 90-114

Caetano-Anolles G., Bassam B.J. and Gresshoff P.M. (1991) *BioTechnology* **9**, 553-557

Caetano-Anolles G., Bassam B.J. and Gresshoff P.M. (1992) *Molecular and General Genetics* **235**, 157-165

Cole D.J. (1988) Lithops: Flowering Stones, Acorn Books, Randburg South Africa, 33-52

Debenham P.G. (1992) *Trends in Biotechnology* **10**, 96-102

Deliaporta S.L., Wood J. and Hicks J.B. (1983) *Plant Molecular Biology Reporter* **1**, 19-21

Dweikat I., Mackenzie S., Levy M. and Ohm H. (1993) *Theoretical and Applied Genetics* **85**, 497-505

Epplen J.T. (1992) in Advances in Electrophoresis, ed. Chrambach A., Dunn M.J. and Radola B.J. Volume 5, VCH, New York, 62-112

Feinberg A.P. and Vogelstein B. (1983) *Analytical Biochemistry* **132**, 6-13

Fleck A. and Munro H.N. (1962) *Biochimica et Biophysica Acta* **55**, 571-583

Gibbs R.A. (1990) *Analytical Chemistry* **62**, 1202-1214

Hartl D.L. and Clark A.G. (1989) Principles of Population Genetics 2nd edition, Sinauer Associates Inc., Massachusetts, 6-30

Heukeshoven J. and Dernick R. (1985) *Electrophoresis* **6**, 103-112

Ikuta S., Takagi K., Wallace R.B. and Itakura K. (1987) *Nucleic Acids Research* **15**, 797-811

Jeffreys A.J, Wilson V. and Thein S.L. (1985) *Nature* **314**, 67-73

Jones S.B. Jr. and Luchsinger A.E. (1979) Plant Systematics, McGraw-Hill, New York. 113-127

Kaemmer D., Afza R., Weising K., Kahl G. and Novak F.J. (1992) *BioTechnology* **10**, 1030-1035

Lander E.S. (1989) *Nature* **339**, 501-505

Lathe R. (1985) *Journal of Molecular Biology* **183**, 1-12

Maniatis T., Jeffrey A. and van de Sande H. (1975) *Biochemistry* **14**, 3787-3794

Maniatis T., Fritsch E.F. and Sambrook J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, 174-177

Meinkoth J. and Wahl G. (1984) *Analytical Biochemistry* **138**, 267-284

Merril C.R., Goldman D., Sedman S.A. and Ebert M.H. (1981) *Science* **211**, 1437-1438

Morris C.J.O.R. and Morris P. (1976) Separation Methods in Biochemistry 2nd edition, Pitman, London, 413-470

Murray M.G. and Thompson W.F. (1980) *Nucleic Acids Research* **8**, 4321-4325

Nakamura Y., Leppert M., O'Connell P., Wolff R., Holm T., Culver M., Martin C., Fujimoto C., Hoff M., Kumlin E., and White R. (1987) *Science* **235**, 1616-1622

Nei M. (1975) Molecular Population Genetics and Evolution, North-Holland, Amsterdam, 211-230

Neto E.F., de Souza C.P., Rollison D., Katz N., Pena S.D.J. and Simpson A.J.G. (1993) *Molecular and Biochemical Parasitology* **57**, 83-88

Newbury H.J. and Ford-Lloyd B.V. (1993) *Plant Growth Regulation* **12**, 43-51

Pena S.D.J., Macedo A.M., Gontijo N.F., Medeiros A.M. and Ribeiro J.C.C. (1991) *Electrophoresis* **17**, 146-152

Puterka G.J., Black W.C. IV, Steiner W.M. and Burton R.L. (1993) *Heredity* **70**, 604-618

Rabilloud T. (1992) *Electrophoresis* **13**, 429-439

Rassmann K., Schlotterer C. and Tautz D. (1991) *Electrophoresis* **12**, 113-118

Richards E. (1990) in Current Protocols, ed. Ausbel F.M., Brent R., Kingston R.E., Moore D., Seidman J.G., Smith J.A. and Struhl K. Suppl. 9, 2.3.1-2.3.3

Rivin C. (1986) *Methods in Enzymology* **118**, 75-86

Robyt J.F. and White B.J. (1987) Biochemical Techniques: Theory and Practice, Brooks/Cole, Monterey California, 278-282

Rogstad S.H., Patton J.C. and Schaal B.A. (1988) *Proceedings of the National Academy of Sciences USA* **85**, 9176-9178

Saghai-Maroof M.A.,Soliman K.M., Jorgensen R.A. and Allard R.W. (1984) *Proceedings of the National Academy of Sciences USA* **81**, 8014-8018

Southern E.M. (1975) *Journal of Molecular Biology* **98**, 503-517

Stenesh J. (1984) Experimental Biochemistry, Allyn and Bacon, Boston, 310-311

Swanson C.P., Merz T. and Young W.J. (1981) Cytogenetics 2nd edition, Prentice Hall Inc., London, 121-138

Tautz D. (1989) *Nucleic Acids Research* **17**, 6463-6471

Thacker J. (1994) *BioTechniques* **16**, 252-253

Thompson J.S. and Thompson M.W. (1986) Genetics in Medicine 4th edition, W.B. Saunders, London, 38-39

Tsanev R. and Markov G.G. (1960) *Biochimica et Biophysica Acta* **42**, 442-452

Weaver K.R., Caetano-Anolles G., Gresshoff P.M. and Callahan L.M. (1994) *BioTechniques* **16**, 226-227

Weising K., Beyermann B., Ramser J. and Kahl G. (1991) *Electrophoresis* **12**, 159-169

Welsh J. and McClelland M. (1990) *Nucleic Acids Research* **18**, 7213-7218

Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A. and Tingey S.V. (1990) *Nucleic Acids Research* **18**, 6531-6535

Wong Z., Wilson V., Jeffreys A.J. and Thien S.L. (1986) *Nucleic Acids Research* **14**, 4605-4616

Yang X. and Quiros C. (1993) *Theoretical and Applied Genetics* **86**, 205-212

Zischler H., Hinkkanen A. and Studer R. (1991) *Electrophoresis* **12**, 141-146

APPENDICES

APPENDIX I

Extraction of Chicken DNA

15 µl of whole chicken blood was suspended in 600 µl of isotonic buffer (0,15 M NaCl, 1 mM EDTA and 50 mM Tris-HCl pH 8,0). The cells were lysed by the addition of 7,5 µl of 25% (w/v) SDS solution. Nucleases were inactivated by incubation with 15 µl proteinase K (10 mg.ml) overnight at 55° C. 500 µl of buffered phenol (pH 8,0) was added and the phases were mixed for 20 minutes and then separated by centrifugation. The aqueous phase was recovered and reextracted with phenol. The aqueous phase was again collected and traces of phenol were removed by extraction with chloroform/isoamyl alcohol (23:1). The DNA is recovered from the aqueous phase by adding 2 volumes of cold (-20° C) ethanol for 2 hours at -20° C. The DNA was collected by centrifugation, vacuum dried and resuspended in TE buffer (10 mM Tris-HCl and 1 mM EDTA pH 8,0).

APPENDIX II

BASIC Program Characterising an Oligonucleotide from its Base Sequence

```
1  REM A260 determinations for oligos of a given sequence
2  CLS
3  PRINT "      A260 DETERMINATIONS FOR OLIGOS OF A GIVEN SEQUENCE"
4  PRINT
5  INPUT "If your sequence is repetitive enter y, otherwise enter n"; P$
6  PRINT
8  IF P$ = "y" THEN 200
9  IF P$ = "Y" THEN 200
10 PRINT "Please enter the sequence of your oligo IN CAPITAL LETTERS"
14 PRINT
15 INPUT C$
16 A=0:C=0:G=0:T=0
17 O=LEN(C$)
18 FOR H=1 TO O
19 A$=MID$(C$,H,1)
20 IF A$="A" THEN A=A+1
21 IF A$="C" THEN C=C+1
22 IF A$="G" THEN G=G+1
23 IF A$="T" THEN T=T+1
24 NEXT H
25 CLS
26 PRINT "      A260 DETERMINATIONS FOR OLIGOS OF A GIVEN SEQUENCE"
27 PRINT
28 PRINT
29 PRINT "The sequence of your oligo is ";C$
30 PRINT
31 PRINT "Its composition is A  G  C  T"
32 PRINT "                  ";A;G;C;T
34 EA=15300: EG=11800: EC=7400: ET=9300
35 CA=EA*A: CG=EG*G: CC=EC*C: CT=ET*T
40 CO=CA+CG+CC+CT
50 MW=330*A+346*G+306*C+321*T+18+17
60 AbG=CO/MW
61 AB=ABS\1
70 PRINT
97 PRINT "The Mw of your oligodeoxynucleotide is";MW
98 PRINT
99 IF (A+C+G+T)<18 THEN M=(G+C)*4+(A+T)*2\1
100 IF (A+C+G+T)>18 THEN M=69.3+.41*(G+C)*2\1
101 PRINT "The melting temperature of its double stranded form is";M"degC"
102 PRINT
103 IF (C+T)=0 THEN PRINT "The ratio of purines to pyrimidines is undefined":GOT
O 105
104 PRINT "The ratio of purines to pyrimidines is " USING "£.ff";(A+G)/(C+T)
105 PRINT
106 PRINT "The (G+C) content is";(G+C)*100\(A+C+G+T) "percent"
107 PRINT
110 PRINT "The absorbance at 260 nm of a 1 mg/ml solution should be";AB
120 PRINT
130 BA=1000\AB
140 POD=1000:BA/MW
141 PO = PI)O\1
150 PRINT 'One A260 unit will equal";BA "ug or as nmol " USING "ff.f";PO
160 PRINT
170 INPUT "If you wish to run again enter y, if not enter n";Q$
180 IF Q$ = "y" THEN 2
181 IF Q$ = "Y" THEN 2
182 PRINT
183 CLS
185 PRINT "                                        CHEERIO"
187 PRINT ""
190 END
200  INPUT "Please enter the repeat sequence IN CAPITAL LETTERS";B$
210  PRINT
220  INPUT "How many times does it repeat";X
230  O=LEN(B$)
240  A=0:G=0:C=0:T=0
250  FOR I=1 TO O
260  A$=MID$(B$,I,1)
270  IF A$="A" THEN A=A+1
280  IF A$="C" THEN C=C+1
290  IF A$="G" THEN G=G+1
300  IF A$="T" THEN T=T+1
310  NEXT I
320  A=A*X:C=C*X:G=G*X:T=T*X
330  CLS
340 PRINT "          A260 DETERMINATIONS FOR OLIGOS OF A GIVEN SEQUENCE"
350 PRINT
360 PRINT "The sequence of your oligo is ";B$ " x";X
370 FOR I=1
```

APPENDIX III

Preparation of the Staining Agar

The staining agar as prepared by dissolving 0,12 g of agarose in 20 ml of buffer (0,9 M NaCl; 0,1 M Tris-HCl pH 9,5 and 50 mM $MgCl_2$). The solution was allowed to cool to about 50° C before 90 µl of NBT (75 mg/ml in 70% dimethyl formamide) and 70 µl X-phosphate (50 mg/ml in 100% dimethyl formamide) were added. The solution was mixed, poured and left to solidify.

Author: Kruger Warren Michael.

Name of thesis: Developing DNA profiling strategies for lithops.

PUBLISHER:

University of the Witwatersrand, Johannesburg
©2015

LEGALNOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page)for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.