GENOME ANALYSIS BY TWO-DIMENSIONAL DNA TYPING

ANDRÉ UITTERLINDEN

GENOME ANALYSIS
BY
TWO-DIMENSIONAL DNA TYPING

This thesis will also be available as a publication of INGENY B.V.

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Uitterlinden, Andreas Gerardus

Genome analysis by two-dimensional DNA typing /
Andreas Gerardus Uitterlinden. - [S.l. : s.n.] (Meppel : Krips Repro). -Ill.
Thesis Leiden. -With ref. - With summary in Dutch.
ISBN 90-9005934-2
NUGI 743
Subject headings: DNA / genetics

STELLINGEN

behorende bij het proefschrift
"Genome analysis by two-dimensional DNA typing"

A.G. Uitterlinden
11 Maart 1993

1. Parallelle multilocus polymorfisme analyse van een genoom is efficiënter dan seriële monolocus polymorfisme analyse op grond van het veel kleinere aantal te verrichten handelingen per genoomanalyse.

2. Polymorfe variaties in het aantal repeterende eenheden van tandem gerepeteerde DNA basepaarvolgorden dienen met Variable Number of Tandem Repeats (VNTR) te worden aangeduid, onafhankelijk van de grootte van de repeterende eenheid.

3. Het genereren van een DNA-fingerprint met "core probes" is in geringe mate afhankelijk van de basepaarvolgorde van de repeterende eenheid in de core probe en sterk afhankelijk van de hybridisatieomstandigheden.

4. Het gebruik van polymorfe "anonieme" DNA basepaarvolgorden in het opsporen van genetische afwijkingen zal vervangen worden door het direct analyseren van genpolymorfismen.

5. Pas wanneer het geheel aan erfelijke informatie van organismen in kaart is gebracht kan de gerontologie als quantitatieve wetenschap haar intrede doen.

6. De hoge mate van inteelt in de Nederlandse rundveestapel maakt intensievere toepassing van DNA-diagnostiek in de fokkerij noodzakelijk.

7. Het voorschrijven van "Il Principe" van Machiavelli als handvest voor het bedrijven van wetenschappelijk onderzoek gaat voorbij aan de historische achtergrond waarin dit epistel in 1524 is opgesteld en is een geringschatting van het synergisme van het collectief.

8. De gronden waarop tot de toelating tot de promotie wordt besloten dienen aan internationaal geldende normen te voldoen.

9. De meest geschikte staatkundige organisatie van Europa is een federatie van bevolkingsgroepen die een taalkundige en culturele eenheid vormen.

10. Privatisering van het openbaar vervoer bevordert het gebruik ervan.

11. Het aan het gezicht van passerende automobilisten onttrekken van de plaatselijke nasleep van verkeersongelukken voorkomt filevorming.

GENOME ANALYSIS BY TWO-DIMENSIONAL DNA TYPING

PROEFSCHRIFT

ter verkrijging van de graad van Doctor
aan de Rijksuniversiteit te Leiden,
op gezag van de Rector Magnificus
Dr. L. Leertouwer,
hoogleraar in de faculteit der
Godgeleerdheid,
volgens besluit van het college van dekanen
te verdedigen op donderdag 11 maart 1993
te klokke 15.15 uur

door

ANDREAS GERARDUS UITTERLINDEN

geboren te Maassluis in 1958

PROMOTIECOMMISSIE

Promotor:	Prof. Dr. Ir. P.H.M. Lohman
Co-promotor:	Dr. J. Vijg (INGENY B.V. en Harvard Medical School, Boston, U.S.A.)
Referent:	Prof. Dr. R.M. Myers (University of California, San Francisco, U.S.A.)
Overige Leden:	Prof. Dr. C.H.C.M. Buys (Rijksuniversiteit Groningen) Prof. Dr. G.J.B. van Ommen Prof. Dr. Ir. A.A. van Zeeland

The research presented in this thesis was performed at the TNO Institute for Experimental Gerontology (Leiden, Netherlands), Lifecodes Corp. (Elmsford/Valhalla, NY, USA) and at INGENY B.V. (Leiden, Netherlands). Financial support for part of the work was received from the International Life Sciences Institute, TOYOBO Corp. Ltd. and INGENY B.V.

Aan mijn ouders

CONTENTS

SECTION I: GENERAL OVERVIEW

SECTION II: ADDENDA

SECTION I: GENERAL OVERVIEW

CHAPTER 1

GENERAL INTRODUCTION: ANALYZING THE HIGHER EUKARYOTIC GENOME FOR DNA SEQUENCE VARIATION

1.1 INTRODUCTION

The phenotype of an organism is ultimately encoded by the sequence of basepairs in the set of DNA molecules that constitute the genome of each species. Although integrity of the basepair sequence is essential to normal development and functioning of the organism this information infrequently undergoes changes each time a cell divides. Errors in the DNA information content occur due to physical or chemical damage, either directly or after aberrant enzymatic processing during replication and repair. Such errors are termed "mutations" and range from rearrangements, sometimes involving large chromosome segments, down to the change of a single basepair. Some mutations will affect the encoded function of the DNA sequence and hence distort the metabolism of the cell. These distortions can have mild effects on the functioning of the cell but occasionally pathological changes, including lethal abnormalities, will result.

Mutations can be stably incorporated in the genome, thereby exerting influence throughout the lifespan of the cell. When they occur in somatic cells they can only affect the individual itself, for example by causing cancer and other age-related deteriorations (Vijg and Gossen, 1993). When they occur in the germ cells, mutations are transmitted to progeny and are responsible for the diversity of phenotypes present in a population of organisms. At a larger time scale mutations are fundamental to the evolution of a species and to the diversification of life forms.

At the basis of this diversification is the length and complexity of the sequence of basepairs in the genome. Small genomes are found in organisms with a relatively simple

TABLE 1 HAPLOID GENOME SIZES OF VARIOUS ORGANISMS.

ORGANISM	GENOME SIZE (bp)
Mycoplasma capricolum	1.0×10^6
Escherichia coli	4.7×10^6
Bacillus amyloliquefaciens	3.0×10^6
Saccharomyces cerevisiae	1.5×10^7
Aspergillus niger	1.5×10^7
Drosophila melanogaster	1.8×10^8
Gallus domesticus	1.2×10^9
Rattus norvegicus	2.9×10^9
Lycopersicum peruvianum	3.0×10^9
Xenopus laevis	3.1×10^9
Bos taurus	3.2×10^9
Homo sapiens	3.4×10^9
Nicotiana tabaccum	3.8×10^9
Allium cepa	1.8×10^{10}
Lilium formosanum	3.6×10^{10}
Amoeba dubia	6.7×10^{11}
Mitochondrial genomes	
Homo sapiens	16,569
Viral genomes	
Bovine Herpes Virus type I	138,000
Bacteriophage Lambda	48,321

organization, such as viruses and bacteria, while longer sequences are found in more complex organisms (Table 1).

For each species the sequence of the DNA molecules is more or less fixed but, as described above, variations at particular positions can occur. The highly specific way in which the genome of each species is organized indicates that only a limited number of nucleotides can be variant. However, because of the considerable size of genomes the task of identifying the variant nucleotides among the multitude of possible positions in such genomes is not trivial. The detailed analysis of genomes is crucial to understand every aspect of the biology of a species. Such an analysis encompasses the delineation of the DNA sequence of the genome, the identification of the variant forms of the sequence, and

the deduction of causal relationships between sequence diversity and observed phenotypes. Defining variant forms of particular DNA sequences is of importance, e.g., to find disease genes, to understand the mechanism of disease, to define mutagenic substances.

1.2 GENOME ANALYSIS

In the past two decades the techniques to analyse DNA sequences have evolved to a point where "genome programs" are now being executed, such as for the human genome, instead of only being contemplated. In genome programs DNA sequence data on the genome of a species of interest are being generated and collected in a coordinated fashion through international collaborations of laboratories.

The human genome project was initiated as a result of a meeting of specialists in the field of human mutation analysis (Mendelsohn et al., 1984). They were addressing the question of how the human mutation rate could best be measured. It was concluded that only when the complete sequence of the human genome was known, measurements at different loci with existing or new mutation detection techniques could provide the most accurate estimates. The perspective of having available the complete human genome sequence made this project transgress the boundaries of human mutation analysis and offers the possibility to unravel the complete molecular biology of Homo sapiens. Due to the previous discovery of DNA polymorphisms (Kan and Dozy, 1978) and their application in linkage analysis of genetic diseases (Botstein et al., 1980), the Human Genome Project found support in different areas of biology. Due to its dimensions and implications, this project has since then quickly gained attention of the public and national parliaments as can be seen from its still increasing budget in all developed countries.

Genome programs have been initiated and already started to bear fruit for a few model organisms having small genomes such as Caenorhabditis (Coulston et al., 1992) and Saccharomyces (Oliver et al., 1992). For the human genome there is still no well coordinated task force to sequence particular parts of the genome. Attention has focussed on the construction of genetic maps of each of the chromosomes and there is discussion on how and what to sequence.

In human genetic research some very detailed knowledge has been gathered on particular loci in the genome in the past 10 years due to the occurrence of hereditary

defects. In the course of the search for the causative DNA sequence variant(s) many new techniques were developed to analyse DNA sequences. However, the techniques to determine DNA sequences and DNA sequence variation are nowadays still limited in their scope and capacity. They can analyse only a small part of the complete sequence of a species' genome at a time and the generation of DNA sequence data is still slow. Most of the initial attention in the human genome program is therefore focussed at the development of techniques which allow more DNA sequence data to be generated simultaneously and analysed in a quicker and cheaper way.

Starting with a complete genome of a higher organism, DNA sequence and sequence organization should be generated in a stratified manner, that is to order fragments of ever smaller size in a linear fashion by using small known sequences as landmarks. An example is shown in Figure 1. This allows for a coordinated approach in which there is room to incorporate desirable or existing efforts to obtain DNA sequence data from selected genomic areas (e.g., those harbouring disease genes).

DNA sequence variation can be studied on a population level to reveal the basis of phenotypic variations observed in different individuals. The most striking variations in this respect are disease-causing gene variants. Attempts to identify such gene variants are time-consuming in view of the still large gap in our knowledge of the human genome in terms of DNA sequence and organization. Especially when there is no a priori knowledge on the biochemical basis of the defect and therefore no clue as to what protein might be involved, these quests encompass little less than looking for a needle in a haystack without knowing what the needle looks like.

The discovery of DNA sequence variants which are more or less randomly dispersed over the genome gave a new impetus to what is known as genetic linkage analysis. In genetic linkage analysis the transmission of DNA sequence variants with a known location on a chromosome (referred to as genetic markers) is followed in pedigrees and evaluated for the coupled inheritance of one or more markers with a particular trait such as a disease. If such coupling or linkage is observed more frequently than statistically expected, the marker and disease gene loci are said to be linked. According to the Mendelian law of independent assortment, variants of two loci have a 50% chance of being inherited together. If the two loci are in close proximity on the same chromosome this frequency increases. The more frequently this genetic linkage is observed the closer

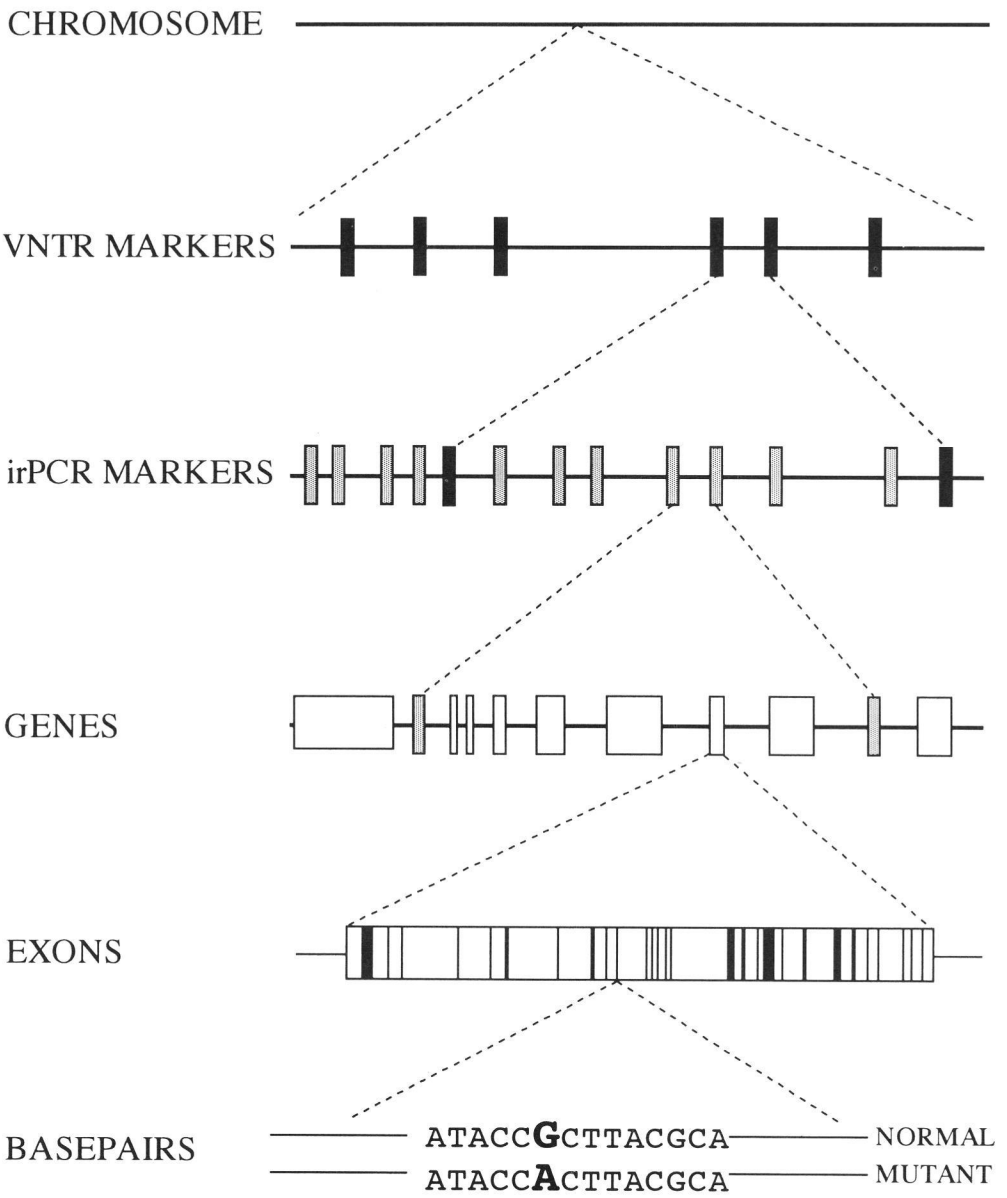


Figure 1 *Schematic representation of a stratified manner in which the organization of a genome can be analysed using known sequences as landmarks.*

the two loci are on the chromosome. Sometimes such linkage is disturbed by the occurrence of recombinations of homologous chromosomes during meiosis (crossing-over) between the marker locus and the disease locus. The chance of such a recombination to occur, is lower for a smaller distance between the disease locus and the marker locus. This phenomenon is exploited to measure the "genetic distance" between the two loci: if no or very few recombinations are observed in pedigrees segregating for the disease it can be assumed that the marker locus and the disease locus are very close to each other on the chromosome. Scoring of such recombination events between a set of defined markers and the phenotype of interest is used to define the chromosomal area harbouring a particular gene of interest.

Another application in which genetic polymorphisms spread over the genome can help to identify the position of genes of interest is the comparison of tumor genomes with their normal counterparts. Due to the frequent occurrence of deletions and amplifications of particular areas in tumor genomes, DNA polymorphisms can identify such areas by observing loss of heterozygosity (LOH) and/or amplification for particular loci. If LOHs or amplifications occur very frequently in tumors from different patients these chromosomal loci are then likely to contain genes contributing to tumor development and/or progression (Vogelstein et al., 1985). Similarly, the scanning of genomes for DNA sequence variation can be used to identify random variations that might occur during the aging process (Vijg and Uitterlinden, 1987).

The many DNA sequence variations spread over the genome, have been exploited for the construction of maps of genetic markers which are referred to as genetic maps. The accuracy and effectiveness in establishing the location of a particular (disease) gene on a chromosome is in part determined by the number of markers on such a map. If large numbers of markers are available, the chance of finding close linkage between a marker and the disease locus of interest will be high. The currently available marker systems are locus-specific, i.e. they each detect DNA sequence variation at only a single site in the genome. An example is a genetic map of the human genome based on a collection of over 800 $(GT)_n$ markers (Chumakov et al., 1992). If a genome is to be scanned for variation this will involve the sequential screening of many markers and, hence, be time- and labour consuming. It is more effective to scan the genome at many sites simultaneously. Such a parallel processing approach to the analysis of DNA sequence variation can consist

of electrophoretic separation of DNA molecules in two dimensions and the subsequent transfer of separation patterns to membranes for hybridization analysis.

1.3 OUTLINE OF THE THESIS

This thesis describes the development of a parallel processing technique, two-dimensional (2-D) DNA typing, for analysing DNA sequence variation. It is based on a 2-D electrophoresis principle developed by Fischer and Lerman (1979) for separating DNA fragments. By combining the 2-D separation of DNA fragments with their subsequent hybridization analysis using repetitive sequences, 2-D displays of DNA sequences derived from areas spread over the genome allow for an increase in the power of analysis as compared to standard one-dimensional separation patterns such as in Souther blot hybridization (Southern, 1975). Two-dimensional DNA typing allows many (up to several hundreds) different sites in a genome to be assessed simultaneously for DNA sequence variation in a single hybridization experiment. Additional hundreds of sites can be assessed for variation by re-hybridizing the membrane with other probes. The resolution obtained depends on the electrophoretic separation principles used to generate the two-dimensional separation pattern, and on the type of probe which is used to detect particular sequences in the genome of interest. The publications included in this thesis (addenda 1 - 4) relate to both these aspects. Addendum 1 contains an example of how the electrophoretic separation principle, Denaturing Gradient Gel Electrophoresis (DGGE), used as the second dimension in the 2-D DNA typing procedure is used to detect sequence variations. Addendum 2 describes the results obtained with the 2-D DNA typing method as applied to the analysis of human genomic DNA. Addendum 3 relates to an important aspect of the micro- and minisatellite repetitive sequences detected in the 2-D DNA typing procedure; their GC-richness determines their suitability for generating 2-D spot patterns of sufficient resolution. Addendum 4 describes one of the potential applications of the 2-D DNA typing method to a biological problem. The latter study illustrates the use of different types of probes in one- and two-dimensional DNA typing and demonstrates the feasibility of analysing cells treated with genotoxic agents for DNA sequence changes.

First, in Chapters 2 and 3 an overview is presented of ways of detecting DNA

sequence variation and their application in the analysis of genomes. The concept of "genome scanning" based on the use of repetitive elements is discussed in Chapter 4. Finally, a discussion on the 2-D DNA typing technique is presented including some of the results which have been obtained with this method in the analysis of biological problems which can be addressed by applying genome scanning techniques (Chapter 5).

CHAPTER 2

MEASURING DNA SEQUENCE VARIATION

2.1 INTRODUCTION

When analysing DNA sequence variation a fundamental distinction should be made between the analysis of known vs. unknown DNA sequence variations. Once a sequence variant is known, one can apply a number of recently developed methods for its rapid and accurate detection. Some of these methods are well suited for the simultaneous analysis of large numbers of samples for one particular mutation and are referred to as "mutation screening" methods. They are, however, generally not capable of assaying a given DNA region for unknown mutations. These techniques, therefore, fall beyond the scope of this thesis and shall not be further discussed. Other methods that, to some extent, allow mutation scanning are forward mutation assays, such as the hypoxanthine phosphoribosyl-transferase (HPRT) method (Albertini et al., 1982). Yet, in order to obtain information on the type and position of the mutation, isolation of the mutant gene is imperative. In both mutation screening and mutation selection methods the use of PCR amplification (Mullis et al., 1987; Saiki et al., 1985) has become indispensible for the detection of particular types of sequence variants. The only disadvantages of PCR in this respect are its sensitivity for contaminants and its relatrively high error rate.

When within a given DNA region previously undefined DNA sequence variations should be found, techniques have to be employed which allow a considerable number of basepairs to be simultaneously analysed or "scanned" for variation. Table 2 lists the presently available DNA scanning methods with their mode of action and their sensitivity in terms of the number of basepairs that can be accurately scanned for variation. Below, the suitability of these methods for serial genome scanning will be discussed.

2.1 DNA-SEQUENCING

The ultimate identification of DNA sequence variation is the comparative determination of the basepair sequence of a wild type molecule and its variants. In

TABLE 2 EFFICIENCIES OF "MUTATION SCANNING" METHODS

METHOD	#BP SCREENED PER LANE	SIZE RANGE TEST FRAGMENT (BP)	EFFICIENCY[1]
DNA SEQUENCING	400	10-400	100 %
RFLP	10-20	500-10000	≤1 %
RNase A	250-500	100-1000	≤40 %
CHEMICAL CLEAVAGE	50-1000	50-1000	≤99 %[2]
DGGE	250-500	100-1000	
unclamped:			
-homoduplex			≤40 %
-heteroduplex			≤50 %
clamped:			
-homoduplex			≤75 %
-heteroduplex			≤99 %
SSCP	250-300	50-500	
-DNA SSCP			≤40 %[2]
-RNA SSCP			≤75 %[2]
HETERODUPLEX MOBILITY	50-1000	100-1000	?[3]

[1] % of all possible base changes detectable in a given fragment
[2] data based on the analysis of a limited number of sequences
[3] no data available yet

essence, DNA sequencing involves the generation of end-labelled fragments of different size which are electrophoretically separated in a polyacrylamide gel. Since 1977 twomethods have been developed and refined to sequence DNA fragments of up to 500 bp. They are based on nucleotide-specific chemical cleavage (hydrazine for T and C, piperidine for A and G; Maxam and Gilbert, 1977; Gilbert, 1982) or the use of chain-terminating inhibitors of DNA polymerases (dideoxynucleotides; Sanger et al., 1977). In the chemical sequencing protocol the differently sized fragments find their origin in reaction conditions that permit on average only one break per molecule at any of the nucleotides for which the chemical is specific. In the enzymatic sequencing protocol the fragments arise due to particular ratios of natural dNTPs to any of the chain-terminating analogs which halt the polymerase at different positions along the molecule. In both protocols the rungs of the ladder of bands which arise after size separation of the fragments in polyacrylamide gel electrophoresis indicate the position of particular nucleotides in the DNA molecule of interest.

Almost all DNA sequencing reactions nowadays are done according to the enzymatic protocol of Sanger, mostly due to its simplicity and accuracy. The protocol has undergone considerable refinement by using modified polymerases such as T7 polymerase or Sequenase (Tabor and Richardson, 1987) and by using fluorescent dideoxynucleotides. The latter are used for non-radioactive DNA sequencing where the detection is based on laser excitation using specially designed instruments so that the whole process of DNA sequencing is semiautomated. Further improvements include the use of the PCR for amplification of the fragment of interest making cloning of the fragment of interest obsolete. These direct DNA sequencing protocols for determining DNA sequence variants (Yandell and Dryja, 1989) have been applied in, e.g., the analysis of Lesch-Nyhan mutations in the HPRT gene (Gibbs et al., 1989), the characterization of mutations in the Factor VIII and IX gene (Higuchi et al., 1990; Koeberl et al., 1990).

DNA sequencing provides the ultimate tool in the detection of DNA sequence variation. However, for measuring variation in large DNA regions ($\geq$ 100,000 bp) it is necessary to first subdivide the region in smaller sections of about 400 bp. Hence, large numbers of sequencing reactions must be performed on each sample. To become applicable for screening purposes, DNA sequencing requires dramatic improvements in speed. Although more rapid methods for sequencing are presently under development

(Hood, personal communication), high speed at low costs can, most likely, only be achieved by methods other than electrophoretic separation. Such methods, e.g. parallel sequencing by hybridization (Drmanac et al., 1989; Strezoska et al., 1991), are currently still under development but hold promise as fast and cheap sequencing methodologies.

From a conceptual point of view DNA sequencing for assaying interindividual variation is unattractive since it provides much unnecessary information; by far the largest part of the sequence to be comparably analysed is identical. It is therefore advantageous to have methods available that only and uniquely identify deviations from a particular wild type sequence.

2.2 RESTRICTION ENZYME DIGESTION: RFLP ANALYSIS

The very start of recombinant DNA research was the discovery of enzymes capable of recognizing and cutting a particular short sequence of basepairs (usually 4-8 basepairs long) in a double stranded DNA molecule, thereby generating staggered or opposite double strand breaks in the DNA molecule. The specificity of these enzymes with respect to the particular basepair sequence recognized, appeared to be determined by the bacterial species from which they were isolated. In the bacteria themselves these enzymes serve as part of a host-defence (host-restriction) system to safeguard the bacteria from intruding non-self DNA molecules from other strains or bacteriophages.

Restriction enzymes were subsequently used in a comparative way, by digesting near-similar DNA molecules in parallel and analysing the different restriction fragments by size after electrophoretic separation. A particular difference in sequence between the 2 or more original DNA molecules is revealed by differences in the length of the fragments, provided the difference is in the recognition site of the enzyme, thereby blocking its activity. This principle has revolutionized genetics when it was combined with so-called Southern hybridization for the analysis of genomes of higher organisms (Southern, 1975, 1979). In this technique the electrophoretic separation patterns of restriction enzyme digests of the DNA molecules of interest are transferred onto a nitrocellulose or nylon (DNA binding) membrane. By hybridization analysis using a radio-actively labelled probe, specific for a particular genomic sequence, homologous fragments within the separation pattern are visualized after autoradiography.

Southern blot hybridization analysis of restriction enzyme digested genomic DNA has been instrumental in the detection of DNA sequence variations, referred to as restriction enzyme length polymorphisms or RFLPs (Flavell et al., 1978; Kan and Dozy, 1978; Jeffreys, 1979). This source of polymorphism has been shown to be very rich even to an extent where it is possible to construct genetic maps of chromosomes based on RFLPs (Botstein et al., 1980; White et al., 1985). The chances of detecting unknown DNA sequence variation purely on the basis of altered restriction enzyme recognition sites are quite slim. Nevertheless, DNA sequence variations have been detected in this way (e.g. Greever et al., 1981; Maddalena et al., 1988; Carothers et al., 1988). Based on such analyses the human genome was estimated to contain 1 in 300 (autosomes) to 1 in 1000 (X chromosome) polymorphic nucleotides (Barker et al., 1984; Hofker et al., 1989). The method was also the first to be used to detect a disease causing point mutation for sickle cell anemia (Chang and Kan, 1982; Orkin et al., 1982). At present it is still used in combination with other techniques, like PCR, for the detection of particular mutations, such as the sickle cell mutation in human DNA (Saiki et al., 1985) and the bovine leukocyte adhesion deficiency (BLAD) mutation in cattle DNA (Shuster et al., 1992).

Although the mutation scanning capacity of RFLP analysis by itself is low, restriction enzyme digestion is useful for the generation of fragments which can be further separated by electrophoretic methods described below. In addition, it can be used in two-dimensional separations in which a second restriction enzyme digestion is performed after the first size separation and followed by a second size separation.

2.3 ANALYSIS OF BASEPAIRING

The structural basis of the double-stranded DNA molecules as discovered by Watson and Crick (1953a and 1953b) is their sequence-dependent basepairing and stacking characteristics: A can only perfectly pair with T, and C only with G. Deviations from this Watson-Crick type of basepairing can form the basis of the detection of DNA sequence variation within a particular sequence. Such sequence-specific deviations from perfect basepairing can be detected as changes in hybridization characteristics of a probe to a target DNA molecule, changes in melting (strand separation) behaviour of a double-stranded molecule or as vulnerability of heteroduplexed DNA molecules to react with

nucleases or particular chemicals.

2.3.1 Hybridization analysis

This method for detecting sequence variation is based upon whether or not there is formation of a heteroduplex molecule (a double-stranded DNA molecule in which there is incomplete basepairing due to a mismatch at one or more positions in the sequence) between a labelled nucleic acid probe and a target under certain hybridization conditions. In the usual format of this assay the target sequence is made single-stranded by denaturation, after which a single-stranded labelled probe molecule is added. If there are more than a certain number of basepair mismatches within the heteroduplex the probe will not bind and hence no signal will remain after removal of unbound probe. This way of detecting sequence variations using polynucleotides is very insensitive (typically more than 70% homology in basepair sequence between probe and target is sufficient for probe hybridization). However, it can for instance be applied for so-called "zoo blot" hybridization analysis where, in the search for homologous DNA sequences in different species, large deviations from the wild type DNA sequence can be expected.

The hybridization principle was later employed in refined form by using oligonucleotides to detect the presence or absence of point mutations (Wallace et al., 1979 and 1981). In this method two oligonucleotides are synthesized, complementary to either the wild type sequence or to a particular mutant sequence, labelled and used as hybridization probes on Southern blots or dried gels of genomic DNA separation patterns (Conner et al., 1983; Piratsu et al., 1983) or PCR-amplified sequences spotted on filter (Saiki et al., 1986, 1988). Examples of mutations detected through this method include Hex A mutations (Paw et al., 1990) and lipoprotein lipase gene mutants (Emi et al., 1990).

Because of the exquisite sequence-specificity of the hybridization reaction only one oligonucleotide can normally be applied. Tetra-alkylammonium salts however, can reduce the dependency of the melting temperature on the base composition (Melchior and Von Hippel, 1973; Wetmur et al., 1981). In this way identical hybridization conditions can be applied for several different oligonucleotides in the so-called reverse dot blot assay (Saiki et al., 1989). This can be especially advantageous when screening genes for several different mutations. It is expected, however, that this will only be possible for the

detection of no more than 30 different point mutations in a single hybridization analysis. As such the method lends itself well for mutation screening purposes, that is, to look for the presence of a given (set of) point mutation(s) at a particular location, rather than for scanning areas of the genome for unknown DNA sequence variations.

One particular application of hybridization analysis using oligonucleotides involves the use of very short oligonucleotides for mapping of ordered libraries of cloned sequences (Michiels et al., 1989; Lehrach et al., 1990; Lennon and Lehrach, 1991) and for sequencing (Drmanac et al., 1989). In these applications the redundancy of short sequence motifs in the genome is exploited to an extent where information on order of fragments and even the actual basepair sequence can be reconstructed on the basis of a "hybridization signature" of a particular set of clones using a (large) collection of motifs. In this respect, such an approach can be extended to comparison of genomes in which slight deviations of a wild type sequence can be detected as aberrations in the hybridization signatures of the respective sets of clones (Lehrach et al., 1990).

2.3.2 Solution melting of heteroduplex molecules

The solution melting method (Smith et al., 1988) exploits the phenomenon that strand separation of duplex molecules, as caused by increased temperature, is sequence dependent and occurs in so-called melting domains (see the section on DGGE). In this technique the target RNA or DNA fragment is solution-hybridized to an RNA probe, and the resulting RNA-RNA or RNA-DNA heteroduplexes are subsequently exposed first to tetraethylammonium bromide, which destabilizes dsDNA and RNA molecules and facilitates helix-coil transitions, and subsequently to a step-wise formamide gradient in different tubes. At particular formamide concentrations the highest melting domain in the duplex molecule will completely dissociate and single stranded fragments will arise. These resulting fragments are then analysed by polyacrylamide gel electrophoresis to distinguish between single-stranded (slow migrating) and double-stranded (fast migrating) RNA or DNA molecules. Sequence variants of the highest melting domain in the fragment of interest can be discerned by the different formamide concentrations at which they will display complete strand dissociation. Under appropriate conditions even single base substitutions can be detected. The method has the advantage that sequence variants in high

melting domains can be detected (of up to 140 bp fragments for RNA-DNA heteroduplexes and up to 260 bp fragments for RNA-RNA duplexes) but considerable ambiguity has been observed in distinguishing homozygotes from heterozygotes, thereby effectively limiting the analysis to X-linked genes (Latham and Smith, 1989).

2.3.3 Basepair mismatch reactivity of heteroduplex molecules

The aberrant conformation of a nucleic acid heteroduplex molecule at the site of a basepair mismatch can be the target for either enzymatic or chemical reactions. These can form the basis of methods for finding DNA sequence variants through the formation of heteroduplexes. One distinguishes between enzymatic reactions, based on S_1 nuclease and RNase A digestion, and chemical reactions, based on carbodiimide modification and chemical cleavage of mismatched basepairs.

2.3.3.1 S_1 nuclease digestion

S_1 nuclease recognizes and digests single-stranded DNA and has also been shown to recognize single-stranded DNA regions in heteroduplex RNA-DNA and DNA-DNA molecules (Shenk et al., 1975). Although sensitivity to single base mismatches has been reported (Shenk et al., 1975), in general, only mismatches encompassing at least 4 basepairs can be reliably detected (Dickson et al., 1984; Chebloune et al., 1984).

2.3.3.2 RNase A digestion

Similar to S_1-nuclease, ribonuclease A (RNase A) has been shown to recognize single-stranded RNA in RNA:DNA duplexes (Myers et al., 1985e and 1989) and in RNA:RNA duplexes (Winter et al., 1985). The method involves the generation of radiolabelled single stranded RNA probes and their subsequent hybridization to a single stranded (denatured) DNA target, followed by digestion by RNase A. The method is sensitive to single base mismatches but some of these, depending on the sequence context of the mismatch, have been shown to be refractory to RNase A cleavage (Myers et al., 1985e). By analysing each of the 2 strands of a particular DNA fragment with the their correspon-

ding RNA probes (of opposite sense) in separate RNAse A reactions, it is possible to detect up to 70% of all possible sequence changes in the DNA fragment. It appears, however, that the presence of purines at the mismatch prevents such mismatches from being cleaved (Cotton, 1993). The method has been successfully used, e.g. for the detection of ß-globin mutations in human genomic DNA (Myers et al., 1985e), cKi-ras variants in RNA from tumour cell lines (Winter et al., 1985), and previously unknown HPRT mutations in Lesch-Nyhan patients (Gibbs and Caskey, 1987) and an unknown OTC mutation in mouse genomic DNA (Veres et al., 1987).

2.3.3.3 Carbodiimide modification

The reactivity of mismatched basepairs with carbodiimide was initially exploited in a mobility shift assay in polyacrylamide gels (Novack et al., 1986). In this method heteroduplexed DNA molecules are allowed to react with carbodiimide, which specifically reacts with unpaired G and T residues and results in the presence of a tag at the site of the mismatch. Sequence variants can then be detected by means of their altered gel electrophoretic mobility caused by the presence of the carbodiimide tag.

More recently, this tagging system of heteroduplexed molecules was combined with the application of antibodies specific for carbodiimide and used to detect mutations in very large fragments (up to 7 kb) by immunomicroscopy (Ganguly et al., 1989). Another variation of the method involves the use of heteroduplexed and carbodiimide-tagged DNA molecules in a subsequent PCR reaction under conditions in which primer extension terminates at the site of the modified bases. This will lead to the production of additional shorter PCR products (Ganguly and Prockop, 1990). This method has been used to detect a new mutation in the human type I procollagen (COLIAI) in a osteogenesis imperfecta patient (Zhuang et al., 1991).

2.3.3.4 Chemical cleavage of basepair mismatches (CCM)

The base-specific reaction of particular chemicals with DNA molecules has been first exploited in DNA sequencing (Maxam and Gilbert, 1977; see section 2.1). This prin-

ciple, however, can also be used to establish the presence of sequence variations in heteroduplexed molecules which contain basepair mismatches. Two reagents, osmium tetroxide and hydroxylamine, were found to modify mispaired T and C, respectively. When followed by piperidine-catalysed cleavage, the size of the resulting fragments can give positional information on the mismatch (Cotton, et al., 1988; Cotton and Campbell, 1989). Although particular mismatches have been shown to be cleaved less efficiently, analysis with probes of opposite sense allows nearly 100% of possible mismatches to be detected (Dianzani et al., 1991; Forrest et al., 1991). The method has been successfully applied in the detection of sequence variations in human genes, e.g. the factor IX gene (Montandon, et al., 1989), the ß-globin gene (Dianzani et al., 1991a), the phenylalanine hydroxylase (PAH) gene (Dianzani et al., 1991b), the ß-hexosaminidase gene (Akli et al., 1991), the OTC gene (Grompe et al., 1991) and the dystrophin gene (Roberts et al., 1992).

2.3.4 Gel electrophoretic mobility assays

Many of the above mentioned methods involve the electrophoretic size-separation of DNA fragments as a secondary means of identifying DNA sequence variation. Below, electrophoretic separation methods will be discussed based on different separation criteria and exploiting sequence-dependent variations in base-pairing which cause differences in conformation and lead to altered electrophoretic mobilities. In DGGE and TGGE, the induced local loss of inter-strand basepairing leads to the arisal of molecules with a conformation which is much different from the normal worm-like rods. Because the migration of DNA molecules in polyacrylamide gels is dependent on their conformation, this in turn determines the position in the gel where DNA molecules undergo severe mobility retardation.

At increased temperature, differences in melting behaviour of duplex molecules can form the basis of separation of DNA sequence variants by gel electrophoresis. Such separations are based on the formation, before or during electrophoresis, of conformation differences between the sequence variants, which will lead to different electrophoretic migratory behaviour of the variants. In this way melting (strand-separation) characteristics of double-stranded DNA molecules, DNA-RNA or RNA-RNA hybrid molecules can be

assessed. Thermosensitive gel electrophoretic separations are based upon the use of temperature gradients applied in the gel, either as a true temperature gradient (Temperature Gradient Gel Electrophoresis: TGGE) or as a gradient of denaturants like urea and formamide (Denaturing Gradient Gel Electrophoresis: DGGE).

In the other techniques altered electrophoretic mobility as such is measured, which results from differences in intra-strand basepairing (as in SSCP), or by the presence of a basepair mismatch (heteroduplex analysis). Gel electrophoretic separation of homoduplex double-stranded DNA molecules under neutral conditions shows very little, if any, sensitivity to the base sequence of the molecules but rather to their length. Under special electrophoretic conditions the analysis of either single-stranded DNA molecules or heteroduplexed DNA molecules can provide the means to detect deviations from a wild type sequence. Examples of the direct dependency of electrophoretic mobility on DNA conformation differences are SSCP and the analysis of heteroduplex molecules.

2.3.4.1 DGGE: Denaturing gradient gel electrophoresis

Principles

DGGE involves separation in a polyacrylamide gel containing a gradient of denaturants (i.e. urea and formamide) at a fixed and elevated temperature (Fischer and Lerman, 1979a and 1979b). The double-stranded DNA fragments will, at first, travel through the gel according to their size, until denaturant concentrations are encountered inducing the strands to leave their helical state and dissociate or melt. The strand dissociation is length-independent (Fischer and Lerman, 1979a, 1980) and highly sequence-dependent, due to differences in basepairing and the influence of neighbouring bases on stacking forces in the double helical chain (Fischer and Lerman, 1979a, 1980 and 1983). Once part of the molecule is melted, the fragment undergoes a change in conformation and a concomitant reduction in electrophoretic mobility (for reviews, see Lerman et al., 1983, 1984 and 1986).

Melting of DNA fragments is not a gradual continuous process but proceeds in discrete steps, due to the presence of so-called melting domains: stretches of roughly 50-400 basepairs with nearly identical melting temperature. Strand separation of the domain with the lowest melting temperature (the lowest- or first-melting domain) initiates the

mobility drop. This migratory reduction, however, is critically dependent on the presence of a juxtaposed higher melting domain (referred to as a "clamp"), the strands of which are still helical at the melting temperature of the first melting domain. At that particular concentration of denaturants a branched molecule will arise with a much lower mobility in the polyacrylamide matrix. Therefore, at a certain position in the denaturing gradient, migration of a particular fragment in the DGGE gel is no longer size-dependent but is dictated by the basepair-sequence of the first melting domain. For calculation of the melting temperature of a particular basepair sequence a computer algorithm is available, developed by Leonard S. Lerman (Fischer and Lerman, 1979; Lerman et al., 1984 and 1986). Based on thermodynamical stability values of a given basepair with its neighbouring basepairs, the computer program can calculate the $T_m(50)$ for each basepair in a particular sequence. The $T_m(50)$ is the temperature at which the melting domain is completely single-stranded in 50 % of the fragments. Since particular concentrations of denaturants in the DGGE gel correspond to particular temperatures, $T_m(50)$ positions of particular fragments in the gel can be calculated using parameters such as buffer concentration, run time, etc. The computer program has been shown to be quite accurate in predicting gel position of particular fragments and their sequence variants, even when they differ by only a single point mutation including A→T and G→C substitutions (Myers et al., 1985a and 1985b; Lerman et al., 1986; Abrams et al., 1991). In reverse, fragments observed at a certain position in the gel can be characterized by a certain $T_m(50)$. This feature, however, has not been investigated to an extent allowing sequence variations to be deduced from particular positions of a given fragment in a DGGE gel.

Thus, the sequence-dependent separation in DGGE gels is based mainly on three characteristics of double-stranded DNA molecules: (A) melting (or strand separation) of contiguous bases in DNA molecules is closely coupled to form so-called melting domains; (B) the temperature (or the concentration of denaturants, such as urea and formamide) at which strands of a melting domain part, is highly dependent on the sequence composition of the domain; and (C) partially melted DNA molecules (containing both double-stranded and single-stranded regions) have a decreased electrophoretic mobility in PAA gels in comparison with native, double-stranded molecules.

Analysis of genomic DNA

An important extension of the DGGE system has involved its application to the analysis of total genomic DNA through the use of PCR amplification (Cariello et al., 1988a; Sheffield et al., 1989) or by transfer of separation patterns to hybridization membranes. The latter can be accomplished by capillary blotting using modified gel media (Börresen et al., 1988) or by means of electroblotting. The latter was first described for the analysis of micro- and minisatellite repeat motifs using core probes (Uitterlinden et al., 1989) and locus-specific VNTR probes (Uitterlinden and Vijg, 1991). More recently, electroblotting has been extended to the detection of single-copy DNA sequences in Drosophila (Gray et al., 1991) and in humans (Burmeister et al., 1991; Gray et al., 1992).

GC-clamping

The attachment of a GC-rich sequence (referred to as a "clamp") to a fragment which normally undergoes immediate strand dissociation upon melting has been shown to alter the melting characteristics of the fragment (Lerman et al., 1984; Myers et al., 1985b). The clamp prevents monodomain fragments to completely dissociate and guarantees the adjacent sequence always to be the first melting domain. When a fragment is "clamped", partially melted molecules can arise upon melting and hence allow the comprehensive detection of mutations in the fragment of interest (Myers et al., 1985c). This principle has been applied in the improvement of detection of sequence variations in genomic DNA through PCR with "clamped" primers (Sheffield et al., 1989 and 1992), through selective attachment of a clamp to genomic restriction fragments (Abrams et al., 1990) or by "chemical clamping" involving the chemical modification of the outer basepair to covalently link the opposite bases (Appligene, France).

Analysis of heteroduplex molecules

Although DGGE is capable of detecting single base differences by comparing the naturally occurring homoduplexes, its sensitivity can be even further increased by introducing base pair mismatches as they occur in heteroduplexes (Lerman et al., 1984). This phenomenon has been exploited in the analysis of heteroduplexes formed between genomic DNA restriction fragments and radio-labelled single-stranded DNA probes for

human DNA (Myers et al., 1985a; Noll and Collins, 1987 and Bewsey et al., 1991) and for maize plant DNA (Riedel et al., 1990). The same principle has also been demonstrated for RNA:DNA heteroduplexes in human DNA (Takahashi et al., 1990) and for DNA:RNA and RNA:RNA heteroduplexes in the analysis of mutation rates of retroviral genomes (Smith et al., 1986; Leider et al., 1988).

CDGE

Increased resolution of sequence variants of particular fragments can be obtained by applying so-called constant denaturing gel electrophoresis (CDGE). Here a constant denaturant concentration is applied in the gel which corresponds to a specific melting domain within the fragment of interest. Particular advantages are that increased separation distances are obtained and that the system is more amenable for automation than DGGE. However, disadvantages are the time-dependency of the separation and suitability for analysing only one melting domain. A high sensitivity of CDGE in detecting point mutations was shown for the HPRT gene in mice and hamster DNA (Hovig et al., 1991) and the p53 gene in human tumour DNA (Börresen et al., 1991).

Analysis of complex populations of molecules

A particular advantage of DGGE in comparison to other mutation detection techniques is its potential to simultaneously identify different sequence variants in a complex population of molecules. Such populations can consist of a wild type molecule and mutants present at different abundancies or of a mixture of closely related molecules derived from different organisms.

One important application of DGGE in characterizing the former population has included the detection of sequence variants after treatment of DNA with damaging agents (Myers et al., 1985d; Lillie et al. 1986; Cariello et al., 1988 and 1991a). In most of these schemes DNA fragments, prepared from cells which have been in contact with a mutagen, are amplified by means of PCR and subsequently separated on a DGGE gel which has been optimized for separation of variants from the fragment of choice. This application makes use of unique advantages offered by DGGE (or TGGE for that matter, see below) which is the possibility to (a) isolate the mutant sequences in an intact form, and (b) detect low frequency mutation events in populations of molecules. The results

obtained sofar indicate that the fidelity of the amplification step is limiting the sensitivity, due to errors which are introduced by the DNA polymerase (Cariello et al., 1991b). Enzymes with higher fidelity than the original Taq polymerase have recently been described, such as Sequenase, Vent, and Pfu polymerase. It appears, however, that even with these high-fidelity enzymes a mutation must be present in about 1% of the cells in the original cell population to be detectable (Cariello et al., 1991b). In a similar way Keohavong et al. (1991) used DGGE to detect point mutations in exon 3 of the HPRT gene in mass cultures of mutant cells (i.e. 6-thioguanine resistant) which were exposed to UV light.

A similar application of DGGE analysis in this respect is the analysis of complex populations of individuals, such as those found in bacterial mats. These mats represent ecological niches in which a variety of different organisms are present in different numbers. Such populations can be typed by means of DGGE analysis of their 16 S small subunit ribosomal RNA (16 S ssu-rRNA) genes. For this purpose, a particular region from the 16 S RNA (the V4 region) in which species-specific variation is expected, is PCR amplified from genomic DNA isolated from a particular mat and subjected to DGGE analysis. Based on species-specific sequence variations in these regions, DGGE-fingerprints, consisting of 10-30 bands of different intensity, can be obtained for these populations (Muyzer, G., de Waal, E.C. and Uitterlinden, A.G., in press). In these DGGE-fingerprints each band corresponds to the presence of a particular species in the mat, whereby the intensity is a reflection of its relative abundance in the population.

Gene analysis

Since its introduction in 1979, DGGE has been used extensively and, as a result of its exquisite sensitivity for differences in basepair composition, it has been widely used for the detection of sequence variations as small as point mutations or even base modifications such as methylation (Collins and Myers, 1987). Genes, allelic variants of which have been detected by DGGE include the Adenine PhosphoRibosyl Transferase (APRT) gene in the mouse (Dlouhy et al., 1989), and in human DNA the Hypoxanthine guanine PhospoRibosyl Transferase (HPRT) gene (Cariello et al., 1988), the Coagulation factor IX gene (Attree et al., 1989), the cHa-ras1 proto-oncogene (Uitterlinden and Vijg, 1990), the Ornithine Transcarbamylase (OTC) gene (Finkelstein et al., 1990), the Factor

VIII gene (Kogan and Gitschier, 1990; Traystman et al., 1990; Higuchi et al., 1991), the cystic fibrosis transmembrane conductance regulator (CFTR) gene (Vidaud et al., 1990; Devoto et al., 1991; Ghanem et al., 1992), the protein C gene (Gandrille and Alach, 1991), the rhodopsin gene (Sheffield et al., 1991), the tumor suppressor gene MCC, p53, prealbumin, rhodopsin, S-antigen, and TGF-α (Sheffield et al., 1992), the insulin receptor gene (Krolewski et al., 1992), the low density lipoprotein (LDL) receptor gene (Top et al., 1992), the beta-globin gene (Losekoot et al., 1990; Rosatelli et al., 1992), and the alpha-globin gene (Fodde et al., 1992).

2.3.4.2 TGGE: Temperature gradient gel electrophoresis

In modifications of the above described electrophoretic system, true temperature gradients in combination with a constant urea and/or formamide concentration have been applied instead of using chemical mimics in PAA gels at a constant temperature (Rosenbaum and Riesner, 1987; Wartell et al., 1990). The method was first applied to the detection of sequence changes in short, double-stranded satellite RNA molecules from viruses (Po et al., 1987; Steger et al., 1987). More recently, however, the system was shown to be capable of detecting single basepair substitutions in DNA molecules (Wartell et al., 1990). Although, in general, the results obtained with TGGE are similar to those found with DGGE, there are differences with respect to the migratory behaviour of DNA molecules. More specifically, an increase of mobility of fragments is observed at temperatures slightly higher than the starting temperature, a phenomenon which is not seen in chemical solvent-gradient gels (Wartell et al., 1990).

2.3.4.3 SSCP: single strand conformation analysis

Under non-denaturing conditions, single-stranded DNA has a folded conformation which is stabilized by intra-strand interactions. The sequence will therefore determine the conformation and hence the electrophoretic mobility. This principle has been exploited in the electrophoretic separation of fully denatured genomic DNA on neutral polyacrylamide gels followed by transfer of the separation pattern to a nylon membrane and hybridization analysis with RNA probes for each of the strands of a target sequence (Orita et al.,

1989a). It was shown that sequence alterations as small as a single base substitution could be detected in genomic DNA after blot hybridization (Orita et al., 1989a) and after PCR amplification of the target sequences (Orita et al., 1989b). Electrophoretic conditions have been shown to be very critical in this method and, according to some reports, only 35 % of all point mutations can be detected (Sarkar et al., 1992; Spinardi et al., 1991). Observed variation in mutation detection efficiency seems to be related to the GC-content of the fragments in that GC-rich fragments allow more different mutations to be detected than GC-poor fragments. Similar to the situation with DGGE, SSCP would gain advantage from a theoretical description the influence of intra-strand basepairing on electrophoretic mobility of double-stranded DNA molecules. The SSCP protocol has nevertheless been successfully applied to detect sequence variants of ras proto-oncogene and individual Alu repeats (Orita et al., 1989b; Suzuki et al., 1990), the ß-hexosaminidase gene (Ainsworth et al., 1991), the PAH gene (Labrune et al., 1991), the CFTR cystic fibrosis gene (Dean et al., 1990; Iannuzi et al., 1991; White et al., 1991; Ivaschenko et al., 1991), the neurofibromatosis gene (Cawthon et al., 1990), the lipoprotein lipase gene (Hata et al., 1990), the factor IX gene (Demers et al., 1990), the KIT-proto-oncogene and insulin-like-growth factor 1 receptor (IGF1R) gene (Poduslo et al., 1991), the rhodopsin gene (Dryja et al., 1991), and the p53 gene (Runnebaum et al., 1991).

Modifications of SSCP

In a modification of SSCP analysis RNA conformation polymorphisms have been detected by neutral gel electrophoresis of RNA molecules generated from PCR amplified genomic DNA fragments (Danenberg et al., 1992; Sarkar et al., 1992a). When applied to the scanning of fragments derived from 2.6 kb of the human Factor IX gene, RNA-SSCP was shown to detect a higher fraction of point mutations in comparison with DNA-SSCP, i.e. 70 % vs. 35 %, respectively (Sarkar et al., 1992a).

The SSCP mutation detection system has also been combined with dideoxy direct sequencing to provide a screening system for mutations: dideoxy fingerprinting or ddF (Sarkar et al., 1992b). In ddF, a ladder of bands of a sequence of interest is generated by performing dideoxy sequencing reactions using one of the four dideoxynucleotides and resolving these on a neutral polyacrylamide gel. Mutation detection can be done along two lines: (a) a position where the dideoxynucleotide can be incorporated is lost or

created resulting in loss or appearance of a band in the ladder, and (b) one but usually more bands which contain the site of the mutation show different migratory behaviour as in SSCP. When mutations in the Factor IX gene were analysed 84 out of 84 different mutations could be detected by ddF (Sarkar et al., 1992b). As the authors stated, however, sensitivity for detecting heterozygosity remains to be established. Furthermore, it was difficult to detect mutations in GC-rich regions due to compression of bands in the sequencing ladder.

2.3.4.4 Analysis of heteroduplex molecules

Direct dependency of electrophoretic mobility on DNA conformation is also exploited in the analysis of heteroduplexed DNA molecules in neutral gel matrices. Aberrant electrophoretic mobilities of heteroduplex molecules has been observed in agarose gels (Shore and Myerowitz, 1990) and in hydrolink gels (Keen et al., 1991). Although few DNA sequences have been analysed by this approach, sensitivity to single basepair substitutions has been reported (Keen et al., 1990; White et al., 1992). Initially, the separation of sequence variants from wild type was rather poor and essentially limited to detection of deletions and/or insertions of several basepairs (see e.g. Paw et al., 1990). More recently, studies have shown that more sequence variants can be detected if gel electrophoretic conditions are optimized by using different gel media and running conditions. These include sequence variants of the HLA locus (Sorrentino et al., 1991; Clay et al., 1991) in which most of the variants differed by more than a single basepair, and variants of the rhodopsin gene and the cystic fibrosis gene (Keen et al., 1991).

2.4 EVALUATION OF MUTATION SCANNING TECHNIQUES

From the previous sections it can be derived that there is presently a wealth of techniques for detecting DNA sequence variation (reviewed in Rossiter and Caskey, 1990; Cotton, 1993). These mutation scanning methods are capable of searching long stretches of DNA for any given type of DNA sequence variation. Apart from the practical differences such as requirements with respect to skill, laboratory equipment and reproducibility, the different methods have different scanning efficiencies in terms of number of

basepairs that can be scanned per assay and the type of DNA sequence variations that can be detected. In Table 2 some characteristics of the mutation scanning methods are listed together with their estimated scanning efficiency. It should be noted that the numbers given for the scanning efficiencies are approximations, since not every method has been tested to the same extent on a number of different genes.

DNA sequencing has the highest efficiency in that it can identify any given sequence variant but it is a typical serial approach for identifying DNA sequence variation. RFLP screening is the least efficient mutation scanning method since very few basepair alterations will fall within the palindromic recognition site of a restriction enzyme. Even when large batteries of restriction enzymes are applied very little sequence is effectively scanned for deviations. Restriction enzyme digestion of DNA fragments can be applied in 2-D separation formats in a consecutive way by using two different enzymes before each electrophoretic separation (see next section on two-dimensional formats).

An advantage of the methods other than sequencing is that they will only identify deviations from the wild type sequence. The RNAse A and Chemical Cleavage methods even have the possibility of localizing the site of the variation in the molecule under study. This is accomplished by digestion and/or cleavage at the site of variation and by subsequently measuring the size of the two resulting fragments. However, by doing so these methods do not leave the variant molecules intact for further analysis. The latter is of importance in, for instance, the screening of complex populations of molecules for sequence variants which occur in low frequencies. Furthermore, although no previous knowledge on the sequence of interest is required for the application of RNase A and the Chemical Cleavage methods, they have been shown to be labour-intensive because several manipulations are required for each sample after PCR amplification and to lack reproducibility, in that not all of the variant molecules are digested/cleaved (Theophilus et al., 1989). However, recent modifications seem to have overcome this problem (Forrest et al., 1991; Cotton, 1993).

Methods in which sequence variants display anomalous electrophoretic mobility have the advantage that the molecules remain intact and can be isolated and purified for further analysis. In this respect, the SSCP method has gained growing popularity despite of the fact that only up to 35-80 % of sequence variants can be detected (Sarkar et al., 1992a). This is most likely due to its ease of use. The heteroduplex mobility assay

remains to be evaluated in comparative studies but is unlikely to be able to detect more than 50 % of possible sequence variations.

DGGE and TGGE have a high potential for detecting sequence variants. Apart from only displaying deviations from the wild type and leaving the variant molecules intact, they have the additional advantage that after separation, the variant and wild type molecules are widely spaced apart (cm rather than mm) in comparison to the other electrophoretic mobility assays. This is of particular importance when this method is incorporated as a separation criterium in a 2-D format since it allows most of the 2-D gel space to be used for displaying fragments (see next section on 2-D formats). In this respect, DGGE (and TGGE) have the additional advantage that fragments are retarded and hence are less dependent on the duration time of electrophoresis, in contrast to CDGE in which the separation is critically time-dependent. Finally, in practical terms, DGGE is cheaper and more flexible than TGGE because of the more expensive apparatus requirements for TGGE which, in addition, allows only one TGGE gel to be run while in DGGE multiple gels can be handled and run simultaneously.

2.5 MEASURING DNA SEQUENCE VARIATION ON A TOTAL GENOME SCALE

All methods reviewed above allow individual variation in the DNA sequences of the genome to be studied but only at few sites simultaneously; such sites should not comprise more than a few hundred basepairs. For simple problems, such as the detection of a single basepair change or the scanning of a small area for sequence variation in a population of organisms, this can be satisfactory. For more complicated problems the scope of the search for variation in the genome, in terms of positions in the genome screened, needs considerable expansion. Examples of such problems are: (1) genetic linkage studies, especially those of multigenic traits which include common diseases; (2) studies aimed at random sampling the genome at as many sites as possible for alterations associated with, e.g., mutagenesis, tumorigenesis, aging; and (4) the detection of any possible sequence variation in a gene, etc. Sequencing complete individual genomes is redundant since approximately 99 out of every 100 basepairs will be identical between individuals. Rather, mutation scanning methods will have to be applied allowing the

identification of only the variant sequences. For this purpose some of the techniques discussed above can be applied to the analysis of DNA fragments. In view of the size of genomes, it is necessary to first generate fragments which are smaller than the original DNA molecules: i.e. the chromosomes. Two enzyme-based methods exist nowadays to generate DNA fragments from a genome under study: (i) restriction enzyme digestion, and (ii) PCR amplification. Fragments generated in this way can be entered into the mutation scanning analysis.

For the subsequent screening of individuals for sequence variation two different strategies can be followed: a serial one and a parallel one. In a serial strategy, the fragments are analysed one by one while in a parallel strategy many fragments are analysed simultaneously. A disadvantage of the serial approach is that large numbers of separate determinations are required. For example, for DNA sequencing it can be calculated that, assuming 300 bp to be determined in a single sequencing reaction, at least 50 million sequencing reactions have to be performed to obtain the sequence of only a single human genome (including correction for overlap but without correction for redundancy). In principle, all of the methods mentioned above allow the scanning of individual genomes for DNA sequence differences. However, their application for this purpose is severly hampered by the fact that they have to be applied serially in large scale studies, such as comprehensive mutation analysis of entire genomes or genomic regions. Indeed, even the scanning of some large genes is already a problem in view of the length of the region to be scanned and the number of different mutations which can occur (see Table 4). An example of the comprehensive analysis of a large gene for mutations is the scanning of the Factor VIII gene region in patients with unknown mutations (Kazazian et al., 1991). In this study 45 different primer sets were necessary to cover 99% of the coding part of the gene (6.9 kb), 41 out of 50 splice junctions and some regulatory sequences. By applying DGGE analysis these authors were capable of detecting 25 out of 29 previously unknown point mutations in this 187 kb large gene.

Studies such as gene-identification by linkage analysis, genome scanning of tumours and other forms of comparative analysis of closely related genomes are even more demanding and represent a considerable number of handlings. This problem can be solved by robotics; in practice this turned out to be hardly faster (G. Evans, personal communication). This type of serial analysis is therefore, by its nature, rather money-,

time- and labour-consuming and allows information only to be compiled after many separate analyses. To assess individual DNA sequence variation on a total genome scale, such as for addressing the problems mentioned above, the parallel processing approach is more efficient. This second strategy seeks to analyze as many fragments as possible in parallel rather than serially. It can maximize the information content of the analysis of a single genome by simultaneously analysing many sites in the genome in parallel with techniques that allow many basepair variants to be detected. One way of achieving comprehensive parallel analysis of a genome is by 2-D gel electrophoretic formats based on the combination of different separation criteria of DNA fragments. These formats will be discussed in the next section. By combining efficient electrophoretic mutation scanning techniques in a two-dimensional display of fragments, the information content of a single analysis can be very high. Complete genomes of lower organisms can be resolved in this way, while simultaneously allowing comprehensive detection of sequence variations at any given site in the genome.

For the analysis of genomes of higher organisms, hybridization or amplification techniques using different types of repetitive sequences can be applied to selectively visualize particular subsets of fragments from the genome of interest. This will be discussed in Chapter 4. By using this approach it is possible to analyse genomes of higher organisms at many sites which can be classified according to the probe and/or primer which is used. Such a classification can be useful in combining the data obtained from the analysis of one individual genome to datasets obtained from the analysis of many genomes. These datasets comprise both physical- and genetic mapping information on the genome of interest.

CHAPTER 3

TWO-DIMENSIONAL SEPARATIONS OF DNA MOLECULES

3.1 INTRODUCTION

In the previous sections several electrophoretic techniques have been discussed which can be used in combination to constitute a two-dimensional separation format for DNA molecules. Two-dimensional separation techniques can be based on a combination of similar or completely independent separation criteria. In the first case, fragments are clustered along a straight diagonal after two-dimensional electrophoresis. If completely independent criteria are combined substantial deviations from the diagonal will arise after separation. For example the use of a second restriction enzyme to digest size-separated fragments generated by digestion with a first restriction enzyme, will result in only the area below the diagonal being used to resolve the fragments.

Initially, 2-D DNA separation methods were applied to resolve enzymatic digests of relatively simple prokaryotic genomes. Usually all fragments could be resolved in agarose or polyacrylamide gels and visualized by ethidium-bromide staining or end-labelling. For more complex molecules only selective visualisation of sets of fragments from the genome of interest will allow interpretable analysis. This can be accomplished by hybridization analysis using probes hybridizing to multiple sites in the genome of interest, or by selective end-labelling of separated fragments. Alternatively, PCR amplification methods can be applied to selectively amplify particular DNA sequences such as those bordered by repetitive sequences (inter-repeat sequences) or multiple fragments of a contiguous stretch of DNA, like a complete gene.

3.2 ENZYMATIC DIGESTION AS SECOND DIMENSION CRITERION

Early applications of two-dimensional separation of DNA fragments involved the use of two subsequent enzymatic treatments of genomic DNA either with two restriction

endonucleases or with a restriction enzyme and S_1-nuclease (reviewed in Yee and Inouye, 1983). In general, the procedure involves restriction enzyme digestion and a first dimensional size-separation, followed by the incubation of a piece of polyacrylamide or agarose gel in a solution, containing the second enzyme, to further digest the DNA fragments. In view of the use of similar size separations in both dimensions, only the area below the diagonal of size-separated fragments can be used for resolution of fragments.

3.2.1 S_1 nuclease digestion

Two-dimensional electrophoresis involving S_1-nuclease (Yee and Inouye, 1984) can be used to detect heteroduplex molecules in complex populations of homoduplex molecules. The method includes digestion of genomic DNA with a restriction enzyme recognizing four bases, followed by a denaturation / renaturation step, separation of the fragments by size in a neutral polyacrylamide gel, incubation of the first dimension lane in a solution containing S_1-nuclease and, finally, a second dimension electrophoresis in a second neutral polyacrylamide gel. As a result only few spots, derived from those fragments which are deviating from the wild type, can be observed which all occur below the expected diagonal of size-separated fragments. These fragments correspond to the smaller-sized fragments resulting from the presence of a stretch of mismatched basepairs (see also the previous section on S_1-nuclease reactivity). This technique of two-dimensional DNA separation can be used for scanning small fragments for sequence variation but is much less applicable to problems relating to genome scanning. Particularly the denaturation/reannealing step is not possible in the analysis of complex genomes. This is a result of the presence of abundant repetitive sequences leading to the formation of aberrant heteroduplexes.

3.2.2 Digestion with a second restriction enzyme

When two consecutive restriction enzyme digestions are applied a two-dimensional display of fragments will occur similar to the one described above for S_1 nuclease. That is, all fragments are below the diagonal. Bands arising in the first dimension separation can be individually excised and treated with the second enzyme

(Chen and Thomas, 1983; Gilroy and Thomas, 1983). This approach has been successfully applied to analyse and isolate repetitive sequences in Drosophila DNA. Alternatively, the whole 1-D gel lane itself or its separation pattern transferred to a DEAE cellulose membrane, can be incubated with the second enzyme.

For the visualisation of the fragments resolved on a 2-D gel three routes can be pursued. First, DNA dyes such as ethidium bromide can be applied to visualize all fragments on the gel. This has been applied to the analysis of simple prokaryotic genomes such as Escherichia coli and Bacillus subtilis and the mycoplasma Acheloplasma laidlawii to analyze repeated DNA sequence families and to estimate genome sizes (Potter and Newbold, 1976; Potter et al., 1977; Yee and Inouye, 1982; Yee and Inouye, 1983; Boehm and Drahovsky, 1984; Poddar and Maniloff, 1986). With appropriate restriction enzyme combinations 180 to 1312 restriction fragments varying in size between 0.350 and 5 kb were observed allowing accurate estimates of the genome size of a variety of prokaryotic species by this approach.

Second, transfer of the separation pattern to a membrane and hybridization analysis using specific probes allows the visualization of particular subsets of fragments. When applied to a more complex genome, such as that of the protozoan Tetrahymena thermophilae which is roughly 200.10^6 bp, hybridization analysis was used to visualize a repetitive subset of fragments (Hüvos et al., 1988). This approach allowed the resolution of 80 members of a repeat family and was capable of detecting inter-strain differences in this organism. Similar types of analyses have been performed with Saccharomyces (Rogan et al., 1991), human genomic DNA (Sakaki et al., 1983), and mouse genomic DNA (Fanning et al., 1985; Au et al., 1990; Sheppard et al., 1991). Polymorphisms were detected in gene families such as the Major HistoCompatibility (MHC) system genes of the mouse (Sheppard et al., 1991) and intra-cisternal A particles (Au et al., 1990; Sheppard et al., 1991) and included the detection of methylation polymorphism (Fanning et al., 1985). In spite of the fact that only half of the gel area could be used for resolution of the fragments detected, still up to 370 spots could be detected using this approach. The use of agarose gels in these separation methods resulting in rather faint and blurry spots limits the resolution (e.g. Sheppard et al., 1991). This can, however, be compensated for by using polyacrylamide gels and electroblotting protocols (Uitterlinden et al., 1989).

A third way of visualizing particular subsets of fragments is by selective end-

labelling or T4 polymerase strand displacement labelling of restriction fragments (see Maniatis et al., 1982; Sambrook et al., 1990). That is, after digestion of genomic DNA with a first restriction enzyme, fragments are labelled and subsequently further digested with a second, more frequently cutting enzyme. This has been applied in the analysis of strains of Escherichia coli (Yi et al., 1990) and in the analysis of the Drosophila and the mouse genome (Hatada et al., 1991). The latter method is referred to as Restriction Landmark Genomic Scanning (RLGS) and allows up to 2000 spots to be analysed in a single gel and was shown to be capable of detecting polymorphisms among different individual mouse DNA samples. The technique is based on the two-dimensional separation of end-labelled DNA restriction fragments in which the first dimension is run in agarose gels and the second dimension is run in PAA (polyacrylamide) gels. By using two different matrices for the separations, efficient use is made of the 30 x 45 cm 2-D gel area resulting in up to a few thousand spots per pattern. The application for mouse genome mapping is based on the analysis of 26 DBAxC57/Bl6 recombinant inbred strains and inter-species backcrosses of C57/Bl6 x Mus spretus. Between C57 and DBA 13.4% (51/378) variant spots were detected while between M. musculus and M. spretus 51% spot variants (188 out of 369) were detected. This illustrates the major drawback of the method, i.e. the very low level of polymorphism that can be detected. As a result it can not be used for human genetic analysis. Indeed, when two unrelated (Japanese) individuals were compared by RLGS only 0.15% spot variants were detected (Hayashizaki, personal communication). This is explained by the low informativeness of restriction enzyme recognition sites in general. The technique has been applied to the analysis of some tumors to detect amplifications (in a breast tumor 10 spot variants showing 4 - 7 times amplification were detected). The most commonly detected genetic change observed in tumors (as shown by other studies; Verwest, A.M., Molijn, A.C., Andersen, T.I., Börresen, A.-L., Uitterlinden, A.G., Mullaart, E., and Vijg, J., submitted), i.e. loss of chromosomal fragments (indicated by the loss of a spot variant), is inherently more difficult to detect by RLGS because intensity decreases of two-fold have to be observed.

3.2.3 Pulsed field electrophoretic separation

The principle of two consecutive restriction enzyme digestions as basis for a two-dimensional DNA separation can be extended to very infrequently cutting enzymes. After digestion of genomic DNA with such an enzyme (for example Not I), separation of the resulting fragments by Pulsed Field Gel Electrophoresis (PFGE; Schwartz and Cantor, 1984) or Field Inversion Gel Electrophoresis (FIGE; Carle et al., 1986) in the first dimension is followed by in situ digestion with a more frequently cutting enzyme and size-separation in the second dimension in normal agarose gels. This approach was applied in the analysis of multigene families such as the murine T-cell receptor complex (Woolf et al., 1988), and Ly-6 multigene family (Kamiura et al., 1992) and the human Immunoglobulin Heavy chain variable region (Walter and Cox, 1989). In addition, FIGE first dimension separation of partial digests, followed by in situ complete digestion of the fragments and separation by FIGE second dimension separation has been used to create macro-restriction maps of micro-organisms such as Mycoplasma mobile (Bautsch, 1988). Although the method is suitable for large scale mapping of genomic regions or even entire genomes of micro-organisms, the effectiveness as a genome scanning method is quite low in view of the low number of fragments obtained and small mutation scanning efficiency. That is, only large insertion or deletions can be detected and variation in the recognition site of the restriction enzyme.

In summary, restriction enzyme digestion is helpful by generating many fragments to be analysed. However, in the absence of an independent second dimension separation criterium of the fragments generated, only the recognition sites of the restriction enzymes used are screened for variation (see Table 2). In general, only the area below the diagonal in a 2-D separation pattern will be used in these 2-D formats. The actual gel area used can be increased by taking different gel media or concentrations of gel media for the subsequent separations (e.g. Hatada et al., 1991).

3.3 CONFORMATION AS SECOND DIMENSION SEPARATION CRITERION

Attempts to combine two independent electrophoretic separation criteria involved the separation on the basis of conformation of DNA molecules under different

circumstances such as the presence of intercalating dyes (Meyer and Hildebrandt, 1986; Vanweye et al., 1991; Snapka et al., 1991) and the use of different electric field strengths, buffers and matrices (Bell and Byers, 1983; Shinomiya and Ina, 1991; Vanwye et al., 1991; Snapka et al., 1991). These methods can reveal differences in GC-content, DNA bending, single strand conformation, etc. which can be applied e.g. for mutation detection when used in combination with SSCP (Kovar et al., 1991). Although these methods can reveal minute differences, the resolution obtained is low in that only few spots will deviate from the diagonal. These methods therefore do not allow large scale scanning of genomic regions.

3.4 DGGE IN THE SECOND DIMENSION

The denaturing gradient principle as applied in DGGE was used in an early stage of development as a second separation criterion in two-dimensional electrophoretic analysis of highly complex mixtures of hundreds of different restriction fragments (Fischer and Lerman, 1979a and b; see also previous section on DGGE). The method could resolve essentially all Eco RI restriction fragments of the Escherichia coli genome (estimated size: 4399 kb) and detect a 50 kb insertion of lambda DNA in the E. coli genome (Fischer and Lerman, 1979a). In addition, the 2-D method has been used to resolve digests of complete genomes of several other microorganisms with sizes ranging from 724 to 1833 kb (Poddar and Maniloff, 1989).

In all the 2-D systems sofar discussed, migration is to a very large extent dependent on fragment size in both first and second dimension. One particularity of the DGGE separations, however, is their considerable independence of the length of a fragment due to different migratory behaviour in PAA gels of partially melted DNA molecules. Initially the PAA gels will sieve the fragments according to their size but upon reaching the $T_m(50)$ of the fragments, this sieving is in most instances overridden by the sieving according to the partially single stranded structures. As a result fragments of different size but sharing a particular melting domain, will migrate to similar positions in the DGGE dimension after 2-D separation. For sheared total genomic DNA, for example, this will result in the appearance of streaks in a horizontal direction at different levels in the second dimension gel pattern. In this way a two-dimensional display of melting domains

present in a given DNA molecule will result, which can also be used for comparative purposes. This feature has been exploited in the separation of sheared pBR322 fragments (Fischer and Lerman, 1983; Lerman et al., 1981 and 1984).

Deviations from the normal position in a 2-D gel can occur due to changes in length or sequence content of the fragment. The sensitivity in this respect is determined by the percentage polyacrylamide (PAA) and the gradient of denaturants, respectively, and by the size of the gel. Normally, 6 % PAA gels are used which allows for the separation of fragments of 0.1 to roughly 2 kb with a resolution of approximately 5 % of the fragment length using standard size equipment (see also chapter 2). Higher percentages PAA will allow smaller length differences to be detected when present in fragments of a few hundred basepairs. For standard DGGE separations a gradient of 0-80 % denaturants is usually applied. This will allow most fragments to reach their melting point and focus in the gel. The sensitivity for detection of sequence variations is lower in this gradient but can be increased by taking gradients spanning a shorter range of denaturants.

The efficient use of the 2-D gel area and the high sensitivity for sequence variations, independent of the restriction enzyme used, make 2-D DNA separation with DGGE in either the first or the second dimension, the ideal method for analysing complex genomes by so-called "genome scanning". It allows the efficient detection of deviations of a normal sequence at many possible sites in the genome of interest. It does not, however, provide information on the location in the genome of the variation. In principle, a two-dimensional DNA separation pattern represents a display of fragments derived from a contiguous sequence of basepairs. The relation of the two-dimensional display of the fragments to their original sequence in the genome of origin is usually not known a priori. In two-dimensional DNA electrophoresis incorporating DGGE, this can be established for each individual fragment on the basis of the size of the fragments and its melting temperature $T_m(50)$.

For small genomes, yielding few restriction enzyme fragments, the relationship between gel position and genomic position can be easily determined by partial digestion experiments. For larger regions or genomes, however, additional characterization is necessary to obtain interpretable 2-D separation patterns. A logical next step in this respect is hybridization with probes. Probes detecting overlap of certain restriction sites in

the genome of interest will identify adjacent restriction fragments and allow for the construction of contigs of spots. When it is known that all possible fragments from a given genome are resolved in a single 2-D gel and that no spots are derived from more than one restriction fragment this approach will allow for the construction of a physical "2-D" map of a genome. Alternatively, such a physical map can be composed by subsequent hybridization analysis with multiple short oligonucleotides (Michiels et al., 1987; Craig et al., 1990; Lehrach et al., 1990). In analogy to this mapping approach, also the sequence of the fragments can be obtained by means of hybridization with multiple overlapping oligo's (Drmanac et al., 1989). For larger genomes such a 2-D genome mapping approach becomes unfeasible due to the great number of fragments; multiple fragments are likely to be present at a certain spot position. Of course for smaller regions contig mapping is possible using unique sequences as probe derived from cosmids or YACs.

The most efficient method of obtaining information from a genome in this respect is by using probes for repetitive sequences. When repetitive sequences are used as a probe in hybridization analysis, multiple sites in the genome can be analysed but information on order of fragments is lost and physical mapping information can no longer be retrieved from the 2-D pattern. However, each spot can be assigned a third characteristic (i.e. partial sequence content, apart from its length and $T_m(50)$) which makes it part of a group of related sequences with a particular function, organization and/or localization on chromosomes. Such repetitive sequences can include "anonymous" repeats such as micro- and minisatellites, retroviral elements occurring in many genomes but may also include motifs specific for groups of protein encoding sequences such as zinc-finger motifs, homeobox motifs etc. Both (anonymous and protein-coding) type of sequences have been shown to be present in numerous species and thereby even allow comparative studies using 2-D DNA typing. A particular form of 2-D DNA typing in which repeats are used but in which the physical and genetic mapping information can be preserved and exploited is the use of primers for repetitive elements in (PCR) amplification reactions. Inter-repeat regions can be selectively amplified from sources with a known physical map location such as cosmids, YACs and somatic cell hybrids and used as probes in comparative studies of different individuals. This allows combined physical- (position on the chromosome) and genetic (variation between individuals) mapping of a particular

region (Uitterlinden et al., 1991a).

CHAPTER 4

REPETITIVE SEQUENCES AS ANCHOR POINTS IN TOTAL GENOME ANALYSIS

4.1 DNA MARKERS

The identification and screening of many marker loci started with the introduction of Southern blot analysis (Southern, 1975) which led to the discovery of so-called restriction fragment length polymorphisms (RFLPs). RFLPs were first detected for the globin gene locus (Kan and Dozy, 1978; Jeffreys, 1979). Soon thereafter RFLPs were described for so-called anonymous DNA segments (Wyman and White, 1980). Especially, the latter type of sequence displays DNA polymorphisms (see below) since as non-coding DNA it is under much less evolutionary selective pressure to preserve its sequence in populations.

The first RFLPs discovered are based on the absence or presence of a restriction enzyme recognition site and can be referred to as Restriction Site Polymorphisms or RSPs. Extensive searches for such RSPs in the human genome showed that on average 1 in 300 to 1 in 1000 basepairs is polymorphic in the population for autosomes and the X-chromosome, respectively (Cooper et al., 1985; Hofker et al., 1987). In this respect CpG sites show a higher frequency of polymorphism due to the susceptibility of methylated cytosine to undergo mutation (Barker et al., 1984).

The frequency of polymorphic basepairs in human DNA measured through RFLP analysis is likely to be an underestimate since it is a very insensitive way of detecting point mutations. Several other methods are now available, such as DGGE analysis, which can detect sequence variation in any given stretch of basepairs much more efficiently. Only recently, these methods have been introduced to exploit DNA sequence variation in genetic linkage analysis of disease. A particular advantage of methods such as DGGE is that they can discern multiple alleles of a given basepair sequence due to the presence of several polymorphic basepairs in the sequence (e.g. Sheffield et al., 1992). Such multi-allelic systems will increase the informativeness of the locus in genetic analysis. Another

set of multi-allelic markers is based on a completely different type of polymorphism involving repetitive sequences which are termed micro- and minisatellite sequences. These sequences are found in several to hundred thousand copies dispersed throughout the genome of many different species. In this respect, they represent ideal candidates for genome scanning studies because a single probe, corresponding to a core sequence that several micro- and/or minisatellite loci share, can detect many genetic marker loci simultaneously. In Table 3 an overview is presented of the different forms of repeat sequence polymorphisms found in the human genome.

4.2 MICRO-AND MINISATELLITE SEQUENCES

Micro- and minisatellite loci consist of short sequence units which are tandemly repeated. Some representatives of these repeat loci show polymorphic variation in the number of repeat units at the locus and can be termed Variable Number of Tandem Repeat or VNTR loci. Microsatellite VNTRs are referred to as either microsatellites, simple sequence repeat motifs (SSRM), short tandem repeats (STR), Variable number of dinucleotide repats (VNDR) or simple sequence length polymorphisms (SSLP).

4.2.1 Minisatellites

Some of the VNTRs have been shown to be extremely polymorphic with more than hundred alleles in the population. Notably, these repeats were originally described as unrelated regions of short tandem repeats in anonymous DNA regions (Wyman and White, 1980) and in the proximity of genes such as the alpha-globin gene (Higgs et al., 1981), the insulin gene (Bell et al., 1982), the zeta-globin gene (Proudfoot et al., 1982; Goodbourn et al., 1983, Ullrich et al., 1982) and the cHa-ras proto-oncogene (Capon et al., 1983). Later, Jeffreys et al. organized them into sets of repeat loci (which were named minisatellites) on the basis of common, short stretches of sequence homology, termed core sequences (Jeffreys et al., 1985a). When such a core sequence was used as a probe in Southern blot hybridization under relaxed hybridization conditions many mini-satellites were detected simultaneously. The resulting banding pattern constitutes an individual-specific "DNA-fingerprint" due to variable size of the large restriction

TABLE 3 HUMAN REPEAT SEQUENCE POLYMORPHISMS

TYPE OF REPEAT	DISTRIBUTION	COPY NUMBER	TYPE OF POLYMORPHISM	POLYMORPHIC FRACTION[1]
Alu I	dispersed	$0.5.10^6$	-insertion/deletion -point mutation -polyA VNTR	infrequent 50% 60%
Kpn I	dispersed	1.10^3- 1.10^5	-insertion/deletion -point mutation -polyA VNTR	infrequent ? ?
alphoid-satellite	centromeric	1.10^4	-VNTR	?
minisatellite	subtelomeric	2.10^4	-VNTR -point mutation	50-75% ?
microsatellite	dispersed/ clustered	1.10^3- 1.10^5	-VNTR -point mutation	50-75% ?
LTR	dispersed	5.10^3	-insertion/deletion -point mutation	infrequent ?

1 = estimates given as a percentage of total number of copies in the human genome

fragments (Jeffreys et al., 1985b and 1991). This RFLP was demonstrated to be due to allelic variations in repeat copy number at the minisatellite loci detected with the core probe (Jeffreys et al., 1985a; Wong et al., 1986; Wong et al., 1987).

Up to now several different core sequences have been described, with homology to minisatellites in several animal species (Georges et al., 1988; Vergnaud et al., 1992). Sequence analysis of repeat units of particular minisatellite loci has revealed variation, which occasionally can ablate or create restriction enzyme recognition sites (Wong et al., 1986; Waye and Fourney, 1990; Uitterlinden and Vijg, 1991). This repeat unit sequence variation is the basis of another (very rich) source of polymorphism which is characteristic for the minisatellite loci and which is detectable by Southern blot hybridization analysis (Jeffreys et al., 1990) or by means of PCR analysis (Jeffreys et al., 1991).

4.2.2 Microsatellites

Different from minisatellites, but a structurally very related type of tandemly repetitive sequences are the microsatellites, the repeat unit of which has arbitrarily been set at 5 bp or smaller. Due to large variations in repeat copy number, similar to those observed for minisatellites, "DNA-fingerprint" patterns can be obtained when microsatellite "core-probes" are used in Southern blot hybridization analysis (Tautz and Renz, 1984; Ali et al., 1986; Schäfer et al., 1988; Kashi et al., 1990). In addition, particular microsatellite loci consisting of $(GT)_n$ dinucleotides (Weber and May, 1989; Litt and Luty, 1989; Tautz, 1989; Smeets et al., 1989) or different tri- and tetranucleotide motifs (Boylan et al., 1990; Edwards et al., 1991 and 1992) can be analysed by using fully denaturing PAA gels in combination with PCR amplification techniques. They display similar VNTR type of polymorphism as minisatellites, albeit with alleles of smaller length. The gel type used allows differences as small as one repeat copy of 1-3 bp to be detected. In a similar way the polydeoxyadenylate (polyA) tract of several Alu repetitive elements has been shown to be of variable size due to a VNTR type of polymorphism of AT-rich microsatellite motifs (Economou et al., 1990; Zuliani and Hobbs, 1990; Epstein et al., 1990).

4.2.3 Large satellites

In addition to the shorter tandemly repetitive sequences also much longer repetitive arrays have been found to display polymorphism very similar to the VNTR type. These include tandemly repetitive alphoid centromeric repeats which are used to create centromere-based maps of a chromosome (Willard et al., 1986; Devilee et al., 1988; Wang-Jabs et al., 1989; Haaf and Willard, 1992) and so-called midisatellites (Nakamura et al., 1987b; Page et al., 1987).

4.2.4 Origin of polymorphism of micro and minisatellites

The majority of minisatellites described to date are composed of GC-rich repeat units which show homology to the chi sequence of Escherichia coli. This has led to speculation that the "core" sequence may serve as a recombination signal to promote unequal crossing over at the minisatellite sequence (Jeffreys et al., 1985a; Steinmetz et al., 1986). In addition, however, several VNTRs have now been identified which are composed of AT-rich motifs (Stoker et al., 1985; Knott et al., 1986; Bowcock et al., 1989; Iwasaki et al., 1992) including AT-rich microsatellite motifs at the 3' ends of Alu repeats (see above). This indicates that the VNTR type of polymorphism is not uniquely associated with a GC-rich chi-like motif.

Analysis of human pedigrees has shown that micro- and minisatellite VNTR loci can vary widely with respect to the frequency of generating new alleles due to mutation (Jeffreys et al, 1988 and 1991; Nürnberg et al., 1989). To determine the origin of the mutation, pedigrees in which mutant minisatellite alleles are observed were analysed with markers flanking the VNTR locus undergoing mutation. This showed that no homologous chromosomal areas were exchanged and thus that unequal crossing-over recombination events are not a common cause for this type of sequence variation (Wolff et al., 1988; Wolff et al., 1989). Instead, polymerase slippage replication errors are a more likely source of the variation observed. Further evidence for this was obtained by the creation of length variants of microsatellite stretches by in vitro synthesis (Schlötterer and Tautz, 1992).

4.2.5 Evolutionary background

Sequences, homologous and probably structurally related to the micro- and minisatellite sequences, have been found in a wide range of taxonomic groups (Tautz and Renz, 1984; Kashi et al., 1990; Uitterlinden et al., 1989; Turner et al., 1990; Schlötterer et al., 1991). Besides in higher eukaryotes, their presence has been demonstrated in lower eukaryotic organisms such as Caenorhabditis elegans (Uitterlinden et al., 1989), Musca domestica and other insects (Blanchetot, 1989) and even unicellular organisms such as Plasmodium spec (Rogstad et al., 1989; van Belkum et al., 1991). It is unknown if the sequences detected over the wide range of taxonomic groups are related with respect to function and origin. It has been speculated that the micro-, mini- and larger satellites play a role in the chromatin folding network (Vogt, 1990).

The widespread occurrence of VNTRs provides an abundant source of DNA markers also for comparative genome analysis. This was demonstrated for a number of human VNTR loci shown to cross-hybridize with mouse minisatellite loci (Julier et al., 1990), for several VNTR loci isolated from bovine DNA (Georges et al., 1991) and for several cetacean (whale) microsatellite loci (Schlötterer et al., 1992).

4.2.6 Coding sequences containing VNTRs

In a search among entries of the DNA sequence databases micro- and minisatellites have been found to be part of coding sequences (Tautz et al., 1986). In particular instances the VNTR type of sequence variation can even be functional as is the case in Plasmodium. Here, tandemly repetitive sequence motifs encode amino acid stretches in immunodominant epitopes of membrane proteins (Galinski et al., 1987; van Belkum et al., 1991 and 1992). In the human genome several examples have been found of coding sequences containing micro- and minisatellite-like repeats and displaying VNTR polymorphism. These include a tumor-associated epithelial mucin coding gene (PUM) containing a 60 bp VNTR (Swallow et al., 1987), and the involucrin gene containing a 30 bp VNTR (Simon et al., 1991). In exon 1 of a sarcoplasmic reticulum protein gene called HRC, $(GAG)_n$ and $(GAT)_n$ VNTRs were found (Hofmann et al., 1991).

Recently discovered and clinically relevant examples of microsatellite polymorp-

hism within the coding regions of genes involve the expansion of particular trinucleotide motifs present in human disease genes. They represent a novel mechanism of genetic disease and involve the genes for X-linked spinal and bulbar muscular atrophy (SBMA or Kennedy's disease; La Spada et al., 1991), Fragile-X syndrome (Verkerk et al., 1991; Fu et al., 1991; Yu et al., 1991) and Myotonic Muscular Dystrophy (DM; Fu et al., 1992; Mahadevan et al., 1992). All three $(CNG)_n$ micro-satellite arrays are found within the mature transcript of the gene: $(CAG)_n$ in the first exon of the androgen receptor gene in Kennedy's disease, $(CGG)_n$ in the 5' untranslated part of the FMR-1 gene in Fragile-X syndrome, and $(CTG)_n$ in the 3' untranslated region of the myotonin gene in DM. Whereas these microsatellite motifs display normal polymorphic variation in the human population (6-54 repeats in Fragile X, 5-30 repeats in DM, and 17-26 repeats in SBMA), a substantial expansion of the array was noted in patients with the disease (200-1,000 repeats in Fragile X, up to 2,000 repeats in DM and 40-52 repeats in SBMA). For Fragile X and DM the expansion was shown to be correlated with the severity of the disease indicating a close functional relationship between the repeat region and gene regulation sites. In addition, the expansion was found to be somatically heterogeneous probably as a result of mitotic instability, i.e. different cells show a different increment in length of the repeat array. This phenomenon has also been noted in telomeric regions containing the $(TTAGGG)_n$ repeat arrays (for a review, see Kipling and Cooke, 1992). Since the discovery of these disease genes, several other genes containing (polymorphic) trinucleotide repeat stretches have been detected in human cDNA libraries and sequence databases (Riggins et al., 1992). This indicates the presence of other human genes which may undergo this new mechanism of mutagenesis giving rise to genetic disease.

4.2.7 Polymorphic fraction of micro- and minisatellites

Since thus far only a fraction of the total number of short tandem repeat loci has been analysed for the VNTR type of polymorphism, it is unknown how many of these sequences are polymorphic in higher animal and plant genomes. Estimates of the polymorphic fraction of both micro- and minisatellite loci vary from 25-75% (Jeffreys et al., 1985a; Nakamura et al., 1987; Uitterlinden et al., 1989; Tautz, 1989; Armour et al., 1991; Weber, 1991). This, in combination with their widespread occurrence in telomeric

(Royle et al., 1988), centromeric (Willard et al., 1986) and interstitial (Luty et al., 1990) regions on the chromosomes, make these repeat sequences good candidates as markers for total genome scanning of eukaryotic genomes. Expressed sequences containing polymorphic micro- and minisatellite motifs also present an ideal source of genetic markers for performing genetic association studies.

4.3 OTHER INTERSPERSED REPETITIVE SEQUENCES

Repetitive sequences can be exploited as a valuable source of polymorphic markers to be employed in a locus-specific manner by serial analysis for their local repeat copy number. The repetitive nature of genomes can also be employed in total genome comparison of individuals by using probes that detect many repetitive sequences simultaneously. Because of sequence similarity of part of the repeat unit many micro- and minisatellite sequences, including VNTRs, can be detected simultaneously by using core probes. Other repetitive sequences can be detected by using probes representing a consensus sequence of the repeat element. Here some general aspects of such repeat families and their analysis will be discussed.

Allmost all genomes studied sofar but especially the eukaryotic genomes have been found to contain a considerable proportion of DNA sequences which are repetitive, i.e., are present in more than two but usually in at least hundreds of copies. The proportion of the genome consisting of repeats can vary widely in different species. A particular class mainly found in lower organisms is formed by the transposons (Shapiro, 1984). These sequences can undergo rapid change in genome position and have been found in organisms such as bacteria, Saccharomyces, Drosophila, and Caenorhabditis and also in plant species such as Zea mais.

The human genome consists of a variety of repetitive sequences that may comprise up to 40% of the complete sequence (Britten and Kohne, 1968). Because the function, if any, of most repeat sequences is unknown, current classification systems are based on empirical characteristics, such as copy number (up to 1,000,000 copies), repeat unit length (up to 6500 bp), distribution through the genome (clustered or dispersed), and local organization (tandem arrays or interspersed). In addition, multigene systems such as the MHC system and gene families such as the globin genes and homeobox genes can be

considered as low copy repetitive sequences. Besides classifying repetitive sequences as being derived from a particular, evolutionary defined repeat unit, particular sequence motifs can also be found throughout the genome. Such motifs can be defined on the basis of function (e.g. zinc-finger motifs) or on the basis of their structure (Z-DNA stretches, palindromic recognition sites for restriction enzymes, etc.).

4.3.1 Polymorphism

Repetitive sequences with an interspersed organization have been shown to exhibit polymorphisms in the human and other mammalian genomes (Table 3). In the human genome such polymorphic sites can involve Kpn I repeats (Katzir et al., 1985; Economou-Pachnis et al., 1985; Lakshmikumaran et al., 1985; Burton et al., 1985; Shyman and Weaver, 1985; Musich and Dykes, 1986; Kazazian et al., 1988; Dombroski et al., 1989; Woods-Samuels et al., 1989), Alu I repeats (Schuler et al., 1983; Lin et al., 1988; Economou-Pachnis and Tsichlis, 1985; Hobbs et al., 1985; Lehrman et al., 1987; Vidaud et al., 1989; Wallace et al., 1991), and members of the Sau3A family (Kominami et al., 1983). These polymorphisms can be the result of the absence or presence of one or more copies of a repeat unit. This will create or destroy a restriction enzyme recognition site (RSP: restriction site polymorphism) or alter the length of a given restriction fragment (insertion/deletion or variable number of tandem repeat (VNTR) polymorphism). RFLPs of interspersed repeats (in which case a single repeat unit is either present or absent at a particular locus) are not essentially different from "classical" RFLPs; only two alleles can be discerned which have varying frequencies in a population. In general this type of polymorphism is very infrequent.

Another class of repeats found in mammalian genomes involve viral-like elements, including the so-called intracisternal A-particles (IAP) in mice and retroviral-like elements in humans. IAP are a closely related set of endogenous proviral-like sequences present in the genome of all inbred mice (Kuff and Lueders, 1988). They have been shown to be polymorphic between strains of mice and probes for IAP sequences have been successfully used to perform genetic linkage studies (Mietz and Kuff, 1992) and to detect a disease-causing duplication in the mouse (Brilliant et al., 1991). In humans different classes of retroviral elements, occurring in copy numbers of 1,000 - 2,000, have been

found including the RTVL-H and the HERV-A elements. The low copy number of these repetitive elements allows them to be used for characterization of somatic cell hybrids and isolation of region-specific cosmid clones (Sugino et al., 1992; Meulenbelt, I.M., Wapenaar, M.C., Patterson, D., Vijg, J., and Uitterlinden, A.G., 1993, in press). Infrequent polymorphisms due to homologous recombination of LTRs have been observed for the RTVL-H elements (Mager and Goodchild, 1989).

Sequence analysis of individual members of many of these repeat families has revealed basepair substitutions within repeat units, suggesting a widespread occurrence of polymorphic basepairs within repeat loci in the human population. Since point mutations will only rarely alter a specific restriction enzyme recognition site, this kind of polymorphism is difficult to detect by Southern blot analysis, especially in view of the sometimes extremely high copy number of these repeat families.

4.3.2 PCR analysis

Single "consensus" primers for repetitive sequences in a PCR reaction can be used for the simultaneous analysis of many inter-repeat regions. One such primer is capable of detecting many loci simultaneously and therefore of great value in parallel genome analysis. Repetitive sequences flank the region of analysis and hence this type of PCR analysis is referred to as inter-repeat PCR (irPCR). The technique was first established for the analysis of human-hamster somatic cell hybrids containing single human chromosomes based on known repetitive sequences in the human genome such as Alu I (Nelson et al., 1989) and Kpn I or LINE repeats (Ledbetter et al., 1990). The primers can be chosen such that only the human Alu I repeats are detected but not the homologous but more distantly related rodent B1 and/or B2 repeats. Similar irPCR primers have been developed for the analysis of somatic cell hybrids of mouse chromosomes on a hamster (Cox et al., 1991; Simmler et al., 1991) or human background (Herman et al., 1991). Besides primers specific for known dispersed repetitive elements also primers homologous to microsatellite sequences such as $(GA[C/T]A)_n$ have been successfully applied (Cox et al., 1991). As a result of the dispersed distribution of the repeats, the collection of fragments obtained from a hybrid containing one human chromosome or part thereof constitutes a specific "fingerprint" or karyotype. By comparing such patterns small variations in chromosome

constitution can be analysed by comparison of electrophoretic separation patterns of the PCR products (Ledbetter et al., 1990) or through in situ hybridisation (Lichter et al., 1990). Such panels of chromosome-specific irPCR fragments have been established for several human chromosomes including chromosome 6 (Meese et al., 1992) and chromosome 10 (Brooks-Wilson et al., 1990).

Two different types of polymorphisms have been detected in the analysis of Alu I repeats. One consists of a microsatellite type of VNTR polymorphism at the 3' end of the Alu repeat unit consisting mainly of $(T[T/A]A)_n$ but also comprising other types of repeat units (Economou et al., 1990; Zuliani and Hobbs, 1990; Epstein et al., 1990). The other type is based on presence or absence of a priming site for the consensus primer either due to basepair polymorphism or presence/absence of a complete repeat unit (Sinnett et al., 1990). Similar types of polymorphisms might be found for the LINE type of repeats since these have also been found to contain polyA stretches at the 3' end (Economou et al., 1990).

Besides using primers specific for known repetitive elements also "random" primers have been applied in irPCR (Williams et al., 1990) including primers as short as 5 bp (Caetano-Anollés et al., 1990). This approach has been used successfully in several species to detect polymorphisms based on the presence or absence of amplified fragments. Since no locus information is available such markers are referred to as random amplified polymorphic DNA (RAPD). In particular segregating populations (i.e. populations of sibs from a single cross) the method has been used to develop markers in a region-specific manner (Michelmore et al., 1991; Giovannoni et al., 1991). Disadvantages of the RAPD system are the limited number of fragments, the low fraction of variant bands detected per primer and the low informativity of the dominant bi-allelic polymorphism. These characteristics require a very large number (several hundreds to several thousands) of primers and PCR reactions per individual for comprehensive genome analysis. Furthermore their lack of species-specificity effectively constrains their application in chromosome fingerprinting/karyotyping. Although Mendelian inheritance has been observed, an excess of mutant bands was found in paternity analysis of primates (Riedy et al., 1992). Further evaluation by cloning and sequence analysis of the polymorphic loci detected by RAPD PCR will allow to determine the origin and location of these sequences and evaluate the usefulness of this approach for genome analysis.

In summary, inter-repeat PCR based on known repetitive elements seems a powerful tool in parallel genome scanning approaches. A single primer can detect hundreds if not thousands of potential marker loci and the method can be applied in a chromosome-specific manner. Combination of several primers specific for different types of repetitive sequences will allow comprehensive genome analysis and thereby total genome comparison among individuals.

4.4 TOTAL GENOME ANALYSIS WITH DNA MARKERS

The essence of genome comparison studies, e.g. linkage mapping of disease genes or the analysis of tumor genomes, is to scan the genome for DNA sequence variation at evenly spaced points along each of the chromosomes. DNA markers, detecting such points of variation, can identify chromosomal regions containing particular genes by following the transmission of specific chromosomal regions in pedigrees, or in studies of tumorigenesis, aging and mutagenesis to identify genome alterations. The efficacy of such an analysis depends in part on the density of markers which constitute the genome map.

Sofar marker analysis in these studies proceeded in a serial way. That is, markers are analysed one by one for a number of samples simultaneously. Three such locus-specific marker systems have been available. Two systems involve Southern blot hybridization analysis of polymorphic loci detected by locus-specific probes. The first markers used were detecting RFLPs based on bi-allelic restriction site polymorphisms while later multi-allelic VNTRs were used (Nakamura et al., 1987; see also previous sections). Although they are preferentially found in subtelomeric regions their high heterozygosity (70-99%) makes them suitable markers.

The third and by now the most widely used system involves the PCR analysis of the $(GT)_n/(CA)_n$ loci (Weber and May, 1989; Litt and Luty, 1989; Weber, 1990; Beckmann and Weber, 1992). Additional microsatellite markers have been developed which consist of repeat units of 3 or 4 basepairs (Edwards et al., 1991) the PCR products of which are easier to interpret. Of all microsatellite markers analysed sofar the $(GT)_n$ loci have been shown to be abundant in many different genomes and display heterozygosity values of usually 70 % and for the most polymorphic ones even up to 85 % (Weber, 1991a and 1991b).

Several genetic maps of the human genome have been published of which the first (Donis-Keller et al., 1987) consisted completely of DNA markers analysed by Southern blot analysis. More recent genetic maps consist entirely of $(GT)_n$ loci, such as for mice (Dietrich et al., 1992), cattle (Georges et al., personal communication) and humans (Weissenbach et al., 1992). When combining the different markers on these maps an ever increasing genetic resolution can be obtained allowing any given particular gene to be mapped in the proximity of one or more of these genetic marker loci. To date there are approximately 1,000 human $(GT)_n$ markers giving an average resolution of more than 1 marker per 5 cM. For being able to extract the necessary genetic information from particular families it is necessary to have a highly informative marker every 1-2 cM. It is anticipated that over 4,000 human polymorphic $(GT)_n$ loci should be available within the next 2-3 years (Todd, 1992). According to this author close to 800 genotypes can be determined per sequencing gel through sequential hybridizations.

A major application of the results to be obtained in the genome mapping and sequencing endeavour will be to compare genome sequence organization in different strains or individuals to identify and isolate specific genes. The approach to assess individuality by sequencing complete individual genomes, however, is unfeasible and cost-ineffective in view of the high level of similarity and the resulting redundancy of the data generated. A more efficient approach is to directly focus on the sequence differences, which can then be used as marker loci in linkage and/or association studies aimed at identifying genes responsible for medical conditions and other traits in humans, animals and plants.

When more complex problems are to be addressed, such as analyzing genetic heterogeneity or polygenetic inheritance, the identification of the DNA sequence variations involves substantial efforts. Such studies require the complete genome of interest to be densely populated with (highly) polymorphic markers spread over the genome and statistical methods that enable exploitation of such a rich source of genetic information (Lander and Botstein, 1986 and 1989). The feasibility of such an approach has been demonstrated by some preliminary studies, resolving particular quantitative traits, influencing e.g. fruit mass in tomato (Paterson et al., 1988) and hypertension in rats (Hilbert et al., 1991), into Mendelian factors and to map these quantitative trait loci (QTL)s on the chromosomes.

In view of the complexity of the higher eukaryote genome it would be advantageous to be able to monitor sequence variation in parallel on a total genome scale, i.e. at multiple sites simultaneously. Even when the number of fragments per individual DNA sample can be increased to several hundreds to a thousand by using two-dimensional electrophoretic separation of fragments, this will only allow small genomes (with genomes of up to 5 Mbp) to be analysed comprehensively i.e. to analyse all restriction fragments. For larger genomes the number of restriction fragments generated will be too high. Therefore techniques have to be employed which visualize selected sets of polymorphic DNA sequences dispersed over the genome of interest. Such sequences can include families of repetitive sequences or can consist of recognition sites for restriction enzymes. Techniques to visualize either of these can be divided in three categories:

(i) Hybridization analysis of electrophoretically separated genomic DNA fragments using repetitive sequences as probes (Southern, 1975; Gusella et al., 1984; Jeffreys et al., 1985; Porteous et al., 1986; Uitterlinden et al., 1989; Sugino et al., 1992; Meulenbelt, I., Wapenaar, M.C., Patterson, D., Vijg, J., and Uitterlinden, A.G., in press).

(ii) Inter-repeat PCR using primers specific for particular repetitive motifs derived either from known repetitive elements (Nelson et al., 1989; Ledbetter et al., 1990; Uitterlinden, A.G., Meulenbelt, I., Patterson, D., and Vijg, J., submitted) or chosen randomly (Williams et al., 1990).

(iii) Selective end-labelling of restriction fragments e.g. by first applying enzymes with relatively few recognition sites (rare cutters) followed by labelling and digestion with a more frequently cutting enzyme. This was first described for analyzing cosmids (Carrano et al., 1989) but has also been applied for the analysis of total mammalian genomes (Yi et al., 1990; Hatada et al., 1991).

These methods can be combined to comprehensively analyse any given genome of interest. Since the occurrence of repetitive sequences varies with different genomes, both the hybridization approach and the irPCR approach are somewhat limited in their

application. This is, however, counterbalanced by the amount of polymorphism they can detect in contrast to the limited DNA sequence polymorphism in restriction enzyme recognition sites. These factors, however, remain to be evaluated in the practice of analysing genomes.

4.5 APPLICATIONS OF TWO-DIMENSIONAL DNA TYPING

Comprehensive genetic analysis of large genomes can be accomplished by combining a two-dimensional electrophoretic method to resolve the maximum number of restriction fragments and the use of particular probes in hybridization analysis or primers in PCR analysis to detect polymorphic dispersed repetitive sequences. The application of the DGGE dimension in the 2-D analysis can effectively circumvent the problem of lack of polymorphism at certain sites in the genome.

Two-dimensional DNA typing is capable of linking detected sequence variants to physical mapping information in several ways. First, all fragments constituting a relatively small genome (or large region from a genome) can be completely resolved and simultaneously analysed for DNA sequence variation. Second, for larger genomes a fraction of sequences which have known physical map locations can be analysed simultaneously as anchor points for DNA sequence variation. Ideal anchor points in this respect are repetitive sequences which have a dispersed distribution through the genome.

Two-dimensional DNA typing can be applied in many areas of genome research. These include:

(i) The direct mutational analysis of DNA fragments (generated by restriction enzyme digestion or by PCR amplification) derived from contiguous stretches of DNA up to a length of several million basepairs. Examples are a single genome in the case of lower organisms (Fischer and Lerman, 1979) or large genomic regions like a gene with introns and exons, in the case of higher organisms.

(ii) Hybridization and/or PCR analysis of large genomes using repetitive sequences specific for polymorphic loci in genetic studies or in studies of somatic instability. Examples are the application of core probes (Uitterlinden et al., 1989) and inter-repeat

PCR (Uitterlinden et al., 1991a).

(iii) analysis of complex mixtures of DNA molecules derived from pcpulations of different individuals or different species. Examples are the analysis of 16 S RNA to fingerprint populations composed of different microbial species (Muyzer, G., de Waal, E.C., and Uitterlinden, A.G., in press), and the analysis of cDNA populations specifying mRNA populations.

CHAPTER 5

DISCUSSION

Delineation of the DNA sequence of the genome of a species allows to understand its molecular biology. By making an inventory of the variant forms of the genome sequence one is able to deduce causal relationships between the sequence diversity and the observed phenotypes. Defining these variant forms of particular DNA sequences can be used, e.g., to track disease genes, understand the mechanism of diseases, study somatic instability of the genome in cancer and aging, define mutagenic substances.

A wealth of techniques is now available which allow for the detection of DNA sequence variation. They have been used to define individual genetic variation and relate this to phenotypic variants but only for some loci. The pace with which these undertakings are currently progressing allows complete sequence information on Homo sapiens to be obtained not before 2005 even according to optimistic estimates. This is mainly due to the nature of the analysis of sequence variation. For most techniques described in the introductory sections, only some variations can be measured and at only one or a few sites in the genome simultaneously. To achieve sufficient genome coverage such analyses are now applied serially, which is time-consuming.

In molecular biology examples of parallel processing approaches can be found in the construction of physical maps of the genome of an organism. Lehrach (Lehrach et al., 1990) has deviced a system whereby many cosmid clones, spotted on a nylon membrane in a regular grid are hybridized subsequently with many different short oligonucleotides. The oligonucleotides have been chosen such that only a fraction of all cosmids will hybridize. By taking several hundreds to thousands of oligo's, overlap among cosmids can be deduced from overlap in hybridization signature of each cosmid. In this way ten thousands of cosmids can be processed simultaneously. In a model system a contig map of overlapping cosmid clones could be generated obtained from the Herpex simplex virus type 1 (HSV-1) genome (Craig et al., 1990). In a similar way the system was applied to construct a complete map of the Schizosaccharomyces pombe genome by hybridizing cosmids to a spotted YAC library (Maier et al., 1992) and to isolate chromosome 21-

specific YACs (Ross et al., 1992).

For parallel detection of DNA sequence variation in complex molecules such as the human genome, no efficient analysis system was available up to now. Two-dimensional electrophoresis of DNA molecules according to independent criteria would be best suited for increasing power of the analysis. The 2-D DNA separation system described in 1979 by Fischer and Lerman is unique in combining two independent separation criteria. The second dimension criterium, DGGE has been shown to be a very powerful method of detecting DNA sequence variation and represents the most comprehensive system for analysis of complex DNA molecules or for complex mixtures of DNA molecules currently available.

The 2-D DNA electrophoretic separation principle according to Fischer and Lerman was considered to be very suitable for assessing DNA sequence variation at many sites in parallel in large genomes. For the analysis of small genomes such as those from bacteria, ethidium-bromide staining of a 2-D DNA separation pattern can visualize all restriction fragments of a given digest. However, for analysis of larger genomes the existing 2-D separation method had to be extended by including a hybridization step. This allows parts of the genome to be displayed which can be selected on the basis of the probe used. The use of consensus probes detecting repetitive elements in this respect results in a parallel analysis system. This approach allows for the first time to simultaneously analyse hundreds of sites in a complex genome for DNA sequence variation.

From the many types of repetitive elements micro- and minisatellites were chosen as probes for hybridization analysis for a number of reasons:

(1) The copy number of the loci detected by a single core probe (i.e. several hundreds) allows all members belonging to the set to be resolved on a standard size 2-D gel (180 x 180 mm).

(2) They display high levels of (VNTR) polymorphism.

(3) When particular (i.e. GC-rich) core probes were applied the homologous sequences were found to display advantageous melting behaviour in the DGGE gel when compared

to other repetitive elements (Uitterlinden et al., 1989; Uitterlinden and and Vijg, 1991a; Uitterlinden and Vijg, 1991b). This is due to the fact the micro- and/or minisatellite sequences detected by these core probes are GC-rich, i.e. have stretches which contain more than 50% GC's. These micro- and/or minisatellite sequences function as a natural clamp, allowing optimal separation in the DGGE gel based on the sequences adjacent to the GC-rich tract which constitute the first melting domain of the restriction fragment containing these sequences.

(4) The sequential use of the many different core probes, each detecting a different set of loci, allows the screening of thousands of loci by (re)hybridization experiments. Thus, genomes can be analysed at different levels of resolution.

The work described in this thesis shows that it is possible to analyse hundreds of micro and/or minisatellite sequences simultaneously by resolving genomic DNA digests by two-dimensional electrophoresis according to the scheme originally designed by Fischer and Lerman (1979), followed by hybridization analysis using consensus or core probes. In keeping with the occurrence of these sequences throughout the animal and plant kingdom it has been possible to generate 2-D spot patterns not only in human genomic DNA (Uitterlinden et al., 1989, 1991a and 1991b), but also of several other species including rodents (Uitterlinden et al., 1991a; Slagboom et al., 1991), cattle (Uitterlinden et al., 1991b; Morolli, B., Molijn, A.C., te Meerman, G., den Daas, J.H.G., Uitterlinden, A.G., Mullaart, E., and Vijg, J., submitted), pigs (unpublished observations), tomato (Uitterlinden et al., 1991b), and yeasts and fungi (unpublished observations). The different distribution of different sets of these sequences throughout almost any genome allows effective scanning for DNA sequence variations at the site of the micro- and/or minisatellite. The advantageous electrophoretic behaviour of micro and/or minisatellite sequences in DGGE gels can be explained by the repeat units being GC-rich relative to the adjacent sequences. It was demonstrated for several cloned human VNTR loci that as a result different size alleles of a VNTR-locus show so-called isothermal electrophoretic behaviour. While their size difference causes them to migrate to different positions in the first dimension neutral sizing gel, they migrate to almost identical positions in the second dimension DGGE. Since a micro- or minisatellite core

probe detects many VNTR loci simultaneously, a 2-D spot pattern is in fact composed of many different size alleles occupying different positions in the denaturing gradient gel due to differences in the VNTR flanking sequences. It follows that alleles of the same locus will migrate to the same position in the second dimension.

5.1 APPLICATIONS OF MICRO- AND MINISATELLITE SEQUENCES IN 2-D DNA TYPING

A particular advantage of two-dimensional DNA typing in genome comparison studies is its efficiency. It is efficient since many genetic markers dispersed over the genome are analysed simultaneously. Increased resolution in defining the region of interest can be relatively easily obtained by rehybridization with additional core probes. It is also efficient since only a single genomic DNA digest is used for the analysis and only spots of interest are analysed further. Particular spot variants can then be identified as markers for a region in the genome. One or more of the genes located within this region are then suspected to be involved in the biological process under study. Another advantage of 2-D DNA typing is that its application does not require the creation of a genetic map. This is especially useful for the analysis of organisms, such as cattle and other economically important animals and plants, for which such maps are not yet available.

A disadvantage is that no a priori information is available on the genomic position of spots of interest. Such information has to be obtained, either by isolating the spot of interest and subsequent mapping, or by studies on co-segregation of the spots, with themselves or with other known markers alleles (te Meerman, G., Mullaart, E., Scheffer, H., Uitterlinden, A.G., and Vijg, J., submitted). A procedure has now been developed that allows spot variants of interest to be directly isolated from a 2-D gel and mapped (Molijn, A.C., Borglum, A., Verwest, A.M., Mullaart, E., Kruse, T., Vijg, J., and Uitterlinden, A.G., submitted). The procedure is based on eluting DNA from a position corresponding to the spot variant of interest in a duplicate 2-D gel and subjecting the DNA to PCR amplification using primers for the core sequence used to generate the spot pattern.

Two-dimensional DNA typing using micro- and minisatellite core probes can be applied in the following areas:

1. Analysis of tumor genomes
2. Mutation analysis
3. Comparison of young vs. old cells
4. Genetic linkage analysis
5. Genetic association analysis

1. Analysis of tumor genomes

It has been shown that a number of different chromosomal regions display allelic imbalance (gain or loss of one of the alleles at a locus) in different types of tumours (Vogelstein et al., 1985). The studies of allelic imbalance are based on the use of RFLP probes which allows one to identify the lost allele in the tumor tissue. The allele then represents a larger area of the chromosome in which genes crucial to the arisal and development of the tumour are likely to be located. Genome searches to identify chromosomal regions altered in the tumour tissue have consisted of the serial application of different locus-specific probes, the number, informativity and location of which determines the efficiency and accuracy of the search. To increase the number of informative cases, locus-specific VNTR probes have been used in several studies. The application of micro- and minisatellite core probes greatly increases the number of sites that can be detected simultaneously in a search for alterations in the tumor genome. By means of 2-D DNA typing the number of sites in the genome which can be simultaneously analysed for variations can be increased to several thousands, simply by hybridization analysis with many different core probes. Two-dimensional DNA typing analysis of human tumours has successfully been performed and is still ongoing for breast tumours (Hovig,E., Mullaart, E., Borresen, A.-L., Uitterlinden, A.G., and Vijg, J., in press; Verwest, A.M., Molijn, A.C., Andersen, T.I., Borresen, A.-L., Uitterlinden, A.G., Mullaart, E., and Vijg, J., submitted; Slagboom, P., personal communication), lung tumors (unpublished observations) and brain tumours (Nürnberg, P., personal communication). Genomic instability of tumors as detected by 2-D DNA typing can be exploited in two ways. First, the 2-D spot pattern itself can be of prognostic value since the changes as a whole can be

of prognostic value. In this respect, the choice of the core probe, which may detect, e.g., telomeric, centromeric, truly dispersed loci, can be important; that is different areas of the human genome can display different frequencies and types of variation. Second, for some specific spot changes locus-specific markers can be developed identifying the region undergoing genetic alterations such as deletions or amplification. Here positional cloning will ultimately identify the gene(s) involved in the pathology of the tumour. In the analysis of breast tumors several specific spot variations have been identifed in more than one breast cancer patient (Verwest, A.M., Molijn, A.C., Andersen, T.I., Börresen, A.-L., Uitterlinden, A.G., Mullaart, E., and Vijg, J., submitted). The chromosomal areas from which these spot variants are derived might therefore contain genes important for tumor progression.

2. Mutation analysis

Current approaches for biomonitoring humans thought to be at risk for DNA damage include mutation analysis at selectable genes such as HPRT and MHC. Although the results thusfar obtained with such systems have greatly contributed to our understanding of the mechanisms underlying mutagenesis in vivo, these methods nevertheless suffer from the disadvantages that (a) selection of cells in culture is involved; (b) due to their inherent stability coding genes could be insensitive indicators for permanent genetic damage; and (c) mutations are measured at only a single site in the genome (See also Neel et al., 1993 for a discussion on this in relation to the study of mutation in atomic bomb survivors). An alternative would be to look directly (i.e. without selection) to large numbers of naturally unstable (hypermutable) mutational target sequences spread over the genome. A category of naturally present DNA sequences that offer such a possibility are micro- and minisatellites. By employing 2-D DNA typing, which is capable of resolve virtually all members belonging to a particular set of micro- and minisatellites, the genome can be "scanned" for variations in micro- and minisatellite DNA sequences.

An important issue in this respect is whether or not micro- and minisatellite regions are subject to similar types of mutation processes as other (e.g. coding) regions in the genome. One way of evaluating this aspect is to analyze induced mutations in cloned cells caused by agents which are known to cause mutations in for example the HPRT gene. An example of the potential use of 2-D DNA typing to detect genetic change as a

result of the challenge of cells with genotoxic agents is presented in addendum 4. Since many different core probes exist detecting many different loci spread over the genome, this approach offers the possibility to provide a better description of the mutational spectrum of genotoxic substances than existing genotoxicity tests do. In addition, the micro- and minisatellite loci in general show higher spontaneous mutation rates, which may make them more susceptible to genotoxic agents. In interpreting the results obtained on mutation frequencies at micro- and minisatellite loci it should be realized that the sequences detected by core probes comprise an array of different loci. This might be the result of differences in sequence composition of the repeat unit of which a locus is composed, the length of alleles of the locus, the location in coding vs. non-coding regions in the genome, and the location on the chromosomes (e.g. subtelomeric vs. interstitial).

3. Comparison of young vs. old cells

2-D DNA typing also offers immediate advantages when looking for unknown changes in genomes derived from young vs. old cells. Somatic cells are suspected in many aging theories of undergoing substantial genetic alterations as a result of continuous exposure to both endogenous and exogenous genotoxic substances (Vijg and Gossen, 1993). With time some of these changes might ultimately damage genes essential for cell metabolism, thereby leading to an increasing number of cells showing departures from normal functioning. Untill recently, however, molecular genetic techniques which could detect unknown changes randomly occurring in the genome were lacking. In this respect, micro- and minisatellite sequences could be sensitive indicators for the detection of changes in genomic instability with age similar to what has been described under "mutation analysis" (Vijg and Uitterlinden, 1987). An example of the analysis of young vs. old cells involving 2-D DNA typing of genomes is provided by the work of Slagboom et al. (1991). They analysed skin fibroblast clones derived from young and old rats and observed mutations to occur at a frequency of $2.7.10^{-3}$ per micro- or minisatellite containing fragment. However, no increase of this frequency was observed in the fibroblast clones derived from aged vs. young rats.

4. Genetic linkage studies

Since the discovery of human DNA polymorphisms in the late seventies, the large

reservoir of DNA markers enabled genetic linkage studies. These studies rely on the observation in pedigrees in which a particular disease is segregating, of co-inheritance of a particular variant of a genetic marker locus together with the disease locus. If such co-inheritance is observed more frequently than statistically expected the two loci are said to be genetically linked. This means that the two loci show very infrequent recombination in meiosis as a result of physical proximity on one of the chromosomes.

Genetic linkage studies require the analysis of many marker loci in several individuals of disease pedigrees. In this respect, 2-D DNA typing allows for the rapid simultaneous screening of hundreds of polymorphic marker loci. For linkage analysis each 2-D spot variant is treated as a dominant marker system in which the other allele is unknown. This reduces the genetic information of the markers scored in a 2-D DNA typing experiment by a factor of about 2 meaning that twice as many individuals have to be analysed to obtain the same statistical likelihood of observing genetic linkage (te Meerman, G., Mullaart, E., Scheffer, H., Uitterlinden, A.G., and Vijg, J., submitted). The lack of information on the loci from which particular 2-D spot variants are derived can be solved by direct isolation and mapping of the spots from the gel, and by co-segregation studies with known markers. The observation that alleles of a VNTR locus show very similar electrophoretic behaviour in the DGGE gel (Uitterlinden and Vijg, 1991a) will facillitate this. Thus, 2-D DNA typing can be applied in genetic linkage analysis in pedigrees segregating for a particular trait. In view of the occurrence of polymorphic micro- and minisatellites in many different species 2-D DNA typing could be widely applicable for this purpose.

5. Genetic association studies

Genetic linkage analysis can be applied to any monogenetic trait in humans, animals or plants, if sufficient pedigree material is available. In multifactorial diseases, however, alleles at a number of different gene loci interact and determine (in combination with environmental factors) an individual's risk of a trait or disease. Many genetic traits, including the most frequent ones, in the human population as well as in economically important animals or plants, are multifactorial. In humans, multifactorial diseases (e.g., heart diseases, cancers, diabetes, arthritis, high blood pressure) are diseases with an unknown genetic background, i.e. which genes interact. Two-dimensional DNA typing

can analyse thousands of DNA polymorphisms by detecting presence or absence of alleles at VNTR loci. The application of 2-D DNA typing for association studies has the potential to reveal presence of particular gene variants through linkage disequilibrium. The analysis proceeds in two phases: an search for variations in allele frequency associations is followed by developing a locus-specific marker for the spot variant with which association is found (Molijn et al., 1993). With this fully informative genetic marker initial associations can be confirmed. However, even when many individuals are analysed, 2-D DNA typing will not always generate statistically significant results since mostly anonymous genetic markers are analysed. In this respect it would be much more advantageous to directly screen only the coding part of the genome in the form of cDNAs (see below: future applications).

5.2 FUTURE APPLICATIONS

Also in other areas of genome research different modalities of 2-D DNA typing can be applied. Three of these will be briefly discussed here. They are in the field of analysis of particular chromosomal regions by inter-repeat PCR (region-specific 2-D DNA typing), analysis of populations of RNA molecules (2-D cDNA typing) and high resolution analysis of large genomic regions (gene scanning).

Region-specific 2-D DNA typing

Particular advantages of repetitive sequences such as Alu I and Kpn I in the human genome and similar repeats in other genomes (higher copy number, dispersed distribution) can also be exploited by the 2-D DNA electrophoresis protocol as described here by using the inter-repeat PCR approach. The irPCR approach allows unique (inter-repeat) sequences bordered by dispersed repeats to be selectively amplified from total genomic DNA and from sources containing reduced amounts of genomic DNA such as somatic cell hybrids and YACs. In a model study it was shown that chromosome-specific irPCR fragments can be detected in 2-D separation patterns (Uitterlinden et al., 1991a; Uitterlinden, A.G., Meulenbelt, I., Patterson, D., and Vijg, J., submitted).

This application allows physical mapping information (dictated by the origin of the irPCR products used as probe, such as YAC or radiation hybrid, etc.) to be combined

with genetic mapping information by following transmission of 2-D spot variants in pedigrees. Although several types of polymorphism for the type of repeat used in this approach have been described in the literature (deletion/insertion of complete repeat units, VNTR type of polymorphism for polyA stretches in the repeat) more extensive studies have to demonstrate the feasibility of such an approach.

2-D cDNA typing

Apart from the analysis of genomic DNA it is also possible to analyse only the protein encoding part of the genome by means of analysing cDNA prepared from mRNA populations. In view of the complexity of mRNA populations in cells, 2-D DNA typing can offer a high resolving capacity to analyse many individual mRNA species (as cDNAs) in parallel from a given cell type. For comparative purposes such cDNA libraries can be prepared from different cell types and from different individuals. The introduction of amplification primers and application of probes (such as particular sequence motifs) in hybridization experiments allows particular subsets to be analysed. Of particular interest might be the analysis of patients from disease pedigrees such as Fragile X and Myotonic Dystrophy. The defect causing these diseases is related to trinucleotide microsatellite polymorphisms (Verkerk et al., 1991; Fu et al., 1992). It thas been shown that patients suffering from these diseases carry extreme length mutations in $[CNG]_n$ stretches in the disease genes. Furthermore, the phenomenon of genetic anticipation (increasing severity of the disease in patients further down the pedigree) has been shown to correlate with increasing length of the trinucleotide stretch. Using core probes for such trinucleotide motifs in the 2-D DNA typing analysis of genomic DNA isolated from these partients and patients in pedigrees of other diseases showing anticipation, it might be possible to directly detect similar mutations. In addition, such a screening can be performed on 2-D separation patterns of total cDNA prepared from mRNA from particular tissues of such patients. In general, two-dimensional cDNA typing allows the simultaneous analysis of polymorphisms in many candidate genes and presents a useful tool to perform genetic association studies.

TABLE 4 MUTATIONS IN HUMAN DISEASE GENES

DISEASE	Fragile-X	AAT*	CF*	DMD*	Haem.A	Haem.B
INCIDENCE	1:1,250(♂)	1:1,600	1:3,600	1:3,500(♂)	1:10,000(♂)	1:30,000(♂)
CARRIER		1:20	1:25			
CHROMOSOME	Xq27.3	14q24.3	7q31	Xp21	Xq28	Xq2.7
GENE	FMR-1	Pi	CFTR	dystrophine	Factor VIII	Factor IX
LENGTH (kb)			250	2,300	186	33
#EXONS			27	76	26	8
mRNA(nucl.)	4,800		6,500	14,000	7,053	1,383
MUTATIONS						
-population frequency				1.10^{-4}	5.10^{-5}	3.10^{-6}
-type:						
% INS/DEL (large)			1%	66%	5%	10%
% POINT MUTATION	1%		99%		95%	90%
% VNTR	99%					
-% spontaneous				33%	50%	
-# different mutations			>300	>300	>100	>115
-prevalent mutations		Z	30-70% δF508			

* AAT = α_1-anti trypsine; CF = Cystic Fibrosis; DMD = Duchenne Muscular Dystrophy

Gene scanning

Finally, also in relation to the ever increasing number of genes that is being identified and of which their molecular genetic involvement in disease processes is being dissected, 2-D DNA typing can be of help in the comprehensive analysis of large regions in the genome. In Table 4 an overview is presented of some of the genes identified until now and specific variants of which are known to be involved in disease processes. High resolution mutation analysis of especially the larger genes is presently impossible but might become feasible by 2-D DNA typing of these regions. These regions can include single but large to very large genes but also the analysis of sets of genes which interact or are involved in particular disease processes. Examples of the latter are the genes involved in cardiovascular disorders such as Apo B, LDL receptor etc.

Accurate mutation analysis of these regions is guaranteed by the inclusion of the DGGE separation which is the most comprehensive mutation scanning method currently available. The format of analysis is likely to involve PCR amplification of fragments of interest (to be selected on the basis of optimal melting behaviour in the DGGE gel), inclusion of a GC-clamp, separation in a neutral sizing gel and subsequent separation in a second dimension DGGE gel. A possibility to allow automated on-line detection of fragments (Carrano et al., 1989) is to switch dimensions and first separate in the DGGE gel and subsequently separate in a neutral sizing gel.

5.3 RECAPITULATION

In conclusion, it can be stated that the 2-D DNA analysis system as described in this thesis will be of help in the study of biological problems which require the scanning of large stretches of DNA for sequence deviations. Examples include tumourigenesis, aging, and mutagenesis, in which chromosomal areas have to be defined which might undergoing alteration upon a challenge. In addition, the method may be of help in performing genetic linkage or association studies where screening many individuals for a large number of polymorphic sites spread over the genome is required. Micro- and minisatellites are particularly useful as probes in the hybridization analysis of 2-D gels since they detect many polymorphic loci spread over the genome. Other repetitive elements, e.g. with higher copy numbers, are also suitable for 2-D DNA typing provided

inter-repeat PCR amplification is applied to selectively visualize regions bordered by these repeats. For all these applications the 2-D DNA typing method presents an increase in speed and efficiency when compared to existing methodology. Together with the development of automated 2-D gel electrophoresis equipment and image analysis software, the high genome-scanning efficiency of 2-D DNA typing is likely to result in the rapid isolation of markers for genetic traits, including diseases with genetic components, in many different species. Future applications of the 2-D DNA analysis method include the analysis of only the coding regions of the genome (2-D cDNA typing) and the high resolution analysis of large contiguous (coding) regions in the genome ("gene scanning"). In concert with the increasing knowledge on genome organization and sequence, these different 2-D DNA typing formats are expected to provide further insights in the causal relationship between genome variation and the individual phenotype.

REFERENCES

Abrams, E.S., Murdaugh, S.E., and Lerman, L.S. (1990) Comprehensive detection of single base changes in human genomic DNA using denaturing gradient gel electrophoresis and a GC clamp. Genomics 7, 463-475.

Ainsworth, P.J., Surh, L.C., and Coulter-Mackie, M.B. (1991) Diagnostic single strand conformational polymorphism (SSCP): a simplified non-radioisotopic method as applied to a Tay-Sachs B1 variant. Nucleic Acids Res. 19, 405.

Akli, S., Chelly, J., Lacorte, J.-M., Poenaru, L., and Kahn, A. (1991) Seven novel Tay-Sachs mutations detected by chemical mismatch cleavage of PCR-amplified cDNA fragments. Genomics 11, 124-134.

Ali, S., Müller, C.R., and Epplen, J.T. (1986) DNA fingerprinting by oligonucleotide probes specific for simple repeats. Hum.Genet. 74, 239-243.

Ali, S., and Wallace, B. (1989) Enzymatic synthesis of DNA probes complementary to a variable number tandem repeat locus. Anal. Biochem. 179, 280-283.

Anderson, R.D., Veigl, M.L., and Sedwick, D. (1990) Rapid mapping of deletion and duplication mutations by the polymerase chain reaction. Biochem. Biophys. Res. Comm. 170, 416-420.

Anderson, N.L., Taylor, J., Scandora, A.E., Coulter, B.P., and Anderson, N.G. (1981) The Tycho system for computer analysis of two-dimensional gel electrophoresis patterns. Clin. Chem. 27, 1807-1820.

Anderson, S., Bankier, A.T., Barrell, B.G., deBruyn, M.H.L., Coulson, A.R., Droin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schrier, P.H., Smith, A.J.H., Staden, R., and Young, I.G. (1981) Sequence and organization of the human mitochondrial genome. Nature 290, 457-461.

Attree, O., Vidaud, D., Amselem, S., Lavergne, J.-M., and Goossens, M. (1989) Mutations in the catalytic domain of human coagulation factor IX: rapid characterization by direct genomic sequencing of DNA fragments displaying an altered melting behaviour. Genomics 4, 266-272.

Au, L.-C., Ts'o, P.O.P., and Yi, M. (1990) Monitoring mammalian genome rearrangements with mid-repetitive sequences as probes. Anal. Biochem. 190, 326-330.

Barker, D., Green, P., Knowlton, R., Schumm, J., Lander, E., Oliphant, A., Willard, H.F., Akots, G., Brown, V., Gravius, T., Helms, C., Nelson, C., Parker, C., Rediker, K., Rising, M., Watt, D., Weiffenbach, B., and Donis-Keller, H. (1987) Genetic linkage map of human chromosome 7 with 63 DNA markers. Proc, Natl. Acad. Sci. USA 84, 8006-8010.

Bautsch, W. (1988) Rapid physical mapping of the Mycoplasma mobile genome by two-dimensional field inversion gel electrophoresis. Nucleic Acids Res. 16, 24, 11461-11467.

van Belkum, A., Trommelen, G.J.J.M., and Uitterlinden, A.G. (1991) Minisatellite-like repeat sequences in the genome of rodent malaria parasites. Acta Leidensia 60, 1, 83-91.

van Belkum, A., Ramesar, J., Trommelen, G.J.J.M., and Uitterlinden, A.G. (1992) Mini- and micro-satellites in the genome of rodent malaria parasites. Gene 118, 81-86.

Bell, L., and Byers, B. (1983) Separation of branched from linear DNA by two-dimensional gel electrophoresis. Anal. Biochem. 130, 527-535.

Bell, G.I., Selby, M.J., and Rutter, W.J. (1982) The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. Nature 295, 31-35.

Bellanné-Chantelot, C., Barillot, E., Lacroix, B., Le Paslier, D., and Cohen, D. (1991) A test case for physical mapping of human genome by repetitive sequence fingerprints: construction of a physical map of a 420 kb YAC subcloned into cosmids. Nucleic Acids Res. 19, 505-510.

Bewsey, K.E., Johnson, M.E., and Huff, J.P. (1991) A method for the rapid detection of single base polymorphisms in genomic DNA. Biotechniques 10, 725.

Blanchetot, A. (1991) A Musca domestica satellite sequence detects individual polymorphic regions in insect genome. Nucl. Acids Res. 19, 929-932

Boehm, T.L.J., and Drahovsky, D. (1984) Two-dimensional restriction mapping by digestion with restriction endonucleases of DNA in agarose and polyacrylamide gels. J. Biochem. Biophys. Meth. 9, 153-161.

Börresen, A.-L., Hovig, E., and Brogger, A. (1988) Detection of base mutations in genomic DNA using denaturing gradient gel electrophoresis (DGGE) followed by transfer and hybridization with gene-specific probes. Mut. Res. 202, 77-83.

Börresen, A.-L., Hovig, E., Smith-Sorensen, B., Malkin, D., Lystad, Andersen, T.I., Nesland, J.M., Isselbacher, K.J., and Friend, S.H. (1991) Constant denaturant gel electrophoresis as a rapid screening technique for p53 mutations. Proc. Natl. Acad. Sci. USA 88, 8405-8409.

Botstein, D., White, R., Skolnick, M., and Davis, R. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32, 314-331.

Boylan, K.B., Ayres, T.M., Popko, B., Takahashi, N., Hood, L., and Prusiner, S.B. (1990) Repetitive DNA $(TGGA)_n$ to the human myelin basic protein gene: a new form of oligonucleotide repetitive sequence showing length polymorphism. Genomics 6, 16-22.

Brilliant, M.H., Gondo, Y., and Eicher, E.M. (1991) Direct molecular identification of the mouse pink-eyed unstable mutation by genome scanning. Science 252, 566-569.

Britten, R.J., and Kohne, D.E.(1968) Repeated sequences in DNA. Science 161, 529-540.

Burghes, A.H.M., Logan, C., Hu, X., Bellfall, B., Worton, R.G., and Ray, P.N. (1987) A cDNA clone from the Duchenne/Becker muscular dystrophy gene. Nature 328, 434-437.

Burke, D.T., Carle, G.F., and Olson, M.V. (1987) Cloning of large segments of DNA into yeast by yeast artificial chromosome vectors. Science 236, 806-812.

Burmeister, M., diSibio, G., Cox, R., and Myers, R.M. (1991) Identification of polymorphisms by genomic denaturing gradient gel electrophoresis: application to the proximal region of human chromosome 21. Nucleic Acids Res. 19, 1475-1481.

Burton, F.H., Loeb, D.D., Chao, S.F., Hutchinson III, C.A., and Edgell, M.H. (1985) Transposition of a long member of the L1 major interspersed DNA family into the mouse beta globin gene locus. Nucl. Acids Res. 13, 5071-5084.

Caetano-Anollés, G., Bassam, B.J., and Gresshoff, P.M. (1991) DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Bio/technology 9, 553-556

Cantor, C.R. (1990) Orchestrating the Human Genome Project. Science, 248, 49-51.

Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H., and Goeddel, D.V. (1983) Complete nucleotide sequences of the T24 human bladder carcinoma oncogene and its normal homologue. Nature, 302, 33-37.

Cariello, N.F., Scott, J.K., Kat, A.G., and Thilly, W.G. (1988a) Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: $HPRT_{Munich}$. Am. J, Hum. Genet. 42, 726-734.

Cariello, N.F., Keohavong, P., Sanderson, B.J.S., and Thilly, W.G. (1988b) DNA damage produced by ethidium bromide staining and exposure to ultraviolet light. Nucleic Acids Res. 16, 4157.

Cariello, N.F., Swenberg, J.A., and Skopek, T.R. (1991) Fidelity of Thermococcus litoralis DNA polymerase (Vent™) in PCR determined by denaturing gradient gel electrophoresis. Nucleic Acids Res. 19, 4193-4198.

Cariello, N.F., Swenberg, J.A., De Bellis, A., and Skopek, T.R. (1991a) Analysis of mutations using PCR and Denaturing gradient gel electrophoresis. Environ. Mol. Mutagen. 18, 249-254.

Carle, G.R., Frank, M., and Olson, M.V. (1986) Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. Science 232, 65-68.

Carothers, A.M., Urlaub, G., Grunberger, D., and Chasin, L.A. (1988) Mapping and characterization of mutations induced by benzo[*a*] pyrene diol epoxide at dihydrofolate reductase locus in CHO cells. Som. Cell Mol. Genet. 14, 169-183.

Carrano, A.V., Lamerdin, J., Ashworth, L.K., Watkins, B., Branscomb, E., Slezak, T., Raff, M., De Jong, P.J., Keith, D., McBride, L., Meister, S., and Kronick, M. (1989) A high-resolution, fluorescence-based, semiautomated method for DNA-fingerprinting. Genomics 4, 129-136.

Caskey, C.T. (1987) Disease diagnosis by recombinant DNA methods. Science 236, 1223-1229.

Cawthon, R.M., Weiss, R., Xu, G., Viskochil, D., Culver, M., Stevens, J., Robertson, M., Dunn, D., Gesteland, R., O'Connell, P., and White, R. (1990) A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure, and point mutations. Cell 62, 193-201.

Chamberlain, J.S., Gibbs, R.A., Ranier, J.E., Ngyuen, P.N., and Caskey, C.T. (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Res. 16, 11141-11156.

Chang, J.C., and Kan, Y.W. (1982) A sensitive new prenatal test for sickle cell anemia. N. Engl. J. Med. 307,30-32.

Chebloune, Y. Trabuchet, G., Poncet, D., Cohen-Solal, M., Faure, C., Verdier, G., and Nigon, V.M. (1984) A new method for detection of small modifications in genomic DNA, applied to the human δ-ß globin gene cluster. Eur. J. Biochem. 142, 473-480.

Chen, C.W., and Thomas, C.A. (1983) A rapid two-dimensional fractionation of DNA restriction fragments. Anal. Biochem. 131, 397-403.

Church, G.M., and Gilbert, W. (1984) Genomic sequencing. Proc. Natl. Acad. Sci. USA 81, 1991-1995.

Collins, M., and Myers, R.M. (1987) Alterations in DNA helix stability due to base modifications can be evaluated using denaturing gradient gel electrophoresis. J. Mol. Biol. 737-744.

Conner, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L., and Wallace, R.B. (1983) Detection of sickle cell ß-globin allele by hybridization with synthetic oligonucleotides. Proc. Natl. Acad. Sci. USA 80, 278-282.

Cotton, R.G.H., Rodrigues, N.R., and Campbell, R.D. (1988) Reactivity of cyosine and thymine in single-basepair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc. Natl. Acad. Sci. USA 85, 4397-4401.

Cotton, R.G.H., and Campbell, R.D. (1989) Chemical reactivity of matched cytosine and thymine bases near mismatched and unmatched bases in a heteroduplex bewteen DNA starnds with multiple differences. Nucleic Acids Res. 17, 4223-4233.

Cotton, R.G.H. (1993) Current methods of mutation detection. Mut. Res. 285, 125-144.

Cox, R.D., and Lehrach, H. (1991) Genome mapping: PCR based meiotic and somatic cell hybrid analysis. BioEssays 13, 193-198.

Cox, R.D., Copeland, N.G., Jenkins, N.A., and Lehrach, H. (1991) Interspersed repetitive element polymerase chain reaction product mapping using a mouse interspecific backcross. Genomics 10, 375-384.

Craig, A.G., Nizetic, D., Hoheisel, J.D., Zehetner, G., and Lehrach, H. (1990) Ordering of cosmid clones covering the herpes simplex virus type 1 (HSV-1) genome: A test case for fingerprinting by hybridisation. Nucleic Acids Res. 18, 2653-

Danenberg, P.V., Horikoshi, T., Volkenandt, M., Danenberg, K., Lenz, H.-J., Shea, L.C.C., Dicker, A.P., Simoneau, A., Jones, P.A., and Bertino, J.R. (1992) Detection of point mutations in human DNA by analysis of RNA conformation polymorphism(s). Nucleic Acids Res. 20, 573-579.

Dean, M., White, M., Amos, J., Gerrard, B., Stewart, C, Khaw, K.-T., and Leppert, M. (1990) Multiple mutations in highly conserved residues are found in mildly affected cystic fibrosis patients. Cell 61, 863-870.

Demers, D.B., Odelberg, S.J., and Fisher, L.M. (1990) Identification of a factor IX point mutation using SSCP analysis and direct sequencing. Nucleic Acids Res. 18, 5575.

Devilee, P., Kievits, T., Waye, J.S., Pearson, P.L., and Willard, H.F. (1988) Chromosome-specific alpha satellite DNA: isolation and mapping of a polymorphic alphoid repeat from human chromosome 10. Genomics 3, 1-7.

Devoto, M., Ronchetto, P., Fanen, P., Telleria Orriols, J.J., Romeo, G., Goossens, M., Ferrari, M., Magnani, C., Seia, M., and Cremonesi, L. (1991) Screening for non-deltaF508 mutations in five exons of the cystic fibrosis transmembrane conductance regulator (CFTR) gene in Italy. Am. J. Hum. Genet. 48, 1127-1132.

Dianzani, I., Camaschella, C., Saglio, G., Forrest, S.M., Ramus, S., and Cotton, R.G.H. (1991a) Simultaneous screening for ß-Thallasemia mutations by chemical cleavage of mismatch. Genomics 11, 48-53.

Dianzani, I., Forrest, S.M., Camaschella, C., Saglio, G., Ponzone, A., and Cotton, R.G.H. (1991b) Screening for mutations in the phenylalanine hydroxylase gene from italian patients with phenylketonuria by using the chemical cleavage method: a new splice mutation. Am J. Hum. Genet. 48, 631-635.

Dickson, L.A., Pihlajaniemi, P., Deak, S., Pope, F.M., Nicholss, A., Prockop, D.J., and Myers, J.C. (1984) Nuclease S_1 mapping of a homozygous mutation in the carboxylpropeptide-coding region of the proα2(I)collagen gene in a patient with osteogenesis imperfecta. Proc. Natl. Acad. Sci. USA 81, 4524-4528.

Dietrich, W., Katz, H., Lincoln, S.E., Shin, H.-S., Friedman, J., Dracopoli, N., and Lander, E. (1992) A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131, 423-447.

Dlouhy, S.R., Schaff, D.A., Trofatter, J.A., Liu, H.-S., Stambrook, P., and Tischfield, J.A. (1989) Denaturing gradient gel analysis of single-base substitutions at a mouse Adenine Phosphoribosyl transferase splice acceptor site. Mol. Carc. 2, 217- 225.

Dombroski, B.A., Kazazian, H.H., and Scott, A.F. (1989) Isolation and characterization of a putative functional L1 retroposon. Am. J. Hum. Genet. 45, A183.

Donis-Keller, H., Green, P., Helms, C. et al. (1987) A genetic linkage map of the human genome. Cell 51, 319-337.

Drmanac, R., Labat, I., Brukner, I., and Crkvenjakov, R. (1989) Sequencing of megabase plus DNA by hybridisation: Theory of the method. Genomics 4, 114-128.

Dryja, T.P., Hahn, L.B., Cowley, G.S., McGee, T.L., and Berson, E.L. (1991) Mutation spectrum of the rhodopsin gene among among patients with autosomal dominant retinis pigmentosa. Proc. Natl. Acad. Sci. USA 88, 9370-9374.

Economou-Pachnis, E., Bergen, A.W., Warren, A.C., and Antonarakis, S.E. (1990) The polydeoxyadenylate tract of Alu repetitive element is polymorphic in the human genome. Proc. Natl. Acad. Sci. USA 87, 2951-2954.

Economou-Pachnis, A., Lohse, M.A., Furano, A.V., and Tsichlis, P.N. (1985) Insertion of long interspersed repeated elements at the *Igh* (immunoglobulin heavy chain) and *Mlvi-2* (Moloney leukemia virus integration 2) loci of rats. Proc. Natl. Acad. Sci. USA 82, 2857-2861.

Economou-Pachnis, A., and Tsichlis, P.N. (1985) Insertion of an Alu SINE in the human homologue of the Mlvi-2 locus. Nucleic Acids Res. 13, 8379-8387.

Edwards, A., Civitello, A., Hammond, H.A., and Caskey, C.T. (1991) DNA typing and genetic mapping with trimeric and tetrameric repeats. Am. J. Hum. Genet. 49, 746-756.

Edwards, A., Hammond, H.A., Jin, L., Caskey, C.T., and Chakraborty, R. (1992) Genetic variation at five trimeric and tetrameric repeat loci in four human population groups. Genomics 12, 241-253.

Emi, M., Hata, A., Robertson, M., Iverius, P.-H., Hegele, R., and Lalouel, J.-M. (1990) Lipoprotein lipase deficiency resulting from a nonsense mutation in exon 3 of the lipoprotein lipase gene. Am. J. Hum. Genet. 47, 107-111.

Epstein, N., Nahor, O., and Silver, J. (1990) The 3' ends of Alu repeats are highly polymorphic. Nucleic Acids. Res. 18, 4634.

Fanning, T.G., Hu, W.-S., and Cardiff, R.D. (1985) Analysis of tissue-specific methylation patterns of mouse mammary tumor virus by two-dimensional Southern blotting. J. Virol. 54, 3, 726-730.

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

Feinberg, A.P., and Vogelstein, B. (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. Addendum: 137, 266-267.

Finkelstein, J.E., Francomano, C.A., Brusilow, S.W., and Traystman, M.D. (1990) Use of denaturing gradient gel electrophoresis for detection of mutation and prospective diagnosis in late onset ornithine transcarbamylase deficiency. Genomics 7, 167-172.

Fischer, S.G., and Lerman, L.S. (1979a) Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. Cell 16, 191-200.

Fischer, S.G., and Lerman, L.S. (1979b) Two-dimensional electrophoretic separation of restriction enzyme fragments of DNA. Methods Enzymol. 68, 183-191.

Fischer, S.G., and Lerman, L.S. (1980) Separation of random fragments of DNA according to properties of their sequences. Proc. Natl. Acad. Sci. USA 77, 4420-4424.

Fischer, S.G., and Lerman, L.S. (1983) DNA fragments differing by single basepair substitutions are separated in denaturing gradient gels: correspondence with melting theory. Proc. Natl. Acad. Sci. USA 80, 1579-1583.

Flavell, R.A., Kooter, J.M., De Boer, E., Little, P.F.R., and Williamson, R. (1978) Analysis of the ß-δ-globin gene loci in normal and Hb Lepore DNA: Direct determination of gene linkage and intergene distance. Cell 15, 25-41.

Forrest, S.M., Dahl, H.-H., Howells, D.W., Dianzani, I., and Cotton, R.G.H. (1991) Mutation detection in phenylketonuria using the chemical cleavage of mismatch method: importance of using probes from both normal and patient samples. Am. J. Hum. Genet. 49, 175-183.

Fu, Y.-H., Kuhl, D.P.A., Pizutti, A., Pieretti M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M.H., Holden, J.J.A., Fenwick, R.G., Warren, S.T., Oostra, B.A., Nelson, D.L., and Caskey, C.T. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67, 1047-1058.

Fu, Y.-H., Pizutti, A., Fenwick, R.G., King, J., Rajnarayan, S., Dunne, P.W., Dubbel, J., Nasser, G.A., Ashizawa, T., de Jong, P.J., Wieringa, B., Korneluk, R., Perryman, M.B., Epstein, H.F., and Caskey, C.T. (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 255, 1256-1258.

Galinski, M.R., Arnot, D.E., Cochrane, A.H., Barnwell, J.W., Nussenzweig, R.S., and Enea, V. (1987) Cell 48, 311-319.

Ganguly, A., Rooney, J.E., Hosomi, S., Zeiger, A.R., and Prockop, D.J. (1989) Detection and location of single-base mutations in large DNA fragments by immunomicroscopy. Genomics 4, 530-538.

Ganguly, A., and Prockop, D.J. (1990) Detection of single-base mutations by reaction of DNA heteroduplexes with a water-soluble carbodiimide followed by primer extension: application to products from the polymerase chain reaction. Nucleic Acids Res. 18, 3933-3939.

Gardiner, K., Horisberger, M., Kraus, J., Tantravahi, U., Korenberg, J., Rao, V., Reddy, S., and Patterson, D. (1990) Analysis of human chromosome 21: correlation of physical and cytogenetic maps; gene and CpG island distributions. EMBO J. 9, 25-34.

Garrels, J.I. (1989) The QUEST system for quantitative analysis of two-dimensional gels. J. Biol. Chem. 264, 5283-5299.

Geever, R.F., Wilson, L.B., Nallaseth, F.S., Milner, P.F., Bittner, M., and Wilson, J.T. (1981) Direct identification of sickle cell anemia by blot hybridization. Proc. Natl. Acad. Sci. USA 78, 5081-5085.

Georges, M., Lequarré, A.-S., Castelli, M., Hanset, R., and Vassart, G. (1988) DNA fingerprinting in domestic animals using four different minisatellite probes. Cyt. Cell. Gent. 47, 127-131.

Ghanem, N., Fanen, P., Martin, J., Conteville, P., Yahia Cheriff, Z., Vidaud, M., and Goossens, M. (1992) Exhaustive screening of exon 10 CFTR gene mutations and polymorphisms by denaturing gradient gel electrophoresis: applications to genetic counseling in cystic fibrosis. Mol. Cell. Probes 6, 27-31.

Gibbs, R.A., and Caskey, C.T. (1987) Identification and localization of mutations at the Lesch-Nyhan locus by ribonuclease A cleavage. Science 236, 303-305.

Gibbs, R.A., Nguyen, P.-N., McBride, L.J., Koepf, S.M., and Caskey, C.T. (1989a) Identification of mutations leading to the Lesch-Nyhan syndrome by automated direct DNA sequencing of in vitro amplified cDNA. Proc. Natl. Acad. Sci. USA 86, 1919-1923.

Gibbs, R.A., Nguyen, P.N., and Caskey, C.T. (1989b) Detection of single DNA base differences by competitive oligonucleotide priming. Nucleic Acids Res. 17, 2437-2448.

Gilbert, W. (1981) DNA sequencing and gene structure (Nobel lecture). Science 214, 1305-1312.

Gill, P., Sullivan, K., and Werret, D.J. (1990) The analysis of hypervariable DNA profiles: problems associated with the objective determination of the probability of a match. Hum. Genet. 85, 75-79.

Gilroy, T.E., and Thomas, C.A. (1983) The analysis of some new Drosophila repetitive DNA sequences isolated and cloned from two-dimensional agarose gels. Gene 23, 41-51.

Giovannoni, J.J., Wing, R.A., Ganal, M.W., and Tanksley, S.D. (1991) Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. Nucleic Acids Res. 19, 6553-6558.

Goodbourn, S.E.Y., Higgs, D.R., Clegg, J.B., and Weatherall, D.J. (1983) Molecular basis of length polymorphism in the human zeta-globin gene complex. Proc. Natl. Acad. Sci. USA 80, 5022-5026.

Gray, M., Charpentier, A., Walsh, K., Wu, P., and Bender, W. (1991) Mapping point mutations in the Drosophila rosy locus using denaturing gradient gel blots. Genetics 127, 139-149.

Gray, M. (1992) Detection of DNA sequence polymorphisms in human genomic DNA by using denaturing gradient gel blots. Am. J. Hum. Genet. 50, 331-346.

Grompe, M., Caskey, C.T., and Fenwick, R.G. (1991) Improved molecular diagnostics for ornithine transcarbamylase deficiency. Am. J. Hum. Genet. 48, 212-222.

Hata, A., Robertson, M., Emi, M., and Lalouel, J.-M. (1990) Direct detection and automated sequencing of individual alleles after electrophoretic strand separation: identification of a common nonsense mutation in exon 9 of the human lipoprotein lipase gene. Nucleic Acids Res. 18, 5407-5411.

Hatada, I., Hayashizaki, Y., Hirotsune, S., Komatsubara, H., and Mukai, T. (1991) A genomic scanning method for higher organisms using restriction sites as landmarks. Proc. Natl. Acad. Sci. USA 88, 9523-9527.

Higgs, D.R., Goodbourn, S.E.Y., Wainscoat, J.S., Clegg, J.B., and Weatherall, D.J. (1981) Nucleic Acids Res. 9, 4213-4214.

Higuchi, M., Wong, C., Kochhan, L., Olek, K., Aronis, S., Kasper, C.K., Kazazian, H., and Antonarakis, S.E. (1990) Characterization of mutations in the factor VIII gene by direct sequencing of amplified genomic DNA. Genomics 6, 65-71.

Higuchi, M., Antonorakis, S., Kasch, L., Oldenburg, J., Economou-Petersen, E., Olek, K., Arai, M., Inaba, H., and Kazazian, H.H. (1991) Molecular characterization of mild-to-moderate hemophilia A: detection of the mutation in 25 of 29 patients by denaturing gradient gel electrophoresis. Proc. Natl. Acad. Sci. USA 88, 8307-8311.

Hilbert, P., Lindpaintner, K., Beckmann, J.S., Serikawa, T., Soubrier, F., Dubay, C., Cartwright, P., De Gouyon, B., Julier, C., Takahashi, S., Vincent, M., Ganten, D., Georges, M., and Lathrop, G.M. (1991) Chromosomal mapping of two genetic loci associated with blood-pressure regulation in hereditary hypertensive rats. Nature 353, 521-529.

Hobbs, H.H., Lehrman, M.A., Yamamoto, T., and Russell, D.W. (1985) Polymorphism and evolution of Alu sequences in the human low density lipoprotein receptor gene. Proc. Natl. Acad. Sci USA 82, 7651-7655.

Hofmann, S.L., Topham, M., Hsieh, C.-L., and Francke, U. (1991) cDNA and genomic cloning of HRC, a human sarcoplasmic reticulum protein, and loclization of the gene to human chromosome 19 and mouse chromosome 7. Genomcis 9, 656-669.

Hovig, E., Smith-Sorensen, B., Brogger, A., and Borresen, A.-L. (1991) Constant denaturant gel electrophoresis, a modification of denaturing gradient gel electrophoresis, in mutation detection. Mut. Res. 262, 63-71.

Hüvos, P.E., Aquiles Sanchez, J., Kramer, K.M., Karrer, K.M., and Wangh, L.J. (1988) Two-dimensional DNA gel electrophoresis as a method for analysis of eukaryotic genome structure: evaluation using Tetrahymena thermophila DNA. Bioch. Biophys. Acta 949, 325-333.

Iannuzi, M.C., Stern, R.C., Collins, F.S., Tom Hon, C., Hidaka, N., Strong, T., Decker, L., Drumm, M.L., White, M.B., Gerrard, B., and Dean, M. (1991) Two frame-shift mutations in the cystic fibrosis gene. Am. J. Hum. Gen. 48, 227-231.

Ivaschenko, T.E., White, M.B., Dean, M., and Baranov, V.S. (1991) A deletion of two nucleotides in exon 10 of the CFTR gene in a Soviet family with cystic fibrosis causing early infant death. Genomics 10, 298-299.

Jarman, A.P., Nicholls, R.D., Weatherall, D.J., Clegg, J.B., and Higgs, D.R. (1986) Molecular characterisation of a hypervariable region downstream of the human alpha-globin gene cluster. EMBO J. 5, 1857-1863.

Jeffreys, A.J. (1979) DNA sequence variants in the $^{G}\gamma$, $^{A}\gamma,\delta$- and ß-globin genes of man. Cell 18, 1-15.

Jeffreys, A.J., Wilson, V., and Thein, S.L. (1985a) Hypervariable "minisatellite regions in human DNA. Nature 314, 67-73.

Jeffreys, A.J., Wilson, V., and Thein, S.L. (1985b) Individual-specific "fingerprints" of human DNA. Nature 316, 76-79.

Jeffreys, A.J., Wilson, V., Thein, S.L., Weatherall, D.J., and Ponder, B.A.J. (1986) DNA "fingerprints" and segregation analysis of multiple markers in human pedigrees. Am. J. Hum. Genet. 39, 11-24.

Jeffreys, A.J., Royle, N.J., Wilson, V., and Wong, Z. (1988) Spontaneous mutation rates to new length alleles at tandem-repetitive hybpervariable loci in human DNA. Nature 332, 278-281.

Jeffreys, A.J., Neumann, R., and Wilson, V. (1990) Repeat unit sequence variation in minisatellites: A novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. Cell 60, 473-485.

Jeffreys, A.J., Turner, M., and Debenham, P. (1991) The efficiency of multilocus DNA fingerprint probes for individualization and establishment of family relationships, determined from extensive casework. Am. J. Hum. Genet. 48, 824-840.

Jeffreys, A.J., MacLeod, A., Tamaki, K., Neil, D.L., and Monckton, D.G. (1991) Minisatellite repeat coding as a digital approach to DNA typing. Nature 354, 204-209.

Julier, C., de Gouyon, B., Georges, M., Guénet, J.-L., Nakamura, Y., Avner, P., and Lathrop, G.M. (1990) Minisatellite linkage maps in the mouse by cross-hybridization with human probes containing tandem repeats. Proc. Natl. Acad. Sci. USA 87, 4585-4589.

Kamiura, S., Nolan, C.M., and Meruelo, D. (1992) Long-range physical map of the Ly-6 complex: mapping the Ly-6 multigene family by Field-Inversion and Two-dimensional gel electrophoresis. Genomics 12, 89-105.

Kan, Y.W. and Dozy, A.M. (1978) Polymorphisms of DNA sequence adjacent to human ß-globin structural gene: Relationship to sickle mutation. Proc. Natl. Acad. Sci. USA 75, 5631-5635.

Kashi, Y., Tikochinsky, Y., Iraqi, F., Nave, A., Beckman, J.S., Gruenbaum, Y., and Soller, M. (1990) Large restriction fragments containing poly-TG are highly polymorphic in a variety of vertebrates. Nucleic Acids Res. 18, 1129-1132.

Katzir, N., Rechavi, G., Cohen, J.B., Unger, T., Simoni, F., Segal, S., Cohen, D., and Givol, D., (1985) "Retroposon" insertion into the cellular oncogene c-*myc* in canine transmissible venereal tumor. Proc. Natl. Acad. Sci. USA 82, 1054-1058.

Kazazian, H.H., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and Antonorakis, S.E. (1988) Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanims for mutation in man. Nature 332, 164-166.

Keen, T.J., Inglehearn, C.F., Lester, D.H., Bashir, R., Jay, M., Bird, A.C., Jay, B., and Bhattacharya, S.S. (1991) Autosomal dominant retinis pigmentosa: four new mutations in rhodopsin, one of them in the retinal attachment site. Genomics 11, 199-205.

Keith, T.P., Green, P., Reeders, S.T., Brown, V.A., Phipps, P., Bricker, A., Falls, K., Rediker, K.S., Powers, J.A., Hogan, C., Nelson, C., Knowlton, R., and Donis-Keller, H. (1990) Genetic linkage map of 46 DNA markers on human chromosome 16. Proc. Natl. Acad. Sci. USA 87, 5754-5758.

Keohavong, P., Liu, V.F., and Thilly, W.G. (1991) Analysis of point mutations induced by ultraviolet light in human cells. Mut. Res. 249, 147-159.

Koeberl, D.D., Bottema, C.D., Ketterling, R.P., Bridge, P.J., Lillicrap, D.P., and Sommer, S.S. (1990) Mutations causing Hemophilia B: Direct estimate of the underlying rates of spontaneous germ-line transitions, transversions, and deletions in a human gene. Am. J. Hum. Genet. 47, 202-217.

Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C., Kunkel, L.M. (1987) Complete cloning of the Duchenne Muscular Dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell 50, 509-517.

Kogan, S., and Gitschier, J. (1990) Mutations and a polymorphism in the factor VIII gene discovered by denaturing gradient gel electrophoresis. Proc. Natl. Acad. Sci. USA 87, 2092-2096.

Kominami, R., Urano, Y., Mishima, Y., and Muramatsu, M. (1983) Novel repetitive sequence families showing size and frequency polymorphisms in the genomes of mice. J. Mol. Biol. 165, 209-228.

Kominami, R., Muramatsu, M., and Moriwaki, K. (1983) A mouse type 2 Alu sequence (M2) is mobile in the genome. Nature 301, 87-89.

Kovar, H., Jug, G., Auer, H., Skern, T., and Blaas, D. (1991) Two-dimensional single-strand conformation polymorphism analysis: a useful tool for the detection of mutations in long DNA fragments. Nucleic Acids Res. 19, 13, 3507-3510.

Krolewski, A.S., Keolewski, B., Gray, M., Stanton, V., Warram, J.H., and Housman, D. (1992) High-frequency DNA sequence polymorphisms in the insulin receptor gene detected by denaturing gradient gel blots. Genomics 12, 705-709.

Kuff, E.L., and Leuders, K.K. (1988) Intracisternal A-particle gene family: structural and functional aspects. Adv. Cancer Res. 51: 183-276.

Labrune, P., Melle, D., Rey, F., Berthelon, M., Caillaud, C., Rey, J., Munnich, A., and Lyonnet, S. (1991) Single-strand conformation polymorphism for detection of mutations and base substitutions in phenylketonuria. Am. J. Hum. Genet. 48, 1115-1120.

Lakshmikumaran, M.S., D'Ambrosio, E., Laimins, L.A., Lin, D.T., and Furano, A.V. (1985) Long interspersed repeated DNA (LINE) causes polymorphism at the rat insulin 1 locus. Mol. Cell. Biol. 5, 2197-2203.

Landegren, U., Kaiser, R., Sanders, J., and Hood, L. (1988) A ligase-mediated gene detection technique. Science 241, 1077-1080.

Lander, E.S., and Botstein, D. (1986) Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. In:"Molecular Biology of Homo sapiens", pp. 49-62.

Lander, E.S., and Botstein, D. (1989) Mapping Mendelian Factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185-199.

Latham, T., and Smith, F. (1989) Detection of single-base mutations in DNA molecules using the solution melting method. DNA 8, 223-231.

Lathrop, M., Nakamura, Y., Cartwright, P., O'Connell, P., Leppert, M., Jones, C., Tateishi, H., Bragg, T., Lalouel, J.-M., and White, R. (1988) A primary genetic map of markers for human chromosome 10. Genomics 2, 157-164.

Ledbetter, S.A., Nelson, D.L., Warren, S.T., and Ledbetter, D.H. (1990) Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. Genomics 6, 475-481.

Ledbetter, S.A., Garcia-Heras, J., and Ledbetter, D.H. (1990) "PCR-karyotype" of human chromosomes in somatic cell hybrids. Genomics 8, 614-622.

Lehrach, H., Drmanac, R., Hoheisel, J., Larin, Z., Lennon, G., Monaco, A.P., Nizetic, D., Zehetner, G., and Poustka, A. (1990) Hybridization fingerprinting in genome mapping and sequencing. In: "Genome Analysis Vol. 1: Genetic and Physical Mapping" (eds. K. E. Davies and S.M. Tilghman), 39-81, Cold Spring Harbor Laboratory Press, New York.

Lehrman, M.A., Goldstein, J.L., Russell, D.W., and Brown, M. (1987) Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. Cell 48, 827-835.

Leider, J.M., Palese, P., and Smith, F.I. (1988) Determination of the mutation rate of a retrovirus. J. Virol. 62, 3084-3091.

Lennon, G.C., and Lehrach, H. (1991) Hybridization analysis of arrayed cDNA libraries. Trends Genet. 7, 10, 314-317.

Lerman, L.S., Fischer, S.G., Bregman, D.B., and Silverstein, K.J. (1981) Base sequence and melting thermodynamics determine the position of pBR322 fragments in two-dimensional gel electrophoresis. In " Biomolecular stereodynamics" (R.H. Sarma ed.), pp.459-470, Adenine Press, Albany, New York.

Lerman, L.S., Fischer, S.G., and Lumelsky, N. (1983) Inferring variations in genotype from properties of DNA. In" Recombinant DNA and medical genetics" (A. Messer, I. Porter, eds.) Vol. 13, 157-186, Academic Press, New York.

Lerman, L.S., Fischer, S.G., Hurley, I., Silverstein, K., and Lumelsky, N. (1984) Sequence-determined DNA separations. Ann. Rev. Biophys. Bioeng. 13, 399-423.

Lerman, L.S., Silverstein, K., and Grinfield, E. (1986) Searching for gene defects by denaturing gradient gel electrophoresis. In "Cold Spring Harbor Symposia on Quantitative Biology" Vol 51, 285-297.

Lichter, P., Ledbetter, S.A., Ledbetter, D.H., and Ward, D.C. (1990) Fluorescence in situ hybridization with Alu I and L1 polymerase chain reaction probes for rapid characterization of human chromosomes in hybrid cell lines. Proc. Natl. Acad. Sci. USA 87, 6634-6638.

Lillie, J.W., Green, M., and Green, M.R. (1986) An adenovirus E1a protein region required for transformation and transcriptional repression. Cell 46, 1043-1051.

Lin, C.S., Goldthwait, D.A., and Samols, D. (1988) Identification of Alu transposition in human lung carcinoma cells. Cell 54, 153-159.

Litt, M., and Luty, J.A. (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet. 44, 397-401.

Losekoot, M., Fodde, R., Harteveld, C.L., van Heeren, Giordano P.C., Bernini, L.F. (1990) Denaturing gradient gel electrophoresis and direct sequencing of PCR amplified genomic DNA: a rapid and reliable diagnostic approach to ß-thalassemia. Br. J. Haematol. 76, 269-274.

Luca Cavalli-Sforza, L.(1990) Opinion: How can one study individual variation for 3 billion nucleotides of the human genome? Am J. Hum. Genet.46, 649-651.

Luty, J.A., Guo, Z., Willard, H.F., Ledbetter, D.H., Ledbetter, S., and Litt, M. (1990) Five polymorphic microsatellite VNTRs on the human X chromosome. Am. J. Hum. Genet. 46, 776-783.

Maddalena, A., Spence, J.E., O'Brien, W.E., and Nussbaum, R.L. (1988) Characterization of point mutations in the same arginine codon in three unrelated patients with ornithine transcarbamylase deficiency. J. Clin. Invest. 82, 1353-1358.

Mager, D.L., and Goodchild, N.L. (1989) Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings. Am J. Hum. Genet. 45, 848-854.

Maier, E., Hoheisel, J.D., McCarthy, L., Mott, R., Grigoriev, A.V., Monaco, A.P., Larin, Z., and Lehrach, H. (1992) Complete coverage of the Schizosaccharomyces pombe genome in yeast artificial chromosomes. Nature genetics 1, 273-277.

Maniatis, T., Fritsch, E.F., and Sambrook, J.(1982) Molecular Cloning: a Laboratory Manual Cold Spring Har bor Laboratory.

Mariat, D., and Vergnaud, G. (1992) Detection of polymorphic loci in complex genomes with synthetic tandem repeats. Genomics 12, 454-458.

Maxam, A.M., and Gilbert, W. (1977) A new method of sequencing DNA. Proc. Natl. Acad. Sci. USA 74, 560-564.

Melchior, W.B., and Von Hippel, P.H. (1973) Alteration of the relative stability of dA.dT and dG.dC base pairs in DNA. Proc. Natl. Acad. Sci. USA 70, 298-302.

Meulenbelt, I.M., Wapenaar, M.C., Patterson, D., Vijg, J., and Uitterlinden, A.G. (1993) Isolation and mapping of human chromosome 21 cosmids using a probe for RTVL-H retrovirus-like elements. Genomics, in press

Meyer, G., and Hildebrandt, A. (1986) Two-dimensional gele analysis of repetitive nuclear DNA sequences in the genome of Physarum polycephalum. Eur. J. Biochem. 157, 507-512.

Michelmore, R.W., Paran, I., and Kesseli, R.V. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88, 9828-9832.

Michiels, F., Craig, A., Zehetner, G., Smith, G.P., and Lehrach, H. (1987) Molecular approaches to genome analysis: A strategy for the construction of ordered overlapping clone libraries. CABIOS 3, 203-210.

Mietz, J.A., and Kuff, E.L. (1992) Intracisternal A-particle specific oligonucleotides provide multilocus probes for genetic linkage studies in the mouse. Mamm. Genome 3, 447-451.

Monaco, A.P., Neve, R., Colletti-Feener, C., Bertelson, C.J., Kurnit, D.M., and Kunkel, L.M. (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature 323, 646-648.

Montandon, A.J., Green, P.M., Giannelli, F., and Bently, D.R. (1989) Direct detection of point mutations by mismatch analysis: application to haemophilia B. Nucleic Acids Res. 17, 3347-3358.

Morse, B., Rothberg, P.G., South, V.J., Spandorfer, J.M., and Astrin, S.M. (1988) Insertional mutagenesis of the myc-locus by a LINE-1 sequence in a human breast carcinoma. Nature 333, 87-89.

Mullis, K.B., and Faloona, F. (1987) Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. In "Methods in Enzymology" (Wu, R., ed.) Vol. 155, 335-350, Academic Press, New York.

Musich, P.R., and Dykes, R.J. (1986) A long interspersed (LINE) DNA exhibiting polymorphic patterns in human genomes. Proc. Natl. Acad. Sci. USA 83, 4854-4858.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by Denaturing Gradient Gel Electrophoresis of Polymerase Chain Reaction amplified genes coding for 16S ribosomal DNA. Appl. Environm. Microbiol., in press.

Myers, R.M., Lumelsky, N., Lerman, L.S., and Maniatis, T. (1985a) Detection of single base substitions in total genomic DNA. Nature 313, 495-498.

Myers, R.M., Fischer, S.G., Maniatis, T., and Lerman, L.S. (1985b) Modification of the melting properties of duplex DNA by attachment of a GC-rich DNA sequence as determined by denaturing gradient gel electrophoresis. Nucleic Acids Res. 13, 3111-3129.

Myers, R.M., Fischer, S.G., Lerman, L.S., and Maniatis, T. (1985c) Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. Nucleic Acids Res. 13, 3131-3145.

Myers, R.M., Lerman, L.S., and Maniatis, T. (1985d) A general method for saturation mutagenesis of cloned DNA fragments. Science 229, 242-247.

Myers, R.M., Larin, Z., and Maniatis, T. (1985e) Detection of single base substitutions by Ribonuclease cleavage at mismatches in RNA:DNA duplexes. Science 230, 1242-1246.

Myers, R.M., Sheffield, V.C., and Cox, D.R. (1989) Detection of single base changes in DNA: ribonuclease cleavage and denaturing gradient gel electrophoresis. In "Genome analysis:A practical approach"(Davies, K.E., ed.), 95-139, IRL Press, Oxford.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987a) Variable Number of Tandem Repeat (VNTR) markers for human gene mapping. Science 235, 1616-1622.

Nakamura, Y., Julier, C., Wolff, R., Holm, T., O'Connell, P., Leppert, M., and White, R. (1987b) Characterization of a human "midisatellite" sequence. Nucleic Acids Res. 15, 2537-2547.

Nakamura, Y., Carlson, M., Krapcho, K., Kanamori, M., and White, R. (1988a) New approach for isolation of VNTR markers. Am. J. Hum. Genet. 43, 854-859.

Nakamura, Y., Lathrop, M., O'Connell, P., Leppert, M., Barker, D., Wright, E., Skolnick, M., Kondoleon, S., Litt, M., Lalouel, M., and White, R. (1988b) A mapped set of DNA markers for human chromosome 17. Genomics 2, 302-309.

Neel, J.V., Satoh, C., and Myers, R. (1993) Report of a workshop on the application of molecular genetics to the study of mutation in the children of atomic bomb survivors. Mut. Res. 291, 1-20.

Nelson, D.L., Ledbetter, S., Corbo, L., Victoria, M.F., Ramirez-Solis, R, Webster, T.D., Ledbetter, D.H., Caskey, C.T. (1989) Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. Proc. Natl. Acad. Sci. USA 86, 6686-6690.

Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C., and Markham, A.F. (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res. 17, 2503-2516.

Nizetic, D., Zehetner, G., Monaco, A., Craig, A., Hoheisel, J., Lennon, G., Young, B., and Lehrach, H. (1989) Am. J. Hum. Genet. 45, A154.

Noll, W.W., and Collins, M. (1987) Detection of human DNA polymorphisms with a simplified denaturing gradient gel electrophoresis technique. Proc. Natl. Acad. Sci. USA 84, 3339-3343.

Novack, D.F., Casna, N.J., Fischer, S.G., and Ford, J. (1986) Detection of single base-pair mismatches in DNA by chemical modification followed by electrophoresis in 15% polyacrylamide gel. Proc. Natl. Acad. Sci. USA 83, 586-590.

Nürnberg, P., Roewer, L., Neitzel, H., Sperling, K., Pöpperl, A., Hundrieser, J., Pöche, H., Epplen, C., Zischler, H. and Epplen, J.T. (1989) DNA fingerprinting with the oligonucleotide probe $(CAC)_5/(GTG)_5$: somatic stability and germline mutations. Hum. Genet. 84, 75-78.

O'Connell, P., Lathrop, G.M., Law, M., Leppert, M., Nakamura, Y., Hoff, M., Kumlin, E., Thomas, W., Elsner, T., Ballard, L., Goodman, P., Azen, E., Sadler, J.E., Cai, G.Y., Lalouel, J.-M., and White, R. (1987) A primary genetic linkage map for human chromosome 12. Genomics 1, 93-102.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989a) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. USA 86, 2766-2770.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5, 874-879.

Orkin, S.H., Little, P.F.R., Kazazian, H.H.,Jr., and Boehm, C.D. (1982) Improved detection of the sickle mutation by DNA analysis application to prenatal diagnosis. N.Engl. J. Med. 307, 32-36.

Page, D.C., Bieker, K., Brown, L.G., Hinton, S., Leppert, M., Lalouel, J.-M., Lathrop, M., Nystrom-Lahti, M., de la Chapelle, A., and White, R. (1987) Linkage, physical mapping, and DNA sequence analysis of pseudoautosomal loci on the human X and Y chromosome. Genomics 1, 243-256.

Parker, R.C., and Seed, B. (1980) Two-dimensional agarose gel electrophoresis "Sea Placque" agarose dimension, in Grossman, L. and Moldave, K. (eds.), Methods in Enzymology, Vol. 65, Academic Press, New York, 1980, 358-363.

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln, S.E., and Tanksly, S.D. (1988) Resolution of quatitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335, 721-726.

Patterson, D. (1991) Report of the second international workshop on human chromosome 21 mapping. Cytogen. Cell genet. 57, 168-174.

Paw, B.H., Tieu, P.T., Kaback, M.M., Lim, J., and Neufeld, E.F. (1990) Frequency of three Hex A mutant alleles among Jewish and non-Jewish carriers identified in a Tay-Sachs screening program. Am. J. Hum. Genet. 47, 698-705.

Peacock, A.C., Bunting, S.L., Cole, S.P.C., and Seidman, M. (1985) Two-dimensional electrophoretic display of restriction fragments from genomic DNA. Anal. Biochem. 149, 177-182.

Piratsu, M., Kan, Y.W., Cao, A., Conner, B.J., Teplitz, R.L., and Wallace, R.B. (1983) Prenatal diagnosis of ß-Thallasemia: detection of a single nucleotide mutation in DNA. N.Engl. J. Med. 309, 284-287.

Po, T., Steger, G., Rosenbaum, V., Kaper, J., and Riesner, D. (1987) Double-stranded cucumovirus associated RNA 5: experimental analysis of necrogenic and non-necrogenic variants by temperature-gradient gel electrophoresis. Nucleic Acids Res. 15, 5069-5083.

Poddar, S.K., and Maniloff, J. (1986) Chromosome analysis by two-dimensional fingerprinting. Gene 49, 93-102.

Poddar, S.K., and Maniloff, J. (1989) Determination of microbial genome sizes by two-dimensional denaturing gradient gel electrophoresis. Nucleic Acids Res. 17, 2889-2895.

Poduslo, S.E., Dean, M., Kolch, U., and O'Brien, S.J. (1991) Detecting high-resolution polymorphisms in human coding loci by combining PCR and single-atrand conformation polymorphism (SSCP) analysis. Am. J. Hum. Gent. 49, 106-111.

Potter, S.S., and Newbold, J.E. (1976) Multi-dimensional restriction enzyme analysis of complex genomes. Anal. Biochem. 71, 452-458.

Potter, S.S., Bott, K.F., and Newbold, J.E. (1977) Two dimensional restriction analysis of the Bacillus subtilis genome: gene purification and ribosomal ribonucleic acid gene organization. J. Bacteriol. 129, 492-500.

Proudfoot, N.J., Zil, A., and Maniatis, T. (1982) The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell 31, 553-563.

Riedel, G.E., Swanberg, S.L., Kurunda, K.D., Marquette, K., LaPan, P., Bledsoe, P., Kennedy, A., and Lin, B.-Y. (1990) Denaturing gradient gel electrophoresis identifies genomic DNA polymorphism with high frequency in maize. Theor. Appl. Genet. 80, 1-10.

Riggins, G.J., Lokey, L.K., Chastain, J.L., Leiner, H.A., Sherman, S.L., Wilkinson, K.D., and Warren, S.T. (1992) Human genes containing polymorphic trinucleotide repeats. Nature genetics 2, 186-191,

Riordan, J.R., Rommens, J.M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M.L., Ianuzzi, M.C., Collins, F.S., and Tsui, L.-C. (1989) Identification of the Cystic Fibrosis gene: Cloning and characterization of complementary DNA. Science 245, 1066-1073.

Roberts, L. (1990) An animal genome project? Science 248, 550- 552.

Roberts, R.G., Bobrow, M., and Bently, D.R. (1992) Pointmutations in the dystrophin gene. Proc. Natl. Acad. Sci. USA 89, 2331-2335.

Rogan, P.K., Lemkin, P.L., Klar, A.J.S., Singh, J., and Strathern, J.N. (1991) Two-dimensional agarose gel electrophoresis of restriction-digested genomic DNA. Methods: A Companion to Methods in Enzymology 3, 2, 91-97.

Rogstad, S.H., Herwaldt, B.L., Schlesinger, P.H., and Krogstad, D.J. (1989) The M13 repeat probe detects RFLPs between two strains of the protozoan malaria parasite Plasmodium falciparum. Nucleic Acids Res. 17, 3610.

Rosatelli, M.C., Dozy, A., Faà, V., Meloni, A., Sardu, R., Saba, L., Kan, Y.W., and Cao, A. (1992) Molecular characterization of ß-Thalassemia in the Sardinian population. Am. J. Hum. Genet. 50, 422-426.

Rosenbaum, V., and Reissner, A.S. (1987) Biophys. Chem. 26, 235-246.

Ross, M.T., Nizetic, D., Nguyen, C., Knights, C., Vatcheva, R., Burden, N., Douglas, C., Zehetner, G., Ward, D.C., Baldini, A., and Lehrach, H. (1992) Selection of a human chromosome 21 enriched YAC sub-librery using a chromosome-specific composite probe. Nature genetics 1, 284-290.

Rossiter, B.J.F., and Caskey, C.T. (1990) Molecular scanning methods of mutation detection. J. Biol. Chem. 265, 22, 12753-12756.

Royle, N.J., Clarkson, R.E., Wong, Z., and Jeffreys, A.J. (1988) Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. Genomics 3, 352-360.

Runnebaum, I.B., Nagarajan, M., Bowman, M., Soto, D., and Sukumar, S. (1991) Mutations in p53 as potential molecular markers for human breast cancer. Proc. Natl. Acad. Sci. USA 88, 10657-10661.

Ryskov, A.P., Jincharadze, A.G., Prosnyak, M.I., Ivanov, P.L., and Limborska, S.A. (1988) FEBS Lett. 233, 388-392.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G., Erlich, H.A., and Arnheim, N. (1985) Enzymatic amplification of ß-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230, 1350-1354.

Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1986) Analysis of enzymatically amplified ß-globin and HLA-DQ DNA with allele-specific oligonucleotide probes. Nature 324, 163-166.

Saiki, R.K., Chang, C.A., Levenson, C.H., Warren, T.C., Boehm, C.D., Kazazian, H.H., Jr., and Erlich, H.A. (1988) Diagnosis of sickle cell anemia and ß-Thallesemia with enzymatically amplified DNA and non-radioactive allel-specific oligonucleotide probes. N.Engl. J. Med. 319,537-541.

Saiki, R.K., Walsh, P.S., Levenson, C.H., and Erlich, H.A. (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc. Natl. Acad. Sci. USA 86, 6230-6234.

Sakaki, Y., Kurata, Y., Miyake, T., and Saigo, K. (1983) Two dimensional gel electrophoretic analysis of the Hind III 1.8 kb repetitive-sequence family in the human genome. Gene 24, 179-190.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edition, Cold Spring Harbor Press, New York.

Sanger, F., Nicklen, S., Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Sarkar, G., Yoon, H.-S., and Sommer, S.S. (1992a) Screening for mutations by RNA single-strand conformation polymorphism (rSSCP): comparison with DNA-SCCP. Nucleic Acids Res. 20, 871-878.

Sarkar, G., Yoon, H.-S., and Sommer, S.S. (1992b) Dideoxy Fingerprinting (ddF): A rapid and efficient screen for the presenc of mutations. Genomics 13, 441-443.

Schäfer, R., Zischler, H., Birnser, U., Becker, A., Epplen, J.T. (1988) Optimized oligonucleotide probes for DNA fingerprinting. Electrophoresis 9, 369-374.

Schäfer, R., Zischler, H., and Epplen, J.T. (1988) $(CAC)_5$, a very informative oligonucleotide probe for DNA fingerprinting. Nucleic Acids Res. 16, 5196.

Schlötterer, C., Amos, B., and Tautz, D. (1991) Conservation of polymorphic simple sequence loci in catacean species. Nature 354, 63-65.

Schlötterer, C., and Tautz, D. (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 20, 211-215.

Schuler, L.A., Weber, J.L., and Gorski, J. (1983) Polymorphism near the rat prolactin gene caused by insertion of an Alu-loke element. Nature 305, 159-160.

Schwartz, D.C., and Cantor, C.R. (1984) Separation of yeast chromosome-sized DNAs by pulsed field gel electrophoresis. Cell 37, 67-75.

Sheffield, V.C., Cox, D.R., Lerman, L.S., and Myers, R.M. (1989) Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc. Natl. Acad. Sci. USA 86, 232-236.

Sheffield, V.C., Fishman, G.A., Beck, J.S., Kimura, A.E., and Stone, E.M. (1991) Identification of novel rhodopsin mutations associated with retinis pigmentosa by GC-clamped denaturing gradient gel electrophoresis. Am. J. Hum. Genet. 49, 699-706.

Sheffield, V.C., Beck, J.S., Stone, E., and Myers, R.M. (1992) A simple and efficient method for attachement of a 40-Base Pair, GC-rich sequence to PCR-amplified DNA. BioTechniques 12, 3, 386-387.

Shenk, T.E., Rhodes, C., Rigby, W.J., and Berg, P. (1975) Biochemical method for mapping mutational alterations in DNA with S_1 nuclease: the location of deletions and temperature- sensistive mutations in Simian virus 40. Proc. Natl. Acad. Sci. USA 72, 989-993.

Sheppard, R.D., Montagutelli, X., Jean, W.C., Tsai, J.-Y., Rose, A., Guénet, J.-L., Cole, M.D., and Silver, L.M. (1991) Two-dimensional gel analysis of complex DNA families: Methodology and apparatus. Mamm. Genome 1, 104-111.

Shinomiya, T., and Ina, S. (1991) Analaysis of chromosomal replicons in early embryo's of Drosophila melanogaster by two-dimensional gel electrophoresis. Nucleic Acids. Res. 19, 14, 3935-3941.

Shore, S., and Myerowitz, R. (1990) Polymerase chain reaction generated heteroduplexes from Ashkenazi Tay-Sachs carriers with an insertion mutation can be detected on agarose gels. Am. J. Hum. Genet. 47, 169.

Shyman, S., and Weaver, S. (1985) Chromosomal rearrangements associated with LINE elements in the mouse genome. Nucleic Acids Res. 13, 5085.

Simmler, M.-C., Cox, R.D. and Avner, P. (1991) Adaptation of the interspersed repetitive sequence polymerase chain reaction to the isolation of mouse DNA probes from somatic cell hybrids on a hamster background. Genomics 10, 770-778.

Simon, M., Phillips, M., and Green, H. (1991) Polymorphism due to variable number of repeats in the human involucrin gene. Genomics 9, 576-580.

Sinnett, D., Deragon, J.-M., Simard, L.R., Labuda, D. (1990) Alumorphs-human DNA polymorphisms detected by the Polymerase Chain Reaction using Alu-specific primers. Genomics 7, 331-334.

Slagboom, P.E., Mullaart, E., Droog, S., and Vijg, J. (1991) Somatic mutations and cellular aging: two-dimensional DNA typing of rat fibroblast clones. Mut. Res. 256, 311-321.

Smeets, H.J.M., Brunner, H.G., Ropers, H., Wieringa, B. (1989) Use of variable simple sequence motifs as genetic markers: application to study of myotonic dystrophy. Hum. Genet. 83, 245-251.

Smith, F.I., Parvin, J.D., and Palese, P. (1986) Detection of single base substituions in Influenza virus RNA molecules by denaturing gradient gel electrophoresis of RNA-RNA or DNA-RNA heteroduplexes. Virology 150, 55-64.

Smith, F.I., Latham, T.E, Ferrier, J.A., and Palese, P. (1988) Novel method of detecting single base substitutions in RNA molecules by differential melting behaviour in solution. Genomics 3, 217-223.

Snapka, R.M., Permana, P.A., Marquit, G., and Shin, C.-G. (1991) Two-dimensional agarose gel analysis of simian virus 40 DNA replication intermediates. Methods: A Companion to Methods in Enzymology 3, 2, 73-82.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Southern, E.M. (1979) Gel electrophoresis of restriction fragments. Methods Enzymol. 68, 152-176.

Spinardi, L., Mazars, R., and Theillet, C. (1991) Protocols for an improved detection of point mutations by SSCP. Nucleic Acids Res. 19, 4009.

Stallings, R.L., Torney, D.C., Hildebrand, C.E., Longmire, J.L., Deaven, L.L., Jett, J.H., Doggett, N.A., and Moyzis, R.K. (1990) Physical mapping of human chromosomes by repetitive sequence fingerprinting. Proc. Natl. Acad. Sci. USA 87, 6218-6222.

Steger, G., Po, T., and Riesner, D. (1987) Double-stranded cucumovirus associated RNA 5: which sequence variations may be detected by optical melting and temperature-gradient gel electrophoresis? Nucleic Acids Res. 15, 5085-5103.

Strezoska, Z., Paunesku, T., Radosavljevic, D., Labat, I., Dramanac, R., and Crkvenjakov, R. (1991) DNA sequencing by hyb ridization: 100 bases read by a non-gel-based method. Proc. Natl. Acad. Sci. USA 88, 10089-10093.

Suzuki, Y., Orita, M., Shiraishi, M., Hayashi, K., and Sekiya, T. (1990) Detection of ras gene mutations in human lung can cers by single-strand conformation polymorphism analysis of polymerase chain reacion products. Oncogene 5, 1037-1043.

Swallow, D.M., Gendler, S., Griffiths, B., Corney, G., Papadimitriou, J., and Bramwell, M.W. (1987) The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. Nature 328, 82-84.

Tabor, S., and Richardson, C.C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA 84, 4767-4771.

Takahashi, N., Hiyama, K., Kodaira, M., and Satoh, C. (1990) An improved method for the detection of genetic variations in DNA with denaturing gradient gel electrophoresis. Mut. Res. 234, 61-70.

Tautz, D., and Renz, M.(1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res. 12, 4127-4138.

Tautz, D., Trick, M., and Dover, G.A. (1986) Cryptic simplicity in DNA is a major source of genetic variation. Nature 322, 652-656.

Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. 17, 6463-6471.

Theophilus, B.D.M., Latham, T., Grabowski, G.A., and Smith, F.I. (1989) Comparison of RNase A, a chemical cleavage and GC-clamped denaturing gradient gel electrophoresis for the detection of mutations in exon 9 of the human acid-glucosidase gene. Nucleic Acid Res. 17, 7707-7722.

Thilly, W.G. (1985) Potential use of gradient denaturing gel electrophoresis in obtaining mutational spectra from human cells. In "Carcinogenesis" (Huberman, E., and Barr, S.H., eds.) Vol. 10, 511-528, Raven Press, New York.

Todd, J.A. (1992) La carte de microsatellites est arrivée! Hum. Mol. Genet. 1, 9, 663-666.

Top, B., Uitterlinden, A.G., van der Zee, A., Kastelein, J.J.P., Gevers Leuven, J.A., Havekes, L.M., and Frants, R. (1992) Absence of mutations in the promotor region of the low density lipoprotein receptor gene in a large number of familial hypercholesterolemia patients as revealed by denaturing gradient gel electrophoresis. Hum. Genet. 89, 561-565.

Traystman, M.D., Higuchi, M., Kasper, C.K., Antonarakis, S.E., and Kazazian, Jr, H.H. (1990) Use of denaturing gradient gel electrophoresis to detect point mutations in the factor VIII gene. Genomics 6, 293-301.

Turner, B.J., Elder, J.F., Laughlin, T.F., and Davis, W.P. (1990) Genetic variation in clonal vertebrates detected by simple sequence DNA fingerprinting. Proc. Natl. Acad. Sci. USA 87, 5653-5657.

Uitterlinden, A.G., Slagboom, P.E., Knook, D.L., and Vijg, J. (1989a) Two-dimensional DNA fingerprinting of human individuals. Proc. Natl. Acad. Sci. USA 86, 2742-2746.

Uitterlinden, A.G., Slagboom, P.E., Johnson, T. E., and Vijg, J. (1989b) The Caenorhabditis elegans genome contains monomorphic minisatellites and simple sequences. Nucleic Acids Res. 17, 9527-9530.

Uitterlinden, A.G., and Vijg, J. (1989) Two-dimensional DNA typing. TIBTECH 7, 336-341.

Uitterlinden, A.G., and Vijg, J. (1990) Denaturing gradient gel electrophoretic analysis of the human cHa-ras1 proto-oncogene. Appl. Theor. Electr. 1, 175-179.

Uitterlinden, A.G., and Vijg, J. (1991a) Denaturing gradient gel electrophoretic analysis of minisatellite alleles. Electrophoresis 12, 12-16.

Uitterlinden, A.G., and Vijg, J. (1991b) Method for the simultaneous detection of DNA sequence variations at a large number of sites, and a kit suitable therefore. United States Patent 5,068,176

Uitterlinden, A.G., Slagboom, P.E., Mullaart, E., Meulenbelt, I., an Vijg, J. (1991a) Genome scanning by two-dimensional DNA typing: the use of repetitive DNA sequences fro rapid mapping of genetic traits. Electrophoresis 12, 119-134.

Uitterlinden, A.G., Mullaart, E.M., Morolli, B., and Vijg, J. (1991b) Genome scanning of higher eukaryotes by two-dimensional DNA typing using micro- and minisatellite core probes. Methods: A Companion to Methods in Enzymology 3, 2, 83-90.

Ullrich, A., Dull, T.J., Gray, A., Philips, J.A., and Peter, S. (1982) Variation in the sequence and modification state of the human insulin gene flanking regions. Nucleic Acids Res. 10, 2225-2240.

VanWye, J.D., Bronson, E.C., and Anderson, J.N. (1991) Species-specifc patterns of DNA bending and sequence. Nucleic Acids Res. 19,19, 5253-5261.

Veres, G., Gibbs, R.A., Scherer, S.F., and Caskey, C.T. (1987) The molecular basis of the sparse fur mouse mutation. Science 237, 415-417.

Vergnaud, G. (1989) Polymers of random short oligonucleotides detect polymorphic loci in the human genome. Nucleic Acids Res. 17, 7623-7630.

Verkerk, A.J.H.M., Pieretti, M., Sutcliffe, J.S., Fu, Y.-H., Kuhl, D.P.A., Pizutti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B., van Ommen, G.J.B., Blonden, L.A.J., Riggins, G., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A., and Warren, S.T. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in Fragile X syndrome. Cell 65, 905-914.

Vidaud, M., Vidaud, D., Siguret, V., Lavergne, J.M., and Goossens, M. (1989) Mutational insertion of an Alu sequence causes hemophilia B. Am. J. Hum. Genet. 45, A226.

Vidaud, M., Fanen, P., Martin, J., Ghanem, N., Nicolas, S., and Goossens, M. (1990) Three point mutations in the CFTR gene in French cystic fibrosis patients: identification by denaturing gradient gel electrophoresis. Hum. Genet. 85, 446-449.

Vijg, J., and Uitterlinden, A.G. (1987) A search for DNA alterations in the aging mammalian genome: an experimental strategy. Mech. Aging Dev. 41, 47-63.

Vijg, J., and Gossen, J. (1993) Somatic mutations and cellular aging. Comp. Biochem. Physiol., in press.

Vogelstein, B., Fearon, E.R., Hamilton, S.R., and Feinberg, A.P. (1985) Use of restriction fragment length polymorphisms to determine the clonal origin of human tumors. Science 227, 642-645.

Vrieling, H., Simons, J.W.I,M, and van Zeeland, A.A. (1988) Nucleotide sequence determination of point mutations at the mouse HPRT locus using in vitro amplification of HPRT mRNA sequences. Mut. Res. 198, 107-113.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., and Itakura, K. (1979) Hybridization of synthetic oligodeoxyribonucleotides to ϕX174 DNA: the effect of single base pair mismatch. Nucleic Acid Res. 6, 3543-3557.

Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Hawashima, E.H., and Itakura, K. (1981) The use of synthetic oligonucleotides as hybridization probes. II: hybridization of oligonucleotides of mixed sequence to rabbit ß-globin DNA. Nucleic Acid Res. 9, 879-894.

Wallace, M.R., Anderson, L.B., Saulino, A.M., Gregory, P.E., Glover, T.W., and Collins, F.S. (1991) A de novo Alu insertion results in neurofibromatosis type 1. Nature 353, 864-866.

Walter, M.A., and Cox, D.W. (1989) A method for two dimensional DNA electrophoresis (2D-DE): Application to the immunoglobulin heavy chain variable region. Genomics 5, 157-159.

Wang-Jabs, E., Goble, C.A., and Cutting, G.R. (1989) Macromolecular organization of human centromeric regions reveals high-frequency, polymorphic macro DNA repeats. Proc. Natl. Acad. Sci. USA 86, 202-206.

Wartell, R.M., Hosseini, S.H., and Moran, C.P., Jr. (1990) Detecting basepair substitutions in DNA fragments by temperature-gradient gel electrophoresis. Nucleic Acid Res. 18, 2699-2705.

Watson, J.D. (1990) The Human Genome Project: Past, Present and Future. Science 248, 44-49.

Watson, J.D., and Crick, F.H.C. (1953a) Molecular structure of nucleic acids: A structure for deoxyribonucleic acid. Nature 171, 737-738.

Watson, J.D., and Crick, F.H.C. (1953b) Genetic implications of the structure of desoxyribonucleic acid. Nature 171, 964-969.

Waye, J.S., and Fourney, R.M. (1990) Identification of complex DNA polymorphisms based on variable number of tandem repeats (VNTR) and restriction site polymorphism. Hum. Genet. 84, 223-227.

Weber, J.L. Informativeness of human $(dC\text{-}dA)_n.(dG\text{-}dT)_n$ polymorphisms. Genomics 7, 524-530.

Weber, J.L., and May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. 44, 388-396.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morisette, J., Millasseau, Vaysseix, G., and Lathrop, M. (1992) A second-generation linkage map of the human genome. Nature 359, 794-801.

Welsh, J., and McClelland, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18, 7213-7218.

Welsh, J., Petersen, C., and McClelland, M. (1991) Polymorphisms generated by arbitrarily primed PCR in the mouse: application to strain identification and genetic mapping. Nucleic Acids Res. 19, 303-306.

Wetmur, J.G., Ruyechan, W.T., and Douthart, R.J. (1981) Denaturation and renaturation of Penicillin chrysogenum mycophage double-stranded ribonucleic acid in tetraalkylammonium salt solutions. Biochemistry 20, 2999-3002.

White, M.B., Krueger, L.F., Holsclaw, D.S., Gerrard, B., Stewart, C., Quittel, L., Dolganov, G., Baranov, V., Ivaschenko, T., Kapranov, N.I., Sebastio, G., Castiglione O., and Dean, M. (1991) Detection of three rare frame-shift mutations in the cystic fibrosis gene in an African American (CF444delA), an Italian (CF2522insC), and a Soviet (CF3821delT). Genomics 10, 266-269.

White, M.-B, Carvalho, M., Derse, D., O'Brien, S.J. and Dean, M. (1992) Detecting single base substitutions as heteroduplex polymorphisms. Genomics 12, 301-306.

White, R., Lippert, M., Bishop, D.T., Barker, D., Berkowitz, J., Brown, C., Callahan, P., Holmes, T., and Jerominsky, L. (1985) Construction of linkage maps with DNA markers for human chromosomes. Nature 313, 101-105.

White, R., and Caskey, C.T. (1988) The human as an experimental system in molecular genetics. Science 240, 1483-1488.

Willard, H.F., Waye, J.S., Skolnick, M.H., Schwartz, C.E., Powers, V.E., England, S.B. (1986) Detection of restriction fragment length polymorphisms at the centromeres of human chromosomes by using chromsome-specific α-satellite DNA probes: Implications for the development of centromere-based linkage maps. Proc. Natl. Acad. Sci. USA 83, 5611-5615.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531-6535.

Winter, E.E., Yamamoto, E., Almoguera, C, and Perucho, M. (1985) A method to detect and characterize point mutations in transcribed genes: amplification and over-expression of the mutant c-Ki-ras allele in tumor cells. Proc. Natl. Acad. Sci. USA 82, 7575-7579.

Wolff, R.K., Nakamura, Y., and White, R. (1988) Molecular characterization of a spontaneously generated new allele at a VNTR locus: no exchange of flanking DNA sequence. Genomics 3, 347-351.

Wolff, R.K., Plaetke, R., Jeffreys, A.J., and White, R. (1989) Unequal crossing over between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. Genomics 5, 382-384.

Wong, Z., Wilson, V., Patel, I., Povey, S., and Jeffreys, A.J. (1987) Characterization of a panel of highly variable minisatellites cloned from human DNA. Ann. Hum. Genet. 51, 269-288.

Wong, Z., Wilson, V., Jeffreys, A.J., and Thein, S.L. (1986) Cloning of a selected fragment from a human DNA "fingerprint": isolation of an extremely polymorphic minisatellite. Nucleic Acids Res. 14, 4605-4616.

Woods-Samuels, P., Wong, C., Mathias, S.L., Scott, A.F., Kazazian, H.H., and Antonorakis, S.E. (1989) Characterization of a non-deleterious L1 insertion in an intron of the human factor VIII gene and further evidence of open reading frames in functional L1 elements. Genomics 4, 290-296.

Woolf, T., Lai, E., Kronenberg, M., and Hood, L. (1988) Mapping genomic organization by field inversion and two-dimensional gel electrophoresis: application to the murine T-cell receptor τ gene family. Nucleic Acids Res. 16,9, 3863-3275.

Wu, D.Y. and Wallace, R.B. (1989a) Specificity of the nick closing activity of bacteriophage T4 DNA ligase. Gene 76, 245-254.

Wu, D.Y., and Wallace, R.B. (1989b) The ligation amplification reaction (LAR)-amplification of specific DNA sequences using sequential rounds of template-dependent ligation. Genomics 4, 560-569.

Wyman, A.R., and White, R. (1980) A highly polymorphic locus in human DNA. Proc. Natl. Acad. Sci. USA 77, 6754-6758.

Yandell, D.W., and Drja, T.P. (1989) Detection of DNA sequence polymorphisms by enzymatic amplification and direct genomic sequencing. Am. J. Hum. Genet. 45, 547-555.

Yee, T., and Inouye, M. (1982) Two-dimensional DNA electrophoresis applied to the study of DNA methylation and the analysis of genome size in Myxococcus xanthys. J. Mol. Biol. 154, 181-196.

Yee, T., and Inouye, M. (1983) Two-dimensional DNA electrophoresis methods utilizing in situ enzymatic digestions. In: "Experimental manipulation of gene expression" (M. Inouye, Ed.), pp. 279-290, Academic Press, New York.

Yee, T., and Inouye, M. (1984) Two-dimensional S1 nuclease heteroduplex mapping: detection of rearrangements in bacterial genomes. Proc. Natl. Acad. Sci. USA. 81, 2723-2727.

Yi, M., Au, L.-C., Ichikawa, N., and Ts'o, P.O.P. (1990) Enhanced resolution of DNA restriction fragments: a procedure by two-dimensional electrophoresis and double-labelling. Proc. Natl. Acad. Sci. USA 87, 3919-3923.

Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, J.C., Warren, S.T., Schlessinger, D., Sutherland, G.R., and Richards, R.I. (1991) Fragile X genotype characterized by an unstable region of DNA. Science 252, 1179-1181.

Zechner, R., Newman, T.C., Steiner, E., and Breslow, J.L. (1991) The structure of the mouse lipoprotein lipase gene: A B1 repetitive element is inserted into the 3' untranslated region of the mRNA. Genomics 11, 62-76.

Zhuang, J., Constantinou, C.D., Ganguly, A., and Prockop, D.J. (1991) A single base mutation in type I procollagen (COLIAI) that converts glycine αI-541 to aspartate in a lethal variant of osteogenesis imperfecta: detection of the mutation with a carbodiimide reaction of DNA heteroduplexes and direct sequencing of products of the PCR. Am. J. Hum. Genet. 48, 1186-1191.

Zischler, H., Nanda, I., Schäfer, R., Schmid, M., and Epplen, J.T. (1989) Digoxigenated oligonucleotide probes specific for simple repeats in DNA fingerprinting and hybridization in situ. Hum. Genet. 82, 227-233.

Zuliani, G., and Hobbs, H.H. (1990) A high frequency of length polymorphisms in repeated sequences adjacent to Alu sequences. Am. J. Hum. Genet. 46, 963-969.

SECTION II: ADDENDA

ADDENDUM 1

Denaturing gradient gel electrophoretic analysis
of the human cHa-ras1 proto-oncogene

[Appl. Theor. Electrop. (1990) 1, 175-179]

Applied and Theoretical Electrophoresis (1990), **1**, 175 - 179

Denaturing gradient gel electrophoretic analysis of the human cHa-*ras*1 proto-oncogene

André G. Uitterlinden & Jan Vijg

Department of Molecular Biology, TNO Institute for Experimental Gerontology, PO Box 5815, 2280 HV Rijswijk, The Netherlands

Activation of *ras* proto-oncogenes frequently involves point mutations at different sites in the gene. Here we describe the application of denaturing gradient gel electrophoresis to identify and characterize such mutations in the cHa-*ras*1 proto-oncogene. We calculated a melting map of the cHa-*ras*1 gene and designed optimal conditions to separate mutant from wild type sequences in the first exon. As an example we examined the T24 guanosine to thymidine point mutation in the first exon which has been found in T24 bladder carcinoma cells. Denaturing gradient gel analysis of both homoduplex as well as heteroduplex molecules resulted in separation of the wild type sequence from the mutant sequence. On the basis of the melting map we present a general scheme for screening the cHa-*ras*1 proto-oncogene sequence for the occurrence of point mutations.

Keywords: Denaturing gradient gel electrophoresis; T24 mutation; DNA polymorphism; heteroduplex analysis.

Introduction

Activated transforming genes (oncogenes) have been found in a number of human tumours by use of assays in which transformed foci result from transfection of tumour DNA into NIH3T3 cells. In several cases the transforming genes detected in this way are members of the *ras* gene family and the mechanism by which they have been activated usually involves a point mutation (Bos *et al.*, 1985; Cohen & Levinson 1988). The first example was a point mutation in the first exon of the cHa-*ras*1 proto-oncogene which was present in T24 bladder carcinoma cells (Premkumar Reddy *et al.*, 1982). In this case a guanosine to thymidine substitution was found at codon 12. In general, an alteration of amino acid residue 12 or 61 of the protein encoded by the *ras* gene is involved (Bos *et al.*, 1987). However, *in vitro* mutagenesis experiments have shown that in transformed cells also codons 13, 59 and 63 of the *ras* genes can contain point mutations (Fasano *et al.*, 1984).

The presence of these point mutations has been shown by sequence analysis and oligonucleotide hybridization analysis. These methods are either not suitable for screening purposes or they require knowledge of the sequence alteration in advance and are therefore not suitable for detecting novel sequence alterations that might be involved in the mutational process. An alternative and potentially fully informative method for screening for sequence alterations is denaturing gradient gel electrophoresis (DGGE; Fischer & Lerman 1979). This method is based on the sequence dependent melting behaviour of so-called low melting domains within DNA molecules (see Lerman *et al.*, 1984 for a review). In practice, the lowest melting domains within restriction fragments can be screened for sequence alterations as small as point mutations (Fischer & Lerman 1983). In order to analyse the highest melting domains too, a so-called GC-clamp should be attached (Myers *et al.*, 1985a). The DGGE method has been successfully applied in the analysis of sequence variations in the β-globin gene region (Myers *et al.*, 1985a), minisatellite sequences (Uitterlinden *et al.*, 1989) and various other sequences (Cariello *et al.*, 1988; Leider *et al.*, 1988; Borresen *et al.*, 1988; Uitterlinden & Vijg, 1989).

Here we describe the application of this technique to the analysis of the T24 point mutation in the cHa-*ras*1 proto-oncogene. We have constructed a melting map of the entire cHa-*ras*1 gene and show that the mutated site is present in a low melting domain. Using cloned sequences we show that this guanosine to thymidine substitution can be detected by homoduplex and heteroduplex analysis with DGGE. By using labelled, strand-specific probes we characterize the heteroduplex molecules and we describe a general strategy for detecting sequence variations by DGGE within this gene region in genomic DNA.

Materials and methods

Plasmids

The first exon of the normal and the T24 mutant *ras* oncogene were obtained in the form of a 327 bp SmaI-KpnI restriction fragment, derived from the HOMER 6N and HOMER 6T cosmid inserts, respectively (Spandidos & Wilkie 1984). The 327 bp fragments were subcloned in HincII-KpnI digested pTZ18R plasmids (Pharmacia). The resulting plasmids were termed pN (containing the normal cHa-*ras*1 proto-oncogene) and pT (containing the T24 cHa-*ras*1 oncogene). When the subclones were digested with PstI, a single 234 bp fragment was generated in which the point mutation was present at position 79.

Melting maps

For the construction of the melting map of the cHa-*ras*1 proto-oncogene region, we used the computer algorithm designed by Dr L.S. Lerman. Basepairs 0 to 4500 of the GENBANK entry were used for the calculation. Separate melting maps were calculated for the 234 bp PstI fragments of plasmids pN and pT. These fragments correspond to position 1640 to 1853 (213 bp) of the gene sequence of the GENBANK entry and include a 21 bp piece of the multiple cloning site of the pTZ18R vector.

Correspondence: A.G. Uitterlinden
Received 19 October 1989; accepted in revised form 16 January 1990

Heteroduplex analysis and single strand probe preparation

Heteroduplex analysis of the double stranded pN and pT plasmids was performed by mixing PstI digests of both plasmids, boiling for 5 min and annealing at 50°C for 1 h. After this the samples were analysed by DGGE.

Single-stranded radiolabelled probe was prepared from pN and pT, essentially as described by Myers *et al.* (1985b). Single stranded DNA was prepared from the pTZ18R plasmids according to the manufacturers' instructions (Pharmacia). A synthetic 20-mer primer which anneals 3′ of the first exon of the cHa-*ras*1 proto-oncogene (60 bp downstream of the first PstI site 3′ of exon 1) was used for primer extension synthesis. Annealing of the primer (30 ng) to the single strand phage DNA (1 μg) was performed at 40°C for 10 min in 50 mM NaCl, 6 mM Tris.Cl (pH 7.4) and 6 mM $MgCl_2$. Complementary strand synthesis was performed at 37°C for 30 min in a volume of 30 μl containing 1 unit Klenow enzyme (Boehringer Mannheim) and 0.1 μM α[^{32}P]-dATP, α[^{32}P]-dCTP, α[^{32}P]-dTTP and α[^{32}P]-dGTP (3000 mCi mM^{-1}). The reaction was chased by adding 1 mM dNTP and incubating for another 20 min. The DNA was recovered by ethanol precipitation and dissolved in a volume of 20 μl containing 20 units HindIII restriction enzyme and the buffer recommended by the manufacturer (Boehringer Mannheim), to truncate the probe upstream of the PstI site in the pTZ18R plasmid. The labelled strand was purified by electrophoresis in a 6% polyacrylamide/5.6 M Urea/32% formamide gel run at 60°C. Heteroduplex analysis was performed by mixing single stranded pN or pT DNA (50 ng) with single stranded probe (1.10^4 cpm) prepared from either pN or pT in a solution containing 0.6 M NaCl, boiling for 5 min and annealing at 50°C for 1 h. After ethanol precipitation the annealing products were digested with PstI resulting in the 234 bp fragment that was analysed by DGGE.

Denaturing gradient gel electrophoresis

Samples were analysed on either parallel or perpendicular denaturing gradient gels. Perpendicular gel analysis was performed in 2.5 mm thick 6% polyacrylamide gels containing a 0–80% denaturant concentration gradient (80% = 5.6 M Urea and 32% deionized formamide) perpendicular to the direction of electrophoresis. The gels were run at 60°C at 150 V for 2.5 h in 1 × TAE. Parallel gel analysis was performed in 1 mm thick 6% polyacrylamide gels containing a 35–65% denaturant concentration gradient. Electrophoresis was performed for 8 h at 150 V at 60°C in 1 × TAE. Separation patterns were visualized either by ethidium bromide (EtdBr) staining at 0.1 μg ml^{-1} or, when labelled probe was used, by autoradiography at room temperature using Kodak XAR film.

Results

Melting maps

In Figure 1 the schematic organization is shown of the cHa-*ras*1 proto-oncogene together with the melting map which was calculated from basepair position 0 to 4500. Several features can be discerned in the T_m(50) plot, such as the HTF island representing a very high melting

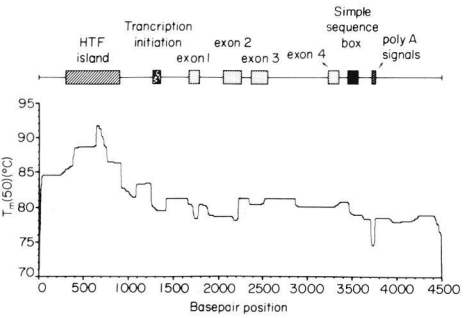


Figure 1 Calculated melting map of the human cHa-*ras*1 proto-oncogene using basepairs 0 to 4500 from the GENBANK entry (HUMHRAS1). HTF = HpaII tiny fragments

domain and the polyadenylation signal region which stands out as a small dip in T_m(50). Strikingly, large parts of all four exons are low melting domains. The melting maps constructed for the 234 bp PstI inserts of pN and pT (spanning positions 1640 to 1853 of the gene shown in Figure 1) is shown in Figure 2. The lowest melting domain spans base position 20 to 150 for both plasmids but the T_m50 of this domain was 0.4°C lower for the pT PstI insert due to the presence of the G to T transition at position 79 in this domain.

Denaturing gradient gel electrophoresis

Figure 3 shows the perpendicular DGGE analysis of the PstI insert of pN. Two inflections can be observed: a major one at 48% denaturant concentration which corresponds with the lowest melting domain and a minor one at 52% denaturant concentration, which indicates melting of the highest melting domain and beyond which the strands will be completely separated.

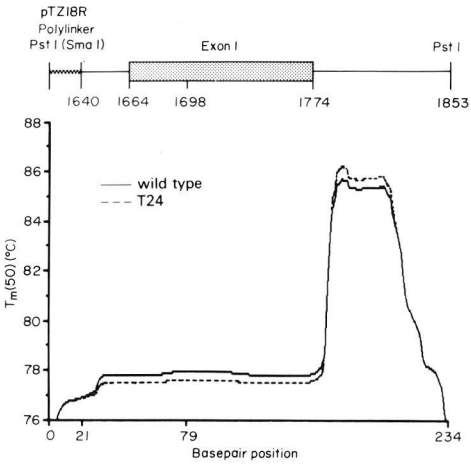


Figure 2 Calculated melting map of the first exon of the human cHa-*ras*1 proto-oncogene and the T24 variant sequence as present in the 234 bp PstI fragments used for DGGE analysis

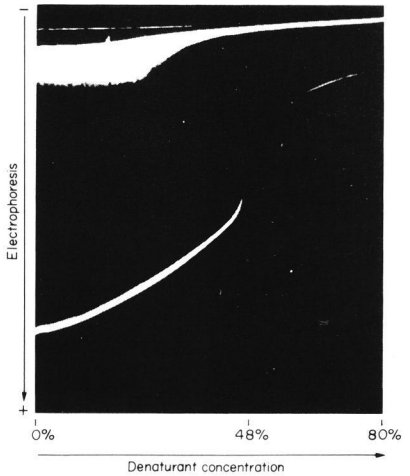


Figure 3 Perpendicular denaturing gradient gel analysis of the cloned 234 bp PstI fragment containing the first exon of the human cHa-*ras*1 proto-oncogene. The thick band in the top of the gel is the plasmid DNA from which the PstI insert was excised

We subsequently analysed pN and pT PstI fragments on a 35–55% parallel DGGE gel. As shown in Figure 4 the normal and T24 derived fragments migrate to slightly different positions in the gradient after prolonged electrophoresis. The pN migrated 2 mm further in the gel than the pT PstI fragment which was sufficient to separate both fragments in a mixture (right lane in Figure 4).

Since the DGGE gel system has been reported to be exquisitely sensitive for mismatches in heteroduplexed molecules we performed heteroduplex analysis of the two PstI inserts. Figure 5 shows the parallel DGGE gel of the heteroduplexed mixture. In addition to the homoduplex bands we observed two extra bands which were halted earlier in the gel due to the presence of mismatches. In order to identify these bands we performed heteroduplex analysis with labelled single strands derived from both pN and pT plasmids. Figure 6a shows a schematic representation of the production and sequence composition of the

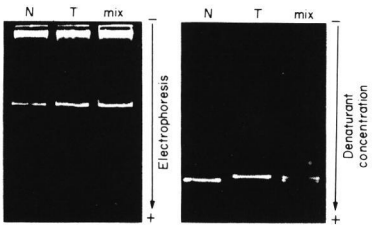


Figure 4 Parallel electrophoresis of the 234 bp PstI fragment containing the first exon of the human cHa-*ras*1 proto-oncogene. N stands for the wild type sequence and T for the T24 variant. On the left results are shown for neutral 6% polyacrylamide gel electrophoresis, run at 150 V for 1 h at 60°C, and on the right the denaturing gradient gel electrophoretic separation pattern. Mix is a mixture of the wild type and T24 derived 234 bp PstI fragment

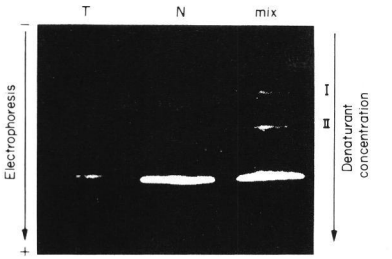


Figure 5 Parallel denaturing gradient gel analysis of the heteroduplex molecules of the 234 bp PstI fragments. N stands for the wild type homoduplex sequence, T for the T24 homoduplex sequence and mix stands for the heteroduplex mixture. Note that in this experiment the two homoduplex molecules are not separated probably due to overloading. I indicates the CT-mismatch containing heteroduplex and II the GA-mismatch containing heteroduplex

labelled strands and the annealing products. In Figure 6b the DGGE analysis of the heteroduplex experiment is shown. The CT-mismatch is less stable than the GA-mismatch, as demonstrated by the position of the CT-mismatch band at lower denaturant concentration (i.e. higher in the gel) than the GA-mismatch. From this we infer the upper band in the plasmid heteroduplex mixture

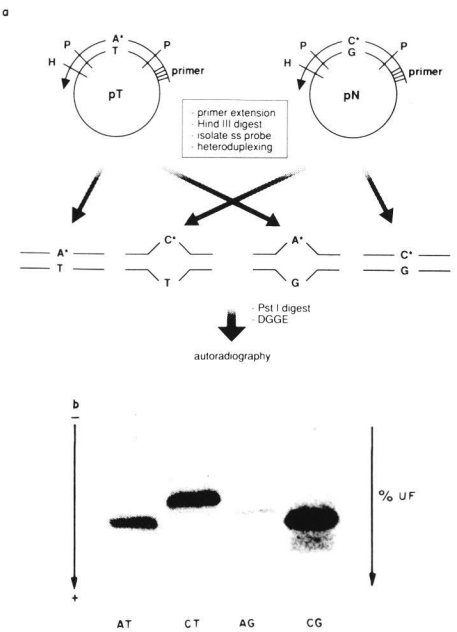


Figure 6 (**a**) Schematic representation of the procedure followed to obtain single stranded, radiolabelled probe derived from the wild type and T24 mutant 234 bp PstI fragment. H = HindIII recognition site; P = PstI recognition site. (**b**) Autoradiograph of the denaturing gradient gel analysis of the homo- and heteroduplex molecules formed using the labelled single-strand probes for the 234 bp PstI fragment. Lettering below the lanes refers to the basepairs at position 79

(I in Figure 5) to be composed of CT-mismatch containing molecules and the lower band (II in Figure 5) to contain the GA-mismatch containing molecules.

Discussion

Much of the molecular genetic analysis of *ras*-proto-oncogenes has focussed on the identification of point mutations which are associated with the acquisition of transforming capacities of the ras proteins. Most of these point mutations have been discovered by sequence analysis of cloned genes and screened for by the oligo dot-blot hydridization method in combination with the PCR method (Bos *et al.*, 1987).

Here we describe the calculation of a melting map for the entire cHa-*ras*1 proto-oncogene sequence, which was used to determine the position of low melting domains amenable to DGGE analysis. From the calculation of the melting map for the first exon containing the T24 point mutation, a lower $T_m(50)$ was predicted for the lowest melting domain in this sequence than for the wild type sequence. In accordance with this prediction, the T24 point mutation in the first exon could be detected by DGGE analysis of both homoduplex molecules as well as heteroduplex molecules. Separation of the homoduplex wild type molecule from the T24 mutant form was very poor when compared to separation of heteroduplex molecules. By using labelled single stranded DNA we could identify the heteroduplex bands and characterize the CT-mismatch as a more thermolabile mismatch in DGGE.

The melting map of the entire cHa-*ras*1 proto-oncogene revealed that large parts of all four exons represent low melting domains. This feature makes the exons of this gene amenable to screening for sequence variations by DGGE. The results presented here and elsewhere (Myers *et al.*, 1985; Cariello *et al.*, 1988; Uitterlinden & Vijg 1989) demonstrate that there is good correllation between predictions by the melting program on the influence of sequence alterations on melting temperature and the detection of sequence variants with altered electrophoretic mobility in DGGE gels. This can be used to predict where particular sequence variants will be in a DGGE gel after electrophoresis (Lerman *et al.*, 1986). In reverse, this could also be used to identify bands with an abnormal electrophoretic mobility in DGGE gels, as particular sequence variants. As demonstrated for other sequences, DGGE can also be applied to the analysis of genomic DNA, either directly by heteroduplexing (Myers *et al.*, 1985), blotting of separation patterns to hybridization membranes (Borresen *et al.*, 1988; Uitterlinden *et al.*, 1989) or by PCR amplification of the corresponding genomic fragments (Cariello *et al.*, 1988). In addition, the attachment of a G + C-rich fragment to the primers used in the PCR reaction also allows the higher melting domains of a particular sequence to be analysed by DGGE (Sheffield *et al.*, 1989). It has been shown that DGGE is also sensitive to the presence of methylated bases in melting domains (Collins & Myers 1987). A disadvantage of the PCR approach would therefore be that possible 5-methyl-cytosine modifications would not be detected. An alternative approach could be to transfer separation patterns of genomic DNA digests to hybridization membranes. Subsequent multiplex hybridization analysis with probes covering all regions of interest could then reveal shifts, including those due to methylation, in DGGE gels of homoduplex molecules.

Since the LMDs of the four cHa-*ras*1 exons have different $T_m(50)$ values it is theoretically possible to screen these sequences for mutations in all exons simultaneously by using only a single lane of a denaturing gradient gel. Either of the aforementioned methods therefore represent efficient and complementary ways of screening large numbers of genomic DNA samples for unknown sequence alterations.

Acknowledgements

We thank Dr D.A. Spandidos (Beatson Institute, Glasgow, United Kingdom) for his kind gift of cosmids HOMER 6N and HOMER 6T, Dr L.S. Lerman (MIT, Massachusetts, USA) for the use of the computer algorithm and Mr E.J. van der Reyden for the preparation of the figures.

References

Borresen, A.-L., Hovig, E. & Broger, A. (1988). Detection of base mutations in genomic DNA using denaturing gradient gel electrophoresis (DGGE) followed by transfer and hybridization with gene-specific probes. *Mut. Res.*, **202,** 77–83.

Bos, J.L., Toksoz, D., Marshal, C.J., Verlaan-de Vries, M., Veeneman, G.H., Van der Eb, A.J., Van Boom, J.H., Janssen, J.W.G. & Steenvoorden, A.C.M. (1985). Amino acid substitutions at codon 13 of the N-*ras* oncogene in human acute myeloid leukaemia. *Nature*, **315,** 726–730.

Bos, J.L., Fearon, E.R., Hamilton, S.R., Verlaan-de Vries, M., Van Boom, J.H., Van der Eb, A.J. & Vogelstein, B. (1987). Prevalence of *ras* gene mutations in human colorectal cancers. *Nature*, **327,** 293–297.

Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H. & Goeddel, D.V. (1983). Complete nucleotide sequences of the T24 human bladder carcinoma oncogene and its normal homologue. *Nature*, **302,** 33–37.

Cariello, N.F., Scott, J.K., Kat, A.G., Thilly, W.G. & Keohavong, P. (1988). Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using *in vitro* DNA amplification: $HPRT_{Munich}$. *Am. J. Hum. Genet.*, **42,** 726–734.

Cohen, J.B. & Levinson, A.D. (1988). A point mutation in the last intron responsible for increased expression and transforming activity of the c-Ha-*ras* oncogene. *Nature*, **334,** 119–124.

Collins, M. & Myers, R.M. (1987). Alterations in DNA helix stability due to base modifications can be evaluated using denaturing gradient gel electrophoresis. *J. Mol. Biol.*, **198,** 737–744.

Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M. & Wigler, M. (1984). Analysis of the transforming potential of the human H-*ras* gene by random mutagenesis. *Proc. Natl. Acad. Sci. USA*, **81,** 4008–4012.

Fischer, S.G. & Lerman, L.S. (1979). Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. *Cell*, **16,** 191–200.

Fischer, S.G. & Lerman, L.S. (1983). DNA fragments differing by single base-pair substitution are separated in denaturing gradient gels: Correspondence with melting theory. *Proc. Natl. Sci. USA*, **80,** 1579–1583.

Leider, J.M., Palese, P. & Smith, F.I. (1988). Determination of the mutation rate of a retrovirus. *J. Virol.*, **62,** 3084–3091.

Lerman, L.S., Fischer, S.G., Hurley, I., Silverstein, K. & Lumelsky, N. (1984). Sequence-determined DNA separations. *Ann. Rev. Biophys, Bioeng.*, **13,** 399–423.

Lerman, L.S., Silverstein, K. & Grinfield, E. (1986). Searching for gene defects by denaturing gradient gel electrophoresis. In: Cold Spring Harbor Symposia on Quantative Biology, Vol. LI *Molecular Biology of Homo sapiens*. Cold Spring Harbor Laboratory: New York, pp. 285–298.

Myers, R.M., Fischer, S.G., Lerman, L.S. & Maniatis, T. (1985a). Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Res.*, **13,** 3131–3145.

Myers, R.M., Lumelsky, N., Lerman, L.S. & Maniatis, T. (1985b). Detection of single base substitutions in total genomic DNA. *Nature*, **313,** 495–498.

Premkumar Reddy, E., Reynolds, R.K., Santos, E. & Barbacid, M. (1982). A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. *Nature*, **300,** 149–152.

Sheffield, V.C., Cox, D.R., Lerman, L.S. & Myers, R.M. (1989). Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single base changes. *Proc. Natl. Acad. Sci. USA*, **86,** 232–236.

Spandidos, D.A. & Wilkie, N.M. (1984). Malignant transformation of early passage rodent cells by a single mutated human oncogene. *Nature*, **310,** 469–475.

Uitterlinden, A.G., Slagboom, P.E., Knook, D.L. & Vijg, J. (1989). Two-dimensional DNA fingerprinting of human individuals. *Proc. Natl. Acad. USA*, **86,** 2742–2746.

Uitterlinden, A.G. & Vijg, J. (1989). Two-dimensional DNA typing. *TIBTECH*, **7,** 336–341.

ADDENDUM 2

Two-dimensional DNA fingerprinting of human individuals

[Proc. Natl. Acad. Sci. USA (1989) 86, 2742-2746]

Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 2742–2746, April 1989
Genetics

Two-dimensional DNA fingerprinting of human individuals

(DNA polymorphisms/minisatellites/variable number of tandem repeat sequences/denaturing gradient gel electrophoresis/gene mapping)

ANDRÉ G. UITTERLINDEN, P. ELINE SLAGBOOM, DICK L. KNOOK, AND JAN VIJG

Department of Molecular Biology, TNO Institute for Experimental Gerontology, PO Box 5815, 2280 HV Rijswijk, The Netherlands

Communicated by Leonard S. Lerman, November 28, 1988 (received for review June 9, 1988)

ABSTRACT **The limiting factor in the presently available techniques for the detection of DNA sequence variation in the human genome is the low resolution of Southern blot analysis. To increase the analytical power of this technique, we applied size fractionation of genomic DNA restriction fragments in conjunction with their sequence-dependent separation in denaturing gradient gels; the two-dimensional separation patterns obtained were subsequently transferred to nylon membranes. Hybridization analysis using minisatellite core sequences as probes resulted in two-dimensional genomic DNA fingerprints with a resolution of up to 625 separated spots per probe per human individual; by conventional Southern blot analysis, only 20–30 bands can be resolved. Using the two-dimensional DNA fingerprinting technique, we demonstrate in a small human pedigree the simultaneous transmission of 37 polymorphic fragments (out of 365 spots) for probe 33.15 and 105 polymorphic fragments (out of 625 spots) for probe 33.6. In addition, a mutation was detected in this pedigree by probe 33.6. We anticipate that this method will be of great use in studies aimed at (*i*) measuring human mutation frequencies, (*ii*) associating genetic variation with disease, (*iii*) analyzing genomic instability in relation to cancer and aging, and (*iv*) linkage analysis and mapping of disease genes.**

The possibility to identify DNA sequence heterogeneity is of major importance for the analysis of genetic diseases and genomic instabilities. This identification depends on the availability of probes that detect variable sites in the genome and on the resolution of electrophoretic separation techniques for the analysis of DNA restriction fragments.

The discovery of hyperpolymorphic VNTR (variable number of tandem repeat) DNA sequences or minisatellites has greatly facilitated studies of genetic variation in the human population (1–6). It has been demonstrated that so-called core probes derived from minisatellites can be used to simultaneously detect a large number of hyperpolymorphic VNTR loci, dispersed in the genome, to provide genetic fingerprints of human individuals (7, 8). Core sequences have been successfully applied in the analysis of tumors for genetic instability (9) and in linkage analysis of genetic diseases (3). Such applications rely on the resolution of Southern blot hybridization analysis, which is based on the gel electrophoretic separation of genomic DNA restriction fragments according to size (10). One-dimensional separation of DNA fragments allows only ≈30 hypervariable minisatellite fragments to be resolved (ref. 3; this publication).

It has been demonstrated that by combining ordinary size separation with sequence separation in denaturing gradient gels, all restriction fragments in an *Eco*RI digest of the *Escherichia coli* genome can be resolved (11–13). To investigate whether this principle can be applied to the analysis of DNA sequence variation in the mammalian genome, we separated genomic DNA restriction fragments of human individuals according to size and base-pair composition by neutral and denaturing gradient polyacrylamide gel electrophoresis, respectively. By subsequent transfer of the electropherograms to nylon membranes and hybridization with radiolabeled minisatellite core sequences as probes, high-resolution DNA fingerprints were obtained. We show that many polymorphic DNA restriction fragments can be detected in the molecular size region of fragments <3 kilobase pairs (kbp), which is not accessible for conventional Southern hybridization analysis. We demonstrate the applicability of this technique in genetic studies on humans by showing that a large number of transmitted polymorphic spots can be simultaneously followed in a two-generation human pedigree of three members.

MATERIALS AND METHODS

DNA Isolation and Restriction Enzyme Digestion. Genomic DNAs, isolated from peripheral blood lymphocytes according to standard procedures, were a kind gift from E. Bakker. DNAs were digested with the restriction endonucleases *Hae* III or *Hin*fI (BRL) under conditions recommended by the manufacturer.

Electrophoretic Separations. Agarose gel electrophoresis of 5 μg of DNA restriction fragments was performed in a horizontal 1.2% gel in 1× TAE (40 mM Tris·HCl, pH 7.4/20 mM sodium acetate/1 mM NaEDTA) at 65 V for 14–30 hr. Gels were stained for 10 min in a solution containing ethidium bromide (EtdBr) at 0.1 μg/ml followed by destaining for at least 30 min. Two-dimensional separations of 5 μg of DNA restriction fragments were performed in 1-mm-thick polyacrylamide gels (acrylamide/bisacrylamide, 37:1) using a gel apparatus that was essentially the same as the one described by Fischer and Lerman (12). The first dimension was run in a neutral 6% gel at 50°C for 2 hr at 250 V in 1× TAE. The separation patterns were visualized by staining the gel in the dark with EtdBr (0.1 μg/ml) for 10 min, followed by destaining for at least 30 min. The 0.34- to 2.8-kbp region (probe 33.15) or the 0.54- to 10-kbp region (probe 33.6) of the lane was quickly cut out of the gel and applied to a 6% polyacrylamide gel containing a 10–75% linear concentration gradient of denaturant (100% denaturant = 7.0 M urea/40% formamide) parallel to the direction of electrophoresis. This gradient was found to give optimal separation patterns for the VNTR sequences. Gels were poured by mixing two solutions, containing the desired boundary denaturant concentrations, in a linear gradient maker with a peristaltic pump. After electrophoresis for 12 hr at 225 V and 60°C, the gel was stained with EtdBr as described above.

Transfer of Separation Patterns. DNA separation patterns in agarose gels were capillary transferred to a nylon membrane (Nytran 13N, Schleicher & Schuell; Zetaprobe, Bio-Rad) in 0.4 M NaOH/0.6 M NaCl for 12 hr. After transfer, the filter was rinsed in 2× SSC (1× SSC = 150 mM NaCl/15 mM

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EtdBr, ethidium bromide; VNTR, variable number of tandem repeats.

sodium citrate), air dried, baked for 1 hr at 80°C, and irradiated with 302-nm UV light (Transilluminator, UVP Products, San Gabriel) for 45 s, which was found to be optimal for cross-linking DNA to the filter (results not shown). DNA separation patterns in denaturing gradient polyacrylamide gels were fragmented by irradiating the gel with 302-nm UV light (Transilluminator, UVP Products) for 4 min, which was found to be optimal (results not shown). Before transfer, the gel was boiled for 5 min in 1× TBE (89 mM Tris·borate, pH 8.0/89 mM boric acid/2 mM NaEDTA) and subsequently placed in an identical solution at room temperature. Transfer to a nylon membrane (Nytran 13N, Schleicher & Schuell; Zetabind, Bio-Rad) was achieved by semidry electroblotting at 400 mA (6–28 V) between horizontal graphite plates. Electrophoresis was performed twice for 45 min between 10 Whatman 3MM paper sheets, which were soaked in fresh 1× TBE between the two transfers. After transfer, the filter was rinsed in 2× SSC, air-dried, baked for 1 hr at 80°C and irradiated for 45 s with 302-nm UV light to cross-link the DNA fragments to the filter.

Probe Preparation and Labeling. The probe was prepared by using T4 kinase (Boehringer Mannheim) to individually phosphorylate the 5′-hydroxyl groups of two partially complementary and overlapping oligonucleotides, each representing the complete 33.15 (5′-AGAGGTGGGCAGGTGG-3′ and 5′ CCACCTCTCCACCTGC-3′) or 33.6 (5′-AGGGCTG-GAGG-3′ and 5′-AGCCCTCCTCC-3′) minisatellite core sequence (7). Subsequently, the two 33.15 or 33.6 oligonucleotides were mixed and allowed to anneal at 57°C for 1 hr, followed by ligation according to standard procedures (14). The synthetic probes thus prepared had an average length of 500 bp or more. Occasionally, T4 kinase phosphorylation and ligation were repeated to increase the average length of the probe. The ligation products (20 ng) were [α-^{32}P]dCTP-labeled either by the random-primer oligolabeling method (Boehringer Mannheim) or by self-priming, after boiling for 5 min and reannealing at 37°C in the presence of 1 unit of Klenow enzyme (Boehringer Mannheim)/2 μM dNTP/50 mM Tris·HCl, pH 7.2/10 mM $MgCl_2$. Specific activities of 3 × 10^8–8 × 10^8 cpm/μg were obtained.

Hybridization Analysis. Filters were prehybridized in 5× SSC/20 mM sodium phosphate, pH 7.2/1% SDS/1 mM NaEDTA/heparin (50 μg/ml) for 2 hr at 65°C. After adding denatured probe at a concentration of 1 × 10^6 cpm/ml, hybridization was performed for 12 hr at 65°C. The filter was washed three times for 5 min at room temperature and three times for 20 min at 65°C in 2.5× SSC/0.1% SDS. Autoradiography was performed for 12–48 hr at −80°C using fine intensifying screens and XAR-5 film (Kodak). For subsequent rehybridizations of the filters, the probe was removed by boiling the filter for 20 min in a solution containing 0.01× SSC and 0.1% SDS. Filters were rinsed in 2.5× SSC and hybridization analysis was performed as described above. Two-dimensional spot patterns were interpreted by eye examination, using grids to quantitate the spots and to match individual two-dimensional fingerprints.

RESULTS

In Fig. 1, Southern blot autoradiographs are shown of a human pedigree (of six members) obtained after prolonged electrophoresis in agarose gels and by using the synthetic probes derived from minisatellite core sequences 33.15 and 33.6 (7) on *Hae* III- and *Hin*fI-digested genomic DNA. A

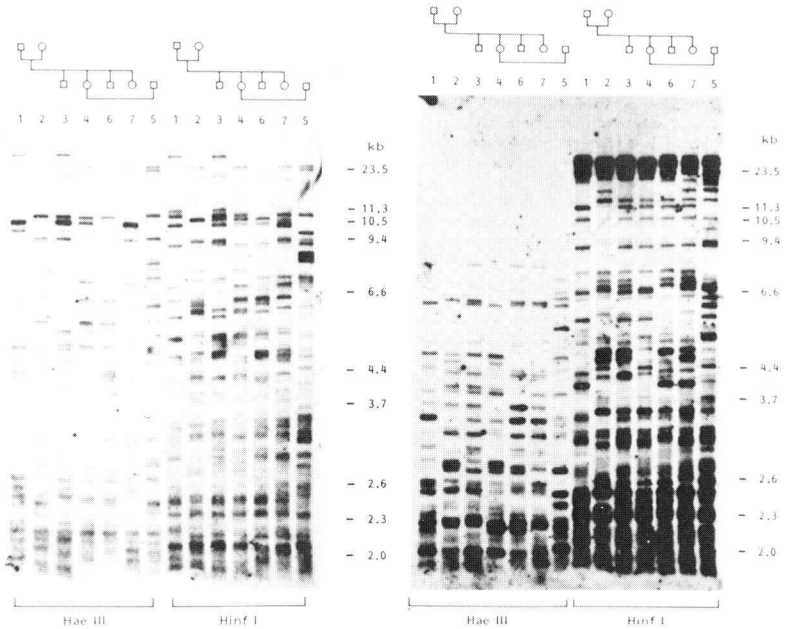


FIG. 1. Southern blot hybridization analysis of DNAs isolated from six members of a human pedigree (DL63; schematically depicted above the lanes in the autoradiographs) and one from an unrelated individual (DL 63.5), digested with *Hae* III or *Hin*fI, and using 35.15 (*Left*) or 33.6 (*Right*) minisatellite core sequences as probe. The autoradiograph on the *Right* was obtained by rehybridizing the filter that was used to obtain the autoradiograph on the *Left*.

number of simultaneously transmitted polymorphic bands were detected in this pedigree in the high molecular weight area of restriction fragments. For the restriction enzyme *Hin*fI, this number was 28 for probe 33.15 and 33 for probe 33.6. With *Hae* III, the number was 28 for probe 33.15 and 22 for probe 33.6. After *Hae* III digestion, the size range of bands detected by probe 33.6 was found to be smaller than after digestion with *Hin*fI, which could be an indication for the presence of *Hae* III recognition sites in the minisatellites homologous to this probe.

When genomic DNA was digested with *Hae* III and subjected to a two-dimensional separation (on the basis of size and base-pair sequence in neutral and denaturing gradient polyacrylamide gels, respectively), clusters of restriction fragments were observed in the EtdBr-stained two-dimensional gel as shown in Fig. 2 (*Left*). A large cluster is seen in the upper part of the gel, containing low concentrations of denaturant, and a smaller cluster is located near the bottom of the gel, which contains the highest concentration of denaturant.

Hybridization analysis of VNTR sequences detected with probe 33.6 resulted in a more or less evenly spread spot pattern. The total number of spots observed for probe 33.6 under the stringency of washing applied here (2.5× SSC) was 545 for this individual (DL 63.4 from the pedigree shown in Fig. 1). Slightly smaller numbers were obtained when the stringency was increased to 1× SSC (results not shown).

We subsequently analyzed three members of the pedigree described above (mother, father, and a son) by two-dimensional DNA fingerprinting. For optimal comparisons a 30-cm-wide version of the gel apparatus originally described by Fischer and Lerman (13) was constructed and used in these experiments so that up to six individuals could be compared on one denaturing gradient gel. Close inspection and comparison of individual spot patterns of the two parents obtained with probes 33.6 (Fig. 3) revealed a total number of 569 spots for the father, 607 for the mother, and 625 for the son. Between the two parents, 150 spot polymorphisms were observed, 105 (70%) of which were transmitted to the son (52 of maternal and 53 of paternal origin). Details of the separation patterns are shown in Fig. 3 (*Lower Left*). Using probe 33.6, we detected a fragment in the son that was not present in the mother or the father (Fig. 3 *Lower Right*) and four fragments, common to the parents, could not be detected in the son (two examples are shown in Fig. 3 *Lower Right*). With probe 33.15, considerably less VNTR-containing fragments were detected than with probe 33.6 (Fig. 4). Among the 372, 290, and 365 spots for the father, mother, and son, respectively, 50 spot polymorphisms could be detected. Among the 37 transmitted spot polymorphisms (74%), 17 were of paternal and 20 were of maternal origin. Some of these spot polymorphisms are shown in detail in Fig. 4 (*Lower*).

DISCUSSION

With the two-dimensional DNA fingerprinting system presented here, we were able to distinguish up to 625 spots per individual for probe 33.6 and 372 for probe 33.15. Since we did not observe severe clustering of spots, it is likely that we have resolved all VNTR sequences belonging to the sets of sequences detected with these probes. The difference between the two probes in number of spots detected could therefore be due to a different copy number of these sets of repetitive sequences. In addition to 33.15 and 33.6, we have also used other core sequences (5) as probes in two-dimensional DNA fingerprinting and obtained comparable results. In this respect, VNTR sequences appear to exhibit an exceptional distribution of spots over a considerable part of the denaturing gradient in the second-dimension gel. By contrast, as illustrated in Fig. 2 (*Left*), total genomic DNA digests have a tendency to cluster in the second dimension.

Most spots were found in the 20–50% denaturant range and should therefore have a high to medium AT/GC ratio of their lowest melting domain. Although the number of repeat units of any VNTR locus is high enough to generate a melting domain, both 33.15 and 33.6 are G+C-rich. Therefore, the gradient level of the majority of homologous restriction fragments in the two-dimensional pattern cannot be determined by the VNTR sequences themselves. Instead, the position of most of the spots in the gradient will be determined by sequences adjacent to the VNTR regions. Virtually all VNTR alleles with a particular locus are therefore iso-

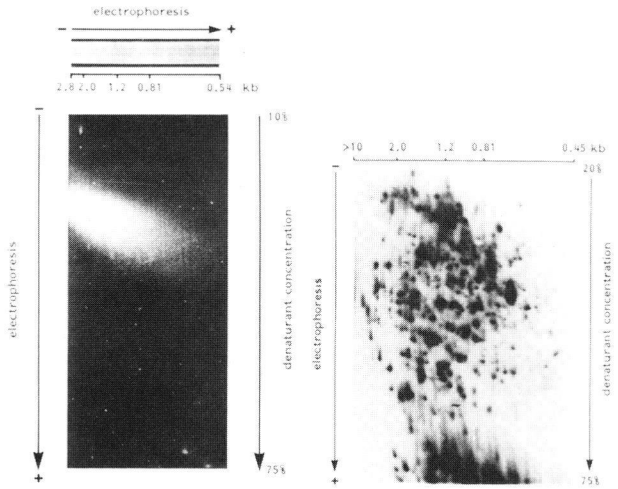


FIG. 2. Two-dimensional DNA fingerprint analysis of a human individual (DL 63.4). The two-dimensional separation pattern is shown after EtdBr staining (*Left*) and after hybridization analysis of the nylon replica filter with probe 33.6 (*Right*).

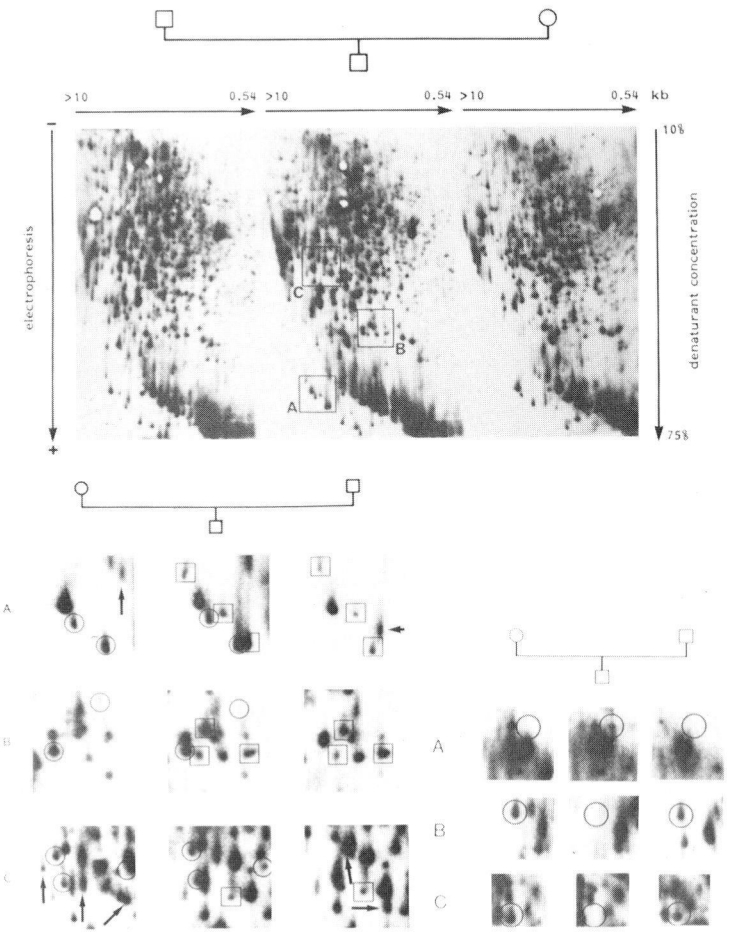


FIG. 3. (*Upper*) Two-dimensional DNA fingerprint analysis of *Hae* III-digested genomic DNA from three members of the human pedigree DL63 (DL63.1, -2, and -3), using probe 33.6. (*Lower Left*) Details from three different areas indicated in *Upper*, showing the transmission of particular spot polymorphisms. ○, Maternal fragments; □, paternal fragments. Arrows indicate nontransmitted polymorphic fragments. (*Lower Right*) Details from *Upper*, demonstrating the presence of a VNTR fragment in the son that is absent from the parents (circles in row A) and two cases in which both parents share a spot that is absent in the son (circles in rows B and C).

thermal—i.e., reach identical positions in the denaturing gradient. This characteristic and the fact that in the two-dimensional system virtually no sequences detected with a particular core probe are lost in a smear allows one to identify them.

When individuals from a human pedigree were analyzed by two-dimensional DNA fingerprinting, the majority of spots detected with both probes were found to be monomorphic (75–85%). However, a large number of spot polymorphisms (up to 150 for probe 33.6) were observed. Evidence for heterozygosity of the parents at particular VNTR loci detected by probe 33.6 was provided in four cases in which the parents shared a spot that could not be demonstrated in the son (Fig. 3 *Lower Right*). The percentage of spot polymorphism transmitted to the son was found to be 70% for both probes 33.6 and 33.15. This phenomenon could be due to clustering in the genome of VNTR loci or to the presence of one or more *Hae* III restriction sites in the minisatellites themselves. In the latter case, several polymorphic spots could stem from the same minisatellite locus, thereby resulting in a number of spots being cotransmitted as minisatellite haplotypes (3). This possibility is supported by the data in Fig. 1 (*Right*), which indicate the presence of *Hae* III sites in large alleles detected by probe 33.6 in these individuals. The spot present in the son but not in the parents is likely to have arisen by unequal exchange between two VNTR regions rather than by point mutation in a VNTR region (6).

On the basis of the results obtained with VNTR core sequence probes, we anticipate that two-dimensional DNA fingerprinting can be applied in a number of different areas in genetic research. The method should be useful in the analysis of large parts of the genome for measuring mutation frequencies (15) and for detecting putative changes in VNTR loci or other unstable DNA regions during tumor induction and growth (9, 16) or during aging (17).

Two-dimensional DNA fingerprinting can be used in association studies and to extend and improve linkage analysis and gene mapping. In this respect, it should be noted that the introduction of the denaturing gradient separation principle offers a solution to the problem of not being able to identify the same VNTR locus in different pedigrees, which effectively constrains widespread application of one-dimensional

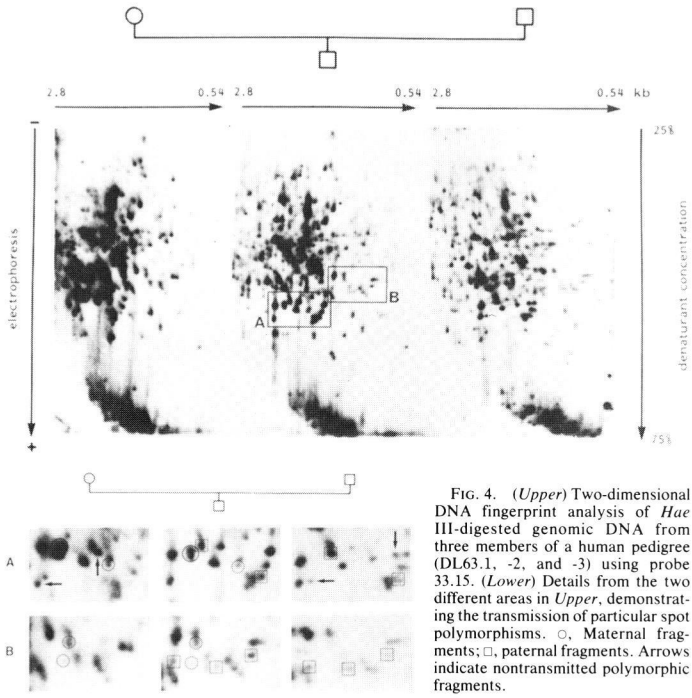


FIG. 4. (*Upper*) Two-dimensional DNA fingerprint analysis of *Hae* III-digested genomic DNA from three members of a human pedigree (DL63.1, -2, and -3) using probe 33.15. (*Lower*) Details from the two different areas in *Upper*, demonstrating the transmission of particular spot polymorphisms. ○, Maternal fragments; □, paternal fragments. Arrows indicate nontransmitted polymorphic fragments.

DNA fingerprinting to genetic analysis. Sequence-specific separation allows one to identify alleles with a particular locus in different pedigrees on the basis of their position on the same isotherm (see above). Furthermore, by using locus-specific probes in parallel experiments, each VNTR core homologous spot in a two-dimensional gel can be identified by comparison.

We especially acknowledge Dr. Stuart Fischer's expert advice and kind hospitality during the initial stages of this work when one of us (A.G.U.) stayed in his laboratory. We thank Dr. Leonard S. Lerman for many helpful comments on the manuscript and Drs. J. A. Gossen, J. H. J. Hoeijmakers, H. Schellekens, and G. J. J. M. Trommelen for helpful discussions. We also thank Dr. E. Bakker (Department of Human Genetics, State University of Leiden, Leiden, The Netherlands) for his kind gift of DNAs from human pedigree DL63 and Mr. E. J. van de Reyden for the preparation of the photographs. Part of this research was supported by Senetek PLC. This work is the subject of a patent application.

1. Wyman, A. R. & White, R. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 6754–6758.
2. Reeders, S. T., Breuning, M. H., Davies, K. E., Nicholls, R. D., Jarman, A. P., Higgs, D. R., Pearson, P. L. & Weatherall, D. J. (1985) *Nature (London)* **317,** 542–544.
3. Jeffreys, A. J., Wilson, V., Thein, S. L., Weatherall, D. J. & Ponder, B. A. J. (1986) *Am. J. Hum. Genet.* **39,** 11–24.
4. Donis-Keller, H., Green, P., Helms, C., *et al.* (1987) *Cell* **51,** 319–337.
5. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. & White, R. (1987) *Science* **235,** 1616–1622.
6. Jeffreys, A. J., Royle, N. J., Wilson, V. & Wong, Z. (1988) *Nature (London)* **332,** 278–281.
7. Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985) *Nature (London)* **314,** 67–73.
8. Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985) *Nature (London)* **316,** 76–79.
9. Thein, S. L., Jeffreys, A. J., Gooi, H. C., Cotter, F., Flint, J., O'Connor, N. T. J., Weatherall, O. J. & Wainscoat, J. S. (1987) *Br. J. Cancer* **55,** 353–356.
10. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
11. Fischer, S. G. & Lerman, L. S. (1979) *Cell* **16,** 191–200.
12. Fischer, S. G. & Lerman, L. S. (1979) *Methods Enzymol.* **68,** 183–191.
13. Fischer, S. G. & Lerman, L. S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1579–1583.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
15. Delehanty, J., White, R. L. & Mendelsohn, M. L. (1986) *Mutat. Res.* **167,** 215–232.
16. Colb, M., Yang-Feng, T., Francke, U., Mermer, B., Parkinson, D. R. & Krontiris, T. G. (1986) *Nucleic Acids Res.* **14,** 7929–7937.
17. Vijg, J. & Uitterlinden, A. G. (1987) *Mech. Ageing Dev.* **41,** 47–63.

ADDENDUM 3

Denaturing gradient gel electrophoretic analysis
of minisatellite alleles

[Electrophoresis (1991) 12, 12-16]

Electrophoresis 1991, *12*, 12–16

André G. Uitterlinden
Jan Vijg

Medscand Ingeny, Rijswijk

Denaturing gradient gel electrophoretic analysis of minisatellite alleles

By two-dimensional DNA fingerprinting, an electrophoretic method which combines separation according to size with separation in a denaturing gradient, virtually all minisatellite sequences detected with a minisatellite core probe can be resolved (Uitterlinden *et al.*, *Proc. Natl. Acad. Sci. USA* 1989, *86*, 2742–2746). To investigate the electrophoretic behavior in denaturing gradient gels of allelic restriction fragments containing minisatellite sequences, we analyzed alleles of the two highly polymorphic minisatellite loci D7S22 and D2S44. The results obtained indicate that for these loci, depending on the restriction enzyme used to digest genomic DNA, alleles of different sizes migrate to regions of similar denaturant concentration, *i.e.* to isothermal positions in the denaturing gradient. Denaturing gradient gel electrophoresis also allows for the discrimination of restriction fragments which are the result of the presence of internal recognition sites in the minisatellite and, therefore, to distinguish between VNTR and restriction site polymorphisms.

1 Introduction

An important set of highly informative DNA markers is the category of minisatellite loci that contain a variable number of tandem repeat DNA sequences (VNTRs), a large number of which has been isolated [1, 2]. Due to extensive allelic variation in the number of repeat units per locus, the VNTR loci display high to very high heterozygosity values (70–99 %). By using minisatellite core probes, it has been demonstrated that several such loci can be simultaneously detected to provide a DNA fingerprint [3]. By combining size separation in neutral gels with sequence-dependent separation in denaturing gradient gels [4], we have demonstrated that it is possible to obtain high resolution two-dimensional DNA fingerprints by hybridization with minisatellite core probes [5]. Recently, two-dimensional separation patterns were obtained by using other repeat sequence probes as well, such as simple sequence motifs and inter-Alu fragments (manuscripts in preparation). Therefore, two-dimensional DNA typing could permit the analysis of genetic variations among individuals on a total genome scale. A major disadvantage, however, is the difficulty in identifying individual spots in the separation pattern as alleles of a specific locus.

We now report that alleles of two highly polymorphic VNTR loci, D7S22 and D2S44, detected by probes plambdaG3 [6] and pYNH24 [7], respectively, display similar and locus-specific electrophoretic behavior in denaturing gradient gels. Depending on the restriction enzyme used to digest genomic DNA, VNTR alleles, which range in size from 800 to as much as 14 000 basepairs (bp), migrate to isothermal positions in the denaturing gradient gel. This locus-specific behavior may facilitate the identification of minisatellite alleles in a two-dimensional DNA fingerprint.

Correspondence: Dr. A. G. Uitterlinden, Genetic Diagnostics Department, Medscand Ingeny, 2280 AJ Rijswijk, The Netherlands

Abbreviations: bp, basepairs; **DGGE**, denaturing gradient gel electrophoresis; **HMD**, high melting domain; **LMD**, low melting domain; **RFLP**, restriction fragment length polymorphisms; **SBA**, Southern blot analysis; **VNTR**, variable number of tandem repeats

2 Materials and methods

2.1 DNA isolation and restriction enzyme digestion

Genomic DNA was isolated from whole peripheral blood according to standard procedures. Restriction enzyme digestion of 5 µg of DNA was done for 1 h in a buffer recommended by the manufacturer (BRL).

2.2 Locus-specific VNTR probes and probe labelling

The locus-specific VNTR probe plambdaG3, detecting locus D7S22, [6] was a kind gift of Dr. A. J. Jeffreys. This probe contains an *Sau*3A/*Bam*HI insert, schematically depicted in Fig. 1. The probe pYNH24, detecting locus D2S44 [7], was a kind gift of Dr. Y. Nakamura. Plasmids were radiolabeled by the random prime labeling method.

2.3 Gel electrophoresis and hybridization analysis

For Southern blot analysis (SBA), restriction enzyme digests were electrophoresed in 1 % agarose gels in 1 × TAE (150 mM Tris-acetic acid, 150 mM sodium acetate, 1 mM Na_2EDTA) and transferred to Gene-Screen (NEN Biolabs) by vacu-blotting (LKB) in 0.4 N NaOH, 0.6 M NaCl for 45 min. Denaturing gradient gel electrophoresis (DGGE) and electrotransfer to Zetaprobe membranes (Bio-Rad) was performed as described [5]. For parallel denaturing gradient gels a 30 cm wide version of the gel apparatus was used. Hybridization analysis was performed in glass tubes in a hybridization oven (GFL) at 65 °C. Filters were prehybridized in 7 % sodium dodecyl sulfate (SDS), 0.5 M sodium phosphate buffer and 1 mM Na_2EDTA for 30 min and hybridization was carried out overnight at 65 °C in the same solution. In order to suppress cross-hybridization of the probes to other minisatellite loci and, in the case of plambdaG3, to *Alu*I repetitive elements, hybridization was performed in the presence of 100 µg/mL denatured, sheared human genomic DNA. The filters were washed once in 1 % SDS, 100 mM sodium phosphate buffer for 2 min at room temperature and once in 100 mM sodium phosphate buffer for 45 min at 65 °C. The filters were rehybridized by stripping at 65 °C in 50 % formamide, 0.5 × SSC (saline-sodium citrate), 25 mM/mL heparine, 10 mM sodium phosphate buffer, 0.5 mM Na_2EDTA and 0.5 % SDS for 30 min. Filters were exposed to Kodak XAR film in cassettes with intensifying screens for 2–48 h at −80 °C.

0173-0835/91/0101-0012 $3.50+.25/0

3 Results

3.1 Melting map of the D7S22 minisatellite region

By using a computer algorithm [8], melting maps were constructed for a number of G+C-rich minisatellites. The melting map of an allele of the D7S22 locus, detected by plambdaG3, containing 14 repeat units (Fig. 1) shows that the minisatellite region represents a high melting domain (HMD), whereas the adjacent sequences are low melting domains (LMD). Of the LMDs, the one 5′ to the minisatellite, between the *Alu*I repeat and the simple sequence region, represents the lowest melting domain. This domain is included in an *Hae*III restriction fragment encompassing the minisatellite region. Therefore, it can be assumed that this domain will determine the position of the *Hae*III restriction fragment after DGGE, irrespective of the size of the allele. *Hin*fI sites are outside the sequence shown in Fig. 1; therefore, it is unknown whether *Hin*fI fragments share the 5′ lowest melting domain with *Hae*III fragments. For the D2S44 locus, detected by pYNH24, no melting maps could be constructed due to a lack of sequence information. However, in addition to the plambdaG3 locus, several other G+C-rich minisatellite sequences were found to represent HMDs. It can therefore be predicted that, in general, G+C-rich minisatellite alleles will migrate to the same position in a denaturing gradient gel on the basis of their common flanking sequences, which represent lower melting domains.

3.2 Parallel DGGE analysis

In order to experimentally test the theoretical predictions made above, we analyzed genomic DNA, digested with the most commonly used DNA-fingerprinting enzymes *Hin*fI and *Hae*III. A total of 33 unrelated human individuals were analyzed with respect to the electrophoretic behavior in denaturing gradient gels of minisatellite-containing fragments from the two loci detected by plambdaG3 and pYNH24. For that purpose genomic digests were separated by DGGE and electrotransferred to a nylon membrane, which was subsequently hybridized with the locus-specific probes.

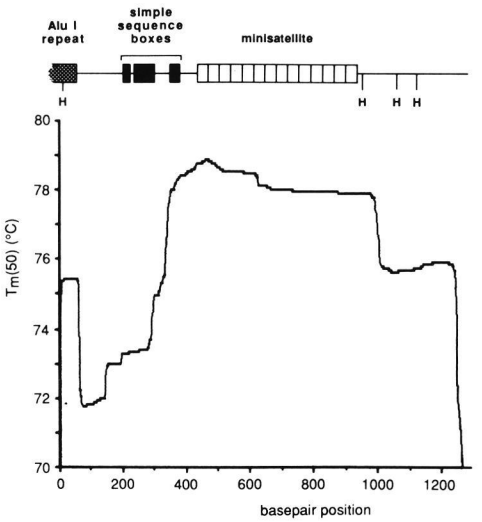


Figure 1. Calculated melting map of the D7S22 VNTR locus detected by plambdaG3. The sequence was taken from [6] with the minisatellite consisting of 14 repeat units of 37 bp (the first five repeat units as published and the following nine according to the consensus sequence with Y=T and R=A). The schematically depicted sequence at the top is included in the probe used for hybridization analysis.

In Fig. 2 results are shown for the *Hin*fI digests of genomic DNA from several individuals hybridized with plambdaG3 and pYNH24. For both these polymorphic loci, alleles of different sizes were observed after neutral agarose gel electrophoresis and SBA of the individuals tested. After DGGE, however, only a single hybridizing band could be observed at 38 % and 39 % denaturant concentration for plambdaG3 and pYNH24, respectively. The position of this band was identical for pYNH24 among the individuals, but varied slightly for plambdaG3, within a 3 % denaturant concentration range. When more narrow gradients were used (*i.e.* from 35–50 % denaturants) it was possible to separate the alleles of this locus by parallel DGGE (results not shown).

When human genomic *Hae*III digests were analyzed by DGGE, we obtained more complex banding patterns. For plambdaG3 (see Fig. 3) an intense band was observed in the DGGE gel at 38 % denaturant concentration, which we infer to be derived from the *Hae*III fragments containing the 5′ LMD plus the minisatellite sequence or a part thereof. The allelic fragments containing the 5′ domain (which range in size from 2.5 kb to about 14 kb) comigrate except for 9 individuals (15 %), whose short common allele (SCA, see [6]) is visible as a sharp band at 42 % denaturant concentration (note homozygous individual #31). This different electrophoretic behavior might be due to a particular sequence difference in the 5′ LMD, which all SCAs have in common, in contrast to the 5′ LMD in larger alleles. The variations in gradient position of the larger alleles detected in *Hae*III digests were also observed in *Hin*fI digests, albeit within a more narrow denaturant range (Fig. 2).

In *Hae*III digests of genomic DNA, probed with plambdaG3, we observed more than two strong bands per lane after neutral agarose gel electrophoresis and SBA in 12 out of 33 cases (18 %). This indicates the presence of *Hae*III recognition sites within the minisatellite region, a phenomenon which has been observed previously for this locus [6, 9]. This was also reflected by the presence of one or two hybridizing bands in the DGGE gel at about 50 % denaturant concentration (Fig. 3). These fragments most likely only contain the G+C-rich minisatellite sequences and no adjacent sequences. The presence of internal melting domains in the minisatellite region (Fig. 1) results in their position at higher denaturant concentrations in the denaturing gradient gel. The slight variation in the position of the bands, derived from internal fragments in the denaturing gradient, is independent of size and suggests differences in sequence composition of minisatellite repeat units. We could discern only six variants, a–f, differing in their position in the gradient (Fig. 3). The low frequency at which these variants occur, together with the low incidence of internal *Hae*III sites in the minisatellite region, indicate that the repeat units of VNTR alleles from this locus have a relatively high level of sequence homogeneity.

Neutral agarose gel electrophoresis and hybridization analysis of *Hae*III-digested genomic DNA, probed with pYNH24,

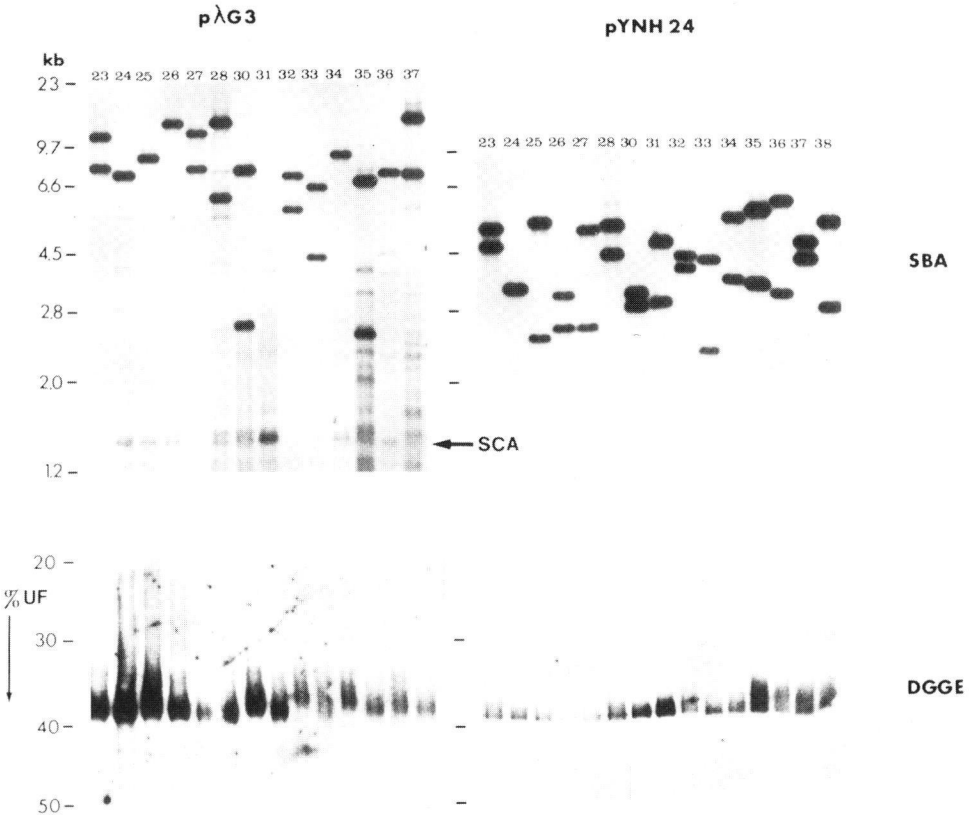


Figure 2. Hybridization analysis of the D7S22 and D2S44 loci, detected by plambdaG3 and pYNH24, respectively, in genomic DNA from 14 unrelated individuals. Genomic DNA was digested with *Hin*fI and separated by neutral agarose gel electrophoresis (SBA) and by DGGE. Results for pYNH24 were obtained by rehybridizing the filters used for the plambdaG3 hybridization. SCA, short common allele detected by plambdaG3 [6].

resulted in banding patterns which were almost identical to those obtained after *Hin*fI digestion, albeit with a slightly lower allele size (results not shown). After DGGE of the *Hae*III digestions, however, we observed uninterpretable banding patterns with pYNH24. One hybridizing fragment was observed at a denaturant concentration which was identical among the individuals but, in addition, we also observed two more fragments, at different and higher denaturant concentrations, in each individual (results not shown). In view of the smaller allele size after *Hae*III digestion in comparison to *Hin*fI digests, we infer this aberrant DGGE behavior to be due to the loss of a stable LMD in pYNH24 *Hae*III alleles.

3.3 Two-dimensional electrophoretic analysis

In order to obtain information on the origin of restriction fragments observed in parallel DGGE, we analyzed genomic digests by two-dimensional electrophoresis. Figure 4 shows a two-dimensional gel analysis of the plambdaG3 locus in *Hae*III and *Hin*fI-digested genomic DNA from two unrelated individuals. First the digests were separated according to size in neutral polyacrylamide, after which the lanes were perpendicularly applied to a denaturing gradient gel. As demonstrated by one-dimensional parallel DGGE (Fig. 3), the short common allele (0.8 kb) migrates to a slightly different gradients position than the *Hae*III fragments derived from the much larger alleles (7.5 and 9.5 kb). The two extra *Hae*III fragments of 2.5 and 2.6 kb in individual #26 migrate further into the gradient. As suggested by the data shown in Fig. 3, these fragments most likely represent restriction fragments which are generated as a consequence of two internal *Hae*III sites in the minisatellite sequences of the large allele in this individual (note that the two fragments appear as a single band of about 2.5 kb in Southern blot hybridization. This was confirmed by using *Hin*fI to digest the same genomic DNA samples. In that case only two spots could be observed under identical hybridization conditions. These spots correspond to the large and small allele; no internal fragments were observed.

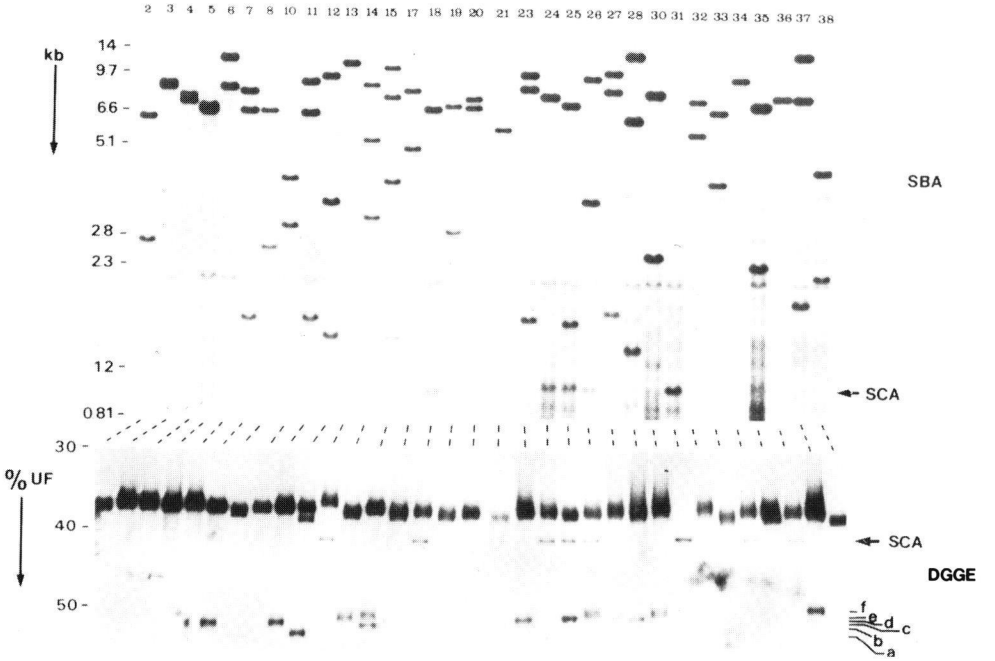


Figure 3. Hybridization analysis of the D7S22 locus in genomic DNA from 33 unrelated individuals. Genomic DNA was digested with *Hae*III and separated in neutral agarose gels (SBA) and in parallel DGGE. The letters a–f indicate different variants of internal *Hae*III fragments of the locus (see Section 3.2). SCA, short common allele detected by plambdaG3 [6].

4 Discussion

The results presented in this paper indicate that different size alleles of a VNTR locus, which belongs to the G+C-rich set of minisatellite sequences, can display similar electrophoretic behavior in denaturing gradient gels. We observed such isothermal electrophoretic behavior in DGGE gels for two such VNTR loci, D7S22 – detected by plambdaG3, and D2S44 – detected by pYNH24, in *Hin*fI digests of human genomic DNA. For VNTR locus D7S22, we could demonstrate that this size-independence is caused by the fact that alleles from the locus share a relatively monomorphic sequence, representing an LMD which lies outside of the highly polymrohic minisatellite. The position in the denaturing gradient gel of a restriction fragment containing a minisatellite allele is determined by this LMD since the minisatellite sequence itself can effectively function as a GC-clamp (an HMD) to form stable, partially single-stranded/partially double-stranded molecules. The slight deviations from this isothermal electrophoretic behavior of VNTR alleles are most likely due to sequence differences in the LMD. Separation of sequences which vary only by point mutations in their LMD can be expected to be minimal when broad gradients (*i.e.* from 10–75 % denaturant concentrations) are used.

The presence of internal restriction sites in the minisatellite region of D7S22 generates fragments which migrate to a higher denaturant concentration than the fragments containing the 5′ and/or 3′ ends of the minisatellite. This feature of the separation by DGGE allows for the discrimination of these internal fragments in a single digest and, therefore, to distinguish between VNTR and restriction site polymorphisms. In this respect, hybridization analysis of minisatellite loci resolved by DGGE offers an alternative to neutral agarose gel electrophoresis, and to SBA of multiple restriction enzyme digests of the same DNA sample, for the determination of allelic fragments in *Hae*III digests for this locus [9].

The isothermal electrophoretic behavior in DGGE gels of VNTR alleles may facilitate the identification of alleles in a two-dimensional DNA fingerprint when a minisatellite core probe is used [5] or when a pool of several locus-specific VNTR probes is used. In that case, however, it is necessary to use a restriction enzyme lacking recognition sites within the minisatellite regions, detected by one or more given core- or locus-probes. In addition, the enzyme should preserve the LMD outside the minisatellite sequences detected. For both the loci detected by plambdaG3 and pYNH24 there were no internal *Hin*fI recognition sites in the chromosomes of the 33 individuals we investigated, and isothermal DGGE behavior of alleles of these loci was demonstrated using this enzyme. Preliminary results obtained for other loci suggests that not all minisatellites lack *Hin*fI recognition sites (see also [2]). Therefore, several other enzymes or combinations of enzymes

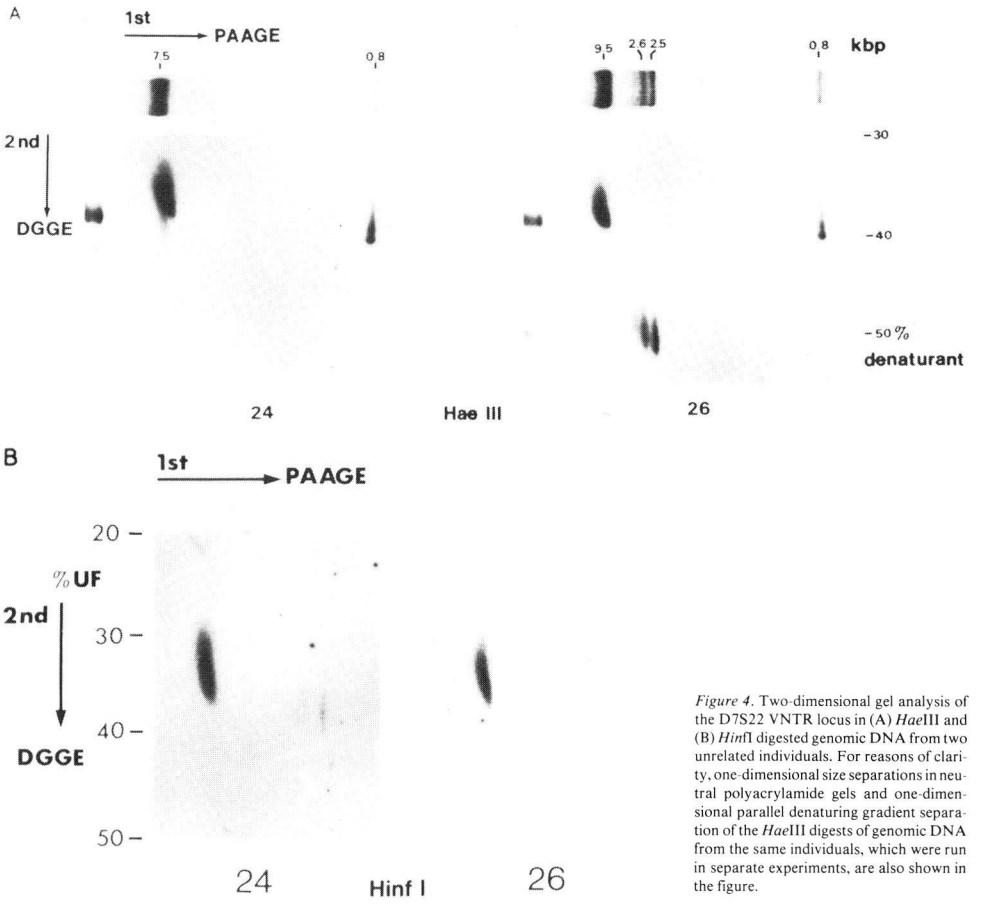


Figure 4. Two-dimensional gel analysis of the D7S22 VNTR locus in (A) *Hae*III and (B) *Hin*fI digested genomic DNA from two unrelated individuals. For reasons of clarity, one-dimensional size separations in neutral polyacrylamide gels and one-dimensional parallel denaturing gradient separation of the *Hae*III digests of genomic DNA from the same individuals, which were run in separate experiments, are also shown in the figure.

have to be tested in order to demonstrate the general applicabilitiy of sorting alleles of a minisatellite locus in denaturing gradient gels.

This is publication No. 1 of Medscand Ingeny. We thank Dr. Leonard Lerman (Department of Biology, MIT, Massachusetts, USA) for his helpful suggestions and discussions and for the use of the melting program. We also thank Dr. A. J. Jeffreys (University of Leicester, Leicester, United Kingdom) for providing the locus-specific minisatellite probe plambdaG3. We thank Dr. Y. Nakamura (Howard Hughes Medical Institute, Utah, USA) for providing pYNH24. Furthermore, we thank Dr. A. Kloosterman (Forensic Laboratories, Rijswijk, The Netherlands) for his kind gifts of genomic DNA samples and Mrs. A. A. Glaudemans and E. J. van der Reijden for help in the preparation of the figures.

Received July 2, 1990

5 References

[1] Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. and White, R., *Science* 1987, *235*, 1616–1622.
[2] Wong, Z., Wilson, V., Patel, I., Povey, S. and Jeffreys, A. J., *Ann. Hum. Genet.* 1987, *51*, 269–288.
[3] Jeffreys, A. J., Wilson, V. and Thein, S. L., *Nature* 1985, *314*, 67–73.
[4] Fischer, S. G. and Lerman, L. S., *Cell* 1979, *16*, 191–200.
[5] Uitterlinden, A. G., Slagboom, P. E., Knook, D. L., and Vijg, J., *Proc. Natl. Acad. Sci. USA* 1989, *86*, 2742–2746.
[6] Wong, Z., Wilson, V., Jeffreys, A. J. and Lay Thein, S., *Nucleic Acids Res.*, 1986, *14*, 4605–4616.
[7] Nakamura, Y., Gilliam, S., O'Connell, P., Leppert, M., Lathrop, G. M., Lalouel, J.-M. and White, R., *Nucleic Acids Res.* 1987, *15*, 10073.
[8] Lerman, L. S. and Silverstein, K., *Methods Enzymol.* 1987, *155*, 501–527.
[9] Waye, J. S. and Fourney, R. M., *Hum. Genet.* 1990, *84*, 223–227.

ADDENDUM 4

One- and two-dimensional DNA-typing as a tool
for human genome monitoring

[Submitted]

ONE- AND TWO-DIMENSIONAL DNA-TYPING AS A POTENTIAL TOOL FOR HUMAN GENOME MONITORING

Bruno Morolli[1*], André G. Uitterlinden[2], Fred J. van Dam[1], Ad D. Tates[1], Jan Vijg[2], and Paul H.M. Lohman[1].

[1] MGC-Department of Radiation Genetics and Chemical Mutagenesis
Sylvius Laboratory
State University of Leiden
Wassenaarseweg 72
2333 AL Leiden
The Netherlands

[2] INGENY B.V.
PO Box 685
2300 AR Leiden
The Netherlands

[*] Present address:
INGENY B.V.
PO Box 685
2300 AR Leiden
The Netherlands

ABSTRACT

One- and two-dimensional DNA-fingerprinting was applied to monitor genome stability of cultivated lymphoid cells. Heterogeneity among cells was assessed for spontaneous variations by analysing sub-clones derived from the lymphoblastoid cell line TK6 and T cell clones from normal human individuals. Using five different core probes, each detecting several hundred micro- and/or minisatellite loci, more than 1200 non-overlapping restriction fragments derived from loci dispersed over the human genome could be analysed. In sub-clones from both the lymphoblastoid cell line TK6 and the normal T cell clones, several spontaneous variations were detected. The results obtained indicate that the system presented here can be used to study the induction of genomic variation in unselected cells after treatment with genotoxic agents.

INTRODUCTION

The human genome is estimated to contain several thousand micro- and minisatellite DNA sequences more or less evenly spread over the different chromosomes (Jeffreys et al., 1985; Ali et al., 1986; Nakamura et al., 1987). Together these sequences roughly comprise about 1 % of the total number of basepairs. It has been demonstrated that a fraction of these sequences (the so-called Variable Number of Tandem Repeats or VNTRs) represent a relatively "unstable" part of the human genome displaying both somatic instability, such as in tumours (Thein et al., 1987; de Jong et al., 1988; Pakkala et al., 1988; Lagoda et al., 1989; Nürnberg et al., 1991), and germ line instability (Jeffreys et al., 1988; Nürnberg et al., 1989). The latter is considered to be responsible for the extensive allelic variation of these sequences in the human population. Micro- and minisatellite sequences can be found in non-coding regions as well as expressed DNA sequences. Examples of recently discovered human microsatellite polymorphisms within genes involve the expansion of particular trinucleotide motifs in the coding regions of the genes for X-linked spinal and bulbar muscular atrophy (Kennedy's disease; La Spada et al., 1991), Fragile-X syndrome (Fu et al., 1991) and Myotonic Muscular Dystrophy (MMD; Fu et al., 1992; Mahadevan et al., 1992).

By using so-called core probes for micro- and minisatellites in Southern blot

hybridization analysis, genomic DNA-fingerprints can be generated (Jeffreys et al., 1985; Ali et al., 1986). Each core probe generally detects DNA sequences derived from hundreds of different loci, only few of which (the largest) can be resolved as restriction fragments of 2-20 kb by one-dimensional DNA fingerprint analysis. By using 2-D DNA fingerprinting, however, all but the largest DNA restriction fragments containing such micro- and minisatellite sequences can be resolved (Uitterlinden et al., 1989 and 1991). By employing these two DNA-typing techniques, the genome can be "scanned" for variations in micro- and minisatellite DNA sequences.

In the current study sub-clones from lymphoblastoid cells, i.e. the established cell line TK6 and T cell clones derived from a normal individual were analysed, to monitor genome stability in cultured cells. A number of spontaneous variations were detected indicating that the method is sensitive for the detection of molecular changes in the genome of unselected cells. The results obtained suggest the possible future application of one- and two-dimensional DNA-typing for the measurement of mutation rates in cloned human cells after treatment with genotoxic agents.

MATERIAL AND METHODS

Cells and culture conditions

T Cells were obtained from normal donors and from the established lymphoblastoid cell line TK6. Cell culture was essentially as described previously (Tates et al., 1991). In brief, lymphocytes, isolated from whole blood with Ficoll density centrifugation, were plated at a density of 2.10^6 cells/ml in RPMI 1640 plus supplements to start initial growth (Tates et al, 1991). After 44 h of incubation at 37 °C and 5-6% CO_2, lymphocytes were plated for sub-cloning in 96-well microtiter plates at 2 cells/well in non-selective medium (200 μl/well; Tates et al., 1991). In addition, 6-thioguanine (6-TG) resistant cells were selected by plating 2.10^4 cells/well in selection medium containing 2.5 μg/ml 6-TG. During culturing each well also contained 1.10^4 irradiated (40 Gy) TK6 lymphoblastoid cells as feeder cells. After 12 to 14 days the wells were scored for colony growth using an inverted microscope. Stock cultures of the established human lymphoblastoid TK6 cell line were grown in RPMI 1640 + 10% calf serum. Cloning of TK6 cells was performed according to the method of Furth et al. (1981). Each well

contained 200 µl RPMI medium + 10% calf serum. After seeding, the cells were allowed to grow in the microtiter plates for 7 days at 37 °C.

DNA isolation and restriction enzyme digestion

For isolation of genomic DNA, cells were harvested, washed with phosphate-buffered saline (PBS) and lysed in 10 mM Tris-HCl, pH 8.2, 400 mM NaCl, 2 mM Na_2EDTA, 1 % SDS and 100 µg/ml proteinase K (Merck) for 20 h at room temperature. Subsequently, the DNA was extracted by adding an equal volume of a phenol-chloroform (1:1) mixture. Thereafter, the water layer containing the DNA was twice extracted with a chloroform-isoamylalcohol (24:1) mixture. Finally, 2 M Na-acetate, pH 5.5 (0.1 volume) was added to the remaining water-phase and the DNA was precipitated at room temperature after adding an equal volume of isopropanol. The precipitated DNA was dissolved in 10 mM Tris-HCl, pH 8.0 to a concentration of 1 µg/µl. Digestion of the DNA with the restriction endonuclease Hae III (Pharmacia) was performed under conditions recommended by the manufacturer.

Electrophoretic separation

In one-dimensional DNA fingerprinting, agarose gel electrophoresis of 7 µg of DNA restriction fragments was performed in a horizontal 1% gel in 1 x TAE (40 mM Tris-HCl pH 7.4 / 20 mM Na-acetate/1 mM Na_2EDTA) at 65 V for 22 h. Marker fragments (500 ng of the "1-kb ladder" [Gibco-BRL]) were included for pattern comparison. Alternatively, one-dimensional DNA fingerprints were prepared by electrophoresis in 1 % gels in 1 x TBE (89 mM Tris-borate, pH 8.0; 89 mM boric acid; 2 mM Na_2EDTA) at 65 V for 44 h. Gels were stained for 15 min in a solution containing ethidium bromide (EtdBr) at 0.1 µg/ml followed by destaining in water for at least 30 min. In two-dimensional DNA fingerprinting (Uitterlinden et al., 1989, 1991), first size-separation of 10 µg of DNA restriction fragments was performed in a 1 mm thick neutral 6% polyacrylamide gel (acrylamide/bisacrylamide, 37:1) for 2.5 h at 200 V in a vertical gel apparatus completely surrounded by vigourously stirred 1 x TAE buffer kept at 50 °C. The separation patterns were visualized by staining the gel with EtdBr (0.1 µg/ml) for 10 min followed by destaining for at least 30 min. For the second dimension separation the 0.54 to 10 kbp region of the size-separated DNA fragments was cut out of

the 1-D gel and applied horizontally on top of a 6% poly-acrylamide gel, containing a 10-75% linear gradient concentration of denaturant (100% denaturant = 7.0 M urea and 40% formamide) parallel to the direction of electrophoresis. Gels were poured by mixing two solutions, containing the desired boundary denaturant concentrations, in a linear gradient maker using a peristaltic pump. For separation in the second dimension a gel apparatus wider than the one used in the first dimension was used (Uitterlinden et al., 1991b). After electrophoresis for 14 h at 200 V and 60 °C, the gel was stained with EtdBr as described above.

Blotting of electrophoretic separation patterns

DNA fragments were capillary transferred from one-dimensional agarose gels to a nylon membrane (ZetaProbe, BioRad; Hybond N plus, Amersham) in 0.4 N NaOH / 0.6 N NaCl for 12 h. After transfer the nylon membrane was rinsed in 2 x SSC (1 x SSC = 150 mM NaCl /15 mM Na-citrate) air-dried, baked for 1 h at 80°C and irradiated with 254 nm UV light for 1 min. For two-dimensional gels, DNA fragments were, after staining and destaining, fragmented by irradiating the gel with 302 nm UV light for 4 min. Before transfer the gel was boiled for 5 min in 1 x TBE in order to denature the DNA and subsequently allowed to cool in the same buffer at room temperature. Transfer of DNA to the nylon membrane was performed by semi-dry electroblotting at 400 mA (7-30 V) between horizontal graphite plates. Electrophoresis was carried out for 1-2 h between 12 Whatman 3MM paper sheets which were soaked in 1 x TBE. After transfer the filter was rinsed in 2 x SSC, air-dried, baked for 1-2 h at 80°C and irradiated with 254 nm UV light for 1 min. Membranes were marked with fluorescent dye to aid in comparing rehybridization patterns obtained from the same membrane.

Probe preparation and labelling

The following core probes were used for preparation of one- and two-dimensional DNA fingerprints of the cloned human T cells (see also Uitterlinden et al., 1991a): minisatellite core probes 33.15, 33.6, and microsatellite core probes $(CAC)_n$, $(TCC)_n$, and $(AGC)_n$. The core probes used in this study were prepared as described previously (Uitterlinden et al., 1989 and 1991a and b). The ligation products (25 ng) were α-^{32}P dCTP-labelled either by the random primed oligolabelling method (BRL) or by self-

priming after boiling for 5 min and re-annealing at 37 °C in the presence of 1 unit Klenow enzyme (Pharmacia), 2 mM dNTP, 50 mM Tris-HCl pH 7.2 and 10 mM $MgCl_2$. Marker fragments (25 ng of the 1 kb ladder or lambda DNA) were also α-^{32}P dCTP-labelled using the random primed oligolabelling method.

Hybridization analysis and pattern interpretation

Pre-hybridization of the filters was performed in 500 mM phosphate buffer pH 7.2, 1 mM Na_2EDTA and 7 % SDS for 2 h at 65°C. After addition of the denatured probe at a concentration of 1.10^6 cpm/ml, the hybridization was carried out for 12 h at 65°C. The filter was washed two times for 5 min at room temperature and two times for 30 min at 65°C in 2.5 x SSC and 0.1 % SDS. Autoradiography was performed for 12-48 h at -80°C using fine intensifying screen and XAR-5 film (Kodak). Rehybridization of membranes was performed after first removing the hybridized radioactive probe by placing the membrane in boiling 0.1 x SSC, 0.1% SDS, letting the solution cool down to room temperature for 30 min and rinsing the membrane in 2.5 x SSC. Membranes were checked for residual radioactivity by inspection with a monitor (Berthold type LB 122; Wildbad, Germany). If more than the background signal (> 30 cps) was measured the stripping procedure was repeated until no residual activity could be measured. Band- and spot-patterns of the one- and two-dimensional DNA typing gels were analysed by eye by at least two persons. For determining overlap in one-dimensional patterns, the 1-kb ladder in the Etbr stained gel separation pattern was compared to the 1 kb ladder hybridization pattern, and the fluorescent marker points on the membrane were used to align the two hybridization patterns. Overlap was concluded if a band occurred at the same position in the gel in the two hybridization patterns. For 2-D patterns, grids (dividing the 2-D gel pattern in 16 equal parts) were made on transparent overlays to ease scoring of spots.

RESULTS

Number of fragments analysed

The frequency of changes as observed by the DNA fingerprinting approach was expressed as the number of variant 1-D bands and 2-D spots detected among clones versus the total number of alleles detected in all clones together. It was assumed that each

band/spot is corresponding to one allele of a micro- or minisatellite locus. Mutation frequencies were then expressed as the number of variant alleles (mutants) per total number of alleles analysed in all clones together. Table 1 lists the number of fragments detected by core probes 33.15, 33.6, $(CAC)_n$, $(TCC)_n$, and $(AGC)_n$ in cells derived from a human individual (GK) after one- and two-dimensional DNA fingerprinting. Examples of the hybridization patterns obtained with the different core probes used are shown in Figure 1. The number of fragments detected in one-dimensional DNA fingerprint analysis varied between 14 and 27, while the two-dimensional analysis allowed between 175 and 390 fragments to be analysed per core probe. The total number of fragments detected by 2-D analysis was over 1600 with 5 core probes (Table 1). When a correction is made for position overlap in the spot pattern of hybridizing restriction fragments, the total number of fragments is lower, i.e. 1260 (see Table 1). More overlap among core probes was detected in one-dimensional analysis (11-54 %) than in two-dimensional analysis (3-22 %), which is most likely due to the higher resolution of two-dimensional gels. When all five core probes were taken into consideration, 49 % ([113-58]/113 x 100 %) of the bands in the 1-D analysis were detected by more than one core probe. This decreases the number of loci screened. For 2-D analysis this percentage was only 24 % ([1660-1260]/1660 x 100 %).

Spontaneous variations

The background mutation frequency observed in one-dimensional analysis was measured by comparing clonally expanded T cells from individual FF and from the established cell line TK6. In total 12 clones were analysed from individual FF and 34 from the TK6 cell line using three different core probes (Table 2). In the T cell clones from individual FF one mutant 33.6 allele of 11.5 kb was detected in a single clone, among a total of 440 alleles analysed. In several clones of the TK6 cell line identical mutant alleles were observed. Rather than representing independent mutation events, these are likely to be derived from subpopulations of cells. Three subpopulations could be discerned (Figure 2). The normal population constituted 76 % of the total number of cells while 15 % lacked a 33.15 allele of 2.9 kb and 9 % had an extra mutant 4.7 kb $(AGC)_n$ allele. Thus, in total 2 mutations had occurred in 1870 alleles analysed.

TABLE 1

A. Percentage of overlapping bands detected in 1-D analysis of Hae III digested genomic DNA from individual GK using different core probes. In brackets the absolute number of overlapping bands detected by two core probes is given.

PROBES	$(AGC)_n$	$(TCC)_n$	$(CAC)_n$	33.6	33.15	TOTAL # OF BANDS
$(AGC)_n$	100	29 (4)	21 (3)	43 (6)	21 (3)	14
$(TCC)_n$	20 (4)	100	25 (5)	45 (9)	25 (5)	20
$(CAC)_n$	11 (3)	19 (5)	100	41 (11)	52 (14)	27
33.6	23 (6)	35 (9)	42 (11)	100	39 (10)	26
33.15	12 (3)	19 (5)	54 (14)	39 (10)	100	26
TOTAL						58[A]

[A] Corrected for overlap among the different core probes

B. Percentage of overlapping spots detected in 2-D analysis of Hae III digested genomic DNA from individual GK using different core probes. In brackets the absolute number of overlapping bands detected by two core probes is given.

PROBES	$(AGC)_n$	$(TCC)_n$	$(CAC)_n$	33.6	33.15	TOTAL # OF SPOTS
$(AGC)_n$	100	-[A]	-[A]	3 (9)	-[A]	175
$(TCC)_n$	-[A]	100	9 (36)	21 (68)	13 (50)	375
$(CAC)_n$	-[A]	10 (36)	100	15 (48)	11 (42)	395
33.6	5 (9)	18 (68)	12 (48)	100	19 (73)	325
33.15	-[A]	13 (50)	11 (42)	22 (73)	100	390
TOTAL						1260[B]

[A] Not determined

[B] Corrected for overlap among spots for the different core probes

TABLE 2 Background mutation frequency of one-dimensional DNA fingerprints[A]

PROBE	FF	TK6
33.15	0/290	4/850[B]
33.6	1/260	0/884
$(AGC)_n$	-	3/612[B]
TOTAL	1/550	2/2346[B]
TOTAL[C]	1/440 (0.23%)	2/1870 (0.11%)

[A] The mutation frequency is estimated from the number of mutant alleles detected in 1-D analysis of separate, clonally expanded T cells from individual FF (12 clones) and established lymphoblastoid cell line TK6 (34 clones from two independent cloning experiments).

[B] five identical 33.15 mutant fragments were detected indicating the presence of two subpopulations. Three identical mutant (AGC)n fragments were detected which were not overlapping with the subpopulations defined by the 33.15 mutant allele. This indicates the presence of three subpopulations of clones A (the 33.15 mutant allele) constituting 15% of the number of clones analysed, B (the (AGC)n mutant allele, constituting 9% of the total number of clones analysed, and the "wild type", constituting 76% of the total number of clones analysed

[C] Total number of non-overlapping fragments

DISCUSSION

In this present paper we describe an approach to measure DNA mutations by the direct and simultaneous analysis of hundreds of micro- and minisatellite loci by using core probes in one- and two-dimensional DNA-typing of clonally expanded cells. To evaluate the usefulness of this approach a strategy was chosen whereby lymphocyte populations were cloned and subsequently compared.

The results indicate that by a combination of one- and two-dimensional DNA-typing, a considerable number of alleles can be analyzed in a single DNA sample. Using only 5 core probes more than 1200 restriction fragments are analysed from an estimated number of 900 loci (assuming at least 25% variant fragments among the total number of fragments detected; Uitterlinden et al., 1989, 1991a). This number can be increased even further by using additional core probes (Uitterlinden et al., 1991a), depending on the overlap among different core probes detecting different loci. We observed less overlap

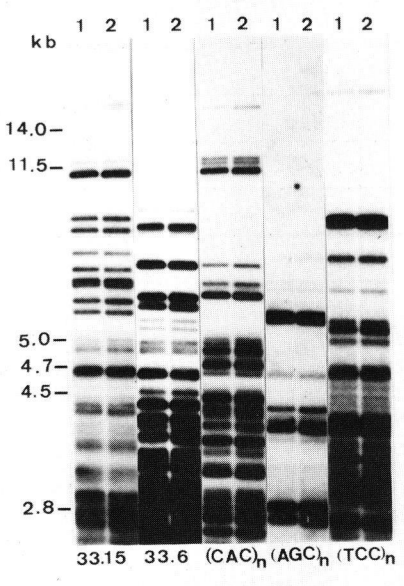
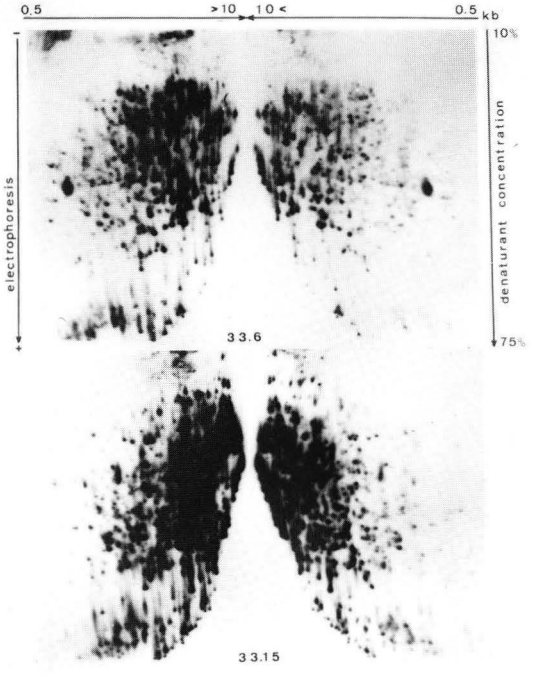


Figure 1 *(A) One-dimensional DNA fingerprints of human T cell clones derived from individual GK obtained using core probes 33.15, 33.6, $(CAC)_n$, $(AGC)_n$ and $(TCC)_n$ obtained by rehybridization of a single membrane containing the electrophoretic separation patterns.*

in 2-D analysis than in 1-D analysis, which is most likely due to the higher resolution of the 2-D system. Overlap can be the result of fortuitous co-migration of 2 or more different restriction fragments, but in view of the high resolving power of this 2-D electrophoretic system it is more likely to be the result of cross-hybridization of core probes to alleles of a single micro- and/or minisatellite locus.

The possibility to detect random genomic variations by this approach is demonstrated by the detection of three subpopulations of cells in the TK6 lymphoblastoid cell-line, characterized by particular mutant micro- or minisatellite fragments. The presence of these mutations can be explained by the relatively high spontaneous mutation frequency of some minisatellite loci, especially the ones which have long alleles



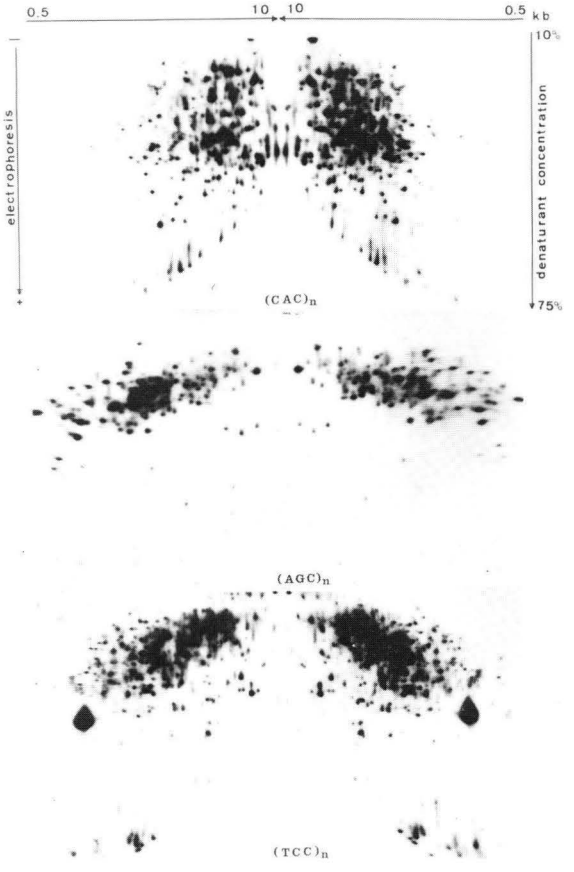


Figure 1 (continued) *(B) Two-dimensional DNA fingerprints of human T cell clones derived from individual GK using core probes 33.15, 33.6, $(CAC)_n$, and $(TCC)_n$ obtained by rehybridization of a single membrane containing the electrophoretic separation pattern.*

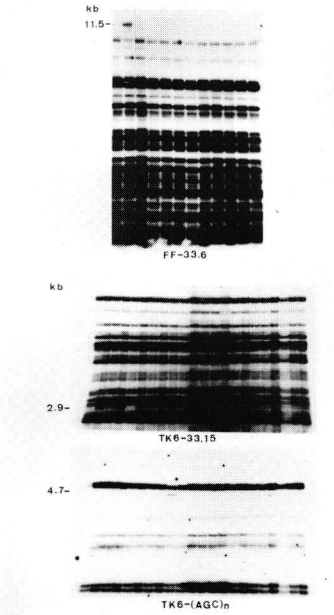


Figure 2 *Spontaneous mutant alleles found by one-dimensional DNA fingerprint analysis of T cell clones (FF) and TK6 sub clones. A represents a mutant 33.6 allele found in a T cell subclone from individual FF and B shows mutant 33.15 and $(AGC)_n$ alleles found in TK6 sub clones. The two latter mutant alleles were detected in both TAE and TBE agarose gels (see also Table 2).*

(Jeffreys et al., 1988). The micro- and minisatellite type of DNA sequences detected by core probes in fact comprises an array of different loci with respect to variability and spontaneous mutation frequencies (Nakamura et al., 1987; Jeffreys et al., 1985, 1988). This might be the result of differences in sequence composition of the repeat unit of which a locus is composed, the length of alleles of the locus, the location in coding vs. non-coding regions in the genome, and the location on the chromosomes (e.g. sub-telomeric vs. interstitial). Some loci (the non-VNTR loci) are known to be almost completely "silent" and have a mutation frequency very close or identical to that of other DNA sequences (i.e. 1.10^{-6}; Jeffreys et al., 1988; Nürnberg et al., 1989). Other loci (the

VNTR loci) are known to be mutable within a very broad range of mutation frequencies but sometimes to an extreme extent. For example, the human VNTR locus detected by probe pMS 1 (D1S8) is known from pedigree studies to have a (meiotic) mutation frequency of 1/20, meaning that 5% of the gametes of an individual have a mutated allele at this locus (Wong et al., 1987; Jeffreys et al., 1988).

In comparison to other tests for genotoxicity such as the HPRT test (Tates et al., 1991), the approach described here has the particular advantages that genomic DNA is directly analysed at a large number of loci for sequence variation in unselected cells. Assuming the spontaneous mutation frequency of the micro- and minisatellite loci to be on average much higher (i.e. 1.10^{-4}) than that of e.g. the HPRT gene, concomitantly less loci would have to be screened to detect induction of sequence changes above this level. In an HPRT test 10 million T cell clones must be screened to detect 10 mutants (Tates et al., 1991) equaling the analysis of 10 million times 1 kb of coding sequence ($=1.10^{10}$ bp) for mutations. In the (1-D and 2-D) DNA typing test using five core probes a single clone is analysed for 1,200 fragments of, on average, 1,000 bp equalling approximately $1.2.10^{6}$ bp analysed (1 core probe equals approximately 250,000 bp analysed). This would mean that the sensitivity of the 1-D and 2-D DNA typing method has to increase to reach a sensitivity that equals that of the HPRT test, assuming equal spontaneous mutation frequencies of the HPRT locus and the micro- and minisatellite loci for mutation. It can be argued (see above) that the latter is not the case so if we assume the average spontaneous mutation frequency of micro- and minisatellites to be 1.10^{-4}, i.e. 100 times greater than that of the HPRT gene, 100 times less bp would have to be analysed, i.e. 1.10^{7} bp, to detect induction above background levels of spontaneous mutations. Compared with the 1,200,000 bp obtained in this study for the analysis of a single clone with five core probes, this can be accomplished by analysing a single clone using approximately 25 core probes (assuming each core probe detects approximately 250 non-overlapping fragments of on average 1,000 bp), or by analysing 10 clones with five core probes.

The spectrum of mutations that can be detected by the 2-D DNA fingerprinting method includes size changes, deletions but also 40 % of all possible point mutations in the flanking sequence of the micro- and/or minisatellite region (Uitterlinden and Vijg, 1991). In this respect, the 2-D DNA typing approach as described here represents a

powerful tool to perform genome scanning studies of somatic DNA sequence variation. It can be applied in mutation studies as is proposed here or in studies of genome variation during aging (Slagboom et al., 1991) and tumorigenesis (Hovig et al., in press).

ACKNOWLEDGMENTS

We thank P.E. Slagboom (TNO-Institute of Ageing and Vascular Research, Leiden, The Netherlands) for helpful discussions. This study was supported by grant# D437638 from ILSI (International Life Sciences Institute) Risk Science Institute to PHML and BM, and by research funds of INGENY B.V..

REFERENCES

Albertini, R.J., Castle, K.L., and Borcherding, W.R. (1982) T cell cloning to detect the mutant 6-thioguanine resistant lymphocytes present in human peripheral blood. Proc. Natl. Acad. Sci. USA 79, 6617-6621.

Ali, S., Müller, C.R., and Epplen, J.T. (1986) DNA fingerprinting by oligonucleotide probes specific for simple repeats. Hum.Genet. 74, 239-243.

Armour, J.A.L., Povey, S., Jeremiah, S., and Jeffreys, A.J. (1990) Systematic cloning of human minisatellites from ordered array charomid libraries. Genomics 8, 501-512.

Fischer S.G. and Lerman L.S. (1979), Two-dimensional electrophoretic separation of restriction enzyme fragments of DNA. Methods Enzymol. 68, 183-191.

Fu, Y.-H., Kuhl, D.P.A., Pizutti, A., Pieretti M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M.H., Holden, J.J.A., Fenwick, R.G., Warren, S.T., Oostra, B.A., Nelson, D.L., and Caskey, C.T. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67, 1047-1058.

Fu, Y.-H., Pizutti, A., Fenwick, R.G., King, J., Rajnarayan, S., Dunne, P.W., Dubbel, J., Nasser, G.A., Ashizawa, T., de Jong, P.J., Wieringa, B., Korneluk, R., Perryman, M.B., Epstein, H.F., and Caskey, C.T. (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 255, 1256-1258.

Furth, E.M., Thilly, W.G., Penham, B.A., Liber, H.L. and Rand, W.M. (1981), Quantitative assay for mutation in diploid human lymphoblasts using microtiterplates. Anal. Biochem., 110, 1-8.

Hovig, E., Mullaart, E., Borresen, A.-L., Uitterlinden, A.G., and Vijg, J. (1993) Genome scanning of human breast carcinoma's using micro- and minisatellite core probes. Genomics, in press.

Jeffreys A.J., V. Wilson and S.L. Thein (1985), Individual-specific 'fingerprints' of human DNA. Nature 314, 67-73.

Jeffreys, A.J., Royle, N.J., Wilson, V., and Wong, Z. (1988) Spontaneous mutation rates to new length alleles at tandem repetitive hypervariable loci in human DNA. Nature 332, 278-281.

de Jong, D., Voetdijk, B.M.H., Kluin-Nelemans, J.C., van Ommen, G.J.B., and Kluin, Ph.M. (1988) Somatic changes in B-lymphoproliferative disorders (B-LPD) detected by DNA-fingerprinting. Br. J. Cancer 58, 773-775.

Kuick, R.D., Neel, J.V., Strahler, J.R., Chu, E.H.Y., Bargal, R., Fox, D.A., and Hanash, S.M. (1992) Similarity of spontaneous germinal and in vitro somatic cell mutation rates in humans: implications for carcinogenesis and for the role of exogeneous factors in "spontaneous" germinal mutagenesis. Proc. Natl. Acad. Sci. USA 89, 7036-7040.

Lagoda, P.J.L., Seitz, G., Epplen, J.T., and Issinger, O.-G. (1989) Increased detectability of somatic changes in the DNA from human tumours after probing with "synthetic" and "genome-derived" hypervariable multilocus probes. Hum. Genet. 84, 35-40.

Lohman, P.H.M., Vijg, J., Uitterlinden, A.G., Slagboom, P., Gossen, J.A., and Berends, F. (1987) DNA methods for detecting and analyzing mutations in vivo. Mutation Res. 181, 227-234.

Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., Leblond, S., Earle-MacDonald, J., de Jong, P.J., Wieringa, B., Korneluk, R.G. (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science 255, 1253-1255.

Maniatis T., E.F. Fritsch and J. Sanbrook (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York

Mohrenweiser, H.W., and Jones, I.M. (1990) Review of the molecular characteristics of gene mutations of the germline and somatic cells of the human. Mut. Res. 231, 87-108.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987) Variable Number of Tandem Repeat (VNTR) markers for human gene mapping. Science 235, 1616-1622.

Neel, J.V., Rosenblum, B.B., Sing, C.F., Skolnick, M.M., Hanash, S.M., and Steinberg, S. (1984) Adapting two-dimensional gel electrophoresis to the study of human germ-line mutation rates, in: J.E. Celis and R. Bravo (Eds.), Two-dimensional Gel Electrophoresis of Proteins, Academic Press, New York, pp. 259-306.

Nürnberg P., L. Roewer, H. Neitzel, K. Sperling, A. Pöppert, J. Hundrieser, H. Pöche, C. Epplen, H. Zischler and J.T. Epplen (1989) DNA fingerprinting with the oligonucleotide probe $(CAC)_5/(GTG)_5$: somatic stability and germline mutation. Hum. Genet. 84, 75-78.

Nürnberg, P., Zischler, H., Fuhrman, E., Thiel, G., Losanova, T., Kinzel, D., Nisch, G., Witkowski, R., and Epplen, J.T. (1991) Coamplification of simple repetitive DNA fingerprint fragments and the EGFR gene in human gliomas. Genes, Chrom. Canc. 3, 79-88.

Pakkala, S., Helminen, P., Saarinen, U.M., Alitalo, R., and Peltonen, L. (1988) Differences in DNA-fingerprints between remission and relapse in childhood acute lymphoblastic leukemia. Leukemia Res. 12, 757-762.

Slagboom, P.E., Mullaart, E., Droog, S., and Vijg, J. (1991) Somatic mutations and cellular aging: two-dimensional DNA typing of rat fibroblast clones. Mut. Res. 256, 311-321.

la Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E., and Fischbeck, K.H. (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352, 77-79.

Tates, A.D., van Dam, F.J., van Mossel, H., Schoemaker, H., Thijssen, J.C.P., Woldring, V.M., Zwinderman, A.H., and Natarajan, A.T. (1991) Use of the clonal assay for the measurement of frequencies of HPRT mutants in T-lymphocytes from five control populations. Mut. Res. 253, 199-213.

Thein, S.L., Jeffreys, A.J., Gooi, H.C., Cotter, F., Flint, J., O'Connor, N.T.J., Weatherall, D.J., and Wainscoat, J.S. (1987) Detection of somatic changes in human cancer DNA by DNA fingerprint analysis. Br. J. Cancer 55, 353-356.

Uitterlinden A.G., E. Slagboom, D.L. Knook and J. Vijg (1989), Two-dimensional DNA fingerprinting of human individuals. Proc. Natl. Acad. Sci. 86, 2742-2746.

Uitterlinden A.G., Slagboom, P.E., Mullaart, E., Meulenbelt, I., and Vijg, J. (1991a) Genome scanning by two-dimensional DNA typing: the use of repetitive DNA sequences for rapid mapping of genetic traits. Electrophoresis 12, 119-134.

Uitterlinden, A.G., Mullaart, E.M., Morolli, B., and Vijg, J. (1991b) Genome scanning of higher eukaryotes by two-dimensional DNA typing using micro- and minisatellite core probes. Methods: A companion to methods in enzymology 3, 83-90.

Vijg, J., and Uitterlinden, A.G. (1987) A search for DNA alterations in the aging mammalian genome: an experimental strategy. Mech. Ageing Dev. 41, 47-63.

Wong, Z., Wilson, V., Patel, I., Povey, S., and Jeffreys, A.J. (1987) Characterization of a panel of highly variable minisatellites cloned from human DNA. Ann. Hum. Genet. 51, 269-288.

SUMMARY

The phenotype of an organism is ultimately encoded in the nucleus of each cell by the sequence of basepairs in the set of DNA molecules that constitute the genome of each species. This information infrequently undergoes changes ("mutations") due to physical or chemical damage which can distort the metabolism of the cell. When occurring in somatic cells mutations can give rise to diseases such as cancer and other age-related deteriorations. When occurring in germ cells they can lead to hereditary disorders. At a larger time scale mutations are fundamental to the evolution of species. To gain a better insight in these processes it is important to be able to measure variations in the DNA sequence. The comparison of DNA sequences can also reveal the underlying molecular causes of genetic disorders including cancer.

This thesis describes a method, two-dimensional (2-D) DNA typing, to analyse the genome of higher organisms including humans for sequence variation at many sites simultaneously. Untill recently, the existing techniques could only measure such variation at very few sites in the genome simultaneously and such genome scanning studies were therefore time-consuming. 2-D DNA Typing allows to scan the genome at hundreds of sites in parallel and requires only little manipulation of the DNA. This facilitates particular studies such as genetic linkage analysis with DNA markers and analysing tumors for genome alterations.

2-D DNA Typing is based on separation of short DNA fragments in a two-dimensional electrophoretic system. The first dimension separation is by size of fragments while the second dimension encompasses denaturing gradient gel electrophoresis (DGGE) in a polyacrylamide gel containing a linear gradient of denaturants. DGGE is based on the decreased electrophoretic mobility of partially single-stranded DNA fragments. Such molecules are formed during electrophoresis in DGGE gels due to the presence of denaturants which cause the molecule to "melt". Melting of DNA molecules is dictated by the basepair sequence and allows to discriminate molecules with subtle sequence differences but of a similar size. **Addendum 1** describes DGGE and its application to detect a point mutation in the human cHa-ras1 proto-oncogene. Optimal separation of sequence variants in DGGE is only achieved if the molecules become *partially* single stranded

while maintaining regions where the molecule is double stranded. Such regions are ususally GC-rich and function as a "clamp".

Addendum 2 describes the 2-D DNA typing of human genomic DNA using probes for repetitive elements that occur dispersed through the genome. This allows to analyse the genome at hundreds of sites simultaneously. To achieve this, human genomic DNA is digested with the restriction enzyme *Hae III* and separated by 2-D electrophoresis. After transferring the separation pattern to a nylon membrane, hybridization analysis using micro- or minisatellite core probes is performed. These probes detect repetitive elements thousands of which occcur in the genomes of many species and which also display a high degree of variability in the population (polymorphism). Using these probes spot patterns can be generated consisting of up to 600 spots depending on which probe is applied. A considerable part of the spots (25%) is variant among individuals and in a pedigree study the transmission of 150 spot variants could be followed. The observed Mendelian transmission of spot variants indicates that these spots are derived from independently segregating loci and thus that 2-D DNA typing can be applied in genetic studies.

In general, micro- and minisatellites are characterized by their GC-richness. This feature makes these sequences to function as a clamp in the restriction fragments in which they are present en thereby ensure an optimal melting behaviour in the second dimension separation of 2-D DNA typing. This is demonstrated in **Addendum 3** by the DGGE and 2-D analysis of two highly polymorphic variable number of tandem repeat (VNTR) loci in human DNA. Alleles with these loci differ greatly in length but upon separation by DGGE, all alleles migrate to a single, locus-specific position in the DGGE gel. This is corroborated by computersimulation of the melting behaviour of a VNTR allele which shows that the minisatellite repeat unit sequences function as a clamp. This means that after 2-D analysis two alleles of the same locus but which differ greatly in length will migrate to the same position in the denaturing gradient: the isotherm.

Many different core probes have been described which can each detect many different GC-rich micro- and/or minisatellites. Such probes are therefore very suitable to screen hundreds of VNTR loci in many different species by 2-D DNA typing. **Addendum 4** describes a potential application of 2-D DNA typing in the mutational analysis of human cells. 2-D DNA typing of clonally expanded cells exposed to genotoxic agents can

detect possible alterations in the genome caused by the induction of damage in the DNA. The analysis of unselected cells at many different loci in the genome is an important advantage in this respect in comparison with other mutation analysis systems. The sequences detected by the core probes comprise an array of loci with different spontaneous mutation frequencies. This allows a better assessment of the genetic damage after exposure to genotoxic agents.

The most important applications of 2-D DNA typing using core probes are in the field of "genome-scanning: the detection of unknown changes and the measurement of many polymorphic sites in the genome. Examples include mutational analysis and the search for genome alterations which might occur during tumorigenesis and aging of cells. The measurement of many polymorphic sites is key in genetic linkage analysis. Here the location of disease genes is determined by the analysis of genetic markers in pedigrees segregating for a genetic disease. 2-D DNA typing is also very suitable for the analysis of multifactorial traits, including some of the most common human diseases such as heart disease. In concert with the development of automated gel electrophoresis equipment and image analysis software, the high "genome-scanning" efficiency of 2-D DNA typing is likely to result in the rapid isolation of markers for genetic traits including diseases with genetic components.

SAMENVATTING

De volgorde van baseparen in het geheel van DNA moleculen in de kern van iedere cel (het genoom) is de basis van de verschijningsvorm van ieder organisme. Door chemische en fysische interacties kunnen er soms veranderingen (mutaties) optreden in deze volgorde waardoor een cel anders kan gaan functioneren. Als mutaties in de lichaamscellen optreden kan dit aanleiding zijn tot ziektes zoals kanker en wellicht ook tot veroudering. Treden dergelijke veranderingen op in de voortplantingsscellen, dan kunnen er erfelijke afwijkingen ontstaan. Op de lange termijn ligt het optreden van mutaties zelfs ten grondslag aan de evolutie van soorten. Om een beter inzicht in dergelijke processen te kunnen krijgen is het dus van belang veranderingen in de basepaarvolgorde van het DNA te kunnen waarnemen. Tevens kan het vergelijken van de basepaarvolgorde tussen individuen een beter inzicht verschaffen in de onderliggende oorzaken van erfelijke ziekten, inclusief kanker.

Dit proefschrift beschrijft een methode, twee-dimensionale (2-D) DNA typering, om het genoom van hogere organismen inclusief de mens op vele plaatsen tegelijkertijd te analyseren op het voorkomen van individuele variatie in de basepaarvolgorde. Voorheen kon met de bestaande technieken DNA sequentievariatie slechts op een beperkt aantal plaatsen in het genoom worden vastgesteld. Studies waarbij het genoom moet worden afgezocht op het voorkomen van dergelijke variaties (genoom-"scanning" studies) zijn dan ook tijdrovend. Met 2-D DNA typering kan het genoom op honderden plaatsen tegelijkertijd worden onderzocht op variatie, na slechts een klein aantal bewerkingen aan het DNA. Hierdoor kunnen een aantal studies sterk versneld worden zoals het aantonen van genetische koppeling met behulp van DNA merkers en het analyseren van tumoren op het voorkomen van veranderingen in het genoom.

2-D DNA Typering is gebaseerd op het scheiden van korte DNA fragmenten in een twee-dimensionaal electroforese systeem. De eerste dimensie bestaat uit een scheiding op grootte van DNA fragmenten. De tweede dimensie bestaat uit Denaturerende Gradiënt Gel Electroforese (DGGE): electroforese in een polyacrylamide gel waarin zich een lineaire gradiënt bevindt van denaturantia. DGGE is gebaseerd op de sterk verminderde electroforetische mobiliteit van een DNA fragment zodra een gedeelte van het molecuul

enkelstrengs is geworden. Dit gebeurt tijdens de electroforese van het fragment in de DGGE gel ten gevolge van de denaturantia die het DNA doen "uitsmelten". Het uitsmelten van DNA is sterk afhankelijk van de basepaarvolgorde van de fragmenten en DGGE is in staat zeer kleine verschillen aan te tonen tussen fragmenten van dezelfde grootte maar van een verschillende basepaarcompositie. **Addendum 1** beschrijft DGGE en haar toepassing om een puntmutatie in het menselijk cHa-ras1 proto-oncogen aan te tonen. Hieruit blijkt onder andere dat optimale scheiding met behulp van DGGE slechts wordt bereikt indien het DNA molecuul *gedeeltelijk* enkelstrengs wordt en de twee strengen plaatselijk nog aan elkaar zijn verbonden. Zo'n plaats is doorgaans GC-rijk en fungeert door de sterke baseparing van GC's als een soort klem ("clamp").

In **Addendum 2** wordt de 2-D DNA typering beschreven van menselijk DNA. Het gebruik van probes voor repeterende elementen die verspreid door het genoom voorkomen maakt het mogelijk honderden plaatsen tegelijkertijd te analyseren. Hiertoe wordt het 2-D scheidingspatroon van menselijk DNA geknipt met het restrictie-enzym *Hae III* wordt overgedragen op een nylon membraan en onderworpen aan hybridisatieanalyse met probes voor micro- of minisatellieten. Deze repeterende elementen komen bij duizenden verspreid door het genoom van zeer veel species voor en vertonen een hoge mate van variabiliteit (polymorfisme) in de populatie. In 2-D DNA typering resulteert het gebruik van deze probes in een spotpatroon dat per individu uit ongeveer 600 spots bestaat afhankelijk van welke probe gebruikt wordt. Een aanzienlijk deel (25%) van de spots is variant tussen individuen en in een familiestudie werd de transmissie van 150 spotvarianten van ouders op kind aangetoond. De waargenomen Mendeliaanse transmissie van de spotvarianten duidt erop dat de spots afkomstig zijn van onafhankelijk overervende loci en dus dat 2-D DNA typering gebruikt kan worden om genetische studies te doen.

Micro- en minisatellieten worden in het algemeen gekenmerkt door een bijzonder hoog gehalte aan GC-baseparen. Deze karakteristiek zorgt ervoor dat juist de micro- en minisatellieten als "clamp" fungeren in de betreffende restrictiefragmenten en deze dus een optimaal uitsmeltgedrag vertonen in de tweede dimensie scheiding van de 2-D DNA typering. Dit wordt in **Addendum 3** gedemonstreerd aan de hand van DGGE en 2-D analyse van twee zeer polymorfe Variable Number of Tandem Repeat (VNTR) loci in menselijk DNA. Allelen van deze loci vertonen grote verschillen in lengte door het VNTR polymorfisme. Na DGGE analyse echter komen alle allelen op een voor het locus

specifieke plaats in de denaturerende gradiënt terecht. Uit computersimulatie van het uitsmeltgedrag van een VNTR allel blijkt dat dit te wijten is aan de GC-rijke repeterende eenheden van de minisatelliet die als "clamp" fungeren. Na 2-D analyse zullen de twee allelen van een locus van een individu zich op eenzelfde lijn in een denaturerende gradiënt bevinden: de isotherm.

Er zijn verschillende sets van GC-rijke micro- en minisatellieten beschreven die ieder met een andere core probe kunnen worden aangetoond. Dergelijke probes zijn dus uitstekend te gebruiken in 2-D DNA typering om tegelijkertijd honderden VNTR loci in het genoom van een groot aantal species te analyseren. In **Addendum 4** wordt een potentiële toepassing beschreven van 2-D DNA typering met een aantal van dergelijke core probes in mutatie-analyse van menselijke cellen. Na klonale uitgroei van gekweekte cellen die aan genotoxische agentia zijn blootgesteld kan 2-D DNA typering mogelijke veranderingen in het genoom aantonen die zijn ontstaan na inductie van schade in het DNA. Een belangrijk voordeel van een dergelijke mutatie-analyse is dat, in tegenstelling tot andere mutatie-analysesystemen, niet-geselecteerde cellen worden geanalyseerd op veranderingen op een groot aantal loci in het genoom. De geanalyseerde loci worden gekenmerkt door een spectrum aan spontane mutatiefrequenties zodat een beter inzicht kan worden verkregen in de genetische schade na blootstelling aan genotoxische agentia.

De belangrijkste toepassingen van 2-D DNA typering met core probes liggen op het gebied van genoom-"scanning": het opsporen van onbekende veranderingen en het doormeten van veel polymorfismen in het genoom. Voorbeelden hiervan zijn de eerder genoemde mutatie-analyse en het opsporen van veranderingen zoals die tijdens tumorgenese en veroudering van cellen kunnen optreden. Het doormeten van veel polymorfismen is van groot belang in genetische koppelingstudies om ziektegene te isoleren. Door de analyse van genetische merkers in stambomen waarin een genetische ziekte optreedt wordt de plaats bepaald waar zich het ziektegen moet bevinden. 2-D DNA typering is tevens bij uitstek geschikt voor de analyse van multifactoriële ziekten waaronder enkele van de meest voorkomende ziektes bij de mens zoals hart- en vaatziekten. Tezamen met de toepassing van geautomatiseerde gelelectroforeseapparatuur en beeldanalyse-software, zal de grote genoom-"scanning" efficiëntie van 2-D DNA typering leiden tot een versnelling in het isoleren van merkers voor genetische eigenschappen waaronder ziektes met erfelijke componenten.

ACKNOWLEDGMENTS

I am especially indebted to Dr. Stuart Fischer for introducing me into denaturing gradient gel electrophoresis and 2-dimensional DNA electrophoresis, his kind hospitality in the laboratory of Actagen Inc./Lifecodes Corp. (Elmsford/Valhalla, New York, USA) from June 1984 till July 1985 and for his stimulating discussions on science, gels and autoradiographs. I am grateful to Prof. Dr. D.L. Knook for his enthusiastic support during my stay at the TNO-IvEG and my promotor Prof. Dr. Ir. P.H.M. Lohman for the continuous and enthusiastic support during the course of this work. I would like to thank all my colleagues for their help and support at the laboratories where the research described in this thesis was performed. I owe many thanks to the people at Actagen Inc./Lifecodes Corp. for their part in making my year in New York a happy one, an affordable one and a productive one. A special word of thanks is for Eline Slagboom with whom part of the experiments were performed with great perseverance and enthusiasm at TNO-IvEG. I thank Gerjan Trommelen for helpful discussions. I thank my colleagues Erik Mullaart for helpful discussions and expert help, and Jan Gossen and Wiljo de Leeuw for their help and support during my work at TNO-IvEG and at INGENY B.V. I thank Bruno Morolli for his help in the collaborative work at the TNO-MBL and at the Department of Radiation Genetics and Chemical Mutagenesis of the Leiden University, and for help and support during my work at INGENY B.V. I thank Ingrid Meulenbelt for her enthusiastic help in work done at the TNO-IvEG and at INGENY B.V.

Finally, I especially thank my co-promotor Dr. Jan Vijg for his unconventional and inspiring coaching, editing of my texts and thoughts, and persistence in getting things done.

CURRICULUM VITAE

De auteur van dit proefschrift werd op 29 oktober 1958 geboren te Maassluis. Na het behalen van het VWO diploma aan de Christelijke Scholengemeenschap "Blaise Pascal" te Spijkenisse in 1977 werd in datzelfde jaar begonnen met de studie Biologie aan de Rijksuniversiteit Leiden. In 1981 werd hiervoor het kandidaatsexamen afgelegd en in 1984 het doctoraalexamen met als hoofdvak Biologie en als tweede hoofdvak Scheikunde.

In mei 1984 werd een dienstverband aangegaan met het TNO Instituut voor Experimentele Gerontologie (TNO-IvEG) te Rijswijk. Onder leiding van Dr. J. Vijg werd aangevangen met het in dit proefschrift beschreven onderzoek dat van juni 1984 tot en met juli 1985 uitgevoerd werd bij het bedrijf Actagen Inc./Lifecodes Corp. (Elmsford/Valhalla, New York, USA), onder begeleiding van Dr. S.G. Fischer. Van 1985 tot en met 1990 werd bij de afdeling Moleculaire Biologie van het TNO-IvEG het onderzoek voortgezet. In oktober 1990 volgde een aanstelling als hoofd van de afdeling genetische diagnostiek bij het bedrijf INGENY B.V. te Leiden.

LIST OF PUBLICATIONS

REFEREED JOURNALS

Published:

1. Uitterlinden, A.G., Vijg, J., Giphart, M.J., and Knook, D.L. (1985) Variation in restriction fragment length and methylation pattern of rat MHC class I genes. Exp. Clin. Immunogenet. 2, 215-222.

2. Vijg, J., and Uitterlinden, A.G. (1987) A search for DNA alterations in the aging mammalian genome: an experimental strategy. Mech. Ageing Dev. 41, 47-63.

3. Lohman, P.H.M., Vijg, J., Uitterlinden, A.G., Slagboom, P., Gossen, J.A., and Berends, F. (1987) DNA methods for detecting and analyzing mutations in vivo. Mutation Res. 181, 227-234.

4. Uitterlinden, A.G., P.E. Slagboom, D.L. Knook, and J. Vijg. (1989) Two-dimensional DNA fingerprinting of human individuals. Proc. Natl. Acad. Sci. USA 86, 2742-2746.

5. Uitterlinden, A.G., P.E. Slagboom, T.E. Johnson, and J. Vijg. (1989) The Caenorhabditis elegans genome contains monomorphic minisatellites and simple sequences. Nucleic Acids Res. 17, 9527-9530.

6. Uitterlinden, A.G., and Vijg, J. (1990) Denaturing gradient gel electrophoretic analysis of the human cHa-ras 1 proto-oncogene. Appl. Theor. Electrophoresis 1, 175-179.

7. Uitterlinden, A.G., and Vijg, J. (1991) Denaturing gradient gel electrophoretic analysis of human minisatellites. Electrophoresis 12, 12-16.

8. Slagboom, P.E., Uitterlinden, A.G., and Vijg, J. (1991) Methylation status of cKi-ras and MHC genes in rat pituitary glands during aging and tumorigenesis. Aging 3, 141-146.

9. Top, B., Uitterlinden, A.G., van der Zee, A., Kastelein, J.J.P., Gevers Leuven, J.A., Havekes, L.M., and Frants, R.R. (1992) Absence of mutations in the promotor region of the low density lipoprotein receptor gene in a large number of familial hypercholesterolemia patients as revealed by denaturing gradient gel electrophoresis. Hum. Genet. 89, 561-565.

10. Van Belkum, A.J., Ramesar, J., Trommelen G.J.J.M. and Uitterlinden, A.G. (1992) Mini- and microsatellites in the genome of rodent malaria parasites. Gene 118, 81-86.

In press:

1. Hovig, E., Mullaart, E., Borresen, A.-L., Uitterlinden, A.G., and Vijg, J. Genome scanning of human breast carcinomas using micro- and minisatellite core probes. (Genomics)

2. Meulenbelt, I., Wapenaar, M.C., Patterson, D., Vijg, J., and Uitterlinden, A.G. Isolation of human chromosome 21-specific cosmids using a probe for RTVL-H retroviral sequences. (Genomics)

3. Strauss, B.H., Macleod D.C., de Feyter, P.J., van Suylen, R.-J., Uitterlinden, A.G., de Leeuw, W.J.F., Trommelen, G.J.J.M, and Serruys, P.W. Analysis of VNTR loci amplified by the Polymerase Chain Reaction to investigate the origin of intimal smooth muscle cells in a coronary artery lesion after heart transplantation in man. (American Heart Journal)

4. Trommelen, G.J.J.M., den Daas, Vijg, J., and Uitterlinden, A.G. Identity- and paternity testing in cattle: Application of a Desoxyribonucleic acid profiling protocol. (J. Dairy Science).

5. Muyzer, G., de Waal, E., and Uitterlinden, A.G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rDNA. (Applied and Environmental Microbiology).

Submitted:

1. Morolli, B., Uitterlinden, A.G., van Dam, H., Tates, A., Vijg, J. and Lohman, P.H.M. One- and two-dimensional DNA typing as a potential tool for human genome monitoring.

2. Trommelen, G.J.J.M., den Daas, J.H.G., Vijg, J., and Uitterlinden, A.G. DNA-profiling of cattle using micro- and minisatllite core probes.

3. Verwest, A.M., Molijn, A.C., Andersen, T.I., Borresen, A.-L., Uitterlinden, A.G., Mullaart, E., and Vijg, J. Identification of specific genetic changes in breast tumors by two-dimensional DNA typing.

4. Molijn, A.C., Borglund, A., Verwest, A.M., Mullaart, E., Kruse, T., Vijg, J., and Uitterlinden, A.G. Direct isolation of micro- and minisatellite flanking sequences from two-dimensional DNA typing gels.

5. te Meerman, G.J., Mullaart, E., Scheffer, H., Uitterlinden, A.G., and Vijg, J. Genetic analysis of two-dimensional DNA typing patterns.

6. Uitterlinden, A.G., Meulenbelt, I., Patterson, D., and Vijg, J. Identifcation and mapping of human chromosome 21-specific inter-repeat PCR amplification products.

In preparation:

1. Meulenbelt, I., Zeegers, P., Bakker, A.Q., Vijg, J., and Uitterlinden, A.G.
10 RFLP reports on human DNA probes:
-A human VNTR polymorphism (MICSf31) at chromosome 3 (D3S1245)
-An Eco RI polymorphism (MICSf2.17) at human chromosome 4 (D4S252)
-An Rsa I polymorphism (pMICS17) at human chromosome 6 (D6S237)
-An Eco RI polymorphism (pMICS13) at human chromosome 7 (D7S475)
-An Eco RI polymorphism (pMICS14a) at human chromosome 8 (D8S253)
-A VNTR polymorphism (pMICS30) at human chromosome 17
-Two adjacent human VNTR polymorphisms (pMICS7a and 7b) at chromosome 21q22.3 (D21S255 and D21S256)
-An Eco RI polymorphism (pMICS22) at human chromosome 21q11.2-q21 (D21S257)
-An Eco RI polymorphism (pMICS19) at human chromosome 21q11.2-q21
-A Msp I polymorphism (pMICS10) at human chromosome 21q11.2-q21

INVITED ARTICLES IN REFEREED JOURNALS

Published:

1. Uitterlinden, A.G. and J. Vijg. (1989) Two-dimensional DNA typing. TIBTECH 7, 336-341.

2. Uitterlinden, A.G., Slagboom, P.E., Mullaart, E., and Vijg, J. (1991) Genome scanning by two-dimensional DNA typing: The use of repetitive DNA sequences for rapidly mapping genetic traits. Electrophoresis 12, 119-136.

3. Uitterlinden, A.G., Mullaart, E., Morolli, B., and Vijg, J. (1991) Genome scanning of higher eukaryotes by two-dimensional DNA typing using micro- and minisatellite core probes. Methods, A companion to Methods in Enzymology 3, 83-90.

4. Hovig, E., Smith-Sorensen, B., Uitterlinden, A.G., and Börresen, A.-L. (1992) Detection of DNA variation in Cancer. Pharmacogenetics 2, 317-328.

CHAPTERS IN BOOKS AND INVITED ARTICLES

Published:

1. Vijg, J., Uitterlinden, A.G., and Knook, D.L. (1984) Arrangement and methylation status of ras oncogenes in liver hyperplastic nodules. In: Pharmacological, morphological and physiological aspects of liver ageing. Edited by C.F.A. van Bezooyen. EURAGE, Rijswijk, 49-55.

2. Uitterlinden, A.G., Vijg, J., and Knook, D.L. (1985) Variation in restriction fragment length and methylation pattern of rat major histocompatibillity complex class I genes. Transplantation Proceedings Vol. XVII, 1805-1807.

3. Vijg, J., Uitterlinden, A.G., Mullaart, E., Lohman, P.H.M., and Knook, D.L. (1985) Processing of DNA damage during ageing: induction of genetic alteration. Edited by: R.S. Sohal, L.S. Birnbaum, and R.G. Cutler. Raven Press, New York, 155-171.

4. Slagboom, P.E., Uitterlinden, A.G., and Vijg, J. (1986) Screening for age-related changes in gene expression in the rat liver. In: Liver, drugs, and ageing. Edited by: C.F.A. van Bezooyen, F.Miglio and Knook, D.L.. EURAGE, Rijswijk, 127-133.

5. Beyreuther, K.T., Cerutti, P.A., Clark, B.F.C., Delabar, J.M., Esser, K., Franceschi, C., Kirkwood, T.B.L., Rattan, S.I.S., Treton, J.A., Uitterlinden, A.G., Vandenberghe, A.M., and Vijg, J. (1988) Molecular Biology of ageing: Research Programme of the EURAGE Molecular Biology Group, EURAGE, Rijswijk.

6. Uitterlinden, A.G. (1989) DNA-fingerprinting en rechtspraak. Analyse 44, 13-8.

7. Uitterlinden, A.G., Giessen, R., den Daas, N., and Vijg, J. (1989) DNA-fingerprinting: een alternatief voor het bloedgroepen onderzoek? Veeteelt 6, 19, 1002-1004.

8. Vijg, J., Gossen, J.A., Slagboom, P.E., and Uitterlinden, A.G. (1990) New methods for the detection of DNA sequence variation. In: Early human retroviruses. UCLA symposia on molecular and cellular biology. New series 123. Edited by M. Cleggand S. O'Brien. A.R. Liss, New York .

9. Vijg, J., Gossen, J.A., Slagboom, P.E., and Uitterlinden, A.G. (1990) New methods for the detection of DNA sequence variation: Applications in molecular genetic studies on aging. In: Molecular Biology of Aging, Edited by Finch, C.E. & Johnson, T.E., A.R. Liss, New York, 103-119.

10. Vijg, J., Gossen, J.A., de Leeuw, W.J.F., Mullaart, E., Slagboom, P., and Uitterlinden, A.G. (1990) New methods for detecting DNA sequence variation in releation to aging. In: Molecular Mechanisms of Aging, Edited by Beyreuther, K. & Schettler, G., Springer Verlag, Heidelberg, 77-88.

11. Slagboom, P., Mullaart, E., Uitterlinden, A.G. and Vijg, J. (1990) Two-dimensional DNA typing as a tool for detecting somatic instability in aging rats. In: From Gene to Man: gerontological research in the Netherlands (Edited by van Bezooijen, C.F.A., Ravid, R., and Verhofstad, A.A.), Stichting Gerontologie en Geriatrie, Rijswijk, 345-348.

12. Uitterlinden, A.G., Meulenbelt, I. and Vijg, J. (1990) Isolation of polymorphic DNA markers for the familial Alzheimer's disease region on chromosome 21. In: From Gene to Man: gerontological research in the Netherlands (Edited by van Bezooijen, C.F.A., Ravid, R., and Verhofstad, A.A.), Stichting Gerontologie en Geriatrie, Rijswijk, 353-356.

13. Lindhoudt, P., Verkerk, R., Mullaart, E., Uitterlinden, A.G., and Vijg, J. (1991) Cultivar identification of tomato by DNA analysis with repetitive probes. TGC-Report 41.

14. De Ruiter, J.R., Scheffrahn, W., Trommelen, G.J.J.M., Uitterlinden, A.G., Martin, R.D., Van Hooff, J.A.R.A.M. (1992) Male social rank and reproductive success in Wild long-tailed macaques: paternity exclusions by blood protein analysis and DNA fingerprinting. In: Paternity in Primates (Scheffrahn and Hoperti, Eds.) Karger publishers, 11, 74-96.

In press:

1. Uitterlinden, A.G., Mullaart, E., Molijn, A.C., Trommelen, G.J.J.M., and Vijg, J. DNA diagnostics in several species. In: Advances in Molecular Genetics

BOOKS

In preparation:

1. Uitterlinden, A.G., and Vijg, J. Two-dimensional DNA typing. Ellis Horwood publishers (expected end 1993).

2. Vijg, J., and Uitterlinden, A.G. Gene mapping. Elsevier Scientific Publications (expected 1994)

PATENTS:

1. J. Vijg and A.G. Uitterlinden: "A method for the simultaneous determination of DNA sequence variations at a large number of sites, and a kit suitable therefore". United States patent # 5,068,176, European patent # 89/201-095 (pending in Japan)

2. A.G. Uitterlinden and J. Vijg: "A method for the simultaneous determination of DNA sequence variation at a large number of sites on specific chromosomes or defined parts thereof, and a kit suitable therefore". (pending in USA, Europe, and Japan)