POLYMERASE CHAIN REACTION AS A SUCCESSFULL BIOTECHNOLOGICAL APPLICATION. WAYS WE USE PCR IN THE FIELDS OF BIOINFORMATICS FORENSICS AND GENETICS.

(ΑΛΥΣΙΔΩΤΗ ΑΝΤΙΔΡΑΣΗ ΤΗΣ ΠΟΛΥΜΕΡΑΣΗΣ ΩΣ ΕΠΙΤΥΧΗΜΕΝΗ ΒΙΟΤΕΧΝΟΛΟΓΙΚΗ ΕΦΑΡΜΟΓΗ. ΤΡΟΠΟΙ ΧΡΗΣΗΣ ΤΗΣ ΣΤΑ ΠΕΔΙΑ ΤΗΣ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗΣ, ΙΑΤΡΟΔΙΑΣΤΙΚΗΣ ΚΑΙ ΓΕΝΕΤΙΚΗΣ)

Σπυριδοφλα Γ. Λάζαρη

Επιβλέποντες: Γιώργος Ματσόπουλος
Επίκουρος καιηγητής, Ε.Μ.Π
Επίκουρα Ερευνητικό Προσωπικό Ιωάννης Μακρής
Επιβλέποντες: Νικόλαος Ουζούνογλου
Καθηγητής, Ε.Μ.Π
Επιβλέποντες: Διονύσιος-Δημήτριος
Καθηγητής, Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 5η Μαίου 2011.

..................................................  ..................................................  ..................................................
Γιώργος Ματσόπουλος  Νικόλαος Ουζούνογλου  Διονύσιος-Δημήτριος
Επίκουρος καθηγητής, Ε.Μ.Π  Καθηγητής  Κουτσούρης
..................................................
Καθηγητής

Αθήνα Μάιος 2011
Copyright © Σπυριδούλα Γ. Λάζαρη 2011

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντίγραφη, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

.................................

Σπυριδούλα Γ. Λάζαρη
Διπλωματούχος Ψηφιακών Συστημάτων
Introduction ........................................................................................................................................... 5

1. Stages of PCR procedure ................................................................................................................. 6

2. Successful strategy for PCR – design of suitable primer-contamination ................................. 8

3. Design and choice of a suitable primer .......................................................................................... 9

4. PCR techniques ................................................................................................................................. 10

   a. Allele-specific PCR ..................................................................................................................... 10
   b. Assembly PCR .......................................................................................................................... 10
   c. Asymmetric PCR ....................................................................................................................... 10
   d. Helicase-dependent amplification ............................................................................................. 10
   e. Hot-start PCR ............................................................................................................................. 10
   f. Intersequence-specific PCR ....................................................................................................... 10
   g. Inverse PCR ............................................................................................................................... 11
   h. Ligation-mediated PCR ............................................................................................................. 11
   i. Methylation-specific PCR .......................................................................................................... 11
   k. Miniprimer PCR ........................................................................................................................ 11
   l. Multiplex Ligation-dependent Probe Amplification ................................................................... 11
   m. Multiplex-PCR .......................................................................................................................... 11
   n. Nested PCR ............................................................................................................................... 11
   o. Overlap-extension PCR .............................................................................................................. 12
   p. Quantitative PCR (Q-PCR) ......................................................................................................... 12
   q. Reverse Transcription PCR (RT-PCR) ....................................................................................... 12
   r. Solid Phase PCR ........................................................................................................................ 13
   s. Thermal asymmetric interlaced PCR (TAIL-PCR) ................................................................... 13
   t. Touchdown PCR ......................................................................................................................... 13
   u. PAN-AC ....................................................................................................................................... 13
   v. Universal Fast Walking ............................................................................................................... 13

5. SNPs related to possible development of a particular illness ...................................................... 14

   5.1. SNPs related to possible development of Alzheimer ............................................................... 15
      5.1a. SNPs related to the PS1 gene ............................................................................................ 15
      5.1b. Detection of apolipoprotein E genotypes in Alzheimer disease .................................. 17
      5.1c. Alzheimer and SNPs related to PS2 gene ....................................................................... 19
      5.1d. Missense Mutations in APP and Alzheimer .................................................................... 22
   5.2 Cancer and PCR ......................................................................................................................... 26
      5.2a Quantitative real-time PCR for cancer detection ............................................................... 26
5.2.b Detection of prostate cancer using methylation-specific PCR .................................. 29
5.2.c. Quantitative real-time PCR for cancer detection: the lymphoma case .................. 30

6. Applications of PCR in genetic analysis ........................................................................ 33

6.1. Genetic testing companies ......................................................................................... 33
6.2. Genetic testing ........................................................................................................... 34
  6.2.a Newborn screening .................................................................................................. 35
  6.2.b Diagnostic testing .................................................................................................... 35
  6.2.c. Carrier testing ........................................................................................................ 35
  6.2.d. Prenatal testing ...................................................................................................... 35
  6.2.e. Predictive and presymptomatic testing ................................................................. 35
  6.2.f. Forensic testing ...................................................................................................... 35
  6.2.g Research testing ...................................................................................................... 35
  6.2.h. Genetic testing in a shoe box .............................................................................. 38
6.3. DNA Paternity Testing ............................................................................................ 39
6.4. DNA testing and forensics ....................................................................................... 46
  6.5 From DNA fingerprinting to DNA profiling in forensic studies .......................... 49

7. Recombinant DNA technology ..................................................................................... 50
  7.a. Rheumatoid arthritis treatment with a recombinant human TNF ......................... 53

8. Ancient DNA and the polymerase chain reaction ....................................................... 55
  8.a. The Polymerase Chain Reaction in Molecular Archaeology ............................... 56
  8.b. Ancient DNA Sequences Revealed via PCR ......................................................... 57

Conclusion ......................................................................................................................... 58
References .......................................................................................................................... 60
Introduction

The increased interest to molecular diagnostics the last years is a result of the great evolution and development of molecular genetics. Molecular genetics use molecular methods that amplify specific fragments of DNA. Today, the molecular techniques which were developed for amplification and detection of specific sequences of nucleic acids helped in a great deal to understand the structure of many diseases.

Polymerase chain reaction or PCR is a technique that is used for isolation and amplification of a specific sequence of DNA. PCR is an in vitro method that exploits the in vivo procedure of replication of DNA. DNA polymerase is used to amplify a piece of DNA by in vitro enzymatic replication. As PCR progresses, the DNA thus generated is itself used as a template for replication. This sets in motion a chain reaction in which the DNA template is exponentially amplified. With PCR it is possible to amplify a single or few copies of a piece of DNA across several orders of magnitude, generating millions or more copies of the DNA piece. PCR was designed and presented by Dr Kary Mullis (1983). From the Human Genome Project to the search for targets to the development of gene tests, there are few areas of genetic research today that do not depend on PCR. Today PCR has a great range of implementations related to 1) the cloning and the study of gene’s expression 2) the detection of mutations that are responsible for hereditary diseases 3) Criminology, Toxicology and Forensics.

Essentials components to succeed polymerase chain reaction

Replication of a piece of DNA via PCR can be successful only if it is known the exact sequence of nucleotides that is synthesized. Almost all PCR applications employ a heat-stable DNA polymerase, such as Taq-polymerase, an enzyme originally isolated from the bacterium Thermus aquaticus. This DNA polymerase enzymatically assembles a new DNA strand from DNA building blocks, the nucleotides, using single-stranded DNA as template and DNA oligonucleotides (also called DNA primers) required for initiation of DNA synthesis.

Generally, PCR procedure requires several components as:

1) DNA template that contains the DNA region (target) to be amplified
2) Two primers, which are complementary to the DNA regions at the 5’ (five prime) or 3’ (three prime) ends of the DNA region.
3) A DNA polymerase such as Taq-polymerase or another DNA polymerase with a temperature at 70°C
4) Building blocks called Deoxynucleoside triphosphates (Dntps) from which DNA polymerases synthesize a new DNA strand.
5) Buffer solution, providing a suitable chemical environment for optimum activity and stability of the DNA polymerase.
6) Divalent cations, magnesium or manganese ions;
7) Monovalent cation potassium ions.
The PCR is commonly carried out in a reaction volume of 10-200 ml in small reaction tubes (0.2-0.5 ml volumes) in a thermal cycler. The thermal cycler heats and cools the reaction tubes to achieve the temperatures required at each step of the reaction. Many modern thermal cyclers make use of the Peltier effect which permits both heating and cooling of the block holding the PCR tubes simply by reversing the electric current. Thin-walled reaction tubes permit favorable thermal conductivity to allow for rapid thermal equilibration. Most thermal cyclers have heated lids to prevent condensation at the top of the reaction tube. Older thermal cyclers lacking a heated lid require a layer of oil on top of the reaction mixture or a ball of wax inside the tube. [1]

**Stages of PCR procedure**

The polymerase chain reaction serves to copy DNA. The temperatures used and the length of time they are applied in each cycle depend on a variety of parameters. These include the enzyme used for DNA synthesis, the concentration of divalent ions and dNTPs in the reaction, and the melting temperature (Tm) of the primers. It uses repeated cycles, each of which consists of three steps:

1. The reaction solution containing DNA molecules (to be copied), polymerases (which copy the DNA), primers (which serve as starting DNA) and nucleotides (which are attached to the primers) is heated to 94°C for 20-30 seconds. This causes the two complementary strands to separate, a process known as **denaturing or melting**.

2. Lowering the temperature to 54°C for 20-45 seconds causes the primers to bind to the DNA, a process known as **hybridization or annealing**. The resulting bonds are stable only if the primer and DNA segment are complementary, i.e. if the base pairs of the primer and DNA segment match. The polymerases then begin to attach additional complementary nucleotides at these sites, thus strengthening the bonding between the primers and the DNA.

3. **Extension**: The temperature is again increased, this time to 72°C. This is the ideal working temperature for the polymerases used, which add further nucleotides to the developing DNA strand. At the same time, any loose bonds that have formed between the primers and
DNA segments that are not fully complementary are broken. Each time these three steps are repeated the number of copied DNA molecules doubles. After 20 cycles about a million molecules are cloned from a single segment of double-stranded DNA. The temperatures and duration of the individual steps described above refer to the most commonly used protocol. A number of modifications have been introduced that give better results to meet specific requirements. [2]

**Figure 2**: schematic drawing of the PCR cycle. (1) denaturing at 94°C - 96°C. (2) Annealing at 54°C (3) elongation at 72°C.

To check whether PCR generated the anticipated DNA fragment, agarose gel electrophoresis is employed for size separation of the PCR products. The size of PCR products is determined by comparison with a DNA ladder (a molecular weight marker), which contains DNA fragments of known size, run on the gel alongside the PCR products. As the number of copies increases, the products consist of a greater proportion of fragments with just the target DNA. After four cycles, half of the fragments consist of just target DNA, and half of the fragments also contain flanking DNA. [3]
**Figure 3:** with each additional cycle the number of copies of our target sequence doubles. At the end of cycle 35 there are more than 68 billion copies of this double stranded target region.

**PCR procedure**

The PCR process can be divided into three stages:

- **Exponential amplification:** at every cycle, the amount of product is doubled (assuming 100% reaction efficiency). The reaction is very specific and precise.

- **Leveling off stage:** The reaction slows as the DNA polymerase loses activity and as consumption of reagents such as dNTPs and primers causes them to become limiting.

- **Plateau:** No more product accumulates due to exhaustion of reagents and enzyme.

**Successful strategy for PCR – design of suitable primer-contamination**

Every successful PCR reaction starts with a preparation of the nucleic acids that contains the DNA target we intend to amplify. Primers define this sequence so it is not necessary the isolation of this fragment of DNA. Great sensitivity of PCR requires procedures of isolation of nucleic acids that have to:

1) Avoid the contamination from other DNA/RNA  
2) Isolate as it is the DNA or the RNA  
3) Avoid the exposure to chemicals (e.g. phenol )  
4) To be quick  
5) To have the ability to process many samples simultaneously.
The last years there are available, isolation kit of nucleic acids that isolate and clean up selectively DNA or RNA, with the help of special membranes and simple stages of centrifugal process. These are quick and simple automated procedures, a fact that is very important for the process of many clinical samples.

**Design and choice of a suitable primer**

The design and the choice of a special primer is a very important factor for the success of PCR. The molecular proportion of the primers related to the DNA-target have to be bigger. Ideally, PCR primers should have GC content similar to the fragment of DNA that is going to be amplified. It would be wise to be avoided primers that contain polypurines, polypyrimidines or repeated sequences. It is essential not to choose sequences with important secondary structure especially at the 3’ end of the primer. It has to be checked the complementation of the primers and not to choose primers that have similarity at the 3’ end, because it is very possible the primer to be annealed using as a template itself or the other primer. This has the result to give a wrong product. Several useful computer programs have been developed to aid in efficient primer design (OLIGO, National Biosciences, Gene Works and PCGene, Intelligenetics).

Generally:

1) Primer sequences need to be those that are exact or very nearly exact matches for the desired target to be amplified and should have virtually no homology to any other sequences in the template mixtures
2) The reaction conditions during the first few cycles should be those that minimize the possibility of primers initiating an extension reaction from any part of the template mixture except the desired sequences.
3) The reaction conditions after the first few cycles should be those that allow all newly synthesized molecules to be perfectly replicated at high efficiency so that they double in number at each cycle.
4) The reaction should be stopped before the major product becomes saturated and can no longer increase exponentially while the contaminant molecules continue to increase.

**Contamination**

Contamination is a serious consideration when carrying out PCR reactions. To prevents contamination of samples by previously amplified products, considerable thought and care should be devoted to the actual logistics of the experiments. A set of micropipettes should be devoted exclusively to preparation of the reaction mixtures at a designed PCR workstation set apart from the rest of the laboratory. Products from complete amplifications should be handled with a separate set of micropipettes at a different location in the laboratory. The key concern is “carry over” of amplified target to unamplified reaction
mixtures. Negative controls (i.e. complete reaction mixtures lacking template DNA) should be included in each experiment as an assay for overt contamination. [4]

**PCR techniques**

**Allele-specific PCR:** Selective PCR amplification of one of the alleles to detect Single Nucleotide Polymorphism (SNP). Selective amplification is usually achieved by designing a primer such that the primer will match/mismatch one of the alleles at the 3'-end of the primer. [5]

**Assembly PCR** or **Polymerase Cycling Assembly (PCA):** is the artificial synthesis of long DNA sequences by performing PCR on a pool of long oligonucleotides with short overlapping segments. The oligonucleotides alternate between sense and antisense directions and the overlapping segments determine the order of the PCR fragments, thereby selectively producing the final long DNA product.

**Asymmetric PCR:** preferentially amplifies one DNA strand in a double-stranded DNA template. It is used in sequencing and hybridization probing where amplification of only one of the two complementary strands is required. PCR is carried out as usual, but with a great excess of the primer for the strand targeted for amplification. Because of the slow (arithmetic) amplification later in the reaction after the limiting primer has been used up, extra cycles of PCR are required. A recent modification on this process, known as Linear-After-The-Exponential-PCR (LATE-PCR), uses a limiting primer with a higher melting temperature (Tm) than the excess primer to maintain reaction efficiency as the limiting primer concentration decreases mid-reaction.

**Helicase-dependent amplification:** similar to traditional PCR, but uses a constant temperature rather than cycling through denaturation and annealing/extension cycles. DNA helicase, an enzyme that unwinds DNA, is used in place of thermal denaturation.

**Hot-start PCR:** is a technique that reduces non-specific amplification during the initial set up stages of the PCR. It may be performed manually by heating the reaction components to the melting temperature (e.g., 95°C) before adding the polymerase. Specialized enzyme systems have been developed that inhibit the polymerase's activity at ambient temperature, either by the binding of an antibody or by the presence of covalently bound inhibitors that only dissociate after a high-temperature activation step. Hot-start/cold-finish PCR is achieved with new hybrid polymerases that are inactive at ambient temperature and are instantly activated at elongation temperature.

**Intersequence-specific PCR** (ISSR): a PCR method for DNA fingerprinting that amplifies regions between simple sequence repeats to produce a unique fingerprint of amplified fragment lengths.
**Inverse PCR**: is commonly used to identify the flanking sequences around genomic inserts. It involves a series of DNA digestions and self ligation, resulting in known sequences at either end of the unknown sequence.

**Ligation-mediated PCR**: uses small DNA linkers ligated to the DNA of interest and multiple primers annealing to the DNA linkers; it has been used for DNA sequencing, genome walking, and DNA footprinting.

**Methylation-specific PCR** (MSP): developed by Stephen Baylin and Jim Herman at the Johns Hopkins School of Medicine, and is used to detect methylation of CpG islands in genomic DNA. DNA is first treated with sodium bisulfite, which converts unmethylated cytosine bases to uracil, which is recognized by PCR primers as thymine. Two PCRs are then carried out on the modified DNA, using primer sets identical except at any CpG islands within the primer sequences. At these points, one primer set recognizes DNA with cytosines to amplify methylated DNA, and one set recognizes DNA with uracil or thymine to amplify unmethylated DNA. MSP using qPCR can also be performed to obtain quantitative rather than qualitative information about methylation.

**Miniprimer PCR**: uses a thermostable polymerase (S-Tbr) that can extend from short primers ("smalligos") as short as 9 or 10 nucleotides. This method permits PCR targeting to smaller primer binding regions, and is used to amplify conserved DNA sequences, such as the 16S (or eukaryotic 18S) rRNA gene.

**Multiplex Ligation-dependent Probe Amplification** (MLPA): permits multiple targets to be amplified with only a single primer pair, thus avoiding the resolution limitations of multiplex PCR (see below).

**Multiplex-PCR**: consists of multiple primer sets within a single PCR mixture to produce amplicons of varying sizes that are specific to different DNA sequences. By targeting multiple genes at once, additional information may be gained from a single test run that otherwise would require several times the reagents and more time to perform. Annealing temperatures for each of the primer sets must be optimized to work correctly within a single reaction, and amplicon sizes, i.e., their base pair length, should be different enough to form distinct bands when visualized by gel electrophoresis.

**Nested PCR**: increases the specificity of DNA amplification, by reducing background due to non-specific amplification of DNA. Two sets of primers are used in two successive PCRs. In the first reaction, one pair of primers is used to generate DNA products, which besides the intended target, may still consist of non-specifically amplified DNA fragments. The product(s) are then used in a second PCR with a set of primers whose binding sites are completely or partially different from and located 3’ of each of the primers used in the first reaction. Nested PCR is often more successful in specifically amplifying long DNA fragments than conventional PCR, but it requires more detailed knowledge of the target sequences.

Procedure:

1. The target DNA undergoes the first run of polymerase chain reaction with the first set of primers, shown in green. The selection of alternative and similar primer binding sites gives a selection of products, only one containing the intended sequence.
2. The product from the first reaction undergoes a second run with the second set of primers, shown in red. It is very unlikely that any of the unwanted PCR products contain binding sites for both the new primers, ensuring the product from the second PCR has little contamination from unwanted products of primer dimers, hairpins, and alternative primer target sequences. [6]

![Diagram of nested PCR stages](image)

**Figure 4**: Diagram of the specific stages of nested PCR

**Overlap-extension PCR**: a genetic engineering technique allowing the construction of a DNA sequence with an alteration inserted beyond the limit of the longest practical primer length.

**Quantitative PCR (Q-PCR)**: used to measure the quantity of a PCR product (commonly in real-time). It quantitatively measures starting amounts of DNA, cDNA or RNA. Q-PCR is commonly used to determine whether a DNA sequence is present in a sample and the number of its copies in the sample. *Quantitative real-time PCR* has a very high degree of precision. QRT-PCR methods use fluorescent dyes, such as Sybr Green, EvaGreen or fluorophore-containing DNA probes, such as TaqMan, to measure the amount of amplified product in real time. It is also sometimes abbreviated to RT-PCR (*Real Time PCR*) or RQ-PCR. QRT-PCR or RTQ-PCR are more appropriate contractions, since RT-PCR commonly refers to reverse transcription PCR, often used in conjunction with Q-PCR.

**Reverse Transcription PCR (RT-PCR)**: for amplifying DNA from RNA. Reverse transcriptase reverse transcribes RNA into cDNA, which is then amplified by PCR. RT-PCR is widely used in expression profiling, to determine the expression of a gene or to identify the sequence of an RNA transcript, including transcription start and termination sites. If the genomic DNA sequence of a gene is known, RT-PCR can be used to map the location of exons and introns.
in the gene. The 5' end of a gene (corresponding to the transcription start site) is typically identified by RACE-PCR (*Rapid Amplification of cDNA Ends*).

**Solid Phase PCR**: encompasses multiple meanings, including Polony Amplification (where PCR colonies are derived in a gel matrix, for example), Bridge PCR (primers are covalently linked to a solid-support surface), conventional Solid Phase PCR (where Asymmetric PCR is applied in the presence of solid support bearing primer with sequence matching one of the aqueous primers) and Enhanced Solid Phase PCR (where conventional Solid Phase PCR can be improved by employing high Tm and nested solid support primer with optional application of a thermal 'step' to favour solid support priming).

**Thermal asymmetric interlaced PCR (TAIL-PCR)**: for isolation of an unknown sequence flanking a known sequence. Within the known sequence, TAIL-PCR uses a nested pair of primers with differing annealing temperatures; a degenerate primer is used to amplify in the other direction from the unknown sequence.

**Touchdown PCR** (*Step-down PCR*): a variant of PCR that aims to reduce nonspecific background by gradually lowering the annealing temperature as PCR cycling progresses. The annealing temperature at the initial cycles is usually a few degrees (3-5°C) above the $T_m$ of the primers used, while at the later cycles, it is a few degrees (3-5°C) below the primer $T_m$. The higher temperatures give greater specificity for primer binding, and the lower temperatures permit more efficient amplification from the specific products formed during the initial cycles.

**PAN-AC**: uses isothermal conditions for amplification, and may be used in living cells.

**Universal Fast Walking**: for genome walking and genetic fingerprinting using a more specific 'two-sided' PCR than conventional 'one-sided' approaches (using only one gene-specific primer and one general primer - which can lead to artefactual 'noise') by virtue of a mechanism involving lariat structure formation. Streamlined derivatives of UFW are LaNe RAGE (lariat-dependent nested PCR for rapid amplification of genomic DNA ends), 5'RACE LaNe and 3'RACE LaNe. [7]

**Practical Applications**

The polymerase chain reaction has become one of the most widely used techniques in biomedical research, enabling the fast inexpensive production of large quantities of DNA from minute amounts of starting material. Although the basic procedure is very simple, many variations of it have been developed for different applications, for use in such diverse areas as forensics and diagnostic medicine, genetic engineering and the food industry.[8]
**SNPs related to possible development of a particular illness**

Single nucleotide polymorphisms, or SNPs (pronounced "snips"), are DNA sequence variations that occur when a single nucleotide (A, T, C, or G) in the genome sequence is altered. For example a SNP might change the DNA sequence AAGGCTAA to ATGGCTAA. For a variation to be considered a SNP, it must occur in at least 1% of the population. SNPs, which make up about 90% of all human genetic variation, occur every 100 to 300 bases along the 3-billion-base human genome. Two of every three SNPs involve the replacement of cytosine (C) with thymine (T). SNPs can occur in coding (gene) and noncoding regions of the genome. Many SNPs have no effect on cell function, but scientists believe others could predispose people to disease or influence their response to a drug.

Although more than 99% of human DNA sequences are the same, variations in DNA sequence can have a major impact on how humans respond to disease; environmental factors such as bacteria, viruses, toxins, and chemicals; and drugs and other therapies. This makes SNPs valuable for biomedical research and for developing pharmaceutical products or medical diagnostics. SNPs are also evolutionarily stable—not changing much from generation to generation—making them easier to follow in population studies.

Scientists believe SNP maps will help them identify the multiple genes associated with complex ailments such as cancer, diabetes, vascular disease, and some forms of mental illness. These associations are difficult to establish with conventional gene-hunting methods because a single altered gene may make only a small contribution to the disease.

SNPs do not cause disease, but they can help determine the likelihood that someone will develop a particular illness.

**Variations in genes related to Alzheimer**

One of the genes associated with Alzheimer’s disease, apolipoprotein E or ApoE, is a good example of how SNPs affect disease development. ApoE contains two SNPs that result in three possible alleles for this gene: E2, E3, and E4. Each allele differs by one DNA base, and the protein product of each gene differs by one amino acid.

Each individual inherits one maternal copy of ApoE and one paternal copy of ApoE. Research has shown that a person who inherits at least one E4 allele will have a greater chance of developing Alzheimer’s disease. Apparently, the change of one amino acid in the E4 protein alters its structure and function enough to make disease development more likely. Inheriting the E2 allele, on the other hand, seems to indicate that a person is less likely to develop Alzheimer’s.

Of course, SNPs are not absolute indicators of disease development. Someone who has inherited two E4 alleles may never develop Alzheimer’s disease, while another who has inherited two E2 alleles may. ApoE is just one gene that has been linked to Alzheimer’s. Like most common chronic disorders such as heart disease, diabetes, or cancer, Alzheimer’s is a disease that can be caused by variations in several genes. The polygenic nature of these disorders is what makes genetic testing for them so complicated.[9]
Single nucleotide polymorphism associated with Late-onset Alzheimer's disease

Alzheimer’s disease is a neurodegenerative disorder characterized by progressive memory loss and is the most common cause of dementia in the elderly. Although the age of onset varies, most of the cases of AD occur at a late age (late-onset AD). Efforts to identify the genetic causes of certain rare familial forms of AD, representing a small fraction of all known AD cases, have been largely successful.

By the use of linkage mapping methods, mutations in the genes for amyloid precursor protein (APP), apolipoprotein E (ApoE), presenilin-1 (PSEN1) and presenilin-2 (PSEN2) have been found to underlie the majority of familial cases. E4 allele of ApoE locus on chromosome 19 has been associated with both familial and sporadic late-onset AD. However, only half of all patients with AD have E4 allele and a substantial number of people having E4 allele do not develop AD. This is a reason that leads to the conclusion that there are other genes as we mentioned, associated with AD. Firstly, PS1 has been suggested as a potential risk factor for late onset AD.

Therefore, a study was carried out to search for Single Nucleotide Polymorphism (SNP) in PS1 in a number of Saudi AD patients. This study describes the association of SNPs with late-onset AD in patients referred for diagnostic testing at King Fahad Guard Hospital and King Khalid University Hospital, Riyadh, Saudi Arabia over a period of eight years. [10]

**Description of the procedure in order to discover possible SNPs related to the PS1 gene**

Saudi patients with Late-Onset Alzheimer’s Disease (LOAD) were recruited from those who visited King Fahad National Guard Hospital and King Khalid University Hospital during 1994-2001. All patients in the study had a probable diagnosis of AD
according to DSM IV criteria [11] and were found to have AD at more than 65 years of age indicating the late-onset form of AD. The age of the patients at the time of testing ranged from 67 to 80 years with a mean age of 70 years. A number of age and sex-matched normal controls were also recruited from the same population.

DNA was extracted from the blood with standard procedures utilizing proteinase-K/phenol/chloroform extraction. Primers were designed on the basis of the sequence data for various exons of PS1 available in the GenBank to amplify the coding sequence of respective exons and flanking intronic regions. PCR was performed using PuRe Taq Ready-To-Go PCR Beads with different sets of primers. A 200 to 300 ng of genomic DNA was used as a template in 25 ml reaction. Genomic DNA was amplified for 40 cycles. Each cycle consisted of 94 °C for 30 sec, 53 °C for 30 sec, 72 °C for 1 min.

PCR products obtained were separated by electrophoresis on 1.5 % agarose gel visualized by ethidium bromide fluorescence. Fragments with the expected size were cut from the gel, purified using GFX PCR DNA Gel band purification kit. Purified DNA was sequenced using forward and/or backward primers. The DNA sequence was compared with PS1 genomic sequences from normal individuals and data available in Genbank.

Amplified DNA when sequenced and compared with the sequences of the respective exon from normal and also with the data available in GenBank, showed two SNPs at the 5′ and 3′ intronic regions of exon 8 (encoding for amino acid 184 to 256). There were T56240G and T56454C in the 5′ and 3′ intronic region of exon 8, respectively (Fig. 5A and 5B).

Fig. 5

Comparison of electrophorograms of PS1 intronic DNA sequences of AD patients with those of normal human. (A. Showing the T56454C SNP. B. Showing the T56240G SNP).
These variations noticed in the intronic region seem to segregated with late – onset AD. No other variations were found in the coding as well as non coding sequences of exon 8 of all other AD patients, which was further confirmed by the comparison of their electrophorograms.

Generally, the search in the PS1 gene in sporadic LOAD Saudi patients discovered 2 SNPs, the T56454C SNP and the T56240G SNP. Furthermore, no mutations in PS1 coding sequences were found in sporadic LOAD patients. This agrees with the fact that these patients were over 65 yrs old. At this age group, as in the case of other ethnic group populations, most inherited AD-causing mutations are scattered in PS2 and ApoE specific genotype. Intronic polymorphism in PS1 gene has also been reported in exon 8 of late-onset American AD patients. Unfortunately, due to lack of information regarding family history, we could not follow the inheritance pattern. Finally the cause of disease in the rest of the patients where no mutations/polymorphism can be detected is unknown and it may be result of certain mutations on other loci or other genetic and/or environmental factors as suggested by Axelman et al [12],[13].

Detection of apolipoprotein E genotypes in Alzheimer disease

Apolipoprotein E (APOE) gene on chromosome 19q13.2 is encoded by three common alleles designated as ε2, ε3 and ε4. In Alzheimer’s disease (AD) the ε4 allele is overrepresented and is considered to be a major genetic risk factor. Several methods have been developed to determine APOE genotypes. At this point we are going to present a study that we used PCR-SSCP method for determining APOE genotypes in 42 cases of AD patients, 40 cases of non-AD dementia patients, and 49 cases of age-matched controls. DNA from the target sequence on APOE was amplified by PCR from peripheral blood genomic DNA. PCR products were electrophoresed in a non-denaturing polyacrylamide gel and visualized by silver staining.

It is now well-established that some cases of AD, mainly early-onset, show autosomal dominant inheritance patterns due to the presence of mutated genes, including those encoding amyloid precursor protein, presenilin 1 and presenilin 2. However, the late-onset disease, which is more challenging for genetic analysis, has been associated with the presence of apolipoprotein E type 4 (APOE ε4) [14]. The APOE gene encodes a 299 amino acid secreted protein (Blacker, 1998). There are three major isoforms of ApoE (E2, E3 and E4) that are three products of three allelic forms (ε2, ε3 and ε4) of this single gene locus. Three homozygous phenotypes (ApoE2/E2, E3/E3 and E4/E4) and three heterozygous phenotypes (ApoE3/2, E4/3 and E4/2) arise from the expression of any two of three alleles. The three isoforms differ by interchange of cysteine (Cys) and arginine (Arg) residues at positions 112 and 158 of the mature ApoE. ApoE2 has Cys residues in both of these positions, ApoE3 has Cys-112 and Arg-158, and ApoE4 has Arg in both positions.

PCR-SSCP vs PCR-RFLP analysis to determine APOE genotypes

Several methods have been developed to determine APOE genotypes, e.g. polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) [15] and polymerase chain reaction-single strand conformation polymorphism (PCRSSCP) [16]. The former is commonly used to detect the APOE genotype but it is rather time-consuming.
and there is an association with incomplete cleavage. The latter method does not require any restriction enzyme and can be used successfully to distinguish the three common APOE alleles that differ from each other by either one or two single-base substitutions. However, PCR-SSCP as previously reported is not very convenient, since it requires radiolabelled primers followed by autoradiography (Orita et al, 1989; Hayashi and Yandell, 1993). In this study we have developed a rapid and nonisotopic PCR-SSCP method for determining the distribution of APOE genotypes in Thai Alzheimer’s patients.

**Subjects**

Subjects were patients from the Department of Neurology, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand. Written informed consent was obtained from all patients or their primary caregivers. All subjects were diagnosed according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, DSM-IV by American Psychiatric Association, 1994 (Boller and Traykov, 1999). One hundred and thirty-one of subjects were obtained in our study. They were classified into 3 groups: 42 Alzheimer’s patients, 40 non-AD demented patients, and 49 age-matched controls. DNA was extracted from peripheral blood leukocytes using the method of Blin and Stafford (1976).

**Results**

Clinical characteristics including age and sex distribution of patients with AD, non-AD dementia and aged-matched controls are summarized in Table below. The mean age of all groups does not differ significantly. Eighty percent of patients had onset of the disease after the age of 60. The PCR-SSCP gel system provided results within 8 hours. The APOE polymorphism patterns of PCR products were silver-stained on 12% non-denaturing polyacrylamide gel. We detected that the APOE ε3/ε3 was the most common genotype in all groups followed by ε3/ε4, ε4/ε4 and ε2/ε3, respectively. The frequency of APOE ε4 allele was significantly higher in AD patients (33.3%) than in age matched controls (13.3%) (χ2 = 10.43; p = 0.001) and in non-AD dementia (10%) (χ2 = 13.02; p < 0.001) which was mainly due to an increase of APOE ε4 homozygotes. The frequency of APOE ε3 was significantly lower in AD patients (66.7%) than in age matched controls (85.7%) (χ2 = 10.68; p = 0.002) and in non-AD dementia (90%) (χ2 = 14.53; p < 0.001).

<table>
<thead>
<tr>
<th>Subject (n)</th>
<th>Median onset age (range)</th>
<th>APOE genotype</th>
<th>APOE allele frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD (42)</td>
<td>73 (51-93)</td>
<td>ε2/ε2 0 0 19 18 5 0</td>
<td>66.7 33.3</td>
</tr>
<tr>
<td>M (18) : F (24)</td>
<td>72 (63-90)</td>
<td>ε2/ε3 0 0 32 8 0 0</td>
<td>90 10</td>
</tr>
<tr>
<td>Non-AD (40)</td>
<td></td>
<td>ε2/ε4 ε3/ε3 ε3/ε4 ε4/ε4 ε4/ε4</td>
<td></td>
</tr>
<tr>
<td>M (12) : F (28)</td>
<td>70 (54-81)</td>
<td>0 0 1 0 35 13 0 1</td>
<td>85.7 13.3</td>
</tr>
<tr>
<td>Control (49)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M (14) : F (35)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

M, male; F, female ; n, numbers of cases.

**Figure 6:** results for ApoE genotypes for AD, non AD and age matched control subjects
**Generally**

The APOE e4 allele frequency to the AD patients was increased as it was expected related to non AD patients. In addition, homozygosity for the e4 was found only in patients with AD. We observed a protective effect of the e2 allele but it was limited to the e2/e3 genotype which was detected only in the control group. These data confirm previous reports (Saunders *et al*, 1996; Farrer *et al*, 1997) showing that the e2/e3 genotype is associated with a lower risk of AD. It is noteworthy that none of the subjects are homozygous for the e2 allele. Similarly, Farrer *et al* (1997) reported that the influence of the rare e2/e2 genotype on AD risk could not be discerned even in a very large sample. Our method PCR-SSCP is faster than PCR-RFLP, which needs at least 6 hours for the restriction-enzyme digestion step.

**Alzheimer and SNPs related to PS2 gene**

Approximately 50% of all familial forms of Alzheimer’s disease (FAD) cases are linked to the presenilin genes where missense mutations are generally found in residues that are conserved between the two proteins with the rare exceptions of in-frame splice deletions and premature truncations. The mechanisms by which mutations in PS and b APP genes cause AD are not known, although mutations in these genes appear somehow interconnected as they increase amyloidogenic Ab fragment accumulation (reviewed by Hardy, 1997). Thus far, at least 22 missense mutations have been identified in the PS1 gene and two missense mutations have been described in the PS2 gene in early-onset FAD. Although the primary structure of presenilins suggests that they are integral membrane proteins, their physiological functions remain presently unknown. It has been suggested that presenilin may participate in APP trafficking or processing and ultimately lead to the abnormal production of amyloid, which accumulates in senile plaques and degenerating neurons in AD.

Recently, a partial cDNA ALG-3, which is a mouse homologue of PS2, has been shown to rescue a T cell hybridoma from T receptor- and Fas-induced apoptosis. It is unclear, however, whether the full length PS2 cDNA has a similar or inverse function, because it is not known whether this partial cDNA ALG-3 conferring resistance to cell death is translated into a polypeptide. Indeed, neuronal death is a prominent feature of AD and several lines of evidence suggest that apoptosis occurs in AD. Thus, it is possible that presenilins may be associated with neuronal cell death.

To date, no data regarding the expression pattern of PS2 and its correlation with the pathology of AD are available. Therefore, the characterization of expression pattern of PS2 in AD brain is prerequisite to understanding the role of PS2 in sporadic AD. In the present study, nonradioactive in situ hybridization was used in combination with immunohistological techniques to examine PS2 mRNA expression and correlation with classical pathological lesions in sporadic AD and aged control brain.

**Materials and methods**

**cDNA cloning and probe preparation**

To obtain a human PS2 cDNA, 1 l.tl of aliquot from the human brain cDNA library (Stratagene) was amplified by polymerase chain reaction (PCR) with the following primers: forward GCCAAGAATTCCGTTGGTCTCAGAGGCA (containing the EcoRI restriction site) and reverse GCAGCTCAATCCCTTCAGCAGGCCA (containing the Xbal restriction site). PCR was performed under the following conditions: 94°C, 1 min; 60°C, 1 min; and 72°C, 1 min for 30 cycles with the Pwo DNA polymerase (Boehringer Mannheim). PCR products were
digested by *EcoRI* and *XbaI* and ligated to a pcDNA3 vector (Invitrogen). The positive insert was sequenced on double strands. For in situ hybridization, digoxigenin-labeled antisense and sense riboprobes were prepared using a Genius Kit (Boehringer Mannheim). Briefly, *EcoRI*- or *XhaI*-linearized PS2 cDNA in a pcDNA3 vector was transcribed in vitro with either SP6 or T7 RNA polymerase to generate antisense and sense probes. After transcription, template DNA was digested by RNasefree DNase and the digoxigenin-labeled transcripts were precipitated using ethanol.

**Preparation of tissue**

Six cases of clinically and neuropathologically defined AD (average age 74 years, average postmortem delay 4.5 h) and four age-matched control cases (average age 77 years, average postmortem delay 5 h) were used in this study. Frozen brain tissue from frontal cortex was fixed in 4% paraformaldehyde in 100 mM Tris buffer (pH 7.4) for 12 h and stored in 20% sucrose in 100 mM Tris buffer overnight until the tissue sunk. Fixed tissue was sectioned on a cryostat (30 μm) sections were collected in Tris buffer (100 mM, pH 7.4) and digested with proteinase K (Boehringer Mannheim) at a final concentration of 1 lag/ml in 10 mM Tris, pH 8.0 at 37°C for 20 min. This reaction was stopped using 100 mM glycine in 100 mM Tris buffer for 30 s, followed by a 4% paraformaldehyde for 10 min. Sections were washed in 100 mM Tris buffer for subsequent hybridization.

**In situ hybridization**

**Immunohistochemistry**

To determine whether PS2 gene expression had any correlation with pathological lesions in AD, double-labeling technique for PS2 gene expression and immune reactivity was employed. Briefly, subsequent to in situ hybridization by the above methods, sections were subject to immunostaining by the following antibodies. Two well-characterized monoclonal antibodies, AT8 (Innogenetics, Belgium, 1:5000) and PHF-1 (kindly provided by Dr. S. Greenberg; 1:200), were used to determine whether PS2 mRNA expression neurons were colocalized with neurofibrillary tangles (NFTs). Both antibodies recognize abnormal phosphorylation of tau protein without showing crossreactivity with normal tau (for review see: [17]). The AT8 antibody stains intracellular NFTs as well as pretangle neurons, whereas the PHF-1 antibody stains both intracellular NFTs and extracellular NFTs. Additionally, a polyclonal antibody (1342) directed against residues 1~12 of the 13-amyloid protein (1:200) was used to detect PS2 mRNA expression within plaques. After first labeling for PS2 mRNA, the sections were then double-labeled with AT8 and PHF-1 or 13A4 antibodies. Labeling for the second antigen was detected using FITC-conjugated IgG (green fluorescence, Sigma).

**Results**

**Distribution of PS2 mRNA in AD and control brain**

In the present study, the expression patterns of mRNA for human PS2 from frontal cortex of AD and aged control brain by in situ hybridization analysis revealed several findings. First, message for PS2 showed a laminar distribution in cerebral cortex and was primarily detectable in neurons. This result is consistent with previous reports from normal human temporal lobes for PS2 expression visualized by radioactive hybridization and is also comparable with PS1 expression pattern in murine brain observed using non-radioactive hybridization and immunohistochemistry. Second, stronger hybridization signal was found in
large neurons while mild or weak signal appeared in small neurons. It is well-established that the frontal cerebral cortex, one of the most vulnerable regions in AD, exhibited specific patterns of neuropathology in specific layers of cortex. NFTs are predominantly present in the cell bodies of neurons of layers III and V, although there are striking regional differences in the proportion of NFTs in these two layers and neuritic plaques predominantly affect layers II and III. In addition, large neurons in neocortical and hippocampal areas are particularly vulnerable to cell death or NFT formation, suggesting that either there are factors present in some neurons which protect them from neurodegeneration or that susceptible neuronal populations have some unique properties that lead to neurodegeneration. Our results showing higher PS2 expression in large neurons and within neuronal laminar layers (II, III and V) of cerebral cortex suggests that PS2 may participate in events that lead to neurodegeneration in AD via presently unknown pathways. It was recently reported that a partial cDNA ALG-3, which is a mouse homologue of PS2, rescues T cell hybridoma from T cell receptor- and Fas-induced apoptosis. Furthermore, programmed cell death occurs during neuronal development, aging and in the AD brain. Our previous immunohistochemical evidence showed that apoptotic changes are present in AD brain and the number of neuronal nuclei displaying the distinct morphology of apoptosis is greater in AD than in aged control brain. In the present study, no significant difference in PS2 mRNA expression in AD and control brain was observed. It seems that PS2 mRNA levels in AD do not correlate with DNA damage. Our explanations for this result are as follows. First, eventual development of sporadic pathology may result from the other alterations, such as the generation of specific variant transcripts and truncated transcripts of PS2 gene or its abnormal protein expression, rather than PS2 mRNA level changes. Indeed, previous reports have demonstrated that specific splice variants and truncates of PS1 transcripts were found in sporadic cases of A. Second, relative PS2 mRNA level may be influenced by postmortem which displays varying degrees of RNA degradation. This is unlikely however, since postmortem delays are short and matched across groups. Third, other AD susceptibility genes may contribute to AD. To define the exact correlation between DNA damage and PS2 expression, double staining for PS2 and DNA damage is needed. Our results suggest that there is no simple relationship between PS2 expression and tangle and plaque formation. Sequence comparison of presenilins and Spe 4 has lead to the hypothesis that presenilins may have some role in 13-amyloid trafficking or processing. There is some evidence supporting this hypothesis. For example, a fibroblast cell line isolated from a patient with PS1 defect produced abnormally high 13-amyloid and PS1 protein appeared in AD neuritic plaques. Indeed 13-amyloid is able to induce apoptosis in primary culture neurons. The pathological function(s) of PS2 and its relationship with 13-amyloid remains to be determined. PS-2 as suggested by homology to ALG-3 may play a role in neuronal vulnerability to various stimuli that can initiate neuronal damage and eventually program cell death. Lipofuscin accumulation has been considered to be age-related, morphological marker of nervous system aging. Lipofuscin autofluorescent granules were widely observed in brain regions such as cerebral cortex and hippocampus in aging rat and human, including AD tissue. It has been suggested that lipofuscin deposits represent the accumulation of the peroxidative action of free radicals on membranes. A significant colocalization of PS2 mRNA expression and lipofuscin-related autofluorescence in a large subset of neurons from cerebral cortex of AD and aged control brain suggests that PS2 may have some relationship with lipofuscin deposits via some as yet unknown pathway such as lipid peroxidation of neuronal membrane. Their precise relationships between PS2 and lipofuscin deposits needs to be further investigated.[18]
Missense Mutations in APP: A Very Rare Cause of Familial AD

APP is a single transmembrane polypeptide that is cotranslationally translocated into the endoplasmic reticulum via its signal peptide and then posttranslationally modified (“matured”) through the secretory pathway. Its acquisition of N- and O-linked sugars occurs rapidly after biosynthesis, and its half-life is relatively brief (45–60 min in most cell types tested). Both during and after the trafficking of APP through the secretory pathway, it can undergo a variety of proteolytic cleavages to release secreted derivatives into vesicle lumens and the extracellular space.

The first proteolytic cleavage identified, that made by an activity designated \( \alpha \)-secretase, occurs 12 amino acids NH\(_2\)-terminal to the single transmembrane domain of APP. This processing results in the release of the large soluble ectodomain fragment (\( \alpha \)-APPs) into the lumen/extracellular space and retention of an 83-residue COOH-terminal fragment (CTF) in the membrane. Alternatively, some APP molecules not subjected to \( \alpha \)-secretase cleavage can be cleaved by an activity designated \( \beta \)-secretase. The first specific genetic cause of AD to be identified was the occurrence of missense mutations in APP. Despite extensive genetic surveying, such mutations have only been confirmed in some two dozen or so families worldwide. Nevertheless, the location of the mutation and the subsequent delineation of their genotype-to-phenotype relationships have provided critical insights into the mechanism of AD. The mutations are strategically located either immediately before the \( \beta \)-secretase cleavage site, shortly after the \( \alpha \)-secretase site, or shortly COOH-terminal to the \( \gamma \)-secretase cleavage site. The fact that, despite substantial investigation, no other mutations in the large APP protein that cause AD have been discovered strongly suggests that these missense mutations lead to AD by altering proteolytic processing at the three secretase sites in subtly different ways. This hypothesis has been confirmed by analysis of each of the mutations, initially in transfected cells or primary cells from patients and then in transgenic mouse models. Families harboring APP missense mutations that cause AD generally have the onset of the disorder before age 65, often in their 50s. There is another way that alterations in the APP gene can predispose to the development of AD. The overexpression of structurally normal APP owing to elevated gene dosage in trisomy 21 (Down’s syndrome) almost invariably leads to the premature occurrence of classical AD neuropathology (neuritic plaques and neurofibrillary tangles) during middle adult years. A life-long increase in APP expression due to duplication of all of chromosome 21 or, in the case of translocation Down’s syndrome, that portion of 21q containing the APP gene results in overproduction of Ab\(_{40}\) and Ab\(_{42}\) peptides dating from birth. This is assumed to be responsible for the strikingly early appearance of many Ab\(_{42}\) diffuse plaques, which can occur as soon as age 12 yr. Down’s subjects often display diffuse plaques composed solely of Ab\(_{42}\) in their teens and 20s, with accrual of Ab\(_{40}\) peptides onto these plaques and the appearance of associated microgliosis, astrocytosis, and surrounding neuritic dystrophy usually beginning in their late 20s or 30s. This observation underscores the importance of Ab\(_{42}\) accumulation as a seminal event in the development of AD-type brain pathology.
Figure 7: b-APP mutations genetically linked to familial Alzheimer’s disease or related disorders. The sequence within APP that contains the Ab and transmembrane region is expanded and shown by the single-letter amino acid code. The underlined residues represent the Ab1–42 peptide. The vertical broken lines indicate the location of the transmembrane domain. The bold letters below the line indicate the currently known missense mutations identified in certain patients with familial Alzheimer’s disease and/or hereditary cerebral hemorrhage with amyloidosis. Three-digit numbers refer to the residue number according to the b-APP770 isoform.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gene Defect</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>b-APP mutations</td>
<td>↑ Production of all Ab peptides or Aβ42 peptides</td>
</tr>
<tr>
<td>15</td>
<td>ApoE4 polymorphism</td>
<td>↑ Density of Ab plaques and vascular deposits</td>
</tr>
<tr>
<td>14</td>
<td>Presenilin 1 mutations</td>
<td>↑ Production of Aβ23 peptides</td>
</tr>
<tr>
<td>1</td>
<td>Presenilin 2 mutation</td>
<td>↑ Production of Aβ39 peptides</td>
</tr>
</tbody>
</table>

b-APP, β-amyloid precursor protein; Aβ, amyloid β-protein.

Figure 8: the relation between the chromosomes and mutations and their affect on phenotype

APP duplication is sufficient to cause early onset Alzheimer’s dementia with cerebral amyloid angiopathy

Evidence is accumulating that an increase in genetic expression of the amyloid precursor protein (APP) gene, encoding the amyloid b precursor protein [MIM 104760], can cause dementia of the Alzheimer type [MIM 104300] through increased production of its pathogenic 42 amino acids proteolysis product, amyloid b42 (Ab42). Not only missense mutations at or near the a-, b- and g-secretase cleavage sites of APP and in the presenilins (Dermaut et al., 2005) but also increased expression of APP per se can cause an increase in concentration of Ab42 and subsequent Ab deposition. It has long been recognized that triplication of APP in patients with Down’s syndrome leads to Alzheimer’s dementia symptoms early in life through-over-expression of APP (Rumble et al., 1989), followed by
deposition of Ab and neurodegeneration (Wisniewski et al., 1985). In a recent study, genomic duplications in the APP locus were reported in families segregating early onset Alzheimer’s dementia with concurrent cerebral amyloid angiopathy (CAA) (Rovelet-Lecrux et al., 2006), suggesting that increased expression of APP can give rise to Alzheimer dementia pathology in the absence of a full trisomy 21. Along this line, previous studies have shown higher levels of APP mRNA in brains of patients with Alzheimer-type (Rovelet-Lecrux et al., 2006), suggesting that increased expression of APP can give rise to Alzheimer dementia pathology in the absence of a full trisomy 21. Along this line, previous studies have shown higher levels of APP mRNA in brains of patients with Alzheimer-type dementia (Cohen et al., 1988; Higgins et al., 1988; Lewis et al., 1988; Clark and Parhad, 1989; Vitek, 1989; Theuns and Van Broeckhoven, 2000) also suggesting that genetic variation in APP transcription might play a role in the pathomechanism of the disease. Indeed, in patients with early onset Alzheimer’s dementia, we recently detected three mutations in the APP proximal promoter that caused a neuron-specific, nearly 2-fold increase in APP transcriptional activity in vitro, mimicking over-expression of APP as observed in trisomy 21 (Theuns et al., 2006). With the evidence that an increase in APP expression through APP genomic duplication or mutations in the APP promoter leads to early onset Alzheimer’s disease, the hypothesis that Ab plays a pivotal role in the disease aetiology gained strength. To assess the impact of APP locus duplications in early onset Alzheimer’s disease, we performed a study in a Dutch population-based sample of patients with early onset Alzheimer-type dementia.

Patients

For genetic studies, 111 patients were available with a mean onset age of 56.1 years (range = 33–65), and a positive family history (i.e. at least one first-degree relative affected with either early or late-onset dementia) in 75 patients, of which 10 were probands of multigenerational families with an autosomal dominant inheritance. The criteria for autosomal dominant inheritance were (i) at least three patients with clinically diagnosed Alzheimer’s disease in two or more generations and (ii) detailed medical records available on the clinical diagnosis of Alzheimer’s disease in at least two affected relatives.

Real-time polymerase chain reaction (PCR) allele quantification

Genomic DNA (gDNA) was extracted from peripheral lymphocytes. APP alleles were quantified using a SYBR Green real-time PCR assay on the ABI Prism 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA). Primers for exons 5, 11 and 18 and the promoter region of APP, for GABPA and ATP5J and for ubiquitin C (hUBC) and b2-microglobulin (hB2M) (one amplicon each) were designed with PrimerExpress software (Applied Biosystems). Primer sequences are available upon request. Human APP alleles were normalized for the housekeeping genes hUBC and hB2M. Twenty nanograms of gDNA were amplified in a 30 ml reaction containing 1· qPCR Mastermix Plus for SYBR Green I (without UNG) (Eurogentec, Seraing, Belgium) and 400 nM of the respective forward and reverse primers using the universal amplification protocol (Applied Biosystems). Duplicate samples were quantified for patients and six cognitively healthy age-matched individuals to calculate dosage quotients (DQs). Samples of two patients with full trisomy 21 were included as positive controls.
Fluorescence in situ hybridization (FISH)

To confirm the presence of a genomic duplication of APP, FISH of both interphase nuclei and mechanically stretched metaphase chromosomes was performed, the latter permitting detection of duplications at an increased resolution (>100 kb) (Haaf and Ward, 1994; Laan et al., 1995).

Multiplex amplicon quantification

To detect and delineate sizes of APP locus duplications, an in-house-developed technique for multiplex amplicon quantification (MAQ) was used (D. Goossens, in preparation), consisting of a multiplex PCR amplification of several fluorescently labelled target and reference amplicons, followed by fragment analysis on an ABI 3730 DNA analyser (Applied Biosystems). The comparison of normalized peak areas between a patient and control individuals results in a DQ of the target amplicon. Five fluorescently labeled amplicons in APP and 17 labelled amplicons in 11 surrounding genes (FLJ42200, c21orf42, MRPL39, JAM2, CYYR1, ADAMTS1, ADAMTSS, USP16, CCT8, BACH1 and GRIK1) were simultaneously amplified with 15 reference fragments randomly located on different chromosomes, using 50 ng gDNA and optimized reaction conditions (available upon request). DQs were calculated using an in-house-developed MAQ software (MAQs) package. A DQ between 1.3 and 1.75 was considered indicative of a heterozygous duplication.

Real-time PCR quantification of APP in 10 probands revealed a DQ > 1.5 for all three APP exons 5, 11 and 18 in the proband of one family, 1104, pointing to a heterozygous duplication of APP. Family 1104 is a four-generation pedigree including 10 relatives (77.8% female) with dementia, and a mean onset age of 52.1 years [standard deviation (SD) = 7.9; range = 40–62] and mean disease duration of 8.2 years (SD = 4.7).

In this Dutch population-based sample of early onset dementia we identified an APP duplication in a family with a segregation pattern compatible with autosomal dominant inheritance, and neuropathology compatible with Alzheimer’s disease with extensive CAA. Our findings are in agreement with the recent observation of APP locus duplications in French families with autosomal dominant early onset dementia of the Alzheimer type and concurrent CAA (Rovelet-Lecrux et al., 2006). In this study a genomic duplication in the APP locus was present in 5 out of 65 families (7.8%) with autosomal dominant early onset Alzheimer’s disease. We detected the APP duplication in 1 out of 10 (10%) of the autosomal dominant early onset Alzheimer’s disease families in our sample. Although numbers are small, together these data indicated that investigation of genomic duplications in the APP locus in multiplex early onset families is warranted when simple mutations in known Alzheimer’s disease genes have been excluded. Fast and reliable screening for these genomic duplications is now feasible owing to the development of our MAQ assay, which not only allows detection of a genomic duplication but also provides information of the size of the duplication and genes involved. In 65 patients with a positive family history, we detected one other similar sized genomic duplication. In this Dutch population-based sample, genomic duplications occurred at a frequency of 2/111 (1.8%) overall, and at a frequency of 2.7% (2/75) in familial Alzheimer’s disease patients. With a frequency of 10% in multiplex early onset families, the frequency increases with increasing evidence of a genetic background. In 36 patients with sporadic early onset Alzheimer’s disease, we did not detect duplications, suggesting that de novo genomic APP duplications are unlikely to be a frequent cause of early onset Alzheimer-type dementia.[20]
Cancer and PCR

Quantitative real-time PCR for cancer detection

The sequence for most of the human genome is now publicly available and can be applied to understand, characterize, and treat complex diseases such as cancer. Normal cell regulation is affected by factors such as viral infections, DNA methylation and sequence alterations. Cancers may develop when these changes affect genes controlling cell division, cell repair, growth control and apoptosis. Advances in the biological sciences and technology are providing molecular targets for diagnosis and treatment of cancer. Gene profiling opens new possibilities to classify the disease into subtypes and guide a differentiated treatment. Real-time PCR is characterized by high sensitivity, excellent precision and large dynamic range, and has become the method of choice for quantitative gene expression measurements. For accurate gene expression profiling by real-time PCR, several parameters must be considered and carefully validated. Real-time quantitative PCR is a homogeneous method that includes both amplification and analysis with no need for slab gels, radioactivity, or sample manipulation. There are now several platforms commercially available for combining thermal cycling with fluorescence acquisition (Committee on DNA Technology in Forensic Science, National Research Council, DNA Technology in Forensic Science (1992)). The fluorescence of DNA dyes or probes is monitored each cycle during PCR. At a certain point during cycling, the product accumulates enough to increase fluorescence above background. The point where fluorescence rises above background noise is best quantified as the second derivative maximum [crossing point (Cp)] of the curve and correlates to the amount of starting copies within a PCR reaction. As the number of initial template copies increases, fluorescence appears sooner and the Cp is lower. The relative copy number between two samples (experimental and control) can be determined by the difference in their Cp values (Fig. 9A). Because PCR is an exponential process, the relative copy number is equal to the PCR efficiency raised to the power, ΔCp. Because it may be difficult to know the total amount of DNA present in different samples, results of the test gene are often normalized to a reference gene presumed to be invariant (Fig. 9B).

![Fig. 9. Relative quantification by real-time PCR.](image-url)
(A), the amounts of target in an experimental sample and a control sample are compared after PCR amplification and fluorescence monitoring each cycle. For example, genomic DNA may be analyzed to assess gene amplification or deletion. Expression of mRNA may also be studied after reverse transcription. The sample with the greater amount of DNA (or cDNA) will show an earlier increase in fluorescence. The second derivative maxima of the curves (vertical dotted lines) are determined as fractional cycle numbers. The relative copy number between samples is the PCR efficiency (eff) raised to the difference between fractional cycle numbers (ΔC). The calculation assumes that the PCR efficiency is the same between samples. The PCR efficiency is usually between 1.7 and 2.0. As a first approximation, an efficiency of 2 is often assumed. This analysis assumes that the starting amount of material (DNA or cDNA) in each sample is the same. (B), another option is to use a test target normalized to a reference target. The amount of starting material in each sample is normalized to a reference (Ref) or housekeeping gene. Both experimental and control samples are amplified for both the test and reference targets. Any difference in the amount of starting material is normalized by the results of the reference target amplification. This method assumes that the reference target is invariant between samples and that the PCR efficiency for each target does not vary between samples. As a first approximation, an efficiency of 2 is often assumed for both targets and has become known as the ΔC method.

Regions of genomic stability (i.e., not altered) for a given tumor type can be identified by CGH studies and can be used for DNA controls. An inexpensive and common method to validate microarray experiments is relative quantification by real-time PCR using SYBR Green I as a fluorescent indicator of double stranded DNA production (H Nailis, T Coenye, F Van Nieuwerburgh, D Deforce, HJ Nelis, Development and evaluation of different normalization strategies for gene expression studies in Candida albicans biofilms bu real time PCR, BMC Mol Biol.7:25, (2006)). The correlation between Cp and starting copy and the use of these values to calculate PCR amplification efficiency are shown in panels A and B, respectively, of Fig. 10. The PCR amplification curves were generated with hybridization probes specific for the serum albumin gene. Genomic leukocyte DNA was used at decreasing concentrations in a series of 10-fold dilutions. The fluorescence was acquired every cycle and then plotted against cycle number to construct an amplification curve at each dilution (Fig. 10A). Linear regression through a plot of the Cps at each dilution against the log of genomic DNA gives the average efficiency of the PCR reaction (Fig. 10B). This is calculated using the formula:

Efficiency = 10 ^ -1/slope

![Figure 10: establishing PCR efficiency](image)
(A), the overall efficiency of PCR within a run can be found by use of a dilution series of DNA (e.g., leukocyte) to generate amplification curves at different concentrations of template. (B), the overall efficiency of PCR for a given primer and probe set can be calculated from the linear regression through a plot of Cp vs log ng DNA. For albumin, the slope of the trend line was -3.52, yielding an efficiency of 1.92.

**Fig. 11:** Determining efficiency and gene copy numbers using three-color multiplexing.

There is greater variation in PCR efficiency between runs than within a run. We established the efficiency of amplification for three targets within a run using color multiplexing and used the calculated efficiencies to determine gene amplification in breast tumors. After color compensation and correction for differences in PCR efficiency, Cps of HER-2, topo IIα, and albumin were compared in healthy leukocytes and breast tumor cells. All three genes had nearly the same Cp for wild-type leukocyte DNA (A). However, the DNA from the breast tumor cells showed an early three-cycle shift for both HER-2 and topo IIα, corresponding to an eightfold amplification over albumin (B).

In conclusion, real-time PCR will play an increasingly important role in clinical testing because it can provide information about gene expression, gene amplification or loss, and small alterations (e.g. point mutations). Moreover, it can be applied to detect and quantify viral causes of cancer, such as Epstein- Barr virus and human papillomavirus. The use of real-time PCR for molecular diagnostics is attractive because it is objective, rapid, versatile and cost effective and can be performed on small tissue samples.
Efforts to detect cancer by specific PCR methods

Detection of prostate cancer using methylation-specific PCR

Prostate cancer is the most common cancer in men in western societies. Early detection is essential for effective treatment. Commonly used detection methods are rectal examination and measurement of prostate-specific antigen (PSA) in serum. However, determination of PSA levels alone is not sensitive or specific enough for a definite diagnosis of prostate cancer. Development of an effective molecular method could provide earlier diagnosis of prostate cancer.

For that reason, Methylation-specific PCR (MSP) technology is used in order to detect GSTP1 hypermethylation. The most frequent genetic alteration in prostatic carcinoma is hypermethylation of the glutathione S-transferase P1 (GSTP1) promoter. GSTP1 is involved in intracellular detoxification reactions and its hypermethylation results in loss of gene expression. It is a candidate tumor suppressor gene in prostate cancer, although its inactivation by hypermethylation has not been proven to have a causative role in the disease. However, hypermethylation has been found in >90% of prostatic carcinomas, including early disease stages, and has not been detected in normal tissues or in benign hyperplastic prostatic tissue. GSTP1 is therefore a potentially useful marker for detection and molecular staging of prostate cancer. Both intracellular and extracellular tumor DNA has been detected in the bodily fluids of cancer patients. We therefore isolated DNA from a range of clinical samples from both prostate cancer patients and control individuals with benign prostatic hyperplasia, and investigated the methylation status of the GSTP1 gene promoter by methylation-specific PCR. The QIAamp DNA Blood Mini Kit and QIAamp DNA Mini Kit were used for DNA isolation, and MSP was carried out using HotStarTaq™ DNA polymerase from QIAGEN. The use of QIAGEN® systems resulted in the development of a highly sensitive and specific MSP assay. Malignant cells were detected not only in tissue and ejaculate samples, but also in plasma and serum, which contain much lower amounts of DNA.

Methylation-specific PCR (MSP) technology

In higher-order eukaryotes, DNA is methylated only at certain cytosines located 5' to a guanosine. This occurs especially in GC-rich regions, known as CpG islands. To distinguish the methylation state of a sequence, MSP relies on differential chemical modification of cytosine residues in DNA. Treatment with sodium bisulfite converts unmethylated cytosine residues into uracil, leaving the methylated cytosines unchanged. This modification thus creates different DNA sequences for methylated and unmethylated DNA. PCR primers can then be designed so as to distinguish between these different sequences. Two sets of primers are designed: one set with sequences annealing to unchanged (methylated in the genomic DNA) cytosines and the other set with sequences annealing to the altered (unmethylated in the genomic DNA) cytosines. A comparison of PCR results using the two sets of primers reveals the methylation state of the DNA. If the primer set with the altered sequence gives a PCR product, then the indicated cytosine was unmethylated. If the primer set with the
unchanged sequence gives a PCR product, then the cytosines were methylated and thus protected from alteration.

Results and discussion
Tumor DNA in body fluids is always accompanied by normal DNA from nonmalignant cells. We therefore tested the sensitivity of the MSP assay by diluting LNCaP cells into blood samples from healthy donors. DNA was isolated from 200 μl buffy coat, and MSP detected 200 prostate cancer cells among 2.2 x 10^7 nonmalignant leukocytes. Of the samples tested, GSTP1 promoter hypermethylation was found in 94% of tissue samples, 30% of buffy coat samples, 72% of plasma or serum samples, and 50% of ejaculates. In contrast, GSTP1 hypermethylation was not found in any samples tested from the 22 BPH patients. The MSP assay was therefore 100% specific. Since hypermethylation of the GSTP1 promoter is a cancer-specific event, this assay might avoid the false-positive signals sometimes found in RT-PCR–based techniques. The assay successfully detected hypermethylation of the GSTP1 promoter, not only in tissue samples from patients with prostate cancer, but also in a high proportion of serum and plasma samples, which contain much lower amounts of DNA. [21].

Quantitative real-time PCR for cancer detection: the lymphoma case
Quantitative real-time PCR detection involves several steps prior to the actual real-time PCR measurement. These include sample collection, nucleic acids isolation, reverse transcription (RT) and sometimes DNase treatment. For genomic analysis, only the first two steps are needed.

Real-time PCR
Through real-time PCR, the amount of any DNA target in a biologic sample can be determined with unprecedented accuracy over an essentially unlimited dynamic range. At a certain point during cycling, enough product accumulates to increase the fluorescence signal significantly above the background noise. This point, typically defined by an arbitrarily set threshold line or by the second derivative maximum of the amplification response curve, is referred to as the cycle of threshold (Ct) or crossing point (Cp). A typical real-time PCR amplification curve is shown in Figure 12. The Ct value relates to the number of initial template molecules as:

\[ N_{Ct} = N_0 \cdot (1 + E)^{Ct} \]

where N_{Ct} is the number of double-stranded DNA molecules after Ct amplification cycles, N_0 is the initial number of double-stranded target molecules, and E is the PCR efficiency. E is assumed to be independent of N_0 over the studied concentration range, and can be estimated from dilution series of mRNA, cDNA or genomic DNA, or from the real-time PCR amplification response curve (Figure 12).
Relative gene expression analysis & normalization

Extensive evidence indicates that all genes are regulated under some conditions. The expression of many genes is often out of control, particularly in tumor tissue. The field may have to face up to the fact that no universal reference gene with a constant expression in all tissues exists. Due to this uncertainty, any system relying on reference genes should be carefully validated. Generally finding appropriate reference genes for data normalization is currently one of the most challenging problems. Appropriate reference genes may be identified by the GeNorm software.

Quantitative gene expression analysis is most commonly performed as a relative measurement between two genes. The expression ratio is given by:

\[
\frac{N_A}{N_B} = K_{RS} \frac{\eta_B (1 + E_B)^{C_{T_B} - 1}}{\eta_A (1 + E_A)^{C_{T_A} - 1}}
\]

where \(N_i\) is the initial number of mRNA molecules of gene \(i\), \(K_{RS}\) is the relative sensitivity of the detection chemistries of the two assays, and \(\eta_i\) is the cDNA synthesis yield of gene \(i\), defined as the fraction of mRNA molecules that are transcribed to cDNA in the RT reaction. The exponent Ct-1 accounts for the production of double-stranded DNA in the first PCR cycle from the single-stranded cDNA template generated by the RT reaction. Typically \(\eta\) is assumed to be independent of both total and target mRNA concentrations. In most applications the expression ratio of two genes is compared in two or more samples (i.e., comparative quantification). Typically, one is the reporter gene whose expression is expected to be affected by the physiological change studied and the other is a reference gene whose expression should be constant. Assuming the same RT yields in the samples, \(K_{RS}\) and \(\eta\) are eliminated, and the comparative expression ratio of the two samples, assuming the same PCR efficiencies in the two samples and 100% PCR efficiency is given by:

\[
\frac{\text{Sample}_1}{\text{Sample}_2} = 2^{\frac{(C_{T_{B1}} - C_{T_{B2}})}{2(C_{T_{A1}} - C_{T_{A2}})}} = 2^{(C_{T_{B1}} - C_{T_{B2}})-(C_{T_{A1}} - C_{T_{A2}})} = 2^{\Delta \Delta Ct}
\]

which is commonly referred to as the \(\Delta \Delta Ct\)-method.

Due to the exponential nature of PCR, errors caused by inhibition accumulate with the product. The common procedure to account for any differences in PCR efficiencies between samples is to amplify a reference gene in parallel to the reporter gene and relate their expression levels. This approach assumes that the two assays are inhibited to the same degree. However, this may not always be the case, particularly not for complex tissue samples, and the procedure, although very common, may lead to false classification of sample.
The response curve may be divided into three regions. The first part is characterized by background noise, where the target specific florescence is too low to be detected. The second part is the exponential phase where the fluorescence rises significantly above the noise level and is proportional to the amount of product formed. Cycle of threshold values should be determined in the exponential phase where extrapolation to determine initial template amount is most reliable. The exponential region is readily identified by plotting the fluorescence signal in logarithmic scale. Eventually, the PCR runs out of chemicals and the response curve enters the plateau phase.

Cancer detection: the lymphoma case

Diffuse large B-cell lymphoma (DLBCL), which is one of the most common subtypes of non-Hodgkin’s lymphoma, has received particular attention due to its highly diverse clinical courses. Based on microarray data 36 genes has been selected, whose expression may be used to predict survival in DLBCL patients. Expression of these genes was measured with quantitative real-time PCR in 66 patients. Out of the 36 genes, six (LMO2, BCL6, FN1, CCND2, SCYA3 and BCL2) were found to be strong predictors for survival. Other potential lymphoma markers that have recently been validated by quantitative real-time PCR include CCND1, CDC37, glutathione-S-transferase pi (GST-pi) and the cytokines. However, no gene has yet been identified whose expression is constant in all lymphoma samples.

Conclusions

The high sensitivity of real-time PCR makes it possible to detect even a single molecule. Diagnostics becomes feasible with much lower amounts of biologic material than what is required by traditional methods. Fine needle aspirates may replace traditional surgical biopsies in many situations. This reduces patient trauma and also allows many more samples to be collected. Despite the high reproducibility and sensitivity of real-time PCR, robust protocols are needed for single cell analysis before the technique can be used as a standard
tool in cancer research. A typical single cell contains only 0.5–1.0 pg mRNA, which accounts to approximately 300,000 mRNA transcripts. This makes single cell real-time PCR analysis sensitive to contamination and statistical uncertainty. Before setting up quantitative real-time PCR-based assays for routine diagnostic purposes, some important parameters should be carefully validated. Most importantly, the reporter genes used for prediction must be statistically validated on a substantial number of patient samples that are characterized and classified by traditional methods. The main technical obstacle in real-time PCR is inhibition caused by unpurified interfering components from the biologic sample. In most cases, this can be accounted and corrected for by proper normalization of all experimental steps including sample collection, nucleic acid extraction, RT and real-time PCR. In addition to gene expression profiling, real-time PCR is also useful to detect chromosomal aberrations. Most chromosomal translocations in non-Hodgkin’s lymphomas can be detected by genomic PCR. However, due to the very large spread of chromosomal breakpoints, only those in narrow regions can in practice be tested for in routine diagnostics. Compared with quantitative gene expression measurements, which require real-time PCR, specific translocations can usually be detected by conventional genomic PCR against zero background.[22]

Applications of PCR in genetic analysis

Genetic testing allows the genetic diagnosis of vulnerabilities to inherited diseases, and can also be used to determine a person’s ancestry. Normally, every person carries two copies of every gene, one inherited from their mother, one inherited from their father. The human genome is believed to contain 20,000-25,000 genes. In addition to studying chromosomes to the level of individual genes, genetic testing in a broader sense includes biochemical tests for the possible presence of genetic diseases, or mutant forms of genes associated with increased risk of developing genetic disorders. Genetic testing identifies changes in chromosomes, genes, or proteins. Most of the time, testing is used to find changes that are associated with inherited disorders. The results of a genetic test can confirm or rule out a suspected genetic condition or help determine a person’s chance developing or passing on a genetic disorder. Several hundred genetic tests are currently in use, and more are being developed.

Since genetic testing may open up ethical or psychological problems, genetic testing is often accompanied by genetic counseling.

Genetic testing companies

There are many companies that claim to be developing new methods and technologies that will enable consumers to understand their own genetic information. As we have mentioned, the types of genetic tests are the following: Newborn screening, diagnostic testing, carrier testing (carrier testing is used to identify people who carry one copy of a gene mutation that, when present in two copies, causes a genetic disorder.), Prenatal testing, preimplantation testing, predictive and presymptomatic testing, forensic testing. Some of the most known companies that offer services for genetic analysis are:
Genetic testing has potential benefits whether the results are positive or negative for a gene mutation. Test results can provide a sense of relief from uncertainty and help people make informed decisions about managing their health care. For example, a negative result can eliminate the need for unnecessary checkups and screening tests in some cases. A positive result can direct a person toward available prevention, monitoring, and treatment options. Some test results can also help people make decisions about having children. Newborn screening can identify genetic disorders early in life so treatment can be started as early as possible.

**Types**

Genetic testing is « the analysis of RNA, chromosomes (DNA), proteins, and certain metabolites in order to detect heritable disease-related genotypes, mutations, phenotypes, or karyotypes for clinical purposes.» (Holtzman & Watson 1997). It can provide information about a person’s genes and chromosomes throughout life. Available types of testing include:

- **Newborn screening**: Newborn screening is used just after birth to identify genetic disorders that can be treated early in life. The routine testing of infants for certain disorders is the most widespread use of genetic testing—millions of babies are tested each year in the United States. All states currently test infants for phenylketonuria (a genetic disorder that causes mental retardation if left untreated) and congenital hypothyroidism (a disorder of the thyroid gland).
• **Diagnostic testing:** Diagnostic testing is used to diagnose or rule out a specific genetic or chromosomal condition. In many cases, genetic testing is used to confirm a diagnosis when a particular condition is suspected based on physical mutations and symptoms. Diagnostic testing can be performed at any time during a person’s life, but is not available for all genes or all genetic conditions. The results of a diagnostic test can influence a person’s choices about health care and the management of the disease.

• **Carrier testing:** Carrier testing is used to identify people who carry one copy of a gene mutation that, when present in two copies, causes a genetic disorder. This type of testing is offered to individuals who have a family history of a genetic disorder and to people in ethnic groups with an increased risk of specific genetic conditions. If both parents are tested, the test can provide information about a couple’s risk of having a child with a genetic condition.

• **Prenatal testing:** Prenatal testing is used to detect changes in a fetus’s genes or chromosomes before birth. This type of testing is offered to couples with an increased risk of having a baby with a genetic or chromosomal disorder. In some cases, prenatal testing can lessen a couple’s uncertainty or help them decide whether to abort the pregnancy. It cannot identify all possible inherited disorders and birth defects, however.

• **Predictive and presymptomatic testing:** Predictive and presymptomatic types of testing are used to detect gene mutations associated with disorders that appear after birth, often later in life. These tests can be helpful to people who have a family member with a genetic disorder, but who have no features of the disorder themselves at the time of testing. Predictive testing can identify mutations that increase a person’s chances of developing disorders with a genetic basis, such as certain types of cancer. For example, an individual with a mutation in BRCA1 has a 65% cumulative risk of breast cancer.[23]

• **Forensic testing:** Forensic testing uses DNA sequences to identify an individual for legal purposes. Unlike the tests described above, forensic testing is not used to detect gene mutations associated with disease. This type of testing can identify crime or catastrophe victims, rule out or implicate a crime suspect, or establish biological relationships between people (for example, paternity).

• **Research testing:** Research testing includes finding unknown genes, learning how genes work and advancing our understanding of genetic conditions. The results of testing done as part of a research study are usually not available to patients or their healthcare providers.

**Medical procedure**

Genetic testing is often done as part of a genetic consultation. Once a person decides to proceed with genetic testing, a medical geneticist, genetic counselor, primary care doctor, or specialist can order the test after obtaining informed consent.
Genetic tests are performed on a sample of blood, hair, skin, amniotic fluid (the fluid that surrounds a fetus during pregnancy), or other tissue. For example, a medical procedure called a buccal smear uses a small brush or cotton swab to collect a sample of cells from the inside surface of the cheek. The sample is sent to a laboratory where technicians look for specific changes in chromosomes, DNA, or proteins, depending on the suspected disorder. The laboratory reports the test results in writing to a person’s doctor or genetic counselor.

Routine screening tests are done on a small blood sample, using a ‘heel stick’, by pricking the baby’s heel and blotting the blood on a special paper. Unlike other types of genetic testing, a parent will usually only receive the result if it is positive.

Interpreting results

The results of genetic tests are not always straightforward, which often makes them challenging to interpret and explain. When interpreting test results, healthcare professionals consider a person’s medical history, family history, and the type of genetic test that was done.

A positive test result means that the laboratory found a change in a particular gene, chromosome, or protein of interest. Depending on the purpose of the test, this result may confirm a diagnosis, indicate that a person is a carrier of a particular genetic mutation, identify an increased risk of developing a disease (such as cancer) in the future, or suggest a need for further testing. Because family members have some genetic material in common, a positive test result may also have implications for certain blood relatives of the person undergoing testing. It is important to note that a positive result of a predictive or presymptomatic genetic test usually cannot establish the exact risk of developing disorder. Also, health professionals typically cannot use a positive test result to predict the course or severity of a condition.

A negative test result means that the laboratory did not find a dangerous copy of the gene, chromosome, or protein under consideration. This result can indicate that a person is not affected by a particular disorder, is not a carrier of a specific genetic mutation, or does not have an increased risk of developing a certain disease. It is possible, however, that the test missed a disease-causing genetic alteration because many tests cannot detect all genetic changes that can cause a particular disorder. Further testing may be required to confirm a negative result.

In some cases, a negative result might not give any useful information. This type of result is called uninformative, indeterminate, inconclusive, or ambiguous. Uninformative tests results sometimes occur because everyone has common, natural variations in their DNA, called polymorphisms that do not affect health. If a genetic test finds a change in DNA that has not been associated with a disorder in other people, it can be difficult to tell whether it is a natural polymorphism or a disease-causing mutation. An uninformative result cannot confirm or rule out a specific diagnosis, and it cannot indicate whether a person has an increased risk of developing a disorder. In some cases, testing other affected and unaffected family members can help clarify this type of result.
Risks and limitations

The physical risks associated with most genetic tests are very small, particularly for those tests that require only a blood sample or buccal smear (a procedure that samples cells from the inside surface of the cheek). The procedures used for prenatal testing carry a small but real risk of losing the pregnancy (miscarriage) because they require a sample of amniotic fluid or tissue from around the fetus. Many of the risks associated with genetic testing involve the emotional, social, or financial consequences of the test results. People may feel angry, depressed, anxious, or guilty about their results. In some cases, genetic testing creates tension within a family because the results can reveal information about other family members in addition to the person who is tested. The possibility of genetic discrimination in employment or insurance is also a concern. Some individuals avoid genetic testing out of fear it will impede their ability to purchase insurance or find a job [24]. Health insurers do not currently require applicants for coverage to undergo genetic testing, and when insurers encounter genetic information, it is subject to the same confidentiality protections as any other sensitive health information.[25]. Legislation in the United States called the Genetic Information Non-discrimination Act prohibits group health plans and health insurers from denying coverage to a healthy individual or charging that person higher premiums based solely on a genetic predisposition to developing a disease in the future. The legislation also bars employers from using individual genetic information when making hiring, firing, job placement, or promotion decisions.[26] It was signed into law by the President on May 21, 2008.[27],[28]. Genetic testing can provide only limited information about an inherited condition. The test often can’t determine if a person will show symptoms of a disorder, how severe the symptoms will be, or whether the disorder will progress over time. Another major limitation is the lack of treatment strategies for many genetic disorders once they are diagnosed. A genetics professional can explain in detail the benefits, risks, and limitations of a particular test. It is important that any person who is considering genetic testing understand and weigh these factors before making a decision. Many people are also concerned about the privacy implications of genetic testimony. In the United States, federal law requires that this kind of medical information to be kept confidential.

Genetic testing in a shoe-box

Canadian scientists have succeeded in building the least expensive portable device for rapid genetic testing ever made.[29]

The cost of carrying out a single genetic test currently varies from hundreds to thousands of pounds, and the wait for results can be weeks. Now a group led by Christopher Backhouse,
University of Alberta, Edmonton, have developed a re-useable microchip-based system that costs just £500 to build, is small enough to be portable, and can be used for point-of-care medical testing.

The well-known techniques reverse transcription, polymerase chain reaction and capillary electrophoresis have been developed over recent decades to take tiny amounts of genetic material and grow and amplify them. These handling techniques make detecting genes possible, and have previously been miniaturised so they fit on a microchip that uses small channels, valves and reaction chambers.

**Figure 11**: The device is about the size of a shoe-box with the optics and supporting electronics filling the space around the microchip.

The team have redesigned the gene handling microchip, and used a different detection method, to develop a system that is 'comparable in performance to much bigger and more expensive machines', explained Backhouse. The size of the device is reduced to that of a shoe-box with optics instrumentation and supporting electronics filling the space around the microchip.

To keep costs down, 'instead of using the very expensive confocal optics systems currently used in these types of devices we used a consumer-grade digital camera', Backhouse explained. The device can be adapted for used in many different genetic tests. 'By making small changes to the system you could test for a person's predisposition to cancer, carry out pharmacogenetic tests for adverse drug reactions or even test for pathogens in a water supply,' said Backhouse. The group strives to make genetic testing accessible to everyone in the same way computers are now. 'It's not long ago that computers were inaccessible to most people but now we all carry more than one on our person. This was made possible by integration and cost reductions,' said Backhouse. He said he plans to cut the manufacturing costs of this device to £50 in the very near future by integrating more of the electronics and further miniaturizing the microfluidics.
DNA Paternity Testing

DNA fingerprinting (also called DNA typing) allows for the identification of the source of DNA samples. The method has become very important to provide evidence in paternity and criminal cases. In contrast to the more conventional methodologies, such as blood typing, which can only exclude a suspect, DNA fingerprinting can provide positive identification with great accuracy.

Paternity determination based on DNA analysis (genetic DNA fingerprinting) has become an important procedure for matching children with biological fathers and mothers. Examples of recent court cases that have utilized this procedure have included rape, incest, immigration, citizenship of children to the United States and matching of children with parents who were mismatched at birth due to hospital errors. This type of testing is also used during unrest as in nations in civil war where children are often separated from parents and subsequently reunited.

For paternity DNA fingerprinting, samples obtained from the mother, the child, and possible fathers are analyzed. A child’s DNA is a composite of its parent DNAs. Therefore, comparison of DNA fragmentation patterns obtained from the mother and child will give a partial match. Bands in the child’s DNA fingerprint that are not present in the mother's must be contributed by the father. Because of allelic differences, the DNA bands present in the child’s fingerprint must be found in either the father’s or mother’s fingerprint.

Prior to the advent of the Polymerase Chain Reaction (PCR), DNA fingerprinting involved the electrophoretic analysis of DNA fragment sizes generated by restriction enzymes followed by Southern Blot Analysis. Restriction enzymes are endonucleases which catalyze the cleavage of the phosphate bonds within both strands of DNA. They require Mg+2 for activity and generate a 5 prime (5’) phosphate and a 3 prime (3’) hydroxyl group at the point of cleavage. The distinguishing feature of restriction enzymes is that they only cut at very specific sequences of bases called recognition sites. Restriction enzymes are produced by many different species of bacteria (including blue-green algae). Over 3,000 restriction enzymes have been discovered and catalogued.

**Figure 12**

Restriction enzymes recognize specific double stranded sequences in DNA. Most recognition sites are 4 to 8 base pairs in length. Cleavage occurs within or near the site. The cleavage positions are indicated by arrows. With some exceptions, recognition sites are frequently symmetrical, i.e., both DNA strands in the site have the same base sequence when read 5’ to 3’. Such sequences are called palindromes. It is these sites in DNA that are substrates for restriction enzymes. In DNA paternity and fingerprinting laboratories, the commonly used
restriction enzymes are \textit{Hae} III (GG'CC) and \textit{Hinf} I (G'ANTC), which are 4-base and 5-base cutting enzymes respectively.

In this experiment, the DNAs from a hypothetical paternity case are cut by a restriction enzyme, which is a six-base cutting enzyme. Examples of six-base cutting enzymes include \textit{Bam} HI and \textit{Pst} I. The recognition sites for these restriction enzymes are:

\begin{align*}
\text{\textit{Bam} HI} & \quad \text{\textit{Pst} I} \\
\downarrow & \quad \downarrow \\
5'\text{-GGATCC-3'} & 5'\text{-CTGCAG-3'} \\
3'\text{-CCTAGG-5'} & 3'\text{-GACGTC-5'}
\end{align*}

\textit{Figure 13}

The size of the DNA fragments generated by restriction enzyme cleavage depends on the distance between the recognition sites. No two individuals have exactly the same pattern of restriction enzyme recognition sites. There are several reasons for this fact. A large number of alleles exist in the population. Alleles are alternate forms of a gene. It is estimated that about 25\% of all human genes occur in multiple alleles which are called polymorphisms. Alleles result in alternative expressions of genetic traits which can be dominant or recessive and are inherited in a Mendelian pattern just as genes.

Chromosomes occur in matching pairs, one of maternal and the other of paternal origin. The two copies of a gene (which can be alleles) at a given chromosomal locus, and which represent a composite of the parental genes, constitutes the unique genotype for an offspring. It follows that alleles have differences in their base sequences which consequently creates differences in the distribution and frequencies of restriction enzyme recognition sites. Other differences in base sequences between individuals can occur because of mutations and deletions. Such changes can also create or eliminate a recognition site.

The example in Figure 14 shows how a silent mutation can eliminate a recognition site but leave a protein product unchanged. Individual variations in the distances between recognition sites in chromosomal DNA are often caused by intervening repetitive base sequences. Repetitious sequences constitute a large fraction of the mammalian genome (Mammalian Genome focuses on experimental, theoretical, and technical aspects of genomics and genetics in mouse, human, and other species, particularly those which bear on studies of gene function[30] and have no known genetic function. These sequences can occur between genes or are adjacent to them. They are also found within introns. Ten to fifteen percent of mammalian DNA consists of sets of repeated, short sequences of bases that are tandemly arranged in arrays. The length of these arrays (the amount of repeated sets) varies between individuals at different chromosomal loci.

\text{\text{\texttt{TGT}}TT\text{\texttt{A}} | \text{\texttt{T}}\text{\texttt{GTT}}TT\text{\texttt{A}} | \text{\texttt{T}}\text{\texttt{GTT}}TT\text{\texttt{A}} | \text{variable number}}

When these arrays are flanked by recognition sites, the length of the repeat will determine the size of the restriction enzyme fragment generated. Variations in the length of these fragments between different individuals, in a population, are known as restriction fragment

40
length polymorphisms, RFLPs. Several hundred RFLPs have been mapped on all 23 chromosomes. RFLPs are a manifestation of the unique molecular genetic profile, or “fingerprint”, of an individual’s DNA. As shown in Figure 15, there are several types of these short, repetitive sequences that have been cloned and purified. In Southern blot analysis, DNA probes are used to detect the length differences between these repetitive sequences. DNA probes are short fragments of single stranded DNA that are isotopically or non-isotopically labeled. DNA probes will complement and hybridize (attach) to single stranded DNA. Southern blot analysis requires electrophoresis, denaturation of the DNA fragments, transfer of DNA to a membrane, and exposure to probes to detect DNA Fingerprints.

Figure 14

There are two types of probes commonly used for genetic identification. The single-locus probes (SLPs) which detect a single segment of the repetitive DNA located at a specific site on a single chromosome. This will result in one or two DNA bands corresponding to one or both chromosome segments recognized. If the segments on the chromosome pairs are the same, then there will be one band. On the other hand, if they are different, it will appear as
two bands. Several SLPs are available and are used less frequently since more than one person can exhibit the same exact pattern for a specific SLP. Multiple-Locus probes (MLPs) detect multiple repetitive DNA segments located on many chromosomes yielding 20-30 bands. Because of the multi-band patterns, the chances of two people chosen at random having the same pattern is enormously remote. For example, it is calculated that two unrelated individuals having the identical DNA pattern detected by MLPs as an average is 1 in 30 billion. It should be kept in mind that the total human population on earth is between 5-6 billion.

Currently, the polymerase chain reaction (PCR) is routinely used in forensics to analyze DNA (Figure). This technique requires about 500-fold less DNA than Southern blot RFLP analysis and is less time-consuming. PCR amplification uses an enzyme known as Taq polymerase. This enzyme originally was purified from a bacterium that inhabits hot springs and is stable at very high (near boiling) temperatures. Also included in the PCR reaction mixture are two synthetic oligonucleotides known as “primers” and the extracted DNA. The region of DNA to be amplified is known as the “target”. In the first step of the PCR reaction, the template complementary DNA strands are separated (denatured) from each other at 94°C, while the Taq polymerase remains stable. In the second step, known as annealing, the sample is cooled to 40°-65°C, to allow hybridization of the two primers, one to each of the two strands of the template DNA. In the third step, known as extension, the temperature is raised to 72°C and the Taq polymerase adds nucleotides to the primers to synthesize the new complementary strands. These three steps - denaturation, annealing, and extension - constitute one PCR “cycle”. This process is typically repeated for 20-40 cycles, amplifying the target sequence within DNA exponentially. PCR is performed in a thermal cycler, an instrument that is programmed to rapidly heat, cool and maintain samples at designated temperatures for varying amounts of time. The PCR products are separated by agarose gel electrophoresis and DNA fingerprints are analyzed.

In forensics, PCR is used to amplify and examine highly variable (polymorphic) DNA regions. These are regions that vary in length from individual to individual and fall into two categories: 1) variable number of tandem repeats (VNTR) and 2) STR (short tandem repeats). A VNTR is a region that is variably composed of a 15-70 base pair sequence, typically repeated 5-100 times. An STR is similar to a VNTR except that the repeated unit is only 2-4 nucleotides in length. By examining several different VNTRs or STRs from the same individual, investigators obtain a unique DNA profile for that individual which is unlike that of any other person (except for identical twins).

In this simulation experiment, DNA was extracted from samples obtained from the mother, child and two possible fathers. The objective is to analyze and match the DNA fragment patterns after agarose gel electrophoresis and determine if Father 1 or Father 2 is the biological parent of the child.
Figure 16: Amplification of DNA by Polymerase Chain Reaction

EXPERIMENT OVERVIEW: FLOW CHART
Figure 17: During electrophoresis, the DNA samples migrate through the agarose gel towards the positive electrode. Before loading the samples, make sure the gel is properly oriented in the apparatus chamber.

Just like in forensic testing which originally relied upon blood groups and then moved into DNA fingerprinting, DNA profiling is increasingly used in kinship testing as well. A child must be a combination of the DNA of its mother and father. Using DNA profiles of all the people involved, usually mother, child and two men questioning their fatherhood, a probability can
be given for one of the men being the father. A man can be excluded as the father if he does not match with the child on at least two STR loci. Inclusion as a father is preferably with 99.9% probability.

<table>
<thead>
<tr>
<th>STR marker in</th>
<th>D3</th>
<th>VWA</th>
<th>FGA</th>
<th>D8</th>
<th>D21</th>
<th>D16</th>
<th>D5</th>
<th>D13</th>
<th>D7</th>
<th>AMEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>X, X</td>
</tr>
<tr>
<td>Child</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>X, X</td>
</tr>
<tr>
<td>? Father 1</td>
<td>16</td>
<td>18</td>
<td>17</td>
<td>13</td>
<td>15</td>
<td>28</td>
<td>14</td>
<td>11</td>
<td>11</td>
<td>X, Y</td>
</tr>
<tr>
<td>? Father 2</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>11</td>
<td>15</td>
<td>28</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>X, Y</td>
</tr>
</tbody>
</table>

**Figure 18**: Example of DNA profiling used in paternity testing. Possible Father 1 would be excluded but there is a high probability that Father 2 is the true father.

For parentage analysis, a paternity index (PI) is produced. This is a measure of the probability that compares the chance that the man is the father in the mother-child-father combination to the chance that the man was randomly chosen from the population. The PI will either exclude a man as the father or demonstrate that there is a high probability that he is the father of the child.[32]

**Overview of Genetic Discrimination**

While most Americans are optimistic about the use of genetic information to improve health, many are concerned that genetic information may be used by insurers to deny, limit or cancel health insurance, and by employers to discriminate in the workplace. They are worried that some insurers may choose not to insure people who are healthy but genetically pre-disposed to future disease onset: such people incur more health-related costs for the insurance company than individuals who are not predisposed. Similarly, they fear that some employers might only employ or retain individuals who are not pre-disposed to future disease onset, since healthy individuals are more productive. Therefore, many lawmakers, scientists and health advocacy groups believe that there is a need for federal legislation to prevent genetic discrimination.[33]

**Dna testing and forensics** [34]

Over 100 years ago fingerprinting was the most accurate way to place a suspect at the scene of the crime. The technique was revolutionary and put many criminals behind bars. Now, however, a newer and more accurate method is being used: DNA testing and it’s just as revolutionary, if not more so. Although DNA testing takes time, usually several weeks because of the legwork involved in matching base pair sequences it is highly accurate; in one case, the odds that the match was
incorrect were 350 million to 1! This makes DNA tests the most accurate piece of scientific evidence a lawyer can have.

The process is simple: two samples are taken, one from the suspect and one from the scene of the crime. The samples can be skin tissue, hair, blood, semen or vaginal fluid, and really anything else with cells in it (the two samples don’t even have to be the same material because all cells of the same organism have identical DNA). Then, the DNA of both samples is extracted, studied, and compared. If the DNA matches, then the suspect was at the scene of the crime. Two different DNA tests exist, the RFLP and the PCR. Both are very accurate, but they’re conducted in different ways. At this point we are going to describe PCR procedure.

The test can be performed with minute crime scene samples, which helps investigators who have little physical evidence. The DNA doesn’t have to be recently collected, either; the PCR test can still be performed even years after decades after-the fact, and still be just as accurate.

Generally, Genetic testing involves the analysis of the information in the ‘coding DNA’ that makes up genes located on specific chromosomes. If a chromosome is considered to be like a string of beads where the beads are the genes, the DNA ‘string’ between the genes is called ‘non-coding’ DNA. ‘Non-coding’ DNA does not contain the coded messages that the cells use and sometimes has been erroneously referred to as ‘junk’ DNA. It is increasingly clear that the role of this form of DNA is important, perhaps in controlling how the genes work. Almost 75% of the total DNA in the human genome is ‘non-coding DNA between and within the genes. Analysis of the non-coding DNA between the genes has applications in forensic studies and biological relationship testing.

The genetic test utilizes the many small differences that are present in this DNA (called polymorphisms) between individuals to create a fingerprint for each person based on their genetic code. The genetic code in all DNA in humans is made up of a string of 6 billion or so ‘letters’: A, T, C and G. These letters are chemical ‘bases’ of the DNA molecule. In the ‘coding DNA’ which makes up most of the DNA in the genes, the letters are combined into groups of three to produce a message made up of a sequence of three letter words. The genetic information instructs the cells to enable the body to grow, develop and function. The non-coding DNA is of course also made up of a long string of the letters A, T, C and G. An important feature of this non-coding DNA is that it contains large numbers of repeated sequences of letters along its length e.g. ATTCGATTCGATTCG. The number of times that a sequence is repeated within a length of non-coding DNA can range from just a few to hundreds. Studies of these repeated sequences are proving very useful in tracking our evolutionary past.

As we have evolved, changes have built up in our non-coding DNA as well as our genes. While changes in the coding DNA sequences in the genes may or may not cause a problem with how the genes work, many changes to the non-coding DNA have occurred with no impact on the individual.

Despite these changes that have built up in the non-coding DNA, most humans are overwhelmingly alike in their sequence of letters in their coding DNA i.e. in their genes, regardless of race or ethnicity although each person’s sequence is unique (except for identical twins). DNA samples from two unrelated people differ, on average, at only one letter or ‘base’ per thousand that works out to be 1/1,000th of 6 billion that equals 6 million: i.e. in their total DNA, two unrelated people have 6 million differences in their sequence of letters. This is enough to produce all the genetic differences between these two people. The chance that two unrelated people have an identical total DNA gene sequence is at least 1 in 6 million. The closer two people are related to each other, the greater the chance that their DNA sequence will be similar. Identical twins have exactly the same DNA sequence. The small differences and similarities in the DNA sequence
between people are used in the tests to determine the identification of individuals as described below.

**Testing the non-coding DNA for non-medical purposes**

The tests look at specific areas (called loci singular locus) in the non-coding DNA that have nothing to do with neither how our bodies grow and develop nor our health. These loci are found at a number of sites on each chromosome. The loci that are most useful are those that have been found to have different numbers of the repeated sequences in different people in the population. These sites are described as being ‘polymorphic’ that means ‘many forms’. Each person has two copies of each chromosome. So each person will have a specific number of repeats at each of these loci on one of their chromosomes and the same, or different, number of repeats, at the partner locus on the other chromosome. These different forms at the same sites on the partner chromosomes are called alleles of the locus. This information is used to create a DNA pattern or DNA fingerprint that is as unique as possible for each person as described below.

**Creating a DNA fingerprint for a person**

**Step1:**

- In the laboratory, using enzymes that are chemical ‘scissors’, the DNA is cut into hundreds of small pieces at sites where there are specific sequences of the DNA letters (usually four to six letters in length)
- As everyone’s DNA has some small differences, the sites may be at different places in people’s non-coding DNA and so the enzymes will cut the DNA into different sizes in different people.

**Step 2:**

- The cut DNA is placed into a slab of ‘jelly’ (a gel matrix) and an electrical current is applied so that the ‘jelly’ becomes electrically charged and has a ‘positive’ (+) end at the top and a negative (-) end at the bottom, just like the positive and negative ends of a battery
- As the DNA is a chemical which has a negative charge, the DNA is pulled towards the positive end of the gel or from the top to the bottom
- The pieces of DNA separate along the gel according to size: the biggest pieces move the slowest and so will remain near to the top of the gel
- The gel now contains all of the person’s DNA spread from the top to the bottom of the gel
Step 3:
- To select out the pieces of DNA that need to be looked at, the pieces of DNA that have spread through the gel are covered with special DNA ‘probes’.
- The probes are made in the laboratory and contain a match for the DNA sequence that the test is looking for.
- The probes in fact have the opposite sequence to the sequence being tested for. They match up because of the ability of the letters A and T, and C and G to pair.
- If the person’s DNA on the gel contains the matching sequence, the probe will combine (hybridize) with the person’s DNA at the site of the matching sequences.
- Chemicals such as fluorescent dyes are attached to the probes. When the gel is exposed to a certain type of light, the DNA on the gel that has been detected as containing a particular sequence of ‘letters’ will show up. The remaining DNA will not be visible. When a number of different probes are used to detect different repeats of sequences; the gel will look like a series of bands as shown below. Each band represents a site or locus on the non-coding DNA at which a specific repeated sequence is located.

![Diagram of DNA fingerprinting process]

**Figure 19**: Genetic testing to create a DNA fingerprint of a person

**Figure 20**: DNA fingerprints of three people created using seven different DNA probes manufactured in the laboratory.
• Some of the bands appear darker than others. This is because that person has the same number of repeated sequences of letters on each of their two chromosomes where the locus is located. The band is showing the same pattern from each chromosome and so the two bands are lying on top of each other on the gel.
  • A person can only ever have one or two bands for each probe.
  • Persons 1 and 2, however, have the same banding pattern for probe 5 and persons 2 and 3 have the same pattern for probe 7
  • That is why it is necessary to use a number of different probes to enable differentiation between different people
  • The bands at the bottom of the gel represent shorter sequences of letters than at the top. For example, a band at the top of the gel may represent a sequence of five letters repeated 30 times so it will be 150 letters or bases long. At the bottom the sequence may be made up of only four letters repeated 10 times so it will be only 40 letters or bases long. These patterns of bands have become known as a person’s DNA fingerprint, an analogy with the classical fingerprint system of identification.

From DNA fingerprinting to DNA profiling in forensic studies

While DNA fingerprinting enabled matching of samples with a high probability from crime scenes with those of suspects in many cases, there were often difficulties in visually separating bands that were very close together. The loci used were called variable number of tandem repeats (VNTRs), usually 8-80 letters in length so that separation was sometimes not optimal. The analysis of VNTRs is now being replaced by the analysis of repeats of shorter sequences of letters called short tandem repeats (STRs) that are usually three to five letters in length.
  • STRs can be used even when the sample of DNA is very small such as that from a cigarette, hairs, coffee cups or the saliva from a postage stamp
  • Their interpretation is also usually less uncertain and the analysis takes only a few days instead of weeks using VNTRs
  • Automatic systems are being developed to make the analysis even more rapid. The more regions of the DNA where the STRs are located (loci) are used, the greater the chance that variability will be found between unrelated people. In addition, analysis of markers on the two sex chromosomes (X and Y) immediately identifies the sex of the person. Further information can be gained from the analysis of the DNA in the mitochondria, small units in the cells that contain genes involved in energy production.

Figure 21: The FBI has selected 13 core STR loci that must be run in all DNA tests in order to provide a common currency with DNA profiles.
• Each STR locus has a number of different possible patterns according to the different number of repeats of the sequence, but each person will only have two of all the possible patterns in their DNA
• Rather than looking at the bands generated from these patterns, and creating a visual fingerprint, the number of repeats in each pattern for each person is measured.
• This gives a numerical **DNA profile** for these STRs for each person.

**Examples**

1)  

<table>
<thead>
<tr>
<th>STR marker in Profile Plus</th>
<th>D3</th>
<th>VWA</th>
<th>FGA</th>
<th>D8</th>
<th>D21</th>
<th>D18</th>
<th>D5</th>
<th>D13</th>
<th>D7</th>
<th>AMEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person 1</td>
<td>14, 14</td>
<td>15, 18</td>
<td>19, 20</td>
<td>11, 13</td>
<td>31,2, 34,2</td>
<td>16, 19</td>
<td>11, 11</td>
<td>10, 12</td>
<td>10, 10</td>
<td>X, X</td>
</tr>
<tr>
<td>Person 2</td>
<td>17, 18</td>
<td>17, 17</td>
<td>18, 22</td>
<td>13, 15</td>
<td>29, 30</td>
<td>10, 14</td>
<td>11, 11</td>
<td>9, 14</td>
<td>11, 13</td>
<td>X, Y</td>
</tr>
</tbody>
</table>

**Figure 22**: DNA profiles for two people. Person 1 is a female and person 2 is a male, identified by the AMEL STR locus on the sex chromosomes.

2)  

<table>
<thead>
<tr>
<th>STR marker in Profile Plus</th>
<th>D3</th>
<th>VWA</th>
<th>FGA</th>
<th>D8</th>
<th>D21</th>
<th>D18</th>
<th>D5</th>
<th>D13</th>
<th>D7</th>
<th>AMEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Victim</td>
<td>14, 15</td>
<td>16, 18</td>
<td>19, 20</td>
<td>11, 13</td>
<td>30, 30</td>
<td>16, 19</td>
<td>11, 12</td>
<td>9, 12</td>
<td>10, 10</td>
<td>X, X</td>
</tr>
<tr>
<td>Semen specimen</td>
<td>15, 16</td>
<td>17, 17</td>
<td>18, 22</td>
<td>13, 15</td>
<td>28, 30</td>
<td>10, 14</td>
<td>11, 11</td>
<td>9, 14</td>
<td>11, 13</td>
<td>X, Y</td>
</tr>
<tr>
<td>Suspect 1</td>
<td>14, 15</td>
<td>16, 16</td>
<td>19, 20</td>
<td>11, 13</td>
<td>28, 30</td>
<td>19, 19</td>
<td>11, 12</td>
<td>12, 12</td>
<td>10, 11</td>
<td>X, Y</td>
</tr>
<tr>
<td>Suspect 2</td>
<td>16, 18</td>
<td>17, 17</td>
<td>18, 22</td>
<td>13, 16</td>
<td>28, 30</td>
<td>10, 14</td>
<td>11, 11</td>
<td>9, 14</td>
<td>11, 13</td>
<td>X, Y</td>
</tr>
<tr>
<td>Suspect 3</td>
<td>14, 15</td>
<td>16, 18</td>
<td>19, 20</td>
<td>11, 15</td>
<td>28, 29</td>
<td>14, 16</td>
<td>14, 12</td>
<td>12, 12</td>
<td>10, 13</td>
<td>X, Y</td>
</tr>
</tbody>
</table>

**Figure 23**: Identification of suspect 2 as an offender in a rape case.

**RECOMBINANT DNA TECHNOLOGY**

Recombinant DNA is DNA that has been created artificially. DNA from two or more sources is incorporated into a single recombinant molecule. Treat DNA from both sources with the same restriction endonuclease (BamHI for example). BamHI cuts the same site on both molecules 5' GGATCC 3' 3' CCTAGG 5'. The ends of the cut have an overhanging piece of single-stranded DNA. These are called "sticky ends" because they are able to base pair with any DNA molecule containing the complementary sticky end. In this case, both DNA preparations have complementary sticky ends and thus can pair with each other when mixed.
DNA ligase covalently links the two into a molecule of recombinant DNA. To be useful, the recombinant molecule must be replicated many times to provide material for analysis, sequencing, etc. Producing many identical copies of the same recombinant molecule is called cloning. Cloning can be done in vitro, by a process called the polymerase chain reaction (PCR).

Figure 24: Making Recombinant DNA (rDNA): An Overview

In vivo cloning

Cloning in vivo can be done in: unicellular microbes like E. coli, unicellular eukaryotes like yeast and, in mammalian cells grown in tissue culture. In every case, the recombinant DNA must be taken up by the cell in a form in which it can be replicated and expressed. This is achieved by incorporating the DNA in a vector. A number of viruses (both bacterial and of mammalian cells) can serve as vectors. But here let us examine an example of cloning using E. coli as the host and a plasmid as the vector.

Basic genetic engineering (GE) takes donor DNA from one organism or type of cell and places it into the DNA of another organism or type of cell. It includes following steps:
1. Isolation of gene
2. Preparation of target DNA
3. Insertion of DNA into plasmid
4. Insertion of plasmid back into cell
5. Plasmid multiplication
6. Target cells reproduction
7. Cells produce proteins

Specifically

1. Isolation of Gene: The gene for producing a protein is isolated from a cell. The gene is on the DNA in a chromosome. Special DNA cutting proteins are used to cut out certain sections of DNA. The gene can be isolated and then copied so that many genes are available to work with.
2. Preparation of Target DNA: In 1973, two scientists named Boyer and Cohen developed a way to put DNA from one organism into the DNA of bacteria. This process is called recombinant DNA technology. First, a circular piece of DNA called a plasmid is removed from a bacterial cell. Special proteins are used to cut the plasmid ring to open it up.
3. **Insertion of DNA into Plasmid:** The host DNA that produces the wanted protein is inserted into the opened plasmid DNA ring. Then special cell proteins help close the plasmid ring.

4. **Insertion of Plasmid back into cell:** The circular plasmid DNA that now contains the host gene is inserted back into a bacteria cell. The plasmid is a natural part of the bacteria cell. The bacteria cell now has a gene in it that is from a different organism, even from a human. This is what is called recombinant DNA technology.

5. **Plasmid multiplication:** The plasmid that was inserted into the bacteria cell can multiply to make several copies of the wanted gene. Now the gene can be turned on in the cell to make proteins.

6. **Target Cells Reproduction:** Many recombined plasmids are inserted into many bacteria cells. While they live, the bacteria's cell processes turn on the inserted gene and the protein is produced in the cell. When the bacteria cells reproduce by dividing, the inserted gene is also reproduced in the newly created cells.

7. **Cells Produced Proteins:** The protein that is produced can be purified and used for a medicine, industrial, agricultural, or other uses.[35]

For example:

![Production of Recombinant DNA Using a Human Gene & Bacterial Plasmids](image)

**Figure 25 :** Recombinant DNA and signal transduction

Signal transduction involves the conversion by a cell of an external stimulus to one or more internal events by a sequence of processes that take place in the cell plasma membrane.

Signal transduction can be recognized as three steps: reception, transduction, induction.

- **Reception** – the binding of the signal molecule (e.g. a hormone) to its specific receptor.
- **Transduction** – the second messenger is formed in or released into the cytosol (the second messenger amplifies the stimulus and initiates the cell’s response).
- **Induction** – activation of the cellular process.
Signal transduction involves gated (protein) channels in the cell’s plasma membrane, the nature of which depends on the genes that control their production. The genetic code therefore controls, for example, the cell’s response to hormones, growth factors, cytokines, and neurotransmitters, the ability to perceive odours, the immune response, the response of slime moulds to cyclic adenosine monophosphate (cAMP) and the mating response in yeasts (among others).

In 1986 the concept of a molecular cascade or linear signal transduction pathway that was linked to disease was published. It is now known that the end products of signal transduction pathways are often transcription factors that act on DNA and affect gene expression.

It was then discovered that multiple linear pathways operate in networks. It was shown, for example, that the cyclic AMP (cAMP) pathway – a second messenger that can transmit a signal across a cell membrane – could repress a pathway and interfere with the cancer transformation of some cells.

This means that a change in the genetic code that results in the change in a plasma membrane protein may cause a signal transduction process change. This, in turn, may result in a disorder, for example diabetes, cystic fibrosis, cancers rheumatoid arthritis and inflammatory diseases.[36]

**Rheumatoid arthritis – treatment with a recombinant human tumor necrosis factor receptor**

Rheumatoid arthritis is a common disease, and it produces substantial morbidity as well as an increase in mortality. Although the causes of rheumatoid arthritis are not fully understood, laboratory and clinical evidence suggests that proinflammatory cytokines, particularly tumor necrosis factor (TNF), have an important role in its pathogenesis. TNF induces the release of matrix metalloproteases from neutrophils, fibroblasts, and chondrocytes induces the expression of endothelial adhesion molecules involved in the migration of leukocytes to extravascular sites of inflammation and stimulates the release of
other proinflammatory cytokines. TNF concentrations are increased in the synovial fluid of persons with active rheumatoid arthritis, and increased plasma levels of TNF are associated with joint pain. Administration of TNF antagonists to patients with rheumatoid arthritis has been shown to reduce symptoms.

There are two distinct cell-surface TNF receptors (TNFRs), designated p55 and p75. Soluble, truncated versions of membrane TNFRs, consisting of only the extracellular, ligand-binding domain, are present in body fluids and are thought to be involved in regulating TNF activity. Soluble TNFRs have been detected in synovial tissue and at the junction between cartilage and pannus. Their levels are increased in serum and synovial fluid in rheumatoid arthritis and in many other autoimmune and inflammatory conditions.

Recombinant TNF receptor

A recombinant human TNFR p75–Fc fusion protein (TNFR:Fc) has been developed for therapeutic neutralization of TNF. DNA encoding the soluble portion of human TNFR p75 was linked to DNA encoding the Fc portion of a human IgG1 molecule, and the combined DNA was then expressed in a mammalian cell line. The resulting immunoglobulin-like dimer, composed exclusively of human amino acid sequences, acts as a competitive inhibitor of TNF and prevents binding of TNF to the cell-surface TNFR, thereby reducing the biologic activity of TNF. Safety studies in normal human volunteers revealed no adverse effects after the intravenous administration of TNFR:Fc. There were trends toward a reduction in disease activity in a safety and dose-finding study of TNFR:Fc administered for four weeks to a small number of patients with refractory rheumatoid arthritis. On the basis of these findings, we undertook a multicenter, randomized, double-blind, placebo-controlled trial of TNFR:Fc in patients with active, refractory rheumatoid arthritis.

Characteristics of the Patients

Forty-eight men and 132 women were enrolled in the trial. Their mean age was 53 years, and 77 percent had disease of more than 5 years’ duration. No significant differences between groups were detected in pretreatment characteristics or base-line disease activity. Seventy-six percent of the patients completed TNFR:Fc treatment (61 percent in the 0.25-mg group, 78 percent in the 2-mg group, and 93 percent in the 16-mg group), as compared with 52 percent of the patients assigned to placebo. The primary reason for withdrawal was inadequate control of arthritis symptoms. Among the patients receiving TNFR:Fc, the proportions of patients who withdrew because of inadequate symptom control were 35 percent in the 0.25-mg group, 17 percent in the 2-mg group, and 5 percent in the 16-mg group; among the patients receiving placebo, it was 43 percent.

Efficacy

TNFR:Fc produced significant improvement in all measures of disease activity. A clear dose–response relation was observed in the numbers of swollen or tender joints, and patients who received the highest dose of TNFR:Fc had the greatest improvement. In the 16-mg group, the mean percent reduction in the total count at three months was 61 percent, as compared with 25 percent in the placebo group (P <0.001). In the placebo and 0.25-mg groups, there was an initial response, but no improvement was noted thereafter. The 16-mg dose of TNFR:Fc was associated with the greatest reduction in the number of swollen or tender joints. This difference was apparent by the end of week 2 and was most pronounced at the end of treatment (at three months). TNFR:Fc treatment was also associated with significant reductions in pain and duration of morning stiffness, significant improvement in
the quality of life and physician’s and patient’s global assessments, and significant reductions in disease activity as assessed by objective laboratory measures (erythrocyte sedimentation rate and C-reactive protein level). According to American College of Rheumatology criteria, at three months 57 percent of the 16-mg group had at least 50 percent improvement, as compared with 7 percent of the placebo group (P<0.001); 75 percent of the 16-mg group had at least 20 percent improvement, as compared with 14 percent of the placebo group (P <0.001). Measures of disease activity moved toward base-line levels after the cessation of TNFR:Fc therapy.[37]

Ancient DNA and the Polymerase Chain Reaction

Antigenic procedure

Students of molecular evolution suffer from the frustration of trying to reconstruct this historic process from only knowledge of the present-day structure of genes. Until recently, there has been no hope of escaping this “time trap.” However, advances in molecular biological techniques have enabled us to retrieve and study ancient DNA molecules and thus to catch evolution red-handed. In consequence, we can now study the genealogical relationships of extinct species and vanished populations. The first indications that molecular genetic information might persist in ancient materials were early demonstrations that the peptide bond can last for up to $10^8$ years in fossil shells and bones and that subcellular detail implying the survival of ribosomes and chromatin is evident in insects from 40 million-year-old amber. Indeed, these findings inspired the hope that genetic information should be retrievable from the amino acid sequences in ancient remains, and substantial efforts over the past two decades went into such endeavors. Unfortunately, the major proteins in bone (collagen) and shell (conchiolin) are likely to be genetically uninformative because collagen has a repetitious primary structure and is encoded by multiple genes (5) whereas conchiolin is a complex mixture of proteins whose genetic basis is unknown. Second, the proteins in ancient remains are structurally heterogeneous because of post-mortem modifications. Even in exceptionally well preserved remains, such as frozen muscle from an extinct Siberian mammoth, extensive modifications were evident from elemental analysis, electron microscopy, and amino acid analysis of the 40,000-year-old tissue. In the case of albumin, one of the most stable globular proteins known in animal tissues, only about 2% of the mammoth molecules could dissolve in water, and 80% of the latter were modified in charge, size, or antigenicity.

The first comparisons of the primary structure of a gene product of an extinct species to that of living species were achieved indirectly by using polyclonal antisera raised against a homogenate of mammoth muscle. Such an antiserum was tested for its ability to form precipitates and complement-fixing lattices with a panel of native albumins from several living species. The reactions were strong with Indian and African elephant albumins, weak with sea cow albumin, and weaker still with other mammalian albumins. Since the cross-reaction specificity of the antimammoth serum matched exactly those of antisera to the pure albumins of elephants, the native albumins of these three species (mammoth, Indian elephant, and African elephant) are nearly identical in primary sequence. Immunological methods are especially likely to give misleading results when employing antisera that are raised and later also tested against mixtures of poorly defined antigens.
For this reason, we consider most antigenic studies reported on ancient materials to be of questionable genetic value.

The Polymerase Chain Reaction in Molecular Archaeology

The polymerase chain reaction (PCR) can amplify preselected segments of DNA up to quantities which permit direct sequencing, starting from extremely small amounts of DNA or even single molecules. The amplification is done with two synthetic oligodeoxynucleotide primers, each about 25 bases long, a thermostable DNA polymerase, and the four deoxyribonucleotide triphosphates. The first primer matches part of the Watson strand at one end of the segment, while the second primer matches the Crick strand at the other end. Since the 3' ends of the two primers point toward each other, repeated cycles of heating and cooling lead to a chain reaction, i.e. an exponential synthesis of many copies of the specific segment bounded by the two primers. PCR is an ideal tool to amplify a small number of intact ancient DNA molecules present in a vast excess of damaged molecules. In addition to its ability to detect extremely small quantities of DNA, PCR has the advantage of being an in vitro system, which has no capacity for repair or misrepair. For example, in bacterial clones of quagga mitochondrial DNA, two replacement substitutions were found when the quagga sequences were compared to other vertebrate sequences. These positions in the quagga were later shown not to differ from the general vertebrate sequence when they were directly sequenced from amplification products. The replacements observed in the cloned sequences were thus due to cloning artifacts, possibly resulting from misrepair of damaged molecules. In contrast, during enzymatic amplification, most damaged molecules will either not be replicated at all, e.g. due to interior intramolecular cross-links, or will be at a replicative disadvantage because many lesions, such as baseless sites, slow down the DNA polymerase. Intact molecules will thus amplify preferentially. Some damaged molecules will of course have only minor lesions, such as deaminated bases, which can generate replication errors without retarding replication. These errors as well as those introduced by the polymerase at undamaged sites will be present in the final population of molecules produced by PCR. However, since each error is specific to one molecule in the starting population (and its descendants), its contribution to the result, the sequence determined on the whole descendant population of amplified molecules, is likely to be negligible. Such errors would be encountered only if one were to clone individual molecules from the final population before carrying out the sequencing reactions. An additional advantage of PCR is that its speed allows for easy reproduction of results.

Ancient DNA Sequences Revealed via PCR

Using PCR and direct sequencing, it has been possible to obtain mitochondrial sequences from a 7000-year-old brain excavated in Florida. Preserved brains exist in association with human skeletal remains at several sites in Florida and owe their excellent state of preservation to anaerobic and neutral conditions in the waters of Florida peat bogs. Extracts from the archaic brain contained DNA that allowed amplification of 140-bp-long mtDNA fragments. Longer amplifications were unsuccessful. This is in sharp contrast to contemporary DNA, where fragments of up to 1 kilobase and longer amplify efficiently. It seems that the damage present in old DNA causes a strongly inverse correlation between
amplification efficiency and size of the amplification product, which sets a limit to the length of amplifications that can be performed.

The same procedures revealed DNA sequences from ancient humans that are preserved in the form of mummies. For example, a 4000-year-old Egyptian priest was shown to carry an unusual D-loop sequence. It is not yet known if this was a common genotype in ancient Egypt and whether it is represented today in Egypt or elsewhere.

Another area where paleomolecular biology is producing information is zoology. The 40,000-year-old Siberian mammoth has yielded PCR products contaminated with human sequences, presumably originating from handling of the mammoth after its discovery. Fortunately, the mammoth sequences were separable from the human sequences and, as expected, proved to be closely related to those of elephants. Since mtDNA is present in many copies/nucleated cell, it can be assumed that this high copy number facilitates its survival and retrieval. In fact, attempts to amplify specific nuclear single-copy genes from ancient remains (e.g. part of the β-globin gene from human mummies) have been unsuccessful, presumably due to their lower abundance in DNA extracts. Furthermore, the fast evolution and maternal mode of inheritance of mtDNA make it ideal for studying ancestor descendant relationships.

**Authenticity of Amplified Sequences**

Contamination by modern DNA. The main concern pertinent to the amplification of ancient DNA sequences is contamination of the DNA extracts or reagents by contemporary DNA. Accordingly, the following additional criteria must be set up and adhered to in order to detect any source of contaminating DNA. We have found it imperative always to do control extracts (i) in parallel with the extracts of the old specimens in order to detect contamination in solutions and reagents. Furthermore, several independent extracts (ii) from every individual should be prepared, and the sequences obtained should be unambiguous and identical. The strong inverse correlation between amplification efficiency and size of the amplification product (iii), which is observed for ancient but not modern DNA, can to some extent serve as a further criterion of authenticity. In our experience, archaeological remains generally do not yield any products above 150 bp in size, whereas better preserved specimens such as museum skins may allow for the amplification of up to 500-bp pieces. Sequences longer than this have invariably proved to originate from contamination of the specimens by modern DNA. When the three above criteria (i,ii and iii) are fulfilled, a given sequence is considered likely to be of ancient origin.

The recently achieved ability to study DNA from museum specimens and archaeological finds via PCR opens up the possibility of studying molecular evolution by actually going back in time and directly approaching DNA sequences that are ancestral to their present-day counterparts. This allows us to address many questions involving the identity and relationship of extinct species to other extinct and extant species. In these cases, even rare examples of single preserved specimens are of great value. Second, the evolution of populations can be studied and will most often involve the comparison of ancestral and descendant populations, attributing changes in gene frequencies to factors such as variations in population size, migration, selection, and genetic drift.

In addition to organismal and population questions, promising avenues of research on molecular evolution may now be approachable. We refer especially to the study of those molecular evolution genetic processes that may occur rapidly enough to be studied on a time scale of $10^4$ years, in contrast to the slow accumulation of base substitutions, which in nuclear and organelle genomes is generally less than 1% of sequence divergence/$10^5$ years. Faster processes include recombination in minisatellite sequences which cause apparent mutation frequencies on the order of $10^{-2}$/locus/generation. Other cases where rapid...
evolution may allow molecular processes to be observed directly can be provided by human parasites (especially viruses) and the evolution of domestic plants. A particularly important goal for the future is to refine the polymerase chain reaction so that it becomes possible to study ancient nuclear single copy genes as well as DNA sequences from bones, teeth, shells, and amber.[38]

**Conclusion**

Summing up, the major pros and cons of the polymerase chain reaction technique are the following:

**Speed and ease of use**

DNA cloning by PCR can be performed in a few hours, using relatively unsophisticated equipment. Typically, a PCR reaction consists of 30 cycles containing a denaturation, synthesis and reannealing step, with an individual cycle typically taking 3-5 min in an automated thermal cycler. This compares favorably with the time required for cell-based DNA cloning, which may take weeks. Clearly, some time is also required for designing and synthesizing oligonucleotide primers, but this has been simplified by the availability of computer software for primer design and rapid commercial synthesis of custom oligonucleotides. Once the conditions for a reaction have been tested, the reaction can then be repeated simply.

**Sensitivity**

Pcr is capable of amplifying sequences from minute amounts of target DNA, even the DNA from a single cell. Such exquisite sensitivity has afforded new methods of studying molecular pathogenesis and has found numerous applications in forensic science, in diagnosis, in genetic linkage analysis using single-sperm typing and in molecular paleontology studies, where samples may contain minute numbers of cells. However, the extreme sensitivity of the method means that great care has to be taken to avoid contamination of the sample under investigation by external DNA, such as from minute amounts of cells from the operator.

**Robustness**

Pcr can permit amplification of specific sequences from material in which the DNA is badly degraded or embedded in a medium from which conventional DNA isolation is problematic. As a result, it is again very suitable for molecular anthropology and paleontology studies, for example the analysis of DNA recovered from archaeological remains. It has also been used successfully to amplify DNA from formalin – fixed tissue samples, which has important applications in molecular pathology and, in some cases, genetic linkage studies.

The major disadvantages of PCR are the general requirement for prior target sequence information, short size and limiting amounts of product and infidelity of DNA replication.

Despite its huge popularity, PCR has certain limitations as a method for selectively cloning specific DNA sequences.
Need for target DNA sequence information. In order to construct specific oligonucleotide primers that permit selective amplification of a particular DNA sequence, some prior sequence information is necessary. This normally means that the DNA region of interest has been partly characterized previously, often following cell-based DNA cloning. However, a variety of techniques have been developed that reduce or even exclude the need for prior DNA sequence information concerning the target DNA, when certain aims are to be met. For example, previously uncharacterized DNA sequences can sometimes be cloned using PCR with degenerate oligonucleotides if they are members of a gene or repetitive DNA family at least one of whose members has previously been characterized. In some cases, PCR can be used effectively without any prior sequence information concerning the target DNA to permit indiscriminate amplification of DNA sequences from a source of DNA that is present in extremely limited quantities. Therefore, although PCR can be applied to ensure whole genome amplification, it does not have the advantage of cell-based DNA cloning in offering a way of separating the individual DNA clones comprising a genomic DNA library.

Short size and limiting amounts of PCR product

A clear disadvantage of PCR as a DNA cloning method has been the size range of the DNA sequences that can be cloned. Unlike cell-based DNA cloning where the size of cloned DNA sequences can approach 2 Mb, reported DNA sequences cloned by PCR have typically been in the 0.1-5 kb size range, often at the lower end of this scale. Although small segments of DNA can usually be amplified easily by PCR, it becomes increasingly more difficult to obtain efficient amplification as the desired product length increases. Recently, however conditions have been identified for effective amplification of longer targets, including a 42-kb product from the bacteriophage λ genome. Often, the conditions for long range PCR involve a combination of modifications to standard conditions with a two-polymerase system. This provides optimal levels of DNA polymerase and 3’→5’ exonuclease activity which serves as a proofreading mechanism.

The amount of PCR product obtained in a single reaction is also much more limited than the amount that can be obtained using cell-based cloning where scale-up of the volumes of cell cultures is possible. The efficiency of a PCR reaction will vary from template to template and according to various factors that are required to optimize the reaction but typically only comparatively small amounts of product are achieved.
References


2. Pdf PCR: an outstanding method


4. PCR 2: a practical approach

5. http://www.changbioscience.com/primo/prc/eAllelesSpecificsPCRs.htm


r+general+pdf&ots=Ed6LUjaZMq&sig=i7JVmJcDDn1U0mlUF9LrEzNayIM#v=onepage&q&f=false


16. Tsai et al, 1993; Wilton and Lim, 1995


18. Gene expression of Alzheimer-associated presenilin-2 in the frontal cortex

19. of Alzheimer and aged control brain.pdf

20. APP duplication is sufficient to cause early onset Alzheimer’s dementia with cerebral amyloid Angiopathy.pdf

21. Detection of prostate cancer using methylation-specific PCR

22. Quantitative real-time PCR for cancer detection: the lymphoma case.pdf


32. DNA Paternity Testing Simulation. pdf


34. http://library.thinkquest.org/28599/courtroom.htm

35. Recombinant DNA technology and genetic engineering a safe and effective meaning for production valuable biologicals

36. Signalling_molecules_Biology_Online_Article.doc

37. Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc Fusion.pdf

38. Ancient DNA and the polymerase chain reaction