**Genetic Engineering**

Asst. Prof. Dr. **Inam Jasim Lafta**

**College of Veterinary Medicine**

**Baghdad University**

2017/2018

**References**

**1- Mullis K B (1990). Recombinant DNA technology and molecular cloning.**

**Scientific American. 262:36.**

**2- Russell P J (2010). iGenetics: A Molecular Approach. Beth Wilbur Editor, 3rd Edition. Pearson Benjamin Cummings, 1301Sansome St., San Francisco, CA 94111.**

**3- Turner P *et al.* (2005). Instant Notes: Molecular Biology. Hames B. D. Editor, 3rd Edition. Taylor and Francis Group.**

**4- Lodish H, Berk A, Zipursky S L *et al.* (2000). Molecular Cell Biology. 4th edition. New York:** [**W. H. Freeman**](http://www.whfreeman.com/)**.**

**5- Winn W C *et al.* (2006). Koneman’s Color Atlas and Textbook of Diagnostic Microbiology. Peter Darcy Editor, 6th Edition. Lippincott Williams and Wilkins. Philadelphia, USA.**

**6- Brown T A (2016). Gene Cloning and DNA Analysis: An Introduction. Wiley Blackwell, 7th Edition. Oxford, UK.**

**7- Brandenberg *et al.* (2011). Introduction to Molecular Biology and genetic engineering.**

**Molecular Biology**

Molecular biology is a branch of science concerning biological activity at the molecular level. The field of molecular biology overlaps with biology and chemistry and in particular, genetics and biochemistry. A key area of molecular biology concerns understanding how various cellular systems interact in terms of the way DNA, RNA and protein synthesis function.

Molecular biology looks at the molecular mechanisms behind processes such as replication, transcription, translation and cell function. The basis of molecular biology concerns understanding how genes are transcribed into RNA and how RNA is then translated into protein.

**Genetic Engineering**

**An Introduction:** Genetic engineering refers to the techniques whereby recombinant DNA, hybrid DNA made by artificially joining pieces of DNA from different sources, is produced and utilized. The term ‘genetic engineering’ encompasses virtually any process involving DNA manipulation. The applications of genetic engineering are now so widespread and well established within the biomedical sciences. Currently, a huge number of articles utilize genetic engineering as an important component of their investigation.

The publication of the mouse genome sequence in late 2002 highlighted the enormous importance of the mouse as a model for human disease. Of the roughly 30 000 genes present in both human and mouse, 99% of the human genes have homologues in the mouse genome (and vice versa). In practice this means that nearly all the genes that contribute to human disease can be studied in the mouse, although gene function is not necessarily identical in the two contexts. Deleting genes out of mice (“making knockouts” in laboratory jargon) and over-expressing genes (“making transgenics”) has been used for more than a decade. In “knock-in” strategies, point mutations can be introduced into a specific gene enabling an excellent level of specificity in structure-function analysis. In cancer studies a mouse can be engineered to develop a tumour. By tagging tumour genes with a fluorescent probe the growth and remission of tumours can be readily assessed using whole body imaging. It is also now possible to introduce defined chromosomal rearrangements into the mouse genome by first genetically engineering them in embryonic stem cells. Other new technologies are enabling genomic DNA in bacterial artificial chromosomes to be directly modified and subcloned by a new approach known as “recombineering”.

An example of genetic engineering application, to identify novel pharmaceutical targets, is using drosophila (fruit fly) to identify genes involved in cell cycle regulation. By systematically deleting the approximately 14 000 or so genes in the drosophila genome using interfering RNA (si-RNA) technology it is possible to work out which gene products control the cell cycle. Since around 40% of human genes are identifiable in the drosophila genome and genes that control the cell cycle are highly conserved in evolution, this approach is expected to reveal human gene products that could become drug targets in the treatment of cancer and other disorders involving uncontrolled cell proliferation.

The applications of genetic engineering to medicine may be broadly divided into two subdivisions, involving either diagnosis or treatment. Applications in diagnosis may be prenatal or postnatal. In the case of preimplantation, genetic diagnosis is typically offered to parents who are already known carriers of genetically lethal mutations, including those involving familial predisposition to cancer. After in vitro fertilization, one or two cells are removed from the very early embryo on day 3 at the 8–12 cell stage. Subsequently PCR amplification is carried out on DNA derived from a single cell and the DNA sequence is then investigated for the presence of the mutation. By amplifying several different sequences simultaneously (multiplex PCR), including the sequence known to contain the mutation and one or more containing polymorphic markers that are closely linked to that mutation, the possibility of misdiagnosis is decreased. In the case of X-linked disorders to screen a specific mutant gene, sex determination can be carried out to ensure the implantation of female embryos only. Typically, this is used for the prenatal diagnosis of fragile X syndrome and haemophilia. Fluorescence in situ hybridization (FISH) is used whereby DNA probes that are complementary to sequences on the X and Y chromosomes, as well as a non-sex chromosome sequence as control, are hybridized to the nuclear DNA from a single embryonic cell.

It is apparent that preimplantation genetic diagnosis has the potential for abuse. Indeed, it is not allowed in many countries (for example, Germany, Austria, Switzerland, Argentina), and has only recently been allowed in France, whereas in other countries it is virtually unregulated. In the UK it is a procedure regulated by license from the Human Fertilization and Embryology Authority (HFEA) under the terms of the Human Fertilization and Embryology Act (1990). The central aim of the procedure (as a current UK regulations) is to prevent the birth of children affected with “very serious, life threatening conditions”. The use of preimplantation genetic diagnosis for sex determination outside of this aim is forbidden by the HFEA. As with any new medical technology, ethical decisions are particularly controversial. This was highlighted by the use of preimplantation genetic diagnosis to ensure the birth of a baby boy tissue typed so that he could become a donor of haematopoietic stem cells for his sister who suffered from Fanconi’s anaemia. In the UK, the HFEA recommended in 2001 that preimplantation tissue typing should only be used when an embryo was being screened for an inherited genetic disorder. Ethically it seems wise to focus the use of preimplantation genetic diagnosis on the prevention of births involving lethally destructive genetic mutations.

The postnatal diagnosis of genetic diseases will be greatly facilitated by the sequencing of the human genome. There are already around 1000 documented disease genes out of the approximately 30 000 genes in the genome. Identification of mutant genes that predispose towards disease but do not guarantee it are difficult to handle, particularly if there are no known environmental changes that will lower the risk. A potential abuse of genetic engineering is to give people genetic information about which they can do nothing. There is also the continued risk of creating a “genetic underclass” who are less able to obtain life insurance or loans.

Fortunately, the applications of genetic engineering in the treatment of genetic disease is at last yielding some positive results, albeit modest and not without setbacks. In April 2002, after a gene therapy trial that occurred two years previously, French researchers announced that the immune systems of several children with X-linked severe combined immunodeficiency (SCID) were nearly normal. Out of the 11 children in the trial, nine were cured. Unfortunately, two of the SCID patients later developed a leukaemia-like condition due to T-cell hyperproliferation, caused by integration of the vector into the *LMO-2* gene, mutations “insertional mutagenesis” which are known to be involved in childhood cancers.

The future potential for somatic cell gene therapy remains enormous. The possibility of germ line therapy has also frequently been mooted, but the increasing success of preimplantation genetic diagnosis appears to render this approach unnecessary.

An important focus for genetic engineering continues to be the diagnosis and healing of human disease. Fortunately, the complexity of the genome itself represents a natural defense against such interventions. The techniques of genetic engineering, if used wisely, can continue to bring enormous benefits to humankind.

**Recombinant DNA Technology**

Following the elucidation of the DNA structure and the genetic code, it became clear that many biological secrets were hidden in the sequence of bases in DNA. Technical and biological discoveries in the 1970s led to a new era of DNA analysis and manipulation. Key among these was the discovery of two types of enzymes that made **DNA cloning** possible.Cloning refers to the isolation and amplification of defined pieces of DNA. One enzyme type, called ***restriction enzymes****,* cut the DNA from any organism at specific sequences of a few nucleotides, generating a reproducible set of fragments.

Restriction enzymes occur naturally in many bacteria, where they serve as defense mechanisms against bacteriophage infection by cutting the bacteriophages genome upon its entry into the cell. The other enzyme type, called *DNA ligases,* can covalently join DNA fragments at their termini that have been created by restriction enzymes. Thus, ligases can insert DNA restriction fragments into replicating DNA molecules such as plasmids (bacterial, circular DNA molecules), resulting in recombinant DNA molecules. The recombinant DNA molecules can then be introduced into appropriate host cells, most often bacterial cells. All descendants from such a single cell, called a clone, carry the same recombinant DNA molecule. Once a clone of cells bearing a desired segment of DNA has been isolated, unlimited quantities of this DNA sequence can be prepared. Furthermore, in case the DNA fragment contains protein-coding genes, the recombinant DNA molecule introduced into a suitable host can direct the expression of these genes, resulting in the production of the proteins within the host. These developments, DNA cloning and the production of **recombinant proteins**, were major breakthroughs in molecular biology and set the stage for modern biological research.

**DNA Cloning**

DNA cloning facilitates the isolation and manipulation of fragments of an organism’s genome by replicating them independently as part of an autonomous vector.

Because every organism’s genome is large and complex, and any sequence of interest usually occurs only once or twice per cell. Hence, standard chemical or biochemical methods cannot be used to isolate a specific region of the genome for study, particularly as the required sequence of DNA is chemically identical to all the others. The solution to this problem is to place a relatively short fragment of a genome, which might contain the gene or another sequence of interest, in an autonomously replicating piece of DNA, known as a vector, forming recombinant DNA, which can be replicated independently of the original genome, and normally in another host species altogether. Propagation of the host organism containing the recombinant DNA forms a set of genetically identified organisms, or a clone. This process is known as DNA cloning.

Among the increasing numbers of applications of DNA cloning, often collected together under the term genetic engineering, are as below:

* DNA sequencing, and hence the derivation of protein sequence.
* Isolation and analysis of gene promoters and other control sequences.
* Investigation of protein/ enzyme/ RNA function by large-scale production of normal and altered forms.
* Identification of mutations, for instance gene defects leading to disease.
* Biotechnology; the large scale commercial production of proteins and other molecules of biological importance, for example human insulin and growth hormone.
* Engineering animals and plants, and gene therapy.
* Engineering proteins to alter their properties.

**DNA cloning** involves the following steps:

1- Isolate DNA from an organism.

2- Cut the DNA into pieces with a restriction enzyme-an enzyme that recognizes and cuts within a specific DNA sequence-and insert (ligate) each piece individually into a cloning vector cut with the same restriction enzyme to make a recombinant DNA molecule, a DNA molecule constructed in vitro containing sequences from two or more distinct DNA molecules.

3- Introduce (transform) the recombinant DNA molecule into a host such as *E. coli.* Replication of the recombinant DNA molecule-the process of molecular cloning-occurs in the host cell, producing many identical copies called clones. As the host organism reproduces, the recombinant DNA molecules are passed on to all the progeny, giving rise to a population of cells carrying the cloned sequences.

**Cutting and joining DNA**Two major categories of enzymes are important tools in the isolation of DNA and the preparation of recombinant DNA: restriction endonucleases and DNA ligases. Restriction endonucleases recognize a specific, rather short, nucleotide sequence on a double-stranded DNA molecule, called a restriction site, and cleave the DNA at this recognition site or elsewhere, depending on the type of enzyme. DNA ligase joins two pieces of DNA by forming phosphodiester bonds.

**Some Applications of Genetic Engineering/Applications of Cloning**

There are many applications of gene cloning including: recombinant protein production as therapeutics and for nontherapeutic use, genetically modified organisms especially for improved food production, DNA fingerprinting, diagnostic kits, and gene therapy in an attempt to correct genetic disorders.

**Recombinant Protein**

Many proteins that are normally produced in very small amounts are known to be missing or detective in various disorders. These include growth hormone, insulin in diabetes, interferon in some immune disorders and blood clotting factor VIII in hemophilia. Prior to the advent of gene cloning and protein production via recombinant DNA techniques, it was necessary to purify these molecules from animal tissues or donated human blood. Both sources have drawbacks, including slight functional differences in the nonhuman proteins and possible viral contamination e.g. HIV. Production of protein from a cloned gene in a defined, nonpathogenic organism would circumvent these problems, and so pharmaceutical and biotechnology companies have developed this technology. Initially, production in bacteria was the only route available and cDNA clones were used as they contained no introns. The cDNA had to be linked to prokaryotic transcription and translation signals and inserted into multicopy plasmids. However, often the overproduced proteins, which could represent up to 30% of total cell protein, were precipitated or insoluble and they lacked eukaryotic post-translational modifications. Sometimes these problems could be overcome by making fusion proteins, which were later cleaved to give the desired protein, but the subsequent availability of eukaryotic cells for production (yeast or mammalian cell lines) has helped greatly.

The human factor VIII protein, which is administered to hemophiliacs, is produced in a hamster cell line which has been transfected with a 186 kb human genomic DNA fragment. Such a cell line has been **genetically modified** and it is now possible to genetically modify whole organisms. Recombinant proteins can also be modified by introducing amino acid substitutions by mutagenesis. This can result in improvements such as more stable enzymes for inclusion in washing powders.

**Genetically Modified Organisms**

Genetically modified organisms (GMOs) are created when cloned genes are introduced into single cells, or cells that give rise to whole organisms. In eukaryotes, if the introduced genes are derived from another organism, the resulting transgenic plants or animals can be propagated by normal breeding. Several types of transgenic plant have been created and tested for safety in the production of foodstuffs. One example of GMO is a tomato that has had a gene for a ripening enzyme inactivated. The strain of tomato takes longer to soften, and ultimately rot, due to the absence of the enzyme, and so has a longer shelf life and other improved qualities.

Transgenic sheep have been produced with the intention of producing valuable proteins in their milk. The desired gene requires a sheep promoter (e.g. from casein or lactalbumin) to be attached to ensure expression in the mammary gland. Purification of the protein from milk is easier than from cultured cells or blood. The definition of a transgenic organism is one containing a foreign gene, but the term is often now applied to organisms that have been genetically manipulated to contain multiple foreign genes, extra copies of an endogenous gene or that have had a gene disrupted (gene knockout). The term is not usually applied to the most extreme form of adding foreign genes seen in some forms of animal cloning (production of identified individuals) where all the genes are replaced by those from another nucleus. This procedure (**nuclear transfer**) of replacing the nucleus of an egg with the nucleus from an adult cell was used to create Dolly the sheep in 1997. This famous example of animal cloning created much controversy because it raised the possibility of human cloning from adult cells. Many countries have now introduced laws to ban most types of human cloning and this may well hinder the development of replacement cells/ organs for therapeutic use. Identical twins are human clones that arise naturally.

**DNA Fingerprinting (DNA Typing)**

DNA typing, or DNA fingerprinting, is done to distinguish individuals based on the concept that no two individuals of a species, save for identical twins, have the same genome sequence. The variations are manifested in restriction fragment length polymorphisms and length variations resulting from different numbers of short tandemly repeated sequences.

Cloning and genomic sequencing projects have identified many repetitive sequences in the human genome. Some of these are simple nucleotide repeats that vary in number between individuals but are inherited (VNTRs). If Southern blots of restriction enzyme-digested genomic DNA from members of a family are hybridized with a probe that detects one of these types of repeats, each sample will show a set of bands of varying lengths. Some of these bands will be in common with those of the mother and some with those of the father and the pattern of bands will be different for an unrelated individual. The different patterns in individuals at each of these kinds of simple repeat sites mean that, by using a small number of probes, the likelihood of two individuals having the same pattern becomes vanishingly small. This is the technique of DNA fingerprinting which is used in forensic science to eliminate the innocent and convict criminals. It is also applied for maternity and paternity testing in humans. Additionally, it can be used for analysis of old or ancient DNA. It can also be used to show pedigree in animals bred commercially, and to discover mating habits in wild animals.

DNA fingerprinting can be carried out on small DNA samples such as blood spot or hair follicle left in a crime scene. Instead of digesting the DNA with restriction enzyme at each end of the VNTR and Southern blotting, a pair of PCR primers can be designed based on the unique sequences flanking the repeats. The VNTRs can thus be amplified and directly visualized by staining after agarose gel electrophoresis.

**Strain Typing**

Hospital microbiologists and epidemiologists may need to determine relatedness among bacteria or fungi of the same species to determine if transmission of the organism of interest has taken place in a manner that may be addressed by infection control. For example, if *Pseudomonas aeruginosa* has been isolated from four babies in a neonatal intensive care unit, the hospital epidemiologist may wonder if these are the same bacterial strain that has been transferred from a point source or if they are different strains and the infection has occurred merely by chance. The determination that the bacteria are all of the same strain strengthens the evidence that there is a point source of transmission of the bacteria and that there is a breakdown in techniques or personal hygiene, all of which may be corrected.

Phenotypic methods of strain typing preceded genotypic methods, but many of these lack the high degree of discrimination that may be achieved by genotypic strain typing. Some others, such as bacteriophage typing, require a high degree of specialized expertise and the maintenance of considerable resources (i.e., libraries of bacteriophages). Antimicrobial susceptibility profiling, although not as definitive as some of the genotyping methods of profiling, is valuable, and uses information that is often already available because of susceptibility testing performed to guide therapy. Bacteria that have very different antimicrobial susceptibility profiles are likely not identical and therefore probably do not warrant more expensive genotypic strain typing.

The earliest molecular attempts at demonstrating relatedness among different groups of bacteria were accomplished with simple DNA-DNA hybridization studies. Groups of bacteria with high degrees of DNA homology, as determined through DNA-DNA hybridization studies, were considered to be more closely related compared with groups that had poor DNA homology when compared with one another. For example, the DNA-DNA hybridization studies between *Staphylococcus* and *Micrococcus*, which are in the same family, would demonstrate greater homology than would a comparison between *Staphylococcus* and *Pseudomonas*.

Plasmids are small, circular DNA molecules that are present in many bacteria and exist separately from the chromosomal DNA. Plasmids are important because they may contain genes important for virulence, including genes that impart resistant to antimicrobial agents. Plasmids may also spread from one bacterium to another via conjugation and thereby disseminate these genes. Developments in electrophoresis that allowed for the separation of plasmids from the chromosomal DNA, and the differentiation of plasmids from one another, made possible plasmid profiling as an early means of strain typing. Plasmid profiling was one of the earliest genetic methods used to determine strain relatedness.

There have been many genetic methods described for analyzing the relationship of microorganisms. Methods include ribo-typing, random amplified polymorphic DNA (RAPD) analysis, and spoligotyping, among others. Molecular typing methods may be categorized as those that require nucleic-acid amplification for the typing process (amplification-based methods) and those that do not (non-amplification-based methods).

Amplification-based methods such as 1- PCR-RFLP combines the strengths of PCR with the typing of restriction-fragment-length polymorphism, and 2- Rep-PCR, a relatively new, commercially available technology that uses, for the purpose of typing, repetitive genetic elements that occur naturally in bacteria.

Non-amplification-based methods are exemplified by pulsed-field gel electrophoresis, which is considered by many to be the standard for microbial typing.

**Non-Amplification-Based Typing**

**Pulsed-Field Gel Electrophoresis**

There are many strain typing methods that vary in the degree of distinction that may be detected, cost, and ease of use. Many of the current methods of molecular strain typing rely on the use of a group of enzymes called restriction endonucleases. These enzymes are naturally occurring enzymes that cleave double-stranded DNA based on particular sequences termed restriction sites. Different enzymes cleave based on different sequences, and a variety of restriction endonucleases are commercially available. When the enzymes cleave the DNA, the two ends of the DNA produced may be either blunt or may have overhanging, so-called sticky ends. When microbial chromosomal or plasmid DNA is extracted from the organisms and exposed to one or more restriction endonucleases, it will be cleaved depending on the restriction endonucleases used and the number of restriction sites present in the DNA. The fragments of DNA that result will vary in size, and may be separated from one another by gel electrophoresis and visualized by staining with ethidium bromide. This is an RFLP analysis. Microorganisms with the same RFLP pattern are likely to be closely related, whereas those with very different RFLP patterns are less likely to be closely related. Although RFLP patterns are highly reproducible and very accurate, problems may arise following restriction enzyme digestion. PFGE was devised as a means to simplify RFLP typing.

PFGE is considered by name to be the standard form of strain typing, to which newer methods should be compared. Briefly, this technique begins with the extraction of the chromosomal bacterial DNA of the strain of interest, which is exposed to a variety of restriction endonucleases. The type of restriction endonucleases used is usually determined by the type of bacteria under analysis. The restriction endonucleases cut the chromosomal DNA into a variety of pieces, depending on the number of restriction sites present in that particular strain. The fragments of chromosomal DNA are then separated using a specialized type of electrophoresis. The chromosomal restriction profile resulting from PFGE usually results in 5-20 fragments that range in size from 10-800 kb. These are visualized, photographed, and submitted for image analysis. Computer programs are then used to determine if the strains under analysis are indistinguishable or distinguishable and the degree to which these may be distinguished. The process has been termed DNA fingerprinting when the process of RFLP is combined with Southern blot transfer and probe hybridization.

**Amplification-Based Typing**

There are a number of post-amplification methods of analysis that may be used for microbial typing. These include single-stranded conformation polymorphism analysis, amplified fragment-length polymorphisms, arbitrarily primed PCR, RFLP analysis of the PCR product (PCR-RFLP), and analysis of PCR products of repetitive elements present in bacterial genomes (rep-PCR).

**PCR-RFLP**

**Restriction fragment length polymorphism (RFLP).** In 1980, Mark Skolnick, Ray White, David Botstein, and Ronald Davis created a restriction fragment length polymorphism (RFLP, pronounced “rif-lip”) marker map of the human genome. A RFLP is defined by the existence of alternative alleles associated with restriction fragments that differ in size from each other. RFLPs are visualized by digesting DNA from different individuals with restriction endonucleases, followed by gel electrophoresis to separate fragments according to size, then Southern blotting, and hybridization to a labeled probe that identifies the locus under investigation. A RFLP is demonstrated whenever the Southern blot pattern obtained with one individual is different from the one obtained with another individual. These variable regions do not necessarily occur in genes, and the function of most of those in the human genome is unknown. An exception is a RFLP that can be used to diagnose sickle cell anemia. In individuals with sickle cell anemia, a point mutation in the β-globin gene has destroyed the recognition site for the restriction endonuclease *Mst*II. This mutation can be distinguished by the presence of a larger restriction fragment on a Southern blot in an affected individual, compared with a shorter fragment in a normal individual.

Restriction endonucleases can be used to cleave the amplicon. The selection of the endonuclease(s) used with PCR-RFLP must be chosen with care, since the amplicon is smaller than the bacterial chromosome and is less likely to randomly contain restriction sites. The enzymes are often chosen based on previous knowledge of restriction sites that are variably present in the amplicon and, therefore, may be useful for the comparison of strains. The size of the amplicon and the number of potential restriction sites that it may contain imposes limitations on the differentiating capabilities of this technology.

RFLP analysis of the PCR product is commonly directed against the rDNA. When the PCR product targeting the rDNA is analyzed by RFLP, it may be referred to as PCR ribotyping. Even when the rDNA is analyzed, it is important to target particular areas, to achieve type-specific information. PCR-RFLP has been used to type all different kinds of microorganisms, including bacteria such as *Campylobacter* and *Borrelia* species and atypical fungi such as *Pneumocystis jiroveci*. In addition to strain comparisons, PCR-RFLP patterns may be used for organism identification. The use of this technology in this manner relies on the use of highly specific PCR primers. This technique, for example, has been used to identify fastidious bacteria, such as *Bartonella* species, and difficult-to-identify parasites, such as hookworms and filarial worms. This technology has also proven very useful for the identification of *Nocardia* species, which is clinically important given the differences in antimicrobial susceptibilities within this group. The non-ribosomal genes that may be used for PCR-RFLP include either housekeeping genes, such as *rpoB*, or any gene that may contain strain-specific information for a particular group of organisms (i.e., genes that encode for the enzymes responsible for hippurate hydrolysis for *Campylobacter jejuni*).

**RFLP can serve as marker of genetic diseases**

By carefully examining the DNA of members of families that carry genetic diseases, it has been possible to find forms of particular RFLPs that tend to be inherited with particular diseases. The simplest RFLPs are those caused by single base pair substitutions. However, RFLPs can also be generated by the insertion of genetic material such as transposable elements, or by tandem duplications, deletions, translocations, or other chromosomal rearrangements. In linkage analysis, families in which individuals are at risk for a genetic disease are identified (i.e. both parents are heterozygous for an autosomal recessive mutation associated with a particular disease). DNA samples from various family members are then analyzed to determine the frequency with which specific RFLP markers segregate with the mutant allele causing the disease. This frequency is a measure of the distance between the markers and the mutation-defined locus.

Generally, the fragment size differences occur not because a restriction site was created or disrupted by the disease itself, but rather because the nucleotide sequence differences just happen to be near the gene involved. A particular form of a polymorphism that is close to a diseased gene tends to stay with that gene during crossing-over (recombination) in meiosis. Relatively large segments of chromosomes are involved in crossing-over, so markers close together on a given chromosome are more likely to be transmitted together (not separated during recombination) than those that are far apart. Linkage thus refers to the likelihood of having one marker transmitted with another through meiosis. Markers that are transmitted together frequently are said to be closely linked. Thus RFLPs can serve as markers of disease, even when the RFLP is not within the disease gene. RFLPs have been useful for detecting genetic diseases, such as cystic fibrosis, Huntington’s disease, and hemophilia.

RFLPs were the predominant form of DNA variation used for linkage analysis until the advent of PCR. The main advantage of RFLP analysis over PCR-based protocols is that no prior sequence information or oligonucleotide synthesis is required. However, when a PCR assay for typing a particular locus is developed, it is generally preferable to RFLP analysis. In some cases, a combined approach of PCR and RFLP is used for analysis.

**REP-PCR**

Repetitive DNA elements in prokaryotic genomes form the basis for rep-PCR. These repetitive elements are the hybridization sites for the primers used in rep-PCR. These primers, which are complementary to the interspersed repetitive sequences, generate through PCR amplification variably sized DNA fragments, which may be separated by gel electrophoresis. The electrophoretic patterns may be used to determine the relatedness of bacterial strains that are being compared. A single primer set may be used for a variety of gram-positive and gram-negative bacteria, or more than one primer set may be used to target different repetitive elements. Rep PCR has been used to type entero-hemorrhagic *E. coli* isolates to investigate animal-to-human transmission. It, like other typing techniques, has been used to track both gram-positive and gram-negative bacteria that are resistant to antimicrobial agents and of importance to the hospital epidemiologist. The information generated through rep-PCR may also be used for microorganism identification. This technology has been used to identify organisms and examine diversity with the *Streptomyces* genus and to identify *Bartonella* species.

**Clinical Applications of Microbial Typing**

Genetic comparisons of microorganisms have been used to investigate outbreaks in hospitals, in clinics, and even across the nation. It is common for potential nosocomial outbreaks to be investigated using genetic typing methods. A possibility of an outbreak is often suspected when the frequency of isolation of a pathogenic microorganism exceeds its usual baseline recovery rate. This may be in a particular place, such as the number of MRSA blood isolates in the hospital in general or a ward in particular, or it may be the number of *Salmonella* isolates reported to a county health department.

Given the fragility of premature babies, these techniques are commonly used to investigate outbreaks in neonatal intensive-care units. Molecular typing methods are also commonly used to investigate infections caused by bacteria that are resistant to antimicrobial agents. In some instances, genetic typing methods have been used to successfully determine the point source of an outbreak. Outbreaks and the transmission of *M. tuberculosis* are often studied using molecular typing methods.

**Medical Diagnosis**

A great variety of medical conditions arise from mutation. In genetic disorders such as muscular dystrophy or cystic fibrosis, individuals are born with faulty genes that cause the symptoms of the disorder. Many cancers arise due to spontaneous mutations in somatic cells in genes whose normal role is the regulation of cell growth. Cloning of the genes involved in both genetic disorders and cancers has shown that certain mutations are more common and some correlate with more aggressive disorders. By using sequence information to design PCR primers and probes, many tests have been developed to screen patients for these clinically important mutations. Using these tests, parents who are both heterozygous for a mutation can now be advised whether an unborn child is going to suffer from a genetic disorder such as muscular dystrophy or cystic fibrosis (by inheriting one faulty gene from each parent) and can consider termination. Checking for the presence of mutations in a gene can confirm a diagnosis that is based on other clinical presentations. In cancer cases, knowing which oncogene is mutated, and in what way, can help decide the best course of treatment as well as providing information for the development of new therapies.

**Gene Therapy**

**Gene therapy is the curing of a genetic disorder by introducing into the individual a normal gene to replace or overcome the effects of a mutant gene. For ethical reasons, only somatic gene therapy is being developed for humans. There are few examples of successful somatic gene therapy in humans, but there is great hope for treating many genetic diseases in this way in the future.**

Attempts have been made to treat some genetic disorders by delivering a normal copy of the defective gene to patients. It is possible to modify the genome to treat genetic diseases. Theoretically, two types of gene therapy are possible: somatic cell therapy, in which somatic cells are modified genetically to prevent a genetic defect in the individual receiving the therapy; and germ-line cell therapy, in which germ-line cells are modified to correct a genetic defect. Somatic cell therapy results in a treatment for the genetic disease in the individual, but progeny could still inherit the mutant gene. Germ-line therapy, however, could prevent the disease because the mutant gene can be replaced by the normal gene and that normal gene would be inherited by the offspring. Both somatic cell therapy and germ-line therapy have been used successfully in nonhuman organisms, including mice, but only somatic cell therapy has been used in humans because of ethical issues raised by germ-line cell therapy.

The most offspring candidates for somatic cell therapy are genetic disorders that result from a simple defect of a single gene and for which the cloned normal gene is available. Gene therapy involving somatic cells proceeds as follows. A sample of the individual’s cells carrying the defective gene is taken. Then normal, wild-type copies of the mutant gene are introduced into the cells, and the cells are re-introduced into the individual. There, it is hoped, the cells will produce a normal gene product and the symptoms of the genetic disease will be completely or partially reversed.

The source of the cells varies with the genetic disease. For example, blood disorders, such as thalassemia or sickle-cell anemia, require modification of bone marrow cells that produce blood cells. For genetic diseases affecting circulating proteins, cells that are constituents of the dermis (the lower layer of the skin). Modified fibroblasts can easily be implanted back into the dermis, where blood vessels invade the tissue, allowing gene products to be distributed.

A cell that has had a gene introduced into it by artificial means is said to be **transgenic,** and the gene involved is called a **transgene.** The introduction of normal genes into a mutant cell poses several problems. First, procedures to introduce DNA into cells (transformation, although actually called transfection for eukaryotic cells) typically are inefficient; perhaps only one in 1,000 or 100,000 cells will receive the gene of interest. Thus, a large population of cells is needed to attempt gene therapy. Present procedures use special virus-related vectors to introduce the transgene. Second, in cells that take up the cloned gene, the fate of the foreign DNA cannot be predicted. In some cases, the mutant gene is replaced by the normal gene, and in others the normal gene integrates into the genome elsewhere. In the first case, the gene therapy is successful provided that the gene is expressed. In the second case, successful treatment of the disease results only if the introduced gene is expressed and the resident mutant gene is recessive, so that it does not interfere with the normal gene.

Successful somatic gene therapy has been demonstrated repeatedly in experimental animals such as mice, rats, and rabbits. However, in humans, there have been more failures than successes. In addition, a recent concern is the development of leukemia in therapy patients as a result of the viral vectors used for introducing the transgene.

One of successful human somatic gene therapy treatments was done in 1990 with a 4-year-old girl suffering from severe combined immunodeficiency caused by a deficiency in adenosine deaminase (ADA), an enzyme needed for normal function of the immune system. T cells (cells involved in the immune response) were isolated from the girl and grown in the laboratory, and the normal ADA gene was introduced using a viral vector. The engineered cells were then reintroduced into the patient. Since T cells have a finite life in the body, continued infusions of engineered cells have been necessary. The introduced ADA gene is expressed, probably throughout the life of the T cell. As a result, the patient’s immune system is functioning more normally, and she now gets more than the average number of infections. The gene therapy treatment has enabled her to live a more normal life. Recently, some patients who received gene therapy for ADA have developed leukemia for reasons unknown.

With time, many other genetic diseases are expected to be treatable with somatic gene therapy, including thalassemia, phenylketonuria, cancer, Duchenne muscular dystrophy, and cystic fibrosis. For example, after successful experiments with rats, human clinical trials are under way for transferring the normal CF gene to patients with cystic fibrosis. Using methods for targeting genes to replace their mutant counterparts and regulating the expression of the introduced genes, increasing success in treating genetic diseases is expected. However, many scientific, ethical, and legal questions must be addressed before the routine implementation of gene therapy.

**Commercial Products: Biotechnology**

The development of cloning and other DNA manipulation techniques has initiated the formation of many biotechnology companies, some of which focus on using DNA manipulations for making a wide array of commercial products. Although the details vary, the general approach to making a product is to express a cloned gene or cDNA in an organism that will transcribe the cloned sequence and translate the mRNA. The gene or cDNA is placed into an expression vector appropriate for the organism into which it will be transformed. Many different organisms are used, from *E. coli* to mammals, so the expression vectors differ in the promoters used for transcription, in the translation start signals, and in the selectable markers. In *E. coli,* for example, the promoter must be recognized by that bacterium’s RNA polymerase, and there must be a Shine-Dalgarno sequence so that ribosomes will read the mRNA from the correct AUG. In mammals such as goats or sheep, the simplest way to isolate the product is to have it secreted into the milk. The milk is easy to collect and the protein product can then be extracted. The production of recombinant protein products in transgenic mammals (in this case sheep) is illustrated in the Figure below.

Here the gene of interest (GOI) has been manipulated so that it is adjacent to a promoter that is active only in mammary tissue, such as the ß-lactoglobulin promoter. The recombinant DNA molecules are microinjected into sheep ova, and each ovum is then implanted into a foster mother. Transgenic offspring are identified using PCR to detect the recombinant DNA sequences. When these transgenic animals mature, the ß-lactoglobulin promoter begins to express the associated gene in the mammary tissue, the milk is collected, and the protein of interest is obtained by biochemical separation techniques.

A few examples of the many products produced by biotechnology companies are as follows:

1. Tissue plasminogen activator (TPA), used to prevent or dissolve blood clots, therefore preventing strokes, heart attacks, or pulmonary embolisms.
2. Human growth hormone, used to treat pituitary dwarfism.
3. Tissue growth factor-beta (TGF- ß), which promotes new blood vessel and epidermal growth and thus is potentially useful for wound and burn healing.
4. Human blood clotting factor VIII, used to treat hemophilia.
5. Human insulin (“humulin”), used to treat insulin-dependent diabetes.
6. DNase, used to treat cystic fibrosis.
7. Recombinant vaccines, used to prevent human and animal viral diseases (such as hepatitis B in humans).
8. Bovine growth hormone, used to increase cattle and dairy yields.
9. Platelet-derived growth factor (PDGF), used to treat chronic skin ulcers in patients with diabetes.
10. Genetically engineered bacteria and other microorganisms used to improve production of, for instance, industrial enzymes (such as amylases to break down starch to glucose), citric acid (flavoring), and ethanol.
11. Genetically engineered bacteria that can accelerate the degradation of oil pollutants or certain chemicals in toxic wastes (such as dioxin).

**Genetic Engineering of Plants**

For many centuries the traditional genetic engineering of plants involved selective breeding experiments in which plants with desirable traits were selectively allowed to produce offspring. As a result, humans have produced hardy varieties of plants (e.g., corn, wheat, and oats) and increased yields, all using long-established plant breeding techniques.

Similar techniques have also been used with animals, such as dogs, cattle, and horses, to produce desired breeds. Now, vectors developed by recombinant DNA technology are available for transforming cells of crop plants; this has made possible the genetic engineering of plants for agricultural use.