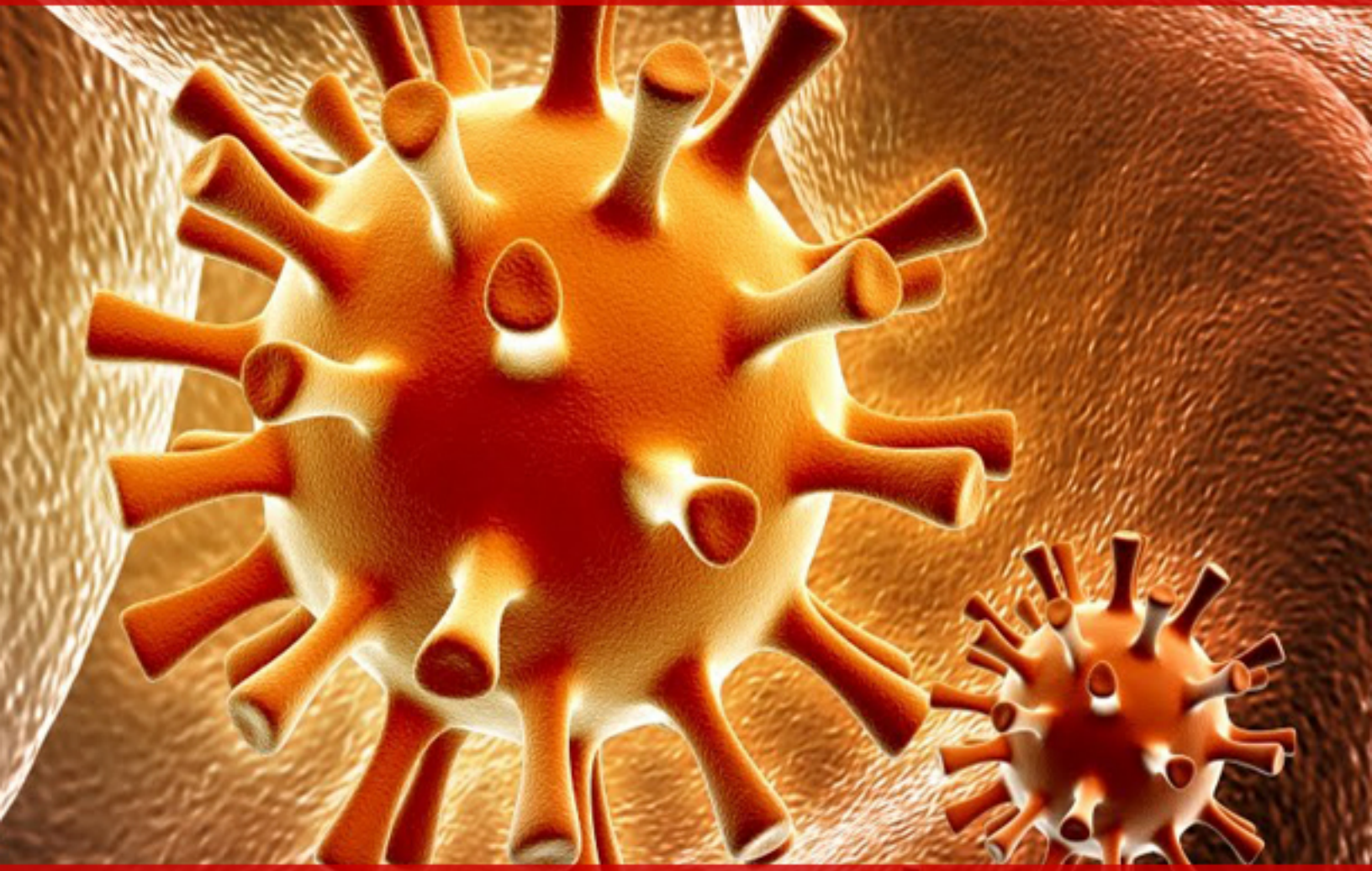
Handbook of Genetic Engineering



Stanford Glover

Gisela Crumpton

First Edition, 2012

ISBN 978-81-323-0914-7

Published by:
Academic Studio
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter- 1

Genetic Engineering

Genetic engineering, also called **genetic modification**, is the direct human manipulation of an organism's genetic material in a way that does not occur under natural conditions. It involves the use of recombinant DNA techniques, but does not include traditional animal and plant breeding or mutagenesis. Any organism that is generated using these techniques is considered to be a genetically modified organism. The first organisms genetically engineered were bacteria in 1973 and then mice in 1974. Insulin producing bacteria were commercialized in 1982 and genetically modified food has been sold since 1994.

The most common form of genetic engineering involves the insertion of new genetic material at an unspecified location in the host genome. This is accomplished by isolating and copying the genetic material of interest, generating a construct containing all the genetic elements for correct expression, and then inserting this construct into the host organism. Other forms of genetic engineering include gene targeting and knocking out specific genes via engineered nucleases such as zinc finger nucleases or engineered homing endonucleases.

Genetic engineering techniques have been applied in numerous fields including research, biotechnology, and medicine. Medicines such as insulin and human growth hormone are now produced in bacteria, experimental mice such as the oncomouse and the knockout mouse are being used for research purposes and insect resistant and/or herbicide tolerant crops have been commercialized. Genetically engineered plants and animals capable of producing biotechnology drugs more cheaply than current methods (called pharming) are also being developed and in 2009 the FDA approved the sale of the pharmaceutical protein antithrombin produced in the milk of genetically engineered goats.

Definition

Genetic engineering alters the genetic makeup of an organism using techniques that introduce heritable material prepared outside the organism either directly into the host or into a cell that is then fused or hybridized with the host. This involves using recombinant nucleic acid (DNA or RNA) techniques to form new combinations of heritable genetic material followed by the incorporation of that material either indirectly through a vector system or directly through micro-injection, macro-injection and micro-encapsulation techniques. Genetic engineering does not include traditional animal and plant breeding, in vitro fertilisation, induction of polyploidy, mutagenesis and cell fusion techniques that do

not use recombinant nucleic acids or a genetically modified organism in the process. Cloning and stem cell research, although not considered genetic engineering, are closely related and genetic engineering can be used within them. Synthetic biology is an emerging discipline that takes genetic engineering a step further by introducing artificially synthesized genetic material from raw materials into an organism.

If genetic material from another species is added to the host, the resulting organism is called transgenic. If genetic material from the same species or a species that can naturally breed with the host is used the resulting organism is called cisgenic. Genetic engineering can also be used to remove genetic material from the target organism, creating a knock out organism. In Europe genetic modification is synonymous with genetic engineering while within the United States of America it can also refer to conventional breeding methods.

History

Humans have altered the genomes of species for thousands of years through artificial selection and more recently mutagenesis. Genetic engineering as the direct manipulation of DNA by humans outside breeding and mutations has only existed since the 1970s. The term "genetic engineering" was first coined by Jack Williamson in his science fiction novel *Dragon's Island*, published in 1951, one year before DNA's role in heredity was confirmed by Alfred Hershey and Martha Chase, and two years before James Watson and Francis Crick showed that the DNA molecule has a double-helix structure.

In 1972 Paul Berg created the first recombinant DNA molecules by combined DNA from the monkey virus SV40 with that of the lambda virus. In 1973 Herbert Boyer and Stanley Cohen created the first transgenic organism by inserting antibiotic resistance genes into the plasmid of an *E. coli* bacterium. A year later Rudolf Jaenisch created a transgenic mouse by introducing foreign DNA into its embryo, making it the world's first transgenic animal. In 1976 Genentech, the first genetic engineering company was founded by Herbert Boyer and Robert Swanson and a year later and the company produced a human protein (somatostatin) in *E.coli*. Genentech announced the production of genetically engineered human insulin in 1978. In 1980, the U.S. Supreme Court in the Diamond v. Chakrabarty case ruled that genetically altered life could be patented. The insulin produced by bacteria, branded humulin, was approved for release by the Food and Drug Administration in 1982.

The first field trials of genetically engineered plants occurred in France and the USA in 1986, tobacco plants were engineered to be resistant to herbicides. The People's Republic of China was the first country to commercialize transgenic plants, introducing a virus-resistant tobacco in 1992. In 1994 Calgene attained approval to commercially release the Flavr Savr tomato, a tomato engineered to have a longer shelf life. In 1994, the European Union approved tobacco engineered to be resistant to the herbicide bromoxynil, making it the first genetically engineered crop commercialized in Europe. In 1995, Bt Potato was approved safe by the Environmental Protection Agency, making it the first pesticide producing crop to be approved in the USA. In 2009 11 transgenic crops were grown

commercially in 25 countries, the largest of which by area grown were the USA, Brazil, Argentina, India, Canada, China, Paraguay and South Africa.

In 2010, scientists at the J. Craig Venter Institute, announced that they had created the first synthetic bacterial genome, and added it to a cell containing no DNA. The resulting bacterium, named Synthia, was the world's first synthetic life form.

Process

Isolating the Gene

Elements of genetic engineering

First, the gene to be inserted into the genetically modified organism must be chosen and isolated. Presently, most genes transferred into plants provide protection against insects or tolerance to herbicides. In animals the majority of genes used are growth hormone genes. Once chosen the genes must be isolated. This typically involves multiplying the gene using polymerase chain reaction (PCR). If the chosen gene or the donor organism's genome has been well studied it may be present in a genetic library. If the DNA sequence is known, but no copies of the gene are available, it can be artificially synthesized. Once isolated, the gene is inserted into a bacterial plasmid.

Constructs

The gene to be inserted into the genetically modified organism must be combined with other genetic elements in order for it to work properly. The gene can also be modified at this stage for better expression or effectiveness. As well as the gene to be inserted most constructs contain a promoter and terminator region as well as a selectable marker gene. The promoter region initiates transcription of the gene and can be used to control the location and level of gene expression, while the terminator region ends transcription. The selectable marker, which in most cases confers antibiotic resistance to the organism it is expressed in, is needed to determine which cells are transformed with the new gene. The constructs are made using recombinant DNA techniques, such as restriction digests, ligations and molecular cloning.

Gene Targeting

The most common form of genetic engineering involves inserting new genetic material randomly within the host genome. Other techniques allow new genetic material to be inserted at a specific location in the host genome or generate mutations at desired genomic loci capable of knocking out endogenous genes. The technique of gene targeting uses homologous recombination to target desired changes to a specific endogenous gene. This tends to occur at a relatively low frequency in plants and animals and generally requires the use of selectable markers. The frequency of gene targeting can be greatly enhanced with the use of engineered nucleases such as zinc finger nucleases, engineered homing endonucleases, or nucleases created from TAL effectors. In addition to enhancing gene targeting, engineered nucleases can also be used to introduce mutations at endogenous genes that generate a gene knockout.

Transformation

A. tumefaciens attaching itself to a carrot cell

About 1% of bacteria are naturally able to take up foreign DNA but it can also be induced in other bacteria. Stressing the bacteria for example, with a heat shock or an electric

shock, can make the cell membrane permeable to DNA that may then incorporate into their genome or exist as extrachromosomal DNA. DNA is generally inserted into animal cells using microinjection, where it can be injected through the cells nuclear envelope directly into the nucleus or through the use of viral vectors. In plants the DNA is generally inserted using *Agrobacterium*-mediated recombination or biolistics.

In *Agrobacterium*-mediated recombination the plasmid construct must also contain T-DNA. *Agrobacterium* naturally inserts DNA from a tumor inducing plasmid into any susceptible plant's genome it infects, causing crown gall disease. The T-DNA region of this plasmid is responsible for insertion of the DNA. The genes to be inserted are cloned into a binary vector, which contains T-DNA and can be grown in both *E. Coli* and *Agrobacterium*. Once the binary vector is constructed the plasmid is transformed into *Agrobacterium* containing no plasmids and plant cells are infected. The *Agrobacterium* will then naturally insert the genetic material into the plant cells.

In biolistics particles of gold or tungsten are coated with DNA and then shot into young plant cells or plant embryos. Some genetic material will enter the cells and transform them. This method can be used on plants that are not susceptible to *Agrobacterium* infection and also allows transformation of plant plastids. Another transformation method for plant and animal cells is electroporation. Electroporation involves subjecting the plant or animal cell to an electric shock, which can make the cell membrane permeable to plasmid DNA. In some cases the electroporated cells will incorporate the DNA into their genome. Due to the damage caused to the cells and DNA the transformation efficiency of biolistics and electroporation is lower than agrobacterial mediated transformation and microinjection.

Selection

Not all the organism's cells will be transformed with the new genetic material; in most cases a selectable marker is used to differentiate transformed from untransformed cells. If a cell has been successfully transformed with the DNA it will also contain the marker gene. By growing the cells in the presence of an antibiotic or chemical that selects or marks the cells expressing that gene it is possible to separate the transgenic events from the non-transgenic. Another method of screening involves using a DNA probe that will only stick to the inserted gene. A number of strategies have been developed that can remove the selectable marker from the mature transgenic plant.

Regeneration

As often only a single cell is transformed with genetic material the organism must be regrown from that single cell. As bacteria consist of a single cell and reproduce clonally regeneration is not necessary. In plants this is accomplished through the use of tissue culture. Each plant species has different requirements for successful regeneration through tissue culture. If successful an adult plant is produced that contains the transgene in every cell. In animals it is necessary to ensure that the inserted DNA is present in the embryonic stem cells. When the offspring is produced they can be screened for the

presence of the gene. All offspring from the first generation will be heterozygous for the inserted gene and must be mated together to produce a homozygous animal.

Confirmation

Further tests using PCR, Southern Blots and Bioassays are needed to confirm that the gene is expressed and functions correctly. The organism's offspring are also tested to ensure that the trait can be inherited and that it follows a Mendelian inheritance pattern.

Applications

Genetic engineering has applications in medicine, research, industry and agriculture and can be used on a wide range of plants, animals and micro organism.

Medicine

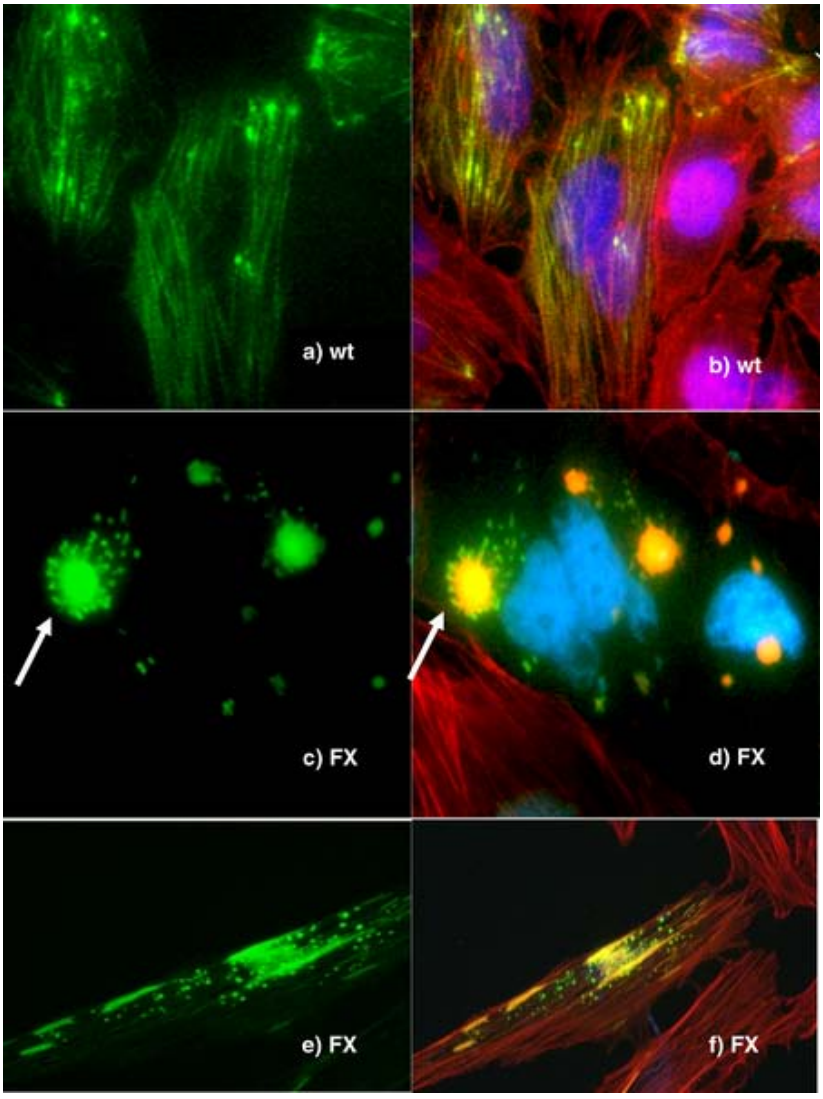
In medicine genetic engineering has been used to mass-produce insulin, human growth hormones, follistim (for treating infertility), human albumin, monoclonal antibodies, antihemophilic factors, vaccines and many other drugs. Vaccination generally involves injecting weak live, killed or inactivated forms of viruses or their toxins into the person being immunized. Genetically engineered viruses are being developed that can still confer immunity, but lack the infectious sequences. Mouse hybridomas, cells fused together to create monoclonal antibodies, have been humanised through genetic engineering to create human monoclonal antibodies.

Genetic engineering is used to create animal models of human diseases. Genetically modified mice are the most common genetically engineered animal model. They have been used to study and model cancer (the oncomouse), obesity, heart disease, diabetes, arthritis, substance abuse, anxiety, aging and Parkinson disease. Potential cures can be tested against these mouse models. Also genetically modified pigs have been bred with the aim of increasing the success of pig to human organ transplantation.

Gene therapy is the genetic engineering of humans by replacing defective human genes with functional copies. This can occur in somatic tissue or germline tissue. If the gene is inserted into the germline tissue it can be passed down to that person's descendants. Gene therapy has been used to treat patients suffering from immune deficiencies (notably Severe combined immunodeficiency) and trials have been carried out on other genetic disorders. The success of gene therapy so far has been limited and a patient (Jesse Gelsinger) has died during a clinical trial testing a new treatment. There are also ethical concerns should the technology be used not just for treatment, but for enhancement, modification or alteration of a human beings' appearance, adaptability, intelligence, character or behavior. The distinction between cure and enhancement can also be difficult to establish. Transhumanists consider the enhancement of humans desirable.

Research

Knockout mice



Human cells in which some proteins are fused with green fluorescent protein to allow them to be visualised

Genetic engineering is an important tool for natural scientists. Genes and other genetic information from a wide range of organisms are transformed into bacteria for storage and modification, creating genetically modified bacteria in the process. Bacteria are cheap,

easy to grow, clonal, multiply quickly, relatively easy to transform and can be stored at -80°C almost indefinitely. Once a gene is isolated it can be stored inside the bacteria providing an unlimited supply for research.

Organisms are genetically engineered to discover the functions of certain genes. This could be the effect on the phenotype of the organism, where the gene is expressed or what other genes it interacts with. These experiments generally involve loss of function, gain of function, tracking and expression.

- **Loss of function experiments**, such as in a gene knockout experiment, in which an organism is engineered to lack the activity of one or more genes. A knockout experiment involves the creation and manipulation of a DNA construct *in vitro*, which, in a simple knockout, consists of a copy of the desired gene, which has been altered such that it is non-functional. Embryonic stem cells incorporate the altered gene, which replaces the already present functional copy. These stem cells are injected into blastocysts, which are implanted into surrogate mothers. This allows the experimenter to analyze the defects caused by this mutation and thereby determine the role of particular genes. It is used especially frequently in developmental biology. Another method, useful in organisms such as Drosophila (fruit fly), is to induce mutations in a large population and then screen the progeny for the desired mutation. A similar process can be used in both plants and prokaryotes.
- **Gain of function experiments**, the logical counterpart of knockouts. These are sometimes performed in conjunction with knockout experiments to more finely establish the function of the desired gene. The process is much the same as that in knockout engineering, except that the construct is designed to increase the function of the gene, usually by providing extra copies of the gene or inducing synthesis of the protein more frequently.
- **Tracking experiments**, which seek to gain information about the localization and interaction of the desired protein. One way to do this is to replace the wild-type gene with a 'fusion' gene, which is a juxtaposition of the wild-type gene with a reporting element such as green fluorescent protein (GFP) that will allow easy visualization of the products of the genetic modification. While this is a useful technique, the manipulation can destroy the function of the gene, creating secondary effects and possibly calling into question the results of the experiment. More sophisticated techniques are now in development that can track protein products without mitigating their function, such as the addition of small sequences that will serve as binding motifs to monoclonal antibodies.
- **Expression studies** aim to discover where and when specific proteins are produced. In these experiments, the DNA sequence before the DNA that codes for a protein, known as a gene's promoter, is reintroduced into an organism with the protein coding region replaced by a reporter gene such as GFP or an enzyme that catalyzes the production of a dye. Thus the time and place where a particular protein is produced can be observed. Expression studies can be taken a step further by altering the promoter to find which pieces are crucial for the proper

expression of the gene and are actually bound by transcription factor proteins; this process is known as promoter bashing.

Industrial

By engineering genes into bacterial plasmids it is possible to create a biological factory that can produce proteins and enzymes. Some genes do not work well in bacteria, so yeast, a eukaryote, can also be used. Bacteria and yeast factories have been used to produce medicines such as insulin, human growth hormone, and vaccines, supplements such as tryptophan, aid in the production of food (chymosin in cheese making) and fuels. Other applications involving genetically engineered bacteria being investigated involve making the bacteria perform tasks outside their natural cycle, such as cleaning up oil spills, carbon and other toxic waste.

Agriculture

Bt-toxins present in peanut leaves (bottom image) protect it from extensive damage caused by European corn borer larvae (top image).

One of the best-known and controversial applications of genetic engineering is the creation of genetically modified food. There are three generations of genetically modified crops. First generation crops have been commercialized and most provide protection from

insects and/or resistance to herbicides. There are also fungal and virus resistant crops developed or in development. They have been developed to make the insect and weed management of crops easier and can indirectly increase crop yield.

The second generation of genetically modified crops being developed aim to directly improve yield by improving salt, cold or drought tolerance and to increase the nutritional value of the crops. The third generation consists of pharmaceutical crops, crops that contain edible vaccines and other drugs. Some agriculturally important animals have been genetically modified with growth hormones to increase their size while others have been engineered to express drugs and other proteins in their milk.

The genetic engineering of agricultural crops can increase the growth rates and resistance to different diseases caused by pathogens and parasites. This is beneficial as it can greatly increase the production of food sources with the usage of fewer resources that would be required to host the world's growing populations. These modified crops would also reduce the usage of chemicals, such as fertilizers and pesticides, and therefore decrease the severity and frequency of the damages produced by these chemical pollution.

Ethical and safety concerns have been raised around the use of genetically modified food. A major safety concern relates to the human health implications of eating genetically modified food, in particular whether toxic or allergic reactions could occur. Gene flow into related non-transgenic crops, off target effects on beneficial organisms and the impact on biodiversity are important environmental issues. Ethical concerns involve religious issues, corporate control of the food supply, intellectual property rights and the level of labeling needed on genetically modified products.

Other uses

In materials science, a genetically modified virus has been used to construct a more environmentally friendly lithium-ion battery. Some bacteria have been genetically engineered to create black and white photographs while others have potential to be used as sensors by expressing a fluorescent protein under certain environmental conditions. Genetic engineering is also being used to create BioArt and novelty items such as blue roses, and glowing fish.

Opposition and criticism

A 2010 study of Canola found transgenes in 80% of wild (uncultivated or "feral") varieties in North Dakota, meaning 80% of the plants which had established themselves in the area were genetically engineered varieties. The researchers stated that "we found the highest densities of [such transgene-containing] plants near agricultural fields and along major freeways, but we were also finding plants in the middle of nowhere" adding that "over time,..the build-up of different types of herbicide resistance in feral [natural] canola and closely related weeds, like field mustard, could make it more difficult to manage these plants using herbicides."

Chapter- 2

Gene Targeting

A chimeric mouse gene targeted for the agouti coat color gene, with its offspring

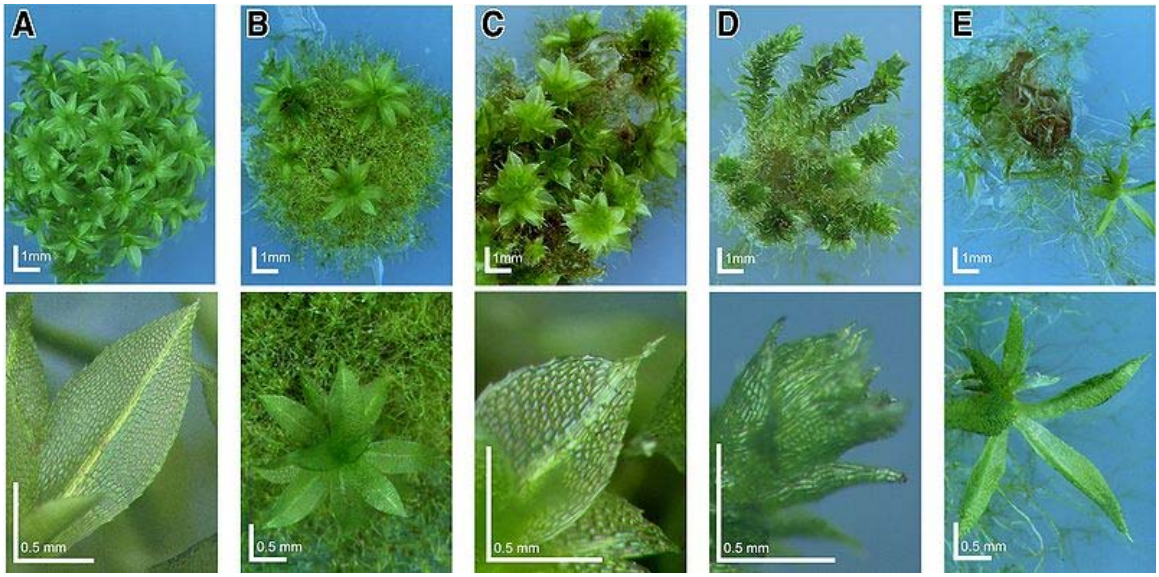
Gene targeting (also, replacement strategy based on homologous recombination) is a genetic technique that uses homologous recombination to change an endogenous gene. The method can be used to delete a gene, remove exons, add a gene, and introduce point

mutations. Gene targeting can be permanent or conditional. Conditions can be a specific time during development / life of the organism or limitation to a specific tissue, for example. Gene targeting requires the creation of a specific vector for each gene of interest. However, it can be used for any gene, regardless of transcriptional activity or gene size.

Methods

Gene targeting methods are established for several model organisms and may vary depending on the species used. In general, a targeting construct made out of DNA is generated in bacteria. It typically contains part of the gene to be targeted, a reporter gene, and a (dominant) selectable marker.

To target genes in mice, this construct is then inserted into mouse embryonic stem cells in culture. After cells with the correct insertion have been selected, they can be used to contribute to a mouse's tissue via embryo injection. Finally, chimeric mice where the modified cells made up the reproductive organs are selected for via breeding. After this step the entire body of the mouse is based on the previously selected embryonic stem cell.



Wild-type Physcomitrella and knockout-mosses: Deviating phenotypes induced in gene-disruption library transformants. Physcomitrella wild-type and transformed plants were grown on minimal Knop medium to induce differentiation and development of gametophores. For each plant, an overview (upper row, scale bar corresponds to 1 mm) and a close-up (bottom row, scale bar equals 0.5 mm) is shown. A, Haploid wild-type moss plant completely covered with leafy gametophores and close-up of wild-type leaf. B-D, Different Mutants.

To target genes in moss, this construct is incubated together with freshly isolated protoplasts and with Polyethylene glycol. As mosses are haploid organisms, regenerating moss filaments (protonema) can directly be screened for gene targeting, either by

treatment with antibiotics or with PCR. Unique among plants, this procedure for reverse genetics is as efficient as in yeast. Using modified procedures, gene targeting has also been successfully applied to cattle, sheep, swine, and many fungi.

Comparison with gene trapping

Gene trapping is a high-throughput approach that is used to introduce insertional mutations across the mammalian genome. It is performed with gene trap vectors whose principal element is a gene trapping cassette consisting of a promoterless reporter gene and/or selectable genetic marker flanked by an upstream 3' splice site (splice acceptor; SA) and a downstream transcriptional termination sequence (polyadenylation sequence; polyA). When inserted into an intron of an expressed gene, the gene trap cassette is transcribed from the endogenous promoter of that gene in the form of a fusion transcript in which the exon(s) upstream of the insertion site is spliced in frame to the reporter/selectable marker gene. Since transcription is terminated prematurely at the inserted polyadenylation site, the processed fusion transcript encodes a truncated and non-functional version of the cellular protein and the reporter/selectable marker. Thus, gene traps simultaneously inactivate and report the expression of the trapped gene at the insertion site, and provide a DNA tag (gene trap sequence tag, GTST) for the rapid identification of the disrupted gene. An international public consortium International Gene Trap Consortium is centralizing the data and cell lines can be requested from them.

Applications

Gene targeting has been widely used to study human genetic diseases by removing "knock-out", or adding "knock-in", specific mutations of interest to a variety of models. Previously used to engineer rat cell models, advances in gene targeting technologies are enabling the creation of a new wave of isogenic human disease models. These models are the most accurate in-vitro models available to researchers to date, and are facilitating the development of new personalised drugs and diagnostics, particularly in the field of cancer.

2007 Nobel prize

Mario R. Capecchi, Martin J. Evans and Oliver Smithies were declared laureates of the 2007 Nobel Prize in Physiology or Medicine for their work on "principles for introducing specific gene modifications in mice by the use of embryonic stem cells", or gene targeting.

Chapter- 3

Transformation

In molecular biology **transformation** is the genetic alteration of a cell resulting from the direct uptake, incorporation and expression of exogenous genetic material (exogenous DNA) from its surrounding and taken up through the cell membrane(s). Transformation occurs most commonly in bacteria and in some species occurs naturally. Transformation can also be effected by artificial means. Bacteria that are capable of being transformed, whether naturally or artificially, are called competent. Transformation is one of three processes by which a exogenous genetic material may be introduced into bacterial cell, the other two being conjugation (transfer of genetic material between two bacterial cells in direct contact), and transduction (injection of foreign DNA by a bacteriophage into the host). Transformation may also used to describe the insertion of new genetic material into nonbacterial cells including animal and plant cells, however, because transformation has a special meaning in relation to animal cells indicating progression to a cancerous state, the term should be avoided for animal cells when describing introduction of exogenous genetic material. Introduction of foreign DNA into eukaryotic cells is usually called "transfection".

History

Transformation was first demonstrated in 1928 by Frederick Griffith, an English bacteriologist searching for a vaccine against bacterial pneumonia. Griffith discovered that a harmless strain of *Streptococcus pneumoniae* could be made virulent after being exposed to heat-killed virulent strains. Griffith hypothesized that some "transforming factor" from the heat-killed strain was responsible for making the harmless strain virulent. In 1944 this "transforming factor" was identified as being genetic by Oswald Avery, Colin MacLeod, and Maclyn McCarty. They isolated DNA from a virulent strain of *S. pneumoniae* and using just this DNA were able to make a harmless strain virulent. They called this uptake and incorporation of DNA by bacteria "transformation."

The results of Avery et al.'s experiments were at first sceptically received by the scientific community and it was not until the development of genetic markers and the discovery of other methods of genetic transfer (conjugation in 1947 and transduction in 1953) by Joshua Lederberg that Avery's experiments were accepted. Transformation did not become routine procedure in laboratories until 1972 when Stanley Cohen, Annie Chang and Leslie Hsu successfully transformed *Escherichia coli* by treating the bacteria with

calcium chloride. This created an efficient and convenient procedure for transforming bacteria and opened the way for molecular cloning in biotechnology and research.

Transformation using electroporation was developed in the late 1980s thus increasing the efficiency and number of bacterial strains that could be transformed. Transformation of animal and plant cells was also investigated with the first transgenic mouse being created by injecting a gene for a rat growth hormone into a mouse embryo in 1982. In 1907 a bacterium that caused plant tumors, *Agrobacterium tumefaciens*, was discovered and in the early 1970s the tumor inducing agent was found to be a DNA plasmid called the Ti plasmid. By removing the genes in the plasmid that caused the cancer and adding in novel genes researchers were able to infect plants with *A. tumefaciens* and let the bacteria insert their chosen DNA into the genomes of the plants. Not all plant cells are susceptible to infection by *A. tumefaciens* so other methods were developed including electroporation and micro-injection. Particle bombardment was made possible with the invention of the Biolistic Particle Delivery System (gene gun) by John Sanford in 1990.

Mechanisms

Bacteria

Bacteria transformation may be referred to as a stable genetic change brought about by the uptake of naked DNA (DNA without associated cells or proteins) and competence refers to the state of being able to take up exogenous DNA from the environment. Two forms of competence exist: natural and artificial.

Natural competence

In microbiology, genetics, cell biology and molecular biology, **competence** is the ability of a cell to take up extracellular ("naked") DNA from its environment. Competence is distinguished into *natural competence*, a genetically specified ability of bacteria that is thought to occur under natural conditions as well as in the laboratory, and *induced* or *artificial competence*, arising when cells in laboratory cultures are treated to make them transiently permeable to DNA. This article is mainly about natural competence in bacteria. Information about artificial competence is provided in the Transformation (genetics) article.

History

Natural competence was discovered by Frederick Griffith in 1928, when he showed that a preparation of killed cells of a pathogenic bacterium contained something that could transform related non-pathogenic cells into the pathogenic type. In 1944 Oswald Avery, Colin MacLeod, and Maclyn McCarty demonstrated that this 'transforming factor' was pure DNA. This was the first compelling evidence that DNA carries the genetic information of the cell.

Since then, natural competence has been studied in a number of different bacteria, particularly *Bacillus subtilis*, *Streptococcus pneumoniae* (Griffith's "pneumococcus"), *Neisseria gonorrhoeae* and *Haemophilus influenzae*. Areas of active research include the mechanisms of DNA transport, the regulation of competence in different bacteria, and the evolutionary function of competence.

Mechanisms of DNA uptake

In the natural world DNA usually becomes available by death and lysis of other cells, but in the laboratory it is provided by the researcher, often as a genetically engineered fragment or plasmid. During uptake, DNA is transported across the cell membrane(s), and the cell wall if one is present. Once the DNA is inside the cell it may be degraded to nucleotides, which are reused for DNA replication and other metabolic functions. Alternatively it may be recombined into the cell's genome by its DNA repair enzymes. If this recombination changes the cell's genotype the cell is said to have been transformed. Artificial competence and transformation are used as research tools in many organisms.

Almost all naturally competent bacteria use components of extracellular filaments called type 4 pili (a type of fimbria) to create pores in their membranes and pull DNA through the pores into the cytoplasm. Some bacteria cut the DNA into short pieces before transporting it; others can take up very long intact fragments and circular plasmids. The details of the uptake machinery are not yet well characterized in any system.

Regulation of competence

In laboratory cultures natural competence is usually tightly regulated and often triggered by nutritional shortages or adverse conditions. However the specific inducing signals and regulatory machinery are much more variable than the uptake machinery, and little is known about the regulation of competence in the natural environments of these bacteria. In bacteria capable of forming spores, conditions inducing sporulation often overlap with those inducing competence. Thus cultures or colonies containing sporulating cells often also contain competent cells. Recent research by Süel et al. has identified an excitable core module of genes which can explain entry into and exit from competence when cellular noise is taken into account.

Most competent bacteria are thought to take up all DNA molecules with roughly equal efficiencies, but bacteria in the families Neisseriaceae and Pasteurellaceae preferentially take up DNA fragments containing short DNA sequences that are very frequent in their own genomes. Neisserial genomes contain thousands of copies of the preferred sequence ATGCCGTCTGAA, and Pasteurellacean genomes contain either AAGTGCGGT or ACAAGCGGT, with corresponding bases of their DNA uptake machinery.

Evolutionary functions and consequences of competence

The evolutionary functions of natural competence are controversial. Competence has conventionally been viewed as a mechanism that cells evolved to provide themselves

with novel genetic information. However the theoretical difficulties associated with the evolution of sex suggest that this explanation is problematic. Cells that take up DNA inevitably acquire the nucleotides the DNA consists of, and, because nucleotides are needed for DNA and RNA synthesis and are expensive to synthesize, these may make a significant contribution to the cell's energy budget. In principle, competence could also allow cells to replace heavily damaged DNA in the cell's genome if needed.

Regardless of the nature of selection for competence, the composite nature of bacterial genomes provides abundant evidence that the lateral gene transfer caused by competence contributes to the genetic diversity that makes evolution possible.

Artificial competence

Artificial competence is induced by laboratory procedures and involves making the cell passively permeable to DNA by exposing it to conditions that do not normally occur in nature.

Calcium chloride transformation is a method of promoting competence. Chilling cells in the presence of divalent cations such as Ca^{2+} (in $CaCl_2$) prepares the cell membrane to become permeable to plasmid DNA. The cells are incubated on ice with the DNA and then briefly heat shocked (e.g., 42°C for 30–120 seconds) thus allowing the DNA to enter the cells. This method works very well for circular plasmid DNA. An excellent preparation of competent cells will give ~10^8 colonies per microgram of plasmid. A poor preparation will be about 10^4/μg or less. Good, non-commercial preparations should give 10^5 to 10^6 transformants per microgram of plasmid. The method, however, usually does not work well for linear DNA, such as fragments of chromosomal DNA, probably because the cell's native exonuclease enzymes rapidly degrade linear DNA. Interestingly, cells that are naturally competent are usually transformed more efficiently with linear DNA than with plasmid DNA.

Electroporation is another method of promoting competence. In the method the cells are briefly shocked with an electric field of 10-20 kV/cm that creates holes in the cell membrane through which the plasmid DNA enters. This method is amenable to the uptake of large plasmid DNA. After the electric shock the holes are rapidly closed by the cell's membrane-repair mechanisms.

The efficiency with which a competent culture can take up exogenous DNA and express its genes is known as Transformation efficiency.

Plasmid transformation

In order to be stably maintained in the cell a plasmid DNA molecule must contain an origin of replication, which allows it to be replicated in the cell independent of the replication of the cell's own chromosome. Because transformation usually produces a mixture of relatively few transformed cells and an abundance of non-transformed cells a method is needed to identify the cells that have acquired the plasmid. The method usually consists of using a plasmid that contains a gene that gives the bacterial cells resistance to an antibiotic that they are naturally sensitive to. The mixture of cells are then plated on media that contains the antibiotic thus only the transformed cells are able to grow. Cells that did not take up the plasmid are killed in the media.

Another selection method called blue-white screen uses a plasmid that contains an antibiotic resistance gene and the *lacZ* gene. The *lacZ* gene codes for the *lacZ-α* subunit of the enzyme β-galactosidase, a homo-tetramer with each monomer composed of one *lacZ-α* subunit and one *lacZ-ω* subunit. The method also requires an *E. coli* strain that possesses in its genome the code for only the *lacZ-ω* subunit and not the *lacZ-α* subunit. One of the first steps in any transformation is the production of a recombined plasmid obtained by the successful ligation of the gene of interest into its corresponding vector, which in this method results in the disruption of *lacZ* because the gene of interest is inserted within the *lacZ* code. A cell that takes up a recombined plasmid would thus not be able to express the *lacZ-α* subunit and would, in turn, not be able to produce a functional β-galactosidase. Conversely, a cell that has taken up non-recombined plasmid (perhaps one formed by the ligation of the vector's own two ends) will express the *lacZ-α* subunit and thus produce a functional β-galactosidase. A cell that does not take up any plasmid is not conferred with antibiotic resistance and will die upon plating. Consequently, the blue-white screen method allows for the ready detection of not just transformed cells, but, most importantly, cells that have been transformed by a successfully recombined plasmid. Selection occurs as a result of the action of β-galactosidase on its substrate X-gal, which is included in the media along with the appropriate antibiotic. X-gal is a colorless, modified galactose sugar whose hydrolysis by β-galactosidase produces galactose and the pre-chromophore 5-bromo-4-chloro-3-hydroxyindole. The latter is subsequently oxidized to 5,5'-dibromo-4,4'-dichloro-indigo, an insoluble, blue product that is readily seen by the naked eye. Colonies of cells that have been transformed by a successfully recombined plasmid will thus appear white whereas those that have been transformed by non-recombined plasmid will appear blue.

Plants

A number of mechanisms are available to transfer DNA into plant cells:

- *Agrobacterium* mediated transformation is the easiest and most simple plant transformation. Plant tissue (often leaves) are cut into small pieces, e.g. 10x10mm, and soaked for 10 minutes in a fluid containing suspended *Agrobacterium*. Some cells along the cut will be transformed by the bacterium, that inserts its DNA into the cell. Placed on selectable rooting and shooting

media, the plants will regrow. Some plants species can be transformed just by dipping the flowers into suspension of *Agrobacterium* and then planting the seeds in a selective medium. Unfortunately, many plants are not transformable by this method.

- Particle bombardment: Particles of gold or tungsten are coated with DNA and then shot into young plant cells or plant embryos. Some genetic material will stay in the cells and transform them. This method also allows transformation of plant plastids. The transformation efficiency is lower than in agribacterial mediated transformation, but most plants can be transformed with this method.
- Electroporation: make transient holes in cell membranes using electric shock; this allows DNA to enter as described above for Bacteria.
- Viral transformation (transduction): Package the desired genetic material into a suitable plant virus and allow this modified virus to infect the plant. If the genetic material is DNA, it can recombine with the chromosomes to produce transformant cells. However genomes of most plant viruses consist of single stranded RNA which replicates in the cytoplasm of infected cell. For such genomes this method is a form of transfection and not a real transformation, since the inserted genes never reach the nucleus of the cell and do not integrate into the host genome. The progeny of the infected plants is virus free and also free of the inserted gene.

Chapter- 4

Isogenic Human Disease Models

Isogenic human disease models are a family of cells that are selected or engineered to accurately model the genetics of a specific patient population, in vitro ("within glass"; or, commonly, "in the lab", in an artificial environment). They are provided with a genetically matched ‘normal cell’ to provide an isogenic system to research disease biology and novel therapeutic agents. They can be used to model any disease with a genetic foundation. Cancer is one such disease for which isogenic human disease models have been widely used.

Historical models

Human isogenic disease models have been likened to ‘patients in a test-tube’, since they incorporate the latest research into human genetic diseases and do so without the difficulties and limitations involved in using non-human models.

Historically, cells obtained from animals, typically mice, have been used to model cancer related pathways. However, there are obvious limitations inherent in using animals for modelling genetically determined diseases in humans. Despite a large proportion of genetic conservation between humans and mice, there are significant differences between the biology of mice and humans that are important to cancer research. For example, major differences in telomere regulation enable murine cells to bypass the requirement for telomerase upregulation, which is a rate-limiting step in human cancer formation. As another example, certain ligand-receptor interactions are incompatible between mice and humans. Additionally, experiments have demonstrated important and significant differences in the ability to transform cells, compared with cells of murine origin. For these reasons, it remains essential to develop models of cancer that employ human cells.

Targeting vectors

Isogenic cell lines are created via a process called homologous gene-targeting. Targeting vectors that utilize homologous recombination are the tools or techniques that are used to knock-in or knock-out the desired disease causing mutation or SNP (single nucleotide polymorphism) to be studied. Although disease mutations can be harvested directly from cancer patients, these cells usually contain many background mutations in addition to the specific mutation of interest, and a matched normal cell line is typically not obtained. Subsequently, targeting vectors are used to ‘knock-in’ or ‘knock out’ gene mutations

enabling a switch in both directions; from a normal to cancer genotype; or vice versa; in characterized human cancer cell lines such as HCT116 or Nalm6.

There are several gene targeting technologies used to engineer the desired mutation, the most prevalent of which are briefly described, including key advantages and limitations, in the summary table below.

Technique	Gene Knock-In	Gene Knock-out
rAAV (recombinant adeno-associated virus vectors)	Targeted insertions or modifications are created within endogenous genes; and so are subject to: 1. The correct gene-regulation mechanisms; and 2. Accurately reflect the disease events found in real patients. rAAV can introduce subtle point mutations, SNPs as well as small insertions with high efficiency. Moreover, many peer reviewed studies have shown that rAAV does not introduce any confounding off target genomic events. Appears to be the preferred method being adopted in academia, Biotech and Pharma on a precision versus time versus cost basis.|	Gene knockouts are at the endogenous locus, and thus are definitive, stable and patient relevant. No confounding off-target effects are elicited at other genomic loci. It requires a 2- step process: 1. Generate a heterozygous KO 2. Generate a bi-allelic knockout by targeting the second allele. This process can therefore generate 3 genotypes (+/+; -/+ and -/-); enabling therefore the analysis of haplo-insufficient gene function. Current limitation is the need to sequentially target single alleles making generation of knock-out cell lines a two-step process.|
Plasmid based homologous recombination	Insertion is at the endogenous locus and has all the above benefits, but it is very inefficient. It also requires a promoterless drug selection strategy entailing bespoke construct generation. A large historical bank of cell-lines has been generated using this method which has been displaced by other methods since the mid 1990s.	Deletion is at endogenous locus and has all the above benefits, but it is inefficient. It also requires a promoterless drug selection strategy that entails bespoke construct generation
Flip-in	This is an efficient technique that allows the directed insertion of 'ectopic' transgenes at a single pre-	Not applicable

	defined genomic locus (integration via a FLP recombinase site). This is not a technique for modifying an endogenous locus. Transgenes will usually be under the control of an exogenous promoter, or a partially defined promoter-unit in the incorrect genomic location. Their expression will therefore not be under the same genomic and epigenetic regulation as the endogenous loci, which limits the utility of these systems for studying gene- function. They are however, good for eliciting rapid and stable exogenous gene expression.	
Zinc-Finger Nucleases (ZFNs)	ZFNs have been reported to achieve high rates of genetic knock-outs within a target endogenous gene. If ZFNs are co-delivered with a transgene construct homologous to the target gene, genetic knock-in's or insertions can also be achieved. However, few peer reviewed publications exemplifying this application can be found. One drawback is the potential for ZFNs to produce off-target double strand breaks which may also lead to random off-target gene insertions, deletions and wider genomic instability; confounding the resulting genotype. Whole genome sequencing may also be required as a standard characterization step to determine whether such off-target genomic alterations are present in any resulting cell-line.	ZFNs are sequence-directed endonucleases which enable the rapid and highly efficient (up to 90% in a bulk cell population) disruption of both alleles of a target gene. User- defined or patient relevant loss of-function alterations appear more difficult to achieve. Off target deletions or insertions elsewhere in the genome cannot be controlled for or readily defined. The speed advantage of obtaining a biallelic KO in one step is also partially mitigated if one still needs to derive a clonal cell-line to study gene function in a homogenous cell-population. ZFN are likely more selective than RNAi, however, are not as amenable for high- throughput functional genomic studies.
Meganucleases	Meganucleases are operationally analogous to ZFN's. They have better site cleavage specificity due to their larger DNA-recognition footprint, and subsequently avoid the off-target gene insertions associated with ZFN's. There are limitations inherent in their use such	

as the meganuclease vector design which can take up to 9 months and cost tens of thousands of dollars. This makes meganucleases more attractive in high-value applications such as gene therapy, agrobiotechnology and engineering of bioproducer lines.

Homologous recombination in cancer cell disease models

Homologous recombination (HR) is a kind of genetic recombination in which genetic sequences are exchanged between two similar segments of DNA. HR plays a major role in eukaryotic cell division, promoting genetic diversity through the exchange between corresponding segments of DNA to create new, and potentially beneficial combinations of genes.

HR performs a second vital role in DNA repair, enabling the repair of double-strand breaks in DNA which is a common occurrence during a cell's lifecycle. It is this process which is artificially triggered by the above technologies, and bootstrapped in order to engender 'knock-ins' or 'knockouts' in specific genes5, 7.

A recent key advance was discovered using AAV-homologous recombination vectors, which increases the low natural rates of HR in differentiated human cells when combined with gene-targeting vectors-sequences.

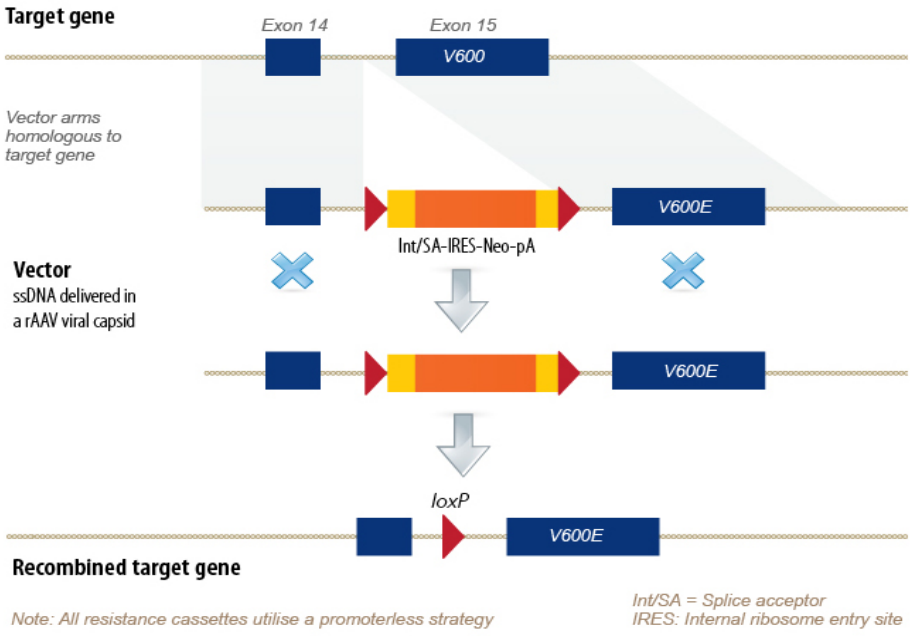


Diagram of a typical rAAV vector

Commercialization

Factors leading to the recent commercialization of isogenic human cancer cell disease models for the pharmaceutical industry and research laboratories are twofold.

Firstly, successful patenting of enhanced targeting vector technology has provided a basis for commercialization of the cell-models which eventuate from the application of these technologies.

Secondly, the trend of relatively low success rates in pharmaceutical RnD and the enormous costs have created a real need for new research tools that illicit how patient sub-groups will respond positively or be resistant to targeted cancer therapeutics based upon their individual genetic profile.

There are several companies working to address this need, a list of the key players and their technology offering is provided below.

- Cellectis:Meganucleases
- Horizon Discovery: Genesis (rAAV)
- Invitrogen: FLP
- Sigma-Aldrich: Zinc Fingers

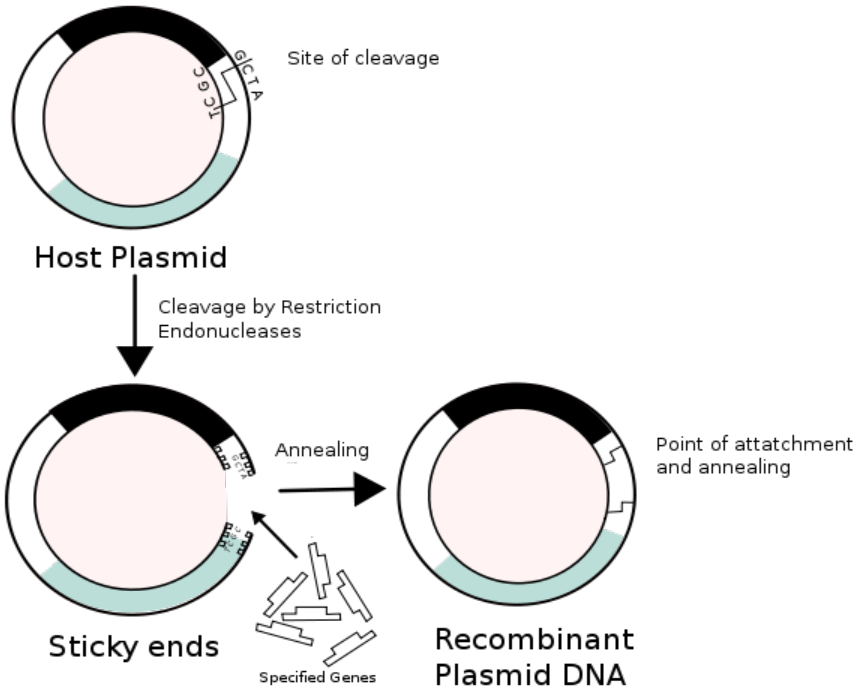
Chapter- 5

Recombinant DNA

Recombinant DNA (rDNA) is a form of artificial DNA that is created by combining two or more sequences that would not normally occur together through the process of gene splicing. In terms of genetic modification, it is created through the introduction of relevant DNA into an existing organismal DNA, such as the plasmids of bacteria, to code for or alter different traits for a specific purpose, such as antibiotic resistance. It differs from genetic recombination in that it does not occur through natural processes within the cell, but is engineered. A **recombinant protein** is a protein that is derived from recombinant DNA.

Methods

Cloning and relation to plasmids



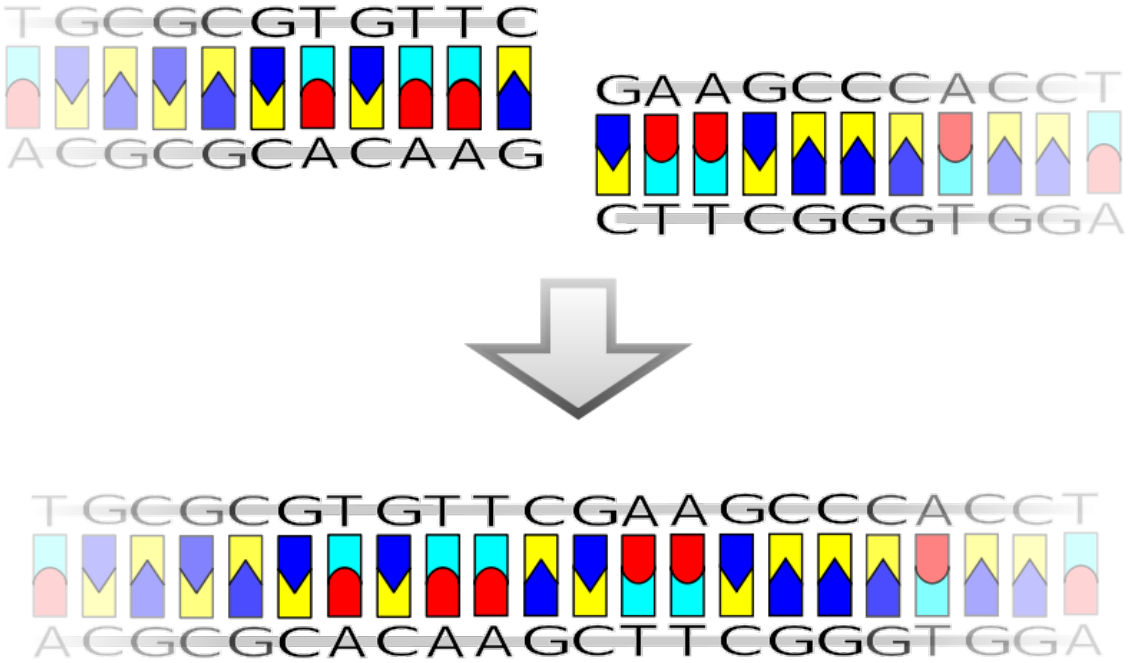
A simple example of how a desired gene is inserted into a plasmid. In this example, the gene specified in the white color becomes useless as the new gene is added.

The use of cloning is interrelated with recombinant DNA in classical biology, as the term "clone" refers to a cell or organism derived from a parental organism, with modern biology referring to the term as a collection of cells derived from the same cell that remain identical. In the classical instance, the use of recombinant DNA provides the initial cell from which the host organism is then expected to recapitulate when it undergoes further cell division, with bacteria remaining a prime example due to the use of viral vectors in medicine that contain recombinant DNA inserted into a structure known as a plasmid.

Plasmids are extrachromosomal self-replicating circular forms of DNA present in most bacteria, such as *Escherichia coli* (E. Coli), containing genes related to catabolism and metabolic activity, and allowing the carrier bacterium to survive and reproduce in conditions present within other species and environments. These genes represent characteristics of resistance to bacteriophages and antibiotics and some heavy metals, but can also be fairly easily removed or separated from the plasmid by restriction endonucleases, which regularly produce "sticky ends" and allow the attachment of a selected segment of DNA, which codes for more "reparative" substances, such as peptide hormone medications including insulin, growth hormone, and oxytocin. In the introduction of useful genes into the plasmid, the bacteria are then used as a viral vector, which are encouraged to reproduce so as to recapitulate the altered DNA within other cells it infects, and increase the amount of cells with the recombinant DNA present within them.

The use of plasmids is also key within gene therapy, where their related viruses are used as **cloning vectors** or carriers, which are means of transporting and passing on genes in recombinant DNA through viral reproduction throughout an organism. Plasmids contain three common features—a **replicator**, **selectable marker** and a **cloning site**. The replicator or "ori" refers to the origin of replication with regard to location and bacteria where replication begins. The marker refers to a particular gene that usually contains resistance to an antibiotic, but may also refer to a gene that is attached alongside the desired one, such as that which confers luminescence to allow identification of successfully recombined DNA. The cloning site is a sequence of nucleotides representing one or more positions where cleavage by restriction endonucleases occurs. Most eukaryotes do not maintain canonical plasmids; yeast is a notable exception. In addition, the Ti plasmid of the bacterium *Agrobacterium tumefaciens* can be used to integrate foreign DNA into the genomes of many plants. Other methods of introducing or creating recombinant DNA in eukaryotes include homologous recombination and transfection with modified viruses.

Chimeric plasmids



An example of chimeric plasmid formation from two "blunt ends" via the enzyme, T4 Ligase.

When recombinant DNA is then further altered or changed to host additional strands of DNA, the molecule formed is referred to as "chimeric" DNA molecule, with reference to the mythological chimera, which consisted as a composite of several animals. The presence of chimeric plasmid molecules is somewhat regular in occurrence, as, throughout the lifetime of an organism, the propagation by vectors ensures the presence of hundreds of thousands of organismal and bacterial cells that all contain copies of the original chimeric DNA.

In the production of chimeric(from chimera) plasmids, the processes involved can be somewhat uncertain, as the intended outcome of the addition of foreign DNA may not always be achieved and may result in the formation of unusable plasmids. Initially, the plasmid structure is linearised to allow the addition by bonding of complementary foreign DNA strands to single-stranded "overhangs" or "sticky ends" present at the ends of the DNA molecule from staggered, or "S-shaped" cleavages produced by restriction endonucleases.

A common vector used for the donation of plasmids originally was the bacterium Escherichia coli and, later, the EcoRI derivative, which was used for its versatility with addition of new DNA by "relaxed" replication when inhibited by chloramphenicol and spectinomycin, later being replaced by the pBR322 plasmid. In the case of EcoRI, the plasmid can anneal with the presence of foreign DNA via the route of sticky-end ligation,

or with "blunt ends" via blunt-end ligation, in the presence of the phage T_4 ligase, which forms covalent links between 3-carbon OH and 5-carbon PO_4 groups present on blunt ends. Both sticky-end, or overhang ligation and blunt-end ligation can occur between foreign DNA segments, and cleaved ends of the original plasmid depending upon the restriction endonuclease used for cleavage.

Applications

There are multitudinous proteins that are created from recombinant DNA and used as medications. Some can alternatively be produced from animal extracts or harvested from humans, such as human growth hormone (rhGH), human insulin, follicle-stimulating hormone (FSH) and factor VIII. Other proteins, when used as medication, only has recombinant DNA as a source, such as with erythropoietin.

History

The recombinant DNA technique was first proposed by Peter Lobban, a graduate student, with A. Dale Kaiser at the Stanford University Department of Biochemistry. The technique was then realized by Lobban and Kaiser; Jackson, Symons and Berg; and Stanley Norman Cohen, Chang, Herbert Boyer and Helling, in 1972–74. They published their findings in papers including the 1972 paper "*Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of Escherichia coli*", the 1973 paper "*Enzymatic end-to-end joining of DNA molecules*" and the 1973 paper *"Construction of Biologically Functional Bacterial Plasmids* in vitro*"*, all of which described techniques to isolate and amplify genes or DNA segments and insert them into another cell with precision, creating a transgenic bacterium.

Recombinant DNA technology was made possible by the discovery, isolation and application of restriction endonucleases by Werner Arber, Daniel Nathans, and Hamilton Smith, for which they received the 1978 Nobel Prize in Medicine. Cohen and Boyer applied for a patent on the Process for producing biologically functional molecular chimeras which could not exist in nature in 1974. The patent was granted in 1980.

A breakthrough in recombinant DNA technology occurred in 1977 when Herbert Boyer produced biosynthetic "human" insulin in the lab. The specific gene sequence, or polynucleotide, that codes for insulin production in humans was introduced to a sample colony of the *E. coli* bacteria. It was the first medicine made via recombinant DNA technology to be approved by the FDA and commercially available under the brand name Humulin. The vast majority of insulin currently used worldwide is now biosynthetic recombinant "human" insulin or its analogs.

Chapter- 6

Genetically Modified Organism

GloFish, the first genetically modified animal to be sold as a pet

A **genetically modified organism (GMO)** or **genetically engineered organism (GEO)** is an organism whose genetic material has been altered using genetic engineering techniques. These techniques, generally known as recombinant DNA technology, use DNA molecules from different sources, which are combined into one molecule to create a new set of genes. This DNA is then transferred into an organism, giving it modified or novel genes. Transgenic organisms, a subset of GMOs, are organisms which have inserted DNA that originated in a different species.

Production

Genetic modification involves the insertion or deletion of genes. When genes are inserted, they usually come from a different species, which is a form of horizontal gene transfer. In nature this can occur when exogenous DNA penetrates the cell membrane for any reason. To do this artificially may require attaching the genes to a virus or just physically inserting the extra DNA into the nucleus of the intended host with a very small syringe, or with very small particles fired from a gene gun. However, other methods exploit natural forms of gene transfer, such as the ability of *Agrobacterium* to transfer genetic material to plants, or the ability of lentiviruses to transfer genes to animal cells.

History

The general principle of producing a GMO is to add new genetic material into an organism's genome. This is called genetic engineering and was made possible through the discovery of DNA and the creation of the first recombinant bacteria in 1973; an existing bacterium *E. coli* expressing an exogenic Salmonella gene. This led to concerns in the scientific community about potential risks from genetic engineering, which were thoroughly discussed at the Asilomar Conference. One of the main recommendations from this meeting was that government oversight of recombinant DNA research should be established until the technology was deemed safe. Herbert Boyer then founded the first company to use recombinant DNA technology, Genentech, and in 1978 the company announced creation of an *E. coli* strain producing the human protein insulin.

In 1986, field tests of bacteria genetically engineered to protect plants from frost damage (ice-minus bacteria) at a small biotechnology company called Advanced Genetic Sciences of Oakland, California, were repeatedly delayed by opponents of biotechnology. In the same year, a proposed field test of a microbe genetically engineered for a pest resistance protein by Monsanto Company was dropped.

Uses

GMOs are used in biological and medical research, production of pharmaceutical drugs, experimental medicine (e.g. gene therapy), and agriculture (e.g. golden rice). The term "genetically modified organism" does not always imply, but can include, targeted insertions of genes from one species into another. For example, a gene from a jellyfish, encoding a fluorescent protein called GFP, can be physically linked and thus co-expressed with mammalian genes to identify the location of the protein encoded by the GFP-tagged gene in the mammalian cell. Such methods are useful tools for biologists in many areas of research, including those who study the mechanisms of human and other diseases or fundamental biological processes in eukaryotic or prokaryotic cells.

To date the broadest and most controversial application of GMO technology is patent-protected food crops which are resistant to commercial herbicides or are able to produce pesticidal proteins from within the plant, or *stacked trait* seeds, which do both. The largest share of the GMO crops planted globally are owned by the US firm Monsanto. In

2007, Monsanto's trait technologies were planted on 246 million acres (1,000,000 km^2) throughout the world, a growth of 13 percent from 2006.

In the corn market, Monsanto's triple-stack corn—which combines Roundup Ready 2 weed control technology with YieldGard Corn Borer and YieldGard Rootworm insect control—is the market leader in the United States. U.S. corn farmers planted more than 32 million acres (130,000 km^2) of triple-stack corn in 2008, and it is estimated the product could be planted on 56 million acres (230,000 km^2) in 2014–2015. In the cotton market, Bollgard II with Roundup Ready Flex was planted on approximately 5 million acres (20,000 km^2) of U.S. cotton in 2008.

According to the International Service for the Acquisition of Agri-Biotech Applications (ISAAA), of the approximately 14 million farmers who grew biotech crops in 2009, some 90% were resource-poor farmers in developing countries. These include some 7 million farmers in the cotton-growing areas of China, an estimated 5.6 million small farmers in India (Bt cotton), 250,000 in the Philippines, South Africa (biotech cotton, maize and soybeans often grown by subsistence women farmers) and the other twelve developing countries which grew biotech crops in 2009. 10 million more small and resource-poor farmers may have been secondary beneficiaries of Bt cotton in China.

The global commercial value of biotech crops grown in 2008 was estimated to be US$130 billion.

In the United States, the United States Department of Agriculture (USDA) reports on the total area of GMO varieties planted. According to National Agricultural Statistics Service, the states published in these tables represent 81–86 percent of all corn planted area, 88–90 percent of all soybean planted area, and 81–93 percent of all upland cotton planted area (depending on the year).

USDA does not collect data for global area. Estimates are produced by the International Service for the Acquisition of Agri-biotech Applications (ISAAA) and can be found in the report, "Global Status of Commercialized Transgenic Crops: 2007".

Transgenic animals are also becoming useful commercially. On February 6, 2009 the U.S. Food and Drug Administration approved the first human biological drug produced from such an animal, a goat. The drug, ATryn, is an anticoagulant which reduces the probability of blood clots during surgery or childbirth. It is extracted from the goat's milk.

Detection

Testing on GMOs in food and feed is routinely done by molecular techniques like DNA microarrays or qPCR. The test can be based on screening elements (like p35S, tNos, pat, or bar) or event-specific markers for the official GMOs (like Mon810, Bt11, or GT73). The array-based method combines multiplex PCR and array technology to screen samples for different potential GMOs, combining different approaches (screening elements, plant-specific markers, and event-specific markers). The qPCR is used to detect

specific GMO events by usage of specific primers for screening elements or event-specific markers.

To avoid any kind of false positive or false negative testing outcome, comprehensive controls for every step of the process is mandatory. A CaMV check is important to avoid false positive outcomes based on virus contamination of the sample.

Transgenic microbes

Bacteria were the first organisms to be modified in the laboratory, due to their simple genetics. These organisms are now used for several purposes, and are particularly important in producing large amounts of pure human proteins for use in medicine.

Genetically modified bacteria are used to produce the protein insulin to treat diabetes. Similar bacteria have been used to produce clotting factors to treat haemophilia, and human growth hormone to treat various forms of dwarfism.

Transgenic animals

Some chimeras, like the blotched mouse shown, are created through genetic modification techniques like gene targeting.

Transgenic animals are used as experimental models to perform phenotypic and for testing in biomedical research. Other applications include the production of human hormones such as insulin.

Fruit flies

In biological research, transgenic fruit flies (*Drosophila melanogaster*) are model organisms used to study the effects of genetic changes on development. Fruit flies are often preferred over other animals due to their short life cycle, low maintenance requirements, and relatively simple genome compared to many vertebrates.

Mammals

Genetically modified mammals are an important category of genetically modified organisms. Transgenic mice are often used to study cellular and tissue-specific responses to disease.

In 1999, scientists at the University of Guelph in Ontario, Canada created the genetically engineered Enviropig. The Enviropig excretes from 30 to 70.7% less phosphorus in manure depending upon the age and diet. In February 2010, Environment Canada determined that Enviropigs are in compliance with the Canadian Environmental Protection Act and can be produced outside of the research context in controlled facilities where they are segregated from other animals.

In 2009, scientists in Japan announced that they had successfully transferred a gene into a primate species (marmosets) and produced a stable line of breeding transgenic primates for the first time.

Cnidarians

Cnidarians such as *Hydra* and the sea anemone Nematostella vectensis have become attractive model organisms to study the evolution of immunity and certain developmental processes. An important technical breakthrough was the development of procedures for generation of stably transgenic hydras and sea anemones by embryo microinjection.

Fish

Genetically modified fish have promoters driving an over-production of "all fish" growth hormone. This resulted in dramatic growth enhancement in several species, including salmonids, carps and tilapias.

Gene therapy

Gene therapy, uses genetically modified viruses to deliver genes that can cure disease into human cells. Although gene therapy is still relatively new, it has had some successes. It has been used to treat genetic disorders such as severe combined immunodeficiency,

and treatments are being developed for a range of other currently incurable diseases, such as cystic fibrosis, sickle cell anemia, and muscular dystrophy. Current gene therapy technology only targets the non-reproductive cells meaning that any changes introduced by the treatment can not be transmitted to the next generation. Gene therapy targeting the reproductive cells—so-called "Germ line Gene Therapy"—is very controversial and is unlikely to be developed in the near future.

Transgenic plants

Kenyans examining insect-resistant transgenic Bt corn

Transgenic plants have been engineered to possess several desirable traits, including resistance to pests, herbicides, or harsh environmental conditions; improved product shelf life, and increased nutritional value. Since the first commercial cultivation of genetically modified plants in 1996, they have been modified to be tolerant to the herbicides glufosinate and glyphosate, to be resistant to virus damage as in Ringspot virus resistant GM papaya, grown in Hawaii and to produce the Bt toxin, a potent insecticide. Most of transgenic varieties grown today are known as first generation transgenics, because the transgenic trait provides benefits to farmers. Plants of the second generation should directly benefit the consumer with nutritional enhancement, taste, texture, etc. Transgenic plants of the second generation are being developed by both public research institutions and private companies. Currently there is no such transgenic variety on the market. Genetically modified sweet potatoes have been enhanced with protein and other nutrients,

while golden rice, developed by the International Rice Research Institute, has been discussed as a possible cure for Vitamin A deficiency. In January 2008, scientists altered a carrot so that it would produce calcium and become a possible cure for osteoporosis; however, people would need to eat 1.5 kilograms of carrots per day to reach the required amount of calcium.

The coexistence of GM plants with conventional and organic crops has raised significant concern in many European countries. Since there is separate legislation for GM crops and a high demand from consumers for the freedom of choice between GM and non-GM foods, measures are required to separate foods and feed produced from GMO plants from conventional and organic foods. European research programs such as Co-Extra, Transcontainer, and SIGMEA are investigating appropriate tools and rules. At the field level, biological containment methods include isolation distances and pollen barriers.

Cisgenic plants

Cisgenesis, sometimes also called **Intragenesis**, is a product designation for a category of genetically engineered plants. A variety of classification schemes have been proposed, that order genetically modified organisms based on the nature of introduced genotypical changes rather than the process of genetic engineering.

While some genetically modified plants are developed by the introduction of a gene originating from distant, sexually incompatible species into the host genome, cisgenic plants contain genes which have been isolated either directly from the host species or from sexually compatible species. The new genes are introduced using recombinant DNA methods and gene transfer. Some scientists hope that the approval process of cisgenic plants might be simpler than that of proper transgenics, but it remains to be seen.

Controversy

Biological process

The use of genetically modified organisms has sparked significant controversy in many areas. Some groups or individuals see the generation and use of GMO as intolerable meddling with biological states or processes that have naturally evolved over long periods of time, while others are concerned about the limitations of modern science to fully comprehend all of the potential negative ramifications of genetic manipulation.

Foodchain

The safety of GMOs in the foodchain has been questioned by some environmental groups, with concerns such as the possibilities that GMOs could introduce new allergens into foods, or contribute to the spread of antibiotic resistance. All studies published to date have shown no adverse health effects resulting from humans eating genetically modified foods, environmental groups still discourage consumption in many countries, claiming that GM foods are unnatural and therefore unsafe. Such concerns have led to the

adoption of laws and regulations that require safety testing of any new organism produced for human consumption.

GMOs' proponents note that because of the safety testing requirements imposed on GM foods, the risk of introducing a plant variety with a new allergene or toxin using genetic modification is much smaller than using traditional breeding processes. An example of an allergenic plant created using traditional breeding is the kiwi. One article calculated that the marketing of GM salmon could reduce the cost of salmon by half, thus increasing salmon consumption and preventing 1,400 deaths from heart attack a year in the United States.

Trade in Europe and Africa

In response to negative public opinion, Monsanto announced its decision to remove their seed cereal business from Europe, and environmentalists crashed a World Trade Organization conference in Cancun that promoted GM foods and was sponsored by Committee for a Constructive Tomorrow (CFACT). Some African nations have refused emergency food aid from developed countries, fearing that the food is unsafe. During a conference in the Ethiopian capital of Addis Ababa, Kingsley Amoako, Executive Secretary of the United Nations Economic Commission for Africa (UNECA), encouraged African nations to accept genetically modified food and expressed dissatisfaction in the public’s negative opinion of biotechnology.

Agricultural surpluses

Patrick Mulvany, Chairman of the UK Food Group, accused some governments, especially the Bush administration, of using GM food aid as a way to dispose of unwanted agricultural surpluses. The UN blamed food companies and accused them of violating human rights, calling on governments to regulate these profit-driven firms. It is widely believed that the acceptance of biotechnology and genetically modified foods will also benefit rich research companies and could possibly benefit them more than consumers in underdeveloped nations.

Labeling

While some groups advocate the complete prohibition of GMOs, others call for mandatory labeling of genetically modified food or other products. Other controversies include the definition of patent and property pertaining to products of genetic engineering. According to the documentary Food, Inc. efforts to introduce labeling of GMOs has repetedly met resistance from lobbyists and politicians affiliated with companies like Monsanto.

Testing

Bruce Stutz's article, “Wanted: GM Seeds for Study,” highlights a story of two dozen scientist who spoke out against the research restrictions put forth by companies producing

genetically modified (GM) seeds such as DuPont, Monsanto, and Syngenta. In February 2009, after scientist warned the U.S. Environmental protection Agency (EPA) “that industry influence had made independent analyses of transgenic crops impossible,” the American Seed Trade Association (ASTA) agreed that they “would allow researchers greater freedom to study the effects of GM food crops.” This agreement left many scientist optimistic about the future, but there is little optimism as to whether this agreement has the ability to “alter what has been a research environment rife with obstruction and suspicion.”

Impoverished nations

Some groups believe that impoverished nations will not reap the benefits of biotechnology because they do not have easy access to these developments, cannot afford modern agricultural equipment, and certain aspects of the system revolving around intellectual property rights are unfair to "undeveloped countries". For example, The CGIAR (Consultative Group of International Agricultural Research) is an aid and research organization that has been working to achieve sustainable food security and decrease poverty in undeveloped countries since its formation in 1971. In an evaluation of CGIAR, the World Bank praised its efforts but suggested a shift to genetics research and productivity enhancement. This plan has several obstacles such as patents, commercial licenses, and the difficulty that third world countries have in accessing the international collection of genetic resources and other intellectual property rights that would educate them about modern technology. The International Treaty on Plant Genetic Resources for Food and Agriculture has attempted to remedy this problem, but results have been inconsistent. As a result, "orphan crops", such as teff, millets, cowpeas, and indigenous plants, are important in the countries where they are grown, but receive little investment.

Private investments

The development and implementation of policies designed to encourage private investments in research and marketing biotechnology that will meet the needs of poverty-stricken nations, increased research on other problems faced by poor nations, and joint efforts by the public and private sectors to ensure the efficient use of technology developed by industrialized nations have been suggested. In addition, industrialized nations have not tested GM technology on tropical plants, focusing on those that grow in temperate climates, even though undeveloped nations and the people that need the extra food live primarily in tropical climates. Some European scientists are concerned that political factors and ideology prevent unbiased assessment of GM technology in some EU countries, with a negative effect on the whole community.

Transgenic organisms

Another important controversy is the possibility of unforeseen local and global effects as a result of transgenic organisms proliferating. The basic ethical issues involved in genetic research are discussed in genetic engineering.

Some critics have raised the concern that conventionally-bred crop plants can be cross-pollinated (bred) from the pollen of modified plants. Pollen can be dispersed over large areas by wind, animals, and insects. In 2007, the U.S. Department of Agriculture fined Scotts Miracle-Gro $500,000 when modified genetic material from creeping bentgrass, a new golf-course grass Scotts had been testing, was found within close relatives of the same genus (*Agrostis*) as well as in native grasses up to 21 km (13 miles) away from the test sites, released when freshly cut grass was blown by the wind.

GM proponents point out that outcrossing, as this process is known, is not new. The same thing happens with any new open-pollinated crop variety—newly introduced traits can potentially cross out into neighboring crop plants of the same species and, in some cases, to closely related wild relatives. Defenders of GM technology point out that each GM crop is assessed on a case-by-case basis to determine if there is any risk associated with the outcrossing of the GM trait into wild plant populations. The fact that a GM plant may outcross with a related wild relative is not, in itself, a risk unless such an occurrence has negative consequences. If, for example, an herbicide resistance trait was to cross into a wild relative of a crop plant it can be predicted that this would not have any consequences except in areas where herbicides are sprayed, such as a farm. In such a setting the farmer can manage this risk by rotating herbicides.

The European Union funds research programs such as Co-Extra, that investigate options and technologies on the coexistence of GM and conventional farming. This also includes research on biological containment strategies and other measures to prevent outcrossing and enable the implementation of coexistence.

If patented genes are outcrossed, even accidentally, to other commercial fields and a person deliberately selects the outcrossed plants for subsequent planting then the patent holder has the right to control the use of those crops. This was supported in Canadian law in the case of Monsanto Canada Inc. v. Schmeiser.

"Terminator" and "traitor"

An often cited controversy is a "Technology Protection" technology dubbed 'Terminator'. This uncommercialized technology would allow the production of first generation crops that would not generate seeds in the second generation because the plants yield sterile seeds. The patent for this so-called "terminator" gene technology is owned by *Delta and Pine Land Company* and the United States Department of Agriculture. Delta and Pine Land was bought by Monsanto Company in August 2006. Similarly, the hypothetical trait-specific Genetic Use Restriction Technology, also known as 'Traitor' or 'T-gut', requires application of a chemical to genetically modified crops to reactivate engineered traits. This technology is intended both to limit the spread of genetically engineered plants, and to require farmers to pay yearly to reactivate the genetically engineered traits of their crops. Genetic Use Restriction Technology is under development by companies including Monsanto and AstraZeneca.

In addition to the commercial protection of proprietary technology in self-pollinating crops such as soybean (a generally contentious issue), another purpose of the terminator gene is to prevent the escape of genetically modified traits from cross-pollinating crops into wild-type species by sterilizing any resultant hybrids. Some environmentalist groups, while considering outcrossing of GM plants dangerous, felt the technology would prevent re-use of seed by farmers growing such terminator varieties in the developing world and was ostensibly a means to exercise patent claims. However other environmental groups welcomed the terminator gene as a means of preventing GM crops from mixing with natural crops.

Hybrid seeds were commonly used in the developed countries long before the introduction of GM crops. Hybrid seeds cannot be saved, so purchasing new seed every year is already a standard agricultural practice.

There are technologies evolving which contain the transgene by biological means and still can provide fertile seeds using fertility restorer functions. Such methods are being developed by several EU research programs, among them Transcontainer and Co-Extra.

Governmental support and opposition

Australia

Several states of Australia had placed bans on planting GM food crops, beginning in 2003. However, in late 2007 the states of New South Wales and Victoria lifted their bans. Western Australia lifted their state's ban in December 2008, while South Australia continues its ban. Tasmania has extended its moratorium until November 2014. The state of Queensland has allowed the growing of GM crops since 1995 and has never had a GM ban.

Canada

In 2005, a standing committee of the government of Prince Edward Island (PEI) in Canada assessed a proposal to ban the production of GMOs in the province. The ban was not passed. As of January 2008, the use of genetically modified crops on PEI was rapidly increasing. Mainland Canada is one of the world's largest producers of GM canola.

Japan

As of 2009, Japan has no commercial farming of any kinds of genetically modified food. Consumers have strongly resisted both imports and attempts to grow GMO in the country. Campaigns by consumer groups and environmental groups, such as Consumers Union of Japan and Greenpeace Japan, as well as local campaigns, have been very successful. In Hokkaido, a special bylaw has made it virtually impossible to grow GMOs, as the No! GMO Campaign collected over 200,000 signatures to oppose GMO farming. Consumers Union of Japan participated together with other Japanese NGOs at the Planet Diversity conference in Bonn, Germany on May 12–16, 2008, a global congress on the

future of food and agriculture, with a demonstration to celebrate biodiversity, to oppose GMOs. “We don’t only need networks between people, but between people and plants, and people and planet earth,” noted Koketsu Michiyo from CUJ.

Cross-pollination has commonly occurred in Japan, as canola seed (rape seed) is imported from Canada. Around ports and the roads to major food oil companies, GE canola has now been found growing wild. Imported canola seeds have been found to be GMO varieties, including the Roundup Ready and Liberty Link types not grown in Japan. Activists and local groups, as well as the No! GMO Campaign and others, are alarmed that imported GMOs may harm the biodiversity and cause irreversible damage. A report from the Japanese National Institute for Environmental Studies (NIES) confirms that herbicide-resistant genetically engineered canola plants were identified in five of the six Japanese ports where samples were collected.

A number of Japanese groups have been making submissions to Western Australia’s Review of the Genetically Modified Crops Free Areas Act 2003. These include the Seikatsu Club Consumers’ Cooperative Union and the Consumers Union of Japan. Seikatsu—an umbrella group of 29 Seikatsu Club Consumers’ Co-Operatives—and its oil crushers Okamura Oil Mill Ltd and Yonezawa Oil Co. Ltd., all have non-GE canola policies. The groups stopped importing canola from Canada after the introduction of GE canola, when cross-pollination made it impossible to guarantee GE-free canola from Canada.

Pakistan

The government supports the use of hybrid seeds. However, Monsanto once tried to sell their hybrid seeds of such important crops as wheat and rice via the government. Even though yields would have increased, it would have made the Pakistani population dependent on the seeds of one company. The contract was never given.

New Zealand

In New Zealand, no genetically modified food is grown and no medicines containing live genetically-modified organisms have been approved for use. However, medicines manufactured using genetically modified organisms that do not contain live organisms have been approved for sale, and imported foods with genetically modified components are sold.

United States

In 2004, Mendocino County, California became the first county in the United States to ban the production of GMOs. The measure passed with a 57% majority. In California, Trinity and Marin counties have also imposed bans on GM crops, while ordinances to do so were unsuccessful in Butte, Lake, San Luis Obispo, Humboldt, and Sonoma counties. Supervisors in the agriculturally-rich counties of Fresno, Kern, Kings, Solano, Sutter, and Tulare have passed resolutions supporting the practice.

In 2007, with reference to US negotiations with the EU on agricultural biotechnology, US diplomatic cables recommended that 'we calibrate a target retaliation list that causes some pain across the EU'.

Zambia

The Zambian government has launched a campaign to educate and increase awareness of the benefits of biotechnology, including genetically modified crops, in order to change negative public opinion.

Other Africa

In 2010, after nine years of talks, the Common Market for Eastern and Southern Africa (COMESA) produced a draft policy on GM technology. This proposed policy was sent to all 19 national governments for consultation in September 2010. Under the policy, a member country which wants to grow a new GM crop would inform COMESA who would have sufficient scientific expertise to make the decision as to whether the crop was safe for the environment and for humans. At the moment, few countries have the resources to make their own decisions. Once COMESA had made their decision, permission would be granted for the crop to be grown in all 19 member countries. Member countries would retain the power not to grow the crop in their own country if they wanted.

France

The cultivation of Monsanto's MON 810 corn was forbidden in France on February 9, of 2008. It was the only GMO authorized in France. The safeguard measure is taken as far as side effects on human health will be known. In 2010 Marion Guillou, president of the National Institute for Agronomical Research and one of France's top farm researcher, said she can no longer work on developing new GMOs due to widespread distrust and even hostility by European consumers.

Germany

Germany placed a ban on the cultivation and sale of GMO maize in April 2009.

Other European Countries

MON 810 (maize) was the first GMO crop to be cultivated in Europe. The initial lines of maize were approved in 1997 and, by 2009, 76,000 hectares of GM maize were grown in Spain (20% of Spain's maize production). Smaller amounts were produced in the Czech Republic, Slovakia, Portugal, Romania and Poland. However, in addition to France and Germany, other European countries that have placed bans on the cultivation and sale of GMOs include Austria, Hungary, Greece, and Luxembourg. Ireland has also banned GMO cultivation, and has instituted a voluntary label for GMO-free food products.

Poland has also tried to institute a ban, with backlash from the European Commission. Bulgaria effectively banned cultivation of genetically modified organisms on March 18, 2010.

On 2 March 2010 a second species of GMO, a potato named Amflora, was approved for cultivation for industrial applications in the EU by the European Commission and was grown in Germany, Sweden and the Czech Republic that year. On 13 July 2010, the European Commission issued a recommendation that in future individual states in the EU should be able to ban the growing of specific GM crops that had been scientifically approved at the EU level. A ban could be justified on cultural, economic or ethical grounds. The EU approval process for imports of GM crops and labelling of GM food products remained in place.

Chapter- 7

Pharming

Pharming is a portmanteau of farming and "pharmaceutical" and refers to the use of genetic engineering to insert genes that code for useful pharmaceuticals into host animals or plants that would otherwise not express those genes. As a consequence, the host animals or plants then make the pharmaceutical product in large quantity, which can then be purified and used as a drug product. Some drug products and nutrients may be able to be delivered directly by eating the plant or drinking the milk. Such technology has the potential to produce large quantities of cheap vaccines, or other important pharmaceutical products such as insulin.

The products of pharming are recombinant proteins or their metabolic products. Drugs made from recombinant proteins potentially have greater efficacy and fewer side effects than small organic molecules (which are often screened as potential drugs) because their action can be more precisely targeted toward the cause of a disease rather than treatment of symptoms. Recombinant proteins are most commonly produced using bacteria or yeast in a bioreactor, but pharming offers the advantage to the producer that it does not require expensive infrastructure, and production capacity can be quickly scaled to meet demand. It is estimated that the expense of producing a recombinant protein drug via pharming will be less than 70% of the current cost.

In the United States, Transgenic plants including but not limited to those that produce pharmaceuticals, are regulated by three government agencies, which comprise the Coordinated Framework for Regulation of Biotechnology established in 1986.

- United States Department of Agriculture Animal and Plant Health Inspection Service - evaluates potential agricultural impacts such as gene flow and 'weediness'
- United States Environmental Protection Agency - evaluates potential environmental impact intergenic microorganisms under the Toxic Substances Control Act
- United States Department of Health and Human Services Food and Drug Administration (FDA) - evaluates human health risk if the plant or one of its proteins is intended for human consumption

Pharming in mammals

Expression in the milk of a mammal, such as a cow, sheep, or goat, is a common application, as milk production is plentiful and purification from milk is relatively easy. Hamsters and rabbits have also been used in preliminary studies because of their faster breeding.

One approach to this technology is the creation of a transgenic mammal that can produce the biopharmaceutical in its milk (or blood or urine). Once an animal is produced, typically using the pronuclear microinjection method, it becomes efficacious to use cloning technology to create additional offspring that carry the favorable modified genome. In February 2009 the US FDA granted marketing approval for the first drug to be produced in genetically modified livestock. The drug is called ATryn, which is antithrombin protein purified from the milk of genetically-modified goats. Marketing permission was granted by the European Medicines Agency in August 2006.

Pharming in plants

Arabidopsis is often used as a model organism to study gene expression in plants, while actual production may be carried out in maize, rice, potatoes, tobacco, flax or safflower. The advantage of rice and flax is that they are self-pollinating, and thus gene flow issues (see below) are avoided. However, human error could still result in pharm crops entering the food supply. Using a minor crop such as safflower or Tobacco, avoids the greater political pressures and risk to the food supply involved with using staple crops such as beans or rice. Despite these risks, corn and soybeans are currently the most common crops used to produce pharmaceuticals.

Plant-Made Pharmaceuticals (PMPs), also referred to as Biopharming, is a sub-sector of the biotechnology industry that involves the process of genetically engineering plants so that they can produce certain types of therapeutically important proteins and associate molecules such as peptides and secondary metabolites. The proteins and molecules can then be harvested and used to produce pharmaceuticals.

There is much debate over the practicality of using plants to produce proteins. Some groups fear that contamination of conventional crops might occur; in several instances, companies have been fined for violating protocols, resulting in potential contamination. This leads to the question of "Why would biotechnology companies use plants to produce proteins?"

Conventional production methods for pharmaceutical proteins involve substantial investments of both time and finances. Not only are there manufacturing challenges involved with conventional production methods, but there are also considerable regulatory challenges that must be met. There are currently ~200 protein-based medicines (vaccines, monoclonal antibody drugs, and other therapeutic proteins and peptides) on the market, and ~400 in development. (Statistics are from BIO, the biotechnology trade

group) Consequently, companies are motivated to provide a wider range of options for production of proteins used in these treatments.

Biopharm proponents claim that using plants can offer an easily controllable, safe, and cost-effective method for manufacturing proteins, provided that proper regulatory safeguards are put into place to ensure that no outcrossing can occur. It is also important to note, that the global demand for particular pharmaceutical protein can easily be met from just a few acres of pharma-crop, which can be grown under high containment conditions (e.g. in the greenhouse). Some scientists even think that the term "gardening" is more appropriate than farming. Opponents are concerned that there are too many ways in which contamination of the food supply and the environment can occur to make this form of production socially desirable, or even economically feasible.

Compared to conventional production methods, plant-made pharmaceuticals could save substantial time, money, and provide a system for producing proteins that could solve current production challenges.

Although no drugs from pharm crops are currently on the market, open field growing trials of these crops began in the United States in 1992 and have taken place every year since. The United States Department of Agriculture has approved planting of pharma crops in every state, with most testing taking place in Hawaii, Nebraska, Iowa, and Wisconsin.

These pharmaceutical crops could become extremely beneficial in developing countries. The World Health Organization estimates that nearly 3 million people die each year from vaccine preventable disease, mostly in Africa. Diseases such as measles and hepatitis lead to deaths in countries where the people cannot afford the high costs of vaccines, but pharm crops could help solve this problem.

Companies in this industry hope that proteins made from plants can be used to develop treatments for some of the most serious diseases and conditions such as cancer, diabetes, HIV, heart disease, Alzheimer's disease, cystic fibrosis, multiple sclerosis, Hepatitis C, and arthritis, but no such products have as yet been approved.

Controversy over pharming

Those opposed to pharming fear that through either mishandling or gene flow, potentially dangerous pharmaceuticals may inadvertently enter the food supply. Precedents involving non-pharmaceutical genetically modified crops include the Starlink controversy, and trade war over genetically modified food between the European union and the USA. A similar reaction to pharmed rice is feared from Japan.

In 2002, ProdiGene was fined $250,000 and ordered by the USDA to pay over $3 million in cleanup costs after allowing a fraction of a bushel of volunteer pharm corn to comingle with the soybean crop later planted in that field. Although the chance of gene flow between species is claimed to be low and there was in this case no threat to consumers,

the USDA has a zero tolerance policy. ProdiGene has since revised its protocols and resumed operations in Nebraska. In 2005, Anheuser-Busch threatened to boycott rice grown in Missouri because of plans by Ventria Bioscience to grow pharm rice in the state. A compromise was reached, but Ventria has withdrawn its 2006 permit to plant in Missouri due to unrelated circumstances. The company's field trials in North Carolina are expected to continue.

Chapter- 8

Genetically Modified Food

Genetically modified (GM) foods are foods derived from genetically modified organisms. Genetically modified organisms have had specific changes introduced into their DNA by genetic engineering techniques. These techniques are much more precise than mutagenesis (mutation breeding) where an organism is exposed to radiation or chemicals to create a non-specific but stable change. Other techniques by which humans modify food organisms include selective breeding (plant breeding and animal breeding), and somaclonal variation.

GM foods were first put on the market in the early 1990s. Typically, genetically modified foods are transgenic plant products: soybean, corn, canola, and cotton seed oil. Animal products have also been developed, although as of July 2010 none are currently on the market. In 2006 a pig was controversially engineered to produce omega-3 fatty acids through the expression of a roundworm gene. Researchers have also developed a genetically-modified breed of pigs that are able to absorb plant phosphorus more efficiently, and as a consequence the phosphorus content of their manure is reduced by as much as 60%.

Critics have objected to GM foods on several grounds, including possible safety issues, ecological concerns, and economic concerns raised by the fact that these organisms are subject to intellectual property law.

Method

Genetic modification involves the insertion or deletion of genes. In the process of cisgenesis, genes are artificially transferred between organisms that could be conventionally bred. In the process of transgenesis, genes from a different species are inserted, which is a form of horizontal gene transfer. In nature this can occur when exogenous DNA penetrates the cell membrane for any reason. To do this artificially may require attaching genes to a virus or just physically inserting the extra DNA into the nucleus of the intended host with a very small syringe, or with very small particles fired from a gene gun. However, other methods exploit natural forms of gene transfer, such as the ability of Agrobacterium to transfer genetic material to plants, and the ability of lentiviruses to transfer genes to animal cells.

Development

The first commercially grown genetically modified whole food crop was a tomato (called FlavrSavr), which was modified to ripen without softening, by Calgene, later a subsidiary of Monsanto. Calgene took the initiative to obtain FDA approval for its release in 1994 without any special labeling, although legally no such approval was required. It was welcomed by consumers who purchased the fruit at a substantial premium over the price of regular tomatoes. However, production problems and competition from a conventionally bred, longer shelf-life variety prevented the product from becoming profitable. A tomato produced using similar technology to the Flavr Savr was used by Zeneca to produce tomato paste which was sold in Europe during the summer of 1996. The labeling and pricing were designed as a marketing experiment, which proved, at the time, that European consumers would accept genetically engineered foods. Currently, there are a number of food species in which a genetically modified version exists (percent modified are mostly 2009/2010 data).

Food	Properties of the genetically modified variety	Modification	Percent Modified in US	Percent Modified in world
Soybeans	Resistant to glyphosate or glufosinate herbicides	Herbicide resistant gene taken from bacteria inserted into soybean	93%	77%
Corn, field	Resistant to glyphosate or glufosinate herbicides. Insect resistance via producing Bt proteins, some previously used as pesticides in organic crop production. Vitamin-enriched corn derived from South African white corn variety M37W has bright orange kernels, with 169x increase in beta carotene, 6x the vitamin C and 2x folate.	New genes, some from the bacterium Bacillus thuringiensis, added/transferred into plant genome.	86%	26%
Cotton (cottonseed oil)	Pest-resistant cotton	Bt crystal protein gene added/transferred into plant	93%	49%

		genome		
Alfalfa	Resistant to glyphosate or glufosinate herbicides	New genes added/transferred into plant genome.	Planted in the US from 2005–2007; no longer planted currently due to court decisions	
Hawaiian papaya	Variety is resistant to the papaya ringspot virus.	New gene added/transferred into plant genome	80%	
Tomatoes	Variety in which the production of the enzyme polygalacturonase (PG) is suppressed, retarding fruit softening after harvesting.	A reverse copy (an antisense gene) of the gene responsible for the production of PG enzyme added into plant genome	Taken off the market due to commercial failure.	Small quantities grown in China
Rapeseed (Canola)	Resistance to herbicides (glyphosate or glufosinate), high laurate canola	New genes added/transferred into plant genome	93%	21%
Sugar cane	Resistance to certain pesticides, high sucrose content.	New genes added/transferred into plant genome		
Sugar beet	Resistance to glyphosate, glufosinate herbicides	New genes added/transferred into plant genome	95% (2010); planting in the US is halted as of 13 Aug. 2010 by court order	9%
Rice	Genetically modified to contain high amounts of Vitamin A (beta-carotene)	"Golden rice" Three new genes implanted: two from daffodils and the third from a bacterium	Forecast to be on the market in 2012	

Squash (Zucchini)	Resistance to watermelon, cucumber and zucchini yellow mosaic viruses	Contains coat protein genes of viruses.	13%	
Sweet Peppers	Resistance to virus	Contains coat protein genes of the virus.		Small quantities grown in China

In addition, various genetically engineered micro-organisms are routinely used as sources of enzymes for the manufacture of a variety of processed foods. These include alpha-amylase from bacteria, which converts starch to simple sugars, chymosin from bacteria or fungi that clots milk protein for cheese making, and pectinesterase from fungi which improves fruit juice clarity.

Growing GM crops

Between 1997 and 2009, the total surface area of land cultivated with GMOs had increased by a factor of 80, from 17,000 km^2 (4.2 million acres) to 1,340,000 km^2 (331 million acres).

Although most GM crops are grown in North America, in recent years there has been rapid growth in the area sown in developing countries. For instance in 2009 the largest increase in crop area planted to GM crops (soybeans) was in Brazil (214,000 km^2 in 2009 versus 158,000 km^2 in 2008.) There has also been rapid and continuing expansion of GM cotton varieties in India since 2002. (Cotton is a major source of vegetable cooking oil and animal feed.) In 2009 84,000 km^2 of GM cotton were harvested in India.

In India, GM cotton yields in Andhra Pradesh were no better than non-GM cotton in 2002, the first year of commercial GM cotton planting. This was because there was a severe drought in Andhra Pradesh that year and the parental cotton plant used in the genetic engineered variant was not well suited to extreme drought. Maharashtra, Karnataka, and Tamil Nadu had an average 42% increase in yield with GM cotton in the same year. Drought resistant variants were developed and, with the substantially reduced losses to insect predation, by 2009 87% of Indian cotton was GM. Though disputed the economic and environmental benefits of GM cotton in India to the individual farmer have been documented.

In 2009, countries that grew 95% of the global transgenic crops were the United States (46%), Brazil (16%), Argentina (15%), India (6%), Canada (6%), China (3%), Paraguay (2%) and South Africa (2%). The Grocery Manufacturers of America estimate that 75% of all processed foods in the U.S. contain a GM ingredient. In particular, Bt corn, which produces the pesticide within the plant itself, is widely grown, as are soybeans genetically designed to tolerate glyphosate herbicides. These constitute "input-traits" are aimed to

financially benefit the producers, have indirect environmental benefits and marginal cost benefits to consumers.

In the US, by 2009/2010, 93% of the planted area of soybeans, 93% of cotton, 86% of corn and 95% of the sugar beet were genetically modified varieties. Genetically modified soybeans carried herbicide-tolerant traits only, but maize and cotton carried both herbicide tolerance and insect protection traits (the latter largely the *Bacillus thuringiensis* Bt insecticidal protein). In the period 2002 to 2006, there were significant increases in the area planted to Bt protected cotton and maize, and herbicide tolerant maize also increased in sown area.

Legal issues in the US

Alfalfa

On 21 June 2010, the US Supreme Court issued its first ruling in regard to a GM crop. This was a ruling in regard to Roundup Ready alfalfa. The case goes back to 2006, when organic farmers, concerned about the impact of GM alfalfa on their crops, sued Monsanto. In response, the California Northern District Court ruled that the United States Department of Agriculture (USDA) was in error when it approved the planting of Roundup Ready alfalfa. According to the presiding judge, the law required the USDA to first conduct a full environmental study, which it had not done. It was the concern of the organic growers that the GM alfalfa could cross-pollinate with their organic alfalfa, making their crops unsalable in countries that forbid the growing of GM crops.

The impact of the current US Supreme Court ruling is somewhat unclear, with both sides appearing to claim victory. While Monsanto can claim technical victory in the case, various other issues still remain open, and will likely be litigated in the future. Meanwhile, the planting of GM alfalfa currently remains halted in the US, and it is unclear when it may resume.

Sugar beets

Between 2009 and 2010, the United States District Court for the Northern District of California considered the case involving the planting of genetically modified sugar beets. This case involves Monsanto's breed of pesticide-resistant sugar beets. Earlier in 2010, Judge Jeffrey S. White allowed the planting of GM sugar beets to continue, but he also warned that this may be blocked in the future while an environmental review was taking place. Finally, on 13 August 2010, Judge White ordered a halt to the planting of the genetically modified sugar beets in the US. He indicated that "the Agriculture Department had not adequately assessed the environmental consequences before approving them for commercial cultivation." The decision was the result of a lawsuit organised by the Center for Food Safety, a US non-governmental organisation that is a critic of biotech crops.

Crop yields

A 1999 study by Charles Benbrook, Chief Scientist of the Organic Center, found that genetically engineered Roundup Ready soybeans did not increase yields. The report reviewed over 8,200 university trials in 1998 and found that Roundup Ready soybeans had a yield drag of 5.3% across all varieties tested. In addition, the same study found that farmers used 2-5 times more herbicide (Roundup) on Roundup Ready soybeans compared to other popular weed management systems.

However research published in Science in 2003 has shown that the use of genetically modified Bt cotton in India increased yields by 60% over the period 1998-2001 while the number of applications of insecticides against bollworm were three times less on average.

A 2008 Soil Association report found that some scientific studies claimed that genetically modified varieties of plants do not produce higher crop yields than normal plants.

In 2009 the Union of Concerned Scientists summarized numerous peer-reviewed studies on the yield contribution of genetic engineering in the United States. This report examined the two most widely grown engineered crops—soybeans and maize (corn). Unlike many other studies, this work separated the yield contribution of the engineered gene from that of the many naturally occurring yield genes in crops.

The report found that engineered herbicide tolerant soy and maize did not increase yield at the national, aggregate level. Maize engineered with Bt insect resistance genes increased national yield by about 3 to 4 percent. Engineered crops increased net yield in all cases.

The study concluded that in the United States, other agricultural methods have made a much greater contribution to national crop yield increases in recent years than genetic engineering. United States Department of Agriculture data record maize yield increases of about 28 percent since engineered varieties were first commercialized in the mid 1990s. The yield contribution of engineered genes has therefore been a modest fraction—about 14 percent—of the maize yield increase since the mid 1990s.

A 2010 article summarised the results of 49 peer reviewed studies on GM crops worldwide. On average, farmers in developed countries experienced increase in yield of 6% and in underdeveloped countries of 29%. Tillage was decreased by 25-58% on herbicide resistant soybeans, insecticide applications on Bt crops were reduced by 14-76% and 72% of farmers worldwide experienced positive economic results.

Coexistence and traceability

The United States and Canada do not require labeling of genetically modified foods. However in certain other regions, such as the European Union, Japan, Malaysia and Australia, governments have required labeling so consumers can exercise choice between foods that have genetically modified, conventional or organic origins. This requires a

labeling system as well as the reliable separation of GM and non-GM organisms at production level and throughout the whole processing chain.

For traceability, the OECD has introduced a "unique identifier" which is given to any GMO when it is approved. This unique identifier must be forwarded at every stage of processing. Many countries have established labeling regulations and guidelines on coexistence and traceability. Research projects such as Co-Extra, SIGMEA and Transcontainer are aimed at investigating improved methods for ensuring coexistence and providing stakeholders the tools required for the implementation of coexistence and traceability.

Detection

Testing on GMOs in food and feed is routinely done using molecular techniques like DNA microarrays or qPCR. These tests can be based on screening genetic elements (like p35S, tNos, pat, or bar) or event-specific markers for the official GMOs (like Mon810, Bt11, or GT73). The array-based method combines multiplex PCR and array technology to screen samples for different potential GMOs, combining different approaches (screening elements, plant-specific markers, and event-specific markers).

The qPCR is used to detect specific GMO events by usage of specific primers for screening elements or event-specific markers. Controls are necessary to avoid false positive or false negative results. For example, a test for CaMV is used to avoid a false positive in the event of a virus contaminated sample.

PLU codes

A Price Look-Up code beginning with the digit *8* indicates genetically modified food.

Controversy

While it is evident that there is a food supply issue, the question is whether GM can solve world hunger problems, or even if that would be the best way to address the issue. Several scientists argue that in order to meet the demand for food in the developing world, a second Green Revolution with increased use of GM crops is needed. Others argue that there is more than enough food in the world and that the hunger crisis is caused by problems in food distribution and politics, not production. Recently some critics and environmentalists have changed their minds on the issue with respect to the need for additional food supplies. Further, it has been widely noted that there are those who consider over-population the real issue here, and that food production is adequate for any reasonable population size.

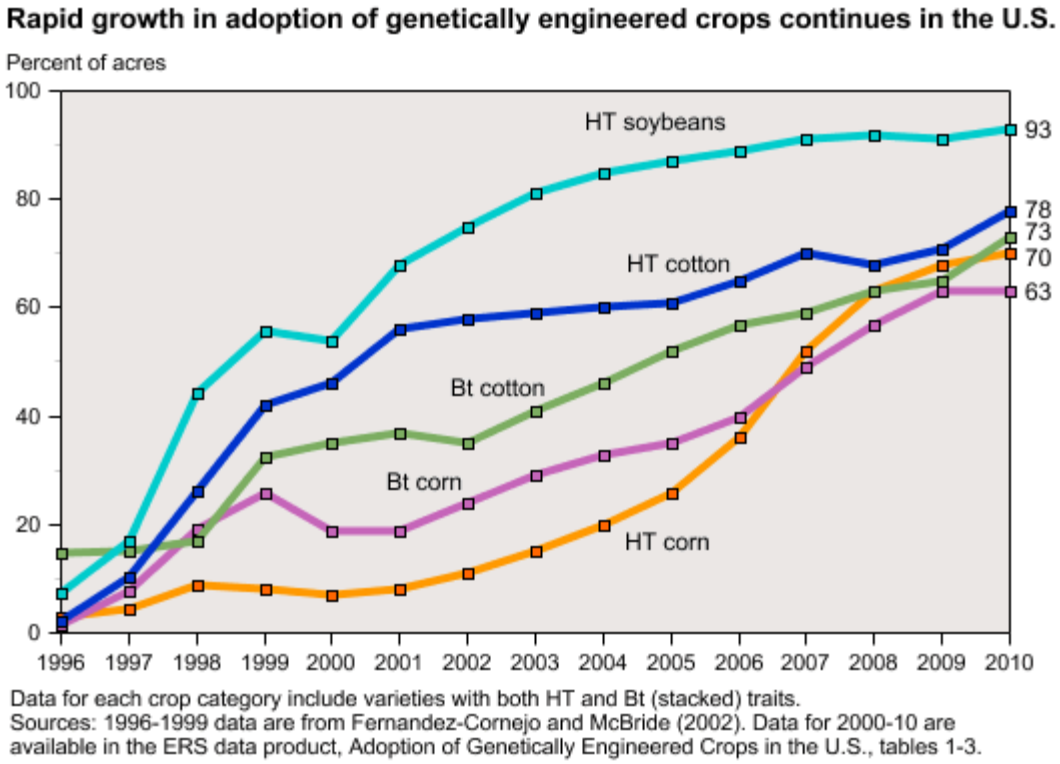
“Genetic modification is analogous to nuclear power: nobody loves it, but climate change has made its adoption imperative,” says economist Paul Collier of Oxford University. "Declining genetic modification makes a complicated issue more complex. Genetic

modification offers both faster crop adaptation and a biological, rather than chemical, approach to yield increases."

On the other hand, many believe that GM food has not been a success and that we should devote our efforts and money into another solution. “We need biodiversity intensification that works with nature’s nutrient and water cycles, not against them,” says Vandana Shiva, the founder of Navdanya, the movement of 500,000 seed keepers and organic farmers in India, argues that GMF’s have not increased yields. Recently, Doug Gurian-Sherman, a member of the Union of Concerned Scientists, a nonprofit science advocacy group, published a report called “Failure to Yield”, in which he stated that in a nearly 20 year record, genetically engineered crops have not increased yields substantially of food and livestock feed crops in the United States.

Some claim that genetically modified food help farmers produce, despite the odds or any environmental barriers. “While new technology must be tested before it is commercially released, we should be mindful of the risks of not releasing it at all,” says Per Pinstrup-Andersen, professor of Food, Nutrition and Public Policy at Cornell University. Per Pinstrup-Anderson argues, “Misguided anti-science ideology and failure by governments to prioritize agricultural and rural development in developing countries brought us the food crisis.” He clearly states the challenge we face is not the challenge of whether we have enough resources to produce, but whether we will change our behavior.

Economic and environmental effects



Adoption of genetically-engineered crops in the United States

- Many proponents of genetically engineered crops claim they lower pesticide usage and have brought higher yields and profitability to many farmers, including those in developing nations. For example, a 2010 study by US scientists, found that the economic benefit of Bt corn to farmers in five mid-west states was $6.9 billion over the previous 14 years. They were surprised that the majority ($4.3 billion) of the benefit accrued to non-Bt corn. This was speculated to be because the European Corn Borers that attack the Bt corn die and there are fewer left to attack the non-GM corn nearby.

- The United States has seen a widespread adoption of genetically-engineered corn, cotton and soybean crops since 1996 (see figure).

- In 2010, the U.S. National Academy of Sciences reported that genetically engineered crops had resulted in reduced pesticide application and reduced soil erosion from tilling. The report also stated that the advent of glyphosate-herbicide resistant weeds—that have developed because of the use of engineered crops—could cause the genetically engineered crops to lose their effectiveness unless farmers also use other established weed management strategies.

- In a study by Scientists at the University of Arkansas published in 2010 showed that about 83 percent of wild or weedy canola they tested contained genetically modified herbicide resistance genes, and they also found some plants that contained resistance to both herbicides, a combination of transgenic traits that had not been developed in canola crops. That leads us to believe that these wild populations that contain modified genes have become established populations.

Bans

- In 2002, Zambia cut off the flow of Genetically Modified Food (mostly maize) from UN's World Food Programme. This left a famine-stricken population without food aid.
- In December 2005 the Zambian government changed its mind in the face of further famine and allowed the importation of GM maize. However, the Zambian Minister for Agriculture Mundia Sikatana has insisted that the ban on genetically modified maize remains, saying "We do not want GM (genetically modified) foods and our hope is that all of us can continue to produce non-GM foods."
- In April 2004 Hugo Chávez announced a total ban on genetically modified seeds in Venezuela.
- In January 2005, the Hungarian government announced a ban on importing and planting of genetic modified maize seeds, which was subsequently authorized by the EU.
- On August 18, 2006, American exports of rice to Europe were interrupted when much of the U.S. crop was confirmed to be contaminated with unapproved engineered genes, possibly caused by cross-pollination with conventional crops.
- On February 9, 2010, Indian Environment Minister, Jairam Ramesh, imposed a moratorium on the cultivation of GMF "for as long as it is needed to establish

public trust and confidence". His decision was made after protest from several groups responding to regulatory approval of the cultivation of Bt brinjal, a GM eggplant in October, 2009.

U.S. government reaction to European ban

In recent years, France and several other European countries banned Monsanto's MON-810 corn and similar genetically modified food crops. In late 2007, the U.S. ambassador to France recommended "moving to retaliation" against France and the European Union in an attempt to fight the French ban and changes in European policy toward genetically modified crops, according to a U.S. government diplomatic cable obtained by WikiLeaks. The U.S. ambassador to France recommended retaliation to cause "some pain across the EU."

Intellectual property

Traditionally, farmers in all nations saved their own seed from year to year. It should be noted that this does not apply in more agriculturally developed countries for some crops. Corn is one example where producers generally have not saved seed since the early 1900s with the advent of hybrid corn through selective breeding. Seed producers grow the seed corn instead due to the effort needed to produce hybrids. The offspring of the hybrid corn, while still viable, lose the beneficial traits of the parents, resulting in the loss of hybrid vigor. In these cases, the use of hybrid plants has been the primary reason for growers not saving seed, not intellectual property issues, and has been in practice well before genetically-modified seed was developed. However, the practice of not saving seed in non-hybrid crops, such as soybean, is mainly due to intellectual property regulations. Allowing to follow this practice with genetically modified seed would result in seed developers losing the ability to profit from their breeding work. Therefore, genetically-modified seed is subject to licensing by their developers in contracts that are written to prevent farmers from following this practice.

Enforcement of patents on genetically modified plants is often contentious, especially because of gene flow. In 1998, 95-98 percent of about 10 km^2 planted with canola by Canadian farmer Percy Schmeiser were found to contain Monsanto Company's patented Roundup Ready gene although Schmeiser had never purchased seed from Monsanto. The initial source of the plants was undetermined, and could have been through either gene flow or intentional theft. However, the overwhelming predominance of the trait implied that Schmeiser must have intentionally selected for it. The court determined that Schmeiser had saved seed from areas on and adjacent to his property where Roundup had been sprayed, such as ditches and near power poles.

Although unable to prove direct theft, Monsanto sued Schmeiser for piracy since he knowingly grew *Roundup Ready* plants without paying royalties (Ibid). The case made it to the Canadian Supreme Court, which in 2004 ruled 5 to 4 in Monsanto’s favor. The dissenting judges focused primarily on the fact that Monsanto's patents covered only the gene itself and glyphosate resistant *cells*, and failed to cover transgenic plants in their

entirety. All of the judges agreed that Schmeiser would not have to pay any damages since he had not benefited from his use of the genetically modified seed.

In response to criticism, Monsanto Canada's Director of Public Affairs stated that "It is not, nor has it ever been Monsanto Canada's policy to enforce its patent on Roundup Ready crops when they are present on a farmer's field by accident...Only when there has been a knowing and deliberate violation of its patent rights will Monsanto act."

Future developments

Future envisaged applications of GMOs are diverse and include drugs in food, bananas that produce human vaccines against infectious diseases such as Hepatitis B, metabolically engineered fish that mature more quickly, fruit and nut trees that yield years earlier, foods no longer containing properties associated with common intolerances, and plants that produce new plastics with unique properties. While their practicality or efficacy in commercial production has yet to be fully tested, the next decade may see exponential increases in GM product development as researchers gain increasing access to genomic resources that are applicable to organisms beyond the scope of individual projects. Safety testing of these products will also, at the same time, be necessary to ensure that the perceived benefits will indeed outweigh the perceived and hidden costs of development. Plant scientists, backed by results of modern comprehensive profiling of crop composition, point out that crops modified using GM techniques are less likely to have unintended changes than are conventionally bred crops.

Health risks

In the United States, the FDA Center for Food Safety and Applied Nutrition reviews summaries of food safety data developed and voluntarily submitted by developers of engineered foods, in part on the basis of comparability to conventionally-produced foods. There are no specific tests required by FDA to determine safety. FDA does not approve the safety of engineered foods, but after its review, acknowledges that the developer of the food has asserted that it is safe. The table below shows the foods that have been reviewed by FDA as of 2002.

Table 1: GM Foods for Human Consumption Evaluated by FDA

Modified Attribute	Insect Resistance	Viral Resistance	Herbicide Tolerance	Modified Oil	Plant Reproductive Sterility	Delayed Ripening/ Softening
GM Plant Product—# of Plant Varieties	Corn—8 Tomato—1 Potato—4 Cotton—2	Squash—2 Papaya—1 Potato—2	Corn—9 Rice—1 Canola—8 Sugar Beet—2 Flax—1 Cotton[a]—4 Radish—1 Soybean—2	Soybean—1 Canola—1	Corn—3 Canola—3 Radish—1	Cantaloupe—1 Tomato—4
Total[b]	15	5	28	2	7	5

[a]Cotton seed has been used as a protein source in candy.

[b]Fifty products have been evaluated, as of April 2002. The total number of modified attributes is 62 because several products were modified with multiple attributes.

Source: GAO analysis of FDA data.

A 2008 review published by the Royal Society of Medicine noted that GM foods have been eaten by millions of people worldwide for over 15 years, with no reports of ill effects. Similarly a 2004 report from the US National Academies of Sciences stated: "To date, no adverse health effects attributed to genetic engineering have been documented in the human population." There have, however, been no epidemiological studies to determine whether engineered crops have caused any harm to the public. Without such studies, it is unlikely that harm, if it occurred, would be detected or attributed to engineered foods. Worldwide, there are a range of perspectives within non-governmental organizations on the safety of GM foods. For example, the US pro-GM pressure group AgBioWorld has argued that GM foods have been proven safe, while other pressure groups and consumer rights groups, such as the Organic Consumers Association, and Greenpeace claim the long term health risks which GM could pose, or the environmental risks associated with GM, have not yet been adequately investigated.

In 1998 Rowett Research Institute scientist Árpád Pusztai reported that consumption of potatoes genetically modified to contain lectin had adverse intestinal effects on rats. Pusztai eventually published a paper, co-authored by Stanley Ewen, in the journal, The Lancet. The paper claimed to show that rats fed on potatoes genetically modified with the snowdrop lectin had unusual changes to their gut tissue when compared with rats fed on non modified potatoes. The experiment modified potatoes to add a toxin (snowdrop lectin), but the experiment failed to include a control for the toxin alone or a control for genetic modifications alone (without added toxin); therefore, no conclusion could be made about the safety of the genetic engineering. The experiment has been criticised by other scientists on the grounds that the unmodified potatoes were not a fair control diet and that all the rats may have been sick, due to them being fed a diet of only potatoes.

In 2009 three scientists (Vendômois et al) published a statistical re-analysis of three feeding trials that had previously been published by others as establishing the safety of genetically modified corn. The new article claimed that their statistics instead showed that the three patented crops (Mon 810, Mon 863, and NK 603) developed and owned by Monsanto cause liver, kidney, and heart damage in mammals. A 2007 analysis of part of this data by the same group of scientists funded by Greenpeace was assessed by a panel of independent toxicologists in a study funded by Monsanto and published in the journal

Food and chemical toxicology. The reviewers reported that the study was statistically flawed and providing no evidence of adverse effects. The French High Council of Biotechnologies Scientific Committee reviewed the 2009 Vendômois et al study and concluded that it "..presents no admissible scientific element likely to ascribe any haematological, hepatic or renal toxicity to the three re-analysed GMOs. An evaluation by the European Food Safety Authority of the 2009 and 2007 studies noted that most of the results were within natural variation and they did not consider any of the effects reported biologically relevant. A review by Food Standards Australia New Zealand of the 2009 Vendômois et al study concluded that the results were due to chance alone.

Gene transfer

As of January 2009 there has only been one human feeding study conducted on the effects of genetically modified foods. The study involved seven human volunteers who had previously had their large intestines removed. These volunteers were to eat GM soy to see if the DNA of the GM soy transferred to the bacteria that naturally lives in the human gut. Researchers identified that three of the seven volunteers had transgenes from GM soya transferred into the bacteria living in their gut before the start of the feeding experiment. As this low-frequency transfer did not increase after the consumption of GM Soy, the researchers concluded that gene transfer did not occur during the experiment. In volunteers with complete digestive tracts, the transgene did not survive passage through intact gastrointestinal tract. Anti-GM advocates believe the study should prompt additional testing to determine its significance.. Other studies have found DNA from M13 virus, GFP and even ribulose-1,5-bisphosphate carboxylase (Rubisco) genes in the blood and tissue of ingesting animals (reviewed by

Two studies on the possible effects of feeding genetically modified feeds to animals found that there was no significant differences in the safety and nutritional value of feedstuffs containing material derived from genetically modified plants. Specifically, the studies noted that no residues of recombinant DNA or novel proteins have been found in any organ or tissue samples obtained from animals fed with GMP plants.

Allergies

In the mid 1990s Pioneer Hi-Bred tested the allergenicity of a transgenic soybean that expressed a Brazil nut seed storage protein in hope that the seeds would have increased levels of the amino acid methionine. The tests (radioallergosorbent testing, immunoblotting, and skin-prick testing) showed that individuals allergic to Brazil nuts were also allergic to the new GM soybean. Pioneer has indicated that it will not develop commercial cultivars containing Brazil nut protein because the protein is likely to be an allergen.

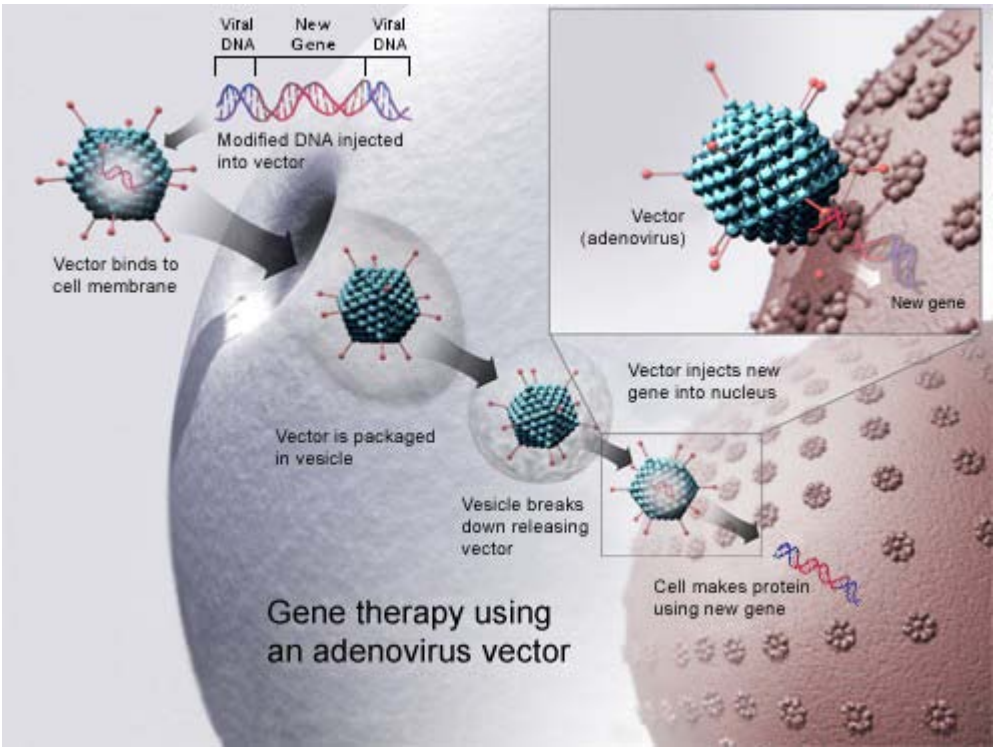
Traceability

In a January 2010 paper by Costa et al. the extraction and detection of DNA along a complete industrial soybean oil processing chain was described to monitor the presence

of Roundup Ready (RR) soybean: "The amplification of soybean lectin gene by end-point polymerase chain reaction (PCR) was successfully achieved in all the steps of extraction and refining processes, until the fully refined soybean oil. The amplification of RR soybean by PCR assays using event-specific primers was also achieved for all the extraction and refining steps, except for the intermediate steps of refining (neutralisation, washing and bleaching) possibly due to sample instability. The real-time PCR assays using specific probes confirmed all the results and proved that it is possible to detect and quantify genetically modified organisms in the fully refined soybean oil. To our knowledge, this has never been reported before and represents an important accomplishment regarding the traceability of genetically modified organisms in refined oils."

Chapter- 9

Gene Therapy



Gene therapy

Gene therapy is the insertion, alteration, or removal of genes within an individual's cells and biological tissues to treat disease. The most common form of gene therapy involves the insertion of functional genes into an unspecified genomic location in order to replace a mutated gene, but other forms involve directly correcting the mutation or modifying normal gene that enables a viral infection. Although the technology is still in its infancy, it has been used with some success. Scientific breakthroughs continue to move gene therapy toward mainstream medicine.

Approach

Scientists have taken the logical step of trying to introduce genes directly into human cells, focusing on diseases caused by single-gene defects, such as cystic fibrosis,

haemophilia, muscular dystrophy and sickle cell anemia. However, this has proven more difficult than modifying bacteria, primarily because of the problems involved in carrying large sections of DNA and delivering them to the correct site on the gene. Today, most gene therapy studies are aimed at cancer and hereditary diseases linked to a genetic defect. Antisense therapy is not strictly a form of gene therapy, but is a related, genetically-mediated therapy.

The most common form of genetic engineering involves the insertion of a functional gene at an unspecified location in the host genome. This is accomplished by isolating and copying the gene of interest, generating a construct containing all the genetic elements for correct expression, and then inserting this construct into a random location in the host organism. Other forms of genetic engineering include gene targeting and knocking out specific genes via engineered nucleases such as zinc finger nucleases, engineered I-CreI homing endonucleases, or nucleases generated from TAL effectors. An example of gene-knockout mediated gene therapy is the knockout of the human CCR5 gene in T-cells in order to control HIV infection. This approach is currently being used in several human clinical trials.

The biology of human gene therapy remains complex and many techniques need further development. Many diseases and their strict genetic link need to be understood more fully before gene therapy can be used appropriately. The public policy debate surrounding the possible use of genetically engineered material in human subjects has been equally complex. Major participants in the debate have come from the fields of biology, government, law, medicine, philosophy, politics, and religion, each bringing different views to the discussion.

Types of gene therapy

Gene therapy may be classified into the two following types:

Germ line gene therapy

In the case of germ line gene therapy, germ cells, i.e., sperm or eggs, are modified by the introduction of functional genes, which are ordinarily integrated into their genomes. Therefore, the change due to therapy would be heritable and would be passed on to later generations. This new approach, theoretically, should be highly effective in counteracting genetic disorders and hereditary diseases. However, many jurisdictions prohibit this for application in human beings, at least for the present, for a variety of technical and ethical reasons.

Somatic gene therapy

In the case of somatic gene therapy, the therapeutic genes are transferred into the somatic cells of a patient. Any modifications and effects will be restricted to the individual patient only, and will not be inherited by the patient's offspring or later generations.

Vectors in gene therapy

Viruses

All viruses bind to their hosts and introduce their genetic material into the host cell as part of their replication cycle. This genetic material contains basic 'instructions' of how to produce more copies of these viruses, hijacking the body's normal production machinery to serve the needs of the virus. The host cell will carry out these instructions and produce additional copies of the virus, leading to more and more cells becoming infected. Some types of viruses insert their genes into the host's genome, but do not actually enter the cell. Others penetrate the cell membrane disguised as protein molecules and enter the cell.

There are two main types of virus infection: lytic and lysogenic. Shortly after inserting its DNA, viruses of the lytic cycle quickly produce more viruses, burst from the cell and infect more cells. Lysogenic viruses integrate their DNA into the DNA of the host cell and may live in the body for many years before responding to a trigger. The virus reproduces as the cell does and does not inflict bodily harm until it is triggered. The trigger releases the DNA from that of the host and employs it to create new viruses. HIV is a lysogenic infection. Some scientists believe that if they find the origin of its trigger, they will be able to stop the virus from ever reproducing throughout the body.

Retroviruses

The genetic material in retroviruses is in the form of RNA molecules, while the genetic material of their hosts is in the form of DNA. When a retrovirus infects a host cell, it will introduce its RNA together with some enzymes, namely reverse transcriptase and integrase, into the cell. This RNA molecule from the retrovirus must produce a DNA copy from its RNA molecule before it can be integrated into the genetic material of the host cell. The process of producing a DNA copy from an RNA molecule is termed reverse transcription. It is carried out by one of the enzymes carried in the virus, called reverse transcriptase. After this DNA copy is produced and is free in the nucleus of the host cell, it must be incorporated into the genome of the host cell. That is, it must be inserted into the large DNA molecules in the cell (the chromosomes). This process is done by another enzyme carried in the virus called integrase.

Now that the genetic material of the virus has been inserted, it can be said that the host cell has been modified to contain new genes. If this host cell divides later, its descendants will all contain the new genes. Sometimes the genes of the retrovirus do not express their information immediately.

One of the problems of gene therapy using retroviruses is that the integrase enzyme can insert the genetic material of the virus into any arbitrary position in the genome of the host; it randomly inserts the genetic material into a chromosome. If genetic material happens to be inserted in the middle of one of the original genes of the host cell, this gene will be disrupted (insertional mutagenesis). If the gene happens to be one regulating cell

division, uncontrolled cell division (i.e., cancer) can occur. This problem has recently begun to be addressed by utilizing zinc finger nucleases or by including certain sequences such as the beta-globin locus control region to direct the site of integration to specific chromosomal sites.

Gene therapy trials using retroviral vectors to treat X-linked severe combined immunodeficiency (X-SCID) represent the most successful application of gene therapy to date. More than twenty patients have been treated in France and Britain, with a high rate of immune system reconstitution observed. Similar trials were restricted or halted in the USA when leukemia was reported in patients treated in the French X-SCID gene therapy trial. To date, four children in the French trial and one in the British trial have developed leukemia as a result of insertional mutagenesis by the retroviral vector. All but one of these children responded well to conventional anti-leukemia treatment. Gene therapy trials to treat SCID due to deficiency of the Adenosine Deaminase (ADA) enzyme continue with relative success in the USA, Britain, Italy and Japan.

Adenoviruses

Adenoviruses are viruses that carry their genetic material in the form of double-stranded DNA. They cause respiratory, intestinal, and eye infections in humans (especially the common cold). When these viruses infect a host cell, they introduce their DNA molecule into the host. The genetic material of the adenoviruses is not incorporated (transient) into the host cell's genetic material. The DNA molecule is left free in the nucleus of the host cell, and the instructions in this extra DNA molecule are transcribed just like any other gene. The only difference is that these extra genes are not replicated when the cell is about to undergo cell division so the descendants of that cell will not have the extra gene. As a result, treatment with the adenovirus will require readministration in a growing cell population although the absence of integration into the host cell's genome should prevent the type of cancer seen in the SCID trials. This vector system has been promoted for treating cancer and indeed the first gene therapy product to be licensed to treat cancer, Gendicine, is an adenovirus. Gendicine, an adenoviral p53-based gene therapy was approved by the Chinese FDA in 2003 for treatment of head and neck cancer. Advexin, a similar gene therapy approach from Introgen, was turned down by the US FDA in 2008.

Concerns about the safety of adenovirus vectors were raised after the 1999 death of Jesse Gelsinger while participating in a gene therapy trial. Since then, work using adenovirus vectors has focused on genetically crippled versions of the virus.

Adeno-associated viruses

Adeno-associated viruses, from the parvovirus family, are small viruses with a genome of single stranded DNA. The wild type AAV can insert genetic material at a specific site on chromosome 19 with near 100% certainty. But the recombinant AAV, which does not contain any viral genes and only the therapeutic gene, does not integrate into the genome. Instead the recombinant viral genome fuses at its ends via the ITR (inverted terminal repeats) recombination to form circular, episomal forms which are predicted to be the

primary cause of the long term gene expression. There are a few disadvantages to using AAV, including the small amount of DNA it can carry (low capacity) and the difficulty in producing it. The production problem however has recently been solved by Amsterdam Molecular Therapeutics. This type of virus is being used, however, because it is non-pathogenic (most people carry this harmless virus). In contrast to adenoviruses, most people treated with AAV will not build an immune response to remove the virus and the cells that have been successfully treated with it. Several trials with AAV are on-going or in preparation, mainly trying to treat muscle and eye diseases; the two tissues where the virus seems particularly useful. However, clinical trials have also been initiated where AAV vectors are used to deliver genes to the brain. This is possible because AAV viruses can infect non-dividing (quiescent) cells, such as neurons in which their genomes are expressed for a long time.

Envelope protein pseudotyping of viral vectors

The viral vectors described above have natural host cell populations that they infect most efficiently. Retroviruses have limited natural host cell ranges, and although adenovirus and adeno-associated virus are able to infect a relatively broader range of cells efficiently, some cell types are refractory to infection by these viruses as well. Attachment to and entry into a susceptible cell is mediated by the protein envelope on the surface of a virus. Retroviruses and adeno-associated viruses have a single protein coating their membrane, while adenoviruses are coated with both an envelope protein and fibers that extend away from the surface of the virus. The envelope proteins on each of these viruses bind to cell-surface molecules such as heparin sulfate, which localizes them upon the surface of the potential host, as well as with the specific protein receptor that either induces entry-promoting structural changes in the viral protein, or localizes the virus in endosomes wherein acidification of the lumen induces this refolding of the viral coat. In either case, entry into potential host cells requires a favorable interaction between a protein on the surface of the virus and a protein on the surface of the cell. For the purposes of gene therapy, one might either want to limit or expand the range of cells susceptible to transduction by a gene therapy vector. To this end, many vectors have been developed in which the endogenous viral envelope proteins have been replaced by either envelope proteins from other viruses, or by chimeric proteins. Such chimera would consist of those parts of the viral protein necessary for incorporation into the virion as well as sequences meant to interact with specific host cell proteins. Viruses in which the envelope proteins have been replaced as described are referred to as pseudotyped viruses. For example, the most popular retroviral vector for use in gene therapy trials has been the lentivirus Simian immunodeficiency virus coated with the envelope proteins, G-protein, from Vesicular stomatitis virus. This vector is referred to as VSV G-pseudotyped lentivirus, and infects an almost universal set of cells. This tropism is characteristic of the VSV G-protein with which this vector is coated. Many attempts have been made to limit the tropism of viral vectors to one or a few host cell populations. This advance would allow for the systemic administration of a relatively small amount of vector. The potential for off-target cell modification would be limited, and many concerns from the medical community would be alleviated. Most attempts to limit tropism have used chimeric envelope proteins

bearing antibody fragments. These vectors show great promise for the development of "magic bullet" gene therapies.

Replication-Competent Vectors

A replication-competent vector called ONYX-015 is used in replicating tumor cells. It was found that in the absence of the E1B-55Kd viral protein, adenovirus caused very rapid apoptosis of infected, p53(+) cells, and this results in dramatically reduced virus progeny and no subsequent spread. Apoptosis was mainly the result of the ability of EIA to inactivate p300. In p53(-) cells, deletion of E1B 55kd has no consequence in terms of apoptosis, and viral replication is similar to that of wild-type virus, resulting in massive killing of cells.

A replication-defective vector deletes some essential genes. These deleted genes are still necessary in the body so they are replaced with either a helper virus or a DNA molecule.

Cis and trans-acting elements

Replication-defective vectors always contain a “transfer construct”. The transfer construct carries the gene to be transduced or “transgene”. The transfer construct also carries the sequences which are necessary for the general functioning of the viral genome: packaging sequence, repeats for replication and, when needed, priming of reverse transcription. These are denominated cis-acting elements, because they need to be on the same piece of DNA as the viral genome and the gene of interest. Trans-acting elements are viral elements, which can be encoded on a different DNA molecule. For example, the viral structural proteins can be expressed from a different genetic element than the viral genome.

Herpes Simplex Virus

Herpes Simplex Virus is a human neurotropic virus. This is mostly examined for gene transfer in the nervous system. The wild type HSV-1 virus is able to infect neurons. Infected neurones are not rejected by the immune system. Though the latent virus is not transcriptionally apparent, it does possess neurone specific promoters that can continue to function normally. Antibodies to HSV-1 are common in humans, however complications due to herpes infection are somewhat rare.

Non-viral methods

Non-viral methods present certain advantages over viral methods, with simple large scale production and low host immunogenicity being just two. Previously, low levels of transfection and expression of the gene held non-viral methods at a disadvantage; however, recent advances in vector technology have yielded molecules and techniques with transfection efficiencies similar to those of viruses.

Injection of Naked DNA

This is the simplest method of non-viral transfection. Clinical trials carried out of intramuscular injection of a naked DNA plasmid have occurred with some success; however, the expression has been very low in comparison to other methods of transfection. In addition to trials with plasmids, there have been trials with naked PCR product, which have had similar or greater success. Cellular uptake of naked DNA is generally inefficient. Research efforts focusing on improving the efficiency of naked DNA uptake have yielded several novel methods, such as electroporation, sonoporation, and the use of a "gene gun", which shoots DNA coated gold particles into the cell using high pressure gas.

Physical Methods to Enhance Delivery

Electroporation

Electorporation is a method that uses short pulses of high voltage to carry DNA across the cell membrane. This shock is thought to cause temporary formation of pores in the cell membrane, allowing DNA molecules to pass through. Electroporation is generally efficient and works across a broad range of cell types. However, a high rate of cell death following electroporation has limited its use, including clinical applications.

More recently a newer method of electroporation, termed electron-avalanche transfection, has been used in gene therapy experiments. By using a high-voltage plasma discharge, DNA was efficiently delivered following very short (microsecond) pulses. Compared to electroporation, the technique resulted in greatly increased efficiency and less cellular damage.

Gene Gun

The use of particle bombardment, or the gene gun, is another physical method of DNA transfection. In this technique, DNA is coated with gold particles and loaded into a device which generates a force to achieve penetration of DNA/gold into the cells.

Sonoporation

Sonoporation uses ultrasonic frequencies to deliver DNA into cells. The process of acoustic cavitation is thought to disrupt the cell membrane and allow DNA to move into cells.

Magnetofection

In a method termed magnetofection, DNA is complexed to a magnetic particles, and a magnet is placed underneath the tissue culture dish to bring DNA complexes into contact with a cell monolayer.

Chemical Methods to Enhance Delivery

Oligonucleotides

The use of synthetic oligonucleotides in gene therapy is to inactivate the genes involved in the disease process. There are several methods by which this is achieved. One strategy uses antisense specific to the target gene to disrupt the transcription of the faulty gene. Another uses small molecules of RNA called siRNA to signal the cell to cleave specific unique sequences in the mRNA transcript of the faulty gene, disrupting translation of the faulty mRNA, and therefore expression of the gene. A further strategy uses double stranded oligodeoxynucleotides as a decoy for the transcription factors that are required to activate the transcription of the target gene. The transcription factors bind to the decoys instead of the promoter of the faulty gene, which reduces the transcription of the target gene, lowering expression. Additionally, single stranded DNA oligonucleotides have been used to direct a single base change within a mutant gene. The oligonucleotide is designed to anneal with complementarity to the target gene with the exception of a central base, the target base, which serves as the template base for repair. This technique is referred to as oligonucleotide mediated gene repair, targeted gene repair, or targeted nucleotide alteration.

Lipoplexes and polyplexes

To improve the delivery of the new DNA into the cell, the DNA must be protected from damage and its entry into the cell must be facilitated. To this end new molecules, lipoplexes and polyplexes, have been created that have the ability to protect the DNA from undesirable degradation during the transfection process.

Plasmid DNA can be covered with lipids in an organized structure like a micelle or a liposome. When the organized structure is complexed with DNA it is called a lipoplex. There are three types of lipids, anionic (negatively charged), neutral, or cationic (positively charged). Initially, anionic and neutral lipids were used for the construction of lipoplexes for synthetic vectors. However, in spite of the facts that there is little toxicity associated with them, that they are compatible with body fluids and that there was a possibility of adapting them to be tissue specific; they are complicated and time consuming to produce so attention was turned to the cationic versions.

Cationic lipids, due to their positive charge, were first used to condense negatively charged DNA molecules so as to facilitate the encapsulation of DNA into liposomes. Later it was found that the use of cationic lipids significantly enhanced the stability of lipoplexes. Also as a result of their charge, cationic liposomes interact with the cell membrane, endocytosis was widely believed as the major route by which cells uptake lipoplexes. Endosomes are formed as the results of endocytosis, however, if genes can not be released into cytoplasm by breaking the membrane of endosome, they will be sent to lysosomes where all DNA will be destroyed before they could achieve their functions. It was also found that although cationic lipids themselves could condense and encapsulate DNA into liposomes, the transfection efficiency is very low due to the lack of ability in

terms of “endosomal escaping”. However, when helper lipids (usually electroneutral lipids, such as DOPE) were added to form lipoplexes, much higher transfection efficiency was observed. Later on, it was figured out that certain lipids have the ability to destabilize endosomal membranes so as to facilitate the escape of DNA from endosome, therefore those lipids are called fusogenic lipids. Although cationic liposomes have been widely used as an alternative for gene delivery vectors, a dose dependent toxicity of cationic lipids were also observed which could limit their therapeutic usages.

The most common use of lipoplexes has been in gene transfer into cancer cells, where the supplied genes have activated tumor suppressor control genes in the cell and decrease the activity of oncogenes. Recent studies have shown lipoplexes to be useful in transfecting respiratory epithelial cells, so they may be used for treatment of genetic respiratory diseases such as cystic fibrosis.

Complexes of polymers with DNA are called polyplexes. Most polyplexes consist of cationic polymers and their production is regulated by ionic interactions. One large difference between the methods of action of polyplexes and lipoplexes is that polyplexes cannot release their DNA load into the cytoplasm, so to this end, co-transfection with endosome-lytic agents (to lyse the endosome that is made during endocytosis, the process by which the polyplex enters the cell) such as inactivated adenovirus must occur. However, this isn't always the case, polymers such as polyethylenimine have their own method of endosome disruption as does chitosan and trimethylchitosan.

Dendrimers

A dendrimer is a highly branched macromolecule with a spherical shape. The surface of the particle may be functionalized in many ways and many of the properties of the resulting construct are determined by its surface.

In particular it is possible to construct a cationic dendrimer, i.e. one with a positive surface charge. When in the presence of genetic material such as DNA or RNA, charge complimentarity leads to a temporary association of the nucleic acid with the cationic dendrimer. On reaching its destination the dendrimer-nucleic acid complex is then taken into the cell via endocytosis.

In recent years the benchmark for transfection agents has been cationic lipids. Limitations of these competing reagents have been reported to include: the lack of ability to transfect a number of cell types, the lack of robust active targeting capabilities, incompatibility with animal models, and toxicity. Dendrimers offer robust covalent construction and extreme control over molecule structure, and therefore size. Together these give compelling advantages compared to existing approaches.

Producing dendrimers has historically been a slow and expensive process consisting of numerous slow reactions, an obstacle that severely curtailed their commercial development. The Michigan based company Dendritic Nanotechnologies discovered a method to produce dendrimers using kinetically driven chemistry, a process that not only

reduced cost by a magnitude of three, but also cut reaction time from over a month to several days. These new "Priostar" dendrimers can be specifically constructed to carry a DNA or RNA payload that transfects cells at a high efficiency with little or no toxicity.

Hybrid methods

Due to every method of gene transfer having shortcomings, there have been some hybrid methods developed that combine two or more techniques. Virosomes are one example; they combine liposomes with an inactivated HIV or influenza virus. This has been shown to have more efficient gene transfer in respiratory epithelial cells than either viral or liposomal methods alone. Other methods involve mixing other viral vectors with cationic lipids or hybridising viruses.

Major developments in gene therapy

1970s and earlier

In 1972 Friedmann and Roblin authored a paper in Science titled "Gene therapy for human genetic disease?" They cite Rogers S for proposing "that exogenous 'good'" DNA be used to replace the defective DNA in those who suffer from genetic defects. They also cite the first attempt to perform gene therapy as [New York Times, 20 September 1970].

1990

The first approved gene therapy case in the United States took place on September 14, 1990, at the National Institute of Health. It was performed on a four year old girl named Ashanti DeSilva. It was a treatment for a genetic defect that left her with an Immune System deficiency. The effects were only temporary, but successful (Boylan 313).

New gene therapy approach repairs errors in messenger RNA derived from defective genes. This technique has the potential to treat the blood disorder thalassaemia, cystic fibrosis, and some cancers. Researchers at Case Western Reserve University and Copernicus Therapeutics are able to create tiny liposomes 25 nanometers across that can carry therapeutic DNA through pores in the nuclear membrane.

Sickle cell disease is successfully treated in mice.

in 1992 Doctor Claudio Bordignon working at the Vita-Salute San Raffaele University, Milan, Italy performed the first procedure of gene therapy using hematopoietic stem cells as vectors to deliver genes intended to correct hereditary diseases. In 2002 this work led to the publication of the first successful gene therapy treatment for adenosine deaminase-deficiency (SCID). The success of a multi-center trial for treating children with SCID (severe combined immune deficiency or "bubble boy" disease) held from 2000 and 2002 was questioned when two of the ten children treated at the trial's Paris center developed a leukemia-like condition. Clinical trials were halted temporarily in 2002, but resumed

after regulatory review of the protocol in the United States, the United Kingdom, France, Italy, and Germany.

In 1993 Andrew Gobea was born with severe combined immunodeficiency (SCID). Genetic screening before birth showed that he had SCID. Blood was removed from Andrew's placenta and umbilical cord immediately after birth, containing stem cells. The allele that codes for ADA was obtained and was inserted into a retrovirus. Retroviruses and stem cells were mixed, after which they entered and inserted the gene into the stem cells' chromosomes. Stem cells containing the working ADA gene were injected into Andrew's blood system via a vein. Injections of the ADA enzyme were also given weekly. For four years T-cells (white blood cells), produced by stem cells, made ADA enzymes using the ADA gene. After four years more treatment was needed.

1995-2000

The 1999 death of Jesse Gelsinger in a gene therapy experiment resulted in a significant setback to gene therapy research in the United States. The pivotal event resulted in the FDA's suspension of several clinical trials as ethical and procedural practices in the field were reevaluated.

2001-2005

In 2003 a University of California, Los Angeles research team inserted genes into the brain using liposomes coated in a polymer called polyethylene glycol. The transfer of genes into the brain is a significant achievement because viral vectors are too big to get across the blood-brain barrier. This method has potential for treating Parkinson's disease.

RNA interference or gene silencing may be a new way to treat Huntington's disease. Short pieces of double-stranded RNA (short, interfering RNAs or siRNAs) are used by cells to degrade RNA of a particular sequence. If a siRNA is designed to match the RNA copied from a faulty gene, then the abnormal protein product of that gene will not be produced.

2005 to present

Scientists at the National Institutes of Health (Bethesda, Maryland) have successfully treated metastatic melanoma in two patients using killer T cells genetically retargeted to attack the cancer cells. This study constitutes one of the first demonstrations that gene therapy can be effective in treating cancer.

In March 2006 an international group of scientists announced the successful use of gene therapy to treat two adult patients for a disease affecting myeloid cells. The study, published in Nature Medicine, is believed to be the first to show that gene therapy can cure diseases of the myeloid system.

In May 2006 a team of scientists led by Dr. Luigi Naldini and Dr. Brian Brown from the San Raffaele Telethon Institute for Gene Therapy (HSR-TIGET) in Milan, Italy reported a breakthrough for gene therapy in which they developed a way to prevent the immune system from rejecting a newly delivered gene. Similar to organ transplantation, gene therapy has been plagued by the problem of immune rejection. So far, delivery of the 'normal' gene has been difficult because the immune system recognizes the new gene as foreign and rejects the cells carrying it. To overcome this problem, the HSR-TIGET group utilized a newly uncovered network of genes regulated by molecules known as microRNAs. Dr. Naldini's group reasoned that they could use this natural function of microRNA to selectively turn off the identity of their therapeutic gene in cells of the immune system and prevent the gene from being found and destroyed. The researchers injected mice with the gene containing an immune-cell microRNA target sequence, and the mice did not reject the gene, as previously occurred when vectors without the microRNA target sequence were used. This work will have important implications for the treatment of hemophilia and other genetic diseases by gene therapy.

In November 2006 Preston Nix from the University of Pennsylvania School of Medicine reported on VRX496, a gene-based immunotherapy for the treatment of human immunodeficiency virus (HIV) that uses a lentiviral vector for delivery of an antisense gene against the HIV envelope. In the Phase I trial enrolling five subjects with chronic HIV infection who had failed to respond to at least two antiretroviral regimens, a single intravenous infusion of autologous CD4 T cells genetically modified with VRX496 was safe and well tolerated. All patients had stable or decreased viral load; four of the five patients had stable or increased CD4 T cell counts. In addition, all five patients had stable or increased immune response to HIV antigens and other pathogens. This was the first evaluation of a lentiviral vector administered in U.S. Food and Drug Administration-approved human clinical trials for any disease. Data from an ongoing Phase I/II clinical trial were presented at CROI 2009.

On 1 May 2007 Moorfields Eye Hospital and University College London's Institute of Ophthalmology announced the world's first gene therapy trial for inherited retinal disease. The first operation was carried out on a 23 year-old British male, Robert Johnson, in early 2007. Leber's congenital amaurosis is an inherited blinding disease caused by mutations in the RPE65 gene. The results of the Moorfields/UCL trial were published in New England Journal of Medicine in April 2008. They researched the safety of the subretinal delivery of recombinant adeno associated virus (AAV) carrying RPE65 gene, and found it yielded positive results, with patients having modest increase in vision, and, perhaps more importantly, no apparent side-effects.

In September 2009, the journal Nature reported that researchers at the University of Washington and University of Florida were able to give trichromatic vision to squirrel monkeys using gene therapy, a hopeful precursor to a treatment for color blindness in humans. In November 2009, the journal Science reported that researchers succeeded at halting a fatal brain disease, adrenoleukodystrophy, using a vector derived from HIV to deliver the gene for the missing enzyme.

A paper by Komáromy *et* al. published in April 2010, deals with gene therapy for a form of achromatopsia in dogs. Achromatopsia, or complete color blindness, is presented as an ideal model to develop gene therapy directed to cone photoreceptors. Cone function and day vision have been restored for at least 33 months in two young dogs with achromatopsia. However, the therapy was less efficient for older dogs.

Problems and ethics

For the safety of gene therapy, the Weismann barrier is fundamental in the current thinking. Soma-to-germline feedback should therefore be impossible. However, there are indications that the Weissman barrier can be breached. One way it might possibly be breached is if the treatment were somehow misapplied and spread to the testes and therefore would infect the germline against the intentions of the therapy.

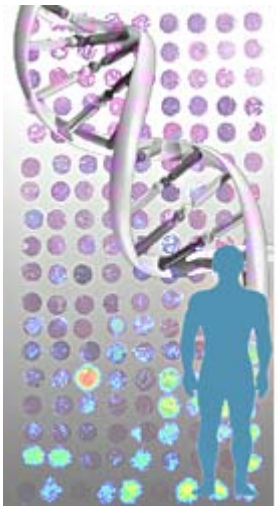
Some of the problems of gene therapy include:

- Short-lived nature of gene therapy – Before gene therapy can become a permanent cure for any condition, the therapeutic DNA introduced into target cells must remain functional and the cells containing the therapeutic DNA must be long-lived and stable. Problems with integrating therapeutic DNA into the genome and the rapidly dividing nature of many cells prevent gene therapy from achieving any long-term benefits. Patients will have to undergo multiple rounds of gene therapy.
- Immune response – Anytime a foreign object is introduced into human tissues, the immune system has evolved to attack the invader. The risk of stimulating the immune system in a way that reduces gene therapy effectiveness is always a possibility. Furthermore, the immune system's enhanced response to invaders that it has seen before makes it difficult for gene therapy to be repeated in patients.
- Problems with viral vectors – Viruses, the carrier of choice in most gene therapy studies, present a variety of potential problems to the patient —toxicity, immune and inflammatory responses, and gene control and targeting issues. In addition, there is always the fear that the viral vector, once inside the patient, may recover its ability to cause disease.
- Multigene disorders – Conditions or disorders that arise from mutations in a single gene are the best candidates for gene therapy. Unfortunately, some of the most commonly occurring disorders, such as heart disease, high blood pressure, Alzheimer's disease, arthritis, and diabetes, are caused by the combined effects of variations in many genes. Multigene or multifactorial disorders such as these would be especially difficult to treat effectively using gene therapy.
- Chance of inducing a tumor (insertional mutagenesis) - If the DNA is integrated in the wrong place in the genome, for example in a tumor suppressor gene, it could induce a tumor. This has occurred in clinical trials for X-linked severe combined immunodeficiency (X-SCID) patients, in which hematopoietic stem cells were transduced with a corrective transgene using a retrovirus, and this led to the development of T cell leukemia in 3 of 20 patients.

Deaths have occurred due to gene therapy, including that of Jesse Gelsinger.

Chapter- 10

Human Genetics



A small piece of human DNA

Human genetics describes the study of inheritance as it occurs in human beings. Human genetics encompasses a variety of overlapping fields including: classical genetics, cytogenetics, molecular genetics, biochemical genetics, genomics, population genetics, developmental genetics, clinical genetics, and genetic counseling. Genes can be the common factor of the qualities of most human-inherited traits. Study of human genetics can be useful as it can answer questions about human nature, understand the diseases and development of effective disease treatment, and understand genetics of human life.

Genetic differences and inheritance patterns

Inheritance of giant woods for black humans are based upon Gregor Mendel's model of inheritance. Mendel deduced that inheritance depends upon discrete units of inheritance, called factors or genes.

Autosomal dominant inheritance

Autosomal traits are associated with a single gene on an autosome (non-sex chromosome)—they are called "dominant" because a single copy—inherited from either

parent—is enough to cause this trait to appear. This often means that one of the parents must also have the same trait, unless it has arisen due to a new mutation. Examples of autosomal dominant traits and disorders are Huntington's disease, and achondroplasia.

Autosomal recessive inheritance

Autosomal recessive traits is one pattern of inheritance for a trait, disease, or disorder to be passed on through families. For a recessive trait or disease to be displayed two copies of the trait or disorder needs to be presented. The trait or gene will be located on a non-sex chromosome. Because it takes two copies of a trait to display a trait, many people can unknowingly be carriers of a disease. From an evolutionary perspective, a recessive disease or trait can remain hidden for several generations before displaying the phenotype. Examples of autosomal recessive disorders are albinism, Cystic Fibrosis, Tay-Sachs disease.

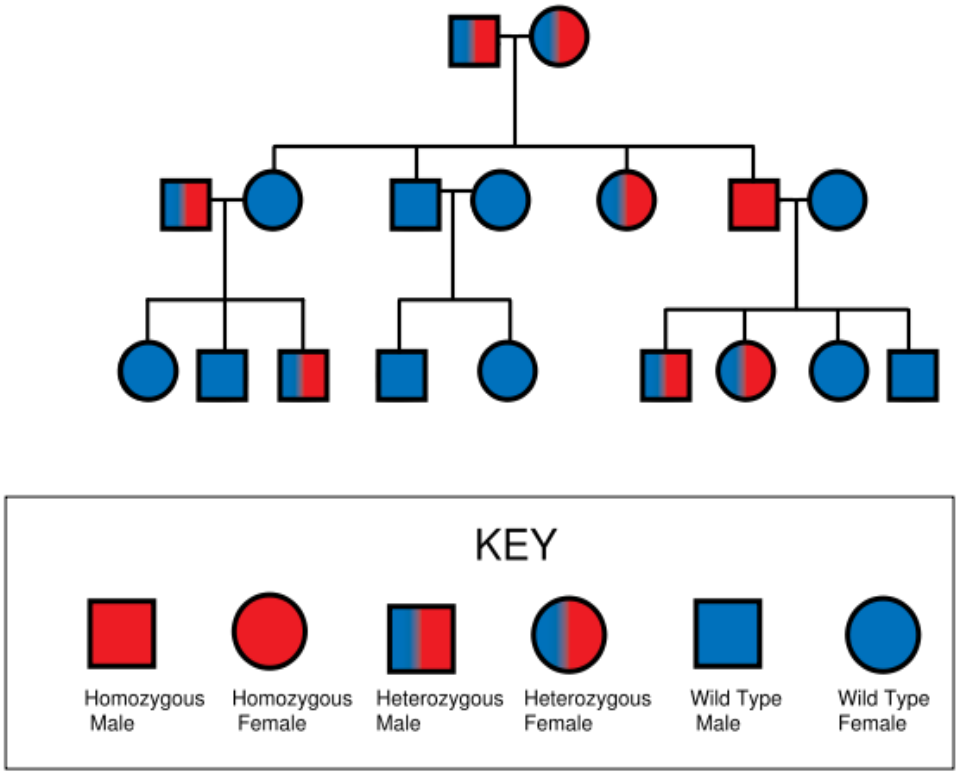
X-linked and Y-linked inheritance

X-linked genes are found on the sex X chromosome. X-linked genes just like autosomal genes have both dominant and recessive types. Recessive X-linked disorders are rarely seen in females and usually only affect males. This is because males inherit their X chromosome and all X-linked genes will be inherited from the maternal side. Fathers only pass on their Y chromosome to their sons, so no X-linked traits will be inherited from father to son. Females express X-linked disorders when they are homozygous for the disorder and become carriers when they are heterozygous. X-linked dominant inheritance will show the same phenotype as a heterozygote and homozygote. Just like X-linked inheritance, there will be a lack of male-to-male inheritance, which makes it distinguishable from autosomal traits. One example of a X-linked trait is Coffin-Lowry syndrome, which is caused by a mutation in ribosomal protein gene. This mutation results in skeletal, craniofacial abnormalities, mental retardation, and short stature.

X chromosomes in females undergo a process known as X inactivation. X inactivation is when one of the two X chromosomes in females is almost completely inactivated. It is important that this process occurs otherwise a woman would produce twice the amount of normal X chromosome proteins. The mechanism for X inactivation will occur during the embryonic stage. For people with disorders like trisomy X, where the genotype has three X chromosomes, X-inactivation will inactivate all X chromosomes until there is only one X chromosome active. X inactivation is not only limited to females, males with Klinefelter syndrome, who have an extra X chromosome, will also undergo X inactivation to have only one completely active X chromosome.

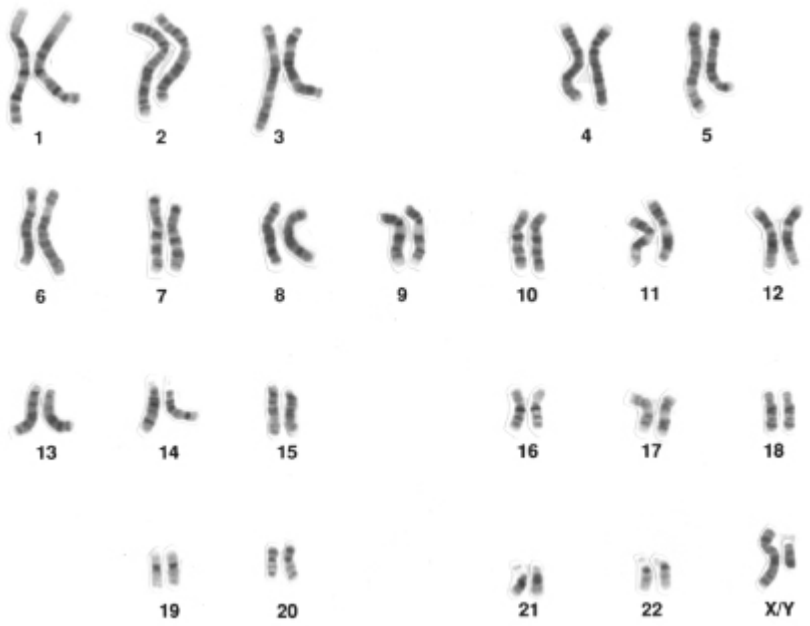
Y-linked inheritance occurs when a gene, trait, or disorder is transferred through the Y chromosome. Since Y chromosomes can only be found in males, Y linked traits are only passed on from father to son. The testis determining factor, which is located on the Y chromosome, determines the maleness of individuals. Besides the maleness inherited in the Y-chromosome there are no other found Y-linked characteristics.

Pedigrees



An example of a family pedigree displaying an autosomal recessive trait

A pedigree is a diagram showing the ancestral relationships and transmission of genetic traits over several generations in a family. Pedigrees are used to help detect many different genetic diseases. A pedigree can also be used to help determine the chances for a parent to produce an offspring with a specific trait. Four different traits can be identified by pedigree chart analysis: autosomal dominant, autosomal recessive, x-linked, or y-linked. Partial penetrance can be shown and calculated form pedigrees. Penetrance is the percentage expressed frequency with which individuals of a given genotype manifest at least some degree of a specific mutant phenotype associated with a trait. Inbreeding, the mating between closely related organisms of traits can clearly be seen on pedigree charts. Pedigree charts of royal families have a high degree of inbreeding, because it was customary and preferable for royalty to marry another member of royalty. Genetic councilors commonly use pedigrees to help couple determine if the parents will be able to produce healthy children.



A karyotype of a human male, showing 46 chromosomes including XY sex chromosomes.

Karyotype

A karyotype is a very useful tool in cytogenetics. A karyotype is picture of all the chromosomes in the metaphase stage arranged according to length and centromere position. A karyotype can also be useful in clinical genetics, due to its ability to diagnose genetic disorders. On a normal karyotype, aneuploidy can be detected by clearly being able to observe any missing or extra chromosomes. Giemsa banding, g-banding, of the karyotype can be used to detect deletions, insertions, duplications, inversions, and translocations. G-banding will stain the chromosomes with light and dark bands unique to each chromosome. A FISH, fluorescent in situ hybridization, can be used to observe deletions, insertions, and translocations. FISH uses fluorescent probes to bind to specific sequences of the chromosomes that will cause the chromosomes to fluoresce a unique color.

Genomics

Genomics refers to the field of genetics concerned with structural and functional studies of the genome. A genome is all the DNA contained within an organism or a cell including nuclear and mitochondrial DNA. The human genome is the total collection of genes in a human being contained in the human chromosome, composed of over three billion nucleotides. In April 2003, the Human Genome Project was able to sequence all the DNA in the human genome, to discover the human genome was composed around 20,000 protein coding genes.

Population genetics

Population genetics is the branch of evolutionary biology responsible for investigating processes that cause changes in allele and genotype frequencies in populations based upon Mendelian inheritance. Four different forces can influence the frequencies: natural selection, mutation, gene flow (migration), and genetic drift. A population can be defined as a group of interbreeding individuals and their offspring. For human genetics the populations will consist only of the human species. The Hardy-Weinberg principle is a widely used principle to determine allelic and genotype frequencies.

Hardy-Weinberg principle

The Hardy-Weinberg principle states that when no evolution occurs in a population the allele and genotype frequencies do not change from one generation to the next. No evolution refers to no mutation, no gene flow, no natural selection, and no genetic drift. To be in equilibrium two more assumptions need to be made that random mating occurs and there are discrete, non-overlapping generations.

Mitochondrial DNA

In addition to nuclear DNA, humans (like almost all eukaryotes) have mitochondrial DNA. Mitochondria, the "power houses" of a cell, have their own DNA because they are descended from a proteobacterium that merged with eukaryotic cells over 2 billion years ago—an assertion known as the endosymbiotic hypothesis. Mitochondria are inherited from one's mother, and its DNA is frequently used to trace maternal lines of descent. Mitochondrial DNA is only 16kb in length and encodes for 62 genes.

Genes and human characteristics

Genes are a fundamental unit of inheritance. Genes can be defined as a sequence of DNA in the genome that is required for production of a functional product. Genes have both minor and major effects on human characteristics. Human genes have become prominent in the nature versus nurture debate.

Genes and behavior

Genes have a strong influence on human behavior. IQ is largely heritable. However, this has been questioned. The stance that humans inherit substantial behavioral characteristics is called psychological nativism, compared to the stance that human behavior and culture are virtually entirely constructed (tabula rasa).

In the early 20th century, eugenics was policy in parts of the United States and Europe. The goal was to reduce or eliminate traits that were considered undesirable. One form of eugenics was compulsory sterilization of people deemed mentally unfit. Hitler's eugenics programs turned social consciousness against the practice, and psychological nativism became associated with racism and sexism.

Genes and gender

Evolutionary psychology

Evolutionary psychology explains many human behaviors as more or less moderated by genes that evolved in the hunter-gatherer stage of human cultural development.

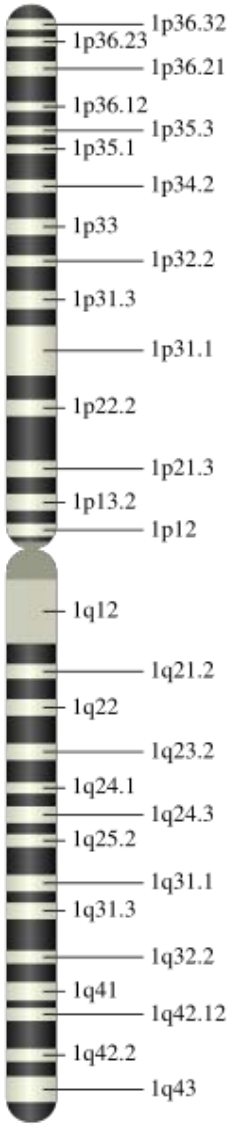
Human traits with simple inheritance patterns

Dominant	Recessive
Widow's peak	No Widow's peak
Facial Dimples	No Facial Dimples
Able to taste PTC	Unable to taste PTC
Unattached earlobe	Attached earlobe
Cleft chin	No Cleft chin
Brunette iris (anatomy)	Blue Iris (anatomy)
Color Vision	Color blindness
Brown Hair	Blonde Hair
normal	turned up nose
Ability to roll tongue (Able to hold tongue in a U shape)	No ability to roll tongue
Normal Pinkies	Crooked Pinkies
Normal Thumb	Hitchhiker's Thumb
Freckles	No Freckles
Wet-type earwax	Dry-type earwax
Curly Hair	Straight Hair

Chapter- 11

Human Chromosomes

Chromosome 1 (Human)



Map of Chromosome 1

Chromosome 1 is the designation for the largest human chromosome. Humans have two copies of chromosome 1, as they do with all of the autosomes, which are the non-sex chromosomes. Chromosome 1 spans about 247 million nucleotide base pairs, which are the basic units of information for DNA. It represents about 8% of the total DNA in human cells.

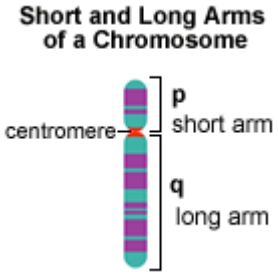
Identifying genes on each chromosome is an active area of genetic research. Chromosome 1 is currently believed to have 4,220 genes, exceeding previous predictions based on its size. It was the last completed chromosome, sequenced two decades after the beginning of the Human Genome Project.

The number of single nucleotide polymorphisms (SNPs) is about 740,000.

Genes

The following are some of the genes located on chromosome 1:

p-arm



Short and long arms

- ACADM: acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain
- COL11A1: collagen, type XI, alpha 1
- CPT2: carnitine palmitoyltransferase II
- DBT: dihydrolipoamide branched chain transacylase E2
- DIRAS3: DIRAS family, GTP-binding RAS-like 3
- ESPN: espin (autosomal recessive deafness 36)
- GALE: UDP-galactose-4-epimerase
- GJB3: gap junction protein, beta 3, 31kDa (connexin 31)
- HMGCL: 3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (hydroxymethylglutaricaciduria)
- KCNQ4: potassium voltage-gated channel, KQT-like subfamily, member 4
- KIF1B: kinesin family member 1B
- MFN2: mitofusin 2
- MTHFR: 5,10-methylenetetrahydrofolate reductase (NADPH)
- MUTYH: mutY homolog (E. coli)
- NGF: Nerve Growth Factor
- PARK7: Parkinson disease (autosomal recessive, early onset) 7

- PINK1: PTEN induced putative kinase 1
- PLOD1: procollagen-lysine 1, 2-oxoglutarate 5-dioxygenase 1
- TSHB: thyroid stimulating hormone, beta
- UROD: uroporphyrinogen decarboxylase (the gene for porphyria cutanea tarda)

q-arm

- ASPM: a brain size determinant
- F5: coagulation factor V (proaccelerin, labile factor)
- FMO3: flavin containing monooxygenase 3
- GBA: glucosidase, beta; acid (includes glucosylceramidase) (gene for Gaucher disease)
- GLC1A: gene for glaucoma
- HFE2: hemochromatosis type 2 (juvenile)
- HPC1: gene for prostate cancer
- IRF6: gene for connective tissue formation
- LMNA: lamin A/C
- MPZ: myelin protein zero (Charcot-Marie-Tooth neuropathy 1B)
- MTR: 5-methyltetrahydrofolate-homocysteine methyltransferase
- PPOX: protoporphyrinogen oxidase
- PSEN2: presenilin 2 (Alzheimer disease 4)
- SDHB: succinate dehydrogenase complex subunit B
- TNNT2: cardiac troponin T2
- USH2A: Usher syndrome 2A (autosomal recessive, mild)

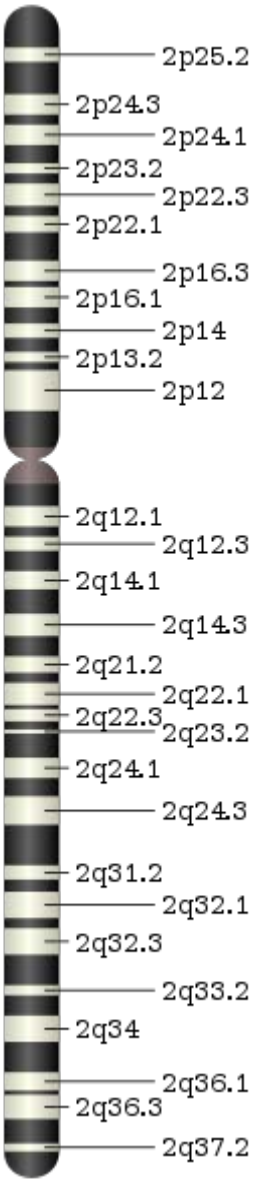
Diseases & disorders

There are 890 known diseases related to this chromosome. Some of these diseases are deafness, Alzheimer disease, glaucoma and breast cancer. Rearrangements and mutations of chromosome 1 are prevalent in cancer and many other diseases. Patterns of sequence variation reveal signals of recent selection in specific genes that may contribute to human fitness, and also in regions where no function is evident. The following diseases are some of those related to genes on chromosome 1 (which contains the most known genetic diseases of any human chromosome):

- 1q21.1 deletion syndrome
- 1q21.1 duplication syndrome
- Alzheimer disease
- Alzheimer disease, type 4
- Breast cancer
- Brooke Greenberg Disease (Syndrome X)
- Carnitine palmitoyltransferase II deficiency
- Charcot-Marie-Tooth disease
- Charcot-Marie-Tooth disease, type 1
- Charcot-Marie-Tooth disease, type 2
- collagenopathy, types II and XI

- congenital hypothyroidism
- Deafness, autosomal recessive deafness 36
- Ehlers-Danlos syndrome
- Ehlers-Danlos syndrome, kyphoscoliosis type
- Factor V Leiden thrombophilia
- Familial adenomatous polyposis
- galactosemia
- Gaucher disease
- Gaucher disease type 1
- Gaucher disease type 2
- Gaucher disease type 3
- Gaucher-like disease
- Gelatinous drop-like corneal dystrophy
- Glaucoma
- Hemochromatosis
- Hemochromatosis, type 2
- Hepatoerythropoietic porphyria
- Homocystinuria
- Hutchinson Gilford Progeria Syndrome
- 3-hydroxy-3-methylglutaryl-CoA lyase deficiency
- Hypertrophic cardiomyopathy, autosomal dominant mutations of TNNT2; hypertrophy usually mild; restrictive phenotype may be present; may carry high risk of sudden cardiac death
- maple syrup urine disease
- medium-chain acyl-coenzyme A dehydrogenase deficiency
- Microcephaly
- Muckle-Wells Syndrome
- Nonsyndromic deafness
- Nonsyndromic deafness, autosomal dominant
- Nonsyndromic deafness, autosomal recessive
- Oligodendroglioma
- Parkinson disease
- Pheochromocytoma
- porphyria
- porphyria cutanea tarda
- popliteal pterygium syndrome
- prostate cancer
- Stickler syndrome
- Stickler syndrome, COL11A1
- trimethylaminuria
- Usher syndrome
- Usher syndrome type II
- Van der Woude syndrome
- Variegate porphyria

Chromosome 2 (Human)



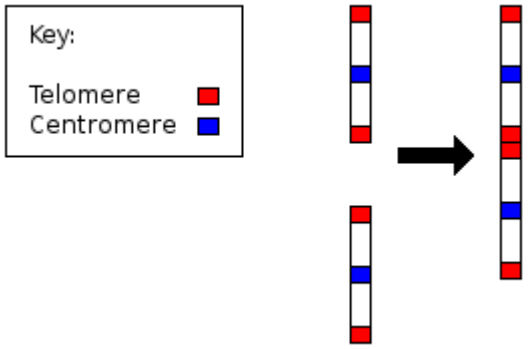
Chromosome 2 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 2 is the second largest human chromosome, spanning more than 237 million base pairs (the building material of DNA) and representing almost 8% of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome,

the estimated number of genes varies. Chromosome 2 likely contains 1,491 genes, including those of the HOXD homeobox gene cluster.

Evolution

All members of Hominidae except humans have 24 pairs of chromosomes. Humans have only 23 pairs of chromosomes. Human chromosome 2 is widely accepted to be a result of an end-to-end fusion of two ancestral chromosomes.



Fusion of ancestral chromosomes left distinctive remnants of telomeres, and a vestigial centromere

The evidence for this includes:

- The correspondence of chromosome 2 to two ape chromosomes. The closest human relative, the chimpanzee, has near-identical DNA sequences to human chromosome 2, but they are found in two separate chromosomes. The same is true of the more distant gorilla and orangutan.
- The presence of a vestigial centromere. Normally a chromosome has just one centromere, but in chromosome 2 there are remnants of a second centromere.
- The presence of vestigial telomeres. These are normally found only at the ends of a chromosome, but in chromosome 2 there are additional telomere sequences in the middle.

Some argue that chromosome 2 presents very strong evidence in favour of the common descent of humans and other apes. According to researcher J. W. IJdo, "We conclude that the locus cloned in cosmids c8.1 and c29B is the relic of an ancient telomere-telomere fusion and marks the point at which two ancestral ape chromosomes fused to give rise to human chromosome 2."

Genes

The following genes are located on chromosome 2:

- ABCA12: ATP-binding cassette, sub-family A (ABC1), member 12
- ABCG5 and ABCG8: ATP-binding cassette, subfamily A, members 5 and 8

- AGXT: alanine-glyoxylate aminotransferase (oxalosis I; hyperoxaluria I; glycolicaciduria; serine-pyruvate aminotransferase)
- ALMS1: Alstrom syndrome 1
- ALS2: amyotrophic lateral sclerosis 2 (juvenile)
- BMPR2: bone morphogenetic protein receptor, type II (serine/threonine kinase)
- COL3A1: collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant)
- COL4A3: collagen, type IV, alpha 3 (Goodpasture antigen)
- COL4A4: collagen, type IV, alpha 4
- COL5A2: collagen, type V, alpha 2
- HADHA: hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (trifunctional protein), alpha subunit
- HADHB: hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (trifunctional protein), beta subunit
- MSH2: mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli)
- MSH6: mutS homolog 6 (E. coli)
- NR4A2: nuclear receptor subfamily 4, group A, member 2
- OTOF: otoferlin
- PAX3: paired box gene 3 (Waardenburg syndrome 1)
- PAX8: paired box gene 8
- PELI1: Ubiquitin ligase
- SLC40A1: solute carrier family 40 (iron-regulated transporter), member 1
- TPO: thyroid peroxidase
- TBR1: T-box, brain, 1

Related diseases & disorders

The following diseases are related to genes located on chromosome 2:

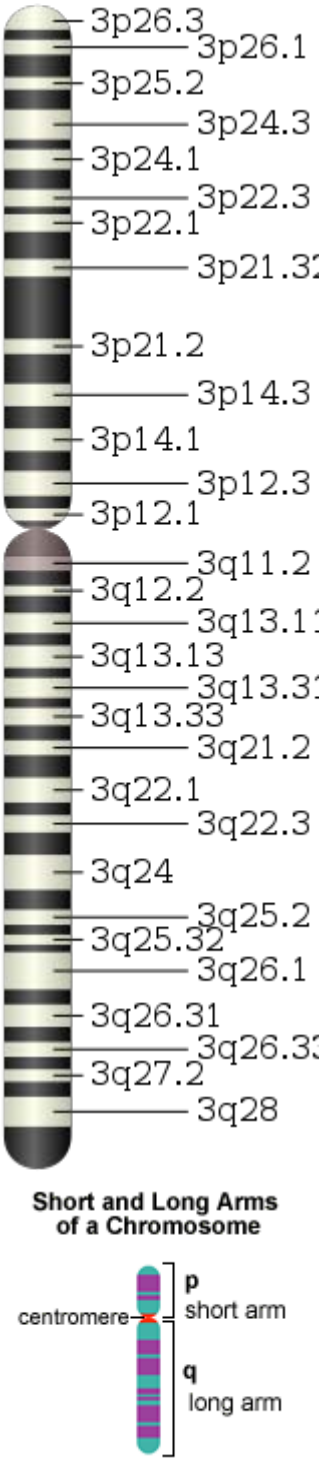
- Autism
- Alport syndrome
- Alström syndrome
- Amyotrophic lateral sclerosis
- Amyotrophic lateral sclerosis, type 2
- Congenital hypothyroidism
- Ehlers-Danlos syndrome
- Ehlers-Danlos syndrome, classical type
- Ehlers-Danlos syndrome, vascular type
- Fibrodysplasia ossificans progressiva
- Harlequin type ichthyosis
- Hemochromatosis
- Hemochromatosis, type 4
- Hereditary nonpolyposis colorectal cancer
- Infantile-onset ascending hereditary spastic paralysis
- Juvenile primary lateral sclerosis
- Long-chain 3-hydroxyacyl-coenzyme A dehydrogenase deficiency

- Maturity onset diabetes of the young type 6
- Mitochondrial trifunctional protein deficiency
- Nonsyndromic deafness
- Nonsyndromic deafness, autosomal recessive
- Primary hyperoxaluria
- Primary pulmonary hypertension
- Sitosterolemia (knockout of either ABCG5 or ABCG8)
- Synesthesia
- Waardenburg syndrome

Intelligence

Recent studies suggest that genes on chromosome 2 may play an important role in human intelligence.

Chromosome 3 (Human)



Short and long arms

Chromosome 3 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 3 spans almost 200 million base pairs (the building material of DNA) and represents about 6.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 3 likely contains between 1,100 and 1,500 genes.

Genes

The following are some of the genes located on chromosome 3:

p-arm

- ALAS1: aminolevulinate, delta-, synthase 1
- BTD: biotinidase
- CCR5: chemokine (C-C motif) receptor 5
- CNTN4: Contactin 4
- COL7A1: Collagen, type VII, alpha 1 (epidermolysis bullosa, dystrophic, dominant and recessive)
- MITF: microphthalmia-associated transcription factor
- MLH1: mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)
- OXTR: oxytocin receptor
- PTHR1: parathyroid hormone receptor 1
- SCN5A: sodium channel, voltage-gated, type V, alpha (long QT syndrome 3)
- SLC25A20: solute carrier family 25 (carnitine/acylcarnitine translocase), member 20
- TMIE: transmembrane inner ear
- VHL: von Hippel-Lindau tumor suppressor

q-arm

- CPOX: coproporphyrinogen oxidase (coproporphyria, harderoporphyria)
- HGD: homogentisate 1,2-dioxygenase (homogentisate oxidase)
- MCCC1: methylcrotonoyl-Coenzyme A carboxylase 1 (alpha)
- PCCB: propionyl Coenzyme A carboxylase, beta polypeptide
- PDCD10: programmed cell death 10
- PIK3CA: phosphoinositide-3-kinase, catalytic, alpha polypeptide
- RAB7: RAB7, member RAS oncogene family
- RHO: rhodopsin visual pigment
- SOX2: transcription factor
- USH3A: Usher syndrome 3A
- ZNF9: zinc finger protein 9 (a cellular retroviral nucleic acid binding protein)

Diseases & disorders

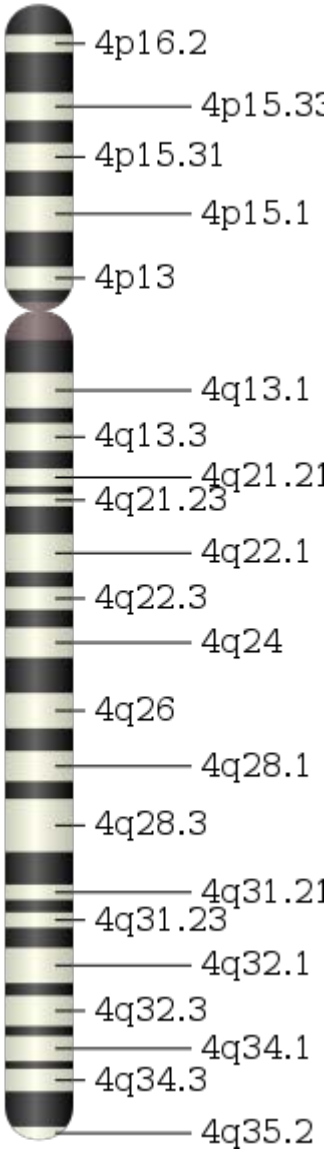
This list is incomplete; you can help by expanding it.

The following diseases are some of those related to genes on chromosome 3:

- 3-methylcrotonyl-CoA carboxylase deficiency
- 3q29 microdeletion syndrome
- Alkaptonuria
- Arrhythmogenic right ventricular dysplasia
- Atransferrinemia
- Autism
- Biotinidase deficiency
- Blepharophimosis, epicanthus inversus and ptosis type 1
- Breast/colon/lung/pancreatic cancer
- Brugada syndrome
- Carnitine-acylcarnitine translocase deficiency
- Cataracts
- Cerebral cavernous malformation
- Charcot-Marie-Tooth disease, type 2
- Charcot-Marie-Tooth disease
- Chromosome 3q duplication syndrome
- Coproporphyria
- Deafness
- Diabetes
- Dopamine receptor
- Dystrophic epidermolysis bullosa
- Endplate acetlycholinesterase deficiency
- Essential tremors
- Glaucoma, primary open angle
- Glycogen storage disease
- Hailey-Hailey disease
- Harderoporphyrinuria
- Heart block, progressive/nonprogressive
- Hereditary coproporphyria
- Hereditary nonpolyposis colorectal cancer
- HIV infection, susceptibility/resistance to
- Hypobetalipoproteinemia, familial
- Leukoencephalopathy with vanishing white matter
- Long QT syndrome
- Lymphomas
- Malignant hyperthermia susceptibility
- Metaphyseal chondrodysplasia, Murk Jansen type
- Moebius syndrome
- Moyamoya disease
- Mucopolysaccharidosis
- Muir-Torre family cancer syndrome

- Myotonic dystrophy, type 2
- Myotonic dystrophy
- Neuropathy, hereditary motor and sensory, Okinawa type
- Night blindness
- Nonsyndromic deafness, autosomal recessive
- Nonsyndromic deafness
- Ovarian cancer
- Porphyria
- Propionic acidemia
- Protein S deficiency
- Pseudo-Zellweger syndrome
- Retinitis pigmentosa
- Romano-Ward syndrome
- Septo-optic dysplasia
- Short stature
- Spinocerebellar ataxia
- Sucrose intolerance
- T-cell leukemia translocation altered gene
- Usher syndrome type III
- Usher syndrome (Finland)
- Usher syndrome
- von Hippel-Lindau syndrome
- Waardenburg syndrome
- Xeroderma pigmentosum, complementation group c

Chromosome 4 (Human)



Chromosome 4 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 4 spans more than 186 million base pairs (the building material of DNA) and represents between 6 and 6.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 4 likely contains between 700 and 1,100 genes.

Genes

The following are some of the genes located on chromosome 4:

- ANK2: ankyrin 2, neuronal
- CRMP1: Collapsin response mediator protein 1, a member of CRMP family
- CXCL1: chemokine (C-X-C motif) ligand 1, *scyb1*
- CXCL2: chemokine (C-X-C motif) ligand 2, *scyb2*
- CXCL3: chemokine (C-X-C motif) ligand 3, *scyb3*
- CXCL4: chemokine (C-X-C motif) ligand 4, Platelet factor-4, PF-4, *scyb4*
- CXCL5: chemokine (C-X-C motif) ligand 5, *scyb5*
- CXCL6: chemokine (C-X-C motif) ligand 6, *scyb6*
- CXCL7: chemokine (C-X-C motif) ligand 7, PPBP, *scyb7*
- CXCL8: chemokine (C-X-C motif) ligand 8, interleukin 8 (IL-8), *scyb8*
- CXCL9: chemokine (C-X-C motif) ligand 9, *scyb9*
- CXCL10: chemokine (C-X-C motif) ligand 10, *scyb10*
- CXCL11: chemokine (C-X-C motif) ligand 11, *scyb11*
- CXCL13: chemokine (C-X-C motif) ligand 13, *scyb13*
- DUX4: Thought to be inactive but 2010 research shows a key role in FSHD
- EVC: Ellis van Creveld syndrome
- EVC2: Ellis van Creveld syndrome 2 (limbin)
- FGFR3: fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism, bladder cancer)
- FGFRL1: fibroblast growth factor receptor-like 1
- Complement Factor I: Complement Factor I
- HTT (Huntingtin): huntingtin protein (Huntington's disease)
- MMAA: methylmalonic aciduria (cobalamin deficiency) cblA type
- PHOX2B: codes for a homeodomain transcription factor
- PKD2: polycystic kidney disease 2 (autosomal dominant)
- PLK4
- QDPR: quinoid dihydropteridine reductase
- SNCA: synuclein, alpha (non A4 component of amyloid precursor)
- UCHL1: ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)
- WFS1: Wolfram syndrome 1 (wolframin)
- FGF2: Fibroblast growth factor 2 (basic fibroblast growth factor)
- KDR: Kinase insert domain receptor (Vascular endothelial growth factor receptor 2)
- IGJ: linker protein for immunoglobulin alpha and mu polypeptides
- HCL2 (also called RHA or RHC): related to red hair

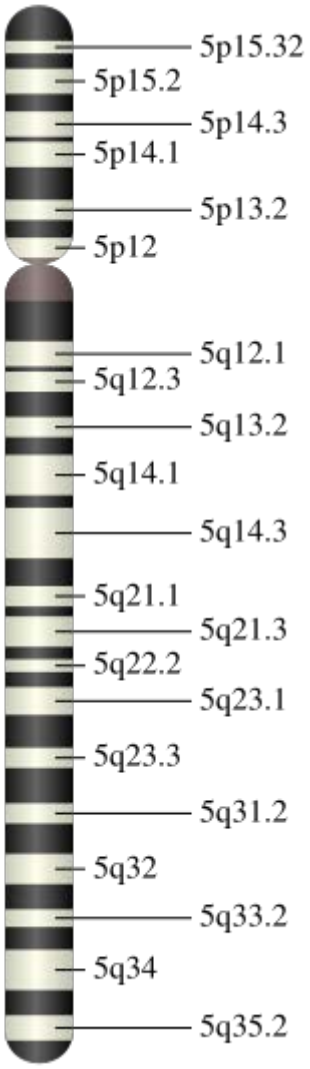
Diseases & disorders

The following are some of the diseases related to genes located on chromosome 4:

- achondroplasia
- bladder cancer

- Crouzonodermoskeletal syndrome
- Chronic Lymphocytic Leukemia
- Ellis-van Creveld syndrome
- Facioscapulohumeral muscular dystrophy
- Fibrodysplasia ossificans progessiva FOP
- Hemophilia C
- Huntington's disease
- Hemolytic Uremic Syndrome
- Hirschprung's disease
- hypochondroplasia
- methylmalonic acidemia
- Muenke syndrome
- nonsyndromic deafness
- nonsyndromic deafness, autosomal dominant
- Ondine's Curse
- Parkinsons disease
- polycystic kidney disease
- Romano-Ward syndrome
- SADDAN
- tetrahydrobiopterin deficiency
- thanatophoric dysplasia
- thanatophoric dysplasia, type 1
- thanatophoric dysplasia, type 2
- Wolfram syndrome

Chromosome 5 (Human)



Chromosome 5 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 5 spans about 181 million base pairs (the building blocks of DNA) and represents almost 6% of the total DNA in cells. **Chromosome 5** is one of the largest human chromosomes, yet has one of the lowest gene densities. This is partially explained by numerous gene-poor regions that display a remarkable degree of non-coding and syntenic conservation with non-mammalian vertebrates, suggesting they are functionally constrained.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 5 likely contains between 900 and 1,300 genes.

Genes

The following are some of the genes located on chromosome 5:

- ADAMTS2: ADAM metallopeptidase with thrombospondin type 1 motif, 2
- APC: adenomatosis polyposis coli
- EGR1: early growth response protein 1
- DTDST: diastrophic dysplasia sulfate transporter
- ERCC8: excision repair cross-complementing rodent repair deficiency, complementation group 8
- FGFR4: fibroblast growth factor receptor 4
- GM2A: GM2 ganglioside activator
- HEXB: hexosaminidase B (beta polypeptide)
- MASS1: monogenic, audiogenic seizure susceptibility 1 homolog (mouse)
- MCCC2: methylcrotonoyl-Coenzyme A carboxylase 2 (beta)
- MTRR: 5-methyltetrahydrofolate-homocysteine methyltransferase reductase
- NIPBL: Nipped-B homolog (Drosophila)
- NSD1: Transcription coregulator protein
- Pikachurin: Responsible for the functioning of the ribbon synapses; allows the eye to track moving objects
- SLC22A5: solute carrier family 22 (organic cation transporter), member 5
- SLC26A2: solute carrier family 26 (sulfate transporter), member 2
- SMN1: survival motor neuron 1, telomeric
- SMN2: survival motor neuron 2, centromeric
- SNCAIP: synuclein, alpha interacting protein (synphilin)
- TGFBI: keratoepithelin
- TCOF1: Treacher Collins-Franceschetti syndrome 1
- FGF1: fibroblast growth factor 1 (acidic fibroblast growth factor)

Diseases & disorders

The following are some of the diseases related to genes located on chromosome 5:

- Achondrogenesis type 1B
- Atelosteogenesis, type II
- Cockayne syndrome
- Cornelia de Lange syndrome
- Corneal dystrophy of Bowman layer, type I
- Corneal dystrophy of Bowman layer, type II
- Cri du Chat
- Diastrophic dysplasia
- Ehlers-Danlos syndrome
- Ehlers-Danlos syndrome, dermatosparaxis type
- Familial adenomatous polyposis
- Granular corneal dystrophy type I
- Granular corneal dystrophy type II

- GM2-gangliosidosis, AB variant
- Homocystinuria
- 3-Methylcrotonyl-CoA carboxylase deficiency
- Nicotine dependency
- Parkinson's disease
- Primary carnitine deficiency
- Recessive multiple epiphyseal dysplasia
- Sandhoff disease
- Spinal muscular atrophy
- Sotos Syndrome
- Survival motor neuron spinal muscular atrophy
- Treacher Collins syndrome
- Usher syndrome
- Usher syndrome type II

Chromosomal conditions

The following conditions are caused by changes in the structure or number of copies of chromosome 5:

- Cri-du-chat syndrome is caused by a deletion of the end of the short (p) arm of chromosome 5. This chromosomal change is written as 5p-. The signs and symptoms of cri-du-chat syndrome are probably related to the loss of multiple genes in this region. Researchers have not identified all of these genes or determined how their loss leads to the features of the disorder. They have discovered, however, that a larger deletion tends to result in more severe mental retardation and developmental delays in people with cri-du-chat syndrome.

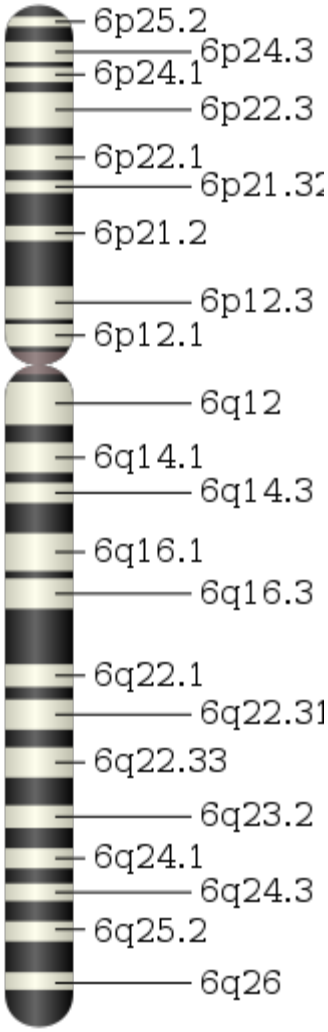
  Researchers have defined narrow regions of the short arm of chromosome 5 that are associated with particular features of cri-du-chat syndrome. A specific region designated 5p15.3 is associated with a cat-like cry, and a nearby region called 5p15.2 is associated with mental retardation, small head size (microcephaly), and distinctive facial features.

- Familial Adenomatous Polyposis is caused by a deletion of the APC tumor suppressor gene on the long (q) arm of chromosome 5. This chromosomal change results in thousands of colonic polyps which gives the patient a 100% risk of colon cancer if total colectomy is not done.

- Other changes in the number or structure of chromosome 5 can have a variety of effects, including delayed growth and development, distinctive facial features, birth defects, and other medical problems. Changes to chromosome 5 include an extra segment of the short (p) or long (q) arm of the chromosome in each cell (partial trisomy 5p or 5q), a missing segment of the long arm of the chromosome in each cell (partial monosomy 5q), and a circular structure called ring

chromosome 5. A ring chromosome occurs when both ends of a broken chromosome are reunited.

Chromosome 6 (Human)



Chromosome 6 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 6 spans more than 170 million base pairs (the building material of DNA) and represents between 5.5 and 6% of the total DNA in cells. It contains the Major Histocompatibility Complex, which contains over 100 genes related to the immune response, and plays a vital role in organ transplantation.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome,

the estimated number of genes varies. Chromosome 6 likely contains between 1,100 and 1,600 genes.

Genes

The following are some of the genes located on chromosome 6:

- BCKDHB: branched chain keto acid dehydrogenase E1, beta polypeptide (maple syrup urine disease)
- CNR1: cannabinoid 1 receptor
- COL11A2: collagen, type XI, alpha 2
- CYP21A2: cytochrome P450, family 21, subfamily A, polypeptide 2
- DSP: Desmoplakin gene linked to cardiomyopathy
- EYA4: eyes absent homolog 4 (Drosophila)
- HFE: hemochromatosis
- HLA-A, HLA-B, HLA-C: major histocompatibility complex (MHC), class I, A, B, and C loci.
- HLA-DQA1 and HLA-DQB1 form HLA-DQ heterodimer MHC class II, DQ: Celiac1, IDDM
- HLA-DRA, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5 forms HLA-DR, heterodimer MHC class II, DR
- HLA-DPA1 and HLA-DPB1 forms HLA-DR, MHC class II, DP
- MUT: methylmalonyl Coenzyme A mutase
- MYO6: myosin VI
- PARK2: Parkinson disease (autosomal recessive, juvenile) 2, parkin
- PKHD1: polycystic kidney and hepatic disease 1 (autosomal recessive)
- TNXB: tenascin XB
- VEGF: vascular endothelial growth factor A (angiogenic growth factor)
- IGF2R: insulin-like growth factor 2 receptor
- HLA-Cw*0602: gene variation related to psoriasis
- PLG: plasminogen (6q26)

BGHS

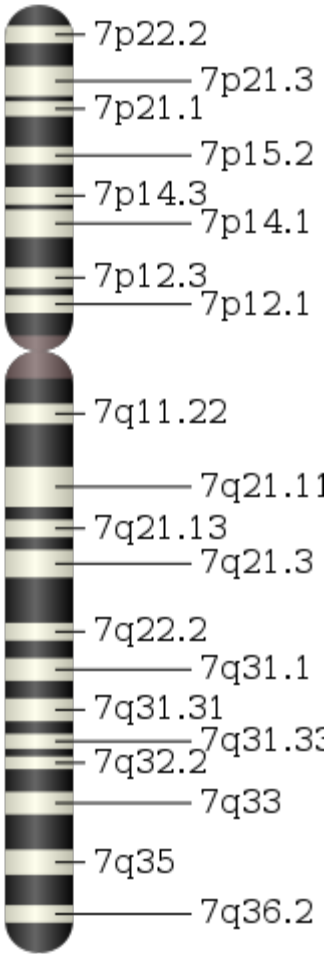
Diseases & disorders

The following diseases are some of those related to genes on chromosome 6:

- ankylosing spondylitis, HLA-B
- collagenopathy, types II and XI
- Coeliac disease, HLA-DQA1 & DQB1
- Ehlers-Danlos syndrome, classical, hypermobility, and Tenascin-X types
- Hashimoto's thyroiditis
- hemochromatosis
- Hemochromatosis type 1
- 21-hydroxylase deficiency

- maple syrup urine disease
- methylmalonic acidemia
- Autosomal nonsyndromic deafness
- otospondylomegaepiphyseal dysplasia
- Parkinson disease
- polycystic kidney disease
- porphyria
- porphyria cutanea tarda
- Rheumatoid arthritis, HLA-DR
- Stickler syndrome, COL11A2
- Systemic lupus erythematosus
- Diabetes mellitus type 1, HLA-DR, DQA1 & DQB1
- X-linked sideroblastic anemia
- Epilepsy

Chromosome 7 (Human)



Chromosome 7 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 7 spans more than 158 million base pairs (the building material of DNA) and represents between 5 and 5.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 7 is likely to contain between 1,000 and 1,400 genes. It also contains the Homeobox A gene cluster.

Diseases & disorders

The following diseases are some of those related to genes on chromosome 7:

- argininosuccinic aciduria
- cerebral cavernous malformation
- Charcot-Marie-Tooth disease
- Charcot-Marie-Tooth disease, type 2
- citrullinemia
- congenital bilateral absence of vas deferens
- cystic fibrosis
- distal spinal muscular atrophy, type V
- Ehlers-Danlos syndrome
- Ehlers-Danlos syndrome, arthrochalasia type
- Ehlers-Danlos syndrome, classical type
- hemochromatosis
- hemochromatosis, type 3
- hereditary nonpolyposis colorectal cancer
- lissencephaly
- maple syrup urine disease
- maturity onset diabetes of the young type 2
- mucopolysaccharidosis type VII or Sly syndrome
- myelodysplastic syndrome
- nonsyndromic deafness
- nonsyndromic deafness, autosomal dominant
- nonsyndromic deafness, autosomal recessive
- osteogenesis imperfecta
- osteogenesis imperfecta, type I
- osteogenesis imperfecta, type II
- osteogenesis imperfecta, type III
- osteogenesis imperfecta, type IV
- p47-phox-deficient chronic granulomatous disease
- Pendred syndrome
- Romano-Ward syndrome
- Shwachman-Diamond syndrome
- Schizophrenia
- Tritanopia or tritanomaly color blindness
- Williams syndrome

Chromosomal disorders

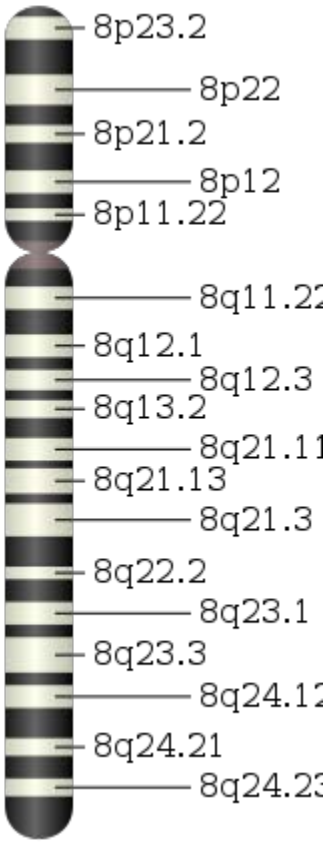
The following conditions are caused by changes in the structure or number of copies of chromosome 7:

- Williams syndrome is caused by the deletion of genetic material from a portion of the long (q) arm of chromosome 7. The deleted region, which is located at position 11.23 (written as 7q11.23), is designated as the Williams syndrome critical region. This region includes more than 20 genes, and researchers believe that the characteristic features of Williams syndrome are probably related to the loss of multiple genes in this region.

While a few of the specific genes related to Williams syndrome have been identified, the relationship between most of the genes in the deleted region and the signs and symptoms of Williams syndrome is unknown.

- Other changes in the number or structure of chromosome 7 can cause delayed growth and development, mental retardation, characteristic facial features, skeletal abnormalities, delayed speech, and other medical problems. These changes include an extra copy of part of chromosome 7 in each cell (partial trisomy 7) or a missing segment of the chromosome in each cell (partial monosomy 7). In some cases, several DNA building blocks (nucleotides) are deleted or duplicated in part of chromosome 7. A circular structure called ring chromosome 7 is also possible. A ring chromosome occurs when both ends of a broken chromosome are reunited.

Chromosome 8 (Human)



Chromosome 8 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 8 spans about 145 million base pairs (the building material of DNA) and represents between 4.5 and 5.0% of the total DNA in cells.

The chromosome has two arms, 8p and 8q. The short arm, 8p, has about 45 million base pairs, about 1.5% of the genome, and includes 484 genes and 110 pseudogenes; about 8% of its genes are involved in brain development and function, and about 16% are involved in cancer. A unique feature of 8p is a big region of about 15 megabases that appears to have a high mutation rate, and which shows an immense divergence between human and chimpanzee, suggesting that its high mutation rates have contributed to the evolution of the human brain.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 8 is likely to contain between 700 and 1,000 genes.

Genes

The following are some of the genes located on chromosome 8:

- AEG1 : Astrocyte Elevated Gene (linked to hepatocellular carcinoma and neuroblastoma)
- Arc/Arg3.1
- FGFR1: fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome)
- GDAP1: ganglioside-induced differentiation-associated protein 1
- LPL: lipoprotein lipase
- MCPH1: microcephaly, primary autosomal recessive 1
- NDRG1: N-myc downstream regulated gene 1
- NEF3: neurofilament 3 (150kDa medium)
- NEFL: neurofilament, light polypeptide 68kDa
- SNAI2: snail homolog 2 (Drosophila)
- TG: thyroglobulin
- TPA: tissue plasminogen activator
- VMAT1: vesicular monoamine transporter protein
- WRN: Werner syndrome
- GULOP pseudogene: responsible for human inability to produce our own Vitamin C

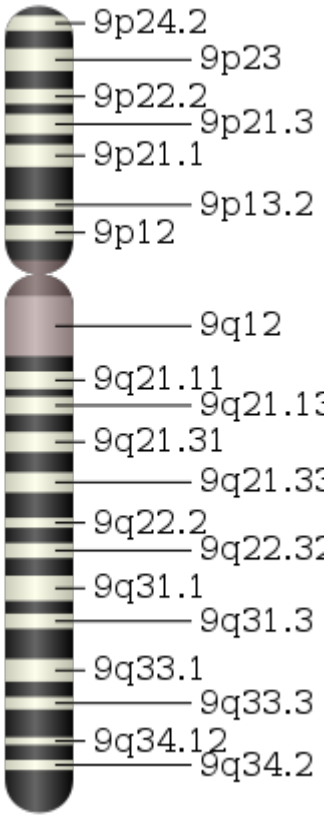
Diseases & disorders

The following diseases are some of those related to genes on chromosome 8:

- 8p23.1 duplication syndrome
- Burkitt's lymphoma
- Charcot-Marie-Tooth disease
- Charcot-Marie-Tooth disease, type 2
- Charcot-Marie-Tooth disease, type 4
- Cleft lip and palate

- Congenital hypothyroidism
- Lipoprotein lipase deficiency, familial
- Primary microcephaly
- Hereditary Multiple Exostoses
- Pfeiffer syndrome
- Rothmund-Thomson syndrome, or poikiloderma congenitale
- Schizophrenia, associated with 8p21-22 locus
- Waardenburg syndrome
- Werner syndrome
- Pingelapese blindness
- Langer-Giedion syndrome
- Roberts Syndrome

Chromosome 9 (Human)



Chromosome 9 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome, as they normally do with all chromosomes. Chromosome 9 spans about 145 million base pairs of nucleic acids (the building blocks of DNA) and represents between 4 and 4.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 9 likely contains between 800 and 1,200 genes.

Genes

The following are some of the genes located on chromosome 9:

- ABO: ABO histo-blood group glycosyltransferases
- ADAMTS13: ADAM metallopeptidase with thrombospondin type 1 motif, 13
- ALAD: aminolevulinate, delta-, dehydratase
- ALS4: amyotrophic lateral sclerosis 4
- ASS: argininosuccinate synthetase
- CCL21: chemokine (C-C motif) ligand 21, SCYA21
- CCL27: chemokine (C-C motif) ligand 27, SCYA27
- COL5A1: collagen, type V, alpha 1
- ENG: endoglin (Osler-Rendu-Weber syndrome 1)
- FXN: frataxin
- GALT: galactose-1-phosphate uridylyltransferase
- GLE1L: Nucleoporin GLE1
- GRHPR: glyoxylate redasductase/hydroxypyruvate reductase
- IKBKAP: inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein
- TGFBR1: transforming growth factor beta, receptor type I
- TMC1: transmembrane channel-like 1
- TSC1: t

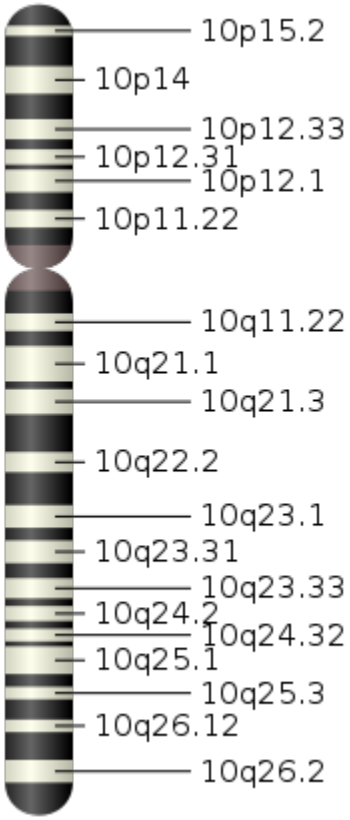
Diseases & disorders

The following diseases are some of those related to genes on chromosome 9:

- acytosiosis
- ALA-D deficiency porphyria
- citrullinemia
- Ehlers-Danlos syndrome
- Ehlers-Danlos syndrome, classical type
- familial dysautonomia
- Friedreich ataxia
- galactosemia
- Gorlin syndrome or Nevoid Basal Cell Carcinoma syndrome
- hereditary hemorrhagic telangiectasia
- lethal congenital contracture syndrome
- Nail-patella syndrome (NPS)
- nonsyndromic deafness
- nonsyndromic deafness, autosomal dominant

- nonsyndromic deafness, autosomal recessive
- porphyria
- primary hyperoxaluria
- Tangier's disease
- tetrasomy 9p
- thrombotic thrombocytopenic purpura
- trisomy 9
- tuberous sclerosis
- VLDLR-associated cerebellar hypoplasia

Chromosome 10 (Human)



Chromosome 10 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 10 spans about 135 million base pairs (the building material of DNA) and represents between 4 and 4.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 10 likely contains between 800 and 1,200 genes.

Genes

The following are some of the genes located on chromosome 10:

- ALOX5: Arachidonate 5-Lipoxygenase (processes essential fatty acids to leukotrienes, which are important agents in the inflammatory response; also facilitates development and maintenance of cancer stem cells, slow-dividing cells thought to give rise to a variety of cancers, including leukemia);
- CDH23: cadherin-like 23
- CXCL12: chemokine (C-X-C motif) ligand 12, SDF-1, *scyb12*
- EGR2: early growth response 2 (Krox-20 homolog, Drosophila)
- ERCC6: excision repair cross-complementing rodent repair deficiency, complementation group 6
- FGFR2: fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1, Crouzon syndrome, Pfeiffer syndrome, Jackson-Weiss syndrome)
- PCBD1: 6-pyruvoyl-tetrahydropterin synthase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1)
- PCDH15: protocadherin 15
- PTEN gene: phosphatase and tensin homolog (mutated in multiple advanced cancers 1)
- RET: ret proto-oncogene (multiple endocrine neoplasia and medullary thyroid carcinoma 1, Hirschsprung disease)
- UROS: uroporphyrinogen III synthase (congenital erythropoietic porphyria)

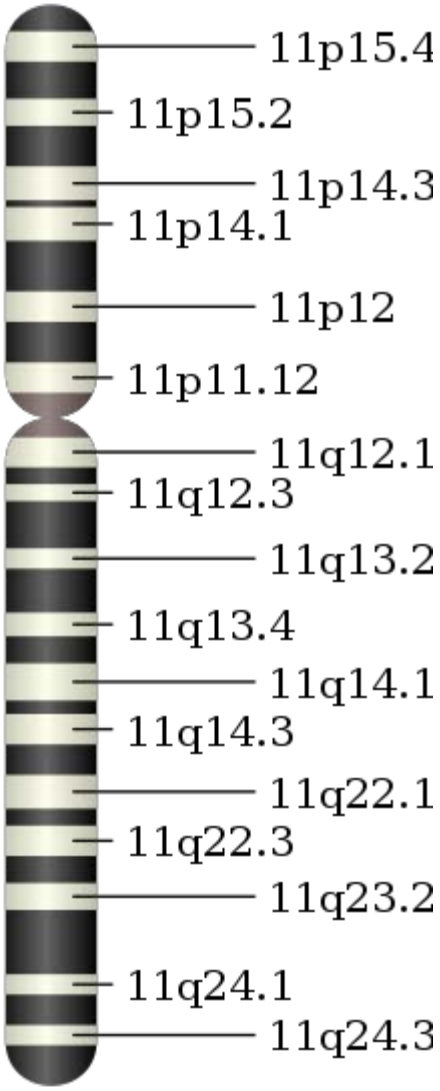
Diseases & disorders

The following diseases are some of those related to genes on chromosome 10:

- Apert syndrome
- Beare-Stevenson cutis gyrata syndrome
- Charcot-Marie-Tooth disease
- Charcot-Marie-Tooth disease, type 1
- Charcot-Marie-Tooth disease, type 4
- Cockayne syndrome
- congenital erythropoietic porphyria
- Cowden syndrome
- Crouzon syndrome
- Hirschprung disease
- Jackson-Weiss syndrome
- multiple endocrine neoplasia type 2

- nonsyndromic deafness
- nonsyndromic deafness, autosomal recessive
- Pfeiffer syndrome
- porphyria
- tetrahydrobiopterin deficiency
- Thiel-Behnke corneal dystrophy
- Usher syndrome
- Usher syndrome type I
- Wolman syndrome

Chromosome 11 (Human)



Chromosome 11 Chart

Chromosome 11 is one of the 23 pairs of chromosomes in humans. Humans normally have two copies of this chromosome. Chromosome 11 spans about 134.5 million base pairs (the building material of DNA) and represents between 4 and 4.5 percent of the total DNA in cells. It is one of the most gene- and disease-rich chromosomes in the human genome.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome,

the estimated number of genes varies. Chromosome 11 likely contains between 1,300 and 1,700 genes.

A recent study shows that 11.6 genes per megabase, including 1,524 protein-coding genes and 765 pseudogenes can be found on chromosome 11.

More than 40% of the 856 olfactory receptor genes in the human genome are located in 28 single- and multi-gene clusters along this chromosome.

Genes

The following are some of the genes located on chromosome 11:

- ACAT1: acetyl-Coenzyme A acetyltransferase 1 (acetoacetyl Coenzyme A thiolase)
- APOA4: apolipoprotein A-IV
- ATM: ataxia telangiectasia mutated (includes complementation groups A, C and D)
- CD81
- WT1
- C11orf1
- CPT1A: carnitine palmitoyltransferase 1A (liver)
- DHCR7: 7-dehydrocholesterol reductase
- HBB: hemoglobin, beta
- HMBS: hydroxymethylbilane VIIA
- PAX6
- PTS: 6-pyruvoyltetrahydropterin synthase
- SAA1: serum amyloid A1
- SBF2: SET binding factor 2
- SMPD1: sphingomyelin phosphodiesterase 1, acid lysosomal (acid sphingomyelinase)
- TECTA: tectorin alpha (nonsyndromic deafness)
- TH: tyrosine hydroxylase
- USH1C: Usher syndrome 1C (autosomal recessive, severe)

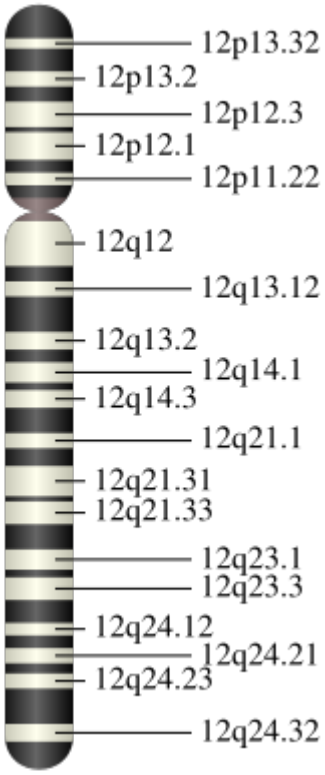
Diseases & disorders

The following diseases are some of those related to genes on chromosome 11:

- autism (neurexin 1)
- aniridia
- acute intermittent porphyria
- albinism
- ataxia-telangiectasia
- Beckwith-Wiedemann syndrome
- Best's disease

- beta-ketothiolase deficiency
- beta thalassemia
- bladder cancer
- breast cancer
- carnitine palmitoyltransferase I deficiency
- Charcot-Marie-Tooth disease
- Charcot-Marie-Tooth disease, type 4
- Denys-Drash syndrome
- familial Mediterranean fever
- Hereditary angioedema OMIM: 106100
- Jacobsen syndrome
- Jervell and Lange-Nielsen syndrome
- Meckel syndrome
- methemoglobinemia, beta-globin type
- Mixed Lineage Leukemia
- multiple endocrine neoplasia type 1
- Hereditary Multiple Exostoses
- Niemann-Pick disease
- nonsyndromic deafness
- nonsyndromic deafness, autosomal dominant
- nonsyndromic deafness, autosomal recessive
- porphyria
- Romano-Ward syndrome
- sickle cell anemia
- Smith-Lemli-Opitz syndrome
- tetrahydrobiopterin deficiency
- Usher syndrome
- Usher syndrome type I
- WAGR syndrome

Chromosome 12 (Human)



Chromosome 12 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 12 spans about 143 million base pairs (the building material of DNA) and represents between 4 and 4.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 12 likely contains between 1,000 and 1,300 genes. It also contains the Homeobox C gene cluster.

Genes

The following are some of the genes located on chromosome 12:Obed M. Ha

- ACVRL1: activin A receptor type II-like 1
- CBX5: chromobox homolog 5
- COL2A1: collagen, type II, alpha 1 (primary osteoarthritis, spondyloepiphyseal dysplasia, congenital)
- HPD: 4-hydroxyphenylpyruvate dioxygenase

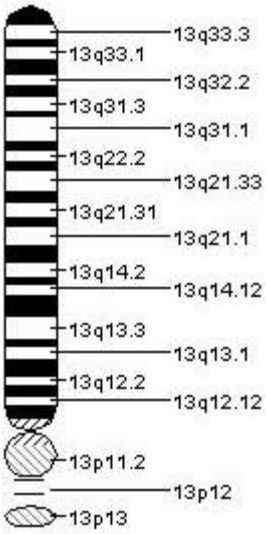
- KERA: keratocan
- LRRK2: leucine-rich repeat kinase 2
- MMAB: methylmalonic aciduria (cobalamin dewrong ficiency) cblB type
- MYO1A: myosin IA
- NANOG: NK-2 type homeodomain gene
- PAH: phenylalanine hydroxylase
- PPP1R12A: protein phosphatase 1, regulatory (inhibitor) subungfdit 12A
- PTPN11: protein tyrosine phosphatase, non-receptor type 11 (Noonan syndrome 1)
- KRAS: V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog

Diseases & disorders

The following diseases are some of those related to genes on chromosome 12:

- achondrogenesis type 2
- collagenopathy, types II and XI
- cornea plana 2
- hereditary hemorrhagic telangiectasia
- hypochondrogenesis
- ichthyosis bullosa of Siemens
- Kniest dysplasia
- maturity onset diabetes of the young type 3
- methylmalonic acidemia
- narcolepsy
- nonsyndromic deafness
- nonsyndromic deafness, autosomal dominant
- Noonan syndrome
- Parkinson disease
- Pallister-Killian syndrome (tetrasomy 12p)
- phenylketonuria
- spondyloepimetaphyseal dysplasia, Strudwick type
- spondyloepiphyseal dysplasia congenita
- spondyloperipheral dysplasia
- Stickler syndrome
- Stickler syndrome, COL2A1
- Stuttering
- Triose Phosphate Isomerase deficiency
- tyrosinemia
- Von Willebrand Disease

Chromosome 13 (Human)



Chromosome 13 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 13 spans about 114 million base pairs (the building material of DNA) and represents between 3.5 and 4 % of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 13 likely contains between 300 and 700 genes.

Genes

The following are some of the genes located on chromosome 13:

- ATP7B: ATPase, Cu++ transporting, beta polypeptide (Wilson disease)
- BRCA2: breast cancer 2, early onset
- CARKD: Carbohydrate Kinase Domain Containing Protein (Unknown Function)
- EDNRB: endothelin receptor type B
- GJB2: gap junction protein, beta 2, 26kDa (connexin 26)
- GJB6: gap junction protein, beta 6 (connexin 30)
- HTR2A: 5-HT_{2A} receptor
- PCCA: propionyl Coenzyme A carboxylase, alpha polypeptide
- RB1: retinoblastoma 1 (including osteosarcoma)
- FLT1: Fms related tyrosine kinase 1 (Vascular endothelial growth factor receptor 1)
- SLITRK1: mutation in this gene causes some (although very few) cases of Tourette syndrome and trichotillomania

- SOX21: Transcription factor SOX-21 is a protein that in humans is encoded by the SOX21; its disruption can lead to types of alopecia in mice.

Diseases and disorders

The following diseases are some of those related to genes on chromosome 13:

- Bladder cancer
- Breast cancer
- Heterochromia
- Hirschsprung's disease
- Maturity onset diabetes of the young type 4
- Nonsyndromic deafness
- Nonsyndromic deafness, autosomal dominant
- Nonsyndromic deafness, autosomal recessive
- Propionic acidemia
- Retinoblastoma
- Waardenburg syndrome
- Wilson disease
- Patau syndrome

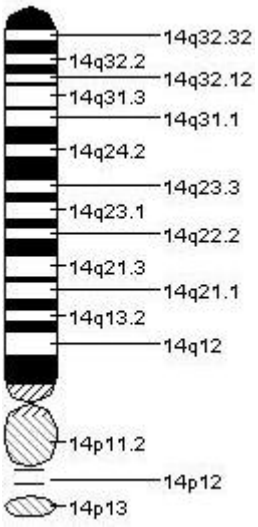
Chromosomal conditions

The following conditions are caused by changes in the structure or number of copies of chromosome 13:

- Retinoblastoma: A small percentage of retinoblastoma cases are caused by deletions in the region of chromosome 13 (13q14) containing the RB1 gene. Children with these chromosomal deletions may also have mental retardation, slow growth, and characteristic facial features (such as prominent eyebrows, a broad nasal bridge, a short nose, and ear abnormalities). Researchers have not determined which other genes are located in the deleted region, but a loss of several genes is likely responsible for these developmental problems.
- Trisomy 13: Trisomy 13 occurs when each cell in the body has three copies of chromosome 13 instead of the usual two copies. Trisomy 13 can also result from an extra copy of chromosome 13 in only some of the body's cells (mosaic trisomy 13). In a small percentage of cases, trisomy 13 is caused by a rearrangement of chromosomal material between chromosome 13 and another chromosome. As a result, a person has the two usual copies of chromosome 13, plus extra material from chromosome 13 attached to another chromosome. These cases are called translocation trisomy 13. Extra material from chromosome 13 disrupts the course of normal development, causing the characteristic signs and symptoms of trisomy 13. Researchers are not yet certain how this extra genetic material leads to the features of the disorder, which include severely abnormal cerebral functions, a small cranium, retardation, non functional eyes and heart defects.

- Other chromosomal conditions: Partial monosomy 13q is a rare chromosomal disorder that results when a piece of the long arm (q) of chromosome 13 is missing (monosomic). Infants born with partial monosomy 13q may exhibit low birth weight, malformations of the head and face (craniofacial region), skeletal abnormalities (especially of the hands and feet), and other physical abnormalities. Mental retardation is characteristic of this condition. The mortality rate during infancy is high among individuals born with this disorder. Almost all cases of partial monosomy 13q occur randomly for no apparent reason (sporadic).

Chromosome 14 (Human)



Chromosome14 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 14 spans about 109 million base pairs (the building material of DNA) and represents between 3 and 3.5% of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 14 likely contains between 700 and 1,300 genes.

Genes

The following are some of the genes located on chromosome 14:

- COCH: coagulation factor C homolog, cochlin (Limulus polyphemus)
- GALC: galactosylceramidase (Krabbe disease)
- GCH1: GTP cyclohydrolase 1 (dopa-responsive dystonia)

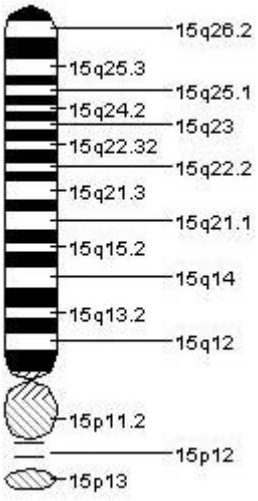
- IGH@: immunoglobulin heavy chain locus
- NPC2: Niemann-Pick disease, type C2
- PSEN1: presenilin 1 (Alzheimer disease 3)
- SERPINA1: serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1
- TSHR: thyroid stimulating hormone receptor

Diseases & disorders

The following diseases are some of those related to genes on chromosome 14:

- alpha-1 antitrypsin deficiency
- Alzheimer disease
- Alzheimer disease, type 3
- congenital hypothyroidism
- dopamine-responsive dystonia
- Krabbe disease
- Machado-Joseph disease
- multiple myeloma
- Niemann-Pick disease
- nonsyndromic deafness
- nonsyndromic deafness, autosomal dominant
- tetrahydrobiopterin deficiency
- Uniparental disomy (UPD) 14

Chromosome 15 (Human)



Human chromosome 15

Chromosome 15 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 15 spans about 106 million base pairs (the building material of DNA) and represents between 3% and 3.5% of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 15 likely contains between 700 and 900 genes.

Genes

The following are some of the genes located on chromosome 15:

- CAPN3: Calpain 3 (limb-girdle muscular dystrophy type 2A)
- CHP: Calcium binding protein P22
- FAH: fumarylacetoacetate hydrolase (fumarylacetoacetase)
- FBN1: fibrillin 1 (Marfan syndrome)
- HEXA: hexosaminidase A (alpha polypeptide)
- IVD: isovaleryl Coenzyme A dehydrogenase
- MCPH4: microcephaly, primary autosomal recessive 4
- OCA2: oculocutsexous albinism II (pink-eye dilution homolog, mouse)
- RAD51: RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)
- STRC: stereocilin
- UBE3A: ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)
- PML: promyelocytic leukemia protein (involved in t(15,17) with RARalpha, predominant cause of acute promyelocytic leukemia.
- SLC24A5: the gene responsible for at least 1/3 of the skin color differences between races, expressed in the brain and the nervous system
- EYCL3 Eye color 3, BROWN - location: 15q11-q15 (note eye colour is a polygenic trait)
- EYCL2 Eye color 2,Determines the positionning of melanocytes on the iris (note eye colour is a polygenic trait)

Chromosomal conditions

The following conditions are caused by mutations in chromosome 15. Two of the conditions (Angelman syndrome and Prader-Willi syndrome) involve a loss of gene activity in the same part of chromosome 15, the 15q11-q13 region. This discovery provided the first evidence in humans that something beyond genes could determine how the genes are expressed.

Angelman syndrome

Angelman syndrome results from a loss of gene activity in a specific part of chromosome 15, the 15q11-q13 region. This region contains a gene called UBE3A that, when mutated or absent, likely causes the characteristic features of this condition. People normally have two copies of the UBE3A gene, one from each parent. Both copies of this gene are active in many of the body's tissues. In the brain, however, only the copy inherited from a person's mother (the maternal copy) is active. If the maternal copy is lost because of a chromosomal change or a gene mutation, a person will have no working copies of the UBE3A gene in the brain.

In most cases (about 70%), people with Angelman syndrome have a deletion in the maternal copy of chromosome 15. This chromosomal change deletes the region of chromosome 15 that includes the UBE3A gene. Because the copy of the UBE3A gene inherited from a person's father (the paternal copy) is normally inactive in the brain, a deletion in the maternal chromosome 15 results in no active copies of the UBE3A gene in the brain.

In 3% to 7% of cases, Angelman syndrome occurs when a person has two copies of the paternal chromosome 15 instead of one copy from each parent. This phenomenon is called paternal uniparental disomy (UPD). People with paternal UPD for chromosome 15 have two copies of the UBE3A gene, but they are both inherited from the father and are therefore inactive in the brain.

About 10% of Angelman syndrome cases are caused by a mutation in the UBE3A gene, and another 3% result from a defect in the DNA region that controls the activation of the UBE3A gene and other genes on the maternal copy of chromosome 15. In a small percentage of cases, Angelman syndrome may be caused by a chromosomal rearrangement called a translocation or by a mutation in a gene other than UBE3A. These genetic changes can abnormally inactivate the UBE3A gene.

Prader-Willi syndrome

Prader-Willi syndrome is caused by the loss of active genes in a specific part of chromosome 15, the 15q11-q13 region. People normally have two copies of this chromosome in each cell, one copy from each parent. Prader-Willi syndrome occurs when the paternal copy is partly or entirely missing. Researchers are working to identify genes on chromosome 15 that are responsible for the characteristic features of Prader-Willi syndrome.

In about 70% of cases, Prader-Willi syndrome occurs when the 15q11-q13 region of the paternal chromosome 15 is deleted. The genes in this region are normally active on the paternal copy of the chromosome and are inactive on the maternal copy. Therefore, a person with a deletion in the paternal chromosome 15 will have no active genes in this region.

In about 25% of cases, a person with Prader-Willi syndrome has two maternal copies of chromosome 15 in each cell instead of one copy from each parent. This phenomenon is called maternal uniparental disomy. Because some genes are normally active only on the paternal copy of this chromosome, a person with two maternal copies of chromosome 15 will have no active copies of these genes.

In a small percentage of cases, Prader-Willi syndrome is caused by a chromosomal rearrangement called a translocation. Rarely, the condition is caused by an abnormality in the DNA region that controls the activity of genes on the paternal chromosome 15.Prader Willi Syndrome is hereditary.

Isodicentric chromosome 15

A specific chromosomal change called an isodicentric chromosome 15 (previously called an inverted duplication 15) can affect growth and development. The patient possesses an "extra" or "marker" chromosome. This small extra chromosome is made up of genetic material from chromosome 15 that has been abnormally duplicated (copied) and attached end-to-end. In some cases, the extra chromosome is very small and has no effect on a person's health. A larger isodicentric chromosome 15 can result in weak muscle tone (hypotonia), mental retardation, seizures, and behavioral problems. Signs and symptoms of autism (a developmental disorder that affects communication and social interaction) have also been associated with the presence of an isodicentric chromosome 15.

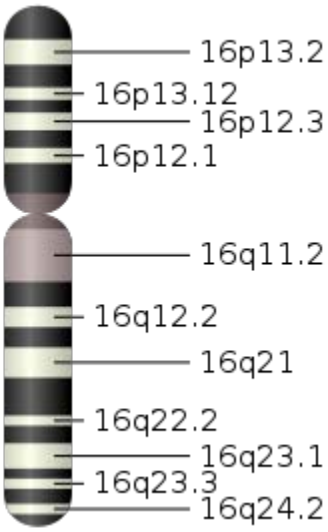
Other chromosomal conditions

Other changes in the number or structure of chromosome 15 can cause mental retardation, delayed growth and development, hypotonia, and characteristic facial features. These changes include an extra copy of part of chromosome 15 in each cell (partial trisomy 15) or a missing segment of the chromosome in each cell (partial monosomy 15). In some cases, several of the chromosome's DNA building blocks (nucleotides) are deleted or duplicated.

The following diseases are some of those related to genes on chromosome 15:

- Bloom syndrome
- Breast cancer
- Isovaleric acidemia
- Marfan syndrome
- Nonsyndromic deafness
- Tay-Sachs disease
- Tyrosinemia

Chromosome 16 (Human)



Chromosome 16 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 16 spans about 90 million base pairs (the building material of DNA) and represents just under 3 % of the total DNA in cells.

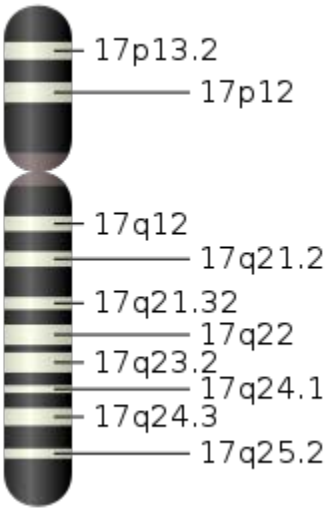
Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 16 probably contains between 850 and 1,200 genes.

In February 2010, a new cause of obesity due to a microdeletion on chromosome 16 was announced. It may explain about 1% of obesity cases. This research was carried out by Professor Froguel, a CNRS researcher, in Lille, and others at Imperial College in London and Vaudois University and was published in *Nature* on February 4, 2010. This defect was identified using DNA microarrays and it leads to the suppression of about thirty genes in one region of chromosome 16. Research showed that this microdeletion is relatively common in obese people but lacking in most non-obese people.

Diseases and Disorders

- Trisomy 16
- Familial Mediterranean fever (FMF)
- Crohn's disease
- Thalassemia

Chromosome 17 (Human)



Chromosome 17 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 17 spans more than 81 million base pairs (the building material of DNA) and represents between 2.5 and 3 % of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 17 likely contains between 1,200 and 1,500 genes. It also contains the Homeobox B gene cluster.

Genes

The following are some of the genes located on chromosome 17:

- ACADVL: acyl-coenzyme A dehydrogenase, very long chain
- ACTG1: actin, gamma 1
- ASPA: aspartoacylase (Canavan disease)
- BRCA1: breast cancer 1, early onset
- CBX1: chromobox homolog 1
- COL1A1: collagen, type I, alpha 1
- CTNS: cystinosin, the lysosomal cystine transporter
- ERBB2 loca leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
- FLCN: folliculin
- GALK1: galactokinase 1
- GFAP: glial fibrillary acidic protein

- KCNJ2: potassium inwardly-rectifying channel, subfamily J, member 2
- MYO15A: myosin XVA
- NF1: neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, Watson disease)
- PMP22: peripheral myelin protein 22
- SHBG: Sex hormone binding globulin
- SLC6A4: Serotonin transporter
- TMC6 and TMC8: Transmembrane channel-like 6 and 8 (epidermodysplasia verruciformis)
- TP53: tumor suppressor protein p53 (Li-Fraumeni syndrome), tumor suppressor gene
- USH1G: Usher syndrome 1G (autosomal recessive)
- RAI1: retinoic acid induced 1
- RAR-alpha: Retinoic acid receptor Alpha (involved in t(15,17) with PML)
- GRB7: Growth factor Receptor-Bound protein 7
- Several CC chemokines: CCL1, CCL2, CCL3, CCL4, CCL5, CCL7, CCL8, CCL11, CCL13, CCL14, CCL15, CCL16, CCL18, and CCL23

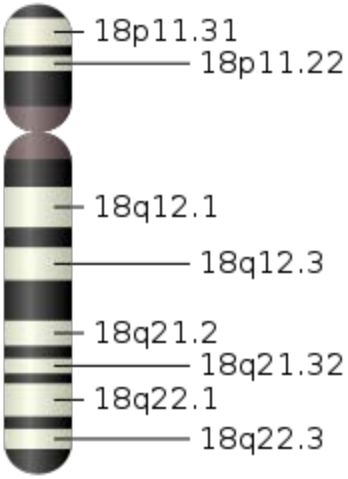
Diseases & disorders

The following diseases are related to genes on chromosome 17:

- Alexander disease
- Andersen-Tawil syndrome
- Birt-Hogg-Dubé syndrome
- Bladder cancer
- Breast cancer
- Camptomelic dysplasia
- Canavan disease
- Charcot-Marie-Tooth disease
- Charcot-Marie-Tooth disease, type 1
- Corticobasal degeneration
- Cystinosis
- Ehlers-Danlos syndrome
- Ehlers-Danlos syndrome, arthrochalasia type
- Ehlers-Danlos syndrome, classical type
- Epidermodysplasia verruciformis
- Galactosemia
- Hereditary neuropathy with liability to pressure palsies
- Li-Fraumeni syndrome
- Maturity onset diabetes of the young type 5
- Miller-Dieker syndrome
- Neurofibromatosis type I
- Nonsyndromic deafness
- Nonsyndromic deafness, autosomal dominant
- Nonsyndromic deafness, autosomal recessive

- Osteogenesis imperfecta
- Osteogenesis Imperfecta, Type I
- Osteogenesis Imperfecta, Type II
- Osteogenesis Imperfecta, Type III
- Osteogenesis Imperfecta, Type IV
- Smith-Magenis syndrome
- Usher syndrome
- Usher syndrome type I
- Very long-chain acyl-coenzyme A dehydrogenase deficiency

Chromosome 18 (Human)



Chromosome 18 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 18 spans about 76 million base pairs (the building material of DNA) and represents about 2.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 18 likely contains between 300 and 400 genes.

Genes

The following are some of the genes located on chromosome 18:

- FECH: ferrochelatase (protoporphyria)

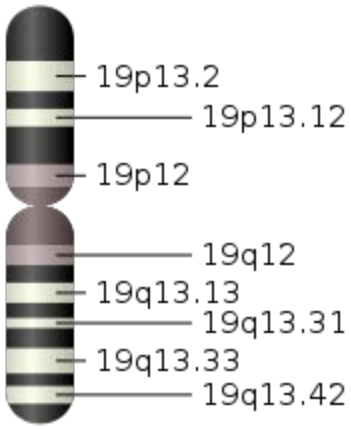
- NPC1: Niemann-Pick disease, type C1
- SMAD4: SMAD, mothers against DPP homolog 4 (Drosophila)
- KC6: Keratoconus gene 6, discovery reported in 2005;

Diseases & disorders

The following diseases are some of those related to genes on chromosome 18:

- erythropoietic protoporphyria
- hereditary hemorrhagic telangiectasia
- Niemann-Pick disease Type C
- porphyria
- Selective Mutism
- Edwards syndrome (Trisomy 18)
- Tetrasomy 18p
- Monosomy 18p
- Pitt Hopkins Syndrome 18q21

Chromosome 19 (Human)



Chromosome 19 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 19 spans more than 63 million base pairs (the building material of DNA) and represents between 2 and 2.5 percent of the total DNA in cells.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome,

the estimated number of genes varies. Chromosome 19 likely contains between 1,300 and 1,700 genes.

Genes

The following are some of the genes located on chromosome 19:

- A1BG: Plasma glycoprotein, unknown function. Gene map locus 19q13.43
- APOE: Apolipoprotein E, gene associated with Alzheimer's disease
- BCKDHA: Branched chain keto acid dehydrogenase E1, alpha polypeptide (maple syrup urine disease)
- CACNA1A: Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit, mutations of which are associated with Familial hemiplegic migraine Type I
- DMPK: Dystrophia myotonica-protein kinase
- EYCL1: EYE COLOR 1; EYE COLOR, GREEN/BLUE; GEY. Gene map locus 19p13.1-q13.11 OMIM: 227240
- GCDH: Glutaryl-Coenzyme A dehydrogenase
- HAMP: Hepcidin antimicrobial peptide
- HCL1: HAIR COLOR 1; BROWN HAIR COLOR; BRHC. Gene map locus 19p13.1-q13.11 OMIM: 113750
- MCPH2: microcephaly, primary autosomal recessive 2
- NOTCH3: Notch homolog 3 (Drosophila)
- NRTN: Neurturin, associated with Hirschsprung's disease
- NWD1: NACHT and WD repeat domain containing 1.
- PEX11G: peroxisomal biogenesis factor 11 gamma
- PRX: Periaxin
- SLC5A5: Solute carrier family 5 (sodium iodide symporter), member 5
- STK11: Serine/threonine kinase 11 (Peutz-Jeghers syndrome)

Diseases & disorders

The following diseases are some of those related to genes on chromosome 19:

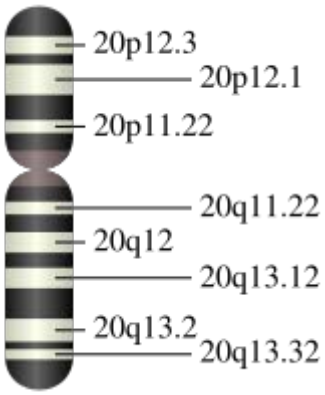
- Alzheimer's disease
- CADASIL
- Centronuclear myopathy autosomal dominant form
- Charcot-Marie-Tooth disease
- Congenital hypothyroidism
- Familial hemiplegic migraine
- Glutaric acidemia type 1
- Hemochromatosis
- Leber's Congenital Amaurosis
- Maple syrup urine disease
- Myotonic dystrophy
- Myotubular myopathy autosomal dominant form
- Marfan Syndrome

- Oligodendroglioma
- Peutz-Jeghers syndrome
- Spinocerebellar ataxia type-6
- X-linked agammaglobulinemia or Bruton's Disease

Proteins

The Human Proteome Project (HPP) has been initiated by Human Proteome Organization (HUPO) aims to sequence the entire human proteome based on the advances in mass spectrometry and will build a global protein capture knowledge base with open sharing of proteome datasets. The proteome sequencing has been initiated with a chromosome centric and gene centric approach. Sequencing of chromosome 19 proteome will be coordinated by Prof. György Marko-Varga, Clinical Protein Science & Imaging Group, Lund University, Sweden.

Chromosome 20 (Human)



Chromosome 20 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. Chromosome 20 spans around 63 million base pairs (the building material of DNA) and represents between 2 and 2.5 percent of the total DNA in cells. Chromosome 20 was fully sequenced in 2001 and contained 59,187,298 base pairs representing 99.4% of the euchromatic DNA.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 20 contains over 900 genes. New discoveries have recently linked this chromosome to the increasing susceptibility to male-pattern baldness.

Genes

The following are some of the genes located on chromosome 20:

- EDN3: endothelin 3
- JAG1: jagged 1 (Alagille syndrome)
- PANK2: pantothenate kinase 2 (Hallervorden-Spatz syndrome)
- PRNP: prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)
- tTG:tissue transglutaminase (Celiac disease)
- AHCY: S-adenosylhomocysteine hydrolase
- ARFGEF2: ADP-ribosylation factor guanine nucleotide-exchange factor 2 (brefeldin A-inhibited)
- GSS: glutathione synthetase
- SALL4: sal-like 4 (Drosophila)
- VAPB: VAMP (vesicle-associated membrane protein)-associated protein B and C
- BMP2: Bone Morphogenetic Protein 2 (osteoblast differentiation)
- GNAS1: Gs alpha subunit (membrane G-protein)

Diseases & disorders

The following diseases are some of those related to genes on chromosome 20:

- Arterial tortuosity syndrome
- Adenosine deaminase deficiency
- Alagille syndrome
- Celiac disease
- Galactosialidosis - CTSA
- Maturity onset diabetes of the young type 1
- Pantothenate kinase-associated neurodegeneration
- Transmissible spongiform encephalopathy (prion diseases)
- Waardenburg syndrome

Chromosome 21 (Human)

Chromosome 21 is one of the 23 pairs of chromosomes in humans. People normally have two copies of this chromosome. The trisomy of the 21st chromosome causes Down's Syndrome. Chromosome 21 is the smallest human chromosome, spanning 47 million nucleotides (the building material of DNA) and representing about 1.5 percent of the total DNA in cells.

In 2000, researchers working on the Human Genome Project announced that they had determined the sequence of base pairs that make up this chromosome. Chromosome 21 was the second human chromosome to be fully sequenced.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 21 likely contains between 300 and 400 genes.

Genes

The following are some of the genes located on chromosome 21:

- APP: amyloid beta (A4) precursor protein (peptidase nexin-II, Alzheimer disease)
- CBS: cystathionine-beta-synthase
- CLDN14: claudin 14
- HLCS: holocarboxylase synthetase (biotin-(proprionyl-Coenzyme A-carboxylase (ATP-hydrolysing)) ligase)
- KCNE1: potassium voltage-gated channel, Isk-related family, member 1
- KCNE2: potassium voltage-gated channel, Isk-related family, member 2
- LAD: leukocyte adhesion deficiency (symbols are ITGB2, CD18, LCAMB)
- SOD1: superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))
- TMPRSS3: transmembrane protease, serine 3
- PCNT: centrosomal pericentrin
- DSCR1: Down Syndrome critical region 1
- DYRK1A: dual specificity tyrosine-(Y)-phosphorylation regulated kinase 1A
- RRP1B: ribosomal RNA processing 1 homolog B
- s100B: calcium binding protein

Diseases & disorders

The following diseases are some of those related to genes on chromosome 21:

- Alzheimer's disease
  - Alzheimer's disease type 1
- Amyotrophic lateral sclerosis
  - Amyotrophic lateral sclerosis type 1
- Down syndrome
- Erondu-Cymet Syndrome
- Holocarboxylase synthetase deficiency
- Homocystinuria
- Jervell and Lange-Nielsen syndrome
- Leukocyte adhesion deficiency
- Majewski osteodysplastic primordial dwarfism type II (MOPD II, or MOPD2)
- Nonsyndromic deafness
  - Nonsyndromic deafness, autosomal recessive

- Romano-Ward syndrome

Chromosomal conditions

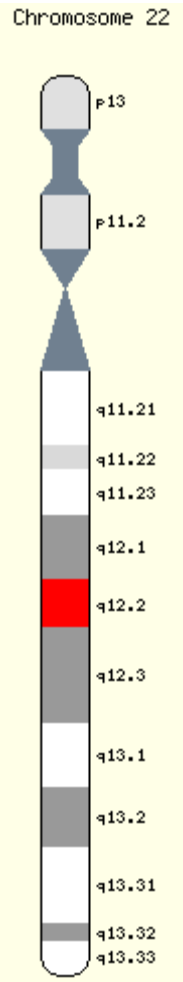
The following conditions are caused by changes in the structure or number of copies of chromosome 21:

- Cancers: Rearrangements (translocations) of genetic material between chromosome 21 and other chromosomes have been associated with several types of cancer. For example, acute lymphoblastic leukemia (a type of blood cancer most often diagnosed in childhood) has been associated with a translocation between chromosomes 12 and 21. Another form of leukemia, acute myeloid leukemia, has been associated with a translocation between chromosomes 8 and 21.

In a small percentage of cases, Down syndrome is caused by a rearrangement of chromosomal material between chromosome 21 and another chromosome. As a result, a person has the usual two copies of chromosome 21, plus extra material from chromosome 21 attached to another chromosome. These cases are called translocation Down syndrome. Researchers believe that extra copies of genes on chromosome 21 disrupt the course of normal development, causing the characteristic features of Down syndrome and the increased risk of medical problems associated with this disorder.

- Other changes in the number or structure of chromosome 21 can have a variety of effects, including mental retardation, delayed development, and characteristic facial features. In some cases, the signs and symptoms are similar to those of Down syndrome. Changes to chromosome 21 include a missing segment of the chromosome in each cell (partial monosomy 21) and a circular structure called ring chromosome 21. A ring chromosome occurs when both ends of a broken chromosome are reunited.
- Duplication in Amyloid precursor protein (APP) locus (duplicated segment varies in length but includes APP) on Chromosome 21 was found to cause early onset familial Alzheimer's disease in a French family set (Rovelet-Lecrux et al.) and a Dutch family set (Sleegers et al.). Compared to Alzheimer's caused by missense mutations in APP, the frequency of the Alzheimer's caused by APP duplications is significant. All patients that have an extra copy of APP gene due to the locus duplication show Alzheimer's with severe Cerebral amyloid angiopathy.

Chromosome 22 (Human)



Chromosome 22 is one of the 23 pairs of chromosomes in humans. Humans normally have two copies of Chromosome 22. Chromosome 22 is the second smallest human chromosome, spanning about 49 million base pairs (the building material of DNA) and representing between 1.5 and 2 % of the total DNA in cells.

In 1999, researchers working on the Human Genome Project announced they had determined the sequence of base pairs that make up this chromosome. Chromosome 22 was the first human chromosome to be fully sequenced.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. Chromosome 22 contains about 693 genes.

Chromosome 22 was originally identified as the smallest chromosome, but after extensive research, researchers concluded that it was indeed chromosome 21. The chromosomes weren't renamed because of the popularity of chromosome 21 (being known as the chromosome that can lead to Down's Syndrome). For this reason, researchers did not rearrange the numbers on these chromosomes.

Genes

The following are some of the genes located on chromosome 22:

Locus	Gene	Description	Condition
22q11.1-q11.2	IGL@	immunoglobulin lambda locus - contains genes for the light chains of antibodies	
22q11.21	TBX1	T-box 1	
22q11	RTN4R	Reticulon 4 receptor	Schizophrenia
22q11.21-q11.23	COMT	catechol-O-methyltransferase gene	
22q12.1-q13.1	NEFH	neurofilament, heavy polypeptide 200kDa	
22q12.1	CHEK2	CHK2 checkpoint homolog (S. pombe)	
22q12.2	NF2	neurofibromin 2	bilateral acoustic neuroma
22q13	SOX10	SRY (sex determining region Y)-box 10	
22q13.2	EP300	E1A binding protein p300	
22q13.3	WNT7B	Wingless-type MMTV integration site family, member 7B	
22q13.3	SHANK3	SH3 and multiple ankyrin repeat domains 3	22q13 deletion syndrome

Diseases & disorders

The following diseases are some of those related to genes on chromosome 22:

- Amyotrophic lateral sclerosis
- Breast cancer
- Desmoplastic small round cell tumor
- 22q11.2 deletion syndrome
- 22q13 deletion syndrome or Phelan-McDermid syndrome
- Li-Fraumeni syndrome
- Neurofibromatosis type 2
- Rubinstein-Taybi syndrome
- Waardenburg syndrome
- Cat eye syndrome

- Methemoglobinemia
- Schizophrenia

Chromosomal conditions

The following conditions are caused by changes in the structure or number of copies of chromosome 22:

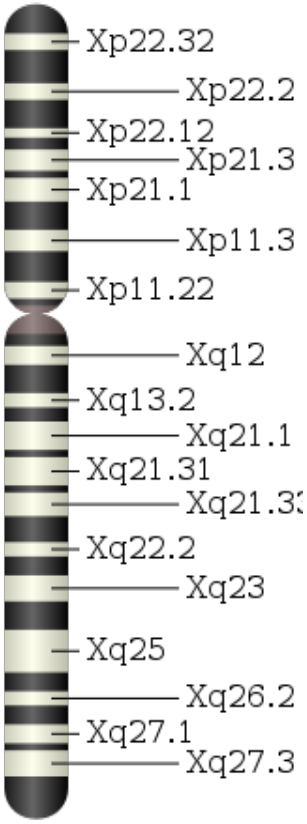
- 22q11.2 deletion syndrome: Most people with 22q11.2 deletion syndrome are missing about 3 million base pairs on one copy of chromosome 22 in each cell. The deletion occurs near the middle of the chromosome at a location designated as q11.2. This region contains about 30 genes, but many of these genes have not been well characterized. A small percentage of affected individuals have shorter deletions in the same region.
  The loss of one particular gene, TBX1, is thought to be responsible for many of the characteristic features of 22q11.2 deletion syndrome, such as heart defects, an opening in the roof of the mouth (a cleft palate), distinctive facial features, and low calcium levels. A loss of this gene does not appear to cause learning disabilities, however. Other genes in the deleted region are also likely to contribute to the signs and symptoms of 22q11.2 deletion syndrome.
- 22q13 deletion syndrome (Phelan-McDermid syndrome): The deletion of the distal tip of the chromosome 22 is related to moderate to severe developmental delay and mental retardation. This region includes the Shank3 gene, thought to be responsible for the neurological deficits of the syndrome (Wilson et al., 2003). Almost all children affected by the 22q13 deletion have absent or severely delayed speech; minor facial dysmorphism; thin, flaky toenails; large, fleshy hands; large feet; prominent, poorly formed ears and other characteristics which are not visually apparent: hypotonia (97%); normal to accelerated growth (95%); increased tolerance to pain (86%); seizures (unknown percentage) .
- Other chromosomal conditions: Other changes in the number or structure of chromosome 22 can have a variety of effects, including mental retardation, delayed development, physical abnormalities, and other medical problems. These changes include an extra piece of chromosome 22 in each cell (partial trisomy), a missing segment of the chromosome in each cell (partial monosomy), and a circular structure called ring chromosome 22 that is caused by the breakage and reattachment of both ends of the chromosome.
- Cat-eye syndrome is a rare disorder most often caused by a chromosomal change called an inverted duplicated 22. A small extra chromosome is made up of genetic material from chromosome 22 that has been abnormally duplicated (copied). The extra genetic material causes the characteristic signs and symptoms of cat-eye syndrome, including an eye abnormality called ocular iris coloboma (a gap or split in the colored part of the eye), small skin tags or pits in front of the ear, heart defects, kidney problems, and, in some cases, delayed development.
- A rearrangement (translocation) of genetic material between chromosomes 9 and 22 is associated with several types of blood cancer (leukemia). This chromosomal abnormality, which is commonly called the Philadelphia chromosome, is found

only in cancer cells. The Philadelphia chromosome has been identified in most cases of a slowly progressing form of blood cancer called chronic myeloid leukemia, or CML. It also has been found in some cases of more rapidly progressing blood cancers (acute leukemias). The presence of the Philadelphia chromosome can help predict how the cancer will progress and provides a target for molecular therapies.

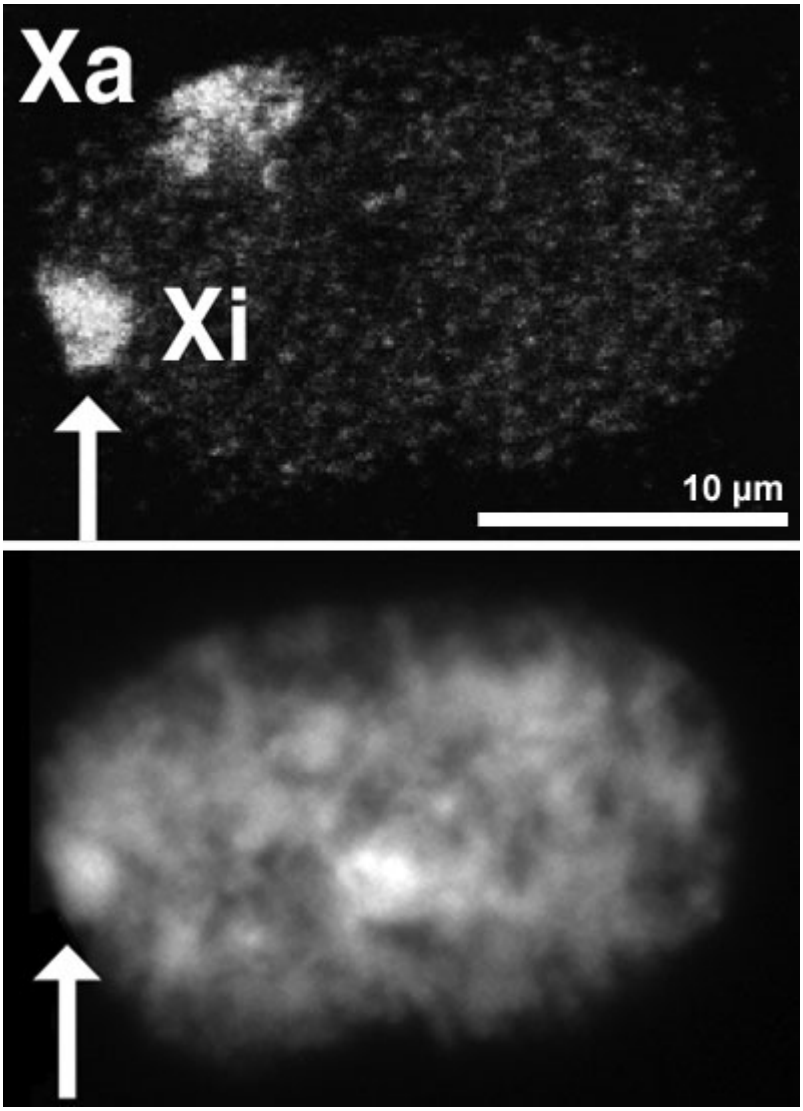
Chapter- 12

X Chromosome & Y Chromosome

X Chromosome



Scheme of the X chromatid



Nucleus of a female amniotic fluid cell. Top: Both X-chromosome territories are detected by FISH. Shown is a single optical section made with a confocal microscope. Bottom: Same nucleus stained with DAPI and recorded with a CCD camera. The Barr body is indicated by the arrow, it identifies the inactive X (Xi).

The **X chromosome** is one of the two sex-determining chromosomes in many animal species, including mammals (the other is the Y chromosome). It is a part of the XY sex-determination system and X0 sex-determination system. The **X** chromosome was named for its unique properties by early researchers, which resulted in the naming of its counterpart **Y** chromosome, for the next letter in the alphabet, after it was discovered later.

In humans

Function

The sex chromosomes X X are one of the 23 homologous pairs of chromosomes in a female. The X chromosome spans more than 153 million base pairs (the building material of DNA) and represents about 5% of the total DNA in women's cells, 2.5% in men's.

Each person normally has one pair of sex chromosomes in each cell. Females have two X chromosomes, whereas males have one X and one Y chromosome. Both males and females retain one of their mother's X chromosomes, and females retain their second X chromosome from their father. Since the father retains his X chromosome from his mother, a human female has one X chromosome from her paternal grandmother (father's side), and one X chromosome from her mother.

Identifying genes on each chromosome is an active area of genetic research. Because researchers use different approaches to predict the number of genes on each chromosome, the estimated number of genes varies. The X chromosome contains about 2000 genes compared to the Y chromosome containing 78 genes, out of the estimated 20,000 to 25,000 total genes in the human genome. Genetic disorders that are due to mutations in genes on the X chromosome are described as **X linked**.

The X chromosome carries a couple of thousand genes but few, if any, of these have anything to do directly with sex determination. Early in embryonic development in females, one of the two X chromosomes is randomly and permanently inactivated in nearly all somatic cells (cells other than egg and sperm cells). This phenomenon is called X-inactivation or Lyonization, and creates a Barr body. X-inactivation ensures that females, like males, have one functional copy of the X chromosome in each body cell. It was previously assumed that only one copy is actively used. However, recent research suggests that the Barr body may be more biologically active than was previously supposed.

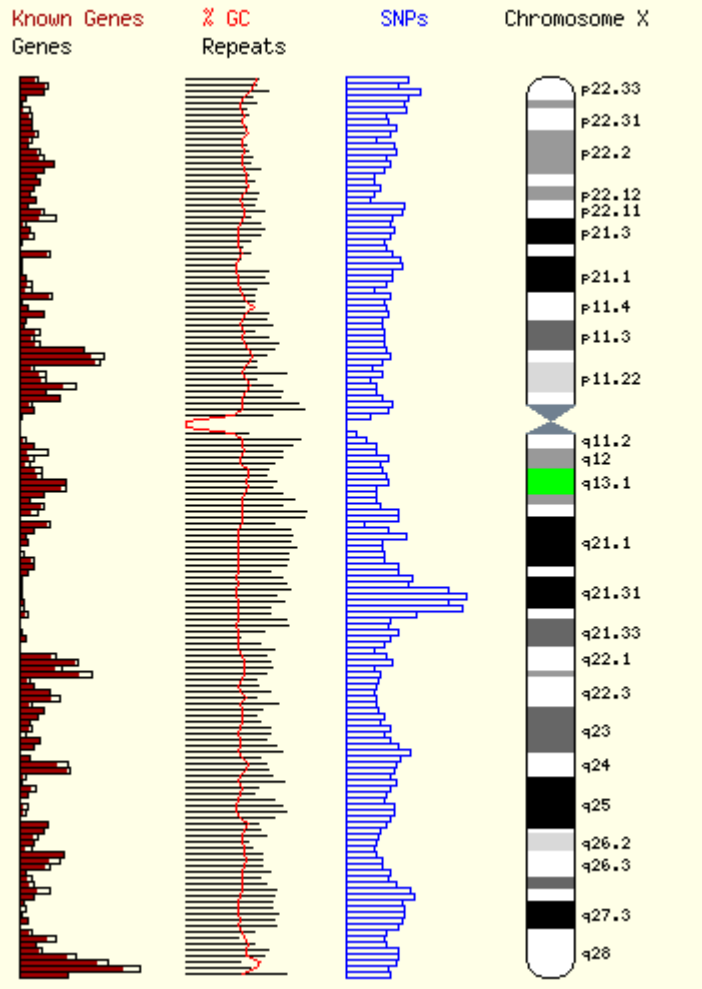
Structure

It is theorized by Ross et al. 2005 and Ohno 1967 that the X-chromosome is at least partially derived from the autosomal (non-sex-related) genome of other mammals evidenced from interspecies genomic sequence alignments.

The X-chromosome is notably larger and has a more active euchromatin region than its Y-chromosome counterpart. Further comparison of the X and Y reveal regions of homology between the two. However, the corresponding region in the Y appears far shorter and lacks regions that are conserved in the X throughout primate species, implying a genetic degeneration for Y in that region. Because males have only one X-chromosome, they are more likely to have an X chromosome-related disease.

It is estimated that about 10% of the genes encoded by the X-chromosome are associated with a family of "CT" genes, so named because they encode for markers found in both tumor cells (in **C**ancer patients) as well as in the human testis (in healthy patients).

Role in diseases



Numerical abnormalities

Klinefelter's syndrome:

- Klinefelter's syndrome is caused by the presence of one or more extra copies of the X chromosome in a male's cells. Extra genetic material from the X chromosome interferes with male sexual development, preventing the testicles from functioning normally and reducing the levels of testosterone.
- Males with Klinefelter's syndrome typically have one extra copy of the X chromosome in each cell, for a total of two X chromosomes and one Y chromosome (47,XXY). It is less common for affected males to have two or three extra X chromosomes (48,XXXY or 49,XXXXY) or extra copies of both the X and Y chromosomes (48,XXYY) in each cell. The extra genetic material may lead

to tall stature, learning and reading disabilities, and other medical problems. The average IQ in Klinefelter syndrome is in the normal range. When additional X and/or Y chromosomes are present in 48,XXXY, 48,XXYY, or 49,XXXXY, developmental delays and cognitive difficulties can be more severe and mild intellectual disability may be present.
- Klinefelter's syndrome can also result from an extra X chromosome in only some of the body's cells. These cases are called mosaic 46,XY/47,XXY.

Triple X syndrome (also called 47,XXX or trisomy X):

- This syndrome results from an extra copy of the X chromosome in each of a female's cells. Females with trisomy X have three X chromosomes, for a total of 47 chromosomes per cell. The average IQ of females with this syndrome is 90, while the average IQ of their normal siblings is 100 . Their stature on average is taller than normal females. They are fertile and their children do not inherit the condition.

- Females with more than one extra copy of the X chromosome (48, XXXX syndrome or 49, XXXXX syndrome) have been identified, but these conditions are rare.

Turner syndrome:

- This results when each of a female's cells has one normal X chromosome and the other sex chromosome is missing or altered. The missing genetic material affects development and causes the features of the condition, including short stature and infertility.
- About half of individuals with Turner syndrome have monosomy X (45,X), which means each cell in a woman's body has only one copy of the X chromosome instead of the usual two copies. Turner syndrome can also occur if one of the sex chromosomes is partially missing or rearranged rather than completely missing. Some women with Turner syndrome have a chromosomal change in only some of their cells. These cases are called Turner syndrome mosaics (45,X/46,XX).

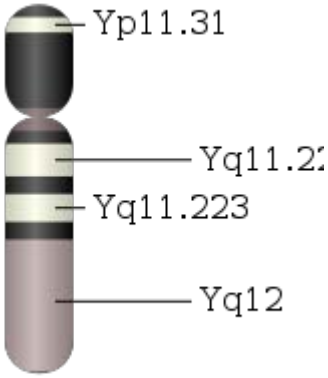
Other disorders

XX male syndrome is a rare disorder, where the SRY region of the Y chromosome has recombined to be located on one of the X chromosomes. As a result, the XX combination after fertilization has the same effect as a XY combination, resulting in a male. However, the other genes of the X chromosome cause feminization as well.

X-linked endothelial corneal dystrophy is an extremely rare disease of cornea associated with Xq25 region. Lisch epithelial corneal dystrophy is associated with Xp22.3.

Megalocornea 1 is associated with Xq21.3-q22

Y Chromosome



Human Y-chromatid

The **Y chromosome** is one of the two sex-determining chromosomes in most mammals, including humans. In mammals, it contains the gene SRY, which triggers testis development if present. The human Y chromosome is composed of about 60 million base pairs. DNA in the Y chromosome is passed from father to son, and Y-DNA analysis may thus be used in genealogy research.

Overview

Most mammals have one pair of sex chromosomes in each cell. Males have one Y chromosome and one X chromosome, while females have two X chromosomes. In mammals, the Y chromosome contains a gene, SRY, which triggers embryonic development as a male. The Y chromosomes of humans and other mammals also contain other genes needed for normal sperm production.

There are exceptions, however. For example, the platypus relies on an XY sex-determination system based on five pairs of chromosomes . Platypus sex chromosomes in fact appear to bear a much stronger homology (similarity) with the avian Z chromosome, and the SRY gene so central to sex-determination in most other mammals is apparently not involved in platypus sex-determination. Among humans, some men have two Xs and a Y, or one X and two Ys, and some women have three Xs or a single X instead of a double X. There are other exceptions in which SRY is damaged (leading to an XY female), or copied to the X (leading to an XX male).

Origins and evolution

Before Y-chromosome

Many ectothermic vertebrates have no sex chromosomes. If they have different sexes, sex is determined environmentally rather than genetically. For some of them, especially reptiles, sex depends on the incubation temperature; others are hermaphroditic (meaning they contain both male and female gametes in the same individual).

Origin

The X and Y chromosomes are thought to have evolved from a pair of identical chromosomes, termed autosomes, when an ancestral mammal developed an allelic variation, a so-called 'sex locus' - simply possessing this allele caused the organism to be male. The chromosome with this allele became the Y chromosome, while the other member of the pair became the X chromosome. Over time, genes which were beneficial for males and harmful to (or had no effect on) females either developed on the Y chromosome, or were acquired through the process of translocation..

Until recently, the X and Y chromosomes were thought to have diverged around 300 million years ago. However recent research, particularly that stemming from the sequencing of the platypus genome, has suggested that the XY sex-determination system wouldn't have been present more than 166 million years ago, at the split of the monotremes from other mammals. This reestimation of the age of the therian XY system is based on the finding that sequences that are on the X chromosomes of marsupials and eutherian mammals are present on the autosomes of platypus and birds. The older estimate was based on erroneous reports that the platypus X chromosomes contained these sequences .

Recombination inhibition

Recombination between the X and Y chromosomes proved harmful - it resulted in males without necessary genes formerly found on the Y chromosome, and females with unnecessary or even harmful genes previously only found on the Y chromosome. As a result, genes beneficial to males accumulated near the sex-determining genes, and recombination in this region was suppressed in order to preserve this male specific region. Over time, the Y chromosome changed in such a way as to inhibit the areas around the sex determining genes from recombining at all with the X chromosome. As a result of this process 95% of the human Y chromosome is unable to recombine.

Shrinking

The human Y chromosome has lost 1,393 of its 1,438 original genes over the course of its existence. With a rate of genetic loss of 4.6 genes per million years, the Y chromosome may potentially lose complete function within the next 10 million years. Comparative genomic analysis, however, reveals that many mammalian species are experiencing a

similar loss of function in their heterozygous sex chromosome. Degeneration may simply be the fate of all nonrecombining sex chromosomes due to three common evolutionary forces: high mutation rate, inefficient selection and genetic drift. On the other hand, recent comparisons of the human and chimpanzee Y chromosomes show that the human Y chromosome has not lost any genes since the divergence of humans and chimpanzees between 6-7 million years ago, providing direct evidence that the linear extrapolation model may be flawed.

High mutation rate

The human Y chromosome is particularly exposed to high mutation rates due to the environment in which it is housed. The Y chromosome is passed exclusively through sperm, which undergo multiple cell divisions during gametogenesis. Each cellular division provides further opportunity to accumulate base pair mutations. Additionally, sperm are stored in the highly oxidative environment of the testis, which encourages further mutation. These two conditions combined put the Y chromosome at a risk of mutation 4.8 times greater than the rest of the genome.

Inefficient selection

Without the ability to recombine during meiosis, the Y chromosome is unable to expose individual alleles to natural selection. Deleterious alleles are allowed to "hitchhike" with beneficial neighbors, thus propagating maladapted alleles in to the next generation. Conversely, advantageous alleles may be selected against if they are surrounded by harmful alleles (background selection). This inability to sort through its gene content, the Y chromosome is particularly prone to the accumulation of "junk" DNA. Massive accumulations of retrotransposable elements are scattered throughout the Y. The random insertion of DNA segments often disrupt encoded gene sequences and render them nonfunctional. However, the Y chromosome has no way of weeding out these "jumping genes". Without the ability to isolate alleles, selection cannot effectively act upon them.

A clear, quantitative indication of this inefficiency is the entropy rate of the Y chromosome. Whereas all other chromosomes in the human genome have entropy rates of 1.5-1.9 bits per nucleotide (compared to the theoretical maximum of exactly 2 for no redundancy), the Y chromosome's entropy rate is only 0.84. This means the Y chromosome has a much lower information content relative to its overall length; it is more redundant.

Genetic drift

Even if a well adapted Y chromosome manages to maintain genetic activity by avoiding mutation accumulation, there is no guarantee it will be passed down to the next generation. The population size of the Y chromosome is inherently limited to 1/4 that of autosomes: diploid organisms contain two copies of autosomal chromosomes while only half the population contains 1 Y chromosome. Thus, genetic drift is an exceptionally strong force acting upon the Y chromosome. Through sheer random assortment, an adult

male may never pass on his Y chromosome if he only has female offspring. Thus, although a male may have a well adapted Y chromosome free of excessive mutation, it may never make it in to the next gene pool. The repeat random loss of well-adapted Y chromosomes, coupled with the tendency of the Y chromosome to evolve to have more deleterious mutations rather than less for reasons described above, contributes to the species-wide degeneration of Y chromosomes through Muller's ratchet.

Gene conversion

In 2003, researchers from MIT discovered a process which may slow down the process of degradation. They found that human Y chromosome is able to "recombine" with itself, using palindrome base pair sequences. Such a "recombination" is called gene conversion or *recombinational loss of heterozygosity* (RecLOH).

In the case of the Y chromosomes, the palindromes are not noncoding DNA; these strings of bases contain functioning genes important for male fertility. Most of the sequence pairs are greater than 99.97% identical. The extensive use of gene conversion may play a role in the ability of the Y chromosome to edit out genetic mistakes and maintain the integrity of the relatively few genes it carries. In other words, since the Y chromosome is single, it has duplicates of its genes on itself instead of having a second, homologous, chromosome. When errors occur, it can use other parts of itself as a template to correct them.

Findings were confirmed by comparing similar regions of the Y chromosome in humans to the Y chromosomes of chimpanzees, bonobos and gorillas. The comparison demonstrated that the same phenomenon of gene conversion appeared to be at work more than 5 million years ago, when humans and the non-human primates diverged from each other.

Future evolution

In the terminal stages of the degeneration of the Y chromosome, other chromosomes increasingly take over genes and functions formerly associated with it. Finally, the Y chromosome disappears entirely, and a new sex-determining system arises. Several species of rodent in the sister families Muridae and Cricetidae have reached these stages, in the following ways:

- The Transcaucasian mole vole, *Ellobius lutescens*, the Zaisan mole vole, *Ellobius tancrei*, and the Japanese spinous country rats *Tokudaia osimensis* and *Tokudaia muenninki*, have lost the Y chromosome and SRY entirely. *Tokudaia* spp. have relocated some other genes ancestrally present on the Y chromosome to the X chromosome. Both genders of *Tokudaia* spp. and *Ellobius lutescens* have an XO genotype, whereas all *Ellobius tancrei* possess an XX genotype. The new sex-determining system for these rodents remains unclear.
- The wood lemming *Myopus schisticolor*, the arctic lemming, *Dicrostonyx torquatus*, and multiple species in the grass mouse genus *Akodon* have evolved

fertile females who possess the genotype generally coding for males, XY, in addition to the ancestral XX female, through a variety of modifications to the X and Y chromosomes.
- In the creeping vole, *Microtus oregoni*, the females, with just one X chromosome each, produce X gametes only, and the males, XY, produce Y gametes, or gametes devoid of any sex chromosome, through nondisjunction.

Outside of the rodent family, the black muntjac, *Muntiacus crinifrons*, evolved new X and Y chromosomes through fusions of the ancestral sex chromosomes and autosomes. Primate Y chromosomes, including in humans, have degenerated so much that primates will also evolve new sex determination systems relatively soon, in about 14 million years in humans.

Human Y chromosome

In humans, the Y chromosome spans about 58 million base pairs (the building blocks of DNA) and represents approximately 2% of the total DNA in a male cell. The human Y chromosome contains 86 genes, which code for only 23 distinct proteins. Traits that are inherited via the Y chromosome are called holandric traits.

The human Y chromosome is unable to recombine with the X chromosome, except for small pieces of pseudoautosomal regions at the telomeres (which comprise about 5% of the chromosome's length). These regions are relics of ancient homology between the X and Y chromosomes. The bulk of the Y chromosome which does not recombine is called the "NRY" or non-recombining region of the Y chromosome. It is the SNPs in this region which are used for tracing direct paternal ancestral lines.

Genes

Not including pseudoautosomal genes, genes include:

- NRY, with corresponding gene on X chromosome
  - AMELY/AMELX (amelogenin)
  - RPS4Y1/RPS4Y2/RPS4X (Ribosomal protein S4)

- NRY, other
  - AZF1 (azoospermia factor 1)
  - BPY2 (basic protein on the Y chromosome)
  - DAZ1 (deleted in azoospermia)
  - DAZ2
  - PRKY (protein kinase, Y-linked)
  - RBMY1A1
  - SRY (sex-determining region)
  - TSPY (testis-specific protein)
  - USP9Y
  - UTY (ubiquitously transcribed TPR gene on Y chromosome)

- ZFY (zinc finger protein)

Y-chromosome-linked diseases

Y-chromosome-linked diseases can be of more common types, or very rare ones. Yet, the rare ones still have importance in understanding the function of the Y-chromosome in the normal case.

More common

No vital genes reside only on the Y chromosome, since roughly half of humans (females) do not have Y chromosomes. The only well-defined human disease linked to a defect on the Y chromosome is defective testicular development (due to deletion or deleterious mutation of *SRY*). However, having two X-chromosomes and one Y-chromosome has similar effects. On the other hand, having Y-chromosome polysomy has other effects than masculinization.

Defective Y-chromosome

This results in the person presenting a female phenotype even though that person possesses an XY karyotype (i.e., is born with female-like genitalia). The lack of the second X results in infertility. In other words, viewed from opposite direction, the person goes through defeminization but fails to complete masculinization.

The cause can be seen as an incomplete Y chromosome: the usual karyotype in these cases is 44X, plus a fragment of Y. This usually results in defective testicular development, such that the infant may or may not have fully formed male genitalia internally or externally. The full range of ambiguity of structure may occur, especially if mosaicism is present. When the Y fragment is minimal and nonfunctional, the child usually is a girl with the features of Turner syndrome or mixed gonadal dysgenesis.

XXY

Klinefelter's syndrome (47, XXY) is not an aneuploidy of the Y chromosome, but a condition of having an extra X chromosome, which usually results in defective postnatal testicular function. The mechanism is not fully understood; the extra X does not seem to be due to direct interference with expression of Y genes.

XYY

47,XYY syndrome is caused by the presence of a single extra copy of the Y chromosome in each of a male's cells. 47, XYY males have one X chromosome and two Y chromosomes, for a total of 47 chromosomes per cell. Researchers have found that an extra copy of the Y chromosome is associated with increased stature and an increased incidence of learning problems in some boys and men, but the effects are variable, often minimal, and the vast majority do not know their karyotype. When chromosome surveys

were done in the mid-1960s in British secure hospitals for the developmentally disabled, a higher than expected number of patients were found to have an extra Y chromosome. The patients were mischaracterized as aggressive and criminal, so that for a while an extra Y chromosome was believed to predispose a boy to antisocial behavior (and was dubbed the "criminal karyotype"). Subsequently, in 1968 in Scotland the only ever comprehensive nationwide chromosome survey of prisons found no overrepresentation of 47,XYY men, and later studies found 47,XYY boys and men had the same rate of criminal convictions as 46,XY boys and men of equal intelligence. Thus, the "criminal karyotype" concept is inaccurate and obsolete.

Rare

The following Y-Chromosome-linked diseases are rare, but notable because of their elucidating of the nature of the Y-chromosome.

More than two Y chromosomes

Greater degrees of Y chromosome polysomy (having more than one extra copy of the Y chromosome in every cell, e.g., XYYYY) are rare. The extra genetic material in these cases can lead to skeletal abnormalities, decreased IQ, and delayed development, but the severity features of these conditions are variable.

XX male syndrome

XX male syndrome occurs when there has been a recombination in the formation of the male gametes, causing the SRY-portion of the Y chromosome to move to the X chromosome. When such an X chromosome contributes to the child, the development will lead to a male, because of the SRY gene.

Genetic genealogy

In human genetic genealogy (the application of genetics to traditional genealogy) use of the information contained in the Y chromosome is of particular interest since, unlike other genes, the Y chromosome is passed exclusively from father to son. Mitochondrial DNA, maternally inherited, is used in an analogous way to trace the maternal line.

Non-mammal Y-chromosome

Many groups of organisms in addition to mammals have Y chromosomes, but these Y chromosomes do not share common ancestry with mammalian Y chromosomes. Such groups include Drosophila, some other insects, some fish, some reptiles, and some plants. In Drosophila melanogaster, the Y chromosome does not trigger male development. Instead, sex is determined by the number of X chromosomes. The *D. melanogaster* Y chromosome does contain genes necessary for male fertility. So XXY *D. melanogaster* are female, and *D. melanogaster* with a single X (X0), are male but sterile. There are some species of Drosophila in which X0 males are both viable and fertile.

ZW-chromosomes

Other organisms have mirror image sex chromosomes: the female is "XY" and the male is "XX", but by convention biologists call a "female Y" a W chromosome and the other a Z chromosome. For example, female birds, snakes, and butterflies have ZW sex chromosomes, and males have ZZ sex chromosomes.

Chapter- 13

Introduction to Genetics

Genetics is the study of genes, and tries to explain what they are and how they work. Genes are how living organisms inherit features from their ancestors; for example, children usually look like their parents because they have inherited their parents' genes. Genetics tries to identify which features are inherited, and explain how these features are passed from generation to generation.

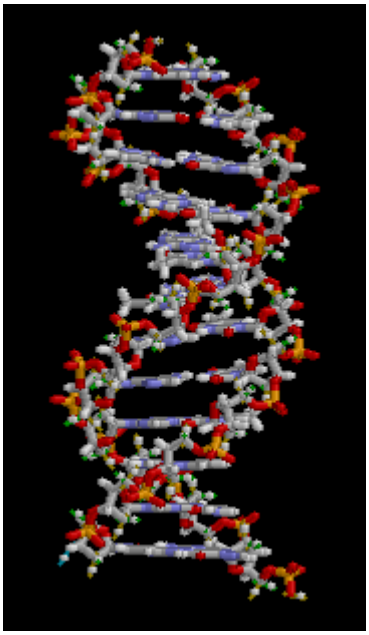
In genetics, a feature of a living thing is called a "trait". Some traits are part of an organism's physical appearance; such as a person's eye-color, height or weight. Other sorts of traits are not easily seen and include blood types or resistance to diseases. Some traits are inherited through our genes, so tall and thin people tend to have tall and thin children. Other traits come from interactions between our genes and the environment, so a child might inherit the tendency to be tall, but if they are poorly nourished, they will still be short. The way our genes and environment interact to produce a trait can be complicated. For example, the chances of somebody dying of cancer or heart disease seems to depend on both their genes and their lifestyle.

Genes are made from a long molecule called DNA, which is copied and inherited across generations. DNA is made of simple units that line up in a particular order within this large molecule. The order of these units carries genetic information, similar to how the order of letters on a page carry information. The language used by DNA is called the genetic code, which lets organisms read the information in the genes. This information is the instructions for constructing and operating a living organism.

The information within a particular gene is not always exactly the same between one organism and another, so different copies of a gene do not always give exactly the same instructions. Each unique form of a single gene is called an allele. As an example, one allele for the gene for hair color could instruct the body to produce a lot of pigment, producing black hair, while a different allele of the same gene might give garbled instructions that fail to produce any pigment, giving white hair. Mutations are random changes in genes, and can create new alleles. Mutations can also produce new traits, such as when mutations to an allele for black hair produce a new allele for white hair. This appearance of new traits is important in evolution.

Inheritance in biology

Genes and inheritance



A section of DNA; the sequence of the plate-like units (nucleotides) in the center carries information.

Red hair is a recessive trait.

Genes are inherited as units, with two parents dividing out copies of their genes to their offspring. You can think of this process like mixing two hands of cards, shuffling them, and then dealing them out again. Humans have two copies of each of their genes and when people reproduce they make copies of their genes and put them into eggs or sperm, but only put in one copy of each type of gene. When an egg joins with a sperm, this gives a child a complete set of genes. This child will have the same number of genes as its parents, but for any gene one of their two copies will come from their father, and one from their mother.

The effects of this mixing depends on the types (the alleles) of the gene you are interested in. If the father has two copies for an allele for red hair, and the mother has two copies for brown hair, all their children will get the two alleles that give different instructions, one for red hair and one for brown. The hair color of these children depends on how these alleles work together. If one allele overrides the instructions from another, it is called the *dominant* allele, and the allele that is overridden is called the *recessive* allele. In the case of a daughter with alleles for both red and brown hair, brown is dominant and she ends up with brown hair.

Although the red color allele is still there in this brown hair girl, it doesn't show. This is a difference between what you see on the surface (the traits of an organism, called its phenotype) and the genes within the organism (its genotype). In this example you can call the allele for brown "B" and the allele for red "b". (It is normal to write dominant alleles with capital letters and recessive ones with lower-case letters.) The brown hair daughter has the "brown hair phenotype" but her genotype is Bb, with one copy of the B allele, and one of the b allele.

Now imagine that this woman grows up and has children with a brown hair man who also has a Bb genotype. Her eggs will be a mixture of two types, one sort containing the B allele, and one sort the b allele. Similarly, her partner will produce a mix of two types of sperm containing one or the other of these two alleles. Now, when the transmitted genes are joined up in their offspring, these children have a chance of getting either brown or red hair, since they could get a genotype of BB = brown hair, Bb = brown hair or bb = red hair. In this generation, there is therefore a chance of the recessive allele showing itself in the phenotype of the children - some of them may have red hair like their grandfather.

Many traits are inherited in a more complicated way than the example above. This can happen when there are several genes involved, each contributing a small part to the end result. Tall people tend to have tall children because their children get a package of many alleles that each contribute a bit to how much they grow. However, there are not clear groups of "short people" and "tall people", like there are groups of people with brown or red hair. This is because of the large number of genes involved; this makes the trait very variable and people are of many different heights. Unlike common belief, the green/blue eye traits are also inherited in this complex inheritance model. Inheritance can also be complicated when the trait depends on the interaction between genetics and the

environment. This is quite common, for example, if a child does not eat enough nutritious food this will not change traits like eye color, but it could stunt their growth.

Inherited diseases

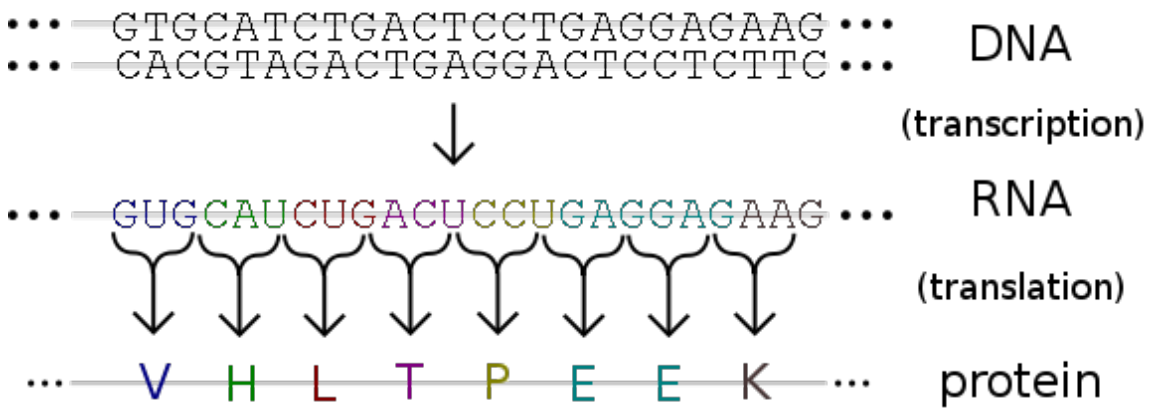
Some diseases are hereditary and run in families; others, such as infectious diseases, are caused by the environment. Other diseases come from a combination of genes and the environment. Genetic disorders are diseases that are caused by a single allele of a gene and are inherited in families. These include Huntington's disease, Cystic fibrosis or Duchenne muscular dystrophy. Cystic fibrosis, for example, is caused by mutations in a single gene called *CFTR* and is inherited as a recessive trait.

Other diseases are influenced by genetics, but the genes a person gets from their parents only change their risk of getting a disease. Most of these diseases are inherited in a complex way, with either multiple genes involved, or coming from both genes and the environment. As an example, the risk of breast cancer is 50 times higher in the families most at risk, compared to the families least at risk. This variation is probably due to a large number of alleles, each changing the risk a little bit. Several of the genes have been identified, such as *BRCA1* and *BRCA2*, but not all of them. However, although some of the risk is genetic, the risk of this cancer is also increased by being overweight, drinking a lot of alcohol and not exercising. A woman's risk of breast cancer therefore comes from a large number of alleles interacting with her environment, so it is very hard to predict.

How genes work

Genes make proteins

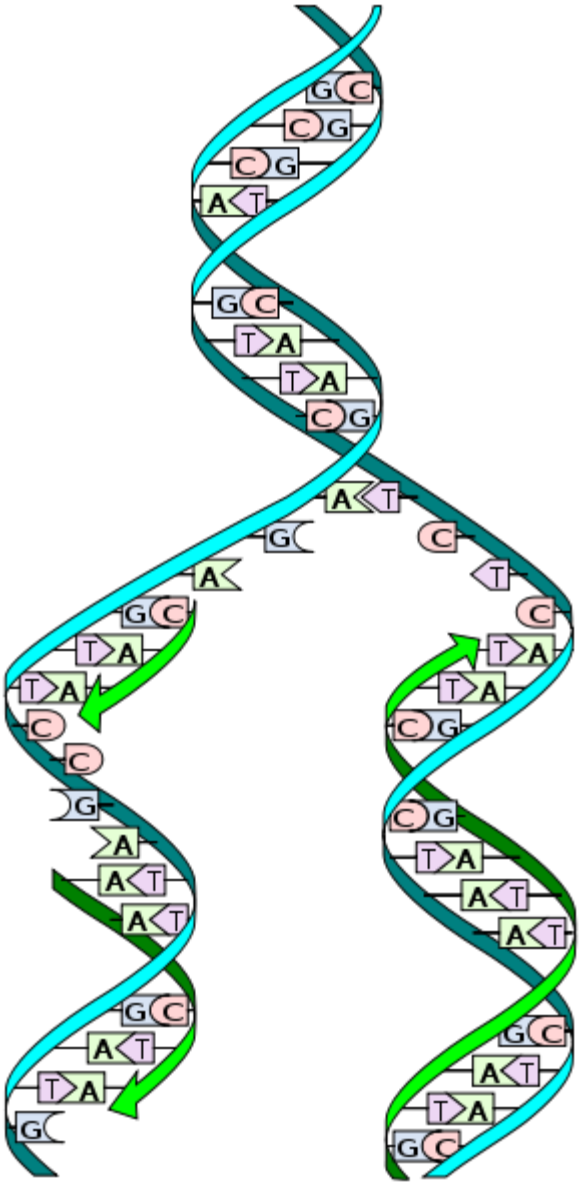
The function of genes is to provide the information needed to make molecules called proteins in cells. Cells are the smallest independent parts of organisms: the human body contains about 100 trillion cells, while very small organisms like bacteria are just one single cell. A cell is like a miniature and very complex factory that can make all the parts needed to produce a copy of itself, which happens when cells divide. There is a simple division of labor in cells - genes give instructions and proteins carry out these instructions, tasks like building a new copy of a cell, or repairing damage. Each type of protein is a specialist that only does one job, so if a cell needs to do something new, it must make a new protein to do this job. Similarly, if a cell needs to do something faster or slower than before, it makes more or less of the protein responsible. Genes tell cells what to do by telling them which proteins to make and in what amounts.



Genes are expressed by being transcribed into RNA, and this RNA then translated into protein.

Proteins are made of a chain of 20 different types of amino acid molecules. This chain folds up into a compact shape, rather like an untidy ball of string. The shape of the protein is determined by the sequence of amino acids along its chain and it is this shape that, in turn, determines what the protein will do. For example, some proteins have parts of their surface that perfectly match the shape of another molecule, allowing the protein to bind to this molecule very tightly. Other proteins are enzymes, which are like tiny machines that alter other molecules.

The information in DNA is held in the sequence of the repeating units along the DNA chain. These units are four types of nucleotides (A,T,G and C) and the sequence of nucleotides stores information in an alphabet called the genetic code. When a gene is read by a cell the DNA sequence is copied into a very similar molecule called RNA (this process is called transcription). Transcription is controlled by other DNA sequences (such as promoters), which show a cell where genes are, and control how often they are copied. The RNA copy made from a gene is then fed through a structure called a ribosome, which translates the sequence of nucleotides in the RNA into the correct sequence of amino acids and joins these amino acids together to make a complete protein chain. The new protein then folds up into its active form. The process of moving information from the language of DNA into the language of amino acids is called translation.



DNA replication. DNA is unwound and nucleotides are matched to make two new strands.

If the sequence of the nucleotides in a gene changes, the sequence of the amino acids in the protein it produces may also change - if part of a gene is deleted, the protein produced will be shorter and may not work any more. This is the reason why different alleles of a gene can have different effects in an organism. As an example, hair color depends on how much of a dark substance called melanin is put into the hair as it grows. If a person has a normal set of the genes involved in making melanin, they make all the proteins needed and they grow dark hair. However, if the alleles for a particular protein have different sequences and produce proteins that can't do their jobs, no melanin will be produced and the hair will be white. This condition is called albinism and the person with this condition is called an albino.

Genes are copied

Genes are copied each time a cell divides into two new cells. The process that copies DNA is called DNA replication. It is through a similar process that a child inherits genes from its parents, when a copy from the mother is mixed with a copy from the father.

DNA can be copied very easily and accurately because each piece of DNA can direct the creation of a new copy of its information. This is because DNA is made of two strands that pair together like the two sides of a zipper. The nucleotides are in the center, like the teeth in the zipper, and pair up to hold the two strands together. Importantly, the four different sorts of nucleotides are different shapes, so in order for the strands to close up properly, an **A** nucleotide must go opposite a **T** nucleotide, and a **G** opposite a **C**. This exact pairing is called base pairing.

When DNA is copied, the two strands of the old DNA are pulled apart by enzymes which move along each of the two single strands pairing up new nucleotide units and then zipping the strands closed. This produces two new pieces of DNA, each containing one strand from the old DNA and one newly made strand. This process isn't perfect and sometimes the proteins will make mistakes and put the wrong nucleotide into the strand they are building. This causes a change in the sequence of that gene. These changes in DNA sequence are called mutations. Mutations produce new alleles of genes. Sometimes these changes stop the gene from working properly, like the melanin genes discussed above. In other cases these mutations can change what the gene does or even let it do its job a little better than before. These mutations and their effects on the traits of organisms are one of the causes of evolution.

Genes and evolution

Mice with different coat colors.

A population of organisms evolves when an inherited trait becomes more common or less common over time. For instance, all the mice living on an island would be a single population of mice. If over a few generations, white mice went from being rare, to being a large part of this population, then the coat color of these mice would be evolving. In terms of genetics, this is called a change in allele frequency—such as an increase in the frequency of the allele for white fur.

Alleles become more or less common either just by chance (in a process called genetic drift), or through natural selection. In natural selection, if an allele makes it more likely that an organism will survive and reproduce, then over time this allele will become more common. But if an allele is harmful, natural selection will make it less common. For

example, if the island was getting colder each year and was covered with snow for much of the time, then the allele for white fur would become useful for the mice, since it would make them harder to see against the snow. Fewer of the white mice would be eaten by predators, so over time white mice would out-compete mice with dark fur. White fur alleles would become more common, and dark fur alleles would become more rare.

Mutations create new alleles. These alleles have new DNA sequences and can produce proteins with new properties. So if an island was populated entirely by black mice, mutations could happen creating alleles for white fur. The combination of mutations creating new alleles at random, and natural selection picking out those which are useful, causes adaptation. This is when organisms change in ways that help them to survive and reproduce.

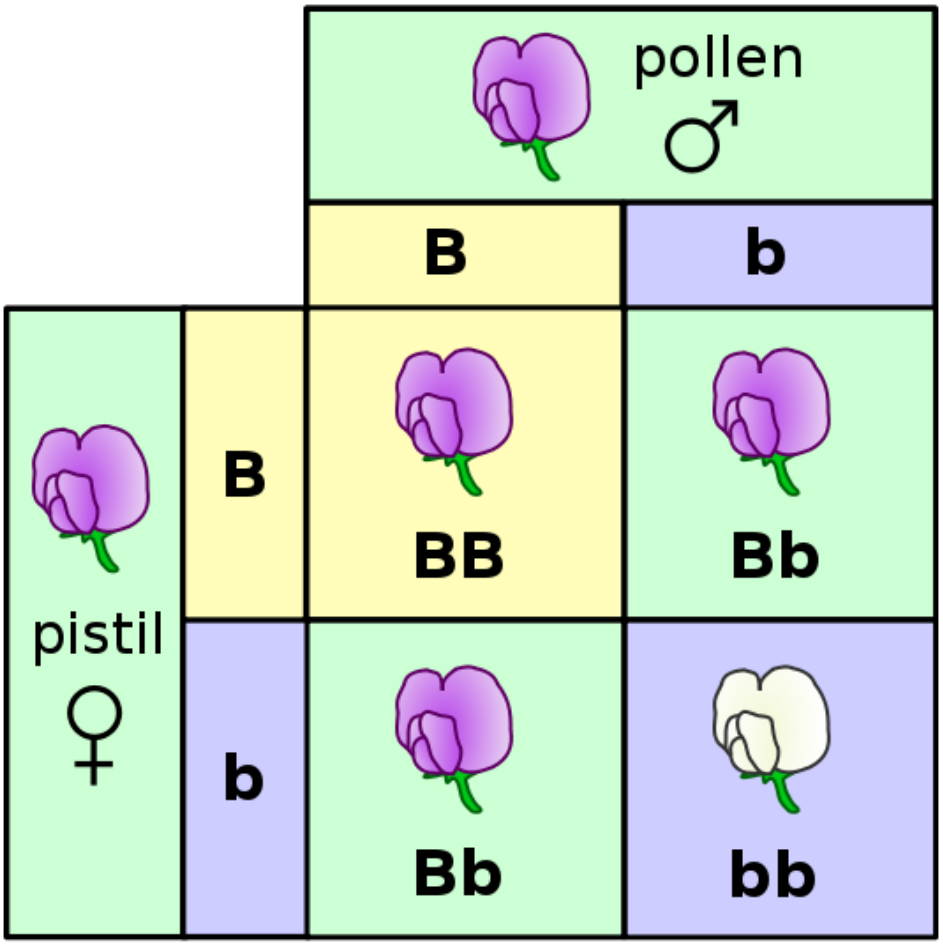
Genetic engineering

Since traits come from the genes in a cell, putting a new piece of DNA into a cell can produce a new trait. This is how genetic engineering works. For example, crop plants can be given a gene from an Arctic fish, so they produce an antifreeze protein in their leaves. This can help prevent frost damage. Other genes that can be put into crops include a natural insecticide from the bacteria *Bacillus thuringiensis*. The insecticide kills insects that eat the plants, but is harmless to people. In these plants the new genes are put into the plant before it is grown, so the genes will be in every part of the plant, including its seeds. The plant's offspring will then inherit the new genes, something which has led to concern about the spread of new traits into wild plants.

The kind of technology used in genetic engineering is also being developed to treat people with genetic disorders in an experimental medical technique called gene therapy. However, here the new gene is put in after the person has grown up and become ill, so any new gene will not be inherited by their children. Gene therapy works by trying to replace the allele that causes the disease with an allele that will work properly.

Chapter- 14

Genotype



Here the relation between genotype and phenotype is illustrated, using a Punnett square, for the character of petal colour in pea. The letters B and b represent genes for colour and the pictures show the resultant flowers.

The **genotype** is the genetic makeup of a cell, an organism, or an individual (i.e. the specific allele makeup of the individual) usually with reference to a specific character under consideration. For instance, the human CFTR gene, which encodes a protein that transports chloride ions across cell membranes, can be dominant (A) as the normal

version of the gene, or recessive (a) as a mutated version of the gene. It is receiving two recessive alleles that individuals will be diagnosed with cystic fibrosis. It is generally accepted that inherited genotype, transmitted epigenetic factors, and non-hereditary environmental variation contribute to the phenotype of an individual.

Non-hereditary DNA mutations are not classically understood as representing the individual's genotype. Hence, scientists and physicians sometimes talk for example about the (geno)type of a particular cancer, that is the genotype of the disease as distinct from the diseased.

Genotype and genomic sequence

One's genotype differs subtly from one's genomic sequence. A sequence is an absolute measure of base composition of an individual, or a representative of a species or group; a genotype typically implies a measurement of how an individual *differs* or is specialized within a group of individuals or a species. So typically, one refers to an individual's genotype with regard to a particular gene of interest and, in polyploid individuals, it refers to what combination of alleles the individual carries. The genetic constitution of an organism is referred to as its genotype. Such as the letters Bb.

Genotype and Mendelian inheritance

The distinction between genotype and phenotype is commonly experienced when studying family patterns for certain hereditary diseases or conditions, for example, haemophilia. Due to the diploidy of humans (and most animals), there are two alleles for any given gene. These alleles can be the same (homozygous) or different (heterozygous), depending on the individual. With a dominant allele, the offspring is guaranteed to inherit the trait in question irrespective of the second allele.

In the case of an albino with a recessive allele (aa), the phenotype depends upon the other allele (Aa, aA or AA). An affected person mating with a heterozygous individual (Aa or aA, also **carrier**) there is a 50-50 chance the offspring will be albino's (phenotype. If a heterozygote mates with another heterozygote, there is 75% chance passing the gene on and only a 25% chance that the gene will be displayed. A homozygous dominant (AA) individual has a normal phenotype and no risk of abnormal offspring. A homozygous recessive individual has an abnormal phenotype and is guaranteed to pass the abnormal gene onto offspring.

In the case of haemophilia, it is sex linked thus only carried on the X chromosome. Only females can be a **carrier** in which the abnormality is not displayed. This woman has a normal phenotype, but runs a 50-50 chance, with an unaffected partner, of passing her abnormal gene on to her offspring. If she mated with a man with hemophilia (another carrier) there would be a 75% chance of passing on the gene.

Genotype and mathematics

Inspired by the biological concept and usefulness of genotypes, computer science employs simulated phenotypes in genetic programming and evolutionary algorithms. Such techniques can help evolve mathematical solutions to certain types of otherwise difficult problems.

Determining Genotype

Genotyping is the process of elucidating the genotype of an individual with a biological assay. Also known as a *genotypic assay*, techniques include PCR, DNA fragment analysis, allele specific oligonucleotide (ASO) probes, DNA sequencing, and nucleic acid hybridization to DNA microarrays or beads. Several common genotyping techniques include restriction fragment length polymorphism (*RFLP*), terminal restriction fragment length polymorphism (*t-RFLP*), amplified fragment length polymorphism (*AFLP*), and multiplex ligation-dependent probe amplification (*MLPA*).

DNA fragment analysis can also be used to determine such disease causing genetics aberrations as microsatellite instability (*MSI*), *trisomy* or aneuploidy, and loss of heterozygosity (*LOH*). MSI and LOH in particular have been associated with cancer cell genotypes for colon, breast and cervical cancer.

The most common chromosomal aneuploidy is a trisomy of chromosome 21 which manifests itself as Down syndrome. Current technological limitations typically allow only a fraction of an individual's genotype to be determined efficiently.